

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

Hondatze Kognitibo Arinaren detekzio goiztiarrerako Hizketa Ezagutza Automatikoan oinarrituriko ekarpenak

Euskal Herriko Unibertsitateko Ingeniaritzan Doktore titulua eskuratzeko aurkeztua

Unai Martínez de Lizarduy Stürtze

Zuzendaria

Miren Karmele López de Ipiña Peña

Donostia, 2017

*Markel, Irati, Marta,
Ekain, Ama eta Aitari.*

Maite zaituztet.

Esker onak

Esker onekin hasi aurretik, ikerketarako baldintza kaxkarrak direla bide emigratu behar izan dutenak eta ikerkuntzan ibili arren ozta-ozta aitorea jasotzen duten “Marie Curie ezezagun” guztiak gogora ekarri nahi nituzke.

Ezagutzen nauenak badaki nire esker ona berari zuzendua doala. Dena den, barren-barrenetik sortzen zaizkidanak hitzez adierazten saiatuko naiz:

Alzheimer gaitzak jota amesten jarraitzen duzuen guztioi eta beren senitartekoei.

Karmele, nire tesi zuzendariari, ikerketaren munduan murgiltzen laguntzeagatik baina batez ere alor pertsonalean beti erakusten duen hurbiltasunagatik.

Lan honi urtetan zehar zimenduak jarri dizkiozuen guztioi: Jirí, Pedro, Marcos, Jordi, Pili, Nora, Aitzol, Victoria, Suni, Carlos ...

CITA-Alzheimer-eko langileei, bereziki, Maite San Sebastián, Miriam Ecay, Pablo Martínez-Lage eta Ainara Estangari. BIODONOSTIAko Alberto Bergarecheri eta Mataró Hospitaleko Pilar Sanz, Pere Clavé eta Weslania Nascimentori. AFAEI Elkarteari.

Nire kuadrilari. Doktore izateko asmoa nuela eta adarra jotzen zidaten horiei. Badakizuen arren, zuentzako ere badut txokoa nire bihotzean. Ezin liteke beste modura izan.

Baldintzarik gabeko lagunei: Bilbotar, Azkoitiar eta Altxatarrari. Beti lagun.

Iñaki lagun fin, zintzo eta eskuzabalari.

Aupa Brother! Nire anaitxo txikia. Beti anaia txikia izango zara niretzat. Sikologo eta lagun fina, beti iparra ez galtzeko gomendioak ematen, animoak ematen. Bizi izan ditudan hilabete gogorak oso ondo ezagutzen dituen. Niretzat, nahiz eta urrun bizi, erreferentea.

“Gran maestro”-ari, nire aholkulari zientifikoari. Inoiz ezagutu dudana pertsona xumeenari eta zientziaren mundua hurbildik ezagutzeko gogoia kutsatu zidanari. Zuri aita. Krak bat haiz!

Amatxooo kuttuna! 30 urterekin Ingeniaritza ikasteko asmoa nuela esan nuenean itsu-itsu aurrera egiteko bultzada handiena eman zidanari. Honaino iritsi gara ama, elkarrekin. Zu gabe ezinezkoa zatekeen bidea, badakizu.

Nire bidelagunari. Marta, hitzak soberan daude. Maratoiaren lehen etapa bukatzera nihoa eta zure arnasa nahitaezkoa izan da. Zure konplizitatearekin dena errazagoa da. Mila esker elkarrekin sortu eta hezten ari garen txiki zoragarriengatik.

Irati eta Markel, ez dago zuengandik sentitzen dudana deskribatu dezakeen hitzik. Nire bizipoza zarete!

Iratitxo, azkenean “liburu potoloa” bukatu dut! Agindu bezala elkarrekin koloreztatuko dugu.

Atatatatata! Jadanik hitz egiten hasi zara Markeltxo! Zu ere liburua kolorezko ametsez tindatzeko prest?

Zuei.

Laburpena

Alzheimerren Gaitzak edo Alzheimer's Disease (AD) eragin lazgarria izan lezake epe luzean gizartean eta horri aurre egiteko diagnostiko goiztiarra eta zehatza lagungarri izan liteke. Hala ere, diagnostikatzeko biomarkatzaile berriek dakartzaten kostuek eta behar teknologikoez ezinezko bihurtzen dute oroimen-arazoak dituzten paziente guztiei testak egitea. Hortaz, aurretiko bahetze optimoa funtsezkoa da eta diagnostikatzeko teknika adimendu ez-inbasiboak ere oso tresna baliagarriak izan daitezke hautemate goiztiarra egiteko. Hori dela eta, biomarkatzaile garesti eta teknika inbasiboen aurrean, alternatiba erabilerrazak izango diren sistema merke eta ez-inbasiboak garatzea da. Ibilbidea luzea da eta lan honek horren aldeko ekarpena izatea duen helburu.

Gaixotasunaren lehen fasean Narriadura Kognitibo Arina edo Mild Cognitive Impairment (MCI) delakoa ager daiteke, baina ez da behar adina larria eguneroko bizitza edo gaitasun independenteak eragozteko. Galera horrek Alzheimer Gaitza garatzeko arriskua areagotu dezake, eta gaur egun, garaiz detektatzea da espezialista medikuen erronketako bat, diagnostiko goiztiar batek gaitzari aurre egiteko neurriak hartzea ahalbidetu eta pazienteen bizimodua hobea baitezake. MCI-ren bereizgarrietako bat eta adierazle goiztiarrena oroimen episodikoaren eta hizketaren nahasmenduak dira. Hortaz, MCI dutenen ahots-seinaleen ikerketak jarraitu beharreko bidea zein den erakuts diezaguke AD garaiz detektatzeko.

MCI-k apurtuta dauden mezueta eramatzen gaitu, disfluentzietara alegia. Hizkera jariakor batean gertatzen den edozein etendura, irregulartasun edo elementu ez-lexikoak dira disfluentziak, eta hizketa eten dezakete. Disfluentzia horiek handitzen joatea narriadura kognitiboaren seinale argia izan daiteke eta neuroendekapenezko gaixotasunetan pairatutako narriadura kognitiboa neurtu eta kuantifikatzeko balio dezakete.

Alzheimerdun gaixoengan, mintzamina ez ezik, erantzun emozionala ere kaltetu egiten da. Emozioak giza gogoaren arkitekturarekin zerikusia dituzten prozesu kognitiboak dira, eta erabakiak hartzarekin eta oroimenaren kudeaketa edota arretarekin zerikusia dute, eta aldi berean ere, horiek hertsiki lotuta dauden komunikazioarekin. Hortaz, erantzun eta kudeaketa emozionalak ere badira gaitzaren hasierako fase horietan nahasten diren beste komunikazio-elementu batzuk, eta disfluentzia bezala, emozio-erantzuna narriadura kognitiboa neurtzeko adierazlea izan daiteke.

Hortaz, zenbait atazaren bidez sortutako ahots-laginen azterketak direla medio, disfluentzia eta emozio-erantzuna jaso daitezke. Hizkuntzarekiko independenteak diren parametroak bildu eta horien hizketaren nahasmenduak ezaugarritu badaitezke, ekarpena lagungarria izan daiteke diagnostikoa egingo duten espezialistentzat.

Lehengaiak ahots-laginak direnez, ingurune kliniko zein etxeko ingurunean egindako atazen bidez grabazioak egin eta datu-baseak osatu dira, osasun-guneen

irizpide etikoak kontuan hartuta eta. Datu-base horien ikerketaren bidez, galera kognitiboaren garapena neurtu, kuantifikatu, balioztatu eta sailkatu nahi da. Gaitzaren etapa desberdinak hautematen laguntzeko ekarpena egin nahi da, eta horretarako, hizkuntzarekiko independenteak diren parametroen azterketa automatikorako teknika eta metodologiak garatu dira. Mintzamen automatikoaren analisisian oinarritutako multihurbilketa ez-lineala egin da, zeinak hizketa-analisisian erabiltzen diren denbora-serieen konplexutasunaren neurtze kuantitatiboa eman diezaguke.

Bestalde, ezaugarrien hautaketa automatikoa burutu da, eta test estatistiko klasikoak ikasketa automatikoko ereduarekin konbinatu dira. Sistema denbora errealerako bideratua dago eta Deep Learning metodologiak ikasketa automatikoko metodologiekin integratzen dira. Lortutako soluzioek Ingeniaritza Biomedikuan oso ohikoak diren eszenatokitik aplikazio zuzena dute. Big Data eta datu konplexu eskaseko kasuetan aplika daitezke.

Azpimarratzekoa da, espezialistentzako euskarri bezala, inguru erreal eta konplexuetan zenbait ataza web bidez egitea erraztu eta ahots-laginen ezaugarritze, hautaketa zein sailkapena egiten duen teknologia merke, naturala, unibertsala, ez-inbasiboa eta zifratua garatu dela. Beti ere aurre-screening erabilgarria aurkitzearen alde. Garatutako metodologiak erabilera orokorreko metodologiak dira eta beraz beste gaitza kognitiboetara aplika daitezke, haien artean, oinarritzko dardara eta Parkinson gaitzera.

Introducción

La enfermedad de Alzheimer o Alzheimer Disease (AD) puede acarrear un fuerte coste a largo plazo en la sociedad y para enfrentarse a ello, un diagnóstico precoz y preciso podría ser de gran ayuda. Sin embargo, los nuevos biomarcadores de diagnóstico requieren unos gastos y exigencias tecnológicas que imposibilitan su aplicación a todos los supuestos pacientes. Por esa razón, un pre-screening es fundamental y las técnicas de diagnóstico no invasivas pueden resultar también ser una herramienta de gran eficacia para el diagnóstico precoz. Como consecuencia, frente a los biomarcadores de alto coste y a las técnicas invasivas, la alternativa puede ser el desarrollo de sistemas de bajo coste y no invasivos de fácil utilización. El camino a recorrer es largo y avanzar en el mismo es el objetivo de esta investigación.

En la primera fase de la enfermedad puede darse un deterioro cognitivo leve o Mild Cognitive Impairment (MCI), que no tiene la suficiente entidad como para impedir el desarrollo de la actividad diaria normal. Sin embargo, ese deterioro puede incrementar el riesgo de desarrollar la enfermedad de Alzheimer y, hoy en día, el diagnóstico precoz es uno de los mayores retos médicos, porque un diagnóstico precoz puede permitir tomar medidas para hacer frente a la enfermedad y mejorar el nivel de vida de los pacientes. Una de las características y de los indicadores más precoces del MCI son los trastornos de memoria episódica y de lenguaje. Por eso mismo, investigar las señales del habla en las personas con MCI es una importante vía para detectar a tiempo la enfermedad de Alzheimer.

MCI supone interrupciones en el habla, es decir, disfluencias. Las disfluencias son interrupciones, irregularidades o elementos no-léxicos que ocurren en el hablar fluido, y que lo pueden llegar a interrumpir. El que estas disfluencias vayan en aumento es clara señal de un deterioro cognitivo leve y puede tener validez a la hora de medir y cuantificar el deterioro cognitivo que traen consigo las enfermedades neurodegenerativas.

Los enfermos de Alzheimer no sólo padecen deterioro en el habla; también en el aspecto emocional. Las emociones son procesos cognitivos relacionados con la arquitectura de la mente humana, y tienen que ver con la gestión en la toma de decisiones y en el recuerdo y también con la atención, y, al mismo tiempo, están estrechamente relacionadas con la comunicación. Consecuentemente, en esa fase inicial de la enfermedad, las reacciones y gestiones emocionales son también elementos comunicativos, y al igual que la disfluencia, pueden ser indicadores a la hora de medir el deterioro cognitivo de la respuesta emocional.

Por esa razón, analizando muestras de voz obtenidas por medio de una serie de tareas, se pueden recoger la disfluencia y la respuesta emocional. Se pueden recopilar parámetros independientes de la lengua y caracterizar los correspondientes deterioros lingüísticos, lo que puede resultar de gran ayuda para los especialistas a la hora de diagnosticar. Dado que la materia prima son las muestras de voz, por medio de una serie de tareas se han realizado grabaciones en entornos clínicos y familiares y

completado bases de datos, siempre teniendo en cuenta los criterios éticos de los centros de salud.

Con el análisis de esas bases de datos, se pretende medir, cuantificar, valorar y clasificar el desarrollo del deterioro cognitivo. El objetivo es realizar una aportación que ayude a diferenciar las diferentes etapas de la enfermedad, para lo cual se han desarrollado técnicas y metodologías automáticas de análisis de parámetros independientes del lenguaje. Se ha realizado una multi-aproximación no lineal basada en el análisis del lenguaje automático, que nos puede dar la medición automática de la complejidad de las series temporales.

Por otro lado, se ha realizado una elección automática de las características, y se han combinado los modelos de aprendizaje automático con los test estadísticos clásicos. El sistema está orientado al tiempo real y las metodologías de Deep Learning se integran con las metodologías de aprendizaje automático. Los resultados obtenidos se pueden aplicar en escenarios muy habituales de la ingeniería biomédica. Es aplicable para casos de Big Data o de carencia de datos complejos.

Hay que poner de relieve que, como soporte para los especialistas, se ha desarrollado una tecnología de bajo coste, natural, universal, no-invasiva y cifrada que, en un entorno real y complejo y por medio de tareas a través de la web, permite caracterizar, elegir y clasificar las muestras de voz. Y, en todo momento, orientado a encontrar un pre-screening útil. Puesto que las metodologías desarrolladas son de uso general, estas son aplicables a otros trastornos como el temblor esencial o el mal de Parkinson.

Abstract

In the future, Alzheimer Disease (AD) is likely to impose a high price on society. In order to ameliorate that probability, a chance of an early and accurate diagnosis could be of the utmost importance. That latest diagnostic biomarkers require expensive equipment, which obviates its application to the increasing number of suspected victims. For this reason, it becomes essential to be able to pre-screen all those who might be affected by the disease. Furthermore, these diagnostic, non invasive techniques could also become a highly useful tool to ascertain early diagnosis. The development of low cost, non invasive and easily applied techniques could become an alternative to high cost, invasive procedures. The road ahead is long. The purpose of this study is to forge ahead and advance in our investigation.

In the first phase of the disease, a mild cognitive impairment (MCI) might occur, but this deficiency does not have sufficient entity to impede the normal function of daily activity. However, that type of deteriorative stage might increase the risk of developing Alzheimer's disease. Nowadays, early diagnosis is one of the greatest and important challenges facing the medical world. If achieved, it could confirm the illness, allow the physician to take the appropriate measures and thus, improve the patient's quality of life. One of the characteristics and early indicators of MCI are episodic memory and language disorders. Therefore, the investigation of perceived language disorders in people suffering from MCI is an important way of detecting Alzheimer on time.

MCI causes speech disfluency. Disfluency means interruptions, irregularities of non-lexical impairments which occur during fluent speech and can even halt it. When this disfluency increases it sends a clear signal of mild cognitive impairment, and can be a trustworthy factor when measuring and quantifying the cognitive deterioration that are part and parcel of neurodegenerative diseases.

Alzheimer's patients not only show speech deterioration; they also suffer in the realm of emotion. Emotions are cognitive processes related to the architecture of the human brain. They deal with decision making, memories and attentiveness. At the same time, emotions are tightly related to communication. Consequently, in that first phase of the illness, emotional reactions and efforts are all elements of communication. Just as it happens with disfluency, they can become an important factor when measuring the cognitive deterioration of the emotional response.

For the reasons stated above, by analyzing a collection of voice samples obtained through a series of various tasks, disfluency and emotional response can be gathered. Parameters independent from language can be collected, and different linguistic deteriorations can be typified, which, in turn, could be of great help to the diagnostician when evaluating the patient.

Given the fact that the basic component of this study is voice samples, we have taped and completed database with the aid of several focused tasks, always respecting the code of ethics of the Health Centers and the various clinical and family environments. Through the analysis of those databases, one attempts to measure, quantify, evaluate and classify the development of the cognitive decline. The objective is to envision, realize and offer a method that will help differentiate the progressive stages of the illness. For this purpose, we have developed techniques and methodologies of parameter analysis independent of the language itself. A non-linear multi-approach based on the analysis of the automatic language has been completed, which can give us the automatic measurement of the time series.

On the other hand, we have carried out an automatic election of its characteristics, and we have combined the models of automatic training with the traditional statistical tests. The system, based on this model is based on actual time, in which the Deep Learning methodologies are combined with that of the Automatic Training. The results we have obtained show great completed, affinity for application on very common cases of biomedical engineering; among others, in Big Data or situations of paucity of complex information.

We wish to emphasize that, intended as a useful support for medical specialists, we have developed a low-cost, natural, universal, non-invasive and coded technology, that in a real and complex environment and through tasks performed through the web, makes it possible to characterize, select and classify voice samples. And at all times, favoring the usage of pre-screening. Finally, the methodologies we have developed are ones that can be applied across the board. Therefore, this technology could also be usefully applied to illnesses such as Essential Tremor or Parkinson.

Aurkibidea

Esker onak.....	V
Laburpena.....	VII
Introducción	IX
Abstract	XI
Aurkibidea.....	XIII
Irudien aurkibidea	XIX
Taulen aurkibidea	XXIII
Laburtzapenak	XXV
1. KAPITULUA: SARRERA.....	1
1.1 Sarrera	1
1.1.1 Alzheimerren Gaitza (AD)	1
1.1.2 Narriadura Kognitibo Arina (MCI).....	2
1.2 Lanaren xede eta motibazioa	2
1.3 Helburuak	4
1.3.1 Helburu nagusia.....	4
1.3.2 Helburu orokorrak	4
1.3.3 Helburu zehatzak.....	4
1.3.4 Helburu gehigarriak	5
1.4 Hipotesiak	5
1.5 Metodologiaren alderdi orokorrak.....	6
1.6 Egitura.....	7
1.6.1 Kapituluak.....	7
1.6.2 Tesiak sorturiko argitarapenak.....	8
2. KAPITULUA: ARTEAREN EGOERA.....	11
2.1 Sarrera	11
2.2 Komunikazio-defizitak AD-n	12
2.2.1 Alzheimerren Gaitza (AD)	13
2.2.2 Narriadura Kognitibo Arina (MCI).....	13
2.2.3 Mintzamenean defizita.....	14
2.2.4 Emozioan eragina	15
2.2.5 Disfluentzian eragina	16

2.2.6 Egitura biomekanikoan eragina.....	17
2.3 Eskuko metodologiak	17
2.3.1 Sarrera	17
2.3.2 Ahozko komunikazio-gaitasunen ebaluaketa.....	17
2.3.2.1 Hitz-jario kategorikorako atazak	17
2.3.2.2 Hizketa irekiagoko atazak.....	19
2.3.2.3 Hizketa ez-murriztuko atazak.....	20
2.3.3 Screening test klasikoak	20
2.3.3.1 Hizkuntza-gaitasuna hautematea.....	21
2.3.3.1.1 MMSE	21
2.3.3.1.2 BNT	21
2.3.3.1.3 CERAD	22
2.3.3.1.4 RAVLT.....	22
2.3.3.1.5 CVF.....	24
2.3.3.2 Emozio-egoera hautematea	24
2.4 Teknika automatikoak	25
2.4.1 Hizketa automatikoaren hautematea	28
2.4.2 Hizketa automatikoaren hautemate ez-lineala.....	31
2.5 Neurona-Sare Konboluzionalak (CNN)	32
2.5.1 ILSVRC aurrekariak	33
2.5.2 Beste aurrekariak.....	38
3. KAPITULUA: MATERIALAK	41
3.1 Sarrera	41
3.2 Datu-baseak.....	42
3.3 Atazak	43
3.3.1 Animaliak izendatzeko jarioaren ataza (CVF)	43
3.3.1.1 Kontrol Taldea edo Control Group (CR)	44
3.3.1.2 Mild Cognitive Impairment (MCI) taldea.....	45
3.3.1.3 Baztertze-irizpideak.....	45
3.3.1.4 PGA-OREKA datu-basea.....	45
3.3.2 Irudien deskripzioari buruzko ataza (PD)	46
3.3.2.1 MINI-PGA datu-basea.....	46
3.3.3 Bat-bateko hizketari buruzko ataza (SS).....	47

3.3.3.1 AZTIAHO datu-basea	48
3.3.3.2 Datu-basearen iragazketa.....	49
3.3.3.3 AZTIAHORE datu-basea	49
3.3.3.4 AZTITXIKI datu-basea.....	49
4. KAPITULUA: METODOAK	51
4.1 Sarrera	51
4.2 Ezaugarrien erauzketa eta hautaketa.....	54
4.2.1 Ezaugarri linealak.....	56
4.2.2 Ezaugarri ez-linealak.....	57
4.2.2.1 Fraktalak	57
4.2.2.1.1 Higuchi-ren Dimentsio Fraktala.....	58
4.2.2.1.2 Katz eta Castiglioni-ren Dimentsio Fraktala	59
4.2.2.2 Entropia	61
4.2.2.2.1 Shannon-en Entropia.....	61
4.2.2.2.2 Gutxi gorabeherako Entropia vs Lagin-Entropia	62
4.2.2.2.3 Permutazio Entropia.....	64
4.2.2.2.4 Eskala anitzeko Permutazio Entropia	64
4.2.2.3 Emozio-tenperatura	66
4.2.2.4 Disfluentzien azterketa.....	66
4.3 Ezaugarrien hautaketa automatikoa	69
4.3.1 Disfluentziaren segmentazio automatikoa (VAD)	69
4.3.2 Proba ez-parametrikoen bidezko ezaugarrien hautaketa automatikoa	69
4.3.3 Ezaugarrien bigarren hautaketa automatikoa	70
4.4 Sistemaren eredutzea: ezagutza automatikorako ereduak	70
4.4.1 Machine Learning paradigmak	70
4.4.2 CNN-ak.....	70
4.4.2.1 Deskripzioa	71
4.4.2.2 Multilayer Perceptron	71
4.4.2.3 Deep Learning: Neurona-sare Konboluzionalak (CNN)	73
4.4.2.3.1 Sare konboluzionalen arkitektura	73
4.4.2.3.2 Geruza konboluzionala	73
4.4.2.3.3 Pooling geruza	76
4.4.2.3.4 Guztiz konektatutako geruza.....	77

4.4.2.3.5 Aktibazio-funtzioak.....	78
4.4.2.3.6 Dropout	79
4.4.2.3.7 Eredua entrenatu	81
4.4.2.3.8 CNN ahotsa analisirako	82
4.5 Sistemaren balioztatzea	83
4.6 Sistemaren konfigurazioaren hautaketa automatikoa.....	84
5. KAPITULUA: ALZUMERIC	85
5.1 Sarrera	85
5.2 Alzumeric.....	87
5.2.1 Arkitektura orokorra.....	89
5.2.1.1 NFC (Near Field Communication) txartela	90
5.2.1.2 Smartphone edo tablet-a	91
5.2.1.3 App-a	91
5.2.1.4 Segurtasuna eta pribatutasuna	93
5.2.1.5 Bakarleri kuantikoa.....	93
5.2.1.6 HTTPS.....	96
5.2.1.7 Webservices	97
5.2.1.8 API.....	97
5.2.1.9 LAMP.....	97
5.2.1.10 MATLAB	98
5.2.1.11 WEKA	98
5.2.1.12 Python	98
5.2.2 Erabiltzailearen interfazea.....	100
6. KAPITULUA: EMAITZAK ETA EZTABAIDA	103
6.1 Sarrera	103
6.2 Erantzun emozionalaren azterketa	105
6.2.1. Materialak eta metodoak.....	105
6.2.2. Esperimentazioa	107
6.2.2.1 Leiho-tamaina aukeratzea.....	107
6.2.2.2 Dimentsio Fraktalaren Analisia.....	107
6.2.2.3 Tenperatura Emozionala	108
6.3 Hizkuntzarekiko independentea den bat-bateko hizketaren analisia.....	111
6.3.1 Materialak eta metodoak.....	111

6.3.2	Hasierako esperimentua.....	114
6.3.3	Bat-bateko hizketaren analisia	116
6.3.3.1	Dimentsio fraktalaren analisia.....	117
6.3.3.2	Ezaugarri-hautatze automatikoa	118
6.3.3.3	Sailkatze Automatikoaren ataza.....	118
6.3.4	Bat-bateko hizketaren azterketa integrala.....	123
6.3.4.1	Materialak eta metodoak.....	123
6.3.4.2	Sailkapen automatikoa	124
6.4	Ataza anitzerako hurbilpen ez-lineala MCI-rako	126
6.4.1	Materialak eta metodoak.....	126
6.4.2	Esperimentazioa	127
6.4.2.1	Ezaugarrien hautaketa automatikoa	127
6.4.2.2	Sailkatze automatikoa	130
6.5	MCI-ren detekzio goiztiarrerako aurrerapenak.....	131
6.5.1	Materialak eta metodoak.....	131
6.5.2	Esperimentazioa	133
6.6	Denbora errealerako sistema	133
6.6.1	Materialak eta metodoak.....	133
6.6.2	Esperimentazioa	133
6.7	Erabilgarritasunaren ebaluaketa	139
7.	KAPITULUA: ONDORIOAK	141
7.1	Ondorioak	141
7.1.1	Ondorio orokorrak.....	141
7.1.2	Ondorio zehatzak.....	143
7.1.3	Etorkizunerako lerroak	146
	BIBLIOGRAFIA	149
	ERANSKINAK	163

Irudien aurkibidea

Irudia 2. 1: Atazetara mugatutako proba klinikoa.....	18
Irudia 2. 2: Atazari murriztutako test klinikoa.....	23
Irudia 2. 3: Bat-bateko hizketa	27
Irudia 2. 4: Ahozko analisi vs analisi automatikoa.....	31
Irudia 2. 5: ILSVRC-en errore-tasa baxuena lortu duten algoritmoen emaitzak.....	34
Irudia 2. 6: AlexNet.....	35
Irudia 2. 7: Deconvolutional Network	36
Irudia 2. 8: Geruzen irudikapen grafikoa.....	36
Irudia 2. 9: ResNet sarea	38
Irudia 2. 10: Eskualdeak erauzi, desitxuratu eta CNN-an sartu.....	39
Irudia 2. 11: CNN-en hitz deskriptiboak aurkitzeko gaitasuna.....	39
Irudia 2. 12: Irudia eraldatu eta aipagarria den eskualdea aukeratu.....	40
Irudia 3. 1: Ataza bakoitzaren ezaugarriak.....	41
Irudia 3. 2: Hizketa murriztuaren test klinikoa	44
Irudia 3. 3: Hizketa murriztuaren test klinikoa II.....	47
Irudia 3. 4: Hizketa murriztuaren etxeko testa	48
Irudia 4. 1: Sistema eta tresna automatikoak	53
Irudia 4. 2: Ahots-laginen erauzketa linealak eta ez-linealak.....	54
Irudia 4. 3: Fonazio-aparatua	55
Irudia 4. 4: Fonazio-aparatua II	55
Irudia 4. 5: Higuchi-ren dimentsio fraktala	61
Irudia 4. 6: Hitz-jario kategorikoaren (CVF) ataza	68
Irudia 4. 7: MLP.....	72
Irudia 4. 8: Irudien datu-matrizeak.....	72
Irudia 4. 9: Konboluzioa.....	74
Irudia 4. 10: CNN-en kernelak	74
Irudia 4. 11: Ezaugarri-mapa	75
Irudia 4. 12: Ezaugarri-mapa II	75
Irudia 4. 13: CNN-en sakontasuna.....	76
Irudia 4. 14: Pooling-a CNN-etan.....	77
Irudia 4. 15: CNN-en egitura orokorra.....	78
Irudia 4. 16: ReLU	79
Irudia 4. 17: Dropout	80
Irudia 4. 18: Overfitting	80
Irudia 4. 19: Overshooting.....	81
Irudia 4. 20: Datu gordinak CNN-ra	82
Irudia 4. 21: Datu aberastuak CNN-ra	83

Irudia 5. 1: Teknika automatiko ez-inbasiboaren prozedura	88
Irudia 5. 2: Alzumeric-en arkitektura orokorra	89
Irudia 5. 3: Sistema automatikoaren front-end eta back-end-a	90
Irudia 5. 4: Smartphone-aren datuak jaso.....	92
Irudia 5. 5: Bakarlari kuantikoa	94
Irudia 5. 6: Bakarlari kuantikoa II.....	94
Irudia 5. 7: Bakarlari kuantikoa III.....	95
Irudia 5. 8: Zifratu-deszifratu simetrikoko eta asimetrikoko	996
Irudia 5. 9: Alzumeric-en monitorizazioa	99
Irudia 5. 10: Alzumeric interfazea	100
Irudia 5. 11: Alzumeric-ek laginak jositzeko modua.....	101
Irudia 5. 12: Ezaugarriak eta sailkapena	101
Irudia 6. 1: Errore Ratioaren Sailkatzea (%).....	106
Irudia 6. 2: Sailkapen Errore Ratioa (%)	107
Irudia 6. 3: Sailkapen Errore Ratioa (%) II.....	108
Irudia 6. 4: Errore Metagarria.....	110
Irudia 6. 5: Bat-bateko hizketa-eboluzioa	113
Irudia 6. 6: Hasierako esperimendazioaren ezaugarria	114
Irudia 6. 7: Ahostun eta ahoskabe (%) AZTITXIKI multzorako	114
Irudia 6. 8: Hizketa Seinalearen zatiak	115
Irudia 6. 9: Kontrol-pertsona baten analisia (CR)	116
Irudia 6. 10: AD duen pertsona baten analisia	117
Irudia 6. 11: Errore Ratioaren Sailkatzea (%).....	118
Irudia 6. 12: Errore Ratioaren Sailkatzea (%) II.....	119
Irudia 6. 13: Klase desberdinetarako Zehaztasuna/Doitasuna	121
Irudia 6. 14: Errore Metagarria.....	121
Irudia 6. 15: Kostu konputazionala	122
Irudia 6. 16: INTEGRAL probarako doitasuna	124
Irudia 6. 17: Sailkapenerako Errore-Tasa Metagarria (%)	125
Irudia 6. 18: Kaxa-diagrama hiru atazetarako	128
Irudia 6. 19: Kaxa-diagrama hiru atazetarako II	128
Irudia 6. 20: Kaxa-diagrama hiru atazetarako III	129
Irudia 6. 21: Ahoskabe segmentuen ezaugarri ez-linealak	129
Irudia 6. 22: Hiru atazen doitasun-tasak	130
Irudia 6. 23: Hiru atazetan lortutako klase bakoitzeko Accuracy-a.....	131
Irudia 6. 24: Ezaugarri bektorea	132
Irudia 6. 25: Acc (%) aukeratutako klase eta sailkatzaileetarako	133
Irudia 6. 26: HF (%) aukeratutako klase eta sailkatzaileetarako	135
Irudia 6. 27: CER (%) aukeratutako klase eta sailkatzaileetarako	136
Irudia 6. 28: Model Time MT (%) aukeratutako klase eta sailkatzaileetarako.....	137
Irudia 6. 29: FP-CR (%) aukeratutako klase eta sailkatzaileetarako	137

Irudia 6. 30: CER (%) hiru atazentzat (SS, PD eta AN).....	138
Irudia 6. 31: Sistemaren balioztatzea	139
Irudia 6. 32: Tipografia tamaina handitu.....	140

Taulen aurkibidea

Taula 2. 1: Dimentsio espazial urritu eta sakontasuna handitzeko CNN-a	37
Taula 3. 1: Erabilitako 3 ataza nagusiak	42
Taula 3. 2: Sortutako datu-base bakoitzaren datu demografikoak.....	43
Taula 3. 3: AN esperimentaziorako azpimultzoen datu demografikoak	46
Taula 3. 4: PD esperimentaziorako azpimultzoen datu demografikoak	46
Taula 3. 5: SS esperimentaziorako azpimultzoen datu demografikoak	49
Taula 3. 6: AZTITXIKI azpimultzoa.....	50
Taula 4. 1: Ezaugarrien inbentarioa.....	56
Taula 5. 1: Keyed-Hash Message Authentication Code	96
Taula 6. 1: Laburtzapenen deskripzioa	104
Taula 6. 2: Ezaugarri-multzoen laburtzapenak.....	104
Taula 6. 3: Accuracy (%).....	122
Taula 6. 4: Accuracy (%).....	123
Taula 6. 5: Sailkatzaileen konfigurazioak.....	134

Laburtzapenak

Laburtzapenak: Ingelesa, euskara

- ACW:** Adaptive Component Weighted Coefficients, *Ponderatutako Osagai Moldagarrien Koefizienteak*
- AD:** Alzheimer's Disease, *Alzheimerren Gaitza*
- AI:** Artificial Intelligence, *Adimen Artifiziala*
- AJAX:** Asynchronous Javascript and XML, *Javascript Asinkronoa eta XML*
- AN:** Animal Naming, *Animalien Izendatzea*
- API:** Application Programming Interface, *Aplikazioetarako Programazio Interfazea*
- APQ:** Amplitude Perturbation Quotient, *Amplitude Perturbazioaren Zatidura*
- AS:** Advanced Stage, *Fase Aurreratua*
- ASA:** Automatic Speech Analysis, *Hizketa Automatikoaren Analisia*
- ASSA:** Automatic Spontaneous Speech Analysis, *Bat-bateko Hizketaren Analisi Automatikoa*
- ASR:** Automatic Speech Recognition, *Hizketa Ezagutze Automatikoa*
- BNT:** Boston Naming Test, *Izendatze Boston Testa*
- CAR:** Classification Accuracy Rate, *Sailkapenaren Doitasun Tasa*
- CE:** Cumulative Error, *Errore Metagarria*
- CER:** Classification Error Rate, *Sailkapenerako Errore-Tasa*
- CERAD:** Consortium to Establish a Registry for Alzheimer's Disease, *Alzheimerren Gaitzaren Erregistroa Ezartzeko Partzuergoa*
- CF:** Classic Features, *Ezaugarri Klasikoak*
- CFD:** Castiglioni Fractal Dimension, *Castiglioni Dimentsio Fraktala*
- CI:** Confidence Interval, *Konfiantza Tarte*
- CL:** Convolution Layer, *Geruza Konboluzionala*
- CNN:** Convolutional Neural Network, *Neurona-sare Konboluzionalak*
- CR:** Control, *Kontrola*
- CSS:** Cascading Style Sheets, *Jauzi Estilo Orriak*
- CST:** Cryptographic And Security Testing, *Kriptografia eta Segurtasun Probak*
- CVF:** Categorical Verbal Fluency, *Hitz-jario Kategorikoa*
- CVFD:** Castiglioni Voice Fractal Dimension, *Ahoskatutako Castiglioni Dimentsio Fraktala*
- CVLT:** California Verbal Learning Test, *Kaliforniako Ahozko Ikaste Testak*
- DL:** Deep Learning, *Ikasketa Sakona*
- DN:** Deconvolutional Network, *Neurona-sare Dekonboluzionala*
- DNC:** Differentiable Neural Computer, *Konputazio Neurona Diferentziagarria*
- EF:** Emotional Features, *Emozio-Ezaugarriak*
- EM:** Experiment matrix, *Esperimentu-Matrizea*
- ERA:** Emotional Request Analysis, *Emozio Erantzunaren Analisiak*
- ES:** Early Stage, *Fase Arina*
- ET:** Emotional Temperature, *Emozio-Tenperatura*
- ETL:** Extraction, Transformation and Loading, *Erauztea, Eraldatzea eta Kargatzea*

EWM: Emotional Working Memory, *Iraupen Laburreko Oroimen Emozionala*

FC: Fully Connected, *Gutziz Konektatua*

FCSRT: Free and Cued Selective Reminding, *Oroimen Selektibo Aske eta Kodetuaren Testa*

FD: Fractal Dimension, *Dimentsio Fraktala*

FIPS: Federal Information Processing Standard, *Informazioaren Prozesatzerako Estandar Federalak*

FP: False Positive, *Positibo Faltsua*

GAN: Generative Adversarial Nets, *Kontrako Sareen Sortzailea*

HFD: Higuchi Fractal Dimension, *Higuchi Dimentsio Fraktala*

HMAC: Keyed-Hash Message Authentication Code, *Gako-Hash Mezu Bidezko Baimen Kodea*

HNR: Harmonic to Noise Ratio, *Armoniko-Zarata arteko Erlazioa*

HTML5: HyperText Markup Language 5, *Hipertesturako Lengoia Markatua 5*

HTTP: Hypertext Transfer Protocol, *Hipertesturako Transferentziarako Protokoloa*

HTTPS: Hypertext Transfer Protocol Secure, *Hipertesturako Transferentziarako Protokolo Segurua*

HVFD: Higuchi Voice Fractal Dimension, *Ahoskatutako Higuchi Dimentsio Fraktala*

HVLT: Hopkins Verbal Learning Test, *Hopkinsen Ahozko Ikaste Testak*

ICC: Inferior Colliculus Coefficients, *Beheko Kolikulu Koefizienteak*

ID: Idea Density, *Idea Dentsitatea*

IG: Information Gain, *Informazio Irabazia*

ILSVRC: ImageNet Large-Scale Visual Recognition Challenge, *ImageNet Eskala Handiko Ikuste-Ikuskapenaren Erronka*

IS: Intermediate Stage, *Fase Ertaina*

KFD: Katz Fractal Dimension, *Katz Dimentsio Fraktala*

KVFD: Katz Voice Fractal Dimension, *Ahoskatutako Katz Dimentsio Fraktala*

k-NN: k-Nearest Neighbors, *k-Bizilagun Hurbilak*

LF: Linear Features, *Ezaugarri Linealak*

LOOCV: Leave-One-Out Cross-Validation, *Bat Kanpoan Geratu Balioeste Gurutzatua*

LPCC: Linear Predictive Cepstral coefficients, *Aurreikuspenerako Koefiziente Lineal Cepstralak*

LPCT: Linear Predictive Cosine Transform coefficients, *Aurreikuspenerako Kosinu Koefiziente Eraldatze Lineala*

LTM: Long Term Memory, *Iraupen Luzeko Oroimena*

MCI: Mild Cognitive Impairment, *Narriadura Kognitibo Arina*

MFCC: Mel Frequency Cepstrum Coefficients, *Mel Eskalako Koefiziente Cepstralak*

MMSE: Mini Mental State Examination, *Buruko Egoeraren Mini Azterketa*

ML: Machine Learning, *Ikasketa Automatikoa*

MLP: Multi Layer Perceptron, *Multigeruza Pertzeptrona*

MMSE: Mini Mental State Examination, *Buruko Egoeraren Mini Azterketa*

MSC: Modulation Spectra coefficients, *Espekro Modulazioen Koefizienteak*

NFC: Near Field Communication, *Komunikazio Eremu Hurbila*

NNHL: Neuron Number in Hidden Layer, *Ezkutuko Geruzaren Neurona Kopurua*

NIST: *National Institute of Standards and Technology, Estandarren eta Teknologiaren Nazio Institutua*

NLF: Non-Linear Features, *Ezaugarri Ez-linealak*

NLP: Natural language processing, *Lengoaia Naturalaren Prozesatzea*

NHR: Noise to Harmonic Ratio, *Zarata-Armonikoaren arteko Erlazioa*

NSA: National Security Agency, *Nazio Segurtasunerako Agentzia*

OASIS: Organization for the Advancement of Structured Information Standards, *Informazio Estandarren Egituratu Aurrerapenerako Antolakundea*

OSI: *Open System Interconnection, Sistema Irekiaren Interkonezioa*

PAN: Personal Area Network, *Eremu Pertsonalerako Sarea*

PD: Picture Description, *Irudien Deskribapena*

PF: Perception Features, *Pertzepzio Ezaugarriak*

PGA: Gipuzkoa-Alzheimer Project, *Gipuzkoa-Alzheimer Proiektua*

PLP: Perceptual Linear Predictive coefficients, *Aurreikuspeneko Pertzepziozko Koefiziente Lineala*

RAVLT: Rey Auditory Verbal Learning Test , *Reyren Entzute-Hitzezko Ikastearen Testa*

RBNT: Revised version of the Boston Naming Test, *Boston Izendatze Test Eraberritua*

ReLU: Rectified linear unit, *Unitate lineal Arteztua*

REST: Representational State Transfer, *Irudikapenezko Estatu Transferentzia*

RMS: Root Mean Square, *Media Koadratikoa*

RNN: Recurrent Neural Networks, *Neurona-sare Atzerakariak*

RPC: Remote Procedure Call, *Urrutiko Prozedura Deia*

SaaS: Software as a Service, *Software-a Zerbitzu moduan*

SHA-256: Secure Hash Algorithm 256, *Hash Algoritmo Segurua 256*

SL: Supervised Learning, *Gainbegiratutako Ikasketa*

SMS: Short Message Service, *Mezu Laburretarako Zerbitzua*

SOAP: Simple Object Access Control, *Objektu Sinpleetarako Sarbiderako Kontrola*

SS: Spontaneous Speech, *Bat-bateko Hizketa*

SSF: Spontaneous Speech Features, *Bat-bateko Hitz-jarioaren Ezaugarriak*

SSL: Secure Sockets Layer, *Sockets Geruza Segurua*

SSSF: Selection of Spontaneous Speech features, *Bat-bateko hitz-jarioaren ezaugarrien hautaketa*

ST: Standard Deviation, *Desbiazio Estandarra*

STN: Spatial Transformer Networks, *Espazio Eraldaketarako Sareak*

SVFD: Selection of Voice Fractal Dimension, *Ahoskatutako Dimentsio Fraktalaren Hautaketa*

SVM: Support Vector Machine, *Euskarri Bektoreen Makina*

TLS: Transmission Layer Security, *Transmisorako Geruza Segurua*

TS: Training Step, *Entrenamendu Pausoa*

VAD: Voice Activity Detection, *Ahots Jarduera Hautematea*

VFD: Voice Fractal Dimension, *Ahoskatutako Dimentsio Fraktala*

VM: Validation Matrix, *Balidazio-Matrizea*

W3C: World Wide Web Consortium, *Munduko Sarerako Patzuergoa*
WM: Work Memory, *Laneko Oroimena*
WRAT-R: Wide Range Achievement Test-Revised, *Tarte zabalekin egindako Lorpenen Probaren Test Eraberritua*
WRAT4: Wide Range Achievement Test 4, *Tarte zabalekin egindako Lorpenen Probaren Testa 4*
WS: Window Size, *Leiho-Tamaina*
WS-I: Web Services Interoperability Organization, *Web Zerbitzuen Barne Eragiketarako Antolakundea*
XML: eXtensible Markup Language, *Lengoaia Markatu Hedagarria*
XSS: Cross-Site Scripting, *Guneen Arteko Scripting-a*
ZCR: Zero-Crossing Rate, *Zero-gurutatzeko abiadura*

1. KAPITULUA: SARRERA

1.1 Sarrera

2016. urteko munduko Alzheimer Txostenak nabarmentzen duenez, adineko populazioaren gehienak garapen-bidean dauden herrialdeetan bizi dira, eta bakarrik Txinan, 200 milioi pertsona har daitezke adineko populaziotzat. Herrialde horretan bakarrik, 2050. urtean, 60 urtetik gorako biztanleria 490 milioikoa izatea aurreikusten da [1]. Are gehiago, txostenak dioenez, bizi-itxaropenak luzatzen jarraitzen du eta heriotza-tasa etengabe ari da murrizten munduko eskualde gehienetan.

Ondorioz, jendea luzeago bizi da, eta dementziako nahasteak eta gaixotasun kronikoak gero eta prebalenteago bilakatzen ari dira bizimodu eta portaera ez osasuntsuak direla eta. Hori nabarmena da erdi-mailako eta maila altuko herrialdeetan eta munduko populazioa dementzia-izurriterako bidean dago, zeinak eragin lazgarria izan lezakeen epe luzean [1].

Egoera hau dela eta, Alzheimer gaitzak eragiten dituen defizit kognitiboen diagnostiko goiztiar, unibertsal eta merkeak ahalbidetuko dituen teknologia garatzeak berebiziko garrantzia dauka, gaixoen bizi-baldintzak hobetzeko bidean, gaitza hasierako etapan atzeman eta eutsi nahi badugu. AD-ren diagnostiko goiztiarra eta zehatza garrantzi handikoa da gaixoei eta beren familiek etorkizuna planifikatu ahal izateko, eta gaitzaren zantzuak tratatzeko aukerarik hoberenak bilatzeko.

1.1.1 Alzheimerren Gaitza (AD)

Alzheimerren Gaitzak (Alzheimer's Disease, AD), zahartzaroan dementziaren arrazoi nagusia denak [2][3], izugarriko kostu sozioekonomikoak ekartzen ditu, eta uste denez, goraka joango da etorkizun hurbilean [1][4]. Alzheimerren eragina handituz doa adinarekin baina 60 urte baino gutxiago dutenen artean ere prebalentea da.

Narriadura kognitibo progresiboa eta atzeraezina da AD-ren ezaugarria, oroimenaren galera eta hizketaren eta judizioaren nahasmenduak barnean hartuta, beste zenbait gabezia kognitibo eta portaerazkorekin batera. Hitzeko mintzaira endekatzeak berehalako eragina du pazientearen gaitasunean bere gizarte-

ingurunearekin modu naturalean elkar eragiteko. Askotan, horrekin batera, aldaketak gerta daitezke bere erantzun emozionaletan [5][6][7].

Defizit kognitiboek eta portaerazko sintomek, orokorrean, pertsona baten gaitasunak eta eguneroko jarduerak lanean, gizartean zein familian mugatzen dituzte. Gaitzak aurrera egin ahala, erion gaitasun eza handituz doa, erabateko mendetasunean egon arte [5][6].

Hortaz, ondoriozta daiteke, AD, gizartean eta ekonomian eragin handia sor dezakeen eta izurri bilakatu daitekeen prebalentzia altuko dementzia-mota arruntena dela, eta eriek, trebetasun eza larriak garatzen dituztela baita erabateko mendetasuna ere.

1.1.2 Narriadura Kognitibo Arina (MCI)

AD-ren lehen fasean Nahasmendu Kognitibo Arina (Mild Cognitive Impairment, MCI) delakoa agertzen da, akats kognitiboa alegia, baina ez da behar adina larria eguneroko bizitza edo gaitasun independenteak eragozteko.

Nahasmendu Kognitibo Arina (MCI) gaixotasun neurobegetatiboen hasierako agerraldietaraino hedatzen da, non pertsona aurreko errutina eta jarduerekin jarraitzeko gai den [8]. Orokorrean, MCI amnesikoa, AD goiztiar edo prodromaltzat hartzen da eta AD duten paziente askoren diagnostikoa MCI dutenean egiten dela kontuan hartu behar da.

1.2 Lanaren xede eta motibazioa

CEAFA-ren arabera, Espainiako estatuan, Alzheimerra 500.000 gaixori diagnostikatu diete. Hala ere, kopurua askoz handiagoa izan daiteke, eta 800.000ra iritsi. Gainera, biztanlerian zahartzea areagotzen ari denez, aurreikusiten da 2025. urterako Alzheimerrek kaltetuak 1.200.000 izan daitezkeela. 5.000.000, gaixoen familiak kontuan hartzen badira.

Egun, patologia honen kontra borrokatzeko merkaturatuak dauden aukera terapeutikoek, bai farmakologiko bai terapeutiko bigunek (test bidezko estimulazioa, jokoak eta abarrek) ez daukate gaixotasunak sortzen duen jatorriari erasotzeko mekanismorik, hori ezezaguna baita, eta zantzuak arintzeko balio dute soilik. Terapia horien bidez, kasu askotan, gaixotasunaren areagotzea moteltzen da, funtzio kognitiboa kitzikatzen bada.

Terapia berriak bilatzen ari diren arren Espainiako estatuan dauden gaixoen artean tratamendua bakarrik % 4-ak jasotzen du, nahiz eta tratamenduan egiten ari den inbertsioa itzela den.

Bestalde, Alzheimerra duen pertsona batek zainketa-ordu asko eskatzen ditu eta gehienetan lan eskerga hori senitartekoek bere egiten dute. Horrek ere asaldura anitzak sor ditzake senitartekoengan egunerokotasunaren higadurak sortuta.

Diagnostiko goiztiarrak eta ondorioz ezarritako tratamenduak gaixotasunaren areagotzea atzeratzea ahalbidetuko luke, eta ondorioz, bai gaixoen bai bere senitartekoen bizitza-kalitatea hobetzea izango litzateke ondoriorik argiena. Horrez gain, urteko ehunka milioi euro aurrezteka ekarriko luke.

MCI-k eragiten dituen aldaketa kognitiboak ordea ez dira gehienetan behar bezala diagnostikatzen, nahiz eta MCI-k Alzheimer edo beste demenziaren bat garatzeko arriskua areagotu dezakeen.

Batzuetan, irristakorra izaten da, batik bat prodromo-fasean, gaitz kognitibo adierazgarriak identifikatzea. Zaila gertatzen zaie batez ere espezialistak ez direneei eta senideei AD goiztiarra edo Narriadura Kognitibo Arina identifikatzea [9][2].

Beraz, gaur egun, garaiz detektatzea da espezialista medikuen erronketako bat [2]. Gainbehera kognitiboa hasierako mailan identifikatzeak berebiziko garrantzia dauka, sintoma kognitiboen hautemate goiztiarra onuragarria izan baitaiteke kudeaketa hobe baterako eta lagungarri izan daitekeelako gaixoengan, senideengan eta gizartean dituen eraginak gutxiagotzeko.

Aipatu behar da, zenbait medikuntza-arloko AD-ren diagnostikoa egiten dutela baina behin betiko diagnostikoak AD-ren aldaketa patologikoak autopsian frogatzea eskatzen duela [5][6][7].

Gaitz honek gizartean sortzen duen asaldura apaltzeko bidean MCI-ren diagnostiko goiztiar, unibertsal, merke eta ez-inbasiborako ekarpena egitea da erronka nagusia. Lan honetan mahaigaineratzen den aportunak eginkizun hori du xede, eta interes handikoa da, bizi dugun bilakaera soziodemografikoa eta AD-ren eragin sozial eta ekonomikoa dela eta. Hirugarren adineko pertsonen bizi-baldintzak hobetzea Eusko Jaurlaritzaren zientzia-politiken planaren lerro nagusi bat da.

1.3 Helburuak

1.3.1 Helburu nagusia

Lanaren helburu orokorra Alzheimerrak eta beste gaitz batzuek dakarzkiten asaldura neurologikoen diagnostiko goiztiarren hobekuntzan laguntzea da, hori guztia, hizketaren azterketa automatiko ez-inbasibo, unibertsal eta merkeen bidez eta osasun-guneen irizpide etikoen baitan.

Helburu hori lortzeko, hastapen-bertsioan izango den nahasmendu neuroendekapenezko antzemate goiztiarrean laguntzeko teknologia automatiko ez-inbasiboak garatu nahi dira.

Horrela, diagnostiko goiztiarrerako teknologia ez-inbasibo berri hauen garapenak, gaixoen inguruneari fidagarriak diren seinaleak jaso eta medikuei helarazteko aukera emango die.

1.3.2 Helburu orokorrak

- Narriadura Kognitibo Arina edo Mild Cognitive Impairment (MCI) prodromo-etapara bideratuta, dementziaren antzemate goiztiarrerako teknikak eta metodologiak garatzea.
- Hizkuntzarekiko independenteak diren parametroen azterketa automatikorako teknika eta metodologiak garatzea.
- Inguru erreal eta konplexuetara bideratutako teknika eta metodologiak garatzea.
- Erabilgarritasunaren estandarretan oinarria duten web-sistema integral merke, unibertsal eta ez-inbasiboaren garapena egitea.
- Ikasketa sakonean (DL) oinarria duten Machine Learning (ML) algoritmoak ikertu eta esperimintatzea.
- Datuen Erauzketa, Eraldaketa eta Karga (ETL: Extraction, Transformation and Loading) prozesu osoa bilduko duen ekosistema sortzea.

1.3.3 Helburu zehatzak

- Bioseinale-baliabideak sortzea osasun-guneekin lankidetzan eta horien irizpide etikoen baitan.
- Medikuntzaren ebaluaketa sendoari erreparatuz, medikuntza-laginen diseinu-metodologia berri eta normalizazioari buruz ikertzea.
- Medikuntza-hautaketan sostengatutako hautaketa automatikoen teknikak ikertzea.
- Seinalean izan daitezkeen patologien aldaketa azkarren hauteman goiztiarrerako ebaluaketa-erakusle berriak ikertzea.

- Hizkeraren disfluentzia, hots, hizketa jariakor batean gertatzen den edozein etendura, irregulartasun edo elementu ez-lexikoren azterketa automatikoa egitea.
- Disfluentzia ez ezik, erantzun emozionalaren azterketa automatikoa egitea.
- Komunikazio seguruetera bideratua dagoen bezero-zerbitzari informazio-transakzio zifratuak garatzea.
- ML algoritmo klasikoek eta ikasketa sakonean oinarritutako algoritmoen emaitzen arteko konparazio eta interpretazioa egitea.

1.3.4 Helburu gehigarriak

- Alzheimerra eta antzeko beste gaixotasunen egoera goiztiar zein Narriadura Kognitibo Arinak ikertzea.
- Inguru konplexuetara bideratutako kalitate desberdinetako ahots-seinaleak erabiltzea.
- Ezaugarrien hautaketa automatiko optimizatu batetarako bideak ikertzea.
- Parametro lineal, ez-linealetan oinarritutako metodologia eta parametro bioinspiratuak ikertzea. Hau da, inguru konplexuen edo eta zatatsuen sendotasuna lortzera bideratuak dauden parametroetan sakontzea.
- Hizkuntzarekiko independenteak diren parametroak aztertzea.
- Diagnostiko-emaitzak hobetzeko DL-en oinarria duten teknikak ikertzea.
- Isilpeko komunikazioak lortzeko algoritmo zifratuak ikertzea.
- Teknologia desberdinak plataforma bakarrean biltzea eta horien arteko komunikazio-protokoloak estandarizatzea.

1.4 Hipotesiak

- Diagnostikorako ez-inbasiboak diren teknika adimenduen erabilera oso baliagarria da dementzien hautemate goiztiarrerako.
- Gaixoen hizketan garunaren dementzia-zeinuen nabaritasun zientifikoa badago.
- Hizkeraren osagai linguistikoen ikerketak egun dauden hizketa zein biomekanikarekiko independenteak diren elementuetan oinarria duten azterketen metodologia hobeto dezake.
- Zenbait atazaren bidez, hizketatik informazio oso garrantzitsua atera daiteke diagnostiko eta eboluzioaren jarraipena egiteko.
- Hizkeraren disfluentzia AD-ren prodromo-fasearen adierazle izan daiteke eta narriadura kognitiboaren zeinu argia izan daiteke.
- Hasierako faseetako andeatzea dela bide, erantzun emozionala ere kaltetu daiteke, giza gogoaren arkitekturarekin zerikusia dituzten prozesu kognitiboak baitira.
- Aurrez aukeratu eta aberastutako datuek Deep Learning sareei ezaugarri-mapak egiteko lana errazten diete.

- Ikasketa sakonak orain arteko Machine Learning algoritmoen diagnostiko-emaitzak hobetzeko gaitasuna izan dezake.
- Prozesu guztia web-teknologiarekin integratzeak eta hizkuntzarekiko independenteak diren parametroekin lan egiteak sistema asko merkeagotu eta bere erabilera unibertsalizatu dezake.

1.5 Metodologiaren alderdi orokorrak

Honako ikerketa-lanean, abiapuntua, artearen egoerari buruz egindako berrikuste sakonak finkatzen du. Ondoren, hainbat lerro metodologikori ekiten zaio ikerketa-lanak dituen fronte desberdinen garapenari aurre egiteko:

- Aztertu nahi den problematikaren ikerketa eta prozesu guztia, hots, datuen Erauzketa, Eraldaketa eta Karga (ETL), bilduko duen sistemaren diseinua:
 - Datuen optimizaziorako metodoen garapena
 - Baliabide gabeziagatiko problematikari irtenbidea emateko modelatua
 - Ingurune konplexuetara bideratutako parametroen moldaketa
 - Sailkapen-emaitzak hobetuko dituen ML algoritmoen erabilera
 - Erabilgarritasuna izpide eta prozesu guztia kontuan hartuko duten lanabes integralen garapena
- Aztertu nahi den ingurune konplexurako baliabideen sorrera:
 - Atazen aukeraketa, grabazio bidezko informazioaren erauzketa eta datu-basearen sorrera
 - Hizketa-ereduen aukeraketa
 - Hizkuntzarekiko independenteak diren ereduen modelatzea
 - Hizketa-ereduen ezaugarri lineal zein ez-linealen hautaketa eta erauzketa automatikoa
 - Atazen ezaugarri aipagarrienen hautaketa automatikoa
 - Ikasketa automatikorako algoritmoen garapena
 - Sistemaren komunikazio zifratuetarako script-en kodetzea
- Ingurune konplexuetara bideratutako teknika berrien modelatzea
- Disfluentzia eta erantzun emozionalak aztertzeke teknika eta metodologien erabilera:
 - Ezaugarri ez-linealen modelatzea
 - Ezaugarri aipagarrienen hautaketarako tekniken (Mann-Whitney U-test ebaluazio ez-parametrikiko eta Support Vector Machine-n ezaugarri-ebaluazioaren bidez) erabilera
 - CNN algoritmoen txertaketa
 - Sistemaren errendimenduari begirako optimizazio-tekniken aplikazioa
- Ebaluaziorako metodologia klasikoen erabilera
- Lanaren garapenean zehar oinarrizko irizpide metodologikoak erabiliko dira:
 - Egungo testuinguruan sistema honen garapenak duen garrantzia

- kontuan hartzea
- Erabilgarritasuna eta unibertsaltasuna izpide izango duen sistemaren inplementazioa burutzea
- Hizkuntzarekiko independentea den sistema eraikitzea
- Erabilitzaile errearen parte-hartze aktiboa bilatzea

1.6 Egitura

1.6.1 Kapituluak

1. KAPITULUA: SARRERA

Kapitulu hau tesiaren memoriaren abiapuntua da, eta bertan, aztertu nahi den problematikaren azalpena, lanaren motibazioa, xede eta helburuak, lan-hipotesiak, metodologiaren alderdi orokorrak eta memoriaren egiturari buruz hitz egiten da.

2. KAPITULUA: ARTEAREN EGOERA

Kapitulu honetan, artearen egoeraren azalpen sakona egin eta ikertuko den problematika hobeto ulertzeko 4 atal nagusitan banatzen da: AD gaitzak giza komunikazioan eragiten dituen arazoak, MCI/AD diagnostikatzeko egun dauden eskuko metodologiak, teknika automatiko aurreratuenak eta Deep Learning algoritmoen azken bost urteetako garapena.

3. KAPITULUA: MATERIALAK

Sistema egoki modelatzeko lehendabizi datuak behar dira. Horretarako hiru atazek (AN, PD eta SS) osatutako datu-baseak sortu dira. Ataza horien bidez lortutako informazioaren deskripzioa egingo da kapitulu honetan. Atal hau sakonago ulertzeko, B.1 eta B.3 eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

4. KAPITULUA: METODOAK

Erabilitako metodologiari buruz hitz egiten da kapitulu honetan. Hasieran, ezaugarri lineal eta ez-linealen hautaketa eta horien azalpenari buruzko sarrera egiten da. Ondoren horien erauzketa eta ezaugarri aipagarrienak aukeratzeko metodologiari buruz hitz egiten da. Bukaeran, emaitzak sailkatzeko ikasketa sakonaren teknikak nabarmentzen ditu. Atal hau sakonago ulertzeko, B.1, B.2, B.3, C.1, C.2, C.3, C.4, D.1, D.2 eta D.3 eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

5. KAPITULUA: ALZUMERIC

Mediku espezialistak zein familiak ahots-laginak metodo ez-inbasiboen bidez hartzeko eta horien sailkapen binomiala (MCI/AD edo CR) egiteko sortutako web-

1. KAPITULUA: SARRERA

plataformaren deskripzioa egiten da kapitulu honetan. Atal hau sakonago ulertzeko, B.1, B.2, B.3, C.3, C.4 eta D.2 eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

Laborategi-ingurunea ardatz, esperimentazioak modelatze klasikoa du oinarri eta aurreko kapituluetan garatutako metodologiaren integrazioaren bidez lortutako emaitzen ebaluazioa egiten du. Atal hau sakonago ulertzeko, B.1, B.2, B.3, C.1, C.2, C.3, C.4, D.1, D.2 eta D.3 eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

7. KAPITULUA: ONDORIOAK

Kapitulu honek, tesiaren garapenaren ondorioz eratorritako ondorio eta etorkizunerako ikerketa-lerro berriei buruzko deskribapena egiten du.

1.6.2 Tesiak sorturiko argitarapenak

A ERANSKINA: SISTEMA ERREALAREN ANALISIAK HIRU ATAZETARAKO (SS, PD, AN)

B ERANSKINA: JCR ARGITARAPENAK

1. K. Lopez-de-Ipiña, J.B. Alonso, C.M. Travieso, J. Solé-Casals , H. Egiraun, M. Faundez-Zanuy, A. Ezeiza, N. Barroso, M. Ecay , P. Martinez-Lage, U. Martinez de Lizardui, On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis, *SENSORS*, 13(5), 6730 -6745, 2013, DOI: 10.3390/s130506730, ISSN: 1424-8220 IF ISI JCR (2013) Q1: 2.245, Instruments and Instrumentation; SNIP: 1.818, SJR: 0.66; 5-Year Impact Factor: 2.457 (2013), (WoK), Scopus: dblp, Google Scholar, kanpoko zitak: 6.
2. K. López de Ipiña, J. Sole-Casals, M. Faundez-Zanuy, P. Calvo, E. Sesa, U. Martinez-de-Lizarduy, P. De-la-Riva, J. F. Marti-Masso, B. Beitia, A. Bergareche, Selection of Entropy Based Features for Automatic Analysis of Essential Tremor, *Entropy*, DOI: 10.1007/s12559-012-9169-9, ISSN, 1866-9956, IF ISI JCR: 1.440, Q2 Computer Science Artificial Intelligence (681123); SNIP: 2.035; SJR: 0.53; 5 Year, impact factor: 1.359, (WoS), Scopus: dblp, Google Scholar.
3. K. López-de-Ipiña, U. Martinez-de-Lizarduy, P. M. Calvo, J. Mekyska, B. Beitia, N. Barroso, P. Gómez-Vilda, A. Estanga, M. Tainta and M. Ecay-Torres, Advances on automatic speech analysis for early detection of Alzheimer Disease: a non-linear multi-task approach, *Current Alzheimer Research*, Onartua. DOI: 1.1007/s12559-012-9169-9, ISSN, 1866-9956, IF ISI JCR: 3.12, Q2 Clinical Neurology, Scopus: dblp, Google Scholar. 1-1-22, 2017, Bentham Science Publishers, USA (accepted).

C ERANSKINA: ARGITARAPEN INDEZATUAK

1. K. Lopez-de-Ipiña, H. Eguraun, J. Sole-Casals, M. Ecay, A. Ezeiza, N. Barroso, P. Martinez-Lage, U. Martinez-de-Lizardui, Feature Extraction Approach Based on Fractal Dimension for Spontaneous Speech Modelling Oriented to Alzheimer Disease Diagnosis, *Advance in Non-linear Speech Processing, Lecture Notes in Computer Science*, Springer, Verlag, Berlin (Alemania) 10.1007/978-3-642-38847-7_19, ISBN 978-3-642-38846-0, ISSN: 0302-9743, ISI Web of Knowledge (WoK), Scopus, SJR (SCImago Journal Rankings) 2012: 0.332, SNIP (Source Normalized Impact per Paper) 2012: 1.190, dblp, Google Scholar, 7911:144-151, 2013.
2. K. López-de-Ipiña, U. Martínez-de-Lizarduy, M. Ecay-Torres, P. Martínez-Lage, F. Torres, N. Barroso, M. Zaundez-Zanuy, Automatic analysis of Categorical Verbal Fluency for Mild Cognitive Impairment detection: a non-linear language independent approach, 4th IEEE IWOB, Proceedings de IEEE, Donostia-San Sebastian, (Spain) 2015, Xplore, ISI WoK.
3. K. López-de-Ipiña, U. Martínez-de-Lizarduy, P. M. Calvo, B. Beitia, J. García-Melero, Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach, Proceedings de IEEE, Xplore, 5th IEEE IWOB, Funchall, (Madeira, Portugal) 2017, ISI WoK (accepted).
4. K López-de-Ipiña, J Solé-Casals, U Martínez de Lizarduy, PM Calvo, J Iradi, M Faundez-Zanuy, A Bergareche, Non-invasive Biosignal Analysis Oriented to Early Diagnosis and Monitoring of Cognitive Impairments, *Converging Clinical and Engineering Research on Neurorehabilitation II*, 867-872, Springer International Publishing.

D. Eranskina: Bestelako argitarapenak

1. K. Lopez-de-Ipiña, J.B. Alonso, J. Solé-Casals, N.Barroso, P.Henriquez, M. Faundez-Zanuy, C. Travieso, M. Ecay-Torres, P.Martinez-Lage, U. Martinez-de-Lizardui, H. Eguraun, A. Ezeiza, Analysis of Spontaneous Speech and Emotional Response oriented to Alzheimer's Disease Diagnosis, *I Jornadas Interdisciplinarias de Voz, Habla y Canto 2013*, Proceedings, Las Palmas de Gran Canaria (Spain) 2013.
2. U.Martinez de Lizarduy, K. López de Ipiña, P. Calvo, P. Gómez-Vilda, M. Ecay-Torres, ALZUMERIC: a decision support System for diagnosis and monitoring of cognitive impairment, *Loquens, CSIC*, 2017 (accepted).
3. Karmele López-de-Ipiña, Harkaitz Eguraun, Jordi Sole-Casals, Miriam Ecay, Aitzol Ezeiza, Nora Barroso, Pablo Martinez-Lage, Unai Martinez-de-Lizardui, Feature extraction approach based on Fractal Dimension for Spontaneous Speech modelling oriented to Alzheimer Disease diagnosis, *NOLISP June 2013*, Mons Belgica.

2. KAPITULUA: ARTEAREN EGOERA

2.1 Sarrera

Dudarik gabe, sintoma kognitiboen hautemate goiztiarra onuragarria da. Horren ondorioz, gaitzaren kudeaketa hobetze aldera, beharrezkoa da sintomak hautemateko bitartekoak merketzea eta horien erabilera orokortzea.

Testuinguru honetan, egun, medikuntza-azterketa zenbaiten (adib. FMRI, PET) baldintza teknologikoak eta kostuak direla kausa, ezinezkoa da markatzaile horiek gaixo guztiei aplikatzea eta, gainera, prozedura inbasiboak eskatzen dituzte (adib. ziztako lunbarra).

Horrez gain, gaurko irizpide eta teknikekin, demenzia agertzen denean eta gainerako kausa posibleak baztertzen direnean, AD-ren diagnostikoa zenbait segurtasun-mailatan adierazten da (litekeena da izatea...). AD-ren behin betiko diagnostikoa egiteko, post-mortem azterketa ezinbestekoa da, burmuinean izan diren aldaketa patologikoen analisia alegia [5]. AD-ren diagnostikoa klinikoa da eta, batetik, demenzia-sindrome progresiboaren berrespena eta, bestetik, osasun-txostena dela bide, gainerako demenzia-motak baztertzea eskatzen ditu.

Diagnostikoa burutzeko honako teknikak erabil daitezke besteak beste: osasun orokorra aztertzea, odol-analisi zorrotza, test neuropsikologiko sakonak, neuroirudietako teknika garestiak eta ziztada lunbarreko teknika inbasiboak [2].

Hala ere, hasierako faseetan diagnostikatzeak zaila izaten jarraitzen du. Izan ere, hala gaixoei nola senideek lehen agerpen klinikoak kontuan ez hartzeko/ezkutatzeko edota adinarekin zerikusia duten ohiko aldaketa kognitiboak leporatzeko [4] joera dute. Ondorioz, pazienteak askotan fase ertainetan edo aurreratuatan daudenean diagnostikatzen dira, eta horrek, sendabide-estrategiak eta horien eraginkortasuna zailtzea [4][6] dakar.

Horrez gain, gaitza aurrera joan ahala, sintoma horiek guztiek errendimendua gutxitzea ekartzen dute eguneroko jardueretan. Azken urteotan, diagnostiko goiztiarra egiteko aurrerapen nabarmenak egin dira, biomarkatzaile kliniko fidagarriak eta diagnostiko/screening edo bahetze-tresna sendoak garatzeari esker. Edonola ere,

markatzaileen baliagarritasuna gorabehera, ezinezkoa da test horiek oroimen-hutsegiteak agertzen dituzten gaixo guztiei edo populazioari orokorrean egitea, dakartzaten kostuak eta eskakizun teknologikoak direla eta. Hortaz, soil-soilik, lehen hautaketan hartutako gaixoei aplikatzen zaizkie, kostu urriko bahetzea egin eta gero AD patologia izan dezaketelakoan [2][87][19].

Testuinguru orokorra hau hobeto ulertzeko artearen egoeraren azalpen sakona egiten da eta 4 azpiataletan banatu da:

Alzheimerrak komunikazioan sortzen dituen defizitetatik hasten da **2.2 Komunikazio-defizitak AD-n** atalean. Giza komunikazioak eguneroko bizitza duen garrantzia eta aztertzen ari garen AD gaitzak eragiten dituen defizitak, batez ere ahozko komunikazioari eta erantzun emozionalari dagokienez.

2.3 Eskuko metodologiak atalean, eskuzko metodologiei buruz hitz egiten da. AD detektatzeko tresna kliniko klasikoei buruz sakontzen da, ahozko komunikazioak balioztatzeke ataza eta screening test klasikoei buruz.

Ondoren, **2.4 Teknika automatikoak** atalean, azken proposamen automatiko berriak azaltzen dira. Test kliniko sendoetatik hasi eta hizkuntzarekin loturarik ez duten metodologia aurreratuetaraino. Hizketa automatikoaren hautemate lineal eta ez-linealen inguruko gaurko tekniken egoerari buruz sakontzen da atal honetan.

Bukatzeke, eta besteak beste, ikerketa honetan erabiliko diren sailkapenerako Deep Learning (DL) algoritmoen inguruko aurrekariari buruz hitz egingo da, **2.5 Neurona-sare Konboluzionalak (CNN)** atalean. Azken bost urteetako ekarpenen azterketa egiten da. Azpiatala eskaini zaie, etorkizunari begira ikerketa-lerro berri oso interesgarriak irekitzen dituztelako.

Mintzamenaren analisi automatikoan izandako aurrekarien zerrenda egiteko, datu-base garrantzitsuenak hartu dira erreferentziatzeko iturritzat, hala nola: PubMed, Scopus edota ISI Web of Knowledge.

2.2 Komunikazio-defizitak AD-n

Komunikazioa ezinbestekoa da gizakiok garatu eta bizirik irauteko. Gizakiok elkar eragiten dugu ahozkoa eta ahozkoa ez den osagaiez osatutako prozesu baten bidez. Ahozkoa ez den informazio horrek (gorputz-hizkuntza, jarrerak, ahots-modulazioak, aurpegi-adierazpenak barne) zerikusi handia du prozesu horrekin eta zehatzago hizketaren ulergarritasunarekin [10][11]. Hala ere, ahozko komunikazioa da, dudarik gabe, gure eboluzioa beste espeziekiko hoberen definitzen duen gaitasuna, eta eguneroko bizitzan gizarte-elkarrekintza gidatzen duena.

Hori dela eta, ahozko hizkuntza da pertsona baten adimena, gizarte-bizitza eta nortasuna definitzeko elementurik garrantzitsuenetako bat. Haren bidez, komunika gaitzke, baina aldi berean gure ezagutzak parteka ditzakegu eta gure nortasun pertsonala eta kulturala adierazi. Hortaz, jendearen arteko harremanetarako modurik intuitiboena eta eraginkorra da. Gai gara mezuek dakartzaten hotsak ekoizteko, gure desirak, ideiak edota emozioak eta gure izatasunaren zatia ezkututzen dituen bioseinale direnak. Oso prozesu konplexua da, gaitasun kognitiboen multzo zabala berarekin dakarrena. Prozesu horretan, burmuinak ezagutza, oroimena, hizkuntza eta informazio semantikoa kudeatu behar ditu. Aurreko ataletan aipatu bezala, lagin biologiko horiek oso erraz erregistra daitezke tresneria inbasibo eta garestirik gabe.

2.2.1 Alzheimerren Gaitza (AD)

Gaitza aurrera joan ahala, giza komunikazioarekin zerikusi zuzena duten zortzi arlo kaltetzen dira [12][13]: Oroimena, hizketa, pertzeptzioa, arreta, eraikuntza-gaitasuna, aholkatze-gaitasuna, problemak ebazteko gaitasunak eta gaitasun funtzionalak.

Beraz, oroimena galtzearen ondoren, AD-ren arazo handienetako bat mintzamen-gaitasunaren galtzea da. Galera horren isla besteekin hitz egiteko eta ulertzeko zailtasunetan antzematen da. Ondorioz, inguruarekin eta gizarte-interakzioarekin komunikazio naturala zailago bilakatzen da, eta horrek, AD dutenen baztertzea eragiten du eta oso eragin ezezkorra du hala eriengan nola horien senideengan [14].

Komunikazio-defizit desberdinak ager daitezke mintzamen-gunean [15][16], besteak beste: afasia, hitz egin eta ulertzeko zailtasuna, afasia anomikoa, gauzak ezagutu eta izendatzeko zailtasuna. Gainera, pazientearen komunikazio-arazo zehatzek gaitzaren fasearekin lotura estua dute [6][8][15][16]:

- Hasierako fasean edo fase arinean (Early Stage, ES), gaixoari zaila egiten zaio hitz egokiak aurkitzea bat-bateko hizketan, eta hori, askotan, ez da detektatzen.
- Bigarren fasean edo fase ertainean (Intermediate Stage, IS) hizketa eta hiztegia pobretu egiten dira eguneroko jardunean.
- Hirugarren fasean edo fase aurreratuan edo larrian (Advanced Stage, AS) erien erantzunak oso hitz guxitara murrizten dira.

2.2.2 Narriadura Kognitibo Arina (MCI)

Nahasmendu Kognitibo Arinari dagokionez, AD-an bizi diren aldaketa horiek sotilak izan daitezke eta, hortaz, gehienetan ez dira detektatzen. Mintzamen-zailtasunak, ordea, komunikatzeko ezintasun hori, gaitza honen hasierako faseetan agertzen da.

Hasierako sintomak batik bat, mintzamen-zailtasunaz gain, hurbileko gertaeren ahazte progresiboa eta ikasteko gaitasuna endekatzea dira, baina badira beste zenbait

seinale adierazgarri batzuk: kognizio-gune batzuen nahasmendua, batez ere hizketa-nahasmendua [17] eta oroimen-narriadura da, dudarik gabe, hasierako sintomarik nabariena.

Azken hamarkadotan, hizketan, mintzamenean eta portaeran izandako zenbait aldaketa jaso dira. Ikerketa askok agerian jarri dute komunikazio-defizita nabarmena eta progresiboa dela, eta gaitzaren fase guztietan nabarmentzen dela. Ulermena eta ahoko diskurtsoan, elkarrizketan, irakurketan eta idazketan [17][18] eragiten duela nabarmendu dute.

2.2.3 Mintzamenean defizita

Mintzamen-defizitak AD duten gaixo askorengan agertzen dira, eta seinale adierazgarritzat jo daitezke hasierako fasean. Azken hamarkadotan, dokumentatu eta berretsi ahal izan da hizketaren disfuntzioa Alzheimer gaitzaren garapenean goiz azal daitekeen sintoma bat dela, eta biomarkatzaile erabilgarria izan daitekeela prebentzio eta detekziorako [2][19][14][20][21]. Lan horietan lantzen diren defizitak honako hauek dira: adierazteko eta jasotzeko trebetasunak, izendatu eta ulertzeko akatsak, maila fonologikoan [20] zein sintaktikoan [22] artikulazio mugatua, zalantza semantiko edota diskurtsiboak, disfluentziak edo hitz-jarioaren narriadurak [21][23].

Izendatzea eta hitz-jarioa beste zenbait alderdi baino lehenago kaltetzen dira, eta dokumentatu denez, jario semantikoa (kategoriak izendatzea, adibidez, animaliak) kaltetuago dago jario fonemikoa baino (letra batez hasten diren hitzak izendatzea) [6][24][25][26][27]. Hortaz, izendatzeko gaitasuna lehen mailako korrelatzailea da hala identifikazio goiztiarrerako nola AD-ren kontrolerako [28][29].

Anomia da gaitzaren aurreko fasean garatzen hasten den hizkuntza-sintomarik nabariena [30]. Ikerketa-sorta zabalak frogatu egiten du defizit hori; alabaina, izendatzeko akatsen tipologiari eta gaitzaren eboluzioan defizit horrek izaten duen eboluzioari buruz eztabaida handiak daude oraindik ere [31][32].

Egile batzuen ustez, defizit horrek oroimen semantikorekin endekatzearekin zerikusia du [31]; beste batzuen iritziz, ordea, arazoa hitzaren forma fonologikora iristeko hutsegiteetan datza [33][34] edota, areago, beste azpisistemetan dauden arazoei leporatzen diete, eta burmuin osoaren atrofiarekin antz handia ez ote duen planteatzen dute [35].

Beste ikertzaile batzuek honako alderdi hauei begira azterzen dituzte defizitok: mintzamen-komplexutasuna (ataza mugatuak, hizketa literala, imitazio-hizketa, parafrase literalak, izenordeak, neologismoak, anbiguotasuna erabakitzea, hizkuntzen ulermena edo bat-bateko hizketa) edo eleaniztasun-gaitasunak [19][36].

Ikerketa berriek, MCI-n, defizit horiek berresten dituzte [37][38], mintzamen-defiziten eta endekatze kognitibo progresiboen [22] arteko konexioak aztertzen dituzte eta eragile fisiologikoetan oinarritutako analisi berriak proposatzen dituzte [39].

2.2.4 Emozioan eragina

Alzheimerdun gaixoengan, mintzamina ez ezik, erantzun emozionala ere kaltetu egiten da, eta, ematen duenez, zenbait fase ditu [40]. Emozioak giza gogoaren arkitekturarekin zerikusia dituzten prozesu kognitiboak dira eta honako prozesu hauekin lotuta daude: erabakiak hartzea eta oroimenaren kudeaketa edota arreta, zeinak hertsiki lotuta dauden komunikazioarekin, ikasi eta ulertzearekin [41][42].

Bestalde, inguruneak, pertzepzio-sistemak eta inguru indibidualen esperientziatik sortutako emozio-oroimenak eragin zuzena dute giza emozioetan [43][44]. Gainera, hasierako faseetako andeatzea dela bide, gizarte-desinhibizioa eta portaeraaldaketak gertatzen dira [45][46]. Hala ere, oroimen emozionalak mantendu egiten dira.

Fase aurreratuagoetan, gaixoak lotsati eta apatiko azal daitezke; sintoma horiek oroimen-galtzeari edota komunikatzeko zailtasunei eta pertzepzio aldatuari egotzen zaizkio. Horrez gain, baliteke emozioak sentitzeko gaitasun-urritze horrek oroimen-galerarekin zerikusia izatea eta horrek apatia eta depresioaren itxura eragitea [47][46].

Iraupen laburreko oroimen emozionalaren (Emotional Working Memory, EWM) eta AD-aren arteko harremanean arreta jartzen duten beste lan berri batzuek azpimarratzen dutenez, EWM osorik izatea lagungarri izan daiteke jarduera orokorraren mailak, motibazioa eta beste gaitasun garrantzitsu zenbait indartzeko eta, ondorioz, AD dutenen bizitza-kalitatea hobetzeko. Hortaz, emozio-gaitasunen endekatzea honako alderdi hauetan nabari daiteke: mantenuan, besteekiko loturan, inhibizioan [48].

Aipatu bezala, erantzun eta kudeaketa emozionalak ere badira gaitzaren hasierako fase horietan nahasten diren beste komunikazio-elementu batzuk eta Mammarella *et al.*ek ondorioztatu zutenek, emozio handiko gertaerak hobeto gogoratzen dira gertaera neutroak baino. Hori bera agertzen da adineko pertsona osasuntsuengan [49][50], eta informazio emozionalaren kodetzea AD duten erien emaitzak ere tankerakoak dira [50]. Berez, zahartze osasuntsuari buruzko ikerlanek [51] agerian jarri dute informazio ezkorra beharrez, informazio baikorra gogoratzeko joera kontrol-prozesuen biltzean datza eta, ondorioz, estimuluek eragin zuzena izan dezakete oroimenean garrantzi emozionalaren garrantziaren arabera. Are gehiago, emozioen eta oroimen emozionalaren (Work Memory, WM) arteko interakzioa AD-n hobeto argitzeko, egileek ikerlan bat aurkeztu dute WM-ren funtzioei, mantentzeari, loturei eta inhibizioari dagokienez [52][53].

Ezagutza- eta komunikazio-defizit eta portaera-sintoma horiek guztiak oso faktore baliagarriak izan daitezke AD ebaluatzeko.

2.2.5 Disfluentzian eragina

Disfluentziak ere hizketa-elementu interesgarriak dira, eta oso baliagarriak izan daitezke MCI behar bezala detektatzeko. Hala isiluneek nola disfluentziek informazio baliagarria izan dezakete ahozkatutako mezua deskodetzeko. Hori dela eta, ikerlan berriek agerian jartzen dute bien garrantzia bat-bateko hizketan MCI eta AD identifikatzeko.

Iradokitzen denez [19], grabaketa-denbora motzetan AD duten gaixoek pertsona osasuntsuek baino ahalegin handiagoa egin behar dute hizketaldi bat burutzeko: AD-ren eriek polikiago hitz egiten dute geldialdi luzeagoekin, denbora gehiago behar dute hitz egokia aurkitzeko eta, ondorioz, hizketa-disfluentzia edo mezu hautsiak izaten dira. Hizketa-disfluentziatzat hartzen dira hizketa arinean gertatzen diren edozein etete, irregulartasun edota hitz ez-lexikal, hizketa hori hasi edo eten dezaketenak. Horien artean, aipa ditzakegu: hasiera faltsuak, behin eta berriro hasi edota errepikatzen diren perpausak, errepikatzen diren silabak, makulu-hitzak edota purrustadak eta hiztunek beraien *lapsus linguae*ak edo ahoskatze okerrak zuzentzea [54].

Elementu horiek helbururik edota esanahi formalik gabeko hizketaren zatitzat jotzen dira. Batzuetan, akatsak zuzentzeko edota ondorengo adierazpenak aurreratzeko erabiltzen dira. Maiztasun handiz agertzen dira hizketan, eta erregulariki errepikatzen dira eguneroko elkarrizketetan. Ikerketa batzuek nabarmendu dute bat-bateko hizketan izaten diren nahasteetan kontuan hartu behar direla jaioterriko hizkuntzen ereduak [55]. Elementu horiek guztiak hizketa hutsarekin batera eboluzio kognitiboaren isla dira ahozko komunikazio-prozesua bitartean, eta beren maiztasuna handiagoa da bat-bateko hizketan. Are gehiago, disfluentzia ez-patologikoez oso informazio aberatsa izan dezakete emoziozko egoeraz edota hizkuntza-planteamentuaz.

Disfluentzia horien handiagotzeak nahaste kognitiboaren isla nabarmena izan daitezke, eta batzuetan, AD-ren gaixoengan, barne-prozesu kognitiboaren hitzezko ahoskatzea bilakatu daitezke, barne-solasketa moduan: “Zer da hori”, Nola du izena?”, /um/, ez naiz gogoratzen”.

Disfluentziak eta isiluneak ugaritzeak gaitzaren okertzearen isla izan daitezke, eta komunikatzeko gaitasun eza nabarmena eragin ditzakete. Hitz batean, komunikatzerako unean, disfluentziak prozesu kognitiboaren isla dira, eta balio handikoa izan daitezke nahaste horiek detektatzeko.

2.2.6 Egitura biomekanikoan eragina

AD batez ere gaixotasun kognitibo den arren, fonazio- zein artikulazio-asaldura biomekanikoak erdietsi ditzake. Hori ezin zaio transmisio- edo modulazio-arazoei egotzi, baizik eta distonia orokor bati. Fonazio disprosodiko eta astenia zein artikulazioen ahultasuna bezala adierazten da. Fenomeno horiek inflexiorik gabeko eta asteniadun-ahots monotono baten antzera jasoko da, baita airepean eroritako itxiera glotiko defizitarioa bezala ere [39][56]. Artikulazioen ahultasunak, batez ere kontsonante plisiboetan, bokalaren hasiera bigundu eta urtua izango ditu ondorio baita hipernasalitate areagotua ere. Garapenaren ondorengo faseetan, laringeko erreflexua aldatua egon daiteke, eta horrek, disfagia eta disfonia zorrotza sor ditzake. Dena den, ia ziur, gaixoak ahoz komunikatzeko gaitasun osoa galdua izango du etapa horretara iritsi orduko.

2.3 Eskuko metodologiak

2.3.1 Sarrera

Ezagutza-ebaluazioak, AD detektatzeko tresna kliniko baliagarrietako bat denak, testak egiteko oso eremu zabala irekitzen du eta ez ohiko zahartze-prozesu baten hasierako faseak ezagutzeko erabil daitezke. Honako hauetan atxikitzen dira: screeningean, jatorriaren diagnostiko diferentzian, nahastearen larritasuna aztertzean, depresioa detektatzean eta gaitzaren progresioa monitorizatzean [57].

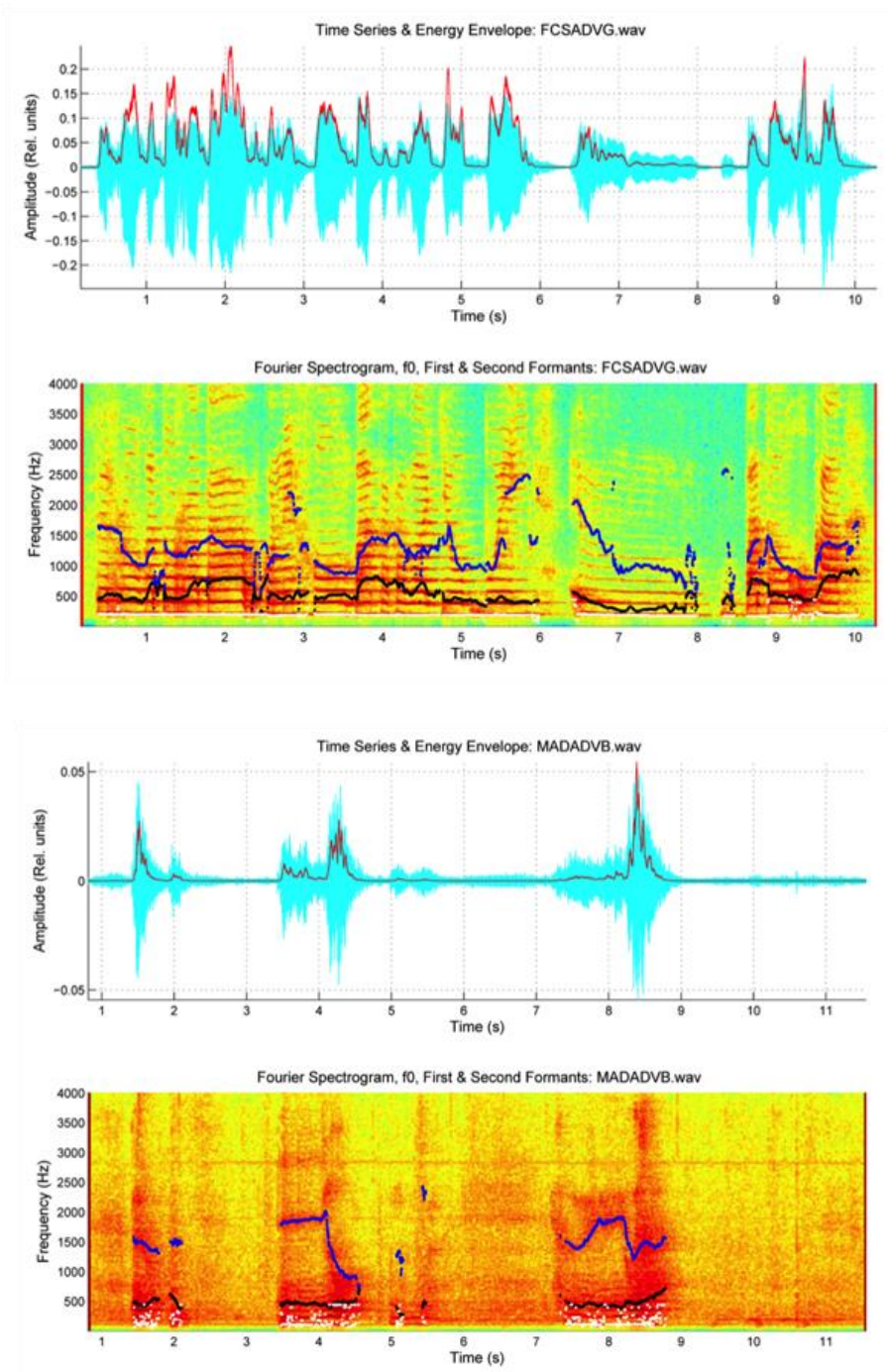
Ebaluatze kognitiboa egitean, hizkuntza eta hizketa dira azterketaren elementu garrantzitsuak, giza komunikazioaren prozesuaren zati ezinbestekoa baitira. Komunikazio-gaitasunen andeatzea honako faktore hauek kontuan hartuta aztertzen dira gehienetan: oroimenaren nahastea, izendatzea, gogoratzea, hitz-jarioa eta emozioen kudeaketa.

2.3.2 Ahozko komunikazio-gaitasunen ebaluaketa

2.3.2.1 Hitz-jario kategorikorako atazak

AD-ren hasierako fasean, oroimenaren nahastearen ondorioz agertzen den lehen sintometako bat izendatzeko defizita da. Ildo horretan, izendatzeko akats motak eta haien eboluzioa gaixotasuna aurrera egin ahala ikertzen ari dira; izendatzeari dagokionez, gehien ikertu diren akatsak honako hauek dira: ahozko parafasia semantikoa, ahozko parafasia fonemikoa, erantzunik eza (anomia garbia), izenaren falta, zirkunlozioa eta loturarik gabeko ahozko parafrasia [17].

2. KAPITULUA: ARTEAREN EGOERA



Irudia 2. 1: Atazetara mugatutako proba kliniko [39]. Animalien izenak. Iraupena 10 segundu kasu bakoitzean. Hizkeraren azterketa BioMetroLing-en bidez. a) Goiko aldean kontrolerako subjektu emakumezkoa. b) Beheko aldean AD duen gaixo gizona. Zian: hizketaren arrastoa. Gorria: energia-zorroa. Zuria: f_0 . Beltza: lehen formantea. Urdina: bigarren formantea. Alderdi garrantzitsuak: Kontrolatutako subjektuak f_0 egonkorra sortzen du eta espektoak maiztasun-laginketa muga arteko armonikoak erakusten ditu. Formanteak ahoskuntza-prozesu biziak erakusten dituzten zoliak dira. Ahoskuntza-segmentuak luzeak dira; isiluneak gutxi batzuk eta motzak. Ahoskuntza asteniko eta aireztatuaren ondorioz, AD duen gaixoak f_0 oso ezegonkorra sortzen du. Armonikoak ia ez dira ikusten. Formanteak trakets mugitzen dira, eta horrek, hondatutako ahoskuntza-prozesu nagia erakusten du. Ahoskuntza-segmentuak laburrak dira eta etenaldiak luzeak. Denbora bitarte berean, kontrolerako pertsonak hamar animaliren izenak esan dituen bitartean (txakurra, katua, arratoia, untxia, behia, zaldia,

idia, mmmm ... ahatea, antzara) gaixoak hiru besterik ez ditu esan (txakurra... eehh...otsoa...eeeh...ardia). Denbora-eskalak fitxategiaren edukiari dagozkio.

Beste sintoma argia eta goiztiarra hitz-jarioaren nahasmendua da. Oroimen semantikoa ebaluatzeko ataza edo ariketa neuropsikologikoen bidez neurtzen da (hitz-jarioaren testak, hitz-zerrendaren sorrera azeleratua). Ariketa horiek erabilitako hizketa murriztu egiten dute lotura semantikoak kontrolatu ahal izateko. Haien artean, AD detektatzeko gehien erabiltzen den testa kategorien arabera hitz-jarioarena da (CVF) [58][59]. Test horrek laneko oroimenarekin zerikusia duten hitz-jarioaren bi azpiko osagai aztertzen ditu: clustering-a (azpikategoria batean segidako hitzak sortzeko gaitasuna) eta switching-a (azpikategoria batetik bestera aldatzeko gaitasuna).

Gehienetan, test hauetako ondoriozko aldagai primarioa ekoiztako hitzen kopuru osoan oinarritzen da. Diagnostikoaren arabera ia ziur AD duten pertsonen fonema- eta kategoria-jarioen hitz-errendimenduak okerrera jotzeko joera du [58][59]; kategoria-jarioaren nahasmendua, ordea, nabarmen larriagoa da fonema-jarioa baino [60][61], nahiz eta ez egokitze horri buruzko lanetan kontraesanak dauden.

Esperimentaziorako erabiliko den Animalien Izendatze (AN) ataza hemen taldeka daiteke.

2.3.2.2 Hizketa irekiagoko atazak

Azken urteotan, aztertzen ari dira hizketa irekiagoan oinarritutako hizketa-egintza konplexuagoak:

- a) Sidtis-ek eta Landker-ek sorterriko hiztunei forman eta esanguran ezagunak zaizkien esapide ez literalen burmuin-prozesatzea ikertu dute [62]; besteak beste, esamoldeak, atsotitzak, biraoak eta hizketaren formulak. Hizketa literal ezagunaren ezaugarriak eta hizketa ez literalenak bereizi egiten dira. Berez, afasian mantendutako hizketa automatikoak eta beste zenbait gaitz neurologikotan narriatutako hizketa ez proposizionalak iradokitzen dute hizketa mota horiek burmuinean gordetzen eta prozesatzen direla, sortu berriko hizketan ez bezala.
- b) Chand *et al.*-ek Ideia Dentsitatea (ID) aplikatu zuten AD detektatzeko [63]. Hasieran, ID proposatu zen narratibetako oroimenaren karga neurtzeko, testuaren edukiaren irudikatze semantikoaren bidez, proposizioak edo ideiak deituriko unitate semantikoen segida moduan, alegia.
- c) Gilmore *et al.*-ek ahozko irakurketaren abiadura balioztatu zuten subjektu osasuntsuekin eta AD dutenen eriekin egindako kontrasteko estimuluen maila-sorta zabala zela bide. Ondorioztatu zuten, AD-ren gaixoen erazketa-ezaugarrien abiadura eta dementziarik gabeko heldu zaharrenarena alderagarriak dira, baina bakar-bakarrik beren arteko kontraste handia duten estimuluen aurrean [64].
- d) Ostberg *et al.*-ek hitz-segida automatikoen atzeranzko oroitzenari buruzko

ikerketa bat aurkeztu zuten, non hitz-segidaren produkzioa aztertzen den. Hiru kategoria ikertzen dira: nahaste kognitibo subjektiboa, nahaste kognitibo arina eta demenzia arina Alzheimerren gaixotasunean.

Kasu horretan, diagnostikatzea helburu, hilabeteko izenak erabili ziren atzeranzko versus aurreranzko hitzen segidaren ekoizpena ikertzeko, demenzian hala abiadura nola zehaztasuna nahastuta daudela kontuan hartuta. Ondorioztatu zutenez, hitzen segidaren ekoizpenak (batez ere atzeranzkoak) informazio garrantzitsua eskaintzen du kostu baxuko demenziaren diagnostikorako [77].

Esperimentaziorako erabiliko den Irudien Deskribapena (PD) ataza hemen taldeka daiteke.

2.3.2.3 Hizketa ez-murriztuko atazak

Berrikitan, zenbait ikerketa berrik konplexutasun handiko egoerak aztertu dituzte, hizketa ez-murriztuko atazetan edo ariketetan oinarrituta (bat-bateko hizketa, elkarrizketak edota hizketaldi libre). Hizketa-egintza konbentzionaletako testek komunikazio-osagai espezifikoak ikertzen dituzte (izendatzea edota hitz-jarioa), baina ez dira kontuan hartzen diagnostiko goiztiarra egiteko garrantzitsuak izan daitezkeen oinarritzko komunikazio-gaitasunak, besteak beste, faktore emozionalak edota hizketaren hitz-ekoizpen erreala. Are gehiago, pertsonak heziketa-maila jaso duenean, baliteke hizketa murriztuaren testak MCI detektatzeko gai ez izatea, eriak behar adina baliabideak baititu izendatze-arazo hipotetikoak saihesteko eta, hortaz, gaitzak ezkutatuta jarraitzen du.

Hori dela eta, gizarte-markatzaile baliotsu bat, adibidez, hizketaldi libre batean ahoskatutako hitzen proportzioa, gaixotasunaren zantzu goiztiarra izan daiteke. Biomarkatzaile berri horiek kognizio normaletik MCI-rako trantsizioekiko sentikor izan daitezke, eta biomarkatzaileei buruzko ikerketak osa ditzakete arriskupeko parte-hartzaileak aukeratzeko edota identifikatzeko lehen sintomak agertu aurreko fasean [38][65].

Esperimentaziorako erabiliko den Bat-bateko Hizketa (SS) ataza hemen taldeka daiteke.

2.3.3 Screening test klasikoak

Ikertzaileak urte luzeetan ibili dira hautemate-teknika garesti eta inbasiboak alternatibak bilatzen. Mintzamen-gaitasunetan edo emoziozko egoeren aldaketetan izaten den andeatzea sakon analizatu dira screening test tradizionalak direla bide.

Ondoren zehaztuko dira gehien erabiltzen diren screening test klasikoak, ahozko komunikazio-analisiarekin batera egiten direnak.

2.3.3.1 Hizkuntza-gaitasuna hautematea

Woodford-ek eta George-k [66] ahozko komunikazio-defizita hautematearen bidez nahaste kognitiboak ebaluatzeko gehien erabiltzen diren tekniken berraztertze zehatza egin dute. Deskribatutako screening test klasikoak era askotakoak izan daitezke: oso iraupen laburreko (minutu bat baino gutxiagoko) screening tresna laburretik zenbait ordu behar dituzten hautemate neuropsikologikoetara. Gehienak oroimen-nahastea edota hiz-jarioa detektatzera zuzenduta daude, oso egintza mugatuen bidez. Test klasiko horien artean, gehien erabiltzen direnak honako hauek dira:

2.3.3.1.1 MMSE

Buruko Egoeraren Mini Azterketa (Mini Mental State Examination, MMSE) [67] edo Folsteinen testa eta bere aldaerak 30 puntuko galdeketa da zeina zabal erabiltzen den klinika- eta ikerketa-testuinguruetan nahaste kognitiboa neurtzeko. Gehienetan, dementzia aztertzeko erabiltzen da eta egoki korrelazionatzen da beste kognizio-screening testen emaitzekin eta test neuropsikologikoekin. Are gehiago, test horren bidez neurtu daiteke nahaste kognitiboaren larritasuna eta progresioa, eta jarraitu daiteke aldaketa kognitiboen bilakaera luzetarako hautemateen bidez edota tratamenduei izandako erantzunaren bidez. Testak gutxi irauten du, 5-10 minutu, eta baliagarria da zenbait funtzio aztertzeko, besteak beste: erregistroa, arreta eta kalkulatzeko, gogoratzea, mintzamina eta agindu sinpleak betetzeko eta orientatzeko gaitasuna [3].

2.3.3.1.2 BNT

Izendatze Boston Test Eraberritua (Revised version of the Boston Naming Test, RBNT) [68] beste erreferentzia klasikoa da. Izendatze Boston Testa [57][68] funtzionamendu kognitiboa hautematen lagundu dezakeen screening tresna da. Test horren bidez diagnostikatu daiteke hitzak aurkitzeko gaitasuna eta egintza horrekin zerikusia duten funtzioak. Jatorrizko testean, 60 lerro beltzen marrazkiak daude, errazentatik zailenera antolaturik, eta testa egiten duenak identifikatu behar ditu. Testaren aurkezleak ordenean aurkezten ditu, eta eriak 20 segundu ditu identifikatzeko. Test hori Alzheimerren Gaitzaren Erregistroa Ezartzeko Partzuergoa (CERAD) izeneko testen bildumaren parte bat da [69].

2.3.3.1.3 CERAD

Alzheimerren Gaitzaren Erregistroa Ezartzeko Partzuergoaren (CERAD) beste testak, CERAD-ren Saila Neruopsikologikoa [70], CERAD-ren Oroitzapen Atzeratueta Hitzen Zerrenda [69], Kaliforniako Ahozko Ikaste Testak (California Verbal Learning Test, CVLT) [71], Hopkinsen Ahozko Ikaste Testak (Hopkins Verbal Learning Test, HVLT) [72], Hitzzerrendak Ikastea edota Tarte zabalekin egindako Lorpenen Probaren Test Eraberritua (Wide Range Achievement Test-Revised, WRAT-R) [73] dira.

Alzheimerren erregistroa ezartzeko partzuergoarentzat garatutako test neuropsikologikoak, gaitz horren ikerketa klinikoetan erabiltzen dira, nahaste kognitiboak neurtzeko gehienbat [66][69]. Gaitzaren zenbait agerpen hautemateko, CERAD-ek tresna estandarizatu batzuk garatu zituen, besteak beste: klinikoa/neuropsikologia, neuropatologia, dementzietan portaera sailkatzeko eskala, familiako historiari buruzko elkarrizketak eta zerbitzuen beharra detektatzea [70].

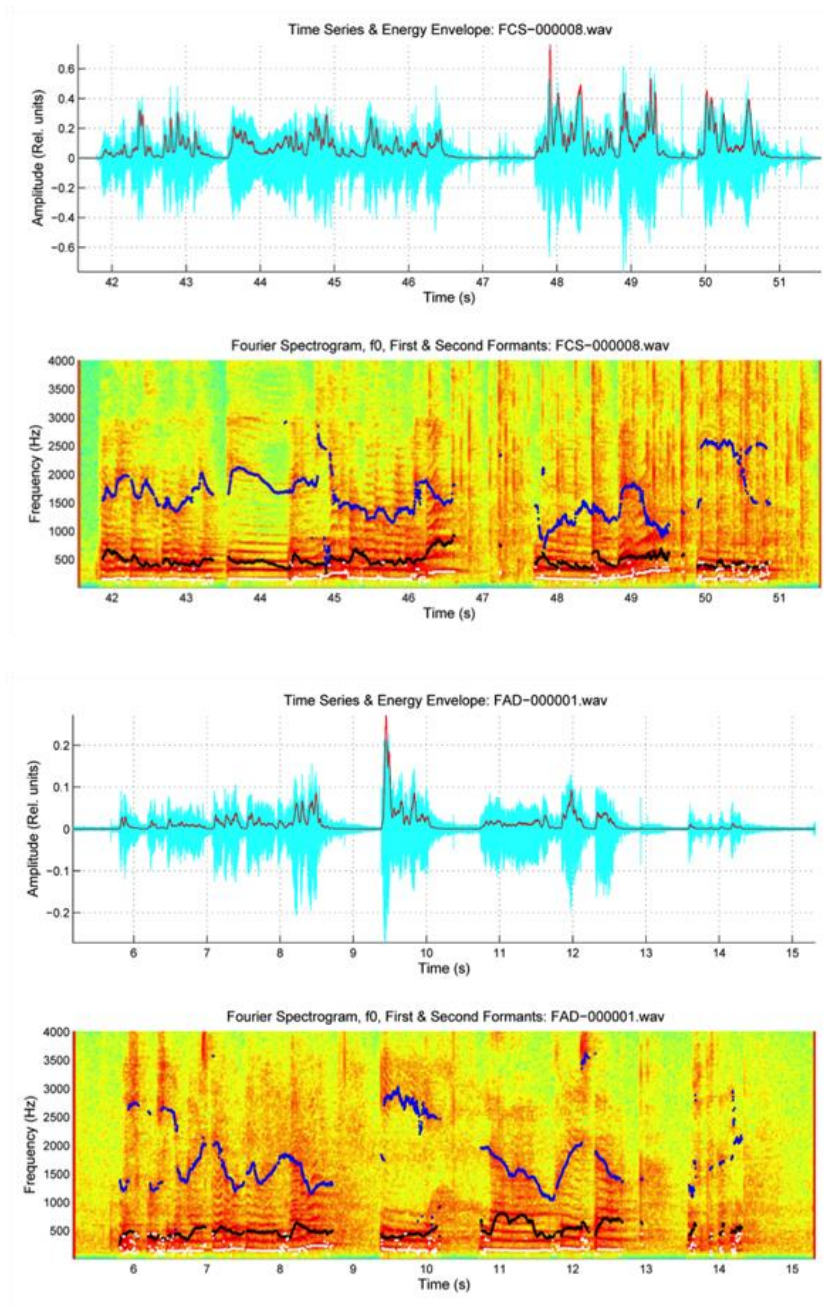
Zehazki, Tarte zabalekin egindako Lorpenen Probaren Testa 4 (WRAT4 ingelesez) [74] errendimendu-proba bat da eta modu eraginkorrean ikasteko, komunikatzeko eta pentsatzeko oinarrizko gaitasun akademikoak neurtzen ditu: irakurtzea, letraka esan/irakurtzea, funtsezko eragiketa matematikoak egitea.

2.3.3.1.4 RAVLT

Reyren Entzute-Hitzeko Ikastearen Testa (RAVLT) delakoa baliagarria gertatu da ahozko ikastea eta oroimena balioztatzeko, honako hauek barne: inibizio proaktiboa, inibizio retroaktiboa, gogoan atxikitzea, kodetzea versus berreskuratzea eta antolatze subjektiboa.

AD-ren kasuak detektatzeko, elkarrizketak eta bat-bateko hautematea Reyren Entzute-Hitzeko Ikastearen Testa dela bide [75][76] sentipen-markatzaile hoberenatarikoak dira. RAVLTek funtzio asko ebaluatzen ditu: entzute-hitzeko oroimena epe motzean, ikasteko gaitasunaren indizea, ikasteko estrategiak, interferentzia retroaktiboak eta proaktiboak, nahastea agertzea oroitze-prozesuetan, informazioa gogoan atxikitzea, eta ikastearen eta berreskuratzearen arteko aldeak.

Parte-hartzaileei elkarren artean ahaidetasunik ez duten hitzen zerrendak ematen zaizkie, zenbait probatan behin eta berriro agertzen direnak, eta errepikatzeko eskatzen zaie, denboratxo bat igaro eta gero.



Irudia 2. 2: Atazari murriztutako test klinikoaren irudiaren deskribapena. Luzapena: 10 segundu inguru kasu bakoitzean. Hizketaren azterketa BioMetroLing-en bidez. a) Goiko aldean kontrolerako emakumezkoa. b) Beheko aldean, AD duen emakume gaixoa. Zian: hizketaldiaren arrastoa. Gorria: energia-zorroa. Zuria: f0. Beltza: lehen formantea. Urdina: bigarren formantea. Alderdi garrantzitsuak: kontrolatutako subjektuak segmentu fonatu bakoitzaren bukaeran inflexio prosodiko argiak erakusten dituen f0 sortzen du (\uparrow) eta espeketroak armonikoen irudikatze ona erakusten ditu. Formanteak ahoskuntza-prozesu biziak erakusten dituzten zoliak dira. Ahoskuntza-segmentuak luzeak dira; isiluneak gutxi batzuk eta motzak. AD duen gaixoak f0-n inflexio prosodikoak sortzen ditu ere: armonikoen irudikatzea onargarria da. Formanteak pixka bat motelagoak dira. Fonazio-segmentuak laburragoak dira. Geldiune gehiago daude. AD duen gaixoaren narriadura, 1. irudikoarekin konparatuz, ez da hain gogorra. Denbora-eskalak fitxategiaren edukiari dagozkie.

2.3.3.1.5 CVF

Beste test baliagarri batzuk, atzeranzko oroitzapena eta Hitz-jario Kategorikoaren azterketa (Categorical Verbal Fluency, CVF) dira. Hitz-segida automatikoen ekoizpenaren atzeranzko oroitzapena honako hauetarako ere aspalditik erabiltzen den testa da: kognizio-nahasmendu subjektiboa, kognizio-nahasmendu arina eta dementzia arina Alzheimerren gaixotasunean [43][77]. Gehienetan, aurreranzko ekoizpena normal mantentzen da kategorietan zehar; atzeranzko abiadura eta zehaztasuna, ordea, nabarmen gutxitzen dira nahasmendu kognitiboan eta nahasmendu kognitibo arinean. Hortaz, atzeranzko produkzioak garrantzi handia du diagnostikatzerakoan.

CVF delakoa test batean oinarritzen da, zeinak nahaste kognitiboaren garapena neurtu eta kuantifikatzen duen gaitz neurodegeneratiboetan [37]. Test hori baliagarria da ere AD gaitzaren hasierako faseetan detektatzeko, prodromo-fasean, alegia [78][79]. Asko erabiltzen da mintzamen-gaitasunak, elkartzeko-oroimena eta funtzio betearazleak neurtzeko.

2.3.3.2 Emozio-egoera hautematea

Emozioek oroimena hobetu edo kaltetu dezakete hala iraupen laburreko oroimenari nola iraupen luzeko oroimenari dagozkien egintzetan. Harreman-oroimena erabiltzen denean, emozio-estimuluen presentziak eragin desberdina izan dezake, baita bi alde edo alde gehiagoren arteko harremanean ere.

Emozioen eragina oroimenaren kudeaketan zenbait testen bidez neurtu daiteke, NYUko (New York Unibertsitateko) istorioen bidez, adibidez [80], zeinak oroimen deklaratiiboaren test estandarizatuak diren. Test horretan, sei hitzezko deskripzio daude, astearen egun jakin batean jazotako gertaera batekin zerikusia dutenak. Testa zuzentzen duenaren prozedura honakoa da: pazienteak gertaeraren deskripzioa ozen irakurtzen du, eta segituan eskatzen zaio ahoz gogoratzea jazoerari buruz gogoratzen duen gutzia.

Kensinger *et al.*ek [81] egindako ikerketak nabarmendu du badirela semantikoki koherenteak diren ahozko estimuluak behar adina emozionalak direnak oroitzaren edertzea edo hobetzea eragiten dutenak heldu osasuntsuengan, gazteago edo zaharrago izan; AD dutenengan, ordea, ez da emozioarekin lotutako modulazio garbi hori hautematen [81].

Priming paradigma afektiboa dela bide, Padovan *et al.*ek aztertu zuten emozio-informazioaren aktibatze inkontziente eta automatikoa iraupen luzeko oroimenean. Haien ikerketan, *target* hitzak (gako-hitzak) aztertzen dira zenbait *primerekin*, positiboak zein negatiboak, batera. Egileek ondorioztatu zuten, AD-ren eriengan bias edo joera negatiboa mantentzen zen; *target* positiboek, ordea, asaldatu egiten zuten *priming* eragin emozionala, baita hasierako faseetan ere. [82].

Emoziozko estimulu ez-neutroek eragin nabarmena dute oroimenean; adibidez, Giffard *et al.*ek iradokitakoaren arabera, emozio kontzeptuek, neutroek direnek bezala, beren nolakotasun bereizgarrietako batzuk galtzen dituzte AD goiztiarrean, eta emozio-prozesuak mantentzen direnez, antzekotasun handiagoa dago emozio kontzeptu itxien artean, kontzeptu neutro itxien artean baino [83].

Tapus *et al.*ek metodo terapeutiko oso interesgarria aurkeztu dute robotekin; izan ere, hori dela bide, dementzia dutenen portaera positiboa areagotu egiten da [84].

Henry *et al.*ek AD dutenen eta kontrol-subjektuen portaera eta emozioak aztertu zituzten. Esperimentu horretan, film-klipak ikusarazten zaizkie bi motatako pazienteei, emozioen bat-bateko adieraztearen, ezabatzearen edo areagotzearen egoerapean. Bi taldeek arazoak izan zituzten jokaera-areagotzearekin. Alabaina, nahita egindako ezabatzearen erabilera osorik mantentzen zen, baita AD dutenengan ere; hori bat dator zahartze-ereduekin, zeinen arabera emozioak kontrolatzeko prozesuak nahiko automatikoak diren zaharrengan [46].

Zhang *et al.*ek ere jakinarazi zuten, AD-ren fase goiztiarrean dauden eriek emozio-kontrol automatikoko mota batzuk burutzeko gai dira, eta oso aldeko jarrera erakusten dute emozio-ezagutzaren eta emozio-oroimenaren aurrean, zeinak ezinbestekoak diren haien funtzionamendu sozio-emozionala mantentzeko, nolabaiteko moldatzearen laguntzaz. [85].

Bergmann *et al.*ek ikertu zituzten hala bi balentzien eraginak (neutroa versus positiboa versus negatiboa) nola kitzikapena (baxua versus altua) inter-item WM (Laneko Oroimena) loteslea eta LTMren (Iraupen Luzeko Oroimenaren, LTM ingelesez) ariketen bidez. Esperimentu horietan, hainbat etiketatutako irudi erakutsi ziren, eta WMri eta LTMri lotutako bi fasetako testak burutu ziren. Ikerketak agerian utzi zituen bi kasuen arteko desberdintasunak maila emozionalei dagokienez. Desberdintasunak txikiagoak ziren LTMren ariketetan [86].

2.4 Teknika automatikoak

Screening testen bidezko kognizio-hautemateak, zuzenean edo zeharka, hizketaren analisia, galderak, elkarrizketak, izendatzea, *clustering* semantikoa edota deskripzioak ekartzen ditu.

Hizketa oso erraz grabatu eta aztertu daiteke, eta oso informazio osagarri baliagarria eskain dezake honako hauen neurtze objektibo eta zehatza direla bide: iraupena, hizketaren ezaugarriak, faktore biomekanikoak edota hizketa-nahastearen maila.

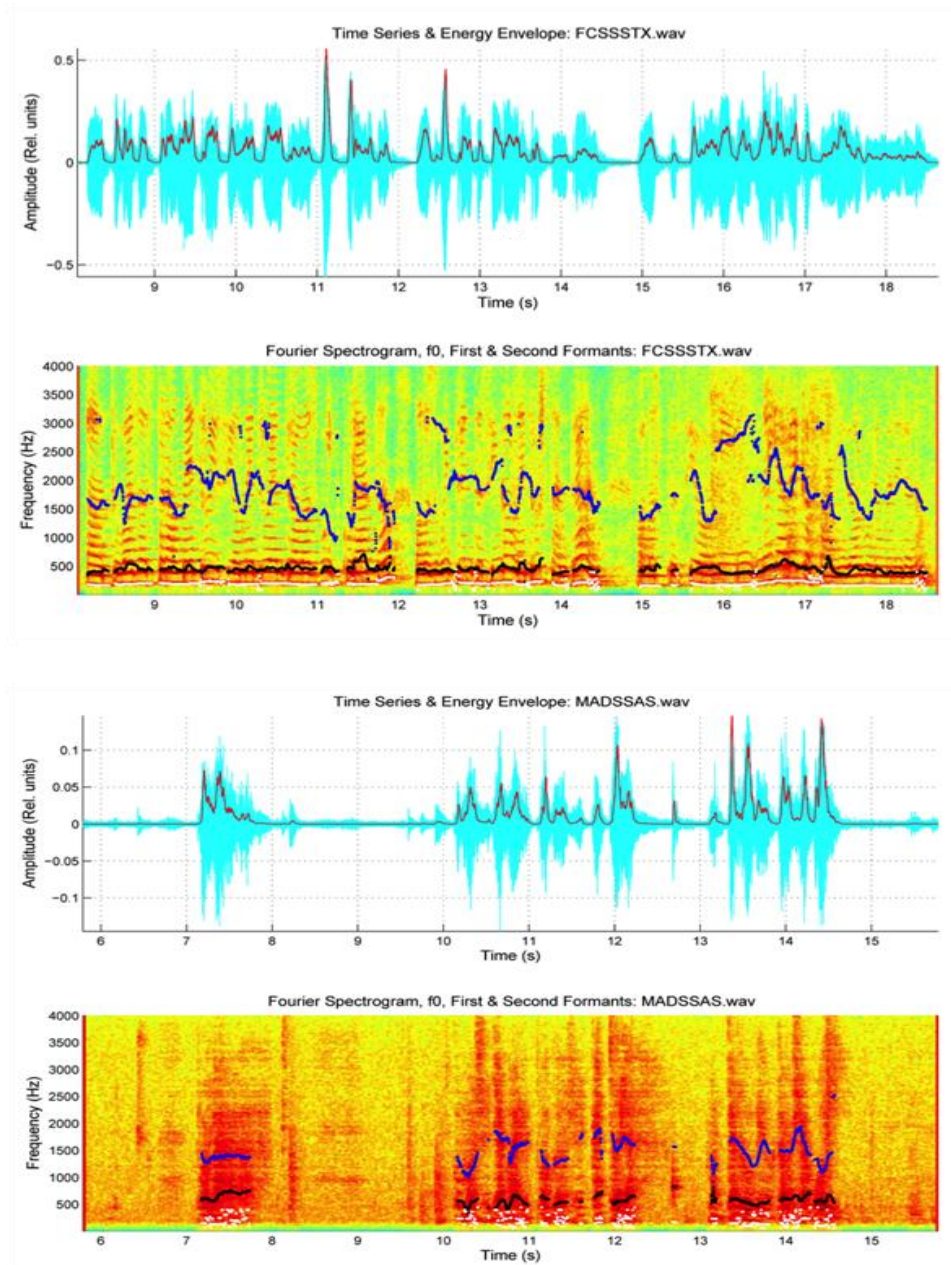
Lortutako informazio hori lagungarri izan daiteke espezialistek zehatzago balioetsi dezaten gaixotasuna hasierako faseetan. Atal honetan, sistema-mota horren garapena deskribatuko da, eta, aldi berean ere, gaur dauden tresnak AD hautemateari begirako hizketa automatikoaren analisirako.

Duela gutxi arte, espezialistek eskuz egin zituzten test horiek; alabaina, azken urteotan, sistema eta tresna automatikoak gero eta gehiago erabiltzen ari dira baita jarduera klinikoan ere.

Gaur egun, aurrez aurreko hizketa-testak eta gaixoekin eta beren senideekin automatikoki egindako elkarrizketa grabatuak aztertzen dira, eta aukera horrek oso tresna baliagarriak eskainzen ditu erabilitako hizkuntza eta erantzun emozionalak aztertzeko. Teknika aurreratu hauen bidez, hiztunen bat-bateko hizketa edota hizketa murriztuaren azterketa egin daiteke, eta baita ahotsaren analisi biomekanikoa ere.

Hizketa automatikoa aztertzeko teknikok zehaztasun handiz neurtu ditzakete ahozko hizkuntzaren nahasteak eta emozio-erantzunetan izandako aldaketak; biak AD-k kaltetzen ditu fase goiztiarretan, eta gaixotasunaren larritasuna balioesten lagun dezakete.

Gainera, diagnostikatzeko prozedura horiek pazientearen ohiko ingurunean egin daitezke (anbulatorioan edo etxean), hau da, giro lasaian eta teknikariak ez diren pertsonak edo senideak direla medio [87][19]. Pazienteek ez dute estresagarritzat jotzen bat-bateko hizketaren edo hizketa murriztuaren analisia [19][88]. Izan ere, patologia horiek diagnostikatu eta ezaugarritzeko teknika horiek ahozko test edo elkarrizketa sinpleak baino ez dituzte behar. Azken urteotan, aurrerapen ugari izan dira AD edo MCI detektatzean eta monitorizatzean.



Irudia 2. 3: Bat-bateko hizketa [39]. Luzapena: 10 segundu inguru kasu bakoitzean. Hizketaren azterketa BioMetroLing-en bidez. a) Goiko aldean kontrolerako emakumezkoa. b) Beheko aldean, AD duen gizonezko gaixoa. Zian: Hizketaldiaren arrastoa. Gorria: energia-zorroa. Zuria: f0. Beltza: lehen formantea. Urdina: bigarren formantea. Alderdi garrantzitsuak: kontrolatutako subjektuak segmentu fonatu bakoitzaren erdian eta bukaeran inflexio prosodiko argiak erakusten dituen f0 sortzen du (\uparrow) eta espektroak armonikoen irudikatze ona erakusten ditu. Formanteak ahoskuntza-prozesu biziak erakusten dituzten zoliak dira. Ahoskuntza-segmentuak luzeak dira; isiluneak gutxi batzuk eta motzak. Ahoskuntza asteniko eta aireztatuaren ondorioz AD duen gaixoak ez du f0 egonkorra sortzerik lortu. Armonikoak ia ez dira ikusten. Formanteak trakets mugitzen dira, eta horrek, hondatutako ahoskuntza-prozesu nagia erakusten du. Ahoskuntza-segmentuak laburrak dira eta etenaldiak luzeak. AD gaixoaren narriadura gogorragoa da [39] kasuan bezala. Denbora-eskalak fitxategiaren edukiari dagozkio.

2.4.1 Hizketa automatikoaren hautematea

Espinoza-Cuadros *et al.*ek hitz egiteko hiru estilo aztertu zituzten: hizketaldiak elkarrizketa estrukturaletan, pasarte motz baten irakurketak eta irudien ahozko deskripzioak.

Ondoren, *Random Forest* sailkagailua aplikatu zuten solasetik ateratako zazpi prosodia-neurritan. Esperimentazio-esparru horrek argi utzi zuen identifikazio automatiko eta objektibo horrek adineko helduengan duen potentziala MCI-ren hasierako sintomak azalerazteko. Hizketaren laginak biltzen joaten dira nahaste kognitiboari buruzko screening testak bitartean. [89].

Gonzalez-Moreira *et al.*ek prosodia automatikoki aztertu zuten dementzia arina dutenei eta adineko pertsona osasuntsuei eginiko irakurketa-ataza batean. Emaitez erakutsi zuten proposaturiko hizketaren analisi konputazionala aukera alternatibo bideragarria dela dementzia-ezaugarriak adineko jendearengan automatikoki identifikatzeko [90].

Romerok eta Kurz-ek luzetarako ikerketa egin zuten, zeinak, Alzheimerra duten gaixoengan bat-bateko hizketaren okerragotzearen ratioa neurtu eta faktore kliniko eta demografikoen eragina ratio horretan zehaztu zuen.

Horrez gain, beste zenbait eredu aztertu ziren: bat-bateko hizketaren nahasmendua komunikazio- eta semantika-alterazio nabarmenekin, hizketa automatikoaren alterazio arinak baina gogoan gordetako egitura fenomenikoekin.

Ikerketak ondorioztatu du dementziatzko historia familiarra duten pazienteek bat-bateko hizketa azkarrago txartzeko joera dutela. Bat-bateko hizketa txartzearen ratioaren eta beste faktore kliniko eta demografikoen arteko zerikusia hasierako bat-bateko hizketa-nahastean bakarrik antzeman zuten [91].

Thomas *et al.*ek metodo automatiko eta objektibo berritzaileak pazienteak diagnostikatzeko bat-bateko hizketaren azterketa dela medio aztertu zituzten, AD detektatu eta neurtzeko arazoarekiko zenbait hurbilketa lexikalez baliatuta. Hurbilketa horiek gramatika-faktoreak ikertu zituzten batez ere metodologia estokastikoen, n-gramen, bidez [92].

Hizketa eta mintzaira aldarazten dituzten zenbait gaixotasun neurodegeneratiboren diagnostiko automatikoari dagokionez, Peintner *et al.*ek oso emaitza akuilagarriak lortu zituzten ikasketa automatikorako paradigmak erabiltzeari esker. Egileok bereizgarriak atera zituzten audio-seinaleetatik eta hitzak lortu zituzten transkripzioko teknologia automatikoak erabilia. Horrela, automatikoki ikasitako ereduak iragartzen dute AD duen pazientearen diagnostikoa [93].

Verucci *et al.*ek testu batetik automatikoki atera zuten ID identifikagarria. Atariko emaitzak aztertu zituzten ondo garatutako kazetaritza-testu batean eta dementziadun pertsonen hizketa-transkripzioetan oinarrituta [94].

Roark *et al.*ek MCI-ren pazienteek azterketa neuropsikologikoetan emandako ahozko erantzunak aztertu zituzten. Adineko subjektu osasuntsuen eta MCI duten pazienteen arteko diskriminazioa egiteko ohiko hautematean, ahozko kontakizun-oroitzapena grabatu zen. Ondoren, hizketa-markatzaile automatiko batzuk, besteak beste, hizketa-ezaugarriak (maiztasuna eta etenaldien iraupena) edota konplexutasun linguistikoaren neurriak, atera zituzten irarketaren audiotik eta transkripziotik. Emaitzek agerian jarri dute neurri automatiko osagarri horien baliagarritasuna MCI-rako [95].

Martinez-Sanchez *et al.*ek tresna erdiautomatikoa aurkeztu dute hizketaren denbora-antolaketa ikertzeko ahozko irakurketa-ataza batean, Alzheimerdun pazienteen test neuropsikologikoetan integratuta. Esperimentu horietan, AD dutenek hizketa- eta artikulazio-ratio murriztuta izaten dute fonazio-denboraren eraginkortasun txikiagorekin eta etenaldien kopuru eta proportzio handienekin batera. Hortaz, seinaleak prozesatzeko algoritmoak irakurketa-jarioarekin erabiliz gero, erabilgarriak izan daitezke hizketa-urritasunen ikerketarako AD-n [96].

Orlandi *et al.*ek metodo berri, automatiko, azkar eta fidagarri berri bat aurkeztu dute ahots-kandidatu delakoak aukeratzeko iraupen luzeko audio-irarketetatik hala klinika- nola etxe-aplikazioetarako [97].

Khodabakhsh *et al.*ek diagnostiko-hurbilketa automatikoa ikertu zuten, hizketa-bereizgarrien erabileran oinarrituta, egoera informaletan subjektuekin egindako bat-bateko elkarrizketak direla medio. Hizketa prosodikoa eta bereizgarri linguistikoak aztertzen dituzte, AD detektatu eta monitarizatzeko ikasketa automatikorako paradigmak erabiltzearen bidez [98][99].

Garrard *et al.*ek honako teknikaren ahalmena esploratu zuten: klinika-saiakuntzen barruan transkribatutako hizketa-ereduak sailkatzea, hiztegi-datuak soilik erabilia. Informazio aberatsena duten hiztegi-bereizgarriak identifikatzeko, informazio-irabazia (Information Gain, IG) delakoaren erabilera aztertzen da maiztasun apaleko edukidun hitzak, hitz generikoak eta adierazpen metanarratiboaren osagaien bidez [100].

Rudzicz *et al.*ek Alzheimerdun pazienteen eta roboten arteko interakzioak eguneroko lanak direla bide ikertu dituzte, etxeko ingurune simulatu batean. Hizketa Ezagutze Automatikoa (Automatic Speech Recognition, ASR) ebaluatu egiten da, solasean nahasteak edota arazoak adierazten dituzten ahozko portaeren indizeekin batera [101].

Hitz-jarioaren kalitatea ebaluatzen duten hizkuntza-bereizgarrien neurketa edota indize semantikoei dagokienez, sistema automatiko berri hauek emaitza azkarragoak eta errazagoak eskain ditzakete eskuzko gainerakoekiko [102][103].

López de Ipiña *et al.*ek lengoaiarekiko independenteak diren metodologiaren integrazioa AD hizketan detektatzeko aztertu zuten, Hitz-jario Kategoriko (Categorical Verbal Fluency, CVF) delakoari buruzko ataza baten bidez [38]; zehatzago, Animaliak Izendatzeko (Animal Naming, AN) ataza zen, honako helburua duen luzetarako ikerketa baten barruan: funtzioaren, garun-egituraren eta osasun-egoeraren biomarkatzaileak aztertzea, eta baita arrisku-faktoreak ere Alzheimerra goiz diagnostikatzeko, prodromo aurreklinikoan bereziki.

Ikerketa horretan, inklusio-irizpideak talde bakoitzaren ezaugarrietan oinarritzen dira [67][104][9].

Berriki, Ahmed *et al* [105] saiatu ziren identifikatzen luzetarako aldaketak profil-nahastean aztertzeko erabilgarriak izan zitezkeen hizketa-bereizgarriak. Ikerketa horrek agerian utzi zuen hizkuntza-osotasunean disrupzio progresiboa detekta daitekeela maila prodromaletik aurrera.

Meilan *et al.*ek [106] isiluneak edota ahotsik gabeko segmentuak aztertu dituzte, eta haien eragin nabarmena test neuropsikologikoan lortutako puntuazio orokorraren dibergentzian erakutsi dute.

Beste lan berri batek [107] agerian jartzen du ASR-ren sendotasuna MCI-ren hautemate klinikoan; egileek erabili dituzten hizketa murriztuaren test klinikoek beren aurreko lanetan jarraitutako bide berak jorratzen dituzte. Kontsulta arrunt batean, parte-hartzaile bakoitzak lau ahozko ataza egin zituen: atzeraka kontatzea, esaldiak errepikatzea, irudiak deskribatzea eta hitz-jarioaren testa. Lortutako emaitzak etorkizun handikoak dira.

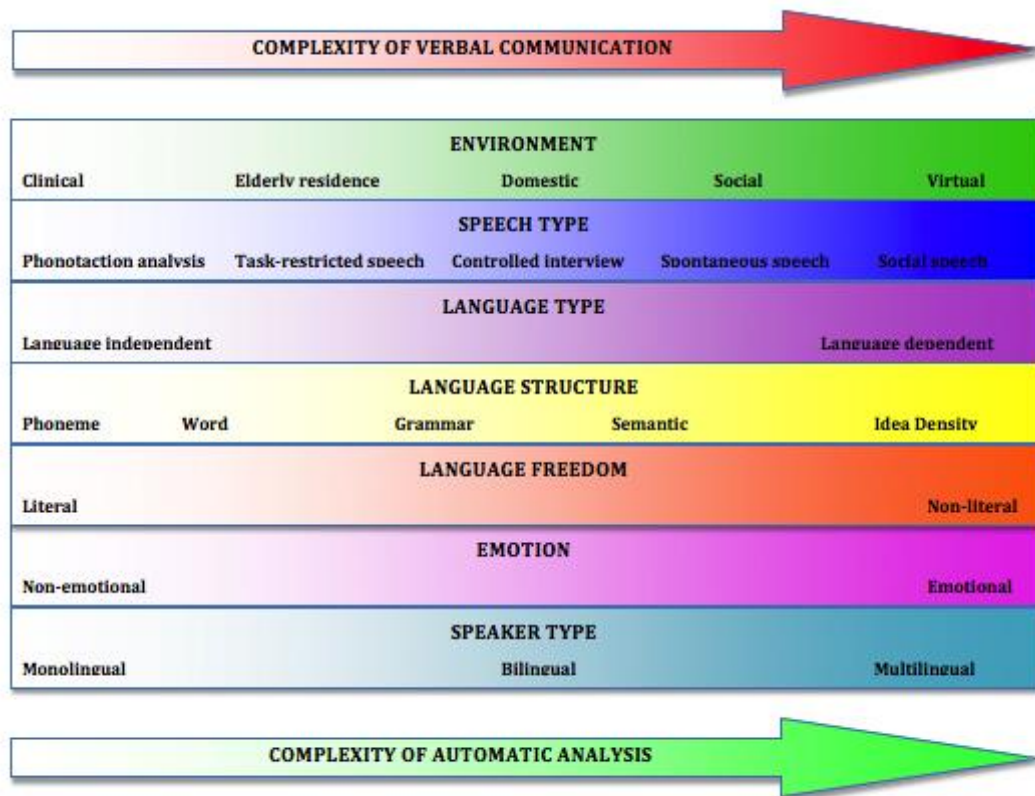
Dodge *et al.*ek [65] gizarte-hizkuntzari buruzko ikerketa interesgarri bat egin dute, nahiz eta laginaren tamaina nahiko txikia izan. Hizketaldian erabilitako formatua urrutiko bideo-hizketaldi bat izan zen, non parte-hartzaileak etxean ziren eta entrenatutako elkarrizketatzaileak estudioan. Ideia hori aurretik egindako ikerketetan oinarritzen da; izan ere, ikerketok erakutsi dute urrutiko teleosasun-komunikazioko formatuak eraginkorrak izan daitezkeela adindunen zenbait antzematearako [108].

MCI definitzeko erabili diren irizpide neuropsikologikoen [109] barne hartu dute gutxienez bi test klasikoren nahasmenduaren proba objektiboa (BNT, CVF eta CERAD testak). Egileek azpimarratzen dute Alzheimerren gaixotasunean, Hitz-Jario Kategorikoa proportziorik gabe nahastuta dagoela; irakurtzeko trebetasuna, ordea, ez da hain nahastuta [110][111], nahiz eta ikerketa guztiak ez dauden ados aurkikuntza horrekin [112].

Hizketaldian oinarritutako formatu sintetiko bati erabateko onura ateratzeko, bat-bateko erantzunetan jarri zuten arreta gehiago, egituratutako erantzunetan baino. Hizketaldi horiek grabatu eta transkribatu ziren, elkarrizketatzaileek versus parte-hartzaileek ahoskatutako hitzen zenbaketak neurtzeko. Esperimentu horrek agerian jarri du plataforma birtual horien garrantzia nahaste kognitiboa detektatzeko.

2.4.2 Hizketa automatikoaren hautemate ez-lineala

Test zehatzen analisiaren automatizaziotik harantzago, oraingo ikerketek aurrerapauso garrantzitsuak eman dituzte hizketa askea aztertzerako bidean, bat-bateko hizketa barne, eta baita hizketa-prozesatze ez-lineala ere kontuan hartuta (dimensio dinamiko, fraktal konplexuaren analisia).



Irudia 2. 4: Ahozko analisi vs analisi automatikoa. Ahozko komunikazio-elementuen konplexutasunaren azterketa vs analisi automatikoa. Honako eskeman, ahozko komunikazio-elementuen konplexutasunaren azterketa gizakien arteko hitzezko komunikazioan parte hartzen duten oinarritzkoak diren hainbat elementu kontuan hartuta aurkezten da. Adibidez: 1) Konplexutasun txikiko eszenatokia: azterketa emozionalik gabeko ingelesezko izen zerrendatzerako proba klinikoa. 2) Konplexutasun handiko eszenatokia: azterketa emozionala eta hizketa askea dakarren bat-bateko elkarrizketa elebiduna Skype bidez.

Bat-bateko Hizketaren Azterketa Automatikoa (BHAA) (Automatic Spontaneous Speech Analysis, ASSA, ingelesez) eta Emozio-Erantzunaren Analisia (EEA) (Emotional Response Analysis, ERA, ingelesez) [19][21][38][40] beste bereizgarrien artean honako hauek aztertzen dituzte: enuntziatuen luzapena, betegarrien tipologia, disfluentziak eta soinudun segmentu eta ahotsik gabekoak

Lan honetan, EEA (*ERA*) aztertu da hala bereizgarri klasikoak nola Emozio-Temperatura (ET) (Emotional Temperature, ET, ingelesez) direla medio, beti ere kontuan hartuta ET bereizgarri sendoa dela ingurune errealean. ET bereizgarri prosodiko eta paralinguistiko gutxi batzuetan oinarritzen da, zeina hizketa-seinalearen denborazko segmentaziotik lortzen den.

Bereizgarri hasiberri eta ez-lineal horiek emandako informazio gehigarria bereziki erabilgarria izan daiteke entrenamendu-datuak murriztuak direnean. Lan honetan, frogatu egiten da hizketa aske lineala eta ez-lineala konbinatzeak duen garrantzia. Egileek maila altuko zehaztasuna lortzen dute AD-rako eta, aldi berean, pazienteak kontrolatzen dituzte kostu konputazionala kontrolatuta denbora errealean eta inguruko baldintza errealean.

Azken ikerketek hobetu egin dute metodologia hori eta MCI-n ere erabiltzen hasi dira [38]. Bi test horiek eta bat-bateko hizketa-metodoek frogatu dute ahalmen ikaragarria dutela Alzheimer gaixotasunaren hasierako diagnostiko-testei dagokienean.

Hizketa-prozesatzean edota portaera-aldaketetan oinarritutako metodoek beste abantaila bat dute, gainera: hizkuntzarekiko independenteak izatearena, eta hortaz, modelatuak izan ez diren edo minorizatuak diren hizkuntzak darabilten hiztunekin erabiltzeko aukera ematen dute.

Hizkera-analisi automatikoa erabiltzeak etorkizun handikoa izan daitezkeen beste ikerketa-lerro batzuk ireki ditzake, jarduera neuromotorearen zenbatetsi alderantzizkoaren bidez ahoskuntza eta artikulazio-disfuntzioa antzemateko hizketa exekuzioaren azterketa zinematikoaren bidez. Azken aurrerapenetan, analisi dinamikoak, sentzore anitzeko analisiak, elkarlaneko sistemak, ingurune birtualak, autoikasketean bideko sistemak edo Deep Learning-en oinarrituriko hurbilpenak, besteak beste, agertzen dira [113].

2.5 Neurona-Sare Konboluzionalak (CNN)

Badaude egun informazioa automatikoki sailkatzeko erabiltzen diren algoritmoak. Adimen Artifizialaren (AI) adar hau Machine Learning (ML) izenez da ezaguna.

Gure kasuan, interfaze ez-inbasiboaren bidez, MCI/AD edo CR gizabanakoen ahots-laginak jaso, eta behin prozesatuak, horien ezaugarrien erauzketa- eta hautaketa-

lanaren ondoren dagokion algoritmoari sartzen dio horrek sailkapena (MCI/AD edo CR) egin dezan.

Egun, informazioa sailkatzeko, hainbat algoritmo estatistiko daude baina horietako batzuek (Convolutional Neural Network, CNN) garrantzi handia hartu dute, azken urteetan erakutsi dituzten gaitasunak medio. Gure kasuan, CNN-ak erabiliko dira besteak beste, sistema unibertsal eta ez-inbasibo honek sailkapena egin zezan egokiak zirela otu dugulako. Arrazoi hori dela eta, atal honetan, CNN-ek azken bosturtekoan Machine Learning-aren diziplinan ekartzen ari diren iraultzari buruz arituko gara.

Memoria honen ikerketa-lerroa, eta beti ere, osasun-guneen irizpide etikoen baitan, AD goiz detektatzeko ekarpena egitera bideratua dago. Algoritmo-mota hauek, ordea, beste hamaika ikerketa-lerro bustitzen dituzte. Horregatik, teknikoki CNN-en arkitekturan mahaigaineratzen ari diren aurrekarietan sakonduko dugu eta ez hainbeste horiek osasunaren sektorean izaten ari diren puntako aplikazioetan.

Machine Learning-en (ML) ataza ohikoenak sailkatzea, erregresioa, clustering-a eta abar dira, baina baldintza batzuetan muga handiak dituzte orain arte erabilitako algoritmoek. Azken urteetan, ordea, Ikasketa Sakonerako (Deep Learning, DL) trebeak diren Neurona-Sare Konboluzionalak (Convolutional Neural Network, CNN) orain arteko ML algoritmoek zituzten mugak gainditzera etorri dira ataza-mota batzuetan, sarreradatuak ezaugarriak hierarkikoki antolatu eta hauetatik ikasteko iaotasuna handia erakutsi baitute. CNN-ak sarrerako datuak modu sakonagoan aztertzen dituzten (horregatik Deep Learning terminoarekin oso lotua dago) neurona-sareak besterik ez dira, baina badituzte desberdin egiten dituzten hainbat ezaugarri.

Adimen Artifizialaren (AI) adar honek, eta bereziki CNN-ek, aurrerapen handiak ekarri dituzte, besteak beste, hizketaren azterketan [114], bideoen ezaugarritzean [115], Lengoia Naturalaren Prozesatzean (NLP) [116] eta irudiak aztertzean [117] atazetan. Aibilezia berezia erakutsi dute datuak ezaugarritu eta ulertzeko. Ekarri duten iraultza dela eta zenbait sektore ekonomikotarako DL aplikazioak asko ugaltu dira: osasun-sektorean irudien azterketa konputazionalan [118], automozio-industrian [119] eta ahotsaren hautematean [120].

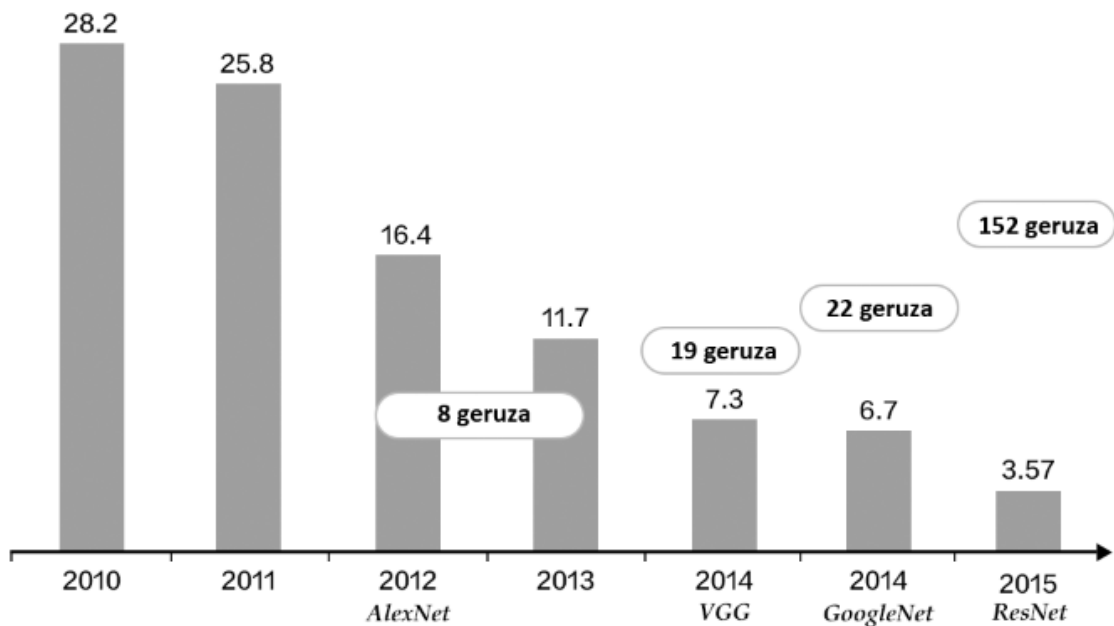
2.5.1 ILSVRC aurrekariak

Neurona-sareek azken urteetan bizi izan duten iraultzarekin zerikusi handia dute neurona-sare konboluzionalak (CNN). Nahiz eta argitalpen aintzindaria Yann Lecun-i [121] atxikitu, Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton-ren [122] argitalpenak eremu honetan izan duen eraginagatik garrantzitsuenentzat hartzen da.

CNN bat erabili zuten 2012. urteko ILSVRC [123] (ImageNet Large-Scale Visual Recognition Challenge) irabazteko. Konputagailuen bidez irudiak taldekatzeko

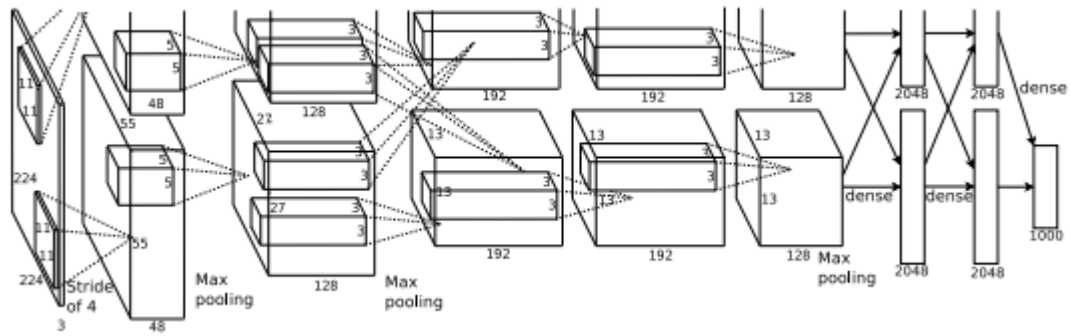
2. KAPITULUA: ARTEAREN EGOERA

txapelketa honek eta berarekin dakartzan argitalpenek punta-puntako ikerketa-lerroak irekitzen ditu eremu honetan. Hori dela eta, artearen egoeran nahitaez urtetik urtera txapelketa honetan parte hartu duten argitalpen interesgarrienei erreferentzia egingo zaie, bereziki 2012. urtetik aurrera, hots, Alex Krizhevsky *et al.*ek erabilitako Deep Learning algoritmotik aurrera.



Irudia 2. 5: ILSVRC-en errore-tasa baxuena lortu duten algoritmoen emaitzak. Honako irudian errore-tasa baxuena lortu duten algoritmoen emaitzak ikus daitezke 2010etik 2015 urtera bitartean. Bereziki aipatu behar da, 2012. urteko txapelketako sarea, horrek ekarri baitzuen inflexio-puntua lortutako errore-tasa beherakada handia zela eta. Emaitza horiek lortzeko, Alex Krizhevsky-k *et al.*ek CNN-a erabili zuten. Ordutik aurrera, Neurona-Sare Konboluzionalak izan dira ILSVRC-en protagonista nagusiak [124].

Alex Krizhevsky-k *et al.*ek erabilitako teknika ez da ezaguna egin txapelketa hau irabazteagatik, baizik eta irudien taldekatzean lortutako errore-tasa baxuarengatik. Ordutik aurrera, irudien konputazio-ulermenerako CNN-ak ikerketa-lerro berrien oinarri bihurtu ziren. Erabilitako sarea ez zen oso konplexua AlexNet izeneko CNN-ean.

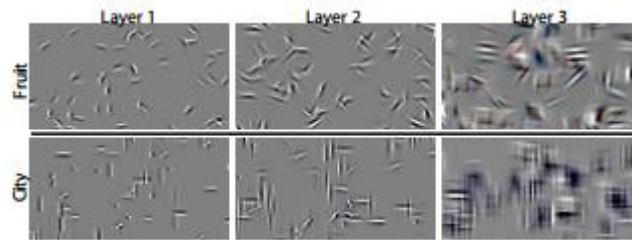


Irudia 2. 6: AlexNet. 5 geruza konboluzional (Convolution Layer) eta guztiz konektatutako (Fully Connected) 3 geruza besterik ez zituen AlexNet izenez ezaguna egin zen CNN-ak. ReLU (Rectified Linear Unit) aktibazio-funtzioak erabili zituzten, sare neuronal sakonek geruza ezkutuetan sarearen pisu eta ereduak azkarrago ikasten dituelako eta datuak ugaltzeko teknikak zein dropout izenez ezagutzen den teknika ereduaren parametroak gutxitzen laguntzeko eta overfitting arriskua txikitzeko. Erabilitako bi GPU-ak sei eguneguz egon ziren entrenatzen [122].

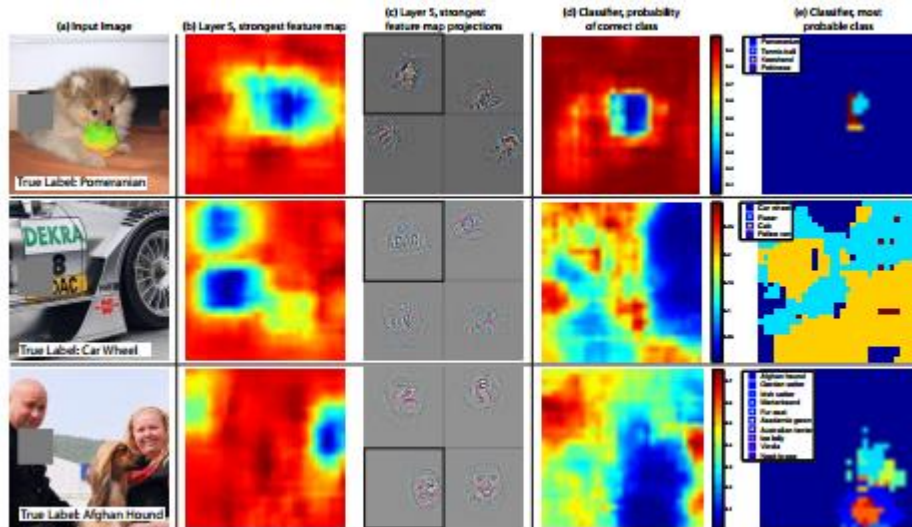
AlexNet-etik aurrera, CNN-en erabilera eta Deep Learning izenez ezagutzen diren ML algoritmoek bultzada handia izan zuten. Ordutik aurrera, ILSVRC txapelketa CNN egituren arteko txapelketa bilakatu zen. Baina ez hori bakarrik, espektrogramen azterketetarako aurrez entrenatutako AlexNet sarea ere erabiltzen da egun. Kasu honetan [125] ahots bidezko emozioen ezagutzarako sistema gisa. Alzheimer gaixotasuna taldekatzeko ere [126] Krizhevsky *et al.*ek entrenatutako sarea erabili da hainbat ikerketa-lerroetan.

Hurrengo urtean Matthew Zeiler eta Rob Fergus-en ZFNet CNN-a izan zen garaile, % 11.7-ko errore-tasarekin. Haren arrakasta CNN-en doikuntzan baino, horiek hobeto ulertzeko egindako lan didaktikoan datza. Iragazki eta pisuak ikusteko teknikek komunitate zientifikoaren arreta sortu zuten bereziki [127]. Lan horrek, 2013. urteko ILSVRC txapelketa irabazteaz gain CNN-en funtzionatzeko modua hobeto ulertzeko ekarpen handia egin zuen eta honen errendimendua eta egiturak hobetzeko hartu beharreko bideari buruz ere arrastoak utzi zituen. Barne-geruzen funtzionamendua ikusteko, Deconvolutional Network (DN) bezala ezaguna egin den teknika garatu zuten [128].

2. KAPITULUA: ARTEAREN EGOERA



Irudia 2. 7: Deconvolutional Network. Deconvolutional Network-en helburua ezaugarri-mapak zein motatako egiturek kitzikatzen dituzten ulertzea da. Horretarako ezaugarrien aktibazioak jaso eta CNN-en iragazki berak dituen DN batetara sartzen dira. Honako irudian lehen geruzan maila baxuko ezaugarria antzematen duen DN-a ikus daiteke eta geruza sakonagoetan sartu ahala, horrek antzematen dituen formak konplexuagoak dira [127].



Irudia 2. 8: Geruzen irudikapen grafikoa. *ImageNet Classification with Deep Convolutional Neural Networks* argitalpenetik hartutako irudi honetan geruzen irudikapen grafikoa ikus daiteke. Horrek CNN-ak hobeto interpretatzen asko lagundu zuen. Entrenatzeko AlexNet-ek baino irudi gutxiago erabili zituen eta iragazkiak ere txikiagoak izan ziren ezaugarri-mapak jatorrizko informazio gehiago gorde zezaten. GPU-ak hamabi egunetan zehar egon ziren sarea entrenatzen [127].

2014. urtean irabazlea GoogleNet izan zen arren VGGNet [129] sareari buruz hitz egitea gomendagarria da, horrek ekarri zituen berrikuntzak direla eta. Eredu horrek xalotasuna eta sakontasuna ekarri zizkien CNN-ei. Dimentsio txikiko kernel eta pooling iragazkiak gehitu zizkieten geruza kantitate desberdinetako CNN-ei. Sare horretatik arnasa hartu duten argitalpenak ugalduta dira soinu-araketarako [130] hau bezala.

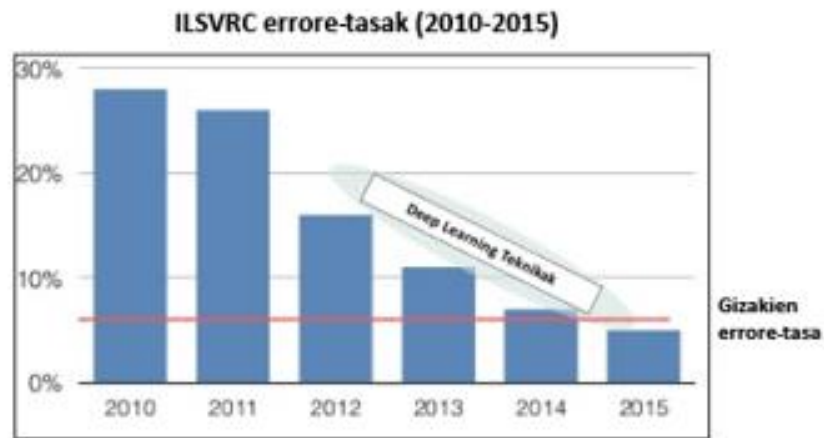
Taula 2. 1: Dimentsio espazial urritu eta sakontasuna handitzeko CNN-a. 3x3-ko iragazki txikiak erabili ziren eta horrek parametro gutxiago eskatzen ditu. Iragazki-kopurua gehitu zuten maxpool geruza bakoitzaren ondoren sakontasuna emateko, nahiz eta dimentsio espazial txikiak erabili. Kontzeptu berria ekarri zuen, hots, dimentsio espazial urritu eta sakontasuna handitzen duten CNN-en erabilera. Datuak gehitzeko teknika eta ReLU aktibazio-funtzioak ere txertatu zituzten eta, azkenean, Nvidia 4 GPU hiru astetan zehar ibili ziren entrenatzen [129].

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

GoogleNet-ek, 2014. urteko ILSVRC irabazteaz gain ikuspuntu berria eman zien CNN-ei eta memoria- eta energia-erabilera nabarmendu zuen. Christian Szegedy *et al.* [131] argitalpenak Inception izeneko teknika gehitu zien sareei baliabide informatikoak modu optimizatuagoan erabiltzeko. Teknika horrek CNN-en geruzak sekuentzialki pilatzeko monopolioa zalantzan jarri zuen eta sormenezko geruzen egituraketa on batek konputazionalki eraginkortasun hobea lor dezaketela agerian jarri zuen. AlexNet-ek baino hamabi bider parametro gutxiago erabili zituen eta GPU gutxi batzurekin astebetetz entrenatu zuten sarea.

Microsoft-ek, ResNet [132] izeneko sarearen eskutik, 152 geruzako CNN-a aurkeztu zuen 2015. urteko txapelketara. Geruza asko behar zituen baina lortutako errore-tasa % 3.6-koa izan zen; gizakiek horrelako atazetan dugun batez besteko errore-tasa baino txikiagoa. Hondar-blokeen teknika txertatu, eta honen bidez datuen tamaina espaziala asko txikitzea lortu zuten. Badira ahotsa aztertzeke ResNet, GoogleNet, VGGNet eta abar

erabili dituzten ikerketa-lerroak, adibidez bideo doinuen ezagutza, ezpainen irakurketa edo hizketaren ulermenarekin zerikusia duten hamaika aplikazio.



Irudia 2. 9: ResNet sarea. Lerro gorriak gizakiak irudiak taldekatzerako orduan dugun errore-tasa adierazten du. ResNet sareak errore-tasa txikitu zuen, tarteko geruza asko gehitu eta hondar-blokeen bidez tamaina espazialak txikituz [133].

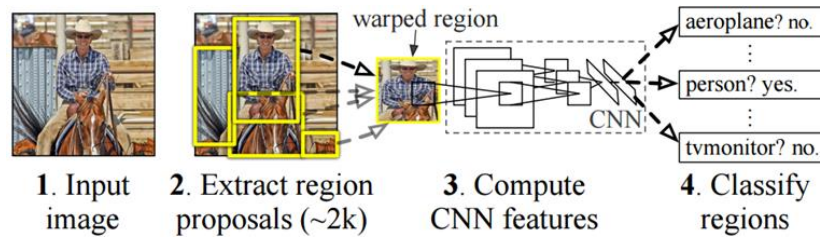
2016. urtean, ILSVRC, CUIImage-k [134] irabazi zuen arren, irudiak taldekatzerako orduan Trimps-Soushen taldeak % 2.991-ko errore-tasa-ko CNN-a eraikitzea lortu zuen. CNN-ak eta teknika hainbat bertsio gehituz errore-tasa hobetzea lortu da nahiz eta portzentualki asko hobetzea zaila den. Gogoan hartu behar da gizakien errore-tasaren azpitik taldekatzen ari direla konboluzio teknika horiek eta lortzen duten doitasuna oso handia dela.

2.5.2 Beste aurrekariak

ImageNet-ek urtero antolatutako txapelketak argitarapen oso interesgarriak ekarri ditu. Horretaz gain, badaude irudien prozesatze eta horien ulermen konputazionalan iraultza ekarri duten hainbat ekarpen.

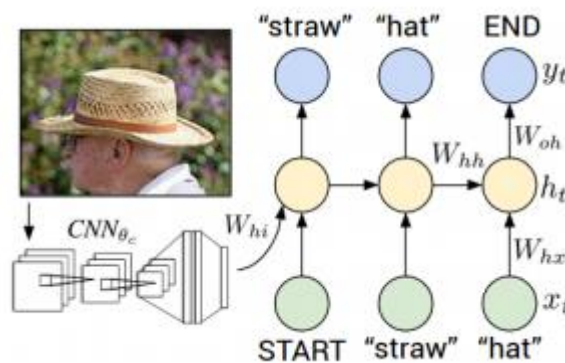
Aipatzekoak dira, adibidez, R-CNN-ek konputagailu bidezko objektuen taldekatzean eta horien kokalekuaren zehaztapenean ekarritako aurrerapenak. Ross Girshick *et al.*-ek [135] objektuen antzematearen buruhaustea ebaztea dute helburu eta erabilitako prozedurak desberdinak diren bi etapa argitu ditu: irudiaren eskualdea proposatu eta ondoren eskualde hori taldekatu.

Lehen etapan, R-CNN-ak objektuek izan ditzaketen eskualdeak proposatzen ditu. Ondoren, eskualde horiek, jadanik entrenatua dagoen CNN batetara sartzen dira. Jakina, eskualdeak, CNN-ak sarrerak jasotzeko egituratuak dauden dimentsiotara desitxuratu behar dira. CNN berriak, adibidez ResNet-ek, eskualde bakoitzaren ezaugarri-bektorea erauzi, eta bukatzeko, sailkapena egiten du.



Irudia 2. 10: Eskualdeak erauzi, desitxuratu eta CNN-an sartu. Eskualdeak erauzi, desitxuratu eta CNN-an sartu, taldekatzea egin dezan. Argitarapen honek zurrunbilo handia sortu zuen, eta horren ondoren, prozesu konputazionala lasterragotzeko beste proposamenak mahaigaineratu ziren [135].

Fast R-CNN [136] eta Faster R-CNN [137] abiaduraren arazoa konpontzeko gai izan ziren. Jakina, Faster R-CNN-a bertsio azkarrena da. Irudietako aurpegiek adierazten duten espresioa erauzteko [138], Faster R-CNN-a erabili dute Jiaxing Li *et al.*ek. Kalkulu-abiadura handitzeko bi teknika erabiltzen dituzte: bata, geruza konboluzionalen kalkulu konputazionala partekatzen du, eta bigarrenak, geruza konboluzionalen azken ezaugarri-mapa begiratuta taldekatzea egiten du.



Irudia 2. 11: CNN-en hitz deskriptiboak aurkitzeko gaitasuna. Kasu honetan, "lasto", "txanoa" eta beste hitz deskriptiboak aurkitzeko gaitasuna du sare honek [139].

Kontrako Sareen Sortzailea (GAN) Ian J. Goodfellow eta haren ikerketa taldeak [140] proposatutako sare honen helburua sarea berari ziria sartzea da. Jatorrizko irudi eta perturbazio txikia duen jatorrizko irudiaren kopia sartzen zaizkio entrenatutako sareari, GoogleNet-i adibidez. Nahiz eta gure begientzako irudi bera izan, gerta daiteke algoritmoak irudi bakoitza ezberdin taldekatu dezakeela. Teknika horrekin sareak aztertzen dituen datuen egitura irudikatzeko gai da, benetako irudi eta horren kopia artifizialaren arteko desberdintasunak aurkitzeko entrenatua izan baita. CNN batek erabili ahal izango dituen funtzioen erauzgailua da nolabait.

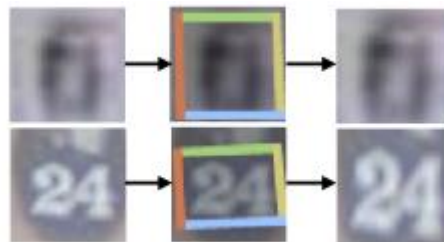
Stanford Unibertsitateko Andrej Karpathy-k eta Li Fei-Fei-ek [139] hainbat sareren nahasketa proposatu zuten Lengoaia Naturalaren Prozesatzearen (NLP) bidez irudien eskualdeen deskripzioak egiteko. Horretarako, CNN-ak eta Neurona-sare Atzerakariak (RNN) konbinatu zituzten. CNN-ek irudien taldekatzerako etiketa argia badute; kasu honetan, etiketa horiek esaldiak dira eta sareak esaldiaren eta deskribatu nahi duen irudiaren arteko erlazioa ondorioztatzen du. NLP eta konputagailu bidezko ikusmen-

2. KAPITULUA: ARTEAREN EGOERA

teknikak nahasten ditu eta ikerketa-lerro berriak irekitzen ditu, hainbat jakintzagai erabiltzen dituen teknika baita.

CNN-ei txertatzen zaizkien teknika berrien eremuan, 2015. urtean Max Jaderberg *et al.* [141], Londreseko Google DeepMind taldeko kideek, Espazio Eraldaketarako Sareak (STN) izeneko proposamen berria mahaigaineratu zuten. Kasu horretan, CNN egitura aldatu baino, datuak sarera sartu aurretik horiek maneiatzen dituzte, sare konboluzionalari lana kentzeko helburuarekin. Ohiko CNN-ek entrenamendu datu asko behar dituzte eskala eta biraketa desberdinak dituzten irudiekin taldekatze ez aldakorrek egiteko.

Teknika horren bidez CNN-ari ezaugarri-mapak eraldatzen irakasten dio entrenamenduaren kostua txikitzeko, baina hori baino garrantzitsuagoa den ideia bat gehitzen du Ikasketa Sakonaren (Deep Learning, DL) ikerketen eremuan: ez da beharrezkoa CNN-en egitura konplexua eraikitzen ibiltzea, errore-tasa eta konputazio-abiadura txikitzeko. CNN-ei lana erraztu besterik ez diegu egin behar, irudiak aurretik eraldatu eta sareen beharretara moldatuz.



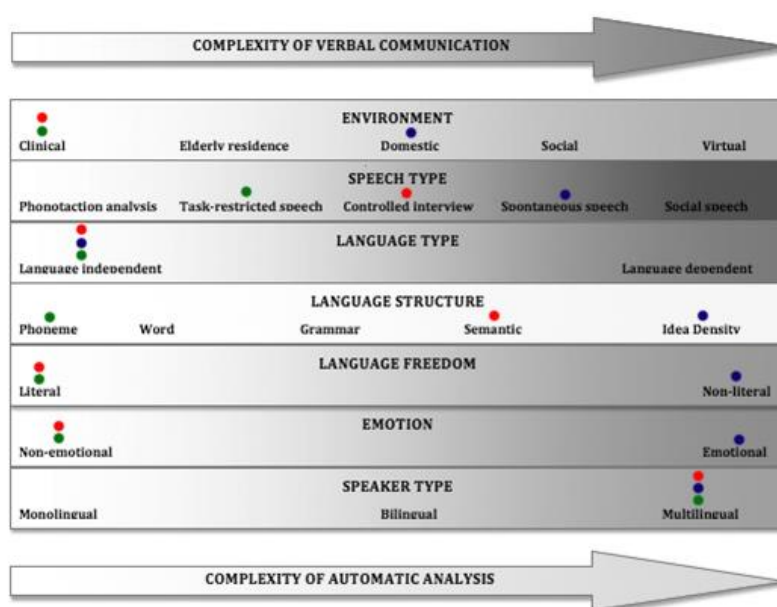
Irudia 2. 12: Irudia eraldatu eta aipagarria den eskualdea aukeratu. Ikus daitekeen bezala, CNN-ra irudi gordina sartu baino, lehenago irudia eraldatu eta aipagarria den eskualdea aukeratzen du. Ondoren, CNN-an txertatzen du sarrerako datu gisa, horrek taldekatu dezan [141].

Ikasketa sakonarekin paradigma eta alternatiba berriak zein neurona-sareekin lan egiteko arkitektura berriak mahaigaineratzen ari dira. Nature-n [142] Alex Graves eta *et al.*ek argitaratu zuten Differentiable Neural Computer (DNC) handitutako memoria duen neurona-sarea aurkeztu zuten. Bere memoria erabiltzeko gaitasuna erakutsi zuen, datu egituratu eta konplexuei erantzuna bilatzeko. DNC kodea eskuragarri dago DeepMind-en GitHub-ean [143].

3. KAPITULUA: MATERIALAK

3.1 Sarrera

Atal hau sakonago ulertzeko, **B.1** eta **B.3** eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira. Ikerketarako erabilitako materialak ahots-laginen datu-baseek osatzen dituzte. Horiek lortzeko, zenbait ataza eta baldintza izan dira eta datu-base orekatuak lortzera bideratu dira. Laginak sortzeko hiru ataza-mota erabili dira hizketa-komplexutasunaren maila desberdinez baliatuta: Hitz-Jario Kategorikoaren (CVF) Animalien Izendatzea (Animal Naming, AN), hizketa irekian oinarria duen Irudien Deskribapena (Picture Description, PD) eta hizketa ez murriztuko bat-bateko Hizketa (Spontaneous Speech, SS) desberdin erabili dira. Orokorrean, laginak ez dira asko baina orekatuak daude.



Irudia 3. 1: Ataza bakoitzaren ezaugarriak. Ataza bakoitzak ezaugarri propioak ditu eta honako irudian AN, PD eta SS atazen ezaugarrien hurbilpen bat egiten da. Berdez AN, gorritz PD eta urdinez SS atazak. Ezaugarrien artean zein inguneatan egin diren, hizketa-mota, hizkuntzarekiko mendetasuna, emozio-eratzuna aztertzeko balioa eta abar azaltzen dira.

3.2 Datu-baseak

Laginen bilduma guztia proiektuan inplikaturako erakundeen arau etikoak kontuan hartuta egin da eta grabazioak hainbat gunetan egiteaz gain, parte-hartzaileak desberdinak izan dira hiru atazetan.

Taula 3. 1: Erabilitako 3 ataza nagusiak. Laginen erauzketarako erabilitako 3 ataza nagusien inguruko informazioa jasotzen da taula honetan: zein inguruetan egin diren eta bakoitzean parte hartu duten gizabanakoen ezaugarriak hain zuzen.

Ataza	Grabazio gunea	Parte hartzaileak
AN: Categorical Verbal Fluency (CVF) edo Hitz-jario Kategorikoaren ataza (Animal Naming, AN, animaliak izendatzeko jarioaren bidez).	Ingurune klinikoan	MCI-ren pazienteak vs CR-ren pertsonak aztertzeko balioko dute.
PD: Picture Description (PD), irudien deskripzioaren ataza.	Ingurune klinikoan	AD-ren gaixoak vs control group (CR) pertsonak aztertzeko balioko dute.
SS: Spontaneous Speech (SS), hizketa espontaneoaren ataza.	Etxeko ingurunean	AD-ren gaixoak vs control group (CR) pertsonak aztertzeko balioko dute.

Ataza bakoitzetik datu-base desberdina sortu da eta horiek esperimentazio-lerroetan erabili dira. Guztira, esperimentaziorako, 4 datu-base sortu dira.

Taula 3. 2: Sortutako datu-base bakoitzaren datu demografikoak. Laginei esker sortutako datu-base bakoitzaren datu demografikoak taulan honetan ikus daitezke zehatz-mehatz.

Datu-baseak	Emakumea	Gizona	Adin-maila	Batez besteko adina	SD Adina
PGA-OREKA. AN ataza-azpimultzoa (CR/MCI)	36/21	26/17	39-74/42-79	56.73/57.15	7.8/8.9
MINI-PGA. PD ataza-azpimultzoa (CR/AD)	8/5	4/1	53-73/62-78	61.51/67.97	7.02/5.2
AZTIAHORE. SS ataza-azpimultzoa (CR/AD)	9/12	11/8	47-80/69-88	57.80/79.90	11.92/5.3
AZTITXIKI. SS ataza- azpimultzoa (CR/AD)	2/2	3/3	80-90/75-95	85.70/87.60	10.2/10.5

3.3 Atazak

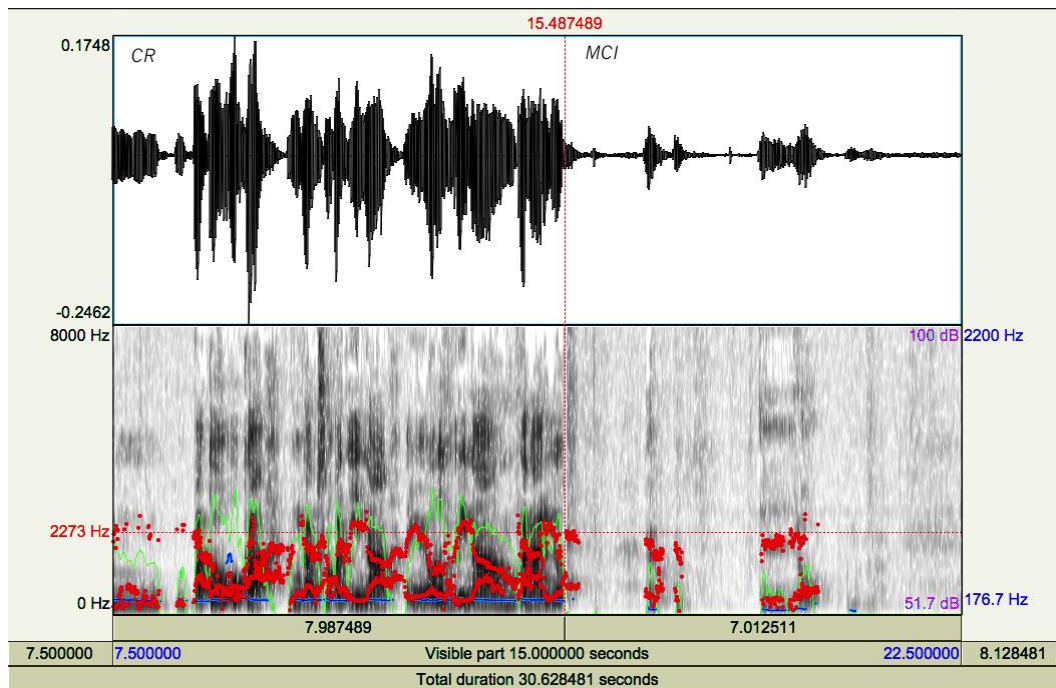
3.3.1 Animaliak izendatzeko jarioaren ataza (CVF)

Hitz-jario Kategorikoaren ataza edo Categorical Verbal Fluency (CVF) Animaliak izendatzeko ataza (Animal Naming, AN) gaitz neurobegetatiboetan izaten den nahaste kognitiboaren progresioa neurtu eta kuantifikatzen duen test bat da [144]. Maiz erabiltzen da hizkuntza-gaitasunak, oroimen semantikoa eta funtzio betearazleak balioesteko [145].

Definizioz, AN atazean, elkarrizketatzaileak pazienteari minutu batean ezagutzen dituen kategoria bateko (animaliak gure kasuan) izen guztiak azaltzeko eskatzen dio. Jardute klinikoan, bakar-bakarrik item-en kopuru osoa erregistratzen da. Hala ere, index espezifikoagoak defini daitezke, zeinek pazientearen egoera kognitiboan izandako aldaketa ñimiñoak detekta baititzakete eta, horrela, gaitz degeneratiboaren arteko bereizketa bereiztuagoa lor baitaiteke [5][12].

Laginean 187 pertsona osasuntsuk eta MCI duten 38 pazienteek parte hartzen dute, CITA-Alzheimer Fundazioko Gipuzkoa-Alzheimer Proiektuaren (PGA) [146] kohortean daudenak. PGA-k, Alzheimerren gaitza goiz detektatzeko, funtzioaren biomarkatzaileak, garunaren egitura, osasunaren egoera eta arrisku-faktoreak aztertzen ditu. PGA-ren helburu nagusia Alzheimerren aldi aurrekliniko eta prodromala ezaugarritzea da.

3. KAPITULUA: MATERIALAK



Irudia 3. 2: Hizketa murriztuaren test kliniko. Hitz-jario kategorikoaren edo Categorical Verbal Fluency-ren (CVF) ataza (Animal Naming, AN, animaliak izendatzea). Ezkerrean, kontrol-subjektu bat; eskuinean, MCI duen pertsona bat. Tonua, lerro urdina; ozentasuna, lerro berdea; formanteak, lerro gorria.

Kontuan hartu behar da, hizketa murriztuaren testek komunikazio-osagai zehatzak aztertzen dituztela, baina ez dituztela kontuan hartzen diagnostiko goiztiarrerako egokiak izan daitezkeen beste zenbait funsezko komunikazio-trebetasun; adibidez, emoziozko faktoreak edota hizketaren ekoizpen espontaneoak. Bestalde, goi-mailako hezkuntza duen pertsona batekin gerta daiteke hizketa murriztuaren testak kalte sotilak detektatzeko gai ez izatea, pazienteak balia bideak baititu izendatze-arazo leunak saihesteko eta, era horretan, gaitzak ezkutuan jarraitzen du. Horren haritik, egokiagoa litzateke beste gizarte-markatzaile bat, adibidez, elkarrizketa askean edo hizketa espontaneoan jaulkitutako hitzen proportzioa, gaitzaren hasierako zantzuen adierazlea izan baitaiteke.

Markatzaile berri horiek sentiberak izan beharko lirateke kognizio normaletik MCI-rako trantsizioekiko, eta beste biomarkatzaileekin erraz integratzeko gai izan, arriskupean izan daitezkeen pertsonak identifikatzeko [38][65].

Azterketa kliniko testuinguru batean egiten da eta talde bakoitzaren ezaugarrietan oinarritutako inklusio-irizpideekin.

3.3.1.1 Kontrol Taldea edo Control Group (CR)

- Parte hartzen dute tarteko adineko edo nagusiagoak, 39 eta 79 urteen artekoak, PGA-ren parte-hartzaile izateko adostasun informatua izenpetu eta gero.

- Oroimen-akatsik ez dute agertzen.
- Funtzio kognitibo globala mantentzen dute 24 ebaki-puntu kopuru altuenarekin Mini Mental State Examination (MMSE – Buruko Egoeraren Mini Azterketa) delakoan [67].
- Oroimen-testean errendimendu-maila normala dute. Puntuazioa 6 baino altuagoa ikaste eta ahozko gaitasuna balioesteko erabiltzen den Buschke-ren Free and Cued Selective Reminding Test delakoan (FCSRT – Oroimen Selektibo Aske eta Kodetuaren Testa) [104].
- Autonomia eguneroko jardueretan. Pfeffer-en jarduera funtzionalei buruzko galdeketan 6 baino gutxiago kalifikazio-eskalan [147].

3.3.1.2 Mild Cognitive Impairment (MCI) taldea

- PGA-n parte hartzeko baimena izenpetu zuten 39 eta 79 urteen arteko gizon-emakumeak.
- Aholkulariak adierazgarritzat jotako oroimen-akatsak.
- Funtzio kognitibo globala mantenduta. 24ko puntuaziorik altuena MMSE delakoan.
- Errendimendua batez bestekoa baina baxuagoa, adina eta hezkuntza-maila kontuan hartuta, oroimena (FCSRT), hizketa (Boston Naming Test [148]), ikuste-espazioari buruzko funtzioa (Judgment Line Orientation [149]), arreta (Letter Number Sequencing [150]) eta funtzio betearazlea (Trail Making Test, Part B [151]). Eskala-puntuazioa 6 edo sei baino gutxiago.
- Autonomia eguneroko jardueretan. Pfeffer-en jarduera funtzionalei buruzko galdeketan 6 baino gutxiago kalifikazio-eskalan.

3.3.1.3 Baztertze-irizpideak

- Dementia agertzea (DSM-IV eta CDR 1).
- Dementziak edo nahasmendu kognitiboak eragin dezaketen edozein motatako gaitz neurologikoaren mediku-historia nabarmena izatea.
- Sintoma psikiatriko arinak (antsietatea, depresioa) dituzten pertsonak ez dira baztertzen.

3.3.1.4 PGA-OREKA datu-basea

Datu-base guztien datu demografikoak 3.2 taulan agertzen dira eta 3.3 taulan PGA-OREKA datu-basearenak ikus daitezke. Azkenean, 100 pertsonaren datu-base orekatua (PGA-OREKA) lortu da. Estatistikoki, adinaren arabera desberdintasun nabarmenak detektatu dira AD eta MCI dutenen artean.

3. KAPITULUA: MATERIALAK

Taula 3. 3: AN esperimentaziorako azpimultzoen datu demografikoak. Honako taulan Animal Naming (AN) atazaren esperimentaziorako aukeratutako azpimultzoen datu demografikoak.

Datu-baseak	Emakumea	Gizona	Adin-maila	Batez besteko adina	SD Adina
PGA-OREKA. AN ataza-azpimultzoa (CR/MCI)	36/21	26/17	39-74/42-79	56.73/57.15	7.8/8.9

3.3.2 Irudien deskripzioari buruzko ataza (PD)

Gipuzkoa-Alzheimer Proiektuak (PGA-k) eta CITA Alzheimer Fundazioak aurrera daramaten ikerketa-pilotu honek barne hartzen ditu Alzheimerri diagnostikatuak duten sei subjektu eta hamabi pertsona osasuntsu (azpidatu-basea MINI-PGA), hau da, laginak 18 pertsonak osatzen zuten, eta demografia-datuak ondoko taulan zehazten dira. Ataza irudi baten ahozko deskripzioan datza [152] 3.2 irudian ikus daitezkeen moduan.

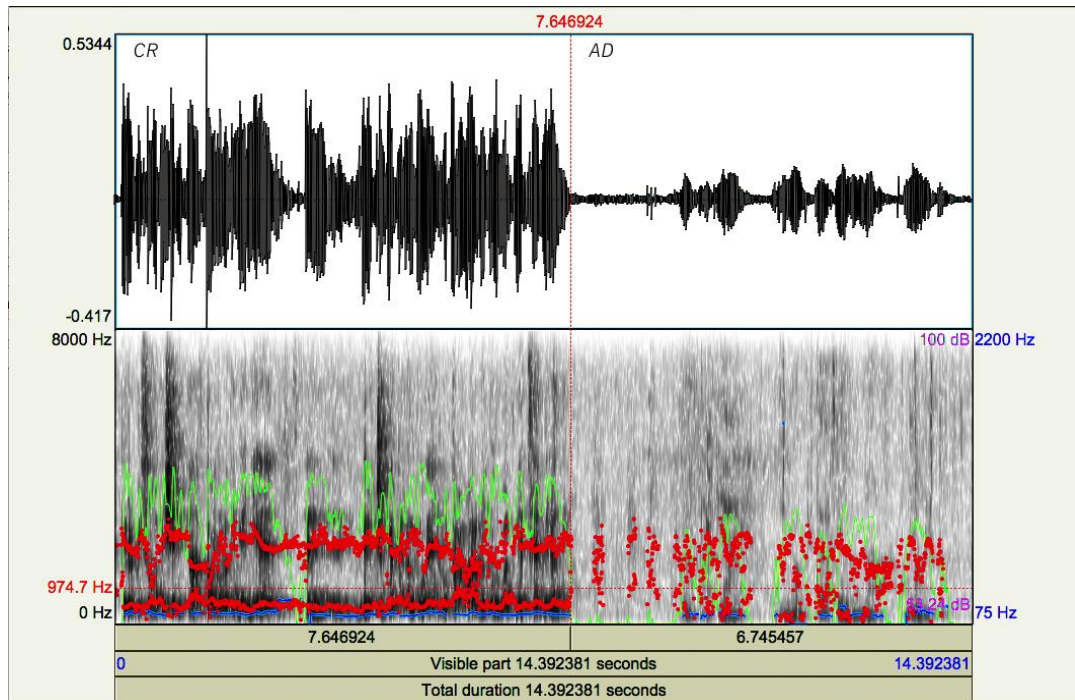
3.3.2.1 MINI-PGA datu-basea

Datu-base guztien datu demografikoak 3.2 taulan agertzen dira eta 3.4 taulan MINI-PGA datu-basearenak ikus daitezke. MINI-PGA, guztira, 18 pertsonak osatzen dute, 13 emakume eta 5 gizonezko:

Taula 3. 4: PD esperimentaziorako azpimultzoen datu demografikoak. Honako taulan Picture Description (PD) atazaren esperimentaziorako aukeratutako azpimultzoen datu demografikoak.

Datu-baseak	Emakumea	Gizona	Adin-maila	Batez besteko adina	SD Adina
MINI-PGA. PD ataza-azpimultzoa (CR/AD)	8/5	4/1	53-73/62-78	61.51/67.97	7.02/5.2

Beste osagai batzuk, adibidez, erantzun emozionala eta emozioen kudeaketa, gaitzaren hasierako estadiotan kaltetuta dauden komunikazio-elementuak izan daitezke, eta baita analisirako faktore erabilgarri bilakatu ere. Osagai horiek antzemateko bidean, PD bezalako atazak lagungarri izan daitezke



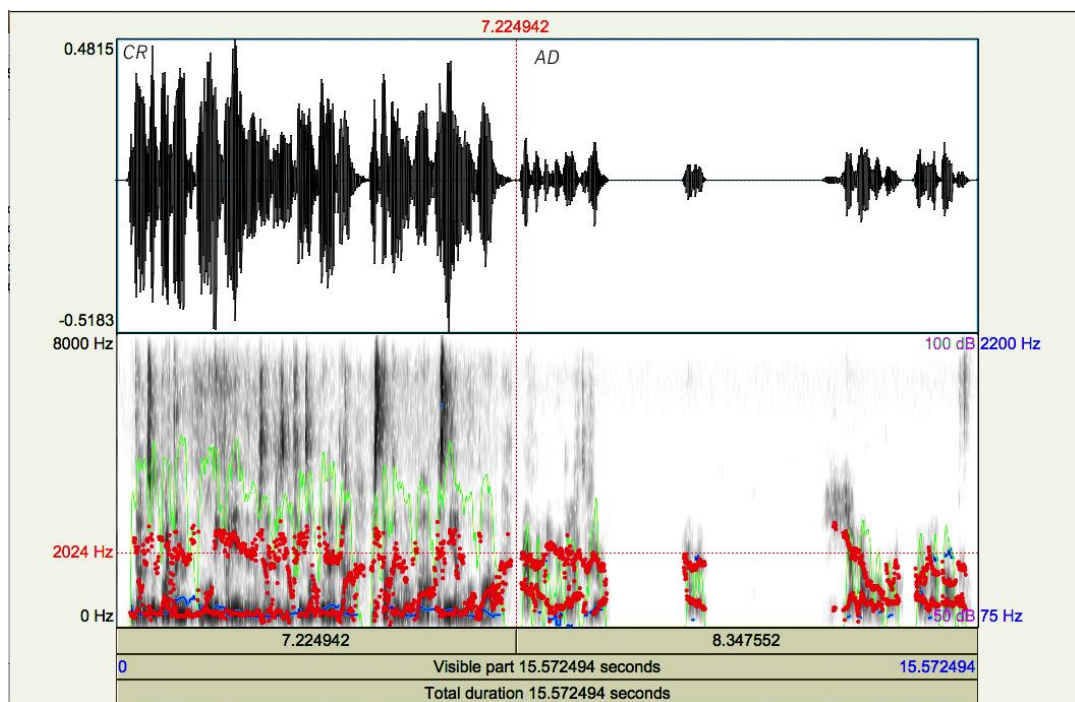
Irudia 3. 3: Hizketa murriztuaren test kliniko II. Irudien deskripzioari (PD) buruzko ataza. Ezkerrean, kontrol-subjektu bat; eskuinean, AD duen pertsona bat. Tonua, lerro urdina; ozentasuna, lerro berdea; formanteak, lerro gorria.

3.3.3 Bat-bateko hizketari buruzko ataza (SS)

AD goiz detektatzeari begira dagoen ikerketa hau etxeko ingurunean egin zen, eta haren helburua AD fase aurreklinikoan (lehen sintomak baino lehenago), eta baita prodromo-fasean ere (hasierako sintoma batzuekin baina demenziarik gabe) identifikatzea da. Lan hau atariko esperimentu osagarria da eta helburu du bat-bateko hizketarekin loturiko biomarkatzaile zenbaiten atalaseak definitzea. Ezaugarrien bilaketa-sistema ebaluazio aurrekliniko baterako diseinatua dago, AD diagnostikatzeko test egokiak definitzeko helburuarekin. Lortutako datuek osatuko dute pertsona bakoitzaren biomarkatzaile-multzoa AD diagnostikatzeko. Lan honetan erabilitako datu-baseak PGA-ren proiektuarekiko independentea da.

Erantzun emozionala eta emozioen kudeaketa, kaltetuta dauden komunikazio-elementuak izan eta informazio baliagarria eman dezaketenez, horiek antzemateko bidean, SS bezalako atazak lagungarri izan daitezke.

3. KAPITULUA: MATERIALAK



Irudia 3. 4: Hizketa murriztuaren etxeko testa. Hizketa espontaneoaren ataza (SS). Ezkerrean, kontrol-subjektu (CR) bat; eskuinean, AD duen pertsona bat. Tonua, lerro urdina; ozentasuna, lerro berdea; formanteak, lerro gorria.

3.3.3.1 AZTIAHO datu-basea

Sexu, adin, hizkuntza eta jatorri sozio-kultural desberdinetako pertsona-multzo zabalarri aplikatzeko metodologia berria sortzeko, datu-base berri bat garatu dugu, kulturanitz eta eleanitza dena, ingelesez, frantsesez, gaztelaniaz, euskaraz, txineraz, arabieraz eta portugesez, 50 pertsona osasuntsuren (12 ordu) eta 20 AD-ren pazienteren, hau da, AD-ren aurretiko diagnostikoa dutenen (8 ordu), bideo-grabaketa daukana. Datu-base osoan, pertsonen adin-tartea 20-98 urtekoa da. Datu-base honen izena AZTIAHO da [19] eta SS atazaren bi laginen (CR/AD) arteko desberdintasuna ikus daiteke 3.4 irudian. Datu-base horretatik ondoko paragrafotan azaltzen den azpidatu-base bat definitu dugu dagozkion datu demografikoekin.

Bideo horietan, bat-bateko hizketak dira oinarria. Bertan, pertsonak istorio atseginak kontatzen zituzten, sentimendu goxoak bergogoratzten zituzten edota haien artean elkarrekintzan zebiltzan adiskidegiroko elkarriketatan. Grabazioen giroa lasaia eta ez-inbasiboa zen. AD dutenekin grabatzeko denborak motzagoak ziren, paziente horien hizketa zailagoa edo traketsagoa baita pertsona osasuntsuena baino, isiluneak luzeagoak eta denbora gehiago behar baitute hitz egokia aurkitzeko eta hizketa-disfluentzia edo mezu hautsiek moldatzeko. Gaitzaren fase aurreratuan, ahalegin hori oso nekagarri zaie eta, askotan, grabazioa eteteko eskatzen dute. Eskaera hori onartu egin zen.

3.3.3.2 Datu-basearen iragazketa

- Bideoa prozesatu egin zen eta audioa wav formatoan (16 bit eta 16 kHz) atera zen.
- Lehen urratsa aztertzeko ezinezkoak diren elementuak kentzea izan zen: barrea, eztula, zarata gogor eta motzak eta hitzunek batera hitz egiten duteneko segmentuak.
- Ondoren, hondoko zarata ezabatu zen denoiser adaptative filtering iragazkiaren bidez.
- Aurreprozesatze horren ondoren, kontrol-taldearen materialaren % 80 eta AD taldearen % 50 baliagarritzat jo ziren azterketarekin jarraitzeko. Datu-baseko hizketa osoa 60 minutukoa da AD taldeari dagokionez, eta 9 ordukoa kontrol-taldeari dagokionez.
- Ondoren, hizketa ondoz ondoko 60 segunduko segmentutan zatitu zen, hitzun guztien segmentu egokiak lortzeko, eta bat-bateko hizketaren 600 segmentuko datu-basea lortu zen.

3.3.3.3 AZTIAHORE datu-basea

Bukatzeko, esperimenduak aurrera eramateko, azpimultzo bat aukeratu da parte-hartzaileen adinarekin eta emoziozko erantzun-mailarekin orekatua dagoena. Azpimultzo honek kontrol-taldeko 20 subjektu hartzen ditu (9 emakumezko eta 11 gizonezko) eta 20 AD paziente (12 emakumezko eta 8 gizonezko). Datu-baseko azpimultzo hau AZTIAHORE deritzo.

Taula 3. 5: SS esperimentaziorako azpimultzoen datu demografikoak. Honako taulan Spontaneous Speech (SS) atazaren esperimentaziorako aukeratutako azpimultzoen datu demografikoak.

Datu-baseak	Emakumea	Gizona	Adin-maila	Batez besteko adina	SD Adina
AZTIAHORE. SS ataza-azpimultzoa (CR/AD)	9/12	11/8	47-80/69-88	57.80/79.90	11.92/5.3

3.3.3.4 AZTITXIKI datu-basea

Hasierako etapan, atariko esperimentazioa AZTITXIKI multzoan oinarrituta burutu da. AD sintomekin zerikusia duten bat-bateko hizketaren ezaugarri zuzenak aztertu dira. Esperimentua, alde batetik, kontrol-taldea (CR) ezaugarritzen duten SS-n izaten diren aldaketak eta ezaugarriak antzemateko diseinatu da. Bestetik, AD-ren maila desberdinetako aldaketak eta ezaugarriak antzemateko ere diseinatu da.

Datu-base honen azpimultzoa AZTIAHORE bera da. Hasierako esperimendu honetan, AZTITXIKI izeneko datu-basearekin egin da esperimentazioa. AZTITXIKI,

3. KAPITULUA: MATERIALAK

AZTIAHOREren azpimultzo bat da. Kontrol-taldeko 5 kidek eta AD diagnostikatuta duten baina gaitzaren faseetan (ES = 1, IS = 2, AS = 2) dauden 5 erik osatzen dute taldea. Kontrol-taldeko kideak tarteko adina (M) edo adin nagusikoak (E) diren 2 emakumezko eta 3 gizonezko osatzen dute. AD diagnostikatutako pertsonak 2 emakumezko eta 3 gizonezko dira eta guztiak adin nagusikoak (E).

Taula 3. 6: AZTITXIKI azpimultzoa.

Datu-baseak	Emakumea	Gizona	Adin-maila	Batez besteko adina	SD Adina
AZTITXIKI. SS ataza-azpimultzoa (CR/AD)	2/2	3/3	80-90/75-95	85.70/87.60	10.2/10.5

4. KAPITULUA: METODOAK

4.1 Sarrera

Atal honetan erabilitako metodoetan sakonduko da eta hobeto ulertzeko **B.1, B.2, B.3, C.1, C.2, C.3, C.4, D.1, D.2** eta **D.3** eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

AD-ren bereizgarri klinikoak eta adierazle goiztiarrena oroimen episodikoaren eta hizketaren nahasmenduak dira. Etapa honetan, ordea, gaixoa bere aurreko errutina eta jarduerekin jarraitzeko gai da [8].

Sintoma klinikoak agertzen direnean, gehienetan beste zenbait defizit kognitibo ere azaltzen dira, funtzio beterrazleen, orientazioaren eta pertzepzioaren gaitasunei dagokienez. Horiekin batera, zenbait sintoma konduktual eta psikologo hauteman daitezke: apatia, suminkortasuna, depresioa, antsietatea, eldarnioak, haluzinazioak, desinhibizioa, oldarkortasuna, jarduera motorretan desgaitasunak, irensteko zailtasunak eta jateko eta lo egiteko ohituretan aldaketak [153][24].

Hortaz, narriadura kognitibo arinaren (MCI-ren) hautemate egokia benetako erronka da komunitate zientifikorako Alzheimer's Disease-ren (AD) eragina arindu nahi bada. Azken hamarkadetan, aurrerapen baliagarriak izan dira hala balioztatze-teknika klasikoetan nola screening edo bahetze metodologia ez-inbasiboetan. Metodo horien artean, mintzamenaren analisi automatikoa (AD dutenengan endekatzen den trebetasunetako bat) oso tresna merke eta erabilgarria da diagnostikatzerakoan.

Dudarik gabe, aurreko screening sakona ezinbestekoa da eta monitorizazio-teknika adimendun ez-inbasiboak oso tresna baliagarriak izan daitezke gaitzaren hautemate goiztiarrerako, diagnostikoa osatzeko edota gaixotasunaren garapena balioztatzeko.

Test horiek ez dute langile espezializaturik behar, ezta laborategi-lanabes berezirik ere; hortaz, pazientearen senide eta inguruak, behar bezalako prestakuntza jasotzerako, horiek egiteko gai izan daitezke, erriaren gaitasunak aldatu edo blokeatu beharrik gabe [2][87][19]. Are gehiago, teknika horiek kostu txikikoak dira eta ez dute azpiegitura estentsiborik behar ezta medikuntza-baliabideak eskura izatea ere. Ondorioz, gauza dira informazioa modu errazean, azkarrean eta kosturik gabekoan [2][87][19] kudeatzeko.

4. KAPITULUA: METODOAK

Metodologia horiek erraz aplikatzen dira, merkeak dira eta protokolo mediku estandarretan txerta daitezke eguneroko atazak aldatu gabe. Bi ezaugarri dituzte:

- Gai dira modu ez-inbasiboan informazioa erraz, azkar eta merke emateko [87][19].
- Pazientearen ohiko inguruneko teknikariak ez diren pertsonak erabil ditzateke teknika horiek pazienteen trebetasunak aldatu edo blokeatu gabe, gaixoei ez baitituzte estresagarritzat jotzen. Oso baliagarriak dira hala balioztatze klinikorako nola etxeko/hasierako arretarako.

Alabaina, teknika horiek gaixoentzat estresagarriak ez direla uste bada ere, zenbaiten iritziz gerta daiteke zaintzaile ez-profesionalak ez izatea medikuntza-langileen doitasuna AD, Narriadura Kognitibo Arina (MCI) edota beste zenbait gaitz garaiz eta zehatz identifikatzeko.

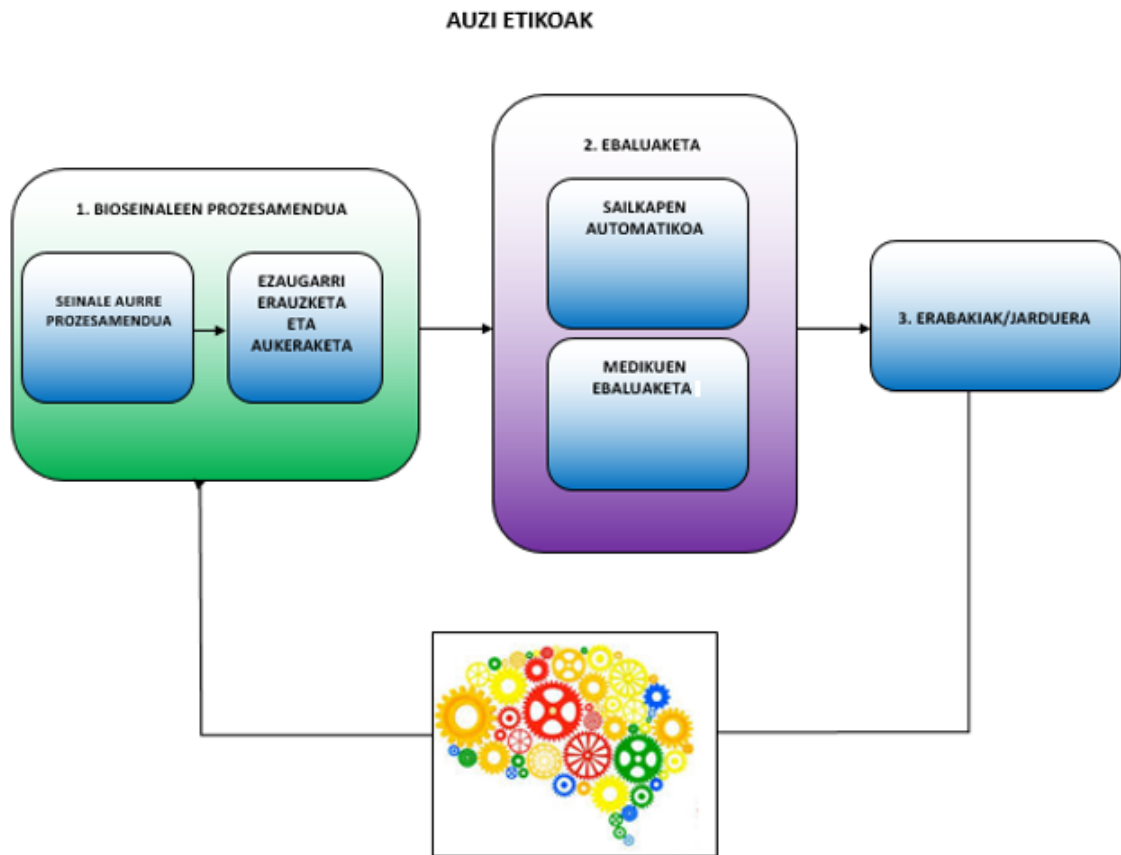
Metodo ez-inbasiboen artean, mintzamenaren edo hizketaren analisi automatikoak (MCI/AD dutenengan endekatzen den gaitasun edo trebetasun lehenetako bat) oso tresna sendoa, naturala eta atsegina eskaintzen du.

Gizakiengan, ahozko komunikazioa da adierazpide garrantzitsuenetako bat eguneroko bizitzan izaten den gizarte-elkarrekintzarako. Prozesu konplexua da. Izan ere, berekin dakar trebetasun kognitiboen sorta oso zabala, zeinaren bidez gai garen mezu txertatua duten soinuak sortzeko. Bioseinaleak barne hartzen ditu gure desioak, ideiak edota emozioak, hots, gure izatasunaren zati bat. Prozesu horretan, garunak ezagutza, oroimena, hizkuntza eta informazio semantikoa kudeatu behar ditu.

Balioeste kognitiboak testen sorta zabala inplikatzeko du, zahartze-prozesu anormal baten hasierako faseak kuantifikatu eta bereizteko [66]. Balioeste horretan, mintzamina eta hizketa analisiaren oso osagai garrantzitsuak dira, giza komunikazio-prozesuaren giltza-faktoreak baitira. Hortaz, komunikazio-trebetasunen andeatzea testen bidez neurtzen da, honako elementu hauek kontuan hartuta, alegia, oroimenaren nahastea, izendatzea, gogoraztea, hitz-jarioa edota kontrol emozionala [5][6][7]. Berrikitan, zenbait ikerketa berri konplexutasun handiko egoerak aztertu dituzte, hizketa ez-murriztuko atazetan edo ariketetan oinarrituta (bat-bateko hizketa, elkarrizketak edota hizketaldi librea) [38].

Bestalde, test horien lagin biologiko horiek badute abantaila bat, alegia, erraz jaso daitezke tresneria inbasibo eta garestirik gabe. Baina ez-inbasiboak izateaz gain, teknika horiek, automatikoak eta erabilerrazak izan behar dute. Lagina erauzitakoan, teknologia horiek gai izan behar dira gaixoak ikusiko ez duen prozesu guztia kudeatzeko: laginak ezaugarritu eta aberastu, diagnostikoak egin eta komunikazioak ahalbidetu, besteak beste.

Testuinguru horrela izanda, hizketaren analisi automatikorako ez-inbasiboak diren teknologiak garatzea ezinbestekoa da, eta lan honek horren aldeko ekarpena izan nahi du.



Irudia 4. 1: Sistema eta tresna automatikoak. Duela gutxi arte, espezialistek eskuz egiten zituzten testak; alabaina, azken urteotan, sistema eta tresna automatikoak gero eta gehiago erabiltzen hasi dira baita jardura klinikoan ere. Testuinguru hori kontuan hartuta garatu da sistema ez-inbasiboa eta seinalea jasotzen denetik erabakia hartzen denerako prozesu osoa biltzen du: (1) Grabazio automatikoa, prozesatzea, ezaugarrien erauzketa eta hautaketa. (2) Sailkapen automatikoa medikuaren ikuskapenarekin eta ebaluketarekin batera. (3) Berrelikadura.

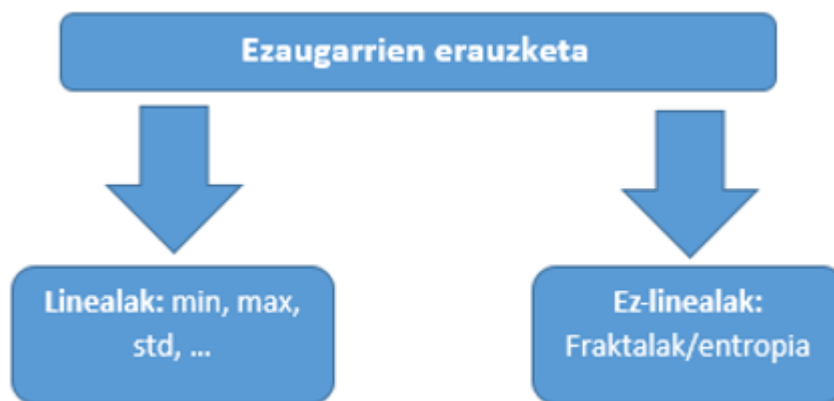
Azkenean, proposatzen den diagnostiko/neuroerrehabilitaziorako kontrol-sistemaren diagrama orokorra 4.1 irudian ikus daiteke. Eskeman, behar diren ezaugarri guztiak garatu ahal izateko, helburuetan oinarritutko hiru bloke nagusi aurki daitezke:

- Seinalearen prozesatzea (seinalearen aurreprozesatzea eta ezaugarrien erauzketa eta hautaketa).
- Azterketa/diagnostikoa (eredu automatikoak, test estatistikoak eta medikuen metodologia estandarrak).
- Erabakiak eta eragiteko elementuak.

Bigarren blokean eta hirugarrenean ez dira soluzio itxiak planteatuko baizik eta sistema malgua non osasun-espezialistak, azken erabiltzaileak eta laguntzaileak zenbait elementurekin elkar eragiteko aukera edukiko duten. Azken hori diseinu unibertsaleko kontzeptuan oinarrituta dago [154].

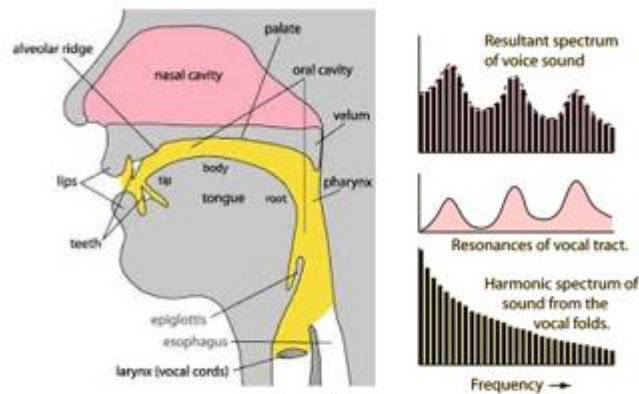
4.2 Ezaugarrien erauzketa eta hautaketa

Hizketa-eredu arruntenetako batzuen erabileran oinarritzen gara, hizketa osasuntsua eta patologikoa bereizteko [38][56][155] eta giza pertzepzioaren bidez diskriminatzeke. Haietako gehienak ezagunak dira hizketa-seinaleen prozesatzeko arloan. Parametro bakoitzerako deskripzio zehatzagoa eta gainerako informazioa aurrerago emango da. Atazen grabazioen datu-basea bildu eta gero, horiek automatikoki aztertzen dira ezaugarriak erauzteko helburuarekin. Bi ezaugarri-mota barne hartuko ditu azterketa automatikoak: linealak eta ez-linealak.

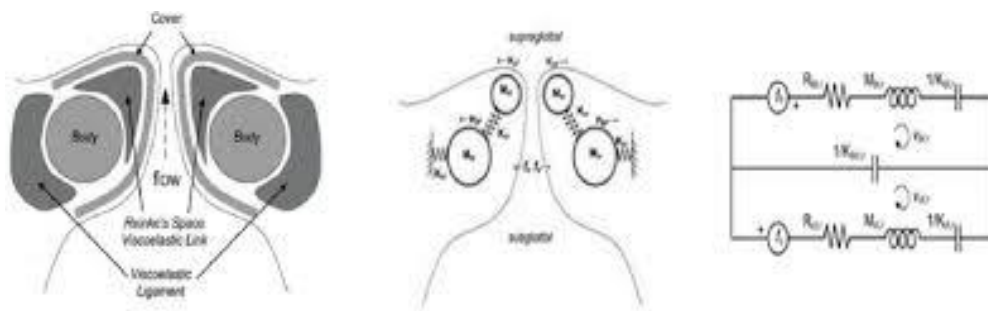


Irudia 4. 2: Ahots-laginen erauzketa linealak eta ez-linealak. Sistema automatikoak ahots-laginen ezaugarri-bektorea sortzeko erauzketa linealak eta ez-linealak biltzen ditu sistemak. Mintzamen automatikoaren analisis oinarritutako multi-hurbilketa ez-linealak hizketa-analisan erabiltzen diren denbora-serieen azterketarako eta denbora-serie dinamikoaren konplexutasunaren neurtze kuantitatiborako baliagarriak izan daitezke.

Bestalde, hurbilketa hori hizketa-patologiaren deskripzioan oinarritzen da, alegia: fonazio, artikulazio, hizketa-kualitate, giza pertzepzio eta sistemaren dinamika konplexuetan.



Irudia 4. 3: Fonazio-aparatua [156]. Fonazioa, fonazio eta arnas aparatuek parte hartzen duten ekintza fisikoa da. Ezkerreko irudian, aparato honetan parte hartzen duten elementuak ikus daitezke. Eskuineko irudian fluxuak traktoan nola eragiten duen eta ondorioz sortzen den ahotsa ikus daitezke.



Irudia 4. 4: Fonazio-aparatua II. Fonazioaren eskema biomekanikoa ikus daiteke irudi honetan [39][157].

Modelatzeari buruzko ikuspuntua guztiz aldatu da eta fonazioa, artikulazioa, hizketaren kalitatea, giza pertzepzioa eta sistemaren dinamika konplexuaren arabera egiten da hizketaren patologiaren deskripzioa. Izan ere, hasieran, ahots patologikoek aldaketa txikiak eta ez-linealtasun finak dituzte. Beraz, ezaugarrien hautaketa ezinbestekoa da diagnostiko zehatza burutzeko. Aipatutako hizketaren azterketarako modelatze- ikuspuntu berri honek hizketa eta disfluentziak modelatzeko zenbait ezaugarri-mota txertatzen ditu: linealak eta ez-linealak.

4. KAPITULUA: METODOAK

Taula 4. 1: Ezaugarrien inbentarioa.

Espektro-arloko ereduak	Harmonizitatea, harmoniko/zarata ratioa (HNR), tonua (pitch), jitter, shimmer, anplitude-nahasmenduaren koizientea (APQ, amplitude perturbation quotient), espektro-zentroidea eta formanteak [38][56][158].
Ezaugarri klasikoak: Denbora domeinuaren ereduak	Ahostun/ahoskabe segmentuak, isiluneak, zero-crossing rate (ZCR, zero puntutik igarotzeko indizea) [38][56][159] eta horien aldaerak.
Ezaugarri klasikoak: Dinamikoak	Energia, intentsitatea, epe motzeko energia, anplitudea, intentsitatea [56][160][161].
Hautemate-ezaugarriak	Giza entzumenak iragazkiek bezala funtzionatzen du, alegia, frekuentzia-osagai jakin batzuetan kontzentratzen da soil-soilik. Iragazki horiek modu ez-uniformean kokatzen dira maiztasun-ardatzean. Iragazki gehiago daude frekuentzia apaleko eremuan eta iragazki-kopuru txikiagoa goi-frekuentziako eremuan. Hori dela eta, Mel eskalako maiztasunen analisia eta espezifikoki Mel Eskalako Koefiziente Zepstralenena (MFCCrena) errendimendu-mota hori simulatzera bideratuta daude [56][160][161].
Ezaugarri aurreratuak	Modulation Spectra Coefficients (MSC) [160]. Perceptual Linear Predictive Coefficients (PLP) [162]. Linear Predictive Cepstral Coefficients (LPCC) [162]. Linear Predictive Cosine Transform Coefficients (LPCT) [66]. Adaptive Component Weighted Coefficients (ACW)[162]. Inferior Colliculus Coefficients (ICC) [163].
Ezaugarri ez-linealak edo Non-Linear Features (NLF)	Dimentsio Fraktala, Shannon-en Entropia eta Permutazio Entropia Multieskalarra kalkulatu dira [38][19][164].

4.2.1 Ezaugarri linealak

Hizkera osasuntsu eta patologikoa zein giza pertzepzioaren bidezko diskriminazioa bereizteko gehien erabiltzen diren ezaugarrietako batzuk erabiliko dira 4.1 Taulan [38][56][157][165]. Horietatik gehienak oso ezagunak dira ahots-seinaleen prozesatzearen eremuan. Ezaugarri guztiak gure ikerketa taldean garatutako software [38] [19], MATLAB [166], Praat [167] eta Alzumeric-en bidez kalkulatu dira.

- Ezaugarri aurreratuaren blokean ezaugarri berriak ageri dira. Besteen artean:
- Adaptive Component Weighted coefficients (ACW): kanalen distortsioa gutxiago nabaritzen dute.

- Modulation Spectra coefficients (MSC): informazio osagarria ematen diete MFCC koefizientei.
- Perceptual Linear Predictive (PLP) coefficients: intentsitate bolumen handienekin egokitzen kontuan hartzen dute.
- Inferior Colliculus coefficients (ICC): beheko Colliculusen eredu bioinspiratua erabiltzen dute.

4.2.2 Ezaugarri ez-linealak

Beraz, mintzamen-konplexutasuneko maila desberdinak dituzten hiru atazen azterketatik abiatuta, mintzamen automatikoaren analisisian oinarritutako multihurbilketa ez-lineala egin da:

Dimentsio fraktala da sistema baten konplexutasuna ezaugarritzeko erabiltzen den neurketa erabilienetako bat. Sistema fraktal gehienek autoantzekotasuna deituriko ezaugarria dute. Objektu bat autoantzekoa da baldin eta bere hurbileko azterketak agerian uzten badu bere buruaren bertsio txikiagoz osatuta dagoela. Autoantzekotasuna kuantifika daiteke patroia bat osatzen duten oinarritzko osagai-kopuruaren neurtze erlatibo gisa, eta neurtze horri Dimentsio Fraktala deitzen zaio. Higuchi [168] eta Castiglioni [169] algoritmoak oso egokiak dira, besteak beste, hizketa-analisisan erabiltzen diren denbora-serieen azterketarako.

Permutazio Entropiak eta Permutazio Entropia Multieskalarrak kontuan hartzen dute denbora-seriek daukaten denbora-informazioa. Gainera, sinpleak eta sendoak dira eta kostu konputazionalaren aldetik kostu txikiak [170][171][172]. Bandt-ek eta Pompe-k [173] metodo sinple eta sendo bat ezartzen dute, Shannon-en entropia-neurketa delakoan oinarrituta. Horrek denbora-kausaltatea hartzen du kontuan, denbora-serie bateko antzeko balioak konparatzea dela medio. Laburbilduz, Permutazio Entropiak zerikusia du denbora-serieen ordenaren egitura lokalarekin, zeinak konplexutasunaren neurtze kuantitatiboa eman dezakeen denbora-serie dinamikoetarako.

4.2.2.1 Fraktalak

Hizketa-seinaleen kuantifikazioa egin eta Alzheimerren sintomak hautemateko Dimentsio fraktalak (FD) erabilgarriak izan daitezke. López-de-Ipiña *et al.*ek [19][21] FD-*ei* buruzko honako deskribapena egiten dute:

Dimentsio fraktalak (FD) denbora-seinale baten garratzasuna kuantifikatu egiten du sinusoide-seinale konplexuaren eta zarata zuriaren artean eta bere askatasun-maila balioesten du, eta hortaz, bere fluktuazioak modelizatzeko tresna baliagarria izan daiteke. Hizketa-aerodinamiketako eremutik, Maragos-ek eta Potamianos-ek motibazioa eta justifikazioa eskaintzen dituzte dimentsio fraktala erabiltzeko turbulentsia-maila

kuantifikatzeko hizketa-seinaleetan [174]. Aldi berean, jardunbide artikulatzaileetan eta ozentasunean dauden desberdintasunak aztertzen dituzte. Ikerketa horretan, Dimentsio Fraktala (FD) eraginkorragoa da *zero-crossing*ak (zero puntutik igarotzeko indizeak) baino. Bioseinaleetan (EEG eta MEGetan) oinarritutako AD-ren ahoskerari buruzko beste lanean bezala [175], metodologia hori erabilgarria izan daiteke Alzheimerren sintomak hautemateko, besteak beste, zalantzak, etenaldiak edota isiluneak hizketa-seinaleetan.

Hurbilketa horrekin sistema hobetu nahi da [19][23], ereduak handiagotzen FD dela bide. FD oso eredu ezaguna eta erabilia da sistema baten konplexutasuna deskribatzeko, eta lagungarri izan daiteke diagnostiko goiztiarrerako erabilgarriak diren aldaketa ñimiñoak detektatzeko. Bestalde, eredu hau gai da sistemaren dinamikak antzemateko eta, hortaz, bariazio adierazgarriak agerian jarri ditzake hizketa-seinaleetan. Zehatzago esateko, Higuchi-ren, Katz-en eta Castiglioni-ren algoritmoen [168][176][169] inplementazioak erabiltzen dira eredu berri hau modeloaren entrenatze-prozesua elikatzen duen multzoari gehitzeko.

FD-ren balioak izan beharko lukeen uhin-forma jakin bati buruzko erreferentzia zehatzik ez dago. Gainera, hizketaren uhin-formak ez dira geldikorrak. Hori dela eta, Hizketa Ezagutze Automatikoaren (Automatic Speech Recognition, ASR) teknika gehienek seinalearen zatiketa motzak erabiltzen dituzte uhin-formatik ezaugarriak ateratzeko. Horrek esan nahi du, fonema desberdinak ezagutu ahal izateko, hizketaren uhin-formatatik ezaugarriak ateratzeko teknika onargarri bat izan daitekeela seinalea zati txikitzen banazea eta zati bakoitzerako ezaugarriak kalkulatzeko.

Beste hitzetan, uhin-formaren segmentu motzen dimentsio fraktala kalkulatu daiteke eta seinale osotik lortutako balioen eboluzioa behatu. Helburua, ahozko mezuaren osagaiak identifikatzen lagungarri izan daitezkeen ezaugarri fraktalak aurkitzea.

4.2.2.1.1 Higuchi-ren Dimentsio Fraktala

Deskribapena egiten da [21] lanean. Dimentsio fraktala neurtzeko badira zenbait algoritmo eta artikulua honetan [21], arreta denbora-serieen analisirako bereziki egokiak diren aukeretan jartzen du arreta, sistemaren aurretiko modelatzea eskatzen ez dutenetan. Algoritmo horietako hiruk beren egileen izena daukate, Higuchi (1988), Katz (1988) eta (2010). Higuchi eta Castiglioni aukeratu dira baliabideak urriak direnetan [177][178] zehatzagoak omen direlako nahiz eta Katz ere algoritmo sendotzat hartzen den dimentsio fraktala kalkulatzeko [179].

Higuchi-k (1988) denbora diskretuko sekuentzien dimentsio fraktala neurtzeko eta denbora-serieetatik zuzenean $x(1), x(2), \dots, x(N)$ hartzeko algoritmo bat proposatu zuen. Algoritmoa, jatorrizkotik eraikitako denbora-serie berri batean, x_m^k oinarritzen da:

$$x_m^k = \left\{ x(m), x(m+k), x(m+2k), \dots, x\left(m + \left\lfloor \frac{N-m}{k} \right\rfloor k\right) \right\},$$

for $m = 1, 2, \dots, k$

(4.1)

non $[a]$ -k a -rako zati osoa, m -k hasierako denbora-balioa eta k -k bi puntuen arteko denbora-tarte diskretua adierazten duen. x_m^k eraikitako denbora-serie bakoitzerako, $L_m(k)$ batez besteko luzera horrela konputatzen da:

$$L_m(k) = \frac{\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |x(m+ik) - x(m+(i-1)k)| (N-1)}{\lfloor \frac{N-m}{k} \rfloor k}$$
(4.2)

Ondoren, kurbaren luzera k denbora-tarterako definitzen da batura-balio bezala $L_m(k)$ ren k multzoen gainean, 3. ekuazioan erakusten den moduan:

$$L(k) = \sum_{m=1}^k L_m(k)$$
(4.3)

Azkenik, $\frac{\ln(L(k))}{\ln(1/k)}$ kurbaren malda kalkulatzeko da karratu minimoetako doitze lineal optimoa erabilita, eta emaitza Higuchi Dimentsio Fraktala da (HFD).

4.2.2.1.2 Katz eta Castiglioni-ren Dimentsio Fraktala

Bestalde, Katz-ek (1988) proposatu zuen dimentsio fraktalaren formula normalizatu bat (ikus 3. Ekuazioa) dimentsio fraktal hori neurtze-unitate partikular horien mende baitago.

Kurba plano baten FD delakoa horrela definitu daiteke:

$$FD = \frac{\log_{10}(L)}{\log_{10}(d)}$$
(4.4)

Kasu horretan, L kurbaren luzera da eta d kurbaren diametroa (hedadura plano). Denbora-serietan, zeinak puntu bikoitien multzo ordenadunak diren (x, y) , L luzera osoa ondoz ondoko puntuen arteko distantzien batuketa baino ez da 5. ekuazioan ikus daitekeen bezala:

$$L = \sum_{i=1}^n l_{i,i+1}$$
(4.5)

non $l_{i,j}$ adierazten duen bi puntu bikoitien (x_i, y_i) eta (x_j, y_j) arteko distantzia euklidiarra 5. ekuazioan agertzen den moduan, eta $n+1$ puntu bikoitien kopuru osoa den.

$$l_{i,j} = \sqrt{(y_i - y_j)^2 + (x_i - x_j)^2} \quad (4.6)$$

Denbora-serieen diametroa kalkulatu daiteke abiapuntuaren eta denbora-serieen edozein beste puntutako distantziaren (i puntua) arteko distantziarik urrutiena bezala, 6. ekuazioan adierazten den moduan:

$$d = \max\{l_{1,i}\} \quad (4.7)$$

FD-k konparatzen du kurba batean une horretan dagoen unitate-kopurua hedadura espazial bereko eredua erreproduzitzeko behar den unitate-multzo txikienarekin; eta, arestian esan bezala, horrek erabilitako neurtze-unitateekin zerikusia izango du, FD kalkulatzeko baita espazio diskretu batean eta ez jarraituan.

Katz-ek irtenbidea ematen dio arazo horri d eta L , a batez besteko urratsaren luzerakin normalizatuz, ondoz ondoko puntuen arteko batez besteko distantzia, $\frac{L}{n}$ moduan definitua, non $n + 1$ puntu bikoitien guztizko kopurua den 5. ekuazioan agertzen den bezala. Hortaz, (4)n distantziak normalizatzean, Katz-en hurbilketa lortu dezakegu denbora-serie bateko FD kalkulatzeko (8):

$$FD = \frac{\log_{10}(n)}{\log_{10}(n) + \log_{10}\left(\frac{d}{L}\right)} \quad (4.8)$$

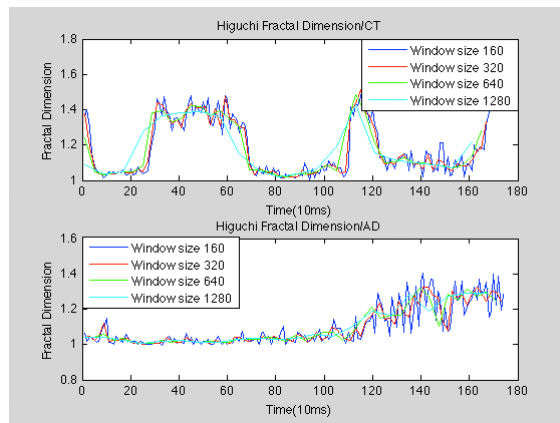
Hala ere, Castiglioni-k (2010) esaten du $X - Y$ planoan, zeinak uhin-forma deskribatzen duen, erabiltzen diren x eta y magnitudeak intrintsekoki desberdinak direla bat etortzen baitira seinalearen magnitudearekin (y) eta denbora-puntuekin (x). Ondorioz, kontuan izanik input seinalea uhin-forma monodimentsionala dela, luzera eta hedadura berridatzi daitezke Mandelbrot-en hurbilketa dela medio. Hori egiteko bide erraza eta eraginkorra zera da, alegia, bi magnitude horiek zuzenean neurtzea bakoitza bere dimentsioan. Bakoitzerako, y -ardatzaren gaineko hedadura y_k ren heina da:

$$d = \max\{y_k\} - \min\{y_k\} \quad (4.9)$$

eta L luzera gehikuntza guztien batura da, modulutan:

$$L = \sum_{k=1}^n |y_{k+1} - y_k| \quad (4.10)$$

Non, beste behin ere, d da kurbaren diametroa, L kurbaren luzera eta $n + 1$ puntu bikoitien guztizko kopurua.



Irudia 4. 5: Higuchi-ren dimentsio fraktala. Higuchi-ren dimentsio fraktala (HFD) kontrol-subjektu bati dagokionez (goian) eta AD duen subjektu bati dagokionez (behean), leiho-tamaina desberdinetarako.

4.2.2.2 Entropia

Entropiak desordena-maila neurtzen du sistema fisikoetan eta, horrez gain, egoera espezifikoaren interpretazio asko dituen oinarrizko kantitatea da. Desordenarekin, egoera-espazioaren bolumenarekin edota informazio-faltarekin lotu izan dute [170][171][180][181]. Maiz, Shannon-en entropia modu klasiko eta naturalentzat jotzen da seinale batean informazioaren aurreikusitako balioa (babezbestekoa) neurtzeko [172][181][182][183]. Richmanek *et al.*ek entropia hori aztertzen dute, eta sistema dinamikoaren arabera izaten da informazio-produkzioaren tasa [184]. Egile batzuek azpimarratzen dute entropia kalkulatzeko datu-multzo oso luzeak eskatzen dituela eta hori, biomedikuntzari dagokionean, oso zaila edo ezinezkoa dela. Bestaldekik, denbora-serie batek adierazitako sistema baten entropia kalkulatzeko metodoak ez dira egokiak biomedikuntza-ikerketetan datu motz eta zaratatsuak aztertzeko [170][184]. Ondorengo azpiataletan, entropia kalkulatzeko erabilitako proposamenak aurkeztuko dira.

4.2.2.2.1 Shannon-en Entropia

X ausazko aldagai diskretu bakar baten entropia $H(X)$ bere batez besteko segurgabetasunaren neurria da. Shannon-en entropia [181][183] honako ekuazio honen bidez kalkulatzen da:

$$H(X) = - \sum_{x_i \in \Theta} p(x_i) \log p(x_i) = - E[\log p(x_i)], \quad (4.11)$$

non X -k adierazten baitu ausazko aldagai bat, Θ balioen multzo bat eta probabilitate-masaren funtzioa dituen $p(x_i) = P_r\{X = x_i\}$, $x_i \in \Theta$, eta E -k itzaropen-eragilea. Kontuan izan $p \log p = 0$, $p = 0$ bada.

Prozesu estokastiko baten irteera adierazten duen denbora-serie bati dagokionean, hau da, n ausazko aldagaien sekuentzia indexatuak, $\{X_i\} = \{X_1, \dots, X_n\}$, hurrenez hurren $\theta_1, \dots, \theta_n$, balioen multzo bat eta $X_i \in \theta_i$, dituenari dagokionean, honako honek definitzen du entropia konposatua:

$$H_n = H(X_1, \dots, X_n) = - \sum_{x_1 \in \theta} \dots \sum_{x_n \in \theta_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n), \quad (4.12)$$

non $p(x_1, \dots, x_n) = P\{X_1 = x_1, \dots, X_n = x_n\}$ da probabilitate konposatua n aldagaietarako, X_1, \dots, X_n .

Ekuaioari katearen araua aplikatzen badiogu (2), entropia konposatua entropia baldintzatuen batura moduan idatz daiteke, non haietako bakoitza kantitate ez-negatiboa den:

$$H_n = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1), \quad (4.13)$$

Hortaz, entropia konposatua n -ren funtzio hazkorra da. Entropia konposatua n -rekin hazten den tasa, hau da, h entropia-tasa, horrela definitu daiteke:

$$h = \lim_{n \rightarrow \infty} \frac{H_n}{n}, \quad (4.14)$$

4.2.2.2 Gutxi gorabeherako Entropia vs Lagin-Entropia

Gutxi gorabeherako entropia neurri estatistiko bat da, zeinak behin-behineko interferentzia moteltzen duen eta zarataren eragina saihestu dezakeen algoritmoaren parametroak behar bezala jarrita. Erabili daiteke seinale estokastikoetan eta baita seinale deterministetan ere [185][186]. Hori erabakigarria da biologia-seinalei dagokienean, zeinak biologia-sare konplexuen irteerak diren eta deterministikoak, estokastikoak edo biak batera izan daitezkeen. Gutxi gorabeherako entropiak seinaleen irregulartasunaren neurketa eskaintzen digu, ereduarekiko independentea den neurketa. Algoritmoak denbora-serie bat zenbaki ez-negatibo batean sintetizatu egiten du, non balio handiagoek sistema irregularragoak adierazten dituzten [185][186].

Metodoak aztertzen ditu denbora-serieak antzeko mementuetarako [187]: mementu sarriagoek eta antzekoagoek gutxi gorabeherako entropiaren balio txikiagoak dituzte. $ApEn(m, r, n)$ honako probabilitate baldintzatu hau neurtzen du, alegia, m

puntuatarako antzekoak diren n luzera duten bi segidek antzeko izaten jarraitzen dute hurrengo lagin-puntuak, r tolerantzia baten barruan. Hortaz, gutxi gorabeherako entropiaren balio baxuak erregulartasunaren maila altua islatzen du. Gutxi gorabeherako entropiaren algoritmoak segida bakoitza kontuan hartzen du bere burua egokitu balu bezala errakuntza-tartea gutxiagotzeko; lagin-entropia $\text{SmEn}(m, r, n)$ garatu zen autoegokitze horiek kontuan ez hartzeko.

Lagin-entropia definitzen da n puntutako denbora-serie baterako. Hasteko, $\mathbf{x}_m(i) = \{u(i+k): 0 \leq k \leq m-1\}$ $n-m+1$ bektoreak definitzen dira $u(i)$ -tik $u(i+m-1)$ -rako m datu-puntuak bektore gisa. Horrelako bi bektoreren arteko distantzia $d[\mathbf{x}_m(i), \mathbf{x}_m(j)] = \max_k \{|u(i+k) - u(j+k)|: 0 \leq k \leq m-1\}$, *i.e.*, bezala definitzen da, haien osagai eskalarren arteko diferentzia maximoa. $\text{SmEn}(m, r, n)$ lagin-entropia horrela definitzen da:

$$\text{SmEn}(m, r, n) = \lim_{n \rightarrow \infty} \{-\ln(A^m(r)/B^m(r))\} = -\ln(A/B), \quad (4.15)$$

non

$$A = [(n-m-1)(n-m)/2]A^m(r), \quad (4.16)$$

eta

$$B = [(n-m-1)(n-m)/2]B^m(r). \quad (4.17)$$

$B^m(r)$ bi sekuentzia m puntutan egokitu daitezen probabilitatea da:

$$B^m(r) = (n-m)^{-1} \sum_{i=1}^{n-m} B_i^m(r), \quad (4.18)$$

non $B_i^m(r)$ da $(n-m-1)^{-1}$ bider $\mathbf{x}_m(j)$ bektore kopurua $\mathbf{x}_m(i)$ ren r tolerantzia baten barruan. Modu berean, $A^m(r)$ da bi sekuentzia $m+1$ puntutan egokitu daitezen probabilitatea:

$$A^m(r) = (n-m)^{-1} \sum_{i=1}^{n-m} A_i^m(r), \quad (4.19)$$

non $A_i^m(r)$ da $(n-m-1)^{-1}$ bider $\mathbf{x}_{m+1}(j)$ bektore-kopurua $\mathbf{x}_{m+1}(i)$ ren r tolerantzia baten barruan. Ikerketa honetan, [187]ean: $m = 2$ eta $r = 0.2$ aholkatutako parametroak erabili dira, eta iturrietako desbideratze estandarra 1ean normalizatu egiten da. Lagin-entropia konplexutasunaren ahalmen handiko zenbatzailea da, adibidez, elektroentzefalografiaren (EEG) seinalei dagokienean [188], eta erabili daiteke markatzaile moduan tresnen presentziarako EEG irarketetan [189]. A/B kantitatea probabilitate baldintzatua da non r tolerantziako puntuen barruko bi sekuentzia elkarrekiko r mantendu daitezen ondorengo puntuak. Gutxi gorabeherako entropiak ez bezala, zeinak probabilitateak txantilo moduan kalkulatu dituen, lagin-entropiak

probabilitate baten logaritmo negatiboa kalkulatu du oro har hartutako denbora-serie batekin lotuta.

4.2.2.2.3 Permutazio Entropia

Permutazio Entropiak (PE) zuzenean analizatzen du denbora-serieek duten behin-behineko informazioa; areago, azpimarratu behar dira eskaintzen dituen erraztasuna, sendotasuna eta kostu konputazional txikia [170][172][182]. Bandt-ek eta Pompe-k [173] Shannon-en entropia-neurketan oinarritutako metodo erraza eta sendoa garatu dute, zeinak kontuan hartzen duen denbora-kausaltatea hurbileko balioak erkatzen denbora-serie batean. Sekuentzia sinboliko egokia sortu egiten da denbora-serietatik era naturalean, aurreko ezagutza asumituarekin gabe [170].

Permutazio entropia denbora-serie jakin baterako $\{x_1, x_2, \dots, x_n\}$ kalkulatu da, ϵ eskala-faktorearen funtzio gisa. X_j denbora berri baten permutazioa konputatu ahal izateko, $S_t = [X_t, X_{t+1}, \dots, X_{t+m-1}]$ sortzen da m dimentsio txertatuaz eta, ondoren, goranzko ordenean antolatuta: $[X_{t+j_1-1} \leq X_{t+j_2-1} \leq \dots \leq X_{t+j_n-1}]$. m balio desberdinak emanda, π $m!$ eredu posibleak izango dira, permutazio izenaz ezagutzen direnak. Baldin eta $f(\pi)$ -k bere frekuentzia adierazten badu denbora-serieetan, bere frekuentzia erlatiboa $p(\pi) = f(\pi)/(L/s - m + 1)$ izango da. Hori horrela, permutazio entropia definitzen da:

$$PE = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i), \quad (4.20)$$

Laburbilduz, permutazio entropia denbora-serieen ordena-estruktura lokalari dagokio, zeinak konplexutasunaren neurri kuantitatiboa eman dezakeen denbora-serie dinamikoetan. Adibidez, Bandt-ek eta Pompe-k [173] $m = 3, \dots, 7$ erabiltzea iradoki zuten, beti ere honako arau hau jarraituta:

$$m! < n. \quad (4.21)$$

Baldin eta m txikiak bada (3 baino txikiagoa), algoritmoak ez du behar bezala funtzionatu egoera desberdin gutxi baino izango ez baitu grabatzeko, baina datuen mende dago. Seinale luzeak erabiltzen direnean, egokiagoa da m -ren balio handiagoa, baina denbora konputazional zabalagoak beharko dira.

4.2.2.2.4 Eskala anitzeko Permutazio Entropia

Costa *et al.*ek proposatu zuten eskala askotariko entropia (Multiscale Entropy, MSE) [172] oso metodo sendoa dela frogatu da ondorio estrukturalak aztertzeko datu erreal

konplexuek dituzten askotariko denbora-eskaletan. Morabito *et al.*-ek azaldu dutenez EEG seinale-prozesatzerako beren aplikazioan [171], ohiko ale larriko (coarse-graining) prozedura honela inplementatu daiteke:

- (i) Jatorrizko denbora-serieetatik, askotariko ale larriko bertsioak atera daitezke hurrenez hurren $y_j(\varepsilon)$, dela medio, non ε eskala-faktorea den. Hona nola kalkulatu den ale larriko denbora-serietako elementu bakoitza:

$$y_j(\varepsilon) = \frac{1}{\varepsilon} \sum_{i=(j-1)\varepsilon+1}^{j\varepsilon} x_i. \quad (4.22)$$

- (ii) Eskala-serie bakoitzerako, PE kalkulatu da.

Morabito *et al.*ek [171], bariazio eta eskala askotariko permutazio entropia (MMSPE) erabiltzen dute zehar-kanaleko aldakortasunarekin eta kanaleko korrelazioarekin zerikusia duen informazioa integratu ahal izateko. Egileok aipatzen dute askotariko eskaletako bariazioa kontuan hartu ez zuten Keller eta Laufer [190]. Ahmed-ek eta Mandic-ek [191] tankerako prosamena analizatu zuten, bariazio eta eskala askotariko entropia (MMSE), lagin-entropiari dagokionez.

Hortaz, zehar-kanalaren konplexutasuna $f_s = 1/T$ -rekin, PE (MPE) askotariko bariazioa adierazten duena, kalkulatu daiteke $s \in [f_s T - m]$ denbora guztietarako p_j -ren permutazio-entropia gisa:

$$H_{MPE}(s) = \sum_{j=1}^{d!} p_j \log_2 p_j. \quad (4.23)$$

Eta MMSPE algoritmoa inplementatu egiten da honako bi urrats hauen arabera:

- (i) Handituz doan luzerako denbora-eskala desberdinak definitzen dira jatorrizko bariazio askotariko denbora-serieak “ale larritzearen” bidez, hau da, *i.e.*, $\{x_i, t\}$, $i = 1, \dots, c$ -rako (non kanalen kopurua den) eta $t = 1, \dots, n$ -rako (non n denbora bakoitzeko serietako lagin-kopurua den). Eskala-faktore baterako, honela deribatu daitezke ale larriko denbora-serieak:

$$y_{i,j}(\varepsilon) = \frac{1}{\varepsilon} \sum_{t=(j-1)\varepsilon+1}^{j\varepsilon} x_{i,t} \text{ for } i = 1, \dots, c, \text{ and, } 1 \leq j \leq \frac{n}{\varepsilon}. \quad (4.24)$$

- (ii) MPE kalkulatu, askotariko bariazio-entropia alegia, askotariko bariazio-entropia $y_{i,j}(\varepsilon)$ “ale-larriko” bakoitzerako eta batez besteko aldaera guztietarako.

4.2.2.3 Emozio-tenperatura

Gogoratu, Alzheimerdun gaixoengan, mintzamina ez ezik, erantzun emozionala ere kaltetu egiten dela, eta, ematen duenez, zenbait fase ditu. Gainera, hasierako faseetako andeatzea dela bide, gizarte-desinhibizioa eta portaera-aldaketak gertatzen dira [45][46].

Fase aurreratuagoetan, gaixoak lotsati eta apatiko azal daitezke; sintoma horiek oroimen-galtzeari edota komunikatzeko zailtasunei eta pertzepzio aldatuari egozten zaizkio.

Aipatu bezala, erantzun eta kudeaketa emozionalak ere badira gaitzaren hasierako fase horietan nahasten diren beste komunikazio-elementu batzuk. Ezagutza- eta komunikazio-defizit eta portaera-sintoma horiek guztiak oso faktore baliagarriak izan daitezke AD ebaluatzeko.

Emozio-tenperaturaren bidez emozio-erantzuna ere aztertuko da. Tenperatura emozionala kalkulatzeko estrategia mistoa erabiltzen da. Batetik, 2 ezaugarri prosodiko eta bestetik, 4 ezaugarri paralinguistiko, pitch eta energiarekin erlazionatzen dira. Ezaugarri horiek azkar eta erraz kalkulatzeko direlako aukeratu ziren. Gainera, emozio-egutzan eraginkorrak eta segmentazio linguistikoarekiko [192] independenteak dira.

4.2.2.4 Disfluentzien azterketa

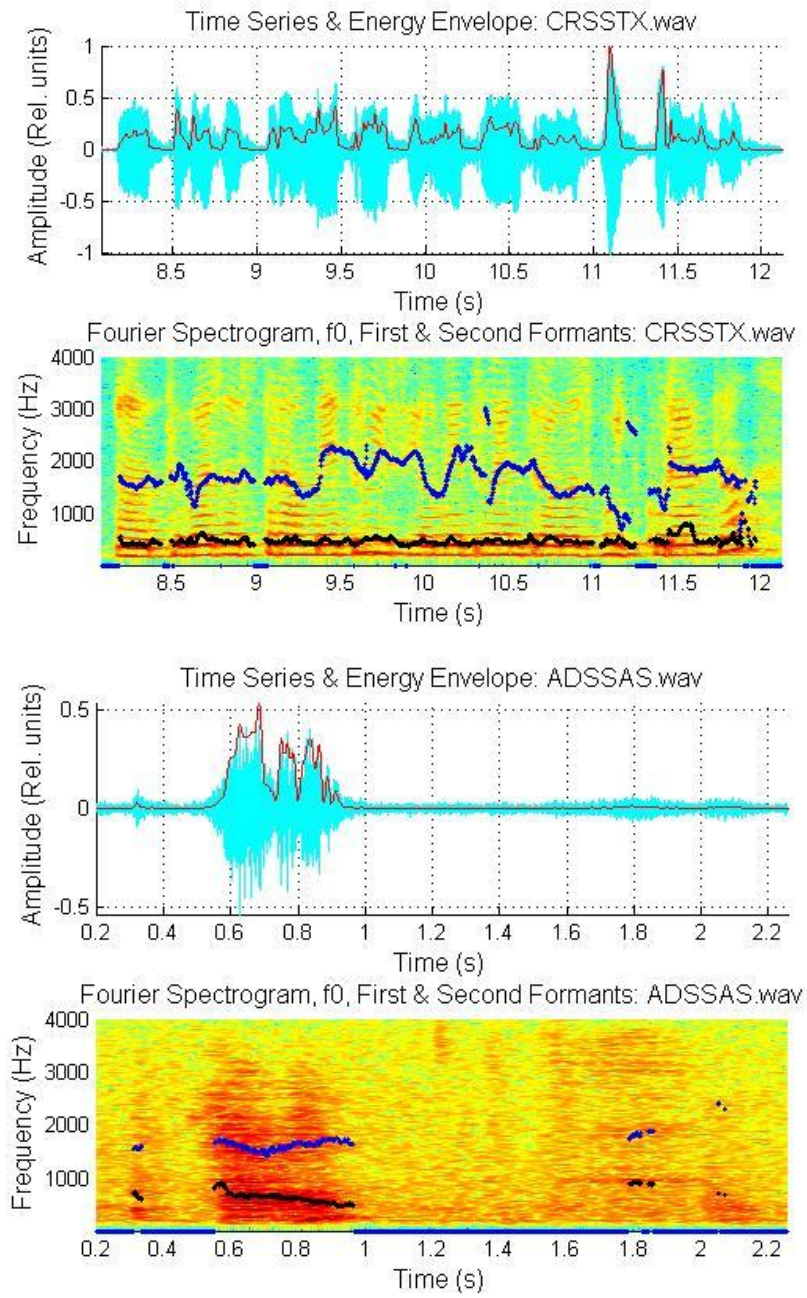
Gure ikerketa AD hasierako fasean, prodromo-fasean alegia, detektatzera bideratua dago, eta hizkuntzatik independenteak diren metodologiaren integrazioa zuzenduta. Horrek guztiak CR zein MCI duten subjektuen disfluentzia aztertzea garrantzitsua. Kontuan hartu behar dugu, memoria galtzarekin batera, AD-ren arazo handienetakoa hizkuntzaren ezagutzaren galera dela. Hitz egiteko eta beste pertsona batzuk ulertzeko zailtasunean islatzen da eta horrek komunikazioaren prozesu naturala zein elkarrekintza soziala are gehiago zailtzen du. Komunikatzeko ez gaitasun hori gaixotasunaren lehen etapetan azaltzen da, hizkuntzaren zailtasuna dela bide. AD dutenen bazterre soziala dakarte eta gaixoenganako zein beren familienganako eragin negatiboa [14]. Gainera, lan berrien arabera, disfluentziak MCI eta AD ezagutzeko garrantzia azpimarratzen dute. Ikerlan horien [19] arabera, adibidez, grabazio laburretan, hizketa sortzeko, AD duten gaixoen osasuntsu daudenak baino lan handiagoa egin behar dutela islatzen dute. AD duten gaixoen motelago hitz egiten dute etenaldi luzeagoz, denbora gehiago behar dute hitzak bilatzen, eta horrek disfluentziekin duen zerikusia eta apurtuta dauden mezuekin.

Hizkeraren disfluentzia hizketa jariatzen batean gertatzen den edozein etenadura, irregulartasun edo elementu ez-lexiko dira, eta hizketa eten dezakete. Besteak beste: hasiera faltsuak, esaldi errepikatuak, silaba errepikatu edo luzatuak, mar-marrak, betetze eta enuntziatu konponduak eta bakoitzaren mingain notak konpontzeko hizlarien instantzia edo mikroahoskatzeak [54].

Disfluentzia horiek areagotzea narriadura kognitiboaren zeinu argia izan daiteke eta gainera hizkuntzatik independentea da. Disfluentzia zein isiluneen gehiagotzeak gaixotasunaren okerragotzearen isla dezake eta komunikazioaren argitasunean eskasia ekar dezake. AD duten gaixoengan, batzuetan, barneko prozesu kognitiboaren hitzezko espresioan bilakatzen dira, hots, bakarrizketa: “Zer da hori?”, “Nola du izena?”, /Uhm/ “Ez dut gogoratzen”.

Ondoriozta daiteke, beraz, disfluentziak prozesu kognitiboaren isla zuzena direla komunikazioan zehar eta asaldura horiek antzematerako ukaezina den ezaugarria igortzen dutela, hizkuntzatik independentea izateaz gain. Ez dugu ahaztu behar AD-k, batez ere, gaixotasun kognitiboa izateaz gain, fonazio- eta artikulazio-aldakuntza biomekanikoak ekar ditzakeela. Hori dela eta, disfluentziak oso elementu interesgarriak dira eta MCI ondo detektatzeko erabilgarriak izan daitezke. Bai ahots isiluneek zein disfluentziak oso informazio garrantzitsua dute ahoskatutako mezuaren esanahia deskodetzeko.

Disfluentzia kalkulatzeko estrategia mistoa erabiltzen da, tenperatu emozionalarekin egiten den bezala. Ahots-laginak automatikoki ahots-seinale eta disfluentzian segmentatu dira Voice Activity Detection (VAD) algoritmoaren bidez. Disfluentziaren segmentutik aterako diren parametroak seinalearen zatitik aterako diren parametro berak dira. Horrela, disfluentzia aztertzeke aukera egongo da eta laginaren narriadura-maila ezagutzeko ezaugarri gehiago eskuratu ahal izango dira.



Irudia 4. 6: Hitz-jario kategorikoaren (CVF) ataza. MCI gizabanakoa. (a) Hitz egiten, (b) Disfluentzia.

4.3 Ezaugarrien hautaketa automatikoa

4.3.1 Disfluentziaren segmentazio automatikoa (VAD)

Ezaugarrien azalpenean aipatu den bezala, hots-laginak, automatikoki, ahots-seinale eta disfluentzian segmentatzen dira. Horretarako, VAD (Voice Activity Detection) [193] [194] algoritmoa erabili da.

Teknika hori hizketaldiaren prozesatzean erabiltzen da eta giza hizketaldia edo horren gabezia atzematen ditu. Ahotsaren kodetze eta horren araketarako erabili ohi da eta ahots-gabezia dagoen HTTP komunikazioetan prozesu batzuk gelditu eta konputazio zein banda zabalera aurrezteko oso erabilia da. Algoritmo horren bidez, AD/MCI/CR dutenengan izango duen eragina balioztatuko da.

4.3.2 Proba ez-parametrikoen bidezko ezaugarrien hautaketa automatikoa

Laginen normalitatea aztertzeko Kolmogorov-Smirnov proba ez-parametrikoa erabili da MATLAB-en [166] bidez, eta laginak normaltasun irizpidea betetzen ez duela ikusi da.

Lehen azterketaren ondoren eta laginak normaltasun-irizpidea betetzen ez dutela ikusita, proba ez-parametrikok erabiliko dira ezaugarrien hautaketa automatikoa egiteko.

Hain zuzen ere, Mann-Whitney-en U testa t-Student probaren baliokidea da: bi proben arteko diferentzia bakarra probarako erabiltzen den estatistikoa da. Proba hori t-Student probaren antzekoa da baina lagin ez-parametrikotarako.

Bestalde, bere zabalpena estatistikan, Kruskal-Wallis proba, k lagin homogenoak edo populazio berekoak diren erabakitzeko kontraste ez-parametrikoa da. Hautaketa automatikoan ohiko adierazleak diren ezaugarri hoberenak aukeratuko dira proba honen bidez.

Aukeraketa hori Mann-Whitney-en U probaren bidez egiten da, non p -balioa $< 0,1$ izango den ezaugarri-talde handiagoa izateko helburuarekin, bigarren hautaketa-etapa izango baita []. Gogoan izan behar dugu, ezaugarrien hautaketa horren helburua adierazgarria den ezaugarri-bektorea lortzea dela. Horrela lortu nahi dena zera da, alegia, ML algoritmoen CNN-ak barnetik aberastua dagoen informazio atxikitzea, haien lana errazteko.

CNN-en kasuan, nolabait lehen ezaugarri-mapa bat txertatzen zaio, ondoren CNN-an, konboluzio-pooling geruza gehitu diezaion jadanik ezaugarritua jasotzen duen matrizeari. CNN-ari sartuko zaion informazioa ez da bektorea izango 2D-tara eraldatzen baitugu neurona-sarean sartu aurretik.

4.3.3 Ezaugarrien bigarren hautaketa automatikoa

Behin Mann-Whitney-en U edo Kruskal-Wallis-en (MATLAB) ezaugarrien hautaketa automatikorako metodoen bidez bektorea atera denean, WEKA softwarea erabiliz, hautaketaren bigarren etapa egingo da. Oraingoan, programa honek eskaintzen duen metodoen bidez.

4.4 Sistemaren eredutzea: ezagutza automatikorako ereduak

4.4.1 Machine Learning paradigmatikak

Machine Learning-en ezagutza automatikorako lau sailkatzaile erabiliko dira klaseak etiketa desberdinetan (AD etapa desberdinak, MCI eta CR) taldekatzeko:

- K-NN: k-Bizilagun Hurbilak
- SVM: Euskarri Bektoreen Makina
- MLP: L geruzadun eta N neuronadun Geruza Aniztunadun Pertzeptroia
- CNN: L geruzadun eta N neuronadun Sare Neuronal Konboluzionala cxc konboluzio eta pxp pooling-arekin. Hurrengo atalean neurona-sare horiei buruz hitz egingo da.

Esperimentazio hori egiteko erabili dira WEKA suite [195], emaitzak ebaluatzeko Taldekatzerako Errore-Tasa (CER%) eta entrenamendu eta emaitzen balioztapenerako K = 10 den K-fold cross-validation delakoak [196].

4.4.2 CNN-ak

CNN-ek AI-aren eremuan ekarri dituzten berrikuntzek ML-aren jakintzagaia eraldatu dute. Ahots-laginak sailkatzeko k-NN, SVM eta MLP-etaz gain CNN-ak ere erabili dira ikerlan honetan eta lortutako emaitzak itxaropentsuak dira. Hori guztiagatik, azpiala osoa eskaini zaie neurona-sare mota horiei.

Gizakion garunak konplexutasun handiko mikroarkitektura dauka, eta bertan, eginkizun desberdinak burutzeko prestatuak dauden neurona-sareen nukleo eta eremuak aurkitu dituzte. Neurozientziak aurkitu du, bat-bateko garun-oinazeen kasu klinikoen azterketa eta irudien diagnostikoan izandako aurrerapenei esker, ikustearen ezaugarri desberdinak (lerroen makurdura, ertzak, simetria eta abar) hautemateko sare berezituak daudela.

Ikasketa automatiko sakonean oinarri duten eredu konputazionalak gizakietan gertatzen den jazoera hori errepikatzen saiatu dira, eta horretarako konfigurazio

zehatza duten neurona-sareak erabili dira. Arkitektura-aldaketa horretan, neurona-sare konboluzionalek edo Convolutional Neural Network-ek (CNN) zer esan handia dute, eta gainera, horrek hautemate konputazionalarekin zerikusia duten atazetan emaitzak nabarmen hobetzea ekarri du.

4.4.2.1 Deskripzioa

Neurona-sare konboluzionalak neurona-sare artifizialak dira eta datu egituratuak prozesatzeko bereziki diseinatu dira. 1 dimentsio, 2 dimentsio eta 3 dimentsiotako (1D, 2D, 3D) datuekin lan egiteko prestatuak daude. Sare konboluzional horiek gizakien pertzepzio-gaitasunera hurbiltzeko trebezia handia erakutsi dute. Neurona-sareak direla bide, errealitate biologiko horretara hurbiltzeko emandako urrats kualitatibo horren ondorioz, ikasketa automatiko sakona (Deep Learning) izeneko jakintzagaia garatzen hasi da.

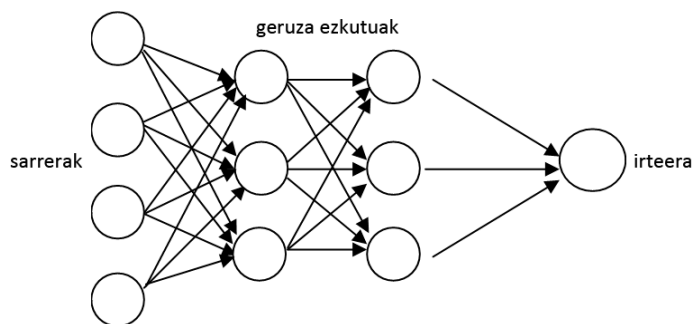
Beraz, Deep Learning-ek ikuspuntu berria dakar, eta erabiltzen diren egitura logiko digitalek ugaztunen nerbio-sistematik oso hurbil daudenez, badituzte hautemandako objektuetatik ezaugarriak antzemateko gaitasuna duten barne-geruza berezituak, ugaztunek bezala. Barruko geruza berezi horietaz gain 2 tarteko geruza baino gehiago dituzten sareak ere sare sakon izenaz ezagutzen dira.

Zenbait eremutan erabiltzen hasi dira, batez ere ikusmen artifizialean [121] edo kortex sentzoriala hobeto ulertzeko lanetan [197], besteak beste. Gainera, badaude egun, GPU-etan [198] exekutatzeko garatu diren Deep Learning liburutegi funtzionalak, eta horrek prozesu konputazionala asko azkartzeko bertutea dauka.

4.4.2.2 Multilayer Perceptron

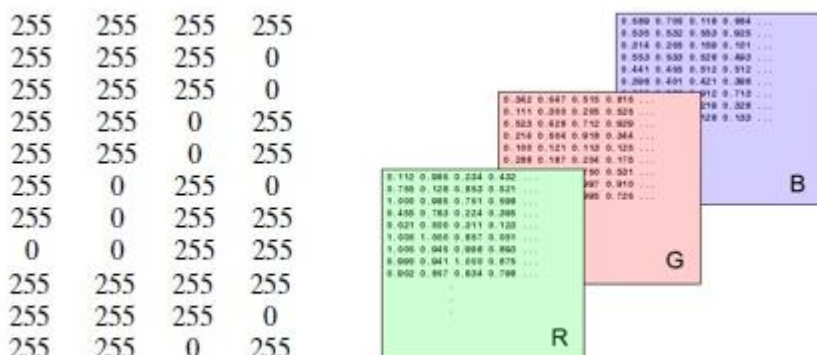
Multilayer Perceptron-a (MLP) edo pertzeptron multigeruza Deep Learning ere bada eta geruza bateko neurona guztiak hurrengo geruzako neurona guztiak elkarren artean konektatuak dituzten sareak osatzen du. Horrek parametro asko eskatzen ditu eta exekutatzerakoan prozesua asko moteldu daiteke.

4. KAPITULUA: METODOAK



Irudia 4. 7: MLP. Oinarritzko MLP baten konexioak geruza bateko neurona hurrengo geruzako neurona guztiekin konektatzen dira [199].

Adibidez 64x64x3 RGB irudiekin lan egiteko 12.000 parametro berri gehitu beharko genizkioke ereduari eta hori ez bideragarria. Horri irtenbidea bilatzeko aukera bat irudiaren erresoluzioa txikitzea da, baina informazio garrantzitsua gal daiteke. Konputazio-kalkuluak asko moteldu daitezke irudien informazio-matrizeekin lanean ari bagara, eta zer esanik ez RGB irudiei MLP sarea txertatzen badiegu.



Irudia 4. 8: Irudien datu-matrizeak. Ezkerreko adibidean 11x4x1 irudi txuri-beltza eta eskuineko adibidean luzera x zabalera x 3 RGB irudi baten matrizeen adibidea ikus daitezke [200].

4.4.2.3 Deep Learning: Neurona-sare Konboluzionalak (CNN)

CNN-k badauka konexio osoa duten pertzepton multigeruza deritzegun neurona-sareekiko desberdintasun handia:

- Entrenamendurako parametro gutxiago behar ditu, beraz, bere entrenamendua azkarragoa da.
- Baina ez du ezaugarri horrek bereizten bakarrik, matrizeen ezaugarrien traslazioarekiko inbariantza aurkezten baitu, eta gaitasun horrek matrize-informazio sastraken artean ezaugarriak aurkitzeko gaitasuna ematen dio.

Teknika horrek hizketaren ikuskapenerako emaitzak nabarmen hobeto ditu [121] eta ikerketa-lerro berriak ireki ditu audio seinaleen ikerketarako [201][202]. Sare konboluzionalak irudi, bideo, ahots eta audioen azterketan aurrerapenak ekarri dituzte eta 2D eta 3D matrizeen azterketan trebezia berezia dutela erakutsi dute.

4.4.2.3.1 Sare konboluzionalen arkitektura

CNN-ak pertzepton multigeruza ere badira, baina, tartean, desberdinak diren ezkutuko geruzak ere badituzte, eta geruza berezi horiei esker:

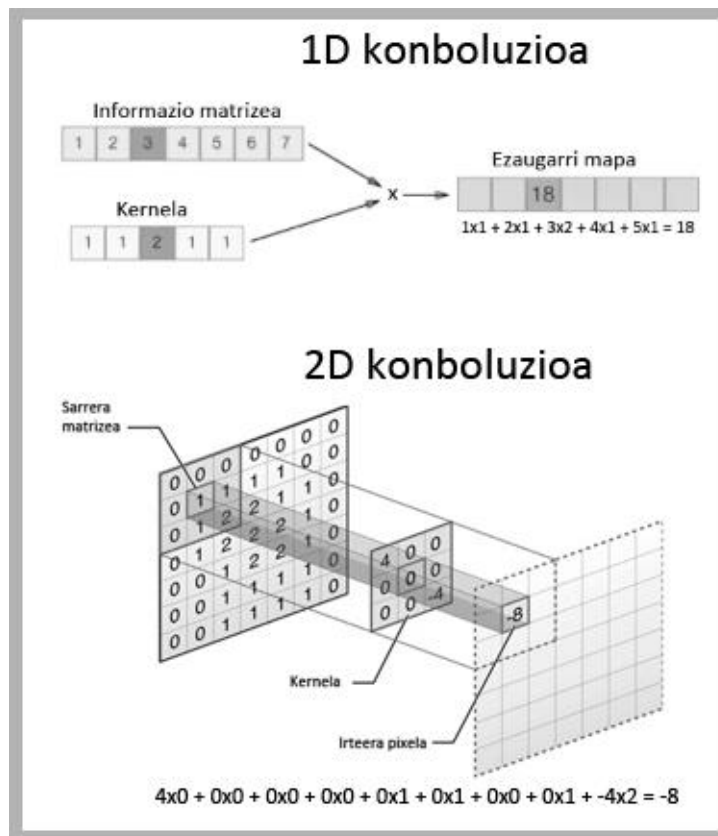
- Konexio-kopurua asko gutxitzen da, eta ondorioz, konputazio-kalkuluak asko azkartzen dira.
- Ezkutuko geruza horiek datuen arteko erlazioa aurkitzeko gaitasuna ematen diote sareari, hots, matrizearen ezaugarriak ezagutzeko gaitasuna eta trebezia.

Beraz, sare konboluzionalak xede desberdinetarako erabiliko diren arren, eta horrek CNN-ak desberdin konfiguratzeko eskatuko duen arren, badaude sare konboluzional guztietan errepikatzen diren geruzak eta bestelakoak: geruza konboluzionala, pooling geruza, guztiz konektatutako geruza, normalizaziorako aktibazio-funtzioak eta dropout teknikak.

4.4.2.3.2 Geruza konboluzionala

Konboluzioa f_1 eta f_2 funtzio matematikoak f_3 -n eraldatzen dituen eragilea da eta f_3 -k nolabait f_1 -en magnitudea irudikatzen du. Geruza konboluzionalak sartutako informazio bektore/matrizearen gain eragiketa matematiko batzuk egiten ditu kernel izeneko iragazkiaren bidez, eta horren ondoren, ezaugarrien mapa osatzeko gaitasuna dute eta geruza konboluzional bat baino gehiago segidan badaude, geruza batek aurrekoaren ezaugarri-mapa izango du sarreran.

Ikerketa-lan batzuek konboluzio horien sakontasunak irteerako doikuntzan eragina izan dezaketela erakutsi dute [129]. Azkenean, sarearen diseinatzaileak aukeratu beharko du konboluziorako erabiliko duen kernela eta sarearen sakontasuna.



Irudia 4. 9: Konboluzioa. Goian 1D-ko bektorea eta 1D-ko kernelaren arteko eragiketa matematikoa eta ondoriozko ezaugarri-mapa [203]. Behean 2D-ko matrize eta kernelaren arteko eragiketaren ondoriozko ezaugarri-mapa [204].

CNN-a konfiguratzerakoan bilatzen den helburuaren arabera, aurreko irudian ikus daitekeen bezala, erabiliko iragazkia desberdina izango da. Beheko irudian bi adibide ikus ditzakegu:

-1	-1	-1
2	2	2
-1	-1	-1

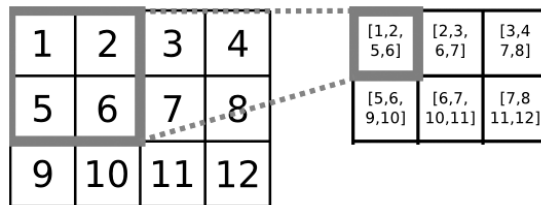
-1	2	-1
-1	2	-1
-1	2	-1

Irudia 4. 10: CNN-en kernelak. Lerro horizontalak eta lerro bertikalak ezaugarritzeko kernelak [205].

Hurbil dauden pixelek ezaugarriak aurkitzen lagundu dezakete konboluzio eragiketaren bidez [206], baina bertute hori ondo kudeatzen jakin behar da overfitting-a ekiditzeko [207]. Eragiketak ezkerretik eskuinera irrist egiten dihoa eta ilara bukatzean berriro hasten da beherago dagoen ilaratik hasita. Prozesua bukatu ondoren jatorrizko matrizea baina txikiagoa den ezaugarri-mapa sortuko da. Irrist-pausoaren tamaina definitzea ere garrantzitsua da kernelaren desplazamenduaren tamaina jakiteko (Normalean 1-ekoa izango da, hau da, zutabez zutabe egingo du salto). Iragazkiak

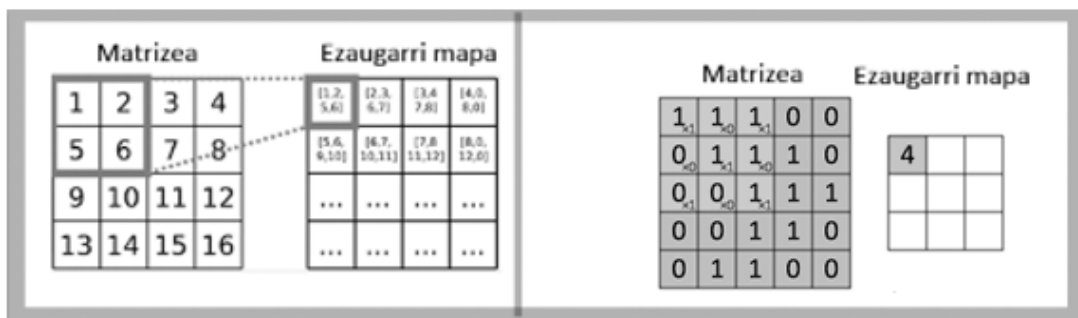
aurreko geruzan zehar egiten duen eskaner mugimenduaren tamaina ere urrats izenez ezagutzen da.

Irteerako maparen tamaina kernelaren tamainaren arabera izango da eta sarearen konexio-kopurua txikitzea lortuko da. 1x1-ekoa baino handiagoa den edozein kernel irteerako ezaugarri-mapa txikituko du. Kernel bat baino gehiago txerta daitezke ezaugarri desberdinak deskribatzeko gaitasuna eman nahi badiogu sareari.



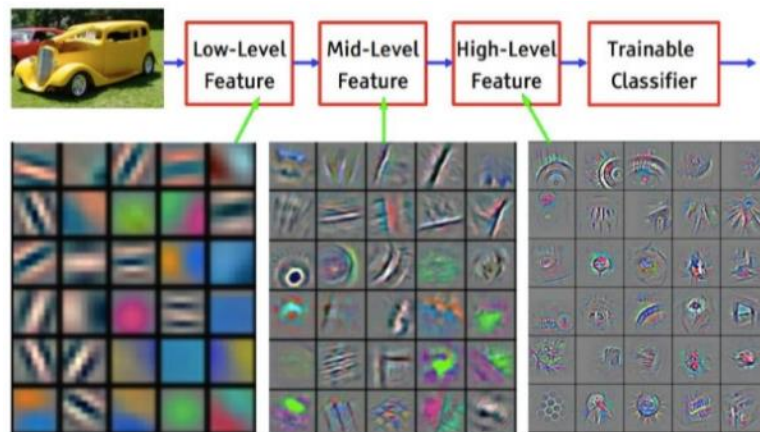
Irudia 4. 11: Ezaugarri-mapa I. 3x4 matrize batetik 2x3-ko ezaugarri-mapa sortzen da [208].

Batzuetan ordea tamaina mantentzea interesgarria izan daiteke eta horretarako zeroen bidezko bigungarria (zero padding) jar daiteke ertzetan (irudia zeroz inguratuta) edo ausazko ertza.



Irudia 4. 12: Ezaugarri-mapa II. Ezkerreko irudian ezaugarri-mapak tamaina mantentzen du konboluzioa posizio guztietan eginez. Hau da, sarrerako matrizeari behar adina gehitzen dizkio ertzetan irteerako ezaugarri-maparen tamaina mantendu dadin. Eskuineko irudian, ordea, ezaugarri-maparen tamaina txikiagotzen da. Sarreran 5x5-eko matrizea izanik irteeran 3x3-koa geratzen da [208][209].

Geruza taldekatzeak geruza konboluzionalen segida batek osatzen dute. Hauen helburua, aurretik definitutako ezaugarri-mapetako ezaugarriak finkatzea da. Ezaugarrien irudikapena orokortu edo eta konprimitzeko balio duen teknikatza har daiteke. Horrela, eredia sortzeko erabilitako datuen gehiegizko doikuntza ekidituko da. Askotan geruza taldekatzeak sinpleak izaten dira eta beren ezaugarri-mapa propioak sortzeko sarrera-datuen batez besteko edo maximoak hartzen dituzte.

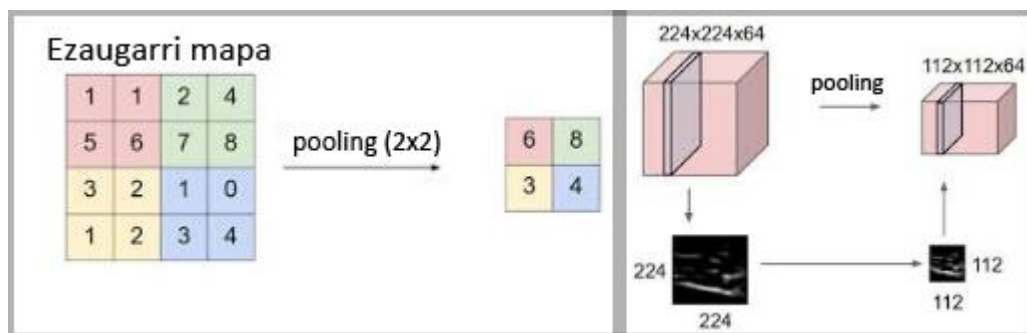


Irudia 4. 13: CNN-en sakotasuna. Lehen geruzan konboluzioa egin eta ezaugarrien mapa atera ondoren beste hiru geruza erabiliko dira kasu honetan ezaugarrietan sakontzeko [210].

4.4.2.3.3 Pooling geruza

Behin konboluzioa bukatu eta gero, ezaugarrien mapa pooling izeneko beste geruza batetik pasako da. Horrek, zonaldeka, *mean-pooling* edo *max-pooling* eragiketak egingo ditu aztertzen ari den ezaugarrien maparen gainean. Pooling iragazkiak tamaina desberdinak izan ditzake, nahiz eta 2x2-koa erabili den. Pooling-arkitektura estatikoa da hori, baina badaude irudietara hobeto egokitzeko helburuarekin piramide bidezko taldekatze espazialerako pooling dinamikoagoak [211].

Pooling delakoaren geruzek ezaugarri-mapatik jasotako informazioa gehiago sinplifikaten dute eta ezaugarrien bertsio laburbildua sortzen dute. Geruza konboluzionalak ematen digun informaziotik abiatuta eremukako (Adibidez 2x2) informazioa hartu eta pooling geruzaren puntu bakarrean sintetizatzen da:



Irudia 4. 14: Pooling-a CNN-etan. 2 kasuetan 2x2 pooling-a egiten da. Ezkerreko irudian zehazki max-pooling egiten du eta eskuineko irudian pooling-ak irudian duen eragina ikus daiteke [212].

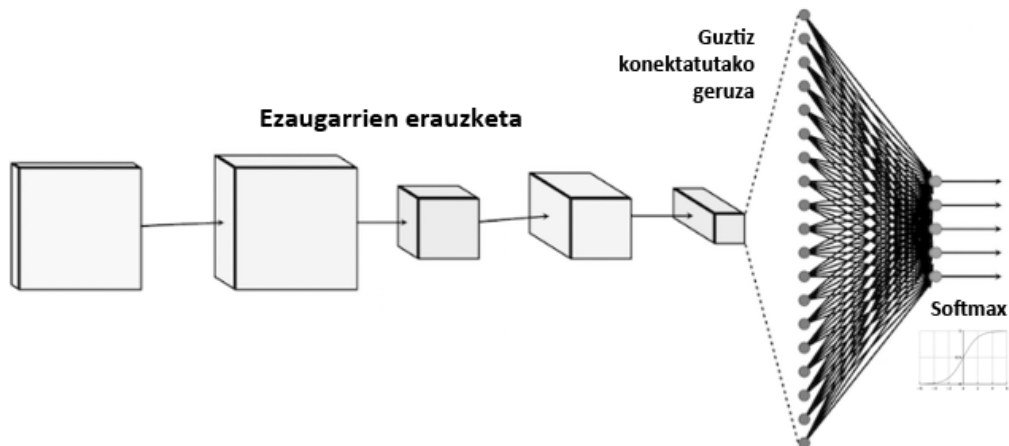
Pooling-aren bidez ezaugarri bat aurkitutakoan, honek beste ezaugarriekiko duen gutxi-gorabeherako kokapena bere posizioa zehatza jakitea baino garrantzitsuagoa da.

4.4.2.3.4 Guztiz konektatutako geruza

Behin konboluzio-pooling (edo subsampling) geruza-segidak behar adina errepikatuz gero (batzuetan konboluzio geruza bat baino gehiago segidan), sailkapenerako algoritmoetan (gure kasuan bezalaxe, hots, MCI edo CR) guztiz konektatua dagoen geruza (densely connected layer) gehitzen zaio.

Irteeran, softmax edo beste aktibazio-funtzio ez-lineala izan dezakete eta entropia gurutzatua (cross-entropy) bezalako optimizatzailea (ataza probabilistikoetarako hobeto lan egiten du errore koadratikoak baino) adibidez. Sailkapenerako Deep Learning sareetan, konbinazio horrek ondo lan egiten du baina badaude beste aukera batzuk.

Softmax, irteerako bektore erreala, probabilitate bektore bihurtzeko diseinatu dago (guztien artean 1 gehitzen dute), hots, datu-matrize bat klase baten kide izateko probabilitateaz ariko ginateke gure kasuan. Mendeko aldagaia klase bateko kide izateko probabilitateaz ari gara, hau da, sailkapen-kasu bati buruz. Taldekatzerako neurona bat erabiltzen da mota bakoitzerako. Gure kasuan bi neurona izango dira irteeran (MCI/CR). Probabilistikoa izan ordez *One-hot encode* motakoa izan daiteke (adibidez: [0 0 1 0 0] bost neuronatako irteera batean).



Irudia 4. 15: CNN-en egitura orokorra. Deep Learning sarearen bukaeran guztiz konektatutako geruza gehitzen da eta sailkapen ereduarentako bektore probabilistikoa sortuko duten funtzioak, softmax besteak beste [213].

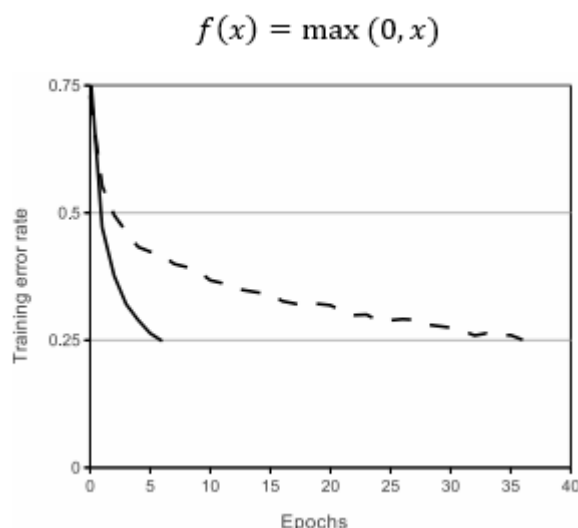
Guztiz konektatutako geruzak sarearen bukaeran erabiltzen dira behin ezaugarriak erazi eta sendotuz gero. Haiak erabiltzen dira ezaugarrien bukaerako konbinazio ez-linealak sortzeko eta sareak aurreikuspenak egiteko.

4.4.2.3.5 Aktibazio-funtzioak

Sarrera batzuk emanda korapilo baten irteera definitzen duen funtzioa da eta zenbait motatakoak daude. Sarrerako geruzan izan ezik beste guztietan burutzen da eragiketa hau. Sarrerako konexioetatik datozen balio guztien gehiketa jaso ondoren, beste balio batean eraldatzen du. Hainbat aukera daude (sigmoide, ReLU, ...) eta horien helburua neurona bakoitzak sortutako balioak arrazoizko eremu batean mantentzea da (0 eta 1 artean dauden balio errealak adibidez). Aktibazio-funtzioek ez-linealak izan behar dute baldin eta ebatzi nahi den kasuistika ez-lineala bada. Lineala zatekeen kasu bat pezetatik eurotarako bihurtzea zatekeen. Gure ikerketa-lerroa, ordea, ez-lineala da guztiz, hortaz, ez-lineala den aktibazio-funtzioa erabili beharko da baita ere.

ReLU (Rectified linear unit) aktibazio-funtzio ez-lineala oso erabilia izaten ari da neurona-sare sakonetako geruza ezkutuetan sarearen pisu eta ereduak azkarrago ikasten dituelako:

Aktibazio-funtzio ez-linealak erabiliko dira ondoren. Zenbait mota daude baina gure kasuan ReLU (Rectified linear unit) erabiliko dugu ikasketa sakonerako neurona-sareetan lan ona egiten baitute:



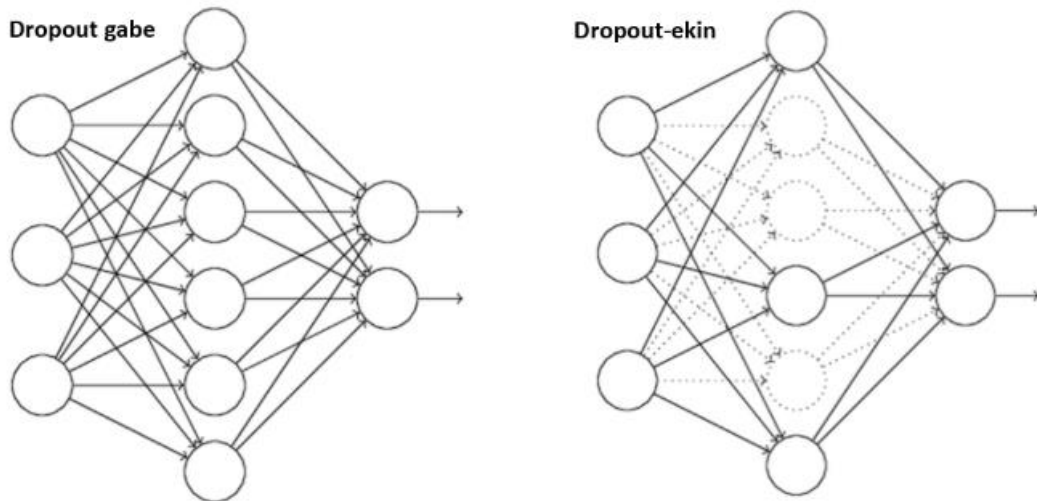
Irudia 4. 16: ReLU. ReLU oso erabilia da, entrenamendua bider 6 azkartzen baitu sigmoide eta tanh aktibazio-funtzioekin alderatuta [122].

Alabaina, gerta daiteke neurona batzuek lanik ez egitea entrenamenduan zehar gradienteak sortutako pisuek ahalbidetu baitezakete neurona batzuk ez aktibatzea. Neurona horientzat gradientea beti zero izango da. Arazo hori ekiditzeko badaude bide batzuk, adibidez, dropout.

4.4.2.3.6 Dropout

Garrantzitsua da ereduaren parametro-kopurua gutxitzea. Konboluzio eta pooling-aren bidez sarearen konputazio-abiadura handitu da baina bada *dropout* izeneko teknika bat ereduaren parametroak gutxitzen lagunduko digunak. Sareko neurona batzuk ezabatzen dira bere sarrera eta irteera arkuekin batera. Sarea osatzen duten neurona eta bere konexioen deuseztapena ausazkoa da.

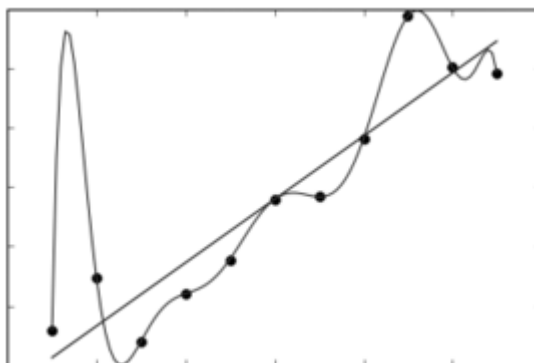
Entrenamendurako nahikoa adibide ez badaude gerta daiteke neurona gutxi batzuek lan guztia egitea, eta gainerako neuronak, ordea, erredundanteak eta errendimendurako kaltegarriak izatea. Horretarako, erregularizazio-teknikak txertatzen zaizkie, ereduaren parametroei murrizketen bidez zarataren ikasketa ekiditzeko helburuz. Iterazio batzuetan, neurona batzuk ezabatu eta besteak lan egitera behartzen ditu, akatsak ekidin eta neurona gutxi batzuekiko mendetasuna desagerrarazteko.



Irudia 4. 17: Dropout. Ezkerreko irudian oinarritzko eta ezkutuko geruza bakarreko MLP sarea ikus daiteke. Eskuineko irudian iterazio batean Dropout-ak lan nola egiten duen ikus daiteke [214].

Teknika horrek bi ahalmen gehituko dizkio CNN-ari:

1. Kalkulurako abiadura konputazionala azkartuko da parametro gutxiago izango baitira iterazio bakoitzean.
2. Gehiegizko doikuntzaren arriskua txikiagotuko du. Oso adierazkorrak diren geruza ezkutuek dituzten eredueta gehiegizko doikuntza gertatzeko arriskua handitzen da eta ausazko zarata edo akatsaren zati bat modelizatu daiteke. Horri *overfitting* deitzen zaio, eta dropout-aren bidez sarearen aurreikuspen pobretze hori ekiditu nahi da. Gehiegizko doikuntza horrek orokortzeko gaitasuna mugatu dezake. Entrenamendurako erabilitako datu-taldeak ikasi (zarata barne) dezake modu ezin hobean baina azpian dakarren prozesua ulertu gabe.



Irudia 4. 18: Overfitting. Oraingoan, overfitting-a argia da. Ereduak entrenamendu datuei buruz dena ikasi du baina ez dauka orokortzeko gaitasunik. Oinarritzko erregresio lineal batek hobeto orokortuko luke [215].

4.4.2.3.7 Eredua entrenatu

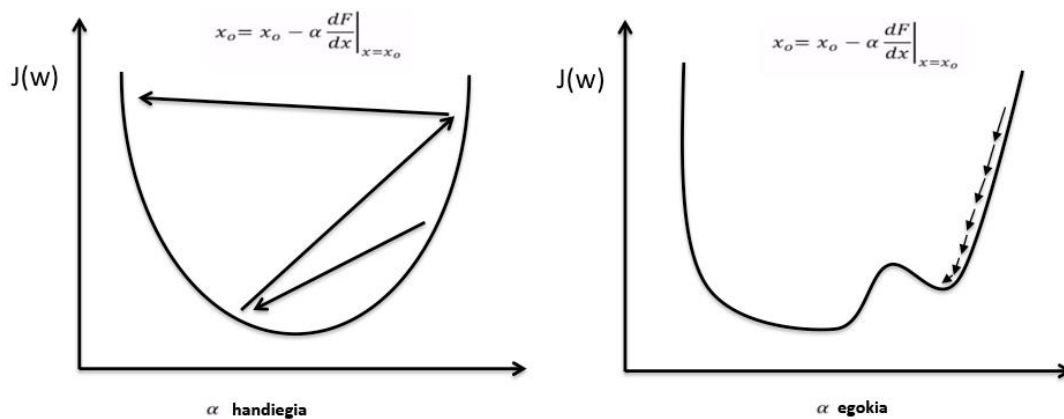
Behin CNN-aren konfigurazioa modelizatuz gero, algoritmoa entrenatzen hasi daiteke. Etiketatuak dauden mendeko aldagaietatik abiatuta, bilatzen den emaitza hobetzeko helburuz, geruza konboluzionalen parametroak doitzeko lan franko dugu. Horretarako, bai WEKA bai python erabili daitezke. Python-en kasuan, keras liburutegiak [216] modelizatzeko honako argumentu hauek eskaintzen ditu, besteak beste:

- `batch_size`: Entrenamendu adibide kopurua iterazioko
- `num_epoch`: Iterazio kopurua
- `kernel_size`: Geruza konboluzionalak nukleoaren tamaina
- `pool_size`: Pooling geruzen tamaina

Entrenamendurako optimizatzaile batzuk daude (gradient descent, ADAM eta abar). Ondorengo Gradiente (Gradient Descent) erabiliena da erabiltzeko erraza delako. CNN sareetan ADAM optimizatzaileak abantailak eskaintzen ditu. Neurona-sareen ikasketa-prozesua errorearen minimizatzeko funtzioetat azaltzen da. Normalean funtzio horrek 2 zati izaten ditu:

- Erroreak zehaztea: Neurona-sarea zenbateraino doitzen den dauden datuetara.
- Erregularizatzea zehaztea: Gehiegizko ikasketa ekiditzeko.

Noski, sareak izango duen akatsa zuzenki erlazionatua egongo da dituen sinapsien pisuekin. Normalean errore funtzioa ez da lineala eta iterazioen bidez errore hori minimizatzeko bidea bilatu behar da.



Irudia 4. 19: Overshooting. Ondorengo gradientearen iterazio bakoitzean sartuko dugun aldaketa α -k zehaztuko du. Handiegia bada overshooting-a izango da (dibergentzia) eta ez du errorea minimizatzeko balioko. Ezkerreko irudian ikus daiteke nola minimotik urruntzen den. Eskuineko irudian, α txikiago batek minimorako bidea erakusten digu [217].

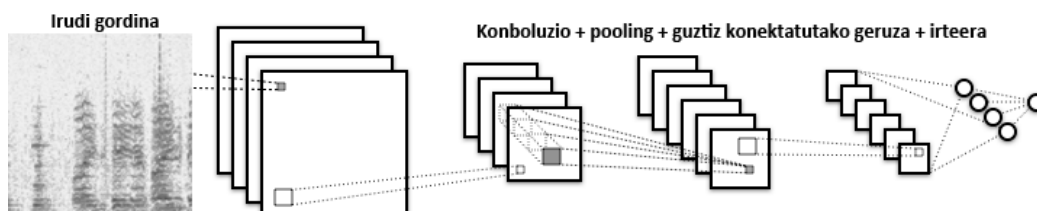
Behin doitu eta entrenatuz gero, sarearen gaitasunak ebaluatu beharko dira eta erabilitako beste algoritmoekin konparatu (SVM, k-NN eta MLP).

4.4.2.3.8 CNN ahotsa analisirako

Badakigu, azkenean, 100 pertsonaren datu-base orekatua (PGA-OREKA) lortu dela eta hori da neurona-sare konboluzionala entrenatzeko eta bere zehaztasun-maila neurtzeko erabiliko duguna.

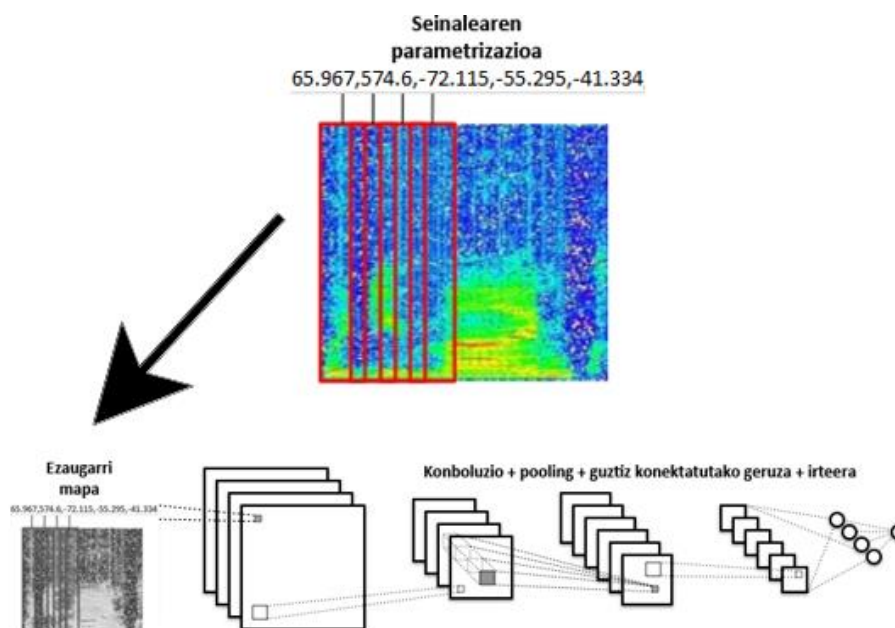
MCI eta CR sailkatzeko orain arte SVM eta KNN bezalako sailkapenerako algoritmo klasikoekin egin dugu lan. Badakigu [122][127][129][131][132], CNN-ek oso erantzun onak ematen dituztela irudi matrizeen sailkapenekin eta gure 100 laginetako ahotsekin lortutako parametro bektoreekin ondo funtziona dezaketela ematen du.

Kontuan hartu behar da neurona-sare konboluzionalean sartuko diren datuak ez direla espektrogramaren irudien pixelak. CNN-ra sartuko dugun informazioa lehen ezaugarri-mapa bat da, .wav fitxategietan egindako iragazketa-lanak seinale horiek parametrizatzeko balio izan baitu.



Irudia 4. 20: Datu gordinak CNN-ra. Aurreko iragazketa-lana egingo ez balitz espektrogramaren matrize-datu gordinak sartuko lirateke CNN-ra [218].

Deep Learning-ek bere trebezia erakutsi badu [122][127][129][131][132] datu-matrizeak ezaugarritzeko eta horiek konprimitzeko, sartzen diren datuak gordinak izan ordez, ezaugarri-mapak badira lor daitezkeen emaitzak hobeak izango direla espero daiteke. Horrek esan nahi du CNN-ra txertatuko den informazioa aberastua dela eta bektore parametrikokoak direnak ezaugarrien bektoreak balira bezala ulertuko dituela sareak. Ezaugarritze-lan handia ematen zaio sareari. Ikuspuntu guztiz berritzaile honek emaitza onak eman ditzakeela uste da.



Irudia 4. 21: Datu aberastuak CNN-ra. Irudi gordina sartu ordez, ahots-seinalea aurrez tratatu eta parametrizatzen da eta lagin bakoitzetik parametroen segida (bektorea) lortzen da. Behin segida horren balio adierazgarrienak aukeratuta (Mann-Whitney U-test ebaluazio ez-parametrikoko eta Support Vector Machine-n (SVM) bidez), segida hori CNN-ra sartzen da. Datu aberastuak sartzen dira, lehen ezaugarri-mapa balira bezala [218].

4.5 Sistemaren balioztatzea

Modelatzearen atal garrantzitsu bat bere ebaluazioa da. Hizketaren ezagutza automatikoaren lana neurtzeko era zuzen eta intuitiboena tasak dira: *Accuracy*, *Classification Error Rate* (CER) eta honen Errore-Tasa Metagarria edo *Cumulative Error* (CE). Bi kasuetarako 4.1 eta 4.2an, (TP = True Positive, TN = True Negative, FP = False Positive eta FN = False Negative).

$$Accuracy = \frac{TP + TN}{TP + TN + F + FN} \quad (4.25)$$

$$CER = \frac{FP + FN}{TP + TN + F + FN} \quad (4.26)$$

K-folds-etan oinarritzen den balidazio gurutzatuko teknika (*k-fold cross-validation*) Esperimentu-Matrizean (Experiment Matrix, EM) zuzenean aplikatzen den sailkatzailearen doitasuna estimatzeko balidazio-metodoa da. Kasu honetan, EM matrizea *k* zatitan banatu egiten da, tamaina bereko *k* zati lortuz (*k*-folds). Sailkatzailea *k* aldiz sortzen da, multzoetako bat baztertuz iterazio bakoitzean. Baztertutako multzoaren laginek Balidazio-Matrizea (Validation Matrix, VM) osatuko dute. Aldi bakoitzean, EM-an geratzen diren laginek iragarle bat sortzen dute, eta iragarle hori VM-aren laginekin testeatuko da. Erroreen baturak teknikaren errore globala ematen digu.

Metodo hau oso eraginkorra da lagin-kopurua handia denean. Azken iragarlean lagin guztiak erabili ohi dira.

Bat-Soilik-Kendu teknika (*Leave-One-Out Cross-Validation*, LOOCV) EM-aren gainean aplikatzen den beste balidazio metodo bat da. Metodo hori oso erabilgarria da lagin-kopuru txikiekin lan egiten denean. Kasu honetan, n lagin duen multzo bat badugu, n iragarle sortuko dira $n-1$ laginekin, eta baztertutako laginak VM-a osatuko du (elementu bakarreko matrizea) errore-tasa kalkulatzeko. LOOCV k -fold cross-validation teknikaren kasu partikularra da, k -fold-etako k koefizientea eta n lagin kopurua berdina den kasua, alegia [43][195].

Bukatzeko, komenigarria da egiaztatzea sortutako iragarleak ausaz sortutako iragarle batek baino hobeto sailkatzen duela. Horretarako *Permutation p -value for cross-validated misclassification error rate* izeneko indizea erabiltzea egokia izan daiteke. Indize honek errore-tasa txikiagoa duen ausaz lortutako sailkatzaile bat aurkitzeko probabilitatea kalkulatu du [219][220][221].

4.6 Sistemaren konfigurazioaren hautaketa automatikoa

Azkenik, soluzio optimoa (ezaugarri-multzo hoberena) Helburu-Funtzio (HF) baten bitartez aukeratu egiten da.

Biz $X = \{x_i\}_{i=1,\dots,n}$ n dimentsioko ezaugarri-bektore eta $Y = \{y_j\}_{j=1,\dots,p}$ p dimentsiotako aukeratutako ezaugarri-bektorea. Era berean helburu-funtzioa definitzen da:

$$HF(Y) = \sum_{i=1}^p \omega_i \cdot y_i \quad (4.27)$$

non ω_i y_i osagairi dagokion pisua da. Soluzio optimoa, helburu-funtzioa maximizatzen duen aukera izanik:

$$z = \max_k \{HF(Y)\} \quad (4.28)$$

Metodo horrek honako modelatze hauek kontuan hartuko ditu: Accuracy, klaseen True Positives, Positibo Faltsuak, ereduaren konplexutasuna, segurtasun-maila, denbora erreala, ROC area esperimentazioaren arabera, eta osasun-espezialisten interesen arabera egokituko da.

5. KAPITULUA: ALZUMERIC

5.1 Sarrera

Atal hau sakonago ulertzeko **B.1, B.2, B.3, C.3, C.4** eta **D.2** eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

AD-ren lehen fasean agertzen da patologikoa den Nahasmendu Kognitibo Arina (MCI), akats kognitiboa alegia, baina ez da behar bezain larria eguneroko bizitza edo gaitasun independenteak eragozteko. Aldaketa horiek ez dira gehienetan behar bezala diagnostikatzen. Are gehiago, MCI-k Alzheimer edo beste dementziaren bat garatzeko arriskua areagotu dezake, eta gaur egun, garaiz detektatzea da espezialista medikuen erronketako bat [2].

Mintzamen-defizitak AD duten gaixo askorengan agertzen dira, eta seinale adierazgarritzat jo daitezke hasierako faseetan. Azken hamarkadotan, dokumentatu eta berretsi ahal izan da hizketaren disfuntzioa Alzheimerren gaitzaren garapenean goiz azal daitekeen sintoma bat dela, eta biomarkatzaile erabilgarria izan daitekeela prebentzio eta detekzioarako [2][19][14][20][21].

Narriadura Kognitibo Arinean (MCI-n) eta AD-n lehen kaltetzen den gaitasunetako bat denaren, alegia, hitzezko komunikazioaren, analisi automatikoa daukagu bahetze metodologia guztiz natural eta erabilgarria. Ahozko komunikazioaren azterketa automatikoak bahetze-metodologia erabilgarria eskaintzeaz gain, dudarik gabe, oso baliagarria izango da balioztatze klinikorako baina baita ere etxeko/hasierako arretarako.

Alzheimerdun gaixoengan, mintzamina ez ezik, zenbait fase dituela dirudien [40] erantzun emozionala ere kaltetu egiten da, eta hori, mintzamenaren bidez azter daiteke. Emozioak giza gogoaren arkitekturarekin zerikusia dituzten prozesu kognitiboak dira eta lotuta daude honako prozesu hauekin: erabakiak hartzea, oroimenaren kudeaketa edota arreta, zeinak hertsiki lotuta dauden komunikazioarekin, ikasi eta ulertzearekin [41][42].

Disfluentziak ere hizketa-elementu interesgarriak dira, eta oso baliagarriak izan daitezke MCI detektatzeko. Hizketa-disfluentziatzat hartzen dira hizketa arinean gertatzen diren edozein etete, irregulartasun edota hitz ez-lexikal, hizketa hori hasi edo

eten dezaketenak. Hala isiluneek nola disfluentziek informazio baliagarria izan dezakete ahozkatutako mezua deskodetzeko. Hori dela eta, ikerlan berriek agerian jartzen dute bien garrantzia bat-bateko hizketan MCI eta AD identifikatzeko. Are gehiago, disfluentzia ez-patologikoek oso informazio aberatsa izan dezakete emoziozko egoeraz edota hizkuntza-planteamentuaz. Disfluentzia horien handiagotzeak nahaste kognitiboaren isla nabarmena izan daitezke.

Ikertzaileak urte luzeetan ibili dira hautemate-teknika garesti eta inbasiboei alternatibak bilatzen mintzamen-gaitasunetan edo emoziozko egoeren aldaketetan izaten den andeatzea sakon aztertzeko helburuarekin. Horretarako monitorizazio-teknika automatiko ez-inbasiboak oso tresna baliagarriak izan daitezke detekzio goiztiarrerako diagnostikoa osatzeko edota gaixotasunaren garapena balioztatzeko eta eriengan nola senideengan inpaktua arintzeko eta gizartean gaitzaren progresioa moteltzeko. Osasunaren adarretan jadanik garatzen hasi dira teknologia ez-inbasiboak [222].

Medizina klasikoak bitarteko teknologikoak bilatzearen arrazoia egin beharreko proba medikuen murrizketa eta osasun-baliabideen hobekuntzan dago. Espezialistei laguntza handia suposatzen die horrelako baliabideak eskuragarri izatea diagnostiko fidagarri bat egiteko beharrezkoak diren denborak murrizten baititu. Alzumeric-en kasuan, gainera, oso kostu baxuko teknologia da.

Gaur egungo teknologiarekin datu digitalen prozesatze abiadura geroz eta handiagoa da. Tabletak zein smartphoneak bezalako gailu adimendunak herrialde garatuetako gizartearen gehiengoarentzat eskuragarri badaude, teknologia merke eta ez-inbasiboak garatzea ekarpen interesgarria izan daiteke mediku zein erien senitartekoentzat. Bestalde, interneterako konexio abiadura hobetzen ari da egunetik egunera, eta honek, geroz eta denbora txikiagoan datu eskerga elkartrukatzeko aukera berriak irekitzen ditu.

Erabilera kliniko baterako mahai gaineko ordenagailu zein mugikorrek diren terminalen erabileraren zabalpena handituko dela ondoriozta daiteke, azken urteotan, iraultza ekarri baitute sektore honetara [223]. Gainera, haien presentzia eta eskuragarritasuna orokortua denez, plataformak ikasketa-kurba oso laburra izan dezakeela aurreikusi daiteke.

Errealitate teknologiko berri honek orain arte pentsaezinak ziren bezero-zerbitzari arkitektura berriak garatzeko aukerak ireki ditu. Arkitektura berri hauek, konkurrentzia altuko eskaerak zentzuzko denboran prozesatzeko gaitasuna, eta gainera, erabiltzeko errazak eta ez-inbasiboak izan behar dute.

5.2 Alzumeric

Badakigu ahots-laginetatik ezaugarri-mota asko erauzi daitezkeela, linealak eta ez-linealak, eta horiek oso informazio aberasgarria eman dezaketela eriazen egoera hobeto ulertzeko. Alzumeric-ek prozesu guztia, atazaren grabaziotik hasi eta emaitzak bueltan bidali arte, automatikoki eta metodologia ez-inbasiboen bidez kudeatuko duen web-plataforma da.

Gainera, interneterako sarbide abiadura eta informazioaren bereizmena gogoan hartzekoak dira eta komunikazio-mota hauetan HTTP/HTTPS protokoloak dira nagusi. Alzumeric-en kasuan, zerbitzari eta bukaerako erabiltzailearen arteko informazio-elkartrukea artikulatzearen errudunak dira eta internetera sarbidea duten terminal guztietan daude integratuak.

Testuinguru horretan, Alzumeric plataforma sortu da, hots, mintzamen-komplexutasuneko maila desberdinak dituzten atazen grabazioetatik abiatuta (AN, PD eta SS), hizkuntzarekiko independentea den mintzamenaren azterketa automatikoa egingo duen web-sistema ez-inbasibo, merke eta unibertsala.

Ataza (AN, PD edo SS) grabatu eta bukatzean ahots-lagina zerbitzarira bidaltzen da gailu adimenduaren (Smartphone, tablet, PC) bidez eta atazaren ezaugarrien laburbilduma (linealak eta ez-linealak) eta hautaketa (Mann-Whitney U-testaren bidez besteak beste) egiten du. Bukatzeko, laginaren sailkapena (MCI/AD edo CR) eta medikuekin adostutako parametroak bueltan bidaltzen ditu espezialistak aintzat har ditzan.

ETL prozesuko etapak kudeatuko ditu: laginak jaso, parametro adierazgarriak erauzi eta aukeratu, sailkapena egin eta bueltan makinaren ekarpena espezialista edo eriazen senitartekoei bidali.

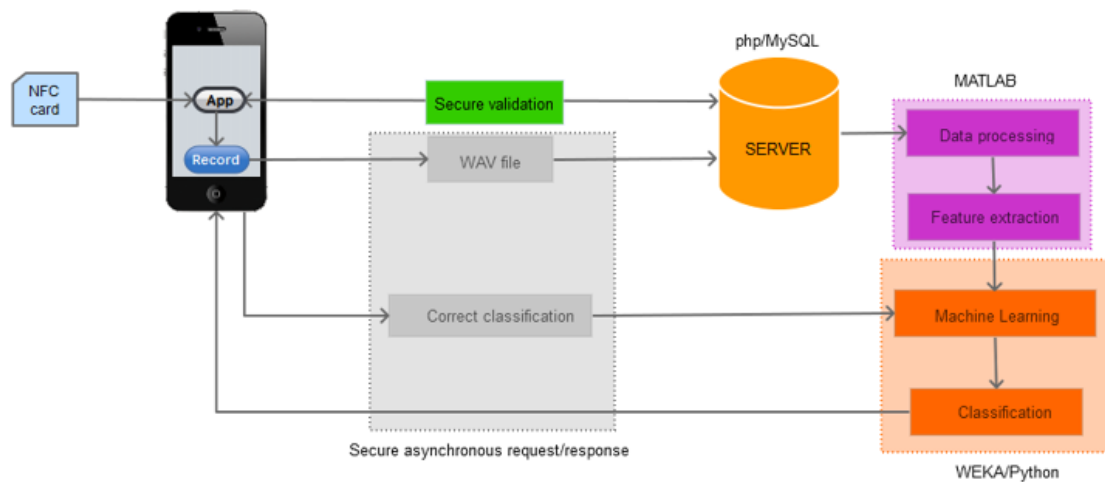
kontuan hartu dira sistemak jaso ditzakeen erasoei aurre egiteko. Informazioaren pribatasuna segurtatzeko helburuarekin komunikazio guztiak zifratuak daude.

Erabiltzaileak terminala (Smartphone eta tableta) besterik ez du beharko sistema erabili ahal izateko. Prozesu guztia ambulatorioan edo pazientearen ohiko ingurunean, etxeko giroan, kudeatu daiteke. Giro lasaian eta teknikariak ez diren pertsonak edo senideak direla medio. Ataza etxeko ingurunean egin bada, erieen senitartekoek informazio hori medikuari helarazi diezaiokete, honek eriazen jarraipena egin dezan. Ataza ingurune klinikoan egin bada, ordea, medikua bera izango da informazioa zuzenean jasoko duena. Osasun-guneen irizpide etikoak kontuan hartuz garatutako sistema da.

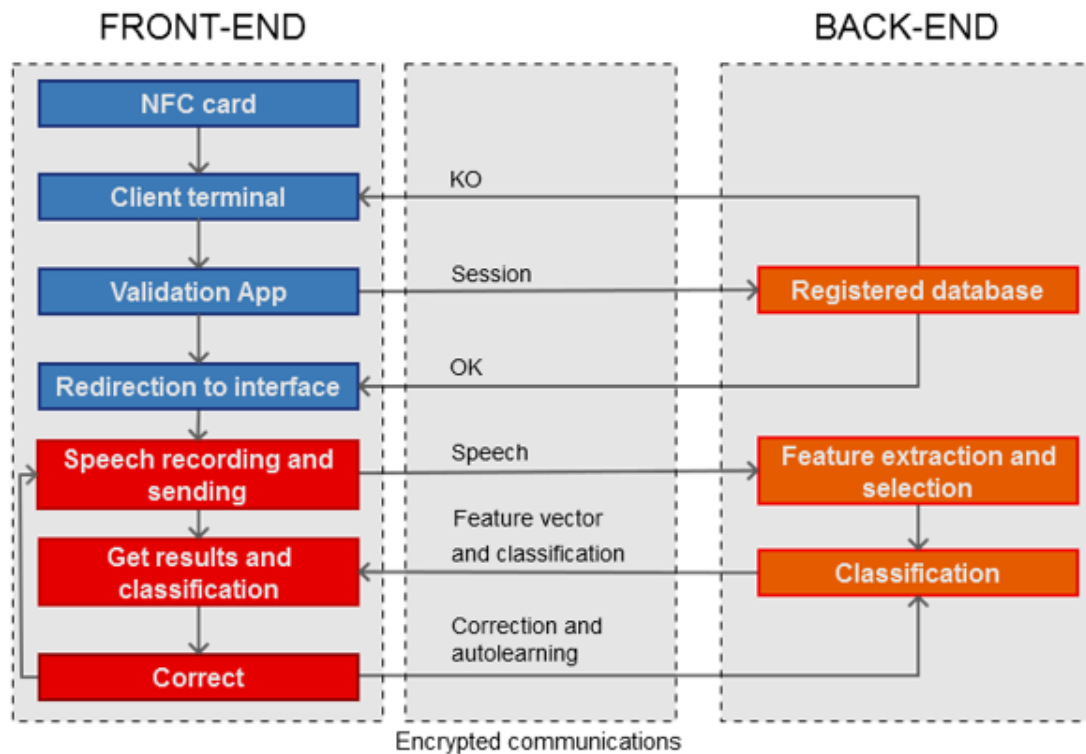
Alzumeric prototipoa ingurune kliniko zein etxekoan balioztatu da osasun-guneen irizpide etikoak kontuan hartuz eta sistemaren lehen bertsio hau balioztatzeke erabiltzaile tekniko, laguntzaile eta osasun-espezialista gutxi batzuek erabiliko dute.

5.2.1 Arkitektura orokorra

Erabiltzaileak zenbat atazaren bidez sistema elikatzen duenetik eta erantzuna jaso arteko bidean, arkitektura teknologiko hibridoa dago. Hori horrela bada, teknologien arteko informazio-transmisioaz eta horren ulermenaz arduratuko diren protokoloek ere osatzen dute sistema.



Irudia 5. 2: Alzumeric-en arkitektura orokorra. Behin mediku edo erieen senitartekoak modu seguruan saioa hasi duenean grabazioa egin diezaiokete gaixoari. Grabazioa bukatu eta zerbitzarira bidaltzen da, honek seinalea parametrizatu eta bere sailkapena egin dezan (MCI/AD edo CR) dezan. Medikuek aurrez adostutako parametroak eta laginaren sailkapena jasoko ditu. Medikuek Alzumeric-ek egindako sailkapena zuzentzeko aukera izango du eta sistemaren ikasketa-prozesuan lagundu.



Irudia 5. 3: Sistema automatikoaren front-end-a eta back-end-a. Osasun-guneen irizpide etikoak kontuan hartu eta komunikazio guztiak zifratzen dira. Medikuei zein euren senitartekoei saioa hasterako orduan jasotzen diren komunikazioek segurtasun-geruza gehiago dituzte. Behin saioa hasi dutenean lagina jaso dezakete ataza desberdinak erabiliz. Lagin horiek zerbitzarira bidali eta honek, automatikoki, ezaugarri lineal eta ez-linealen azterketa, hautaketa eta laginaren sailkapena egin eta informazio guztia erabiltzaileari bidaltzen dio. Eriren senitartekoen kasuan informazio hau medikuei bidaliko diete jarraipena egin dezaten.

5.2.1.1 NFC (Near Field Communication) txartela

NFC kaberik gabeko, irismen txikiko eta maiztasun handiko gailuen arteko datuen elkartruckerako komunikaziorako teknologia da. NFC estandarrak komunikazio-protokolo zein datuen elkartruckerako formatuak barne hartzen ditu eta ISO 14443 [226][227] eta Felica-n [228] dute oinarria. Egun, teknologia mota hau smartphone zein tabletetan modu orokortuan integratzen hasi da kontakturik gabeko diru-transferentziak egiteko.

Hurbileko datu-transferentzietarako protokolo honek 20 cm-tara gehienez dauden gailuen artean komunikatzeko gaitasuna eskaintzen du. Egungo smartphone zein tabletek integratua daramate NFC bidez komunikatzeko zerbitzua, erabiltzaileek modu erraz eta seguru batean dauden zerbitzuak erabiltzeko aukera izan dezaten. Tarte txikietan komunikatzeko aukera honek inplizituki segurtasun handiko komunikazioak ahalbidetzen ditu.

Gailuen arteko komunikazio hori posible egiteko erabiltzen diren terminalak desberdinak izan daitezke. Adibidez, gaur egun gailu elektronikoen txikiagotzea eta haien energia-beharrak oso txikiak direla eta, osasun-egoerari buruz informazioa jasotzeko gorputzean itsatsirik (wearableak) [229] dihoazen gailu elektronikoak garatzen hasi dira. Sarean konektatzeko modu berri hau, Personal Area Network edo PAN [230] bezala ezagutzen da eta maiztasun baxuko eramailea erabiltzen da energia barreiatu ez dadin eta urrutiko zelatarien arriskua gutxiagotuz. Kasu honetan, terminalak pertsonak gara.

Alzumeric-en kasuan ordea, nahiz eta wearableen integrazioak ikerketa-lerro berriak irekiko lituzkeen, lehen etapa honetan, NFC integratua duten txartelak erabiliko dira. Sistemak eskaintzen dituen zerbitzuak erabili ahal izateko, aurrez programatutako NFC txartelak izango dira. Komunikazio-mota hau aukeratzeko arrazoiak segurtasuna izan da. Erabiltzaileak NFC txartela gailura hurbildu ondoren gailuan instalatua behar duen App-a irekiko da.

5.2.1.2 Smartphone edo tablet-a

Bai Tablet edo Smartphone gailu adimendua. Honek NFC irakurgailua, HTML5 Audio API-a duen nabigatzailea, mikrofonoa eta plataformara sartu ahal izateko Alzumeric Sistemaren App-a dauzka.

5.2.1.3 App-a

Android sistema eragilea duten gailu adimenduetan instalatuko den eta .apk extensioa izango duen fitxategia da. Honek, erabiltzaile eta sistemaren artean komunikazioa irekitzeko zubi lana egingo du.

Medikuak NFC txartela gailura hurbiltzean App-a irekiko da. Horren betebeharra NFC-aren lehen segurtasun-geruza behin gaindituta, bigarren segurtasun-geruza bat gehitzea da, hots, saioa irekitzeko bigarren balidazioa.

Behin txartela gailu adimendura hurbilduta irekiko den app-ak Smartphone-aren datu batzuk jaso eta aplikazioa erabiltzeko baimenak dituen segurtatuko du.

```
element.innerHTML = 'Device Name: ' + device.name
                   'Device Platform: ' + device.platform
                   'Device UUID: ' + device.uuid
                   'Device Version: ' + device.version
```

Irudia 5. 4: Smartphone-aren datuak jaso. Terminalaren izena, plataforma, identifikadorea eta bertsioa jasotzen ditu eta datu-basean erregistratua dagoen ikusi. Komunikazioak HTTPS (Hypertext Transfer Protocol Secure) bidez egingo dira eta zuzena bada formularioa betetzen utziko dio.

Dena zuzena bada bete beharreko bi eremu dituen formularioa irekiko da:

- Lehen eremurako, medikuak, NFC txartelaz gain, atxikitua doan pasahitza pertsonala eta simetrikoa [231] dauka. App-ak pasahitza eskatzen du, eta hori balidatu berri den NFC txartelaren izenari lotua dago. Pasahitz hori erabiltzaileari NFC txartelarekin batera beste segurtasun txartel batean emango zaio. Horrek ausazko metodo baten bidez sortutako 16 karaktere dituen (zenbaki + letra larri eta xehea + karaktere bereziak) pasahitza du.
- Bigarren eremua betetzeko, erabiltzaileak NFC telefonora hurbiltzean, automatikoki, zerbitzariak NFC-aren identifikagarria (ID) jasoko du. Balidazio guztiak egin ondoren eta dena zuzena bada, erabiltzaileak lehen eremua betetzen duen bitartean, zerbitzariak SMS-a (*Short Message Service*) bidaliko dio (desberdina saio bakoitzean) txartel horretara lotua duen telefono zenbakira saio berrira egokitutako token-arekin. Patente asko daude eta oso desberdinak segurtasun-neurri horien inguruan [232][233][234]. Gure kasuan, token propioa garatuko dugu. Segurtasun token hori formularioaren bigarren eremuan txertatuko du.

Gako simetrikoa [231] eta tokena txertatuak daudenean, programatutako NFC txartelaren ID-a (identifikagarria, zenbaki eta letraz osatutako string-a izango da) gehituko zaio zerbitzarira bidaliko den string-ari. Hiru argumentuak string bakarrean HTTPS [235][236] bidez bidaliko dira. Saioa hasteko, HTTPS geruzaz gain, Alzumeric-entzat bereziki garatutako eta “bakarlari kuantiko” izena eman zaion bigarren segurtasun-geruza bat ere badu. Segurtasun-geruza horren funtsa informazioaren zifratutako algoritmo segurua [237] da, Bruce Schneier-ek Neal Stephenson-en *Cryptonomicon* (2000) [238] nobelarako sortutakoa da. GNU General Public License lizentziapean irekitako algoritmo horrek [239] legeria aldetik dituen mugei buruz honakoa aurkitu da [240].

Zerbitzariak ID-ari lotetsia doakion gako simetrikoa erabiliz string-aren deskodetzea egin eta dena zuzena dela balioztatu ondoren, saioa zertifikatu guztiakin hasteko oniritzia bidali eta gailuak, automatikoki, sistema erabiltzeko aukera emango dio. Plataforma ireki eta saio berria hasiko da.

5.2.1.4 Segurtasuna eta pribatutasuna

Datuen pribatutasuna eta konfidentzialtasuna garrantzitsuak dira osasunari buruzko ikerketetan, eta horregatik, segurtasunari berebiziko garrantzia eman zaio plataforma honetan.

Alzumeric-entzat bada espreski sortutako zifratu algoritmoa hasierako balidaziorako erabiliko dena. Bakarleri kuantikoa du izena eta hurrengo puntuan gehiago sakonduko da honi buruz.

Especialistak txertatuko duen gako simetrikoaren zifratu eta deszifratuaren algoritmo hori Man in the Middle [241][236] motako erasoaldien aurrean zaurgarritasunak ekiditzeko sortu da. Bakarleri kuantikoak zutarria bakarleriaren algoritmoan du [237].

Datuen segurtasuna gehiago babesteko, SSL zertifikatupeko HTTPS [235] komunikazioak gehituko zaizkio. Aplikazio-geruzan exekututzen den bakarleri kuantikoa saio seguruak segurtatzeko garatu da. Garraio-geruzan ordea, komunikazio guztiak SSL (Secure Sockets Layer) segurtasun ziurtagiriaren bidez kodifikatzen dira. Horrek esan nahi du plataformaren komunikazioa guztiak HTTPS [235] bidez egiten direla.

5.2.1.5 Bakarleri kuantikoa

Alzumeric-en zifratu algoritmo honen jatorria, Bruce Schneier-ek Neal Stephenson-en Cryptonomicon (2000) [238] nobelarako sortutako bakarleriaren zifratu algoritmoan [237] dago. Erwin Schrödinger fisikari austriarraren katuaren esperimenduaren paradoxarekin hitz-jokoa egin eta mekanika kuantikoaren gainjartze printzipioari [242][243] metaforikoki bere lekua eman diogu algoritmoan, aurrerago ikusiko dugun bezala.

Zifratu eta deszifratuak simetrikoak dira. Gako sekretu edo kriptografia simetriko [231] izeneko metodo kriptografikoek gako sekretua erabiltzen dute informazioa zifratu eta deszifratzeko. Komunikatuko diren bi muturrak ados jarri behar dira erabiliko duten gakoari buruz. Behin gakoa dutela, igorleak informazioa gakoa erabiltzen zifratuko du eta hartzaileak alderantzizko prozesua egingo du, datuak irakurri ahal izateko.

Algoritmo horrek *keystream* deitutako zenbaki-segida sortzen du testua zifratzeko eta horrela funtzionatzen du:

- Zifratu nahi den informazioa hartu eta bere zenbaki-segida baliokidea idazten du. $A = 00, B = 01, C = 02, \dots, Z = 25$;
- Alfabetoak 26 letra dituela kontuan hartuz eta gainjartze kuantikoaren [242][243] “metafora” bat eginez, 13 letrako 2 array birtual sortzen ditu. Horrekin hizki bakoitzak bi egoera posible dituela (A eta Z, B eta Y, eta abar) ezaugarritu nahi da, eta

5. KAPITULUA: ALZUMERIC

Schrödinger-ren paradoxa erabiliz, neurtzen den arte, bietako edozein balio behagarri har dezakeela islatu nahi da. Orduan, letra bakoitzak array-ean duen posizioaren baliokide simetrikoa bilatu eta balioa esleitzen dio (adibidez, A-k Z balio du). Pausoa hizki guztiekin errepikatzen du, bai informazioa eta pasahitza simetrikorako.

- Pasahitza simetrikoarekin pauso bera jarraitzen da pasahitza zifratzeko.
- Informazioa eta pasahitzaren zenbaki-segidak gehitzen dira. 26 baino gehiago gehitzen badute emaitzari 26 kentzen zaio. Adibidez: $00 + 25 = 25$; $18 + 25 = 17$;

$$\begin{array}{r} 251400051321081723211223080110062122 \\ + \\ 252525062325252525220707252525252525 \\ = \\ 241325111020071622171904072609052021 \end{array}$$

Irudia 5. 5: Bakarlari kuantikoa. Informazioa eta pasahitza bi digituka gehitzen dira. Elkarren artean 26 baino gehiago gehitzen badute emaitzari 26 kentzen zaio. Adibidez: Ezkerreko lehen bi zenbakien batura $25 + 25 = 50$ da. 26 baino gehiago denez ken 26 eragiketa egiten da; beraz, $50 - 26 = 24$.

- Jadanik zifratua dagoen mezu baliokidea atera eta zerbitzarira bidaltzen da.

ZIFRATUA:

Hasierako mezua: `alzumericrypted`

Mezuaren zenbaki baliokidea: 001125201204170802041302172415190403
Mezuaren superposizioa: `zoafnvirxvmxibkgvw`
Mezuaren superposizioaren zenbaki baliokidea: 251400051321081723211223080110062122

Pasahitza simetrikoa: `a_Ytc34%&dss$$44##`
Pasahitza simetrikorekin zenbaki baliokidea: 0000190200000318180000000
Pasahitza simetrikorekin superposizioa: `zzgxzzzwhhzzzzzz`
Pasahitza simetrikorekin superposizioaren zenbaki baliokidea: 252525062325252525220707252525252525

Mezu zifratuaren zenbaki baliokideak (baturak): 241325111020071622171904072609052021

Mezua zifratua: `ynzlkuhqrtehajfuv`

Irudia 5. 6: Bakarlari kuantikoa II. Adibide honetan bidali nahi den mezua “alzumericrypted” da eta pasahitza simetrikoarekin batera zifratu ondoren, garraio-geruzan zehar zerbitzarira “ynzlkuhqrtehajfuv” mezua bidaltzen da.

- Zerbitzariak, jasoko duen mezuaren zenbaki baliokidea aterako du.
- Transmittitu nahi den informazioa pasahitza simetrikorekin bidez zifratu da eta zerbitzariak pasahitza bera erabiliko du mezua irakurri ahal izateko. Komunikazioak Man in the Middle [241][236] erasoan aurrean gakoaz babestua egongo da. Zifratu

sistema on batek segurtasuna pasahitzan zein algoritmoan jartzen du. Erasotzaileak algoritmoa jakin arren pasahitza ez badaki gauza handirik ezin dezake egin.

- Bi zenbaki segiden arteko kenketa egiten du. Emaizta negatiboa edo zero ateratzen bada 26 gehitzen dio mezua irakurri ahal izateko.
- Bukatzeko, gainjartze kuantikoaren [242][243] “metafora” bat egin eta jatorrizko mezua lortuko du.

DESZIFRATUA:

Mezua zifratua: ynzkuhqwrtehajfuv

Mezu zifratuaren zenbaki baliokideak (baturak): 241325111020071622171904070009052021

Pasahitza simetrikoa: a_Ytc34%&dss\$\$44##

Pasahitza simetrikoaren zenbaki baliokideak: 0000190200000318180000000

Pasahitza simetrikoaren superposizioa: zzzgxzzzzwhhzzzzzz

Pasahitza simetrikoaren superposizioaren zenbaki baliokidea: 252525062325252525220707252525252525

Mezu deszifratuaren zenbaki baliokideak (kenketak): 251426051321081723211223080110062122

Mezu deszifratuaren letra baliokideak: zoafnvirxvmxibkgvw

Mezu deszifratua: alzumericencrypted

Irudia 5. 7: Bakarleri kuantikoa III. Zerbitzariak pasahitza simetrikoa izan behar du jatorrizko mezua irakurri ahal izateko. Adibide honetan jasotzen duen mezu zifratua “ynzkuhqwrtehajfuv” da eta pasahitza simetrikoarekin “alzumericencrypted” mezua irakurtzea lortzen du.

Bakarleriaren algoritmoak poker karta-sortarekin sortzen du segida. 54! ordenamendu posible daude poker karta-sorta batean, 236 bit gutxi gorabehera.

Eredu kriptografiko batek FIPS 140-2 baldintzak bete beharko lituzke eredu kriptografiko balioztatu bezala ziurtagiria jaso dezan. Horretarako, arrakastaz bete behar du algoritmo kriptografikoen balioztatze-prozesu osoa [244].

5. KAPITULUA: ALZUMERIC

Taula 5. 1: Keyed-Hash Message Authentication Code. Adibidez, irudian ikus dezakegu Keyed-Hash Message Authentication Code (HMAC) algoritmoaren inplementazioa balioztatzeko arrakastaz exekutututako ezaugarriak. Proba horiek NVLAP ziurtagiria duen (Cryptographic And Security Testing , CST) laborategiek egin dute [245].

HMAC-SHAx HMAC-SHA3x	HMAC with supporting SHA or SHA3 implementation.
KS	Key Size
BS	Byte Size
SHS: Cert.#[number]	Corresponding Secure Hash Standard validation number on the SHS Validation List .
SHS3: Cert.#[number]	Corresponding Secure Hash Standard 3 validation number on the SHA3 Validation List SHS Validation List .

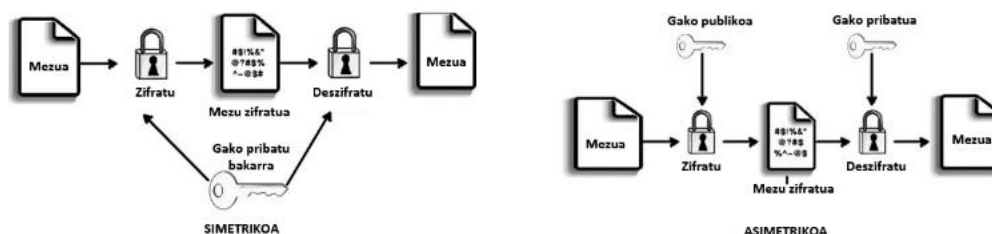
5.2.1.6 HTTPS

Komunikazio guztietarako beste HTTPS [235] segurtasun-geruza bat gehitzen da. Horrek igorpen datu guztiak zifratzen ditu seguruagoak izan daitezen.

HTTPS (Hypertext Transfer Protocol Secure) komunikazioetarako 443 portua erabiltzen da eta SSL/TLS[246] (Secure Sockets Layer/Transmission Layer Security) bidez pakete zifratuak erabiltzen ditu bezero-zerbitzariaren arteko kanal seguruak sortzeko.

Bezero eta zerbitzariaren arteko komunikazioak zifratzerako orduan Man in the Middle [236][241] edo *eavesdropping* [247] bezalako erasoak ekiditzen dira. HTTP OSI [248] ereduaren aplikazio-geruzan (7. geruzan) dago, HTTPS, ordea, geruza baxuenetan. Gako publiko eta pribatuaren, hots, kriptografia asimetrikoaren[231] sistema erabiltzen du. Honek esan nahi du informazio-transakzio guztiek bezero-zerbitzariaren arteko gakoien trukean dutela zutarrria.

Zerbitzarira mezu zifratua bidaltzeko, horren gako publikoarekin zifratzen da, ondoren, zerbitzariak berak gako pribatuarekin deszifra dezan. Datuen transferentzian erabiliko diren xehetasun teknikoak zehaztu ondoren (protokolo bertsioa, zifratutako algoritmoak eta abar), bezeroak, gako publikoa erabiliz, mezu zifratuko du eta bakarrik zerbitzariak duen gako pribatuaren bidez deszifra dezake. Datuen zifratuak egin behar izateak eta honek dakarren datuen bolumenaren hazkundeak garraioa moteltzen du ordea.



Irudia 5. 8: Zifratu-deszifratu simetrikoko eta asimetrikoko. Ezkerreko irudian gako simetrikoa erabiltzen duen zifratu sistema bat (Bakarlarri kuantikoa). Eskuinean zifratu-deszifratu asimetrikoko erabiltzen duen sistema bat (HTTPS) [249].

5.2.1.7 Webservices

Batzuetan, teknologia-mota desberdinak erabiltzen dituzten aplikazioek batetik, eta programazio-lengoiaren garatutako softwareek edota plataformetan exekutatu eta sareetan dauden plataformek bestetik, elkar ulertu behar dute, arkitektura konplexuago baten zati baitira. Jatorri desberdina duten aplikazioek komuna izango den hizkuntza hitz egin behar dute elkar ulertzeko.

Alzumeric-en egitura konplexua da eta teknologia desberdinak biltzen ditu. Plataformaren barruan zenbait zerbitzu eskaintzeko garatutako teknologia horiek elkarren artean ulertzea garrantzitsua da. Horretarako, web zerbitzuak erabili dira.

Web zerbitzuak [250] aplikazioen artean datuak elkartrukatzeko zenbait protokolo eta estandar-talde erabiltzen dituen teknologia da. OASIS[251] eta W3C[252] web zerbitzuen arautegiaren eta arkitekturaren batzorde arduradunak dira. Web zerbitzuen arteko barneko jarduna hobetzeko WS-I[253] erakundea sortu da eta informazioaren elkartruckerako erabiliko diren estandarrak zehazteko betebeharra du.

Alzumeric-ek arkitektura konplexua du, eta hori dela eta, web zerbitzuak (webservices) erabili behar izan dira komunikazio-zubi, azpiegituren artean datu elkartrukea ahalbidetzeko. Horretarako, API-ak (*Application Programming Interface*) erabili dira.

5.2.1.8 API

API-a (Application Programming Interfaces) software bat beste batekin komunikatzeko zehaztaperen formala da, hots, agindu, funtzio eta protokolo informatikoen multzoa da, eta sistema eragileetarako programa zehatzak sortzea ahalbidetzen die garatzaileei. Datuen elkartrukea errazten dute, eta aurretik zehaztutako funtzioak erabiliz, beste sistema eragile zein aplikazioekin elkar eragin dezake. Beste hitzetan, aplikazioak eraiki ahal izateko liburutegi batzuk eskaintzen dituen errutina, datu egituratu, klase eta protokoloen multzoa da.

Bai web zerbitzuak bai API-ak komunikaziorako tresnak dira. Web zerbitzuei dagokienez, HTTP da protokolo erabiliena nahiz eta SOAP, REST eta XML-RPC ere jasaten dituen. API-ek edozein komunikazio bitarteko erabili dezakete, adibidez, dll fitxategiak, Linux-en etenaldiak eta abar.

Alzumeric-ek HTML5-en audio API-a, JSON [254] bidez sortutako diagnostikoen API-a, Jupyter-ren notebook-ak berak integratua duen API-a eta abar erabiltzen ditu.

5.2.1.9 LAMP

Linux pean exekututzen den interneterako azpiegitura da eta lanabes batzuk ditu:

1. Kode irekiko Apache [255] HTTP zerbitzaria
2. MySQL [256], erlaziopeko datu-baseen kudeaketa sistema. GPL/Lizentzia Komertziala lizentzia dualapean garatua.

3. PHP [257], zerbitzariaren aldetiko programazio-lengoaia. Eduki dinamikoak dituzten webguneak garatzeko sortutakoa eta PHP lizentziapean argitaratua. PHP 7 bertsioarekin, beste gauza batzuen artean, prozesu-abiadura azkarragoak lortzen dira.

PHP saioen deszifratuaz eta horien kontrolaz arduratzen da. Grabatutako WAV fitxategiak aurretik erabakitako direktorioan gordetzen ditu. MySQL datu-basean, erabiltzaileen jardunaren, sortutako fitxategien eta warning-en zein sistemaren akatsen historikoa gordetzen du MySQLi [258] konexioen bidez. Cross Site Scripting (XSS) [259] eta SQL injekzioen[260][261] kontrako filtroak ere kudeatzen ditu.

5.2.1.10 MATLAB

Software matematikorako lanabesa da eta garapenerako ingurune integratua (IDE) eskaintzen du. M programazio lengoaia propioa du eta jabegopeko The Mathworks lizentziapean garatua izan da.

Dagokion path-ean audioak jaso eta horien matrizeetatik ezaugarriak atera eta iragazten ditu ezaugarri-bektoreak sortuz. Dozenaka parametro (pitch, intensity, jitterLocal, maenNHR eta abar) aterako ditu, eta hauek, string moduan, koma batez bereziak eta csv hedapena duen fitxategi batean utziko ditu dagokion direktorioan. Aurretik iragazkia pasatzen die adierazgarriak diren parametroek bakarrik osatu dezaten ezaugarri-mapa berria.

5.2.1.11 WEKA

Machine Learning/Deep Learning egiteko Softwarea da. Waikato-ko Unibertsitateak garatua eta GNU-GPL lizentziapean banatua.

MATLAB-ek egindako ezaugarri adierazgarrienen hautaketa-lanean ere laguntzen du. Sortutako ezaugarri-bektore iragaziak hartu eta CNN, KNN, SVM eta MLP-etatik pasa eta sailkatzen ditu (MCI/AD edo CR). Eredua, aurretik, kasu desberdinetarako entrenatua izan da, proiektuan engaiatutako erakundeen arau etikoak kontuan hartuta sortutako datu-baseei esker (PGA-OREKA, MINI-PGA, AZTIAHORE eta AZTITXIKI).

5.2.1.12 Python

Interpretatua den programazio lengoaia eta Python Software Foundation License lizentziapean argitaratua.

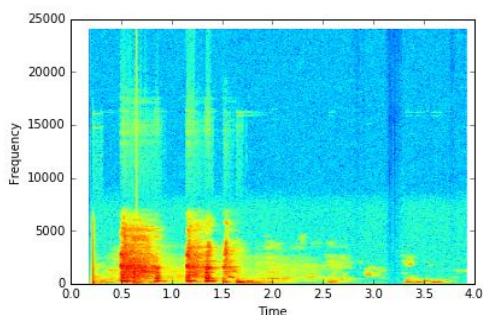
Pooling bidez CPU-ak, WEKA-k (Sailkapena MCI/AD edo CR) eta MATLAB-ek (Ezaugarrien erauzketa eta hautaketa) dagokion path-ean utziko duten informazio berria bilatzen du. Informazio hori guztia (espezialistentzat baliagarriak diren

parametroak eta sistemaren sailkapena) erabiltzailearen interfazera bidaliko du. Informazioa HTTPS protokoloan paketatua bidaliko da. Pooling teknika hori balidaziorako nahikoa izango da, nahiz eta komenigarria izan daiteke etendurak bezalako aukerak aztertzea baliabide gutxiago xurgatuko dituen sistema garatzeko bidean.

Paraleloan beste exekuzio haria, beste thread-a [262], irekia dauka, eta horrek Alzumeric plataformaren jarduera guztia miatzen du. Exekuzio hari horiek zerbitzarian monitorizatzen dira Jupyter-ren[263] shell interaktibo eta bere notebook-aren bidez. Horrela, administratzaileak Alzumeric-en prozesuak kontrolatu ditzake denbora errealean eta paraleloan. Monitorizazioa ipynb fitxategien bidez egiten da, eta plataformaren jarraipenerako, batera, fitxategiak ireki eta itxi daitezke. Prozesuaren historikoak gorde daitezke.

Berez, ipynb extensiopeko notebook-ak JSON [254] dokumentuak dira. Webservice modura, Alzumeric-en API REST-a garatzeko eta beste aplikazioetarako hedagarria izateko erabili daitezke.

```
Frecuencia de muestreo: 48000
Tamaño del array: (188416,)
Número de registros del array: 188416
Número de filas del array: 188416
Valores del array esquematizados: [ 0  0  0 ..., -4  3 11]
50 muestras del array: [4402 6301 7580 8121 7896 7071 5927 4693 3575 2680 2000 1543 1269 1145 1244
1567 2075 2724 3393 3919 4162 4030 3547 2830 2021 1362 1097 1251 1804 2721
3763 4724 5588 6318 6879 7384 7828 8098 8152 7896 7252 6319 5225 4117 3211
2616 2389 2540 3018 3739]
Valor máximo del array: 18015
Valor mínimo del array: -18013
*****
*****
*****
```



Irudia 5. 9: Alzumeric-en monitorizazioa. Alzumeric-en monitorizazioa denbora errealean. Kasu honetan, erabiltzaile batek ataza egin ostean, ahots-lagina zerbitzarira bidali, eta honek, lagin horren ezaugarrien ikerketa egin, monitorizatu eta txostenak gordetzen ditu. Plataformaren sendotasuna helburu izanik, hori garrantzitsua da sistemaren akatsak aurkitu eta zuzentzeko.

Exekuzio denborak ere kontrola daitezke *timeinit* objektuaren bidez. Jarraipen hori egitea garrantzitsua da Alzumeric-en exekuzio denborak koherenteak izatea nahi badugu eta jasotzen dituen eskaera pilaketen arabera zerbitzariaren gaitasunak berkonfiguratzeko.

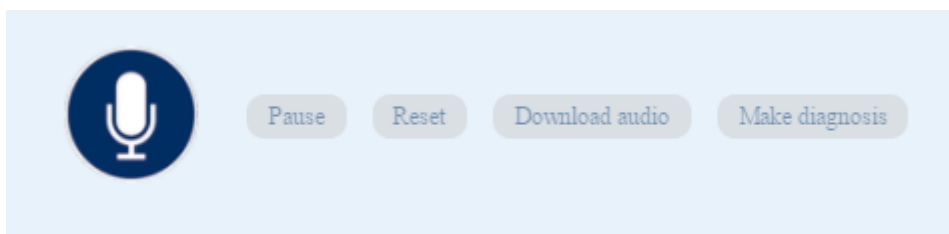
Alzumeric, SaaS (Software as a Service) [264] izeneko software eredia da, non euskarri logikoa eta datuak internet bidez sarbidea duen zerbitzari zentralizatu batean hostatzen diren. Zerbitzariaren administratzailea arduratuko da mantenu, eguneroko eragiketa eta softwarearen euskarriaz.

Machine Learning euskarria, erabiltzailearen interfazea eta kontrol monitorizazioa automatikoki kudeatzen diren zerbitzuak dira. Nabigatzaile xume batekin baimenduta dauden zerbitzu guztiak erabil daitezke. Ez da softwarea makina lokaletan instalatu behar, eta horri esker, edozein eremutara hedagarria eta unibertsala da Alzumeric.

5.2.2 Erabiltzailearen interfazea

Erabiltzailearen interfazea erabiltzailea eta zerbitzariaren arteko komunikazioa ahalbidetuko duen bitartekoa besterik ez da. Makina eta erabiltzailearen arteko exekuzio hari guztiak hartzen ditu kontuan.

Grafikoki gailuen tamainara egokitzen den sistema da eta erabiltzaile interfazea erabilgarritasuna izan du izpide. HTML eta CSS bidez diseinatu eta irudikapen estandarrak betetzeko asmoz garatu da. Erabilienak diren nabigatzaileek eusten duten HTML5/Javascript audio API-a [265] jasateko garatua izan da interfaze hau.



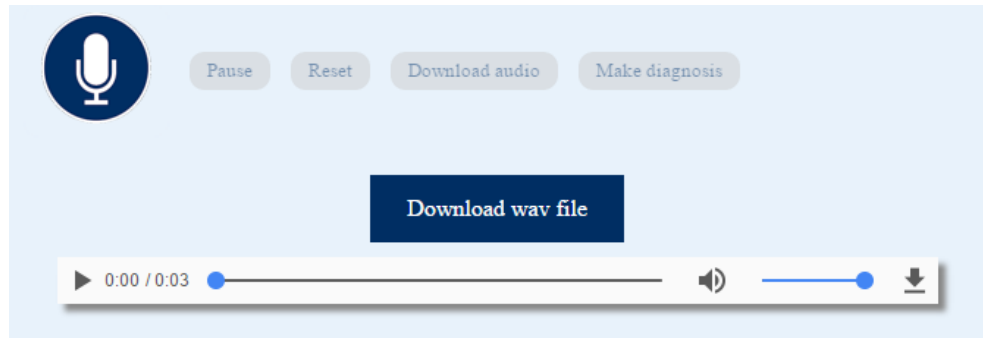
Irudia 5. 10: Alzumeric interfazea. Alzumeric interfazea. Grabaziorako botoia sakatu eta erabiltzailea (espezialista zein erriaren senitartekoa) ataza grabatzen hasi daiteke. Behin hau bukatua, makinak parametro adierazgarrienak zein laginaren sailkapena (MCI/AD eta CR) egiten ditu. Medikuek zein erriaren senitartekoez euskarri bezala erabiltzeko ekarpena izateko garatu da Alzumeric plataforma.

Medikuak botoi-segida batzuen bidez prozesu guztia kontrola dezake inongo konplexutasun teknologikorik gabe. Horien bidez, ahozko komunikazio-gaitasunen ebaluaketarako atazak egin, grabatu eta kontrola daitezke. Grabatzeko aukera hardware bidez dator txertatua, gailu adimendu gehienek audio sentsorea baitute.

Audioaren lagin maiztasuna 48 KHz eta bereizmena 16 bit-ekoa da. Kanal-kopurua software bidez programatua dago hardwarearen baldintza teknologikoak kontuan

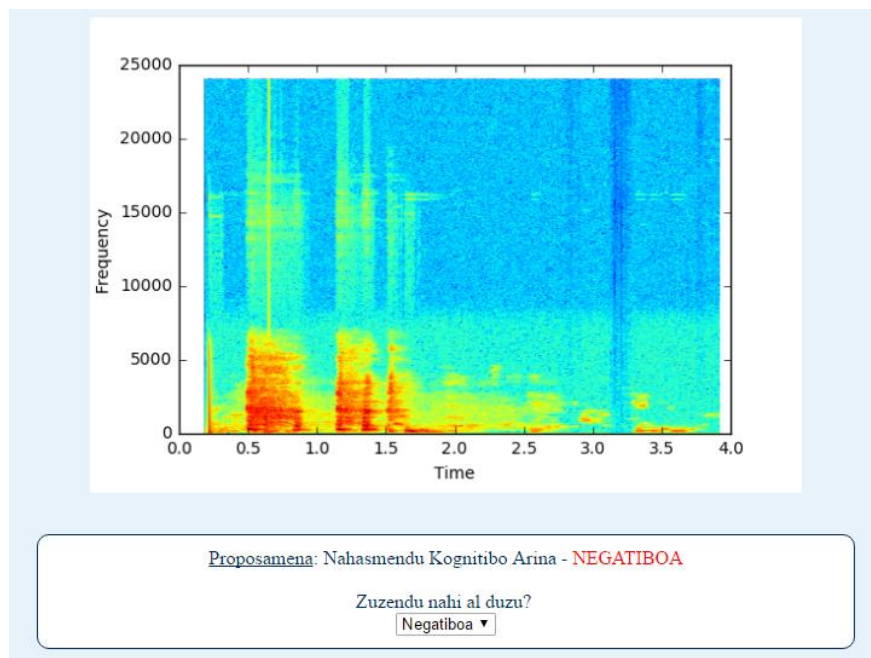
hartuta. Ataza bukatutakoan matrize numerikoak zerbitzarira bidali eta bertan eragiketa guztiak egiten dira.

Ataza horiek Apache zerbitzarira bidaltzen dira WAV formatupean eta AJAX (Asynchronous Javascript and XML) [266] komunikazio asinkronoen bidez jQuery liburutegiarekin.



Irudia 5. 11: Alzumeric-ek laginak jositzeko modua. Behin ataza bukatu eta gero, Alzumeric-ek wav fitxategia gorde eta entzuteko zein deskargatzeko aukera eskaintzen dio erabiltzaileari. Fitxategi guztiak datu-base batean gordeko dira osasun-guneen irizpide etikoak kontuan hartuta.

Makinak egindako sailkapenean asmatu ez badu, medikuak badu aukera modu asinkronoan hori zuzentzeko. Horrek ikasketa-prozesua hobetzeko aukera emango dio plataformari.



Irudia 5. 12: Ezaugarriak eta sailkapena. Sistemak lagina digitalizatu, ezaugarriak erazi eta aipagarrienak direnekin ezaugarri-bektore aberastua sortu ondoren sailkapena (MCI/AD edo CR) egiten du. Medikuentzat interesgarriak diren ezaugarriak eta sailkapena bueltan bidaltzen dizkio erabiltzaileari. Ataza etxeko ingurunean egin bada, erieen senitartekoek informazio hori medikuari helarazi diezaiokete pazientearen jarraipena egin dezan. Ataza ingurune klinikoan egin bada medikua bera izango da informazioa zuzenean jasoko duena.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

6.1 Sarrera

Atal hau sakonago ulertzeko, **B.1, B.2, B.3, C.1, C.2, C.3, C.4, D.1, D.2** eta **D.3** eranskinak irakurtzea gomendatzen da atalaren osagarriak baitira.

Oroimen-galtzearen ondoren, AD-ren arazo handienetako bat mintzamen-gaitasunaren galtzea da. Galera horren isla besteekin hitz egiteko eta ulertzeko zailtasunetan antzematen da. Ondorioz, inguruarekin eta gizarte-interakzioarekin komunikazio naturala zailago bilakatzen da. Mintzamen-zailtasun horiek direla kausa, komunikatzeko ezintasun hori gaixotasun horien hasierako faseetan agertzen da, eta AD dutenen baztertzea eragiten du, eta oso eragin ezezkorra du hala eriengan nola horien senideengan.

Lan honen helburu nagusia ebaluazio estandarraren testen azterketa automatizatu garatzea da, adibidez, narriadura kognitiboaren diagnostiko goiztiarraren testa. Hala, lan honen egitura eta progresioa honako eremu hauetara bideratuta dago:

- Ingurunea: azterketa aberastutako (etxeko) inguruetan eta ingurune klinikoetan.
- Gaitzaren sailkapena: AD vs MCI (Narriadura Kognitibo Arina). Diagnostiko goiztiarra.
- Diagnostiko-teknika ez-inbasiboak.
- Hizketarekiko independentea den sistemaren garapena.

Ondoren, 6.2 atalean, erantzun emozionala SS-n (hitz-jarioan) aztertzen da; 6.3 atalean, hitz-jarioaren ataza aztertzen da aberastutako (etxeko) inguruneetarako; 6.4 atalean, multi-hurbilketaren atazak dira aztergai; 6.5 atalak detekzio goiztiarrerako estrategia aurreratuak garatzen ditu, ikasketa sakonean eta disfluentzietan oinarrituta, aldaketa sotilak eta aplikazio errealak aztertzeko; 6.6 denbora errealeko sistema aztertzen da; bukatzeko 6.6 atalean, sistemaren erabilgarritasunari buruz egiten da azterketa.

Emaitzek gizartearentzat oso onuragarria izan litekeen geroaldi itxaropentsua erakusten dute eremu honetan.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

Bestalde, kapitulu honetan laburtzapen asko erabiliko direnez, 6.1 taulan ikus daiteke horien deskripzioa.

Taula 6. 1: Laburtzapenen deskripzioa

Akronimoak	Deskripzioa
Acc	Accuracy
AD	Alzheimer's Disease
CE	Cumulative Error
CER	Classification Error Rate
CNN	Convolutional Neural Network
CFD	Castiglioni FD
CR	Control
EF	Emotional Features
ET	Emotional Temperature
FD	Fractal Dimension
HFD	Higuchi FD
KFD	Katz FD
KNN	k -Nearest Neighbors
MLP	Multi Layer Perceptron
MLPA500.	NNHL = $\max(\text{Attribute}/\text{Number} + \text{Classes}/\text{Number})$ and TS= NNHL*10
MLP501000	NNHL = 50 and TS = 1000
NNHL	Neuron Number in Hidden Layer
SSF	Spontaneous Speech Features
SSSF	Selection of SSF
SSSF+FD	SSSF set and HFD, KFD or CFD
SVFD	Automatic Selection of VFD
TS	Training Step
VFD	maximum, minimum, standard deviation, median and mode average for full and voiced signal

Erauzitako hainbat ezaugarri-multzo erabiltzen dira esperimendazioan eta horiek deskripzioa izatea ere lagungarria izan daiteke kapitulua hobeto ulertzeko.

Taula 6. 2: Ezaugarri-multzoen laburtzapenak

Multzoaren izena	Deskripzioa
SSF	Bat-bateko hitz-jarioaren ezaugarriak (Spontaneous Speech Features)
SSSF	Bat-bateko hitz-jarioaren ezaugarrien hautaketa automatikoa (Automatic Selection of Spontaneous Speech Features)
SSSF+FD	SSSF multzoa eta dimentsio fraktala (FD)
VFD	Maximoa, minimoa, bariantza, desbideratze estandarra, batez besteko mediana eta moda seinale osorako eta ahostun seinalerako
SVFD	VFD-ren hautaketa automatikoa: maximoa, minimoa, bariantza, desbideratze estandarra, batez besteko mediana eta moda seinale osorako eta ahostun seinalerako

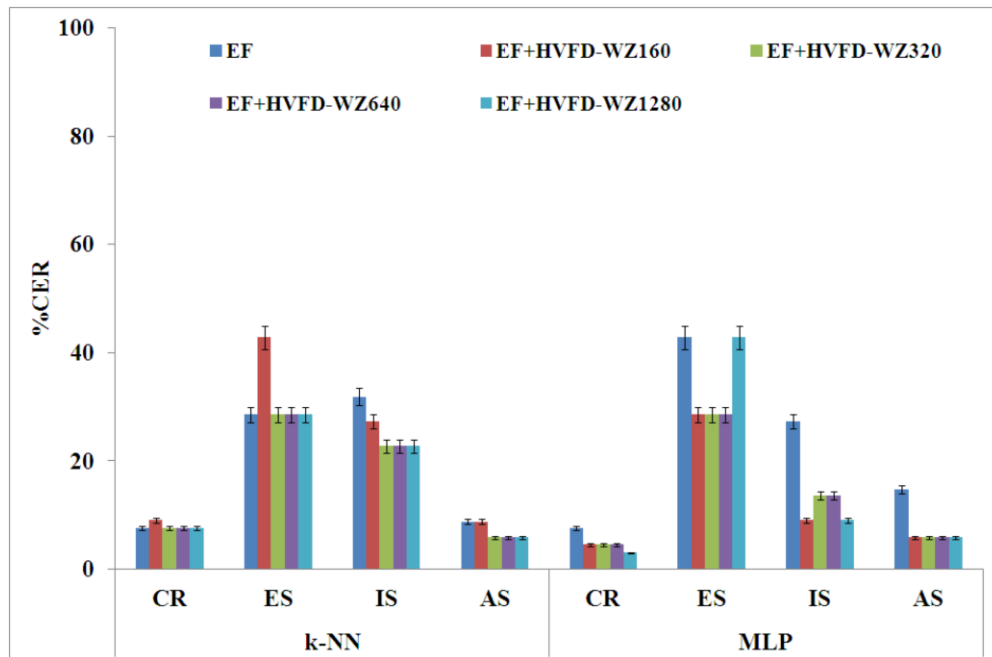
6.2 Erantzun emozionalaren azterketa

6.2.1. Materialak eta metodoak

Esperimentazioa AZTIAHORE datu-multzoaz baliatuta egiten da eta lau analisi-mota burutu dira: leiho-tamaina aukeratzea, ezaugarriak aztertzea, emaitza osoak aztertzea eta klaseen emaitzak aztertzea.

Esperimentazioan hiru ezaugarri-multzo erabiliko dira

- Ezaugarri Emozionalen multzoan (Emotional Features, EF):
 - Ezaugarri akustikoak: pitch (tonua), desbideratze estandarren pitch-a, max eta min pitch-a, intentsitatea, desbideratze estandarren intentsitatea, max eta min intentsitatea, batez besteko periodoa (period mean), periodoaren desbideratze estandarra, eta batez besteko balio koadratikoaren anplitudea (RMS, Root Mean Square).
 - Ahots-kalitatearen ezaugarriak: shimmer, tokiko jitter, Zarata/Harmoniko ratioa (NHR), Harmoniko /Zarata ratioa (HNR) eta autokorrelazioa.
 - Denbora domeinuaren ezaugarriak: ahoskabe segmentuen zatikia, ahots-hausteen maila.
- EF+VFD: EF multzoa eta Higuchi, Katz eta Castiglioni Dimentsio Fraktala (Higuchi, Katz and Castiglioni Fractal Dimension, HFD, KFD and CFD) eta bere aldaerak (Variation, V: maximoa, minimoa, bariantza, desbideratze estandarra, batzestekoa eta moda) seinale osorako eta ahoskatutako seinalerako: HVFD, KVFD, CVFD.
- EF+VFD+ET: EF, VFD eta Tenperatura Emozionala.



Irudia 6. 1: Errore Ratioaren Sailkatzea (%) EF eta Higuchi Dimentsio Fraktalaren algoritmoetarako, zenbait leiho-tamaina eta bi paradigmatan: MLP eta k-NN.

Lan honen helburu nagusia bat-bateko hizketaren ezaugarri-bilaketa da, ebaluaketa aurreklinikoan AD-ren dignostikorako testak definitu ahal izateko. Ezaugarri horiek Kontrol Taldea (CR) eta AD-ren hiru mailek (ES, IS eta AS) definituko dituzte.

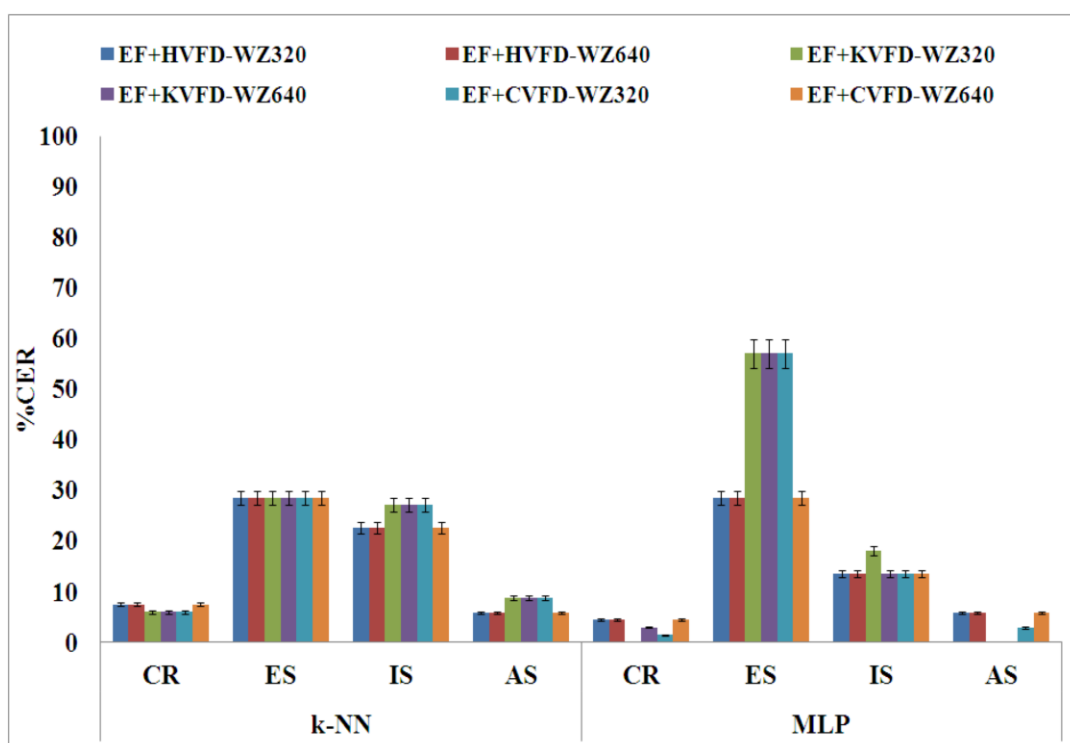
Bigarren helburua, kostu konputazionalaren optimizazioa izango litzateke ingurune errealetan erabilgarriak izan daitezzen denbora errealeko aplikazioetarako. Sailkapen automatikoa hori kontuan hartuta moldatuko da eta bi paradigmen bidez balioetsiko dira emaitzak:

- Multigeruza Perzeptrona (Multi Layer Perceptron, MLP), geruza ezkutuko neurona-kopuruarekin (NNHL) = max (atributu-kopurua, klase-kopurua) eta Training Step (TS - Entrenatze-Urratsa) NNHL*10, MLPA500.
- k-Nearest Neighbors (k-NN, Ondoko Hurbilenak) paradigma

Hasierako fase honetan, Erantzun Emozionalaren analisi automatikorako k-NN eta MLP paradigmatarako erreferentziak lortzen dira. Zehazki CER %-tan eta FD gabe. Honakoak dira: MLP % 14.73 eta k-NN % 13.96 (ikus 6.3 irudia).

Hemendik aurrera, horiek izango dira erreferentzia-balioak. Bi algoritmoetako ezagutze-ratioak onak dira CR eta AS taldeetarako eta eskasak ESrako.

Sailkapen Errore Ratioaren (CER) eta Errore Metagarriaren (CE) bidez balioztatu dira emaitzak.



Irudia 6. 2: Sailkapen Errore Ratioa (%) klaseetan leiho-tamaina ertainerako (320 eta 640) eta bi paradigmak (MLP eta k-NN Dimentsio Fraktaleko Ezaugarri-multzoekin (VFD) hiru aukeratarako HVFD, KVFD eta CVFD.

6.2.2. Esperimentazioa

6.2.2.1 Leiho-tamaina aukeratzea

Fase horretan, EF eta HVFD ezaugarri-multzoak erabili dira. 6.1 irudiak Sailkapen Errore Ratioa %-tan erakusten du, klase guztietarako (CR, ES, IS eta AS), HVFD, zenbait leiho-tamaina eta bi paradigmekin (MLP eta k-NN). Higuchi algoritmorako aztertutako leiho-tamaina honakoak dira: 160, 320, 640 eta 1280 puntu.

Dimentsioa Fraktala barne hartuta erreferentziazko emaitzekiko proposamen berri gehienek hobekuntza agertzen dute. CER minimoa MLP-rekin lortzen da. Tamaina ertaineko leihoekin seinale-eboluzio globala eta lokala (320 eta 640) jasotzeko gai dira. Hortaz, leiho-tamaina horiek hurrengo esperimenduetan erabiliko dira.

6.2.2.2 Dimentsio Fraktalaren Analisia

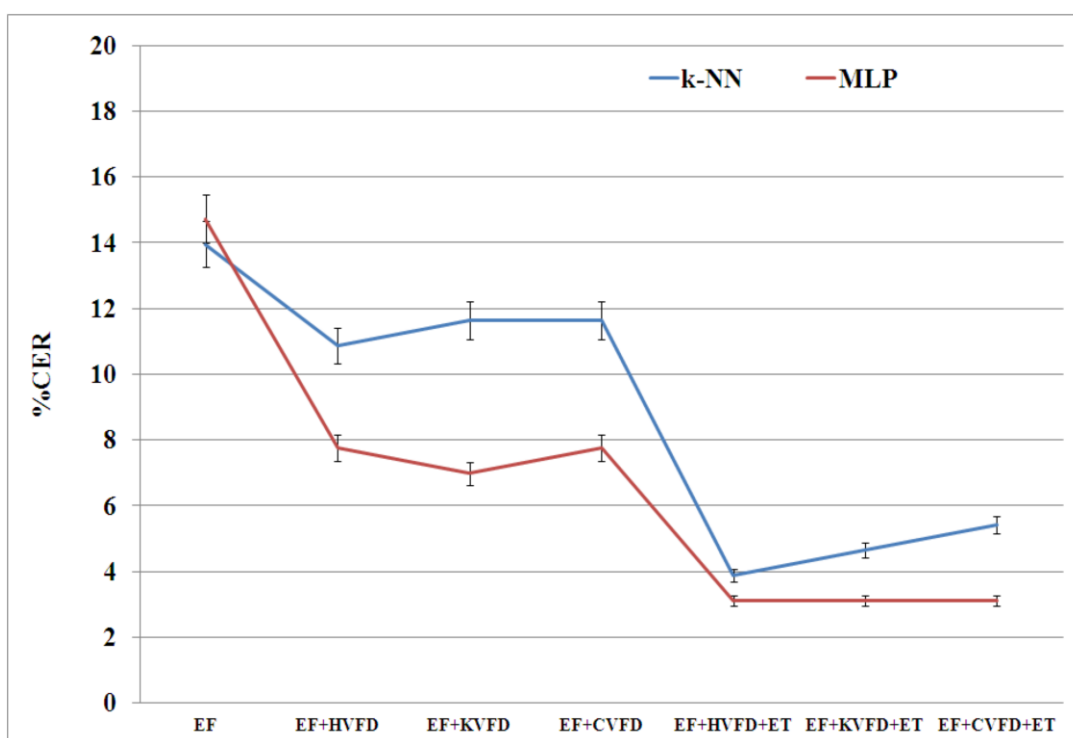
Oraingoan Dimentsio Fraktalaren ezaugarriak, paradigma-modelatzea eta leiho-tamaina aztertu dira.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

6.2 irudiak CER (%) klaseen emaitzak erakusten ditu, tamaina ertaineko leihoekin (320 eta 640), bi paradigmeekin (MLP eta k -NN) eta Dimentsio Fraktaleko hiru algoritmoetarako (HVFD, KVFD eta CVFD) ezaugarri-multzoekin.

k -NN-ri dagokionez, emaitza orokor hoberenak Higuchi algoritmoarekin lortzen dira. Funtzionamendua okerragoa da Katz algoritmoa erabilia. Kasu guztietan, kostu konputazionala gutxiagotzen da paradigma horrekin (ikus 6.15 irudia).

MLP-ren emaitza orokorreki dagokienez, CER gutxiagotzen da kasu guztietan. Emaitzarik onenak MLP algoritmoarekin lortzen dira; alabaina, diagnostiko goiztiarrari begira, 640 puntuko leiho-tamaina eta HFD eta CFD aukera hobeak dira. k -NN paradigma dela eta, horiek ere aukera hobereenen artean daude. Hortaz, 640ko leiho-tamaina erabiliko da hemendik aurrera.



Irudia 6. 3: Saillkapen Errore Ratioa (%) II definitutako hiru Dimentsio Fraktaletako algoritmoetarako, bi paradigmak (MLP eta k -NN), Dimentsio Fraktaleko Ezaugarri-multzoetarako (VFD) eta Temperatura Emozionalerako.

6.2.2.3 Temperatura Emozionala

6.3 irudiak hobekuntza (CER %-tan) erakusten du kasu gehienetan, ET barne hartzen denean. Batez ere bere MLP-rekin. MLP-k antzeko emaitzak lortzen ditu aukera guztietarako (CER = % 3.11).

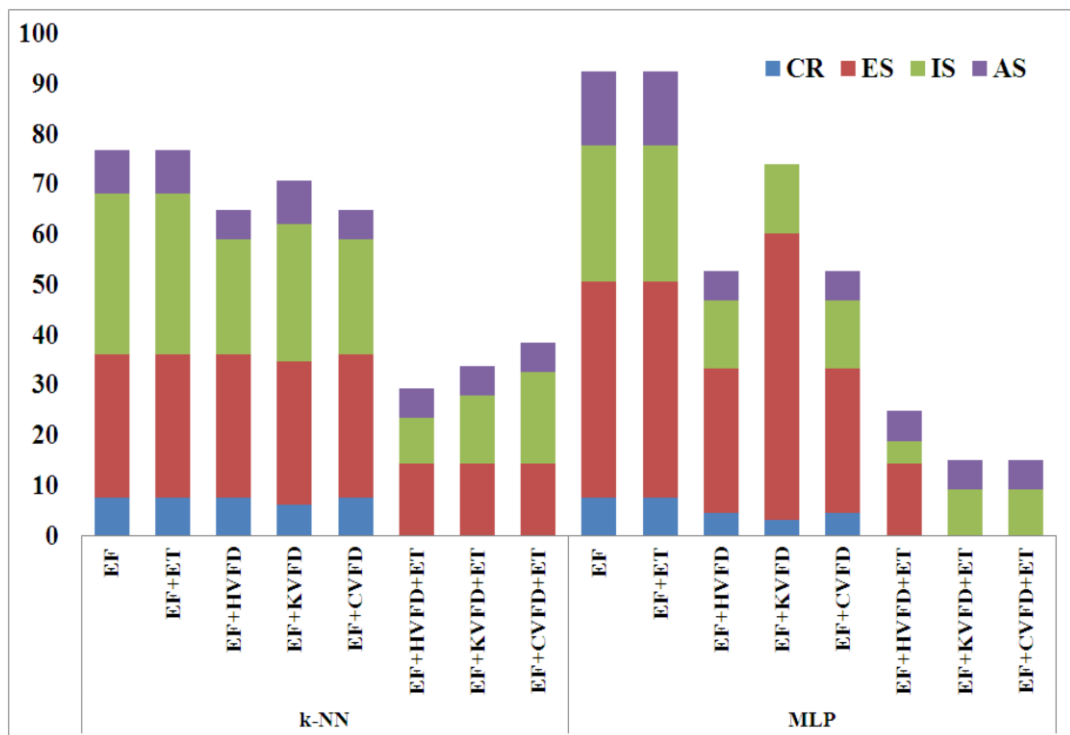
Oro har, k -NN ez da MLP bezain egonkorra eta emaitza orokor hoberenak Higuchi-rako (CER = % 3.98) eta Katz-erako (CER = % 4.66) lortzen ditu. Gainera, k -NN-en kostu

konputazionala txikiena da kasu guztietarako (ikus 6.15 irudia) eta Dimentsio Fraktaleko ezaugarriak ES segmentuak egoki detektatzeko gai da, nahiz eta MLP-ren bidez lortutako emaitzak lortzeko gauza ez izan. k-NN-k emaitzarik hoberenak EF+HVFD+ET-rekin lortzen ditu (ikus 6.3 eta 6.4 irudiak).

Bestalde, MLP-k, denbora errealeko aplikazioetarako aukeratutako konfigurazioarekin, ratio egokiak ematen ditu kostu konputazionalari dagokionean. MLP-rekin, emaitza orokor hobereana (CER = % 3.11) EF+CVFD+ET konbinazioa da gutxigatik bada ere.

Azkenik, 6.4 irudiak, klaseetarako lortutako emaitzak erakusten ditu Errore Metagarria (CE) %-tan oinarri hartuta. MLP paradigmak k-NN paradigma gainditzen du baina ez bere kostu konputazionalari dagokionez. k-NN-ri dagokionez, EF+HVFD+ET da konfiguraziorik hobereana. Edonola ere, MLP paradigma erabiltzen denean, EF+KVFD+ET eta EF+CVFD+ET-k emaitza hobekak dituzte EF+HVFD+ET-k baino. Azken bi multzo horiek dira emaitzarik hoberenak ematen dituztenak klase guztietarako.

Detekzio goiztiarrerako (ES) sailkapena oso ona da. IS klaseak maila ertaineko pazienteak hobeto bereizten ditu. Aldi berean, ezaugarri hori erabilgarria da, hizketa-segmentu patologiko eta ez-patologikoen artean diskriminatzeko gaixo bakoitzean.



Irudia 6. 4: Errore Metagarria (Cumulative Error, CE) klase guztietarako (%), bi paradigmatarako eta 640 puntuko leiho-tamainarako.

640 puntuko leiho-tamainaren aukerak CE baxuena azaltzen du, eta hasiera batean, hobeto harrapatzen ditu seinalearen dinamikak Katz eta Castiglioni-ren Dimentsio Fraktalaren aldaerak kontuan hartuta. Zehaztasuna 96.89-koa da, +3 konfiantza-tarteareekin. AD-ren hautemate goiztiarrari begira ezaugarri modelatu ez-lineala egokia da aldaketa sotilak detektatzeko.

Dimentsio Fraktalaren ezaugarriak sartzeak informazio garrantzitsua ematen du hizketa-seinaleetako ez-linealtasunari dagokionez, eta Erantzun Emozionala zuzen aztertzen laguntzen du. Hori, diagnostiko goiztiarrari begira, lagungarri izan daiteke desberdintasun sotilak ezaugarritzeko, eta ondorioz, Erantzun Emozionalean aldaketak antzemateko. Badirudi, Emozio-Tenperatura erreferentzia egokia dela Emozio-Erantzuna AD-n neurtzeko.

Ezaugarri horrek zenbait abanbaila ditu:

- Azkar eta erraz kalkulatu daiteke.
- Emozioa ezagutzeko sendotasun handikoa da.
- Segmentazio linguistikoarekiko independentea da, eta hortaz, denbora errealeko aplikazioetan eta ingurune errealetan ager daitezkeen arazoak ekidin daitezke. Areago, denbora errealeran bideratutako aplikazioak, 6.2.2-n proposatutako k-NN eta MLP konfigurazioekin, zehaztasun eta kostu konputazional egokiak lortzen dituzte.

ET bezalako ezaugarri berrien erabilerak sendotasuna ematen dio sistemari kostu konputazionala areagotu gabe. Agerian jarri behar da positibo faltsuen ehuneko oso txikia dela. Nahaste horrek ahoskatze-egitean sor daitezkeen zalantzekin zerikusia izan dezake. Segmentu patologiko eta ez-patologikoen azterketa eta sailkapena garrantzitsua da gaitzaren eboluzioa monitorizatzeko garaian.

Osasun-espezialistek sistemaren gaitasuna azpimarratzen dute, aldi berean, biomarkatzaile independenteen azterketa aurrera eramateko, bat-bateko hizketa gisa edota zenbait biomarkatzaileen azterketa integral gisa.

6.3 Hizkuntzarekiko independentea den bat-bateko hizketaren analisia

6.3.1 Materialak eta metodoak

Jatorrizko datu-basetik oinarritutako hasierako esperimendu honetan, 20 AD gaixoen azpimultzo bat aukeratu zen (68-96 adinekoak, 12 emakumezko, 8 gizonezko), AD-ren hiru mailatan banatuta (ES = 4, IS = 10, AS = 6).

Pazienteek ingelesez, espainolez, euskaraz eta portugesez hitz egiten dute. Kontrol-taldea hizketa-erantzun zabalaren ordezkari izan daitezkeen 20 subjektuk osatzen dute (20-28 urte arteko 10 gizonezko eta 10 emakumezko). Aukeratutako hizkuntzak honakoak dira: ingelesa, frantsesa, espainola, katalana, euskara, txinera eta arabiera.

Datu-base horren azpimultzoa AZTIAHORE deitzen da. Emaitzen berri ematen den hasierako esperimenduan, AZTITXIKI izeneko datu-basea (AZTIAHOREren azpimultzo bat) izan da oinarria.

Hasierako etapan, atariko esperimendua AZTITXIKI multzoan oinarrituta burutu zen, AD sintomekin zerikusia duten bat-bateko hizketaren ezaugarri zuzenak aztertzen. Esperimendua, alde batetik, kontrol-taldea ezaugarritzen duten SS-n izaten diren aldaketak eta ezaugarriak antzemateko diseinatu zen. Bestetik, AD-ren maila desberdinetako aldaketak eta ezaugarriak antzemateko diseinatu zen baita ere.

Kontrol-taldeko 5 kidek eta AD diagnostikatuta duten baina gaitzaren fase desberdinetan (ES = 1, SS = 2, AS = 2) dauden 5 erik osatzen dute taldea. AD diagnostikatutako pertsonak (2 emakumezko eta 3 gizonezko) adin nagusikoak dira guztiak.

Kontrol-taldeko kideak (2 emakumezko eta 3 gizonezko) tarteko adinekoak (M) edo adin nagusikoak (E) ziren.

Kostu konputazional apaleko aplikazioak garatu ahal izateko, ezaugarri-multzoak optimizatu behar dira. Lehen hurbilketan, hiru algoritmo erabiliko dira ezaugarri-hautatze automatikorako. Klase bakoitzari dagokion irabazia, ezaugarri edo atributu bakoitza eta haien arteko harremanak kontuan hartzeko, bereziki WEKA software-sortatik, ondoko hauek aukeratu ziren:

- SVMAttributeEval: Ezaugarri baten erabilgarritasuna balioesten du, SVM sailkatzailea erabilita. Ezaugarriak SVM-k esleitutako pisuaren ber bi-aren arabera sailkatzen dira. Multiklase arazoetarako ezaugarrien hautaketa, klase bakoitzerako eta bakarka, sailkapen-ezaugarriek gidatzen dituzte bat-versus-dena metodoa erabiliz eta, ondoren, pila bakoitzeko muturretik banatuz azkeneko sailkapena ateratzen da (WEKA).
- CfsSubsetEval: Atributuen azpimultzo baten erabilgarritasuna ebaluatzen du, ezaugarri bakoitzaren bakarkako iragarpen-trebetasuna kontuan hartuta eta baita beren arteko erreduantzia-maila ere.
- GainRatioAttributeEval: Ezaugarri baten erabilgarritasuna balioztatzen du, irabaziak klasearekiko duen erlazioa neurtzen (WEKA).

Algoritmo horiek hasierako ezaugarri-multzoetan erabiliko dira ezaugarri-multzo txikiagoak eta egokiak aukeratzeko kostu konputazional apala dela kausa.

Esperimentazioan, hiru ezaugarri-multzoko taldeak erabiliko dira (ikus taula 6.2):

- SSF eta SSSF
- SSSF+FD
- SSSF+VFD eta SSSF+SVFD

Hiru talde horietatik bost multzo ateratzen dira, zeinak ondoko esperimentuetan erabiliko diren, 6.2 taulan ikus daitekeenez (ikus baita ere 6.1 taula lan honetan erabiltzen diren laburtzapenen zerrenda osoa izateko).

Lan honen helburu nagusietakoa bat-bateko hizketan ezaugarri-bilaketa da, balioztatze aurreklinikora zuzendua AD-ren diagnostikorako testen definizioan laguntzeko. Ezaugarri horiek CR taldea eta AD-ren hiru mailek definituko dituzte.

Bigarren helburuetako bat kostu konputazionala kontrolatzea da, denbora errealeko aplikazioei zuzenduta baitago. Bi paradigmen bidez balioetsiko dira emaitzak:

- Multigeruza Perzeptrona (MLP), geruza ezkutuko neurona-kopuruarekin (NNHL) = max (atributu-kopurua, klase-kopurua) eta Training Step (TS - Entrenatze-Urratsa) NNHL*10, MLPA500.
- k-Nearest Neighbors (k-NN, Ondoko Hurbilenak) paradigma

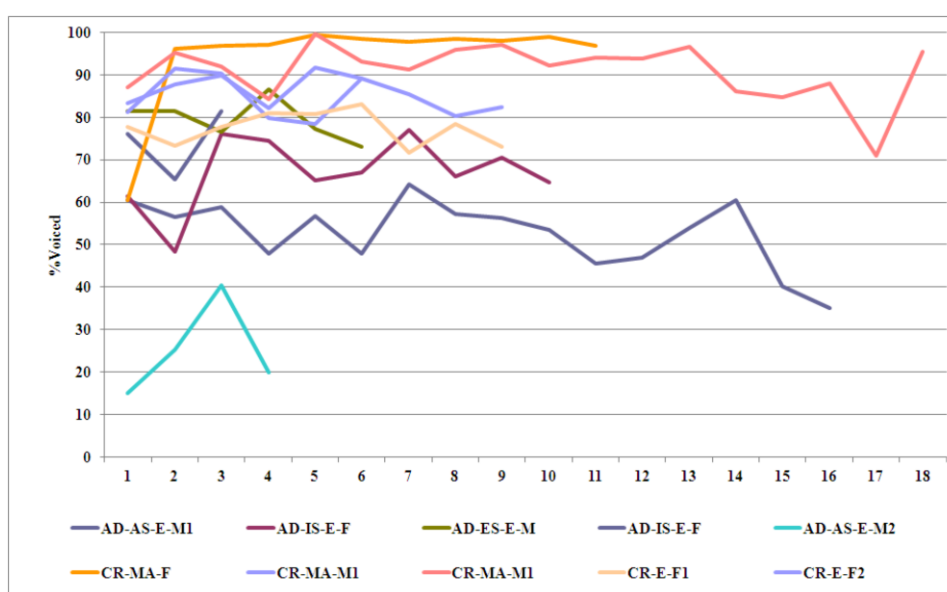
WEKA software-sorta erabili izan da esperimentuetan. Emaitzak horiek direla bide balioztatu ziren: Sailkapen Errore Ratioa (CER), Zehaztasuna/Doitasuna (Acc) eta Errore

Metagarria [267][268][269]. Entrenatze- eta balioztatze-neurketetarako, k-geruzako balioztatze gurutzatua erabili zen $k = 10$ ekin.

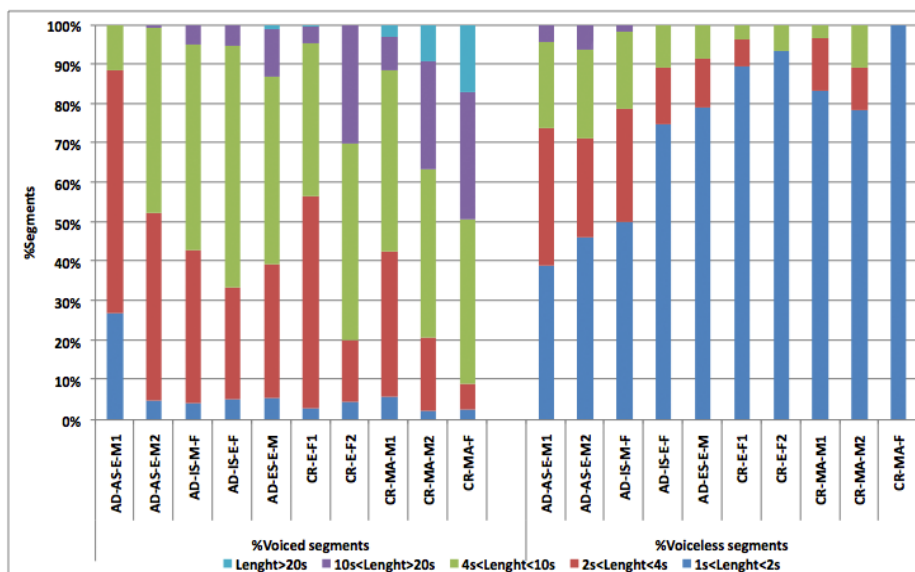
Balioztatze gurutzatu errepikatuak (WEKA ingurunerako kalkulaturik dagoen bezala) oso estatistika-test zehatzak ahalbidetzen ditu. WEKAren "Coverage of cases" (Kasuen estalera, 0.95eko maila) delakoak automatikoki ematen duen aurrez zehaztutako balioa ere erabili dugu, eta Konfiantza-Tartea (CI, Confidence Interval) % 95, % 90 eta % 80 ehunekotan ezarri da.

Esperimentuan leiho-tamaina egokia aukeratzea garrantzi handiko erabakia da. 6.9 irudiak azaltzen du leiho-tamainaren eragin kualitatiboa esperimenteren emaitzetan. Oro har esanda, Dimentsio Fraktala sistemaren dinamikak atzemateko tresna bat da. Leiho labur batekin, balioespena lokala da batik bat eta azkar egokitzen da uhin-formaren aldaketetara.

Leihoa luzeagoa bada, zenbait xehetasun galtzen dira baina Dimentsio Fraktalak hobeto aurreratzen ditu seinalearen ezaugarriak. Gainera, leiho-tamaina kontuan hartzen duten tankerako balioespenetan, kasu batzuetan, leiho luzeagoa erabilgarria izan daitekeela iradokitzen dute [177][179].



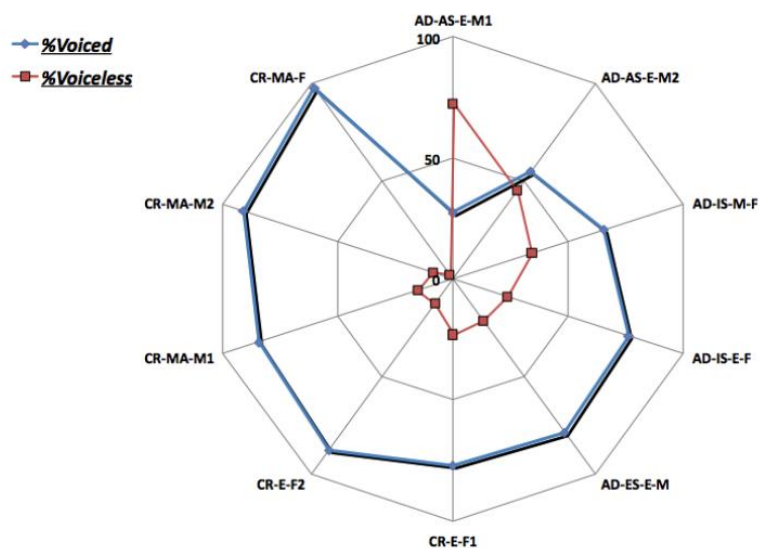
Irudia 6. 5: Bat-bateko hizketa-eboluzioa hizketa ehunekoarekiko denbora-ardatzean, ondoz ondoko segmentuetarako (S1:SN).



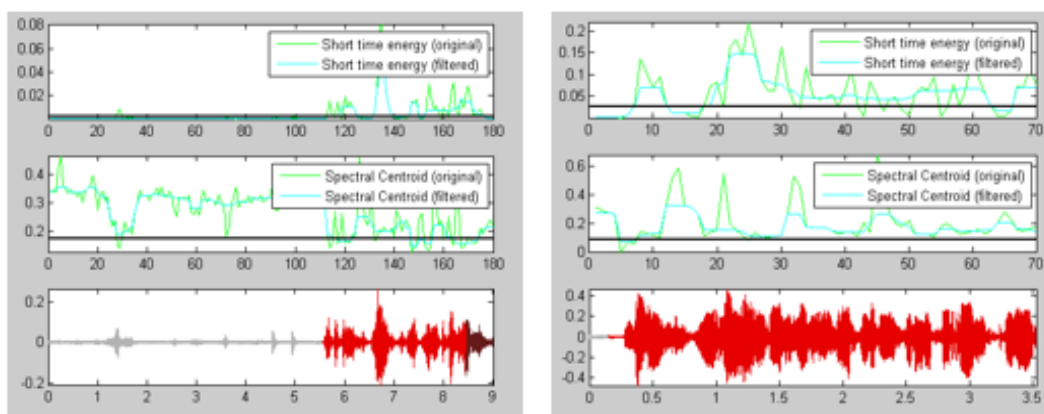
Irudia 6. 6: Hasierako esperimentazioaren ezaugarria ahostun eta ahoskabe segmentuaren (%) analisia ahoskabe segmentuaren luzerarekiko.

6.3.2 Hasierako esperimentua

Lehen proba-multzoa Bat-bateko hizketaren analisiaren (Automatic Spontaneous Speech Analysis, ASSA) esperimentuarekin osatu zen. AZTITXIKI multzoaz baliatuta honakoa ikus zitekeen: bat-bateko hizketan, AD dutenek agertu zuten ahostun ehuneko apalagoa eta ahoskabe ehuneko altuagoa (ikus 6.5) pertsona osasuntsuek baino. Horrek, hizketan, AD dutenen hitz-jarioaren galera nabarmena adierazten du. Azterketak erakusten du AD duten pertsonak ahostun segmentuen kopurua, luzera eta jarioa gutxiagotzeko joera dutela eta ahoskabe segmentuen luzera handiagotzen dutela (ikus 6.6 irudia).



Irudia 6. 7: Ahostun eta ahoskabe (%) AZTITXIKI multzorako. Tarte ahoskabeak handitu (%) eta tarte ahostunak gutxiagotzen (%) dira gaitzak ES-tik AS-era aldatzen doan heinean.



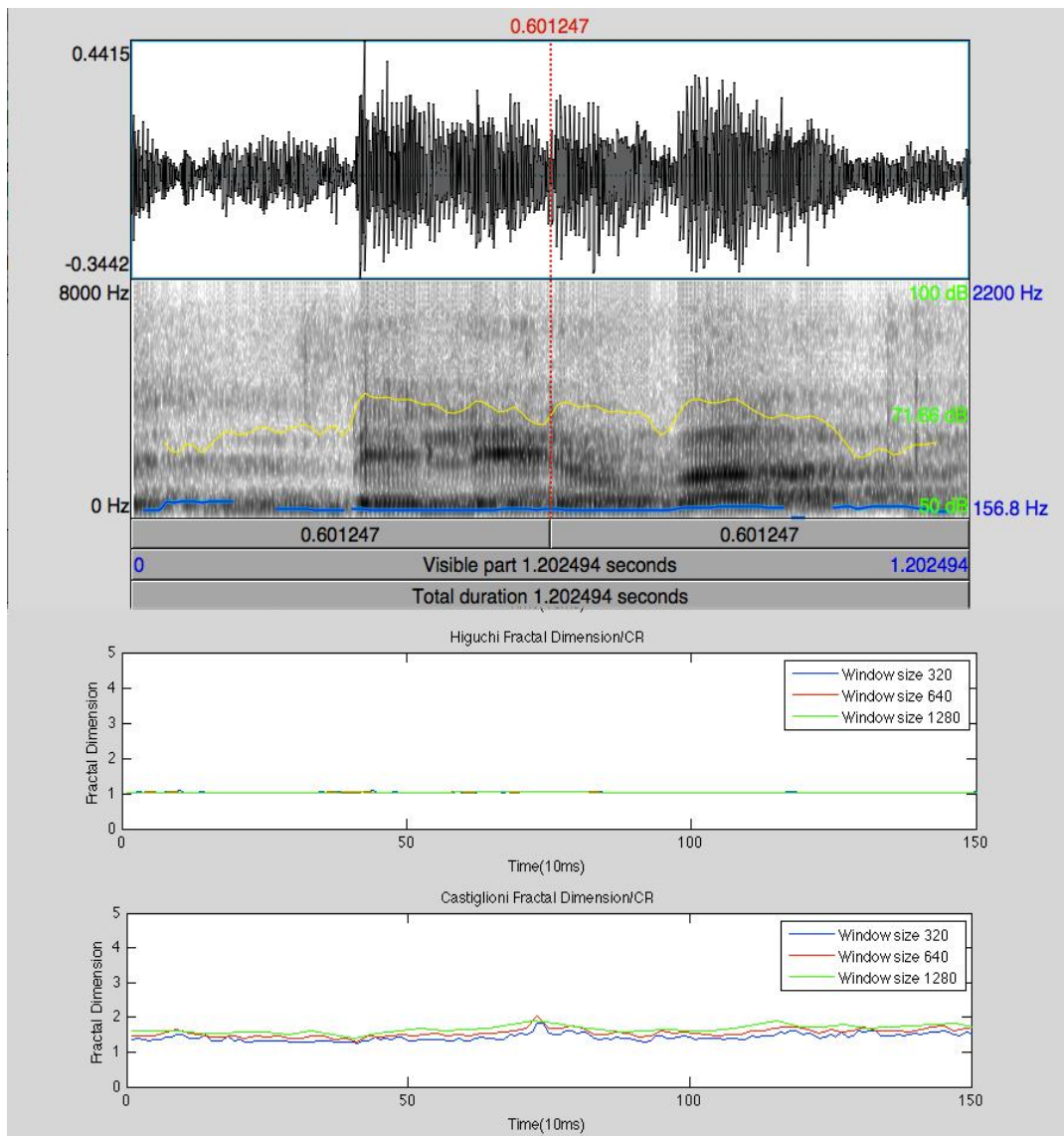
Irudia 6. 8: Hizketa Seinalearen zatiak epe Laburreko Energia eta Espekto Zentroidea, mediana iragazkiaren batez filtratuta, AD duen subjektu batentzat eta kontrol-subjektu batentzat.

Aldi berean, AD dutenen pertsonak beren bat-bateko hizketaren eboluzioan apalduz doan malda erakutsi zuten, eta apalduz doan malda hori oso nabarmena da luzera-segmentuekiko beren ahostun eta ahoskabe segmentuen azterketan. 6.7 irudiak agerian uzten du ahostun segmentu motzak eta ahoskabe segmentu luzeak erabiltzeko joera.

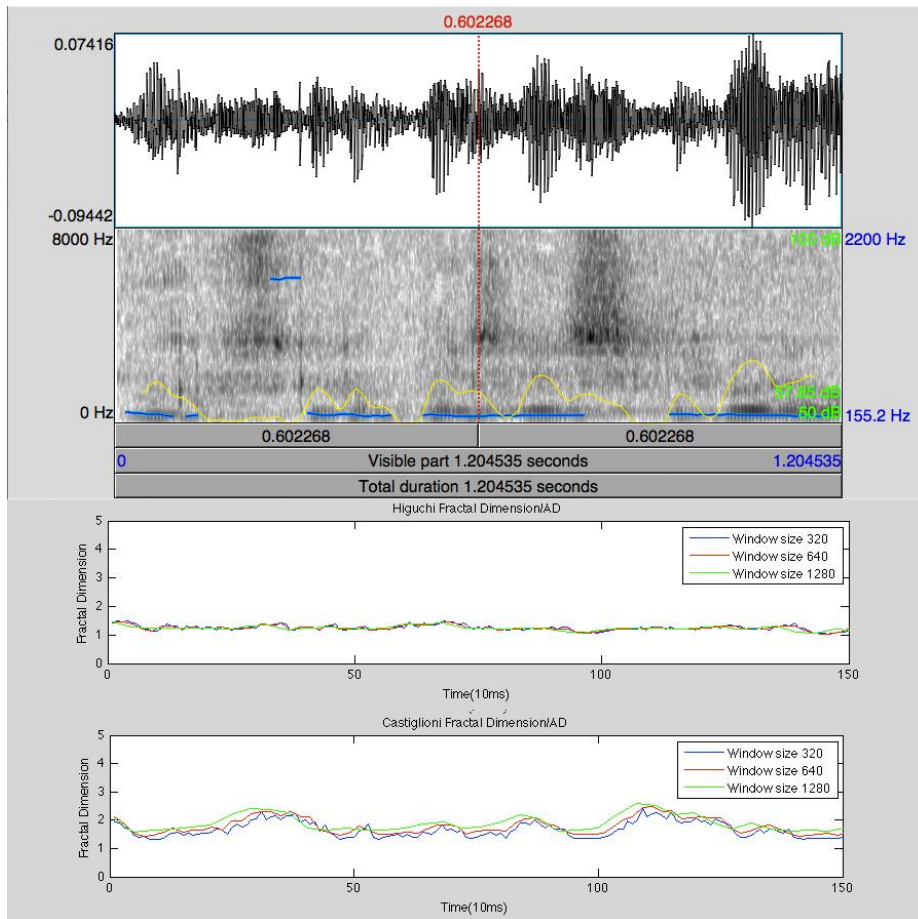
AD dutenen hizketa arina edo jariakorra da oso denbora-tarte motzetan bakarrik, eta 20 segundo baino luzeagoak diren segmentuak oso gutxitan agertzen dira beren bat-bateko hizketan.

6.8. irudiak erakusten du ondoz ondoko segmentuetarako (S1:SN), bat-bateko hizketaren eboluzioa hizketaren ehunekoarekiko denbora-ardatzean zehar. AD dutenek, denbora-ardatzean zehar, ahostun ehunekoa gutxiagotzen dutela ikus daiteke. Aldi berean, emaitzek agerian uzten dute Epe Laburreko Energia altuagoa kontrol-taldearen kideengan AD dutenengan baino eta Espekto Zentroide altuagoa AD dutenengan (ikus 6.8 Irudia).

6.3.3 Bat-bateko hizketaren analisia



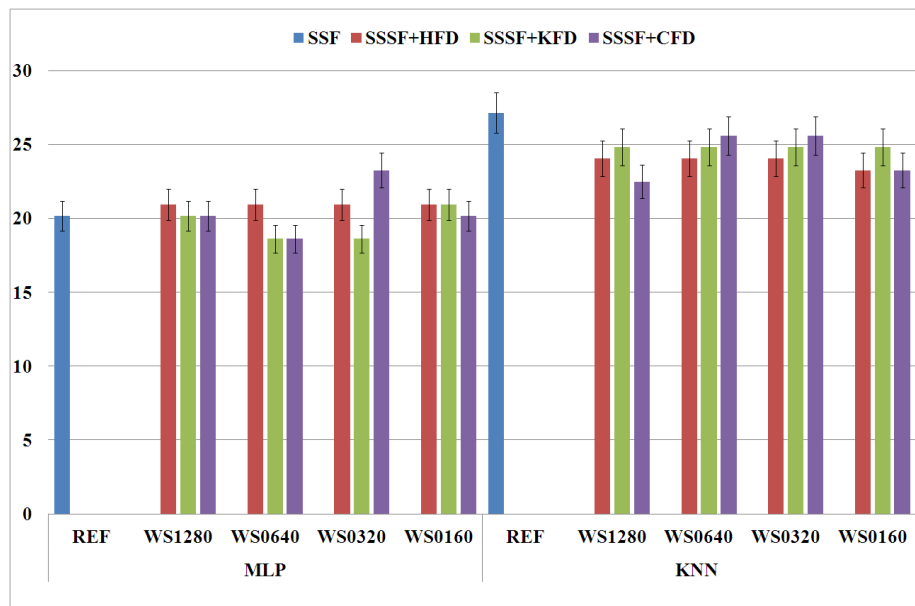
Irudia 6. 9: Kontrol-pertsona baten analisia (CR). Adierazpen bat euskaraz, “...Belero bat...”: (a) hizketa-seinalea, (b) espektrograma , pitch (urdina) eta intentsitatea (horia), (c) Higuchi, (d) Castiglioni-ren dimentsio fraktala (HFD eta CFD).



Irudia 6. 10: AD duen pertsona baten analisia. Adierazpen bat espainolez, “No sé lo que es eso”: (a) hizketa-seinalea, (b) espektrograma pitch (urdina) eta intentsitatea (horia), (c) Higuchi eta (d) Castiglioni-ren dimentsio fraktala (HFD eta CFD).

6.3.3.1 Dimentsio fraktalaren analisia

Esperimentazioa AZTIAHORE datu-multzoaz baliatuta atera zen aurrera. Lehen etapan, 80 bat atributu sortu ziren deskribatu berri diren irizpideak kontuan hartuta: bat-bateko Hizketaren Anlisi Automatikorako (ASSA) aukeratutako ezaugarriak eta dimentsio fraktalean oinarritutakoak.



Irudia 6. 11: Errore Ratioaren Sailkatzea (%) dimentsio fraktaleko algoritmoetarako, lau leiho-tamainetarako eta bi paradigmatarako: MLP eta KNN. Ref barrak dimentsio fraktalen ezaugarriak erabili gabe lortu ziren.

6.3.3.2. Ezaugarri-hautatze automatikoa

4. kapituluan deskribatutako metodologia erabiliz burutu zen ezaugarri-hautatze automatikoa. Hiru dimentsio fraktalen algoritmoetarako eta lau leiho-tamainetarako (160, 320, 640 eta 1280 puntu). Ezaugarri-multzoen 12 antzeko proposamen aztertu ziren. 50 ezaugarri aukeratu ziren jatorrizko multzoetatik. Hautatutako atributuak berdinak dira kasu guztietarako: ASSAren zerikusia duten ezaugarriak eta dimentsio fraktalean oinarritutakoak.

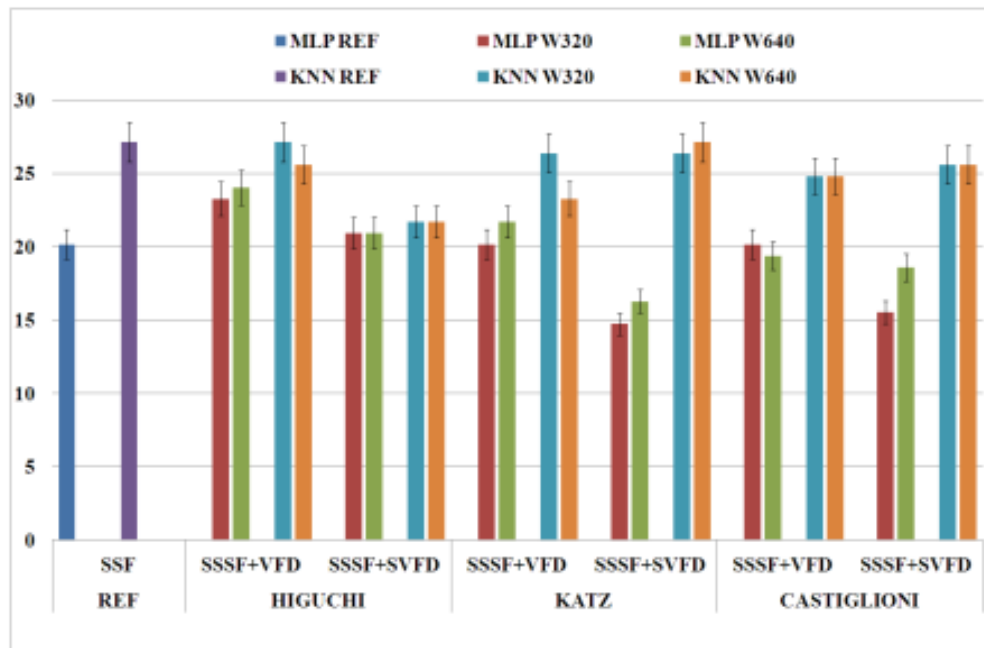
Azkenean, ezaugarri-familia hauek erabili ziren: SSF, SSFren hautaketa (SSSF), SSSF+FD, ezaugarri-multzo osoak (SSSF+VFD) eta hautatutako ezaugarriekiko multzoak (SSSF+SVFD). SSSF multzoak, % 5 batean gairitu egiten ditu gutxi gorabehera CER ratioak SSFren %-kiko (MLP, % 24.81 eta KNN, % 29.46).

Espezifikoki, erreferentziako CER emaitzak %-tan SSFrako, FD gabe, honako hauek dira: MLP, % 20.16 eta KNN, % 27.14 (ikus barra urdina 6.11. irudian). Ezagutze-ratioak bi algoritmoetarako optimoak dira CR taldeari dagokionez, baina oso kaxkarrak ES-ren kasuan. Hortaz, SSSF aukeratu zen ondoren deskribatuko diren esperimenduetan.

6.3.3.3 Sailkatze Automatikoaren ataza

Ondorengo etapan, 4. Kapituluaren deskribatutako metodologia erabili zen. Sailkatze Automatikoa zen egin beharra eta xedeak honakoak ziren: patologia neurologikorik gabeko hiztun osasuntsuak eta AD diagnostikatuta zuten hiztunak bereiztea.

Lau esperimendu-mota burutu ziren: leiho-tamaina hautatzea; ezaugarri-hautatzea; emaitza orokorrak aztertzea; klaseen emaitzak analizatzea.



Irudia 6. 12: Errore Ratioaren Sailkatzea (%) II leiho-tamaina batzuetarako eta bi paradigmoetarako: MLP eta KNN dimentsio fraktaleko multzoekin eta atributu-multzoekin, hautaketa egin eta gero (SVFD). REF barrak dimentsio fraktaleko ezaugarriak erabili gabe lortu ziren.

6.11. irudiak erakusten du Errore Ratioaren Sailkatzea %-tan dimentsio fraktaleko algoritmoetarako eta bi paradigmatarako: MLP eta KNN. Ataza honetan, zenbait leiho-tamaina aztertu ziren: 160, 320, 640 eta 1280 puntu.

6.11 irudiak agerian uzten du atributu-hautaketaren eragina dimentsio fraktaleko algoritmoei, leiho-tamainari eta ezaugarri egiteari dagokienez. MLP-rekin, CER igotzen da %-tan eta kasu guztietan. KNN-ri dagokionez, emaitza orokor hoberenak Higuchi algoritmoarekin lortu ziren eta Castiglioni algoritmoa bakarrik erabilita emaitzak kaxkarragoak ziren. Kasu guztietan, kostu konputazionala nabarmen jaitsi zen KNN paradigmaterekin (ikus 6.15 irudia).

6.12 irudian ikus daitekeen bezala, fase honetan, SSSF multzoa eta dimentsio fraktaleko multzoak erabili ziren. CER minimoa MLP eta tamaina ertaineko leihoekin (320 eta 640) lortu zen. Bi leiho-tamaina horiek erabiliko dira ondorengo esperimenduetan. Dimentsio fraktala barne hartzen duten proposamen berri gehienek erreferentziako metodoekiko hobekuntzak ekarri zituzten.

Hortaz, SSSF+SVFD multzoak erabili ziren ondorengo esperimenduetan, behin ezaugarriak hautatze automatikoaz aukeratu eta gero. Ikerketa horren emaitza orokorrak onak izan dira. Ezaugarri fraktal berriek CER gutxiagotzen dute MLP-n, % 20.16-tik % 14.73-ra, eta KNN-n % 27.14-tik % 21.71-ra (ikus 6.12 irudia).

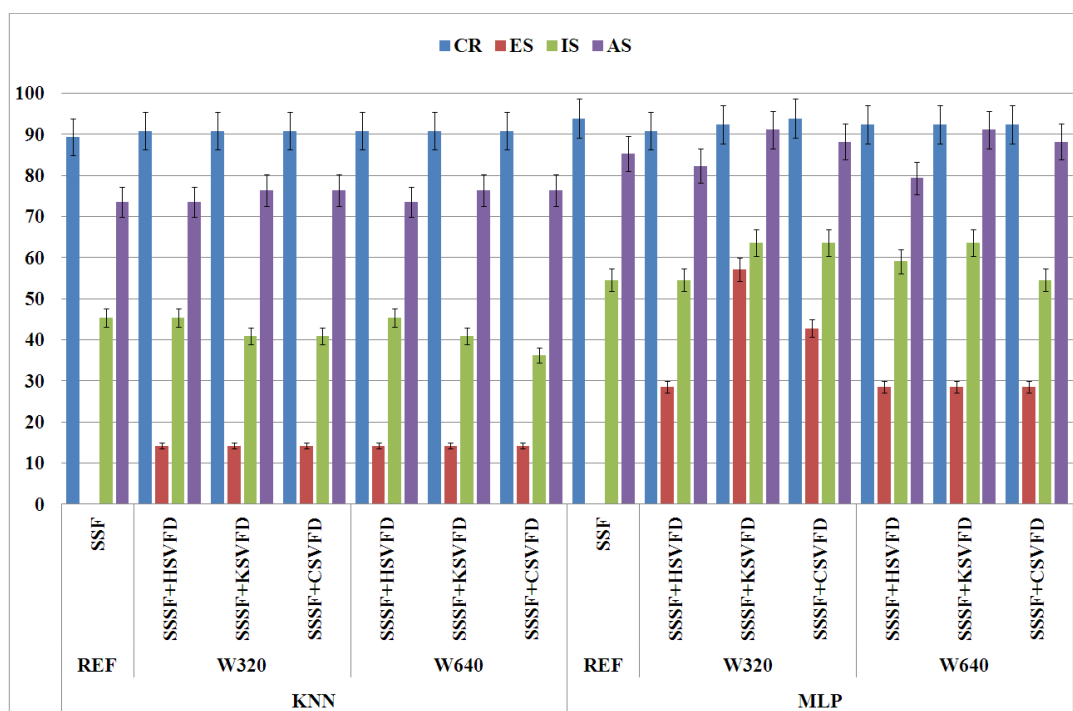
Honek adierazten du dimentsio fraktala, horren aldaketa zehaztuak eta atributu-hautatze automatikoa barne hartzen dituzten multzoak seinale-ezaugarri ez-linealak egoki modelatzeko gai zirela. Bi esperimentuetan, kasuen estaleremua % 94-koa izan zen, gutxi gorabehera. Konfiantza-tarteak $+3$ eta $+5$ dira % 80-tik gorako KNN-ei dagokienez.

6.12 irudiak kasu gehienetan izandako hobekuntza (CER %-tan) erakusten du SVFD barne hartzen denean, batez ere MLP-rako. Emaitzarik hoberenak Katz eta Castiglioni-ren MLP-rekin lortzen dira, leiho-tamaina 320 puntukoak direnean. Oro har, KNN ez da MLP bezain egonkorra. Areago, MLP-ren ratioak onak dira kostu konputazionalari begira (ikus 6.15 irudia).

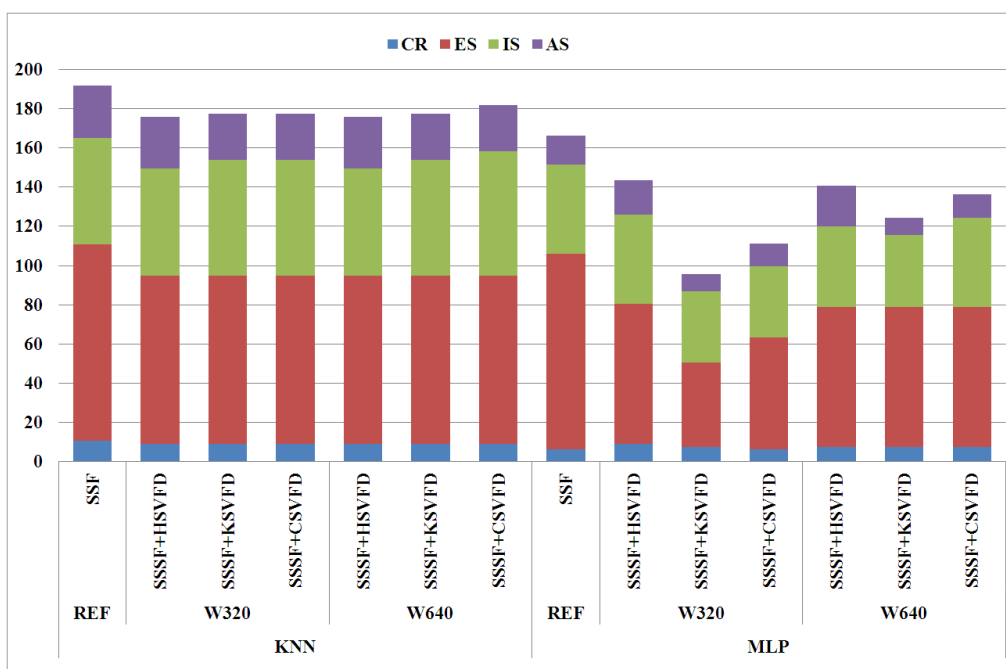
KNN-ren kostu konputazionala apalena da kasu guztietan eta dimentsio fraktaleko ezaugarriekin gai da ES segmentuak detektatzeko, nahiz eta gai ez izan MLP-k lortutako emaitzak erdiesteko. Emaitza orokor hoberenak SSSF+KSVFD-rekin lortzen dira (% 14.73), MLP eta WS = 320 erabilita. KNN-ren emaitzarik hoberenak Higuchi algoritmoaren aldaerez (% 21.71) eta Katz-en dimentsio fraktalaz (% 24.81) erdiesten dira. MLP-ren emaitzetatik urrun (6.12 irudia).

6.13 irudiak emaitzak klaseen arabera (zehaztasuna %-tan) erakusten ditu. SSSF+KSVFD dira, beste behin ere, aukerarik onena. Multzo horrek lortzen ditu emaitzarik onenak klase guztietan eta, gainera, sailkatzea hobetzen du detekzio goiztiarrari begira (ES klasea). IS-ren kasuan, ratio hobea lortu genuen AD-ren erdiko maila identifikatzerakoan.

Bestalde, ezaugarria gai da eri bakoitzaren segmentu patologikoak eta ez-patologikoak diskriminatzeke. Azkenik, klaseen araberrako errore metagarriak %-tan (CE) direla eta, 6.14 irudiak lortutako emaitzak erakusten ditu. W320 delako aukerak, zeinak CE baxuena duen, hobeto atzematen omen ditu seinalearen dinamikak Castiglioni-ren dimentsio fraktalaren aldaerekin. Kontuan hartu behar da positibo faltsuen ehunekoa txikia dela, eta hori gerta daiteke adierazpenak egiterakoan izandako zalantzengatik. Kasuen estaleremua % 95-koa da sistema globalean. ES-ren eta beste fase batzuen arteko desberdintasunak ratioei dagokienean, eriak gaitzaren fasearekin zerikusia duten komunikazio-arazo zehatzengatik dira, 2. kapituluan adierazten den moduan (afasia, anomia). Areago, eriak ES edo AS fasean daudenean, adinarekin loturiko arazoak -ahots sinfonikoa, adibidez- ager daitezke aldi berean ere, adinak aldaketak sortzen baititu ahots-ekoizpenaren sisteman. 6.3 taulak azkeneko sistemaren zehaztasuna (%) agertzen du, MLP eta KNN-rekin, CR eta ADetarako eta baita Konfiantza-Tartea % 95 eta % 90-erako.



Irudia 6. 13: Klaseetarako Zehaztasuna/Doitasuna. Klase desberdinetarako Zehaztasuna/Doitasuna MLP eta KNN-rekin: CR, ES, IS eta AS. REF barrak dimentsio fraktalen ezaugarriak erabili gabe eskuratu dira.



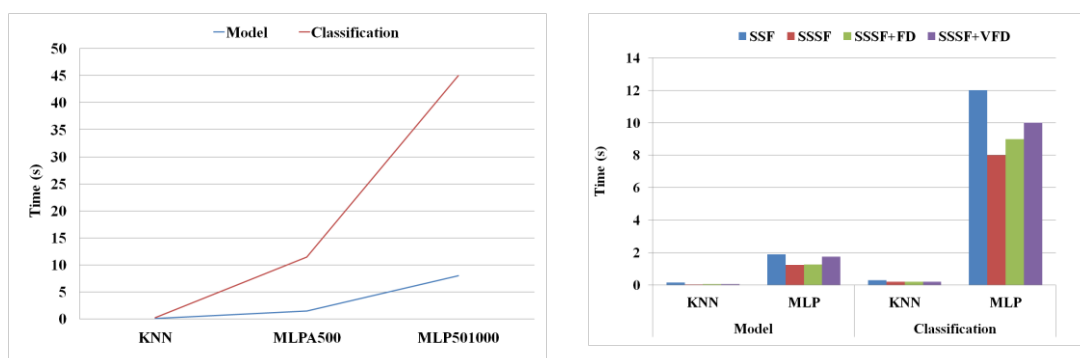
Irudia 6. 14: Errore Metagarria. Klaseetarako (CR, ES, IS eta AS) Errore Metagarria (%) MLP eta KNN-rentzat.

6.15a irudian proposatutako algoritmoaren kostu konputazionala aztertzen da: (a) eraikuntza-prozesuari eta ezaugarrien antzemateari dagokienez: KNN, MLPA500 eta MLP501000 (MLP, NNHL = 500 eta TS = 1000 dituelarik); eta (b) eraikuntza-prozesuari eta ezaugarrien antzemateari dagokienez: KNN eta MLPA500 ezarritako ezaugarriekin.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

6.15a irudian ikus dezakegu, ezaugarria eraikitzean, KNN eta MLPA500-ek antzeko kostu konputazionala dutela eta ondoren MLP501000 dator. Saillapenerako ezaugarritzatzeari dagokionez, aldiz, garbi ikus daiteke KNN-a hoberena dela. 16b irudia zehatz-mehatz aztertzen badugu, ezaugarria eraikitze eta saillapena egiteko kostu konputazionala ikus daitezke, baina KNN eta MLPA500 ezaugarrietarako eta ezaugarri-multzoetarako soilik. Nahiz eta KNN-a MLP-a baino hobea den, ezaugarrien hautaketa automatikoak kostu konputazionala asko gutxitzea ahalbidetzen du (MLP501000 baino lau aldiz gutxiago), eta batera ere, Saillapenerako Errore Ratioaz (CER) ari bagara errendimendu hoberena ematen duen metodoa da.

Beraz, hautaketa automatikoa izango duen MLP-a aukera hoberena da CER txikiagoa duelako eta arrazoizko denbora konputazionala behar duelako antzemate prozesurako (10 segundu baino gutxiago). Egokia da denbora errealeko aplikazioetarako, (ikus 6.15 irudia) .



Irudia 6. 15: Kostu konputazionala. (a) Eraikuntza-prozesu eta ezaugarrien antzematearekiko kostu konputazionala: KNN, MLPA500 eta MLP501000 (MLP, NNHL = 500 eta TS = 1000 dituelarik), b) Kostu konputazionala ereduaren eraikuntza eta ezagutza-prozesua kontuan hartuz: KNN eta MLPA500 taldeetarako ezaugarri-taldeekin.

Taula 6. 3: Accuracy (%) MLP eta KNN azken sisteman CR eta AD klaseetarako (SSSF+KSVFH eta WS = 320) eta Konfiantza tartea (Confidence Interval, CI), % 95 eta 90.

Klasea	KNN			MLP		
	Acc	CI (% 95)	CI (% 90)	Acc	CI (% 95)	CI (% 90)
CR	90.90	[84.09, 97.90]	[85.21, 96.79]	92.43	[86.85, 99.15]	[87.83, 98,16]
AD	87.30	[79.07, 95.52]	[80.39, 94.20]	90.9	[84.10, 98.03]	[85.06, 96.93]

6.3.4 Bat-bateko hizketaren azterketa integrala

6.3.4.1 Materialak eta metodoak

Emozio ezaugarriak eta tenperatu emozionala ere gehituko dira esperimentu honetan bat-bateko hizketaren azterketa integrala egiteko. Aurkeztutako ezaugarriak oinarri hartuta, esperimentaziorako lau ezaugarri-talde sortu dira:

- SSF
- EF, aurrez deskribatutako taldea
- FD1: Higuchi Fractal Dimension (HFD).
- FD2: HFD, HFD maximoa, HFD minimoa, HFD bariantza eta HFD desbiazio estandarra.

Hizkera emozionalaren sailkapen automatikoaren oinarria Multi Layer Pertzepronean (MLP) datza. Horrek 100 neurona dituen ezkutuko geruza bat dauka eta 1000 entrenamendu-pauso. WEKA [195] softwarea erabili da esperimentuak egiteko. Emaitzak Doitasuna (Acc) eta Sailkapenerako Errore-Tasa (CER) erabiliz balioztatzen dira.

Entrenamendu-pauso eta balioesterako k-fold balioeste gurutzatua erabili da emaitza sendoak lortzeko. Ezaugarri horrek CR eta AD hiru mailak definituko dituzte. Gure esperimentuetan $k = 10$ erabiltzen da. Ezaugarri horrek Kontrol Taldea (CR) eta AD hiru mailen artean bereiziko dute.

Taula 6. 4: Accuracy (%), ezaugarri multzoekiko emaitza orokorrak.

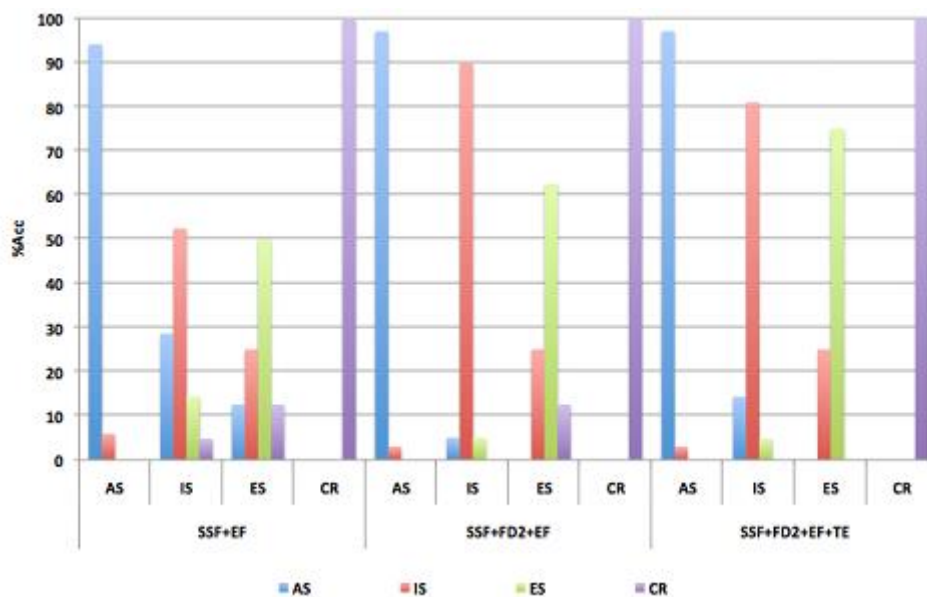
Testa	Ezaugarri tadea	% Acc
Bat-bateko hizketaren jariakortasuna	SSF	75.19
	SSF+FD1	76.74
	SSF+FD2	86.07
Erantzun emozionala	EF	90.69
	EF+TE	97.67
INTEGRALA	SSF+EF	92.24
	SSF+FD2+EF	94.57
	SSF+FD2+EF+TE	94.57

Esperimentazioa AZTIAHORE azpitalde orekatuarekin egin da. Esperimentu horien helburua bat-bateko hizketaren narriaduraren neurketa automatikoa, emozio-erantzuna eta AD duten pertsonengan integrazioa egiteko aukeratutako ezaugarrien ahalmena aztertzea da. Beraz, aurretik definitutako ezaugarri-multzoak balioztatu dira Kontrol Taldeak eta AD mailak ondo definitzeko.

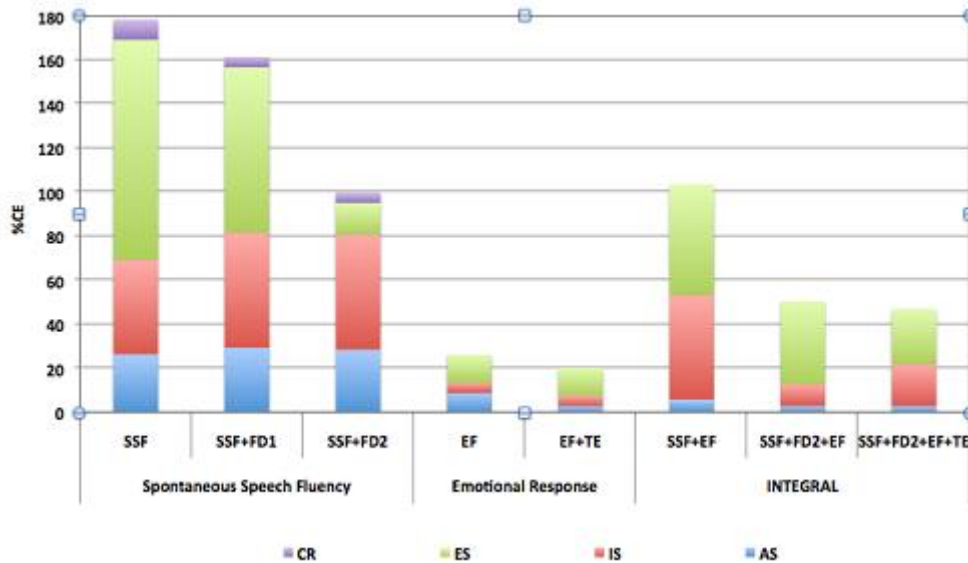
6.3.4.2 Saillkapen automatikoa

Lehen etapan, segmentu bakoitzerako tenperatura emozionala kalkulatu da. Proba pilotuaz baliatuta ikerketa egiteko helburuz, MLP saillkapen automatikoa azaldutako hizketaren ezaugarri-multzoekin burutu da: bat-bateko hizketaren jarioortasun automatikoan, emozio-erantzunean eta bien bilduma izango den hizketa integralean. 6.4 taulak proba aurrekliniko eta ezaugarri-multzoekiko doitasunaren (%) emaitza orokorrak laburbiltzen ditu.

- Lehen proban, zehazki, Hizketaren Bat-Bateko Jarioortasunarenean, bakarrik SSF erabili da eta Dimentsio Fraktalaren (FD) bi multzoekin integratu da. FD-en ezaugarrien erabilerak emaitzak hobetzen ditu. Sistemak, % 10-eko hobekuntza dauka, hainbat FD ezaugarri txertatuak dituen FD2 multzoarekin.
- Bigarren proban, Emozio-Erantzunaren proban, ET ezaugarria txertatzen denean lortzen da emaitzarik hoberena eta Acc-a balio optimotik hurbil dago.
- Bat-bateko hizketaldiko zein emozio-erantzuneko informazioa barne hartzen dituen proban, hots, INTEGRAL probari dagokionean, emaitzek ezaugarri ez-linealen hobekuntza erakusten dute eta emaitza hobeak ET barne hartzen denean.



Irudia 6. 16: INTEGRAL probarako doitasuna (%) klase eta dagozkien ezaugarri-multzoentzat.



Irudia 6. 17: Sailkapenerako Errore-Tasa Metagarria (%) azaldutako hiru probentzat eta dagozkien ezaugarri-multzoentzat.

6.16 irudiak, klaseetan, kontrol-talde (CR) eta AD hiru mailetarako (ES, IS eta AS) lortutako emaitzak erakusten ditu. Klase horietan lortutako emaitzek argi erakusten dute ezaugarri ez-linealen barne hartzeak, hots, FD2 multzoaren erabilerak, emaitza hoberenak lortzen dituela klase guztietarako (kontuan hartuz parametro gehiago integratzen direla). Multzo horrek ere antzemate goiztiarrerako sailkapenaren (ES klasea) emaitzak hobetzen ditu. IS-k ere tarteko AD diskriminatzeke tasa hobea du. Ezaugarri hori pazientez paziente segmentu patologiko zein ez patologikoak diskriminatzeke gai da hiru pilotu probatan.

Sailkapenerako Errore-Tasa Metagarria (% CER) 6.17 irudian aztertzen da. Sailkapenerako Errore-Tasa Metagarria txikiagoa lortzen da ezaugarri ez-linealak txertatuak daramatzaten FD2 eta ET multzoekin. Hau aipagarria da diagnostiko goiztiarrerako, ES klaserako, emaitza hobeak eta sailkapen errore txikiagoa lortzen baita.

6.16 irudian, klase guztietan, Acc-arekiko (%) INTEGRAL probaren azterketa sakonagoa egiten da:

- Hizketa sinpleen ezaugarrien emaitzak ikertzen direnean errendimendu optimoa ikus daiteke kontrol-talderako; sailkapen mistoa ordea AD mailetarako, bereziki IS eta ES-rako.
- Sistema orokorrak oso emaitza onak ditu FD2 txertatzen denean, Acc emaitzak IS tarteko etapan hobetzeaz gain ES tartean ere asko hobetzen baititu.
- Hirugarren esperimentuan, ET txertatzen da eta Acc-k oso emaitza onak lortzen ditu klase guztietarako eta ES gaintzen du kontrol-subjektuen segmentuekiko nahastea desagertzen delako.

Kasu batzuetan, gertatzen den segmentuen arteko bukaerako nahastea gizabanako bereko segmentutan patologia-maila desberdinak agertzeagatik izan daiteke. Aukera hori etorkizuneko ikerlanetan miatuko da proba berrien bidez.

Azkenean, osasun-adituek emaitza horiek aztertu dituzte. Hala bat-bateko hizketa eta emozio-erantzuna bezalako biomarkatzaile independenteen azterketa nola hainbat biomarkatzailearen azterketa integrala egiteko sistemak duen gaitasunaren garrantzia behatzeko aukera izan dute.

Beraz, proba ez-inbasibo horiek oso lanabes erabilgarria izan daitezke mediku-espezialistentzat etorkizuneko diagnostiko goiztiarrerako proba klinikoetan.

6.4 Ataza anitzerako hurbilpen ez-lineala MCI-rako

6.4.1 Materialak eta metodoak

Dakigun bezala esperimentazioa burutzeko azpimultzo orekatuak erabilia egin da: PGA-OREKA, MINI-PGA eta AZTIAHORE. Bigarren esperimentazioko etapan, PGA-OREKA azpi datu-basearekin aritu gara batez ere disfluentzia aztertzeke.

Ezaugarri guztiak gure lantaldeak [270], Matlab-ek [166] eta Praat-ek [167] garatutako softwarearen bidez kalkulatu dira. Datu-basetik ateratako ezaugarri-kopurua 340ekoa da, 16.000 kHz-ko lagin-frekuentzia baterako. Ondoren, ezaugarri-hautaketa automatikoa egiten da. Prozedura horrek bi urrats ditu:

- Hasteko, medikuntza- eta estatistika-informazioa kontuan hartuta, p balioa ateratzen da Matlab-en Mann-Whitney U proba ez-parametrikoa dela bidez. Test hori eta gero, p balio 0.05eko eta p balio 0.1eko adierazgarritasun-maila duten ezaugarriak aukeratzen dira. P balio altuagoa atalase moduan erabiltzen da ezaugarri-multzo zabalago bat aukeratu ahal izateko bigarren urratserako. Ezaugarrien hautaketa automatikorako algoritmoak ezaugarri-multzo optimoa aukeratzeko erabiltzen dira, ezaugarrien integrazio-faktorea ere kontuan hartuta.
- Ondoren kontuan hartzen dira aurretik lortutako ezaugarriak, ezaugarri-hautaketa automatikorako algoritmoak direla bidez. Datuetan oinarritutako ezaugarri-hautaketa burutzeak hiru oinarritzko abantaila ditu:
 - Gehiegizko doitzea murriztea (datuen erreduantzia kentzea).
 - Zehaztasuna hobetzea, datu engainagarri (zarata) gutxiago izateari esker.
 - Kostu konputazionala txikiagotzea.

WEKAtik [195] hartutako bi algoritmo erabili dira ezaugarrien hautaketa eta optimizaziorako. Algoritmo horiek ezaugarri-azpimultzoen bilatze- eta balioztatze-metodoak aztertzen dituzte, eta aztertu egiten dute ezaugarri bakoitzaren eragina eta baita ezaugarrien arteko elkarrekintzaren eragina sailkapenari dagokionez:

- *CfsSubsetEval*: ezaugarrien azpimultzo baten balioa ebaluatzen du, ezaugarri bakoitzaren bakarkako iragarpen trebetasuna kontuan hartuta eta baita beren arteko erredundantzia maila ere.
- *GainRatioAttributeEval*: Klasearekiko duen irabazi erlazioa neurtuz ezaugarriaren balioa balioztatzen du.

Azkenik, sailkatze automatikoa egin da Multilayer Perceptron (MLP edo Multigeruza Perzeptrona) klasikoa eta Deep Learning (Ikasketa Sakona) erabilia Convolutional Neural Networks (CNN edo Neurona-sare konboluzionalak) direla bide. Sare horiek biologian oinarritutako MLP-en aldaerak dira. Kontuan hartu behar da entrenatze-laginen kopuruen zenbatekoa txikia dela eta arazoaren konplexutasuna handia:

- AN ataza (CR/MCI): Deep Learning (Ikasketa Sakona) delakoaren ildoan Convolutional Neural Networks (CNN) erabili da: 40 ezaugarri-mapak, 5x5eko eremu-tamaina eta 2x2ko max-pooling-a.
- PD eta SS atazak (CR/AD): Multilayer Perceptron (MLP edo Multigeruza Perzeptrona). Ezaugarriak: Erabili genituen (ezaugarriak + klaseak)/2 neurona geruzan. 500 entrenatze-iterazio.

WEKA software sorta erabili dugu esperimentuak aurrera ateratzeko. Emaitzak Classification Accuracy Rate (Acc, % edo Sailkapenaren Doitasun Tasa) erabilia ebaluatu dira. Entrenatze- eta balioztatze-neurketetarako, k-fold cross-validation (k geruzako balidazio gurutzatua) erabili da k = 10ekin. Balioztatze gurutzatu errepikatuak (Weka ingurunerako kalkulaturako dagoen bezala) oso estatistika test zehatzak ahalbidetzen ditu. Weka-ren Coverage of cases (0.95eko mailan) delakoak automatikoki ematen duen aurrez zehaztutako balioa ere erabili dugu, hau da, % 95-eko konfiantza-tartea.

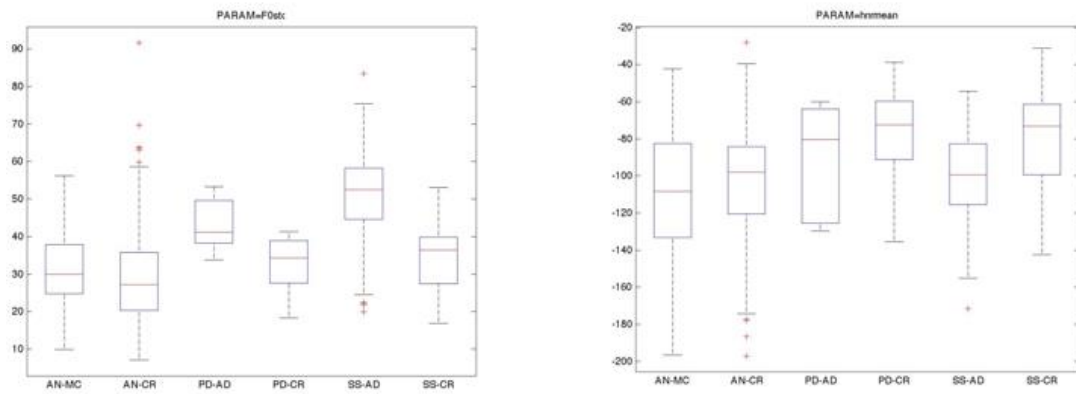
6.4.2 Esperimentazioa

6.4.2.1 Ezaugarrien hautaketa automatikoa

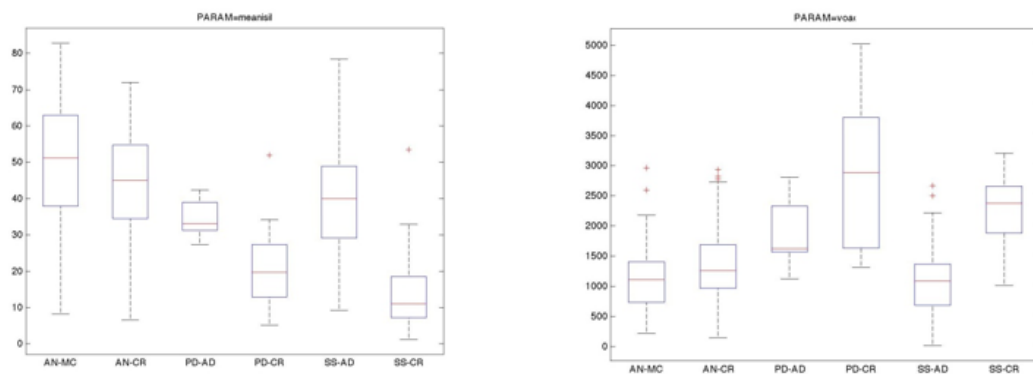
Laginetatik ateratako ezaugarri-bektoreak 340 parametro ditu: 70 klasiko, 60 hautematekoak, 180 hautemate-ezaugarri aurreratuak eta 30 ezaugarri ez-linealak

Ezaugarri askotan desberdintasunak daude aipatutako 4 azpimultzoei dagokienez, baina desberdintasun horietako batzuek ataza-motarekin zerikusia dute.

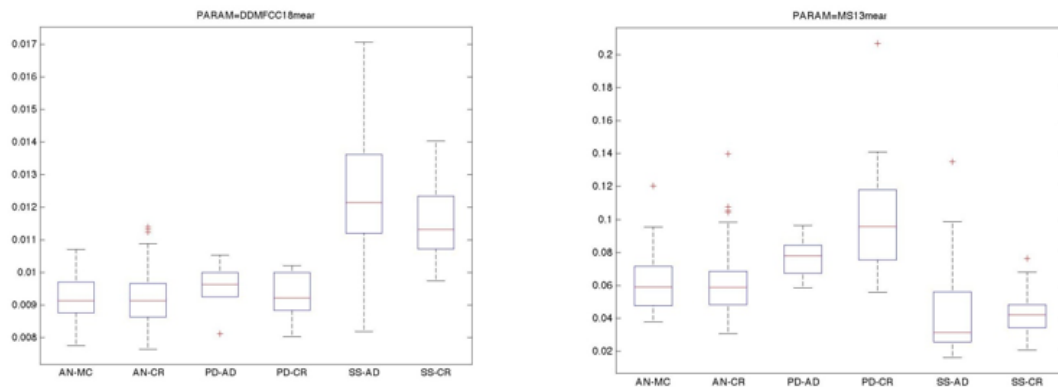
6. KAPITULUA: EMAITZAK ETA EZTABAIDA



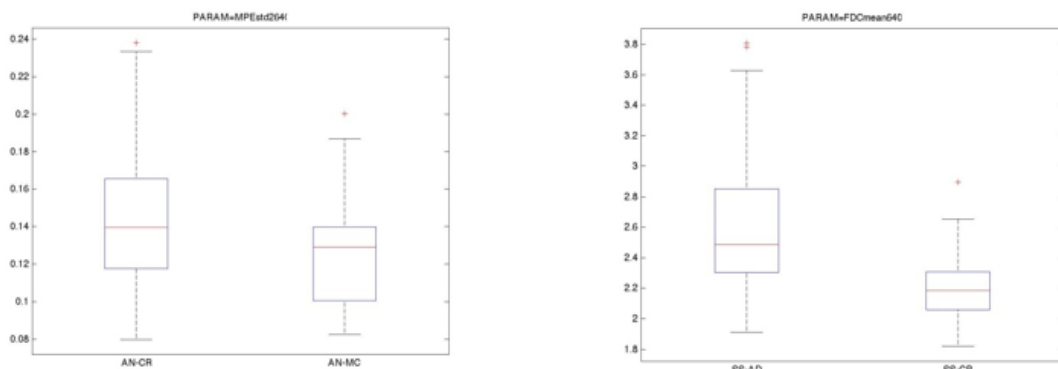
Irudia 6. 18: Kaxa-diagrama hiru atazetarako: animaliak izendatzea (AN), irudien deskripzioa (PD) eta bat-bateko hizketa (SS). (a) FO desbideratze estandarra. (b) harmonizitatea. Irudiak ezaugarri klasiko batzuetarako analisia erakusten du. MCI eta AD dutenengan, FO-ren egonaitasuna garbia da eta harmonizitatea ere txikiagoa da. Bat-bateko hizketan altuagoa den FO desbideratze estandar horren desberdintasuna erabili zitekeen emoziozko erantzuna aztertzeko.



Irudia 6. 19: Kaxa-diagrama hiru atazetarako II: animaliak izendatzea (AN), irudien deskripzioa (PD) eta bat-bateko hizketa (SS). (a) Etenaldiaren iraupenaren batez bestekoa. (b) Ahoskatze-jardueraren batez bestekoa. Ezaugarri batzuek desberdintasunak agertzen dituzte etenaldiei eta ahoskatze-jarduerari dagokienez. Ezaugarri horiek hitz-jarioaren, prozesu kognitiboaren edota egoera emozionalaren adierazle izan daitezke.



Irudia 6. 20: Kaxa-diagrama hiru atazetarako III: animaliak izendatzea (AN), irudien deskripzioa (PD) eta bat-bateko hizketa (SS). (a) $\Delta MFCC18$. (b) MS13ren batez bestekoa. Halaber, badira zenbait joera garbi beste zenbait ezaugarri-azpimultzotan, besteak beste, hautemate-ezaugarri klasikoak, hautemate-ezaugarri aurreratuak edo ez-linealak.



Irudia 6. 21: Ahoskabe segmentuen ezaugarri ez-linealak 640 lagineko leiho batekin. (a) Animalien izendatzea (AN): Multieskaladun Permutazio-Entropiaren desbideratze estandarra (2 eskala). (b) Bat-bateko hizketaren ataza (SS): Castiglioni-ren Fraktalaren batez bestekoa. Halaber, badira zenbait joera garbi beste zenbait ezaugarri-azpimultzotan, besteak beste, hautemate-ezaugarri klasikoak, hautemate-ezaugarri aurreratuak edo ez-linealak.

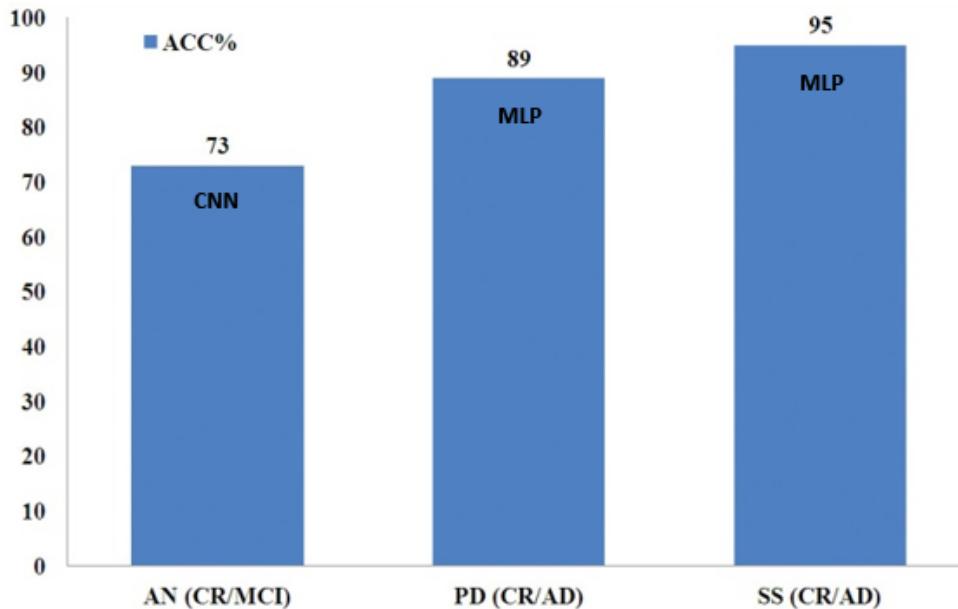
Edonola ere, nahiz zenbait joera garbi egon ezaugarri askotan, desberdintasunak ere agertzen dira ataza bakoitzean lortutako p-balioan (atazarekiko mendetasuna), ataza bakoitzaren nolakotasuna dela kausa: hizketa-komplexutasuna, adina, pertsonaren fasea eta ingurunea.

Hortaz, ezaugarri-hautaketa automatikoa ezinbestekoa da ataza bakoitzerako ezaugarririk hoberenak aukeratu ahal izateko. Bigarren urrats honetan, ezaugarri-kopurua murriztu egin dugu: % 50 animalien izendatzerako, % 50 irudien deskripziorako eta % 40 bat-bateko hizketarako.

Hori dela eta, hurrengo azpiatalean, metodologia automatikoa erabiliko da hala laginak sailkatzeko.

6.4.2.2 Sailkatze automatikoa

Azkenik, Neural Networks (MLP eta CNN) direla bide, sailkatze automatikoa egin dugu aurrerago deskribatutako balidazio gurutzatuen estrategian oinarrituta.

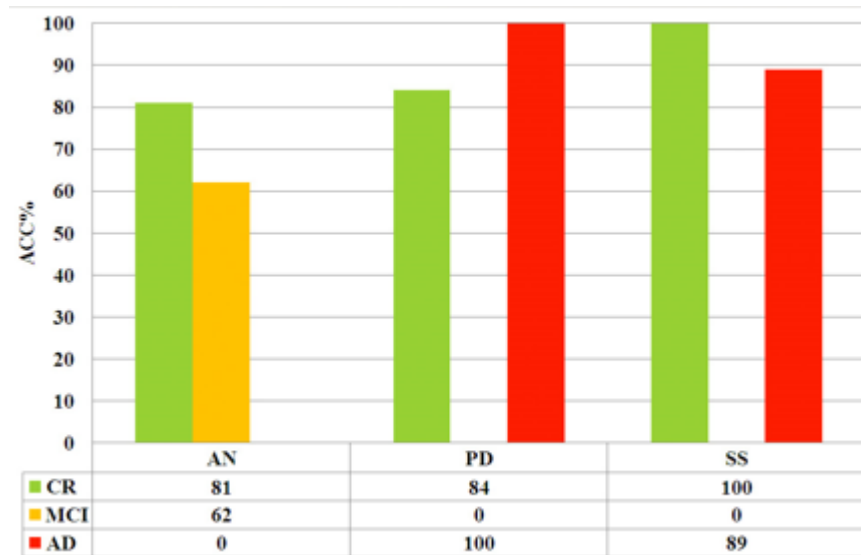


Irudia 6. 22: Hiru atazen doitasun-tasak (Classification Accuracy Rate - Acc, %): animalien izendatzea (AN), irudien deskripzioa (PD) eta bat-bateko hizketa (SS).

6.22 irudian ikus daiteke, emaitza global hoberenak askeena den hizketan lortu dira, bat-bateko hizketan (SS) alegia: % 95. Hori honako honen ondorio da bereziki: grabazioa ingurune lasai batean egitea, aldaketa kognitibo sotilak izatea seinalean hizketa irekiagoa dela kausa eta Alzheimer gaixoekin egitea, Nahasmendu Kognitibo Arina dutenekin egin beharrean.

Animaliak izendatzea (AN) delako atazean emaitzak txarrenak dira: % 73. Hori honako honen ondorio da bereziki: hizketa murriztuaren ataza errazagoa gerta dakieke MCI dutenei AD dutenei baino.

Irudien deskripzioa (PD) delako atazaren emaitzak baikorrak dira, % 89, nahiz eta eremu-tamaina txikia izan. Kontuan hartu behar da Alzheimer gaitza dutenekin egin dela eta ez Nahasmendu Kognitibo Arina dutenekin.



Irudia 6. 23: Hiru atazetan lortutako klase bakoitzeko Accuracy-a (%). Emaitzak hiru atazentzat: animaliak izendatzeko jariora (AN), irudien deskripzioaren ataza (PD) eta hizketa espontaneoaren ataza (SS).

Hortaz, klase desberdinentzat lortutako emaitzarik hoberenak bat-bateko hizketaren (SS) atazarekin lortu dira. Hori honako hauen ondorioa da: AD dutenen pazienteak erabiltzea, MCI dutenak beharrezko eta zalantza kognitibo sotilak detektatzea modu askean hitz egiten dutenean.

Animaliak izendatzearen (AN) atazak emaitza txarrenak lortu ditu MCI dutenak hizketa murriztuaren atazaren bidez aztertu dituztenean: entrenatze-eremu zabalagoa behar da kontrol-taldea (CR) eta nahasmendu kognitibo arina (MCI) dutenen arteko ezaugarriak bereizteko. Zenbait atazatan, MCI/CR seinaleak elkarrengandik hurbilago daude AD/CR artean baino.

Irudiak deskribatzeari (PD) buruzko atazari dagokionez, eremu-tamaina txikia bada ere, emaitzak onak dira: animaliak izendatzeari eta irudien deskripzioari buruzko atazen emaitzak etorkizun handikoak dira eta entrenatze-lagin gehiago beharko lukete.

6.5 MCI-ren detekzio goiztiarrerako aurrerapenak

6.5.1 Materialak eta metodoak

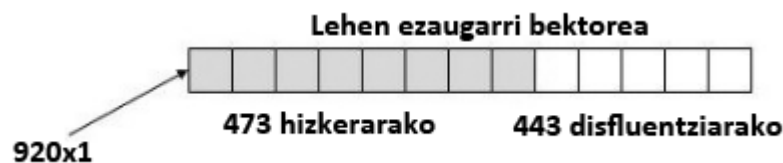
Narriadura Kognitibo Arinaren (MCI) detekzio goiztiarrerako aurrerapenetan hizketaren disfluentziaren azterketa automatikoa gehituko dugu. Saillkatzaile moduan ikasketa sakonaren azterketa egingo dugu baita ere CNN neurona-sareak gehituz esperimentazioari, MLP, KNN eta SVM-ekin batera.

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

CVF atazak AN-ren bidez neuroendekapenezko gaixotasunetako narriadura kognitiboa neurtu eta kuantifikatzen duen proba da [144]. Hizkeraren abileziak, memoria semantikoa eta funtzio exekutiboak [271] balioztatzeko oso erabilia den proba da. Laginak 187 gizabanako osasuntsu eta CITA-Alzheimer [38][272] Fundazioaren Gipuzkoa-Alzheimer Egitasmoko (PGA) Multzokoak diren eta MCI duten 38 pertsonak osatzen dute. Esperimentaziorako PGA-OREKA azpitalde orekatua aukeratu da. Esperimentazioan erabilitako materialak gutxi gorabehera MCI taldeko 40 ahots lagin eta kontrol-taldeko (CR) 60 lagin izan dira.

Grabazioak automatikoki segmentatu dira ahots-seinale eta disfluentzietan Ahots Jarduera Hautematerako (VAD) algoritmoa erabiliz. Saillapen automatikorako 4 saillkatzaile erabili dira: k-NN, SVM, MLP eta CNN-ak. Emaitzak Saillkapenerako Errore-Tasa-ren (CER, %) bidez balioztatu dira eta entrenamendu eta balioztapenerako $k = 10$ -ko balioztatze gurutzatua erabili da.

22 KHz-ko laginketa maiztasunarekin lortutako ezaugarri-kopurua 920 da. CNN-ari sartu beharko geniokeen bektorea 920×1 -ekoa izango litzateke iragazkirik jarriko ez bagenio. 920 ezaugarri horietatik batzuk hizketakoak eta besteak disfluentziakoak dira: 473 hizketarako eta 447 disfluentzietarako.



Irudia 6. 24: Ezaugarri-bektorea. Iragazkitik pasa aurretik ezaugarri-bektoreek 920 ezaugarri biltzen ditu.

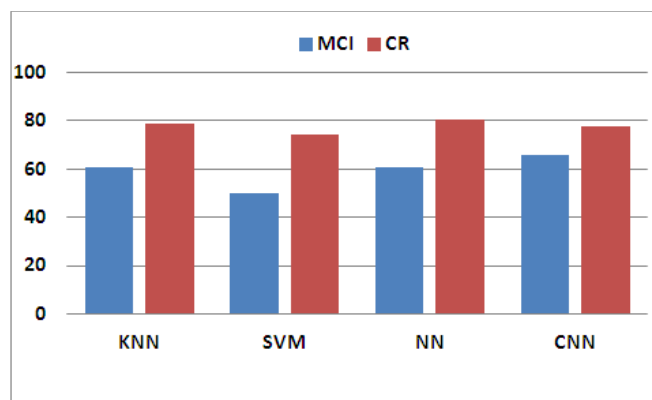
- Ezaugarrien normalizazioa eta hautaketa automatikoa: Ezaugarriak normalizatu ondoren, ezaugarrien hautaketa automatikoa egiten da $p < 0,1$ den Mann-Whitney-ren U-proba ez-parametrikoren bidez. Proba hori bi laginen arteko heterogeneotasuna balioztatzeko erabiltzen da. 920 ezaugarritik 150 ezaugarri adierazgarrienak aukeratzen dira proba honen bidez.
- Ezaugarrien bigarren hautaketa automatikoa: Bigarren iragazketa fasean ezaugarrien hautaketarako WEKA-ren SVMAttributeEval metodoa erabiltzen da, zeinak, 150 ezaugarrietatik 80 ezaugarri aukeratzen ditu. Bai hizketarako bai disfluentzietarako mota guztietako ezaugarriak daude bukaerako saillapen honetan. Nahiko bektore orekatua ateratzen da, eta zer esanik ez, bektore gordina baino aberastuagoa. Lehen ezaugarri-mapa dela ondoriozta daiteke.
- Saillkapena: Datu horiek Neurona-sare Konboluzionalari sartuko zaizkio entrena dadin. Behin entrenatuta, bere doitasuna balioztatuko da.

6.5.2 Esperimentazioa

6.25 irudiak klasekako (MCI eta CR) taldekatze automatikoaren bidezko emaitzak azaltzen ditu. Aipaturako sailkatzaile guztientzat Acc (%) erabiliz balioztatzen da.

Disfluentzia azterketa ikerketa prozesuan txertatzeko proposamenak aurreko esperimentuen emaitzak [38] hobetu ditu sailkatzaile ia guztientzat. Lortutako emaitzak itxaropentsuak, egonkorak eta orekatuak dira.

Gainera, CNN-en kasuan emaitza hoberenak 20 neuronaz osatutako 2 geruzen bidez lortzen dira. Konboluzioa 3x3koa eta pooling-a 2x2-koa. Aukera horrek 2 geruza eta 100 neuronako MLP-ak baino emaitza hobekak lortzen ditu. MCI taldekatzerako emaitzak nabarmen hobetzen ditu sare konboluzionalak.



Irudia 6. 25: Acc (%) aukeratutako klase eta sailkatzaileetarako. k-Bizilagun Hurbilak (k-NN), Euskarri Bektoreen Makina (SVM), L geruza eta N neuronadun Geruza Aniztunadun Pertzeptroia (NN) eta Sare Neuronal Konboluzionalak (CNN).

6.6 Denbora errealerako sistema

6.6.1 Materialak eta metodoak

Hiru datu-baseak erabiliko dira esperimentazio honetan: AZTIAHORE, MINI-PGA eta PGA-OREKA.

Esperimentazioan ezaugarri-multzoa 50-5 tartean aztertuko da denbora errealera bideraturik. Ezaugarri hautaketa egingo da aurreko kapituluetan definituriko teknika automatikoen bidez; batez ere *SVMAttributeEval* (MATLAB, WEKA).

6.6.2 Esperimentazioa

Azken helburua denbora errealean lan egingo duen sistema mediku-laguntzailea garatzea da. Malgua, eraginkorra eta diseinua konfiguragarria izango duen ALZUMERIC

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

sistema garatu da. Erabiltzaileak (osasun-espezialistak, laguntzaileak edo azken erabiltzaileak) sistema konfiguratuko du beharren arabera: ataza, segurtasuna, dohitasuna eta positibo faltsuen tasa zerorantz abiatzea. Sistemaren ezaugarri orokorrak automatikoki aukeratuko dira esperimentazio-prozesuan.

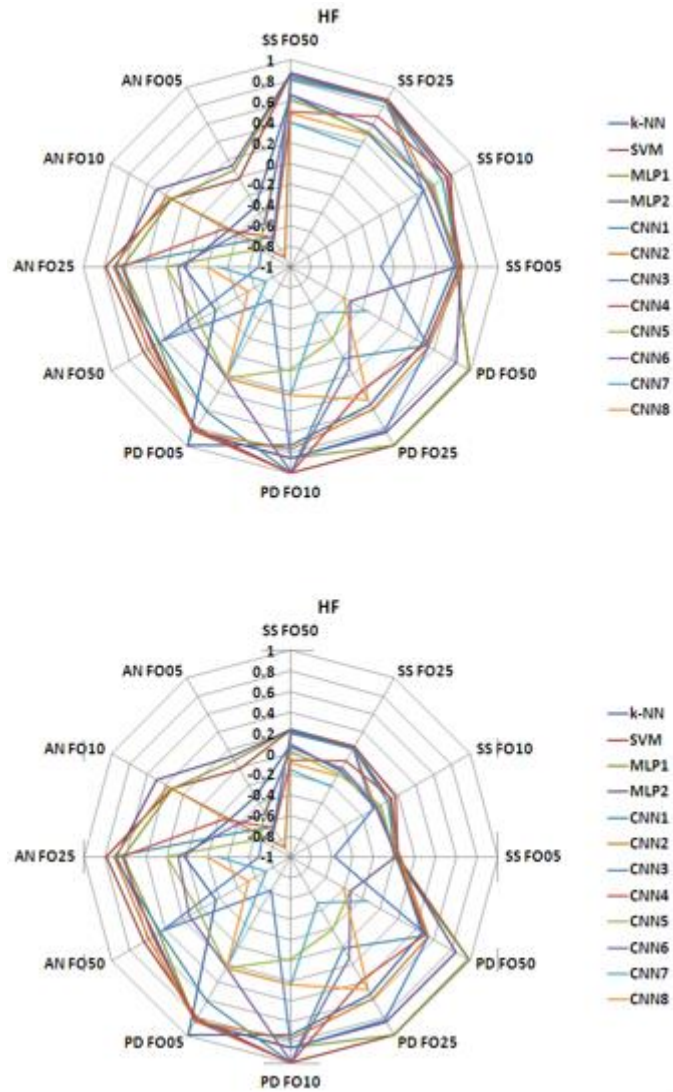
Konfigurazioaren hautaketa erabiliko den Helburu Funtzioaren (HF) bidez egingo da (6.26). Ondoko formularen bidez egingo da:

$$HF = w_1 * ACC - w_2 * FP - CR + w_2 * ROC - AREA - CR - w_2 * MT \quad (6.1)$$

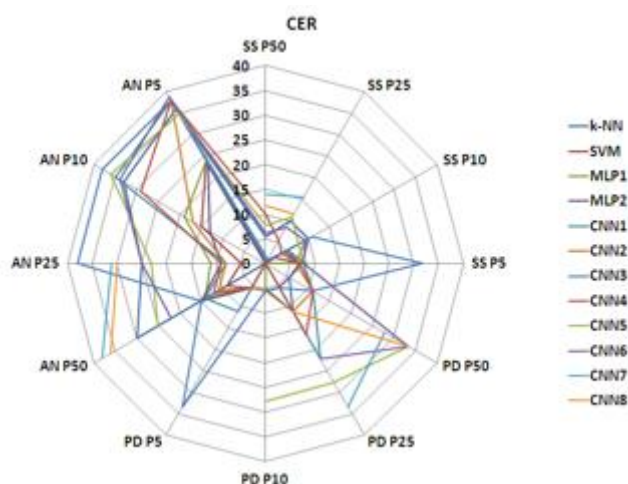
Azkenik, hiru atazetarako (AN, PD eta SS), 50, 25, 10 eta 5 ezaugarrien azpimultzoak aztertuko dira ondoko sailkatzaile automatiko eta konfigurazioetarako (ikus taula 6.5):

Taula 6. 5: Sailkatzaileen konfigurazioak. NNHLren balioa a = ezaugarri-kopurua+klaseak/2, konboluzioa (cxc), pooling(pxp), DI = Dropout Initial, DH = Dropout Hidden.

	NUM-L	cxc	pxp	DI	DH
k-NN					
SVM					
MLP1	a				
MLP2	a,a				
CNN1	a			0.2	0.5
CNN2	a			0.2	0.2
CNN3	a,a			0.2	0.5
CNN4	a,a			0.2	0.2
CNN5	a	2x2	2x2	0.2	0.5
CNN6	a	2x2	2x2	0.2	0.2
CNN7	a,a	2x2	2x2	0.2	0.5
CNN8	a,a	2x2	2x2	0.2	0.2



Irudia 6. 26: HF (%) aukeratutako klase eta sailkatzaileetarako (ikus 6.1 algoritmoa). k-Bizilagun Hurbilak (k-NN), Euskarri Bektoreen Makina (SVM), L geruza eta N neuronadun Geruza Aniztunadun Pertzeptroia (MLP) eta Sare Neuronal Konboluzionalak (CNN).



Irudia 6. 27: CER (%) aukeratutako klase eta sailkatzailetarako. k-Bizilagun Hurbilak (k-NN), Euskarri Bektoreen Makina (SVM), L geruza eta N neuronadun Geruza Aniztunadun Pertzeptroia (MLP) eta Sare Neuronal Konboluzionalak (CNN). SS eta AN kasuetan Cross-validation k-fold = 10-ekin erabiliko da eta MINI-PGA atazarako Leave One Out (LOO) metodologia k- fold = 18-rekin.

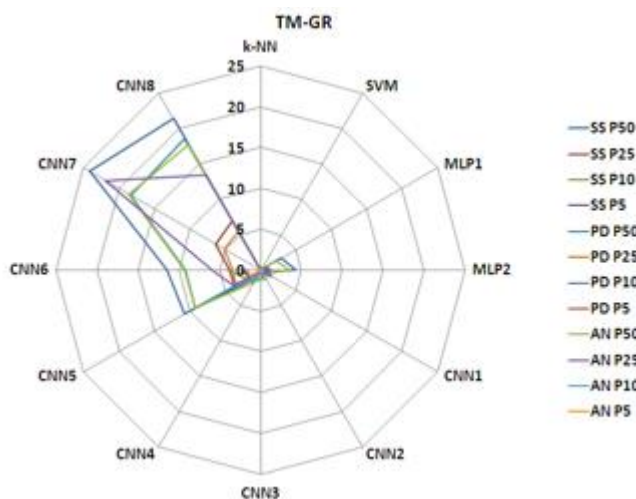
Horrek hautaketa azkarra, naturala eta erreza egiten uzten ditu.

Irudietan hiru atazetarako emaitzak agertzen dira: CER-erako (6.27), Model Time-erako MT (6.28), Positibo Faltsuen kontrolerako FP-CR (6.29) eta Helburu Funtzioaren indizerako HF (6.26).

Aukeraturiko metodologiarekin emaitzak orokorrean oso onak dira eta aurreko ataletan lortutako emaitzak hiru atazetarako gainditzen dituzte. 20 ezaugarri-multzoekin emaitza onenak lortzen dira.

Indizeen arabera, SVM soluzio optimotzat har daiteke SS eta PD atazetarako biek % 100-eko igartzea lortzen baitute 25 ezaugarriekin. AN-ri dagokionez, aukera onenak CNN eta SVM-etarako dira.

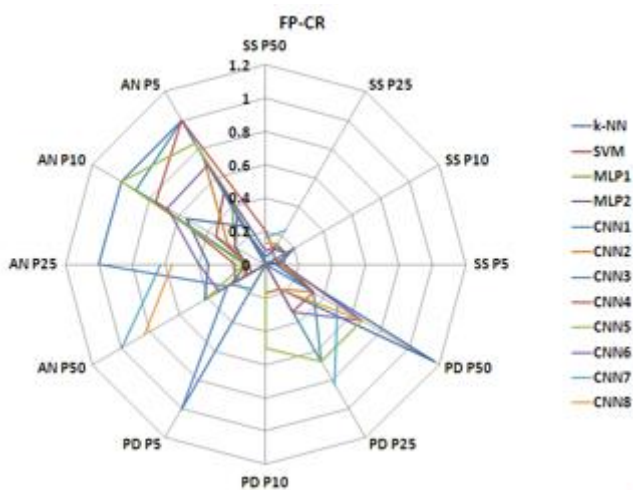
Kasu horietan, % 95 eta % 92 lortzen dira, besteak beste. Parametroetan ikusten denez disfluentziei buruzko ezaugarriak aztertzeagatik izan da. Ereduak, aurreko esperimientuetan kanpoan gelditzen ziren datu garrantzitsuenekin entrenatzen baitira. Esaterako, pertsona horiek duten barruko elkarrizketa.



Irudia 6. 28: Model Time MT (%) aukeratutako klase eta sailkatzaileetarako. k-Bizilagun Hurbilak (k-NN), Euskarri Bektoreen Makina (SVM), L geruza eta N neuronadun Geruza Aniztunadun Pertzeptroia (MLP) eta Sare Neuronal Konboluzionalak (CNN).

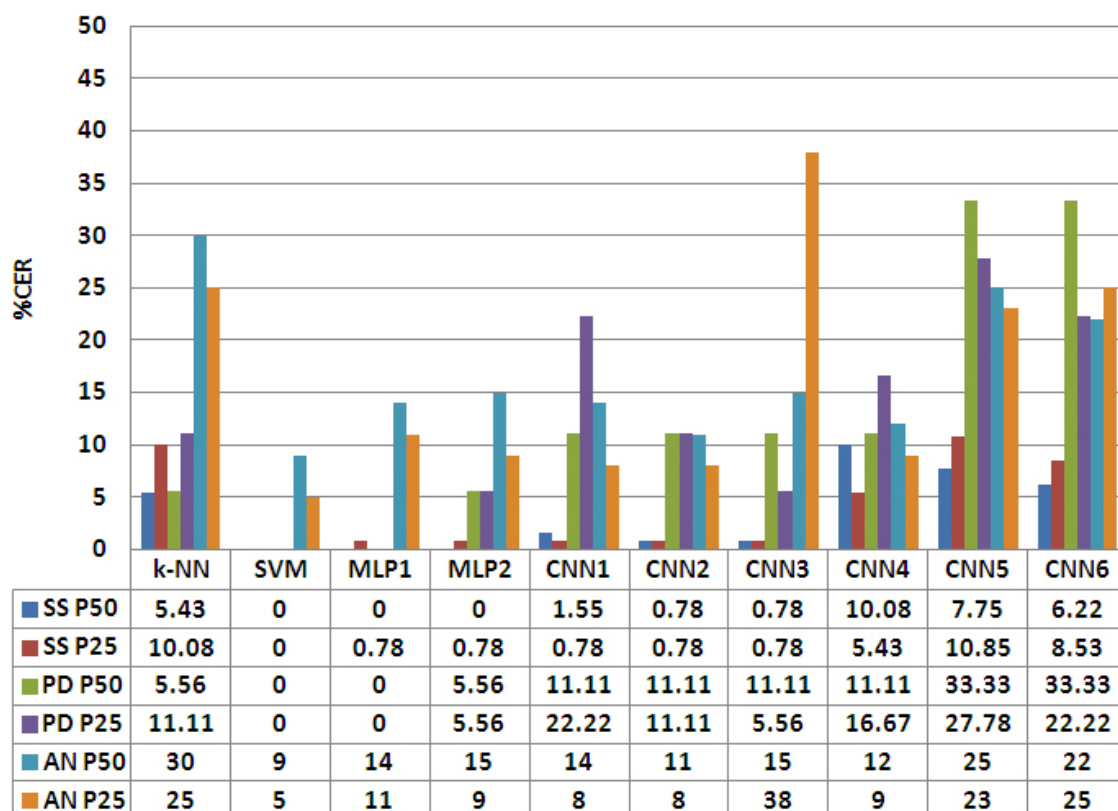
Izan ere, kasu guztietan CNN konplexuak degradatzen dira ezaugarrien multzoa oso txikia denean. Agerian gelditzen da aurretik egindako datuen aberasteak izugarri laguntzen duela sailkatzaileek egoki azertu eta sailkatzeko.

Lortutako emaitzek erakusten dute AN ataza konplexuena dela diferentziak bilatzeko. Gainera, PD-ren datuak konplexuak dira lagin oso txikia delako. Ordea, datu aberasteak eta LOOCV estrategiak egoki funtzionatzen dute.



Irudia 6. 29: FP-CR (%) aukeratutako klase eta sailkatzaileetarako. k-Bizilagun Hurbilak (k-NN), Euskarri Bektoreen Makina (SVM), L geruza eta N neuronadun Geruza Aniztunadun Pertzeptroia (MLP) eta Sare Neuronal Konboluzionalak (CNN).

6. KAPITULUA: EMAITZAK ETA EZTABAIDA

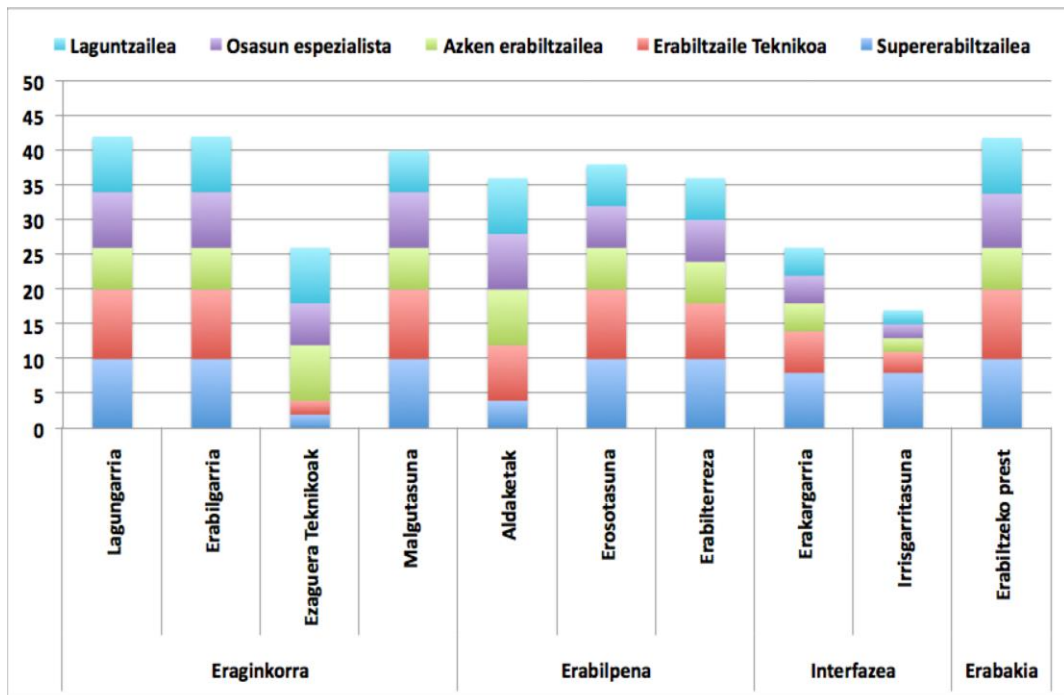


Irudia 6. 30: CER (%) hiru atazentzat (SS, PD eta AN). Ondoko taulan k-NN, SVM, MLP eta CNN-entzat hiru atazetan lortutako emaitzak ikus daitezke.

6.30 irudian ikus daitezkeen bezala, MCI-rako atazean (AN) emaitza oso onak lortu dira, batez ere 25 ezaugarriekin balioztatuta eta SVM zein CNN1, CNN2 eta MLP2-rekin. Hobekuntza adierazgarri horiek ezaugarrien erazketarekin zein horien aukeraketekin zerikusia izan dezakete. Kontuan hartu behar da disfluentzia segmentuetan dugun informazioa txertatzen ari gataizkiola algoritmoari. Aldi berean, garrantzitsua da erabilitako metodologia ezaugarriak aukeratzeko. Adierazgarrienak diren ezaugarriak txertatzeak informazio oso aberastua ematen dio algoritmoari. Thick Data-ren pisu espezifiko garrantzitsua da kasu honetan.

SVM-ek oso emaitza onak ematen ditu ataza guztietan, bereziki adierazgarria % 5 CER-a. Adierazgarria da ere CNN-ek lortzen dituzten emaitzak, batez ere, MCI aztertzerakoan kontuan izanik lagin gutxi daudela. Datu oso ezaugarrituak emateak asko laguntzen die eta % 8-ko CER-a duten emaitzak lortu dira CNN-ekin.

6.7 Erabilgarritasunaren ebaluaketa



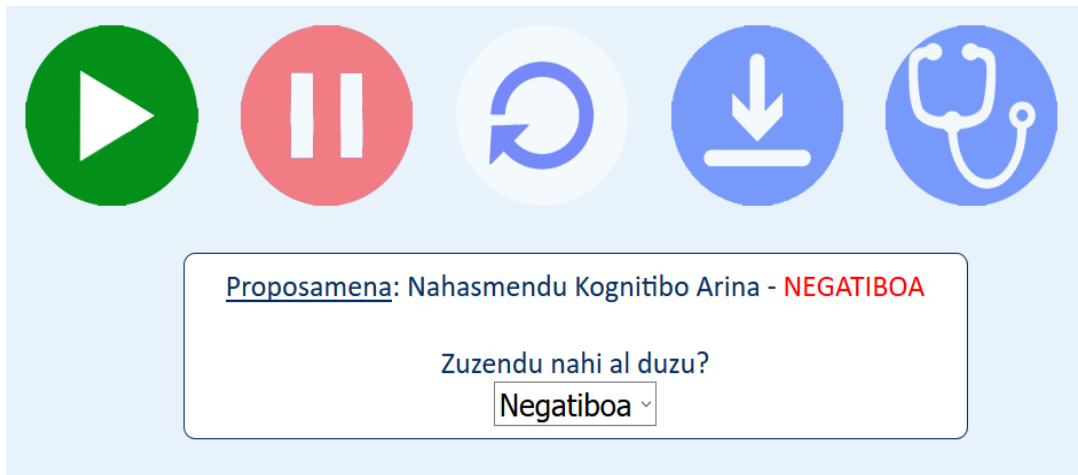
Irudia 6. 31: Sistemaren balioztatzea

Sistema profil desberdineko pertsonekin erabili da eta ondoren sistemaren erabilgarritasunaren inguruko galderak egin zaizkie sistema balioztatzeko helburuarekin. Ingurune kliniko zein etxeko giroan probatu da.

6.31 irudian ikus daiteke sistema probatu duten pertsonen profila: supererabiltzailea, erabiltzaile teknikoa, azken erabiltzailea, osasun-espezialista eta laguntzailea. Sistema erabili ondoren, erabilgarritasuna ebaluatzeko galdera-sorta bat mahaigaineratu zaie 1 eta 5 puntu bitartean balora dezaten (1 puntu baxuena eta 5 altuena). Ondoren taldekako batez bestekoa atera da.

Ikus daitekeen bezala, orokorrean balorazioa ona den arren erabiltzailearen profilaren arabera ikuspuntua asko aldatzen da. Ezagutza teknikoak asko baldintzatu du erabiltzailearen ikuspegia.

Ebaluaketak erakutsi zuen interfazearen lehen bertsioak (5. kapitulua) batez ere erabilgarritasun eta irristagarritasunean hobekuntzak behar zituela. Hau dela eta, bigarren bertsioa garatu da 6.32 irudian ikus daitekeen bezala.



Irudia 6. 32: Tipografia tamaina handitu eta interfazearekin elkar eragiteko botoiak guztiz aldatu dira erabiltzeko errazagoa eta intuitiboagoa izan dadin.

7. KAPITULUA: ONDORIOAK

7.1 Ondorioak

7.1.1 Ondorio orokorrak

1. Gaixoen hizketan garunaren demenzia-zeinuen nabaritasun zientifikoa badago, MCI-n eta AD-n lehen kaltetzen den gaitasunetako bat hitzeko komunikazioa denez, horren analisi automatikoa bahetze metodologia guztiz natural eta erabilgarria da, eta AN atazaren grabazio eta azterketaren bidez, galera kognitiboaren garapena neurtu, kuantifikatu eta balioztatu da. AD-ren hasierako prodromo-fasea detektatzeko balio izan du.

2. Diagnostikorako ez-inbasiboak diren teknika adimendu automatikoen erabilera oso baliagarria izan da demenzien hautemate goiztiarrerako; beraz, ondoriozta daiteke hizketa-analisi automatikoa oso tresna erabilgarria bilakatzen ari dela hautemate kognitiborako. Azterketa ez-inbasibo horiek gaitz kognitiboak hautemateko eta monitorizatzeko metodo lasai, natural, merke eta zehatza eskaintzen dute. Une honetan, diagnostiko goiztiarra ezinbestekoa da .

3. Metodologia eta tresna ez-inbasiboak hainbat eszenatokitan erabili daitezke, diagnostikorako, kontrolerako eta neuroerrehabilitaziorako.

4. Zenbait ataza/protokoloren bidez (estandarrek barne), hizketatik informazio oso garrantzitsua atera daiteke diagnostiko, neuroerrehabilitazio eta eboluzioaren jarraipena egin ahal izateko. Hori dela eta, neurketa-ataza egokien aukeraketa oinarri-oinarrizkoa da. Atazaren hizketa-komplexutasunak zerikusi zuzena dauka sistemak atera dezakeen informazioarekin. Berez, aldaketa sotilagoak neurtzeak hizketa espontaneoagoa izatea beharko luke prozesu kognitiboagoa sortu ahal izateko. Hizketa murriztuaren atazek, ordea, oroimen-nahasmendua detektatzeko doitasun handia eskaintzen dute hitz-jarioa bezalako adierazle sendoak direla bide. Gainera, komunikatzeko ezinbestekoak diren elementuek, erantzun emozionalek adibidez, ez dituzte hizketa murriztuaren atazak behar, behar bezala neurtuak izateko. Lan honetan, hizketa-komplexutasunaren maila desberdinak dituzten hiru ataza (AN, PD eta SS)

7. KAPITULUA: ONDORIOAK

aztertzen dira, eta ebaluazio sakonagoa egitera bultzatzen gaituzten etorkizun handiko emaitzak lortzen dira.

5. Hizketa-komunikazioaren analisi automatikoak oso screening-metodologia naturala eta erabilgarria ematen du, zeina pazienteen ingurunean aplikatu baitaiteke, teknikarien laguntzarik gabe: Hizketa Automatikoaren Analisia (Automatic Speech Analysis, ASA, ingelesez), Hizketa Ezagutze Automatikoa (IEA) (Automatic Speech Recognition, ASR, ingelesez) edota Emozio Erantzunaren Analsiak (Emotional Request Analysis, ERA) eta Fonazio Biomekanikoak.

6. AD-n jario- eta semantika-harremanak aztertzeko tresna baliagarriak dira, baita emozio-erantzunak balioesteko ere. Metodo horiek Hizketa Automatikoaren Analsian eta/edo jarrera aldaketetan dute oinarria. AD detekzio goiztiarrerako etorkizun handiko diagnostikorako lanabesak dira, ez-inbasiboak, merkeak, erabiltzeko errazak, eskuragarriak eta hizkuntzatik independenteak direlako. Beraz, inguru erreal ia guztietarako dira baliozkoak. Kostu gehigarri handirik gabe.

7. Hizketarekiko sasi-independenteak diren osagaien ikerketetatik oso informazio baliagarria ateratzen da. Narriadura Kognitibo Arina edo Mild Cognitive Impairment (MCI) prodromo-etapara bideratuta, hizkuntzarekiko sasi-independenteak diren hizketaren azterketa automatikorako antzemate goiztiarrerako teknika eta metodologiak garatu dira. Honeka baimenduko du proba mediku bera baliagarria izatea hizkuntza gehienetarako.

8. Ez hizketak soilik, isiluneek eta disfluentziek ere galera kognitiboari buruzko informazioa daukate bere baitan biomarkadore erraza eta naturala bihurtuz. Bereziki disfluentziaren azterketara bideratu dira. Emaitzak onak izateaz gain unibertsala den (edozein hizkuntzarako balio duen) teknologia garatu da. Beraz, disfluentziak oso elementu interesgarriak dira eta MCI ondo detektatzeko erabilgarriak izan daitezke. Bai ahots isiluneek zein disfluentziek oso informazio garrantzitsua dute ahoskatutako mezuaren esanahia deskodetzeko.

9. Erabilgarritasunaren estandarretan oinarria duten web-sistema integral merkearen garapena egiteak teknikarien laguntza gutxiagotzea ekartzen du.

10. Beste gaixotasun askok bezala, galera kognitiboak prebalentzia desberdina eduki dezake generoaren arabera. Lortutako emaitzak beste azterketa medikuen osagai bihurtzen da genero-desberdintasunak ere ikusi ahal izateko. Hormonen eragina ahotsaren aldaketan erraz ikus baitaitezke.

11. Heldutasunean galera kognitiboen inpaktuari aurre egiteko, zeinak epidemia bilakatzeko bidean dauden, diagnostiko goiztiarra eta zehatza lagungarria da inpaktu hori gutxiagotzeko hala eriengan nola senide eta gizartearengan ere, eta geroaldia behar bezala antolatzeke aukera emango du eraginkorrakoak diren tratamenduekin. Horrek

ADren eta MCIren patologiek sortzen duten inpaktu ekonomiko negatiboa murrizten lagunduko du.

12. Erantzun emozionalen kuantifikazioak ere informazio garrantzitsua ematen digu gaitzaren prodromo-fasean. Askotan, lehenengo faseetan gaitza depresioarekin batera agertzen delako. Erantzun emozionalak eta honek ezagutza- eta komunikazio-defizit zein portaera-sintometan duen eraginak AD-ren prodromo-fasearen adierazle izan daiteke. Mezuaren apalagotzea somatzen da MCI duten pertsonengan eta gaitzaren lehen etapetan aurkitzen garenaren adierazle izan daiteke. Alzheimerdun gaixoengan, mintzamina ez ezik, erantzun emozionala ere kaltetu egiten da. Emozioak giza gogoaren arkitekturarekin zerikusia dituzten prozesu kognitiboak dira eta honako prozesu hauekin lotuta daude: erabakiak hartzea, oroimenaren kudeaketa edota arreta, zeinak hertsiki lotuta dauden komunikazioarekin, ikaste eta ulertzearekin.

13. Garaturiko tresna beste patologietan eta seinaleetan ere erabil daitezke; besteak beste, Parkinson edo Oinarrizko Dardara delakoetan.

14. Tresna baliagarria izango da diseinu unibertsalei jarraituz, eszenatoki desberdinetan eta erabiltzaile desberdinentzat erabilgarria izateko.

7.1.2 Ondorio zehatzak

1. Hiru ahots hiru datu-base berriak sortu dira AD/MCI era automatikoan aztertzeko. Horietako bi CITA-Alzheimer Fundazioak sortuta.

2. Bat-bateko hizketa irekiagoaz baliatuz lortu dira hizkuntzarekiko independenteak diren ezaugarriak eta baita AD/MCI oso goiz agertzen den seinalea aztertzeko bidea irekitzea. AD/MCI detektatzeko, emozioa aztertea era automatikoan lortu da ezaugarri linealak, ez-linealak eta Emozio-Temperaturaren bidez.

3. Ezaugarri linealak, ez-linealak (fraktalak eta entropia), perzepzioarekiko ezaugarriak eta ezaugarri aurreratuak aldaketa txikiak detektatzeko aztertu eta aplikatu dira. Bestalde, permutazio entropikoaren aldaera berria aztertu da, emaitza hobekia lortuz. Hain zuzen ere, Multieskaladun Permutazio Entropia pisatua.

4. MCIren diagnostikorako, eta bereziki disfluentzien azterketarako proposatzen den modelatze berriak emaitzak hobetzen ditu MCIren diagnostikoan laguntzeko.

5. Hizkera eta disfluentzien azterketa automatikorako proposamen berritzailea mahaigaineratzen da.

6. Ezaugarriak aukeratzeko metodologia automatiko misto berria erabili da proba estatistikoetan (irizpide medikuen baitan) eta automatikoan (ezaugarriak aukeratzeko metodo automatiko klasikoak).

- Ezaugarrien erauzketa (ezaugarri klasikoak, hautemate-ezaugarriak, Castiglioni-ren dimentsio fraktala edo Multieskaladun Permutazio Entropia) egiten da.
- Ezaugarrien hautaketa automatikoak (Mann-Whitney U-testak eta SVM balidazioak) ezaugarri eredugarrienen hautaketan oinarria duen eredu ez-lineal multifaktoriala aurkezten du.
- Proposatutako ezaugarri-bektorea Neurona-sare Konboluzionalak (CNN) barne hartzen ditu eta sailkapena egiten du.

7. Emaitzak oso onak dira, itxaropentsuak dira eta ikerketarako lerro berriak irekitzen dituzte.

8. Laguntzaileei eta pazienteei laguntzen die diagnostikoan, monitorizazioan eta neuroerrehabilitazioan.

9. Neurketa-ataza egokien aukeraketa, parametro egokien hautaketa eta ML algoritmoaren aukeraketak berak eragin zuzena daukate bukaerako aurreikuspenean.

10. Protokolo berriak eta sinpleak diseina daitezke ingurune eta lengoia-mota guztietarako non beste metodologiak eta tresnak fusioa daitezkeen; esaterako, idazkeraren azterketa automatikoa.

11. Sistema malgua behar da dauden protokoloei eta osasun-espezialisten diagnostikorako erabiltzen diren irizpideei aurre egiteko. Hori oso garrantzitsua izan daiteke fase goiztiarretan faltsu positiboak gutxitzeko eta persona osasuntsuak ez medikatzeke.

12. Sistema egokitu da eszenatoki errealetan funtzionatu ahal izateko eta kasu konplexuenetan ezaugarriak aberastearen metodoen bidez eta paradigma egoki eta sendoen bidez emaitza optimoak lortu dira eta kostu konputazionala nabarmen jaitsi da. Denbora errealera bideratutako sistema garatu da.

13. Protokolo bakoitzak sistemaren konfigurazio propioa eskatuko du, bai ezaugarrien aldetik bai atazaren aldetik, bai ereduaren aldetik ere.

14. Azkenik, CNNak tresna itxaropentsu bezala ageri dira aldaera eta sintoma oso arinak direnean MCI kasuan bezala. Gainera, Dropout-ak sistemaren eraginkortasuna eta kostu konputazionala jaitea lortzen du. Eredu horiek duten potentzia kasu zehatzetan erabili behar dira, eta beste paradigmak dituen onuraz eta kostu konputazionalaz baliatuta, emaitza optimoak lor ditzake.

15. Lortutako emaitzak hobetu eta onak izan diren arren, ikerketa-lerro berriak ireki dira. Prozesu osoan zehar justifikatuak dauden erabakiak hartu eta emaitzak hobetu dira. Lan-hipotesi berriak sortu dira.

16. Modelatze linealek zein ez-linealek ezaugarri-bektorea asko aberasten dute. Mintzamenaren konplexutasuna oinarri hartuta, hizketa automatikoaren analisisian funtsatutako multi-hurbilketa ez-lineala egin da. Horrela, jasotako laginarekiko adierazgarriak eta ordezkagarriak diren bektoreak lortzen dira eta sareari sartuko diogun informazioa ezaugarritzea dela eta oso aberatsa izango da.

17. Inguru erreal eta konplexuetara bideratutako teknika eta metodologiaren garapenak emaitza onak lortu ditu. Garatutako modelatze-eredu berriak prozesua zerbait integraltzat hartzen du eta datuen gainean hasieratik egiten du lan, kasuistikaren konplexutasuna kontuan hartuta. Horrek sailkapen-emaitzak hobetzea ekarri du.

18. Parametro lineal, ez-lineal, entropia eta fraktaletan oinarritutako metodologia eta parametro bioinspiratuak ikertzeak inguru konplexuen edota zatatsuen sendotasuna lortzera bideratuak dauden parametroetan sakontzea eraman gaitu.

19. Teknika mistoa aurrez aberastutako datuek Deep Learning sareei ezaugarri-mapak egiteko lana errazten diete. Neurona-sareari datu aberastuak sartzean makinak hobeto ikasten du elementu zehatzak bilatuz.

20. Deep Learning-ek bere trebezia erakutsi badu datu-matrizeak ezaugarritzeko eta horiek konprimitzeko, sartzan diren datuak, gordinak izan ordez, ezaugarri-mapak badira lortzen diren emaitzak hobeak dira. Horrek esan nahi du CNN-ra txertatuko den informazioa aberastua dela eta bektore parametrikokoak direnak ezaugarri-bektoreak balira bezala ulertuko dituela sareak. Ezaugarritze-lan handia egina ematen zaio sareari. Ikuspuntu guztiz berritzaile honek emaitza onak eman ditu.

21. Hala ere, ezaugarrien garrantzia eta datu aberatsen garrantzia ezinbestekoa da eszenatoki konplexuetan eta datu konplexuetan. Ataza bakoitzak datu propioak ditu eta hori kontuan hartuta egokitzapen unibertsala eta malgutasuna bilatu da.

22. Kernelaren tamaina txikiak jatorrizko ezaugarriak mantendu eta ikasketaren prozesua errazten du. Kontuan hartu behar dugu sareari txertatzen diogun matrizea oso aberastua eta konprimitua dela (parametroen hautaketa fase baten emaitza da) eta ezaugarriak oso definituak dituela. CNN-ak sortuko dituen ezaugarri-mapa berriek matrize aberastuaren "pixel"-en arteko harreman eta "korrelazioak" ezaugarritzeko gaitasuna dute. Beraz, sartzan dugun informazioa, konprimitua denez, sortuko diren ezaugarri-mapa berriek jatorrizko informazioaren nortasuna mantentzea interesatzen zaigu. Hori dela eta, konboluziorako kernela 2x2koa izango da emaitzak hobetzeko. Gainera, matrizearen tamaina txikia dela hartu behar dugu kontuan.

23. 2D-ko matrizeak kasu konplexuetan beste paradigmek baino emaitza hobeak ematen ditu askotan. 2D-ko matrizeetarako konboluzio kernelak ondoko pixelekiko duen muga azalera zabalagoa da 1D-ko bektorearekin konparatzen badugu. Hori horrela, mugakide gehiago izateak ezaugarri gehiagoren arteko korrelazioa hobeto aurkitzeko gaitasuna eman dio sareari.

24. Pooling-aren urratsaren zabalera txikia izateak ezaugarri-bektore txikietara hobeto moldatzea dakar. Sarrerako matrizearen tamaina txikia izanik jatorrizko informazioaren nortasuna mantentzea interesatzen zaigu makinak xehetasunik ez galtzeko eta ikasketa-prozesua optimizatzeko. Gainera, ertzen kudeaketa erraztuko da. Horrek emaitzak hobetuko ditu.

25. Ikasketa sakonak orain arte erabilitako beste Machine Learning algoritmoen diagnostiko-emaitzak hobetzeko gaitasuna erakutsi du baldin eta datu aberatsak baditu. Konboluzio zein pooling teknikek eta neurona-sarearen sakontasunak (2 geruza baino gehiago) ahots bidezko fitxategietan zehar ezaugarriak aurkitzeko gaitasun handia erakutsi du eta beste algoritmoekin lortutako emaitzak hobetu ditu.

26. Saillapen automatikoa egin da Multilayer Peceptron (MLP) eta neurona-sare konboluzionalak direla bide (CNN), eta bigarrenak lehenak baina emaitza hobekak lortu ditu zenbait kasutan eta Dropout erabiliz kostu konputazional egokiagoa.

7.1.3 Etorkizunerako lerroak

1. Sistema horiek diagnostiko zehatzagoak lortuko ditu onurak ekarriz ez soilik ikuspuntu teknikitik baizik eta baita ikuspuntu sozioekonomikotik ere.

2. Metodo ez-inbasiboak etorkizunerako oso tresna erabilgarriak izango dira datu-iturri desberdinak, sentzore anitzeko irteerak eta eszenatokiak zabalduz.

3. Beraz, Big data vs Datu Konplexuak eta Thick Data etengabe agertuko dira Ingeniaritza Biomedikuan eta beste zenbait arlotan. Patologia bakoitzak eszenatoki desberdina planteatzen du eta.

4. Data Mining, ikerkuntza operatiboa, Big data estrategiak eta ezaugarrien erauzketa automatikoa datuak aberasteko eta informazioa bilatzeko metodologiak ezinbestekoak izango dira hurrengo urteetan osasun arloan.

5. Sentzore fusioen teknikak aplikatuko dira seinale ez-inbasiboen Big data estrategietan.

6. Beraz, datu-base handiagoekin ebaluatuko da eta beste metodologiekin integratuko.

7. Genero-hurbilpenak ere aztertuko dira, kontuan izanda Ingeniaritza Biomedikuan eta zehazki AD/MCIn prebalentzia desberdina agertzen dela generoren arabera.

8. Ditugun laginak gutxi dira eta horrek Alzheimer-en ikasketa-prozesua mugatu dezake. Hutsune hori saihesteko ezaugarriak laginketa denbora batez jaso ahal izatea proposatzen da. Hau da, bektore asko lortu lagin bakoitzetik. Horrela, datu-basea

handitzea lortuko dugu. Datuak gutxi izanik lagin bakoitzetik azpilaginak ateratzea proba daiteke datu-basea handitu eta CNN-aren entrenamenduan sakondu.

9. Bilatzen den emaitza hobetzeko helburuz geruza konboluzionalen parametroak eta sakontasuna doitzeko lan handia dago egiteko.

10. Emaitzak hobetu diren arren hobetu daitezke eta CNN-en konfigurazioak eta sakontasunak horretan lagun dezakete.

11. Multi-hurbilketa ez-linealak hautemate sakonago bat egitera akuilatzen duten etorkizun handiko emaitzak ematen ditu.

12. Mintzamen-konplexutasuneko maila desberdinak dituzten hiru egintza edo atazen azterketatik abiatuta, mintzamen automatikoaren analisisian oinarritutako multi-hurbilketa ez-lineala dago. Hizkuntza-konplexutasunaren maila desberdinekin egindako hiru atazen (AN, PD eta SS) diagnostikoen arteko alderaketa dago. Azterketa horrek ikerbide berriak irekitzen ditu gaitza kognitiboaren balioeste eta monitorizatzeko sakonagoetarako.

13. Alzumeric-en migrazioa. Sistema osoa biltzeko teknologiak erabili dira eta hori Python-era migratzea aukera ona izan daiteke modelatu, prozesu konputazionala azkartu eta arkitektura sinpleagotzeko.

14. Azterketen metodologia hobe daiteke emaitzak itxaropentsuak diren arren.

15. Hizkeraren osagai linguistikoen ikerketak eta horien hizketa zein biomekanikarekiko independenteak diren elementuetan oinarria duten azterketen metodologia hobe daitezke emaitzak, itxaropentsuak diren arren. Inplementatutako modelatze berriak ikerketa-lerro berriak ekarri ditu.

16. Komunikazio zifratuetan sakondu. Alzumeric-erako garatutako geruza zifratua beta bertsioan dagoela eta ekoizpen eremura salto emateko bidean dagoela eta, aldaketak biziko dituela kontuan hartu behar da.

17. Ikasketa sakonean mahaigaineratzen diren paradigma berriak berrikustea. Neurona-sare bidezko ikasketa sakona aldaketa handiak izaten ari da gaur egun. Adibidez, DNC bezalako arkitekturek datu egituratu eta konplexuekin ondo lan egiten dutela erakutsi dute. Baina ez DNC-a bakarrik. AlexNet-etik hasita GAN-etara aukerak asko dira. Gure kasurako, alternatiba hobereena bilatzeak ikerketa-lerro berriak irekitzen ditu.

18. Beste patologietara zabalduko dira tesis honen metodologiak haien arteko antzekotasunak eta desberdintasunak aztertuz ere.

BIBLIOGRAFIA

- [1] World Alzheimer Report 2016. Available online: www.alz.co.uk/research/world-report-2016 (accessed on 29 May 2017).
- [2] Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, Dauwels J, Soekadar SR, Mueller S, Linnemann C, Bridenbaugh SA, Kanagasingam Y, Martins RN, O'Bryant SE. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer & Dementia* 2015;11(5):561-78.
- [3] Querfurth HW, LaFerla FM. Alzheimer's disease. *N Engl J Med* 2010;362:329-44.
- [4] Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA. Alzheimer disease in the US population: prevalence estimates using the 2000 census. *Arch Neurology* 2003; 60:1119-22.
- [5] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the National Institute on Aging-Alzheimer's Association workgroup on Alzheimer's disease. *Neurology* 1984;24:939-44.
- [6] McKhann G, Knopman DS, Chertkow H, Hyman BT, Jack CR.Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement* 2011;7:263-69.
- [7] Van de Pol LA, Van der Flier WM, Hensel A, Gertz HJ, Scheltens P. The effects of age and Alzheimer's disease on hippocampal volumes A MRI study. *Alzheimer's and Dementia*;1(1).
- [8] Alzheimer's Association. Available online: <http://www.alz.org/>, (accessed on 10 October 2016).
- [9] Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Under diagnosis of dementia in primary care: variations in the observed prevalence and comparisons to the expected prevalence. *Aging Ment Health* 2011;15(8):978-84.
- [10] K napp, M. L. Essentials of nonverbal communication. Holt, Rhinehart & Winston 1980.
- [11] Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience* 1980;1:3-33.
- [12] Morris JC. The Clinical dementia rating (CDR): Current version and scoring rules. *Neurology* 1993;43.
- [13] American Psychiatric Association. Available online: www.psychiatry.org, (accessed on 15 Dec 2015).
- [14] Klimova B, Maresova P, Valis M, Hort J, Kuca K. Alzheimer's disease and language impairments: social intervention and medical treatment. *Clinical Interventions in Aging* 2015;10:1401-08.
- [15] Buiza C. Evaluación y Tratamiento de los Trastornos del Lenguaje; Matia Fundazioa 2010.
- [16] Martinez F, Garcia J, Perez E, Carro J, Anara JM. Patrones de Prosodia expresiva en pacientes con enfermedad de Alzheimer. *Psicothema* 2012;24:16-21.
- [17] Silagi M, Ferreira P, Zazo K. Naming ability in patients with mild to moderate Alzheimer's disease: what changes occur with the evolution of the disease?. *Clinics* 2015;70(6):423-28.
- [18] De Lira J, Ortiz KZ, Campanha AC, Bertolucci PHF, Minett TS C. Microlinguistic Aspects of the oral Narrative in Patients with Alzheimer's Disease. *Int Psychogeriatr*.2011;23(3):404-12.
- [19] Lopez-de-Ipina K, Alonso J-B, Manuel Travieso C, Solé-Casals J, Egiraun H, Faundez-Zanuy M, Ezeiza A., Barroso N., Ecay M., Martinez-Lage P., Martinez de Lizardui U. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors* 2013;13:6730-45.
- [20] Appell J, Kertesz A, Fisman M. A study of language functioning in Alzheimer patients. *Brain Lang* 17 1982:73-91.

- [21] Lopez de Ipiña K, Solé Casals J, Eguiraun H, Alonso JB, Travieso CM, Ezeiza A, Barroso N, Ecay Torres M, Martínez Lage P, Beitia B. Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach. *Computer Speech & Language* 2015; 30(1).
- [22] Schecker M, Kochler C, Schmidtke K, Rauha R. Are There Any Connections between Language Deficits and Cognitive Slowing in Alzheimer's Disease. *Dement Geriatr Cogn Dis Extra* 2014;4(3):442-49.
- [23] López-de-Ipiña K, Alonso JB, Solé-Casals J, Barroso N, Henríquez P, Faundez-Zanuy M, Travieso CM, Ecay-Torres M, Martínez-Lage P, Eguiraun H. On Automatic Diagnosis of Alzheimer's Disease Based on Spontaneous Speech Analysis and Emotional Temperature. *Cognitive Computation* 2013;5(1):136-51.
- [24] Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr, Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement* 2011;7:280-92.
- [25] Verma M, Howard RJ. Semantic memory and language dysfunction in early Alzheimer's disease: a review international. *Journal of Geriatric Psychiatry* 2012; 27(12):1209-17.
- [26] Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia* 2004;42:1212-22.
- [27] Laws KR, Adlington RL, Gale TM, Moreno-Martínez FJ, Sartori G. A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia* 2007;45:2674-82.
- [28] Chenery HJ, Murdoch BE, Ingram CL. An investigation of confrontation naming performance in Alzheimer dementia as a function of disease severity. *Aphasiology* 1996;10:425-41.
- [29] Pollman S, Haupt M, Kurz A. Changes of the relative severity of naming, fluency and recall impairment in the course of dementia of the Alzheimer type. *Dementia*. 1995;6(5):252-7.
- [30] Bayles K, Tamoeda CK. Confrontation naming impairment in dementia. *Brain Lang*. 1983;19(1):98-114.
- [31] Hodges JR, Salmon DP, Butters N. Semantic memory impairment in Alzheimer's disease: failure of access or degraded knowledge? *Neuropsychologia* 1992;30(4):301-14.
- [32] Adlington RL, Laws KR, Galea TM. Visual processing in Alzheimer's disease: Surface detail and colour fail to aid object identification. *Neuropsychologia* 2009;47(12):2574-83.
- [33] Moreaud O, David D, Charnallet A, Pellat J. Are semantic errors actually semantic? Evidence from Alzheimer disease. *Brain Lang* 2001;77(2):176-86.
- [34] Faust ME, Balota DA, Multhaup KS. Phonological Blocking During Picture Naming in Dementia of the Alzheimer Type. *Neuropsychology* 2004;18(3):526-36.
- [35] Schott JM, Crutch SJ, Frost C, Warrington EK, Rossor MN, Fox NC. Neuropsychological correlates of whole brain atrophy in Alzheimer's disease. *Neuropsychologia*. 2008;46(6):1732-7.
- [36] Gold B. Lifelong bilingualism and neural reserve against Alzheimer's disease: a review of findings and potential mechanisms. *Behav Brain Res*. 2015;281:9-15.
- [37] Mueller K, Kosci R, La Rue A, Clark L, Hermann B, Johnson S, Sager M. Verbal Fluency and Early Memory Decline: Results from the Wisconsin Registry for Alzheimer's Prevention. *Arch Clin Neuropsychol* 2015;30(5):448-57.
- [38] Lopez-de-Ipiña K, Martínez-de-Lizarduy U, Barroso N, Ecay-Torres M, Martínez-Lage P, Torres F, Faundez-Zanuy M. Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impairment detection: A non-linear language independent approach. *Bioinspired Intelligence (IWOB)* 2015;101-104.
- [39] Gómez-Vilda P, Rodellar-Biargea V, Nieto-Lluisa V, López de Ipiña K, Álvarez-Marquina A, Martínez-Olalla R, Ecay-Torres M, Martínez-Lage P. Phonation biomechanical analysis of Alzheimer's Disease cases. *Neurocomputing* 2015;167(1):83-93.

- [40] López-de-Ipiña K, Alonso-Hernández JB, Solé-Casals J, Travieso-González CM, Ezeiza A, Faúndez-Zanuy M, Calvo PM, Beitia B. Feature selection for automatic analysis of emotional response based on nonlinear speech modeling suitable for diagnosis of Alzheimer's disease. *Neurocomputing* 2015;150(B):392–401
- [41] John Wiley & Sons. *Handbook of Cognition and Emotion*. Tim Dalgleish, Mick J. Power 1999.
- [42] D Sander D, D Grandjean D, Scherer KR. A systems approach to appraisal mechanisms in emotion. *Neural Networks* 2005;18(4):317-52.
- [43] Picard RW. *Affective Computing*. MIT Press 1997.
- [44] Cowie R, Douglas-Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, Taylor JG. Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine* 2001;18(1):32-80.
- [45] Shimokawa A, Yatomi N, Anamizu S, Torii S, Isono H, Sugai Y, Kohno M. Influence of deteriorating ability of emotional comprehension on interpersonal behaviour in Alzheimer-type dementia. *Brain Cognit* 2001;47:423–33.
- [46] Henry JD, Rendell PG, Scicluna A, Jackson M, Phillips LH. Emotion experience, expression, and regulation in Alzheimer's disease. *Psychol Aging* 2009;24:252-57.
- [47] Goodkind MS, Gyurak A, McCarthy M, Miller BL, Levenson RW. Emotion regulation deficits in frontotemporal lobar degeneration and Alzheimer's disease. *Psychol Aging* 2010;25:30-37.
- [48] Mammarella N. And B. Fairfield. Emotional Working Memory and Alzheimer's Disease. *Dement Geriatr Cogn Dis Extra* 2011;1(1):124-38.
- [49] Mammarella N, Fairfield B. Emotional Working Memory and Alzheimer's Disease. *International Journal of Alzheimer's Disease* 2014;
- [50] Mikels JA, Reuter-Lorenz PA, Beyer JA, Fredrickson BL. Emotion and working memory: evidence for domain-specific processes for affective maintenance. *Emotion* 2008;8(2) 256-66.
- [51] Mather M, M Knight. Goal-directed memory: the role of cognitive control in older adults emotional memory. *Psychol Aging* 2005;20(4) 554-70.
- [52] Schultz RR, de Castro CC, Bertolucci PHF. Memory with emotional content, brain amygdala and Alzheimer's disease. *Acta Neurologica Scandinavica* 2009; 120(2)101-10.
- [53] Zannino GD, Caltagirone C, Carlesimo GA. The contribution of neurodegenerative diseases to the modelling of semantic memory: A new proposal and a review of the literature. *Neuropsychologia* 2015;75:274-90.
- [54] Dingemanse M, Torreira F, Enfield NJ. (2013). Is "Huh?" a Universal Word? *Conversational Infrastructure and the Convergent Evolution of Linguistic Items*. *Plos One* 2013;8(11).
- [55] Fox Tree JE. The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language* 1995;34(6)709-38.
- [56] Mekyska J, Janousová E, Gómez Vilda P, Smékal Z, Rektorová I, Eliasova I, Kostalova M, Mrackova M, Alonso Hernández JB, Faúndez-Zanuy M, López de Ipiña K. Robust and complex approach of pathological speech signal analysis. *Neurocomputing* 2015;167: 94-111.
- [57] Silagi ML, Bertolucci PH, Ortiz KZ. Naming ability in patients with mild to moderate Alzheimer's disease: what changes occur with the evolution of the disease? *Clinics*. 2015;70(6):423-28.
- [58] Clark LR, Schiehser DM, Weissberger GH, Salmon DP, Delis DC, Bondi MW. Specific measures of executive function predict cognitive decline in older adults. *Journal of the International Neuropsychological Society* 2012;18 (1):118–27.
- [59] Diggle K, Mueller R, Kosciak L, LaRue A, Clark L, Hermann B, Johnson S, Sager M. Verbal Fluency and Early Memory Decline: Results from the Wisconsin Registry for Alzheimer's Prevention. *Archives of Clinical Neuropsychology* 2015;30:448-57.
- [60] Barr A, Brandt J. Word-list generation deficits in dementia. *Journal of Clinical and Experimental Neuropsychology* 1996;18 (6):810-22.

BIBLIOGRAFIA

- [61] Troyer AK, Moscovitch M, Winocur G, Leach L, Freedman M. (1998). Clustering and switching on verbal fluency tests in Alzheimer's and Parkinson's disease. *Journal of the International Neuropsychological Society* 1998;4 (2):137-43.
- [62] Van Lancker Sidtis D. Where in the brain is nonliteral language? *Metaphor and Symbol*. 2006;21(4):213-44.
- [63] Chand V, Baynes K, Bonnici L, Tomaszewski S. A Rubric for Extracting Idea Density from Oral Language Samples. *Curr Protoc Neurosci* 2012;
- [64] Gilmore G, Groth K, Thomas C. Stimulus contrast and word reading speed in Alzheimer's disease. *Exp Aging Res* 2005;31(1):15-33.
- [65] Dodge H, Mattek N, Gregor M, Bowman M, Seelye A, Ybarra O, Asgari M, Kaye J. Social Markers of Mild Cognitive Impairment: Proportion of Word Counts in Free Conversational Speech. *Current Alzheimer Research* 2015; 12:513-19.
- [66] Woodford HJ, George J. Cognitive assessment in the elderly: a review of clinical methods. *QJM: An International Journal of Medicine* 2007.
- [67] Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the physician. *J Psychiatr Res* 1975;12:189-98.
- [68] Kaplan E, Goodglass H, Weintraub S. The Boston naming test 1983.
- [69] Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, Mellits ED, Clark C. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part 1. Clinical and Neuropsychological Assessment of Alzheimer's Disease. *Neurology* 1989;39:1159-65.
- [70] CERAD-An overview. Available online: cerad.mc.duke.edu/Library/CERAD.pdf (accessed on 04 April 2016)
- [71] Delis DC, Kramer JH, Kaplan E, Ober BA. The California Verbal Learning Test. New York: Psychological Corporation 1987.
- [72] Brandt J, Benedict RHB. Hopkins Verbal Learning Test-Revised Professional manual. Lutz, FL: Psychological Assessment Resources 2001.
- [73] Johnstone B, Wilhelm KL. The longitudinal stability of the WRAT-R Reading subtest: Is it an appropriate estimate of premorbid intelligence? *Journal of the International Neuropsychological Society* 1996;2:282-85.
- [74] Wilkinson GS, Robertson GJ. Wide Range Achievement Test 4 professional manual. Psychological Assessment Resources 2006.
- [75] Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. *Archives de Psychologie* 1941;28(112)286-340.
- [76] Rey A. L'examen clinique en psychologie. Presses Universitaires de France 1958.
- [77] Ostberg P, Fernaeus SE, Bogdanovic N, Wahlund LO. Word sequence production in cognitive decline: forward ever, backward never. *Logoped Phoniatr Vocol* 2008;33:126-35.
- [78] Pakhomov S., Laura SH, Kelvin OL. Automated semantic indices related to cognitive function and rate of cognitive decline. *Neuropsychologia* 2012;50(9)2165-75.
- [79] Weakley A, Schmitter-Edgecombe M. Analysis of Verbal Fluency Ability in Alzheimer's Disease: The Role of Clustering, Switching and Semantic Proximities. *Archives of Clinical Neuropsychology* 2014;29:256-68.
- [80] Randt CT, Brown ER, Osborne DP. Randt memory test. New York University, Department of Neurology 1981.
- [81] Kensinger EA, Anderson A, Growdonb JH, Corkin S. Effects of Alzheimer disease on memory for verbal emotional information. *Neuropsychologia* 2004;42:791-800.

- [82] Padovan C, Versace R, Thomas-Anterion C, Laurent B. Evidence for a selective deficit in automatic activation of positive information in patients with Alzheimer's disease in an affective priming paradigm. *Neuropsychologia* 2002;40(3):335-39.
- [83] Giffard B, Laisney M, Eustache F, Desgranges B. Can the emotional connotation of concepts modulate the lexico-semantic deficits in Alzheimer's disease?. *Neuropsychologia* 2008;47(1):258-67
- [84] Tapus A, Tapus C, Mataric MJ. The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia. *IEEE International Conference on Rehabilitation Robotics* 2009;1:924-29.
- [85] Zhang F, Ho YW, Fung HH. Learning from Normal Aging: Preserved Emotional Functioning Facilitates Adaptation among Early Alzheimer's Disease Patients. *Aging Dis* 2015;6(3):208-15.
- [86] Bergmann HC, Rijpkema M, Fernández G, Kessels RP. The effects of valence and arousal on associative working memory and long-term memory. *PLoS One* 2012;7(12).
- [87] Faundez-Zanuy M, Hussain A, Mekyska J, Sesa-Nogueras E, Monte-Moreno E, Esposito A, Mohamed Chetouani, Garre-Olmo J, Abel A, Smekal Z, Lopez-de-Ipiña K. Biometric applications related to human beings: there is life beyond security. *Cogn Comput* 2013;5:136-51.
- [88] Forbes KE, Venneri A, Shanks MF. Distinct patterns of spontaneous speech deterioration: an early predictor of Alzheimer's disease. *Brain Cogn* 2002;48: 356-61.
- [89] Espinoza-Cuadros F, Garcia-Zamor MA, Torres-Boza D, Ferrer-Riesgo CA, Montero-Benavides A, Gonzalez-Moreira E, Hernandez-Gomez LA. A Spoken Language Database for Research on Moderate Cognitive Impairment: Design and Preliminary Analysis. *Advances in Speech and Language Technologies for Iberian Languages* 2014;8854:219-28.
- [90] Gonzalez-Moreira E, Torres-Boza D, Kairuz H, Ferrer C, Garcia-Zamora M, Espinoza-Cuadros F, Hernandez-Gomez L. Automatic Prosodic Analysis to Identify Mild Dementia. *BioMed Research International* 2015.
- [91] Romero B, Kurz A. Deterioration of spontaneous speech in AD patients during a 1-year follow-up: Homogeneity of profiles and factors associated with progresión. *Dementia* 1996;7(1):35-40.
- [92] Thomas C, Keselj V, Cercone N, Rockwood K, Gu J, Liu PX. Automatic detection and rating of dementia of Alzheimer type through lexical analysis of spontaneous speech. *Mechatronics and Automation, 2005 IEEE International Conference* 2005;3:1569-74.
- [93] Peintner B, Jarrold W, Vergyri D, Richey R, Tempini ML, Ogar J. Learning diagnostic models using speech and language measures. *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008;4648-51.
- [94] Verucci da Cunha AL, Bender de Sousa L, Mansur LL, Aluisio SM. Automatic proposition extraction from dependency trees: helping early prediction of alzheimer's disease from narratives. *Computer-Based Medical Systems (CBMS), 2015 IEEE 28th International Symposium* 2015; 1:127-30.
- [95] Roark B, Mitchell M, Hosom JP, Hollingshead K, Kaye J. Spoken Language Derived Measures for Detecting Mild Cognitive Impairment. *IEEE Transactions on Audio, Speech, and Language Processing* 2011; 19(1):2081-90.
- [96] Martinez-Sanchez F, Meilan J, Garcia-Sevilla J, Arana JM. Oral reading fluency analysis in patients with Alzheimer disease and asymptomatic control subjects. *Neurologia* 2013;28(6):325-31.
- [97] Orlandi S, Dejonckere PH, Schoentgen J, Lebacqz J, Rruqja N, Manfredi C. Effective pre-processing of long term noisy audio recordings: An aid to clinical monitoring. *Biomedical Signal Processing and Control* 2013;8(6):799-810.
- [98] Khodabakhsh A, Kuscuoglu S, Demiroglu C. AutorÆa conjunta del libro: IEEE. Detection of Alzheimer's disease using prosodic cues in conversational speech. *Signal Processing and Communications Applications Conference (SIU)* 2014;1:1003-6.

- [99] Khodabakhsh A, Demiroglu C. Analysis of Speech-Based Measures for Detecting and Monitoring Alzheimer's Disease. *Methods in Molecular Biology* 2014;1246:159-173.
- [100] Garrard P, Rentoumi V, Gesierich B, Miller B, Gorno-Tempini ML. Machine learning approaches to diagnosis and laterality effects in semantic dementia discourse. *Cortex* 2014;55:122-9.
- [101] Rudzicz F, Wang R, Begum M, Mihailidis A. Speech Interaction with Personal Assistive Robots Supporting Aging at Home for Individuals with Alzheimer's Disease. *Special Issue on Speech and Language Processing for AT* 2015;7(2):6.
- [102] Baldas V, Lampiris C, Capsalis C, Koutsouris D. Early diagnosis of Alzheimer's type dementia using continuous speech recognition. *Lecture notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* 2011;55:105-10.
- [103] Fraser K, Rudzicz F, Graham N, Rochon E. Automatic speech recognition in the diagnosis of primary progressive aphasia. *Proceedings of the Fourth Workshop on Speech and Language Processing for Assistive Technologies* 2013.
- [104] Buschke H. Cued recall in amnesia. *Journal of Clinical and Experimental Neuropsychology* 1984;6:433-40.
- [105] Ahmed S, Haigh AM, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. *Brain* 2013;136(12):3727-37.
- [106] Meilan JJ, Martinez-Sanchez F, Carro J, Sanchez JA, Perez E. Acoustic markers associated with impairment in language processing in Alzheimer's disease. *Span J Psychol* 2012;15:487-94.
- [107] Königa A, Sat A, Sorinc A, Hoory R, Toledo-Ronen O, Derreumaux A, Maner V, Verhey F, Aalten A, Robert PH, David R. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* 2015;1(1):112-24.
- [108] Choi NG, Hegel MT, Marti N, Marinucci ML, Sirrianni L, Bruce ML. Telehealth problem-solving therapy for depressed low-income homebound older adults. *Am J Geriatr Psychiatry* 2014;22:263-71.
- [109] Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, Delis DC. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry* 2009;17:368-75.
- [110] Cerhan JH, Ivnik RJ, Smith GE, Tangalos EC, Petersen RC, Boev BF. Diagnostic utility of letter fluency, category fluency, and fluency difference scores in Alzheimer's disease. *Clin Neuropsychol* 2002;16:35-42.
- [111] Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia* 2004;42:1212-22.
- [112] Brandt J, Manning KJ. Patterns of word-list generation in mild cognitive impairment and Alzheimer's disease. *Clin Neuropsychol* 2009;23:870-79.
- [113] López-de-Ipiña K, Martínez-de-Lizarduy U, Calvo PM, Mekyska J, Beitia B, Barroso N, Estanga A, Tainta M, Ecay-Torres M. Advances on automatic speech analysis for early detection of Alzheimer Disease: a non-linear multi-task approach. *Current Alzheimer Disease* 2017;1-20 (accepted).
- [114] Sainath TN, Weiss RJ, Wilson KW, Li B, Narayanan A, Varianni E, Kim C. Multichannel signal processing with deep neural networks for automatic speech recognition. *Ieee-Acm Transactions on Audio Speech and Language Processing* 2017;25(5):965-979.
- [115] Vega MT, Mocanu DC, Famaey J, Stavrou S, Liotta A. Deep learning for quality assessment in live video streaming. *IEEE Signal Processing Letters* 2017;24(6):736-40.
- [116] Sun S, Luo C., Chen J. A review of natural language processing techniques for opinion mining systems. *Information Fusion* 2017;36:10-25.
- [117] Lingfeng W, Zehao H, Yongchao G, Chunhong P. Ensemble based deep networks for image super-resolution. *Pattern Recognition* 2017;68:191-8.

- [118] Cruz-Roa A, Gilmore H, Basavanthally A, Feldman M, Ganesan S, Shih N, Madabhushi A. Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent. *Scientific Reports* 2017;7:(46450).
- [119] Lu K, An X, Li J, He H. Efficient deep network for vision-based object detection in robotic applications. *Neurocomputing* 2017;245:31-45.
- [120] Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks. *IEEE International Conference on Acoustics, Speech and Signal Processing* 2013;6645-49.
- [121] LeCun Y, Bengio Y, Hinton GE. Deep learning. *Nature: International Weekly Journal of Science* 2005; 521(7553):436-44.
- [122] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Paper presented at the *Advances in Neural Information Processing Systems* 2012;2:1097-1105.
- [123] ILSVRC2012. Available online: <http://image-net.org/challenges/LSVRC/2012/>, (accessed on 20 May 2017).
- [124] Image classification. Available online: <https://goo.gl/cCdqhE> (accessed on 22 February 2017).
- [125] Badshah AM, Ahmad J, Rahim N, Baik SW. Speech emotion recognition from spectrograms with deep convolutional neural network. *International Conference on Platform Technology and Service, PlatCon* 2017.
- [126] Sarraf S, Tofghi G. Deep learning-based pipeline to recognize alzheimer's disease using fMRI data. *Proceedings of Future Technologies Conference* 2017.
- [127] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks 2014;818-33.
- [128] Zeiler MD, Krishnan D, Taylor GW, Fergus R. Deconvolutional networks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2010;2528-35.
- [129] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014.
- [130] Takahashi N, Gygli M, Pfister B, Van Gool L. Deep convolutional neural networks and data augmentation for acoustic event recognition. *Conference of the International Speech Communication Association* 2016;2982-86.
- [131] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Rabinovich A. Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2015;1-9.
- [132] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Conference on Computer Vision and Pattern Recognition* 2016;770-778.
- [133] Deep Learning Dive & Workshop on Convolutional and Recurrent Neural Networks. Available on: <https://www.slideshare.net/nervanasys/sd-meetup-12215>, (accessed on 12 July 2016).
- [134] CUIImage. Available online: <http://image-net.org/challenges/LSVRC/2016/results>, (accessed on 10 May 2017).
- [135] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. *Conference on Computer Vision and Pattern Recognition* 2014;580-87.
- [136] Girshick R. Fast R-CNN. *International Conference on Computer Vision* 2015;1440-48.
- [137] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2017;39(6):1137-49.
- [138] Facial Expression Recognition with Faster R-CNN. Available online: <https://goo.gl/epOSXA> (accessed on 29 May 2017).
- [139] Karpathy A, Li F. Deep visual-semantic alignments for generating image descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2017;39(4):664-76.

BIBLIOGRAFIA

- [140] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Bengio Y. Generative adversarial nets. *Advances in Neural Information Processing Systems* 2014;2672-80.
- [141] Jaderberg M, Simonyan K, Zisserman A, Kavukcuoglu K. Spatial transformer networks. *Advances in Neural Information Processing Systems* 2015;2017-2025.
- [142] Graves A, Wayne G, Reynolds M, Harley T, Danihelka I, Grabska-Barwińska A, Hassabis D. Hybrid computing using a neural network with dynamic external memory. *Nature*, 538(7626):471-76.
- [143] Differentiable neural computers. Available online: <https://github.com/deepmind/dnc> (accessed on 01 June 2017).
- [144] Lezak MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment*. Oxford University Press 2012.
- [145] Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. *Brain Lang* 1997;57: 394-405.
- [146] CITA-Alzheimer Foundation, PGA project. Available online: <https://goo.gl/zh8Sjg> (accessed on 10 June 2016).
- [147] Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *Journal of Gerontology* 1982;37(3):323-29.
- [148] Goodglass H. *Evaluación de la Afasia y de Trastornos Relacionados*. Editorial Médica Paramericana 2005.
- [149] Benton AL, Hannay HJ, Varney N. Visual perception of line direction in patients with unilateral brain disease. *Neurology* 1975;25:907-10.
- [150] Wechsler D. *WAIS-III. Escala de inteligencia de Wechsler para adultos-III*. TEA Ediciones 2001.
- [151] Reitan RM, Wolfson D. *The Halstead - Reitan neuropsychological test battery. Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press 1985.
- [152] Kertesz A. *Batería de Afasias Western (BAW)*. Nau Llibres 2000.
- [153] *Diagnostic and Statistical Manual of Mental Disorders, 4th ed.*; American Psychiatric Association: Washington, DC, USA, 2000.
- [154] National Disability Authority. Available online: <http://universaldesign.ie/Technology-ICT/Universal-Design-for-ICT/> (accessed on 09 March 2016)
- [155] Gomez-Vilda P, Rodellar-Biarge V, Nieto-Lluis V, Munoz-Mulas C, Mazaira- Fernandez L, Martinez-Olalla R, Alvarez-Marquina A, Ramirez-Calvo C, Fernandez-Fernandez M. Characterizing neurological disease from voice Quality biomechanical analysis. *Cognitive Computation* 2013;5(4):399-425.
- [156] Barroso N. *Aportaciones a la gestión de información semántica de audio en entornos complejos*. UPV/EHU. Donostia.
- [157] Gómez-Vilda P, Rodellar-Biarge V, Nieto-Lluis V, López de Ipiña K, Álvarez-Marquina A, Martínez-Olalla R, Eca-Torres M, Martínez-Lage P. Phonation biomechanic analysis of Alzheimer's Disease cases. *Neurocomputing* 2015;167:83-93.
- [158] Little MA, Mcsharry PE, Roberts SJ, Costello DAE, Moroz IM. Exploiting non linear recurrence and fractal scaling properties for voice disorder detection. *Biomed Eng* 2007;6:23.
- [159] Song Y, Wang WH, Guo FJ. Feature extraction and classification for audio information in news video. *International Conference on Wavelet Analysis and Pattern Recognition* 2009;1:43-46.
- [160] Markaki M, Stylianou Y, Arias-Londono J, Godino-Llorente J. Dysphonia detection based on modulation spectral features and cepstral coefficients. *IEEE International Conference on Acoustics Speech and Signal Processing* 2010;1: 5162-65.
- [161] López-de-Ipiña K, Iturrate M, Alonso JB, Rodríguez-Herrera B. Automatic acoustic analysis for biodiversity preservation: a multi-environmental approach. *IWOBI* 2015;1: 43-48.

- [162] Mammone R, Zhang X, Ramachandran R. Robust speaker recognition: a feature-based approach. *IEEE Signal Process Mag* 1996;13(5):58-71.
- [163] Malyska N, Quatieri T, Sturim D. Automatic Disphonia recognition using biologically-inspired amplitude-modulation features. *Proceedings of IEEE International Conference on Acoustics Speech And Signal Processing* 2005;1:873-876.
- [164] López-de-Ipiña K, Satue-Villar A, Faundez-Zanuy M, Arreola V, Ortega O, Clavé P, Sanz-Cartagena MP, Mekyska J, Calvo P. Advances in a Multimodal Approach for Dysphagia Analysis Based on Automatic Voice Analysis. *Advances in Neural Networks* 2016; 201-211.
- [165] López de Ipiña K, Martínez de Lizarduy U, Calvo P, Beita B, García-Melero J, Ecay-Torres M, Estanga A, Faúndez-Zanuy M. Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach. *International Work Conference on Bioinspired Intelligence* 2017;1:1-4.
- [166] Mathworks. Available online: www.mathworks.com (accessed on 18 June 2015).
- [167] Praat: doing Phonetics by Computer. Available online: www.praat.org/ (accessed on 18 June 2015).
- [168] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Non linear Phenomena* 1988;31(2):277-83.
- [169] Castiglioni P. What is wrong in Katz's method? Comments on: "A note on fractal dimensions of biomedical waveforms". *Comput Biol Med* 2010;40:11-12.
- [170] Zanin M, Zunino L, Rosso OA, Papo D. Permutation entropy and its main biomedical and econophysics applications: a review. *Entropy* 2017;14(8):1553-77.
- [171] Morabito FC, Labate D, La Foresta F, Bramanti A, Morabit G, Palamara I. Multivariate Multi-Scale Permutation Entropy for Complexity Analysis of Alzheimer's Disease EEG. *Entropy* 2012;14(7):1186-202.
- [172] Costa M, Goldberger A, Peng CK. Multiscale entropy analysis of biological signals. *Phys Rev* 2005; 71:1-18.
- [173] Bandt C, Pompe B. Permutation entropy: A natural complexity measure for time series. *Phys Rev Lett* 2002;88:1-4.
- [174] Maragos P, Potamianos A. Fractal dimensions of speech sounds: Computation and application to automatic speech recognition. *J of the Acoustical Society of America* 1999;105(3):1925-32.
- [175] Abásolo D, Escudero J, Hornero R, Espino P, Gómez C. Fractal Dimension of the EEG in Alzheimer's Disease. *Encyclopedia of Healthcare Information Systems* 2008.
- [176] Katz M. Fractals and the analysis of waveforms. *Comput Biol Med* 1988;18(3):145-56.
- [177] Ezeiza A, López de Ipiña K, Hernández C, Barroso N. Enhancing the feature extraction process for automatic speech recognition with fractal dimensions. *Cognitive Computation* 2013;5(4):545-50.
- [178] López-de-Ipiña K, Eguiraun H, Solé-Casals J, Ecay M, Ezeiza A, Barroso N, Martínez-Lage P, Martínez-de-Lizardui U. 2013c. Feature Extraction Approach Based on Fractal Dimension for Spontaneous Speech Modelling Oriented to Alzheimer Disease Diagnosis. *NOLISP* 2013;7911:144-151.
- [179] Esteller R, Vachtsevanos G, Echaz J, Litt B. , 2012. A comparison of waveform fractal dimension algorithms. *IEEE Trans Circuits Syst I Fundam Theory Appl* 2001. 48:2:177-183.
- [180] Brissaud JB. The meaning of entropy. *Entropy* 2005;7:68-96.
- [181] Shannon CE. A mathematical theory of communication. *Bell Syst Tech J* 1948;27:379-423.
- [182] Eguiraun H, López-de-Ipiña K, Martínez I. Application of, Entropy and Fractal Dimension Analyses to the Pattern Recognition of Contaminated Fish Responses in Aquaculture. *Entropy* 2014;16:6133-51.
- [183] Shannon C, Weaver W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign. IL 1949.
- [184] Richman JS, Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 2000;278:2039-49.

BIBLIOGRAFIA

- [185] Pincus SM, Huang WM. Approximate entropy, statistical properties and applications. *Commun. Stat. Theory Methods* 1992;21:3061-77.
- [186] Pincus SM. Approximate entropy as a measure of system complexity. *Proc Natl Acad Sci* 1991;8(6):2297-301.
- [187] Yentes JM, Hunt N, Schmid K, Kaipust JP, McGrath D, Stergiou N. The use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng* 2013;41:349-65.
- [188] Ramanand P, Nampoori V.P, Sreenivasan R. Complexity quantification of dense array EEG using sample entropy analysis. *J Integr Neurosci* 2004;3:343-58.
- [189] Solé-Casals J, Vialatte FB. Towards Semi-Automatic Artifact Rejection for the Improvement of Alzheimer's Disease Screening from EEG Signals. *Sensors* 2015;15:17963-76.
- [190] Keller K, Lauffer H. Symbolic analysis of high-dimensional time series. *Int. J. Bifurc. Chaos* 2003;13:2657-68.
- [191] Ahmed MU, Mandic DP. Multivariate multiscale entropy. In *Proceedings of the IEEE Signal Processing Letters* 2012;19:91-95.
- [192] Alonso JB, Cabrera J, Travieso CM, López-de-Ipiña K, Sánchez-Medina A. Continuous tracking of the emotion temperature. *Neurocomputing* 2017.
- [193] Freeman DK, Boyd I. Voice Activity Detection. US5276765A 1994.
- [194] Solé-Casals J, Zaiats V. A non-linear VAD for noisy environments. *Cognitive Computation* 2010;2(3):191-198.
- [195] WEKA. Available online: <http://www.cs.waikato.ac.nz/ml/weka> (accessed on 13 August 2016).
- [196] Picard R, Cook D. Cross-Validation of Regression Models. *Journal of the American Statistical Association* 1984;79(387):575-83.
- [197] Daniel L K Yamins & James J DiCarlo. Using goal-driven deep learning models to understand sensory cortex. *Nature Neuroscience* 2016;356-65.
- [198] Bergstra J, Breuleux O, Lamblin P, Pascanu R, Delalleau O, Desjardins G, Goodfellow I, Bergeron A, Bengio Y, Kaelbling P. Theano: Deep Learning on GPUs with Python. *Journal of Machine Learning* 2001;1:1-48.
- [199] Topology of a multilayer perceptron neural network. Available online: <https://goo.gl/0IEq8v> (accessed on 13 August 2016).
- [200] Ways to Organize data in structure arrays. Available online: <https://goo.gl/dx2Lv3> (accessed on 15 August 2016).
- [201] Hinton G, Deng L, Yu D, Dahl G, Mohamed A, Jaitly N, Senior A, Vanhoucke V, Nguyen P, Sainath T, Kingsbury B. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine* 2012;29:82–97.
- [202] Sainath T, Mohamed AR, Kingsbury B, Ramabhadran B. Deep convolutional neural networks for LVCSR. *Proc Acoustics Speech and Signal Processing* 2013;8614–18.
- [203] Convolutional neural networks in iOS and macOS. Available online: <https://goo.gl/HuSzSv> (accessed on 28 August 2016).
- [204] Performing convolution operations. Available online: <https://goo.gl/KX2jGs> (accessed on 02 September 2016).
- [205] Convolutions. Available online: <https://goo.gl/XJN1st/> (accessed on 12 September 2017).
- [206] Masci J, Meier U, Cireşan D, Schmidhuber J. Stacked Convolutional Auto-Encoders for Hierarchical Feature Extraction. *Artificial Neural Networks and Machine Learning* 2011;52-59.
- [207] Geoffrey EH, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. 2012.

- [208] Deep Learning. Available online: <https://rubenlopezg.wordpress.com/2014/05/07/que-es-y-como-funciona-deep-learning/> (accessed on 12 September 2016).
- [209] Convolutional Neural Networks. Available on: <https://hackernoon.com/visualizing-parts-of-convolutional-neural-networks-using-keras-and-cats-5cc01b214e59> (accessed on 12 January 2017).
- [210] Case study of CNN. Available online: <https://www.slideshare.net/nmhkahn/case-study-of-convolutional-neural-network-61556303> (accessed on 10 February 2016).
- [211] Kaiming H, Xiangyu Z, Shaoqing R, Jian S. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2015;37(9):346-61.
- [212] Convolutional Neural Networks. Available online: <http://cs231n.github.io/convolutional-networks> (accessed on 14 September 2016).
- [213] Deep learning for complete beginners: convolutional neural networks with keras. Available online: <https://goo.gl/KARO03> (accessed on 23 September 2016).
- [214] Improving the way neural networks learn. Available online: <https://goo.gl/48E5dW> (accessed on 14 September 2016).
- [215] Overfitting. Available online: <https://en.wikipedia.org/wiki/Overfitting>, (accessed on 14 September 2016).
- [216] Keras. Available online: <https://keras.io/> (accessed on 14 September 2016).
- [217] Single-Layer Neural Networks and Gradient Descent. Available online: <https://goo.gl/p7vViF> (accessed on 14 September 2016).
- [218] Convolutional Neural Networks (LeNet). Available online: <https://goo.gl/a4Ph1> (accessed on 14 September 2016).
- [219] Le V, Besacier L. Automatic speech recognition for under-resourced languages: application to Vietnamese language. *IEEE Transactions on Audio, Speech, and Language Processing* 2009;17(8):1471-82.
- [220] Seng S, Sam S, Le V, Bigi B, Besacier L. Which Units For Acoustic and Language Modeling For Khmer Automatic Speech Recognition. *Conference on Spoken Language Processing for Under-resourced languages* 2008.
- [221] Kumar V, Steinbach M, Karypis G. A comparison of document clustering techniques. *Proceedings of KDD Workshop on Text Mining* 2000.
- [222] Vegesna A, Tran M, Angelaccio M, Arcona S. Remote Patient Monitoring via Non-Invasive Digital Technologies: A Systematic Review. 2017.
- [223] Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: An overview, with example from eCAALYX. *BioMedical Engineering* 2011.
- [224] Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Computation* 1998 10(7):1895-1923.
- [225] TCP/IP Tutorial and Technical Overview. Available online: <https://goo.gl/GTuu8C> (accessed on 02 June 2017).
- [226] Hancke G. A Practical Relay Attack on ISO 14443 Proximity Cards. 2005.
- [227] Madlmayr G, Langer J, Kantner C, Scharinger J. NFC devices: Security and privacy. *International Conference on Availability, Reliability and Security* 2008.
- [228] Felica. Available online: <https://www.sony.net/Products/felica/> (accessed on 02 June 2017).
- [229] Tang Y. Mobile wireless communications device providing security features based upon wearable near field communication "NFC" device. EP2720444 A1 2014.
- [230] Howitt I, Gutierrez JA. IEEE 802.15.4 low rate - wireless personal area network coexistence issues. Paper presented at the *IEEE Wireless Communications and Networking Conference* 2003;3:1481-86.

BIBLIOGRAFIA

- [231] Fujisaki E, Okamoto T. Secure integration of asymmetric and symmetric encryption schemes. *Journal of Cryptology* 2013;26(1):80-101.
- [232] Grynberg A. Methods for Extending a Security Token Based Identity System. US20080196096 A1 2008.
- [233] Nanopoulos A, Ackerman K, Bowness P, Duane W, Jakobsson M, Kaliski B, Pal D, Rice S, Rivest R. System and method providing disconnected authentication. US20050166263 A1 2005.
- [234] Lee J, Moon K, Sohn S, Park C. Method for electronic commerce using security token and apparatus thereof. US20050097060 A1 2005.
- [235] The HTTPS-Only Standard. Available online: <https://https.cio.gov/faq/> (accessed on 23 September 2016).
- [236] Callegati F, Cerroni W, Ramilli M. Man-in-the-middle attack to the HTTPS protocol. *IEEE Security and Privacy* 2009;7(1):78-81.
- [237] The Solitaire Encryption Algorithm. Available online: <https://goo.gl/jQ5HWK> (accessed on 02 June 2017).
- [238] Stephenson N. *Criptonomicon*. Avon 2007.
- [239] The Solitaire encryption algorithm programmed in C. Available online: <https://goo.gl/VQKJoR> (accessed on 23 May 2017).
- [240] El algoritmo de cifrado "Solitario". Available online: <https://goo.gl/6rcHVd> (accessed on 23 May 2017).
- [241] Desmedt Y. Man in the Middle Attack. *Encyclopedia of Cryptography and Security* 2011;368-368
- [242] Schrödinger E. The present situation in quantum mechanics. *Naturwissenschaften* 1935;23:823-28.
- [243] TRIMMER JD. The present situation in quantum-mechanics - a translation of Schrödingers cat paradox paper. *Proceedings of the American Philosophical Society* 1980;124(5):323-338.
- [244] Algorithm validation lists. Available online: <http://csrc.nist.gov/groups/STM/cavp/validation.html> (accessed on 25 November 2016)
- [245] Keyed-Hash Message Authentication Code (HMAC). Available online: <https://goo.gl/ZUswt3> (accessed on 26 November 2016).
- [246] Lee H. K., Malkin T., Nahum E. Cryptographic strength of SSL/TLS servers: Current and recent practices. *Proceedings of the ACM SIGCOMM Internet Measurement Conference* 2007;83-92.
- [247] Zhang ZJ, Man ZX, Li Y. Improving wojcik's eavesdropping attack on the ping-pong protocol. *Physics Letters A* 2004;333(1-2):46-50.
- [248] Zimmermann H. OSI Reference Model--The ISO Model of Architecture for Open Systems Interconnection. *IEEE Transactions on Communications* 1980;28(4).
- [249] C# and .NET interview question: - What are symmetric and asymmetric algorithms? Available online: <https://goo.gl/SxXh1U> (accessed on 26 November 2016).
- [250] Alonso G, Casati F, Kuno H, Machiraju V. *Web Services. Data-Centric Systems and Applications* pp 123-149.
- [251] OASIS. Available online: <https://www.oasis-open.org/standards> (accessed on 29 November 2016).
- [252] W3C. Available online: <https://goo.gl/owUIGa> (accessed on 19 December 2016).
- [253] WS-I. Available online: <http://www.ws-i.org/> (accessed on 19 September 2016).
- [254] JSON. Available online: <http://www.json.org/json-es.html> (accessed on 20 December 2016).
- [255] Apache. Available online: <https://httpd.apache.org/> (accessed on 19 December 2016).
- [256] MySQL. Available online: <https://www.mysql.com/> (accessed on 19 December 2016).
- [257] PHP. Available online: <http://www.php.net/> (accessed on 19 December 2016).

-
- [258] Kromann FM. Using MySQL Databases. PHP and MySQL Recipes, Apress 2016;333-56.
- [259] Dayal M, Singh N, Raw RS. A comprehensive inspection of cross site scripting attack. International Conference on Computing, Communication and Automation 2016;497-502.
- [260] Singh N, Dayal M, Raw RS, Kumar S. SQL injection: Types, methodology, attack queries and prevention. International Conference on Computing for Sustainable Global Development 2016;2872-76.
- [261] Voitovych OP, Yuvkovetskyi OS, Kupershtein L. M. SQL injection prevention system. International Radio Electronics & Info Communications 2016.
- [262] Dalcín L, Storti M, Paz R. Desarrollo de Aplicaciones Paralelas en Python. Mecánica Computacional 2004;23:3153-63.
- [263] Jupyter. Available online: <http://jupyter.org/> (accessed on 03 January 2017).
- [264] Howells J. Software as a Service (SaaS). Management Information Systems 2015;7.
- [265] Pfeiffer S. HTML5 Audio API. The Definitive Guide to HTML5 Video. Apress 2011;223-45.
- [266] Bruno EJ. Ajax: Asynchronous JavaScript and XML. Dobb's Journal 2006;31(2)32-35.
- [267] Arias-Londoño JD, Godino-Llorente JI, Sáenz-Lechón N, Osma-Ruiz V, Castellanos-Domínguez G. An improved method for voice pathology detection by means of a HMM-based feature space transformation. Pattern Recognition 2010;43(9):3100-12.
- [268] Godino JI, Gómez-Vilda P, Sáenz-Lechón N, Blanco-Velasco M, Cruz-Roldán F, Ferrer-Ballester MA. Support Vector Machines Applied to the Detection of Voice Disorders. Springer Verlag Berlin Heidelberg 2005;219-30.
- [269] Godino JI, Gómez-Vilda P, Blanco-Velasco M. Dimensionality Reduction of a Pathological Voice Quality Assessment System Based on Gaussian Mixture Models and Short Term Cepstral Parameters. IEEE Transactions on biomedical engineering 2006;53(10):1943 -53.
- [270] Elekin. Available online: <http://elekin.net/inicio/> (accessed on 25 May 2016).
- [271] Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. Brain Lang. 57: 394-405 (1997).
- [272] CITA-Alzheimer Foundation, PGA project: <http://www.cita-alzheimer.org/investigacion/proyectos>, (accessed on 10 June 2016).

ERANSKINAK

A ERANSKINA: SISTEMA ERREALAREN ANALISIAK HIRU ATAZETARAKO (SS, PD, AN)

B ERANSKINA: JCR ARGITARAPENAK

1. On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis.
2. Selection of Entropy Based Features for Automatic Analysis of Essential Tremor.
3. Advances on automatic speech analysis for early detection of Alzheimer Disease: a non-linear multi-task approach (accepted).

C ERANSKINA: ARGITARAPEN INDEZATUAK

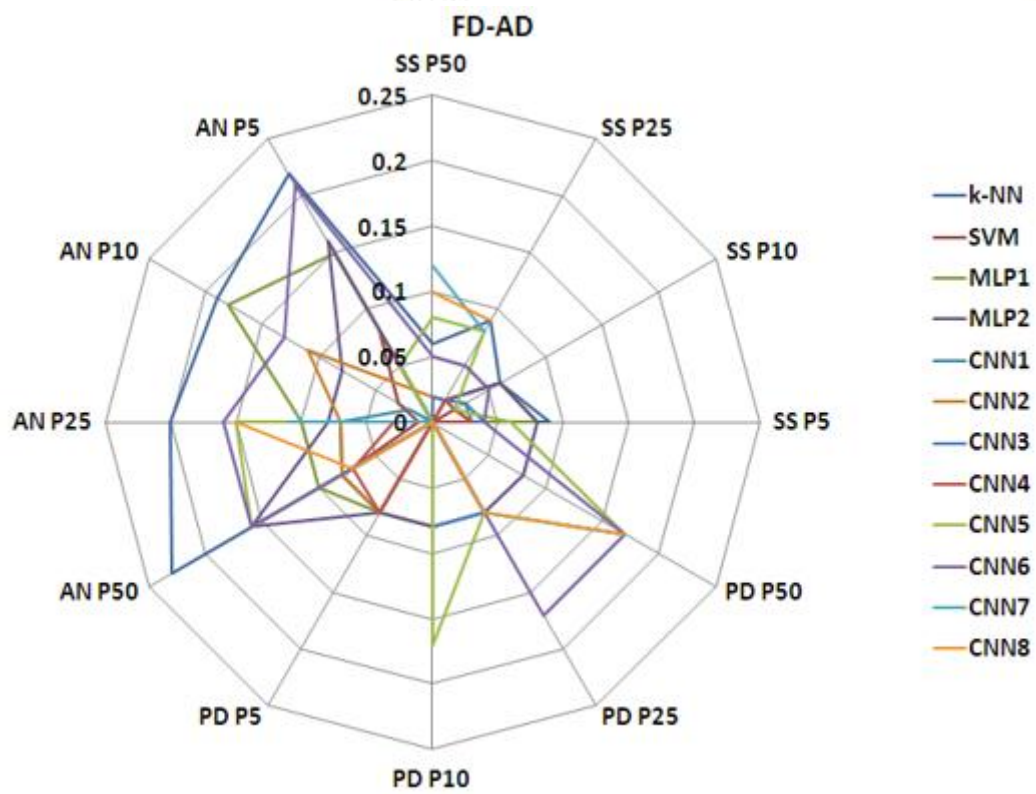
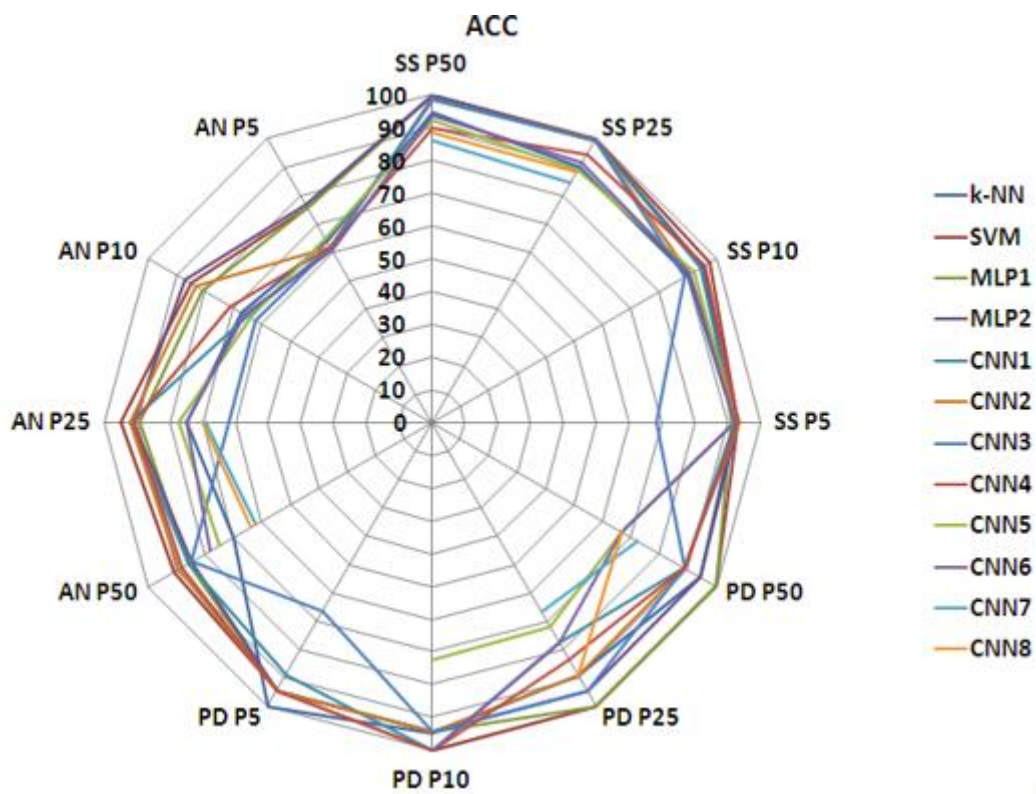
1. Feature Extraction Approach Based on Fractal Dimension for Spontaneous Speech Modelling Oriented to Alzheimer Disease Diagnosis.
2. Automatic analysis of Categorical Verbal Fluency for Mild Cognitive Impairment detection: a non-linear language independent approach.
3. Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach (accepted).
4. Non-invasive Biosignal Analysis Oriented to Early Diagnosis and Monitoring of Cognitive Impairments.

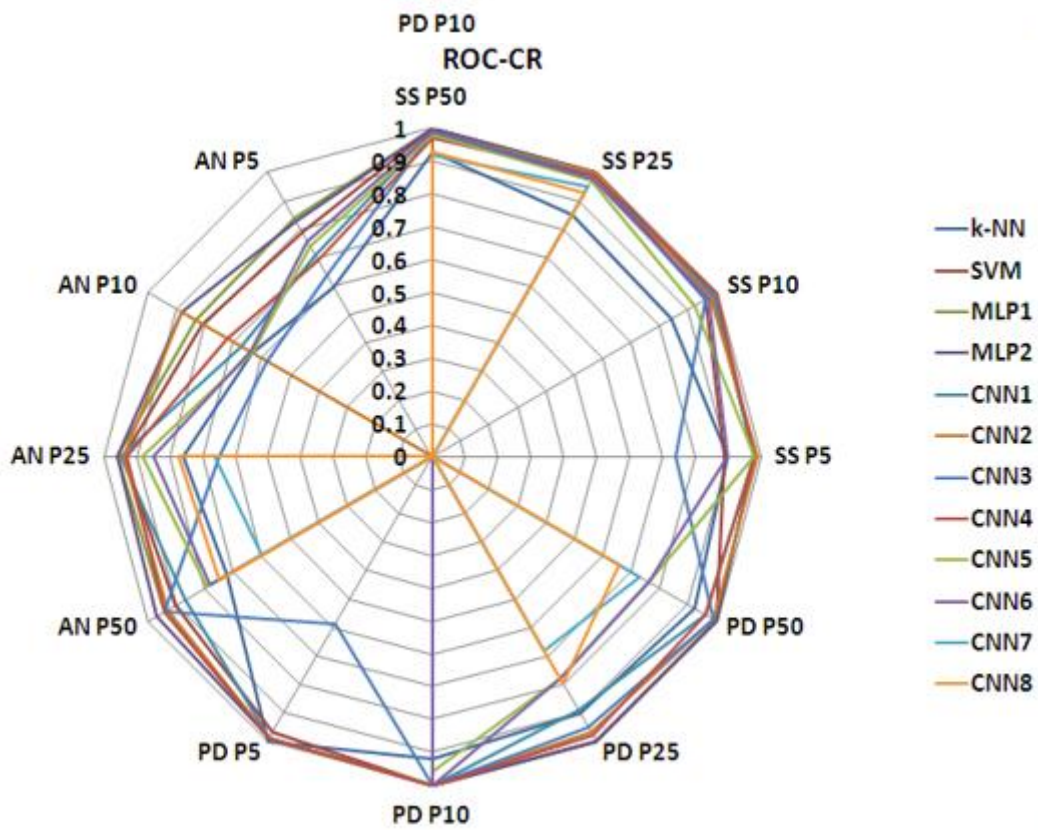
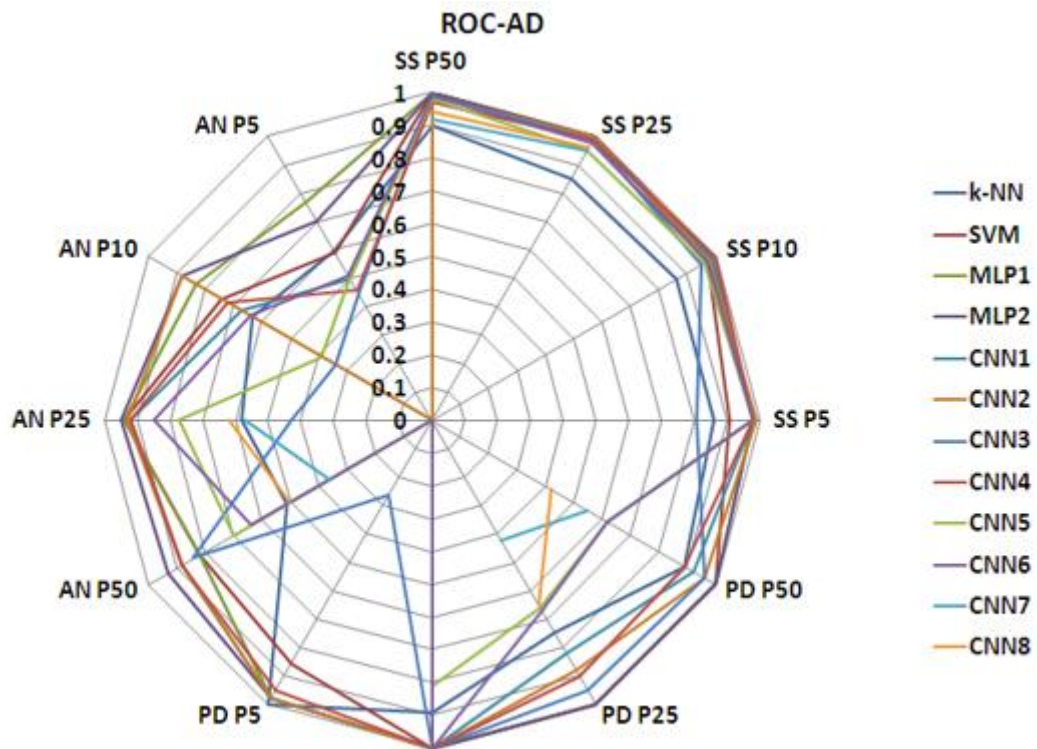
D Eranskina: Bestelako argitarapenak

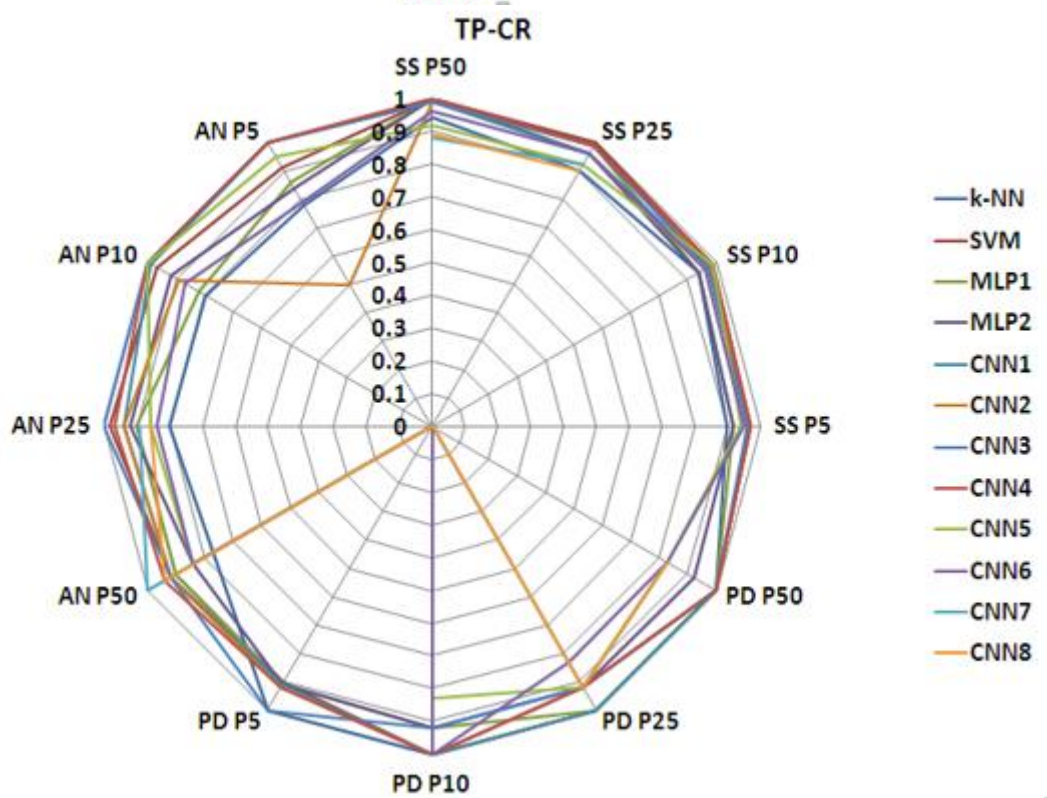
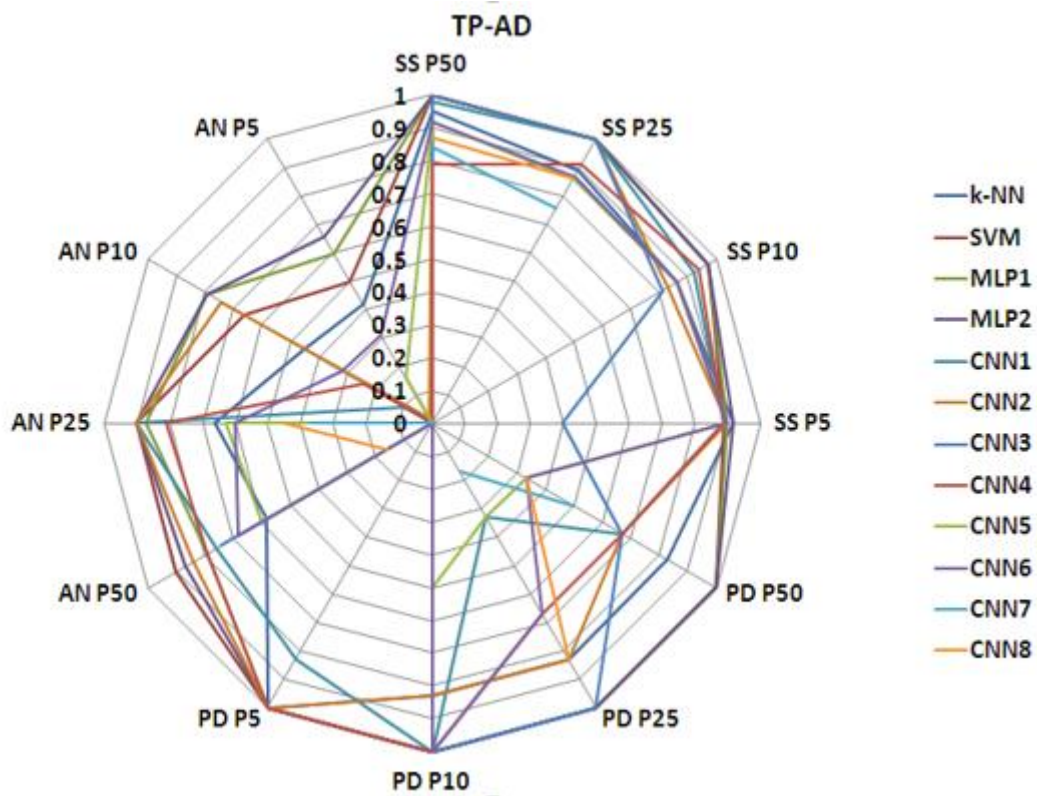
1. Analysis of Spontaneous Speech and Emotional Response oriented to Alzheimer's Disease Diagnosis.
2. ALZUMERIC: a decision support System for diagnosis and monitoring of cognitive impairment (accepted).
3. Feature extraction approach based on Fractal Dimension for Spontaneous Speech modelling oriented to Alzheimer Disease diagnosis. NOLISP, 2013, Mons Belgica (C1 eranskina).

A ERANSKINA

SISTEMA ERREALAREN ANALISIAK HIRU
ATAZETARAKO (SS, PD, AN)







B1 ERANSKINA

On the Selection of Non-Invasive Methods
Based on Speech Analysis Oriented to
Automatic Alzheimer Disease Diagnosis

Article

On the Selection of Non-Invasive Methods Based on Speech Analysis Oriented to Automatic Alzheimer Disease Diagnosis

KarmeLe López-de-Ipiña ^{1,*}, Jesus-Bernardino Alonso ², Carlos Manuel Travieso ², Jordi Solé-Casals ³, Harkaitz Egiraun ^{1,4}, Marcos Faundez-Zanuy ⁵, Aitzol Ezeiza ¹, Nora Barroso ¹, Miriam Ecay-Torres ⁶, Pablo Martinez-Lage ⁶ and Unai Martinez de Lizardui ¹

¹ Systems Engineering and Automation Department, University of the Basque Country UPV/EHU, Donostia 20018, Spain; E-Mails: harkaitz.egiraun@ehu.es (H.E.); aitzol.ezeiza@ehu.es (A.E.); nora.barroso@ehu.es (N.B.); unai.martinezdelizardui@ehu.es (U.M.L.)

² Signal and Communication Department (DSC), Institute for Technological Development and Innovation in Communications (IDeTIC), University of Las Palmas de Gran Canaria (ULPGC), Campus of Tafira, Las Palmas de Gran Canaria 35017, Spain; E-Mails: jalonso@dsc.ulpgc.es (J.-B.A.); ctravieso@dsc.ulpgc.es (C.M.T.)

³ Digital Technologies Group, University of Vic, Sagrada familia 7, Vic 08500, Spain; E-Mail: jordi.sole@uvic.cat

⁴ Research Center for Experimental Marine Biology and Biotechnology, Plentzia Marine Station, University of the Basque Country, Plentzia 48620, Spain

⁵ Escola Universitaria Politècnica de Mataró (UPC), Tecnocampus, Mataró, Barcelona 08302, Spain; E-Mail: faundez@tecnocampus.cat

⁶ CITA-Alzheimer Foundation, San Sebastian 20009, Spain; E-Mails: meca@cita-alzheimer.org (M.E.-T.); pmlage@cita-alzheimer.org (P.M.-L.)

* Author to whom correspondence should be addressed; E-Mail: karmeLe.ipina@ehu.es; Tel.: +34-943-018-667; Fax: +34-943-017-230.

Received: 15 April 2013; in revised form: 8 May 2013 / Accepted: 13 May 2013 /

Published: 21 May 2013

Abstract: The work presented here is part of a larger study to identify novel technologies and biomarkers for early Alzheimer disease (AD) detection and it focuses on evaluating the suitability of a new approach for early AD diagnosis by non-invasive methods. The purpose is to examine in a pilot study the potential of applying intelligent algorithms to speech features obtained from suspected patients in order to contribute to the improvement of diagnosis of AD and its degree of severity. In this sense, Artificial Neural Networks (ANN) have been used for the automatic classification of the two classes (AD and control

subjects). Two human issues have been analyzed for feature selection: Spontaneous Speech and Emotional Response. Not only linear features but also non-linear ones, such as Fractal Dimension, have been explored. The approach is non invasive, low cost and without any side effects. Obtained experimental results were very satisfactory and promising for early diagnosis and classification of AD patients.

Keywords: Alzheimer's disease diagnosis; spontaneous speech; emotion recognition; machine learning; non-invasive diagnostic techniques; dementia

1. Introduction

Alzheimer's disease (AD) is the most common type of dementia among elderly people in Western countries and it has a large socioeconomic cost to society which is expected to increase in the near future. It is characterized by progressive and irreversible cognitive deterioration with memory loss, impaired judgment and language and other cognitive deficits and behavioural symptoms that end up becoming severe enough to limit the ability of an individual to perform professional, social or family activities of daily life. As the disease progresses patients develop increasingly severe disabilities to finally become completely dependent. An early and accurate diagnosis of AD would be of much help for patients and their families, both to plan for the future and to start an early treatment of the symptoms of the disease.

According to current criteria, the diagnosis is expressed with different degrees of certainty as possible or probable AD when dementia is present and other possible causes have been ruled out, but an unambiguous diagnosis of AD requires the demonstration of the typical AD pathological changes in brain tissue by autopsy (*post-mortem* analysis) [1–3]. The clinical hallmark and earliest manifestation of AD is episodic memory impairment. At the time of clinical presentation other cognitive deficits are usually already present in their language, executive functions, orientation, perceptual abilities and constructional skills. Associated behavioural and psychological symptoms include apathy, irritability, depression, anxiety, delusions, hallucinations, disinhibition, aggression, aberrant motor behaviour, as well as eating or sleep behaviour changes [1–5]. All these symptoms lead to impaired performance in family, social or professional activities of daily life as the disease progresses from mild to moderate and to severe.

As already mentioned above, the diagnosis of AD is made on clinical grounds and requires, on one hand, the confirmation of a progressive dementia syndrome and, on the other, the exclusion of other potential causes of dementia by clinical history and examination, complete blood workup tests and brain-imaging analysis test, such as computer tomography (CT) or magnetic resonance imaging (MRI). The criteria to exclude other potential causes have changed in the last years as the interpretation of neuroimaging tests, including functional imaging with Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET), has focused on the positive findings of typical AD changes, such as medial temporal atrophy detected by CT or MRI and temporoparietal hypometabolism by PET [6–9].

Nonetheless, the diagnosis of the early stages of not only mild cognitive impairment but also mild dementia remains problematic, since patients and relatives tend to either ignore the first clinical manifestations or ascribe them to the expected cognitive changes related to age. It usually takes 2 to 3 years to seek medical advice after the onset of the symptoms [1–3]. In addition, physicians may feel uncertain or uncomfortable to establish a diagnosis when the whole picture of dementia is not yet fully present; therefore, they usually feel the need to apply long neuropsychological tests, expensive neuroimaging techniques or invasive tests such as a lumbar puncture to reach a diagnosis [10]. It is consequently not surprising that most of the patients are diagnosed when they have already reached the moderate stage of the disease and have become substantially dependent. At this stage, it is very difficult for any treatment strategy to show significant efficacy to stop or even delay the disease process [2–5].

Significant advances have taken place during the last years in the early diagnosis of AD using clinical biomarkers [10], but the currently high cost and technology requirements make it unfeasible to use these diagnostic procedures on any patient displaying only memory complaints. As a result, they are usually applied to pre-selected patients based on their being highly suspect of suffering an underlying AD pathology; who are then apt to have an invasive lumbar puncture or a very expensive PET performed [10–13].

In this setting, the development of non-invasive intelligent diagnosis techniques would be very valuable for the early detection and classification of different types of dementia. Particularly, because they do not require specialized personnel or laboratory equipment, so that anyone in the habitual environment of the patient could perform (after proper training) without altering or blocking the patient's abilities [14–19]. Automatic Spontaneous Speech Analysis (ASSA) and Emotional Response Analysis (ERA) on speech are two of them [15].

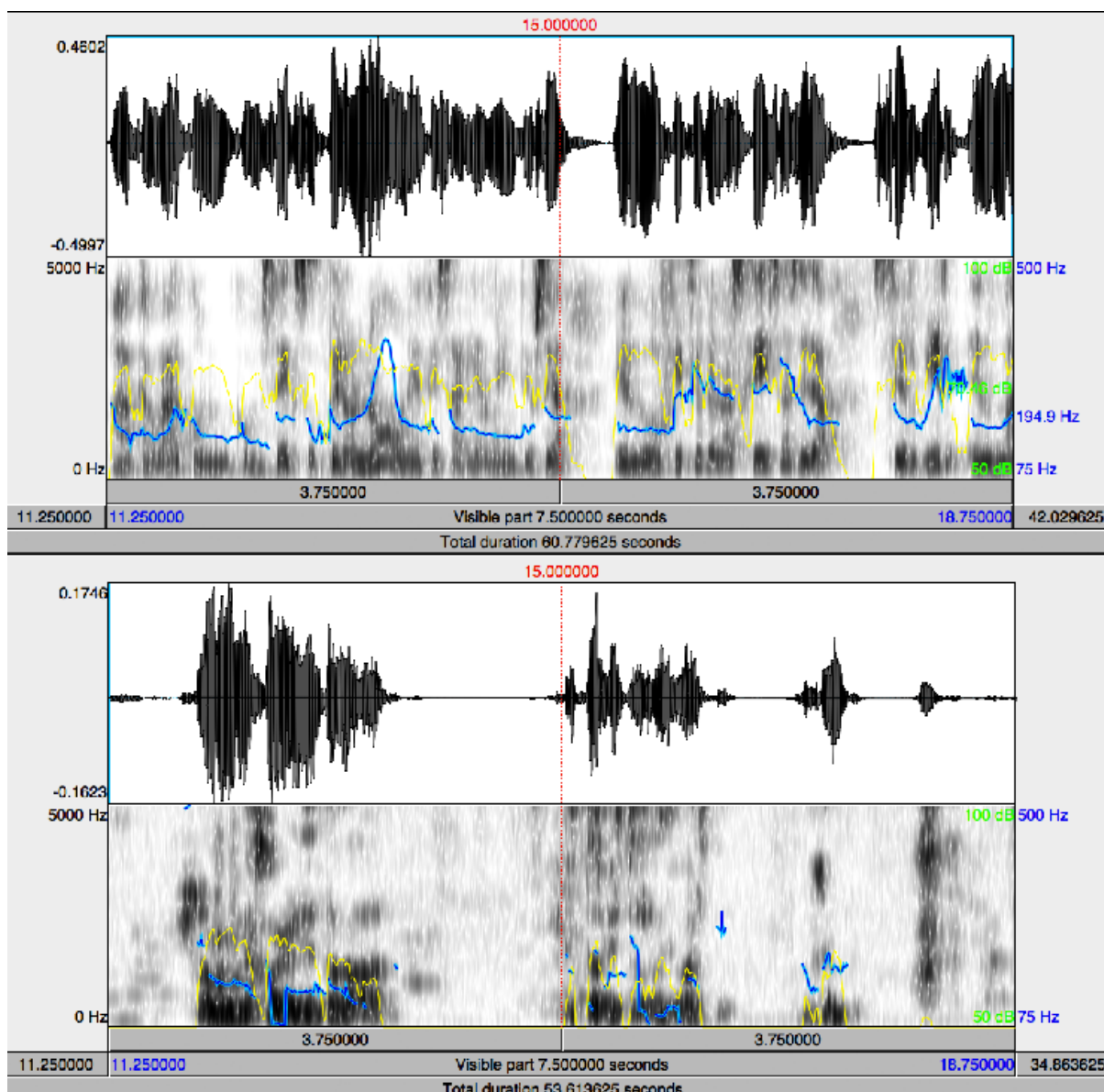
Spoken language is one of the most important elements defining an individual's intellect, his/her social life and personality; it allows us to communicate with each other, share knowledge, and express our cultural and personal identity. Spoken language is the most spontaneous, natural, intuitive and efficient method of communication among people. Therefore, the analysis by automated methods of Spontaneous Speech (SS) or Automatic Speech Analysis (ASSA) which is the freer and more natural expression of communication, possibly combined with other methodologies, has the potential to become a useful non-invasive method for early AD diagnosis [15,20–24].

Emotional Response Analysis (ERA) on speech also has that potential: emotions are cognitive processes related to the architecture of the human mind, such as decision making, memory or attention, closely linked to learning and understanding that arise in intelligent systems when they become necessary to survive in a changing and partially unpredictable world [25–27]. The nonverbal information, which often includes body-language, attitudes, modulations of voice, facial expressions, etc., is essential in human communication as it has a large effect on the communication provision of the partners and on the intelligibility of speech [25]. Human emotions are affected by the environment, the direct interaction with the outside world but also by the emotional memory emerging from the experience of individual and cultural environment, the so called socialized emotion. Emotions use the same components subjective, cultural, physiological and behavioural that the individual's perception expresses with regard to the mental state, the body and how it interacts with the environment [26,27]. In this work ERA has been analyzed by classical features and by Emotional Temperature, described in

Section 3. This feature is based on the analysis of several prosodic and paralinguistic features sets obtained from a temporal segmentation of the speech signal.

Finally, we wished to apply non-invasive methods to estimate the severity of Alzheimer in the patient. In this sense, analysis of spontaneous speech is not perceived as a stressful test and moreover its cost is lower than for other methods. None of these speech analysis based techniques require extensive infrastructure or the availability of medical equipment, and suggest obtaining by these means is easy, quick and inexpensive [14,15].

Figure 1. Signal and spectrogram of a control subject (**Top**) and an AD subject (**Bottom**) during spontaneous speech (pitch in blue, intensity in yellow).



We have focused our work on non-invasive diagnostic techniques based on the analysis of speech and emotions because after the loss of memory, one of the major problems of AD is the loss of language skills, illustrated by the poorer signal and spectrogram during spontaneous speech of the AD patient shown in Figure 1. This loss is reflected in difficulties both to speak and to understand others,

which makes more difficult the natural communication process with the environment. The inability to communicate appears already in the early phases of the diseases. It is possible to find different communication deficits in the area of language, including [28,29] aphasia (difficulty in speaking and understanding) and anomic aphasia (difficulty for recognizing and naming things). The specific communication problems the patient encounters depend on the stage of the disease [2–5,28,29]:

1. *First Stage or early stage (ES)*: difficulty in finding the right word in spontaneous speech. Often remains undetected.
2. *Second Stage or intermediate stage (IS)*: impoverishment of language and vocabulary in everyday use.
3. *Third Stage or advanced stage (AS)*: answers sometimes very limited and restricted to very few words.

Not only the language but also the emotional responses in Alzheimer's patients become impaired and seem to go through different stages. In the early stages, social and even sexual disinhibition appears and behavioural changes are also observed (for example, being angry and not being able to perform common tasks, express themselves or remember) [30–33]. However, the emotional memory remains, and they cry more easily and gratefully acknowledge caresses, smiles and hugs. The Alzheimer's patient reacts aggressively to things that for healthy people are harmless, and perceives a threat or danger where none exists. In more advanced stages they may often seem shy and apathetic, symptoms often attributed to memory loss and/or difficulty in finding the right words and some responses are likely to be magnified due to an alteration in perception.

Alternatively, it has been suggested that the reduced ability to feel emotions is due to memory loss, which may in turn induce the appearance of apathy and depression [31,33]. The work presented here is part of a larger study to identify novel technologies and biomarkers for early AD detection, and it focuses on evaluating the suitability of a new approach for early AD diagnosis based on non-invasive and low cost methods, namely Automatic Spontaneous Speech Analysis and the Emotional Response Analysis, whose results are susceptible to be used for the automatic classification of tested individuals.

2. Materials

2.1. Main Database of Individuals

Trying to develop a new methodology applicable to a wide range of individuals of different sex, age, language and cultural and social background, we have built up a multicultural and multilingual (English, French, Spanish, Catalan, Basque, Chinese, Arabian and Portuguese) database with the video recordings of 50 healthy and 20 AD patients (with a previous diagnosis) recorded for 12 hours and 8 hours respectively. The age span of the individuals in the database was 20–98 years and there were 20 males and 20 females. This database is called AZTIAHO.

All the work was performed strictly following the ethical consideration of the organizations involved in the project. The recordings consisted of videos of Spontaneous Speech where people tell pleasant stories or feelings and interact with each other in a friendly conversation. The recording atmosphere is relaxed and non-invasive. The shorter recording times for the AD group are due to the fact that AD patients find speech more of an effort than healthy individuals: they speak more slowly,

with longer pauses, and with more time spent on efforts to find the correct word and uttering speech disfluencies or break messages. In the advanced stage of the disease, they find this effort tiring and often want to stop the recording. We complied with their requests.

2.2. Pre-Processing

Video has been processed and audio extracted in wav format (16 bits and 16 kHz). Firstly non-analyzable events have been removed: laughter, cough, short hard noises and speaker mixes. Then background noise has been removed by a denoiser adaptive filtering. After the pre-processing, about 80% and 50% of the material from the control and AD groups respectively, is suitable for further analysis. The full database consisted of about 60 minutes for the AD group and of about 9 hours for the control. The speech is divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers. Finally, a database of about 600 segments of spontaneous speech is obtained.

2.3. Individuals Selected for the Study

From the original database, a subset of 20 AD patients (68–96 years of age, 12 women, 8 men, within the three stages of AD: First Stage (ES = 4), Secondary Stage (IS = 10) and Tertiary stage (AS = 6) was chosen. The control group (CR) was made up of 20 individuals (10 male and 10 female, aged 20–98 years) representing a wide range of speech responses. This subset of the database is called AZTIAHORE.

3. Methods

3.1. Feature Extraction

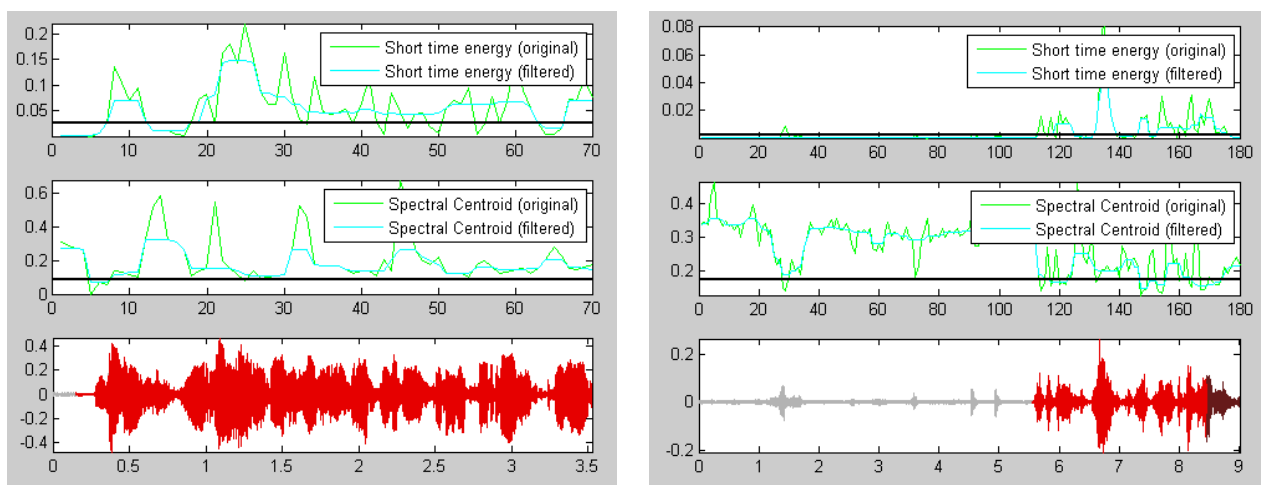
3.1.1. Automatic Spontaneous Speech Analysis (ASSA)

The analysis of spontaneous speech fluency is based on three families of features (SSF set), obtained using Praat software [34]. For that purpose, an automatic Voice Activity Detector (VAD) [35,36] has extracted voiced/unvoiced segments, as parts of an acoustic signal. These three families of features include:

1. *Duration*: this includes the histogram calculated over the most relevant voiced and unvoiced segments, the average of the most relevant voiced/unvoiced, voiced/unvoiced percentage and spontaneous speech evolution along the time, and the voiced and unvoiced segments' mean, max and min.
2. *Time domain*: short time energy.
3. *Frequency domain, quality*: spectral centroid.

The energy of a signal is typically calculated on a short-time basis, by windowing the signal at a particular time, squaring the samples and taking the average. The spectral centroid is commonly associated with the measure of the brightness of a sound. This measure is obtained by evaluating the “centre of gravity” using the Fourier transform’s frequency and magnitude information (Figure 2).

Figure 2. Plots of speech signal, Short Time Energy and Spectral Centroid, all measures filtered by a median filter, for a control subject (**Left**) and an AD subject (**Right**).



3.1.2. Higuchi Fractal Dimension

When appropriate corpora are available, linear systems can be implemented fairly rapidly, as they rely on well-known machine learning techniques to achieve their goals, avoiding complex adjustments to the system. These latter types of tasks often require experimentation with alternative techniques, which can lead to improved systems. One such alternative technique of particular interest is nonlinear analysis, and some works show that combining nonlinear features with linear ones can produce higher recognition accuracies without substituting the whole linear system with novel nonlinear approaches (see [37,38] for examples on nonlinear speech processing). This is especially promising for solving non-typical tasks, since it would be very demanding to design a complete nonlinear system from scratch for solving a task already made difficult by the scarcity of resources.

The fractal dimension is one of the most popular features, which describe the complexity of a system. Most if not all of the fractal systems have a characteristic called self-similarity. An object is self-similar if a close-up examination of the object reveals that it is composed of smaller versions of itself. Self-similarity can be quantified as a relative measure of the number of basic building blocks that form a pattern, and this measure is defined as the fractal dimension. This current work focus on the alternatives, which do not need previous modelling of the system. Higuchi proposed an algorithm for measuring the fractal dimension of discrete time sequences directly from time series, so in our experiments we use the method described in [39] (see Figure 3).

3.1.3. Emotional Speech Analysis (ESA)

In this study we aim to accomplish the automatic selection of emotional speech by analyzing three families of features in speech:

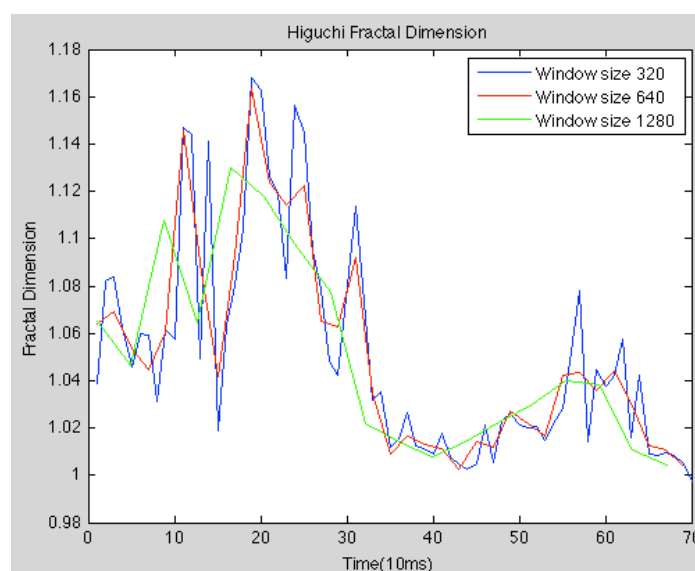
1. *Acoustic features*: pitch, standard deviation of pitch, max and min pitch, intensity, standard deviation of intensity, max and min intensity, period mean, period standard deviation, and Root Mean Square amplitude (RMS);

2. *Voice quality features*: shimmer, local jitter, Noise-to-Harmonics Ratio (NHR), Harmonics-to-Noise Ratio (HNR) and autocorrelation;
3. *Duration features*: fraction of locally unvoiced frames, degree of voice breaks.

Short-term energy is the principal and most natural feature that has been used. Physically, energy is a measure of how much signal exists at any one time. Energy is used in a continuous speech to discover voiced sounds, which have higher energy than silence/un-voiced, as shown in Figure 2.

The energy of a signal is typically calculated on a short-time basis, by windowing the signal at a particular time, squaring the samples and taking the average [36]. The square root of this result is the engineering quantity, known as the root-mean square (RMS) value.

Figure 3. Higuchi Fractal Dimension c of speech signal for an AD subject and different window length.



3.1.4. Emotional Temperature

The Emotional Temperature (ET) is based on the analysis of a few prosodic and paralinguistic features sets obtained from a temporal segmentation of the speech signal [40–42].

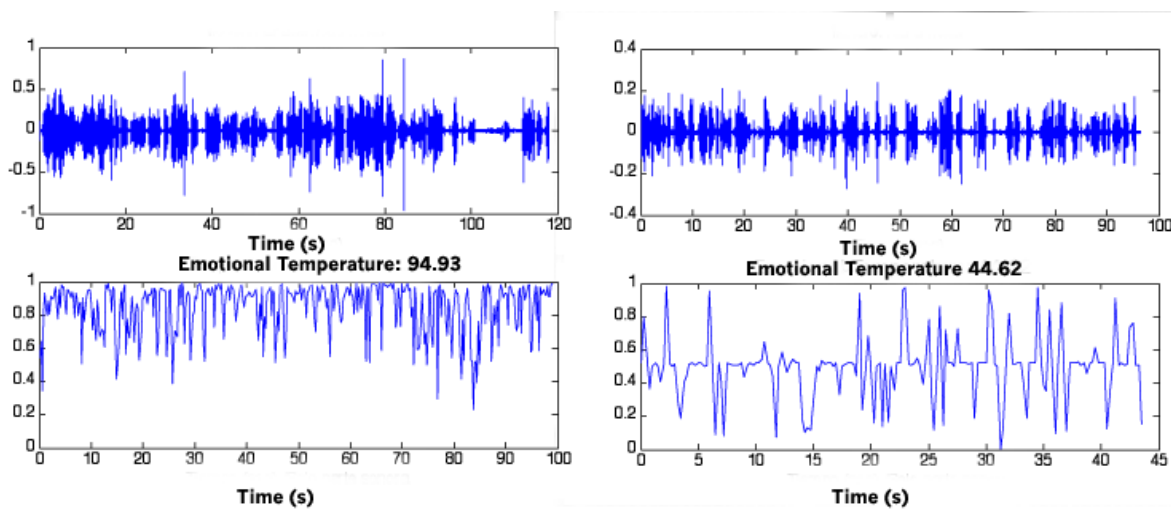
Two prosodic and four paralinguistic features related to the pitch and energy, respectively, were estimated from each frame. These features were chosen because their robustness in emotion recognition has been proven [43–47], they are quickly and easily calculated, and they are independent of linguistic segmentation, which helps us to avoid problems in real time applications on real environments. For prosodic features, a voiced/unvoiced decision is made to each frame and two linear regression coefficients of the pitch contour [43–46] are obtained. For paralinguistic features, voice spectral energy balances [43,47] are calculated from each frame, and quantified using 4 percentages of energy concentration in 4 frequency bands. Then Emotional Temperature is calculated as follows:

1. A Support Vector Machine (SVM) is trained with a balanced segment set extracted from the database (SVMs have been used to quantify the discriminative ability of the proposed measures [45]. We have used a freely available implementation called LIBSVM [48] with a radial basis kernel function).

2. For each speech segment, each temporal frame is classified by the SVM as “pathological or “non-pathological”.
3. The percentage of temporal frames classified as “non-pathological” is calculated. This value, *i.e.*, the number of non-pathological frames is the “emotional temperature”.
4. The “Emotional Temperature” is finally normalized in order to have $ET = 50$ as threshold obtained from the training database, which indicates the limit between pathological and non-pathological frames. This normalization will help medical specialists to easily interpret the data.

Figure 4 shows an example of ET values for a healthy subject ($ET = 94.93$) and for an AD subject ($ET = 44.62$).

Figure 4. Emotional Temperature for a healthy subject (**Left**) and an AD subject (**Right**).



3.1.5. Feature Sets

Based in these presented characteristics, four feature sets have been created for experimentation:

1. SSF.
2. EF, set described in Section 3.1.3.
3. FD1: Higuchi Fractal Dimension (HFD).
4. FD2: HFD, maximum HFD, minimum HFD, variance HFD and standard deviation HFD.

3.2. Automatic Classification

The automatic classification of emotional speech is based on the Multi Layer Perceptron (MLP) neural network with one hidden layer of 100 neurons and 1,000 training steps. WEKA [49] software has been used in carrying out the experiments. The results are evaluated using Accuracy (Acc) and Classification Error Rate (CER) measurements. For the training and validation steps, we used k-fold cross-validation in order to ensure solid results. Cross validation is a robust validation for variable selection [50]. These features will define CR group and the three AD levels. The original sample set was randomly divided into k subsets. Then, a single subset was retained as the validation data set for testing the model, and the remaining k-1 subsets were used as training data. The cross-validation

process was repeated k times, with each of the k subsets used exactly once as the validation data set. The k obtained values from the folds, were then averaged to obtain the final result. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once. In our experiments we use $k = 10$. These features will discriminate among control group (CR) and the three AD levels.

4. Results and Discussion

The experimentation has been carried out with the balanced subset AZTIAHORE. The goal of these experiments was to examine the potential of selected features for automatic measurement of the degradation of Spontaneous Speech, Emotional Response and their integration in people with AD. Thus, previously defined feature sets have been evaluated in order to properly define control and AD level groups.

In a first stage Emotional Temperature is calculated for each segment by the method described in Section 3.1.4. Automatic classification by MLP was performed over the speech features sets described in Section 3.1.5 in order to analyze for the pilot study the tests: Automatic Spontaneous Speech fluency, Emotional Response in speech and both in Integral Speech. Table 1 summarized Accuracy (%) global results with regard to pre-clinical test and feature sets.

Table 1. Accuracy (%), global results with regard to test and feature sets.

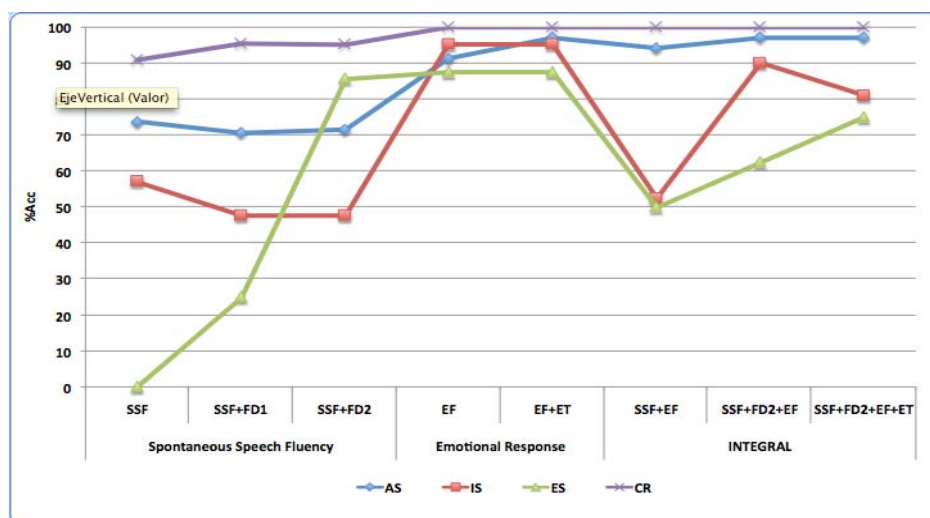
Test	Feature Set	%Acc
Spontaneous Speech Fluency	SSF	75.2
	SSF + FD1	76.7
	SSF + FD2	86.1
Emotional Response	EF	90.7
	EF + TE	97.7
INTEGRAL	SSF + EF	92.2
	SSF + FD2 + EF	94.6
	SSF + FD2 + EF + TE	94.6

- In the first test, *Spontaneous Speech Fluency* test, SSF set alone has been used and also integrated with two different Fractal Dimension (FD) sets. The use of FD features outperforms the results, and the system obtains an improvement of %10 with FD2 set, which includes several FD features.
- In the second test, *Emotional Response* test, the best result is obtained when ET feature is included being Acc near to the optimum value.
- With regard to the *INTEGRAL* test, (which includes information relative to both Spontaneous Speech and Emotional Response) results show also an improvement of non-linear features and better results when ET is included.

Figure 5 shows the obtained results for control group (CR) and the three levels of AD (ES, IS and AS) for classes. In these classes' results the inclusion of non-linear features FD2 set obtains the best results for all classes (Figure 5). This set improves also the classification with regard to early detection (ES class). IS has also better rate to discriminate middle AD level. The model is able also to

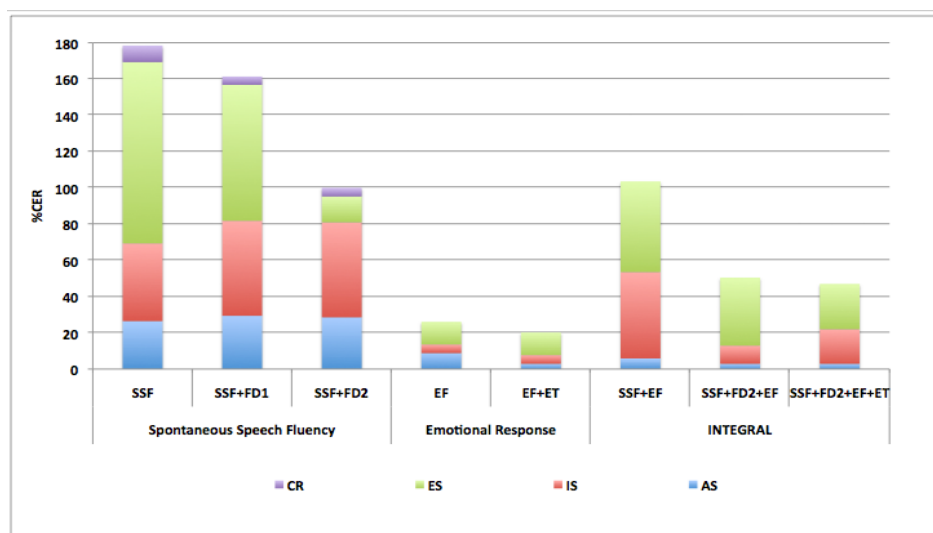
discriminate pathological and non-pathological segments in each patient for the three tests of the pilot study.

Figure 5. Accuracy in % for the three defined tests in the pilot study and each corresponding feature sets.



Accumulative Classification Error Rate (%CER) is detailed in Figure 6. Less accumulative error rate is obtained for experiments which include FD2 non-linear feature set and ET. This is relevant for early diagnosis because in these cases better results are obtained with smaller classification error for ED class.

Figure 6. Accumulative Classification Error Rate (%) for the three defined tests and each corresponding feature sets.

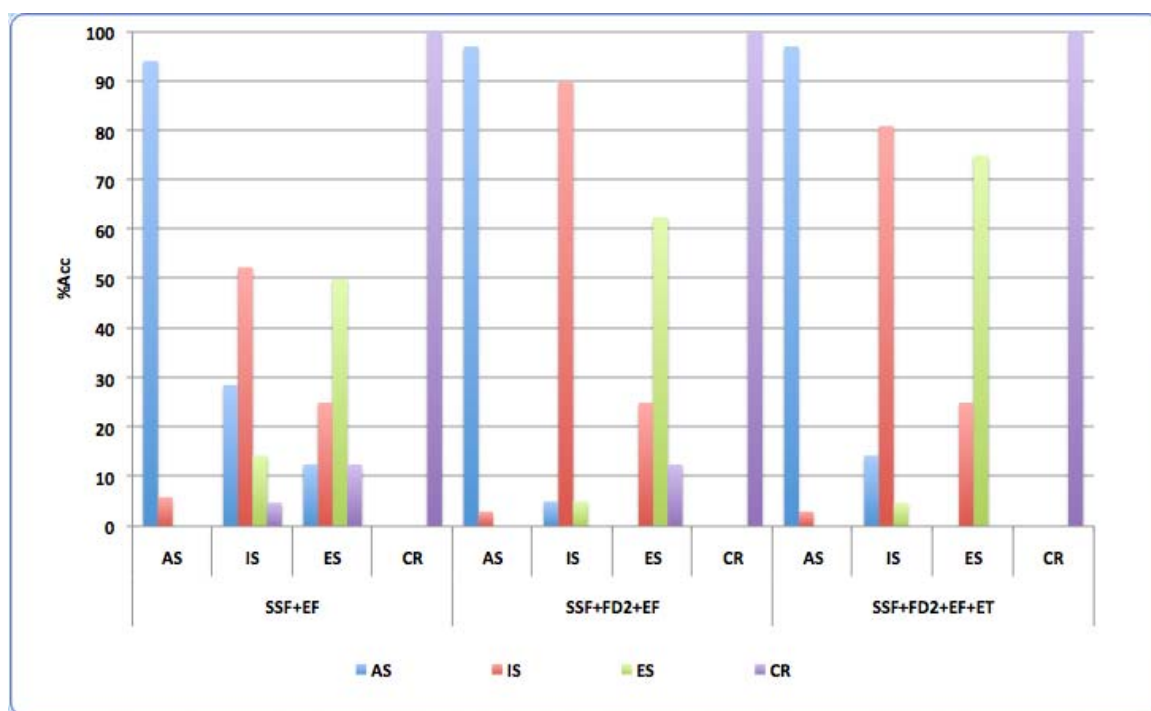


A detailed analysis of *INTEGRAL* test, with regard to Acc (%) for all classes is showed in Figure 7.

- When the results with simple speech features are analyzed, it is observed optimum performance for control group but very mixed classification in AD levels mainly for IS and ES.

- The global system obtains very good results when FD2 set is added because not only outperforms Acc in middle stage IS, but also there is a great improvement in Acc for ES.
- In the third experiment, ET is included and Acc obtains very good results for all classes, and ES outperforms because confusion disappears with regard to the segments of control subjects.

Figure 7. Accuracy (%) of classes for *INTEGRAL* test and each corresponding feature sets.



Finally, the health specialists notice the relevance of the system's ability to carry out both the analysis of independent biomarkers as Spontaneous Speech and Emotional Response features, and/or the integral analysis of several biomarkers. The final confusion among segments could be due, in some cases, to the possible occurrence of segments with different pathological levels in the same individual. This possibility will be explored with new tests in future works. Therefore these non-invasive tests could be a very useful tool for medical specialists' in future clinical AD early diagnosis grounds.

5. Conclusions

The main goal of the present work is the analysis of features in Spontaneous Speech and Emotional Response oriented to pre-clinical evaluation for the definition of appropriate tests for early AD diagnosis. These features are of great relevance for health specialists to define health people and the three AD levels. Features relative to speech duration, time domain, spectral domain and fractal dimension have been analyzed. In this work a first approach including nonlinear features is described. More precisely, an implementation of Higuchi's algorithm in order to add this new feature to the set that feeds the training process of the model. The approach's performance is very satisfactory and promising results for early diagnosis and classification of AD patient groups. Moreover, new features (Higuchi Fractal Dimension and Emotional Temperature) significantly outperform previous results. In

future work we will evaluate this approach with an early diagnosis database and new tests oriented to semantic and memory tasks, and we will also introduce new features relatives to non-linear dynamic.

Acknowledgments

This work has been partially supported by a SAIOTEK from the Basque Government, University of Vic under the research grant R0904, and the Spanish Ministerio de Ciencia e Innovación TEC2012-38630-C04-03. Iciar Martinez (Research Center for Experimental Marine Biology and Biotechnology-Plentziako Itsas Estazioa (PIE), University of the Basque Country & IKERBASQUE, Basque Foundation for Science, is gratefully acknowledged for helpful discussions and for her contribution to the preparation of the manuscript.

Conflict of Interest

The authors declare no conflict of interest.

References

1. McKhann, G.; Drachman, D.; Folstein, M.; Katzman, R.; Price, D.; Stadlan, E.M. Clinical diagnosis of Alzheimer's disease: Report of the National Institute on Aging-Alzheimer's Association workgroup on Alzheimer's disease. *Neurology* **1984**, *24*, 939–944.
2. McKhann, G.; Knopman, D.S.; Chertkow, H.; Hyman, B.T.; Jack, C.R., Jr.; Kawas, C.H.; Klunk, W.E.; Koroshetz, W.J.; Manly, J.J.; Mayeux, R.; *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* **2011**, *7*, 263–269.
3. Alzheimer's Association. Available online: <http://www.alz.org> (accessed on 15 May 2013).
4. *Diagnostic and Statistical Manual of Mental Disorders*, 4th ed.; American Psychiatric Association: Washington, DC, USA, 2000.
5. Family Caregiver Alliance (FGA), National Center on Caregiving. Alzheimer's Disease. Available online: <http://www.caregiver.org/caregiver/jsp/home.jsp> (accessed on 15 May 2013).
6. Morris, J.C. The Clinical dementia rating (CDR): Current version and scoring rules. *Neurology* **1993**, *43*, 2412b–2414b.
7. Van de Pole, L.A.; Van der Flier, W.M.; Hensel, A.; Gertz, H.J.; Scheltens, P. The effects of age and Alzheimer's disease on hippocampal volumes, a MRI study. *Alzheimer's and Dementia. J. Alzheimer's Assoc.* **2005**, *1*(Supplement 1), doi:10.1016/j.jalz.2005.06.205.
8. Laakso, M.P.; Lehtovirta, M.; Partanen, K.; Riekkinen, P.J., Sr.; Soininen, H. Hippocampus in Alzheimer's disease: A 3-year follow-up MRI study. *Biol. Psychiat.* **2000**, *47*, 557–561.
9. Petrella, J.R.; Coleman, R.; Doraiswamy, P. Neuroimaging and early diagnosis of Alzheimer disease: A look to the future. *Radiology* **2003**, *226*, 315–336.
10. Alzheimer's Association/Biomarkers. Available online: http://www.alz.org/research/funding/global_biomarker_consortium.asp (accessed on 15 May 2013).

11. Wernickand, M.N.; Aarsvold, J.N. *Emission Tomography: The Fundamentals of PET and SPECT*; Elsevier: New York, NY, USA, 2004.
12. Pareto, D.; Aguiar, P.; Pavia, J.; Gispert, J.; Cot, A.; Falcon, C.; Benabarre, A.; Lomena, F.; Vieta, E.; Ros, D. Assessment of SPM in per-fusion brain SPECT studies. A numerical simulation study using boot-strap resampling methods. *IEEE Trans. Biomed. Eng. TBE* **2008**, *55*, 1849–1853.
13. Álvarez, I.; Górriz, J.M.; Ramírez, J.; Salas-Gonzalez, D.; López, M.; Segovia, F.; Padilla, P.; Gracia, C. Projecting independent components of SPECT images for computer aided diagnosis of Alzheimer's disease. *Pattern Recognit. Lett.* **2010**, *31*, 1342–1347.
14. Faundez-Zanuy, M.; Hussain, A.; Mekyska, J.; Sesa-Nogueras, E.; Monte-Moreno, E.; Esposito, A.; Chetouani, M.; Garre-Olmo, J.; Abel, A.; Smekal, Z.; *et al.* Biometric applications related to human beings: There is life beyond security. *Cognitive Computation* **2012**, *5*, 136–151.
15. López-de-Ipiña, K.; Alonso, J.B.; Solé-Casals, J.; Barroso, N.; Faundez, M.; Ecay-Torres, M.; Travieso, C.; Ezeiza, A.; Estanga, A. New approaches for Alzheimer's disease diagnosis based on automatic spontaneous speech analysis and emotional temperature. *Ambient Assist. Living Home Care* **2012**, *7657*, 407–414.
16. Hu, W.T.; McMillan, C.; Libon, D.; Leight, S.; Forman, M.; Lee, V.M.Y.; Trojanowski, J.Q.; Grossman, M. Multimodal predictors for Alzheimer disease in nonfluent primary progressive aphasia. *Neurology* **2010**, *75*, 595–602.
17. Grimaldi, G.; Manto, M. Neurological tremor: Sensors, signal processing and emerging applications. *Sensors* **2010**, *10*, 1399–1422.
18. Almeida, A.; López-de-Ipiña, D. Assessing ambiguity of context data in intelligent environments: Towards a more reliable context managing system. *Sensors* **2012**, *12*, 4934–4951.
19. Bravo, J.; Villarreal, V.; Hervás, R.; Urzaiz, G. Using a communication model to collect measurement data through mobile devices. *Sensors* **2012**, *12*, 9253–9272.
20. Horley, K.; Reid, A.; Burnham, D. Emotional prosody perception and production in dementia of the Alzheimer's type. *J. Speech Lang. Hear. Res.* **2010**, *53*, 1132–1146.
21. Pakhomov, S.V.S.; Hemmy, L.S.; Lim, K.O. Automated semantic indices related to cognitive function and rate of cognitive decline. *Neuropsychologia* **2012**, *50*, 2165–2175.
22. Haugrud, N.; Crossley, M.; Vrbancic, M. Clustering and switching strategies during verbal fluency performance differentiate Alzheimer's disease and healthy aging. *J. Speech, Lang. Hear. Res.* **2010**, *53*, 1132–1146.
23. Henry, J.D.; Crawford, J.R.; Pshillips, L.H. Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia* **2004**, *42*, 1212–1222.
24. Meilán, J.J.; Martínez-Sánchez, F.; Carro, J.; Sánchez, J.A.; Pérez, E. Acoustic markers associated with impairment in language processing in Alzheimer's Disease. *Span. J. Psychol.* **2012**, *15*, 487–494.
25. Knapp, M.L. *Essentials of Nonverbal Communication*; Holt, Rinehart & Winston: New York, NY, USA, 1980.
26. Cowie, R.; Douglas-Cowie, E.; Tsapatsoulis, N.; Votsis, G.; Kollias, S.; Fellenz, W.; Taylor, J.G. Emotion recognition in human-computer interaction. *IEEE Signal Process. Mag.* **2001**, *18*, 32–80.
27. Plutchik, R. *Emotion: A Psychoevolutionary Synthesis*; Harper and Row: New York, NY, USA, 1980.

28. Buiza, C. *Evaluación y Tratamiento de los Trastornos del Lenguaje*; Matia Fundazioa: Donostia, Spain, 2010.
29. Martinez, F.; Garcia, J.; Perez, E.; Carro, J.; Anara, J.M. Patrones de Prosodia expresiva en pacientes con enfermedad de Alzheimer. *Psicothema* **2012**, *24*, 16–21.
30. Shimokawa, A.; Yatomi, N.; Anamizu, S.; Torii, S.; Isono, H.; Sugai, Y.; Kohno, M. Influence of deteriorating ability of emotional comprehension on interpersonal behaviour in Alzheimer-type dementia. *Brain Cognit.* **2001**, *47*, 423–433.
31. Goodkind, M.S.; Gyurak, A.; McCarthy, M.; Miller, B.L.; Levenson, R.W. Emotion regulation deficits in frontotemporal lobar degeneration and Alzheimer's disease. *Psychol. Aging* **2010**, *25*, 30–37.
32. Cadieux, N.; Greeve, K. Emotion processing in Alzheimer's disease. *J. Int. Neuropsychol. Soc.* **1997**, *3*, 411–419.
33. Henry, J.D.; Rendell, P.G.; Scicluna, A.; Jackson, M.; Phillips, L.H. Emotion experience, expression, and regulation in Alzheimer's disease. *Psychol. Aging* **2009**, *24*, 252–257.
34. Praat: Doing Phonetics by Computer. Available online: www.fon.hum.uva.nl/praat (accessed on 15 May 2013).
35. Voice Activity Detector Algorithm (VAD). Available online: www.mathwork.com (accessed on 15 May 2013).
36. Solé-Casals, J.; Zaiats, P.V. A non-linear VAD for noisy environments. *Cogn. Comput.* **2010**, *2*, 191–198.
37. Solé-Casals, J.; Zaiats, P.V.; Monte-Moreno, E. Non-linear and non-conventional speech processing: Alternative techniques. *Cogn. Comput.* **2010**, *2*, 133–134.
38. Ezeiza, A.; López-de-Ipiña, K.; Hernández, C. Enhancing the feature Extraction Process for Automatic Speech Recognition with Fractal Dimensions. Available online: <http://link.springer.com/content/pdf/10.1007/s12559-012-9165-0.pdf> (accessed on 15 May 2013).
39. Higuchi, T. Approach to an irregular time series on the basis of the fractal theory. *Phys. D* **1988**, *31*, 277–283.
40. Schuller, B.; Batliner, A.; Seppi, D.; Steidl, S.; Vogt, T.; Wagner, J.; Devillers, L.; Vidrascu, L.; Amir, N.; Kessous, L.; *et al.* The Relevance of Feature Type for the Automatic Classification of Emotional User States: Low Level Descriptors and Functionals. In Proceedings of 8th Annual Conference of the International Speech Communication Association (Interspeech'07), Antwerp, Belgium, 27–31 August 2007; pp. 2253–2256.
41. Vogt, T., André, E. Comparing Feature Sets for Acted and Spontaneous Speech in View of Automatic Emotion Recognition. In Proceeding of IEEE International Conference on Multimedia and Expo (ICME'05), Amsterdam, The Netherlands, 5–8 July 2005; pp. 474–477.
42. Pao, T.L.; Chien, C.S.; Yen, J.H.; Chen, Y.T.; Cheng, Y.M. Continuous Tracking of User Emotion in Mandarin Emotional Speech. In Proceedings of 3th International Conference on International Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), Splendor Kaohsiung, Taiwan, 26–28 November 2007; Volume 1, pp. 47–52.
43. Petrushin, V.A. Emotion in Speech: Recognition and Application to Call Centers. In Proceedings of Conference on Artificial Neural Networks in Engineering (ANNIE 1999), St. Louis, MO, USA, 7–10 November 1999; pp. 7–10.

44. Lee, C.M.; Narayanan, S. Emotion Recognition Using a Data-Driven Fuzzy Interference System. In Proceedings of 8th European Conference on Speech Communication and Technology (ECSCT 2003), Geneva, Switzerland, 1–4 September 2003; pp. 157–160.
45. Kwon, O.W.; Chan, K.; Hao, J.; Lee, T.W. Emotion Recognition by Speech Signals. In Proceedings of 8th European Conference on Speech Communication and Technology (ECSCT 2003), Geneva, Switzerland, 1–4 September 2003; pp. 125–128.
46. De Cheveigné, A.; Kawahara, H. YIN, a fundamental frequency estimator for speech and music. *J. Acoust. Soc. Am.* **2002**, *111*, 1917–1930.
47. Alonso, J.; de León, J.; Alonso, I.; Ferrer, M.A. Automatic detection of pathologies in the voice by HOS base parameters. *J. Appl. Signal Process.* **2001**, *4*, 275–284.
48. Chang, C.C.; Lin, C.J. LIBSVM: A Library for Support Vector Machines. Available online: <http://www.csie.ntu.edu.tw/~cjlin/libsvm> (accessed on 15 May 2013).
49. WEKA. Available online: <http://www.cs.waikato.ac.nz/ml/weka/> (accessed on 15 May 2013).
50. Picard, R.; Cook, D. Cross-validation of regression models. *J. Am. Statist. Assoc.* **1984**, *79*, 575–583.

© 2013 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

B2 ERANSKINA

Selection of Entropy Based Features for
Automatic Analysis of Essential Tremor

Article

Selection of Entropy Based Features for Automatic Analysis of Essential Tremor[†]

Karmele López-de-Ipiña^{1,*}, Jordi Solé-Casals², Marcos Faundez-Zanuy³, Pilar M. Calvo¹, Enric Sesa³, Unai Martínez de Lizarduy¹, Patricia De La Riva⁴, Jose F. Marti-Masso⁴, Blanca Beitia¹ and Alberto Bergareche⁴

¹ Systems Engineering and Automation Department, University of the Basque Country UPV/EHU, Donostia 20018, Spain; pilarmaria.calvo@ehu.eus (P.M.C.); unai.martinezdelizarduy@ehu.eus (U.M.-d.-L.); mariablanca.beitia@ehu.eus (B.B.)

² Data and Signal Processing Research Group, University of Vic—Central University of Catalonia, Vic, Catalonia 08500, Spain; jordi.sole@uvic.cat

³ Escola Superior Politècnica Tecnocampus (UPF), Mataró, Catalonia 08302, Spain; faundez@tecnocampus.cat (M.F.-Z.); sesa@tecnocampus.cat (E.S.)

⁴ BioDonostia Health Institute, Neurology Department Hospital Donostia, Donostia 20014, Spain; patricia.delarivajuez@osakidetza.net (P.D.L.R.); josefelix.martimasso@osakidetza.net (J.F.M.-M.); jesusalberto.bergarecheyarza@osakidetza.net (A.B.)

* Correspondence: karmele.ipina@ehu.es; Tel.: +34-943-01-86-67; Fax: +34-943-01-72-30

† This paper is an extended version of one paper published in the 4th IEEE International Work Conference on Bioinspired Intelligence, Donostia, Spain, 9–12 June 2015.

Academic Editors: Carlos M. Travieso-González and Jesús B. Alonso-Hernández

Received: 8 March 2016; Accepted: 9 May 2016; Published: 16 May 2016

Abstract: Biomedical systems produce biosignals that arise from interaction mechanisms. In a general form, those mechanisms occur across multiple scales, both spatial and temporal, and contain linear and non-linear information. In this framework, entropy measures are good candidates in order provide useful evidence about disorder in the system, lack of information in time-series and/or irregularity of the signals. The most common movement disorder is essential tremor (ET), which occurs 20 times more than Parkinson's disease. Interestingly, about 50%–70% of the cases of ET have a genetic origin. One of the most used standard tests for clinical diagnosis of ET is Archimedes' spiral drawing. This work focuses on the selection of non-linear biomarkers from such drawings and handwriting, and it is part of a wider cross study on the diagnosis of essential tremor, where our piece of research presents the selection of entropy features for early ET diagnosis. Classic entropy features are compared with features based on permutation entropy. Automatic analysis system settled on several Machine Learning paradigms is performed, while automatic features selection is implemented by means of ANOVA (analysis of variance) test. The obtained results for early detection are promising and appear applicable to real environments.

Keywords: permutation entropy; essential tremor; automatic drawing analysis; Archimedes' spiral; non-linear features; automatic feature selection

1. Introduction

Biomedical systems produce biosignals that arise from interaction mechanisms. In a general form, those mechanisms occur across multiple scales, both spatial and temporal, and contain linear and non-linear information. Complex fluctuations are habitually present in the output variables of real systems. These fluctuations are due to noise but also contain information about the dynamics of the system. Linear methods can capture global aspects of the dynamics, but the different approaches are not able to discern all the relevant physical details [1,2]. In this framework, the measurement

of non-linear features (for example the entropy of the system) is an essential and useful tool to investigate the state of the system. This analysis provides the information encoded in the system and the probability distributions of all its possible states [1]. Particular characteristics of data affect the applicability of entropy-based methodologies. In this sense, stationarity property, noise level, length of the time series, *etc.*, are of great relevance, since important information may be present in the temporal dynamics. Habitually, all those aspects are not taken into account [1,3].

Biological and biomedical systems generate time series generated that contain deterministic and stochastic components [4]. Classic methods of signal and noise analysis can deal with part of the interesting features, but they only model linear components without yielding any information about non-linearities, irregularities or stochastic components. When analyzing inconspicuous changes this complex information could be essential. Massimiliano Zanin *et al.* [1] present a review based on biomedical applications, which includes the analysis of heart rhythms, anesthesia, electroencephalography (EEG) or cognitive neuroscience. These last ones, related to neurological diseases, are challenging due to their variability and the tremendous impact they exert on the society.

Essential tremor (ET) affects individuals worldwide and it is 20 times more present than Parkinson's disease. The prevalence of ET in the western countries is around 0.3%–4.0%. Males and females of 40 years old are affected approximately equally (incidence of 23.7 per 100,000 people per year). Several studies suggest prevalence from 3.9% to 14.0% in these patients, and 50% to 70% of the cases of ET seems to have a genetic origin [5]. Essential tremor is a rhythmic tremor (4–12 Hz) that only occurs when the affected muscle is exerting. The amplitude of the tremor increases and varies with age, but there is no gender distinction. Additionally, physical or mental stress could worsen the tremor. Parkinson's disease (PD) and Parkinsonism can occur simultaneously with essential tremor. In fact, the prevalence of PD in people with ET is greater than in the general population. Concerning symptoms, hand tremor is the predominant one (as it is in PD) and occurs in nearly all cases, followed by other sort of tremors like voice, head, face, neck, tongue, leg and trunk. Due to its (partially) genetic origin, PD and ET often occur in individuals of the same family [5].

In order to manage and palliate the symptoms, it is of great importance to clinically detect the earliest manifestations of the disorder. In the past few years, approaches seeking early diagnosis of ET have made significant advances towards the development of consistent clinical biomarkers. Despite the valuableness of those biomarkers, cost and technology requirements make unviable to apply such tests to all patients with motor disorders. In this framework, early detection performed through non-invasive intelligent techniques could be a good alternative. Non-technical staff could use those methodologies without altering the patients' abilities, because handwriting analysis, drawing analysis or speech analysis are not perceived as stressful tests by the patients. Furthermore, these are low-cost techniques and do not require medical equipment nor extensive infrastructures [6–9].

Doctors traditionally use handwritten task in order to diagnose ET. Archimedes' spiral is a well-known and established test [10], therefore we will focus our work on the analysis of such drawings. Traditionally, the analysis of handwriting was performed offline because only the writing itself (strokes on a paper) was available. Currently modern digitizing tablets and pens (with or without ink) can gather data keeping all temporal information and the dynamics of the whole process. In this case, the analysis is referred to as online. Modern digitizing tablets capture the x , y and z coordinates of the movement of the writing process, the applied pressure on the surface and the azimuth and altitude angles (the angles of the pen in the horizontal plane and with respect to the vertical axis, respectively) [11]. This makes it possible to analyze both static (offline) and dynamic (online) features [12].

The results presented here are part of a wider cross study on the diagnosis of ET carried out by the Biodonostia Health Institute with the objective of characterizing this impairment, and it is based on families with identified genetics loci. Among several drawings and handwriting exercises, Archimedes' spiral has been chosen and will be explored to determine the best non-linear biomarkers for early diagnosis and following of ET [13,14]. The following sections analyze classic static and dynamic linear

features and also non-linear ones based on several entropy algorithms. Automatic methodologies will be used in the selection of biomarkers. Finally, the quality of the selected features is measured by ANOVA, multiple comparisons test, and Machine Learning paradigms.

2. Materials

2.1. Acquisition System

The acquisition is carried out by means of an Intuos Wacom 4 digitizing tablet. The pen tablet USB [11] captures the following information (Figure 1) at a sample frequency of 100 Hz [12,13]: the spatial coordinates (x , y), the pressure, and azimuth and altitude angles. Using this set of dynamic data, further information can be inferred such as acceleration, speed, instantaneous trajectory angle, instantaneous movement, tangential acceleration, curvature radius, centripetal acceleration *etc.* [12,15].

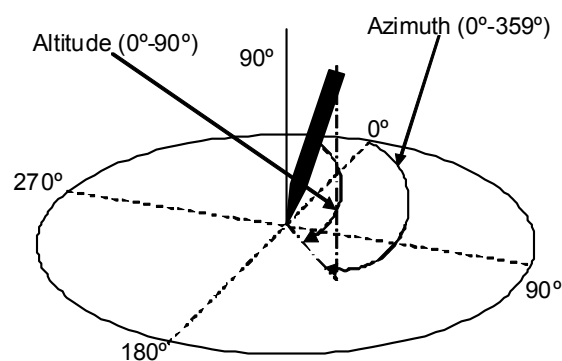


Figure 1. Information extracted from the digitizing tablet.

2.2. Database of Individuals

The database BIODARW consists of 21 healthy subjects (control group, CR) and 29 patients diagnosed with ET (ET group) with identified genetics loci. For all patients, register of electrophysiological test (EPT) and functional magnetic resonance imaging (fMRI) are also available. Both hands are considered in order to perform the handwritten test. Therefore, the total number of samples is $2 \times (21 + 29) = 100$. The handwriting test consists of drawing a line, drawing the Archimedes' spiral and handwriting with dominant hand and non-dominant hand. In this work only Archimedes' spiral is used. The database has variability with regard to: tremor frequency, amplitude and pattern, rating scale values, and demographic data (age and gender). Subjects were recruited from patients of a previous descriptive study that considers familiar and sporadic ET cases and controls from the Movement Disorders Unit at the Donostia University Hospital (San Sebastian, Spain). Written informed consent, fully approved by the local ethics committee at the Donostia University Hospital, was obtained from all participants in this new study. Table 1 summarizes the features of the group with ET with regard to EPT, diagnosis and demography [13,14]. In addition to the standard clinical exploration (Neuropsychological and Electrophysiological Studies), evaluation of ET was carried out by recording the drawing of an Archimedes' spiral. The Montreal Cognitive Assessment [16] with the Fahn–Tolosa–Marin (FTM) tremor rating scale, which assesses different cognitive domains, was used to determine possible cognitive dysfunction; and criteria from the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) of the American Psychiatric Association (APA) were used for the diagnosis of depression and anxiety disorders. Surface electromyography (EMG) was recorded from wrist extensor and flexor muscles using surface electrodes placed over the muscle bellies 3 cm apart. The filters were set with a band-pass of 10–1000 Hz. A triaxial accelerometer was placed over the first dorsal interossei muscle of the hand.

Table 1. Some samples of the electrophysiological test (EPT). Fahn–Tolosa–Marin (FTM) scale values for the selected individuals with ET (ET_X).

ET_X	EPT Features			Diagnosis	Demography	
	Frequency (Hz)	Amplitude (v)	Pattern	FTM Scale	Age	Gender
ET_01	8.5	20	synchronous	1	48	Female
ET_02	6.5	variable	alternating	8	72	Male
ET_03	10.5	200	synchronous	1	46	Male
ET_04	4.5	503.6	synchronous	3	80	Female
ET_05	6.6	298	synchronous	22	68	Female
ET_06	9.5	46	synchronous	2	46	Female
ET_07	5	173	synchronous	50	75	Male
ET_08	6.5	159	synchronous	40	75	Male
ET_09	8	128	asynchronous	9	75	Female

2.3. Individuals Selected for the Study

From the original database of 100 samples, a subset of the samples of Archimedes' spiral is selected. The team of medical doctors carried out this selection. For the control group, the best sample (usually the dominant hand) is kept, but in some small number of cases the non-dominant hand is used as well. For the ET group, only the sample with the best quality is chosen (one hand), but five subjects are discarded due to the poor quality of the samples. Thus, this sub-database BIODARWO consists of 51 samples: 27 samples for the control group and 24 samples for the ET group.

3. Methods

3.1. Online Drawing Applied to Health Analysis

In the medical field, the study of handwriting has proved to be an aid to diagnosing and tracking some diseases of the nervous system. For instance, handwriting skill degradation and Alzheimer's disease (AD) appear to be significantly correlated [17,18], and some aspects of handwriting can be good indicators for its diagnosis [6,17,18] or help to differentiate between mild Alzheimer's disease and mild cognitive impairment [17,18]. Additionally, the analysis of handwriting has proved useful to assess the effects of substances such as alcohol [19], marijuana [20] or caffeine [21]. Thanks to modern acquisition devices, the field of psychology has also benefited from the analysis of handwriting. Rosenblum *et al.* [22] link the proficiency of the writers to the length of the in-air trajectories of their handwritings. The visual inspection of the pen-down image suggests a progressive degree of impairment when drawing becomes more disorganized, and only ET patients in mild states can achieve the three dimensional effect. The visual information provided by the pen-up drawing of pathological individuals also indicates a progressive impairment and disorganization when the individuals try to plan the drawing. It is also important to note that the comparison of pen-up drawings between people with disease and the control group shows noticeable differences as well. In addition to the increased time in-air, there is a higher number of hand movements before putting the pen on the surface to draw. We consider that these graph-motor measures applied to the analysis of drawing and writing functions may be a useful alternative to study the precise nature and progression of the drawing and writing disorders associated with several neurodegenerative diseases.

3.2. Pressure Derived Measures and in-Air Analysis

One of the major advantages of online drawing is the possibility of evaluating the handwriting pressure quantitatively. For instance, people diagnosed with AD produce softer and simpler strokes, and people with ET are less stable. ET and Parkinson's disease are other examples [13,14,23–25]. For example, in [18], they have reported the proficiency of writers with regard to the length of the in-air trajectories and the pressure of their handwritings [26]. A very interesting aspect of modern

online analysis of handwriting is that it can take into account information gathered when the writing device was not exerting pressure on the writing surface. As an example Figure 2 shows the acquisition of the word INEXPUGNABLE: the pen-up (in-air) movement information is represented in red, while the pen-down (on-surface) movement is represented in blue. Our previous experiments on biometric recognition of people revealed that these two kinds of information are complementary [12], and in fact have a similar discriminating capability, even when using a database of 370 users [15,25,26]. Figure 3 shows an example of the Archimedes' spiral performed by a control subject (Figure 3a) and a subject with ET (Figure 3b). The blue line corresponds to pen-down (on surface drawing), and the red line to pen-up (in-air drawing). The drawing is very regular in the case of the control subject (Figure 3a), and with clear irregularities in the case of the ET subject (Figure 3b) not only with regard to Cartesian components but also to the control of the pressure level.

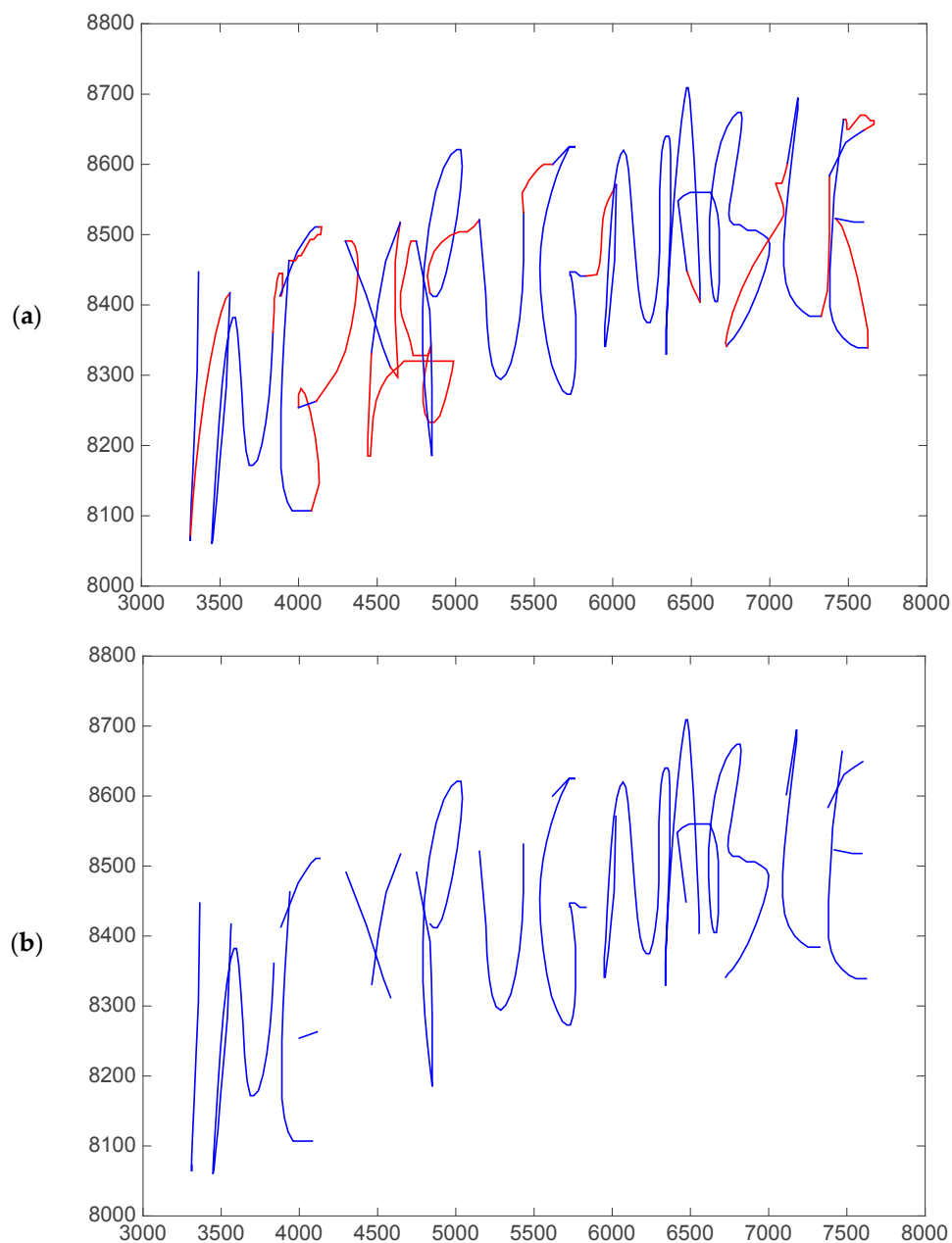


Figure 2. Cont.

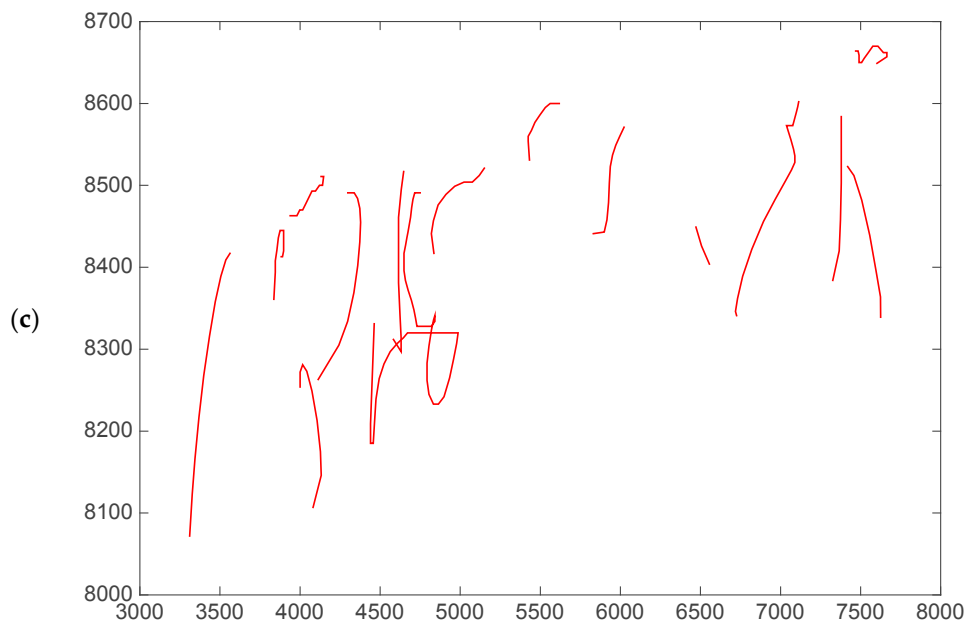


Figure 2. Execution of the word INEXPUGNABLE as captured by the acquisition device: (a) both pen-up and pen-down strokes; (b) pen-down strokes; and (c) pen-up strokes.

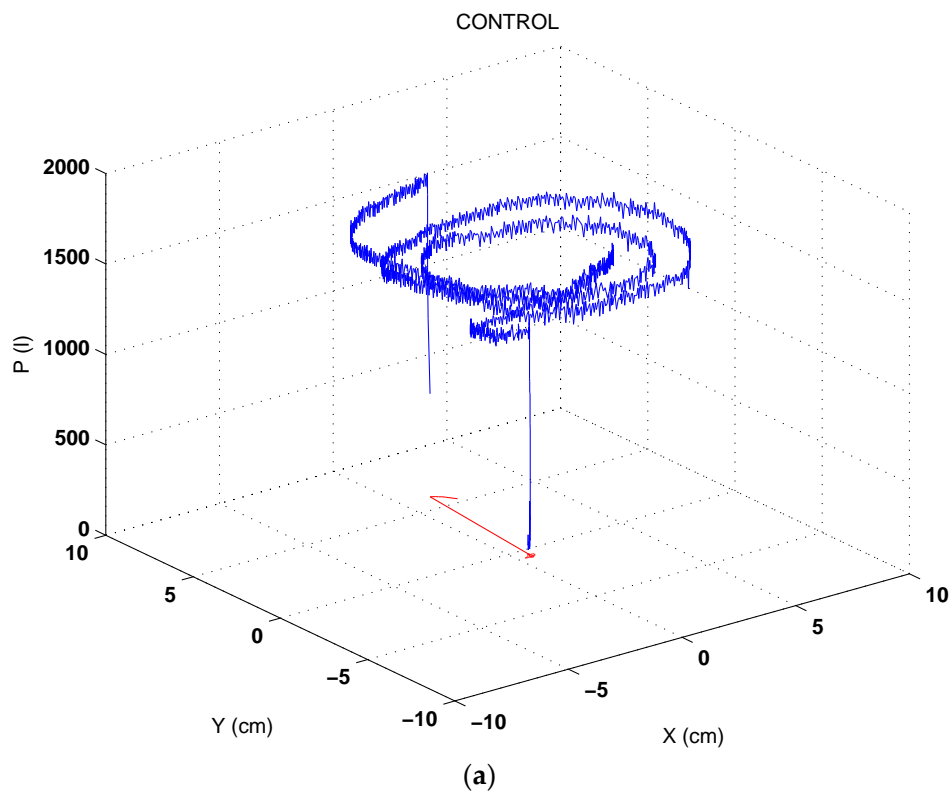


Figure 3. Cont.

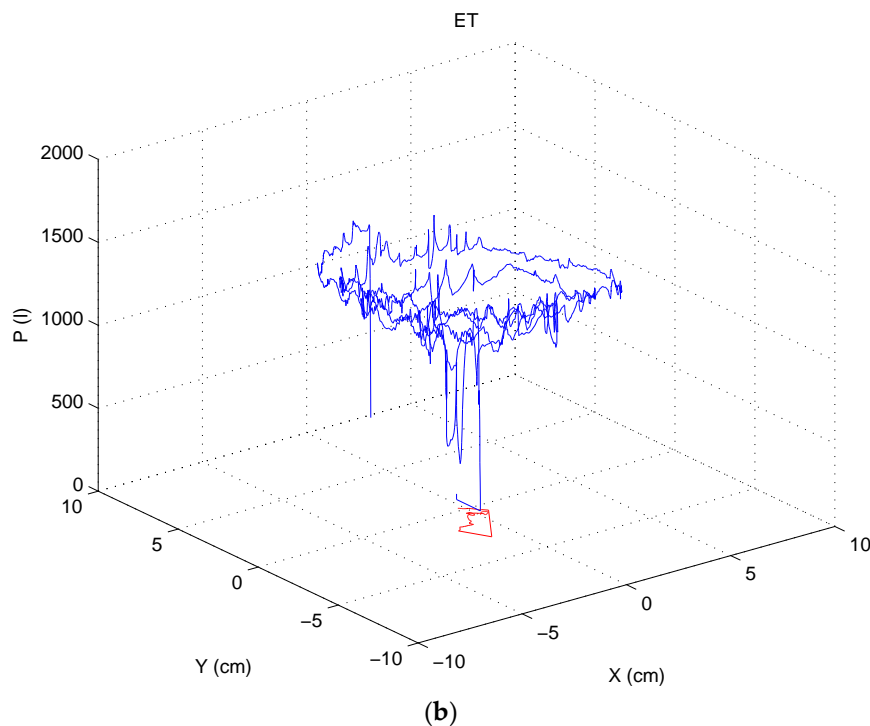


Figure 3. Archimedes' spiral 3D drawings performed by a control subject (a); and a subject with ET (b). Cartesian coordinates X (cm), Y (cm) and Pressure level P (l). The blue line corresponds to pen-down (on surface drawing), and the red line to pen-up (in-air drawing). The drawing is very regular in the case of the control subject (a); and with clear irregularities in the case of the ET subject (b) not only with regard to Cartesian components but also to the control of the pressure level.

3.3. Features Extraction

The research presented here has a preliminary nature; its aim is to define thresholds for a number of biomarkers related to handwriting. It is part of a wider study focused on early ET detection. Feature search in this work aims at preclinical evaluation in order to define useful tests for ET diagnosis [5,13,14].

3.3.1. Linear Features

In this study, the aim is to automatically distinguish handwriting of ET patients from healthy subjects by means of the analysis of different linear features (LF) and their variants (max, min, mean and median) in handwriting, *i.e.*, the following.

1. Time related measures: Time in-air, time on-surface and total time (in-air plus on-surface). Time has been measured as number of samples.
2. Spatial components and their variants: X and Y Cartesian coordinates, altitude (O) and azimuth (A) angles, and angle and modulus polar components (Z and R, respectively) and their projections over a horizontal axis for both pen-down and pen-up signal (see Figure 4 for an example of the distortion of the polar components in the sample of the ET patient).
3. Pressure and its variants.
4. Dynamic features and their variants: Speed and acceleration for both pen-down and pen-up signals.
5. Zero crossing rate: The rate which evaluates the sign-changes along a signal.
6. Frequency domain: Spectral components for both pen-down and pen-up signals (Figure 5).

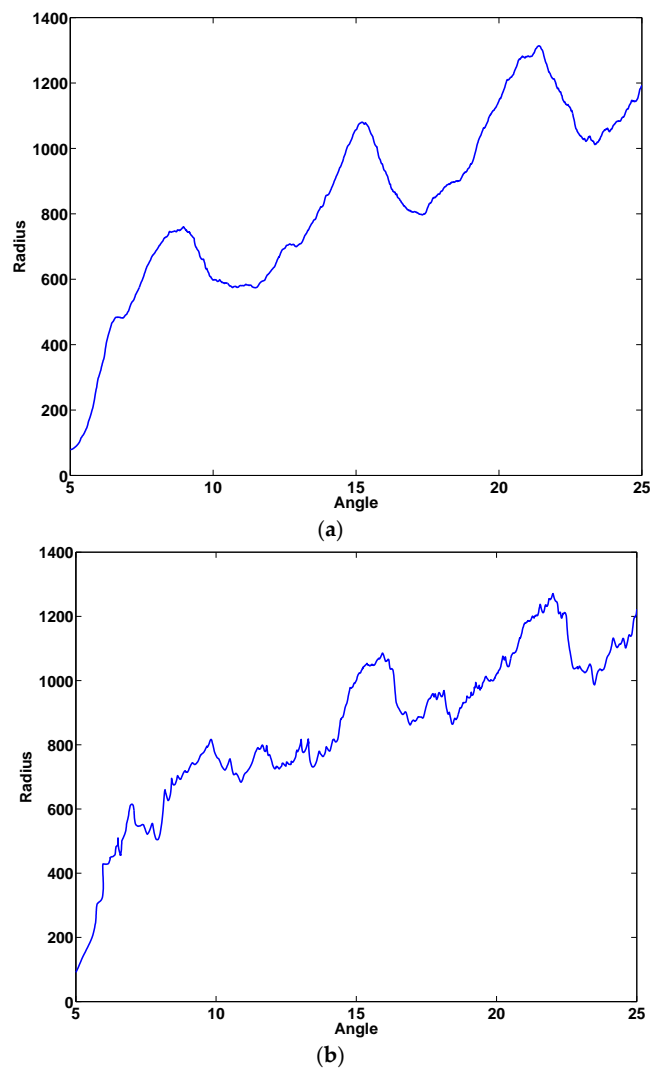


Figure 4. Polar components radius (R) and angle (θ) of the Archimedes' spiral drawing performed by a control subject (a); and a subject with essential tremor (b). The ET patient produces a clear distortion in these components.

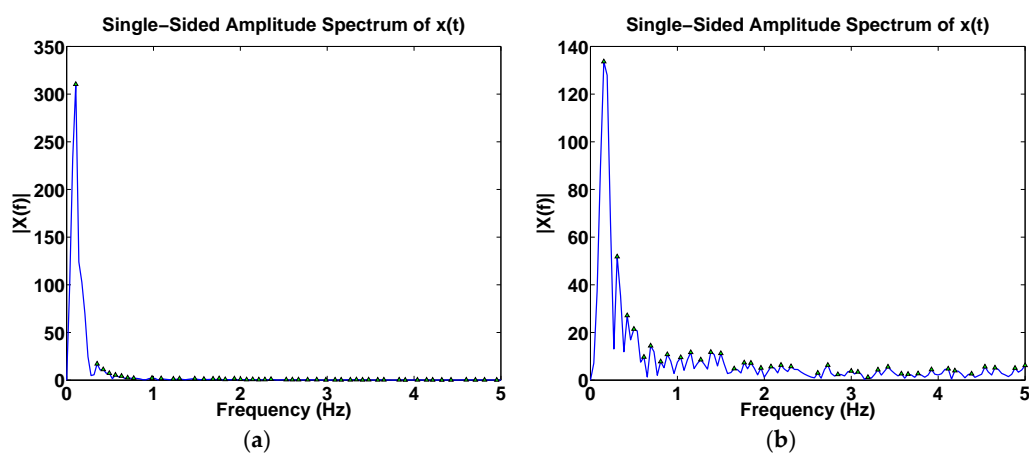


Figure 5. Spectrum of the X component of the *pen-down* (on-surface) signal for a control subject (a); and a subject with essential tremor (b). The spectrum has more components for the subject with ET: it shows many peaks along the whole spectrum, not only in the low frequencies.

3.3.2. Non-Linear Features: Entropy

Entropy is a measure of disorder in physical systems and also a basic quantity with multiple field-specific interpretations. It has been associated with disorder, state-space volume, or lack of information [1,2,27,28]. Shannon entropy is often considered as the classic and most natural way to measure the expected value (average) of the information in a signal [3,4,29,30]. Richman *et al.* analyze that entropy, and in relation to dynamic systems is the rate of information production [31]. On the one hand, some authors point out that calculation of entropy usually requires very long data sets that in the case of biomedical signals can be difficult or impossible to obtain. On the other hand, methods for entropy estimation of a system represented by a time series are not suitable for analyzing the short and noisy data sets of biomedical studies [1,31]. In the following subsections, we present several proposals for calculating entropy used in this work.

Shannon Entropy

The entropy $H(X)$ of a single discrete random variable X is a measure of its average uncertainty. Shannon entropy [29,30] is calculated by the equation:

$$H(X) = - \sum_{x_i \in \Theta} p(x_i) \log p(x_i) = -E[\log p(x_i)] \quad (1)$$

where X represents a random variable with a set of values Θ and probability mass function $p(x_i) = P_r\{X = x_i\}$, $x_i \in \Theta$, and E represents the expectation operator. Note that $p \log p = 0$ if $p = 0$.

For a time series that represents the output of a stochastic process, that is, an indexed sequence of n random variables, $\{X_i\} = \{X_1, \dots, X_n\}$, with a set of values $\theta_1, \dots, \theta_n$, respectively, and $X_i \in \theta_i$, the joint entropy is defined by:

$$H_n = H(X_1, \dots, X_n) = - \sum_{x_i \in \theta} \dots \sum_{x_n \in \theta_n} p(x_1, \dots, x_n) \log p(x_1, \dots, x_n) \quad (2)$$

where $p(x_1, \dots, x_n) = P\{X_1 = x_1, \dots, X_n = x_n\}$ is the joint probability for the n variables X_1, \dots, X_n .

By applying the chain rule to Equation (2), the joint entropy can be written as the addition of conditional entropies, where each of them is a non-negative quantity:

$$H_n = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1) \quad (3)$$

Therefore, as a conclusion, the joint entropy is an increasing function of n . The rate at which the joint entropy grows with n , *i.e.*, the entropy rate h , is defined as:

$$h = \lim_{n \rightarrow \infty} \frac{H_n}{n} \quad (4)$$

Approximate Entropy *versus* Sample Entropy

The approximate entropy is a statistical measure that smoothens transient interference and can suppress the influence of noise by properly setting the parameters of the algorithm. It can be used in the analysis of both stochastic and deterministic signals [32,33]. This is crucial in the case of biological signals, which are outputs of complex biological networks and may be deterministic, stochastic, or both. Approximate entropy provides a model-independent measure of the irregularity of the signals. The algorithm summarizes a time series into a non-negative number, where higher values represent more irregular systems [32,33].

The method examines time series for similar epochs [34]: more frequent and more similar epochs lead to lower values of approximate entropy. ApEn (m, r, n) measures the conditional probability that

two sequences of length n which are similar for m points remain similar at the next sample point, within a tolerance r . Thus, a low value of approximate entropy reflects a high degree of regularity. Approximate entropy algorithm counts each sequence as matching itself to reduce bias; sample entropy $\text{SmEn}(m, r, n)$ was developed in order not to count self-matches.

The sample entropy is defined for a time series of n points. We first define the $n - m + 1$ vectors $\mathbf{x}_m(i) = \{u(i+k) : 0 \leq k \leq m - 1\}$ as the vectors of m data points from $u(i)$ to $u(i+m-1)$. The distance between two such vectors is defined as $d[\mathbf{x}_m(i), \mathbf{x}_m(j)] = \max_k \{|u(i+k) - u(j+k)| : 0 \leq k \leq m - 1\}$, i.e., the maximum difference of their scalar components. The sample entropy $\text{SmEn}(m, r, n)$ is defined as:

$$\text{SmEn}(m, r, n) = \lim_{n \rightarrow \infty} \{-\ln(A^m(r)/B^m(r))\} = -\ln(A/B) \tag{5}$$

where

$$A = [(n - m - 1)(n - m)/2] A^m(r) \tag{6}$$

and

$$B = [(n - m - 1)(n - m)/2] B^m(r) \tag{7}$$

$B^m(r)$ is the probability that two sequences match for m points:

$$B^m(r) = (n - m)^{-1} \sum_{i=1}^{n-m} B_i^m(r) \tag{8}$$

where $B_i^m(r)$ is $(n - m - 1)^{-1}$ times the number of vectors $\mathbf{x}_m(j)$ within a tolerance r of $\mathbf{x}_m(i)$. Similarly, $A^m(r)$ is the probability that two sequences match for $m + 1$ points:

$$A^m(r) = (n - m)^{-1} \sum_{i=1}^{n-m} A_i^m(r) \tag{9}$$

where $A_i^m(r)$ is $(n - m - 1)^{-1}$ times the number of vectors $\mathbf{x}_{m+1}(j)$ within a tolerance r of $\mathbf{x}_{m+1}(i)$. The scalar r is the tolerance for accepting matches. In the present investigation we used the parameters recommended in [35]: $m = 2$ and $r = 0.2$, and standard deviation of the sources is normalized to 1. Sample entropy is a robust quantifier of complexity for instance for electroencephalography (EEG) signals [36], and can be used as a marker for the presence of artifacts in EEG recordings [37]. The quantity A/B is the conditional probability that two sequences within a tolerance r points remain within r of each other at the next point. In contrast to approximate entropy, which calculates probabilities in a template-wise fashion, sample entropy calculates the negative logarithm of a probability associated with the time series as a whole.

Multivariate Multiscale Permutation Entropy

Permutation entropy (PE) directly analyzes the temporal information contained in the time series; furthermore, it has the quality of simplicity, robustness and very low computational cost [1,3,4]. Bandt and Pompe [38] introduce a simple and robust method based on the Shannon entropy measurement that takes into account time causality by comparing neighboring values in a time series. The appropriate symbol sequence arises naturally from the time series with no prior knowledge assumed [1].

Permutation entropy is calculated for a given time series $\{x_1, x_2, \dots, x_n\}$ as a function of the scale factor ϵ . In order to be able to compute the permutation of a new time vector X_j , $S_t = [X_t, X_{t+1}, \dots, X_{t+m-1}]$ is generated with the embedding dimension m and then arranged in increasing order: $[X_{t+j_1-1} \leq X_{t+j_2-1} \leq \dots \leq X_{t+j_n-1}]$. Given m different values, there will be

$m!$ possible patterns π , also known as permutations. If $f(\pi)$ represents its frequency in the time series, its relative frequency is $p(\pi) = f(\pi) / (L/s - m + 1)$. The permutation entropy is then defined as:

$$PE = - \sum_{i=1}^{m!} p(\pi_i) \ln p(\pi_i), \quad (10)$$

Summing up, permutation entropy refers to the local order structure of the time series, which can give a quantitative measure of complexity for dynamic time series. This calculation depends on the selection of the m parameter, which is strictly related to the length n of the analyzed signal. For example, Bandt and Pompe [38] suggested the use of $m = 3, \dots, 7$ following always this rule:

$$m! < n. \quad (11)$$

If m is too small (smaller than 3), the algorithm will work incorrectly because it will only have few different states for recording, but it depends on the data. When using long signals a larger value of m is preferable but it would require larger computational times.

Multiscale Entropy (MSE) was proposed by Costa *et al.* [4] and has been shown to be a robust method for analyzing structural effects at multiple time scales present in complex real data. As Morabito *et al.* explain in their application to EEG signal processing [2], usual coarse-graining procedure can be implemented as follows:

- (i) From the original time series, multiple successive coarse-grained versions are extracted by $y_j(\varepsilon)$, where ε is the scale factor. Each element of the coarse-grained time series is calculated as:

$$y_j(\varepsilon) = \frac{1}{\varepsilon} \sum_{i=(j-1)\varepsilon+1}^{j\varepsilon} x_i \quad (12)$$

- (ii) For each scaled series, the PE is calculated.

In Morabito *et al.* [2], multivariate multiscale permutation entropy (MMSPE) is used as a methodology that can integrate information loss related to relevant cross-channel variability and to channel correlation. These authors refer to Keller and Laufer [39] that didn't take into account variation over multiple scales. In previous works, Ahmed and Mandic [40] analyzed a similar proposal, multivariate multiscale entropy (MMSE), for sample entropy.

Thus the cross-channel complexity with $f_s = \frac{1}{T}$, representing the multivariate PE (MPE), can be calculated for all time $s \in [f_s T - m]$ as the permutation entropy (PE) of p_j :

$$H_{MPE}(s) = \sum_{j=1}^{d!} p_j \log_2 p_j. \quad (13)$$

And the MMSPE algorithm is implemented according to the two following steps:

- (i) Different time scales of increasing length are defined by coarse-graining the original multivariate time series, *i.e.*, $\{x_i, t\}$, for $i = 1, \dots, c$ (where c is the number of channels) and for $t = 1, \dots, n$ (where n is number of samples in each time series). For a scale factor ε , the elements of the multivariate coarse-grained time series can be derived as:

$$y_{i,j}(\varepsilon) = \frac{1}{\varepsilon} \sum_{t=(j-1)\varepsilon+1}^{j\varepsilon} x_{i,t} \text{ for } i = 1, \dots, c, \text{ and, } 1 \leq j \leq \frac{n}{\varepsilon}. \quad (14)$$

- (ii) Calculate the multivariate permutation entropy, MPE, for each coarse-grained multivariate $y_{i,j}(\varepsilon)$ and all variants of average.

3.3.3. Feature Sets

In the experimentation, the following feature sets will be used (see Abbreviations Section for a complete list of acronyms and their meanings):

1. Linear features set (LF), the set described in Section 3.3.1
2. Non-linear features sets (NLF) that consist of LF and the features described in Section 3.3.2: linear features + Shannon entropy (SE), linear features + Approximate Entropy (ApEn), linear features + Sample Entropy (SmEn) and linear features + permutation entropy (PE).
3. Set after selection of features by ANOVA: selection of linear features (SLF), Selection of linear features + Shannon Entropy (SSE), Selection of linear features + Approximate Entropy (SApEn), Selection of linear features + Sample Entropy (SSmEn) and Selection of linear features + permutation entropy (SPE).

3.4. Automatic Selection of Features by ANOVA

In a next stage the feature set will be optimized and the best feature with regard to common significance level will be selected automatically. Thus, automatic feature selection is performed by an ANOVA one-way test [41]. This test analyzes the p -value under the null hypothesis that all samples in a matrix X are drawn from populations with the same mean. If p is near zero, it casts doubt on the null hypothesis, and suggests that at least one mean is significantly different from the others. Common significance level will be less than 0.05 to select a feature as discriminative. In the box-plot graphic the columns of X suggests the size of the F -statistic and the p -value. Large differences in the centre lines of the boxes correspond to large values of F and correspondingly small values of p and therefore useful feature for discrimination tasks. Then, in order to confirm the selection, a *multiple comparison test* of the means of the group is performed by the MATLAB (R2014) [41] function *multcompare*.

3.5. Modeling and Automatic Classification

The main goal of the present work comprises feature search in handwriting aiming at preclinical evaluation in order to define tests for ET diagnosis. These features will define the control group (CR group) and the essential tremor group (ET group). A secondary goal is the optimization of computational cost (measured as execution time of the algorithms) with the aim of making these techniques useful for real-time applications in real environments. Thus, automatic classification will be modeled taking this into consideration. We used three different classifiers based on WEKA (3.7) software suite [42]:

1. A Support Vector Machine (SVM) with polynomial kernel;
2. A Multi Layer Perceptron (MLP) with number of units in the hidden layer given (NNHL) by $= \max(\text{Attribute/Number} + \text{Classes/Number})$ and training step (TS) = NNHL · 10; and
3. A k Nearest Neighbor (k -NN) k -NN algorithm.

The results were evaluated using Accuracy (Acc, in %) and Classification Error Rate (CER, in %) [7,42–44]. For training and validation steps, we used k -fold cross-validation with $k = 10$, where accuracy will be the average of the k iterations. Cross-validation is a robust validation method for variable selection [43]. Repeated cross-validation (as calculated by the WEKA environment) allows robust statistical tests. We also use the measurement automatically provided by WEKA “Coverage of cases” (0.95 level); that is, the confidence interval at 95% level. Despite the small sample size that could point to the use of Leave One Out Cross-Validation (LOOC), we were oriented to the selected k -fold cross-validation by the model number, the sample richness and previous works [14,22].

3.6. General Procedure of the Experimentation

In previous works, both linear and non-linear features (fractal dimension and Shannon entropy) have been used with good results [22]. However, the system was not able to detect subtle changes in the case of level tremor. Therefore, the main goal of these experiments was to examine the potential of other entropy algorithms, and to select the optimum entropy based features for automatic measurement of the degradation of the drawing of Archimedes' spiral with ET. All experiments will be oriented to subtle changes and prodromal stages.

In this work, the general procedure of the experimentation is divided into two phases: a feature selection phase and an optimization phase (Figure 6).

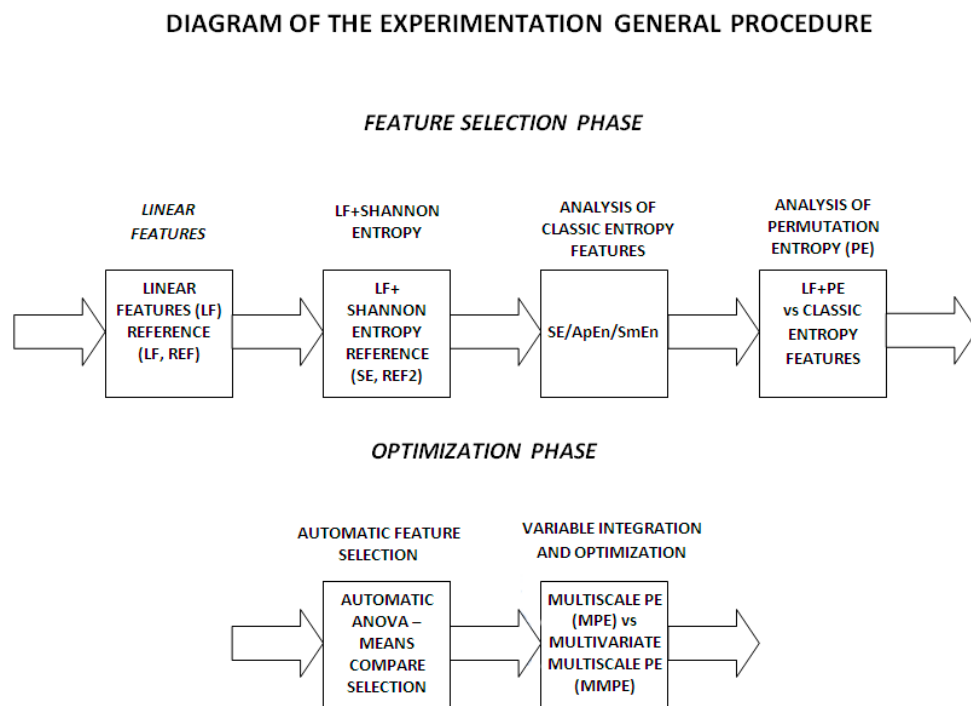


Figure 6. Diagram of the general procedure of the experimentation. The experimentation is divided into two phases. Firstly, an *entropy feature selection phase* is carried out across four steps: (i) analysis of classic linear features (this will be the reference); (ii) analysis of linear features and Shannon entropy; (iii) analysis of classic entropy features ApEn and SmEn and adjustment and selection of optimum parameters, m and r ; and (iv) analysis of previously adjusted classic entropy features against permutation entropy. Secondly, an *optimization phase* is carried out following the next two steps: (i) an automatic selection of the best linear and non-linear features by statistical-medical criteria based on ANOVA and multiple comparison test of group means; and (ii) an optimization analysis of the PE based features based on signal scale decomposition and the integration of signals by the MPE and the MMSPE algorithms.

Firstly, an *entropy feature selection* is carried out following the next four steps:

- (1) *Analysis of classic linear features.* An automatic classification is carried out in order to obtain the reference rates (LF) for linear features.
- (2) *Analysis of linear features and Shannon entropy.* A second reference (SE) is calculated integrating linear and the classic entropy feature, Shannon entropy.
- (3) *Analysis of classic entropy features.* ApEn and SmEn are analyzed and compared in order to adjust and select optimum parameters for the algorithms (m and r).

- (4) *Entropy based classic features vs. permutation entropy.* An analysis of the previous adjusted classic entropy features against permutation entropy is carried out in order to obtain the optimum parameters for permutation entropy.

Then, an *optimization* phase is carried out across two steps:

- (1) *Automatic feature selection.* An automatic feature selection is carried out based on statistical-medical criteria by ANOVA and a multiple comparisons test of the group means. Thus, only the linear and non-linear features with a p -value under a fixed threshold are selected and an optimum feature set is obtained.
- (2) *Optimization.* Finally, an optimization analysis of the PE based features is carried out. Two new algorithms are used based on: (1) scale analysis by multiscale permutation entropy (MPE); and (2) the integration of signals and scale analysis by the novel multivariate multiscale permutation entropy (MMSPE) algorithm. This last step is oriented to integrate signal correlations and to reduce even more the number of features for real-time system purposes.

4. Results and Discussion

The experimentation has been carried out with the balanced subset BIODARWO. In this section, we have used the automatic classification and modeling, the general procedure of experimentation and the feature sets described in Section 3.3.

4.1. Phase of Entropy Feature Selection

In a first stage the reference rates are calculated for both *linear features* (LF, 186 features) and a non-linear proposal that consist of *linear features and Shannon entropy* (SE, 198 features). Thus, integration of Shannon entropy outperforms LF reference for MLP and k -NN with similar results for SVM paradigm but with an increase in the number of features of around 7% (Table 2).

Table 2. Feature sets and feature number (FN). The number following the small letter is the value of that parameter.

	LF	SE	ApEn-m3	SmEn-m3	PE-m7t7	SLF	SSE	SAPEn-m3	SSmEn-m3	SPE-m7t7
FN	186	198	198	198	198	70	76	73	77	78

Then, an analysis of classic entropy features has been carried out. ApEn and SmEn are used and compared in order to adjust and select the optimum parameters (m and r). For both algorithms, $m = 2, 3$ and tolerance $r = 0.2$ have been evaluated. Figure 7 shows the obtained CER (%) values for the three paradigms with regard to the references LF and SE. ApEn cannot improve SE, but promising results are obtained for SmEn and $m = 3$, outperforming the rates for MLP and SVM. However, there is not a clear improvement in performance in the case of k -NN, the option with less computational cost with regard to the model generation and to the classification process.

Secondly, the integration of classic entropy based features is compared with permutation entropy features (PE) for different orders (m) and time delays (t). In our particular case and due to the length of the signals that have a mean of 3497 samples, m parameter was fixed up to $m = 7$. In this step, the previous best option, SmEn with $m = 3$, is evaluated with the references and different options of PE. Figure 8 shows the obtained CER (%) values for the three paradigms. PE based features improve the CER (%) rates in most of the cases. The best results are obtained for $m = 7$ and $t = 7$; this configuration improves the previous references with 198 features for MLP (15.69%) and SVM (17.65%) and it maintains a similar rate for k -NN. Good results are achieved also with k -NN with less computational cost (21.57%).

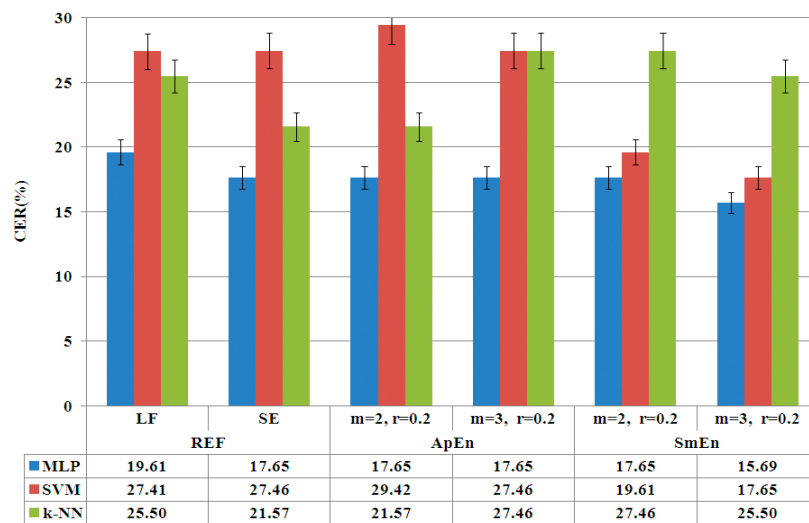


Figure 7. Analysis of classic entropy features. ApEn and SmEn are analyzed and compared in order to adjust and select the optimum parameters (m and r). The graphic presents CER (%) value for the three paradigms compared with the references for linear features and Shannon entropy (SE). The best results are obtained for $m = 3$ and SmEn, but there is not an improvement in the case of k -NN with a worse result with regard to SE [14].

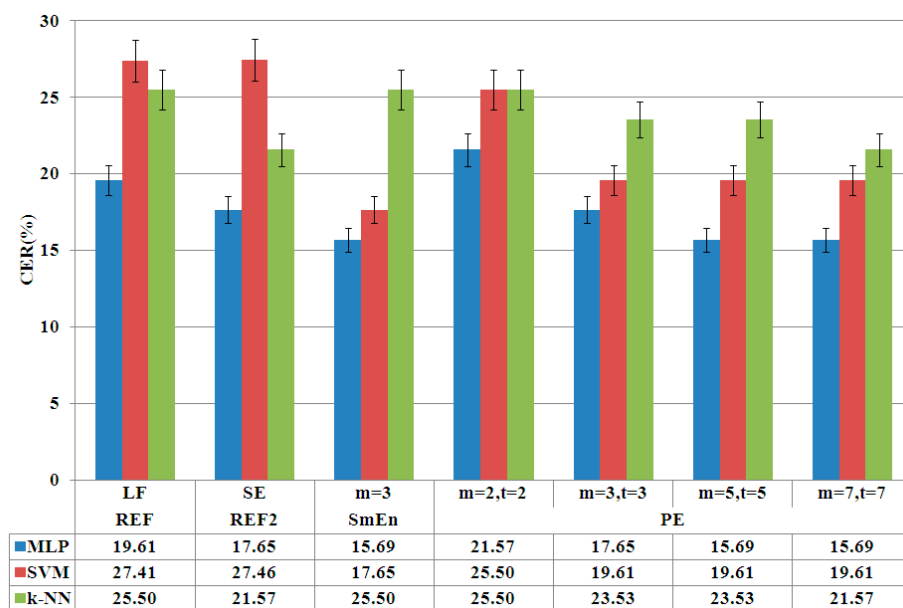


Figure 8. LF and classic entropy based features *vs.* SE and permutation entropy (PE). The best option, LF with SmEn for $m = 3$, is compared with the references and different options of LF and PE. The graphic presents CER (%) value for the three paradigms. The PE based features improve the CER (%) rates in most of the cases. The best results are obtained for $m = 7$ and $t = 7$, which improve the previous references for MLP and SVM and are equal to the k -NN system [14].

4.2. Optimization Phase

In the next stage, during the optimization phase an automatic feature selection is carried out based on statistical-medical criteria by ANOVA and multiple comparison test of the group means. In this automatic process medical and statistical criteria are combined and only the linear and non-linear features with p -value < 0.05 are selected. Thus, the number of features is reduced by around 60%

and an optimum feature set is obtained. Table 2 summarizes the results for the references and the previous best options. In order to show the values considered for the parameters of the algorithms, a number is added after the letter that identifies each parameter. Figure 9 shows an example of the statistical comparison for *pen-up* (in air) time. This feature is relevant with $p < 0.05$ as it can be seen in the analysis of the means of the groups CR and ET, which are significantly different.

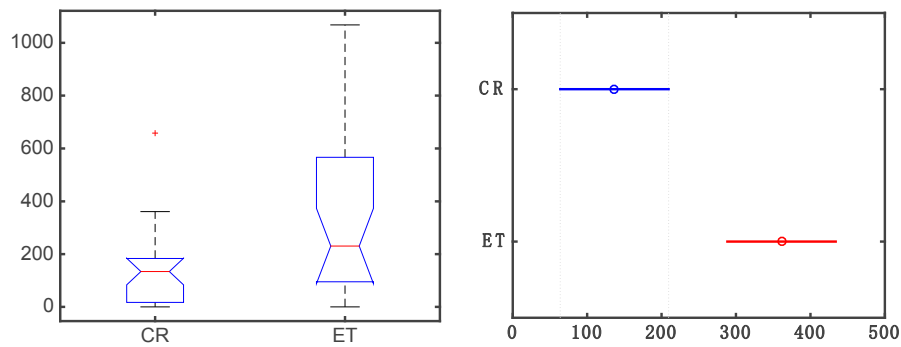


Figure 9. In the optimization phase an automatic feature selection is carried out based on statistical-medical criteria by ANOVA and the analysis of the means of the groups (p -value). In this process, only the linear and non-linear features with a $p < 0.05$ are selected, and an optimum feature set is obtained. The example shows the box-plot for the ANOVA test (left); and analysis of the group means for *pen-up* (in air) time (right). The means of CR and ET groups are significantly different.

After the optimization, an automatic classification is carried out with the selected feature sets in order to analyze the robustness of the generated sets. Figure 10 shows CER (%) for the references and the three paradigms (MLP, SVM and k -NN). There is a clear improved performance of the models after feature optimization for all the options. The best option is clearly MLP with the SPE-mt7t set (3.93%), and also SVM (7.85%) and k -NN obtain promising results for future real time developments (Table 3). Moreover, permutation entropy seems to be a powerful non-linear feature for modeling the non-linear dynamic of the system.

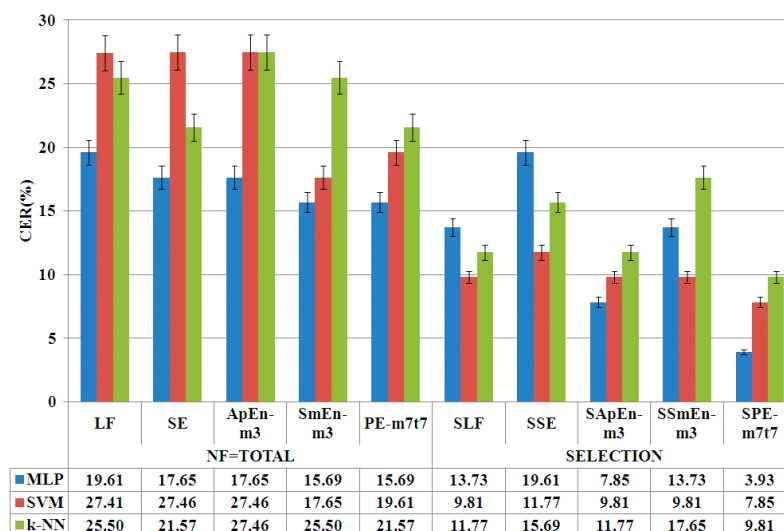


Figure 10. Automatic classification and obtained CER (%) for the references and the three paradigms with feature sets, before and after the automatic selection. The results show a clear improved performance of the models after feature optimization for all the options. The best option after the optimization is MLP with the SPE-m7t7 set (Selection of linear features + permutation entropy for $m = 7$ and $t = 7$).

Table 3. Results summary. CER (%) for the three paradigms with LF, SLF and the optimum feature set.

	MLP	SVM	k-NN
LF	19.61	27.41	25.50
SLF	13.73	9.81	11.77
SPE-m7t7	3.93	7.85	9.81

On the other hand, an analysis for classes has also been carried out. Figure 11 shows in detail the Accumulative Classification Error Rate (% ACCERR) for CR and ET as a measure of optimization impact by class. It can be seen that the optimization and the integration of the novel entropy algorithms improve the ACCERR in all the cases. The best option is SPE-m7t7 for the three paradigms, even for *k*-NN, which yields appropriate rates for both classes with less computational cost in modeling and classification phases. The computational cost is also reduced in both phases with regard to the original feature sets. SPE-m7t7 improves not only the global rates, but also the class rate even for the less powerful model: *k*-NN. Figure 12 shows the analysis for classes with regard to Acc (%) with similar results. SPE-m7t7 is confirmed as the best option.

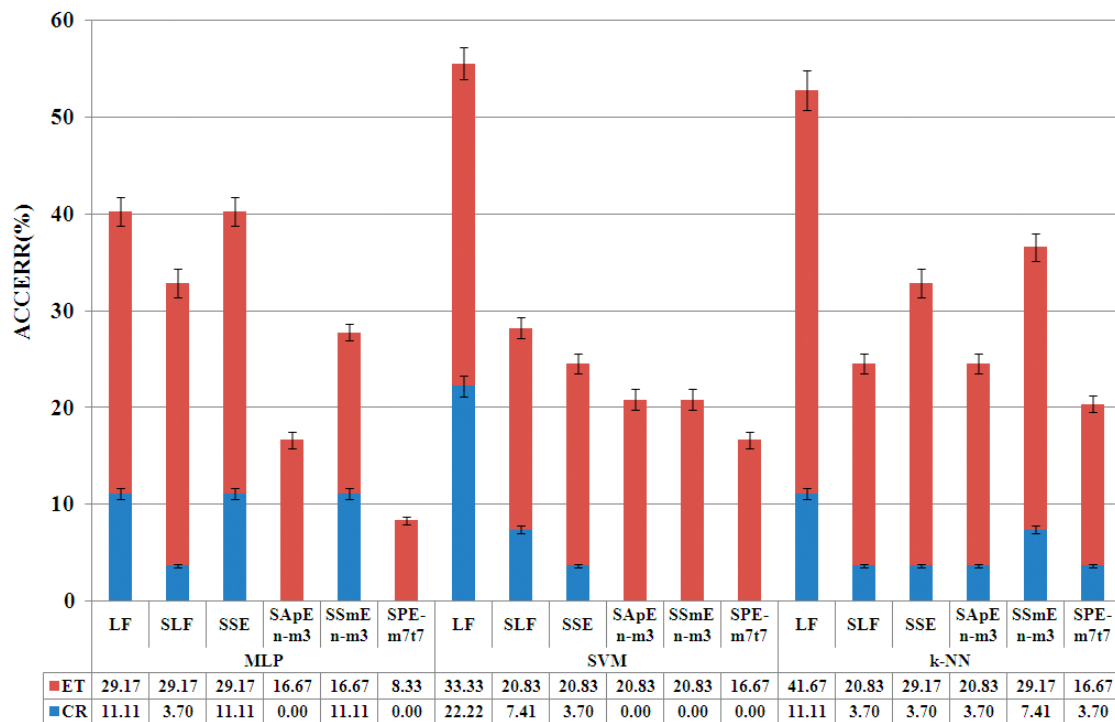


Figure 11. Analysis of the entropy based feature sets and impact on classes by Accumulative Classification Error Rate (% ACCERR). The optimization and the integration of the novel entropy algorithms improve the ACCERR in all the cases. The best option is SPE-m7t7 for the three paradigms even for *k*-NN, which presents appropriate rates for both classes.

In the final stage, an optimization analysis of the PE is carried out. This analysis is based on the signal scale decomposition (MPE) and the integration of signals (*X*, *Y*, pressure) by the novel MMSPE algorithms. This last step is oriented to integrating signal correlations and to further reducing the number of features for real-time system purposes. Figure 13 shows the effect of the introduction of different scales in the PE calculation. These differences in the scales could contribute to the detection of subtle changes in the signal that could be very useful for early ET detection.

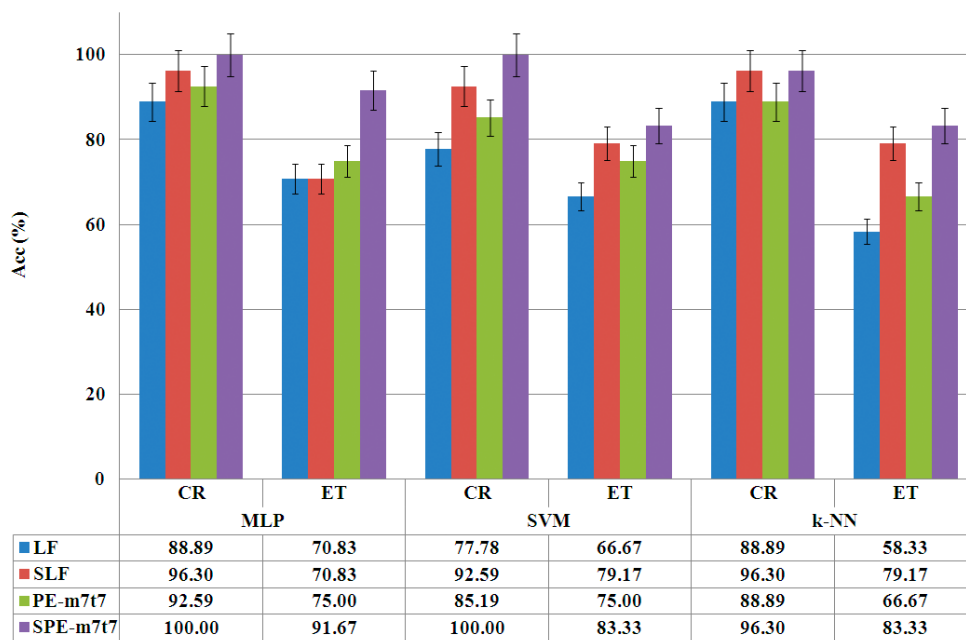


Figure 12. Accuracy (%) of classes for the paradigms for the references and the best options. SPE-m7t7 improves not only the global rates, but also the class rate even for the less powerful model: *k*-NN.

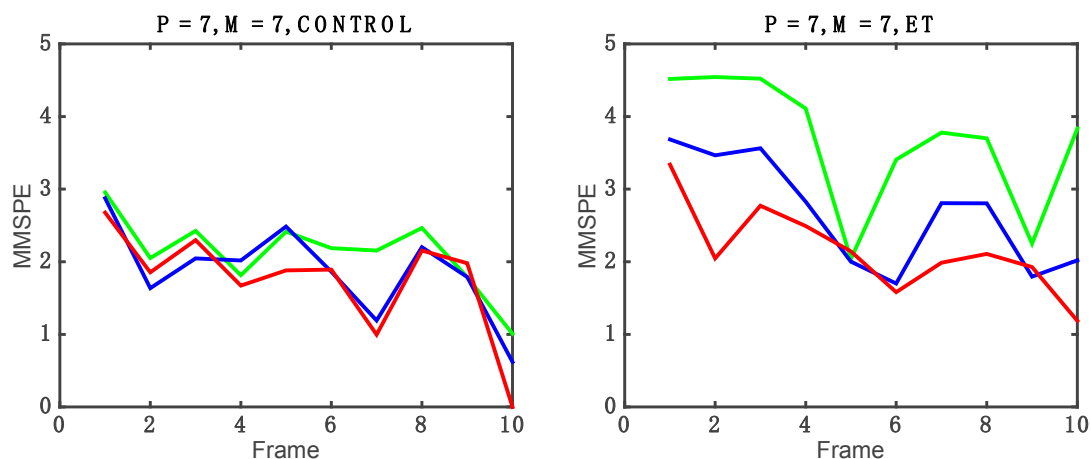


Figure 13. Detail of MMSPE (value: *y*-axis) across the time (frame: *x*-axis) of the drawing of the Archimedes’ spiral performed by a control subject (left) and a subject with essential tremor (right) for three scales ($\epsilon = 3$: green line, $\epsilon = 1$: blue line, $\epsilon = 2$: red line), $p = 7$ and $t = 7$.

These novel approaches provide very promising results. Figure 14 shows the obtained results. On the one hand, the SMPE approach with $\epsilon = 3$, $p = 7$ and $t = 7$, which includes three scales, outperforms SPE-m7t7 with MLP, and obtains similar rates to the ones of the other two paradigms but with an increase in the number of features (84). On the other hand, the option that integrates three signals with $\epsilon = 3$, $p = 7$ and $t = 7$, SMMSPE (76 features) set, decreases the number of features around 10% with regard to SMPE and 6% with regard to SPE-m7p7, but with similar CER for most of the paradigms. Only for MLP the results are worse due to a reduction in the relevant information of independent signals. Moreover, the computational cost decreases with regard to SPE not only in the training phase, but also during the classification process. The SMMSPE methodology seems to be very promising to analyze both interaction between different signals and subtle changes in ET subjects.

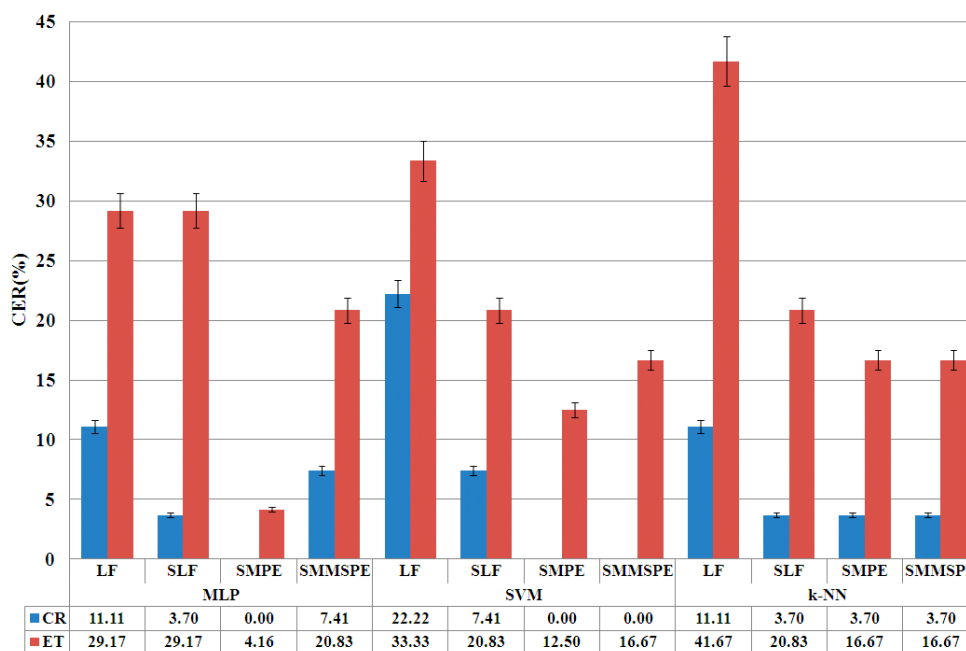


Figure 14. CER (%) for the paradigms with the references LF and SLF and the selected features set SMPE and SMMSPE.

Finally, the health specialists have noticed the relevance and usefulness of the system for selection of handwriting biomarkers for early diagnosis of ET.

5. Conclusions

This work focused on the selection of non-linear biomarkers from drawings and handwriting is part of a wider cross study on the diagnosis of essential tremor, which is developed in the Biodonostia Health Institute. Specifically, the main goal of the present work is the analysis of features in Archimedes' spiral drawing, one of the most used standard tests for clinical diagnosis of ET. In this sense, entropy based features have been added to a set of classic linear features (static and dynamic). Several entropy algorithms have been evaluated: Shannon entropy, approximate entropy, sample entropy and permutation entropy. The automatic analysis system consists of several Machine Learning paradigms, automatic features selection by ANOVA, and multiple comparison test (*multcompare*). The results are optimal even with a reduction of around 60% in the number of features after the selection process. The best option is MLP with permutation entropy. However, for real-time applications, MMSPE appears to be a promising methodology to analyze with less computational cost, both interactions between different signals and subtle changes in ET subjects. In future works, new non-linear features, entropy algorithms and automatic selection methodologies will be evaluated.

Acknowledgments: This work has been partially supported by the University of the Basque Country under project ref. UPV/EHU—58/14, SAIOTEK program and others from the Basque Government, the Spanish Ministerio de Ciencia e Innovación TEC2012-38630-C04-03, the University of Vic—Central University of Catalonia under the research grant R0904, INNPACTO program from the Spanish Government, and UPV/EHU Summer Courses Foundation.

Author Contributions: Karmele López-de-Ipiña, Jordi Solé-Casals, Marcos Faundez-Zanuy, Pilar M. Calvo, Unai Martinez De Lizarduy, Enric Sesa and Blanca Beitia have pre-processed the handwritten signals, and designed, developed and evaluated the system and tools. Alberto Bergareche, Patricia De La Riva and Jose F. Marti-Masso have designed and acquired the signals and collaborated in the system design and evaluation. All authors developed the system and participated writing of the manuscript, which has been read and approved by all of them

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CR	control group
ET	essential tremor
EPT	electrophysiological test
fMRI	functional magnetic resonance imaging
MoCA	Montreal Cognitive Assessment
EMG	electromyography
AD	Alzheimer's disease
LF	linear features (reference)
SLF	selection of linear features
NLF	non-linear features
SE	linear features + Shannon entropy (reference)
ApEn	linear features + Approximate Entropy
SmEn	linear features + Sample Entropy
EEG	electroencephalography
PE	permutation entropy
MPE	multi scale permutation entropy
MSE	multiscale entropy
MMSE	multivariate multiscale entropy
MMSPE	multivariate multiscale permutation entropy
MLP	Multi Layer Perceptron
NNHL	number of hidden layers units
SVM	Support Vector Machine
CER	classification error rate
ACCERR	Accumulative Classification Error Rate
ACC	Accuracy
SE	linear features + Shannon entropy
SSE	Selection of linear features + Shannon Entropy
SAPEn	Selection of linear features + Approximate Entropy
SSmEn	Selection of linear features + Sample Entropy
SPE	Selection of linear features + Permutation Entropy
SMPE	Selection of linear features + Multi scale Permutation Entropy
SMMSPE	Selection of linear features + Multivariate Multiscale Permutation Entropy

References

- Zanin, M.; Zunino, L.; Rosso, O.A.; Papo, D. Permutation entropy and its main biomedical and econophysics applications: A review. *Entropy* **2012**, *14*, 1553–1577. [[CrossRef](#)]
- Morabito, F.C.; Labate, D.; La Foresta, F.; Bramanti, A.; Morabito, G.; Palamara, I. Multivariate Multi-Scale Permutation Entropy for Complexiy Analysis of Alzheimer's Disease EEG. *Entropy* **2012**, *14*, 1186–1202. [[CrossRef](#)]
- Eguiraun, H.; López-de-Ipiña, K.; Martinez, I. Application of, Entropy and Fractal Dimension Analyses to the Pattern Recognition of Contaminated Fish Responses in Aquaculture. *Entropy* **2014**, *16*, 6133–6151. [[CrossRef](#)]
- Costa, M.; Goldberger, A.; Peng, C.K. Multiscale entropy analysis of biological signals. *Phys. Rev.* **2005**, *71*. [[CrossRef](#)] [[PubMed](#)]
- Louis, E.D.; Vonsattel, J.P. The emerging neuropathology of essential tremor. *Mov. Disord.* **2007**, *23*, 174–182. [[CrossRef](#)] [[PubMed](#)]
- Faundez-Zanuy, M.; Hussain, A.; Mekyska, J.; Sesa-Nogueras, E.; Monte-Moreno, E.; Esposito, A.; Chetouani, M.; Garre-Olmo, J.; Abel, A.; Smekal, Z.; *et al.* Biometric Applications Related to Human Beings: There Is Life beyond Security. *Cogn. Comput.* **2013**, *5*, 136–151. [[CrossRef](#)]
- López-de-Ipiña, K.; Alonso, J.B.; Solé-Casals, J.; Barroso, N.; Faundez-Zanuy, M.; Travieso, C.; Eca-Torres, M.; Martinez-Lage, P.; Martinez-de-Lizardui, U. On the selection of non-invasive methods based on speech analysis oriented to Automatic Alzheimer Disease Diagnosis. *Sensors* **2013**, *13*, 6730–6745. [[CrossRef](#)] [[PubMed](#)]
- Laske, C.; Sohrobi, H.R.; Frost, S.M.; López-de-Ipiña, K.; Garrard, P.; Buscem, M.; Dauwels, J.; Soekadar, S.R.; Mueller, S.; Linnemann, C.; *et al.* Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimer Dement.* **2015**, *11*, 561–578. [[CrossRef](#)] [[PubMed](#)]

9. Sesa-Nogueras, E.; Faundez-Zanuy, M. Biometric recognition using online uppercase handwritten text. *Pattern Recognit.* **2012**, *45*, 128–144. [[CrossRef](#)]
10. Pullman, S.L. Spiral Analysis: A New Technique for Measuring Tremor with a Digitizing Tablet. *Mov. Disord.* **1998**, *13*, 85–89. [[CrossRef](#)] [[PubMed](#)]
11. WACOM. Available online: <http://www.wacom.com> (accessed on 10 May 2016).
12. Faundez-Zanuy, M. Online signature recognition based on VQ-DTW. *J. Pattern Recognit.* **2007**, *40*, 981–982. [[CrossRef](#)]
13. López-de-Ipiña, K.; Bergareche, A.; De La Riva, P.; Faundez-Zanuy, M.; Calvo, P.M.; Roure, J.; Sesa-Nogueras, E. Automatic non-linear analysis of non-invasive writing signals, applied to essential tremor. *J. Appl. Log.* **2015**. [[CrossRef](#)]
14. López-de-Ipiña, K.; Iturrate, M.; Calvo, P.M.; Beitia, B.; Garcia-Melero, J.; Bergareche, A.; De La Riva, P.; Marti-Masso, J.F.; Faundez-Zanuy, M.; Sesa-Nogueras, E.; *et al.* Selection of Entropy Based Features for the Analysis of the Archimedes' Spiral Applied to Essential Tremor. In Proceedings of the 4th IEEE International Work Conference on Bioinspired, Intelligence, Donostia, Spain, 9–12 June 2015; pp. 157–162.
15. Ortega-Garcia, J.; Gonzalez-Rodriguez, J.; Simon-Zorita, D.; Cruz-Llanas, S. From Biometrics Technology to Applications Regarding Face, Voice, Signature and Fingerprint Recognition Systems. In *Biometric Solutions for Authentication in an E-World*; Kluwer Academic Publishers: Berlin, Germany, 2002; pp. 289–337.
16. The Montreal Cognitive Assessment (MoCA). Available online: <http://www.mocatest.org/> (accessed on 10 May 2016).
17. Faundez-Zanuy, M.; Sesa-Nogueras, E.; Roure-Alcobé, J.; Garré-Olmo, J.; López-de-Ipiña, K.; Solé-Casals, K. Online Drawings for Dementia Diagnose: In-Air and Pressure Information Analysis. In *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*; Springer International Publishing: New York, NY, USA, 2014; pp. 567–570.
18. Neils-Strunjas, J.; Groves-Wright, K.; Mashima, P.; Harnish, S. Dysgraphia in Alzheimer's disease: A review for clinical and research purposes. *J. Speech Lang. Hear. Res.* **2006**, *49*, 1313–1330. [[CrossRef](#)]
19. Phillips, J.G.; Ogeil, R.P.; Müller, F. Alcohol consumption and handwriting: A kinematic analysis. *Hum. Mov. Sci.* **2009**, *28*, 619–632. [[CrossRef](#)] [[PubMed](#)]
20. Foley, R.G.; Miller, A.L. The effects of marijuana and alcohol usage on handwriting. *Forensic Sci. Int.* **1979**, *14*, 159–164. [[CrossRef](#)]
21. Tucha, O.; Walitza, S.; Mecklinger, L.; Stasik, D.; Sontag, T.; Lange, K.W. The effect of caffeine on handwriting movements in skilled writers. *Hum. Mov. Sci.* **2006**, *25*, 523–535. [[CrossRef](#)] [[PubMed](#)]
22. Rosenblum, S.; Parush, S.; Weiss, P.L. The in Air Phenomenon: Temporal and Spatial Correlates of the Handwriting Process. *Percept. Mot. Skills* **2003**, *96*, 933–954. [[CrossRef](#)] [[PubMed](#)]
23. Sadikov, A.; Groznic, V.; Žabkar, J.; Mozina, M.; Georgiev, D.; Pirtosek, Z.; Bratko, I. Parkinson Check smart phone app. In *Frontiers in Artificial Intelligence and Applications*; IOS Press: Amsterdam, The Netherlands, 2014; Volume 263, pp. 1213–1214.
24. Georgiev, D.; Groznic, V.; Sadikov, A.; Mozina, M.; Guid, M.; Kragelj, V.; Bratko, I.; Ribaric, S.; Pirtosek, Z. Digitalised spirography and clinical examination based decision support system of differentiating between tremors. *Eur. J. Neurol.* **2012**, *19*, 298.
25. Bolle, R.; Pankanti, S. *Biometrics, Personal Identification in Networked Society*; Jain, A., Ed.; Kluwer Academic Publishers: Norwell, MA, USA, 1998.
26. Faundez-Zanuy, M. Privacy issues on biometric systems. In *IEEE Aerospace and Electronic Systems Magazine*; IEEE Xplore: New York, NY, USA, 2005; Volume 20, pp. 13–15.
27. Gray, R.M. *Entropy and Information Theory*; Springer: Berlin/Heidelberg, Germany, 1990.
28. Brissaud, J.B. The meaning of entropy. *Entropy* **2005**, *7*, 68–96. [[CrossRef](#)]
29. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
30. Shannon, C.E.; Weaver, W. *The Mathematical Theory of Communication*; University of Illinois Press: Champaign, IL, USA, 1949.
31. Richman, J.S.; Moorman, J.R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* **2000**, *278*, 2039–2049.
32. Pincus, S.M.; Huang, W.M. Approximate entropy, statistical properties and applications. *Commun. Stat. Theory Methods* **1992**, *21*, 3061–3077. [[CrossRef](#)]

33. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2297–2301. [[CrossRef](#)] [[PubMed](#)]
34. Dragomir, A.; Akay, Y.; Curran, A.K. Investigating the complexity of respiratory patterns during the laryngeal chemoreflex. *J. NeuroEng. Rehabil.* **2008**, *5*. [[CrossRef](#)] [[PubMed](#)]
35. Yentes, J.M.; Hunt, N.; Schmid, K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.* **2013**, *41*, 349–365. [[CrossRef](#)] [[PubMed](#)]
36. Ramanand, P.; Nampoori, V.P.; Sreenivasan, R. Complexity quantification of dense array EEG using sample entropy analysis. *J. Integr. Neurosci.* **2004**, *3*, 343–358. [[CrossRef](#)] [[PubMed](#)]
37. Solé-Casals, J.; Vialatte, F.B. Towards Semi-Automatic Artifact Rejection for the Improvement of Alzheimer's Disease Screening from EEG Signals. *Sensors* **2015**, *15*, 17963–17976. [[CrossRef](#)] [[PubMed](#)]
38. Bandt, C.; Pompe, B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett.* **2002**, *88*. [[CrossRef](#)] [[PubMed](#)]
39. Keller, K.; Lauffer, H. Symbolic analysis of high-dimensional time series. *Int. J. Bifurc. Chaos* **2003**, *13*, 2657–2668. [[CrossRef](#)]
40. Ahmed, M.U.; Mandic, D.P. Multivariate multiscale entropy. In *IEEE Signal Processing Letters*; IEEE Xplore: New York, NY, USA, 2012; Volume 19, pp. 91–95.
41. Mathworks. Available online: <http://www.mathworks.com> (accessed on 10 May 2016).
42. Weka. Available online: <http://www.cs.waikato.ac.nz/ml/weka> (accessed on 10 May 2016).
43. Picard, R.; Cook, D. Cross-Validation of Regression Models. *J. Am. Stat. Assoc.* **1984**, *79*, 575–583. [[CrossRef](#)]
44. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufmann: San Francisco, CA, USA, 2005.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

B3 ERANSKINA

Advances on automatic speech analysis for
early detection of Alzheimer Disease: a
non-linear multi-task approach (accepted)

Advances on automatic speech analysis for early detection of Alzheimer Disease: a non-linear multi-task approach

K. López-de-Ipiña^{*1a}, U. Martínez-de-Lizarduy^{*b}, P. M. Calvo^{2a}, Jiri Mekyska^c, B. Beitia^d, N. Barroso^{3a},
A. Estanga^{1e}, M. Tainta^{2e} and M. Ecay-Torres^{3e}

^aDepartment of Systems Engineering and Automation, Faculty of Engineering-Gipuzkoa, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Spain; ^bDepartment of Electronic Technology, Faculty of Engineering-Gipuzkoa, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Spain; ^cDepartment of Telecommunications, Brno University of Technology, Brno, Czech Republic; ^dDepartment of Applied Mathematics, Faculty of Engineering-Vitoria, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Donostia, Spain; ^eDepartment of Neurology, CITA-Alzheimer Foundation, Donostia, Spain



Abstract: Nowadays proper detection of cognitive impairment has become a challenge for the scientific community. Alzheimer's Disease (AD), the most common cause of dementia, has a high prevalence that is increasing at a fast pace towards epidemic level. In the not-so-distant future this fact could have a dramatic social and economic impact. In this scenario, an early and accurate diagnosis of AD could help to decrease its effects on patients, relatives and society. Over the last decades there have been useful advances not only in classic assessment techniques, but also in novel non-invasive screening methodologies. Among these methods, automatic analysis of speech –one of the first damaged skills in AD patients– is a natural and useful low cost tool for diagnosis. In this paper a non-linear multi-task approach based on automatic speech analysis is presented. Three tasks with different language complexity levels are analyzed, and promising results that encourage a deeper assessment are obtained. Automatic classification was carried out by using classic Multilayer Perceptron (MLP) and Deep Learning by means of Convolutional Neural Networks (CNN) (biologically-inspired variants of MLPs) over the tasks with classic linear features, perceptual features, Castiglioni fractal dimension and Multiscale Permutation Entropy. Finally, the most relevant features are selected by means of the non-parametric Mann-Whitney U-test.

Keywords: 6 to 8 keywords must be provided.

1. INTRODUCTION

We are facing a considerable increase in diseases due to the dramatic aging of the world population and the proliferation of unhealthy lifestyles [1]. In this scenario Alzheimer's Disease (AD), the most common cause of dementia, has a high prevalence that in the long term could reach epidemic level leading to high economic and social impacts all over the world. The diagnosis of AD is carried out by several medical fields, and the definitive diagnosis requires the demonstration of the typical AD pathological changes at autopsy [2][3][4]. Many AD patients receive their diagnosis at a mild-moderate dementia stage. Identifying cognitive decline at an early stage has become increasingly important to physicians. AD patients develop severe disability and sometimes even full dependence.

Mild Cognitive Impairment (MCI) extends back to the first manifestations of the neurodegenerative diseases when the person is able to continue his previous routines and activities [5]. Amnesic MCI typically is considered the prodromal or early AD. Sometimes is still tricky to identify significant cognitive complaints. It is especially challenging for non-specialists and relatives to identify early AD or MCI [6][7]. Early detection of cognitive symptoms would be useful for a better management and could open a new scenario for potential interventions. In this sense, an early and accurate diagnosis of this disease could help to decrease its effects on patients, relatives and society.

This early clinical changes should be consistent and reinforced with the complementary evaluations. Over the last decades there have been useful advances not only in classic assessment techniques such as neuroimaging, but also in novel screening strategies for preliminary assessment or monitoring during the evolution of the disease.

* Department of Systems Engineering and Automation and Department of Electronic Technology, Faculty of Engineering-Gipuzkoa, Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU), Europa Plaza 1, Donostia 20018, Spain; Tel/Fax: +34-943-018-667, +34-943-033-110; E-mails: {karmele.ipina, unai.martinezdelizarduy}@ehu.eus

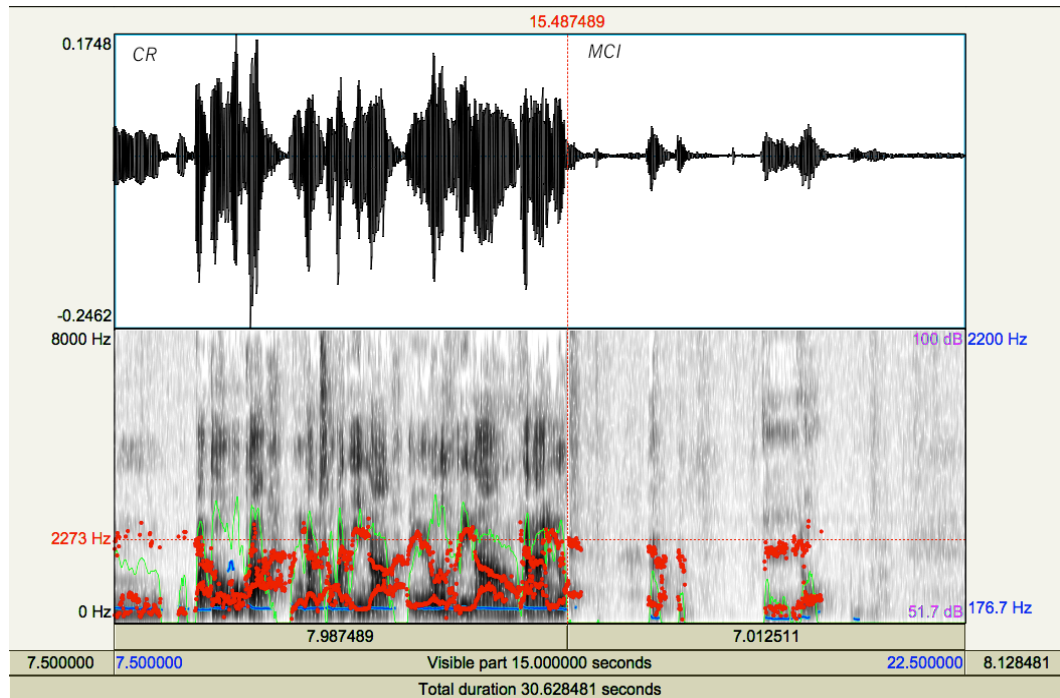


Fig. (1). Task-restricted clinical test, categorical fluency task (*animal naming*, AN). On the left a control subject, on the right a person with MCI. The Pitch is represented in blue, intensity in green, and formants in red (Praat [20]).

In this framework, the cost and technology requirements of some medical examinations (e.g. fMRI, PET) make it impossible to apply such biomarkers to all patients or require invasive procedures (e.g. lumbar puncture). Considering these problems, non-invasive intelligent techniques of diagnosis may become valuable tools for early detection of diseases that can complement diagnosis and/or monitor its progress. These methodologies are easy to apply, low cost and can be introduced in the standard medical protocols without altering daily tasks. Therefore, they are capable of yielding information easily, quickly and inexpensively [8][9]. Thus, non-technical staff in the usual environment of the patient could use these methodologies without altering or blocking the patients' abilities because they are not perceived as stressful tests by patients.

Among non-invasive methods, automatic analysis of speech—which involves verbal communication, one of the first damaged skills in MCI/AD patients—provides a natural and friendly powerful tool. For human beings, verbal communication is one of the most important means of expression during social interaction in daily life. It is a complex process which involves a wide range of cognitive abilities that allows us to produce sounds with an embedded message. The biosignal contains our desires, ideas and/or emotions, a piece of our being. During that process the brain has to manage knowledge, memory, language and semantic information. Moreover, these biological samples have the advantage of being easily recordable without invasive and expensive equipment. In the case of Alzheimer's Disease, the

deterioration of spoken language immediately affects the patient's ability to naturally interact with his/her social environment, and it is usually also accompanied by alterations in the emotional response [2][3][4].

Cognitive assessment involves a wide range of tests in order to quantify and distinguish early stages of an abnormal aging process [10]. During this assessment, language and speech are important components of the analysis as they are key factors of the human communication process. Thus, the deterioration of communication skills is measured by tests taking into account elements such as memory impairment, naming, recalling, verbal fluency and/or emotional management [2][3][4]. In recent novel research studies highly complex scenarios are analyzed by non-restricted language tasks such as spontaneous speech, interviews or free conversation [11][12].

For example, conventional task-restricted tests analyze specific communication components (naming or verbal fluency), but essential communication abilities that could be relevant in early diagnosis are not taken into account, e.g. emotional factors or spontaneous production of speech. Moreover, when the individual is for example highly educated, task-restricted tests could often be unable to detect subtle damages because the patient has enough resources to avoid mild naming problems, and thus the disease remains hidden. In this sense, a valid social marker such as the proportion of spoken words uttered during free conversation or spontaneous speech can indicate early signs of the disease.

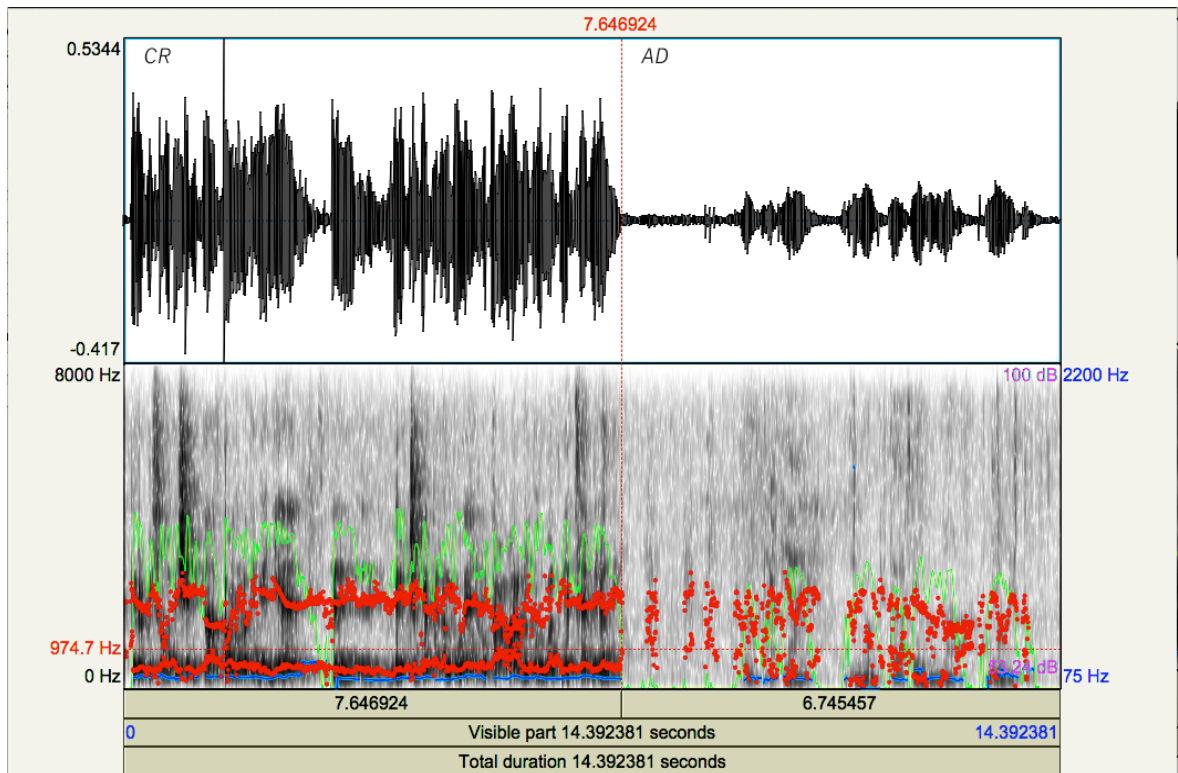


Fig. (2). Task-restricted clinical test, *picture description* task (PD). On the left a control subject, on the right a person with AD. The Pitch is represented in blue, intensity in green, and formants in red (Praat [20]).

These new biomarkers should be sensitive to transitions from normal cognition to MCI, and should be easily integrated with other biomarkers to identify at-risk individuals [11][12][13]. Other components such as the emotional response and the management of emotions are also communication elements that are impaired in early stages of the disease, and they could be valuable analysis factors as well [9][14][15].

The work presented in this paper on selection of biomarkers from speech for the detection of AD is part of a wider cross-study for the diagnosis of Alzheimer. In this work, automatic analysis of speech is applied specifically in a non-linear multi-task approach. The information is extracted from three tasks with different levels of language complexity, and is compared among them: a *categorical verbal fluency task*, a *picture description* task and a *spontaneous speech* task. This research study opens new work-lines towards deeper assessment and monitoring of cognitive diseases.

The hybrid assessment methodology is based on statistical tests (*Mann-Whitney U-test*) and an automatic classification which is carried out by using classic Multilayer Perceptron (MLP) and Deep Learning by means of Convolutional Neural Networks (CNN), which are biologically-inspired variants of MLPs. Speech samples are modeled by classic linear, perceptual and non-linear features.

The paper is organized as follows. Materials are described in section 2. Section 3 presents the used methods. The results and discussion are summarized in section 4, and finally concluding remarks are drawn in section 5.

2. MATERIALS

Three tasks with different levels of language complexity have been used: categorical fluency task (*animal naming*, AN, or *animal fluency* task), *picture description* (PD) task and *spontaneous speech* (SS) task. AN task and PD task have been recorded in a clinical environment and SS task in a domestic environment. The participants are different in the three tasks. PD task and SS task analyze AD patients vs. individuals from a control group (CR), and AN task analyzes MCI patients vs. individuals from CR. All the work was performed strictly following the ethical guidelines of the organizations involved in the project.

2.1 Task of *Categorical Verbal Fluency* (CVF), *animal naming* task

The task of *Categorical Verbal Fluency* (CVF, *animal naming*, AN), or *animal fluency* task, is a test that measures and quantifies the progression of cognitive impairment in neurodegenerative diseases [16], (Fig. (1), [20]). It is widely used to assess language skills, semantic memory and executive functions [17].

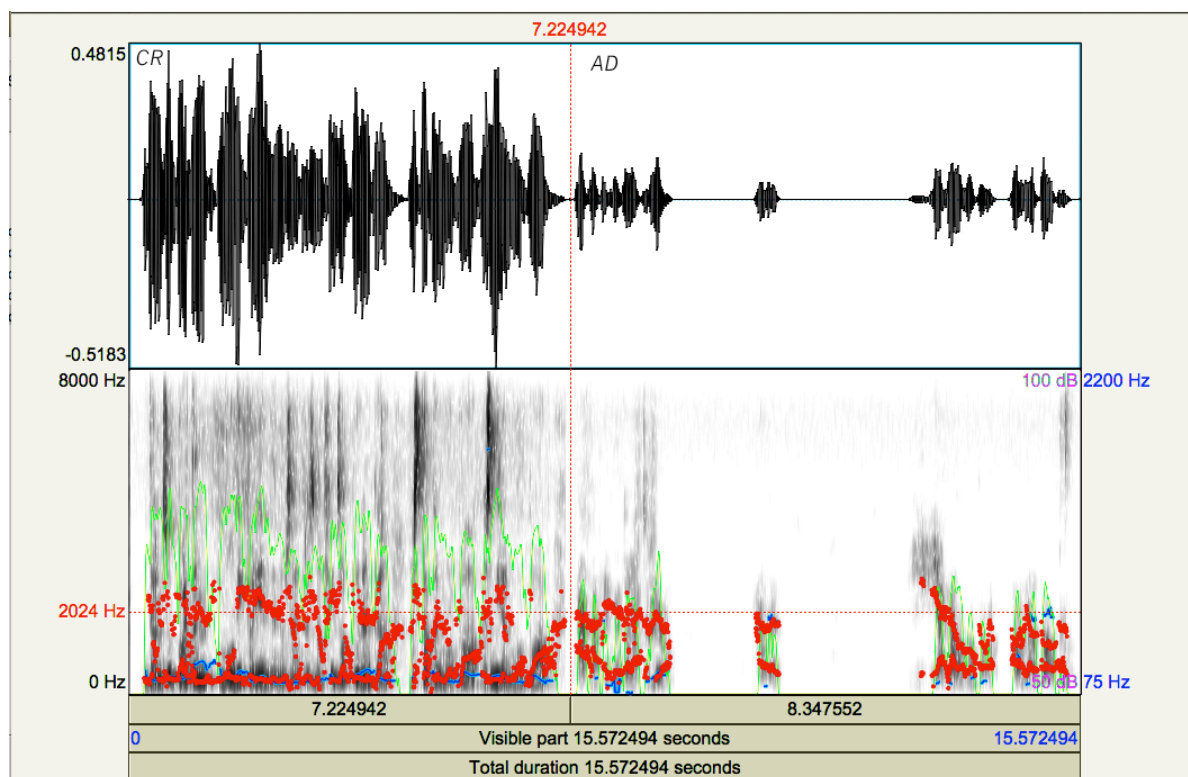


Fig. (3). Non-restricted domestic test, Spontaneous Speech task (SS). On the left a control subject, on the right a person with AD. The Pitch is represented in blue, intensity in green, and formants in red (Praat [20]).

During the CVF task the interviewer asks the patient to list in one minute all the names she/he knows of a category, animals in our case. In clinical practice, only the total number of items listed is recorded. However, more specific indexes could be defined, which could detect subtle changes in the cognitive status of the patient, and thus achieve a more specific discrimination between different degenerative diseases [2][18].

The sample consists of 187 healthy people and 38 patients with MCI belonging to the cohort of *Gipuzkoa-Alzheimer Project* (PGA [19]) of the *CITA-Alzheimer Foundation* (Fig. (1), [20]). The PGA is a longitudinal study focusing on the search for biomarkers of function, brain structure, state of health and risk factors for early diagnosis of Alzheimer's disease. The main objective of PGA is to characterize pre-clinical and prodrome Alzheimer's disease phases. Inclusion criteria based on the characteristics of each group are listed below. The study is carried out in a clinical environment.

1) Control group (CR)

- Men and women of middle age or elderly, between 39 and 79 years inclusive who sign the informed consent to participate in the PGA.

- Lack of memory complaints.

- Global cognitive function preserved with the highest score of 24 in the Mini Mental State Examination (MMSE) [21].

- Performance within the normal range in the memory test. Score greater than 6 in all indexes of Free and Cued Selective Reminding Test by Buschke (FCSRT) [22].

- Autonomy in daily activities. Less than 6 in the scalar score of the functional activities questionnaire by Pfeffer [23].

2) MCI group

- Men and women aged between 39 and 79 years inclusive who sign the informed consent to participate in the PGA.

- Presence of memory complaints considered relevant by the assessor.

- Global cognitive function preserved. Highest score of 24 in the MMSE.

- Performance below the average adjusted by age and educational level in at least one of the tests that assesses memory (FCSRT), language (Boston Naming Test [24]), visual-spatial function (Judgment Line Orientation [25]), attention (Letter Number Sequencing [26]) and executive function (Trail Making Test, Part B [27]). Scalar score equal to or less than 6.

Tab. (1). Demographic data of the subsets selected for the experimentation

	Female	Male	Age-Range	Age-Mean	Age-SD
PGA-OREKA AN task-subset (CR/MCI)	36/21	26/17	39-74/42-79	56.73/57.15	7.8/8.9
MINI-PGA PD task-subset (CR/AD)	8/5	4/1	53-73/62-78	61.51/67.97	7.02/5.2
AZTIAHORE SS task-subset (CR/AD)	9/12	11/8	47-80/69-88	57.80/79.90	11.92/5.3

- Autonomy in daily activities. Less than 6 in the score of the functional activities questionnaire by Pfeffer [23].

Exclusion criteria are: the presence of dementia (DSM-IV and stage CDR \geq 1); significant history of neurological disease of any kind that might cause cognitive impairment or dementia; and history of diagnosed psychiatric illness that could cause dementia or cognitive impairment. People with mild psychiatric symptoms (anxiety, depression) are not excluded. Finally, a balanced database of 100 individuals has been created (PGA-OREKA). Demography data are detailed in Tab. (1). We detect statistically significant differences in age between patients with AD and MCI.

2.2 Picture description task

This pilot study that belongs to the Gipuzkoa-Alzheimer Project (PGA) and the Memory Clinic of the CITA-Alzheimer Foundation includes six subjects with diagnosis of AD and twelve healthy people (sub-database MINI-PGA), that is, the sample consisted of 18 people, and the demography data are detailed in Table 1. The task consists in the verbal description of a picture [28] (Fig. (2), [20]).

2.3 Spontaneous speech task

This study focusing on early detection of AD was carried out in a domestic environment, and its objective is to identify AD in the pre-clinical stage –before first symptoms–, and in the prodromic stage –with some very early symptoms but no dementia–. This research work is a complementary preliminary experiment that seeks to define thresholds for a number of biomarkers related to spontaneous speech. The feature search is designed for pre-clinical evaluation in order to define appropriate tests for AD diagnosis. The obtained data will complement the biomarker set of each person when diagnosing AD [9]. The database used in this task is independent from the PGA project.

In order to develop a new methodology applicable to a wide range of individuals of different sex, age, language, and cultural and social background, we have constructed a multicultural and multilingual database in English, French, Spanish, Catalan, Basque, Chinese, Arabic and Portuguese, that comprises video recordings of 50 healthy people (12 hours) and 20 AD patients (8 hours), i.e., patients with a prior diagnosis of Alzheimer’s Disease. The age span of the individuals in the full database is 20-98 years. This database is called AZTIAHO [9] (Fig. (3), [20]). Within this database we have defined a sub-database explained in the next paragraphs. Table 1 shows the demographic data.

The recordings consisted of videos of spontaneous speech where people tell pleasant stories, recount pleasant feelings,

or interact with each other in friendly conversations. The atmosphere of the recordings was relaxed and non-invasive. The recording times of the AD group are shorter because AD patients find speech more awkward than healthy individuals: they speak more slowly, with longer pauses, and they spend more time looking for the correct word and dealing with speech disfluencies or broken messages. In the advanced stage of the disease, they find this effort tiring and often want to stop the recording. We acceded to their requests.

The video was processed and the audio was extracted in wav format (16 bits and 16 kHz). The first step was to remove non-analyzable events: laughter, coughing, short hard noises, and segments where speakers overlapped. Next, background noise was removed by means of denoiser adaptive filtering. After the pre-processing around 80% of the material from the control group and 50% of the material from the AD group remained suitable for further analysis. The complete speech database consists of about 60 minutes for the AD group and about 9 hours for the control group. Hereafter, the speech was divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers, yielding a database of approximately 600 segments of spontaneous speech, (Fig. (3), [20]).

Finally, in order to perform our experiments a subset was selected, balanced with regard to the age of the participants and the emotional response level. The subset consists of 20 subjects (9 women and 11 men) from the control group and 20 AD patients (12 women and 8 men). This subset of the database is called AZTIAHORE. The demography data are detailed in Table 1.

3. METHODS

3.1 Speech feature extraction

This work implies the use of some of the most common speech features for both differentiation between healthy and pathological speech [11][29][30] and discrimination through human perception. Most of them are usually well known in the field of speech signal processing, and therefore for each parameter a reference with a deeper description and further information will be provided. After the database collection, the voice analyzing approach will integrate two kinds of features: linear and non-linear. Moreover, this approach is based on the description of speech pathologies with regard to phonation, articulation, speech quality, human perception and complex dynamics of the system. In the next subsections the analyzed features are described.

3.2 Classic speech features

Classic speech features comprise:

- Spectral domain features: harmonicity, harmonic to noise ratio (HNR), noise to harmonic ratio (NHR), pitch, jitter, shimmer, amplitude perturbation quotient (APQ), spectrum centroid and formants and their variants (mean, median, min, max, mode, std) [11][29][31].
- Time domain features: voiced/unvoiced segments, breaks, zero-crossing rate (ZCR)[11][29][32] and their variants.

3.3 Classic perceptual features

Classic perceptual features are among others:

- Mel Frequency Cepstrum Coefficients (MFCC): Human ear acts as filters that only concentrates on certain frequency components. These filters are non-uniformly spaced on the frequency axis. There are more filters in the low frequency region and a lower amount of filters in the high frequency region. In this sense, Mel-Frequency analysis and specifically Mel Frequency Cepstrum Coefficients (MFCC) are oriented to simulate this kind of performance [29][33][34].
- Energy, intensity, short time energy [29][33][34].

These features are sometimes extended by their first and second order regression coefficients, Δ and $\Delta\Delta$ respectively.

3.4 Advanced perceptual features

These features imply coefficients which provide information related to voice quality, perception, adaptation or amplitude modulation: Modulation Spectra Coefficients (MSC) [33], Perceptual Linear Predictive coefficients (PLP) [35], Linear Predictive Cepstral Coefficients (LPCC) [36], Linear Predictive Cosine Transform coefficients (LPCT) [29], Adaptive Component Weighted coefficients (ACW) [36], and Inferior Colliculus Coefficients (ICC)[38]. These features sometimes are extended by their first and second order regression coefficients, Δ and $\Delta\Delta$ respectively[29][34].

3.5 Non-linear features

The Fractal Dimension is one of the most common measures used to characterize the complexity of a system. Most fractal systems have a characteristic called self-similarity. An object is self-similar if a close-up examination of the object reveals that it is composed of smaller versions of itself. Self-similarity can be quantified as a relative measure of the number of basic building-blocks that form a pattern, and this measure is defined as the Fractal Dimension. Higuchi [39] and Castiglioni [40] are two algorithms especially suited for the analysis of time series that are used in speech analysis as well.

Permutation Entropy and Multiscale Permutation Entropy take into consideration the temporal information contained in the time series; furthermore, they have the quality of simplicity, robustness and very low computational cost [41][42][43]. Bandt and Pompe [44] introduce a simple and robust method based on the Shannon entropy measurement that takes into account time causality by comparing neighbor values in a time series. Summarizing, Permutation Entropy refers to the local order-structure of the time series, which can give a quantitative measure of complexity for dynamic time series.

3.6 Automatic feature selection

All features are calculated by software developed in our working group [45], Matlab [46] and Praat [20]. The number of features is about 350 for sample frequency of 16,000 kHz.

Then, an automatic feature selection is performed. This process is carried out in two steps:

1. First, taking into consideration medical and statistical information the *p-value* is extracted by the non-parametric Mann-Whitney *U-test* from Matlab [46]. After the test, the features with a significance level of $p\text{-value} < 0.05$ and $p\text{-value} < 0.1$ are selected. A higher $p\text{-value}$ is used as a threshold in order to select a wider feature set, for the second step. In the second phase, algorithms for automatic attribute selection are used to select the optimum feature set taking also into account feature integration factor.
2. Second, the previously obtained features are taken into account by automatic feature selection algorithms. There are three key benefits of performing feature selection on the data: reduction of over-fitting (reduction of data redundancy), improvement in the accuracy thanks to less misleading data, and computational cost reduction. Two algorithms from Weka [46] have been used for attribute selection and optimization. These algorithms analyze search methods and assessment of feature subsets exploring not only the impact of each feature, but also the impact of feature mutual interaction in classification.

- CfsSubsetEval: evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy between them [47].
- GainRatioAttributeEval: evaluates the worth of an attribute by measuring the gain ratio with respect to the class [47].

Finally, automatic classification was carried out by using classic Multilayer Perceptron (MLP) and Deep Learning by means of Convolutional Neural Networks (CNN), which are biologically-inspired variants of MLPs, taking into account the training sample set size and the problem complexity [47]:

1. AN task (CR/MCI): In the line of Deep Learning: Convolutional Neural Networks (CNN). We used one convolutional layer, with 40 feature maps and 5x5 patch-size and 2x2 max-pooling.
2. PD and SP tasks (CR/AD): Multilayer Perceptron (MLP). We used (attributes + classes)/2neurons in the layer and 500 training iterations.

The Weka software suite [46] has been used to carry out the experiments. The results were evaluated using Classification Accuracy Rate (Acc, %). For training and validation steps we used k-fold cross-validation with $k=10$, which is a robust validation method for variable selection [47][48]. Repeated cross-validation (as calculated by the Weka environment) allows robust statistical tests. We also use the default value provided automatically by Weka "Coverage of cases" (0.95 level) that is the confidence interval at 95%.

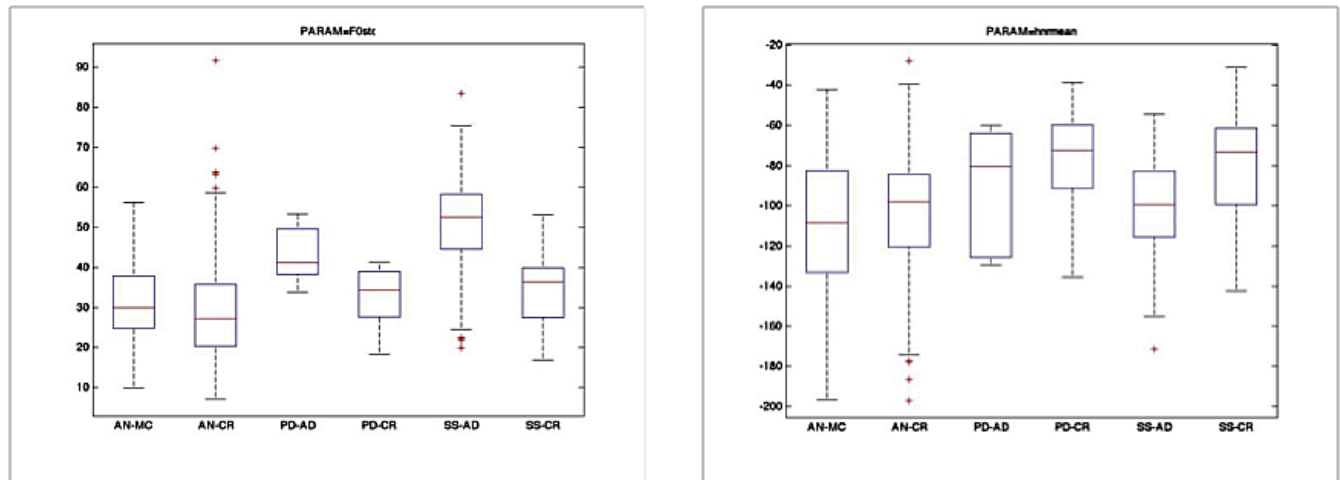


Fig. (4) Boxplot for the three tasks: *animal naming* (AN), *picture description* (PD) and *spontaneous speech* (SS). (a) F0 standard deviation. (b) Harmonicity.

4 RESULTS AND DISCUSSIONS

The experimentation has been carried out with the balanced subsets: PGA-OREKA, MINI-PGA and AZTIAHORE (Table 1).

4.1 Preliminary analysis

In the first stage all the features described in Section 3 are extracted, 340: (a) 70 classic, (b) 60 perceptual, (c) 180 advanced perceptual and (d) 30 non-linear features. There are differences in many features for all the subsets (a, b, c and d), but some of these differences are linked to the task type.

As an example Fig. (4) shows the analysis for some classic features. For MCI/AD people there is a clear instability of F0 and less harmonicity as well. This difference in F0-std, that is higher for the *spontaneous speech* task, could be applied

to analyze emotional response. In Fig. (5) some features detect differences with regard to pauses and voicing activity. These features could be indicators of speech fluency, cognitive process and/or emotional stage.

There are also clear trends in other subsets of features such as classic perceptual features, advanced perceptual features or non-linear ones (Fig. (6) and Fig. (7)).

However, despite some clear trends in many features, there are also differences in the obtained p-value for each task (dependency of the task) due mainly to the nature of each task: language complexity, age factor, individuals' stage and/or environment.

Therefore, an automatic feature selection is required in order to select the best features for each task. In this sense, in the next subsection an automatic methodology is applied for both feature selection and disease diagnosis.

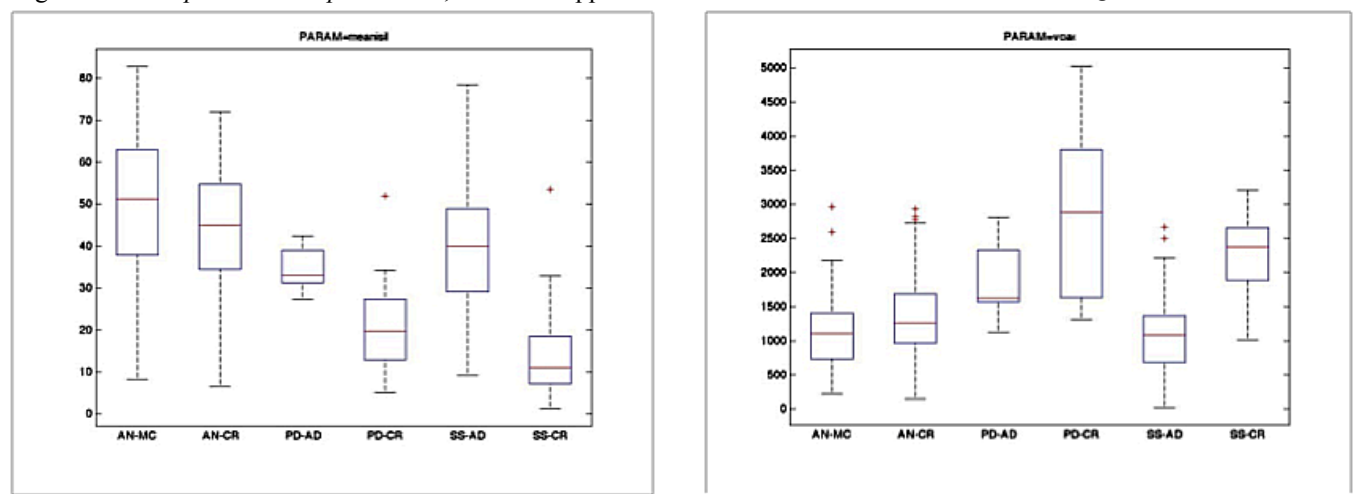


Fig. (5) Boxplot for the three tasks: *animal naming* (AN), *picture description* (PD) and *spontaneous speech* (SS). (a) Mean of pause duration. (b) Mean of voicing activity.

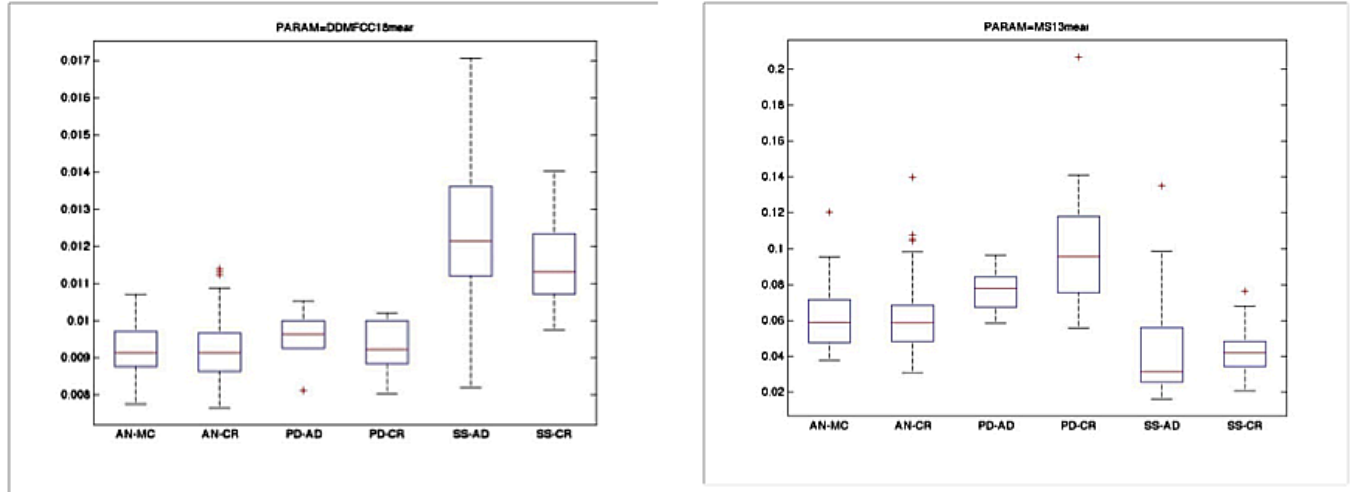


Fig. (6). Boxplot for the three tasks: *animal naming* (AN), *picture description* (PD) and *spontaneous speech* (SS). (a) Mean of ΔMFCC18 . (b) Mean of MS13.

4.2 Automatic feature selection and classification

In this second stage firstly a methodology to automatically select the most discriminative features is applied, as described in subsection 3.6.

After this process the number of features is reduced to around 60% for *animal naming*, 50% for *picture description* and 40% for *spontaneous speech*. Finally, an automatic classification by Neural Networks (NN) is carried out by means of a cross-validation strategy as described in Section 3.

The obtained global classification accuracy rates in % (Acc, %) for the three tasks are shown in Fig. (8):

- The best global results are obtained for the freest language, *spontaneous speech* task: 95%. This is mainly due to: the recording environment within a relaxing atmosphere; the presence of subtle cognitive changes in the signal due to

a more open language; and the use of AD patients instead of MCI subjects.

- The *animal naming* task has the lowest results: 73%. This is mainly due to: the restricted-task that could be easier for people with MCI; the use of MCI subjects instead of AD patients.
- The *picture description* task presents promising results, 89%, despite the small size of the set:

For the three tasks the obtained Acc% for classes is shown in Fig. (9):

- The best results for classes are also obtained for *spontaneous speech* task. This is mainly due to the following factors: The use of AD patients instead of MCI patients, and the detection of subtle cognitive doubts when they speak in a free way.

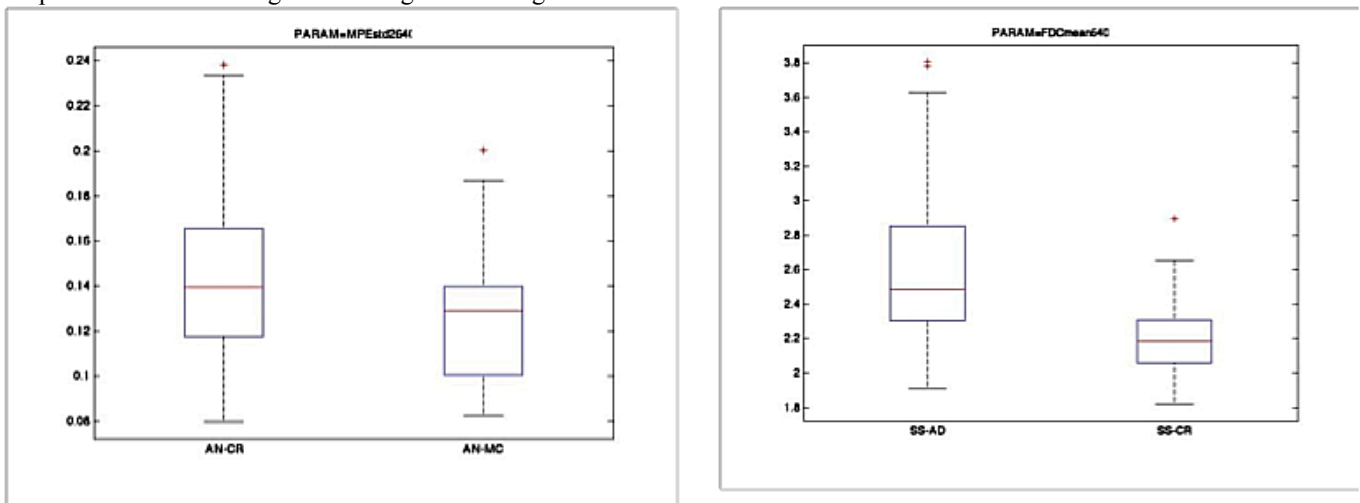


Fig. (7). Non-linear features of unvoiced segments with a window of 640 samples. (a) *Animal naming* task (AN): Standard deviation of Multiscale Permutation Entropy (2 scale). (b) *Spontaneous speech* task (SS): Mean of Castiglioni Fractal Dimension.

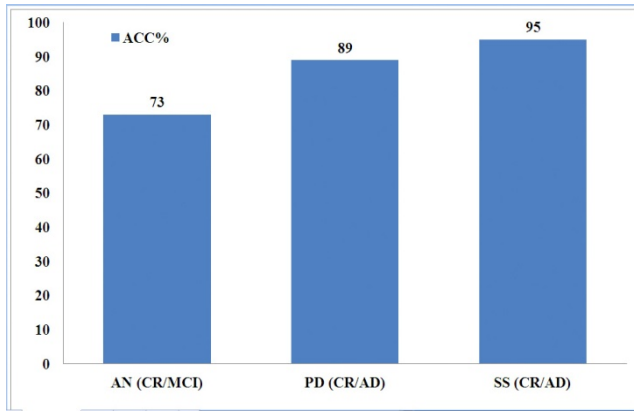


Fig. (8). Results of Accuracy (%) for the three tasks: *animal naming* (AN), *picture description* (PD) and *spontaneous speech* (SS).

- The *animal naming* task has the worst results when MCI subjects have been analyzed with a language restricted task. A larger training set is required because CR and MCI models are closer than CR and AD in the other tasks.
- In the case of the *picture description* task despite the small size of the set, the results are significantly good.
- The results of *animal naming* and *picture description* tasks are promising and require more sample training.

This work presents a novel methodology that could undoubtedly be very useful for health professionals and patients’ relatives. The two complementary methods provide the possibility of analyzing:

1. Novel specific features: which are independent from the language (linear and non-linear) in an independent way by means of statistical tests.
2. The automatic integration of the most relevant features by NN, providing useful information not available from statistical tests.

CONCLUSIONS

Automatic speech analysis is becoming a useful tool for cognitive assessment. This non-invasive analysis provides a friendly, natural, non expensive and accurate method for both assessment and monitoring of cognitive diseases. In this sense, the selection of appropriate measurement tasks is a crucial issue. The language complexity of the task is directly associated with the information that the system can extract. In fact, the measurement of subtle changes could require more spontaneous language in order to produce a more cognitive process. In contrast, language restricted tasks provide high accuracy for memory impairment detection by means of very robust indicators such as verbal fluency. Moreover, relevant communication elements such as emotional response do not need restricted language tasks to be properly measured. In this work three tasks with different levels of language complexity are analyzed, and promising results that encourage a deeper assessment are obtained. Automatic classification was carried out by using classic Multilayer Perceptron (MLP) and Deep Learning by means of Convolutional Neural Networks (CNN), which are biologically-inspired variants of MLPs. over the tasks with classic linear features, perceptual features, Castiglioni Fractal Dimension and Multiscale Permutation Entropy. Finally, the most relevant features are selected by means of the non-parametric Mann-Whitney U-test.

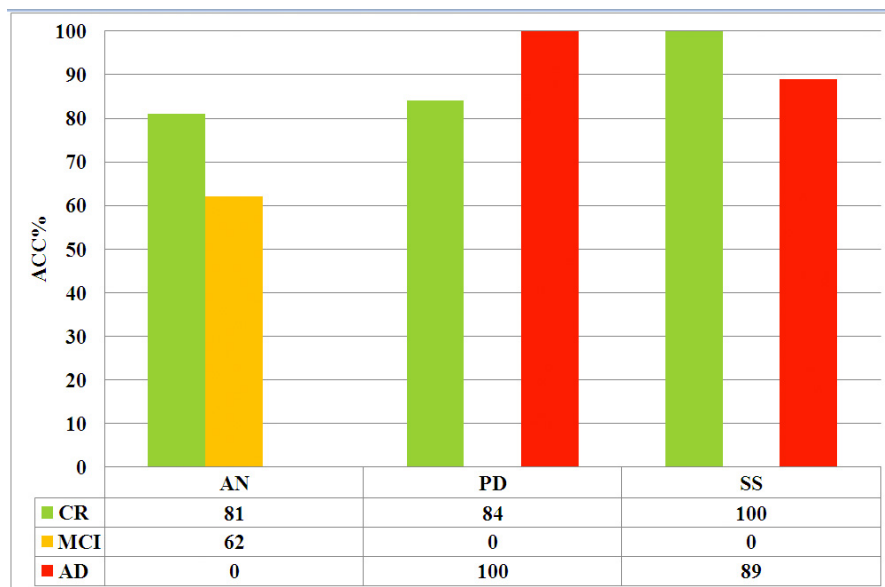


Fig. (9). Results of Accuracy (%) for the three tasks: *animal naming* (AN), *picture description* (PD) and *spontaneous speech* (SS).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGEMENTS

This work has being funded by grants from the Spanish, Basque and Gipuzkoa Government, and the University of the Basque Country: *IPT14/08, UPV/EHU-1548, MOV14/103, DG15/12.*

REFERENCES

[1] World Alzheimer Report 2015. Available online: www.alz.co.uk/research/world-report-2015 (accessed on 10 June 2016).

[2] McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the National Institute on Aging-Alzheimer's Association workgroup on Alzheimer's disease. *Neurology* 24: 939–944 (1984).

[3] McKhann G, Knopman DS, Chertkow H, Hyman BT, Jack CR, Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* 7: 263–269 (2011).

[4] Van de Pole LA, et al. The effects of age and Alzheimer's disease on hippocampal volumes A MRI study. *Alzheimer's and Dementia*, 1(1) S1: S51.

[5] Alzheimer's Association. Available online: <http://www.alz.org/>, (accessed on 10 June 2016).

[6] Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Under diagnosis of dementia in primary care: variations in the observed prevalence and comparisons to the expected prevalence. *Aging Ment Health* 15(8): 978-984 (2011).

[7] Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, Dauwels J, Soekadar SR, Mueller S, Linnemann C, Bridenbaugh S, Kanagasanga Y, Martins RN, O'Bryant SE. Innovative diagnostic tools for early detection of Alzheimer's disease *Alzheimer's and Dementia* 11(5): 561-78 (2015).

[8] Faundez-Zanuy M, et al. Biometric Applications Related to Human Beings: There Is Life beyond Security. *Cognitive Computation* 5(1): 136-151 (2013).

[9] Lopez-de-Ipiña K, Alonso JB, Travieso CM, Solé-Casals J, Egiraun H, Faundez-Zanuy M, Ezeiza A, Barroso N, Ecay M, Martinez-Lage P, Martinez-de-Lizarduy U. On the selection of non-invasive methods based on speech analysis oriented to Automatic Alzheimer Disease Diagnosis. *Sensors* 13 (5): 6730-6745 (2013).

[10] Woodford HJ, George J. Cognitive assessment in the elderly: a review of clinical methods. *QJM: An International Journal of Medicine.* 100(8): 469-84 (2007).

[11] Lopez-de-Ipina K, Martinez-de-Lizarduy U, Barroso N, Ecay-Torres M, Martinez-Lage P, Torres F, Faundez-Zanuy M.

Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impairment detection: A non-linear language independent approach. *Bioinspired Intelligence (IWOB)*, 2015 4th International Work Conference 1: 101–104, Donostia (Spain) (2015).

[12] Kato S, Homma A, Sakuma T. Easy Screening for Mild Alzheimer's Disease and Mild Cognitive Impairment from Elderly Speech Current Alzheimer Research, Special Issue

[13] Dodge H, Mattek N, Gregor M, Bowman M, Seelye A, Ybarra O, Asgari M, Kaye J. Social Markers of Mild Cognitive Impairment: Proportion of Word Counts in Free Conversational Speech. *Current Alzheimer Research* 12: 513-519 513 (2015).

[14] Mammarella N, Fairfield B. Emotional Working Memory and Alzheimer's Disease *International Journal of Alzheimer's Disease* 2014(207698): 1-6 (2014).

[15] Mikels JA, Reuter-Lorenz PA, Beyer JA, Fredrickson BL. Emotion and working memory: evidence for domain-specific processes for affective maintenance. *Emotion* 8(2): 256–266 (2008).

[16] Lezak, MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment* (fifth edition). Oxford University Press (2012).

[17] Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. *Brain Lang.* 57: 394-405 (1997).

[18] Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurol* 43:2412–2414 (1993).

[19] CITA-Alzheimer Foundation, PGA project: <http://www.cita-alzheimer.org/investigacion/proyectos>, (accessed on 10 June 2016).

[20] Praat: doing Phonetics by Computer. Available online: www.praat.org/ (accessed on 10 June 2016).

[21] Folstein M, Folstein S. E., & McHugh P. R. "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* 12(3):189-98 (1975).

[22] Buschke H. Cued recall in amnesia. *Journal of Clinical and Experimental Neuropsychology* 6:433–440 (1984).

[23] Pfeffer RI, Kurosaki TT, Harrah CH Jr, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *Journal of Gerontology* 37(3): 323-329 (1982).

[24] Goodglass H. *Evaluación de la Afasia y de Trastornos Relacionados* (3ªed.). Madrid: Editorial Médica Paramericana (2005).

[25] Benton AL Hannay H.J., Varney N. Visual perception of line direction in patients with unilateral brain disease. *Neurology*, 25: 907-10 (1975).

- [26] Wechsler D. WAIS-III. Escala de inteligencia de Wechsler para adultos-III (2ª ed.). Madrid: TEA Ediciones, S.A. (2001).
- [27] Reitan R.M. & Wolfson D. (1985). The Halstead – Reitan neuropsychological test battery. Theory and clinical interpretation (2nd. Ed.). Tucson, AZ: Neuropsychology Press.
- [28] Kertesz, A. Bateria de Afasias Western (BAW). Nau Llibres (2000).
- [29] Mekyska J, Janousova E, Gomez-Vilda P, Smekal Z, Rektorova I, Eliasova I, Kostalova M, Mrackova M, Alonso-Hernandez JB, Faundez-Zanuy M, López-de-Ipiña K. Robust and complex approach of pathological speech signal analysis. *Neurocomputing* 167(2015): 94–111 2015.
- [30] Gomez-Vilda P, Rodellar-Biarge V, Nieto-Lluis V, Munoz-Mulas C, Mazaira- Fernandez L, Martinez-Olalla R, Alvarez-Marquina A, Ramirez-Calvo C., Fernandez-Fernandez M. Characterizing neurological disease from voice Quality biomechanical analysis. *Cognitive Computation* 5(4): 399–425 (2013).
- [31] Little MA, Mcsharry PE, Roberts SJ, Costello DAE, Moroz IM, Exploiting non linear recurrence and fractal scaling properties for voice disorder detection *Biomed. Eng.* doi: 10.1186/1475-925X-6-23 6: 23 (2007).
- [32] Song Y, Wang WH, Guo FJ. Feature extraction and classification for audio information in news video. *International Conference on Wavelet Analysis and Pattern Recognition, 2009 (ICWAPR 2009)* 1: 43–46 (2009).
- [33] Markaki M, Stylianiou Y, Arias-Londono J, Godino-Llorente J. Dysphonia detection based on modulation spectral features and cepstral coefficients In: 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP) 1: 5162–5165 (2010).
- [34] López-de-Ipiña K, Iturrate M, Alonso JB, Rodriguez-Herrera B. Automatic acoustic analysis for biodiversity preservation: a multi-environmental approach. *IWOBI 2015 Donostia (Spain)* 1: 43-48 (2015).
- [35] Hermansky H. Perceptual linear predictive (PLP) analysis of speech *J. Acoust. Soc. Am.* 87(4): 1738–1752 (1990).
- [36] Mammone R, Zhang X, Ramachandran R. Robust speaker recognition: a feature-based approach. *IEEE Signal Process. Mag.* 13(5): 58–71 (1996).
- [37] Gelzinis A, Verikas A, Bacauskiene M. Automated speech analysis applied to laryngeal disease categorization *Comput. Methods Progr. Biol.* 91(1): 36–47 (2008).
- [38] Malyska N, Quatieri T, Sturim D. Automatic Disphonia recognition using biologically-inspired amplitude-modulation features In: *Proceedings of IEEE International Conference on Acoustics Speech And Signal Processing (ICASSP'05)* 1: 873–876 (2005).
- [39] Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D: Non linear Phenomena* 31(2): 277:283 (1988).
- [40] Castiglioni P. What is wrong in Katz's method? Comments on: "A note on fractal dimensions of biomedical waveforms". *Comput Biol Med* 40:11-12 (2010).
- [41] Zanin M, Zunino L, Rosso OA, Papo D. Permutation entropy and its main biomedical and econophysics applications: a review. *Entropy* 14(8): 1553-1577 (2012).
- [42] Morabito FC, Labate D, La Foresta F, Bramanti A, Morabit G, Palamara I. Multivariate Multi-Scale Permutation Entropy for Complexity Analysis of Alzheimer's Disease EEG. *Entropy* 14(7): 1186-1202 (2012).
- [43] Costa M, Goldberger A, Peng CK. Multiscale entropy analysis of biological signals. *Phys. Rev.* 71: 021906:1–021906:18 (2005).
- [44] Bandt C, Pompe B. Permutation entropy: A natural complexity measure for time series. *Phys. Rev. Lett* 88: 174102:1–174102:4 (2002).
- [45] EleKin Available online: www.elekin.net (accessed on 10 June 2016).
- [46] Matlab. Available online: www.mathworks.com (accessed on 10 June 2016).
- [47] Weka. Available online: <http://www.cs.waikato.ac.nz/ml/weka> (accessed on 10 June 2016).
- [48] Picard R, Cook D. Cross-Validation of Regression Models. *Journal of the American Statistical Association* 79(387): 575–58 (1984).

C1 ERANSKINA

Feature Extraction Approach Based on
Fractal Dimension for Spontaneous Speech
Modelling Oriented to Alzheimer Disease
Diagnosis

Feature extraction approach based on Fractal Dimension for Spontaneous Speech modelling oriented to Alzheimer Disease diagnosis

Karmele López-de-Ipiña¹, Harkaitz Egi rau n¹, Jordi Sole-Casals², Miriam Ecay³, Aitzol Ezeiza¹, Nora Barroso¹, Pablo Martinez-Lage³, Unai Martinez-de-Lizardui¹

¹ Department of System Engineering and Automation, University of the Basque Country, Spain
karmele.ipina@ehu.es

² Digital Technologies Group, Universitat de Vic

³ Neurology Department CITA-Alzheimer Foundation

Abstract. Alzheimer's disease (AD) is the most prevalent form of progressive degenerative dementia; it has a high socio-economic impact in Western countries. The purpose of our project is to contribute to earlier diagnosis of AD and better estimates of its severity by using automatic analysis performed through new biomarkers extracted from non-invasive intelligent methods. The methods selected in this case are speech biomarkers oriented to Spontaneous Speech. Thus the main goal of the present work is feature search in Spontaneous Speech oriented to pre-clinical evaluation for the definition of test for AD diagnosis. Nowadays our feature set offers some hopeful conclusions but fails to capture the nonlinear dynamics of speech that are present in the speech waveforms. The extra information provided by the nonlinear features could be especially useful when training data is scarce. In this work, the Fractal Dimension (FD) of the observed time series is combined with lineal parameters in the feature vector in order to enhance the performance of the original system.

Keywords. Nonlinear Speech Processing, Alzheimer disease diagnosis, Spontaneous Speech, Fractal Dimensions.

1 Introduction

Alzheimer's Disease (AD) is the most common type of dementia among the elderly. It is characterized by progressive and irreversible cognitive deterioration with memory loss and impairments in judgment and language, together with other cognitive deficits and behavioral symptoms. The cognitive deficits and behavioral symptoms are severe enough to limit the ability of an individual to perform everyday professional, social or family activities. As the disease progresses, patients develop severe disability and full dependence. An early and accurate diagnosis of AD helps patients and their families to plan for the future and offers the best opportunity to treat the symptoms of the disease. According to current criteria, the diagnosis is expressed with different degrees of certainty as possible or probable AD when dementia is pre-

sent and other possible causes have been ruled out. The diagnosis of definite AD requires the demonstration of the typical AD pathological changes at autopsy [1,2,3]. The clinical hallmark and earliest manifestation of AD is episodic memory impairment. At the time of clinical presentation, other cognitive deficits are present in areas like language, executive functions, orientation, perceptual abilities and constructional skills [4,5]. All these symptoms lead to impaired performance in everyday activities. Approaches to the early diagnosis of AD have in the past few years made significant advances in the development of reliable clinical biomarkers [6]. However, the cost and technology requirements make it impossible to apply such biomarkers to all patients with memory complaints. Given these problems, non-invasive Intelligent Techniques of diagnosis may become valuable tools for early detection of dementia. Non-technical staff in the habitual environments of the patient could use these methodologies, which include e.g. Automatic Spontaneous Speech Analysis (ASSA) without altering or blocking the patients' abilities, as the spontaneous speech involved in these techniques is not perceived as a stressful test by the patient. Moreover, these techniques are very low-cost and do not require extensive infrastructure or the availability of medical equipment. They are thus capable of yielding information easily, quickly, and inexpensively [7,8]. In addition to the loss of memory, one of the major problems caused by AD is the loss of language skills. We can meet different communication deficits in the area of language, including aphasia (difficulty in speaking and understanding) and anomia (difficulty in recognizing and naming things). The specific communication problems the patient encounters depend on the stage of the disease [3,4,5]:

1. First Stage or early stage (ES): difficulty in finding the right word in spontaneous speech. Often remains undetected.
2. Second Stage or intermediate stage (IS): impoverishment of language and vocabulary in everyday use.
3. Third Stage or advanced stage (AS): answers sometimes very limited and restricted to very few words.

The main goal of the present work is feature search in Spontaneous Speech oriented to pre-clinical evaluation for the definition of test for AD diagnosis. These features will define control group (CR) and the three AD levels. One of the most relevant nonlinear techniques for ASR is the consideration of the Fractal Dimension (FD) of the speech signal as a feature to be used in the training process. The interest on fractals in speech date back to the mid-80's [9], and they have been used for a variety of applications, including consonant/vowel characterization [10-11], speaker identification [12], and end-point detection [13], even for whispered speech [14]. The approach of this work is to improve the system developed in our previous work [8] augmenting the features with FD. The FD is one of the most significant features which describe the complexity of a system and could help in the detection of subtle changes for early diagnosis. Moreover this feature has in the ability to capture the dynamics of the system and thus relevant variations in speech utterances. More precisely, an implementation of Higuchi's algorithm [18] in order to add this new feature to the set that feeds the training process of the model.

The rest of this paper is organized this way: In Section 2, the materials are presented. Section 3 explains the methodology of the experiments, Section 4 shows the experimental results, and finally, conclusions are presented in Section 5.

2 MATERIALS

This study focuses on early AD detection and its objective is the identification of AD in the pre-clinical (before first symptoms) and prodromic (some very early symptoms but no dementia) stages. The research presented here is a complementary preliminary experiment to define thresholds for a number of biomarkers related to spontaneous speech. Feature search in this work is oriented to pre-clinical evaluation for the definition of test for AD diagnosis. Obtained data will complement the biomarkers of each person.

Trying to develop a new methodology applicable to a wide range of individuals of different sex, age, language and cultural and social background, we have built up a multicultural and multilingual (English, French, Spanish, Catalan, Basque, Chinese, Arabian and Portuguese) database with video recordings of 50 healthy and 20 AD patients (with a prior diagnosis of Alzheimer) recorded for 12 hours and 8 hours respectively. The age span of the individuals in the database was 20-98 years and there were 20 males and 20 females. This database is called AZTIAHO. All the work was performed in strict accordance with the ethical guidelines of the organizations involved in the project. The recordings consisted of videos of Spontaneous Speech – people telling pleasant stories or recounting pleasant feelings as well as interacting with each other in friendly conversation. The recording atmosphere was relaxed and non-invasive. The shorter recording times for the AD group are due to the fact that AD patients find speech more of an effort than healthy individuals: they speak more slowly, with longer pauses, and with more time spent on efforts to find the correct word and uttering speech disfluencies or break messages. In the advanced stage of the disease, they find this effort tiring and often want to stop the recording. We complied with their requests. The video was processed and the audio extracted in wav format (16 bits and 16 KHz). The first step was removing non-analyzable events: laughter, coughing, short hard noises and segments where speakers overlapped. Next, background noise was removed using denoiser adaptive filtering. After the pre-processing, about 80% of the material from the control group and 50% of the material from the AD group remained suitable for further analysis. The complete speech database consists of about 60 minutes of material for the AD group and about 9 hours for the control. The speech was next divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers, resulting finally in a database of about 600 segments of Spontaneous Speech. From the original database, a subset of 20 AD patients was selected (68-96 years of age, 12 women, 8 men, with a distribution in the three stages of AD as follows: First Stage [ES=4], Secondary Stage [IS=10] and Tertiary stage [AS=6]). The control group (CR) was made up of 20 individuals (10 male and 10 female, aged 20-98 years) representing a wide range of speech responses. This subset of the database is called AZTIAHORE.

3 Methods

3.1 Feature extraction

3.2.1 Features oriented to Automatic Spontaneous Speech Analysis (ASSA)

Spoken language is one of the most important elements defining an individual's intellect, social life, and personality; it allows us to communicate with each other, share knowledge, and express our cultural and personal identity. Spoken language is the most spontaneous, natural, intuitive, and efficient method of communication among people. Therefore, the analysis by automated methods of Spontaneous Speech (SS – free and natural spoken communication), possibly combined with other methodologies, could be a useful non-invasive method for early AD diagnosis. The analysis of Spontaneous Speech fluency is based on three families of features (SSF set), obtained by the Praat software package and software [15] that we ourselves developed in MATLAB. For that purpose, an automatic Voice Activity Detector (VAD) [16,17] has extracted voiced/unvoiced segments as parts of an acoustic signal.

These three families of features include: 1) *Duration*: this includes the histogram calculated over the most relevant voiced and unvoiced segments, the average of the most relevant voiced/unvoiced, voiced/unvoiced percentage and spontaneous speech evolution along the time dimension, and the voiced and unvoiced segments' mean, max and min; 2) *Time domain*: short time energy; 3) *Frequency domain, quality*: spectral centroid.

The energy of a signal is typically calculated on a short-time basis, by windowing the signal at a particular time, squaring the samples and taking the average. The spectral centroid is commonly associated with the measure of the brightness of a sound. This measure is obtained by evaluating the “center of gravity” using the Fourier transform's frequency and magnitude information.

3.2.2 Fractal Dimension

Most of the fractal systems have a characteristic called self-similarity. An object is self-similar if a close-up examination of the object reveals that it is composed of smaller versions of itself. Self-similarity can be quantified as a relative measure of the number of basic building blocks that form a pattern, and this measure is defined as the Fractal Dimension. There are several algorithms to measure the Fractal Dimension, but this current work focus on the alternatives which are specially suited to time series analysis and which don't need previous modelling of the system. Two of these algorithms are Higuchi [18] and Katz [19], named after their authors. Higuchi was choice because it has been reported to be more accurate in previous works with under-resourced conditions [20]. Higuchi [18] proposed an algorithm for measuring the Fractal Dimension of discrete time sequences directly from the time series $x(1), x(2), \dots, x(n)$. Without going into detail, the algorithm calculates the length $L_m(k)$ (see Equation 1) for each value of m and k covering all the series.

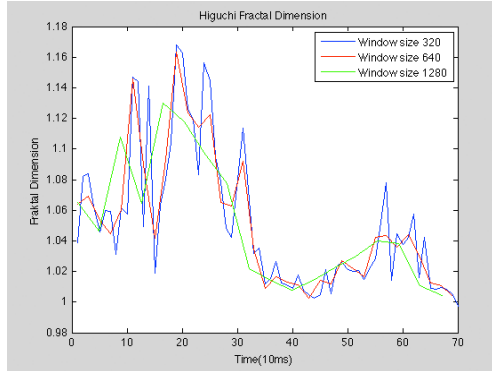


Figure 1. Higuchi Fractal Dimension (HFD) for an AD signal for different window sizes.

$$L_m(k) = \frac{\sum_{i=1}^{\lfloor \frac{N-m}{k} \rfloor} |x_{m+ik} - x_{m+(i-1)k}|}{\lfloor \frac{N-m}{k} \rfloor} \quad (1)$$

After that, a sum of all the lengths $L_m(k)$ for each k is determined with Equation 2.

$$L(k) = \sum_{m=1}^k L_m(k) \quad (2)$$

And finally, the slope of the curve $\ln(L(k))/\ln(1/k)$ is estimated using least squares linear best fit, and the result is the Higuchi Fractal Dimension (HFD).

3.2.3 Feature sets

In the experimentation four feature set will be used: 1) SSF: Spontaneous Speech features; 2) SSF+HFD1: SSF set and Higuchi Fractal Dimension (HFD). 3) SSF+HFD2: SSF set, HFD, maximum HFD, minimum HFD, variance HFD and standard deviation HFD.

4 Automatic Classification method

The automatic classification of speech is based on the Multi Layer Perceptron (MLP). WEKA software [21] has been used in carrying out the experiments. The results were evaluated using Accuracy (Acc) and Classification Error Rate (CER) measurements. For the training and validation steps, we used k -fold cross-validation with $k=10$. Cross validation is a robust validation for variable selection [22]. These features will define the CR group and the three AD levels. The main goal of the present work is feature search in Spontaneous Speech oriented to pre-clinical evaluation for the definition of test for AD diagnosis. These features will define CR group and the three AD levels.

Table 1. Global Accuracy (%) with MLP for Automatic Spontaneous Speech Analysis and different Neuron Number in Hidden Layer (NNHL) and Training Step (TS)

NNHL	TS	SSF	SSF+HFD1	SSF+HFD2
50	500	75.96	76.74	85.27
	1000	75.96	76.44	83.72
	1500	74.41	76.44	82.94
100	500	76.74	79.06	85.27
	1000	75.96	77,51	86.04
	1500	75.96	77,51	86.82
150	500	75.96	79,84	84.49
	1000	75.96	79,06	83.72
	1500	74.41	79,06	82.84

5 Experimental results

The task was Automatic Classification, with the classification targets being: healthy speakers without neurological pathologies and speakers diagnosed with AD. The experimentation is carried out with AZTIAHORE. Engineers and health specialists have analyzed results with regard to: global results and also AD level results.

1. Global system results: The results are satisfactory for this study. The new fractal features improve the system, being SSF+HFD2 the best option. This feature set includes both HFD and its detailed variations, which are able to model non-linear signal features (Table 1).
2. MLP selection: Table 1 shows detailed results with regard to the selection of MLP characteristics. The best results are obtained with one hidden layer of 100 neurons with 1500 training steps.
3. Window size: For HFD algorithm three different window size have been used: 1280, 640 and 320 samples. The best results have been obtained for a window size of 320 samples.



Figure 2. Classification Error Rate (%) with MLP for different classes: CR, ES, IS and AS

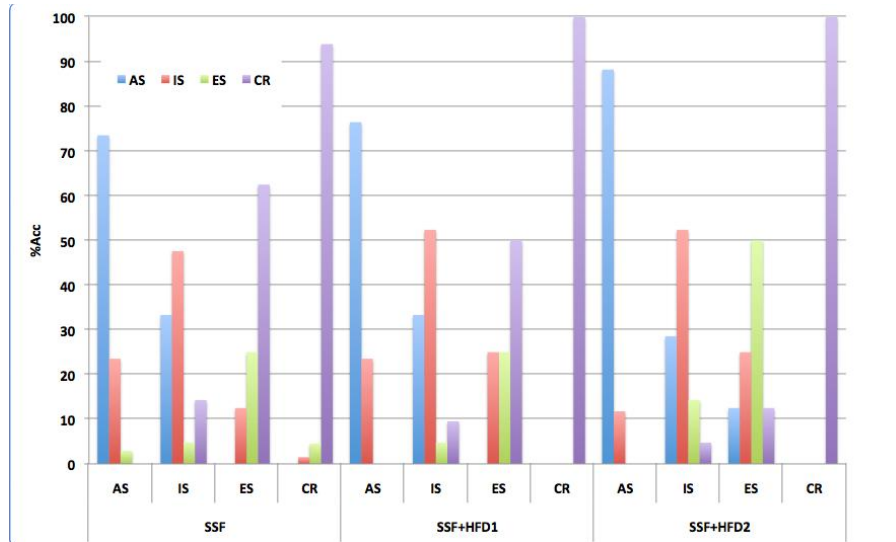


Figure 3. Accuracy (%) with MLP for different classes: CR, ES, IS and AS

4. Classes' results: SSF+HFD2 set obtains the best results for all classes (Figure 2). This set improves also the classification with regard to early detection (ES class, Figure 3). IS has also better rate to discriminate middle AD level. The model is able also to discriminate pathological and non-pathological segments in each patient.

Health specialists note the relevance of the system's ability to carry out both the analysis of independent biomarkers as spontaneous speech and/or the integral analysis of several biomarkers.

6 Conclusions and future work

The main goal of the present project is feature search in Spontaneous Speech oriented to pre-clinical evaluation for the definition of test for AD diagnosis. These features are of great relevance for health specialists to define health people and the three AD levels. The approach of this work is to improve the previous modelling based on Spontaneous Speech features with Fractal Dimensions. More precisely, an implementation of Higuchi's algorithm in order to add this new feature to the set that feeds the training process of the model. In this work, it is described a first approach to the inclusion of nonlinear features. This straightforward approach might be robust in terms of capturing the dynamics of the whole waveform, and it offers many advantages in terms of computability, and it also makes easier to compare the power of the new features against the previous ones. In future works we will introduce new features relatives to speech modelling oriented to standard medical tests for AD diagnosis and to emotion response analysis. We will also model Fractal Dimension by other algorithms based on previous works in Automatic Speech Recognition [20].

References

1. Mc Kahn G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Workgroup on Alzheimer's disease. 1984; 24:939-944.
2. McKhann GM et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011 May;7(3):263-9.
3. Van de Pole, L.A., et al., *The effects of age and Alzheimer's disease on hippocampal volumes, a MRI study*. *Alzheimer's and Dementia*, 2005. 1(1, Supplement 1): p. 51.
4. Morris JC, The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurology*, 1993. 43: p. 2412b-2414b.
5. American Psychiatric Association, 2000. *Diagnostic and Statistical Manual of Mental disorders, 4th Edition Text Revision*. Washington DC.
6. Alzheimer's Association.: <http://www.alz.org/>
7. M. Faundez-Zanuy et al. Biometric Applications Related to Human Beings: There Is Life beyond Security, *Cognitive Computation*, 2012, DOI 10.1007/s12559-012-9169-9
8. K. López de Ipiña, J. B. Alonso, J. Solé-Casals, N. Barroso, M. Faundez, M. Ecay, C. Travieso, A. Ezeiza and A. Estanga Alzheimer Disease Diagnosis based on Automatic Spontaneous Speech Analysis, *Proceedings of NCTA 2012*. Barcelona, 2012.
9. Pickover CA, Khorasani A. Fractal characterization of speech waveform graphs. *Comput Graph* 1986. 10:1:51-61
10. Martinez F, Guillamon A, Martinez JJ. Vowel and consonant characterization using fractal dimension in natural speech. In: *Proceedings of NOLISP' 2003*; 2003.
11. Langi A, Kinsner W. Consonant Characterization Using Correlation Fractal Dimension for Speech Recognition. In: *Proceedings of Communications, Power, and Computing*. Conference Proceedings. IEEE; 1995;
12. Nelwamondo FV, Mahola U, Marwola T. Multi-Scale Fractal Dimension for Speaker Identification Systems. *WSEAS Trans Syst* 2006. 5:5:1152-1157.
13. Li Y, Fan Y, Tong Q. Endpoint Detection In Noisy Environment Using Complexity Measure. In: *Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, China; 2007.
14. Chen X, Zhao H. Fractal Characteristic-Based Endpoint Detection for Whispered Speech. In: *Proceedings of the 6th WSEAS International Conference on Signal, Speech and Image Processing*, Lisbon, Portugal; 2006.
15. Praat: doing Phonetics by Computer. Available online: www.fon.humuva.nl/praat
16. Voice Activity Detector algorithm (VAD). Available online: www.mathwork.com
17. Solé J, Zaiats V., A Non-Linear VAD for Noisy Environment. *Cognitive Computation*. 2010; 2(3):191-198.
18. Higuchi T. Approach to an irregular time series on the basis of the fractal theory. *Physica D* 1988. 31277:283.
19. Katz M. Fractals and the analysis of waveforms. *Comput Biol Med* 1988. 18:3:145:156.
20. A Ezeiza, KL de Ipiña, C Hernández, N Barroso, Enhancing the feature extraction process for automatic speech recognition with fractal dimensions, *Cognitive Computation*, 2012
21. WEKA. Available online: <http://www.cs.waikato.ac.nz/ml/weka/>
22. Picard, Richard; Cook, Dennis (1984). "Cross-Validation of Regression Models". *Journal of the American Statistical Association* 79 (387): 575–583. doi:10.2307/2288403. JSTOR 2288403.

C2 ERANSKINA

Automatic analysis of Categorical Verbal
Fluency for Mild Cognitive Impairment
detection: a non-linear language
independent approach

Automatic analysis of Categorical Verbal Fluency for Mild Cognitive Impairment detection: a non-linear language independent approach

K. López-de-Ipiña, U. Martínez-de-Lizarduy, N. Barroso

Universidad del País Vasco/Euskal Herriko Unibertsitatea
{karme.leipina, unai.martinezdelizarduy}@ehu.eus

M. Ecay-Torres, P. Martínez-Lage, F. Torres
Fundación CITA Alzheimer
{mecay, pmlage, ftorres}@cita-alzheimer.org

Marcos Faundez-Zanuy

Escola Universitaria Politècnica de Mataró (UPF),
Tecnocampus {faundez, sesa}@tecnocampus.cat

Abstract—: Alzheimer's disease (AD) is one the main causes of dementia in the world and the patients develop severe disability and sometime full dependence. In previous stages Mild Cognitive Impairment (MCI) produces cognitive loss but not severe enough to interfere with daily life. This work, on selection of biomarkers from speech for the detection of AD, is part of a wide-ranging cross study for the diagnosis of Alzheimer. Specifically in this work a task for detection of MCI has been used. The task analyzes Categorical Verbal Fluency. The automatic classification is carried out by SVM over classical linear features, Castiglioni fractal dimension and Permutation Entropy. Finally the most relevant features are selected by ANOVA test. The promising results are over 50% for MCI.

Keywords— *Mild Cognitive Impairment, Alzheimer's Disease, Automatic speech analysis; Non.linear features; Entropy; Fractals; automatic selection of features*

I. INTRODUCTION

Alzheimer's disease (AD) is one the main causes of dementia and death in the world and the patients develop severe disability and sometime full dependence. AD is the most common type of dementia among the elderly. It is characterized by progressive and irreversible cognitive deterioration with memory loss and impairments in judgment and language, together with other cognitive deficits and behavioral symptoms. The cognitive deficits and behavioral symptoms are severe enough to limit the ability of an individual to perform everyday life.. An early and accurate diagnosis of AD helps to increase their quality of life and offers the best possibilities of treating the symptoms of the disease. The diagnosis of definite AD requires the demonstration of the typical AD pathological changes at autopsy [1,2].

During social interaction in everyday life, verbal communication is, for human being, one of the most important ways of expression. It is a complex process, which involves a wide range of cognitive abilities. We are able to produce

sounds with a message inside. The biosignal hides our desires, ideas and/or emotions, a piece of our being. In that process the brain has to manage knowledge, memory, language and semantic information. Moreover, these biological samples have the advantage of be very easy to record without invasive and expensive equipments. In the case of Alzheimer's Disease (AD), the deterioration of spoken language immediately affects the patient's ability to naturally interact with his or her al environment, and it is usually also accompanied by alterations in emotional response. [1-3]. In this sense, the cost and technology requirements of some of medical proofs (fMRI, PET) make impossible to apply such biomarkers to all patients or require painful tests. Given these problems, non-invasive intelligent techniques of diagnosis may become valuable tools for early detection of diseases to complement diagnosis and/or monitoring its progress.

Mild cognitive impairment produces cognitive loss but not severe enough to interfere with daily life or independent abilities. These changes not affect daily life and a person with MCI usually don't have an appropriate diagnosis. MCI could increase the risk of developing Alzheimer's or another dementia and currently its detection is one of the challenges of medical specialists [4]. In addition, some references show that non-specialists and even familiars are not able to identify exactly early AD as well as Mild Cognitive Impairment (MCI) [5,6]. Thus non-technical staff in the habitual environment of the patient could use these methodologies, which include speech and voice analysis, without altering or blocking the patients' abilities, as the spontaneous speech involved in these techniques is not perceived as a stressful test by the patient. Moreover, these techniques are very low-cost and do not require extensive infrastructure or the availability of medical equipment. They are thus capable of yielding information easily, quickly, and inexpensively [7,8]. In this sense it is essential the search for clinically useful screening and diagnostic Tools.

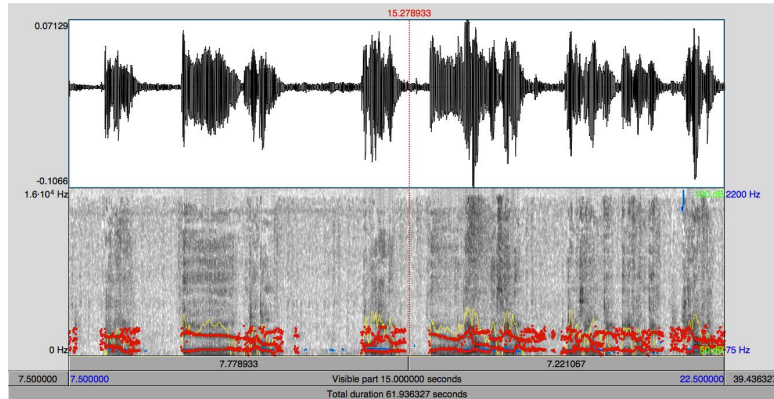


Figure 1. Categorical Verbal Fluency (CVF) task uttered by an individual of the control group

The main goal of this project is the development of an automated analysis of categorical verbal fluency (CVF) through speech therapy techniques that allow obtaining these specific analyses quickly and reliably. There are several works as references during the last years. In this paper we focus our research on the integration of the language independent methodologies to detect AD in speech, to a task of Categorical Verbal Fluency (CVF) [1,9]. Specifically the so-called animals naming task.

The materials are described in section II. Section III presents the used methods. The results and discussion are described in section IV finally concluding remarks are included in section V.

II. MATERIALS

A. Gipuzkoa-Alzheimer Project (PGA)

The sample consists of 187 healthy people and 38 with MCI belonging to the cohort of Gipuzkoa-Alzheimer Project (PGA) of the CITA-Alzheimer Foundation. The PGA is a longitudinal study focusing on the search for biomarkers of function, brain structure and state of health and risk factors for early diagnosis of Alzheimer's disease. The main objective of PGA is to characterize Alzheimer's disease and pre-clinical prodrome. Inclusion criteria based on the characteristics of each group are listed below.

1) Control group

- Men and women of middle age or elderly, between 39 and 79 years (inclusive) to sign the informed consent to participate in the PGA.
- Lack of memory complaints.
- Global cognitive function preserved with have the highest score of 24 in the Mini Mental State, examination (MMSE) [10].
- Performance within the range of normal in the memory test. Greater scale score of 6 in all indexes of free recall test and selectively provided Buschke (FCSRT) [11].

- Autonomy in activities of daily living. Less than 6 in the functional activities questionnaire Pfeffer [12] score.

2) MCI group

- Men and women aged between 39 and 79 years (inclusive) to sign the informed consent to participate in the PGA.
- The presence of memory complaints that the assessor considers relevant.
- Global cognitive function preserved. Highest score of 24 in the MMSE.
- Performance below the average adjusted by age and educational level in at least one of the tests that assess memory, language, visual-spatial function, attention and executive function. Scale score equal to or less than 6.
- Autonomy in activities of daily living. Less than 6 score on the Questionnaire of Pfeffer Functional Activities

Exclusion criteria are the presence of dementia (DSM-IV and stage $CDR \geq 1$); significant history of neurological disease of any kind that might cause cognitive impairment dementia; history of diagnosed psychiatric illness that could cause dementia cognitive impairment; people with mild psychiatric symptoms (anxiety, depression) are not excluded. Finally a balanced database of about 100 individuals has been created PGA-OREKA.

B. Task of Categorical Verbal Fluency (CVF)

The task of categorical verbal fluency (CVF) is a test to measure and quantify the progression of cognitive impairment in neurodegenerative diseases. It is widely used to assess language skills, associative memory and executive functions. During the CVF task the interviewer ask the patient to list all the names you know of a category in one minute (animals). In clinical practice, only the total number of elements emitted is recorded. However, you can get more specific indexes that detect subtle changes in cognitive status of the patient and thus gain in specificity to discriminate between different degenerative diseases [1,9].

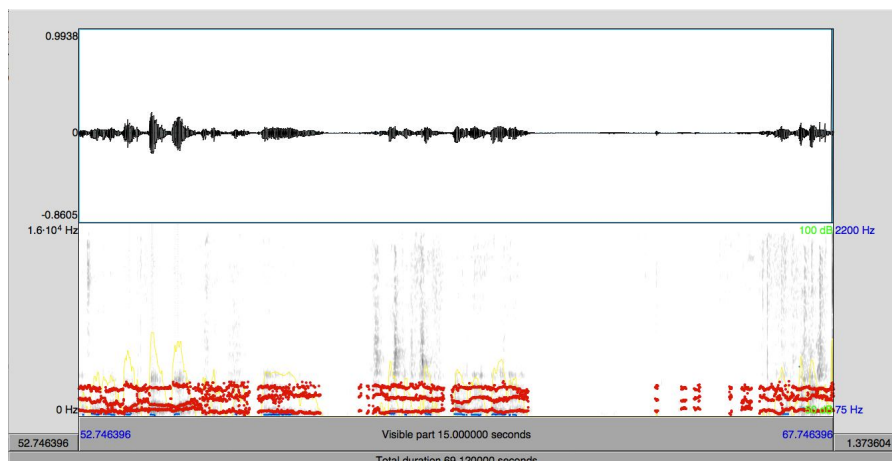


Figure 2. Categorical Verbal Fluency (CVF) task uttered by an individual with MCI

III. METHODS

The analysis by automated methods of this task, possibly combined with other methodologies, could be a useful non-invasive method for early AD diagnosis [8]. The analysis of CVF task is based on three families of features (SSF set), obtained by the Praat software package [13] and software that we ourselves developed in MATLAB. For that purpose, an automatic Voice Activity Detector has extracted voiced/unvoiced segments as parts of an acoustic signal [8].

The families of features include in this experimentation are [8]:

A. Feature extraction

1) Linear features

- *Duration*: the histogram calculated over the most relevant voiced and unvoiced segments, the average of the most relevant voiced/unvoiced, voiced/unvoiced percentage and spontaneous speech evolution along the time dimension, and the voiced and unvoiced segments' mean, max and min;
- *Time domain*: short time energy. Energy, derivative of the energy.
- *Frequency domain, quality*: spectral centroid. If m is too small (smaller than 3) the algorithm will work wrongly because it will only have few distinct states for recording but it depends on the data. When using long signals, a large value of m is preferable but it would require a larger computational time.
- *12 Mel Frequency Cepstral Coefficients MFCC*
- *Acoustic features*: pitch, standard deviation pitch, max and min pitch, intensity, standard deviation intensity, max and min intensity, period mean, period standard deviation, and Root Mean Square amplitude (RMS).
- *Voice quality features*: shimmer, local jitter, Noise-to-Harmonics Ratio (NHR), Harmonics-to-Noise Ratio (HNR) and autocorrelation.
- *Duration features*: fraction of locally unvoiced frames, degree of voice breaks.

2) Non linear features

- Castiglioni fractal dimension
- Permutation entropy

B. Automatic selection of features by ANOVA

Then we will select automatically the best argument with regard to common significance level. Thus automatic feature selection is performed by ANOVA [14]. This function performs balanced one-way ANOVA for comparing the means of two or more columns of data in the matrix X , where each column represents an independent sample containing mutually independent observations. The function returns the p -value under the null hypothesis that all samples in X are drawn from populations with the same mean. If p is near zero, it casts doubt on the null hypothesis and suggests that at least one sample mean is significantly different than the other sample means. Common significance levels are 0.05 or 0.01 in our case, 0.05.

C. Automatic classification

The WEKA software suite [15] has been used in carrying out the experiments. Support Vector Machines have been used for the automatic classification. The results were evaluated using Classification Error Rate (CER). For the training and validation steps, we used k -fold cross-validation with $k=10$. Cross-validation is a robust validation method for variable selection [16]. Repeated cross-validation (as calculated by the WEKA environment) allows robust statistical tests. We also use the measurement provided automatically by WEKA "Coverage of cases" (0.95 level) and Confidence Interval for percentages (CI) for 95%, 90% and 80%.

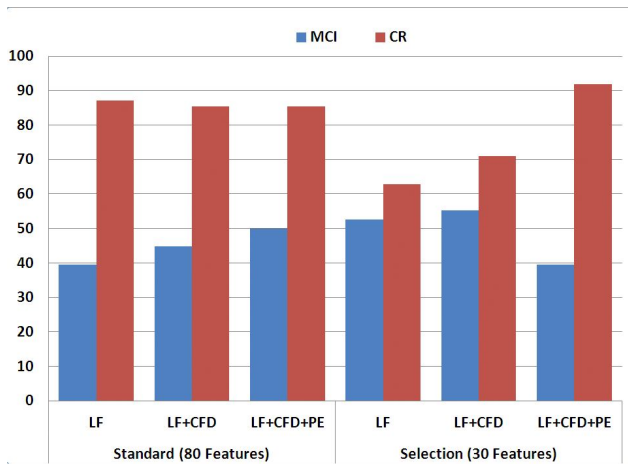


Figure 1. CER (%) for CR and MCI groups with SVM for standard features and selection of features by ANOVA.

IV. RESULTS AND DISCUSSION

The experimentation has been carried out with the balanced subset PGA-OREKA. This preliminary experimentation is divided in two stages.

In the first stage the linear and non-linear features described in Section III.A are used, about 80. Figure 1 shows the results in the left size.

- SVM with Linear Features (LF) obtain good results for CR but the results for MCI are about 40%.
- The inclusion of Castiglioni fractal dimensions improves the results for MCI.
- LF+CFD+PE looks as the best option for both CR and MCI and in this last case the rate achieves is about %50.

In the next stage a selection of features by ANOVA is carried out. The size set is reduced more than 50%. Figure 1 shows the results in the right size.

- The global results are improved in all cases.
- The best global option is LF+CFD+PE. In this case the results for CR are about 90% but for MCI decreases to 40%

V. CONCLUSION

This work, on selection of biomarkers from speech, is part of a wide-ranging cross study for the diagnosis of Alzheimer. Specifically in this work a task for detection of MCI has been used. The task analyzes Categorical Verbal Fluency. The automatic classification is carried out by SVM. Classical linear features, Castiglioni fractal dimension and Permutation Entropy are analyzed. Finally the most relevant features are selected by ANOVA test. The promising results are over 50% for MCI. In ongoing works new non-linear features, other entropy algorithms and semantic indexes will be used.

Acknowledgments

This work has been partially supported by the University of the Basque Country by UPV/EHU—58/14 project, SAIOTEK from the Basque Government, and the Spanish Ministerio de Ciencia e Innovación TEC2012-38630-C04-03.

References

- [1] McKahn G, et al. Clinical diagnosis of AD: report of the NINCDS-ADRDA Workgroup on AD. (1984); vol. 24, pp 939-944.
- [2] McKahn GM et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the NIAA Association workgroups on diagnostic guidelines for AD. *Alzheimers Dement.* (2011) May, vol. 7(3):263-9. [1] C. Pickering, J. Alsiö, J. Morud, M. Ericson, T.W. Robbins, B. Söderpalm, "Ethanol impairment of spontaneous alternation behavior and associated changes in medial prefrontal glutamatergic gene expression precede putative markers of dependence". *Pharmacol Biochem Behav.*, Vol. 132, pp. 63-70, Mar 2015.
- [3] Van de Pole, L.A., et al., The effects of age and Alzheimer's disease on hippocampal volumes, a MRI study. *Alzheimer's and Dementia*, (2005). 1(1, Supplement 1): pp. 51.
- [4] Alzheimer's Association.: <http://www.alz.org/>
- [5] Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Underdiagnosis of dementia in primary care: variations in the observed prevalence and comparisons to the expected prevalence. *Aging Ment Health* 2011; 15(8):978-984.
- [6] Laske C.,H.R. Sohrabi, S.M. Frost, K. López-de-Ipiña, P. Garrard, M. Buscema, J. Dauwels, S.R. Soekadar, S. Mueller, C. Linnemann, S. Bridenbaugh, Y. Kanagasinga, R:N. Martins, S.E. O'Bryant. Innovative diagnostic tools for early detection of Alzheimer's disease, *Alzheimer's and Dementia*, DOI: doi: 10.1016/j.jalz.2014.06.004
- [7] M. Faundez-Zanuy et al. Biometric Applications Related to Human Beings: There Is Life beyond Security, *Cognitive Computation*, 2012, DOI 10.1007/s12559-012-9169-9
- [8] Lopez-de-Ipiña K., J.B. Alonso, C.M. Travieso, J. Solé-Casals, H. Egiaraun, M. Faundez-Zanuy, A. Ezeiza, N. Barroso, M. Ecay, P. Martinez-Lage, and U. Martinez-de-Lizardui, 2013a. On the selection of non-invasive methods based on speech analysis oriented to Automatic Alzheimer Disease Diagnosis, *Sensors*, vol.: 13 (5) pp, 6730-6745
- [9] Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. *Neurol.* 1993;43:2412-2414.
- [10] Folstein, M. F., Folstein, S. E., & McHugh, P. R. "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12,189-198. 1975.
- [11] Buschke H. Cued recall in amnesia. *Journal of Clinical and Experimental Neuropsychology.* 1984;6:433-440.
- [12] Pfeffer, R.I., Kurosaki, T.T., Harrah, C.H. Jr., Chance, J.M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, 37(3), 323-329.
- [13] Praat: doing Phonetics by Computer. Available online: www.fon.hum.uva.nl/praat
- [14] B. Rosado, A. González-Martínez, P. Pesini, S. García-Belenguer, J. Palacio, A. Villegas, M.L. Suárez, G. Santamarina, M. Sarasa, "Effect of age and severity of cognitive dysfunction on spontaneous activity in pet dogs - part I: locomotor and exploratory behaviour". *Vet Journal*, vol. 194(2), pp. 189-95, November 2012.
- [15] MATLAB, www.mathworks.com
- [16] WEKA. Available online: <http://www.cs.waikato.ac.nz/ml/weka>
- [17] Picard, R.; Cook, D. Cross-Validation of Regression Models. *Journal of the American Statistical Association* (1984), vol. 79(387), pp. 575-583.

C3 ERANSKINA

Analysis of disfluencies for automatic
detection of Mild Cognitive Impairment: a
deep learning approach

Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach

K. López-de-Ipiña, U. Martínez-de-Lizarduy, P. M. Calvo, B. Beitia, J. García-Melero
Faculty of Engineering
20018, Donostia-San Sebastian Spain
Universidad del País Vasco/Euskal Herriko Unibertsitatea (UPV/EHU)
{karmele.ipina, unai.martinezdelizarduy, pilarmaria.calvo, mariablanca.beitia, jgarcia}@ehu.eus

M. Ecay-Torres, A. Estanga
Fundación CITA Alzheimer, 20009 Donostia, Spain
{mecay, aestanga}@cita-alzheimer.org

Marcos Faundez-Zanuy
Escola Superior Politècnica de Mataró (UPF),
Tecnocampus 08302 Mataró
faundez@tecnocampus.cat

Abstract—: The so-called Mild Cognitive Impairment (MCI) or cognitive loss appears in a previous stage before Alzheimer’s Disease (AD), but it does not seem sufficiently severe to interfere in independent abilities of daily life, so it usually does not receive an appropriate diagnosis. Its detection is a challenging issue to be addressed by medical specialists. This work presents a novel proposal based on automatic analysis of speech and disfluencies aimed at supporting MCI diagnosis. The approach includes deep learning by means of Convolutional Neural Networks (CNN) and non-linear multifeature modelling. Moreover, to select the most relevant features non-parametric Mann-Whitney U-testt and Support Vector Machine Attribute (SVM) evaluation are used.

Keywords— *Mild Cognitive Impairment, Automatic speech analysis, Deep Learning, Convolutional Neural Networks, Non-linear features, Disfluencies*

I. INTRODUCTION

The World Alzheimer Report 2015 highlights that about 900 million people can be considered as the world’s elderly population, and most of them live in developed countries [1]. AD is characterized by a progressive and irreversible cognitive deterioration including memory loss and impairments in judgment and language, along with other cognitive deficits and behavioral symptoms. An early and accurate diagnosis of AD helps patients and their families to plan the future, and offers the best possibilities of symptoms being treated. The cognitive loss appears in a previous stage, the so-called Mild Cognitive Impairment (MCI), but it does not seem sufficiently severe to interfere in independent abilities of daily life, so it usually does not receive an appropriate diagnosis. Its detection is a challenging issue to be addressed by medical specialists [2]. Along with memory loss, one of the

major problems of AD is the loss of language skills. This loss is reflected in difficulties speaking to and understanding other people, which makes even more difficult the natural communication process and social interactions. This inability to communicate appears in early stages of the disease due to language difficulties, and it leads to social exclusion of people with AD, and to a serious negative impact not only on the sufferers, but also on their relatives [3]. In this sense, disfluencies are interesting language elements, which could be very useful to properly detect MCI. Both speech silences and disfluencies have valuable information for decoding the meaning of the uttered message.

The main aim of this project is the development of an automatic analysis of standard assessment tests such as categorical verbal fluency (CVF) by means of speech therapy techniques that will allow to obtain reliably and quickly these specific analyses [4]. Last years several works in the state of the art have addressed this. In the present paper we focus on the integration of language independent methodologies in order to detect AD in speech, with a task of Categorical Verbal Fluency (CVF), that is called animals naming task.

In section II the materials are described. The used methods are presented in section III. Section IV comprises the results and discussion, and finally in section V concluding remarks are drawn.

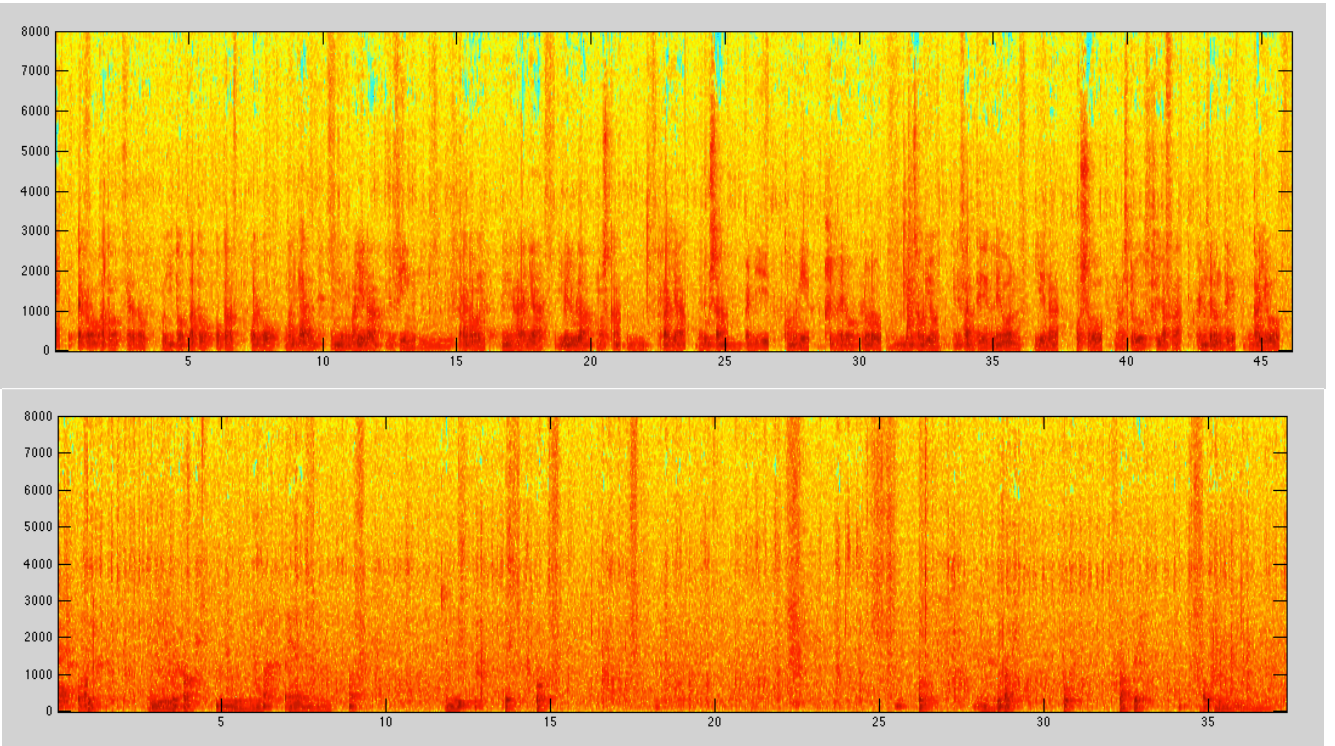


Figure 1. Categorical Verbal Fluency (CVF) task, for an individual of the MCI group (a) speech, (b) disfluencies

II. MATERIALS

Recent works point out the relevance of disfluencies in speech to identify MCI and AD. In [5] it is suggested that shorter recording times reflect that AD patients require a greater effort to produce speech than healthy individuals: AD patients speak more slowly with longer pauses, and they spend more time finding the correct word, which leads to speech disfluencies or broken messages. Speech disfluencies are any break, irregularity or non-lexical element which occurs within the period of fluent speech, and that could start or interrupt it. These include among others: false starts, repeated or re-started phrases, repeated or extended syllables, grunts or non-lexical utterances such as fillers and repaired utterances, and instances of speakers correcting their own slips of the tongue or mispronunciations [6]. An increase in these disfluencies could be a clear sign of cognitive impairment, and in AD patients sometimes they become a verbal utterance of the internal cognitive process or an inner dialogue: “What is that?”, “What was the

name?”, “/uhm/ “I can’t remember”. The increase in the number of disfluencies and silences may point to a worsening of the disease, and could lead to a deficit in clear communication. As a conclusion, disfluencies are a direct reflection of the cognitive process during communication, and convey an unquestionable characteristic for the detection of these disorders. Although AD is mainly a cognitive disease, it may have phonation and articulation biomechanical alterations. The Categorical Verbal Fluency task (CVF, animal naming, AN), or animal fluency task, is a test for neurodegenerative diseases that measures and quantifies the progression of cognitive impairment [7], Fig. (1). It is commonly used in order to assess language skills, semantic memory and executive functions [8]. The sample includes 187 healthy people and 38 MCI patients that belong to the cohort of Gipuzkoa-Alzheimer Project (PGA) of the CITA-Alzheimer Foundation [4,9], Table (1). A balanced subset PGA-OREKA was selected for experimentation.

Tab. (1). Demographic data of the subsets selected for the experimentation

	Female	Male	Range of age	Age-Mean	Age-SD
PGA-OREKA	36/21	26/17	39-74/42-79	56.73/57.15	7.8/8.9

III. METHODS

The new speech analysis approach is based on the integration of several kinds of features in order to model speech and disfluencies: linear and non-linear ones. Moreover, this approach is based on the description of speech pathologies with regard to phonotaxis, articulation, speech quality, human perception, and the complex dynamics of the system. In this work, we will use some of the most used speech features for both differentiation between healthy and pathological speech [4,5,10,11], and discrimination through human perception. Most of them are well known in the field of speech signal processing, and thus for each parameter a reference is provided where a deeper description and further information can be found. All features are calculated using software developed in our research group [4,5], MATLAB [12] and Praat [13].

A. Automatic disfluency segmentation

The recording has been automatically segmented in speech signal and disfluencies by means of a VAD (Voice activity detection) algorithm.

B. Feature extraction

Next, the analyzed features are described.

- *Classical features (CF):*

1. Spectral domain features: harmonic to noise ratio (HNR), noise to harmonic ratio (NHR), harmonicity, pitch, jitter, shimmer, APQ (Amplitude Perturbation Quotient); spectrum centroid and formants and its variants (mean, median, min, max, mode, std) [4-5].
2. Time domain features: Voiced/unvoiced segments, breaks, ZCR (Zero-Crossing Rate) [4-5] and its variants.
3. Energy, intensity, short time energy, and spectrum centroid [7-10, 12, 13].

- *Perceptual features (PF):*

1. Mel Frequency Cepstrum Coefficients (MFCC): Human ear behaves as some filters, it only concentrates on some components of frequency. These filters are not spaced uniformly on the frequency axis. There are more filters at low frequencies,

and fewer filters at high frequencies. This kind of performance is simulated by means of Mel-Frequency analysis and particularly Mel Frequency Cepstrum Coefficients (MFCC) [4,5,14].

2. Coefficients that provide information related to voice quality, perception, adaptation or amplitude modulation: MSC (Modulation Spectra coefficients), PLP (Perceptual Linear Predictive coefficients), LPCC (Linear Predictive Cepstral coefficients), LPCT (Linear Predictive Cosine Transform coefficients) and ACW (Adaptive Component Weighted coefficients), ICC (Inferior Colliculus Coefficients). These features are sometimes extended by their 1st and 2nd order regression coefficients (Δ and $\Delta\Delta$ respectively) [10,14].

- *Non-linear features (NLF):* Fractal Dimension, Shannon Entropy and Multiscale Permutation Entropy have been calculated [4,5,14].

C. Automatic selection of features by Mann-Whitney U-test

In this step the best features are automatically selected with respect to common significance level. Therefore, the automatic selection of features is performed by Mann-Whitney U-test being p-value < 0.1 in order to get a bigger set for the second feature selection phase [14].

D. Automatic selection of features by WEKA

Then a new selection phase is carried in WEKA by the *SVMAttributeEval* algorithm. This provides a selection by analyzing the features group.

E. Feature normalization by WEKA

During the preprocessing of data, all the features will be normalized by WEKA.

F. Automatic classification

Four classifiers will be used: k-nearest neighbors (k-NN), Support Vector Machines, Multilayer Perceptron (MLP) with L layers and N neurons, and a Convolutional Neural Network (CNN) with L layers of N neurons, a convolution of cxc and a pool of pxp . We have used the WEKA software suite [15] to perform the experiments. Classification

Error Rate (CER, %) has been used to evaluate the results. We have used k-fold cross-validation with $k=10$ for training and validation [15,16].

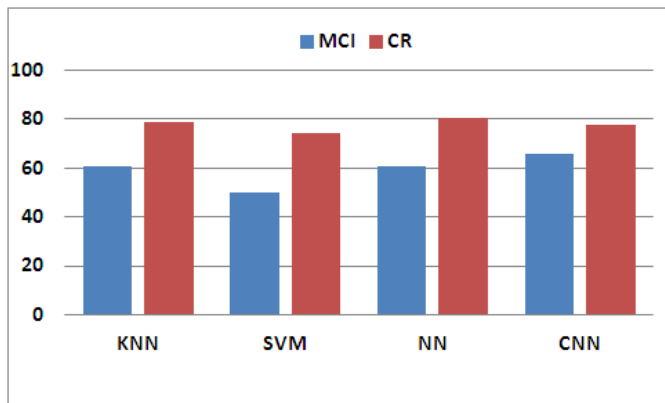


Figure 2. CER (%) for classes and selected classifiers: k-nearest neighbors (k-NN), Support Vector Machines (SVM), Multilayer Perceptron (NN) with L layers and N neurons and a Convolutional Neural Network (CNN)

IV. RESULTS

In the experimentation the used materials are about 40 speech samples for the MCI group and 60 for the control group (CR), PGA-OREKA (Table 1). Initially, the obtained number of features is about 920 (473 for speech and 447 for disfluencies) for a sampling frequency of 22000 KHz. Then, after a normalization test, an automatic feature selection is performed based on a non-parametric Mann-Whitney U-test with a p -value < 0.1 , and about 150 features are selected. In the second optimization step, the attribute selection algorithm *SVMAttributeEval* of WEKA yields about 80 features that are finally selected. The proposed feature set includes features from all the feature types described in subsection III.B for speech and disfluencies. Figure 2 shows the results of the automatic classification by classes, CR and MCI. CER (%) is evaluated for all the classifiers detailed in subsection III.F. The new proposal that includes integration of disfluency analysis outperforms previous works [4] for most of the classifiers. The results are hopeful, stable, good and equilibrated for all of them about %75. The deep learning option with CNN yields the best results for a configuration of 2 layers of 20 neurons, a convolution of 3×3 and a pool of 2×2 . This option outperforms MLP for 2 layers of 100 neurons.

V. CONCLUSIONS

This work presents a novel proposal based on automatic analysis of speech and disfluencies in order to support MCI diagnosis. A non-linear multifeature modeling is presented based on selection of the most relevant features by statistical tests (under medical criteria) and automatic attribute selection: Mann-Whitney U-test and Support Vector Machine Attribute (SVM) evaluation. The approach includes deep learning by means of Convolutional Neural Networks (CNN). The results are hopeful and open a new research line.

ACKNOWLEDGEMENT

This work has been supported by FEDER and MICINN, TEC2016-77791-C4-2-R and UPV/EHU-Basque Research Groups IT11156.

REFERENCES

- [1] World Alzheimer Report 2015. Available online: 2015-12-16, www.alz.co.uk/research/world-report-2015, (accessed on 15 December 2015).
- [2] Laske C., Sohrabi H.R., Frost S.M., López-de-Ipiña K., Garrard P., Buscema M., Dauwels J., Soekadar S.R., Mueller S., Linnemann C., S.A. Bridenbaugh, Y. Kanagasigam, R.N Martins, S.E. O'Bryant, Innovative diagnostic tools for early detection of Alzheimer's disease, *Alzheimer & Dementia*, 11(5):561-78, doi:10.1016/j.jalz.2014.06.004, (2015)
- [3] Klimova B., Maresova P., Valis M., Hort J., Kuca K. Alzheimer's disease and language impairments: social intervention and medical treatment, *Clinical Interventions in Aging, Clinical Interventions in Aging* 2015:10 1401-1408, <http://dx.doi.org/10.2147/CIA.S89714>, (2015).
- [4] Lopez-de-Ipiña K., U. Martinez-de-Lizarduy, N. Barroso, M. Eca-Torres, Martinez-Lage, P., F. Torres, M. Faundez-Zanuy. Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impairment detection: A non-linear language independent approach. *Bioinspired Intelligence (IWOB)*, 2015 4th International Work Conference, pp. 101 - 104, Donostia (Spain) (2015).
- [5] Lopez-de-Ipiña K, Alonso J-B, Manuel Travieso C, Sole-Casals J, Egiraun H, Faundez-Zanuy M, et al. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors* 2013;13:6730-45.
- [6] Dingemans, Mark; Torreira, Francisco; Enfield, N. J. (2013). "Is 'Huh?' a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items". *PLOS ONE* 8 (11): e78273. doi:10.1371/journal.pone.0078273.
- [7] Lezak, MD, Howieson DB, Bigler ED, Tranel D. *Neuropsychological Assessment* (fifth edition). Oxford University Press (2012).
- [8] Ruff RM, Light RH, Parker SB, Levin HS. The psychological construct of word fluency. *Brain Lang.* 57: 394-405 (1997).
- [9] CITA-Alzheimer Foundation, PGA project: <http://www.cita-alzheimer.org/investigacion/proyectos>, (accessed on 10 June 2016).
- [10] Mekyska J et al.. Robust and complex approach of pathological speech signal analysis. *Neurocomputing* 167(2015): 94-111 2015.
- [11] Gómez-Vilda P. et al. Phonation biomechanics analysis of Alzheimer's Disease cases. *Neurocomputing*, 167(1) November 2015, pp 83-93
- [12] MATLAB. www.mathworks.com
- [13] Praat: doing Phonetics by Computer. www.praat.org/
- [14] López-de-Ipiña K et al.. Advances in a Multimodal Approach for Dysphagia Analysis Based on Automatic Voice Analysis. *Advances in Neural Networks* 54 of the series Smart Innovation, Systems and Technologies: 201-211, 2016.

- [15] WEKA. <http://www.cs.waikato.ac.nz/ml/weka>
- [16] Picard, R.; Cook, D. Cross-Validation of Regression Models. *Journal of the American Statistical Association* (1984), vol. 79(387), pp. 575–583

C4 ERANSKINA

Non-invasive Biosignal Analysis Oriented to
Early Diagnosis and Monitoring of
Cognitive Impairments

Non-invasive biosignal analysis oriented to early diagnosis and monitoring of cognitive impairments

K. López-de-Ipiña, J. Solé-Casals, U. Martínez de Lizarduy, P. M. Calvo, J. Iradi, M. Faundez-Zanuy, A. Bergareche

Abstract— In this paper some of novel and low-cost methodologies are analysed for most common cognitive impairments: Alzheimer disease, Parkinson and Essential Tremor. These methodologies allow us to obtain information from the brain in an indirect form, without needing to do neither invasive nor expensive or long-time tests. We will present handwriting and speech as interesting biomarkers for the early detection and monitoring of cognitive impairments and detail possible fields of application, with real experiments carried-out at the Hospital Donostia (Spain) in several past and current projects.

I. INTRODUCTION

LIFE expectancy continues increasing whereas mortality rate in elderly population is falling in most of the world regions [1]. As a consequence, people's life is longer, and dementia disorders and chronic diseases are becoming more prevalent due mainly to a trend towards non-healthy. In this context, world's population seems in transition to a dementia epidemic that could have a dramatic impact in the long term [1]. Alzheimer's disease, Parkinson or Essential Tremor, are some of the most common cause of cognitive impairment among elderly people. These diseases imply large socioeconomic costs, which are expected to increase in the near future.

The diagnosis of diseases is based on clinical criteria and requires, on the one hand, the confirmation of a progressive dementia syndrome, and on the other hand, the exclusion of other potential causes by inspection of the clinical history [2]. The diagnosis is carried out by applying general health examination, complete blood workup tests, long neuropsychological tests, expensive neuroimaging techniques, or invasive and unpleasant tests such as lumbar puncture [3][4][5]. Nonetheless, the diagnosis of early stages of diseases is still problematical, since patients and relatives tend to either ignore/hide the first clinical manifestations, or

The present work is funded by the grant TEC2012-38630-C04-03 from Plan Nacional Economic Affairs and Competitiveness of Spain and from University of the Basque Country, UPV/EHU 16/48.

KLI, UML, PML and JI are with the UPV/EHU (Basque Country University) Donostia, Spain ({karme.le.ipina, unai.martinezdelizarduy, pilar.calvo, jon.iradi}@ehu.eus).

JSC is with the U Science Tech of the University of Vic-Central University of Catalonia, Vic, Catalonia, Spain (jordi.sole@uvic.cat).

MFZ is with the Tecnocampus of Universitat Pompeu Fabra, Barcelona, Catalonia, Spain (faundez@tecnocampus.cat).

AB is with the Hospital Donostia, Biodonostia Health Research Center, Donostia, Spain (jesusalbertobergarecheyarza@osakidetza.net).

ascribe them to typical cognitive changes related to age. Consequently, these patients are often diagnosed when they have already reached moderate or advanced phases, which complicate treatment strategies and their efficacy

In addition, all symptoms lead to impaired performance in everyday activities as the disease progresses from mild to moderate or severe. In the past few years, approaches to early diagnosis of diseases have made significant advances in the development of reliable clinical biomarkers and robust screening/diagnosis and/or monitoring tools [6]. However, despite the usefulness of biomarkers, the cost and technology requirements involved make it impossible to apply such tests to all patients with memory problems or to the general population. As a result, they are applied to preselected patients by low-cost screening based on the probability of suffering an underlying pathology **¡Error!No se encuentra el origen de la referencia.**[8][9].

Undoubtedly, an optimum pre-screening is essential, and non-invasive intelligent monitoring techniques might become valuable tools for early detection in order to complement diagnosis and/or to evaluate disease progression. These tests do not require specialized personnel or laboratory equipment, so that anyone in the habitual environment of the patient could perform them after proper training without altering or blocking the patient's abilities. Moreover, these techniques are very low-cost and do not need extensive infrastructures or the availability of medical resources. They are thus capable of managing information easily, quickly and inexpensively **¡Error!No se encuentra el origen de la referencia.**[8][9].

However, although these techniques are not perceived as stressful by patients, some authors suggest that non-professional caregivers might lack the accuracy of medical staff to identify exactly early cognitive impairment [6].

II. HANDWRITING SIGNALS

In our research, the acquisition is carried out using an Intuos Wacom 4 digitizing tablet. This USB pen tablet captures the information depicted in Figure 1. The device acquires 100 samples per second including the spatial coordinates (x, y) , the pressure, and a couple of angles:

- Position in x-axis.
- Position in y-axis.
- Pressure applied by the pen.
- Azimuth angle of the pen.

- Altitude angle of the pen.

Using this set of dynamic elemental features, further information can be inferred. For instance: acceleration, velocity, instantaneous trajectory angle, instantaneous displacement, tangential acceleration, curvature radius, centripetal acceleration, etc. Figure 1 shows a modern digitizing tablet, along with a graphical depiction of the azimuth and altitude angles, while figure 2 shows examples of drawings and text captured by the device. Notice that all three captures include both in-air (red colour) and on-surface (blue or black colour) trajectories.

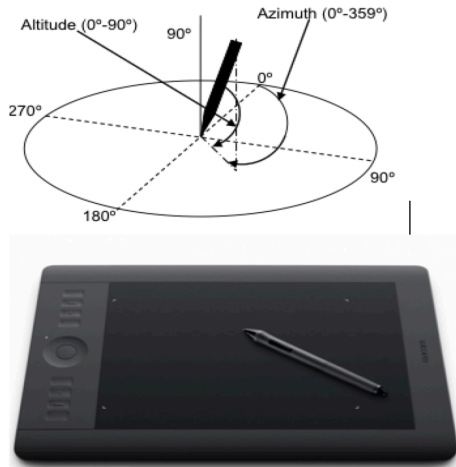


Fig. 1. On the top, information extracted from the digitizing tablet. On the bottom, an Intuos Wacom 4 digitizing tablet.

III. SPEECH SIGNALS

In the case of speech signals, we will record the signals and pre-process them in order to increase SNR ratio. Features extracted from speech frames will be diverse:

- Acoustic features: pitch, standard deviation pitch, max and min pitch, intensity, standard deviation intensity, max and min intensity, period mean, period standard deviation, and Root Mean Square amplitude (RMS).
- Voice quality features: shimmer, local jitter, Noise-to-Harmonics Ratio (NHR), Harmonics-to-Noise Ratio (HNR) and autocorrelation.
- Duration features: voice segment length, pause length, short time energy and centroid, Voiced/unvoiced segment average, voiced/voiceless percentage and spontaneous speech evolution along the time.
- Entropy, fractal dimension, Emotional Temperature.

IV. CONCLUSION

As shown in many studies, most of the cognitive impairments have an effect on the motor areas of the brain. Therefore, any activity involving these areas will be affected. The presented techniques allow us to obtain information of the cognitive state of the subjects. Due to its

non-invasive intrinsic characteristic, its simplicity and low price, we believe can be very useful in helping for the early diagnosis of cognitive impairments. According to the opinion of the health specialists the biomarkers obtained with these techniques are relevant and potentially applicable to the early diagnosis of several diseases. Speech can be useful for detecting Alzheimer's disease and Parkinson. Handwriting can be also useful for Alzheimer's disease, Parkinson but also Essential Tremor. Finally, it must be emphasized that the use of this technology could provide undoubted benefits towards the development of more sustainable, low cost, high quality, non-invasive technologies easily adaptable to the users and their environments. From a social and economic point of view they can be very useful in real complex environments. In future research other non-linear features and automatic selection algorithms will be explored.

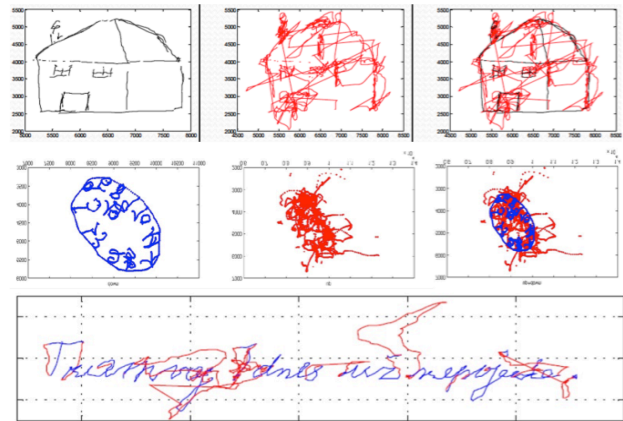


Fig. 2. An example of handwriting captures. On the top and middle, drawings performed by an AD patient. On the bottom, by a Parkinson's disease patient. Red trajectories corresponds to in-air, blue/dark to on-surface.

REFERENCES

- [1] World Alzheimer Report 2015. Available online (accessed on 30 March 2016): <http://www.alz.co.uk/research/world-report-2015>
- [2] McKhann G., Knopman D.S., Chertkow H., et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:263-269.
- [3] Morris J.C. The Clinical dementia rating (CDR): Current version and scoring rules. *Neurology.* 1993;43:2412b-2414b.
- [4] Van de Pole L.A., Van der Flier W.M., Hensel A., et al. The effects of age and Alzheimer's disease on hippocampal volumes, a MRI study. *Alzheimer's and Dementia. J. Alzheimer's Assoc.* 2005; doi: 10.1016/j.jalz.2005.06.205
- [5] Alzheimer's Association/Biomarkers. Available online (accessed on 30 March 2016): http://www.alz.org/research/funding/global_biomarker_consorrtium.asp.
- [6] Laske C., Sohrabi H.R., Frost S.M., et al. O'Bryant, Innovative diagnostic tools for early detection of Alzheimer's disease, *Alzheimers Dement.* 2015 May;11(5):561-78. doi: 10.1016/j.jalz.2014.06.004.
- [7] M.P. Broderick, A.W.A. Van Gemmert, H.A. Shill, G.E. Stelmach. Hypometria and bradykinesia during drawing movements in individuals with Parkinson's disease. *Experimental Brain Research* 2009, Volume 197, Issue 3, pp 223-233
- [8] Lopez-de-Ipina K, Alonso J-B, Manuel Travieso C, et al. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. *Sensors* 2013;13:6730-45.

- [9] López de Ipiña K, Iturrate M, Calvo PM, Beitía B, et al. Selection of entropy based features for the analysis of the Archimedes' spiral applied to essential tremor, 4th International Work Conference on Bioinspired Intelligence (IWOBI), 2015, pp. 157-162 doi: 10.1109/IWOBI.2015.7160160

D1 ERANSKINA

Analysis of Spontaneous Speech and
Emotional Response oriented to
Alzheimer's Disease Diagnosis

Analysis of Spontaneous Speech and Emotional Response oriented to Alzheimer's Disease Diagnosis

K. Lopez-de-Ipiña¹, J.B. Alonso², J. Solé-Casals³, N.Barroso¹, P.Henriquez², M.
Faundez-Zanuy⁴, C. Travieso², M. Ecay-Torres⁵, P.Martinez-Lage⁵, U. Martinez-de-
Lizardui¹, H. Egiraun^{1,5}, A. Ezeiza¹

¹System Engineering and Automation Department, University of the Basque Country, Donostia
20008, Spain,
{karmele.ipina, nora.barroso, unai.martinezdelizarduy, harkaiz.egiraun, aitzol.ezeiza}@ehu.es

²Universidad de Las Palmas de Gran Canaria, IDETIC,
{jalonso,ctravieso,phernandez}@dsc.ulpgc.es

³Digital Technologies Group. University of Vic,
jordi.sole@uvic.cat

⁴Escola Universitaria Politècnica de Mataró (UPC), Tecnocampus ,
faundez@tecnocampus.cat

⁵Research Center for Experimental Marine Biology and Biotechnology,
Plentzia Marine Station, University of the Basque Country, Plentzia, Spain

Abstract. Alzheimer's disease is the most prevalent form of progressive degenerative dementia; it has a high socio-economic impact in Western countries. Therefore it is one of the most active research areas today. Alzheimer's is sometimes diagnosed by excluding other dementias, and definitive confirmation is only obtained through a post-mortem study of the brain tissue of the patient. The work presented here is part of a larger study that aims to identify novel technologies and biomarkers for early Alzheimer disease detection, and it focuses on evaluating the suitability of a new approach for early diagnosis of Alzheimer's disease by non-invasive methods. The purpose is to examine, in a pilot study, the potential of applying Machine Learning algorithms to speech features obtained from suspected Alzheimer sufferers in order help diagnose this disease. Two human capabilities relevant in communication have been analyzed for feature selection: Spontaneous Speech and Emotional Response. A novel parameter is used Emotional Temperature. The experimental results obtained were very satisfactory and promising for the early diagnosis and classification of Alzheimer's disease patients.

D2 ERANSKINA

ALZUMERIC: a decision support System for
diagnosis and monitoring of cognitive
impairment

ALZUMERIC: A DECISION SUPPORT SYSTEM FOR DIAGNOSIS AND MONITORING OF COGNITIVE IMPAIRMENT

U. Martinez de Lizarduy, P. Calvo, P. Gómez Vilda, M. Ecay-Torres, K. López de Ipiña

¹UPV/EHU (Basque Country University) Donostia, [unai.martinezdelizarduy, [karmele.ipina](mailto:karmele.ipina@ehu.eus), [pilarmaria.calvo](mailto:pilarmaria.calvo@ehu.eus)][@ehu.eus](mailto:), ORCID: 0000-0002-2135-2393, 0000-0003-0969-4857

²P. Gómez-Vilda Centro de Tecnología Biomédica. UPM

³ Fundacion CITA-Alzheimer Fundazioa, Donostia-San Sebastian, Spain, mecay@cita-alzheimer.org.

⁴ Department of Basic Psychology, Psychology Faculty. UNED, Madrid, Spain

ABSTRACT

Internet of things or smart cities are becoming a reality. Nowadays, more and more devices are interconnected and to deal with this new situation data processing speeds are increasing to keep the pace. Smart devices like tablets and smartphones are accessible to a wide part of society in developed countries, and Internet connections for data exchange make it possible to handle large volumes of information in less time. This new reality has opened up the possibility of developing client-server architectures focused on clinical diagnosis in real time and at a very low cost. This paper illustrates the design and implementation of the *ALZUMERIC* system that is oriented to the diagnosis of Alzheimer's disease (AD). It is a platform from which the medical specialist can gather voice samples through non-invasive methods from patients with and without mild cognitive impairment (MCI), and the system automatically parameterizes the input signal to make a diagnostic proposal. This type of impairment, although produces a cognitive loss, it is not severe enough to interfere with daily life. The present approach is based on the description of speech pathologies with regard to the following profiles: phonation, articulation, speech quality, analysis of the emotional response, language perception, and complex dynamics of the system. Privacy, confidentiality and information security have also been taken into consideration, as well as possible threats that the system could suffer, so this first prototype of services offered by *ALZUMERIC* has been targeted to a predetermined number of medical specialists.

Keywords: Alzheimer's Disease, Mild Cognitive Impairment, Automatic Speech Processing, Decision Support Systems, Convolutional Neural Networks.

RESUMEN

La Internet de las cosas o las ciudades inteligentes se están convirtiendo en una realidad. Cada vez más dispositivos están interconectados y para hacer frente a esta nueva situación las velocidades de procesamiento de datos se incrementan. Los dispositivos inteligentes como las tabletas y los teléfonos son accesibles para una gran parte de la sociedad en los países desarrollados, y las mejoras en las conexiones a Internet para el intercambio de datos hacen posible manejar grandes volúmenes de información en menos tiempo. Esta nueva realidad ha abierto la posibilidad de desarrollar arquitecturas cliente-servidor centradas en el diagnóstico clínico en tiempo real y a un coste muy bajo. Este trabajo ilustra la concepción del sistema *ALZUMERIC* orientado al diagnóstico de la enfermedad del Alzheimer. Es una plataforma a partir de la cual el médico especialista puede tomar muestras de voz a través de métodos no invasivos a pacientes con y sin deterioro cognitivo leve (MCI), en la cual se parametriza la señal de entrada automáticamente para posteriormente hacer una propuesta de diagnóstico. El MCI es un tipo de deterioro neurológico que produce una pérdida cognitiva no lo suficientemente grave como para interferir en la vida cotidiana. El presente estudio está enfocado en base a la descripción de las patologías del habla con

respecto a los siguientes perfiles: fonación, articulación, calidad del habla, análisis de la respuesta emocional, la percepción del lenguaje, y dinámica de sistemas complejos. También se consideran aspectos de la privacidad, la confidencialidad y la seguridad de la información frente a las posibles amenazas que pudiera sufrir el sistema, por lo que este primer prototipo de servicios ofrecidos por *ALZUMERIC* se ha dirigido a un número predeterminado de médicos especialistas.

Palabras clave: Enfermedad de Alzheimer, deterioro cognitivo leve, procesamiento de voz automática, sistemas de apoyo a la decisión, Redes Neuronales convolucionales.

1. INTRODUCTION

In an ever-changing world, where digital interactions are part of daily life and Internet of things or smart cities are becoming a reality in many fields, health services start demanding new interactive and digital solutions where the final user will play an active role. Besides, classic clinical practice relies increasingly on technological support for making, which serves as a help for specialists, reducing the time to perform a reliable diagnosis, and substantially improves health resources by reducing medical testing time and costs (Alzheimer's Disease International, 2015).

In this scenario, the dramatic aging of the world population is causing cognitive impairment to reach epidemic level. In this sense, its early detection by premature symptoms would be crucial to optimize management, and could open a new scenario for potential interventions. Doubtless, an early and accurate diagnosis of this kind of diseases could help decreasing its social impact and consequences. Thus, early clinical tests should be reinforced and consistent with complementary evaluations, rehabilitation and monitoring. Over the last decades there have been useful advances not only in classic assessment techniques but also in novel screening strategies for preliminary assessment or monitoring during the evolution of the disease. Non-invasive intelligent techniques of diagnosis may become valuable tools for

clinical and domestic environments that can complement diagnosis and/or monitor its progress. These methodologies are easy to apply, and can be introduced in the standard medical protocols without altering daily life. Moreover, they are capable of yielding information easily, quickly and inexpensively (Faundez-Zanuy et al., 2013; López de Ipiña et al., 2013; Laske et al., 2015; Klimova et al., 2015). Thus, even non-technical staff in the usual patient's environment could use these methodologies without altering or blocking the patient abilities, but a medical supervision is required (Laske et al., 2015). Among non-invasive methods, automatic analysis of speech—which involves verbal communication—one of the first damaged skills in MCI/AD patients—, provides a natural, friendly and powerful tool for early diagnose and low cost screening.

This work presents the conception and first evaluation of the *ALZUMERIC* system, a low cost technological platform oriented to the diagnosis of Alzheimer's disease by speech analysis. It is a platform from which the medical specialist can gather voice samples through non-invasive methods from patients with and without mild cognitive impairment (MCI) and AD. The server parameterizes the input signal for medical assessment and it makes an evaluation so that the patient is assigned to the group with MCI or excluded from it. The score of diagnosis has substantially improved with deep learning algorithms. This diagnosis is crucial because it may allow treating the disease at an early stage. The entire process is monitored in real time and is based on protocols included under TCP/IP. Privacy and information security have also been taken into consideration, as well as possible threats that the system could suffer, so this first prototype of services offered by *ALZUMERIC* has been targeted to a predetermined number of medical specialists. In this framework, HTTP/HTTPS clients, on web protocols, are responsible for the coordination and exchanges of information between the final user and the servers, and are integrated into all terminals with Internet access. They are the most important part of the

tool. It is likely that the expansion of desktop and mobile devices for clinical use will keep increasing, and since their use is widespread, the platform has a really short learning curve. The analysis of the information matrix to obtain a reliable diagnosis is based on a set of algorithms of deep learning that achieves substantial improvements on results.

The paper is organized as follows. Materials are described in section 2. Section 3 summarizes the used methods. The protocols and algorithms are discussed in section 4. The results and discussion are summarized in section 5, and finally concluding remarks are drawn in section 6.

2. MATERIALS

In the development and clinical evaluation protocol of *ALZUMERIC*, three tasks with different levels of language complexity have been used: Categorical Verbal Fluency (CVF, based on Animal Naming, AN), Picture Description (PD), and Spontaneous Speech (SS). AN and PD have been recorded in a clinical environment, and SS has been carried on a domestic environment. The participants are different in the three tasks. PD and SS analyze AD patients vs. individuals from a control group (CR), and AN analyzes MCI patients vs. CR. All the work was performed strictly following the ethical guidelines of the organizations involved in the project.

2.1. Task of *Categorical Verbal Fluency* (CVF), based on animal naming

The task of *Categorical Verbal Fluency* (CVF, or Animal naming, AN), is a test that measures and quantifies the progression of cognitive impairment in neurodegenerative diseases (López de Ipiña, 2015). It is widely used to assess language skills, semantic memory and executive functions. During the CVF task the interviewer asks patients to list in one minute all the names they can remember from a category, these being animal names in our case. The sample recruited 187 healthy controls and 38 patients with MCI, from the cohort of Gipuzkoa-Alzheimer Project (PGA, CITA-PGA, 2017) of the CITA-Alzheimer Foundation.

2.2. Picture description task

This pilot study which is part of the Gipuzkoa-Alzheimer Project (PGA) and the Memory Clinic of the CITA-Alzheimer Foundation includes six subjects with diagnosis of AD and 12 healthy controls (sub-database MINI-PGA), that is, the sample included data from 18 subjects. The task consists in the verbal description of a picture (CITA-PGA, 2017).

2.3. Spontaneous speech task

In order to develop a new methodology applicable to a wide range of individuals of different sex, age, language, and cultural and social background, a multicultural and multilingual spontaneous speech database has been built in English, French, Spanish, Catalan, Basque, Chinese, Arabic and Portuguese, comprising video recordings of 50 healthy people (12 hours), and 20 AD patients (8 hours), i.e., patients with a prior diagnosis. The age span of the individuals in the full database is 20-98 years. This database is called *AZTIAHO* (López de Ipiña et al., 2015). The complete speech database consists of about 60 minutes for the AD group, and about 9 hours for the control group. Hereafter, the speech was divided into consecutive segments of 60 seconds in order to obtain appropriate segments for all speakers, yielding a database of approximately 600 segments of spontaneous speech. Finally, in order to perform our experiments a subset was selected, balanced with regard to the age of the participants and the emotional response level. The subset consists of 20 subjects (9 women and 11 men) from the control group, and 20 AD patients (12 women and 8 men). This subset of the database is called *AZTIAHORE*.

3. METHODS

Recent works point out the relevance of dysfluency in speech as a hallmark of MCI and AD. In (López de Ipiña et al., 2015) it is suggested that shorter speech segments reflect that AD patients require a greater effort to produce speech than healthy individuals:

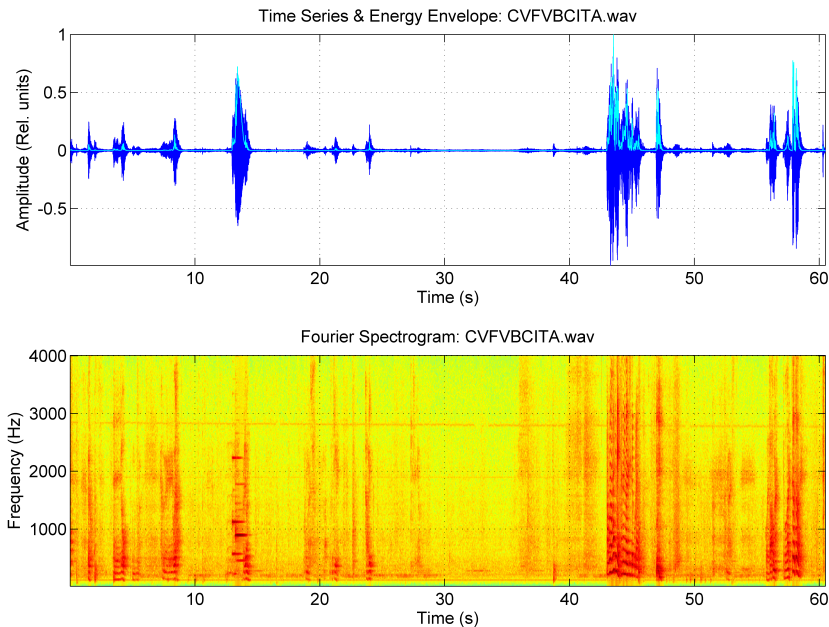


Figure 1. Categorical Verbal Fluency (CVF) task, individual of the MCI group

AD patients speak more slowly with longer pauses, and they spend more time finding the correct word, which leads to speech dysfluencies or broken messages. Speech dysfluencies are any break, irregularity or non-lexical element which occurs within the period of fluent speech, and that could start or interrupt it.

These include among others, false starts, repeated or re-started phrases, repeated or extended syllables, grunts or non-lexical utterances such as fillers and repaired speech, and instances of speakers correcting their own lips or mispronunciations (López de Ipiña et al., 2017; Dingemanse et al. 2013). In AD patients sometimes they become a verbal utterance of the internal cognitive process or an inner dialogue and they produce sentences such as: “What is that?”, “What was the name?”, “/uhm/ “I can’t remember”. The increase in the number of dysfluencies and silences may point to a worsening of the disease, which could lead to a deficit in clear communication. As a conclusion, dysfluencies are a direct reflection of the cognitive process during communication, and become an unquestionable hallmark for the detection of

these disorders. Although AD is mainly a cognitive disease, it may also present phonation and articulation biomechanical alterations. The task of Categorical Verbal Fluency (CVF, animal naming, AN), or animal fluency task, is a test that measures and quantifies the progression of cognitive impairment in neurodegenerative diseases (Lezak et al., 2012; Ruff et al., 1997). The CVF protocol is widely used to assess language skills, semantic memory and executive functions (Lezak et al., 2012). In Fig. 1 two examples collected using this protocol may be seen. In Fig. 1.a the production of a control subject is presented. The subject can produce as many as xx names in 46 s, whereas in Fig. 1.b the production of an MCI subject is of only yy names in 37 s.

Table 1. Age distribution of the subsets selected for the experimentation

	PGA-OREKA AN task-subset (CR/MCI)
Females	36/21
Males	26/17
Age range	39-74/42-79
Age mean	56.73/57.15
Age std. dev.	7.8/8.9

The data sample produced using CVF consists of 187 healthy controls and 38 MCI patients being a subset of the cohort of Gipuzkoa-Alzheimer Project (PGA of the CITA-Alzheimer Foundation (López de Ipiña et al., 2015;CITA-PGA, 2017). Age distributions are given in Table 1. A subset from PGA-OREKA was selected for experimentation.

4. PROTOCOLS AND ALGORITHMS

The speech analysis approach to evaluate the CVF protocol is based on the integration of several kinds of features in order to model speech and disfluencies: linear and non-linear ones. Besides, this approach is based on the description of speech pathologies with regard to phonation, articulation, speech quality, human perception, and the complex dynamics of the system. In this work, we will make use some of the most efficient speech features for both differentiations between healthy and pathological speech in the state of the art (Gómez-Vilda et al., 2015;López de Ipiña et al., 2016; López de Ipiña et al., 2017). Most of them are well known in the field of speech signal processing, and thus for each parameter a reference is provided where a deeper description and further information can be found. All features are calculated using a software tool developed in our research group (López de Ipiña et al., 2017;Mekyska et al., 2015;MATLAB, 2017;Praat, 2017).

4.1. Automatic disfluency segmentation

The recording has been automatically segmented in speech signal and disfluencies by means of a VAD (Voice activity detection) algorithm (Solé-Casals & Zaiats, 2010).

4.2. Feature extraction

In this approach the extracted features are the same for both genders moreover the approach is independent of the language then it is appropriate for multilingual environment and it is also flexible with regard to the task. The final selected feature test will be different with regard to the used task. This is the feature set used in the evaluation of speech quality under the above premises:

- *Classical features (CF)*:

1. Spectral domain features: Harmonic contents, harmonic to noise ratio (HNR), noise to harmonic ratio (NHR), pitch, jitter, shimmer, APQ (Amplitude Perturbation Quotient), spectrum centroids and formants and its variants (mean, median, min, max, mode, standard deviation).
2. Time domain features: Voiced/unvoiced segments, breaks, ZCR (Zero-Crossing Rate) and its variants.
3. Energy, intensity, short time energy, and spectrum centroids.

- *Perceptual features (PF)*:

1. Mel-Frequency Cepstral Coefficients (MFCC): As the *human ear* behaves as a filter concentrating only on certain frequency components, the logarithm of the power spectrum of the signal is processed by a bank of filters simulating this behavior. The cosine transform encodes the frequency contents in a vector which constitutes a compact representation of the signal. These filters are non-uniformly spaced on the frequency axis. There are more filters at low frequencies, and fewer filters at high frequencies. Mel-Frequency analysis and specifically Mel-Frequency Cepstrum Coefficients (MFCC) are oriented to simulate auditory perception.
2. Coefficients that provide information related to voice quality, perception, adaptation or amplitude modulation (Mekyska et al., 2015): MSC (Modulation Spectra coefficients) provide complementary information to MFCC, PLP (Perceptual Linear Predictive coefficients) take into account an adjustment to the equal loudness curve and intensity-loudness power law, LPCC (Linear Predictive Cepstral coefficients), LPCT (Linear Predictive Cosine Transform coefficients), ACW (Adaptive Component Weighted coefficients) their advantage is that these coefficients are less sensitive to channel distortion and ICC (Inferior Colliculus Coefficients) analyze amplitude modulations in voice using a biologically-inspired model of the inferior colliculus. These features are sometimes extended by

their 1st and 2nd order time derivatives (Δ and $\Delta\Delta$ respectively).

- *Non-linear features (NLF):*

1. Fractal Dimension, Shannon Entropy and Multiscale Permutation Entropy have been calculated (López de Ipiña et al., 2017).

4.3. Automatic selection of features by Mann-Whitney U-test

In this step the best features will automatically be selected regarding to common significance level. Thus, automatic feature selection is performed by Mann-Whitney U-test with a p-value < 0.1 in order to get a larger set for the second feature selection phase (MATLAB, 2017).

4.4. Automatic selection of features by WEKA

Then a new selection phase is carried in WEKA, which is a collection of machine learning algorithms for data mining tasks (WEKA, 2017). The functional used is the *SVMAttributeEval* algorithm. This provides a selection by analyzing the feature set.

4.5. Feature normalization by WEKA

During the preprocessing of data, all the features will be normalized by WEKA, which is standard software.

4.6. Automatic classification

Four classifiers have been used: k-nearest neighbors (k-NN), Support Vector Machines (SVM's), Multilayer Perceptrons (MLP's) with L layers and N neurons, and a Convolutional Neural Network (CNN) with L layers of N neurons, a convolution of cxc and a pool of pxp (Eibe et al., 2016).

The WEKA software suite (WEKA, 2016) has been used to carry out the experiments. The results were evaluated using the merit figure known as Accuracy (%). For training and validation steps we used k-fold cross-validation with $k=10$ (Picard & Cook, 1984). The results of the validation process are summarized in Fig. 2.

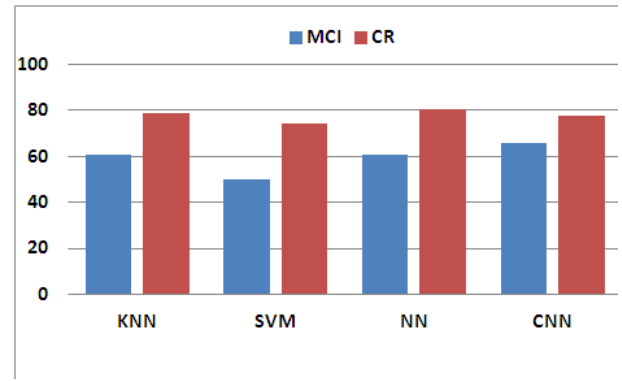


Figure 2. Accuracy (%) for classes and selected classifiers: k-nearest neighbors (k-NN), Support Vector Machines (SVM), Multilayer Perceptron (NN) with L layers and N neurons and a Convolutional Neural Network (CNN)

5. RESULTS AND DISCUSSION

5.1. Animal Names

Around 40 speech samples were used in the experimentation for the MCI group and 60 for the control group (CR), PGA-OREKA (see Table 1). Initially, the number of features included in the study is about 920 (473 for speech and 447 for disfluencies), obtained at a sampling frequency of 22.050 kHz. After a normalization step, an automatic feature selection is performed based on a non-parametric Mann-Whitney U-test with a p-value < 0.1 , after which the feature set was reduced to 150 features. In the second optimization step, the attribute selection algorithm *SVMAttributeEval* of WEKA yields about 80 features that are finally used in the experiments. The proposed feature set includes features from all the feature types described in subsection 4.2 for speech and disfluencies. Figure 2 shows the results of the automatic separation in two classes, CR and MCI. The merit figure Accuracy (%) is evaluated for all the classifiers detailed in subsection 4.6. The integration of disfluency analysis outperforms previous results (López de Ipiña et al., 2015) for most of the classifiers. The results are acceptable, robust, and balanced for all the classifiers about (average of %75). The deep learning option with CNN yields the best results for a configuration of 2 layers of 20 neurons, a convolution of 3×3 and a pool of

2x2. This option outperforms MLP for 2 layers of 100 neurons.

5.2. Comparing classifiers performance on AN, PD and SS

The evaluation of the models has been carried out with the balanced subsets: PGA-OREKA, MINI-PGA and AZTIAHORE.

In the first stage all the 473 features described in Section 3 are extracted: (a) 70 classic, (b) 60 perceptual, (c) 180 advanced perceptual and (d) 30 non-linear. Then, an automatic feature selection is carried out and the number of features is reduced to around 60% for animal naming, 50% for picture description, and 40% for spontaneous speech. Finally, an automatic classification by the classifiers is carried out by means of a cross-validation (SS, AN) and LOO (PD) strategy as described in Section 3. The attained global classification accuracy rates for the three tasks (%Accuracy.) are shown in Table 2.

Table 2: Recognition rate accuracy for the three tasks with the best classifiers

Task	%Acc
SS	95
AN	80
PD	94

Summarizing the results in the table:

- The best global results are obtained for the Spontaneous Speech task (Accuracy=95%). This is mainly due to the relevant emotional level of the recording obtained in a relaxing atmosphere, the presence of subtle cognitive changes in the signal due to a more open language; and the inclusion of AD patients instead of MCI subjects.
- The animal naming task (AN) shows the lowest results (82%). This is mainly due to the restricted-task that could be an easier exercise for people with MCI and the inclusion of MCI subjects instead of AD patients.
- In the case of the picture description task (PD), despite the small size of the set, the results are significantly good 94%.
- The results for AN and PD tasks are promising and require validation on larger sample sets.

6. CONCLUSIONS

The automatic integration of the most relevant features by CNN provides useful information not available from statistical tests. This work presents a novel proposal based on a platform to collect and automatically analyze a speech and disfluencies in order to support MCI and AD diagnosis (ALZUMERIC). Regarding the platform the following conclusions may be drawn:

- The system is friendly and easy to use.
- It is a valuable platform to objectify the patient's state and evolution to the clinician.
- The novel specific features are very useful for a multilingual and multicultural environment, because they are independent from the language (linear and non-linear).
- The configurable options are very useful specially for health specialists.
- The results are easy to interpret and manage.
- In this approach the extracted features are the same for both genders.
- The approach is independent of the language then it is appropriate for multilingual environment.
- It is also flexible with regard to the task.
- The final selected feature test will be different with regard to the used task.

Regarding the analysis methods it may be said that a non-linear multi-feature modeling is presented based on the selection of the most relevant features by statistical tests (under medical criteria) and automatic attribute selection: Mann-Whitney U-test and Support Vector Machine Attribute (SVM) evaluation. The approach includes deep learning by means of Convolutional Neural Networks (CNN). The results are encouraging and open a new research perspective.

ACKNOWLEDGMENTS

This work has been supported by FEDER and MICINN, TEC2016-77791-C4-2-R and UPV/EHU-Basque Research Groups IT11156.

REFERENCES

Alzheimer's Disease International. (2015). World Alzheimer Report 2015. 2017-05-29, de Alzheimer's Disease International,

- <https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf>.
- CITA-Alzheimer Foundation, PGA project. (2017). 2017-05-29, <http://www.cita-alzheimer.org/investigacion/proyectos>, (accessed on 10 June 2016).
- Dingemans, M., Torreira, F., Enfield, N.J. (2013). "Is "Huh?" a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items". PLOS ONE 8 (11): e78273.
- Eibe, F., Mark, A., Hall, I., Witten, H. (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
- Faundez-Zanuy, M, et al. (2013). Biometric Applications Related to Human Beings: There Is Life beyond Security. Cognitive Computation, 5(1), 136-151.
- Klimova, B., Maresova, P., Valis, M., Hort J., Kuca, K. (2015). Alzheimer's disease and language impairments: social intervention and medical treatment. Clinical Interventions in Aging, Clinical Interventions in Aging 2015:10, 1401–1408.
- Laske, C., Sohrabi, H.R., Frost, S.M., López-de-Ipiña, K., Garrard, P., Buscema, M., Dauwels, J., Soekadar, S.R., Mueller, S., Linnemann, C., Bridenbaugh, S.A., Kanagasingham, Y., Martins, R.N, O'Bryant, S.E. (2015). Innovative diagnostic tools for early detection of Alzheimer's disease. Alzheimer & Dementia, 11(5), 561-78
- Lopez-de-Ipiña K., Alonso, J.B., Manuel Travieso, C., Sole-Casals, J., Egiraun, H., Faundez-Zanuy, M., et al. (2013). On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. Sensors, 2013(13), 6730–6745.
- Lopez-de-Ipiña, K., Martinez-de-Lizarduy, U., Barroso, N., Ecay-Torres, M., Martinez-Lage, P., Torres, F., Faundez-Zanuy, M. Automatic analysis of Categorical Verbal Fluency for Mild Cognitive impairment detection: A non-linear language independent approach. (2015). 4th International Work Conference on Bioinspired Intelligence (IWOBI), 2015(1), 101 – 104.
- López-de-Ipiña K et al. (2016). Advances in a Multimodal Approach for Dysphagia Analysis Based on Automatic Voice Analysis. Advances in Neural Networks 54 of the series Smart Innovation, Systems and Technologies, 1, 201-211.
- López de Ipiña, K., Martínez de Lizarduy, U., Calvo, P., Beita, B., García-Melero, J., Ecay-Torres, M., Estanga, A. and Faúndez-Zanuy, M. (2017). Analysis of disfluencies for automatic detection of Mild Cognitive Impairment: a deep learning approach. 4th International Work Conference on Bioinspired Intelligence (IWOBI), 2017(1), 1-4.
- Lezak, M.D., Howieson, D.B., Bigler, E.D., Tranel, D. Neuropsychological Assessment (fifth edition). Oxford University Press (2012).
- Ruff, R.M., Light, R.H., Parker, S.B., Levin, H.S. The psychological construct of word fluency. Brain Lang. 57: 394-405 (1997).
- Mekyska J et al. (2015). Robust and complex approach of pathological speech signal analysis. Neurocomputing, 167(2015): 94–111.
- Gómez-Vilda P. et al. (2015). Phonation biomechanical analysis of Alzheimer's Disease cases. Neurocomputing, 167(1) , 83–93
- MATLAB. 2017. 2017-05-29. MATLAB, www.mathworks.com
- Praat. (2017). Praat: doing Phonetics by Computer. 2017-05-29. www.praat.org/
- Solé-Casals, J., Zaiats, V. (2010). A non-linear VAD for noisy environments. Cognitive Computation, 2(3), 191-198.
- WEKA. 2017-05-29. WEKA <http://www.cs.waikato.ac.nz/ml/weka>
- Picard, R., Cook, D. (1984). Cross-Validation of Regression Models. Journal of the American Statistical Association, 79(387), 575–583.

D3 ERANSKINA

Feature extraction approach based on Fractal Dimension for Spontaneous Speech modelling oriented to Alzheimer Disease diagnosis. NOLISP, 2013, Mons Belgica (C1 eranskina)

