IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 27, 2021, accepted May 5, 2021, date of publication May 11, 2021, date of current version May 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3079375

Online Student Authentication and Proctoring System Based on Multimodal Biometrics Technology

MIKEL LABAYEN[1,3], RICARDO VEA[1], JULIÁN FLÓREZ[2], (Member, IEEE), NAIARA AGINAKO[3], AND BASILIO SIERRA[3]

[1] Smowltech, 20009 Donostia, Spain
[2] Vicomtech Research Center, 20009 Donostia, Spain
[3] Computer Sciences and Artificial Intelligence Department, University of the Basque Country, 20018 Donostia, Spain

Corresponding author: Mikel Labayen (mikel.labayen@ehu.eus)

This work was supported by the Spanish Ministry of Sciences, Research and Universities (Ministerio de Ciencia, Innovación y Universidades (MCIU)/Agencia Estatal de Investigación (AEI)/Fondo Europeo de Desarrollo Regional (FEDER), Unión Europea (UE)) under Grant RTC-2016-5711-7.

ABSTRACT Identity verification and proctoring of online students are one of the key challenges to online learning today. Especially for online certification and accreditation, the training organizations need to verify that the online students who completed the learning process and received the academic credits are those who registered for the courses. Furthermore, they need to ensure that these students complete all the activities of online training without cheating or inappropriate behaviours. The COVID-19 pandemic has accelerated (abruptly in certain cases) the migration and implementation of online education strategies and consequently the need for safe mechanisms to authenticate and proctor online students. Nowadays, there are several technologies with different grades of automation. In this paper, we deeply describe a specific solution based on the authentication of different biometric technologies and an automatic proctoring system (system workflow as well as AI algorithms), which incorporates features to solve the main concerns in the market: highly scalable, automatic, affordable, with few hardware and software requirements for the user, reliable and passive for the student. Finally, the technological performance test of the large scale system, the usability-privacy perception survey of the user and their results are discussed in this work.

INDEX TERMS Biometric authentication, cloud computing, computer vision, data science applications in education, distance education and online learning, machine learning, security, computer vision.

I. INTRODUCTION

There is no doubt that online learning has been gaining popularity throughout the past years. This phenomenon is not surprising given that online learning allows education institutes to operate at a lower cost and with greater reach-out to more students. Educational institutions are offering courses online to leverage the benefits of online learning. This is especially so since the advent of Massive Open Online Courses (MOOC). On the other hand, COVID-19 has been a challenge for traditional institutes offering face-to-face teaching, and these institutions have had to migrate (in a very short period of time) to a fully online education model forced by the pandemic situation. However, online learning implementation presents challenges.

E-learning has a serious deficiency, which is the lack of efficient mechanisms that assure user authentication, in the system login as well as throughout the session. Especially for online certification and accreditation, the training organizations need to verify that the online learners who completed the learning process and received the academic credits are precisely those who registered for the courses. Inadequate methods of identity verification affect the reliability of credentials and certification earned online.

Without certainty of the authenticity of the online learner's identity, the aspiration towards fully online education is stymied and the evaluation of the knowledge and skills obtained by the online learner is unreliable. In order to prevent compromising the credibility of online accreditation,

The associate editor coordinating the review of this manuscript and approving it for publication was Tony Thomas.

validation must be carried out in a constant or continuous manner. At the same time, validation should be non-invasive and non-disruptive, and does not distract the learning process.

Online proctoring, generally refers to proctors (humans) monitoring an exam over the internet through a webcam. It includes as well the processes, occurring at a distance, for authenticating the examinee as the person who should be taking the exam. Online proctoring was first introduced by Kryterion [1], [2] in 2006, marketing it as a technological solution in 2008. Since then, several other organizations have followed Kryterion's lead creating more capable technology-based alternatives, which are gaining attention, such as online proctoring.

Nowadays, there are commercial solutions in the market as well as research publications that try to solve this problem. Some of them only authenticate the identity, others monitor, some in real time, others record the sessions. Some cover only exams or specific activities. Some are totally human based solutions (non-scalable) or fully automatic ones (non-reliable). There are also a few scientific approaches which develop the idea of combining some of the cited functionalities. However, there is no comprehensive and reliable solution which combines multi-biometric continuous authentication with continuous visual and audio monitoring, with device activity monitoring and lock-down options and human supervision (only when required) to guarantee 100% reliable results.

In this work we present a new system which gives commercial solutions to all that was needed. It is based on web applications which offer a continuous authentication identity service of online students through a constant biometric (face, voice, typing) recognition system (biometric traits cannot be lost, stolen, or recreated), as well as automatic continuous proctoring through automatic image and audio processing (device monitoring & lock-down and inappropriate behaviour detection) allowing online courses to gain value of what benefits both institutions and students. This solution is based on a high accuracy biometrics recognition and digital signal processing algorithms and it is complemented with human supervision for those situations in which the automatic algorithms are not able to determine reliable results. It can be used to continuously authenticate the learners, either throughout the entire learning process, or only at certain sensitive stages of e-learning. It is contactless and needs only a low level of user collaboration. In addition, the whole system is based on cloud computing technologies, which removes geographical and technological barriers for online learning providers.

The article is organized as follows. Section II gives an overview of some relevant related works and highlights the main differences with our approach. Section III describes the whole system overview and workflow. Section IV contains a scientific-technical description of core modules. Section V presents system tests to measure the algorithms' performance as well as a survey made for user experience evaluation. Section VI presents the results of the tests. Finally, section VII draws the conclusions and presents future works.

II. RELATED WORK

The ability to authenticate and monitor online users is becoming more important due to the increase of the internet world (e-learning, e-banking, e-gambling, e-government). Since first human based online proctoring systems, various fully or semi-automatic authentication and proctoring technologies based on biometric features have appeared in the last few years. Biometrics has proved itself to be one of the best methods for recognizing people based upon physiological or behavioural characteristics [3]. These technologies can be divided into two categories: those that are based on physical characteristics and those that are based on behaviour characteristics. The former includes face recognition, fingerprint scanners, iris scanners, vein matching, etc. The latter includes voice recognition, handwriting recognition, keystroke dynamics, etc. It is proved that no technology will provide the right answer on its own, but that the combination of different solutions will come up with the appropriate functionality depending on customer needs. In addition, most remote authentication proctoring technologies involve some level of human intervention for fully reliable service, thereby putting limitations on scale.

These biometric technologies have been widely used for various purposes, and they have become more and more common in our daily lives. However, very few of them have been successfully adopted for online learning validation.

A. COMMERCIAL SOLUTIONS

Some initial approaches have been brought to market as commercial solutions. The following is an overview of these services:

1) **Fully Live Online Proctoring:** Students are on video and watched remotely by a live proctor. Live proctoring is a live online service for students taking exams online. After making an appointment, the students are taken to the online proctoring room where they will connect with a live proctor from one of the two online proctoring centres via their web cameras. The students connect their screen to the proctor. This allows the proctor to see their computer screen. The proctor asks them to show a photo ID and to answer a few questions about themselves in order to verify they are in fact the right student. During the exam, the proctor looks at the student directly through a webcam. It is a secure and complete solution for exam proctoring, but since it is a non-automatic solution, it cannot deal with continuous identification during all learning process. Furthermore, it needs a high speed internet channel to transmit video data, probably unaffordable for different parts of the world and it is not passive for students. Some commercial solutions in the market are ProctorU [4], Examity [5] and Software Secure - PSI [6].
2) **Recorded and Reviewed Proctoring:** Sessions are recorded as the computer monitors students. A human can then review the video at any time afterward.

In these systems, students use their own computer and a webcam to record assessment sessions, the student and the surrounding environment are recorded during the entire exam. Instructors can quickly review details of the assessment, and even watch the recorded video. Recorded proctoring has the same limitations as live proctoring. In addition, it is a passive system. However, nobody analyzes the videos, so teachers must watch all of them in order to detect undesirable behaviours and maintain the live proctoring advantages. Some commercial solutions in the market are Kryterion [1], ProctorExam [7], Respondus [8], Remote Proctor [9], ProctorCam [10], B virtual [11] and Learner verified [12].

3) **Fully Automated Solutions:** The computer monitors students, it authenticates them and determines whether they are cheating. These are automatic and passive solutions. They just cover the beginnings of exams and work submission processes. However, users must be totally active in this kind of system (they must type a predefined paragraph and take an ID photo themselves). In addition, this kind of system does not cover all the learning process continuously. Some commercial solutions in the market are Proctorio [13], ProctorTrack [14], Comprobo [15], Sumadi [16], ProctorFree [17], HonorLock [18] and ExamSoft [19].
   a) **Authentication technologies:** Recognition technologies are used to authenticate a student based on a prior examination of some physical feature. They are typically built upon a before/during/after analysis to verify that the same student who initially registered for the course was actually the same student who took the exam. Commonly-known recognition technologies include facial, fingerprint, or voice recognition. In the last year, new biometric procedures such as keystroke dynamics (it recognizes typing patterns based on rhythm, pressure, and style) are gaining popularity. It is likely that recognition technologies will be most effective when used with some combination of other technologies available.
   b) **Monitoring technologies:**
      i) Webcams and microphones are one of the original technologies used to replace a live proctor and are present in most remote exam proctoring solutions on the market. They can record individual students when the camera is part of the computer, or groups when the camera is placed in a classroom. They can monitor the behaviour of the students, whether they are cheating, receiving help from other students, using mobile devices, books…Webcam/Microphone technologies often require significant storage capabilities so that video records can be reviewed if necessary.
      ii) Computer lockdowns are able to monitor the activity carried out by the student within their computer preventing them from ‘‘surfing the internet’’ while taking a test. This monitoring will be done only and exclusively when the student is doing an activity that can be evaluated.

None of the cited commercial solutions provides a multi-biometric authentication solution or continuous authentication/proctoring service (based on automatic analysis) through the whole learning course (not only exams). In addition, this work presents a completely new commercial approach to overcome barriers such as low-speed internet connection (using data samples, not continuous heavy video signals) or costly extra HW/SW requirements (using non-installable and fully integrated in LMS web applications).

B. SCIENTIFIC AND ACADEMIC APPROACHES

1) TECHNICAL WORKS

Nowadays, although there are still some non-biometric based authentication approaches [20], the latest attempts for online student authentication automation tends to use biometric technologies; facial [21]–[26], fingerprints [27] or typing [28], [29]. On the other hand, some approaches try some combination of them, such as face and voice [30] or face, voice and typing [31], [32]. All the approaches are focused mainly on student authentication without providing proctoring service.

It is through facial authentication complemented with other biometrics such as voice or typing recognition, that an opportunity appears in e-learning to verify the absence of frauds while the students do their activities on the platform.

The main novel contribution of the work we present in this article includes a completely new combination workflow of three main biometrics providing a continuous and non-intrusive authentication service. It also adds new automatic and continuous proctoring features based on image and audio signal processing to the system. Furthermore, it integrates computer activity monitoring and lock-down possibility and, finally, it even complements the service with automatic alarms which trigger minimal human supervision, guaranteeing the reliability of results.

Finally, the recent concern for safety and privacy has also provided recent research on this topic related to online proctoring [33].

2) USER EXPERIENCE RELATED WORKS

On the other hand, very few works completed the research about teachers and student user experience with this kind of authentication and proctoring approaches. One of them completed the research about the implementation of facial verification into education with a successful positive result [34]. The objective was to guarantee students authentication and to know exactly the amount of time that they spend in front of the computer reading or realizing their virtual activities.

TABLE 1. Commercial solutions vs SMOWL (solution described in this article). Service characteristics: 1-Authentication during whole exam or session; 2-Multi biometric authentication (at least 2 different); 3-Exam monitoring; 4-Continuous (full course) monitoring; 5-Dishonest behaviour detection; 6-Totally Passive and non-intrusive system; 7-Automatically analyzed results; 8-100% guaranteed and reliable results; 9-Personalised alarms; 10-Human real-time proctor; 11-Device monitoring. Technical features: 12-Scalable system; 13-Flexible access to students - no scheduled; 14-No extra SW/HW installation required for authentication and proctoring; 15-Works with low-speed connection; 16-Fully integrated in institution LMS; 17-Multi-Browser & device. Legal aspects: 18-EU-hosted solution; 19-GDPR compliance. ✓- Yes | X- No.

	Service characteristics											Technical features						Legal asp.	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Fully Live Online Proctoring																			
ProctorU	✓	X	✓	X	✓	X	✓	✓	X	✓	X	X	X	X	X	✓	✓	X	X
Examity	✓	X	✓	X	✓	X	X	✓	X	✓	X	X	X	X	X	X	X	X	X
PSI	✓	X	✓	X	✓	X	X	✓	X	✓	X	X	X	X	X	X	X	X	X
Recorded and Reviewed Proctoring																			
Proctoexam	✓	X	✓	X	✓	X	X	✓	X	✓	X	X	✓	X	X	X	X	✓	✓
Kryterion	✓	X	✓	X	✓	X	X	X	X	✓	✓	X	✓	X	X	X	X	X	X
Remote Proctor	✓	X	✓	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Proctorcam	✓	X	✓	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X
B Virtual	✓	X	✓	X	✓	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Fully Automated Solutions																			
Proctorio	✓	X	✓	X	✓	X	✓	✓	X	X	X	✓	✓	X	X	X	X	X	X
Proctortrack	✓	X	✓	X	X	X	✓	X	X	✓	✓	✓	✓	X	X	X	X	X	X
Respondus	✓	X	✓	X	✓	X	X	X	X	X	X	X	✓	X	X	✓	✓	X	X
Comprobo	✓	X	✓	X	✓	X	✓	X	X	X	X	✓	✓	X	X	✓	X	X	X
Sumadi	✓	X	✓	X	X	X	✓	X	X	X	✓	✓	✓	X	X	X	X	X	X
Proctorfree	✓	X	✓	X	X	X	✓	X	X	X	X	✓	✓	X	X	✓	✓	X	X
HonorLock	✓	X	✓	X	✓	X	✓	X	X	X	✓	✓	✓	X	X	X	X	X	X
ExamSoft	✓	X	✓	X✓	X	✓	X	X	X	✓	✓	✓	X	X	X	X	X	X	X
SMOWL	✓	✓	✓	✓	✓	✓	✓	✓	✓	X	✓	✓	✓	✓	✓	✓	✓	✓	✓

In the same way, a facial authentication mechanism was also presented. This insured that the students are not impersonated to improve their marks in virtual tests [35].

III. SYSTEM OVERVIEW

The system we present in this work aims to provide a practical cyber-security solution for both a) continuous online user identification (using biometric technology) and b) monitoring using automatic signal processing and a computer monitoring system. The authentication process is based on automatic authentication of facial images (captured by webcams), audio clips (captured by the microphone) and keystroke dynamics (captured by the keyboard), checking that it is the person that it really should be during the entire online interaction. The monitoring process is supported by webcams and microphones too, checking continuously that the student is not making any inappropriate behaviour (using forbidden devices and applications, receiving help...). It also locks down the computers (with a previous installation in the learner computer and consent) during exams or training sessions preventing the user from visiting web pages or other documents while performing the course.

The system can be used for any online user authentication but it is specialized in the institutions that offer online courses providing training and degree certification, including verified MOOCs and corporate training for employees. This system can help e-learning providers in their objective to be awarded credit by Quality Educational Agencies for their courses by seeking traceability of evidence of student authenticity and their behaviour. It can be used to track the continuous authentication of the student in all or in sensitive stages of

TABLE 2. State-of-the-art solutions vs SMOWL (solution described in this article). Authentication method: 1-Face recognition; 2-Voice recognition; 3-Typing recognition; 4-Continuous authentication during whole session (not only at the beginning). Proctoring-Monitoring method: 5-Image processing; 6-Audio processing; 7-Screenshots capture; 8-Device information capture (active window, open processes, peripherals devices, copy/paste commands...). Proctoring-Device Lock-Down: 9-Device lock-down. Guarantee: 10-Human supervision to clarify doubts providing 100% guaranteed and reliable results. ✓-Yes | X- No.

	Auth.				Monit. & Proctor.					%
	1	2	3	4	5	6	7	8	9	10
[21] to [26]	✓	X	X	X	X	X	X	X	X	X
[28]	X	X	✓	X	X	X	X	X	X	X
[30]	✓	✓	X	X	X	X	X	X	X	X
[31] [32]	✓	✓	✓	X	X	X	X	X	X	X
SMOWL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

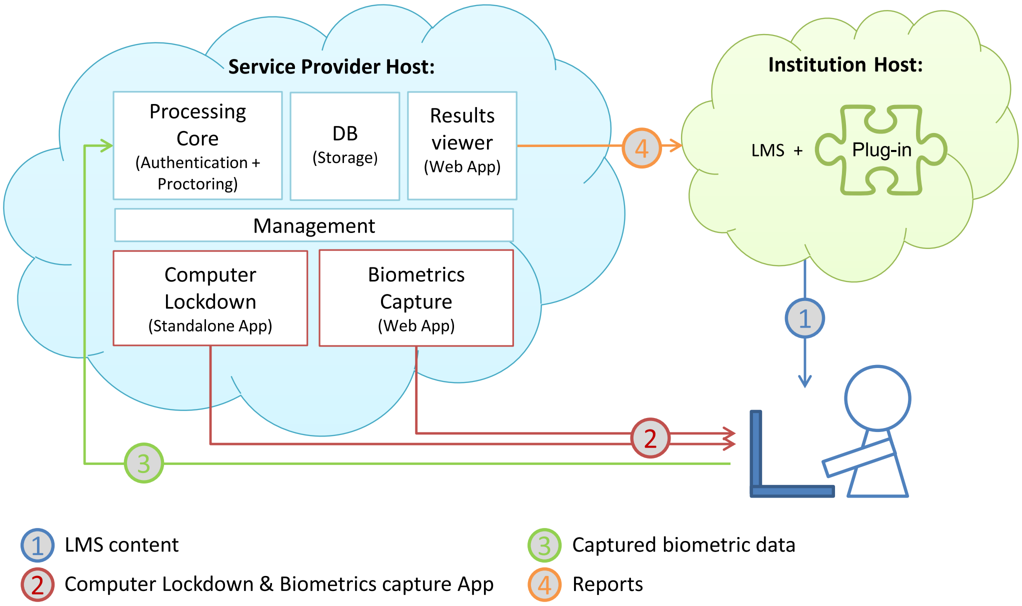

Service Provider Host:
Processing Core (Authentication + Proctoring)
DB (Storage)
Results viewer (Web App)
Management
Computer Lockdown (Standalone App)
Biometrics Capture (Web App)
Institution Host:
LMS + Plug-in
1
2
3
4
1 LMS content
2 Computer Lockdown & Biometrics capture App
3 Captured biometric data
4 Reports

FIGURE 1. Authentication and proctoring system set-up.

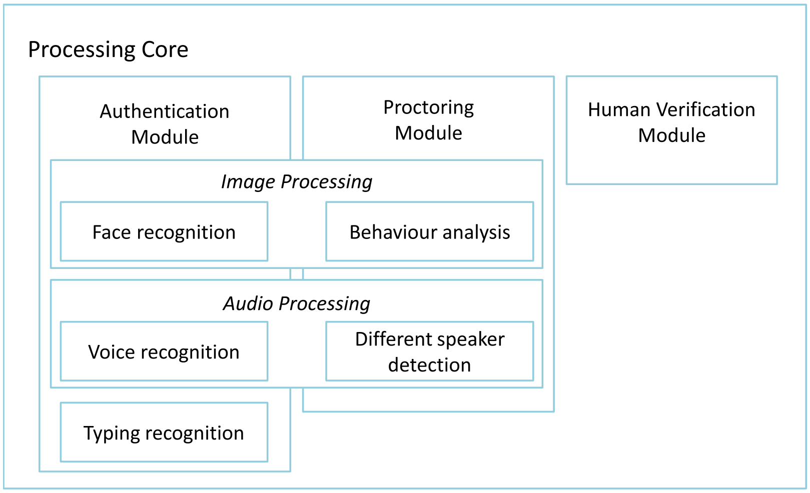


FIGURE 2. Processing core description.

e-learning. Figure 1 shows general set-up of the system and Figure 2 details the processing core description.

The complete system workflow is embedded in cloud computing applications, and can be used anywhere, removing geographical and technological barriers. The general scheme of operation is as follows and is given in more detail in Figure 3:

1) The system is integrated into the virtual campus of the training centre (available for different LMS platforms).
2) The training centre sends a code (unique student identifier) with an image of the student to register in the system. According to system data privacy policy, the system works with images, audio clips…not identities, so it lacks connection with the student personal data such as name, age or address [36].
3) The first time the student enters the virtual campus the system takes biometric samples (picture, short speech, predefined paragraph typing) which will help us create the tracking biometrical model.
4) Thereafter, whenever the student is connected to work, biometric samples will be taken randomly and continuously. This data is sent to servers in the cloud. The online management module stores and analyzes the data which is compared with the biometrical model that has been created previously for authentication purposes and analyzed to detect inappropriate behaviours. All storage, analysis and results report and alarm creation tasks are executed in online servers, making the integration, support and maintenance tasks for institutions easier and more transparent. During this period, the computer lockdown module can be activated for monitoring purposes.
5) The result leads to an individual user report that is updated constantly and to which the training centre has access.

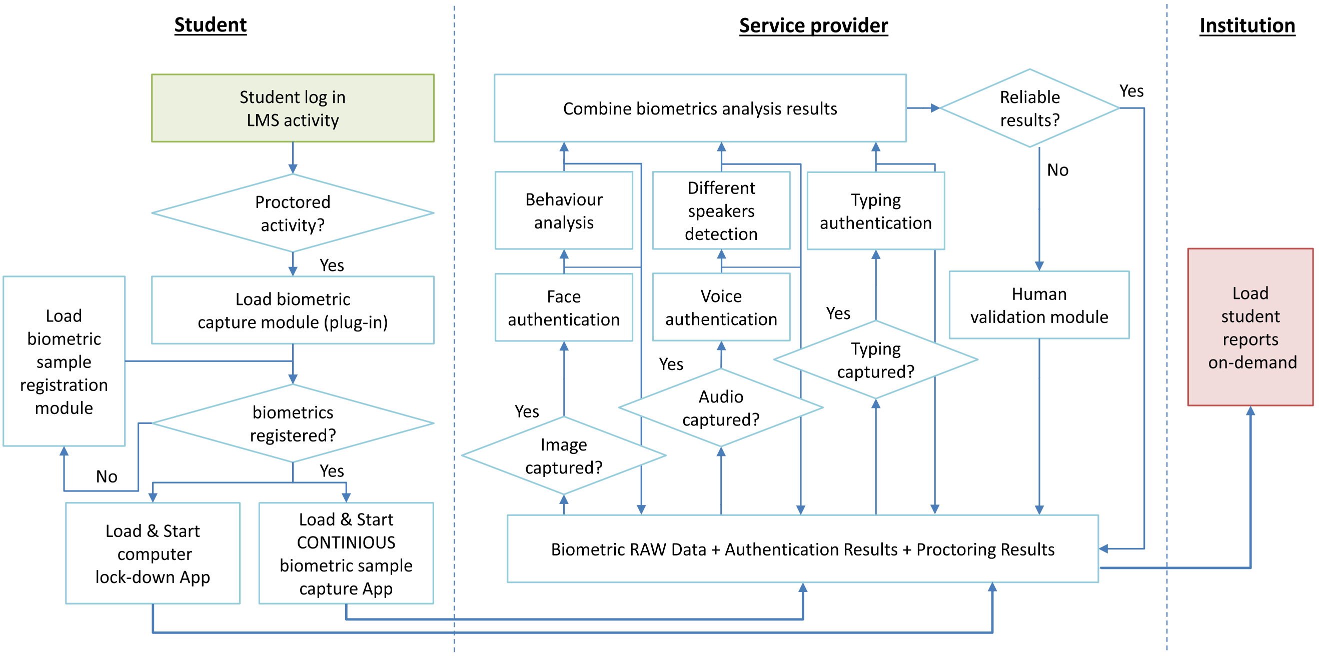

Student
Service provider
Institution
Student log in LMS activity
Proctored activity?
Yes
Load biometric capture module (plug-in)
Load biometric sample registration module
biometrics registered?
No
Yes
Load & Start computer lock-down App
Load & Start CONTINIOUS biometric sample capture App
Combine biometrics analysis results
Reliable results?
Yes
No
Behaviour analysis
Different speakers detection
Typing authentication
Face authentication
Voice authentication
Human validation module
Yes
Typing captured?
Yes
Audio captured?
Yes
Image captured?
Biometric RAW Data + Authentication Results + Proctoring Results
Load student reports on-demand

FIGURE 3. System workflow.

The key characteristics of the system are:

1) **Continuous and not scheduled system.** Proctoring and authentication processes are carried out throughout the entire session, not only when users log in. Furthermore, in the e-learning case, it can follow every session of the course, not only the assessments. It is very flexible. Service is given 24/7, anywhere. Previous schedule is not required.
2) **Passive & non-intrusive system.** The system offers a passive system for students when taking photos, audio clips or keystroke pattern. It does not need the collaboration of the student and it is contactless. For this reason, in the case of images, it properly works when the pose/appearance/complements/expressions of the students or the light conditions of the room are not controlled (in the wild), getting low-contrast images with partial occlusions due to wrong position or the appearance/compliments/expressions variations of the student. Regarding audio clips, the microphone only records when it detects some noise, nothing if the student is in silence. The clips are later analyzed and if voice is detected in the recording it is compared with the data gathered during registration of the student, to validate their identity, or to detect cheating when there are different voices in the recording.
3) **Automatic and scalable:** All capture, verification, data management and monitoring report modules are carried out with cloud computing technology as services in the cloud. Photos and patterns are taken automatically and randomly and compared with the biometric model made during registration. This scalable automatic set-up makes it possible to bring this solution to overcrowded scenarios such as MOOCs.
4) **Few requirements for the end user.** Cloud-based (SaaS) automatic solution. Needed Hardware - Software (HW/SW): basic webcam, microphone, keyboard and any updated browser. Final users do not have to install anything. This system works over any device, platform, OS and browsers with no installation needed.
5) **Automatic analyzed results.** 100% guaranteed results with custom alarms. If automatic validation cannot be confirmed (if the pictures or audio clips do not compile with the quality needed to allow the system to automatically validate the student), a manual checking by staff will be set to certify the results 100%.
6) **Fully integrated in customer LMS.** It can be integrated in any Learning management system (LMS) using a general API but it has a specific plugin for Moodle, Moodlerooms, Blackboard, OpenedX, Canvas, etc. (most used LMS).
7) **Secure.** Data is transmitted under secure internet protocol and stored in safe cloud servers.
8) **Private.** The user's identity remains protected because we only handle data that are not linked to identities but to user codes provided by the online entity.

A. DATA CAPTURE AND STORAGE MODULE

This module captures data from the student webcam, microphone and keyboard. The core of this application has been developed using the latest HTML5 standard implementation in web browsers. The application is downloaded into the student's terminal and executed without any installation needed. Whenever the user is connected to the course, quiz or specific exercise into LMS, pictures, audio clips and keystroke dynamics samples will be taken randomly and continuously with predefined mean periodicity. This data is sent to servers in the cloud, through a SSL encrypted channel, with the user identification code. The system online management module stores and analyzes the images.

B. AUTHENTICATION MODULE

Once all data is stored in cloud servers, it is compared with the biometrical model, linked to student's identification code, which has been created at registration time and has been updated with recent positive data. The result is stored in the system database. The system recognition and training algorithms are developed using the latest algorithms in artificial intelligence (explained in Section IV) which are improving constantly their recognition precision and robustness facing light, position and student appearance (physical changes and complements such as hat, glasses…) change problems, noise in audio clips and variability in typing samples. The authentication result is a combination of each biometric authentication module result (face, voice and typing).

C. PROCTORING AND COMPUTER LOCK-DOWN MODULES

During monitoring sessions, the captured image and audio clips (which have been used for authentication purposes) are processed with different techniques in order to detect inappropriate behaviour of students during e-learning activities. For this reason, the system is able to detect if the student is receiving help (by phone, help from presential friend…) or is checking forbidden documentation (books, other devices connected to the internet…). All these actions can be strictly forbidden in some face-to-face learning activities according to the institution code of honour.

In addition, attempts to cheat are detected and reported if any student tries to trick the system, such as mounting a photograph in front of the camera or replacing the image of the ID card with someone else's. Attempts to insert another image or video signal into the camera are also detected.

On the other hand, the system contains a computer lockdown module. During all the online session, a computer lockdown module (Section IV) will monitor the computer of the student detecting connected peripherals, active windows, computer information (HW/SW), executing programs or processes, browsing history/webs and copy-paste commands. All the information captured in each session is stored in the database.

D. HUMAN VERIFICATION MODULE

As part of the quality warranty, a random data and results auditory must be set. This task will test try the quality assurance mechanism definition and implementation with a huge number of students connected at the same time. It will be based on a random data cross-verification (same images, voice and keystroke patterns validated by different persons) of images, voice and keystroke samples captured during the session with registered data. Besides, when the quality of the photos or audio does not reach the threshold needed, a human verification is made by trained staff delivering a 100% reliable verification of the student.

E. REPRESENTATION MODULE OF THE RESULTS

Final results are presented by the data representation module. It creates graphic charts and tables on demand, 24h/365d, as a dynamic web page. The final reports can be downloaded or printed in different formats. In addition, the data representation module also generates automated alarms when some predefined prohibited behaviour happens.

IV. AUTHENTICATION AND PROCTORING MODULES IMPLEMENTATION

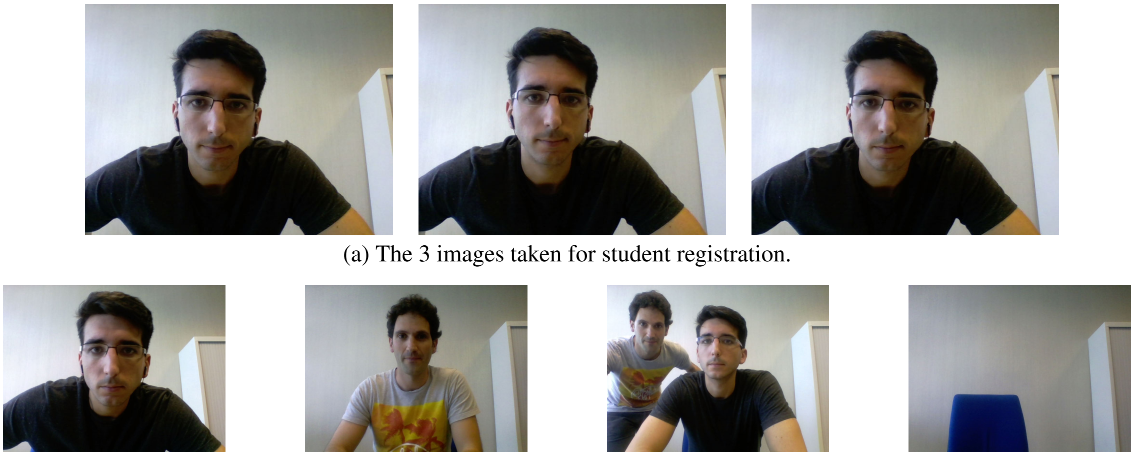
As explained in the previous sections, the system presented in this work contains artificial intelligence-based modules for user authentication as well as computer lockdown technologies for device monitoring. In this section the scientific algorithm behind authentication modules and technology and functionalities of the computer monitoring are explained and referenced in depth.

A. FACE DETECTION AND RECOGNITION

This system includes a facial detection and recognition module through a biometric model created using registration time face pictures. The module output results are clustered in five groups determining: a) If there is someone in front of the webcam or not, b) How many people (if any) are in front of the webcam, c) If one of these people is the person who should be in front of the screen, d) when only one person is in the image, whether this person is the person it should be, e) If the person who it should be is not involved in any inappropriate behaviour (book or electronic device use). Some examples are shown in Figure 4.

There are different approaches for face detection in the literature [37]. However, few of them are robust enough when dealing with variation in pose and lighting of captured images (remember that pictures are taken without student attention and randomly). The facial detection procedure presented in this work is based on the FaceBoxes methodology [38]. This methodology is known for being the most common "Deep Learning" based technique whose optimal deployment is based on use of GPUs. This methodology obtains better results in the Face Detection Data Set and Benchmark (FDDB) benchmark (Jain and Learned-Miller, 2010) than other methodologies tested in the development process of this module.

The image processing and authentication processes takes [39] as the base reference method for the extraction and normalization of facial texture. This algorithm contains the following subtasks: (1) face detection, (2) face characteristic points detection in the facial region and (3) deformable parametric 3D facial model adjustment based on the detected points. However, the requirement of system passiveness makes it necessary to have continuous improvements in the detection and authentication algorithm to deal with high



(a) The 3 images taken for student registration.

(b) Correct student authentication.

(c) Another person.

(d) More than one person.

(e) Nobody in front of the device.

FIGURE 4. Authentication and proctoring system captured and analyzed image examples.

variability of input images. Starting in this reference work, a series of improvements have been added:

1) **Pose and expressions correction:** A new method, called as M3L (Multi- level, Multi-modal, Multi-task Learning) [40], is used to improve efficiency in face points and other facial attributes detection (gestures of the face and eyes). M3L addresses the problem of extracting all these facial and ocular data through a hierarchy of neural networks using existing correlations between the data. Furthermore, a new multi-level deformable 3D facial model adjustment distributes the deformation error in an equitable way, distinguishing three stages with different levels of priority in estimation of (from greater to minor): (1) pose, (2) interpersonal deformations (user-specific facial shape) and (3) intra-personal deformations (deformations due to facial expressions).
2) **Aspect normalization, feature selection and classification:** The extraction of biometric features through a deep neural network [41] has been improved training a database with 10M of images of 100K individuals with great variability of appearances and facial shapes, lighting, facial expressions, accessories and poses (Guo *et al.*, 2016).
3) **Normalization of the lighting:** The procedure of normalization of the lighting has been carried out with a hierarchical method in which the facial region as a whole as well as specific and normalized regions of the face are analyzed. This normalization is performed using the Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm [42], which equalizes the image locally, highlighting the contrasts, applied to each RGB color channel.
4) **Robustness against partial occlusions:** Occlusions detection is based on the MobileNet-SSD neural network [2], [43]. Combining this person detector and the face detector, the system increases its robustness in detection when (at least) partial face occlusion is occurring. This people detector (body) is more robust than the face detector in these cases. Therefore, if a person is detected, but not a face, it is more likely that this face is at least partially occluded. In this case, the face detection alarm is considered. Additionally, the methodology proposed in [44] has been implemented and adapted to the framework of the needs of the project to handle the occluded normalized facial images. The facial detection returns more partially occluded facial cuts than desirable ones. Normally these occlusions are given either by the user's hands in front of the face or because the camera is only pointing to the top-half of the face. This occlusion negatively influences the later stages of facial point detection and biometric vector extraction. This system includes a facial image synthesis from Generative Adversarial Network [45], which fills the occluded part with close facial features obtained from the trained model. In this way, the negative impact of occlusion can be reduced.

B. VOICE DETECTION AND RECOGNITION

This module implements a continuous voice detection and authentication algorithm. The developments are based on the Kaldi tool [46] and the implementation of the method of [47]. Both include tools for the development of the biometric model, the vector representation of each speaker's characteristics. The algorithm works on four tasks:

1) **Analysis, interpretation and normalization of audio by VoIP:** Since the data used in VoIP (technology in which this system is based on) use the G.711 codec with a 64 kbps bit rate, which implies a loss of important information in order to compress the audio signal, all training data from the available acoustic databases are transformed into this encoding and format. In this way, the training and evaluation audio matches were obtained in the different frequency

ranges. Signal pre-processing is integrated to discard that acoustic segments that do not contain speech (silence, music or noise). The final version of the VAD vocal activity detection module has been developed using GMM Gaussian mixture models and processing functions proposed in the Kaldi code tool. A total of 3 model training level were performed. The difference between each of them is based on the transformation of training data for greater robustness versus the high acoustic variability of the application scenario.

2) **Background and speaker modelling:** The speaker modelling is based on d-vectors or speaker embeddings using deep neural networks. This solution offers better performance in terms of robustness and accuracy. The implementation follows the solution presented by Google in 2018 [47]. In this approach, a recurrent neural network based on LSTM cells is generated. It receives an acoustic characteristic of a specific audio (Mel filter bank) as input and returns its d-vector. Once the training is finished, the neural network can be used to generate d-vectors from the acoustic characteristics of the speaker. Then, a centroid is generated, which is considered as the speaker's biometric footprint.
3) **Patterns comparison:** For a verification or identification process, given a vector of acoustic characteristics and its associated d-vector, they are compared with the centroids of each of the speakers in a new similarity matrix.
4) **Speaker segmentation on streaming audio:** This diarization system employs d-vectors or speaker embeddings and an agglutination model based on recurrent neural networks [38]. This approach aims to overcome the traditional agglutination approach problems, which work with the sentences individually and independently, it being difficult to benefit from the information provided by large amounts of labelled data. This system is based on the work presented by [48]. An independent text announcer recognition network is used to extract d-vectors or speaker embeddings from 240 millisecond windows and a 50% overlap. A vocal activity detector based on Gaussian models is used to eliminate speechless parts and split the signal into segments less than 400 milliseconds. These segments are converted to d-vectors and included in the RNN network based diarization system.

C. TYPING RECOGNITION

Keystroke dynamics are an effective behavioural biometric, which captures the habitual patterns or rhythms an individual exhibits while typing on a keyboard. According to neurophysiological analysis [49], these typing styles are idiosyncratic, in the same way as handwriting or signatures, due to their similar governing neuronal mechanisms. For this reason, they can be used to authenticate an individual.

The system presented in this work applies keystroke dynamics in dynamic text, that is, the analysis occurs for any text that is typed by the user and continuously. Keystroke dynamics in static text requires less effort to be implemented and it also reached lower error rates in the literature [50]. However, a dynamic text analysis [51] is necessary to keep final student passiveness in the authentication process without bothering them by asking them to type a predefined paragraph (usually not related to the e-learning activities in progress). This approach considers the fact that the keystroke dynamics of one person may vary in different psycho-emotional states. For example, researches noticed [52] that tired people usually type more slowly and make more mistakes, for this reason, every typed sample is stored to make the recognition model more robust.

Two distinctive processes are involved in the keystroke dynamics analysis module:

1) **Feature extraction:** The extracted features (detailed timing information [53]) are time differences between the instants in which:
   a) DT: A key is pressed and released.
   b) PR: A key is pressed and the next key is released.
   c) FT: A key is released and the next is pressed.
   d) PP: A key is pressed and the next key is pressed.
   e) RR: A key is released and the next key is released.

   Based on different analysis carried out in develop and test cycles, DT (dwell time) and FT (flight time) features are considered the most relevant ones and they are weighted accordingly. In addition, a number of typing mistakes (number of presses of such keys such as "Delete" and "Backspace") are calculated separately as auxiliary parameter.
2) **Classification of the extracted features:** This module employs the CNN+RNN model [54] to learn a more complete personal keystroke input mode to carry out continuous authentication. The sequence length of 30 keystroke data (best performance) is vectorized and then divided into fixed-length keystroke feature sequences in order to enable keystroke sequences to be input into the RNN networks. The fact that the input data is pre-processed by CNN (extract a higher-level keystroke feature) improves the performance of the network model.

D. COMPUTER MONITORING

The needs of online proctoring have evolved. In recent times, the market not only seeks to identify students, but also to verify that they are not performing any type of cheating or behaviour that is not allowed with the device on which students perform the activity. In other words, one of the greatest changes is without any doubt the desire to monitor the activity within the device of the students who are doing evaluable activities.

The objective of this development is to obtain an application which is able to monitor the activity carried out by the student within their computer. This monitoring will be done only and exclusively when the student is doing an activity that can be evaluated and supervised by the proctoring system. Because clients can access exams from different operating systems, the objective is to develop a multi-platform application. The user interface is as small as possible so that it does not bother the student during the performance of the evaluable activity. However, it is large and visible enough so that the students know that they are being monitored. The data obtained through the application will be stored in the database or on the servers of the system, therefore, it is necessary that the application complies with all the standards and legislation related to confidentiality and data protection.

The software is developed using Electron JS, a framework that allows multi-platform application development in a simple way. In addition, it is based on web application technologies (as well as data capture modules) which means it does not need to be installed locally on the device in order to be executed. As far as requirements are concerned, the system monitoring tool complies with the following:

1) **Active window detection:** this functionality is one of the key aspects within the application. Not only does the system gets the name of the active window, but it also gets a screenshot of it.
2) **Detection of open/running processes:** This monitoring enables us to know what programs the students open and at what time they have opened them, as well as when they have closed them if the case arises.
3) **Peripheral devices:** A computer has different types of peripheral devices that can be connected. The system knows how many keyboards, microphones, screens and cameras the student has connected to the computer. In the case of cameras, the system also knows the name of them, in order to detect virtual cameras.
4) **Device Information:** Each computer has specific components that make it unique, such as the motherboard or processor. In order to identify if two users use the same computer, information about the computer and its connection is collected: the processor, the motherboard, the IP, the name of the manufacturer, operating system…
5) **Browsing history:** The tool is used especially during evaluable exams, where the students have to answer questions that are presented to them. The student can use any type of browser to look for these answers to these questions. For this reason, the user will be answering correctly without having the necessary knowledge. To combat this type of behaviour, or at least monitor it, the user's browsing history is collected during the activity. Not only the URL, but also the title of the website and the time of entry are registered in the system.
6) **Copy/Paste commands:** Closely linked to the previous point are copying and pasting events. To prevent the student from cheating and copying the answers or sending the test questions to other people, it is necessary to monitor these events. In particular, every event of copying and pasting of text that the user makes during the evaluable activity is recorded. In addition, the screenshots made by the student are collected, for example, if the student screenshots the quiz page to send the exam questions to another person.
7) **Screenshots:** In order to monitor behaviours that we have not yet contemplated, periodic screenshots are made. These screenshots allow the system to identify new methods of cheating.

TABLE 3. Number of captured samples for each type of biometrics.

Images	Audio clips	Keystroke dynamics
373.410	1.007	653

Taking into account that the online student usually uses the same device/browsers/connection to perform their online activities, the information related to computer HW/SW, as well as IP directions are analyzed and their variability in time for the same user is used to trigger more exhaustive automatic and manual authentication and proctoring analysis.

V. TEST

This system was tested through more than 57 activities in 5 different e-learning institutions (3 universities, 2 training centers) in 3 different countries (Latin America, Europe and Asia).

350 students did their assessment activities with the authentication and monitoring system, in three different generic categories: exams (22), short quizzes (10) and forum discussion (25) activities. These activities were chosen because they allow instructors to design activities that need students to spend more time on the platform and have a more complete experience of the biometric authentication and proctoring system.

The courses containing test activities had 3 types of pages: (1) pages of contents, which included texts, schemes and images about the main topic, (2) pages of short quizzes or more extensive exams where the students had to answer questions about what they had read or visualized before and (3) forum activities where instructors promoted discussion related course content through dynamic questions.

Furthermore, the activities were tested in 3 different LMS platforms: Moodle, Blackboard and OpenEdx in order to check the system's compatibility and integrability in the world's most used LMS platforms.

The average time students spent doing these activities was 1 hour and 42 minutes.

A. TECHNICAL TEST

The system captured images randomly every 5-8 second interval, and audio and typing samples every time one of the students spoke or typed text during the activity. The collected data is presented in Table 3. The image/audio/typing algorithms have been tested in depth in each of the captured samples.

All images contain at least 80% of face area (when a person is in the captured photo) and with enough illumination to distinguish facial features after applying brightness and contrast filters (if necessary). On the other hand, the audio samples signal to noise relation (SNR) is acceptable enough to identify the speaker by humans.

B. USER EXPERIENCE TEST

On the other hand, different surveys have been performed during these tests. The objective of this survey was to analyze the perception of students and teachers about the inclusion of these kinds of systems in order to be accepted in the future.

350 students and 50 teachers during the 2018-19 academic year were surveyed about the suitability of this technology. Once they had finished, the students replied to the questionnaire about their experience. In this work we present the most remarkable questions:

1 *Do you think it is appropriate to apply biometric authentication and proctoring to the learning activities?*
2 *Do you think this biometric authentication and proctoring should be implemented in e-learning?*
3 *Do you think this biometric authentication and proctoring should be implemented in all online universities?*
4 *If you could choose, would you prefer to carry out the activities with the incorporation of this software to demonstrate that you have done your activity and you will not be harmed in front of students who ask other people to do the activity?*
5 *Do you think it is fair to monitor distance education in order to avoid cheating?*
6 *Would you feel comfortable if authentication and the monitoring system was working while doing course activities?*

The most remarkable questions for teachers were the following ones:

7 *Would you like to introduce biometric authentication and proctoring tools in your activities?*
8 *Do you think the use of this kind of system will avoid fraud in e-learning activities?*
9 *In your opinion, would the use of the system increase the value and prestige of your courses?*
10 *Do you think authentication and proctoring systems, transparent applications which do not disturb the student, are needed in e-learning environment?*

The questions of the current research are answered with the seven-point Likert scale: Totally disagree (1), Disagree (2), Slightly Disagree (3), Neither agree nor disagree (4), Slightly Agree (5), Agree (6) and Strongly Agree (7).

VI. RESULTS

A. TECHNICAL RESULTS

In this section, the artificial intelligence modules processing results are presented. Since keystroke dynamics samples taken from students cannot be labelled manually (we cannot see or hear), the Table 4 only show an image and audio processing results. The precision and recall data are calculated based on a fully labelled database.

TABLE 4. Performance of authentication and automatic proctoring modules Vs artificial intelligence technologies: a) Authenticating student identity, b) Determining if student is alone or not c) Detecting inappropriate behaviour such as electronic device or book use during online exercises/exams).

	Image processing		Audio processing	
	Precision	Recall	Precision	Recall
Authentication	0.998	0.865	0.964	0.667
Student alone	0.996	0.985	0.963	0.865
Inappropriate behaviour	0.938	0.375	-	-

On the other hand, an analysis of the false positives and negatives of the automatic system has been carried out. Regarding facial authentication, 78% of the failures are a consequence of an excessive face occlusion due to an inappropriate pose and 12% due to poor lighting, mainly caused by the wrong placing of the student against the light. For voice authentication, 53% of failures are due to the low input amplitude of the signal and 33% due to background noise. When determining whether the student is alone or accompanied, motorization based on image processing has failed in 87% of cases due to occlusions (regarding the proximity between individuals or because part of the person protrudes from the image), and 5% because the non-student person is too far away in the image. Finally, sound monitoring has failed by 85% for confusing the second voice (usually with a lower signal amplitude) with background noise and 4% for those samples in which two or more voices have overlapped in the exact same instant. The rest of the errors (including most of the errors in detecting inappropriate behaviour) have been authentication and in monitoring errors made even when the conditions were acceptable for correct automatic operation.

As results table shows, the high precision and recall rates make human intervention almost unnecessary to guarantee 100% of accuracy in final result report. However, human verification is still required. During the tests, all false positive and false negatives (as well as a low rate of true positives and true negatives) were driven to human cross-verification. This action guarantees 100% accuracy in the given final results.

B. USER EXPERIENCE RESULTS

Among other questions, students they were asked whether this system was appropriate to verify the identity of students and proctoring their activities while learning online, which obtained an average of 6.01 in a seven-point Likert scale. However, the opinion of the teachers surveyed about the effectiveness and suitability of this kind of system in an e-learning environment is not as positive as that of the student.

In table 5, the results of the most remarkable questions are analyzed individually.

If we analyze the perceptions of the students based on the results of the most remarkable questions, most of the survey responses have been very positive and welcome. Firstly, students say that it is fair to have any type of biometric recognition software to monitor whether students cheat. Students give an average of 6,03 points in the seven-point Likert scale, in other words, this means that they think it is

TABLE 5. Technology suitability survey results.

	Totally disagree	Disagree	Slightly disagree	Neither agree nor disagree	Slightly Agree	Agree	Strongly Agree
Q Nº	Student results (%)						
1	0	1.67	1.67	3.33	5	53.33	35
2	0	1.67	0	3.33	11.67	43.33	40
3	1.67	1.67	0	3.33	13.33	38.33	41.67
4	0	6.67	5	6.67	13.33	31.67	36.67
5	0	1.67	1.67	8.33	8.33	35	45
6	3.33	8.33	11.67	8.33	21.67	26.67	20
Q Nº	Teacher results (%)						
7	4	6	18	12	30	22	8
8	6	4	12	8	32	28	10
9	2	10	14	8	30	24	12
10	2	0	2	6	36	50	4

appropriate to rate this question with "I agree". Secondly, students were asked if face-to-face universities with a virtual learning platform should implement a software, and they had a good opinion of this question with an average of nearly 6 (specifically 5,79), which corresponds to "I agree" in the seven-point Likert scale.

The main reason why the implementation of the biometric recognition and proctoring software in education are so favourable for 87% of the student asked is that they are conscious about those students who cheat in their tasks and this is not fair for the rest of them.

In this experience, it is noticeable that there are quite positive average values. Thus, the students think that biometric authentication and proctoring is appropriate (in the range between agree and strongly agree on average) for Moodle lessons when these are used for evaluation, in the range between agree and strongly agree on average. In addition, they considered as a positive experience the one they had with the system presented in this work.

Finally, teachers have been asked (with a free answer type question) what are the main reasons that justify surveying the results of the teachers. It is remarkable that all of the reasons are related to privacy issues; they think student will feel a) observed (83%), b) not comfortable (58%), c) worried with the fact that a computer application is recording/managing their personal data (72%) (not real worries for students according to their survey results). Any given reason arguments lack of suitability, effectiveness or convenience of this kind of system use. Moreover, 78% of them explicitly recognise the need for this kind of application to authenticate and monitor online students in their e-learning activities in the near future.

VII. CONCLUSION AND FUTURE WORK

There is a need to know if the student who enrols in an e-learning course is the same student who completes the learning process and receives academic credit. In this work we present an application which offers a continuous authentication identity service of online student through constant biometrics (face, voice, typing) recognition system and a continuous online proctoring and monitoring system. Allowing online courses to take advantage of something that benefits both institutions and students.

The technical results shows that fully automated, continuous (not scheduled), passive (for students), scalable, fully integrated in LMS (with few HW requirements), secure and private biometric authentication and proctoring solutions are affordable and reliable. Furthermore, they exist in the current e-learning supplier market. As future work, more robust biometric models are needed to avoid undesirable deviations due to variance in face pose and light and noise conditions, and reduce human cross-verification needs only for quality warranty purposes (not to complement automatic system limitations).

The study, based on surveys of the uses of the system shows that the solution presented in this work is recognized as a system which is able to verify the identity of students while doing their activities with the purpose of preventing cheating, and as the system should be integrated in LMS as a needed and appropriate solution. Thus, this type of biometric system is positioned as a promising tool to be used in distance education, opening a variety of possibilities to improve the current LMSs. The results provided qualitative and quantitative data that support the use of this kind of software in distance education in order to prevent students from cheating when they are doing their virtual duties.

Institutions, teachers and students can take advantage of this system in their e-learning experience. Students are interested in better and more reliable academic credit for e-learning courses, despite the necessity of classroom exams, to take advantage with his/her competitor in the real-life professional market. The teachers can manage and take decisions during the subject period without having to wait for classroom exams. Finally, the respect of the institution is based on the quality of its study system and results, which are its students. It is crucial to make sure that the person who gets their academic credit in an e-learning environment is the person who completes all the study plan of the institution.

ACKNOWLEDGMENT

The system presented in this work is a commercial solution named Smiley Owl (SMOWL), which is patented under priority number and title: EP18382363-"Method and System for Verifying the Identity of a User of An Online Service Using Multi-Biometric". The authors would like to thank the Smowltech Company which develops the described system and markets it all over the world. They also thank the Vicomtech Research Center which supports the image and audio processing core algorithms explained in this article.

REFERENCES

[1] Kryterion. (2021). *Kryterion Global Testing Solutions*. [Online]. Available: https://www.kryteriononline.com/

[2] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: http://arxiv.org/abs/1704.04861

[3] A. A. Jain, A. K. Flynn, and P. J. Ross, *Handbook of Biometrics*. Springer, 2008. [Online]. Available: https://www.springer.com/gp/book/9780387710402#aboutBook

[4] ProctorU. (2021). *The Leading Proctoring Solution for Online Exams*. [Online]. Available: https://www.proctoru.com/

[5] Examity. (2021). *Better Test Integrity*. [Online]. Available: https://examity.com/

[6] PSIOnline. (2021). *Certification Testing Services and Programs*. [Online]. Available: https://www.psionline.com/en-gb/certification/

[7] ProctorExam. (2021). *Infrastructure for Online Proctoring & Invigilation*. [Online]. Available: https://proctorexam.com/

[8] (2021). *Assessment Tools for Learning Services*. [Online]. Available: https://web.respondus.com/

[9] RemoteProctor. (2021). *Remote Proctor*. [Online]. Available: https://remoteproctor.com/

[10] OnVUE. (2021). *OnVUE*. [Online]. Available: https://home.pearsonvue.com/Test-Owner/Deliver/Online-proctored.aspx

[11] BVirtual. (2021). *Online Proctoring Redefined*. [Online]. Available: https://bvirtualinc.com/

[12] L. Verified. (2021). *Make Your Online Learning Defensible*. [Online]. Available: https://learnerverified.com/

[13] Proctorio. (2021). *A Comprehensive Learning Integrity Platform*. [Online]. Available: https://proctorio.com/

[14] Proctortrack. (2021). *Trusted Exam Integrity | Remote Online Proctoring*. [Online]. Available: https://www.proctortrack.com/

[15] Comprobo. (2021). *OnlineValidation*. [Online]. Available: https://comprobo.co.uk/

[16] Sumadi. (2021). *AI-Powered Proctoring*. [Online]. Available: https://sumadi.net/

[17] ProctorFree. (2021). *Secure Online Proctoring*. [Online]. Available: http://proctorfree.com/

[18] HonorLock. (2021). *Honorlock On-Demand Online Proctoring Service*. [Online]. Available: https://honorlock.com/

[19] ExamSoft. (2021). *Learning Assessments Tools & Software*. [Online]. Available: https://https://examsoft.com/

[20] Y. Khlifi and H. El-Sabagh, "A novel authentication scheme for e-assessments based on student behavior over e-learning platform," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 4, pp. 62–89, 2017. [Online]. Available: https://online-journals.org/index.php/i-jet/article/view/6478

[21] Z. Zhang, M. Zhang, Y. Chang, S. Esche, and C. Chassapis, "A virtual laboratory system with biometric authentication and remote proctoring based on facial recognition," *Comput. Educ. J.*, vol. 7, no. 4, pp. 74–84, 2016.

[22] Z. Zhang, E.-S. Aziz, S. Esche, and C. Chassapis, "A virtual proctor with biometric authentication for facilitating distance education," in *Online Engineering & Internet of Things*, M. E. Auer and D. G. Zutin, Eds. Cham, Switzerland: Springer, 2018, pp. 110–124.

[23] H. S. G. Asep and Y. Bandung, "A design of continuous user verification for online exam proctoring on M-learning," in *Proc. Int. Conf. Electr. Eng. Informat. (ICEEI)*, Jul. 2019, pp. 284–289.

[24] L. K. Musambo and J. Phiri, "Student facial authentication model based on OpenCV's object detection method and QR code for Zambian higher institutions of learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 5, Jan. 2018.

[25] F. Guillen-Gamez, I. García-Magariño, and G. Palacios, *Comparative Analysis Between Different Facial Authentication Tools for Assessing Their Integration in m-Health Mobile Applications*. Cham, Switzerland: Springer, Mar. 2018, pp. 1153–1161. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-77712-2_110

[26] S. Sawhney, K. Kacker, S. Jain, S. N. Singh, and R. Garg, "Real-time smart attendance system using face recognition techniques," in *Proc. 9th Int. Conf. Cloud Comput., Data Sci. Eng. (Confluence)*, Jan. 2019, pp. 522–525.

[27] A. Alshbtat, N. Zanoon, and M. Alfraheed, "A novel secure fingerprint-based authentication system for student's examination system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 9, pp. 515–519, 2019. [Online]. Available: https://thesai.org/Publications/ViewPaper?Volume=10&Issue=9&Code=IJACSA&SerialNo=68

[28] J. V. Monaco, J. C. Stewart, S.-H. Cha, and C. C. Tappert, "Behavioral biometric verification of student identity in online course assessment and authentication of authors in literary works," in *Proc. IEEE 6th Int. Conf. Biometrics, Appl. Syst. (BTAS)*, Sep. 2013, pp. 1–8.

[29] E. Flior and K. Kowalski, "Continuous biometric user authentication in online examinations," in *Proc. Int. Conf. Inf. Technol.*, Jan. 2010, pp. 488–492.

[30] Y. Atoum, L. Chen, A. X. Liu, S. D. H. Hsu, and X. Liu, "Automated online exam proctoring," *IEEE Trans. Multimedia*, vol. 19, no. 7, pp. 1609–1624, Jul. 2017.

[31] A. Okada, I. Noguera, L. Alexieva, A. Rozeva, S. Kocdar, F. Brouns, T. Ladonlahti, D. Whitelock, and A. Guerrero-Roldán, "Pedagogical approaches for e-assessment with authentication and authorship verification in higher education," *Brit. J. Educ. Technol.*, vol. 50, no. 6, pp. 3264–3282, Nov. 2019.

[32] G. Fenu, M. Marras, and L. Boratto, "A multi-biometric system for continuous student authentication in e-learning platforms," *Pattern Recognit. Lett.*, vol. 113, pp. 83–92, Oct. 2018.

[33] L. Slusky, "Cybersecurity of online proctoring systems," *J. Int. Technol. Inf. Manage.*, vol. 29, no. 3, pp. 56–83, 2020.

[34] F. Guillen-Gamez, J. Bravo, and I. García-Magariño, "Students' perception of the importance of facial authentication software in moodle tools," *Int. J. Eng. Educ.*, vol. 33, pp. 84–90, Jan. 2017.

[35] A. Ullah, H. Xiao, and T. Barker, "A dynamic profile questions approach to mitigate impersonation in online examinations," *J. Grid Comput.*, vol. 17, no. 2, pp. 209–223, Jun. 2019, doi: 10.1007/s10723-018-9442-6.

[36] S. A. Razak, N. H. M. Nazari, and A. Al-Dhaqm, "Data anonymization using pseudonym system to preserve data privacy," *IEEE Access*, vol. 8, pp. 43256–43264, 2020.

[37] L. Li, X. Mu, S. Li, and H. Peng, "A review of face recognition technology," *IEEE Access*, vol. 8, pp. 139110–139120, 2020.

[38] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "FaceBoxes: A CPU real-time face detector with high accuracy," in *Proc. IEEE Int. Joint Conf. Biometrics (IJCB)*, Oct. 2017, pp. 297–309. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0925231219310719

[39] L. Unzueta, W. Pimenta, J. Goenetxea, L. P. Santos, and F. Dornaika, "Efficient generic face model fitting to images and videos," *Image Vis. Comput.*, vol. 32, no. 5, pp. 321–334, May 2014.

[40] X. Liu, X. Ma, J. Wang, and H. Wang, "M3L: Multi-modality mining for metric learning in person re-identification," *Pattern Recognit.*, vol. 76, pp. 650–661, Apr. 2018.

[41] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823. [Online]. Available: https://ieeexplore.ieee.org/document/7298682

[42] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*. 1994. [Online]. Available: https://dl.acm.org/doi/10.5555/180895.180940

[43] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, *SSD: Single Shot Multibox Detector* (Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, 2016. [Online]. Available: https://www.springer.com/gp/book/9783319464770

[44] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, "Semantic image inpainting with deep generative models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6882–6890. [Online]. Available: https://ieeexplore.ieee.org/document/8100211

[45] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst. (NIPS)*, vol. 2. Cambridge, MA, USA: MIT Press, 2014, pp. 2672–2680. [Online]. Available: http://dl.acm.org/citation.cfm?id=2969033.2969125

[46] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlicek, Y. Qian, P. Schwarz, J. Silovsky, G. Stemmer, and K. Vesely, "The kaldi speech recognition toolkit," in *Proc. IEEE Workshop Autom. Speech Recognit. Understand. (ASRU)*, Waikoloa, HI, USA. Piscataway, NJ, USA: IEEE Signal Processing Society, Dec. 2011, p. 30. [Online]. Available: https://dblp.org/db/conf/asru/asru2011.html

[47] L. Wan, Q. Wang, A. Papir, and I. L. Moreno, "Generalized end-to-end loss for speaker verification," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 4879–4883. [Online]. Available: https://ieeexplore.ieee.org/document/8462665

[48] Q. Wang, C. Downey, L. Wan, P. A. Mansfield, and I. L. Moreno, "Speaker diarization with LSTM," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 5239–5243. [Online]. Available: https://ieeexplore.ieee.org/document/8462628

[49] Y. Zhong, Y. Deng, and A. K. Jain, "Keystroke dynamics for user authentication," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2012, pp. 117–123. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6239225

[50] H. Crawford, "Keystroke dynamics: Characteristics and opportunities," in *Proc. 8th Int. Conf. Privacy, Secur. Trust*, Aug. 2010, p. 20108.

[51] A. T. Kiyani, A. Lasebae, K. Ali, M. U. Rehman, and B. Haq, "Continuous user authentication featuring keystroke dynamics based on robust recurrent confidence model and ensemble learning approach," *IEEE Access*, vol. 8, pp. 156177–156189, 2020.

[52] A. F. M. N. H. Nahin, J. M. Alam, H. Mahmud, and K. Hasan, "Identifying emotion by keystroke dynamics and text pattern analysis," *Behav. Inf. Technol.*, vol. 33, no. 9, pp. 987–996, Sep. 2014.

[53] R. Moskovitch, C. Feher, A. Messerman, N. Kirschnick, T. Mustafić, A. Camtepe, B. Löhlein, U. Heister, S. Möller, L. Rokach, and Y. Elovici, "Identity theft, computers and behavioral biometrics," in *Proc. IEEE Int. Conf. Intell. Secur. Informat.*, Jun. 2009, pp. 155–160. [Online]. Available: https://ieeexplore.ieee.org/document/5137288

[54] L. Xiaofeng, Z. Shengfei, and Y. Shengwei, "Continuous authentication by free-text keystroke based on CNN plus RNN," *Procedia Comput. Sci.*, vol. 147, pp. 314–318, Jan. 2019.

MIKEL LABAYEN received the degree in technical telecommunication engineering from the Public University of Navarre, in 2005, with a focus on image and sound, and the degree from the Faculty of Telecommunication Engineering, Public University of Navarre, in 2007. He is currently pursuing the Ph.D. degree in computer vision and machine learning. He completed his undergraduate dissertation at the Electronic Engineering Department, University of Surrey, U.K. He completed his master thesis at the Digital Television and Multimedia Services Department, Vicomtech Research Center. From 2007 to 2012, he started his professional carrier as a Staff Researcher in computer vision-multimedia content analysis with the Vicomtech Research Center. In 2012, he was the Co-Founder and the Research and Development Project Manager with Smowltech, start-up (spin-off of Vicomtech), where he researches in the automatic facial and voice recognition area developing applications for online user authentication based on human biometrics. In the same period, he also was an Associate Teacher with the Electronic Technology Department, University of the Basque Country. He is currently working in the field of intelligent transport systems (ITS) designing computer vision and machine learning-based solutions for autonomous train operations in the railway sector with the CAF Group. His work as a Researcher includes a number of publications and four patents.

RICARDO VEA is currently the CEO and the Business Development Manager of Smowltech. He has worked more than twenty years in the world of HR in different organizations, where he has developed director and consultant roles. He is also a Professor-Collaborator with the Deusto Business School and previously with the UPV. Before entering in Smowltech, he was developing a Cyber-Security Certification Agency for a multinational computing security company. He is an Expert in e-learning and an enthusiast of modern technologies for teaching training. His research in social psychology and the e-learning market knowledge provide Smowltech team with the social and market analysis abilities required that will ensure the success of the project.

JULIÁN FLÓREZ (Member, IEEE) received the degree in industrial engineering from the University of Navarra, in 1980, and the Ph.D. degree in adaptive control from the University of Manchester, Institute of Science and Technology (UMIST), in 1985. From 1985 to 1990, he worked as a Researcher with the Center of Study and Technical Research of Gipuzkoa (CEIT), where he collaborated in several research projects related to electrical and industrial engineering with a marked industrial focus. From 1985 to 1994, he was an Associate Professor with the School of Industrial Engineering, University of Navarra, where he has been a Professor, since 1994. From 1990 to 1997, he worked as a Senior Researcher with CEIT, where he was an In Charge of the Department of Industrial Applications. From 1997 to 2001, he worked as the Director of Corporate Development of AVANZIT-SGT (Servicios Generales de Teledifusión) in the fields of Information Systems, Communications and Broadcasting infrastructure. He has a strong background in digital television infrastructures and was closely involved in the deployment of one of the biggest digital TV organizations in Spain and Europe, Quiero TV. Since 2001, he has been the Principal Researcher with Vicomtech. He holds some patents and has written more than 40 research articles in different areas of industrial and electrical engineering.

NAIARA AGINAKO received the degree in telecommunications engineering from the University of the Basque Country, in 2005, and the Ph.D. degree in image analysis and content-based retrieval of images and video. From 2003 until 2005, she collaborated the Signal Processing and Communication Group, Electronics and Telecommunications Department. She developed and managed research projects with Digital Media Department, Vicomtech, from 2005 to 2015. Since 2015, she has been teaching at the Polytechnic School of Donostia and the Faculty of Informatics. Her work as a Researcher includes a number of publications and two patents.

BASILIO SIERRA received the B.Sc. degree in computer sciences, the M.Sc. degree in computer science and architecture, and the Ph.D. degree in computer sciences from the University of the Basque Country, in 1990, 1992, and 2000, respectively. He is currently a Full Professor with the Computer Sciences and Artificial Intelligence Department, the University of the Basque Country. He is also the Director of the Robotics and Autonomous Systems Group, Donostia-San Sebastian. He is a Researcher in the fields of robotics, computer vision and machine learning, and he is working on the development of different paradigms to improve classification behaviours. He has written more than 25 journal articles in those fields, more than 100 conference contributions, and more than 30 book chapters.