Universidad del País Vasco
Euskal Herriko Unibertsitatea

Máster Universitario en Ingeniería Computacional y Sistemas Inteligentes

---

Prototipo CAD de segmentación automática de cáncer de pulmón en imágenes histopatológicas TMA

Tesis de Máster

Jefferson Jair Arcos

Directores:
Phd. Ignacio Arganda Carreras
Phd. Carlos Ortiz de Solórzano

Computer Vision and Pattern Discovery Group

Donostia, Julio de 2017

Agradecimientos

Agradezco a Dios por darme la visión, fortaleza y ayuda para cursar este máster.

Al profesor Yosu Yurramendi director del programa de máster, gracias por la valiosa oportunidad y la fe depositada en mi.

Estoy muy agradecido con el Dr. Ignacio Arganda director de este trabajo por su orientación, consejo, seguimiento constante y paciencia para conmigo. También agradezco al Dr. Carlos Ortiz de Solórzano director del laboratorio de imagen del cáncer y de la unidad de imagen del CIMA, por el apoyo y colaboración en el desarrollo de este trabajo.

Quiero hacer extensiva mi gratitud al Gobierno de Colombia quien mediante la fundación CEIBA me otorgó ayuda financiera parcial para realizar este estudio.

Por último agradezco profundamente la comprensión, el apoyo y aliento recibido de mi familia y amigos.

Índice general

Índice de figuras

Índice de cuadros

Capítulo 1

Introducción

El cáncer de pulmón es una enfermedad letal que para el 2012 se situó como la quinta causa de muerte a nivel mundial, la tercera en Europa y la primera en España con casi 20.000 nuevos casos cada año; aproximadamente el 85 % de los sujetos que padecen cáncer de pulmón, morirán por esta enfermedad (Alberto *et al.* [1]). El principal obstáculo en la lucha contra esta patología es su detección tardía.

El desarrollo que ha experimentado el campo de la imagen médica en aspectos como la adquisición, almacenamiento y visualización ha contribuido al mejoramiento de la calidad del análisis y diagnóstico de las diferentes patologías (entre ellas el cáncer de pulmón) convirtiéndola actualmente en un componente indispensable en medicina.

En las últimas décadas, se han realizado numerosos esfuerzos para detectar de manera precoz el cáncer de pulmón mediante el desarrollo de distintas tecnologías, entre ellas los sistemas de diagnóstico asistido por computador (CAD), los cuales mediante el análisis automático de la imagen médica brindan al especialista una segunda opinión diagnostica, con el objetivo de obtener diagnósticos mas precisos que permitan formular tratamientos mas adecuados.

La imagen médica histopatológica es el "gold standard.[en] detección temprana de la mayoría de patológicas incluido el cáncer de pulmón. La tarea de detección suele ser bastante tediosa e que implica una importante inversión de tiempo y esfuerzo por parte de los expertos en histopatología. Shazia *et al.* [2] dice que el crecimiento de los bancos de tejidos ya ha superado las habilidades manuales de análisis disponibles. Además, la revisión de patología experta sufre variaciones ínter e intra observador. Lo anterior evidencia la gran necesidad de automatizar el análisis de imagen médica en histopatológica.

En este trabajo se hace una aproximación a la detección de cáncer de pulmón en imagen médica, concretamente abordando el problema de segmentación de tejido tumoral y no tumoral sobre imágenes histopatológicas TMA, mediante el desarrollo de un prototipo de sistema de diagnóstico asistido por computador CAD.

1.1. Objetivos

1.1.1. Objetivo Principal

Construir un sistema CAD que contribuya a la detección temprana y diagnóstico del cáncer de pulmón sobre imágenes histopatológicas

1.1.2. Objetivos Específicos

- Conformar una base de datos de imágenes TMA de cáncer de pulmón, marcadas por expertos para la experimentación.
- Investigar el estado del arte.
- Proponer un método eficaz de detección temprana de cáncer de pulmón.
- Evaluar el rendimiento del método.
- Proponer nuevas líneas de trabajo.

1.1.3. Distribución de la tesis

La tesis se ha organizado en 7 capítulos (incluido la introducción), los 6 restantes se resumen de la siguiente manera:

- El capítulo 2, inicia introduciendo los fundamentos sobre cáncer de pulmón, imagen médica histopatológica y diagnóstico asistido por computador CAD. Después se hace un repaso del estado del arte sobre sistemas CAD para detección de cáncer de pulmón para finalmente presentar una revisión del trabajo más relevante haciendo énfasis en sistemas que hayan empleado imagen histopatológica.
- En el capítulo 3, se detalla el conjunto de datos y se presenta el software empleado para la experimentación.
- El capítulo 4, expone el método propuesto, se inicia mostrando un esquema general que posteriormente se explica paso a paso primero introduciendo los conceptos de las técnicas empleadas y luego describiendo los procedimientos seguidos.
- En el capítulo 5, se presentan los resultados obtenidos siguiendo un orden similar al capítulo 4.
- En el capítulo 6, se hace la interpretación de los resultados, señalando las consecuencias teóricas y posibles aplicaciones, así como también las limitaciones, para finalmente presentar las conclusiones.
- Finalmente en el capítulo 7, se formulan las futuras líneas de trabajo.

Capítulo 2

Estado del Arte

2.1. Fundamentos

2.1.1. Los Pulmones

Los pulmones son órganos esenciales del aparato respiratorio, estos son responsables de proporcionar oxígeno al torrente sanguíneo, además de eliminar el dióxido de carbono. A lo largo de toda la vida, una persona puede usar sus pulmones para respirar más de mil millones de veces. Los plumones están situados en la cavidad torácica y están recubiertos por una doble membrana lubricada llamada pleura (Frank *et al.* [3]). El pulmón derecho consta de tres lóbulos mientras que el pulmón izquierdo es ligeramente más pequeño y consta de sólo dos lóbulos (Figura 2.1). Mediante el movimiento del diafragma, los pulmones aspiran el aire del ambiente para extraer el oxígeno y lo expulsan para eliminar el dióxido de carbono, este proceso se repite continuamente, incluso mientras dormimos.

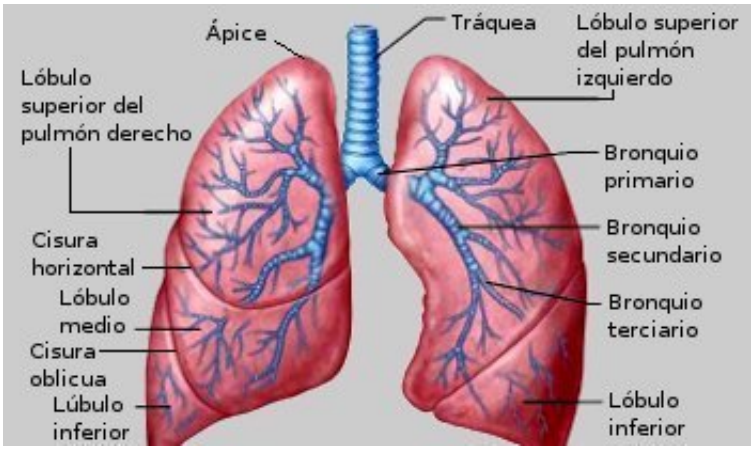


Figura 2.1: Representación de los pulmones.

2.1.2. El cáncer de pulmón

El cáncer pulmonar es un tumor maligno que consiste en el crecimiento anormal de células tanto pulmonares como bronquiales, existen múltiples factores que aumentan la probabilidad de desarrollar cáncer de pulmón, el más importante es el consumo de tabaco ya que se encuentra relacionado con alrededor del 90 % de los tumores de pulmón. Los

síntomas comunes del cáncer de pulmón incluyen:

- Una tos que no desaparece y empeora con el tiempo
- Dolor constante en el pecho
- Tos con expectoración con sangre
- Falta de aire, silbidos al respirar o ronquera
- Problemas repetidos por neumonía o bronquitis
- Inflamación del cuello y la cara
- Pérdida del apetito o pérdida de peso
- Fatiga

Hay dos grandes tipos: el cáncer de pulmón de células pequeñas (CPCP) y el cáncer de pulmón de células no pequeñas (CPCNP), este último es el más frecuente con aproximadamente un 80 % de los casos divididos en dos grupos:

a) **Adenocarcinoma:** Es el más frecuente ya que se desarrolla en células glandulares (secretoras) que están en continua división y que se encuentran en el tejido que reviste los órganos, este tipo de cáncer es el más común en pacientes no fumadores.

b) **Carcinoma escamoso:** Es el más relacionado a pacientes que consumen tabaco, aunque tiene asociado el mejor pronóstico ya que generalmente se detecta y trata tempranamente (Carla *et al.* [4]).

Lastimosamente el cáncer de pulmón suele detectarse cuando está en etapa avanzada generalmente por sus síntomas, cuando ya ha comprometido otros órganos. A decir verdad, esta es la principal dificultad en la lucha con esta enfermedad.

Las sospechas del cáncer de pulmón pueden darse porque el paciente presenta sus síntomas y acude al especialista o bien indirectamente al practicar otros exámenes. Generalmente para el diagnóstico del cáncer de pulmón primero se inicia haciendo una exploración física del paciente y analizando su historia clínica, luego se obtiene una muestra de tejido (biopsia) para confirmar su presencia, finalmente se buscará determinar el grado de extensión a través de un examen de tomografía computarizada (TC) o radiografía de tórax (Julia *et al.* [5]).

El tratamiento para el cáncer de pulmón depende del tipo de cáncer, así como de lo avanzado esté y de cuán saludable se encuentre el sujeto. Los más comunes son:

- La cirugía para extirpar el tumor se puede hacer cuando este no se haya propagado más allá de los ganglios linfáticos cercanos.
- La quimioterapia utiliza medicamentos para destruir las células cancerosas y detener el crecimiento de las nuevas células.

- La radioterapia utiliza potentes rayos X u otras formas de radiación para destruir las células cancerosas.

Finalmente el pronóstico del cáncer pulmonar depende de que tanto se haya diseminado, en general la esperanza de vida está en torno a los 5 años aunque sin tratamiento se reduce a entre 3 y 6 meses por lo cual siempre que se pueda se debería iniciar un tratamiento lo mas pronto posible [6, 7].

2.1.3. Tissue MicroArray Analysis (TMA)

TMA es una técnica muy eficiente para el análisis de tejidos histológicos, actualmente es muy empleada en la detección de múltiples tipos de cáncer, entre ellos el de pulmón (Valentina *et al.* [8]), en la figura 2.2 se muestra el procedimiento que consiste en tomar bloques de parafina en los que se insertan pequeñas muestras de tejido cortado milimétricamente en forma de círculo formando una matriz, posteriormente estas se analizan para obtener bio-marcadores del paciente (Nazar [9]).

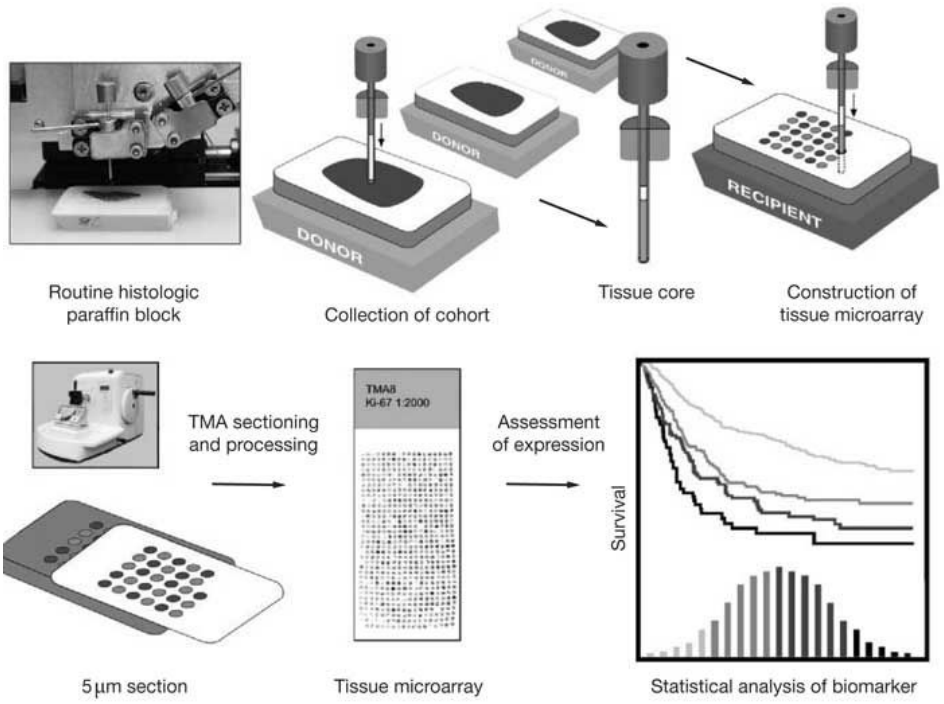


Figura 2.2: Esquema del TMA.

El objetivo de TMA es efectuar estudios en el menor tiempo posible realizando un procesamiento más veloz, simultáneo y estandarizado de múltiples muestras (figura 2.3) además de asegura el tamaño de área de las muestras para que el análisis sea suficiente, con lo cual se superan las limitaciones del análisis tradicional de tejidos (Iqbal *et al.* [10]). Juliana *et al.* [11] menciona que TMA hace posible el uso de muestras randomizadas representativas de una lesión, susceptibles de ser evaluadas por métodos automáticos (reconocimiento de imagen, densidad nuclear, densidad cromática, resultados de tinciones inmunohistoquímicas) lo cual abre un camino al diagnóstico automatizado en anatomía

patológica, se dice que este es uno de los campos diagnósticos con más intervención humana y mayor subjetividad diagnóstica.

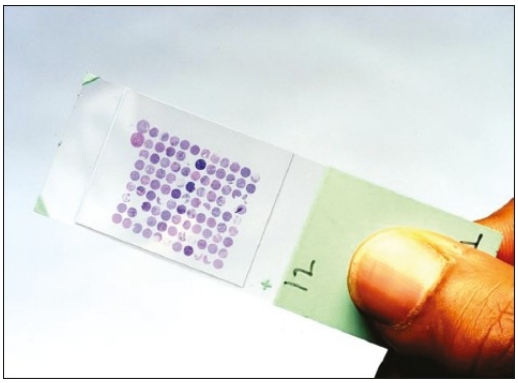


Figura 2.3: Una muestra de TMA

2.1.4. Diagnóstico asistido por Computador (CAD)

Los primeros trabajos en diagnóstico asistido por computador, más comúnmente denominado CAD (Computer Aided Diagnosis), se registran por la década de los 60. Al principio, la comunidad científica los concibió como sistemas de diagnóstico automático con la convicción de que podrían diagnosticar enfermedades sin ayuda de los expertos humanos. Sin embargo con el tiempo estas expectativas se fueron diluyendo teniendo en cuenta que por aquel tiempo aún no se contaba con la tecnología tanto de software como de hardware necesaria. Más tarde por los años 80, surgió otro enfoque mas realista que planteaba la salida del computador como una segunda opinión diagnóstica con el objetivo de aumentar la precisión del diagnóstico aunque sería el experto humano quien tomase la decisión final. Con el tiempo este nuevo enfoque fue evolucionado positivamente y para el año 1990 se presentó el primer trabajo sobre sistemas CAD en donde se afirmaba que su empleo había aportado una mejora estadísticamente significativa al diagnóstico de cáncer de seno (Heang *et al.* [12]), después se publicaría sobre los beneficios de CAD en otro tipo de lesiones como microcalcificaciones agrupadas, masas en mamografías, nódulos pulmonares, aneurismas y radiografía de tórax (Doi, 2007 [13]). De esta forma CAD ha evolucionado positivamente a lo largo del tiempo en cuanto a detección de lesiones en imagen médica.

CAD generalmente es definido como “sistemas que realizan un análisis cuantitativo sobre la imagen médica y su resultado se tiene en cuenta por un especialista al momento de emitir un diagnóstico” (Doi, 2005 [14]). Por otra parte se dice que los sistemas CAD analizan la imagen médica, diferenciando los patrones normales y anormales, con el fin de detectar, cuantificar y caracterizar regiones de interés. Estos al final actúan como una “segunda opinión” que complementa a la del especialista quien como ya se mencionó antes es quien emite el diagnóstico final. Algunos objetivos que se busca con el desarrollo y empleo de sistemas de diagnóstico asistido por computador son:

- Aumentar el rendimiento del experto al enfocar su atención a regiones sospechosas que pudieran haber pasado desapercibidas.

- Disminuir el tiempo de diagnóstico en cada caso.
- Aportar una segunda opinión más objetiva teniendo en cuenta que es hecha por una máquina.
- Contribuir con sistemas de auto-evaluación para radiólogos con diversos grados de experiencia y de aprendizaje para residentes en periodo de formación.
- Ayudar en la detección temprana de la enfermedad.

Un suceso histórico importante fue la aprobación otorgada por la FDA (Food and Drug Administration, (USA)) en 1998 para el uso del R2 Technology un sistema CAD enfocado en la detección de lesiones en mamografía (Julian *et al.* [15]). De ahí en adelante estos sistemas han evolucionado positivamente hasta ser hoy en día ampliamente usados por hospitales, laboratorios y clínicas alrededor del mundo, convirtiéndose en herramientas de uso común para apoyar el diagnóstico en imagen médica. Cada vez son más los estudios enfocados hacia del desarrollo de CAD en múltiples ámbitos, en la revisión hecha en (Doi, 2007 [13]) se reporta que principalmente son el seno, tórax y colon aunque también se trabaja sobre el cerebro, hígado, sistemas esquelético y vascular.

Diagnóstico de cáncer de pulmón asistido por computador

Desde el principio, el diagnóstico de cáncer pulmonar ha sido una de las principales áreas de trabajo en CAD, el primer reporte fue en 1963 (Gwilym *et al.* [16]) en donde se empleó el computador para la detección de nódulos pulmonares en radiografía de tórax, desde entonces la mayoría de esfuerzos se han enfocado en el campo de la radiología aunque con los avances en el área de la imagen médica en la actualidad se trabaja también en campos alternativos como la histopatología en donde la comunidad científica ha observado un gran potencial.

La aparición de la tomografía computarizada supuso una oportunidad para la construcción de sistemas CAD para detección de nódulos pulmonares [17–24], se ha demostrado que este tipo de sistemas han logrado mejoras significativas en el diagnóstico emitido por los radiólogos que los usaron frente al de aquellos radiólogos que no los usaron (Kyung *et al.* [25]).

En la revisión sobre CADs para la detección de nódulo pulmonar hecha por Macedo *et al.* [26] se analiza la precisión del diagnóstico que tienen estos sistemas actualmente, aquí sobresalen dos casos por sus buenos resultados además de que se validaron de forma robusta; el primero es Tan *et al.* [27] en el cual se emplearon 574 nódulos diferentes con una sensibilidad del 87,5 % con 4 falsos positivos por caso, el segundo Kumar *et al.* [24] en donde se emplearon 538 nódulos diferentes y se obtuvo un 86 % de sensibilidad con 2,17 falsos positivos por caso. Con lo anterior se aproxima el estado actual del rendimiento de este tipo de sistemas CAD que a pesar de obtener buenos resultados, han avanzado lentamente debido al alto grado de dificultad y rigurosidad inherente en la realización de esta clase de estudios.

Si bien se logra detectar y diagnosticar el cáncer sobre imágenes TC, en radiología no se está cubriendo un aspecto importante como lo es la detección temprana del cáncer, ya

que el objetivo del examen TC es determinar el grado de extensión de la enfermedad, es decir, cuando esta ya ha avanzado lo suficiente como para ser visualizada por el ojo del radiólogo (ver figura 2.4), esto implica que este tipo de sistemas CAD no se enfocan en la detección temprana del cáncer de pulmón.

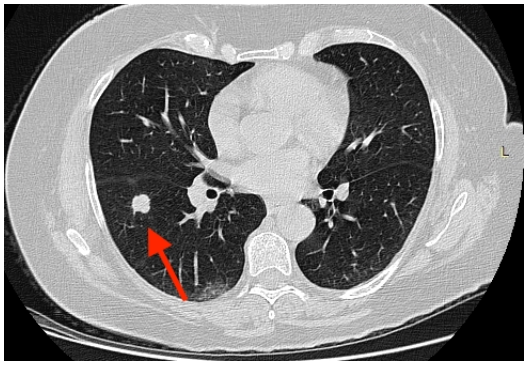


Figura 2.4: Nódulo pulmonar identificado en TC tórax.

2.2. Motivación

Hoy en día el único método definitivo para confirmar la presencia de la mayoría de tipos de cáncer (incluido el de pulmón) así como también para clasificarlos, es según Raphael *et al.* [28] un análisis histopatológico de una biopsia o muestra de tejido, ya que este permite visualizar claramente el estado de la enfermedad además de ayudar con la formulación de un tratamiento más adecuado, en este sentido se han realizado diversas investigaciones sobre sistemas CAD para el análisis de imágenes histopatológicas. Los objetivos en este aspecto son ambiciosos ya que no solo se intenta crear sistemas que ayuden con la detección y clasificación de la enfermedad sino que además puedan cuantificarla, debido a la necesidad de conocer con más precisión el estado del cáncer para mejorar su estimación y predecir su progreso. Un aspecto a diferenciar es que en radiología según Metin N. *et al.* [29] se han construido CAD para detectar y caracterizar el cáncer; sin embargo en la histopatología la simple identificación de la presencia o ausencia de cáncer o incluso su extensión espacial exacta puede no tener tanto interés con respecto a preguntas más sofisticadas como: ¿cuál es el grado del cáncer?, Además a escala histológica (microscópica) se puede empezar a distinguir entre diferentes subtipos histológicos de cáncer, lo cual es absolutamente imposible (o al menos difícil) en la escala radiológica teniendo en cuenta la limitación inherente en la resolución espacial de los datos radiológicos.

A diferencia de radiología en donde las imágenes están en escala de grises, en histopatología las imágenes se obtienen en color. Además los conjuntos de datos son considerablemente mayores y diversos. Por ejemplo: una radiografía de tórax de alta resolución tiene alrededor de 134 millones de voxels, una biopsia de próstata podría contener entre sus muestras unos 4000 millones de pixeles, incluso su digitalización y almacenamiento han sido retos difíciles de sobrepasar. Estas limitaciones que en el pasado podrían haber impedido el avance de CAD en histopatología, en la actualidad ya no suponen ningún obstáculo gracias al gran avance logrado en imagen digital, almacenamiento y sistemas

PACS (sistema de archivado y transmisión de imágenes).

La necesidad de crear sistemas de diagnóstico asistido por computador que aprovechen los avances logrados en imagen médica histopatológica para la detección temprana de las distintas enfermedades es apremiante.

En la siguiente sección se estudiará el trabajo realizado sobre el desarrollo de sistemas CAD para detección de cáncer de pulmón sobre imágenes histopatológicas, esto es el centro del presente estudio.

2.3. Trabajo Relacionado

El análisis de imagen histopatológica es el "Gold standard" en cuanto a detección y diagnóstico del cáncer, la facilidad actual en digitalización y gestión de la imagen médica ha evidenciado la necesidad de crear herramientas para automatizar esta tarea [30,31]. El problema básico es segmentar y clasificar correctamente las regiones de interés. El trabajo en CAD así como los métodos y enfoques empleados son diversos.

Para empezar, la identificación y evaluación de estructuras histológicas como linfocitos, núcleos y glándulas permiten obtener indicadores sobre la presencia o gravedad de la enfermedad; un gran número de estudios [32–43] desarrollaron métodos que lograron automatizar estas tareas, aunque algunos han empleado imágenes TMA, todos usaron imágenes histopatológicas.

Distinguir de forma más global las regiones de interés como las zonas tumorales y no tumorales dentro de la imagen histopatológica es uno de los retos que enfrentan a diario los expertos y es también el principal componente de interés para este estudio. Un método de reconocimiento de cáncer de mama fue desarrollado por Shazia *et al.* [2], este se formula como un problema de clasificación de superpixeles. Al proceso de segmentación mediante SLIC superpixeles se complementa de forma novedosa con la adición de información espacial invariante a la rotación a lo cual los autores denominaron Rotation Invariant Superpixel Pyramid (RISP). Este método se valida con un conjunto de datos de 32 imágenes TMA, obteniendo una precisión del 69.2 %, un resultado mayor comparado con dos métodos del estado del arte en su momento.

Para detectar adenocarcinoma de próstata en (Scott *et al.* [44]), mediante un enfoque bayesiano se inició asignando una probabilidad de tumor a las características extraídas de la imágenes lo cual se tradujo en una mejora para la clasificación con el algoritmo AdaBoost. El sistema fue evaluado en 33 casos de cáncer y su resultado fue comparado con el marcado hecho manualmente por los patólogos. Finalmente se reportó un acierto del 88 %.

Un método para la detección de cáncer de pulmón fue presentado en (Cataldo *et al.* [45]). La clasificación fue hecha a modo de comparativa entre el enfoque no supervisado y el supervisado empleando los algoritmos K-means y Support Vector Machines (SVM) respectivamente. El resultado indicó que K-means alcanzó una tasa de acierto del 90 % superando a SVM en un 8 %.

En (Bilge *et al.* [41]) se hace la segmentación del cáncer de mama empleando un método estadístico basado en la umbralización del color en el espacio CIELab. Esto según

los autores, aprovecha mejor la disimilitud del color entre diferentes regiones. Al contrastar los resultados con imágenes marcadas por expertos, se informa de una especificidad del 95 %.

Como una estrategia de mejora en la detección y clasificación del cáncer de próstata, en (Ali *et al.* [46]) se emplea un marco de clasificación supervisada que compara el rendimiento de los clasificadores: Gaussiano (basado en la función discriminante de Bayes), KNN, y SVM. La imagen es descompuesta en sus 3 canales para obtener características de color, textura y señales morfo-métricas a nivel de los objetos globales e histológicos. Al final en detección del cáncer se obtuvo un acierto del 96.7 % y en clasificación del tipo un 81 %, para ello se contó con dos datasets de 367 y 268 imágenes TMA respectivamente.

La clasificación del tipo de cáncer es también un reto de CAD. En (Ching-Wei *et al.* [47]) se propuso un método de discriminación en tiempo real entre los dos principales tipos de cáncer de pulmón: el de células pequeñas y el de células no pequeñas; primero se extrajeron características densitométricas y de tipo Haralick y se clasificó con varios algoritmos pero el mejor resultado fue para AdaBoost.M1 con un rendimiento del 92.4 % para imágenes TMA (369 muestras) y un 95.4 % en imágenes tradicionales o de corte completo (284 muestras). Aunque lastimosamente no se especifican los tiempos de cómputo, el estudio hace énfasis en la detección en tiempo real lo cual es un aspecto a tener en cuenta para el trabajo a realizar.

En (Sieren *et al.* [48]) se desarrolló un método basado en clustering para segmentar imágenes histológicas de corte transversal de nódulos pulmonares, el resultado es comparado con segmentaciones manuales (73 % precisión de los expertos) alcanzando una precisión máxima del 72 %; la segmentación automática es eficiente (menos de 1 minuto) comparada con la segmentación manual que según los investigadores llevó varias horas.

En los últimos años ha surgido un nuevo enfoque de trabajo sobre imagen denominado Deep Learning, en (Angel *et al.* [49]) se propone un método automático de aprendizaje profundo de detección y análisis de carcinoma ductal invasivo en imágenes enteras histopatológicas de cáncer de mama, el dataset se conforma de 162 imágenes de pacientes diagnosticados con dicha enfermedad, el método es evaluado distribuyendo 113 imágenes para entrenamiento y 49 para test. Los resultados indicaron una precisión del 84.23 % y una medida F de 71.80.

Una sistema CAD llamado TissueMark para el reconocimiento, anotación y análisis de tumores de cáncer de pulmón de células no pequeñas en imágenes H&E, fue desarrollado por Hamilton *et al.* [50] utilizando un conjunto de 136 imágenes, se extrajeron un grupo (patentado) de características optimizadas mediante función de base radial gaussiana para el algoritmo SVM. El sistema logra discriminar entre el tejido de tumor y no tumor, con una precisión del 89.02 %, similar al obtenido por otros métodos del estado del arte.

Por último en un esfuerzo por obtener datos de supervivencia de los pacientes, se han desarrollado métodos que además de detectar el cáncer, pueden obtener indicadores asociados a la supervivencia y el pronóstico de vida de los pacientes. El primer trabajo hecho por Kun-Hsing *et al.* [51] se enfoca en el diagnóstico de cáncer de pulmón. Aquí se observa una comparativa de clasificadores entre los cuales Random Forest obtiene el mejor resultado con una medida AUC (Area Under the Curve) de 0.85. En el segundo

trabajo (Andrew H. *et al.* [52]) se propone un método para detección de cáncer de seno basado en pre-segmentación por superpixeles, a los cuales se les extraen características tipo estadísticos de primer orden, después se emplea un clasificador de regresión logística regulada que consigue una precisión del 89 % en la segmentación. Ambos estudios tienen en común el empleo de datasets de la Stanford TMA database. La cuantificación del cáncer es indispensable dentro de este tipo de estudios. Es muy interesante que se utilicen las segmentaciones para obtener datos de supervivencia de los pacientes, este es un aspecto de suma importancia y es tenido en cuenta en el marco de este trabajo.

Capítulo 3

Materiales

3.1. El conjunto de datos

EL conjunto de datos ha sido construido en colaboración con el CIMA (Centro de Investigación Médica Aplicada) en Pamplona, quienes proporcionaron un total de 10 imágenes TMA de distintos tumores de pulmón del tipo adenocarcinoma, con una tinción de inmunohistoquímica (IHC) para el marcador RRM2. Estas imágenes fueron marcadas manualmente por expertos empleando el software ImageJ. Cabe mencionar que con la intención de hacer una comparativa con el trabajo relacionado, se hizo la búsqueda de conjuntos de datos similares aunque no se encontró ninguno con los requerimientos necesarios (esencialmente imágenes marcadas por expertos) ya que aunque existen iniciativas enfocadas a la conformación de bases de datos abiertas de imágenes TMA como la Stanford TMA database (Robert J. *et al.* [53]), es algo difícil encontrar el respectivo marcado de la imagen por parte del experto.

3.1.1. Imágenes TMA

Las imágenes TMA proporcionadas tienen un aspecto similar a la que se muestra en la figura 3.1. La imagen es cuadrada, con el tejido en forma de círculo diferenciado del fondo color gris claro, con un tamaño aproximado de 6000 x 6000 pixeles en formato JPEG. Por otra parte a cada imagen corresponde un archivo con el marcado correspondiente a las zonas tumorales en formato ROI, realizado con el programa ImageJ empleando su herramienta de selección de polígonos.

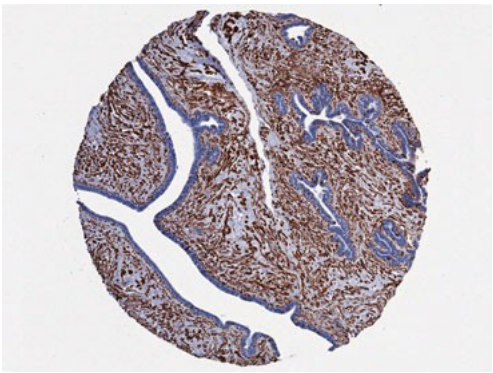


Figura 3.1: Diapositiva TMA digitalizada.

3.1.2. Imágenes TMA etiquetadas

El objetivo propuesto por los expertos fue el de diferenciar 3 regiones de interés: el fondo, la zona tumoral y la zona no tumoral. Como inicialmente la imágenes solo contaban con el marcado del área tumoral, fue necesario crear imágenes marcadas con las 3 zonas a las cuales se les llamó "imágenes de etiquetas". Inicialmente esta tarea pareció trivial ya que al tener marcada la zona tumoral (que es la más importante) solo quedaría marcar la zona no tumoral y el fondo, sin embargo para intentar evitar la subjetividad se optó por automatizar la tarea con el programa de procesado de imágenes GIMP, usando su herramienta de selección difusa, con ella se asignó la etiqueta 0 (blanco) a la zona del fondo, luego a la zona tumoral como ya estaba marcada se asignó la etiquetada 2 (rojo), quedando finalmente la zona no tumoral que se seleccionó haciendo una exclusión con las primeras zonas, asignándole la etiqueta 3 (color verde); este procedimiento podría realizarse de muchas formas por tanto no se entrará en más detalles. En la figura 3.2 se muestra el resultado del etiquetado para una de las imágenes TMA originales.

Figura 3.2: TMA etiquetada, zonas: rojo-tumoral, verde-no tumoral, blanco-fondo.

Es importante mencionar que las imágenes etiquetadas fueron revisadas en busca de posibles fallos en el marcado asegurando de este modo la congruencia con sus respectivas zonas. Por último las imágenes etiquetadas fueron convertidas a 8 bits con la intención de que al revisar su histograma solo se encuentren los 3 valores correspondientes a las etiquetas.

3.2. El software

La implementación del método propuesto se llevó a cabo de la mano de dos grandes plataformas de procesamiento de imagen, estas son la distribucion FIJI de ImageJ empleada para el pre-procesamiento de la imágenes y KNIME para las tareas de extracción de características, clasificación y evaluación del modelo. Ambos programas son libres, multiplataforma, bastante conocidos y su uso está muy extendido en todo el mundo. Es precisamente esa la razón de su elección en un intento por facilitar la reproducibilidad del experimento ante la comunidad científica.

3.2.1. ImageJ

ImageJ es un programa de procesamiento de imagen digital de código abierto, además de multiplataforma al estar escrito en Java, fue desarrollado inicialmente el National Institut of Health (Caroline A. *et al.* [54]). Soporta la mayoría de tipos conocidos de imágenes por lo que sus usos son diversos aunque se enfoca principalmente en el análisis y procesamiento de imagen bio-médica. ImageJ es altamente extensible, con miles de complementos y scripts para realizar una amplia variedad de tareas. Su gran comunidad de usuarios se encarga de mantenerlo y actualizarlo. Otro aspecto que destaca es su integración con herramientas de procesamiento de imagen como Matlab, KNIME o ITK.

FIJI

Figura 3.3: Fiji Is Just ImageJ .

"Fiji es solo ImageJ"(figura 3.3) es una distribución de ImageJ que viene con muchos plúgins incluidos (que por otro lado algunos podrían instalarse manualmente en ImageJ) con el objetivo de facilitar las tareas de análisis y procesamiento de la imagen (Johannes *et al.* [55]). Fiji ha sido el programa que se ha empleado en los experimentos para la etapa de pre-procesamiento. Fiji es fácil de instalar y tiene una función de actualización automática, agrupa una gran cantidad de plugins y ofrece una documentación completa. Se puede encontrar mas información en la web imagej.net/Fiji.

3.2.2. KNIME Analytics Platform

El Konstanz Information Miner (KNIME) es un plataforma modular que permite un fácil montaje visual y ejecución interactiva de flujos de datos. Está diseñado como una herramienta de enseñanza, investigación y colaboración, que permite la integración sencilla

de nuevos algoritmos y herramientas, así como métodos de manipulación o visualización de datos en forma de nuevos módulos o nodos (Michael R. *et al.* [56]). Su interfaz de usuario es potente e intuitiva, permitiendo la exploración interactiva de resultados de análisis o modelos entrenados.

El flujo de trabajo (Workflow) como se presenta en la figura 3.4, se conforma básicamente de nodos (que representan las acciones) y flechas (representan el flujo de datos) que se combinan y ejecuta interactivamente, pudiéndose observar y controlar fácilmente los datos tratados. Uno de los aspectos más relevantes de KNIME es su alta capacidad de integración con la herramientas del arte como Weka, Matlab, ImageJ, R y Python. Las colecciones de nodos se conocen como extensiones, actualmente existen para diversos campos como procesamiento de imágenes, minería de datos, inteligencia de negocios, análisis financiero y computación química.

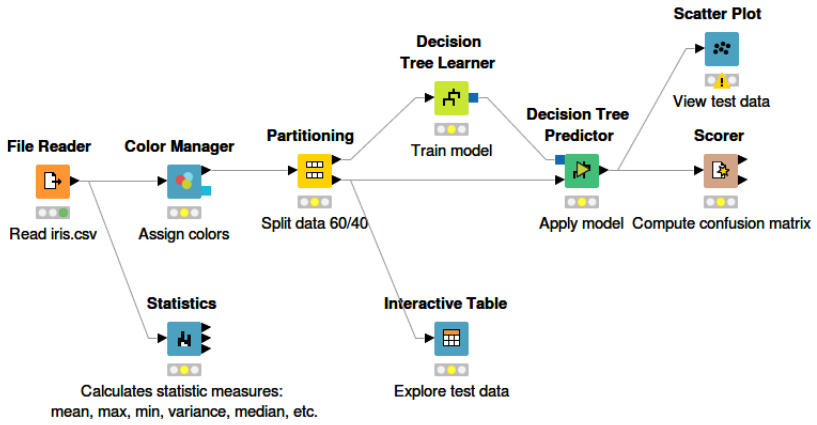


Figura 3.4: Workflow clasificación supervisada dataset Iris

KNIME fue desarrollado por el departamento de bioinformática y minería de datos de la Universidad de Constanza, Alemania, y actualmente está a cargo de la empresa KNIME.com GmbH radicada en Zurich, Zuiza. Esta herramienta puede ser descargada y utilizada gratuitamente bajo los términos de la licencia GPLv3.

Extensión de procesamiento de imágenes de KNIME

KNIME cuenta con un complemento de procesamiento de imagen que contiene alrededor de 100 nodos para tratar diversos tipos de imágenes (2D y 3D) y videos, además de aplicar métodos comunes como pre-procesamiento, segmentación, extracción de características, seguimiento y clasificación. Actualmente esta extensión se puede utilizar con diversas herramientas de procesamiento de imágenes como BioFormats, SCFIO, ImageJ, Omero y SciJava.

Capítulo 4

Método Propuesto

Para facilitar la explicación primero se deben tener en cuenta tres aspectos sobre el método propuesto:

- Primero, se pasa de un plano tradicional de procesamiento de pixeles a uno de superpixeles; estos son desplegados como una estrategia de captura de redundancia de la imagen que consiste en agrupar píxeles (vecinos) con características comunes (brillo, color, textura) facilitando el posterior procesamiento.
- Segundo, el método aborda el problema de segmentación de la imagen como un problema de clasificación supervisada de superpixeles.
- Tercero, el método se construyó de forma dinámica mediante la comparativa entre 3 técnicas de mejoramiento de la imagen y entre 5 algoritmos de clasificación supervisada, donde finalmente se eligió la mejor técnica y el mejor algoritmo para configurar el método final de segmentación.

En la figura 4.1, se presenta el esquema del método propuesto. Este se compone de 3 fases. en la primera (pre-procesamiento) se toman las imágenes TMA originales y se aplican 3 diferentes técnicas de mejoramiento de la imagen (normalización, ecualización del histograma y histogram matching). Luego a partir de las imágenes mejoradas, se aplica el Simple Linear Iterative Clustering (SLIC) para obtener las imágenes de superpixeles. En la segunda fase (Extracción de características) se extrajeron 69 características de textura más 1 variable clase que se obtuvo empleando las imágenes etiquetadas (marcadas manualmente). Cuando el dataset estuvo listo, se pasó a la tercera fase (aprendizaje de máquina) en dónde se clasificó comparando 5 algoritmos diferentes. Los superpixeles clasificados fueron reagrupados obteniendo finalmente las imágenes segmentadas.

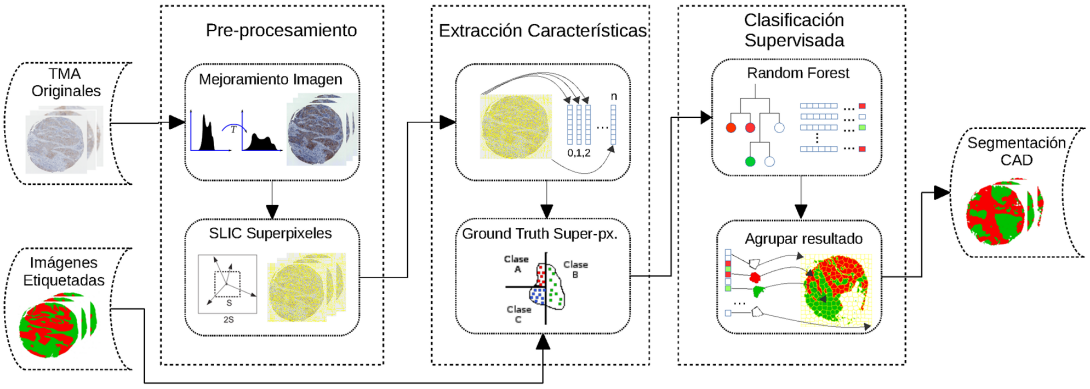


Figura 4.1: Esquema del método propuesto

4.1. Pre-procesamiento

Las imágenes TMA inicialmente eran muy grandes y se comprobó que tenían una carga computacional muy alta sin aportar demasiados beneficios frente a una posible reducción de su dimensión. Entonces evaluando los recursos de hardware disponibles, se decidió aplicarles un escalado al 25 % quedando cada imagen con una resolución de 1500 x 1500 pixeles aproximadamente.

Los métodos de mejoramiento de la imagen se han convertido en un factor indispensable dentro de la imagen médica [57,58]. Teniendo en cuenta la naturaleza difusa de las regiones de tejido en las imágenes histopatológicas, así como también la varianza de los parámetros de adquisición, se compararon 3 técnicas de mejoramiento de la imagen: normalización, ecualización del histograma y histogram matching. Finalmente se seleccionó la mejor para la fase de segmentación, dichas técnicas se aplicaron por separado produciendo 3 nuevos conjuntos de imágenes TMA, 1 por cada transformación.

4.1.1. Normalización de la imagen

La normalización de una imagen cambia el rango de valores de intensidad de los pixeles con el objetivo de obtener mayor consistencia dentro del conjunto de datos. En este caso se normalizó las imágenes TMA aplicando el método de mejoramiento espacial del contraste propuesto en (Jong-Sen [59]) estableciendo valores para su media y desviación estándar de modo que estos valores fueran lo más similar posible entre imágenes. Esta normalización requiere establecer los valores de media y varianza tal que:

$$x'_{i,j} = m_d + \sqrt{\frac{v_d}{v_{i,j}}}(x_{i,j} - m_{i,j}) \tag{4.1}$$

Donde en (4.1), $m_{i,j}$ y $v_{i,j}$ son la media local y la varianza, el principal inconveniente de esta técnica es que tiende a mejorar detalles sutiles a expensas de las características principales que se pierden en el proceso. Los parámetros de normalización que se usaron fueron *media=198* y *varianza=35* para cada canal RGB.

4.1.2. Ecualización del histograma

La ecualización del histograma es una técnica de transformación no lineal que modifica el valor de intensidad de los pixeles distribuyendolo a lo largo de todo el espectro, lo cual produce un histograma más constante.

El histograma es la distribución de cada nivel de intensidad dentro de la imagen. Este brinda un estimado de la probabilidad de ocurrencia de cada nivel de gris (r):

$$p(r_k) = \frac{n_k}{n} \tag{4.2}$$

Donde $p(r_k)$ es la probabilidad del nivel k, n_k es el número de pixeles que toma este valor y n es el número total de pixeles en la imagen. El histograma presenta una descripción global de la imagen y sobre todo da una indicación del contraste en la imagen. De aquí que si se modifica el histograma, se puede controlar el contraste en la imagen. Primero se asume que el nivel de gris de la imagen, r, es una función continua y normalizada (entre 0 y 1). Se desea realizar una transformación de forma que a cada nivel de gris r corresponda un nuevo nivel s:

$$s = T(r) \tag{4.3}$$

Esta transformación debe satisfacer lo siguiente (ver figura 4.2):

- T es una función monotónicamente creciente (mantener el orden).
- $0 \leq \mathrm{T} \leq 1$ (mantener el rango).

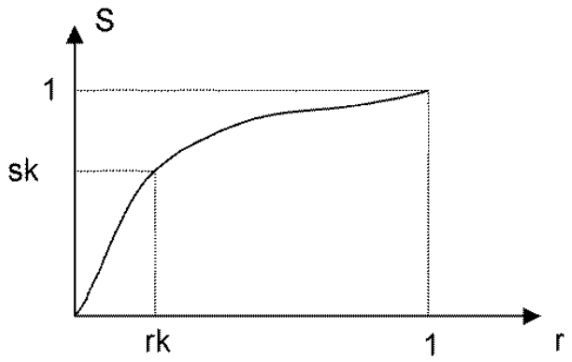


Figura 4.2: Función de transformación.

Se puede considerar las distribuciones de $p(r)$ y $p(s)$ como densidades de probabilidad. Entonces según la teoría de probabilidad:

$$p(s) = [p(r)\frac{dr}{ds}] \tag{4.4}$$

Si se usa como función de transformación la distribución acumulativa de r:

$$s = T(r) = \int p(r)dr \tag{4.5}$$

Entonces, derivando s respecto a r en la ecuación (4.5), se obtiene:

$$\frac{ds}{dr} = p(r) \tag{4.6}$$

Y, sustituyendo (4.6) en la ecuación (4.4), finalmente se llega a que $p(s) = 1$. En el caso discreto, la transformación se convierte en:

$$s(k) = T(r) = \sum_{i=0}^{k} \frac{n_i}{n} \tag{4.7}$$

Para $k = 0, 1, ..., N$, donde N es el número de niveles. Esto considera que ambos r y s están normalizados entre 0 y 1. Para poner la imagen de salida en otro rango hay que multiplicar por una constante (p. ej., 255). Para mas información véase [60].

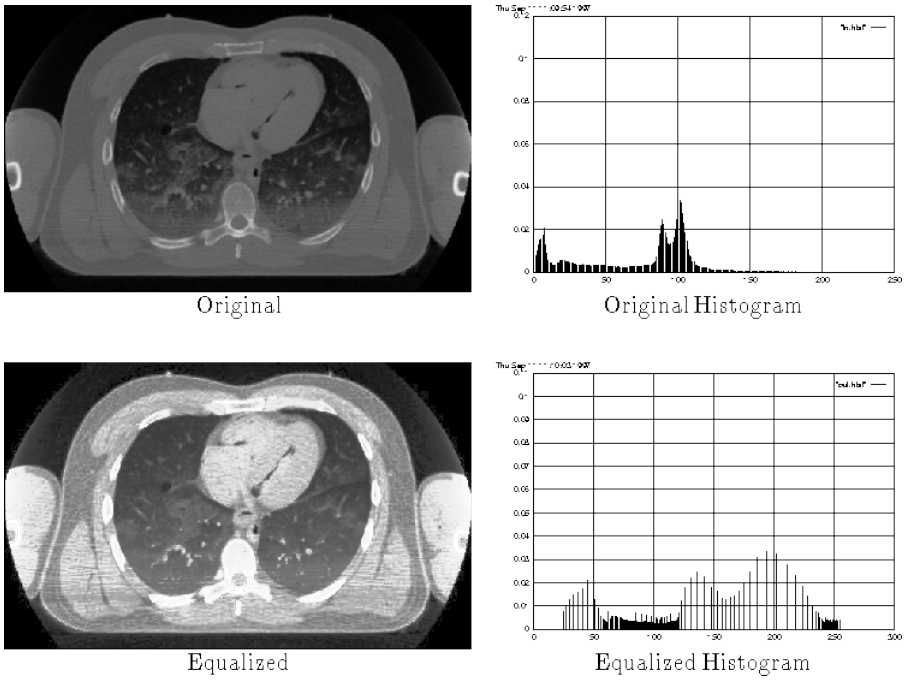


Figura 4.3: Caso de aplicación de la ecualización del histograma.

En la figura 4.3 se presenta un ejemplo de aplicación de cómo la ecualización del histograma puede mejorar el contraste. La parte superior muestra la imagen CT original y su histograma; Abajo aparece la imagen transformada y su histograma ecualizado. La transformación permite realizar una mejor interpretación de la imagen médica realzando aspectos antes más difusos.

4.1.3. Histogram matching

El histogram matching (HM) o correspondencia de histograma en español, es la transformación de una imagen para que su histograma coincida con un histograma especificado [61]. Histogram matching es una operación puntual que transforma una imagen de entrada para que su histograma coincida con una determinada forma definida por una función matemática o el histograma de otra imagen. Es particularmente útil para comparación de imágenes y diferenciación. Si las dos imágenes en cuestión se modifican para que

tengan histogramas similares, la comparación será razonable (Jiang *et al.* [62]).

HM se puede implementar aplicando ecualización del histograma (EH) dos veces. La ecuación (4.7) implica que un histograma ecualizado sólo se decide por el tamaño N de imagen y el rango DN de salida. Las imágenes del mismo tamaño siempre tienen el mismo histograma igualado para una rango de salida fija y, por tanto, EH puede actuar como puente para vincular imágenes del mismo tamaño pero con diferentes histogramas (Figura 4.4). Considere $h_i(x)$ como el histograma de una imagen de entrada y $h_o(y)$ como el histograma de referencia a ser emparejado. Suponiendo que $z = f(x)$ es la función EH para transformar $h_i(x)$ en un histograma ecualizado $h_e(z)$, y $z = g(y)$ la función EH para transformar el histograma de referencia $h_o(y)$ para el mismo histograma ecualizado $h_e(z)$, Entonces:

$$z = g(x) = f(x) \tag{4.8}$$

De este modo:

$$y = g^{-1}(z) = g^{-1}\{f(x)\} \tag{4.9}$$

Trayendo a memoria la fórmula (4.5), se establece del mismo modo a $f(x)$ y $g(y)$ como funciones de distribución acumulativa de $h_i(x)$ y $h_o(y)$ respectivamente. Por tanto, HM puede implementarse fácilmente mediante un LUT de tres columnas que contiene los DN niveles correspondientes de x, z e y. Un nivel DN de entrada x se transformará en un nivel DN de salida y que comparte el mismo valor z.

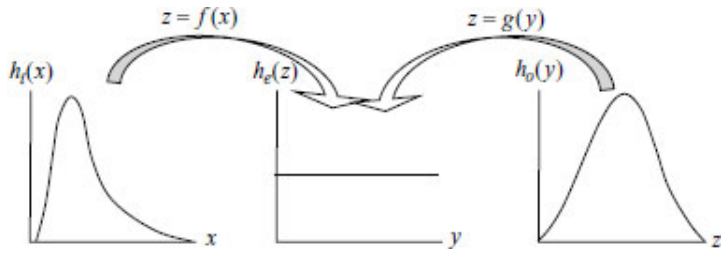


Figura 4.4: Ecualización del histograma como puente para la coincidencia de histogramas.

Como se muestra en el cuadro 4.1, para $x = 5$, $z = 3$, mientras que para $y = 0$, $z = 3$. Así, para una entrada $x = 5$, la LUT se cubre a una salida $y = 0$ y así sucesivamente. La imagen de salida Y tendrá un histograma que coincida con el histograma de referencia $h_o(y)$.

Cuadro 4.1: Un LUT de ejemplo para histogram matching

x	**y**	**z**
5	3	0
6	4	2
7	5	4
8	6	5
...	...	...

En la figura 4.5, se observa como dada una imagen de referencia con buena calidad y

una imagen de entrada de mediana calidad, tras aplicar histogram matching, la imagen resultante mejora notoriamente con respecto a la de entrada, obteniendo una calidad similar a la imagen de referencia.

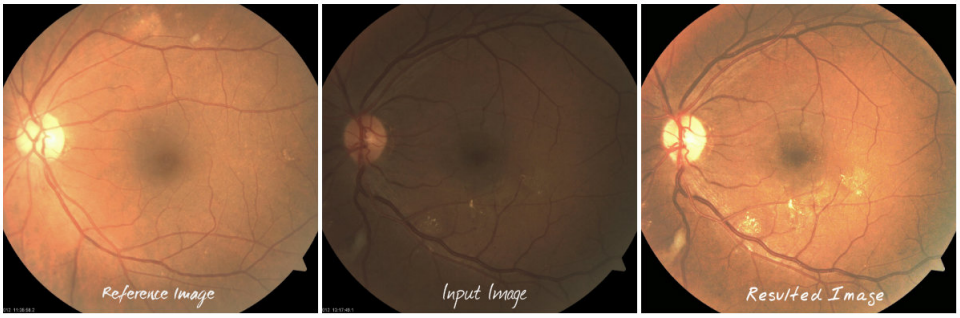


Figura 4.5: Mejoramiento de una imagen de retina con histogram matching

Esta técnica logra que las imágenes terminen siendo más uniformes entre sí, al modificar su histograma para parecerse al especificado, de manera que con ello se superan los problemas de diferencias de iluminación. Debido a que solo se transforman 2 imágenes a la vez, se tomó una de las imágenes TMA (la que visualmente presentaba la mejor distribución de intensidades) y se aplicó esta técnica a las 9 restantes.

4.2. Construcción de SLIC superpixeles

Los superpixeles son agrupaciones de pixeles que reúnen características similares como color, brillo o textura. Estos capturan la redundancia de la imagen, proporcionando una primitiva conveniente desde la cual calcular las características de la imagen además de reducir considerablemente la complejidad de las tareas posteriores de procesamiento de imagen. Su eficacia sobre imágenes histopatológicas ha sido comprobada en (Borovec, 2013 [63]). En la figura 4.6 se presenta una imagen de tejido de mama segmentada vía superpixeles.

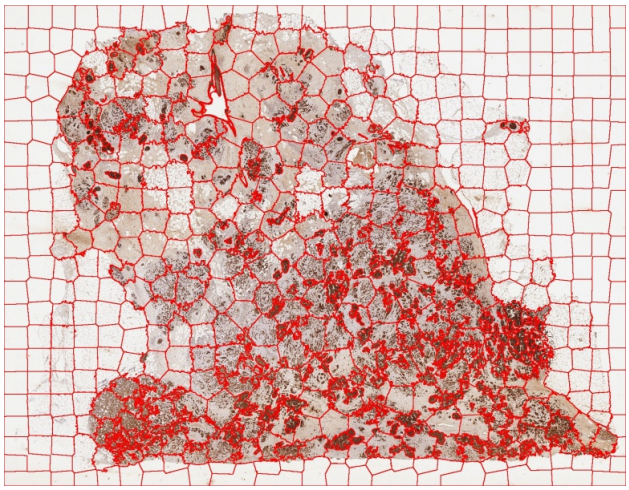


Figura 4.6: Diapositiva de mama teñida por citoqueratina, segmentación por superpixeles

Existen diversos métodos de construcción de superpixeles pero en este estudio se optó

por el Simple Linear Iterative Clustering (SLIC) introducido en (Achanta *et al.*, 2010 [64]) gracias al rendimiento superior mostrado en comparación con otros métodos de referencia (Achanta *et al.*, 2012 [65]).

4.2.1. Segmentación por superpixeles

El SLIC superpixeles agrupa los pixeles en función de su similitud de color y proximidad en el plano de la imagen. Esto se hace en el espacio $[labxy]$ de cinco dimensiones, donde lab es el vector de color de pixeles en el espacio de color CIELAB, que es considerado como perceptualmente uniforme para distancias de color pequeñas, y xy es la posición del pixel. Mientras que la distancia máxima posible entre dos colores en el espacio CIELAB (suponiendo imágenes de entrada sRGB) es limitada, la distancia espacial en el plano xy depende del tamaño de la imagen. No es posible utilizar simplemente la distancia euclidiana en este espacio 5D sin normalizar las distancias espaciales.

Con el fin de agrupar los pixeles en el espacio 5D, se introduce una nueva medida de distancia que considera el tamaño del superpixel. Usándola, se refuerza la similitud de color así como la proximidad de pixeles en el espacio 5D de manera que los tamaños de clúster esperados y su extensión espacial sean aproximadamente iguales.

4.2.2. Medida de distancia

El algoritmo toma como entrada un número deseado para generar superpixeles de igual tamaño K. Para una imagen con N pixeles, el tamaño aproximado de cada superpixel es N/K pixeles. Para superpixeles de igual tamaño hay un centro de superpixel en cada intervalo de cuadrícula $S = \sqrt{N/K}$.

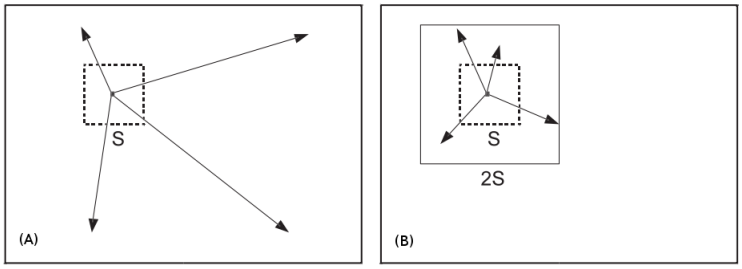


Figura 4.7: (A) En el algoritmo k-means convencional, las distancias se calculan desde cada centro de agrupación a cada píxel de la imagen. (B) SLIC sólo calcula las distancias desde cada centro del clúster a píxeles dentro de una región 2S x 2S. El tamaño de superpixel esperado es sólo SxS, indicado por el cuadrado más pequeño. Este enfoque no sólo reduce los cálculos de distancia, sino que también hace que la complejidad de SLIC sea independiente del número de superpixeles.

En el inicio del algoritmo, se eligen K centros de los clústers de superpixeles $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ con $k = [1, K]$ a intervalos regulares de la cuadrícula S. Dado que la extensión espacial de cualquier superpixel es de aproximadamente S^2 (el área aproximada de un superpixel), se puede asumir con seguridad que los pixeles que están asociados con este centro del clúster, se encuentran dentro de un área $2S \times 2S$ alrededor del centro del

superpixel en el plano xy. Esto se convierte en el área de búsqueda de los pixeles más cercanos a cada centro de clúster.

Las distancias euclídeas en el espacio de color CIELAB son perceptualmente significativas para pequeñas distancias. Si las distancias espaciales de pixeles exceden este límite de distancia de color, entonces empiezan a sobrepasar las similitudes de color de pixeles (resultando en superpixeles que no respetan los límites de la región, sólo la proximidad en el plano de la imagen), Por lo tanto, en lugar de usar una simple norma euclídea en el espacio 5D, se usa una medida de distancia D_s definida como sigue:

$$d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2}$$

$$d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}$$

$$D_s = d_{lab} + \frac{m}{S} d_{xy} \tag{4.10}$$

donde D_s es la suma de la distancia lab y la distancia normalizada del plano xy por el intervalo de cuadrícula S. La variable m se introdujo en D_s , para controlar la compacidad (grado de deformación) de un superpixel. Cuanto mayor sea el valor de m, mayor será la proximidad espacial y más compacto el clúster.

4.2.3. Algoritmo

El simple linear iterative clustering se resume en el Algoritmo 1. Se empieza por muestrear K centros de agrupación regularmente espaciados y moviéndose las semillas que corresponden a la posición del gradiente más bajo en una vecindad 3×3. Esto se hace para evitar colocarlos en un borde y para reducir las posibilidades de elegir un pixel ruidoso. Los gradientes de la imagen se calculan como:

$$G(x, y) = ||I(x + 1, y) - I(x - 1, y)||^2 + ||I(x, y + 1) - I(x, y - 1)||^2 \tag{4.11}$$

Donde $I(x, y)$ es el vector lab correspondiente al pixel en la posición (x, y), y $||.||$ es la norma L_2. Esto toma en cuenta tanto el color como la información de intensidad. Cada pixel de la imagen está asociado con el centro de clúster más cercano, el área de búsqueda se superpone a este pixel. Después de que todos los pixeles están asociados con el centro de clúster más cercano, se calcula uno nuevo como el promedio del vector $labxy$ medio de todos los pixeles pertenecientes al clúster. Se repite iterativamente el proceso de asociar todos pixeles con el centro de clúster más cercano y re-calcular el centro del clúster hasta la convergencia. Al final de este proceso, pueden quedar algunas etiquetas dispersas, es decir, unos pocos pixeles en la proximidad de un segmento mayor que tiene la misma etiqueta pero no está conectado a ella. Aunque es raro, esto puede surgir a pesar de la medida de proximidad espacial ya que este clustering no hace cumplir explícitamente la conectividad. Sin embargo, se refuerza la conectividad en el último paso del algoritmo re-etiquetando segmentos disjuntos con las etiquetas del clúster vecino más grande. Este paso es complejo $O(N)$ y toma menos del 10 % del tiempo total necesario para segmentar una imagen.

Algoritmo 1 Segmentación eficiente de superpixeles SLIC

Entrada: Centros de clusters $C_k = [l_k, a_k, b_k, x_k, y_k]^T$

1: Inicializar centros clusters C_k mediante pasos regulares de la cuadricula S
2: Perturbar C_k en un vecindario $n \times n$, hasta la posición de gradiente más bajo
3: **Mientras** $E \leq umbral$ **Hacer**
4: **Para** cada centro de cluster C_k **Hacer**
5: Asignar los mejores pixeles coincidentes de un cuadrado cuadrado $2S \times 2S$ alrededor C_k de acuerdo con la medida de distancia D_s (Eq. 4.7)
6: **Fin Para**
7: Calcular nuevos C_k
8: Calcular error residual E $\{L1$ distancia entre centros anteriores y centros recalculados$\}$
9: **Fin Mientras**
10: Hacer cumplir la conectividad

4.2.4. Experimentación

El método SLIC fue aplicado sobre las imágenes TMA originales (sin transformar), las normalizadas y las de ecualización del histograma, utilizando ImageJ con el plugin jSLIC (Borovec *et al.*, 2014 [66]) qué corresponde a una variación de SLIC computacionalmente más eficiente pero que conserva el método de original. Para empezar jSLIC se rige básicamente por dos parámetros, estos son:

- Init. grid size: el tamaño inicial promedio de superpixeles (30 por defecto)
- Regularisation: grado de deformación de los superpixeles estimados. El rango es de 0 para superpixeles muy elásticos a 1 para superpixeles casi cuadrados (0.20 por defecto)

Los parámetros por defecto se sugieren por los autores del método aunque estos varían de acuerdo a la necesidad de cada aplicación. En la experimentación se probó con distintos valores y combinaciones, teniendo en cuenta que los superpixeles resultantes deberían presentar un tamaño adecuado a las estructuras histopatológicas además de una apropiada segmentación que diferenciara correctamente las 3 zonas. Finalmente se optó por establecer los valores por defecto ya que se comprobó que produjeron superpixeles que cumplieron con las condiciones mencionadas. Aunque es beneficioso que los superpixeles se hayan ajustado distinguiendo anticipadamente las distintas zonas, es necesario aclarar que son los algoritmos de clasificación supervisada quienes realizan la clasificación y segmentación final, de modo que los superpixeles se corresponden a una pre-segmentación como estrategia para optimizar la fase de clasificación.

La aplicación de jSLIC al conjunto de imágenes TMA, produjo como resultado imágenes de superpixeles que contienen por cada superpixel una etiqueta. A modo ilustrativo se presenta una de estas imágenes en la figura 4.8. El propósito de ellas es señalar el área de cada superpixel la cual fue necesaria para la siguiente la fase de extracción de características.

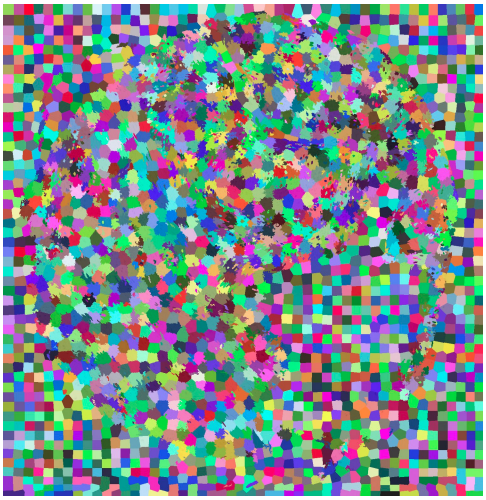


Figura 4.8: Representación de la imagen de superpixeles

4.3. Extracción de características

Esta etapa consistió en obtener una representación de la imagen (en forma de datos) que ayudara a discriminar entre las distintas clases (tumor, no-tumor y fondo). Según Rafabel Llovet [67] el rendimiento de un sistema de CAD depende más de la extracción y selección de características que del método de clasificación.

4.3.1. Estadísticos de primer orden

Las medidas de textura de primer orden son estadísticos calculados a partir de los valores de imagen originales, como la varianza, y no consideran relaciones de vecindario de pixeles. El enfoque bas|ado en histograma para el análisis de textura se enfoca en las concentraciones de intensidad en la totalidad o parte de una imagen representada como histograma. Las características comunes incluyen momentos tales como media, varianza, dispersión, valor cuadrático medio o energía media, entropía, asimetría y curtosis. La variación del nivel de gris en una región cercana a un píxel es una medida de la textura (Fumiaki *et al.* [68]).

Al establecer $\phi(i)$ $(i = 1, ..., n)$ como el número de puntos cuya intensidad es i en la imagen y A el área de la imagen (el número total de puntos en la imagen), la probabilidad de ocurrencia de la intensidad i en la imagen es calculada por:

$$h(i) = \phi(i)/A \tag{4.12}$$

Los siguientes estadísticos (simples) se utilizan con mayor frecuencia para caracterizar el histograma. Algunos son:

- **Media**

$$\sum_{i=1}^{n} ih(i) \tag{4.13}$$

- **Desviación estándar**

$$\sum_{i=1}^{n}(i-\mu)^2 h(i) \tag{4.14}$$

- **Tercer momento**

$$\sum_{i=1}^{n}(i-\mu)^3 h(i) \tag{4.15}$$

- **Entropía**

$$-\sum_{i=1}^{n} h(i)\log h(i) \tag{4.16}$$

Donde μ es la media de las intensidades.

4.3.2. Características Tamura

En Tamura *et al.* [69], se introducen seis características de textura que aprovechan a la percepción visual humana: rigurosidad, aspereza, contraste, direccionalidad, linealidad y regularidad. La efectividad de estas características visuales ha sido demostrada en Francesco *et al.* [70], ayudando a los clasificadores a alcanzar tasas de acierto de hasta 97 % sobre imágenes histopatológicas. Las características de contraste y direccionalidad fueron seleccionadas al ser consideradas "mas relevantes"para obtener una mejor representación de las imágenes TMA.

- **Contraste:** El contraste de la imagen está influenciado por el rango dinámico de los niveles de grises en la imagen, la polarización de la distribución del negro y el blanco, la nitidez de los bordes y el período de los patrones repetitivos. También podría representar la calidad de la imagen en el sentido más estrecho. Se puede calcular el contraste $Icon$ como sigue:

$$I_{con} = \frac{\sigma}{\sqrt[4]{\alpha_4}} \tag{4.17}$$

  Aquí, σ y α_4 son un estimador de la desviación estándar y la cuarta raíz que fue sugerida en Tamura *et al.* [69] basada en sus experimentos, respectivamente. α_4 se calcula como:

$$\alpha_4 = \frac{\mu_4}{\sigma^4} \tag{4.18}$$

- **Direccionalidad:** Este descriptor proporciona una visión de la propiedad de textura global sobre una región midiendo el grado total de dirección de textura (Maroua *et al.* [71]). Se calcula utilizando un histograma de probabilidades de los bordes locales H_D en contra de su ángulo direccional. Mediante la cuantificación de la nitidez de los picos H_D, la direccionalidad de la textura se mide sumando los segundos momentos alrededor de cada pico de acuerdo con la ecuación 4.19.

$$Fdir = 1 - rn_p \sum_{p}^{n_p} \sum \phi_h \epsilon w_p (\Phi_h - \Phi_p)^2 H_D(\Phi_h) \tag{4.19}$$

Donde, n_p, Φ_p, w_p, r y Φ_h denotan el numero de picos de histogramas fijado para 2, la posición p^{th} pico de H_D, el rango de pico p^{th} de valles, el factor de normalización relacionado con los niveles cuantificados de Φ_h y el código de dirección cuantificado (cíclicamente en módulo 2π), respectivamente.

4.3.3. Características extraídas

Para el problema abordado se determinó que las características de textura eran las más adecuadas; se seleccionaron 23 características de textura, divididas en 17 estadísticos de primer orden y 6 de tipo Tamura como sigue:

- **Estadísticos de primer orden**
  - Min
  - Max
  - Mean
  - Geometric Mean
  - Sum
  - Squares of Sum
  - Std Dev
  - Variance
  - Skewness
  - Kurtosis
  - Quantil 25
  - Quantil 50
  - Quantil 75
  - Median absolute deviation (MAD)
  - WeightedCentroid Dim 1
  - WeightedCentroid Dim 2
  - Mass Displacement
- **Tipo Tamura**
  - TamuraGranularity
  - TamuraContrast
  - TamuraKurtosisOfDirectionality
  - TamuraStdDevDirectionality
  - TamuraMaxDirectionality
  - TamuraSkewness

4.3.4. Descomposición del espacio RGB

Debido a que las imágenes TMA originales estaban en el espacio RGB, fue necesario descomponerlas en sus 3 canales R *(rojo)*, G *(verde)*, B *(azul)* como requisito para el cálculo de los diferentes tipos de características. Al final se obtuvieron un total de 69 características por cada superpixel. Para esto se empleó KNIME, con el nodo *Splitter* para separar la imagen en canales y el nodo *Image Segment Features* para extraer las características seleccionadas de cada canal.

4.3.5. Asignación de clases

En este paso se creó y asignó la variable clase (tumor, no-tumor o fondo) a cada vector de características (superpixel). Para ello se emplearon las imágenes de superpixeles, junto con las imágenes de etiquetas. El procedimiento consistió en tomar la imagen de superpixeles recorriendo cada superpixel e interponiéndolo sobre su correspondiente área dentro de la imagen de etiquetas. Esta, a su vez, tiene información de la clase a la que pertenece cada pixel (ver apartado 3.1.2), por lo cual mediante un sistema de votos se cuenta la cantidad de pixeles pertenecientes a cada clase y finalmente se asigna al superpixel la clase que más votos tenga. En otras palabras se asigna la clase de la región mayoritaria a la que pertenece el superpixel. En la figura 4.9, se representa de forma ideal un superpixel ubicado entre tres regiones, la clase asignada a dicho superpixel sería la C.

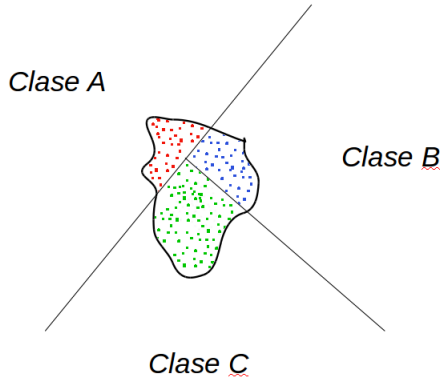


Figura 4.9: Asignación de clase a un superpixel mediante voto mayoritario de las etiquetas de pixel del ground truth.

Una vez se asignaron las clases a todos los superpixeles, estas fueron añadidas como una nueva variable al vector de características dejando listo el dataset para la fase de clasificación.

4.4. Segmentación

Esta etapa consistió en la construcción y comparación de 5 modelos de clasificación supervisada, de los cuales se eligió el mejor para luego emplearse en la segmentación de las

imágenes TMA. Primero se seleccionaron 5 algoritmos de entre los más relevantes según lo evidenciado en el estado del arte, después, una vez se entrenaron los modelos a modo de comparativa, se evaluó su rendimiento y finalmente se aplicó el mejor algoritmo en la segmentación.

4.4.1. Algoritmos de clasificación

Se seleccionaron los siguientes algoritmos de clasificación supervisada: Random Forest, Supourt Vector Machines (SMO), LogiBoost, J48 y BayesNet. Empleando el entorno KNIME que tiene integrada Weka versión 3.7.

Random Forest

Este algoritmo está entre los mejores y más usados, fue introducido por Breiman [72]. También conocido en castellano como "Bosques Aleatorios" es una combinación de árboles predictores tal que cada árbol depende de los valores de un vector aleatorio probado independientemente y con la misma distribución para cada uno de estos. Es una modificación sustancial de bagging que construye una larga colección de árboles no correlacionados y luego los promedia.

Supourt Vector Machines (SMO)

Sequential Minimal Optimization es un algoritmo que sobresale por su eficiencia con grandes conjuntos de datos. Implementa el algoritmo de optimización mínima secuencial de Platt [73] para el entrenamiento de un clasificador de soporte de vectores. Esta implementación reemplaza globalmente todos los valores faltantes y transforma los atributos nominales en binarios. También normaliza todos los atributos por defecto. (En ese caso, los coeficientes de la salida se basan en los datos normalizados, no en los datos originales - esto es importante para interpretar el clasificador).

LogitBoost

Es la implementación en Weka del algoritmo de regresión logística aditiva introducido en (Jerome *et al.* [74]). Este realiza la clasificación usando un esquema de regresión como el base learner, y puede manejar problemas de varias clases. Este algoritmo se seleccionó debido a su rendimiento en comparación a otros algoritmos tipo Boosting para este problema.

J48

Es la implementación de código abierto (Weka) del algoritmo C4.5 desarrollado por Quinlan [75] el cual genera un árbol de decisiones podado o no-podado. Estos árboles son referidos como un clasificador estadístico. C4.5 construye árboles de decisión desde un grupo de datos de entrenamiento, usando el concepto de entropía de información.

BayesNet

Es la implementación del clasificador de probabilidad Naive Bayes. Está estructurado como una combinación de un grafo acíclico dirigido de nodos y enlaces, y un conjunto de cuadros de probabilidad condicional (Remco R. [76]). Este clasificador demostró un mayor rendimiento frente a otros del tipo bayesiano.

4.4.2. Métricas de evaluación

Conforme al problema tratado, el enfoque utilizado para entrenamiento y test, varía un poco con respecto a al tradicional en aprendizaje supervisado, ya que aquí no solo se debe clasificar las filas del dataset (que en este caso corresponden a superpixeles) sino que además se necesita presentar la segmentación final de la imagen. Por tanto cada vez se probó el método con al menos 1 imagen TMA y se entrenó con el resto. Las métricas usadas en los estudios vistos en el arte son diversas. De ellas, han sido seleccionadas 2 de las más frecuentes. Antes de entrar a los detalles de dichas métricas, es importante comprender la matriz de confusión de una evaluación de clasificación binaria (figura 4.10). Las etiquetas de clase en el conjunto de entrenamiento pueden tomar solo dos valores posibles, a los que normalmente podemos referirnos como positivo o negativo. Las instancias positivas y negativas que un clasificador predice correctamente se denominan positivos verdaderos (TP) y negativos verdaderos (TN), respectivamente. De forma similar, las instancias clasificadas incorrectamente se denominan falsos positivos (FP) y falsos negativos (FN).

	Predicted	
	Positive	Negative
Actual True	**TP**	**FN**
Actual False	**FP**	**TN**

Figura 4.10: Matriz de confusión de la clasificación binaria.

A diferencia de la clasificación binaria, en la clasificación multiclase se realiza el cálculo de los positivos y negativos verdaderos, y de los falsos positivos y negativos con un recuento por clase, ya que no existe ninguna clase general positiva o negativa.

Medida F (F-measure o F1-score)

En una tarea de clasificación, la precisión (positive predictive value (PPV)) de una clase es el número de positivos verdaderos (es decir, el número de elementos correctamente etiquetados como pertenecientes a la clase positiva) dividido por el número total de elementos etiquetados como pertenecientes a la clase positiva (es decir, la suma de verdaderos positivos y falsos positivos, que son elementos incorrectamente etiquetados como pertenecientes a la clase).

$$PPV = \frac{TP}{TP + FP} \tag{4.20}$$

El recuerdo (recall, sensitividad, true positive rate (TPR)) se define como el número de verdaderos positivos divididos por el número total de elementos que realmente pertenecen a la clase positiva (es decir, la suma de verdaderos positivos y falsos negativos, que son casos que no fueron etiquetados como pertenecientes a la clase positiva pero deberían haberlo sido.

$$TPR = \frac{TP}{TP + FN} \tag{4.21}$$

En el análisis estadístico de la clasificación binaria, la medida F, puede interpretarse como un promedio ponderado de la precisión y el recuerdo, donde un puntaje F1 alcanza su mejor valor en 1 y el peor en 0. Esta medida es aproximadamente la media de los dos cuando están cerca, y es más generalmente la media armónica, que para el caso de dos números coincide con el cuadrado de la media geométrica dividido por la media aritmética.

$$F_1 = 2\frac{PPV.TPR}{PPV + TPR} = \frac{2TP}{2TP + FP + FN} \tag{4.22}$$

Hay varias razones por las que el F-score (ecuación 4.22) puede ser criticado en circunstancias particulares debido a su sesgo como métrica de evaluación, ya que no toma en cuenta el número de verdaderos negativos TN. Aun así es ampliamente utilizada para medir la eficacia de los modelos de clasificación.

Precisión (Accuracy)

La precisión también se utiliza como una medida estadística de lo bien que una prueba de clasificación binaria identifica o excluye correctamente una condición. Es decir, la precisión es la proporción de resultados verdaderos (tanto los verdaderos positivos como los negativos verdaderos) entre el número total de casos examinados.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.23}$$

La precisión da un panorama general (a nivel multi-clase) del rendimiento del modelo construido y, como se observa en la ecuación 4.23, la precisión sí tiene en cuenta el número de verdaderos negativos TN.

Validación cruzada

La validación cruzada (cross-validation) es una técnica utilizada para evaluar los resultados de un análisis estadístico y garantizar que son independientes de la partición entre datos de entrenamiento y prueba [77]. En la validación cruzada de K iteraciones los datos de muestra se dividen en K subconjuntos. Uno de los subconjuntos se utiliza como datos de prueba y el resto $(K-1)$ como datos de entrenamiento. El proceso de validación cruzada es repetido durante K iteraciones, con cada uno de los posibles subconjuntos de datos de prueba. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado.

Figura 4.11: Validación cruzada 10-fold

Entonces para evaluar y garantizar estadísticamente los resultados se empleó la técnica Cross Validation a 10-fold: en cada iteración el conjunto de entrenamiento lo conformaron 9 imágenes y el de test, 1 imagen, a modo ilustrativo ver la Figura 4.11. Esto finalmente facilita la comparación de los resultados obtenidos con otros del estado del arte ya que muchos de los métodos revisados en la sección 2.3 emplearon esta técnica para validar sus modelos.

4.4.3. Anotación Ground truth

El ground truth es el procedimiento que permite evaluar la precisión de la segmentación del sistema CAD al compararla con la segmentación manual hecha por los expertos. Según (Quian *et al.* [78]) cuando el ground truth (imágenes marcadas) está presente, la segmentación se puede medir en términos de precisión (tasa de acierto) y robustez. La precisión por su parte refleja la exactitud de la segmentación con respecto a las imágenes marcadas.

El ground truth se realizó empleando el mismo procedimiento de asignación de clases descrito en el apartado 4.3.5, esto teniendo en cuenta que el fin del procedimiento es el mismo: asignar clase al superpixel en función de su imagen marcada.

4.4.4. Imágenes segmentadas

Finalmente, a partir de la clasificación se construyeron imágenes segmentadas las cuales son similares a las imágenes de etiquetas pero con la diferencia de que fueron producidas por el sistema CAD. Debe recordarse que Inicialmente la imagen fue descompuesta en sus superpixeles, los cuales fueron clasificados. Por tanto, fue necesario mantener dentro del dataset de clasificación una columna con la información de la posición de cada superpixel en la imagen. Así después de la clasificación se usó la información de esa columna (que no fue tenida en cuenta por el clasificador) para colocar cada superpixel en su posición agrupándolos de acuerdo a su clase.

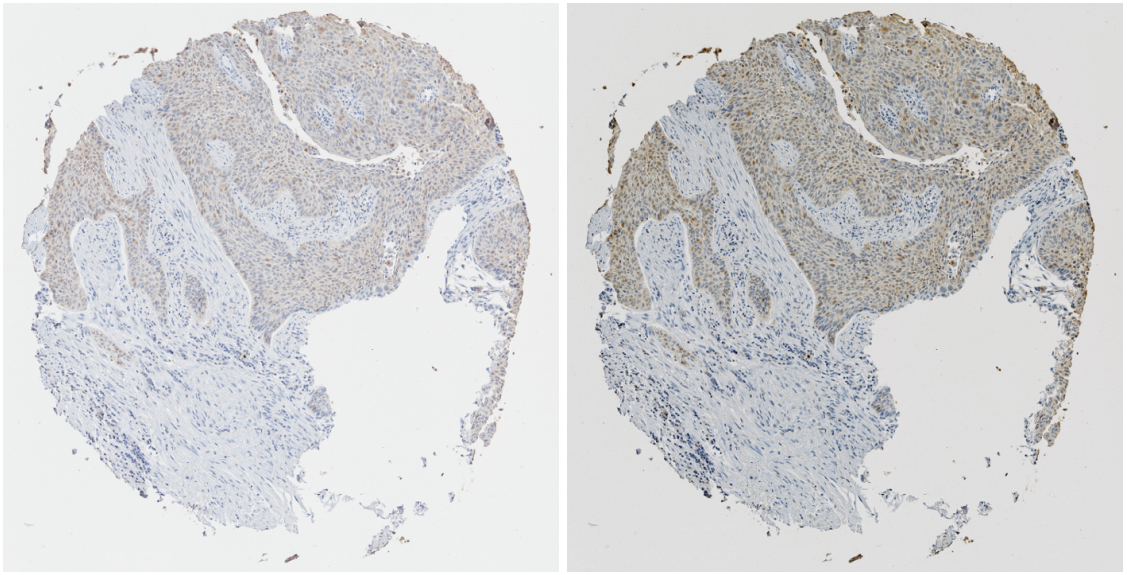
Capítulo 5

Resultados

5.1. Pre-procesamiento

Esta fase inició reduciendo la dimensión de las imágenes con un escalado al 25 %, después se aplicaron tres métodos de mejoramiento de la imagen: normalización, histogram matching y ecualización del histograma con lo cual se obtuvieron 3 datasets de imágenes TMA mejoradas; a modo de ejemplo se presenta solo una por cada transformación.

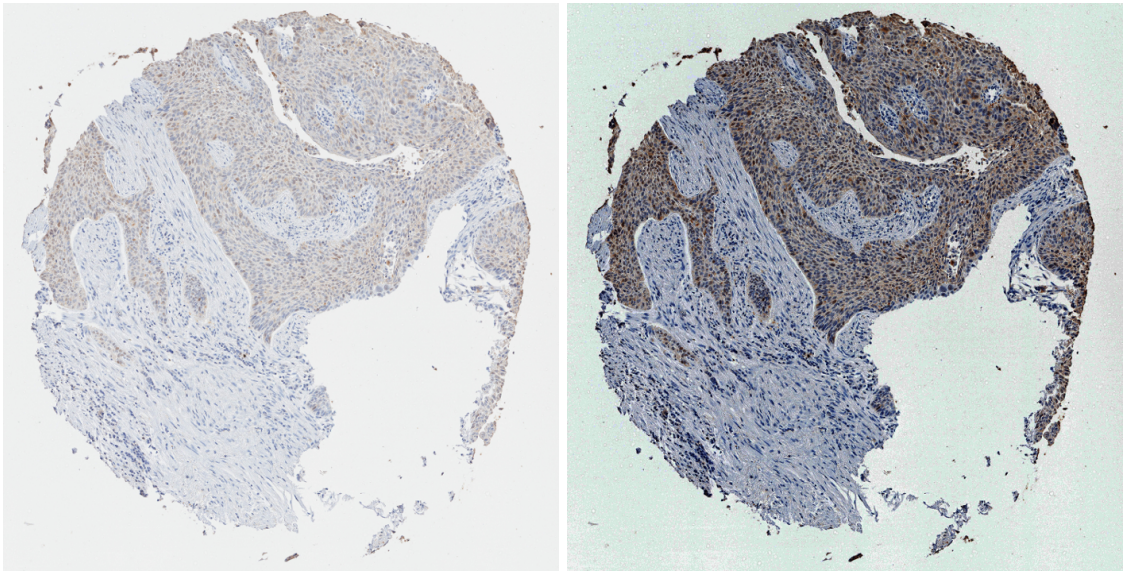
5.1.1. Normalización



(a) (b)

Figura 5.1: Transformada por normalización (b), a partir de la imagen original (a).

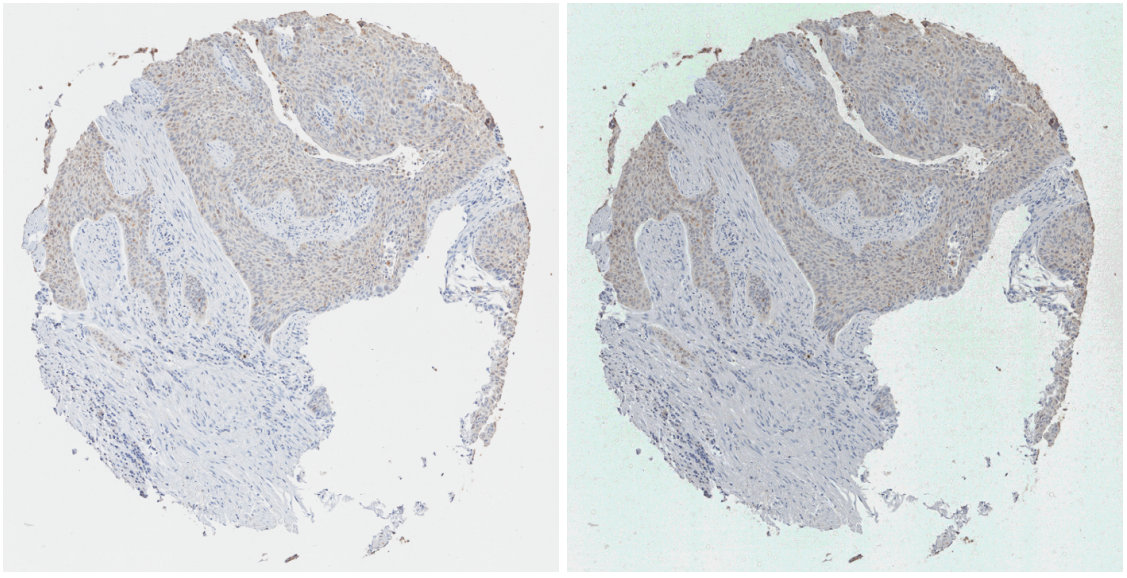
5.1.2. Ecualización del histograma



(a) (b)

Figura 5.2: Transformada por ecualización del histograma (b), a partir de la imagen original (a).

5.1.3. Histogram matching

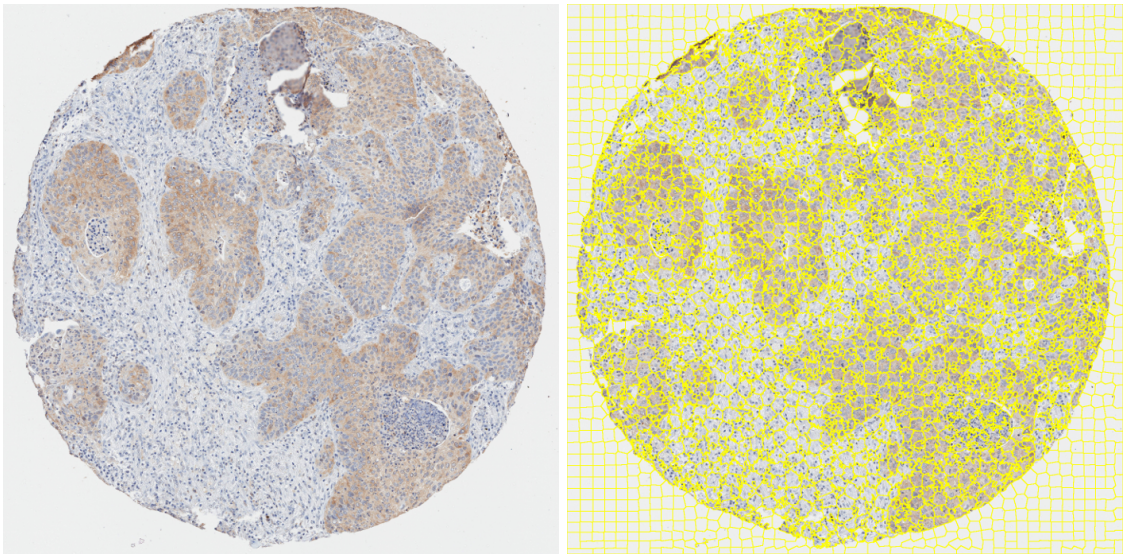


(a) (b)

Figura 5.3: Transformada por histogram matching (b), a partir de la imagen original (a).

5.2. SLIC superpixeles

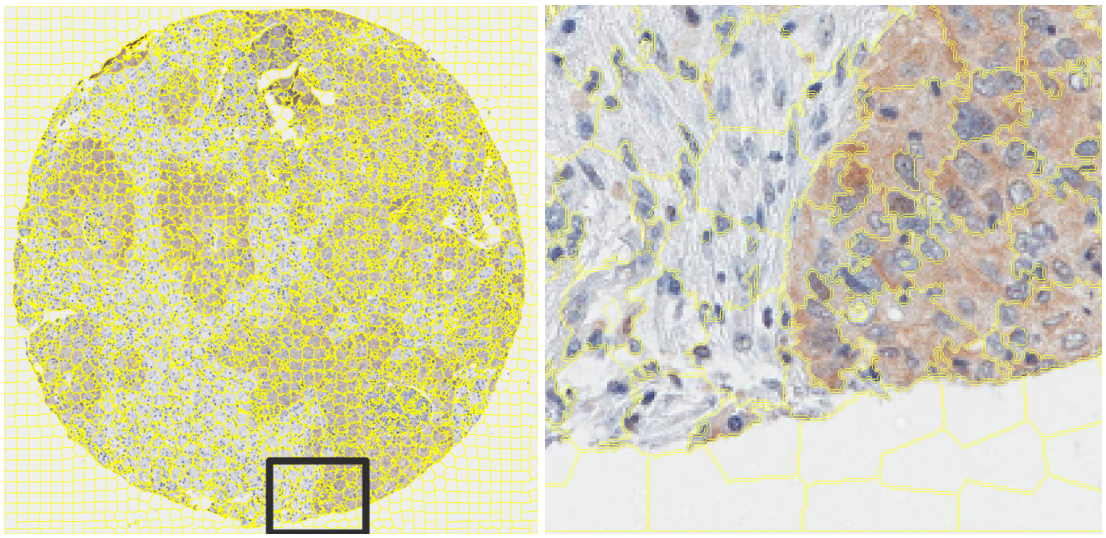
El método jSLIC fue aplicado sobre el conjunto de imágenes TMA iniciales, empleando los parámetros (sugeridos por sus autores), el resultado puede verse en la figura 5.4.



(a) (b)

Figura 5.4: Segmentación por superpixeles (b), a partir de la imagen original (a).

La imagen anterior se compone de 3567 superpixeles por lo que a esta escala es difícil detallar con claridad el ajuste de los superpixeles sobre las estructuras histopatológicas. Sin embargo a continuación (figura 5.5) se hace un acercamiento para visualizar este aspecto con más detalle.



(a) (b)

Figura 5.5: Segmentación de superpixeles (b), Versión recortada de (a).

En la figura 5.5 se aprecia el buen ajuste de los superpixeles a las 3 diferentes regiones, a la izquierda se encuentra tejido sano, a la derecha (color marrón) está el tejido tumoral y abajo (color gris claro) el fondo; las lineas amarillas indican los limites de los superpixeles y su ajuste a las estructuras.

5.3. Características extraídas

En esta fase inicialmente se seleccionaron 23 características de los tipos tamura y estadísticos de primer orden, sin embargo como la imagen estaba en formato RGB, para el cálculo de características tuvo que dividirse en sus 3 canales por tanto se obtuvieron 69 características por cada superpixel a los cuales se les asignó sus respectivas clases. Finalmente se obtuvo un dataset de 70 variables (69 predictoras y 1 clase) y con un total de 37279 casos (superpixeles). la distribución de las clases se muestra a continuación (cuadro 5.1).

Cuadro 5.1: Número de características por clase

Clase superpixeles	**Número**
Tumor	17468
No-tumor	11864
Fondo	7947
Total:	37279

5.4. Clasificación

La estrategia de experimentación en las fases previas de pre-procesamiento, construcción de superpixeles y extracción de características, consistió en probar diferentes configuraciones de parámetros de modo que se pudiese encontrar los más óptimos. En esta fase también se mantuvo dicha estrategia mediante la comparativa de 5 diferentes algoritmos del estado del arte. Estos fueron probados sobre los 3 conjuntos de imágenes TMA mejoradas.

Las métricas para la evaluación del rendimiento de los clasificadores son la precisión (accuracy) y la F-measure (F-score), además se realizó una validación cruzada 10-fold. Estas métricas son un indicador de la bondad del clasificador a nivel general en problemas multiclase, aunque en ocasiones se utiliza una o u otra, fue necesario para este estudio tomar ambas para facilitar la comparación con los métodos de referencia.

5.4.1. Resultados en imágenes normalizadas

La precisión obtenida por imagen (sub-sample) para cada algoritmo empleando como entrada las imágenes normalizadas se muestra en la figura 5.6. En este caso, el algoritmo que muestra el mejor resultado es Random Forest, seguido de SMO con un comportamiento similar. La imagen mejor predicha por todos los algoritmos fue la 9, lo cual indica que en esta iteración el entrenamiento obtuvo una buena capacidad explicativa a partir de las demás imágenes. Por otra parte, la imagen 10 es la que tiene la mayor varianza en su clasificación y aunque es una de las mejor clasificadas por Random Forest y SMO (acercándose al 90 % de acierto), es la peor clasificada por J48 y LogitBoost.

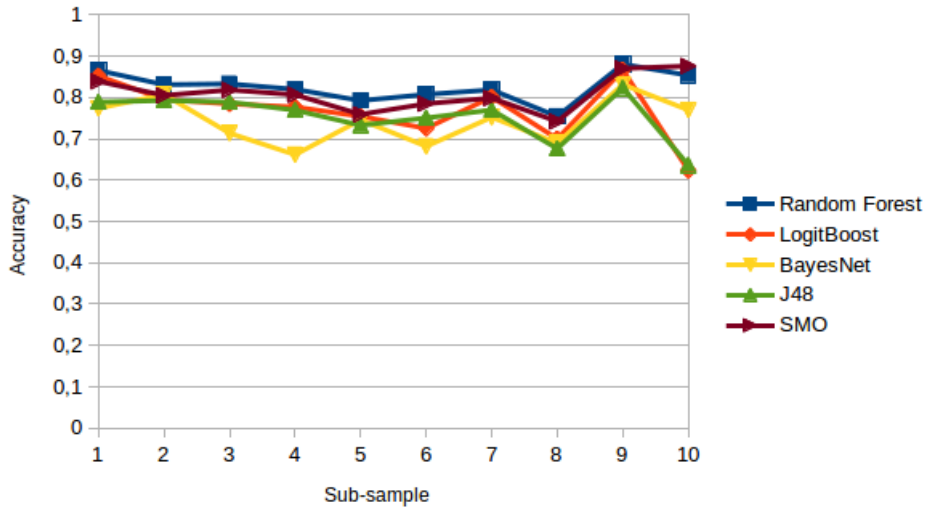


Figura 5.6: Rendimiento de los clasificadores por cada imagen durante validación cruzada a 10-fold.

El rendimiento medio por cada algoritmo se presenta en el cuadro 5.2. Las medidas de bondad indican que el algoritmo que obtuvo el mejor rendimiento es Random Forest con una precisión y medida F del 83 %, seguido de SMO, LogitBoost, J48 y BayesNet.

Cuadro 5.2: Comparación de métricas - rendimiento medio

	Accuracy	**F-measure**
Random Forest	**0,83**	**0,83**
LogitBoost	0,77	0,77
BayesNet	0,74	0,76
J48	0,75	0,76
SMO	0,81	0,82

5.4.2. Resultados en imágenes de ecualización del histograma

Para este dataset el algoritmo Random Forest presenta una precisión por imagen superior a los demás algoritmos (figura 5.7), incluso llegando a sobrepasar el 90 % (imagen 9). La imagen 8 es la peor clasificada por todos los algoritmos. Finalmente BayesNet es el algoritmo que muestra el rendimiento más bajo para la mayoría de imágenes.

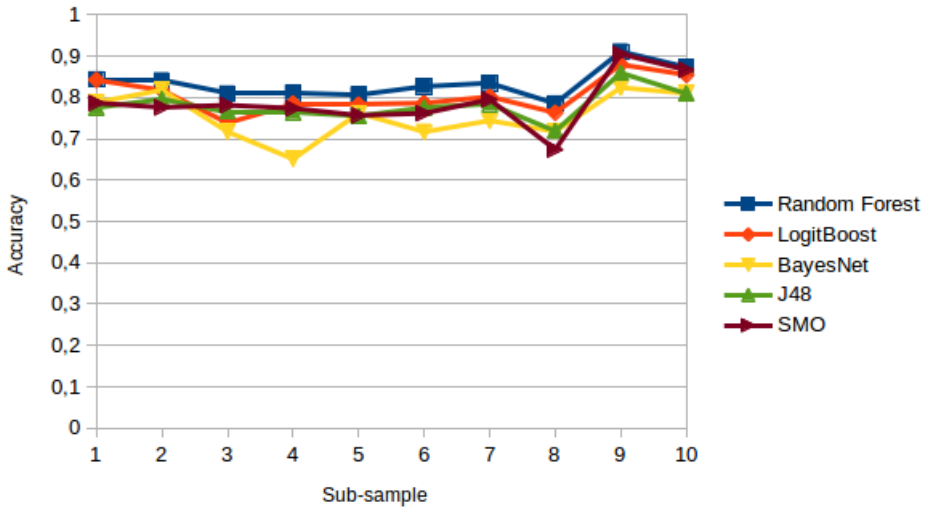


Figura 5.7: Rendimiento de los clasificadores por cada imagen durante validación cruzada a 10-fold.

En cuanto al rendimiento medio general (cuadro 5.3), destaca el algoritmo Random Forest con una precisión del 83 % y una F-measure de 84 %. En este caso el segundo lugar lo obtuvo LogitBoost y después SMO, J48 y BayesNet.

Cuadro 5.3: Comparación de métricas (rendimiento medio)

	Accuracy	**F-measure**
Random Forest	**0,83**	**0,84**
LogitBoost	0,81	0,82
BayesNet	0,75	0,77
J48	0,78	0,79
SMO	0,79	0,80

5.4.3. Resultados en imágenes de histogram matching

Para el dataset de histogram matching, en la figura 5.8, se puede notar el rendimiento superior de Random Forest en la mayoría de imágenes a excepción de la imagen 2 en donde BayesNet obtiene una precisión mayor; la imagen mejor clasificada (por la mayoría de algoritmos) es la 9 con una precisión del 90 % obtenida por SMO junto a Random Forest y la imagen 8 es la peor clasificada.

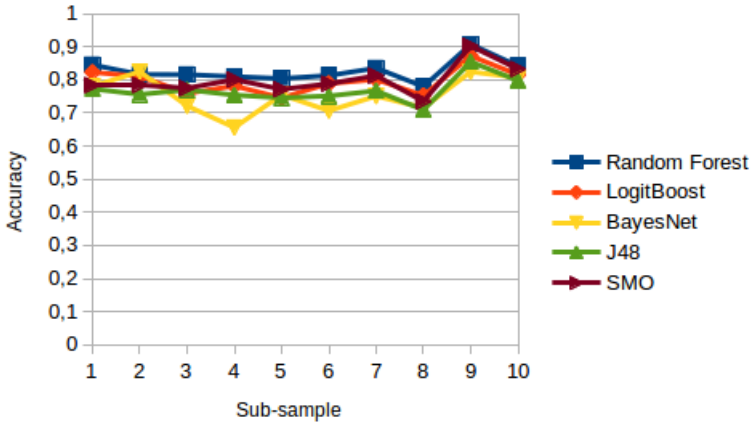


Figura 5.8: Rendimiento de los clasificadores por cada imagen durante validación cruzada a 10-fold.

Las medidas de bondad (cuadro 5.4) indican que Random Forest es el algoritmo con el rendimiento medio más alto para este dataset, con una precisión y medida F del 83 % supera ampliamente a SMO y LogitBoost (mismo resultado) y finalmente quedan J48 y BayesNet.

Cuadro 5.4: Comparación de métricas (rendimiento medio)

	Accuracy	**F-measure**
Random Forest	**0,83**	**0,83**
LogitBoost	0,80	0,81
BayesNet	0,75	0,77
J48	0,77	0,78
SMO	0,80	0,81

5.4.4. Análisis comparativo

Las mejores medidas de bondad (resaltadas en las cuadros 5.2, 5.3 y 5.4) indicaron que el algoritmo que presentó el mayor rendimiento medio durante las pruebas fue Random Forest. Por tanto fue el algoritmo seleccionado para configurar el método final de segmentación.

Cuadro 5.5: Rendimiento medio de Random Forest por dataset de imágenes mejoradas.

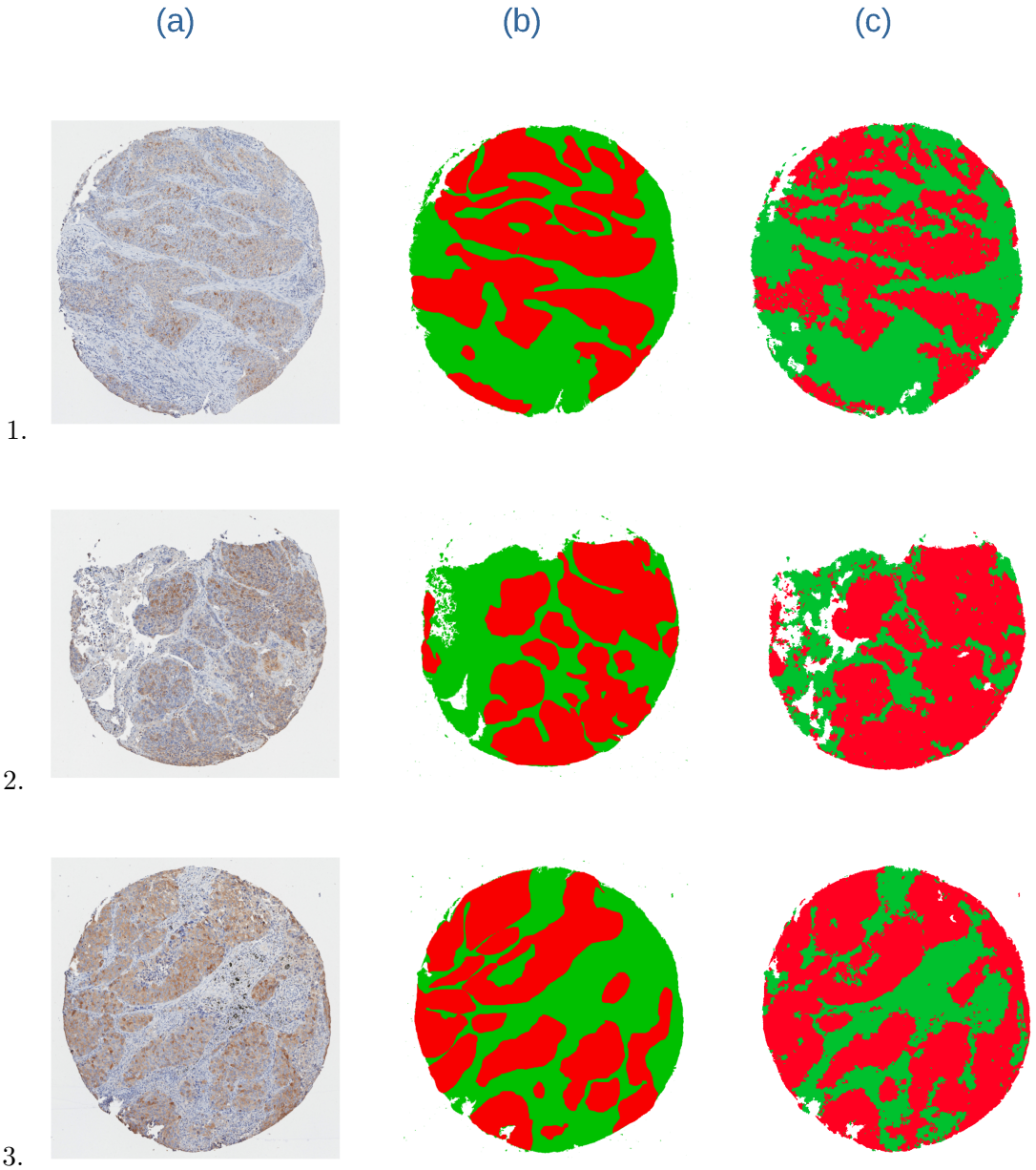
Dataset	**Accuracy (%)**	**F-measure (%)**
Norm.	82.54	83.48
Equal. Hist.	**83.41**	**84.39**
Hist. Match.	82.69	83.48

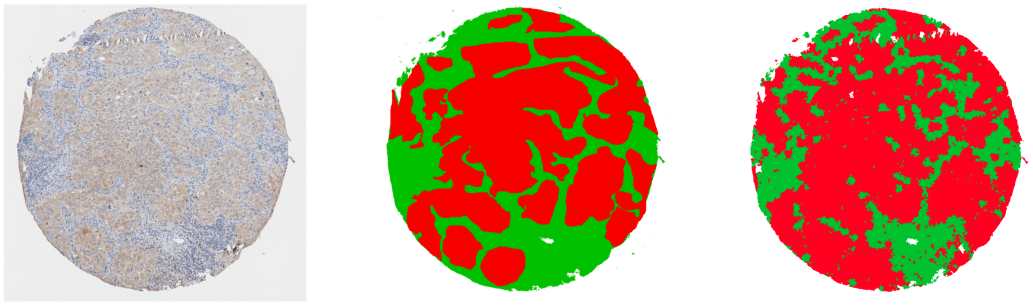
Además de seleccionar el mejor algoritmo, fue necesario seleccionar la mejor técnica de mejoramiento de la imagen, esta fue la ecualización del histograma que como se observa en la cuadro 5.5, en conjunto con Random Forest es la que obtiene el mejor resultado de

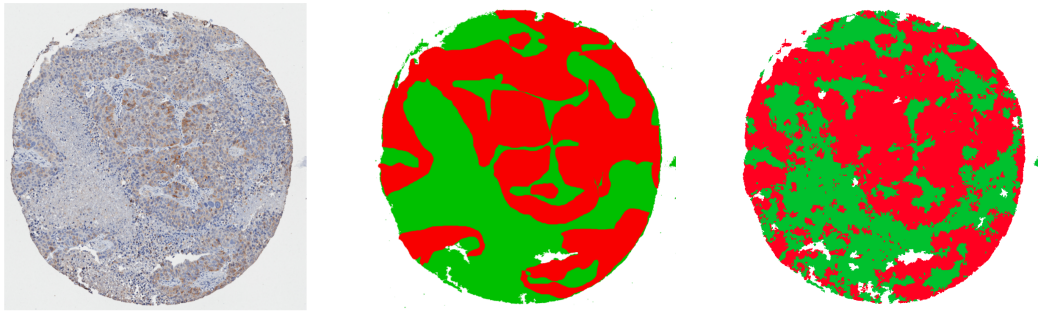
clasificación, a esta técnica le siguen histogram matching y por último normalización con resultados bastante similares.

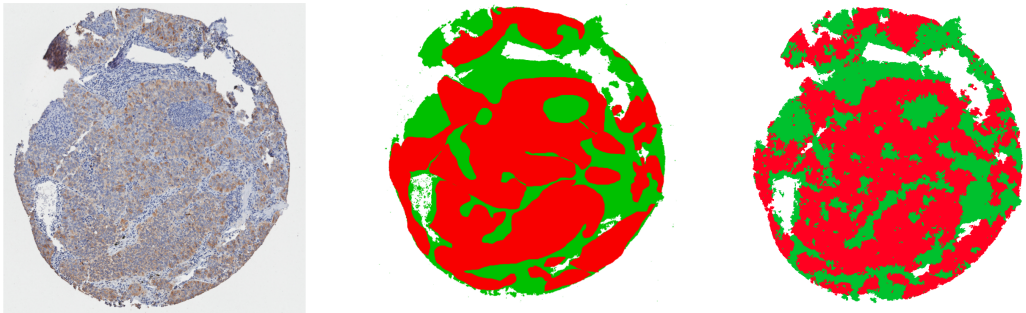
5.5. Imágenes segmentadas

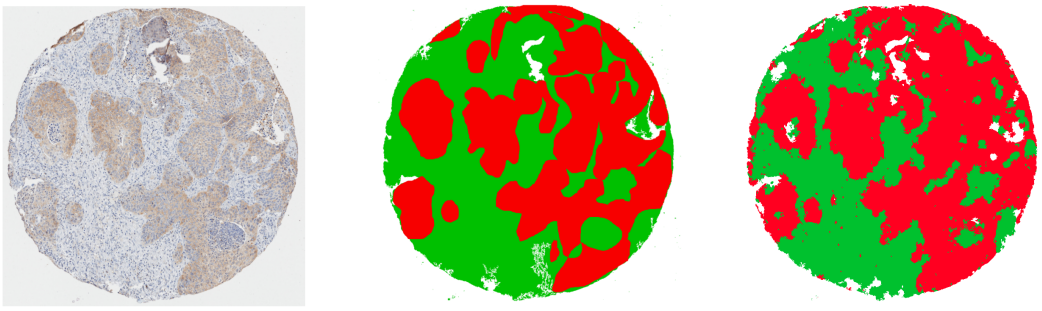
En este estudio se tomó la tarea de segmentación como un problema de clasificación de superpixeles, en donde a través de un análisis comparativo se determinó que la ecualización del histograma fue la mejor técnica de mejoramiento de la imagen y que el mejor algoritmo de clasificación fue Random Forest; juntos obtuvieron una precisión media de 83,4 % y una medida F de 84.4 %, con esta configuración se obtuvo la segmentación final y a continuación (figura 5.9) se presentan los resultados obtenidos con el método propuesto; en las imágenes de las columnas (b) y (c), los colores corresponden a rojo : tumor, verde: no-tumor y blanco: fondo.

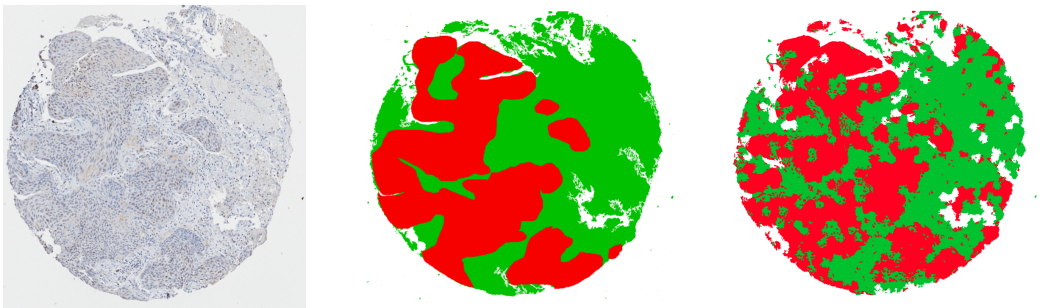


4. 

5. 

6. 

7. 

8. 

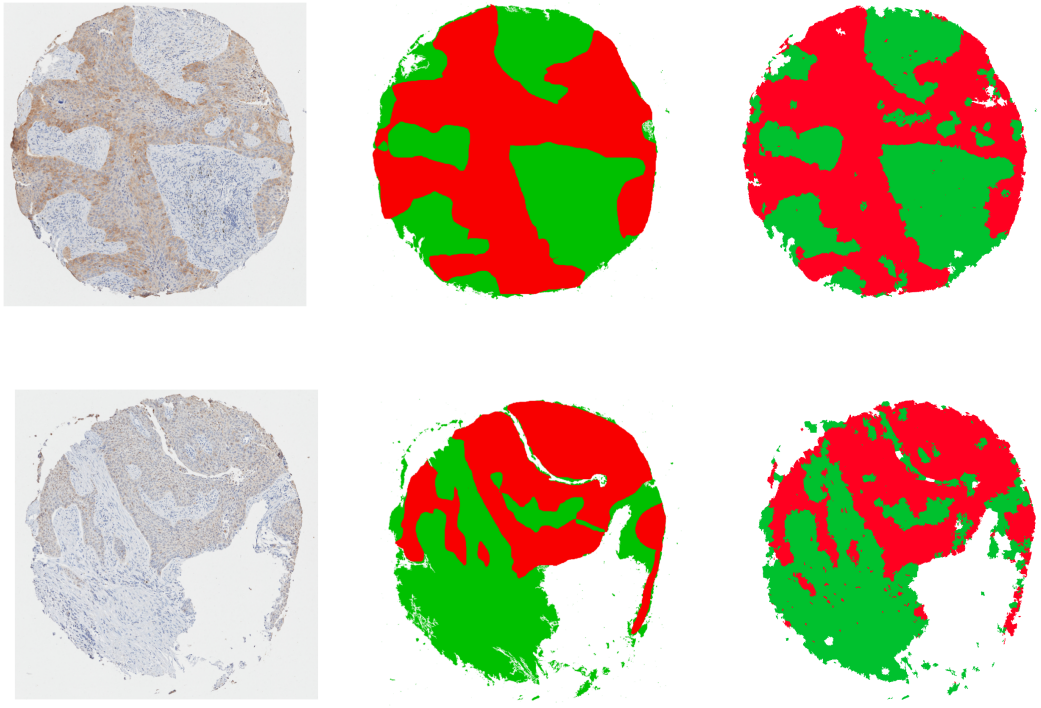


Figura 5.9: Resultado de segmentación del método propuesto. La columna (a) es la imagen original, (b) la imagen marcada por experto y (c) la segmentación automática CAD

Es preciso mencionar que en imágenes marcadas hay implícito un cierto grado de error (humano), sin embargo dicho marcado se toma como verdadero para poder tener una referencia con la cual comparar las imágenes segmentadas automáticamente. Ahora bien, al hacer dicha comparación (figura 5.9), se aprecia que el fondo (en color blanco) es bien diferenciado del tejido para la mayoría de las imágenes aunque al observar la imagen 2 se aprecia que partes de no-tumor son clasificadas como fondo, si se examina con más detalle la imagen inicial (a), se ve que los expertos marcaron estas partes como no-tumor, esto podría ser objeto de revisión por parte de los expertos para confirmar si la segmentación automática es o nó correcta. Aunque se sabe que el error humano es un factor presente dentro de las imágenes etiquetadas. Esto no supone un problema mayor y es bastante común dentro de las tareas de marcado manual.

En cuanto a la zona de tumor, la cual es la más relevante, se observa que en general el método acierta en la segmentación distinguiendo las principales regiones tumorales (color rojo) sin embargo también se evidencia sobre todo en la segmentación de las imágenes (c) 4, 5 y 8 que hay pequeñas áreas clasificadas como tumor que en realidad son no-tumor. A diferencia de la verdadera región tumoral que se distingue por estar agrupada en áreas más grandes, estas pequeñas áreas se caracterizan por ser pequeñas y estar distribuidas en toda el área no tumoral.

La imagen mejor segmentada es la 9 (c), esta se aproxima mucho a la imagen marcada manualmente (b). Las diferencias más notorias son partes de la región tumoral mal clasificadas como no-tumor. Al observar con mas detalle este aspecto, se cree que gran parte de dichas zonas si corresponden a zona no-tumor por tanto el sistema CAD podría haber acertado en esta parte de su segmentación. Por otra parte, la imagen peor segmentada es

la 8 en donde no se aprecia con claridad las formas de las regiones tumorales sino que se pierden entre pequeñas áreas mal clasificadas como zona no tumoral. Esto indicaría que las clases más confusas para el clasificador son tumor y no-tumor.

Capítulo 6

Discusión

Inicialmente se compararon 3 técnicas de mejoramiento de la imagen de las cuales la ecualización del histograma aportó una mejora significativa a la segmentación final del sistema CAD. Aunque en principio las técnicas de mejoramiento de la imagen se crearon para mejorar la interpretación visual por parte de los humanos, en este estudio se evidencia que también son efectivas para mejorar la interpretación por parte de sistemas computacionales. Un aspecto bastante interesante es que la ecualización del histograma a diferencia de las otras 2 técnicas no "normaliza" a nivel del conjunto de datos sino que trabaja independientemente en cada imagen y aunque sus resultados son buenos, no se podría saber el resultado de probar una imagen con una diferencia demasiado alejada al resto del conjunto.

Se evidencia que el trabajo con SLIC superpixeles mediante el cual se pasa de un plano de clasificación de pixeles a uno de superpixeles ha sido efectivo en capturar la redundancia de las imágenes TMA asi como también logrando un buen ajuste a sus diferentes estructuras histopatológicas. Cabe señalar que este proceso podría mejorar en el futuro mediante la optimización de los parámetros del método SLIC. Por otra parte, se ha observado que aunque los superpixeles capturan la redundancia de la imagen, el uso de superpixeles muy grandes no ayuda a obtener buenas características de textura. Por el contrario, introduce un mayor error en el proceso del ground truth. La estimación del error (y su minimización) entre las imágenes marcadas por los expertos y las segmentadas es un aspecto a considerar a futuro.

Las características de textura extraídas fueron apropiadas para construir descriptores que ayudaron a los algoritmos de clasificación a discriminar mejor entre las clases. Aunque el dataset construido tiene un buen tamaño, si se compara con estudios similares, se evidencia que es relativamente pequeño; una posibilidad de aumentar el poder descriptivo podría ser incluir un mayor número de imágenes TMA además de emplear un método automático de selección de variables dado que en muchos casos el rendimiento del modelo depende más de las características seleccionadas que del algoritmo de clasificación.

En cuando a la clasificación, se observó que el algoritmo que presentó el mejor desempeño para los 3 datasets de las 3 transformaciones de mejoramiento de imagen, fue Random Forest, alcanzando una precisión media de 83,4 % y una medida F de 84,4 %. Este resultado se aproxima bastante al obtenido por estudios similares, los cuales han reportado una

precisión de 72 % [48], 89 % [50], 90 % [45] y 92,4 % [47] para cáncer de pulmón, y para otros tipos: 69,2 % [2], 84,2 % [49] y 89 % [52] en cáncer de mama y 88 % [44], 96,7 % [46] para cáncer de próstata. Se observó que para los estudios con tasas de acierto superiores al 90 % un factor común fue el empleo de grandes conjuntos de datos (más de 100 imágenes) esto se traduce en modelos más robustos con una capacidad explicativa más amplia; otro aspecto a tener en cuenta es que no todos los estudios emplearon imágenes TMA, algunos utilizaron cortes de imágenes tradicionales con lo cual sus modelos solo clasificaron zonas de tumor y no tumor. El empleo de imágenes TMA otorgó al presente estudio una ventaja significativa pues con ello se asegura un marco de trabajo común y más reproducible a nivel científico.

En general, el método propuesto realiza una buena segmentación, las imágenes resultantes son bastante similares a las imágenes marcadas y aunque se segmenta incorrectamente pequeñas áreas de las distintas zonas, se puede decir que existe la posibilidad de que algunas de estas áreas "mal clasificadas" sean en realidad (falsos positivos), lo cual implica que con un marcado más exhaustivo por parte de los expertos se pueda optimizar la evaluación del error y mejorar el rendimiento del modelo.

Capítulo 7

Conclusiones

El cáncer de pulmón es el más frecuente en todo el mundo. En este trabajo se hizo una aproximación a su detección, abordando el problema de segmentación de tejido tumoral y tejido no tumoral sobre imágenes histopatológicas TMA, mediante el desarrollo de un prototipo de sistema de diagnóstico asistido por computador que permitió segmentar las zonas de tumor, no-tumor y fondo con una precisión media de 83,4 % y una medida F de 84.4 %.

Un hallazgo muy interesante fue que algunas de las regiones de tejido mal clasificadas podrían deberse principalmente a las anotaciones no detalladas (omitidas) por parte de los expertos; el sistema CAD logró descubrir regiones tumorales que no fueron marcadas por los expertos lo cual lo constituyó en "una segunda opinión diagnóstica".

Se ha desarrollado un método prometedor para enfocar el problema de análisis histopatológico con eficacia y que además disminuye considerablemente el tiempo de diagnóstico.

El método propuesto tiene potencial futuro de cuantificación del cáncer de pulmón, con ello se abre un camino para conocer con más precisión su estado, mejorar su estimación y predecir su progreso.

El pre-procesamiento basado en SLIC superpixeles captura la redundancia de la imagen reduciendo la complejidad del procesado. Las características de textura extraídas de los superpixeles, producen buenos descriptores de las clases. Esto finalmente optimiza la etapa de clasificación.

La necesidad de sistemas CAD en histopatología es apremiante. Se contribuye con un método de detección temprana del cáncer de pulmón considerando primero que la imagen histopatológica es el "gold stadard" en detección de cáncer y segundo que la mayor parte del trabajo hecho en CAD para cáncer de pulmón se ha enfocado en radiología y se ha evidenciado que esta área no se encamina en la detección temprana del cáncer.

Aunque este estudió se enmarcó en la detección del cáncer de pulmón, en teoría se ha construido un método que podría extender su aplicación hacia otros tipos de cáncer dado que la mayoría de ellos se diagnostican usando la imagen histopatológica. Por lo tanto, el empleo de imágenes TMA se constituye como una ventaja que contribuye a lograr este objetivo a futuro.

Capítulo 8

Trabajo Futuro

Existen varios aspectos en los cuales el método propuesto puede mejorarse. Como trabajo futuro se plantea lo siguiente:

- El principal objetivo futuro es implementar la cuantificación de las zonas tumoral y no tumoral, para ofrecer a los patólogos indicadores de estimación y progreso del cáncer de pulmón.
- Se sugiere experimentar con los parámetros de los superpixeles, tanto con el tamaño como también con el grado de deformación, con el objetivo de lograr unos superpixeles más eficaces.
- Extraer nuevas características para aumentar el poder descriptivo del dataset y emplear un método automático de selección de variables como estrategia de reducción de la dimensionalidad.
- El error humano está implícito en el marcado de las imágenes etiquetadas, es importante conocerlo y estimarlo. Una alternativa propuesta sería marcar las imágenes por más de un experto y medir la varianza ínter observador.
- Se necesita refinar el mercado de las imágenes por parte de los expertos.
- El conjunto de imágenes es relativamente pequeño, se plantea aumentar el número de imágenes marcadas como estrategia de aumento de la capacidad explicativa del modelo.
- Como última propuesta se plantea la posibilidad de añadir una nueva fase de post-procesamiento para depurar aquellos superpixeles incorrectamente clasificados.

Bibliografía

[1] Alberto Ruano-Ravina, Mónica Pérez Ríos, and Alberto Fernández-Villar. Cribado de cáncer de pulmón con tomografía computarizada de baja dosis después del national lung screening trial. el debate continúa abierto. *Archivos de Bronconeumología*, 49(4):158–165, 2013.

[2] Shazia Akbar, Lee Jordan, Alastair M Thompson, and Stephen J McKenna. Tumor localization in tissue microarrays using rotation invariant superpixel pyramids. In *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pages 1292–1295. IEEE, 2015.

[3] Frank H Netter. *Netter-Atlas de Anatomia Humana*. Elsevier Brasil, 2008.

[4] Carla R Moctezuma Velasco and Mario Patiño Zarco. Cáncer de pulmón. In *Anales de Radiología, México*, volume 8, pages 33–45, 2009.

[5] Julia Calzas Rodríguez, Isidoro Carlos Barneto Aranda, and José Miguel Sánchez Torres. Cáncer de pulmón. *Bol Neum. Clin*, 8(2):17, 1988.

[6] Dalong Wang, Minghui Zhang, Xuan Gao, and Lijuan Yu. Prognostic value of baseline 18 f-fdg pet/ct functional parameters in patients with advanced lung adenocarcinoma stratified by egfr mutation status. *PloS one*, 11(6):e0158307, 2016.

[7] Ana Esther Jiménez Massa. *Cáncer de pulmón y citocinas: variantes clínicas y genéticas*. PhD thesis, Universidad de Salamanca, 2011.

[8] Valentina Faoro and Anna Sapino. Tissue microarray (tma). In *Guidelines for Molecular Analysis in Archive Tissues*, pages 23–26. Springer, 2011.

[9] Nazar Jawhar. Tissue microarray: a rapidly evolving diagnostic and research tool. *Annals of Saudi medicine*, 29(2):123, 2009.

[10] Iqbal S Shergill, NK Shergill, M Arya, and HRH Patel. Tissue microarrays: a current medical research tool. *Current medical research and opinion*, 20(5):707–712, 2004.

[11] Juliana Escobar Stein, Aurora Astudillo González, Primitiva Menéndez Rodríguez, and Elena Belyakova. Aplicación de los tissue microarrays en el estudio inmunohistoquímico de los tumores. *Revista Española de Patología*, 39(1):11–17, 2006.

[12] Heang-Ping Chan, Kunio Doi, Carl J Vybrony, Robert A Schmidt, Charles E Metz, Kwok Leung Lam, Toshihiro Ogura, Yuzheng Wu, and Heber MacMahon. Improvement in radiologists' detection of clustered microcalcifications on mammograms: The potential of computer-aided diagnosis. *Investigative radiology*, 25(10):1102–1110, 1990.

[13] Kunio Doi. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized medical imaging and graphics*, 31(4):198–211, 2007.

[14] Kunio Doi. Current status and future potential of computer-aided diagnosis in medical imaging. *The British journal of radiology*, 78(suppl_1):s3–s19, 2005.

[15] Julian Marshall, Harlan M Romsdahl, and Ewout A Mante. User interface for computer aided diagnosis system, June 29 1999. US Patent 5,917,929.

[16] Gwilym S Lodwick, Theodore E Keats, and John P Dorst. The coding of roentgen images for computer analysis as applied to lung cancer 1. *Radiology*, 81(2):185–200, 1963.

[17] Samuel G Armato, Maryellen L Giger, Catherine J Moran, James T Blackburn, Kunio Doi, and Heber MacMahon. Computerized detection of pulmonary nodules on ct scans 1. *Radiographics*, 19(5):1303–1311, 1999.

[18] Yongbum Lee, Takeshi Hara, Hiroshi Fujita, Shigeki Itoh, and Takeo Ishigaki. Automated detection of pulmonary nodules in helical ct images based on an improved template-matching technique. *IEEE Transactions on medical imaging*, 20(7):595–604, 2001.

[19] Kenji Suzuki, Samuel G Armato, Feng Li, Shusuke Sone, et al. Massive training artificial neural network (mtann) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. *Medical physics*, 30(7):1602–1617, 2003.

[20] Samuel G Armato III, Geoffrey McLennan, Michael F McNitt-Gray, Charles R Meyer, David Yankelevitz, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, Ella A Kazerooni, Heber MacMahon, et al. Lung image database consortium: Developing a resource for the medical imaging research community 1. *Radiology*, 232(3):739–748, 2004.

[21] K Murphy, A Schilham, H Gietema, M Prokop, and B Van Ginneken. Automated detection of pulmonary nodules from low-dose computed tomography scans using a two-stage classification system based on local image features. In *Medical Imaging*, pages 651410–651410. International Society for Optics and Photonics, 2007.

[22] Xujiong Ye, Xinyu Lin, Jamshid Dehmeshki, Greg Slabaugh, and Gareth Beddoe. Shape-based computer-aided detection of lung nodules in thoracic ct images. *IEEE Transactions on Biomedical Engineering*, 56(7):1810–1820, 2009.

[23] Temesguen Messay, Russell C Hardie, and Steven K Rogers. A new computationally efficient cad system for pulmonary nodule detection in ct imagery. *Medical image analysis*, 14(3):390–406, 2010.

[24] S Aravind Kumar, J Ramesh, PT Vanathi, and K Gunavathi. Robust and automated lung nodule diagnosis from ct images based on fuzzy systems. In *Process Automation, Control and Computing (PACC), 2011 International Conference On*, pages 1–6. IEEE, 2011.

[25] Kyung Nyeo Jeon, Jin Mo Goo, Chang Hyun Lee, Youkyung Lee, Ji Yung Choo, Nyoung Keun Lee, Mi-Suk Shim, In Sun Lee, Kwang Gi Kim, David S Gierada, et al. Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening ct. *Investigative radiology*, 47(8):457, 2012.

[26] Macedo Firmino, Antônio H Morais, Roberto M Mendoça, Marcel R Dantas, Helio R Hekis, and Ricardo Valentim. Computer-aided detection system for lung cancer in computed tomography scans: review and future prospects. *Biomedical engineering online*, 13(1):41, 2014.

[27] Maxine Tan, Rudi Deklerck, Bart Jansen, Michel Bister, and Jan Cornelis. A novel computer-aided lung nodule detection system for ct images. *Medical physics*, 38(10):5630–5645, 2011.

[28] Raphael Rubin, David S Strayer, Emanuel Rubin, et al. *Rubin's pathology: clinicopathologic foundations of medicine*. Lippincott Williams & Wilkins, 2008.

[29] Metin N Gurcan, Laura E Boucheron, Ali Can, Anant Madabhushi, Nasir M Rajpoot, and Bulent Yener. Histopathological image analysis: A review. *IEEE reviews in biomedical engineering*, 2:147–171, 2009.

[30] Jason Hipp, Jerome Cheng, Liron Pantanowitz, Stephen Hewitt, Yukako Yagi, James Monaco, Anant Madabhushi, Jaime Rodriguez-Canales, Jeffrey Hanson, Sinchita Roy-Chowdhuri, et al. Image microarrays (ima): Digital pathology's missing tool. *Journal of pathology informatics*, 2(1):47, 2011.

[31] Jennifer A Hipp, Jason D Hipp, Megan Lim, Gaurav Sharma, Lauren B Smith, Stephen M Hewitt, and Ulysses GJ Balis. Image microarrays derived from tissue microarrays (ima-tma): New resource for computer-aided diagnostic algorithm development. *Journal of pathology informatics*, 3, 2012.

[32] Ashkan Tashk, Mohammad Sadegh Helfroush, Habibollah Danyali, and Mojgan Akbarzadeh. A novel cad system for mitosis detection using histopathology slide images. *Journal of medical signals and sensors*, 4(2):139, 2014.

[33] Cheng Chen, Wei Wang, John A Ozolek, and Gustavo K Rohde. A flexible and robust approach for segmenting cell nuclei from 2d microscopy images using supervised learning and template matching. *Cytometry Part A*, 83(5):495–507, 2013.

[34] Jun Xu, Lei Xiang, Qingshan Liu, Hannah Gilmore, Jianzhong Wu, Jinghai Tang, and Anant Madabhushi. Stacked sparse autoencoder (ssae) for nuclei detection on breast cancer histopathology images. *IEEE transactions on medical imaging*, 35(1):119–130, 2016.

[35] Peter J Schüffler, Thomas J Fuchs, Cheng Soon Ong, Volker Roth, and Joachim M Buhmann. Automated analysis of tissue micro-array images on the example of renal cell carcinoma. In *Similarity-Based Pattern Analysis and Recognition*, pages 219–245. Springer, 2013.

[36] Yousef Al-Kofahi, Wiem Lassoued, William Lee, and Badrinath Roysam. Improved automatic detection and segmentation of cell nuclei in histopathology images. *IEEE Transactions on Biomedical Engineering*, 57(4):841–852, 2010.

[37] Elisa Ficarra, Santa Di Cataldo, Andrea Acquaviva, and Enrico Macii. Automated segmentation of cells with ihc membrane staining. *IEEE Transactions on Biomedical Engineering*, 58(5):1421–1429, 2011.

[38] Michael Chiang, Sam Hallman, Amanda Cinquin, Nabora Reyes de Mochel, Adrian Paz, Shimako Kawauchi, Anne L Calof, Ken W Cho, Charless C Fowlkes, and Olivier Cinquin. Analysis of in vivo single cell behavior by high throughput, human-in-the-loop segmentation of three-dimensional images. *BMC bioinformatics*, 16(1):397, 2015.

[39] Xin Qi, Fuyong Xing, David J Foran, Lin Yang, et al. A fast, automatic segmentation algorithm for locating and delineating touching cell boundaries in imaged histopathology. *Methods of information in medicine*, 51(3):260, 2012.

[40] Ching-Wei Wang. Fast automatic quantitative cell replication with fluorescent live cell imaging. *BMC bioinformatics*, 13(1):21, 2012.

[41] Bilge Karaçali and Aydin Tözeren. Automated detection of regions of interest for tissue microarray experiments: an image texture analysis. *BMC Medical Imaging*, 7(1):2, 2007.

[42] Maqlin Paramanandam, Robinson Thamburaj, Marie Theresa Manipadam, and Atulya K Nagar. Boundary extraction for imperfectly segmented nuclei in breast histopathology images–a convex edge grouping approach. In *International Workshop on Combinatorial Image Analysis*, pages 250–261. Springer, 2014.

[43] Anna Korzynska, Lukasz Roszkowiak, Carlos Lopez, Ramon Bosch, Lukasz Witkowski, and Marylene Lejeune. Validation of various adaptive threshold methods of segmentation applied to follicular lymphoma digital images stained with 3, 3'-diaminobenzidine&haematoxylin. *Diagnostic pathology*, 8(1):48, 2013.

[44] Scott Doyle, Anant Madabhushi, Michael Feldman, and John Tomaszeweski. A boosting cascade for automated detection of prostate cancer from digitized histology. *Medical image computing and computer-assisted intervention–MICCAI 2006*, pages 504–511, 2006.

[45] Santa Cataldo, Elisa Ficarra, and Enrico Macii. Automated discrimination of pathological regions in tissue images: Unsupervised clustering vs. supervised svm classification. *Biomedical Engineering Systems and Technologies*, pages 344–356, 2009.

[46] Ali Tabesh, Mikhail Teverovskiy, Ho-Yuen Pang, Vinay P Kumar, David Verbel, Angeliki Kotsianti, and Olivier Saidi. Multifeature prostate cancer diagnosis and gleason grading of histological images. *IEEE transactions on medical imaging*, 26(10):1366–1378, 2007.

[47] Ching-Wei Wang and Cheng-Ping Yu. Automated morphological classification of lung cancer subtypes using h&e tissue images. *Machine vision and applications*, 24(7):1383–1391, 2013.

[48] JC Sieren, J Weydert, A Bell, B De Young, AR Smith, J Thiesse, E Namati, and Geoffrey McLennan. An automated segmentation approach for highlighting the histological complexity of human lung cancer. *Annals of biomedical engineering*, 38(12):3581–3591, 2010.

[49] Angel Cruz-Roa, Ajay Basavanhally, Fabio González, Hannah Gilmore, Michael Feldman, Shridar Ganesan, Natalie Shih, John Tomaszewski, and Anant Madabhushi. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In *SPIE medical imaging*, volume 9041, pages 904103–904103. International Society for Optics and Photonics, 2014.

[50] Peter W Hamilton, Yinhai Wang, Clinton Boyd, Jacqueline A James, Maurice B Loughrey, Joseph P Hougton, David P Boyle, Paul Kelly, Perry Maxwell, David McCleary, et al. Automated tumor analysis for molecular profiling in lung cancer. *Oncotarget*, 6(29):27938, 2015.

[51] Kun-Hsing Yu, Ce Zhang, Gerald J Berry, Russ B Altman, Christopher Ré, Daniel L Rubin, and Michael Snyder. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nature communications*, 7, 2016.

[52] Andrew H Beck, Ankur R Sangoi, Samuel Leung, Robert J Marinelli, Torsten O Nielsen, Marc J Van De Vijver, Robert B West, Matt Van De Rijn, and Daphne Koller. Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science translational medicine*, 3(108):108ra113–108ra113, 2011.

[53] Robert J Marinelli, Kelli Montgomery, Chih Long Liu, Nigam H Shah, Wijan Prapong, Michael Nitzberg, Zachariah K Zachariah, Gavin J Sherlock, Yasodha Natkunam, Robert B West, et al. The stanford tissue microarray database. *Nucleic acids research*, 36(suppl_1):D871–D877, 2007.

[54] Caroline A Schneider, Wayne S Rasband, and Kevin W Eliceiri. Nih image to imagej: 25 years of image analysis. *Nature methods*, 9(7):671–675, 2012.

[55] Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7):676–682, 2012.

[56] Michael R Berthold, Nicolas Cebron, Fabian Dill, Thomas R Gabriel, Tobias Kötter, Thorsten Meinl, Peter Ohl, Kilian Thiel, and Bernd Wiswedel. Knime-the konstanz information miner: version 2.0 and beyond. *AcM SIGKDD explorations Newsletter*, 11(1):26–31, 2009.

[57] Raman Maini and Himanshu Aggarwal. A comprehensive review of image enhancement techniques. *arXiv preprint arXiv:1003.4053*, 2010.

[58] Karel Zuiderveld. Contrast limited adaptive histogram equalization. In *Graphics gems IV*, pages 474–485. Academic Press Professional, Inc., 1994.

[59] Jong-Sen Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE transactions on pattern analysis and machine intelligence*, (2):165–168, 1980.

[60] L Enrique and Gómez Giovani. Visión computacional. *Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México*, 2011.

[61] Ref. 79, p. 94.

[62] Jian Guo Liu and Philippa J Mason. *Essential image processing and GIS for remote sensing*. John Wiley & Sons, 2013.

[63] Jiří Borovec. Fully automatic segmentation of stained histological cuts. In *International Student Conference on Electrical Engineering*, volume 17, pages 1–7, 2013.

[64] Radhakrishna Achanta, A Shaji, K Smith, A Lucchi, P Fua, and S Susstrunk. Slic superpixels. no. Technical report, EPFL-REPORT-149300, 2010.

[65] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelligence*, 34(11):2274–2282, 2012.

[66] Jiří Borovec and Jan Kybic. jslic: superpixels in imagej. 2014.

[67] Rafael Llobet Azpitarte. *Aportaciones al diagnóstico de cáncer asistido por ordenador*. PhD thesis, 2008, p. 40.

[68] Fumiaki Tomita and Saburo Tsuji. *Computer analysis of visual textures*, volume 102. Springer Science & Business Media, 2013.

[69] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, Man, and Cybernetics*, 8(6):460–473, 1978.

[70] Francesco Bianconi, Alberto Álvarez-Larrán, and Antonio Fernández. Discrimination between tumour epithelium and stroma via perception-based features. *Neurocomputing*, 154:119–126, 2015.

[71] Maroua Mehri, Pierre Héroux, Petra Gomez-Krämer, and Rémy Mullot. Texture feature benchmarking and evaluation for historical document image analysis. *International Journal on Document Analysis and Recognition*, pages 1–35, 2017.

[72] Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[73] J Platt. Fast training of support vector machines using sequential minimal optimization. advances in kernel methods-support vector learning, 185-208, 1999.

[74] Jerome Friedman, Trevor Hastie, Robert Tibshirani, et al. Additive logistic regression: a statistical view of boosting (with discussion and a rejoinder by the authors). *The annals of statistics*, 28(2):337–407, 2000.

[75] J Ross Quinlan. *C4. 5: programs for machine learning*. Elsevier, 2014.

[76] Remco R Bouckaert. Bayesian network classifiers in weka. 2004.

[77] Pierre A Devijver and Josef Kittler. *Pattern recognition: A statistical approach*. Prentice hall, 1982.

[78] Qian Huang and Byron Dom. Quantitative methods of evaluating image segmentation. In *Image Processing, 1995. Proceedings., International Conference on*, volume 3, pages 53–56. IEEE, 1995.

[79] Rafael C Gonzalez and Richard E Woods. *Image processing*, volume 2. 2007, p. 91.