



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

Técnicas de ayuda al diagnóstico de distrofias musculares basadas en procesamiento de imágenes y en sistemas reconfigurables

Autor: Borja Jesús Carrillo Pérez
Tutor: Dr. José Francisco López Feliciano
Fecha: Julio de 2016



t +34 928 451 086 | iuma@iuma.ulpgc.es
f +34 928 451 083 | www.iuma.ulpgc.es

Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

**Técnicas de ayuda al diagnóstico de distrofias
musculares basadas en procesamiento de imágenes y
en sistemas reconfigurables**

HOJA DE FIRMAS

Alumno: Borja Jesús Carrillo Pérez Fdo.:

Tutor: José Francisco López Feliciano Fdo.:

Fecha: Julio de 2016



t +34 928 451 086 | iuma@iuma.ulpgc.es
f +34 928 451 083 | www.iuma.ulpgc.es

Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

Técnicas de ayuda al diagnóstico de distrofias musculares basadas en procesamiento de imágenes y en sistemas reconfigurables

HOJA DE EVALUACIÓN

Calificación:

Presidente Roberto Sarmiento Rodríguez Fdo.:

Secretario Sunil Lalchand Khemchandani Fdo.:

Vocal Sebastián López Suárez Fdo.:

Fecha: Julio de 2016



t +34 928 451 086 | iuma@iuma.ulpgc.es
f +34 928 451 083 | www.iuma.ulpgc.es

Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria

Agradecimientos

Quiero agradecer a mi tutor José López y a los doctores del Materno Infantil, Beatriz Báez y Miguel Ángel Sánchez, su disponibilidad, consejos y ayuda.

Gracias a mis seres queridos por el inestimable apoyo.

Contenido

Índice de figuras	5
Índice de tablas	9
Acrónimos y siglas.....	11
Capítulo 1	13
Introducción y objetivos.....	13
1.1 Antecedentes.....	13
1.2 Motivación	13
1.3 Objetivos.....	14
1.4 Estructura de la memoria.....	15
Capítulo 2	17
Fundamentos teóricos y conceptos básicos	17
2.1 El tejido muscular	17
2.2 Las distrofias musculares.....	19
2.3 La biopsia muscular	20
2.4 Procedimiento tradicional para el diagnóstico de distrofias musculares	22
2.5 Introducción a la segmentación de imágenes.....	24
2.5.1 Espacios de color	24
2.5.2 Segmentación de una imagen.....	25
Capítulo 3	31
Estado del arte en sistemas de apoyo al diagnóstico de distrofias musculares.....	31
3.1 Sistemas que segmentan imágenes de biopsias musculares y cuantifican fibras	31
3.2 Sistemas que segmentan imágenes de biopsias musculares, cuantifican fibras y calculan diámetros celulares.....	34
Capítulo 4	37
Segmentación de imágenes de biopsias musculares.....	37
4.1 Características de la imagen de la biopsia.....	37
4.1.1 Corte histológico.....	37
4.1.2 Tinción de la muestra.....	38
4.1.3 Captura de la imagen de la biopsia muscular	39
4.2 Imágenes analizadas	41
4.3 Métodos analizados	43

4.3.1	Transformada de Watershed.....	44
4.3.2	Segmentación basada en color utilizando K-means.....	47
4.4	Aportaciones a la segmentación de muestras de biopsias musculares.....	51
4.5	Resultados y comparación entre métodos de segmentación.....	55
4.5.1	Resultados para la Imagen I.....	57
4.5.2	Resultados para la Imagen II.....	59
4.5.3	Resultados para la Imagen III	61
4.5.4	Resultados para la Imagen IV	63
Capítulo 5	67
Algoritmos para la obtención de parámetros de interés.....		67
5.1	Aportaciones a la cuantificación y clasificación de células musculares	67
5.2	Aportaciones al cálculo de los diámetros celulares	70
5.2.1	Método de fuerza bruta	70
5.2.2	Método del diámetro de Feret	72
5.2.3	Método de las tensiones principales	74
5.2.4	Resultados y comparaciones de métodos de cálculo de diámetros celulares 79	
Capítulo 6	83
Desarrollo de la interfaz gráfica de usuario: MuCSA.....		83
6.1	Introducción a la creación de interfaces gráficas de usuario con MATLAB .83	
6.2	Desarrollo del entorno: MuCSA.....	85
6.2.1	Ejecución del entorno y cargar una imagen	86
6.2.2	Segmentación de la imagen.....	86
6.2.3	Selección de las células cuyos diámetros son calculados.....	87
6.2.4	Cálculo de los diámetros.....	88
6.2.5	Obtención y tratamiento de los resultados	88
6.2.6	Cálculo de otros parámetros necesarios	90
6.2.7	Exportación de los resultados a EXCEL.....	91
6.3	Tutorial de utilización de MuCSA	91
6.4	Análisis de rendimiento del software.....	101
6.4.1	Análisis temporal de la segmentación	102
6.4.2	Análisis temporal de la clasificación y cuantificación de células	103
6.4.3	Análisis temporal del cálculo de los diámetros.....	104
Capítulo 7	105

Conclusiones y líneas futuras de trabajo	105
7.1 Conclusiones.....	105
7.2 Líneas futuras.....	107
Referencias	109

Índice de figuras

Figura 1: Diagrama del músculo esquelético [1].....	18
Figura 2: Sección transversal de un músculo esquelético humano con tinción de ATPasa A con pH=9.4 (aumento x40).....	22
Figura 3: Ejemplo de diámetro menor (en rojo) de una célula muscular.....	22
Figura 4: Este diagrama, que se encuentra en [2], demuestra la importancia de la medida del diámetro menor de cada fibra. Esta es la única medida que no se ve alterada tanto por una sección oblicua como por fibras retorcidas.....	23
Figura 5: A la izquierda, imagen original en escala de grises. A la derecha, su gradiente (en magnitud), mediante operador Sobel [8].....	26
Figura 6: La T indica el umbral elegido en el histograma [7].....	27
Figura 7: A la izquierda, imagen de un iceberg. En el centro, su histograma. A la derecha, imagen segmentada con dos umbrales y tres clases, siguiendo el método de Otsu [7].	28
Figura 8: Ejemplo de dilatación [7].	29
Figura 9: Imagen de biopsia muscular con H&E-Stain [13].....	31
Figura 10: Misma muestra con ATPasa pH=4.5 y pH=10.4 respectivamente [15]....	32
Figura 11: Diagrama del sistema propuesto por Sertel et al. [15].....	32
Figura 12: Misma muestra de ratón con tres tinciones distintas [17].....	33
Figura 13: Distinción de células musculares con NDICIA y su red [20].....	33
Figura 14: Tinción fluorescente del artículo de Dubach-Powell [25].	35
Figura 15: Ejemplos de pliegues en muestras de biopsias musculares.....	38
Figura 16: Muestras con tinción débil (superiores) frente a muestras con buena tinción (inferiores).....	38
Figura 17: Ejemplo comparativo de la definición de células con aumento x40 y x100. Izquierda x40, derecha x100.	39
Figura 18: Ejemplo de barrido de capturas con el microscopio sobre la muestra (aumento x100). Orden: de izquierda a derecha y de arriba abajo (de 1 a 6).....	40
Figura 19: Imagen completa tras aplicar Image Stitching.	40
Figura 20: Imagen I: Sección transversal sin pliegues y tinción correcta (grupo I)....	42
Figura 21: Imagen II: Sección transversal sin pliegues y tinción débil (grupo II).....	42
Figura 22: Imagen III: Sección transversal con pliegues y tinción correcta (grupo III).	43
Figura 23: Imagen IV: Sección transversal con pliegues y tinción débil (grupo IV). .	43
Figura 24: Diagrama de Watershed [7].....	44
Figura 25: Diagrama de la implementación desarrollada con el algoritmo de Watershed controlado por marcadores.	45
Figura 26: K centroides iniciales (K=3).	47
Figura 27: Distancia de cada objeto a los centroides.....	47
Figura 28: K grupos generados.....	47
Figura 29: Diagrama de la segmentación basada en K-means.	48

Figura 30: Ejemplo de clasificación K-means con $K=3$ para una de las imágenes del Servicio de Anatomía Patológica.....	49
Figura 31: Imagen que muestra en negro los objetos del clúster correspondiente a las células Tipo II. Las Tipo II aparecen de color claro con píxeles en negro mezclados.	50
Figura 32: Algoritmo desarrollado en este trabajo como solución a la segmentación de imágenes de muestras de biopsias musculares.	52
Figura 33: Arriba la imagen normalizada con contraste mejorado. En medio, aplicando el primer filtro gaussiano. Abajo, aplicando el segundo.....	54
Figura 34: Ejemplo de imagen etiquetada. Cada área aislada posee una etiqueta individual.	56
Figura 35: Diagrama de flujo de la cuantificación automática de áreas aisladas (células) en una imagen segmentada con etiquetas.....	56
Figura 36: Imagen I en escala de grises (luminancia).	57
Figura 37: Segmentación de la Imagen I mediante el algoritmo basado en la Transformada Watershed.....	58
Figura 38: Segmentación de la Imagen I mediante el algoritmo basado en el clasificador K-means.	58
Figura 39: Segmentación de la Imagen I mediante el algoritmo aportado en este trabajo.	58
Figura 40: Imagen II en escala de grises (luminancia).....	60
Figura 41: Segmentación de la Imagen II mediante el algoritmo basado en la Transformada Watershed.....	60
Figura 42: Segmentación de la Imagen II mediante el algoritmo basado en el clasificador K-means.	60
Figura 43: Segmentación de la Imagen II mediante el algoritmo aportado en este trabajo.....	61
Figura 44: Imagen III en escala de grises (luminancia).....	62
Figura 45: Segmentación de la Imagen III mediante el algoritmo basado en la Transformada Watershed.....	62
Figura 46: Segmentación de la Imagen III mediante el algoritmo basado en el clasificador K-means.....	62
Figura 47: Segmentación de la Imagen III mediante el algoritmo aportado en este trabajo.....	63
Figura 48: Imagen IV en escala de grises (luminancia).	64
Figura 49: Segmentación de la Imagen IV mediante el algoritmo basado en la Transformada Watershed.....	64
Figura 50: Segmentación de la Imagen IV mediante el algoritmo basado en el clasificador K-means.....	64
Figura 51: Segmentación de la Imagen IV mediante el algoritmo aportado en este trabajo.....	65
Figura 52: Diagrama del algoritmo desarrollado para la cuantificación y clasificación de células.....	67
Figura 53: Histograma de la imagen del paso 3 del diagrama.	68

Figura 54: Clasificación de tipos para la Imagen I.	69
Figura 55: Clasificación de tipos para la Imagen II.	69
Figura 56: Clasificación de tipos para la Imagen III.	69
Figura 57: Clasificación de tipos para la Imagen IV.	69
Figura 58: Cálculo de la distancia entre dos puntos en un plano.	71
Figura 59: Búsqueda mediante fuerza bruta del diámetro más largo, mediante el cálculo de todas las distancias posibles desde todos los píxeles.	71
Figura 60: Ejemplo de diámetro máximo de una célula.	71
Figura 61: Diámetro más largo proyectado sobre nuevos ejes.	72
Figura 62: Diámetro menor de la célula ejemplo, en rojo y sin puntear.	72
Figura 63: Ejemplo de diámetro de Feret.	72
Figura 64: Ejemplo de diámetro de Feret en otra dirección. Se observa un cambio de dirección con respecto a la Figura 63.	73
Figura 65: Píxeles con coordenadas mayor y menor en el eje de abscisa.	73
Figura 66: Ejemplo de un diámetro de Feret, una vez se ha girado la célula.	74
Figura 67: Descomposición de tensiones sobre un paralelepípedo [33].	74
Figura 68: Ejemplo de célula muscular con sus tensiones normales y tangenciales representadas, sobre un eje de coordenadas.	76
Figura 69: Célula muscular proyectada sobre los ejes de las tensiones principales.	76
Figura 70: Diámetros mayor y menor (en rojo) calculados en base a las direcciones de las tensiones principales.	77
Figura 71: Cálculo del diámetro menor (en rojo), mediante el método de las tensiones principales.	78
Figura 72: Primera célula analizada, con diámetros en rojo.	80
Figura 73: Segunda célula analizada, con sus diámetros en rojo.	80
Figura 74: Segunda célula analizada, con sus diámetros en rojo.	81
Figura 75: Ejemplo de interfaz creada con GUIDE [34].	84
Figura 76: Editor GUIDE, durante la creación de la interfaz gráfica de MuCSA.	85
Figura 77: Imagen original en escala de grises fusionada con su segmentada.	86
Figura 78: Casilla de verificación para habilitar la superposición de la imagen segmentada sobre la original en escala de grises.	87
Figura 79: a) Imagen I en escala de grises, b) Imagen I segmentada y con sus células clasificadas en Tipo I (claras) y Tipo II (oscuras), tanto a) como b) se abren en ventanas emergentes tras pulsar los botones pertinentes, c) en rosa las células Tipo I seleccionadas para calcular sus diámetros y d) en rojo las células Tipo II seleccionadas.	87
Figura 80: Aspecto de la ventana secundaria que muestra los resultados en el editor GUIDE.	89
Figura 81: a) células Tipo I etiquetadas y cuyos diámetros han sido calculados y b) las Tipo II.	89
Figura 82: Ejemplo de consulta de un diámetro.	90
Figura 83: Histogramas de diámetros de las células. A la izquierda las Tipo I y a la derecha las Tipo II.	90

Figura 84: Tabla con los parámetros resultantes.....	91
Figura 85: Ventana emergente al pulsar el botón de exportar datos.	91
Figura 86: Aspecto de MuCSA v1.0 al ejecutarse desde MATLAB.....	92
Figura 87: Cargar una nueva imagen.	92
Figura 88: Selección de la imagen desde el directorio en el que se encuentre.	93
Figura 89: Imagen I cargada en la interfaz.....	93
Figura 90: Interfaz con la Imagen I segmentada.	94
Figura 91: Interfaz con la Imagen I segmentada y fusionada con la original en escala de grises.....	94
Figura 92: Aplicación de zoom y navegación por las imágenes.....	94
Figura 93: Ventana secundaria con la imagen original en escala de grises, generada tras pulsar “Open original”	95
Figura 94: Ventana secundaria con la imagen segmentada y con sus células clasificadas, generada tras pulsar “Open classified”. En gris claro, las células Tipo I. En gris oscuro las Tipo II.	95
Figura 95: Células Tipo I seleccionadas, en rosa.	96
Figura 96: Proceso de selección manual de células Tipo I. a) y b) desección de una célula Tipo I, c) selección de una célula Tipo II como si fuera una Tipo I (por decisión del usuario)	97
Figura 97: Ventana de ayuda que muestra la explicación a los colores en el caso de la selección de las células Tipo I.....	98
Figura 98: Células Tipo II seleccionadas, en rojo.	98
Figura 99: Ventana de ayuda que muestra la explicación a los colores en el caso de la selección de las células Tipo I.....	99
Figura 100: Ventana secundaria con los resultados.	99
Figura 101: Ventana que emerge tras pulsar “Show Type I cells”. Muestra las células Tipo I, etiquetadas para las cuales han sido calculados los diámetros menores.....	100
Figura 102: Ventana que emerge tras pulsar “Show Type II cells”. Muestra las células Tipo II, etiquetadas para las cuales han sido calculados los diámetros menores.....	100
Figura 103: Consulta de diámetros. En este caso el 15 de las células Tipo I y el 7 de las Tipo II.	100
Figura 104: Ventana emergente al pulsar el botón de exportar datos.	101
Figura 105: Aspecto de las tres hojas de EXCEL, de izquierda a derecha, del fichero “.xls” generado.	101
Figura 106: Ejemplo de uso de la herramienta MATLAB Profiler.....	102

Índice de tablas

Tabla 1: Tipos de fibras de músculo esquelético [1].....	19
Tabla 2: Límites superiores del valor de los factores de Atrofia e Hipertrofia en adultos varones y mujeres.....	24
Tabla 3: Clasificación de las imágenes recibidas. La marca ✓ indica que el grupo posee esa característica y ✗ indica lo contrario.	41
Tabla 4: Operación bit a bit OR en el paso 10.....	55
Tabla 5: Tabla modelo de comparación cuantitativa entre algoritmos.....	57
Tabla 6: Resultados cuantitativos de la segmentación de la Imagen I.....	59
Tabla 7: Resultados cuantitativos de la segmentación de la Imagen II.	61
Tabla 8: Resultados cuantitativos de la segmentación de la Imagen III.....	63
Tabla 9: Resultados cuantitativos de la segmentación de la Imagen IV.....	65
Tabla 10: Resultados numéricos de la clasificación de tipos de células.....	70
Tabla 11: Resultados para la primera célula analizada.	80
Tabla 12: Resultados para la segunda célula analizada.....	81
Tabla 13: Resultados para la tercera célula analizada.....	81
Tabla 14: Análisis temporal de la segmentación.....	103
Tabla 15: Análisis temporal de la clasificación.....	103
Tabla 16: Análisis temporal del cálculo de los diámetros menores.	104

Acrónimos y siglas

ATPasa	Adenosín Trifosfasa
C.H.U.	Centro Hospitalario Universitario
CPU	Unidad Central de Procesamiento
FPGA	Field Programmable Gate Array
GUI	Graphical User Interface
GUIDE	Graphical User Interface Design and Evaluation
GPU	Unidad de Procesamiento Gráfico
H&E-Stain	Tinción Hematoxilina y Eosina
MHC	Myosin Heavy Chain
ppp	Puntos Por Pulgada
RGB	Red Green Blue

Capítulo 1

Introducción y objetivos

En este capítulo se enmarcan los antecedentes que motivaron el desarrollo de este Trabajo de Fin de Máster, se muestran los objetivos a alcanzar y se describe la estructura que sigue esta memoria.

1.1 Antecedentes

La idea de este Trabajo de Fin de Máster surgió tras una reunión llevada a cabo en el mes de octubre de 2015 con los Dres. Dña. Beatriz Báez Acosta y D. Miguel Ángel Sánchez Ramos, ambos médicos especialistas del Servicio de Anatomía Patológica del Complejo Hospitalario Universitario Insular – Materno Infantil de Las Palmas de Gran Canaria. El objetivo era explorar posibles vías de colaboración en el área del procesamiento digital de imágenes biomédicas.

En el Servicio de Anatomía Patológica, se ocupan del estudio de las causas, desarrollo y consecuencias de las enfermedades, cuyo fin es el diagnóstico correcto de biopsias, piezas quirúrgicas, citologías y autopsias.

En la reunión mencionada, entre los temas propuestos por los Dres., se encontraba el que trata este trabajo. La razón que impulsó elegir esta propuesta frente al resto fue la necesidad, por parte de los patólogos, de automatizar el análisis de muestras transversales de músculo esquelético humano obtenidas por biopsia, hasta ahora realizada mediante un procedimiento manual. El estudio anatómico de la sección transversal de los diferentes tipos de fibras que componen el tejido muscular, es fundamental para determinar la estructura y la funcionalidad del músculo esquelético sano o enfermo.

Periódicamente, se concertaron reuniones con los Dres. del Servicio de Anatomía Patológica para ir mostrando los resultados obtenidos y aclarar las dudas que surgían conforme avanzaba el presente trabajo.

1.2 Motivación

La biopsia muscular tiene un papel importante en el diagnóstico de enfermedades neuromusculares, de las cuales muchas son hereditarias. El diagnóstico rápido y correcto de este tipo de enfermedades permite un manejo óptimo de los pacientes y un posible consejo genético. Por otro lado, la biopsia muscular puede contribuir en el desarrollo y la aplicación de tratamientos nuevos. Dentro del amplio grupo de las enfermedades neuromusculares destacan las distrofias musculares, que son enfermedades crónicas, de origen genético y de naturaleza progresiva, es decir, el daño

que provocan aumenta con el paso del tiempo. Es por esto que un diagnóstico temprano permite que el deterioro muscular del paciente sea el menor posible.

La forma de detectar este tipo de enfermedades se basa en analizar de forma minuciosa el tejido muscular obtenido en biopsias de músculo esquelético. Estas muestras, tras ser congeladas con nitrógeno líquido y procesadas en el laboratorio, son estudiadas mediante microscopía óptica. A continuación, comienza una tarea ardua y tediosa que se basa en diferenciar un tipo de fibras frente a otro (Tipo I y Tipo II), y extraer a partir de las mismas una serie de características relacionadas con su geometría, así como otros valores estadísticos. En la mayoría de los servicios de anatomía patológica de los hospitales españoles, estas tareas se realizan de forma manual, empleándose para ello una gran cantidad de tiempo y personal altamente cualificado. que podría estar plenamente dedicado a la interpretación de los datos para realizar un diagnóstico preciso. El tiempo que dedica un patólogo experimentado en la obtención de estos datos es de 30 a 60 minutos por muestra. Posteriormente, el patólogo debe proceder a realizar los cálculos de los parámetros necesarios y la generación de tablas e histogramas pertinentes.

De la revisión del estado del arte en sistemas de apoyo al diagnóstico de distrofias musculares, que se muestra en el capítulo 3, se concluye que ninguno de los sistemas se adapta a los requisitos del Servicio de Anatomía Patológica.

En este Trabajo de Fin de Máster, se propone acelerar el proceso de diagnóstico de enfermedades musculares, liberando al personal médico de tareas que por otro lado son mecánicas y rutinarias, mediante la utilización de una herramienta aquí desarrollada.

1.3 Objetivos

El objetivo principal de este Trabajo de Fin de Máster es reducir de forma considerable los tiempos empleados en la actualidad para analizar imágenes de muestras de biopsias musculares y extraer los parámetros que permitan hacer una diagnosis correcta. Para ello, los patólogos del Servicio de Anatomía Patológica suministraron imágenes de casos reales de biopsias musculares.

Se han analizado las tareas llevadas a cabo manualmente por los patólogos para ser posteriormente automatizadas mediante su implementación en MATLAB. En concreto, estas tareas son:

- Cuantificación de fibras Tipo I, Tipo II y la suma de ambas.
- Selección de 100 fibras aleatorias de cada tipo y medición del diámetro menor de cada una. En términos geométricos, el diámetro menor de una célula puede ser definido como la perpendicular más larga al diámetro más largo (o mayor).

Tras la cuantificación de fibras y medición de diámetros, los patólogos continúan el procedimiento realizando cálculos como los porcentajes de cada tipo de fibras, el cálculo

del diámetro medio de cada tipo y total, generación de histogramas o cálculos de coeficientes.

Para conseguir automatizar las tareas descritas, previamente se investigarán y desarrollarán técnicas de segmentación de imágenes mediante MATLAB, para conseguir obtener los bordes de las células de forma eficaz.

Una vez los algoritmos desarrollados presenten la robustez suficiente, se creará una interfaz gráfica de usuario a medida para que su utilización resulte sencilla para los patólogos y les permita tomar decisiones.

Por último, se realizará un estudio del código generado y de los procesos del mismo, de los cuales se analizará el tiempo de ejecución para determinar aquellos computacionalmente más complejos y con mayor carga temporal. Además, se valorará la posibilidad de paralelizar los algoritmos.

1.4 Estructura de la memoria

En este capítulo, se ha realizado una introducción con los antecedentes y la motivación de este trabajo, y se definen los objetivos marcados para este Trabajo de Fin de Máster.

En el capítulo 2, se realiza un estudio teórico de los conceptos que conciernen a este trabajo. Entre ellos se encuentran el tejido muscular, las distrofias musculares, la biopsia como herramienta para el diagnóstico de distrofias y, por último, los conceptos básicos sobre procesamiento digital de imágenes, necesarios para el desarrollo de este trabajo.

En el capítulo 3, se estudian los sistemas dedicados al estudio de células musculares, cuyas características son de interés para la realización de este trabajo. Se muestran los trabajos desarrollados a partir de 2007 hasta la actualidad. Es posible dividir los sistemas revisados en dos clases: sistemas que segmentan imágenes de biopsias musculares y realizan la cuantificación de fibras; y sistemas que realizan lo anterior, así como el cálculo de diámetros celulares.

En el capítulo 4, se explica cuáles son las características que deben tener las imágenes para un análisis óptimo por parte del software. Seguidamente, se muestran cuatro de las imágenes del Servicio de Anatomía Patológica que han sido analizadas para demostrar las bondades de los algoritmos. Consecutivamente, se expone el estudio de dos algoritmos desarrollados en base a dos métodos de segmentación conocidos: Watershed y K-means. Posteriormente, se propone un algoritmo novedoso, desarrollado en este trabajo, como solución a la segmentación. Por último, se presentan los resultados de la segmentación mediante los tres métodos mencionados en las cuatro imágenes de referencia.

En el capítulo 5, se muestra el algoritmo desarrollado para cuantificar y clasificar células en sus correspondientes tipos. A continuación, se comparan y se discuten tres

algoritmos desarrollados para el cálculo del diámetro menor, con el objetivo de comparar sus resultados y elegir la mejor solución.

En el capítulo 6, se describe el desarrollo de la interfaz gráfica que actúa como herramienta de apoyo al diagnóstico de distrofias musculares. En primer lugar, se muestran las posibilidades que ofrece MATLAB a la hora de crear interfaces gráficas. En segundo lugar, se describen las funcionalidades implementadas en la interfaz, las cuales principalmente parten de los algoritmos desarrollados anteriormente, así como de otras características añadidas. En tercer lugar, se describe la utilización de la herramienta a modo de tutorial. En cuarto y último lugar, se analiza del código generado y de los procesos del mismo, y se escogen aquellos que son computacionalmente más complejos y con mayor carga temporal.

En el capítulo 7, se exponen las conclusiones a las que se ha llegado tras la realización de este Trabajo Fin de Máster, teniendo en cuenta los resultados obtenidos en capítulos anteriores. Además, se muestran las líneas futuras que formarían parte de la continuación de este trabajo.

Por último, se muestra una lista con las fuentes bibliográficas consultadas durante el desarrollo de este Trabajo de Fin de Máster, y la elaboración de la memoria.

Capítulo 2

Fundamentos teóricos y conceptos básicos

En este capítulo se tratan los conceptos teóricos que conciernen a este trabajo. Se introducen conceptos de anatomía patológica para el diagnóstico de distrofias musculares. Se parte de la definición de tejido muscular, para continuar mostrando la biopsia muscular como herramienta de diagnóstico y el procedimiento posterior seguido por los patólogos. Por último, se introducen conceptos teóricos sobre procesamiento digital de imágenes, necesarios en el desarrollo de este trabajo, entre ellos los términos relacionados con la segmentación de imágenes.

2.1 El tejido muscular

Muchas de las células de los organismos multicelulares tienen la capacidad de contracción, aunque de forma limitada. Las células musculares, sin embargo, están especializadas en la contracción, y esta habilidad es la que permite el movimiento en animales. Mediante la contracción de las células musculares y la disposición de los componentes extracelulares, los músculos realizan la locomoción, constricción, bombeo y otros movimientos de propulsión [1].

Las células musculares son alargadas y conforman el denominado tejido muscular, del cual pueden diferenciarse el músculo liso y el estriado:

- El músculo liso, cuya contracción es involuntaria, se localiza en las paredes de vasos sanguíneos y vísceras, así como en la dermis de la piel [1].
- Las células del músculo estriado muestran de modo característico dos tipos alternados de bandas transversales (claras y oscuras), que no existen en el músculo liso. Existen dos tipos de músculo estriado:
 - o Cardíaco: se limita al tejido muscular del corazón.
 - o Esquelético (Figura 1): constituye la mayor parte de la masa muscular voluntaria del cuerpo. Está compuesto de células multinucleadas largas y cilíndricas que facilitan el movimiento voluntario del cuerpo o sus partes.

En este trabajo es de interés el tejido muscular estriado esquelético humano, concretamente la sección transversal del mismo, y es por eso que sus células se describen con mayor detalle a continuación.

Como se observa en la Figura 1, todo el músculo está rodeado de epimisio, una capa de tejido conectivo denso e irregular. El perimisio, formado por tejido conjuntivo menos denso, organiza las células musculares en fascículos. El endomisio, se compone de fibras reticulares y una lámina externa, que circunda cada célula muscular. Las células que comparten lámina externa hacen de regenerativas entre sí y son llamadas células satélite. Asimismo, cada fibra está constituida por su membrana citoplásmica o sarcolema, el cual contiene miofibrillas cilíndricas, diversas proteínas o enzimas, organelas como las mitocondrias o el retículo sarcoplásmico y por varios núcleos, que se encuentran en la periferia.

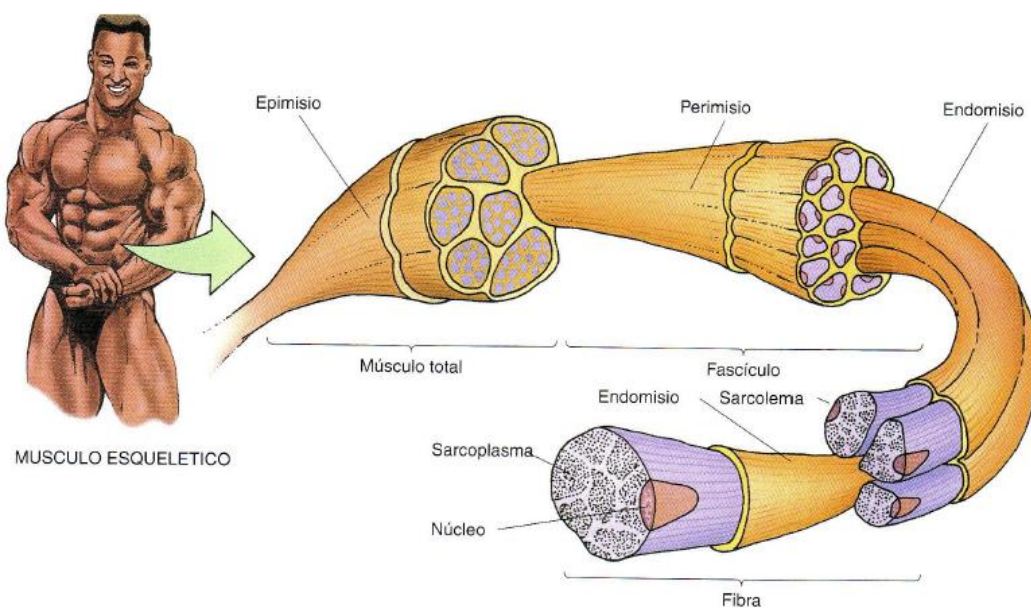


Figura 1: Diagrama del músculo esquelético [1].

Debido a que las células musculares poseen formas mucho más alargadas que anchas, a menudo son llamadas fibras musculares. Las fibras musculares están dispuestas de manera paralela entre sí. Entre las células se disponen espacios intercelulares intermedios, que incluyen conjuntos paralelos de capilares continuos. En este trabajo se tratan los términos fibra y célula muscular indistintamente.

La fuerza relativa de una fibra muscular depende directamente de su diámetro [2], en tanto que la fuerza de todo el músculo depende del número y grosor de fibras que lo componen. Según sea el diámetro de la fibra, la cantidad de mioglobina, número de mitocondrias, extensión del retículo sarcoplásmico, concentración de diversas enzimas e índice de contracción, las fibras musculares pueden clasificarse como rojas (Tipo I) o blancas (Tipo II). En la Tabla 1 se pueden observar algunas de las diferencias entre estos tipos de fibras.

Características	Fibras rojas (Tipo I)	Fibras blancas (Tipo II)
Vascularización	Irrigación vascular abundante	Riego más escaso
Inervación	Fibras nerviosas pequeñas	Fibras nerviosas grandes
Diámetro de la fibra	Más pequeño	Más grande
Contracción	Lenta	Rápida
Retículo sarcoplásmico	No extenso	Extenso
Mitocondrias	Numerosas	Varias
Mioglobina	Abundante	Poca
Enzimas	Deficiente adenosín trifosfa (ATPasa)	Abundante adenosín trifosfa (ATPasa)

Tabla 1: Tipos de fibras de músculo esquelético [1].

Gracias a la microscopía óptica, fue posible determinar la importancia funcional y morfológica de las estriaciones transversales del músculo esquelético, es decir, contribuyó a precisar que la forma y las dimensiones de las células poseen un alto valor, como se verá en las secciones posteriores.

2.2 Las distrofias musculares

El conjunto de las distrofias musculares lo conforman más de 30 enfermedades neurológicas [3]. Son crónicas, de origen genético y de naturaleza generalmente progresiva, es decir, el daño que provocan aumenta con el paso del tiempo. Se encuentran dentro del grupo de las denominadas enfermedades raras [4] y su aparición puede producirse en cualquier etapa de la vida, tanto en el nacimiento como en la adolescencia y la edad adulta.

Se estima que las distrofias afectan a 250.000 niños en todo el mundo, dándose más de 20.000 nuevos casos anuales [3], lo cual implica que uno de cada 3.500 nacimientos se ve afectado, independientemente del lugar de procedencia o la raza. En el caso de España, se estiman 4.000 afectados y 130 nuevos casos anuales [5].

Las principales características de las distrofias son la debilidad y la degeneración progresiva de los músculos esqueléticos usados durante el movimiento voluntario [5]. La mayoría de los pacientes pierden la capacidad de caminar y otras funciones motoras básicas.

La degeneración muscular progresiva se produce en la membrana protectora de las fibras musculares, el sarcolema. Cuando comienzan a perder la proteína creatina quinasa, que es necesaria para catalizar las reacciones químicas que producen la energía necesaria para la contracción muscular, y comienzan a captar calcio excesivo, se producen daños en las fibras musculares que llevan a la muerte de las propias células [3].

Existen numerosos grupos principales de distrofias musculares [5]. Los trastornos se clasifican por el alcance y la distribución de la debilidad muscular, edad al inicio, velocidad de evolución, gravedad de los síntomas, y antecedentes familiares (incluido cualquier patrón de herencia). Aunque algunas formas de distrofia muscular se hacen aparentes en la infancia o la niñez, otras pueden no aparecer hasta la mediana edad o más tarde. Las tasas de incidencia y gravedad varían, pero, en general, cada una de las distrofias causa deterioro progresivo de los músculos esqueléticos, e incluso algunos tipos afectan también al músculo cardíaco.

La microscopía óptica, mediante el análisis de un corte histológico transversal de músculo esquelético, permite identificar cambios en los componentes subcelulares de las fibras musculares. También puede identificar cambios morfológicos en las fibras que, con la correlación clínica adecuada, harán posible el diagnóstico. Entre dichos cambios morfológicos se encuentran las alteraciones en la geometría, en los diámetros, en la distribución de los tipos de fibras, en el aumento del tejido conectivo, en la presencia de necrosis y regeneración o inflamación en las fibras.

2.3 La biopsia muscular

Una biopsia de músculo esquelético es el procedimiento que se lleva a cabo cuando existen indicios de enfermedad neuromuscular en un paciente. Consiste en la obtención de un corte histológico de una sección transversal de músculo esquelético, el cual es congelado en nitrógeno líquido y en el que los patólogos se apoyan para examinar de forma minuciosa, mediante microscopía óptica, las células musculares. Es la principal herramienta en el diagnóstico de enfermedades musculares ya que permite observar de forma directa las alteraciones del músculo [5].

La biopsia muscular fue introducida en la práctica médica por Duchenne, en 1868, y fue Griesinger, en la misma época, quien obtuvo un fragmento de tejido para el estudio de un paciente con distrofia muscular [6].

Existen dos técnicas de biopsias de músculo esquelético, según el método de extracción de la muestra:

- La biopsia muscular abierta es el método utilizado en la mayoría de los hospitales españoles. Se utiliza principalmente en niños menores de cuatro años. Para estos, además de la anestesia local se requiere sedación en quirófano [6]. Para obtener la muestra de tejido, se puede seccionar directamente el músculo a cielo abierto, a través de una incisión no menor de 2 cm sin pinzamientos ni ligaduras.
- La biopsia muscular por punción percutánea es el método utilizado en los centros hospitalarios más especializados. No requiere hospitalización y no es doloroso ni deja cicatriz. Esto aporta la ventaja de poder repetir el procedimiento en caso

de obtener material inadecuado. Además, al no ser tan invasivo, es más aceptado por los pacientes y permite realizar estudios en personas sanas.

En este procedimiento, se realiza una incisión en la piel de 2 a 3 mm bajo anestesia local, se introduce una aguja de 4 mm hasta alcanzar la masa muscular, se extrae y se introduce el émbolo que secciona la parte extraída [6].

El tamaño y la forma de las fibras musculares son factores importantes para el diagnóstico de enfermedades neuromusculares. Uno de los mayores desafíos del estudio es la obtención de las secciones transversales del músculo mediante una biopsia. Como se ha mostrado en el capítulo anterior, el músculo esquelético se compone de células alargadas (fibras), estrechamente agrupadas en fascículos. Esta organización estructural puede provocar que durante el seccionamiento se produzcan pliegues en las células y los límites de estas se confundan al observar la muestra en el microscopio.

Una vez extraída la muestra de músculo, es observada mediante una lupa para realizar la orientación manual de las fibras musculares mediante bisturíes finos. Tras la orientación de la muestra, ésta es congelada en nitrógeno líquido. Así la biopsia muscular queda ya preparada para la obtención de las secciones histológicas y la aplicación de diversas técnicas de laboratorio.

Una vez seccionado el tejido, se pasa a la aplicación de técnicas histológicas, entre las que se encuentran las tinciones de las muestras. En este trabajo se muestra una de las técnicas utilizada por el personal del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil de Las Palmas de Gran Canaria, que se muestra a continuación.

La técnica que el personal del Servicio de Anatomía Patológica emplea es la tinción histoquímica enzimática de adenosín trifosfato, o comúnmente denominada ATPasa [2], aunque existen otras tinciones de tejido muscular esquelético. El adenosín trifosfato miofibrilar dependiente de calcio es una enzima que forma parte de la molécula de miosina [6]. ATPasa es la técnica que los patólogos eligen para evaluar el tamaño, la distribución y el número de fibras Tipo I y Tipo II (ver Tabla 1).

La muestra puede preincubarse en medio ácido o alcalino, lo que da por resultado reacciones inversas en cada tipo de fibra, es decir, el color de la reacción depende del pH utilizado. Por ejemplo, a pH=9.4 (ATPasa A) las fibras Tipo I son claras y a pH=4.6 (ATPasa B) son oscuras [6].

Una vez se tiñe la muestra de músculo, este se estudia mediante microscopía. El patólogo utiliza un microscopio para realizar el procedimiento a través del cual obtiene parámetros de interés para el diagnóstico de enfermedades neuromusculares. Este procedimiento se detalla en la siguiente sección.

2.4 Procedimiento tradicional para el diagnóstico de distrofias musculares

Cuando el patólogo posee la sección transversal de músculo esquelético, ya teñida y en un portaobjetos, y procede a trabajar con el microscopio, lo hace sobre una imagen parecida a lo que se observa en la Figura 2: (aumento x40). En esta muestra, la tinción realizada fue la de ATPasa A con pH=9.4, por tanto, las células claras se corresponden con las de Tipo I y las oscuras con las de Tipo II.

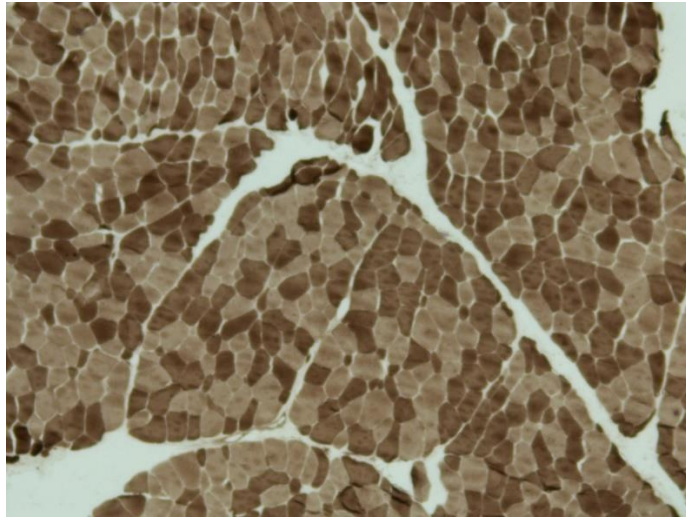


Figura 2: Sección transversal de un músculo esquelético humano con tinción de ATPasa A con pH=9.4 (aumento x40).

Antes de mostrar los pasos del procedimiento, es necesario definir un nuevo concepto. El “diámetro menor” [2] es el parámetro esencial en células musculares y se define como el máximo diámetro a lo largo del aspecto menor de la fibra. Geométricamente se puede traducir como el diámetro más largo perpendicular al diámetro más largo de la fibra (ver Figura 3).

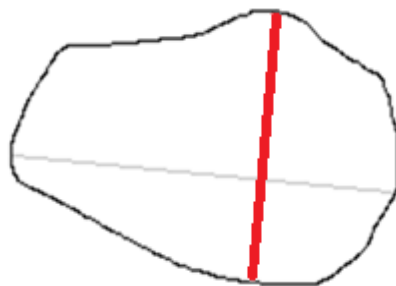


Figura 3: Ejemplo de diámetro menor (en rojo) de una célula muscular.

La importancia de la medida de este diámetro reside en la eliminación de la distorsión, que ocurre cuando la fibra es cortada de forma oblicua o está retorcida. A menos que se mida el diámetro menor, la medida resultará incorrecta [2], como se muestra en la Figura 4.

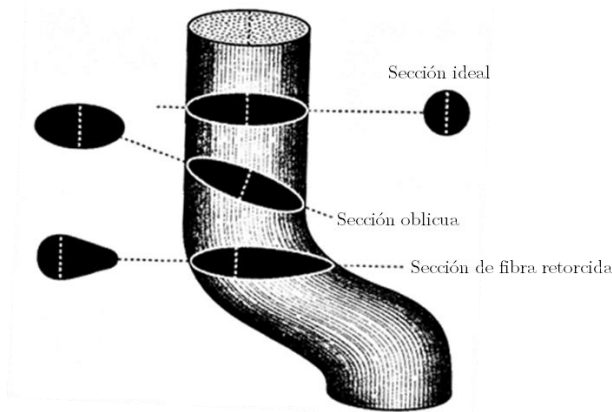


Figura 4: Este diagrama, que se encuentra en [2], demuestra la importancia de la medida del diámetro menor de cada fibra. Esta es la única medida que no se ve alterada ni por una sección oblicua ni por fibras retorcidas.

El patólogo realiza los siguientes pasos en el procedimiento:

1. Cuantificación de fibras Tipo I, Tipo II y la suma total.
2. Selección de 100 células aleatorias de cada tipo y medición del diámetro menor de cada una.

Este procedimiento lleva a un patólogo experimentado a emplear un tiempo estimado de 30 a 60 minutos por muestra. Tras la cuantificación de fibras y medición de diámetros, el procedimiento continúa con los pasos siguientes:

- Cálculo del porcentaje de fibras Tipo I y Tipo II frente al total.
- Cálculo del diámetro medio de cada tipo y del diámetro medio total.
- Generación de dos histogramas con los diámetros, uno por cada tipo de célula.
- Evaluación de la variabilidad mediante el Coeficiente de Variabilidad [2]:

$$\text{Coeficiente de variabilidad} = \frac{\text{Desviación estándar} \times 100}{\text{Diámetro medio}} \quad (1)$$

En un músculo normal este coeficiente es menor de 250 [2].

- Determinación de los Factores de Atrofia e Hipertrofia para cada tipo de fibra: en el músculo de un adulto normal, el diámetro modal se sitúa entre 40 y 80 μm . El factor de atrofia se calcula de la siguiente forma [2]:

$$\text{Factor de Atrofia} = \frac{(N_{\text{fibra}_{30\mu\text{m}<d<40\mu\text{m}}} + 2 \cdot N_{\text{fibra}_{20\mu\text{m}<d<30\mu\text{m}}} + 3 \cdot N_{\text{fibra}_{10\mu\text{m}<d<20\mu\text{m}}} + 4 \cdot N_{\text{fibra}_{d<10\mu\text{m}}})}{N_{\text{total de fibras en el histograma}}} \cdot 1000 \quad (2)$$

donde se multiplica el número de fibras en el histograma con diámetro (d) entre 30 y 40 μm por 1, el número de fibras con diámetro entre 20 y 30 μm por 2, el número de aquellas entre 10 y 20 μm por 3 y finalmente aquellas cuyo diámetro sea menor de 10 μm por 4. Tras sumar estos productos, se dividen entre el

número total de fibras presentes en el histograma y se multiplica el resultado por 1000.

El factor de hipertrofia se calcula expresando la proporción de fibras mayores de 80 μm :

$$\text{Factor de Hipertrofia} = \frac{N_{\text{fibras}_{d>80\mu\text{m}}}}{N \text{ total de fibras en el histograma}} \cdot 1000 \quad (3)$$

donde se divide el número de fibras en el histograma con diámetro mayor a 80 μm entre el número total de fibras presentes en el histograma y se multiplica el resultado por 1000.

La Tabla 2 muestra un resumen de los valores límite de los Factores de Atrofia e Hipertrofia en músculo sano tanto para varones como para mujeres.

		Fibras Tipo I		Fibras Tipo II	
		F. Atrofia	F. Hipertrofia	F. Atrofia	F. Hipertrofia
Bíceps	Varones	150	300	150	500
	Mujeres	100	200	150	150
Vastus	Varones	150	150	150	400
	Mujeres	100	400	200	150

Tabla 2: Límites superiores del valor de los factores de Atrofia e Hipertrofia en adultos varones y mujeres.

Para acelerar todo el procedimiento nombrado, se hace necesaria la creación de un entorno que segmente imágenes de muestras de biopsias musculares, para obtener las células de forma individual, y automatizar posteriormente las tareas descritas.

2.5 Introducción a la segmentación de imágenes

En esta sección se da una visión general de la segmentación de imágenes para introducir los conceptos que son posteriormente utilizados en los siguientes capítulos. La segmentación de imágenes consiste en la detección de grupos en una imagen en función de lo que se quiere determinar. En este trabajo la segmentación se utiliza para obtener el contorno de las células musculares, y es por eso que se incluye esta sección teórica.

2.5.1 Espacios de color

El espacio de color RGB (Red/Green/Blue) es el más común en sistemas de procesamiento digital de imágenes. Sin embargo, existen otros espacios de color cuyas características pueden ser convenientes o apropiadas a la hora de trabajar con ciertas imágenes digitales. Todos los espacios de color se pueden obtener mediante transformaciones bilineales del espacio RGB.

Entre los espacios de color se encuentra el YCbCr, cuya descripción se destaca debido a su importancia para este trabajo. La Y indica luminancia, Cb es la diferencia entre

la componente azul y un valor de referencia y la componente Cr es la diferencia entre la componente roja y un valor de referencia [7]. La transformación bilineal de RGB a YCbCr se define de la siguiente forma:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4)$$

Entre otros espacios de color se encuentran:

- YIQ (Luminance Hue Saturation).
- HSV (Hue Saturation Value).
- CMYK (Cian Magenta Yellow Key).
- HSI (Hue Saturation Intensity).

2.5.2 Segmentación de una imagen

Segmentar una imagen implica dividirla en regiones u objetos, los cuales están correlacionados con la realidad. El nivel al que llega esta división depende del problema que se intenta resolver. Esto significa, que la segmentación debería aplicarse hasta que el objeto o área de interés haya sido aislado [7]. Por tanto, al segmentar una imagen, se intenta que los píxeles correspondientes a un mismo objeto tengan el mismo valor.

Los algoritmos de segmentación generalmente están basados en dos propiedades básicas de los valores de intensidad de las imágenes [7]: discontinuidad y similitud. En la primera categoría, la meta es conseguir dividir la imagen basándose en cambios abruptos en la intensidad, como por ejemplo en los bordes de los objetos. En la segunda, se trata de dividir la imagen en regiones que son similares de acuerdo a un criterio predefinido.

A continuación, se describen conceptos básicos de segmentación de imágenes, los cuales son necesarios para la comprensión de este trabajo. Los dos primeros se relacionan con los cambios abruptos en la intensidad de la imagen y los siguientes con la similitud entre regiones.

2.5.2.1 Detección de bordes

La detección de bordes en una imagen es de las tareas más comunes en la segmentación de imágenes [7]. Consiste en la detección de discontinuidades significativas en los valores de intensidad mediante las derivadas de primer y segundo orden. La derivada de primer orden utilizada en el procesamiento de imágenes es el gradiente, que se define para imágenes de dos dimensiones como el vector:

$$\nabla f(x, y) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (5)$$

La magnitud de este vector es:

$$\text{mag}(\nabla f(x,y)) = [G_x^2 + G_y^2]^{\frac{1}{2}} = \left[\left(\frac{\partial f(x,y)}{\partial x} \right)^2 + \left(\frac{\partial f(x,y)}{\partial y} \right)^2 \right]^{\frac{1}{2}} \quad (6)$$

Las derivadas de segundo orden en procesamiento de imágenes se calculan, generalmente, usando el Laplaciano:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \quad (7)$$

Como se ha visto, la idea básica de la detección de bordes es encontrar zonas en la imagen donde la intensidad varíe bruscamente. Esto puede conseguirse mediante uno de los dos criterios siguientes:

- Si se encuentran lugares donde la primera derivada de la intensidad sea mayor en magnitud que en un umbral específico.
- Si se encuentran lugares donde la segunda derivada de la intensidad tenga un cruce de ceros.

La clave reside en cómo estimar las derivadas G_x y G_y computacionalmente. Es por ello que existen varios métodos de detección de bordes en función de la estimación de la derivada que realizan: Sobel (ver Figura 5), Prewitt, Roberts, Laplaciano del Gausiano, Zero-Crossings y Canny, son algunos de ellos [7].



Figura 5: A la izquierda, imagen original en escala de grises. A la derecha, su gradiente (en magnitud), mediante operador Sobel [8].

2.5.2.2 Filtrado espacial

Las imágenes pueden tener información no deseada que la contaminan. Este ruido puede estar presente tanto en el proceso de adquisición, en el de transmisión o el de procesamiento de la imagen [9]. Los tipos de ruidos existentes en una imagen son:

- Ruido impulsivo o sal y pimienta.
- Ruido aditivo.
- Ruido multiplicativo.
- Ruido frecuencial.

Para eliminar el ruido se utilizan los filtros espaciales. Los utilizados en este trabajo son [9]:

- Filtros de suavizado: para eliminar pequeños detalles antes de la extracción de un objeto, rellenar espacios vacíos y eliminar ruido. Se clasifican como lineales y no lineales. Para este trabajo se han utilizado los siguientes filtros lineales:
 - o Filtro de la media: se reemplaza el valor de cada píxel por la media de los valores de los píxeles vecinos.
 - o Filtro gaussiano: modeliza la función gaussiana [9], para disminuir nitidez, aumentar borrosidad o perder detalles.
- Filtros de realce: para realzar los detalles de una imagen, por tanto, están asociados con la detección de bordes.

2.5.2.3 Umbralización

La umbralización es otro de los conceptos esenciales que abarca la segmentación de imágenes. Supóngase que el histograma mostrado en la Figura 6 se corresponde con una imagen, $f(x, y)$, compuesta por objetos claros en un fondo oscuro, donde los píxeles tienen niveles de intensidad agrupados en dos modos dominantes.

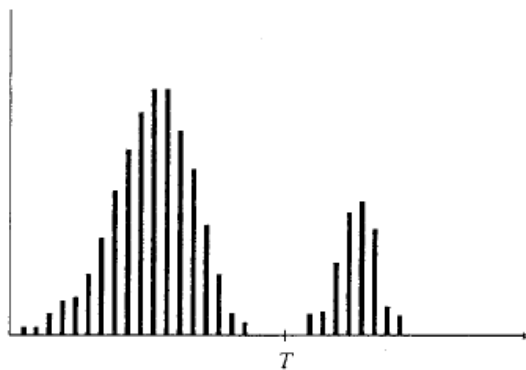


Figura 6: La T indica el umbral elegido en el histograma [7].

Seleccionando el umbral T se podrían separar los dos modos. De esta forma, cualquier punto (x, y) que cumpla con la condición $f(x, y) \geq T$ se correspondería con un píxel del objeto (valor binario 1), y en caso contrario con uno del fondo (valor binario 0). En otras palabras, la imagen binaria $g(x, y)$ se definiría como:

$$g(x, y) = \begin{cases} 0, & \text{si } f(x, y) < T \\ 1, & \text{si } f(x, y) \geq T \end{cases} \quad (8)$$

Con el fin de automatizar la selección del umbral (o umbrales) y que este varíe en función de las propiedades de la imagen, se utiliza el método de Otsu [10].

Cuando se trata de calcular un umbral global en la imagen, este método, comienza normalizando el histograma para tratarlo como una función de densidad de probabilidad discreta:

$$p_r(r_q) = \frac{n_q}{n}; \quad q = 0, 1, 2, \dots, L - 1 \quad (9)$$

siendo n el número total de píxeles en la imagen, n_q el número de píxeles que tienen nivel de intensidad r_q , y L es el número total de niveles de intensidad posibles en la imagen, suponiendo que el umbral k se elige de manera que C_0 es el conjunto de píxeles con niveles de intensidad en $[0, 1, \dots, k - 1]$ y C_1 es el conjunto de píxeles con niveles de intensidad en $[k, k + 1, \dots, L - 1]$.

El método de Otsu es un procedimiento en el que se selecciona el umbral óptimo k que maximiza la varianza entre clases $\sigma_B^2(k)$, que viene definida como:

$$\sigma_B^2(k) = \omega_0(\mu_0 - \mu_r)^2 + \omega_1(\mu_1 - \mu_r)^2 \quad (10)$$

siendo ω_0 y ω_1 las distribuciones de probabilidad de ambas clases, μ_0 y μ_1 las medias de las clases y μ_r la intensidad media total de la imagen [7].

A continuación, se dispersan lo máximo posible los niveles de intensidad de las dos clases mediante la maximización de la varianza entre clases σ_B^2 , y se obtiene el valor del umbral que maximiza dicha varianza.

$$\mu = \frac{\max \sigma_B^2(k)}{\sigma_G^2} \quad (11)$$

siendo σ_G^2 la varianza global de la imagen y μ el umbral que se obtiene como respuesta, cuyo valor está entre 0 y 1.

El método de Otsu puede ser extendido a un número arbitrario de clases. Por ejemplo, para tres clases, los dos umbrales serían k_1 y k_2 los cuales maximizarían la varianza entre las clases $\sigma_B^2(k_1, k_2)$. La imagen segmentada vendría dada por:

$$g(x, y) = \begin{cases} a, & \text{si } f(x, y) \leq k_1 \\ b, & \text{si } k_1 < f(x, y) \leq k_2 \\ c, & \text{si } f(x, y) > k_2 \end{cases} \quad (12)$$

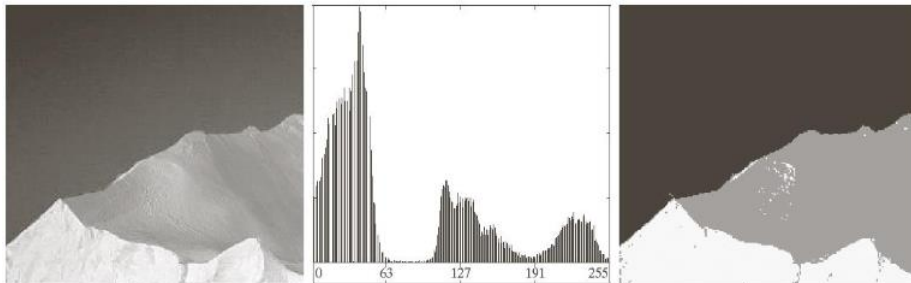


Figura 7: A la izquierda, imagen de un iceberg. En el centro, su histograma. A la derecha, imagen segmentada con dos umbrales y tres clases, siguiendo el método de Otsu [7].

2.5.2.4 Operaciones morfológicas

Las operaciones morfológicas son útiles para aumentar, reducir o eliminar en la imagen algunos objetos o regiones que tengan alguna forma y tamaño particular [11]. Estas operaciones dependen de un elemento estructurante, que se encarga de limitar el área en el cual se realiza la operación morfológica. En MATLAB, por ejemplo, este puede adoptar formas como un cuadrado, un rombo o un rectángulo, entre otros [12].

Las operaciones morfológicas que se consideran en este trabajo son dilatación, erosión, apertura y cierre de la imagen [7]:

- En el caso de la dilatación, el valor de salida de un píxel es el máximo valor de todos los píxeles en la vecindad del píxel en la imagen de entrada, donde la vecindad está definida por el elemento estructural. Esto quiere decir que agranda o engrosa los objetos de una imagen binaria.

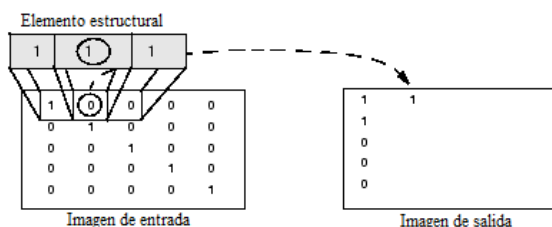


Figura 8: Ejemplo de dilatación [7].

- Para la erosión el procedimiento es similar [7], sin embargo, en este caso el valor de salida es el mínimo de todos los píxeles de dentro de la vecindad. Su función es la de encoger o afinar objetos en una imagen.
- La apertura y el cierre pueden ser definidos como:

$$f \circ b = (f \ominus b) \oplus b \quad (13)$$

$$f \bullet b = (f \oplus b) \ominus b \quad (14)$$

siendo apertura la ecuación superior y cierre la inferior, y donde:

$$\begin{aligned} f \ominus b &\rightarrow \text{Erosión} \\ f \oplus b &\rightarrow \text{Dilatación} \end{aligned}$$

donde f es la imagen original y b el elemento estructurante.

De esta forma, la apertura y cierre consisten en la aplicación de erosiones y dilataciones consecutivas. La primera se utiliza para suavizar contornos, eliminar pequeñas protuberancias o romper conexiones débiles y la segunda para rellenar detalles o conectar objetos próximos entre sí.

Capítulo 3

Estado del arte en sistemas de apoyo al diagnóstico de distrofias musculares

En este capítulo se estudian sistemas dedicados al estudio de células musculares, cuyas características son de interés para la realización de este trabajo. Se muestran los trabajos desarrollados a partir de 2007 hasta hoy.

Es posible dividir los sistemas revisados en dos clases: sistemas que segmentan imágenes de biopsias musculares y realizan la cuantificación de fibras y sistemas que realizan lo anterior, así como el cálculo de diámetros celulares.

3.1 Sistemas que segmentan imágenes de biopsias musculares y cuantifican fibras

En 2007, Tzekis et al. [13] presentaron un sistema implementado en lenguaje de programación C para la segmentación de imágenes de biopsias y cuantificación de fibras. La tinción que utilizan para las muestras es de Hematoxilina-Eosina. Esta, también llamada H&E-Stain, colorea de rosa fuerte el músculo, de rosa pálido el colágeno del endomisio y en azul oscuro la cromatina de los núcleos [14].

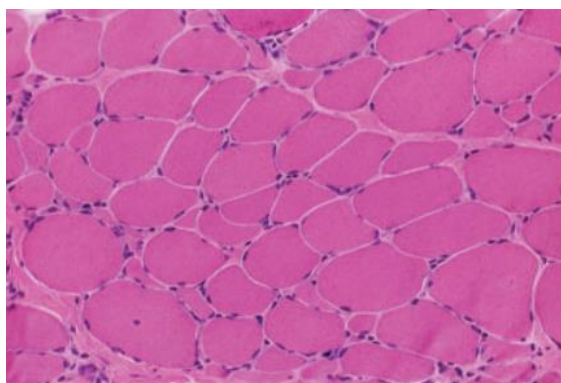


Figura 9: Imagen de biopsia muscular con H&E-Stain [13].

Como se puede observar en la Figura 9, en este tipo de tinción no es apreciable la diferencia entre los dos tipos de células (Tipo I y Tipo II).

En este artículo, Tzekis et al. comienzan comparando la detección de bordes mediante filtros (Sobel y Laplaciano) y un algoritmo basado en el color RGB de cada píxel. Este asigna blanco o negro a cada píxel, teniendo en cuenta su cercanía al color RGB modal

de la imagen. Posteriormente combinan los resultados de cada segmentación para obtener una mejorada.

El equipo de Sertel et al. propuso en 2011 [15] un sistema para el análisis cuantitativo de fibras musculares con el objetivo de caracterizar la composición de las mismas. Su sistema se basa en aplicar varias tinciones de ATPasa (con $\text{pH}=4.5$ y $\text{pH}=10.4$) a una misma muestra (ver Figura 10).

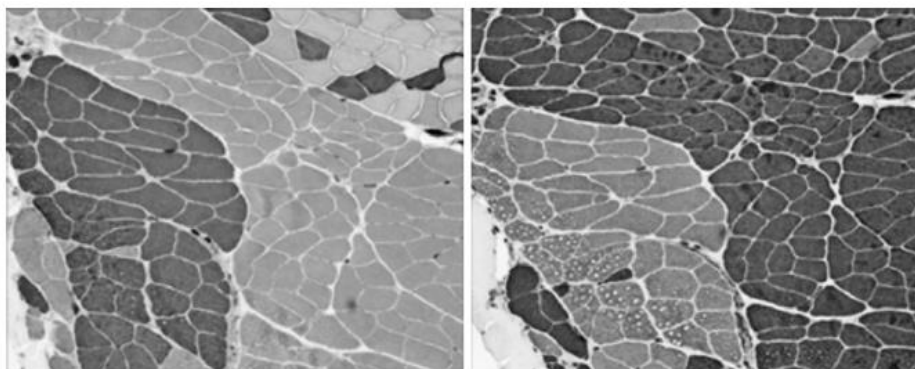


Figura 10: Misma muestra con ATPasa $\text{pH}=4.5$ y $\text{pH}=10.4$ respectivamente [15].

A partir de imágenes de la muestra teñida, obtienen la segmentación y cuantificación siguiendo el diagrama de flujo de la Figura 11. En la clasificación de los tipos de fibras también realizan la cuantificación de las mismas.

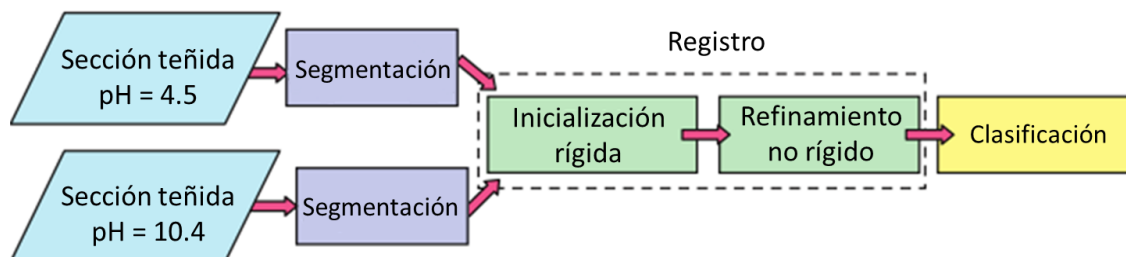


Figura 11: Diagrama del sistema propuesto por Sertel et al. [15].

En 2013, Liu et al. [16] propusieron un algoritmo de segmentación de imágenes de biopsias musculares teñidas con H&E-Stain (ver Figura 9). Su método se lleva a cabo en dos pasos:

- 1) Detección de semillas para encontrar automáticamente los centros geométricos de las fibras musculares.
- 2) Utilización del modelo de activación del contorno o snakes [7], que adopta formas iterativamente en función del gradiente de color hasta obtener la forma deseada.

Una vez obtienen la imagen segmentada plantean una cuantificación automática del número de fibras, pero sin hacer distinción entre los tipos.

Paralelamente, en 2013, Pertl et al. en [17] estudiaron la distrofia muscular de Duchenne en ratones. Practicaron a las muestras tres tipos de tinción (ver Figura 12).

La primera es de Myosin Heavy Chain (MHC) y las otras dos son tinciones fluorescentes verde y azul, respectivamente.

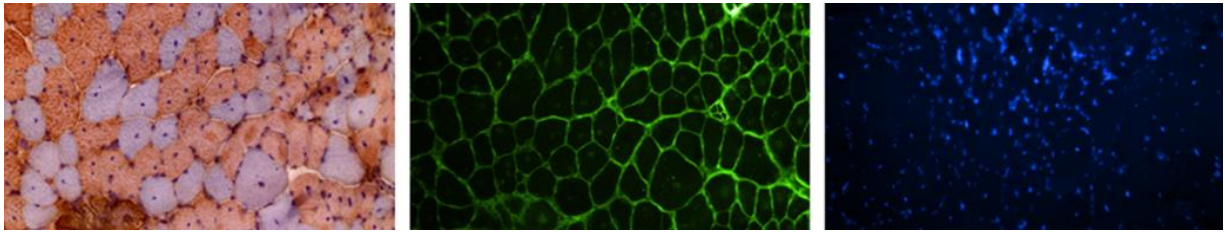


Figura 12: Misma muestra de ratón con tres tinciones distintas [17].

En su artículo, Pertl et al. superponen las tres imágenes con el software Cognition Network Technology [18]. De la imagen resultante, extraen automáticamente las estructuras relevantes como objetos individuales (fibras) y asignan un ID a cada uno, mediante el software web S.CORE [19]. Destacan en su artículo la importancia de que la muestra no posea pliegues, artefactos o tinción pálida que impida distinguir los límites de las fibras.

También en 2013, Sáez et al. [20] desarrollaron NDICIA (Neuromuscular Diseases Computerized Image Analysis). Este parte de una muestra con tinción inmunohistoquímica, la cual diferencia ambos tipos de células musculares, así como el endomisio que las separa, mediante fluorescencia.

Extraen el canal verde de la imagen RGB, debido al alto contraste entre células y endomisio. Como las fibras son más oscuras que el espacio que las separa, aplican la transformada H-Minima [7] buscando valles de intensidad. Esta transformada elimina los mínimos de intensidad de la imagen cuyo valor sea menor de un escalar dado.

Por otro lado, aplican el algoritmo de Watershed [7]. Este les da resultados con sobre-segmentación. Aprovecharon esta sobre-segmentación para crear una red en la que cada fibra se corresponde con un nodo y dos nodos están conectados si se trata de células adyacentes, como se observa en Figura 13.

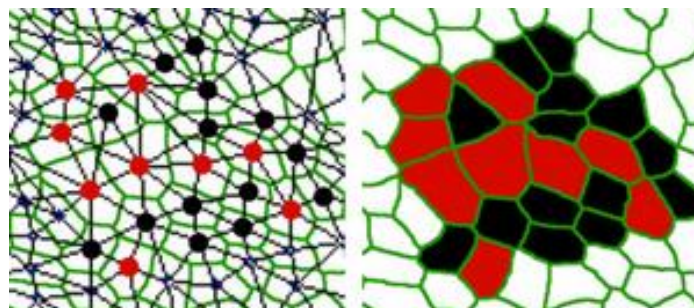


Figura 13: Distinción de células musculares con NDICIA y su red [20].

3.2 Sistemas que segmentan imágenes de biopsias musculares, cuantifican fibras y calculan diámetros celulares

En 2007, Kim et al. [21] presentaron una técnica de segmentación automática, la cual parte de imágenes con tinción H&E-Stain (ver Figura 9), por tanto, no permite diferenciación entre tipos de células.

El algoritmo que utilizan en su artículo realiza una segmentación de bordes basada en activación del contorno, extendida con segmentación basada en regiones. Esta extensión permite explotar tres fuentes de información adicionales: el color, la textura y los bordes.

Para la medida de los diámetros de las fibras musculares, propusieron utilizar el software ImageJ [22], que es un programa de procesamiento de imagen de dominio público. La medida la realizan mediante una rutina incorporada en dicho software.

El equipo de Karen et al., en 2009, propuso un software para la clasificación y análisis de fibras musculares [23]. El software de Karen et al. se divide en dos programas “parciales” desarrollados en C++:

- El primero, MuscRegM, para segmentar la imagen con tinción de MHC (sí se diferencian los tipos de fibras). El usuario previamente debe marcar de 4 a 8 puntos de control de forma manual.
- Es posible que sea necesario un paso intermedio para marcar manualmente bordes mal segmentados. Para ello proponen la utilización del software Ellipse, desarrollado por ViDiTo [24]. En sus resultados exponen que no llegaron a métodos de segmentación concretos; implementaron activación del contorno y “Region Growing”, que consiste en el agrandamiento de un grupo de píxeles mediante un criterio predefinido, pero ambos necesitaban posterior corrección manual.
- El segundo, FibClasM, realiza la diferenciación de tipos de células, las cuantifica y calcula sus diámetros.

En el trabajo de Dubach-Powell [25], revisado por última vez en 2014, se propone un método de segmentación automática con tiempo de computación de 30 minutos cada 2000 fibras. La tinción de la imagen de la biopsia de la que parte es fluorescente, como se observa en la Figura 14.

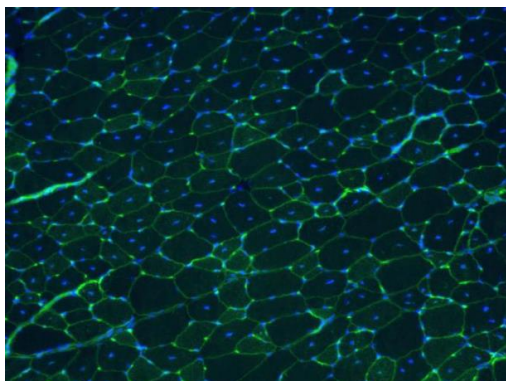


Figura 14: Tinción fluorescente del artículo de Dubach-Powell [25].

Para realizar la segmentación, primero toma el canal verde de la imagen RGB. Con el software CellP [26], utiliza el “separator filter” del programa para segmentar la imagen. Una vez obtiene la imagen segmentada, se da una etiqueta a cada célula de forma individual para cuantificarlas. Posteriormente, es posible realizar el cálculo automático de cada diámetro con el mismo programa.

Por último, en 2014, Smith y Barton lanzaron la herramienta semi-automática con interfaz gráfica creada con MATLAB: SMASH [27].

SMASH segmenta imágenes con tinción fluorescente. El usuario debe seleccionar manualmente los filtros, funciones y parámetros de segmentación a utilizar. Tras esto, el programa devuelve la imagen segmentada mediante Watershed. Es necesario corregir manualmente la segmentación y marcar los bordes que no estén marcados.

Tras la segmentación, SMASH da la posibilidad de cuantificar las células, obtener los diámetros de las mismas y mostrarlos en un histograma.

Tras la revisión del estado del arte en sistemas de apoyo al diagnóstico de distrofias musculares, se observa que ninguno de los sistemas se adapta a los requisitos del Servicio de Anatomía Patológica del C.H.U Insular – Materno Infantil de Las Palmas de Gran Canaria. Las razones son las siguientes:

- No utilizan la misma tinción que la de las muestras del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil de Las Palmas de Gran Canaria. Algunos trabajos revisados incluso parten de varias tinciones sobre la misma muestra para obtener una segmentación completa, lo cual incrementa el precio y el tiempo.
- Muchos no unifican las tareas en un mismo software, sino que recomiendan el uso de varios paquetes software, lo cual al final puede ser tedioso, además de incrementar el precio.
- Ninguno muestra tiempos de computación en segmentación y cálculo de diámetros.

- No expresan resultados concretos que ofrecer al patólogo, como informes con la toda información reunida, por ejemplo.

Por todo ello, se propone unificar las tareas en un único entorno para simplificar la labor del patólogo, reducir el coste y los tiempos de obtención de resultados.

Capítulo 4

Segmentación de imágenes de biopsias musculares

En los capítulos 2 y 3 se han mostrado los fundamentos teóricos en los que se basa este trabajo y se ha hecho un repaso del estado del arte en sistemas de ayuda al diagnóstico de distrofias musculares. En este capítulo se muestra cómo se ha llevado a cabo la segmentación de las imágenes de muestras de músculo que han sido recibidas por parte de los patólogos del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil.

En primer lugar, se explica cuáles son las características que deben tener las imágenes para un análisis óptimo por parte del software. En segundo lugar, se muestran cuatro de las imágenes del Servicio de Anatomía Patológica que han sido analizadas para demostrar las bondades de los algoritmos. En tercer lugar, se mostrarán dos algoritmos desarrollados en base a dos métodos de segmentación conocidos: Watershed y K-means. En cuarto lugar, se propone un algoritmo novedoso, desarrollado en este trabajo, como solución a la segmentación. Por último, se presentan los resultados de la segmentación mediante los tres métodos mencionados en las cuatro imágenes de referencia.

Todas las implementaciones de los algoritmos han sido plenamente desarrolladas con MATLAB.

4.1 Características de la imagen de la biopsia

En esta sección se describe qué características deben poseer las imágenes de las biopsias musculares para poder ser procesadas de forma eficiente por el software desarrollado en este trabajo.

4.1.1 Corte histológico

La importancia de la muestra histológica de músculo esquelético reside en el estudio individual del tamaño y forma de cada fibra. Resulta trascendental que el corte histológico se realice de forma correcta, para obtener así una sección transversal correcta del músculo. La organización estructural de las fibras, mostrada en el Capítulo 2, puede provocar que durante el seccionamiento se produzcan pliegues en las células y los límites de estas se confundan al observar la muestra en el microscopio.

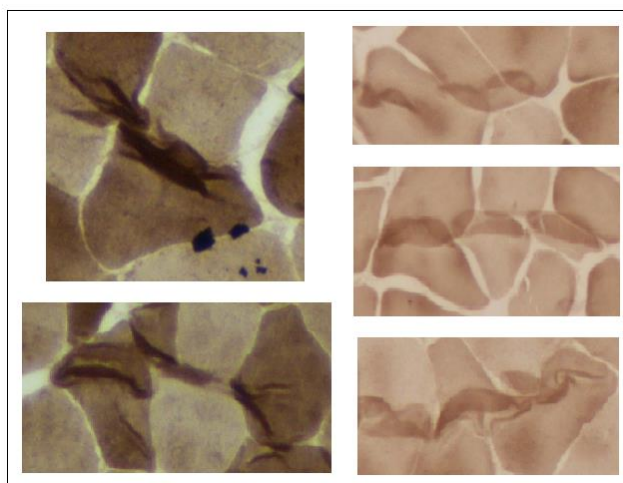


Figura 15: Ejemplos de pliegues en muestras de biopsias musculares.

Como se puede observar en la Figura 15, la aparición de pliegues en la muestra, provoca solape entre células y dificulta el tratamiento tanto por un software automático como manualmente por el patólogo.

4.1.2 Tinción de la muestra

La técnica de tinción utilizada en las muestras procesadas en este trabajo es la que emplean en el laboratorio del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil de Las Palmas de Gran Canaria. Esta es la técnica de la ATPasa A con $\text{pH}=9.4$ (ver sección 2.3).

La ATPasa A muestra las células Tipo I de color claro y las Tipo II de color oscuro (ver Figura 2). Es importante realizar la tinción de forma cuidadosa, así como la captura de las imágenes con el microscopio digital rápidamente, antes de que se pierda la tonalidad de la tinción con el paso de los días, debido a la pérdida de la actividad enzimática en las fibras. En la Figura 16, se observan muestras con tinción débil en las cuales resulta difícil distinguir entre los dos tipos de células, frente a muestras con mejor tinción.

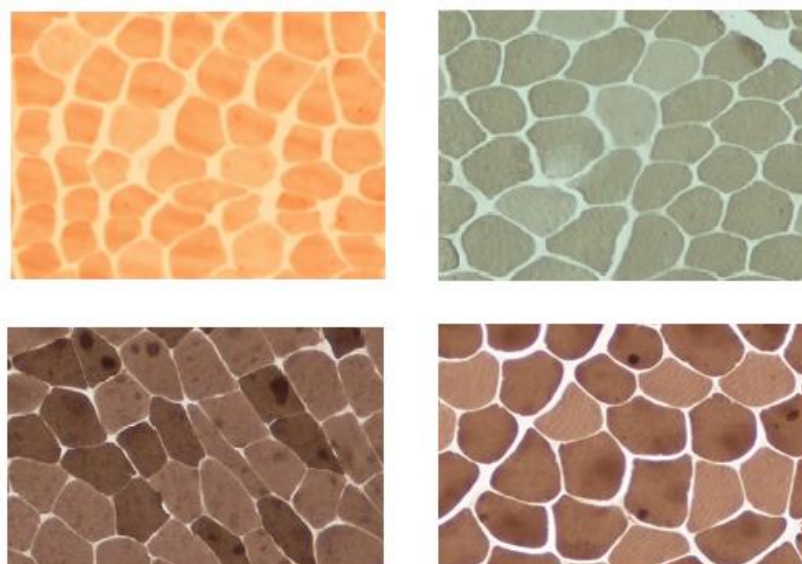


Figura 16: Muestras con tinción débil (superiores) frente a muestras con buena tinción (inferiores).

4.1.3 Captura de la imagen de la biopsia muscular

Para realizar la captura de la imagen de la biopsia muscular y que esta después pueda ser procesada por el software, es necesario emplear un microscopio digital. En el Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil de Las Palmas de Gran Canaria utilizan el modelo de NIKON Eclipse 80i, el cual está conectado a una cámara Digital Sight DS-2Mv y un ordenador NIKON Digital Sight DS-L2.

En la etapa inicial se comenzó pidiendo al patólogo capturas de imágenes de muestras con aumento x40 (Figura 17 izquierda), pero dada la calidad deficiente y los malos resultados obtenidos, se optó por el aumento inmediatamente posterior, el de x100 (Figura 17 derecha).

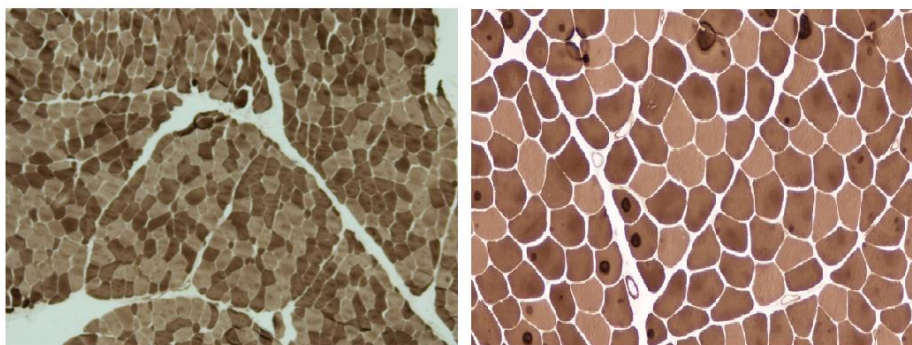


Figura 17: Ejemplo comparativo de la definición de células con aumento x40 y x100. Izquierda x40, derecha x100.

Dado que con una captura con aumento x100 no se alcanza a observar la muestra entera, y recordando que se necesita realizar la cuantificación de las células de la muestra, así como la medida de 100 diámetros aleatorios de cada tipo de fibra (200 en total), es necesario que el patólogo realice un barrido de capturas sobre la muestra. Es preciso recalcar que cada captura debe contener un solape cercano al 20% entre distintas imágenes para, de esta forma, componer todas en una única imagen.

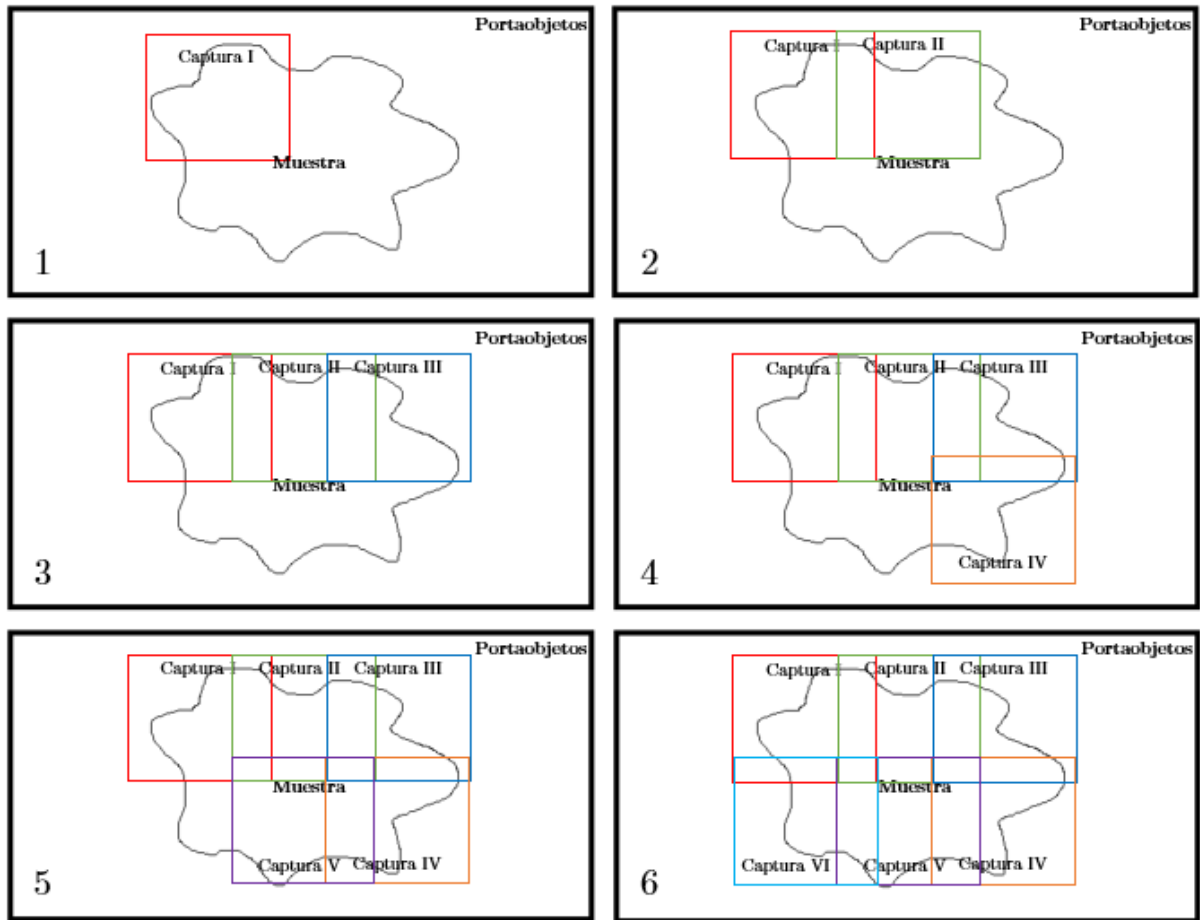


Figura 18: Ejemplo de barrido de capturas con el microscopio sobre la muestra (aumento x100). Orden: de izquierda a derecha y de arriba abajo (de 1 a 6).

En la Figura 18 se observa un ejemplo de barrido de capturas sobre una muestra. Para obtener una imagen completa de la misma, es posible aplicar la técnica “Image Stitching” [28], que es el proceso de combinar imágenes que contienen partes superpuestas para conseguir una imagen panorámica. Para el caso anterior se obtendría una imagen completa de la muestra como la que se observa en la Figura 19.

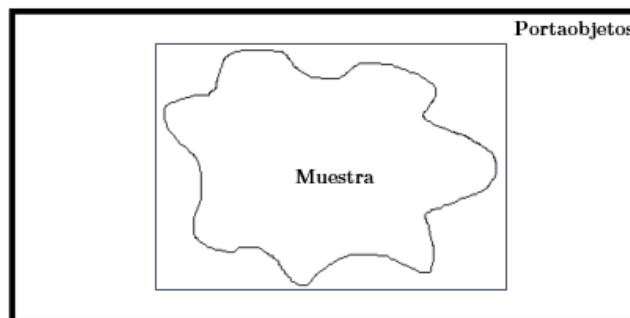


Figura 19: Imagen completa tras aplicar Image Stitching.

Por último, cabe comentar que el enfoque con el microscopio, antes de tomar la captura, debe ser preciso para evitar que esta quede borrosa o con efecto de movimiento.

4.2 Imágenes analizadas

Durante los meses de noviembre de 2015 y mayo de 2016 se recibieron nueve imágenes completas de muestras por parte del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil de Las Palmas de Gran Canaria. Algunas de las imágenes recibidas cumplen de forma precisa las especificaciones nombradas en la sección anterior. Sin embargo, otras de ellas no lo hacen. En ellas se encuentran numerosas células solapadas debido a un corte histológico incorrecto, o no se distinguen células debido a una tinción débil. No obstante, estas resultaron positivas para proveer de mayor robustez a los algoritmos desarrollados en este trabajo y que estos funcionen lo mejor posible en todos los casos.

Las imágenes recibidas cuentan con una resolución tanto horizontal como vertical de 300 ppp (puntos por pulgada), en el espacio de representación RGB. Las dimensiones de las imágenes son variables, dependiendo del tamaño de la muestra o del número de capturas realizadas por el patólogo.

Tras tratar todas las imágenes, se ha observado que es posible dividir las en 4 grupos. Su pertenencia a estos grupos se define en función de:

- El buen corte histológico realizado, dando lugar a una sección transversal de músculo sin pliegues.
- La correcta tinción que posean (fuerte o débil).

Cuanto más alejadas de las características idóneas se encuentren las imágenes, mayor dificultad ofrece a la actuación del software, tanto para segmentar la imagen como para calcular posteriormente los parámetros necesarios. La Tabla 3 muestra esta clasificación:

Grupo de imágenes	Sección transversal sin pliegues	Tinción correcta
I	✓	✓
II	✓	✗
III	✗	✓
IV	✗	✗

Tabla 3: Clasificación de las imágenes recibidas.

La marca ✓ indica que el grupo posee esa característica y ✗ indica lo contrario.

Dada la clasificación mostrada, se han seleccionado 4 imágenes características, una en representación de cada grupo. Los resultados que se muestran en secciones posteriores se han obtenido para dichas imágenes.

A continuación, se describen las imágenes seleccionadas.

Imagen I: Sección transversal sin pliegues y tinción correcta

Esta imagen posee una tinción correcta (se diferencian de forma favorable los tipos de fibras) y el corte fue realizado sin dar lugar a pliegues.

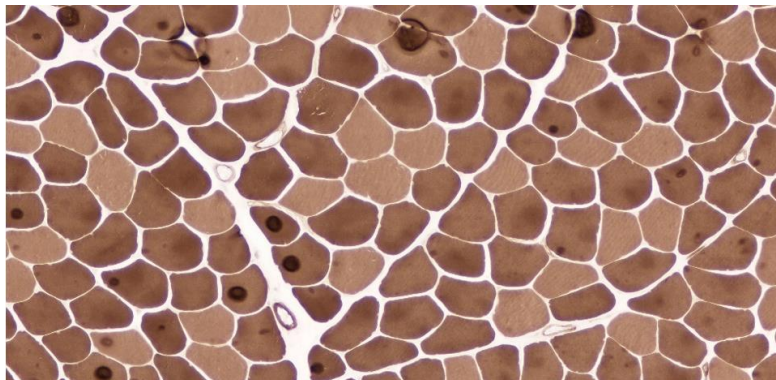


Figura 20: Imagen I: Sección transversal sin pliegues y tinción correcta (grupo I).

Imagen II: Sección transversal sin pliegues y tinción débil

En la Imagen II el corte fue realizado sin producir pliegues en la muestra, pero la tinción fue algo débil.



Figura 21: Imagen II: Sección transversal sin pliegues y tinción débil (grupo II).

Imagen III: Sección transversal con pliegues y tinción correcta

La Imagen III posee una tinción correcta. Sin embargo, el corte produjo pliegues que producen el solape entre grupos de células.

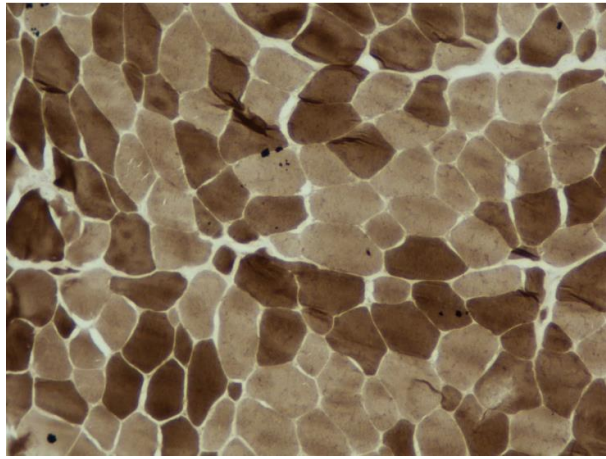


Figura 22: Imagen III: Sección transversal con pliegues y tinción correcta (grupo III).

Imagen IV: Sección transversal con pliegues y tinción débil

Por último, en la Imagen IV, tanto la tinción como el corte fueron incorrectos, dando lugar a multitud de confusiones entre células debido al solapamiento. La tinción además ha producido cambios de intensidad en el interior de las células, lo que dificulta aún más su clasificación.

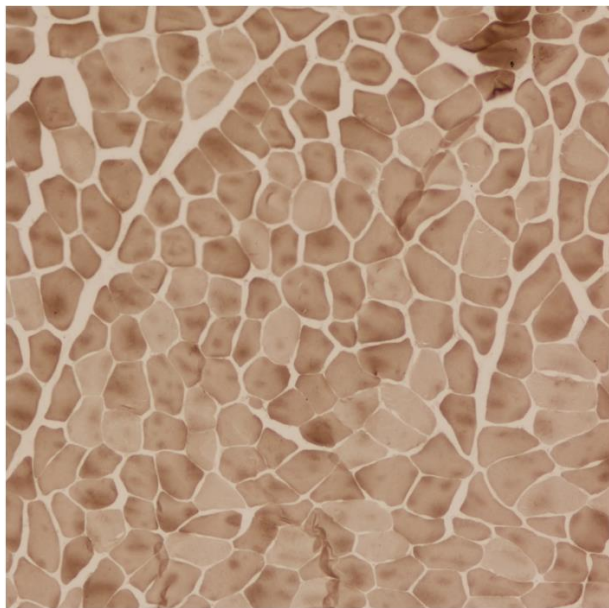


Figura 23: Imagen IV: Sección transversal con pliegues y tinción débil (grupo IV).

4.3 Métodos analizados

En esta sección se muestra el desarrollo de dos métodos de segmentación de imágenes sobre escala de grises como posible solución a la segmentación de imágenes de muestras musculares. En primer lugar, se ha analizado la estrategia de segmentación basada en la Transformada de Watershed, que fragmenta la imagen con líneas divisorias en función de determinadas características. El segundo método analizado aprovecha un clasificador no supervisado que divide la imagen en un determinado número de grupos en función del color, denominado K-means, para segmentar la imagen.

4.3.1 Transformada de Watershed

La palabra Watershed, que en español significa “línea de división de aguas”, da una pista sobre el funcionamiento de esta técnica de segmentación. Una cuenca hidrográfica es el área geográfica que drena un río o embalse. La Transformada de Watershed aplica esta idea a la escala de grises del procesamiento digital de imágenes de forma que pueda ser útil en la segmentación.

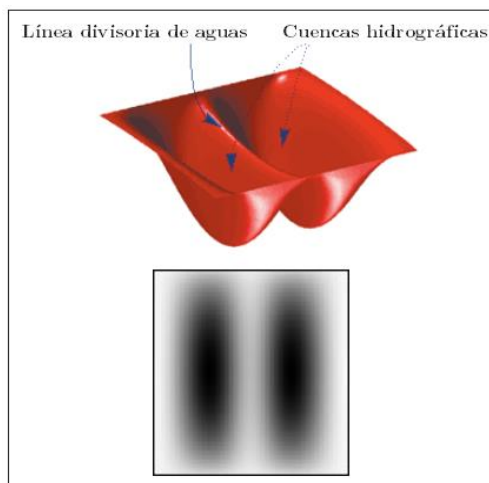


Figura 24: Diagrama de Watershed [7].

La Figura 24 da una visión de la base de funcionamiento de la Transformada de Watershed. En dicha figura, las áreas “altas” están representadas con mayor claridad y las más “bajas” con zonas más oscuras.

Imaginando la imagen en escala de grises como una superficie topológica, la división se realiza mediante la asignación de pesos a los píxeles de la imagen en función de su valor en la escala de gris. El peso que se le asigna al píxel depende de su textura, intensidad y color.

El proceso de segmentación morfológica suele partir del gradiente (ver Figura 5, ejemplo de gradiente) y es a este al que se le aplica el algoritmo de Watershed. Se aplica Watershed al gradiente y no a la imagen original en escala de grises porque lo que se pretende es identificar el contorno mediante mínimos globales. El resultado en imágenes reales suele conducir al fenómeno de sobre-segmentación (aparición de demasiadas líneas divisorias), por la presencia de mínimos locales en las imágenes.

Para aprovechar Watershed y conseguir segmentar imágenes de biopsias musculares, se ha evaluado la segmentación Watershed controlada por marcadores [8] ya que reduce el efecto de sobre-segmentación. Se ha desarrollado un algoritmo que obtiene los mejores resultados posibles manteniendo un compromiso entre todas las imágenes recibidas desde el Servicio de Anatomía Patológica. Dado que se busca automatizar el sistema, es necesario que el algoritmo resulte lo más eficiente posible para todos los tipos de imágenes.

El diagrama de la Figura 25 muestra el procedimiento desarrollado en este trabajo para obtener segmentación mediante Watershed controlado por marcadores. Las imágenes que acompañan al algoritmo parten de una muestra real de músculo esquelético y se exponen a modo descriptivo para apoyar al diagrama.

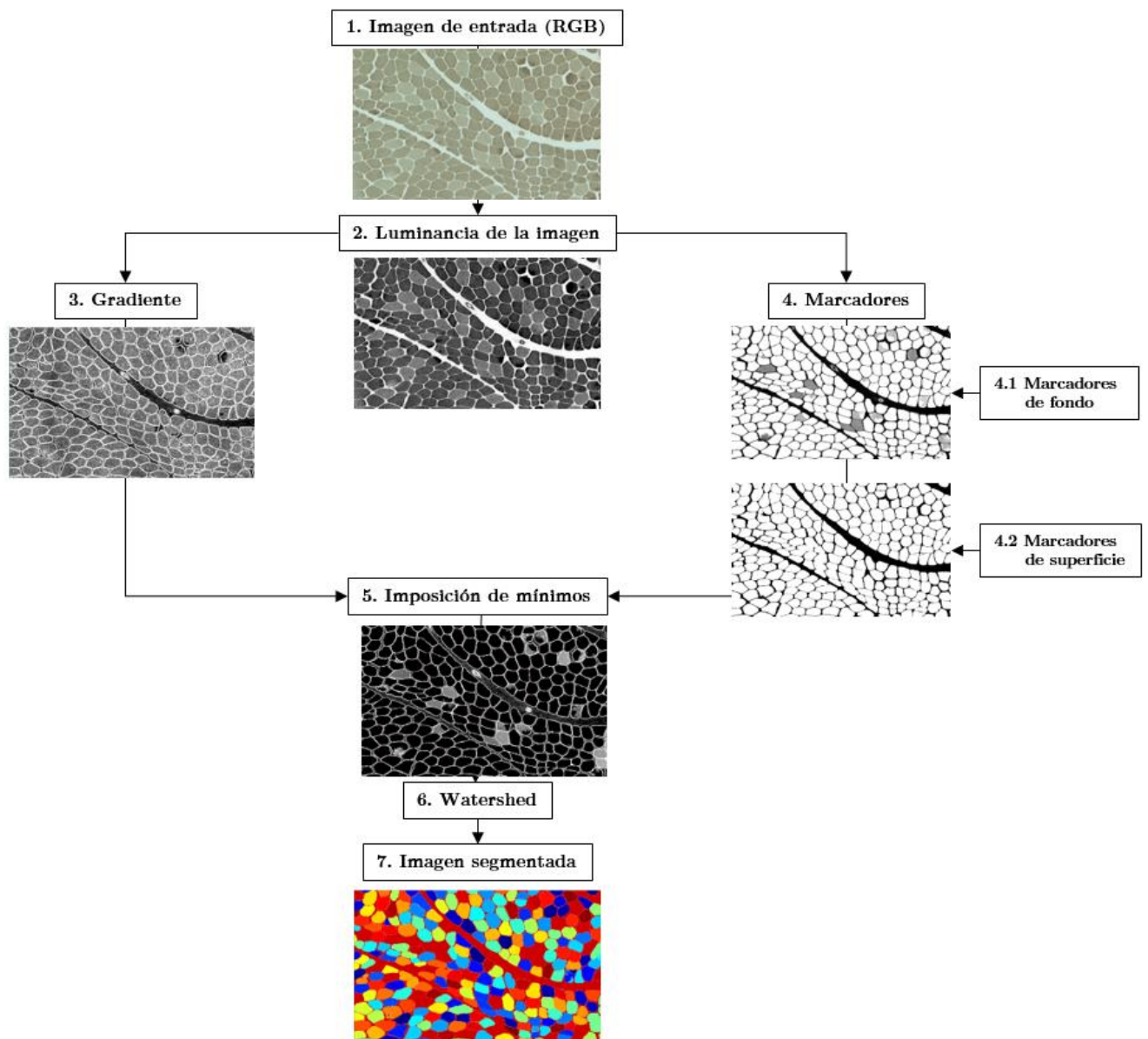


Figura 25: Diagrama de la implementación desarrollada con el algoritmo de Watershed controlado por marcadores.

Los pasos de funcionamiento que sigue el algoritmo son los siguientes:

1. La imagen de entrada es una de las recibidas del Servicio de Anatomía Patológica, en espacio de color RGB.
2. Se obtiene la luminancia cambiando de RGB al espacio YCbCr como se indica en la sección 2.5. El canal de la luminancia ofrece mejor contraste en escala de grises para las imágenes de muestras de músculo. También se realiza un ajuste de los valores de intensidad de la imagen en escala de grises para aumentar aún más el contraste.

3. Por un lado, se calcula la magnitud del gradiente con el operador Sobel como estimador de la derivada, ya que este se ha considerado que es el que mejores resultados ofrece. El gradiente es alto en los bordes de las células y todo lo bajo posible en el interior. En el interior de las células el gradiente no es nulo debido a la aparición de píxeles cuyos cambios de intensidad provocan un aumento de la frecuencia espacial. En MATLAB, el cálculo del gradiente se representa en una imagen en escala de grises con valores tipo “double” en el intervalo $[0,1]$, donde el 0 indica el máximo gradiente (color blanco) y el 1 indica mínimo gradiente (negro).
4. Por otro lado, se generan marcadores de primer plano y de fondo:
 - o Los primeros consisten en separar los grupos de píxeles que se encuentren dentro de células para marcar el primer plano. Esto se lleva a cabo mediante técnicas morfológicas llamadas “apertura y reconstrucción” y “cierre y reconstrucción”.

La diferencia entre la apertura vista en la sección 2.5 y la “apertura y reconstrucción” está en que la primera consiste en una erosión seguida por una dilatación y la segunda en una erosión seguida de una reconstrucción morfológica. Igual ocurre con el caso del cierre y “cierre y reconstrucción”. En MATLAB existe la función “imreconstruct()”, que implementa la reconstrucción morfológica.
 - o Los segundos marcadores se generan para señalar el fondo de la imagen. Partiendo de los primeros marcadores, que indicaban los píxeles que pertenecían al interior de células, se impone que el resto de píxeles pertenecen al fondo de la imagen, mediante binarización. Para evitar un posible solape entre marcadores del fondo de la imagen con bordes de células, se estrecha el fondo. Para ello, se realiza una transformada Watershed auxiliar a la transformada de la distancia de la imagen binaria obtenida. En MATLAB, es “bwdist()” la función que computa la transformada de la distancia Euclídea, es decir, mide la distancia entre cada píxel cuyo valor es 0 y el píxel más cercano cuyo valor es 1.
5. Para la imposición de mínimos se utiliza la función de MATLAB “imimposemin()”. Esta, modifica la magnitud del gradiente para que los mínimos locales sólo ocurran en los píxeles que son marcadores.
6. Finalmente se aplica Watershed con la función “watershed()” de MATLAB.
7. Por último, se etiqueta cada área aislada (células en teoría) para su posterior cuantificación y clasificación. En el diagrama se utilizan etiquetas RGB para poder diferenciar visualmente unas células de otras. No obstante, el etiquetado utilizado en este trabajo para la cuantificación y el tratado individual de cada

célula se ha realizado sobre la imagen segmentada binaria. Las funciones que MATLAB aporta para el etiquetado son “bwlabel()” para etiquetar las áreas aisladas de la imagen binaria y la función “label2rgb()” para darle color.

Los resultados que se obtienen a partir de la utilización de este algoritmo se muestran en comparación con los demás en la sección 4.5.

4.3.2 Segmentación basada en color utilizando K-means

K-means es un algoritmo de agrupamiento de datos, y como tal, su objetivo es identificar “K” grupos o clusters dentro de los datos basándose en alguna medida de similitud [29]. Se trata de un algoritmo no supervisado habitualmente utilizado en procesamiento de imágenes para clasificación basada en color.

Inicialmente, se determina el número de grupos K. Posteriormente y de forma aleatoria, se determinan los K centroides de cada grupo. Luego, el algoritmo ejecuta los siguientes tres pasos hasta que alcance el criterio de convergencia, es decir, que los objetos no se muevan de grupo [30]:

1. Se determina el o los centroides iniciales de acuerdo al número de grupo esperado.

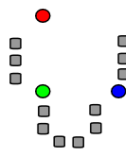


Figura 26: K centroides iniciales (K=3).

2. Se determina la distancia de cada objeto con relación a los centroides.

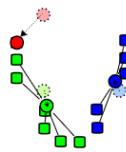


Figura 27: Distancia de cada objeto a los centroides.

3. Se agrupan los objetos en función de la distancia mínima. Se calculan nuevos centros y se repiten los pasos 2 y 3 hasta alcanzar convergencia.

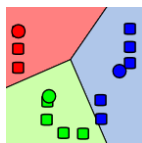


Figura 28: K grupos generados.

Por tanto, K-means busca minimizar la suma de las distancias entre todos los puntos del grupo y su centro.

En lo que concierne a este trabajo, se ha evaluado K-means como posible solución a la segmentación de los bordes de las células. El diagrama de la Figura 29 muestra el procedimiento seguido en el algoritmo desarrollado. En este diagrama se muestra cómo además de hacer uso del clasificador K-means, ha sido necesario implementar un procesado posterior para completar la segmentación. El diagrama se apoya en una imagen de un caso real a modo descriptivo.

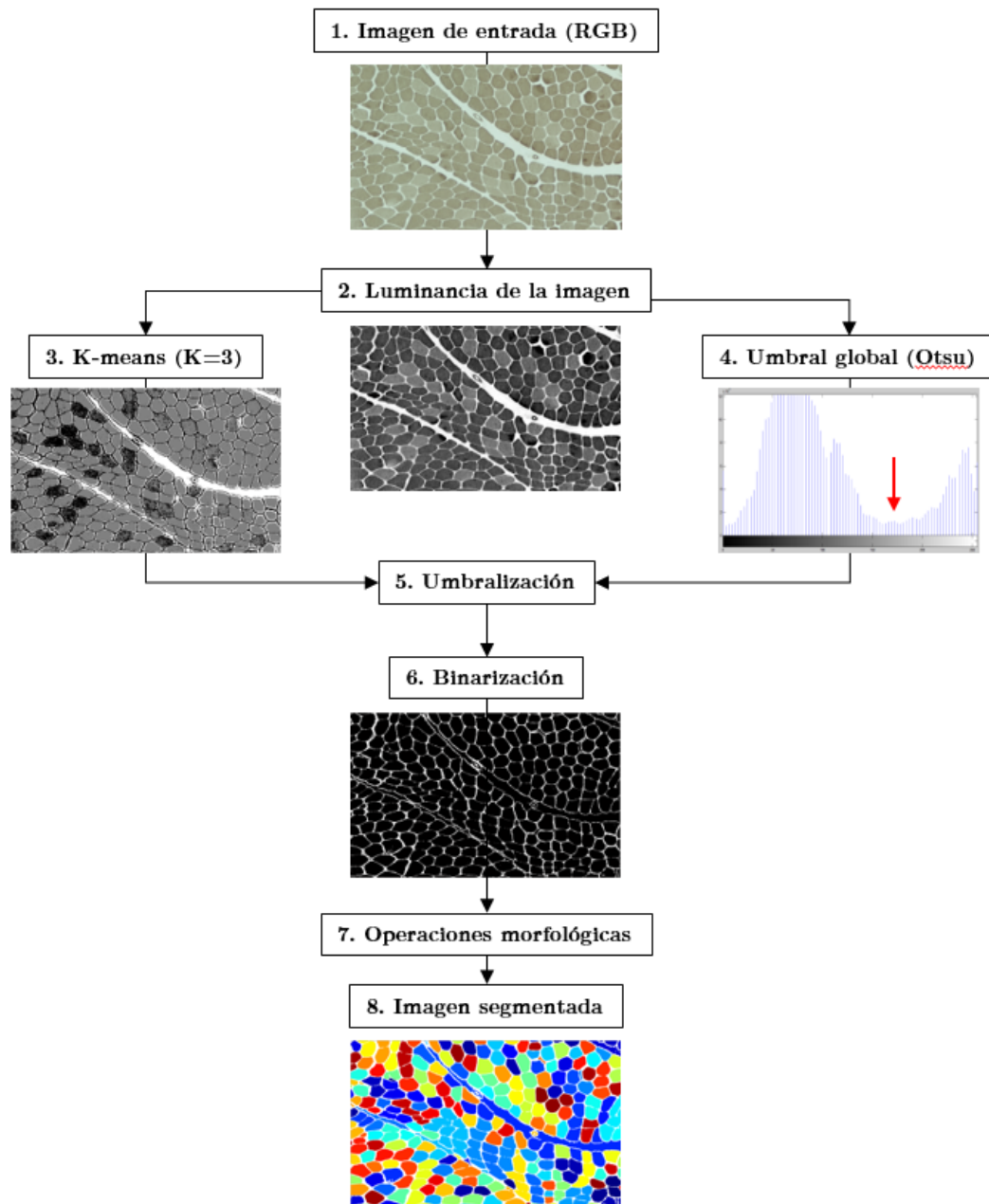


Figura 29: Diagrama de la segmentación basada en K-means.

Los pasos de funcionamiento que sigue el algoritmo son los siguientes:

1. La imagen de entrada es una de las recibidas del Servicio de Anatomía Patológica, en espacio de color RGB.
2. Se obtiene el canal de la luminancia cambiando de RGB al espacio YCbCr como se indica en la sección 2.5.

3. Para el óptimo funcionamiento de K-means, se ha decidido que el número de grupos a clasificar es $K=3$. La elección de K se debe a los tres tipos de elementos que contienen las imágenes a analizar, que son: células Tipo I, Tipo II y el endomisio que poseen a su alrededor. Dicho número se ha elegido también manteniendo un compromiso entre todas las imágenes, dando el mejor resultado posible para todas. Es importante recalcar que el número de grupos es fijo para todas las imágenes ya que se pretende automatizar el programa al máximo.

Una vez aplicado K-means, se obtiene una imagen que contiene las tres clases, el endomisio en blanco, las células Tipo II en gris (células oscuras en la imagen original) y las Tipo I (claras en la imagen original) en negro (ver Figura 30).

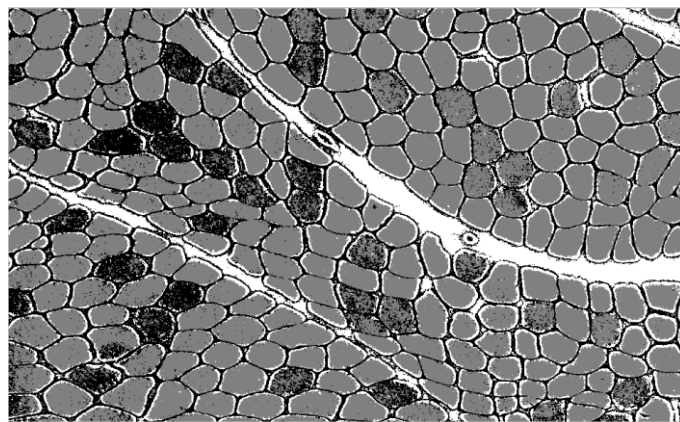


Figura 30: Ejemplo de clasificación K-means con $K=3$ para una de las imágenes del Servicio de Anatomía Patológica

- La clase blanca indica lo que el algoritmo ha detectado como zonas de endomisio, elemento que rodea las células.
- La clase gris indica las células oscuras (Tipo II). Se observa como en el interior de dichas células hay píxeles que se confunden con endomisio.
- La clase en negro indica las células claras (Tipo I). Es observable que esta clasificación es la que peor selecciona los píxeles correctos. Esto se debe a que habitualmente, en el interior de las células claras, existen píxeles demasiado claros que son confundidos con endomisio, y demasiado oscuros que son confundidos con píxeles de la clase de células oscuras.

En los siguientes pasos se busca mejorar la selección de las células claras para mejorar la segmentación. Para ello se selecciona la imagen que muestra en negro los objetos de la clase de las células Tipo II (ver Figura 31).

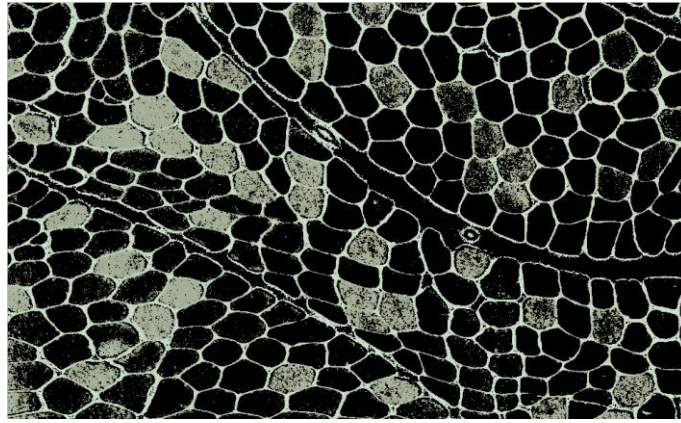


Figura 31: Imagen que muestra en negro los objetos del clúster correspondiente a las células Tipo II. Las Tipo II aparecen de color claro con píxeles en negro mezclados.

4. Por otro lado, se calcula el histograma de la imagen de la luminancia del paso 2 para, aplicando el método de Otsu descrito en la sección 2.5, calculado el umbral global de la imagen (ver flecha roja en el paso 4 del diagrama). Lo que queda a la izquierda de este umbral son píxeles pertenecientes a células mayoritariamente, y a la derecha lo que pertenece a endomisio, también mayoritariamente.
5. Una vez calculado el umbral global, se umbraliza la imagen seleccionada en el paso 3 (ver Figura 30). Para ello, se impone que los píxeles que pertenecen a células claras se conviertan en el color de su clase (negro en este caso) si el valor de ese píxel en la imagen de la luminancia es menor que el umbral calculado. Se obtiene así una mejor definición de las células claras.
6. Posteriormente, se binariza la imagen, quedando en blanco los bordes y en negro el fondo.
7. Mediante operaciones morfológicas se eliminan los pequeños grupos de píxeles conectados que no forman parte de los bordes. La función “`bwareaopen()`” de MATLAB es idónea para esta tarea, ya que elimina objetos pequeños en imágenes binarias.
8. Por último, se etiqueta cada área aislada (células en teoría) para su posterior cuantificación y clasificación. En el diagrama se utilizan etiquetas RGB para poder diferenciar visualmente unas células de otras. No obstante, el etiquetado utilizado en este trabajo para la cuantificación y el tratado individual de cada célula se ha realizado sobre la imagen segmentada binaria.

Los resultados que se obtienen a partir de la utilización de este algoritmo se muestran en comparación con los demás en la sección 4.5.

4.4 Aportaciones a la segmentación de muestras de biopsias musculares

Para alcanzar una segmentación más precisa de la imagen inicial, a fin de mejorar los resultados obtenidos con los dos algoritmos de segmentación explicados, se ha desarrollado un algoritmo comenzando desde cero. En este, se ha decido combinar mediante una operación OR bit a bit, las imágenes resultantes de dos detectores de bordes. El diagrama de la Figura 32 presenta el flujo que sigue este algoritmo.

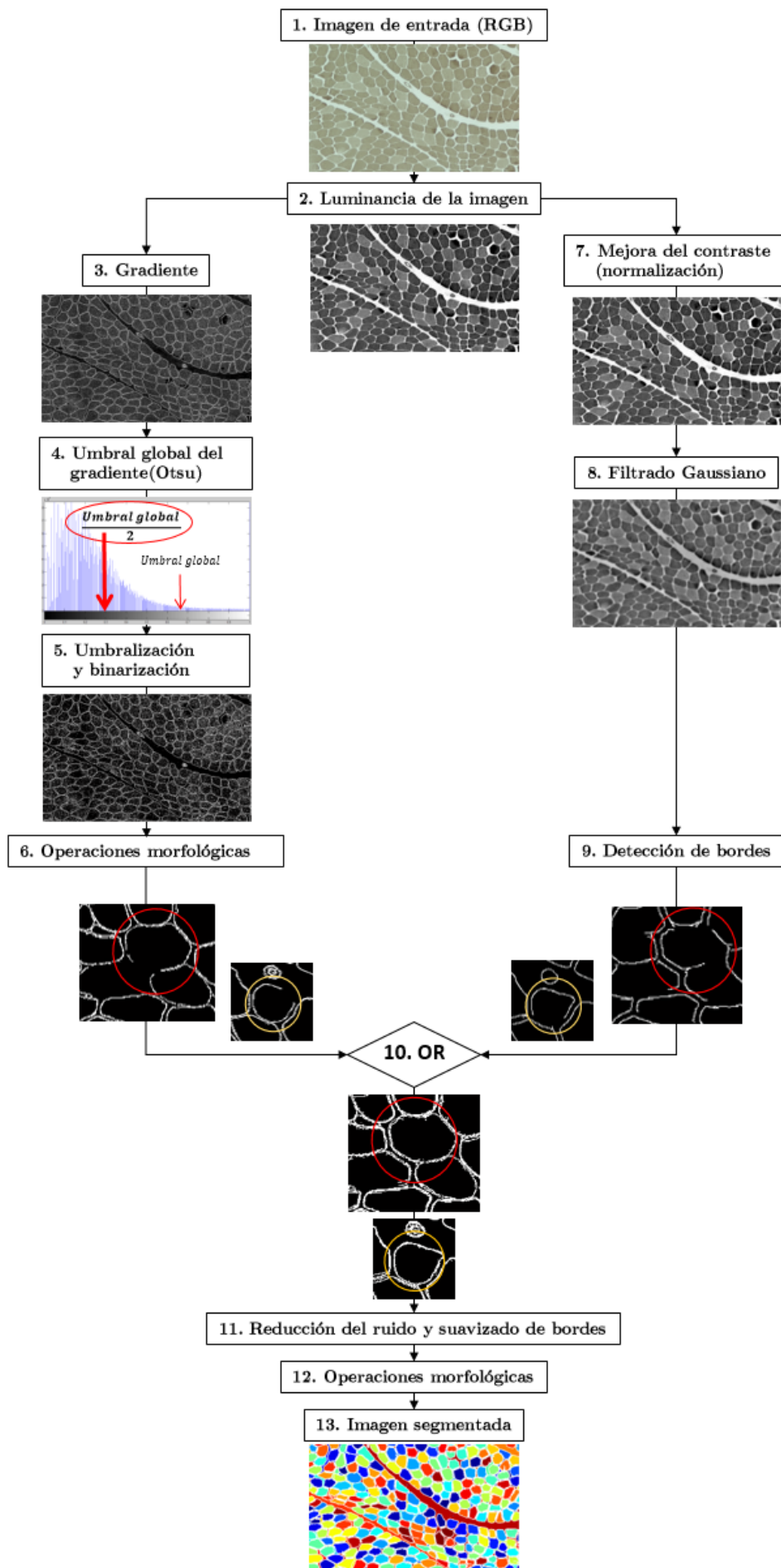


Figura 32: Algoritmo desarrollado en este trabajo como solución a la segmentación de imágenes de muestras de biopsias musculares.

Los pasos que sigue el algoritmo son los siguientes:

1. La imagen de entrada es una de las recibidas del Servicio de Anatomía Patológica, en espacio de color RGB.
2. Se obtiene el canal de la luminancia cambiando de RGB al espacio YCbCr como se indica en la sección 2.5.
3. Por un lado, se calcula la magnitud del gradiente con el operador Sobel como estimador de la derivada, ya que este se ha considerado que es el que mejores resultados ofrece. El gradiente es alto en los bordes de las células y más bajo en el interior. El gradiente en el interior de las células no es nulo debido a la aparición de píxeles cuyas intensidades provocan un aumento de la frecuencia espacial. En MATLAB, el cálculo del gradiente se representa en una imagen en escala de grises con valores tipo “double” en el intervalo $[0,1]$, donde el 0 indica el máximo gradiente (color blanco) y el 1 indica mínimo gradiente (negro).
4. El objetivo siguiente es binarizar la imagen del gradiente. Idealmente, los píxeles con valor 0 deberían ser los de los bordes de las células. Sin embargo, conservar sólo los píxeles con valor de gradiente nulo puede producir pérdida de información. Para ello primero se calcula el umbral global de la imagen del gradiente mediante el método de Otsu (ver paso 4 en el diagrama).
5. Sin embargo, utilizar el umbral global calculado también puede producir pérdida de información. Después de analizar todas las imágenes, se ha decidido que el umbral idóneo es la mitad del calculado $\left(\frac{\text{Umbral global}}{2}\right)$. Utilizando este umbral para binarizar, se da valor 0 todo lo que esté a la izquierda (negro) y 1 todo lo que esté a la derecha (blanco). Se consigue obtener bordes más aceptables, aunque con pequeños elementos conectados que contaminan la imagen. Estos elementos provienen del notable gradiente que contienen las células en su interior y se eliminan en el siguiente paso.
6. Las operaciones morfológicas que se realizan en este paso con la imagen binaria son aquellas que permiten eliminar pequeños grupos de elementos interconectados. En este caso se ha utilizado “bwareaopen()”. La imagen binaria de salida de este paso contiene la imagen segmentada, aunque con errores en la segmentación de los bordes de algunas células, que se subsanan en el paso 10.
7. Por otro lado, se ha realizado una mejora del contraste. En esta parte del algoritmo se han normalizado los valores de la intensidad de la imagen (valores “doblé” entre 0 y 1). A continuación, el algoritmo mejora el contraste de la imagen mediante la nivelación de su histograma.
8. J. Tan emplea la resta de dos filtros gaussianos para reducir cambios de intensidades de fondo en imágenes de contraste mejorado [31]. Esta resta se ha

modificado en este trabajo para eliminar, de la imagen de muestra muscular con contraste mejorado, los cambios de intensidades de las zonas interiores de células. La diferencia de filtros gaussianos funciona como un filtro paso-banda. El primer filtro deja la imagen con menos nitidez que la original y el segundo la deja muy poco nítida.

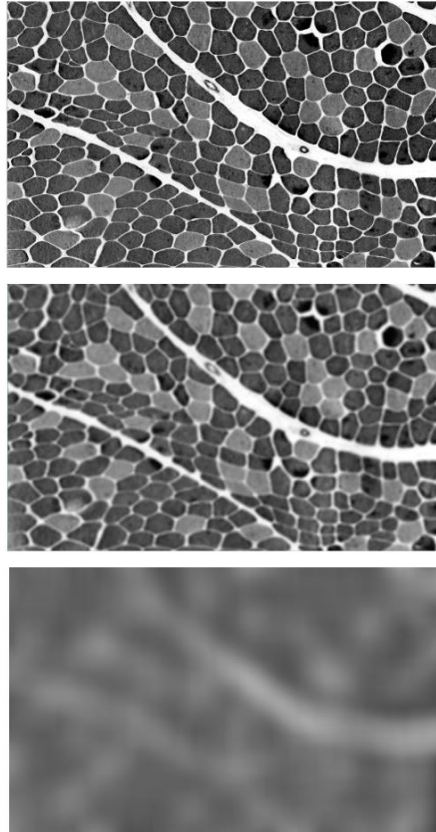


Figura 33: Arriba la imagen normalizada con contraste mejorado. En medio, aplicando el primer filtro gaussiano. Abajo, aplicando el segundo.

9. Mediante un filtro Sobel se obtiene la segmentación de los bordes de la imagen resultante de la resta de filtros gaussianos del paso 8. Se logra así una imagen cuyos bordes contienen detalles que se omiten en el cálculo de bordes del paso 6.
10. Con la utilización de una operación OR bit a bit con las imágenes binarias de los pasos 6 y 9, se consigue una detección de bordes más sofisticada que en los dos casos por separado. El operador OR funciona como una puerta lógica OR de 1 bit, siendo el 0 el color negro y el 1 blanco. Por tanto, la imagen de salida del paso 10 contiene los píxeles blancos de ambas.

Bit paso 6	Bit paso 9	Bit de salida en paso 10
0	0	0
0	1	1
1	0	1
1	1	1

Tabla 4: Operación bit a bit OR en el paso 10.

Es observable, antes de la operación OR, en el diagrama, que en ambas partes existen zonas donde los bordes no están bien definidos. Tras la operación estos se resuelven notablemente.

11. En este paso se aplica un filtro de la media para reducir el ruido y suavizar irregularidades en los bordes de las células.
12. Se vuelve a realizar operaciones morfológicas para obtener unos resultados óptimos. En concreto se emplean erosión, cierre, dilatación y eliminación de pequeños grupos de píxeles conectados.
13. Por último, se etiqueta cada área aislada (células en teoría) para su posterior cuantificación y clasificación. En el diagrama se utilizan etiquetas RGB para poder diferenciar visualmente unas células de otras. No obstante, el etiquetado utilizado en este trabajo para la cuantificación y el tratado individual de cada célula se ha realizado sobre la imagen segmentada binaria.

4.5 Resultados y comparación entre métodos de segmentación

En las secciones anteriores se ha mostrado el desarrollo de tres algoritmos automáticos para la obtención de la segmentación de imágenes de muestras de biopsias musculares.

A continuación, en esta sección, se muestran los resultados obtenidos tras aplicar los tres algoritmos de segmentación a las cuatro imágenes seleccionadas en la sección 4.2. El objetivo es contemplar qué algoritmo ofrece los mejores resultados, para ser llevado al software que utilizará el patólogo.

Cada algoritmo ha sido aplicado bajo las mismas condiciones y sin ningún tipo de modificación entre imágenes. Lo que se desea obtener es un método de segmentación automático, que no necesite de participación del patólogo para segmentar la imagen. Tras aplicar los algoritmos de segmentación a cada imagen, se muestran resultados visuales y cuantitativos a modo de comparación entre algoritmos.

Para comparar el resultado de la segmentación de las imágenes, se han cuantificado automáticamente el número de áreas aisladas obtenidas tras segmentar y se ha comprobado si se corresponden con células en la imagen original. Para realizar la cuantificación, se toma la matriz de la imagen de salida de los algoritmos, la cual

contiene la imagen segmentada cuyas áreas aisladas (de color blanco) poseen sus etiquetas individuales correspondientes. Las etiquetas están numeradas de 1 a n, siendo n la última área etiquetada. El valor 0 en la imagen se corresponde con los bordes segmentados y se representan en negro.

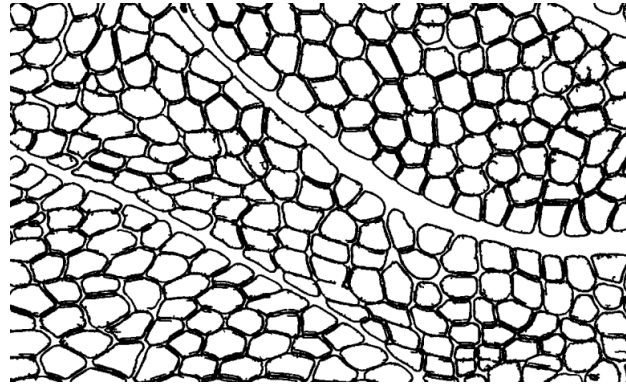


Figura 34: Ejemplo de imagen etiquetada. Cada área aislada posee una etiqueta individual.

Sólo se cuentan aquellos contornos aislados cuya área sea mayor que la cincuentava parte del área media de todas las áreas aisladas. Se ha elegido esta condición para no contar como células pequeños elementos de píxeles conectados que hayan podido quedar tras la segmentación. Esta decisión será comprobada posteriormente, cuando se comparen los resultados con respecto a la cuantificación de células manual. El diagrama de la Figura 35 muestra el flujo seguido para cuantificar áreas aisladas.

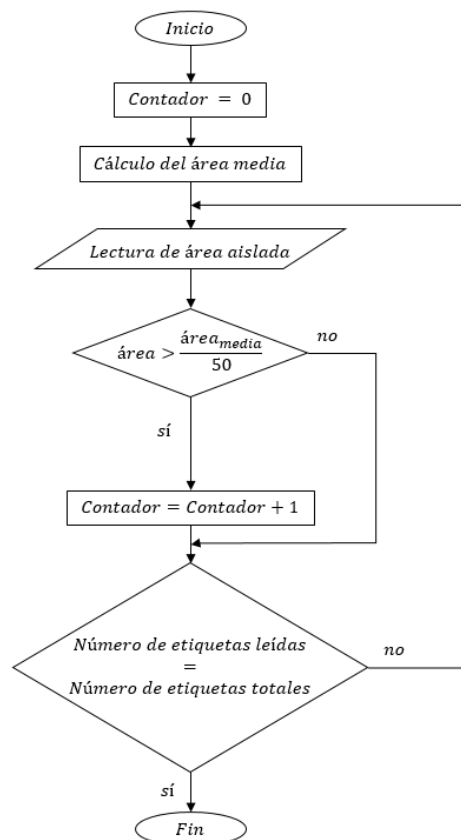


Figura 35: Diagrama de flujo de la cuantificación automática de áreas aisladas (células) en una imagen segmentada con etiquetas.

Para realizar las comparaciones cuantitativas entre algoritmos, para todas las imágenes se utilizarán tablas como la del ejemplo (Tabla 5). Dichas tablas, muestran en la primera columna los tres métodos utilizados para segmentar la Imagen I. En la segunda, se muestra el total de áreas aisladas conseguidas tras segmentar la imagen y aplicar el diagrama de la Figura 35. En la tercera, se muestra cuantas de esas áreas aisladas se corresponden con células en la imagen original, lo cual se ha realizado mediante inspección visual. La última columna muestra una figura de mérito calculada para mostrar cuantitativamente qué método es mejor. Esta Figura de Mérito (FdM) se calcula de la siguiente forma:

$$\text{FdM}_{\text{Segmentación}} = \frac{2 \cdot (\text{Áreas aisladas que son células})}{\text{Áreas aisladas totales} + \text{Número total}} \cdot 100 \quad (15)$$

Donde “Áreas aisladas totales” es el número de áreas aisladas calculadas en la imagen segmentada después de aplicar el método correspondiente, “Áreas aisladas que son células” indica cuales de las áreas anteriores se corresponden con células realmente y “Número total” indica el número total de células contadas manualmente. El valor de esta FdM va de 0 a 100, siendo 100 el mejor valor posible, en cuyo caso los valores de las tres variables son el mismo. Lo que se ha buscado con esta FdM es evaluar, tras aplicar los tres métodos de segmentación a las imágenes, cuanto se parece el resultado al valor perfecto (100).

Método	Áreas aisladas	Correspondencia con células	FdM
Watershed	-	-	-
K-means	-	-	-
Algoritmo propuesto	-	-	-

Tabla 5: Tabla modelo de comparación cuantitativa entre algoritmos.

4.5.1 Resultados para la Imagen I

La Imagen I (ver Figura 20) pertenece al grupo categorizado como imágenes que poseen un corte histológico transversal del músculo correcto y sin pliegues, además de una tinción correcta (ver Tabla 3). En las siguientes figuras es posible observar las tres segmentaciones, mostrando previamente la imagen original en el canal de la luminancia (escala de grises) para facilitar la comparación visual:

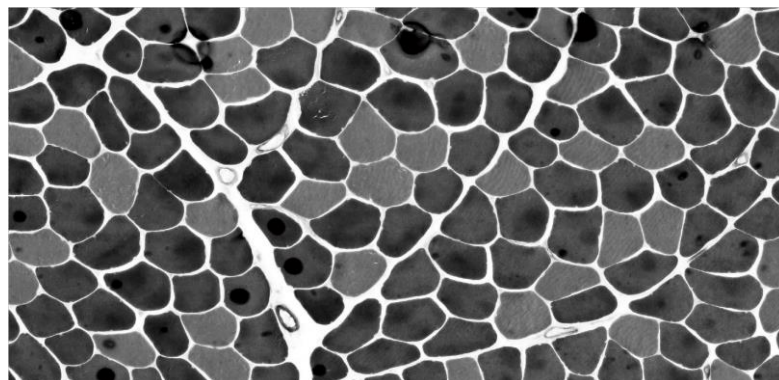


Figura 36: Imagen I en escala de grises (luminancia).

Es observable que, en la Imagen I en escala de grises, algunas células poseen cambios de intensidad en su interior. Esto se debe a las reacciones químicas provocadas por la tinción de la ATPasa A.

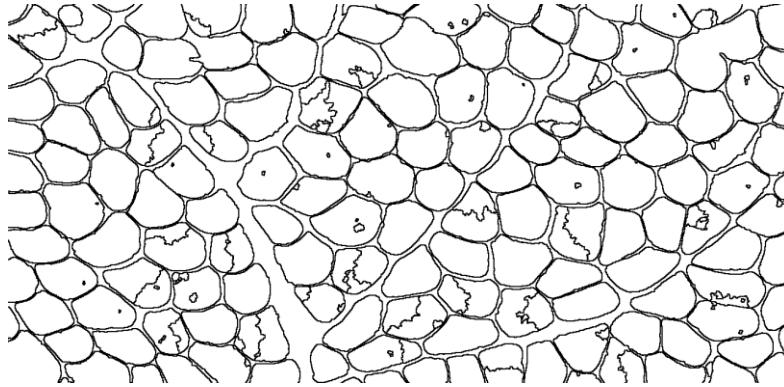


Figura 37: Segmentación de la Imagen I mediante el algoritmo basado en la Transformada Watershed.

A priori, es perceptible que la segmentación a través el algoritmo basado en Watershed provoca sobre-segmentación. Esto se debe a los ya comentados cambios de intensidad que se producen dentro de algunas células.

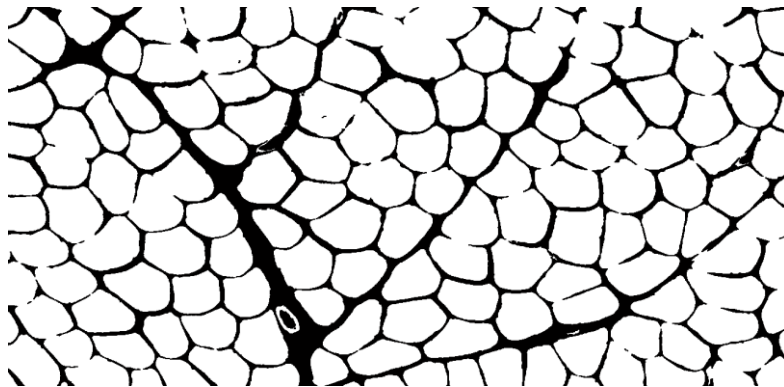


Figura 38: Segmentación de la Imagen I mediante el algoritmo basado en el clasificador K-means.

En el caso del algoritmo desarrollado en base al clasificador K-means, puede observarse que existen grupos de células interconectados, que no deberían estarlo.

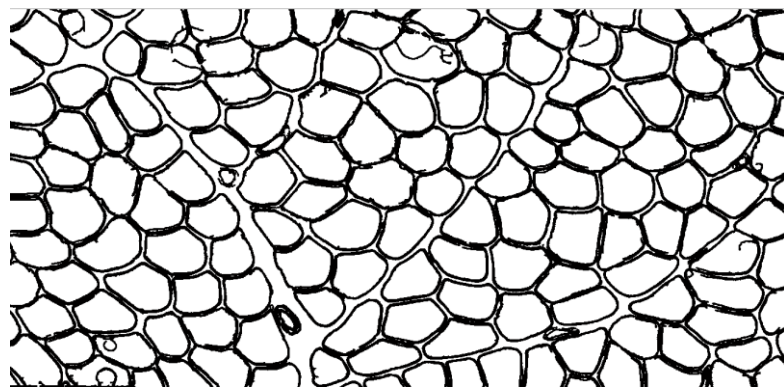


Figura 39: Segmentación de la Imagen I mediante el algoritmo aportado en este trabajo.

A primera vista, tras aplicar el algoritmo aportado en este trabajo a la Imagen I, se observa que el contorno de las células está bien definido. Sin embargo, también se

observan algunas líneas no deseadas debido a los cambios de intensidad, aunque notablemente menos que en el caso de Watershed.

Cuantitativamente se ha decidido calcular el número de áreas aisladas en cada método para posteriormente comprobar si de verdad estas áreas aisladas se corresponden con células de la imagen original. Por otro lado, el número real de células cuantificadas en la Imagen I de forma manual es:

Número total = 153 células

La Tabla 6 muestra los resultados obtenidos para la Imagen I tras aplicar los tres métodos:

Método	Áreas aisladas	Correspondencia con células	FdM
Watershed	195	97	55.8
K-means	87	58	48.3
Algoritmo propuesto	155	151	98.0
Número total = 153			

Tabla 6: Resultados cuantitativos de la segmentación de la Imagen I.

Como se puede observar, el algoritmo desarrollado para este trabajo en la sección 4.4, además de dar la FdM mayor (98.0), su valor está muy próximo a 100, lo cual indica un correcto funcionamiento. Además, las células detectadas con este algoritmo son 151 para 155 áreas aisladas, sólo dos menos que las 153 que posee la imagen.

Por otro lado, los métodos de Watershed y K-means no dejan resultados tan aceptables. Watershed produce una clara sobre-segmentación de la Imagen I, creando demasiadas áreas aisladas que fragmentan células, provocando su pérdida. K-means sin embargo, genera menos áreas aisladas de las que debiera, debido a la existencia de grupos de células interconectados.

4.5.2 Resultados para la Imagen II

La Imagen II, cuya representación en escala de grises puede verse en la Figura 40, pertenece al grupo categorizado como imágenes que poseen un corte histológico transversal del músculo correcto y sin pliegues, pero con una tinción débil o incorrecta (ver Tabla 3). En la Figura 40 el contraste ha sido mejorado mediante ajuste de su histograma para poder diferenciar mejor los dos tipos de células.

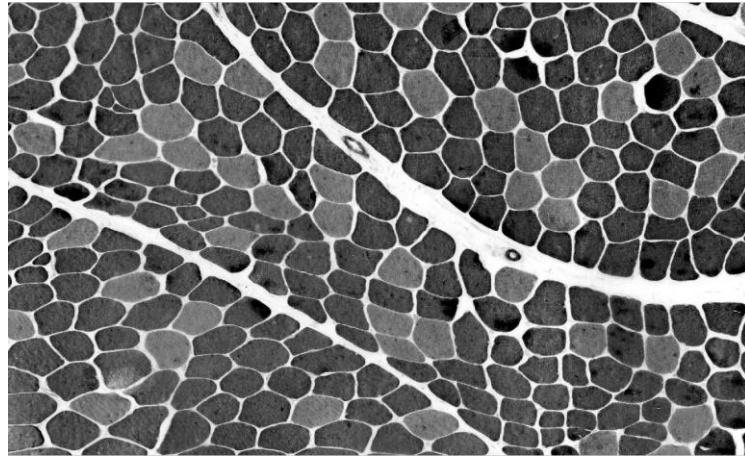


Figura 40: Imagen II en escala de grises (luminancia).

En las siguientes figuras es posible observar la segmentación mediante los tres métodos propuestos:

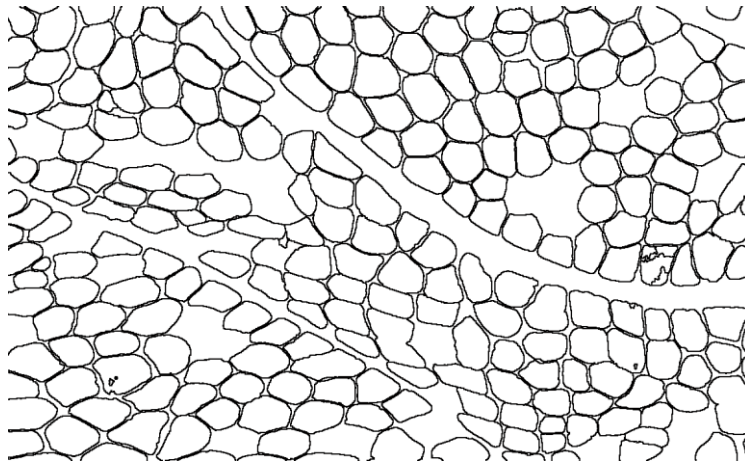


Figura 41: Segmentación de la Imagen II mediante el algoritmo basado en la Transformada Watershed.

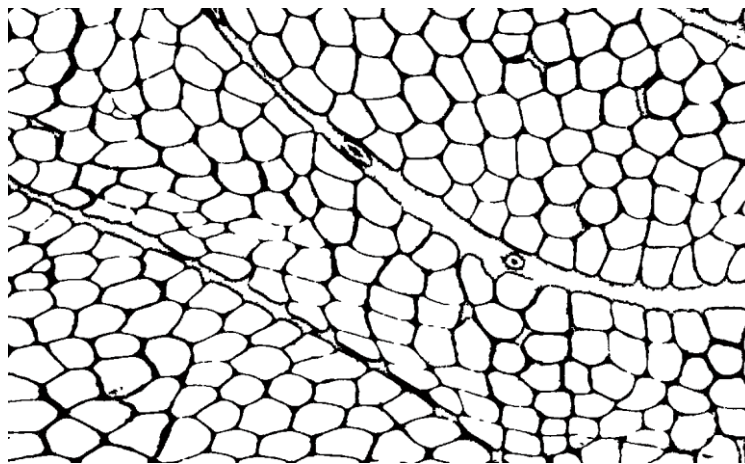


Figura 42: Segmentación de la Imagen II mediante el algoritmo basado en el clasificador K-means.

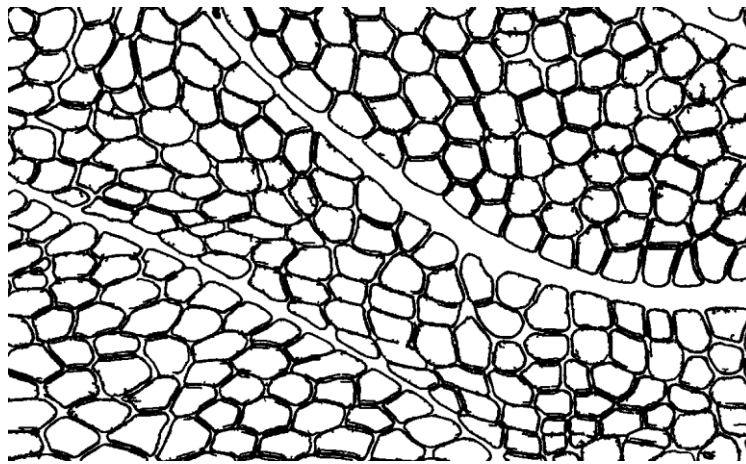


Figura 43: Segmentación de la Imagen II mediante el algoritmo aportado en este trabajo.

El número total de células cuantificadas manualmente ha sido:

Número total = 280 células

En la Tabla 7 se muestran los resultados relativos a la segmentación en la Imagen II:

Método	Áreas aisladas	Correspondencia con células	FdM
Watershed	258	222	82.5
K-means	199	156	65.1
Algoritmo propuesto	290	278	97.5
Número total = 280			

Tabla 7: Resultados cuantitativos de la segmentación de la Imagen II.

En el caso de la Imagen II, con el algoritmo desarrollado para este trabajo se vuelve a obtener un buen resultado, con la FdM=97.5 muy próximo a 100. Sólo se han detectados dos células menos de las que posee la Imagen II.

Con Watershed, se obtiene un resultado más apropiado que en el caso de la Imagen I, aunque alejado del último método. En cuanto a K-means, es con la que peor FdM se obtiene.

4.5.3 Resultados para la Imagen III

La Imagen III (ver Figura 22), pertenece al grupo categorizado como imágenes que poseen un corte histológico transversal del músculo deficiente y con aparición de numerosos pliegues, aunque con tinción correcta que permite diferenciar correctamente entre tipos de células (ver Tabla 3).

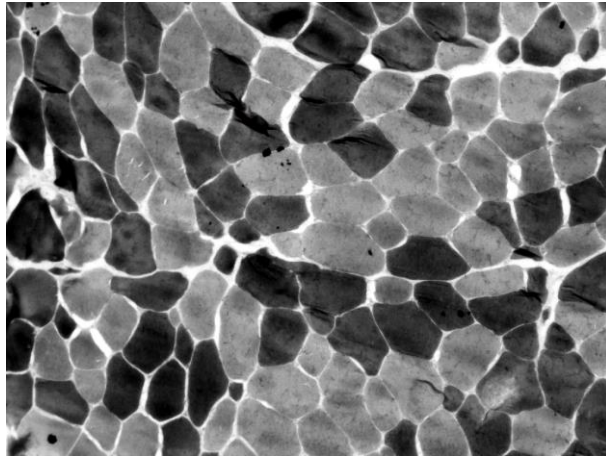


Figura 44: Imagen III en escala de grises (luminancia).

En las siguientes figuras es posible observar la segmentación mediante los tres métodos propuestos:

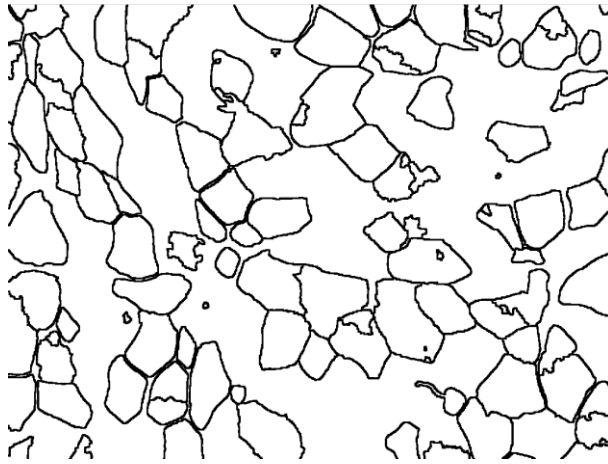


Figura 45: Segmentación de la Imagen III mediante el algoritmo basado en la Transformada Watershed.

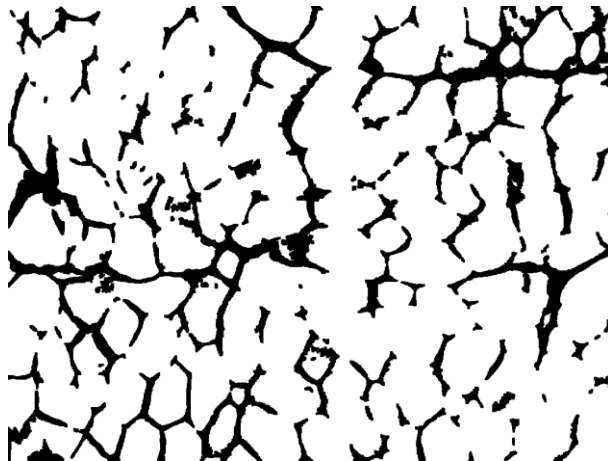


Figura 46: Segmentación de la Imagen III mediante el algoritmo basado en el clasificador K-means.

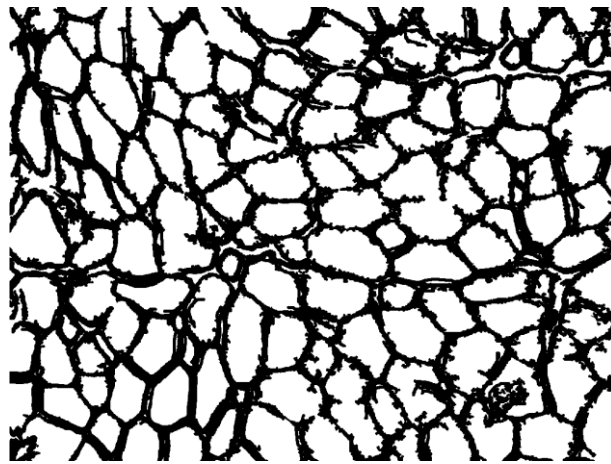


Figura 47: Segmentación de la Imagen III mediante el algoritmo aportado en este trabajo.

El número total de células cuantificadas manualmente ha sido:

Número total = 156 células

En la Tabla 8 se muestran los resultados relativos a la segmentación en la Imagen II:

Método	Áreas aisladas	Correspondencia con células	FdM
Watershed	75	45	38.9
K-means	15	10	11.7
Algoritmo propuesto	135	120	82.4
Número total = 156			

Tabla 8: Resultados cuantitativos de la segmentación de la Imagen III.

El algoritmo de segmentación propuesto resulta aceptable en el caso de la Imagen III, con una FdM=82.4, volviendo a ser la mayor de las tres obtenidas. No obstante, se hace patente la aparición de pliegues como impedimento para alcanzar una buena segmentación. En cuanto a los otros dos métodos, los resultados no ofrecen valores aceptables, siendo K-means el peor de todos, con una FdM=11.7.

4.5.4 Resultados para la Imagen IV

Por último, se muestran los resultados para la Imagen IV (ver Figura 23). Esta imagen pertenece al grupo de imágenes cuyo corte histológico es incorrecto, produciendo la aparición de pliegues, y con tinción débil e incorrecta (ver Tabla 3).

La Imagen IV es especialmente característica, debido a sus amplias zonas con pliegues que solapan numerosas células. Además, la tinción, ha provocado una distinción débil entre tipos de células y la aparición de grandes cambios de intensidad en el interior de las propias células. Tanto los cambios de intensidad como los pliegues pueden ser observados en la Figura 48, la cual se corresponde con el canal de la luminancia de Imagen IV, con contraste mejorado e histograma ajustado.

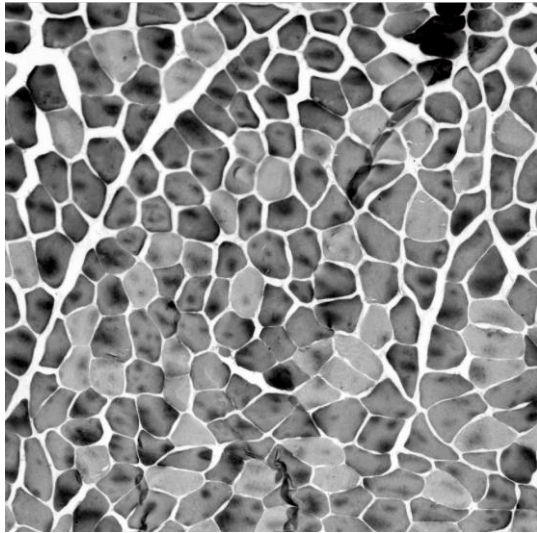


Figura 48: Imagen IV en escala de grises (luminancia).

En las siguientes figuras es posible observar la segmentación mediante los tres métodos propuestos:

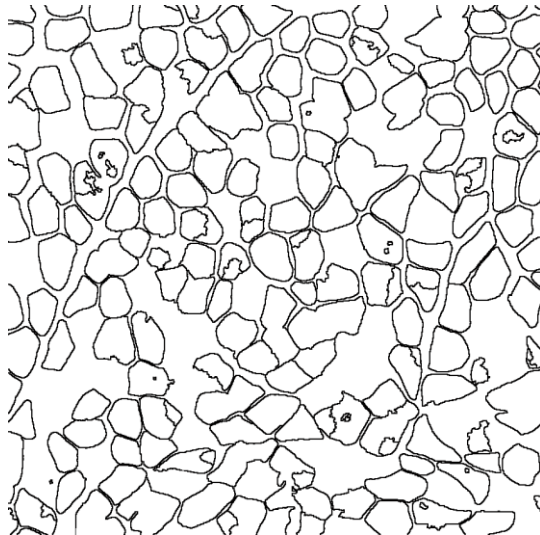


Figura 49: Segmentación de la Imagen IV mediante el algoritmo basado en la Transformada Watershed.

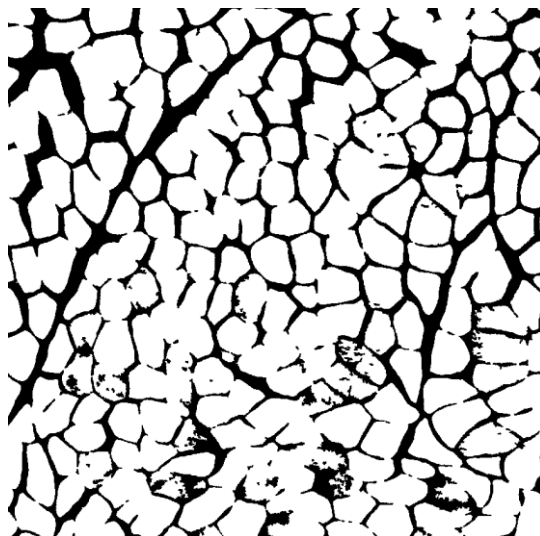


Figura 50: Segmentación de la Imagen IV mediante el algoritmo basado en el clasificador K-means.

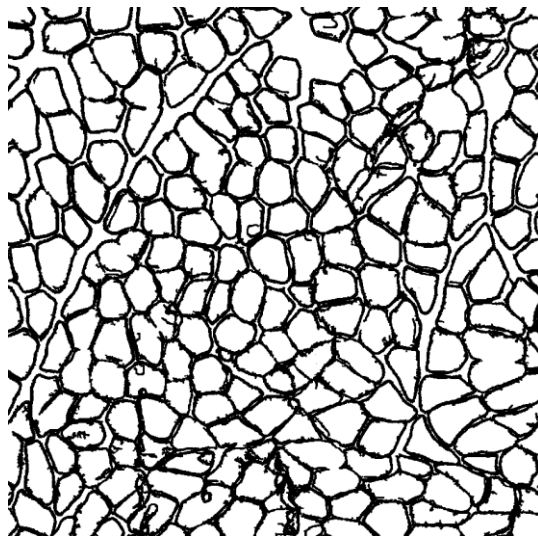


Figura 51: Segmentación de la Imagen IV mediante el algoritmo aportado en este trabajo.

El número total de células cuantificadas manualmente ha sido:

Númeto total = 226 células

En la Tabla 8 se muestran los resultados relativos a la segmentación en la Imagen II:

Método	Áreas aisladas	Correspondencia con células	FdM
Watershed	194	109	53.7
K-means	56	25	18.7
Algoritmo propuesto	246	195	82.6
Número total = 226			

Tabla 9: Resultados cuantitativos de la segmentación de la Imagen IV.

Se observa como con el algoritmo propuesto en este trabajo se consigue una aceptable FdM=82.6, la mayor de los tres. K-means vuelve a ser el que resulta con una FdM menor y Watershed ofrece un resultado inaceptable.

En la sección 4.1.1, en la cual se habla de la importancia del corte histológico, se recalca que el corte debe realizarse de la forma más transversal posible. Resulta notable cómo la aparición de pliegues provoca una disminución en el número de células detectadas por los algoritmos. Pese a esto, la importancia radica en determinar si la existencia de células solapadas provoca peores resultados en el software o no.

La realidad es que, cuando existen varias células solapadas, algunas partes de las mismas se encuentran en el interior de otras, por lo que no hay manera de separarlas mediante segmentación. Un ser humano sí es capaz de distinguirlas visualmente, y es por eso que se han considerado en la cuantificación manual como células separadas. Por tanto, el número de células correctas detectadas por el algoritmo va a ser siempre menor al número de células cuantificadas manualmente, cuando algunas de estas se encuentren solapadas.

Sin embargo, queda patente que el método desarrollado en este trabajo proporciona una mejor solución a la segmentación que los otros, y es por eso que ha sido seleccionado para ser incluido en el software final.

Capítulo 5

Algoritmos para la obtención de parámetros de interés

Tras decidir qué algoritmo de segmentación de imágenes de muestras musculares se va a utilizar en la futura herramienta a proporcionar al patólogo, se procede a realizar el cálculo de algunos de los parámetros necesarios para realizar un apoyo al diagnóstico.

En este capítulo se muestra el algoritmo desarrollado en este trabajo para cuantificar y clasificar células en sus correspondientes tipos. Posteriormente, se estudian los tres algoritmos desarrollados para el cálculo del diámetro menor de cada célula de la muestra, con el objetivo de compararlos después. El mejor de ellos ha sido seleccionado para la herramienta de apoyo al diagnóstico de distrofias musculares, que se muestra en el capítulo 6.

5.1 Aportaciones a la cuantificación y clasificación de células musculares

En la sección 2.4 se ha mostrado el procedimiento tradicional para el diagnóstico de enfermedades musculares. En esta sección se describe el procedimiento seguido en este trabajo para obtener la cuantificación de fibras Tipo I, Tipo II y las totales, que son parámetros necesarios en el diagnóstico. El diagrama de la Figura 52, muestra el flujo que sigue el algoritmo de cuantificación y clasificación desarrollado.

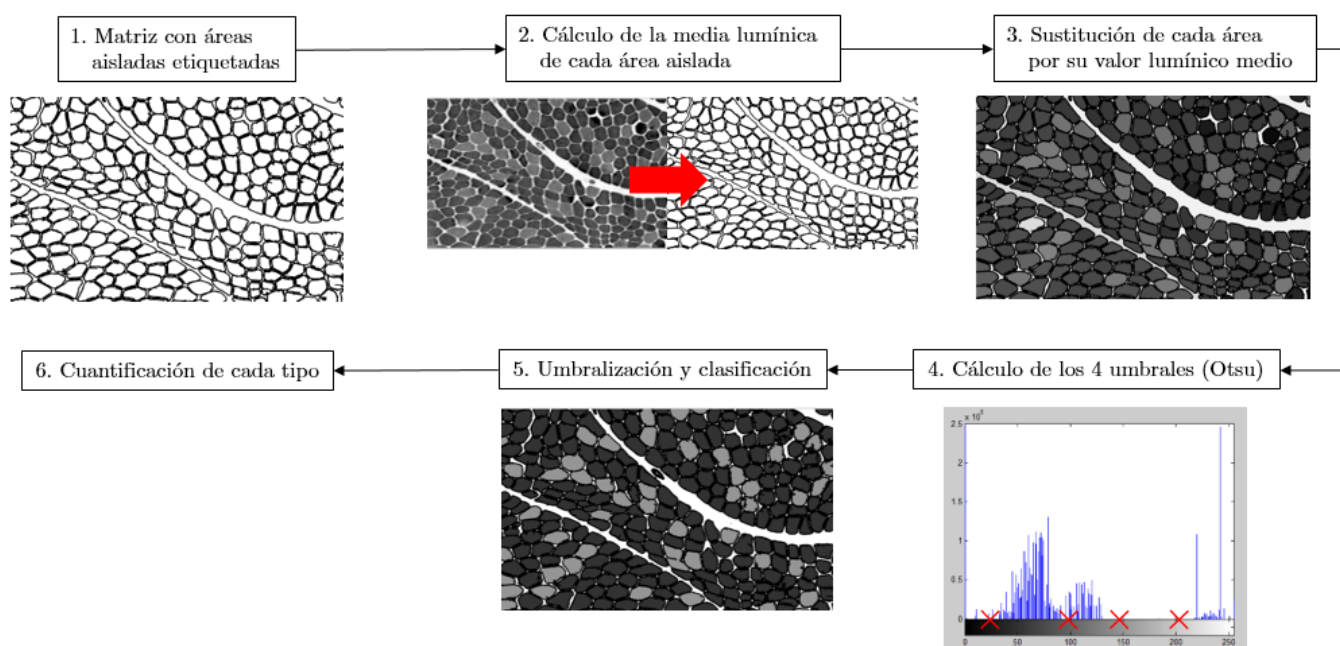


Figura 52: Diagrama del algoritmo desarrollado para la cuantificación y clasificación de células.

Los pasos que sigue el algoritmo son los siguientes:

1. En primer lugar, se toma la matriz de la imagen de salida de los algoritmos de segmentación, la cual contiene la imagen segmentada cuyas áreas aisladas (de color blanco) poseen sus etiquetas individuales correspondientes (de 1 a n). Los bordes se representan de color negro y con valor 0.
2. Posteriormente, se mide la media lumínica de cada área aislada sobre la imagen en escala de grises.
3. Se sustituye cada área aislada por su media lumínica, como se observa en la imagen que acompaña al punto 3 del diagrama.
4. El histograma de la imagen del paso anterior, es similar para todas las imágenes. Aparecen dos picos: el de la izquierda se corresponde con las células oscuras (Tipo II) y el de la derecha con las claras (Tipo I). Se aplica el método de Otsu, para calcular cuatro umbrales de manera automática. En la Figura 53, se observan cuatro marcas que indican los cuatro umbrales automáticos.

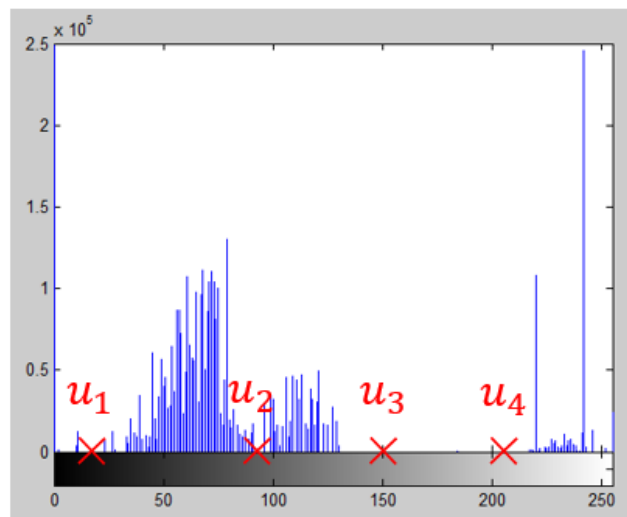


Figura 53: Histograma de la imagen del paso 3 del diagrama.

5. Se umbraliza la imagen del paso 3, teniendo en cuenta entre qué umbrales se encuentra la media lumínica de cada área. Si esta se encuentra entre el umbral u_1 y u_2 , se clasifica como célula oscura (Tipo II), con un valor de 50 en escala de grises, y si se encuentra entre u_2 y u_3 , como clara (Tipo I), con un valor de 150 en escala de grises. Además, para que un área aislada sea clasificada como célula, se debe cumplir el criterio, utilizado anteriormente, de que su área sea mayor que la cincuentava parte del área media (ver sección 4.5).
6. Por último, se realiza la cuantificación automática de las células de cada tipo, ya que del paso 5 es posible conocer cuántas células de cada tipo han sido umbralizadas.

Las figuras siguientes muestran la luminancia de las cuatro imágenes utilizadas de test en este trabajo, junto con sus imágenes segmentadas con sus respectivas fibras clasificadas. Las células claras representan las Tipo I y las oscuras las Tipo II.

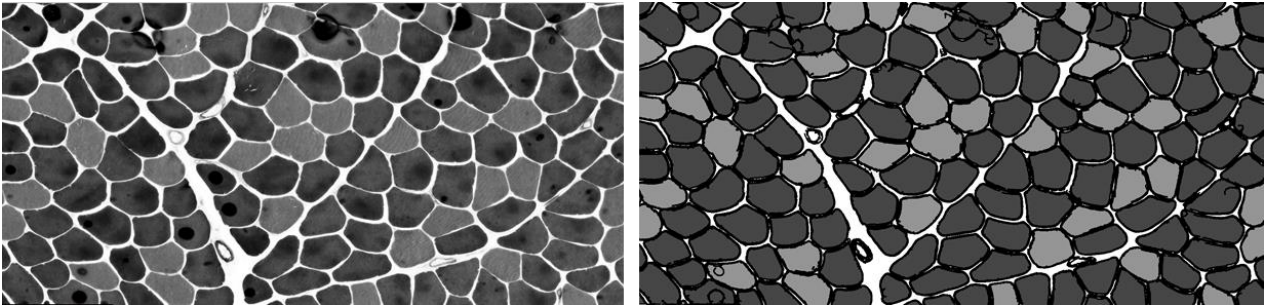


Figura 54: Clasificación de tipos para la Imagen I.

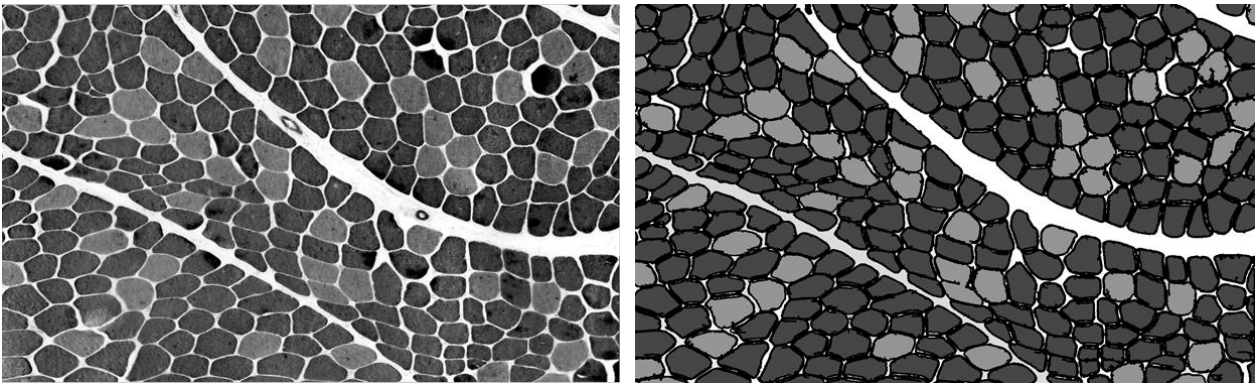


Figura 55: Clasificación de tipos para la Imagen II.

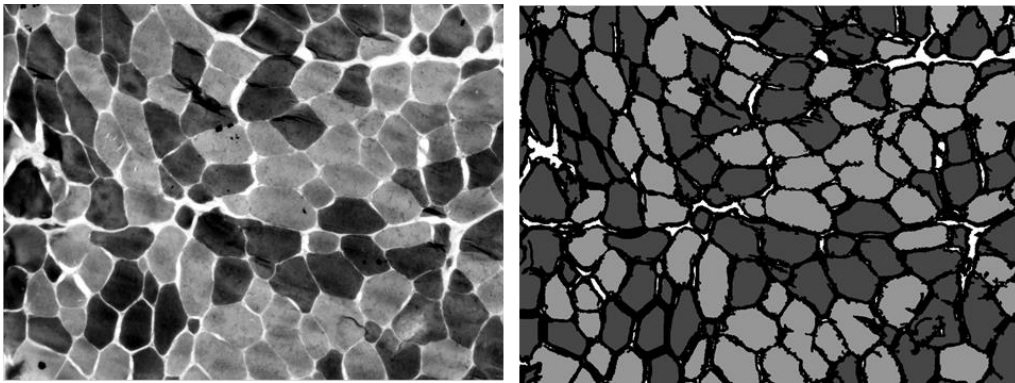


Figura 56: Clasificación de tipos para la Imagen III.

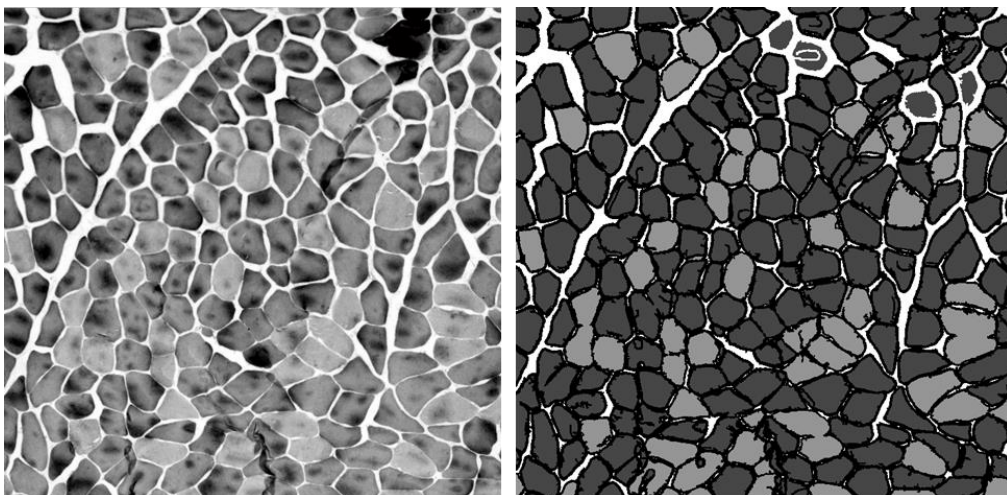


Figura 57: Clasificación de tipos para la Imagen IV.

En la Tabla 10, se muestran los resultados obtenidos para en la clasificación de tipos de células para las cuatro imágenes representativas. El porcentaje de exactitud, sobre una cuantificación manual como la realizada en el capítulo anterior, indica con qué acierto se ha determinado dicha clasificación células.

Imagen	Tipo I (manual)	Tipo I (software)	% de exactitud	Tipo II (manual)	Tipo II (software)	% de exactitud
I	36	34	94.4 %	117	122	95.9 %
II	49	51	96.1 %	231	239	96.7 %
III	70	55	78.5 %	86	80	93 %
IV	56	50	89.3 %	170	192	88.5 %

Tabla 10: Resultados numéricos de la clasificación de tipos de células.

Se observa que para las Imágenes III y IV, las que poseían numerosos pliegues y células solapadas, el porcentaje de exactitud es menor, como ya se predijo en el capítulo anterior.

Una vez efectuada la clasificación y cuantificación de fibras, es posible realizar el cálculo de algunos de los pasos que sigue el patólogo en el diagnóstico (ver sección 2.4):

- Cálculo del porcentaje de fibras Tipo I y Tipo II frente al total.
- Cálculo del diámetro medio de cada tipo y diámetro medio total.

5.2 Aportaciones al cálculo de los diámetros celulares

Una vez se han clasificado y cuantificado las células musculares, se continúa el resto del procedimiento. En este caso, con la medida del diámetro menor de las células.

En la sección 2.4, en la cual se presenta el concepto de diámetro menor de una fibra, se muestra la importancia de esta medida para el diagnóstico de distrofias musculares. El procedimiento seguido por el patólogo consiste en seleccionar 100 células aleatorias de cada tipo y medir sus correspondientes diámetros menores, para posteriormente realizar los cálculos de parámetros y representaciones pertinentes.

En este trabajo, se ha implementado el cálculo de dichos diámetros de forma automática. En las siguientes subsecciones se muestra cómo han sido desarrollados tres posibles métodos para resolver el problema, para posteriormente elegir el que mejor resultados ofrece.

5.2.1 Método de fuerza bruta

Como su propio nombre indica, este algoritmo se ha implementado bajo la idea de realizar una búsqueda exhaustiva del diámetro mayor y menor, mediante la enumeración sistemática de todos los posibles candidatos. Se basa en el cálculo de la distancia entre dos puntos en un plano:

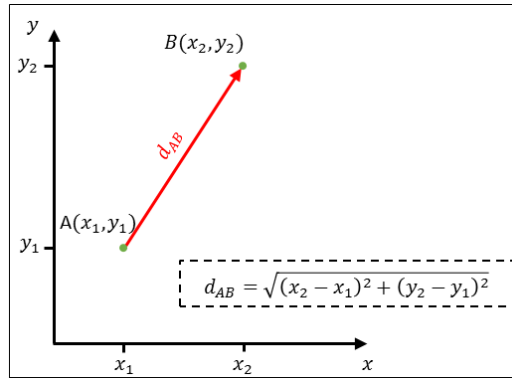


Figura 58: Cálculo de la distancia entre dos puntos en un plano.

En este trabajo, este algoritmo ha sido llevado a cabo para encontrar el diámetro más largo en la célula, para posteriormente calcular su perpendicular más larga (a la cual se la denomina diámetro menor).

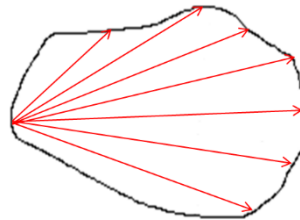


Figura 59: Búsqueda mediante fuerza bruta del diámetro más largo, mediante el cálculo de todas las distancias posibles desde todos los píxeles.

En primer lugar, se realiza un barrido de todos los píxeles de la célula y se almacenan todas las coordenadas de dichos píxeles en una matriz de dimensiones 2xN:

$$M_{2 \times N} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \end{bmatrix} \quad (16)$$

En esta matriz se almacenan todas las coordenadas (x en la primera fila e y en la segunda) de los píxeles de la célula en cuestión, siendo N el último píxel.

A continuación, se realiza un barrido columna por columna, sobre la matriz M, en búsqueda de la distancia más larga posible en la célula.

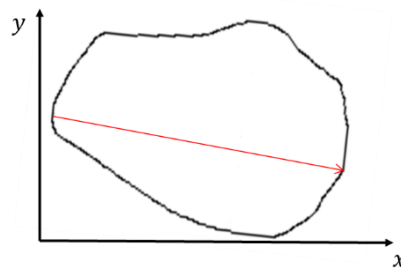


Figura 60: Ejemplo de diámetro máximo de una célula.

Seguidamente, para calcular el diámetro menor, es decir, el diámetro máximo perpendicular al diámetro más largo, se proyecta el diámetro mayor calculado (ver Figura 60) sobre dos nuevos ejes, x' e y', de forma que el diámetro mayor quede paralelo al nuevo eje x' (ver Figura 61).

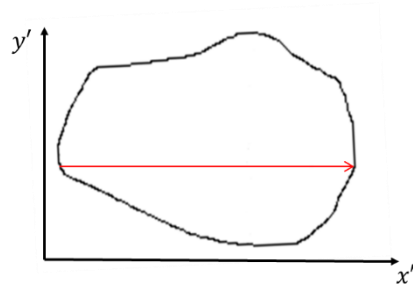


Figura 61: Diámetro más largo proyectado sobre nuevos ejes.

Una vez que el diámetro más largo se encuentra paralelo al nuevo eje de abscisa, se calcula el punto más alto y bajo posible (en el eje y' , sobre un valor de x') de entre todas las perpendiculares del diámetro, dando así con el diámetro perpendicular máximo, el denominado diámetro menor.

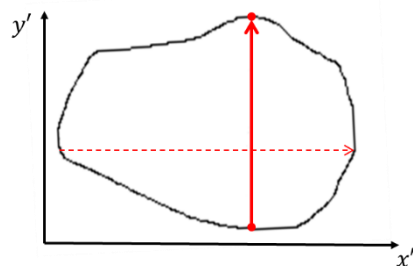


Figura 62: Diámetro menor de la célula ejemplo, en rojo y sin puntear.

Los resultados obtenidos para este método se muestran en la sección 5.2.4.

5.2.2 Método del diámetro de Feret

El diámetro de Feret [32] representa la distancia entre dos líneas paralelas que son tangenciales al contorno de la célula.

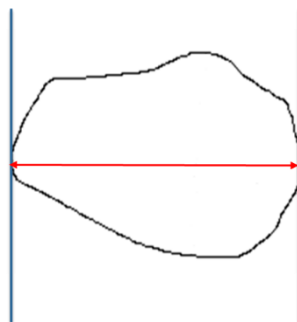


Figura 63: Ejemplo de diámetro de Feret.

Dichas líneas paralelas, separan el par de puntos que hacen de antípodas en esa dirección. Es posible determinar otros diámetros de Feret si se modifica la dirección con la que se trazan las tangentes.

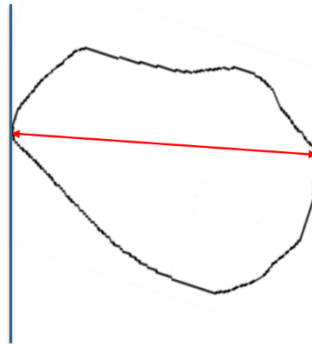


Figura 64: Ejemplo de diámetro de Feret en otra dirección. Se observa un cambio de dirección con respecto a la Figura 63.

Se ha desarrollado un algoritmo que primero calcula el diámetro mayor en una célula y seguidamente realiza el cálculo del diámetro menor o perpendicular más largo.

Para calcular el diámetro de Feret de mayor longitud, se utilizan las coordenadas de los puntos antípodas de la célula:

- En primer lugar, se toman las coordenadas, de los píxeles de la célula, con mayor y menor valor sobre el eje de abscisa (ver Figura 65).

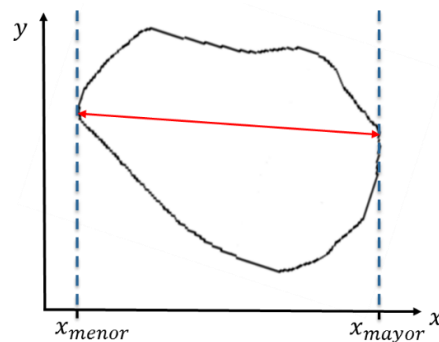


Figura 65: Píxeles con coordenadas mayor y menor en el eje de abscisa.

- En segundo lugar, se calcula la distancia entre dichos puntos antípodas, de igual forma que en la Figura 58.
- Seguidamente, se almacena el diámetro calculado y se cambia la dirección de la célula, girándola en el sentido antihorario. Un nuevo diámetro de Feret es calculado y comparado con el anterior, quedándose almacenado el mayor de los dos. Se continúa este proceso hasta que se ha girado la célula de forma completa, obteniendo entonces el diámetro más largo obtenido.

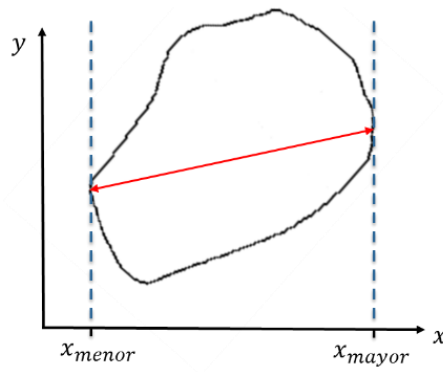


Figura 66: Ejemplo de un diámetro de Feret, una vez se ha girado la célula.

Para calcular el diámetro máximo perpendicular al diámetro más largo, se proyecta el diámetro más largo calculado, tras el giro completo de la célula, sobre dos nuevos ejes x' e y' , de forma que el diámetro mayor quede paralelo al nuevo eje x' (ver Figura 61), tal y como se ha planteado en el método anterior.

Una vez que el diámetro más largo se encuentra paralelo al nuevo eje de abscisa, se calcula el punto más alto y bajo posible (en el eje y' , sobre un valor de x') de entre todas las perpendiculares del diámetro, dando así con el diámetro perpendicular máximo, el denominado diámetro menor (ver Figura 62).

Los resultados obtenidos para este método se muestran en la sección 5.2.4.

5.2.3 Método de las tensiones principales

Finalmente, se ha adaptado un método habitualmente utilizado en mecánica tensorial como solución a un problema de procesamiento digital de imágenes y medicina.

Es necesario introducir conceptos sobre tensiones en paralelepípedos y posteriormente, en planos [33]. Sobre cada una de las caras planas del paralelepípedo de la Figura 67, existe una tensión normal (σ) y una tensión cortante o tangencial (τ). Si se descomponen dichas tensiones, además se obtienen sus correspondientes componentes sobre las direcciones de los tres ejes (x, y, z).

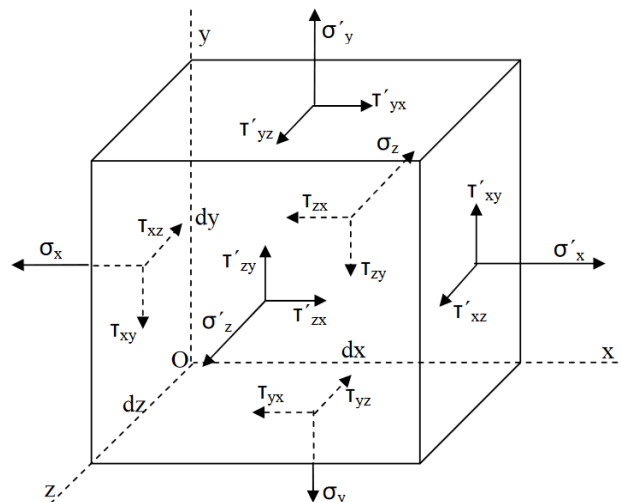


Figura 67: Descomposición de tensiones sobre un paralelepípedo [33].

Para las tensiones normales, σ_x indica que esta tensión está sobre una superficie normal al eje x e igual ocurre con las demás tensiones normales. En el caso de las tensiones tangenciales, τ_{xy} indica que esta tensión pertenece a una superficie normal al eje x con la dirección del eje y , e igual ocurre con las demás.

Las tensiones en las tres caras del paralelepípedo que no pasan por el origen de los ejes ($\sigma'_x, \sigma'_y, \sigma'_z, \tau'_{xy}, \tau'_{yz}, \tau'_{zx}$), pueden ser expresados matemáticamente en función de las tensiones en las otras tres caras que sí pasan por el origen ($\sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz}, \tau_{zx}$) por desarrollo de Taylor:

$$\sigma'_x = \sigma_x + \frac{d\sigma_x}{dx} \cdot dx, \quad \tau'_{xy} = \tau_{xy} + \frac{d\tau_{xy}}{dx} \cdot dx, \quad \tau'_{xz} = \tau_{xz} + \frac{d\tau_{xz}}{dx} \cdot dx \quad (17, 18, 19)$$

$$\sigma'_y = \sigma_y + \frac{d\sigma_y}{dy} \cdot dy; \quad \tau'_{yx} = \tau_{yx} + \frac{d\tau_{yx}}{dy} \cdot dy, \quad \tau'_{yz} = \tau_{yz} + \frac{d\tau_{yz}}{dy} \cdot dy \quad (19, 20, 21)$$

$$\sigma'_z = \sigma_z + \frac{d\sigma_z}{dz} \cdot dz, \quad \tau'_{zx} = \tau_{zx} + \frac{d\tau_{zx}}{dz} \cdot dz, \quad \tau'_{zy} = \tau_{zy} + \frac{d\tau_{zy}}{dz} \cdot dz \quad (22, 23, 24)$$

Dado que el paralelepípedo es semejante en todas sus caras y se encuentra en condiciones de equilibrio, se puede considerar que:

$$\begin{aligned} \sigma'_x &= \sigma_x; & \sigma'_y &= \sigma_y; & \sigma'_z &= \sigma_z; \\ \tau_{xy} &= \tau'_{xy} = \tau_{yx} = \tau'_{yx}; & \tau_{xz} &= \tau'_{xz} = \tau_{zx} = \tau'_{zx}; & \tau_{yz} &= \tau'_{yz} = \tau_{zy} = \tau'_{zy}; \end{aligned}$$

Por tanto, son sólo seis las tensiones distintas que actúan sobre las caras del paralelepípedo:

$$\sigma_x; \quad \sigma_y; \quad \sigma_z; \quad \tau_{xy}; \quad \tau_{xz}; \quad \tau_{yz}$$

Es posible definir el tensor de tensiones, como la representación matemática en forma de matriz del estado tensional al que se encuentra sometido un punto en el sistema, la cual es simétrica:

$$\mathbf{T} = \begin{bmatrix} \sigma_x & \tau_{xy} & \tau_{xz} \\ \tau_{xy} & \sigma_y & \tau_{yz} \\ \tau_{xz} & \tau_{yz} & \sigma_z \end{bmatrix} \quad (25.1)$$

Una vez conocido el tensor de tensiones definido sobre un paralelepípedo, es posible definirlo para un sistema plano. Se considera que el tensor de tensiones en este caso cumple:

$$\begin{aligned} \sigma_z &= 0; & \tau_{xz} &= 0; & \tau_{yz} &= 0 \\ \rightarrow \mathbf{T} &= \begin{bmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{bmatrix} \quad (25.2) \end{aligned}$$

Si se traslada este procedimiento al caso que atañe este trabajo, las células musculares son sistemas planos que pueden ser tratados sobre dos ejes, x e y . Es por esto que a continuación se tratan las tensiones principales sobre sistemas planos. En la Figura 68

se observa una célula ejemplo representada sobre dos ejes, junto con la representación de sus tensiones normales y tangenciales y su correspondiente matriz del tensor de tensiones.

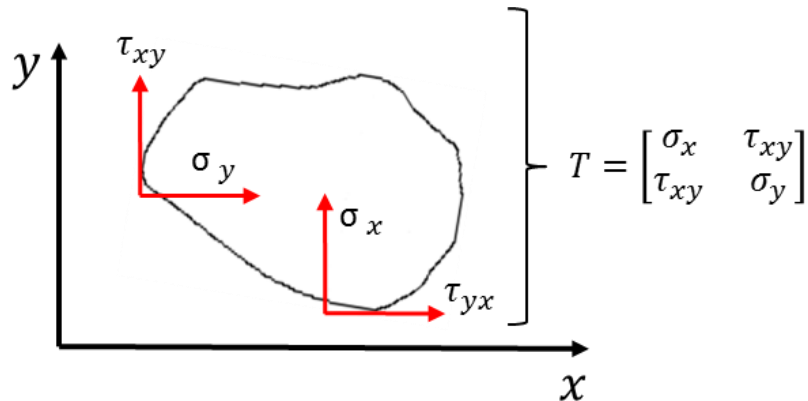


Figura 68: Ejemplo de célula muscular con sus tensiones normales y tangenciales representadas, sobre un eje de coordenadas.

La matriz del tensor de tensiones calculado para un plano, se apoya en el sistema de referencia original con ejes x e y . En dicho plano, también es posible definir otro sistema de referencia particular, con ejes llamados 1 y 2, en el cual la matriz del tensor de tensiones es diagonal. Esta matriz se denomina matriz de las tensiones principales, y se dice que es diagonal porque en ella las componentes tangenciales son nulas y sólo existen tensiones normales, σ_1 y σ_2 (ver Figura 69).

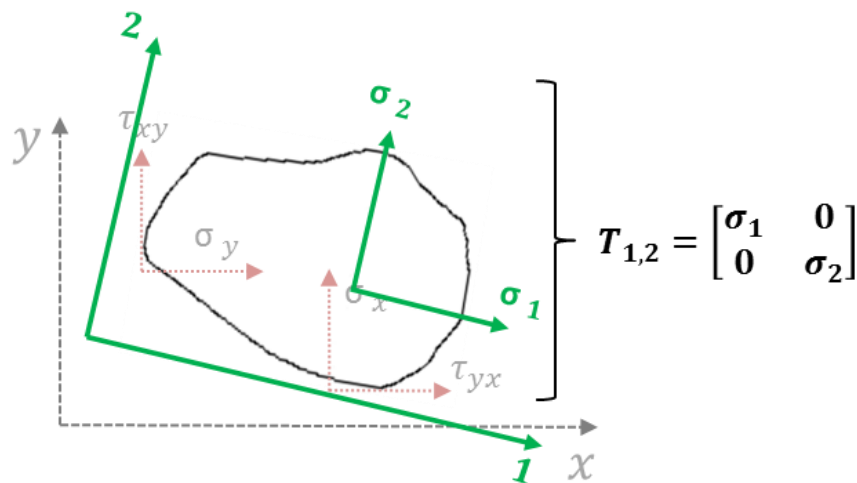


Figura 69: Célula muscular proyectada sobre los ejes de las tensiones principales.

Al proyectar el plano de la célula sobre los ejes de las tensiones principales, las direcciones principales apuntan hacia el diámetro más largo de la célula y su perpendicular, hacia el diámetro menor. Por tanto, en este trabajo se ha implementado el cálculo de las direcciones de las tensiones principales para obtener el diámetro menor requerido. En el caso de la Figura 70, el diámetro menor se corresponde con la dirección de la tensión normal σ_2 .

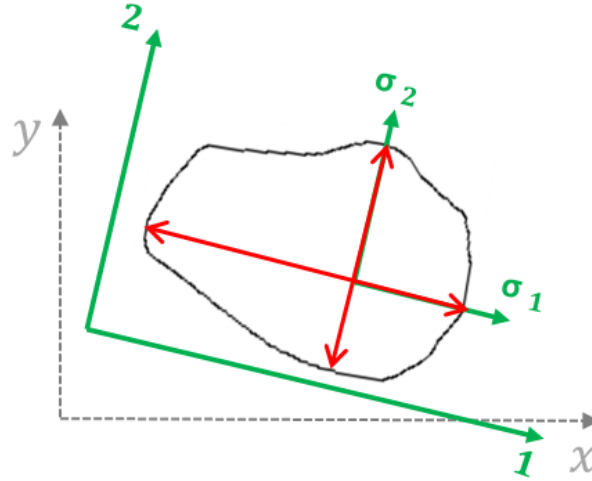


Figura 70: Diámetros mayor y menor (en rojo) calculados en base a las direcciones de las tensiones principales.

Para obtener matemáticamente lo anterior, primero se calcula la matriz del tensor de tensiones de la célula, sobre el sistema de referencia origen, con ejes x e y :

$$T = \begin{bmatrix} \sigma_x & \tau_{xy} \\ \tau_{xy} & \sigma_y \end{bmatrix}$$

Para realizar este cálculo, se realiza un barrido de todos los píxeles de la célula y se almacenan todas las coordenadas de los píxeles en una matriz de dimensiones $2 \times N$:

$$M_{2 \times N} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \end{bmatrix}$$

En esta matriz se almacenan todas las coordenadas (x en la primera fila e y en la segunda) de los píxeles de la célula en cuestión, siendo N el último píxel.

A continuación, se calculan las tensiones normales y tangenciales del tensor de tensiones, T , como covarianzas:

$$\tau_{xy} = \text{Cov}(x, y) = \sum_{i=1}^N \frac{(x_i - \bar{x})(y_i - \bar{y})}{N} \quad (26)$$

$$\sigma_x = \text{Cov}(x, x) = \sum_{i=1}^N \frac{(x_i - \bar{x})^2}{N} \quad (27)$$

$$\sigma_y = \text{Cov}(y, y) = \sum_{i=1}^N \frac{(y_i - \bar{y})^2}{N} \quad (28)$$

donde $x_{i...N}$ e $y_{i...N}$ se corresponden con las coordenadas de los píxeles de la célula, \bar{x} e \bar{y} con la media de la primera y segunda fila de la matriz $M_{2 \times N}$ respectivamente, y N es el número total de píxeles de la célula.

Una vez se obtiene la matriz del tensor de tensiones, se diagonaliza. Con esto, se consigue la matriz del tensor de tensiones principales, así como las direcciones de dichas tensiones. La función de MATLAB “eigs(T)” devuelve:

- Una matriz diagonal que se corresponde con la nueva matriz del tensor de tensiones:

$$T_{1,2} = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \quad (29)$$

- Otra matriz 2x2, cuyas columnas se corresponden con los autovalores o valores propios de la matriz T. Los autovalores de la matriz del tensor de tensores original son, por definición, las direcciones de las tensiones principales. Por tanto, esta matriz proporciona las direcciones de las tensiones principales, es decir, las direcciones de los ejes 1 y 2 (ortogonales) para los cuales la matriz de tensiones principales es nula.

$$V_{\text{propios}} = [\text{Dirección}_1 \text{ Dirección}_2] = \begin{bmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{bmatrix} \quad (30)$$

Como se observa, en la matriz cuyas columnas tiene contenidos los vectores de los valores propios de T, dichos vectores son ortogonales.

En base a las direcciones de las tensiones principales, se calcula el diámetro más largo y el diámetro perpendicular más largo, o diámetro menor. Para ello, se multiplica la matriz de los valores propios (V_{propios}) por la matriz que contiene las coordenadas de los píxeles ($M_{2 \times N}$) de la célula:

$$M_{\text{coordenadas proyectadas}} = V_{\text{propios}} \cdot M_{2 \times N} = \begin{bmatrix} v_1 & -v_2 \\ v_2 & v_1 \end{bmatrix} \cdot \begin{bmatrix} x_1 & x_2 & \dots & x_N \\ y_1 & y_2 & \dots & y_N \end{bmatrix} \quad (31)$$

Se obtienen así las coordenadas de los píxeles de la célula, proyectadas sobre los ejes de las tensiones principales, tal y como se muestra en la Figura 69. Una vez situada la célula sobre los nuevos ejes, se calcula el diámetro menor, tomando el mayor y menor valor en el eje 2 para un mismo valor del eje 1, es decir, buscando la perpendicular más larga al diámetro mayor.

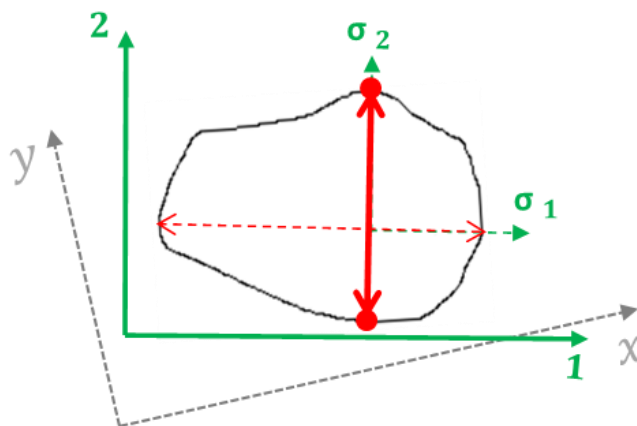


Figura 71: Cálculo del diámetro menor (en rojo), mediante el método de las tensiones principales.

La figura anterior muestra el resultado de la implementación del método de las tensiones principales para calcular el diámetro menor de una célula.

5.2.4 Resultados y comparaciones de métodos de cálculo de diámetros celulares

Para poder decidir cuál de los métodos mostrados ofrece mejores resultados, para incluirlo en el software que se va a proporcionar a los patólogos, se hace necesario compararlos.

La comparación que se ha realizado tiene en cuenta tanto la exactitud con la que se obtienen los diámetros mayor y menor, como el tiempo de computación de los cálculos. El tiempo de computación del cálculo de los diámetros es importante, ya que el objetivo es obtener un método que permita ahorrar tiempo al patólogo.

Para presentar los resultados, se han analizado tres células a escala real. A dichas células se les han aplicado los tres métodos descritos anteriormente para compararlos con una medida manual, la cual se toma como referencia. En el caso del método del diámetro de Feret, se ha estudiado cómo varía este método en el cálculo de los diámetros, girando 1° ó 5° . En las tablas que contienen los resultados, se muestran:

- Los resultados del cálculo del diámetro mayor y menor de cada célula. Estos diámetros se muestran utilizando píxeles como unidad. Sin embargo, en el capítulo 6 se cambiará la unidad a μm .
- El porcentaje de exactitud de las medidas con respecto a los diámetros medidos manualmente.
- El tiempo de cómputo empleado para el cálculo de los dos diámetros. Este tiempo ha sido calculado con un ordenador cuyo procesador es un Intel Core i7-5500U CPU @ 2.40 GHz y en MATLAB R2014a.
- El tiempo estimado que se tardaría en computar 200 células, como realiza un patólogo, mediante la multiplicación del tiempo de cómputo de los dos diámetros por un factor de 200.
- Una Figura de Mérito (Fdm), cuyo valor aumenta cuanto mejor sea el método empleado. La ecuación que sigue esta FdM es la siguiente:

$$\text{FdM}_{\text{diámetros}} = \frac{\% \text{ exactitud diámetro mayor} + \% \text{ exactitud diámetro menor}}{100 \cdot \text{Tiempo de cómputo}} \quad (32)$$

donde en el numerador se encuentra la suma de los porcentajes de exactitud en el cálculo de los dos diámetros, y en el denominador el tiempo de cómputo empleado en calcular dichos diámetros, por un factor de 100 para eliminar el tanto por ciento del numerador.

La primera célula analizada se muestra en la Figura 72, cuyos diámetros se encuentran representados. En primer lugar, se ha realizado la medida manual de ambos diámetros:

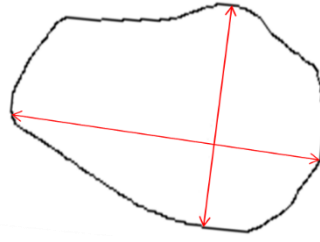


Figura 72: Primera célula analizada, con diámetros en rojo.

$$\text{Diámetro}_{\text{mayor}} = 205.9 \text{ píxeles}$$

$$\text{Diámetro}_{\text{menor}} = 162.8 \text{ píxeles}$$

La siguiente tabla muestra los resultados obtenidos:

	Medida		% exactitud		Tiempo de cómputo	Tiempo estimado para 200 células	FdM
	Mayor	Menor	Mayor	Menor			
Fuerza Bruta	205.2	163	99.7%	99.9%	19.9 s	66.6 min	0.10
Feret (1º)	205.1	163	99.6%	99.9%	15.4 s	51.3 min	0.13
Feret (5º)	205	164	99.5%	99.3%	3.2 s	10.7 min	0.62
Tensiones	204.9	161.8	99.5%	99.4%	0.48 s	1.6 min	4.14

Tabla 11: Resultados para la primera célula analizada.

La FdM calculada, tras ejecutar los tres métodos sobre la primera célula, indica que el mejor es el método de las tensiones principales, con una FdM de 6 a 40 veces mayor que las que aportan los otros dos métodos. Esto se debe a que el error cometido al calcular los diámetros óptimos es despreciable. Además, el método de las tensiones principales ofrece un resultado en tiempo de cómputo total estimado aceptable, con 1.6 minutos de cómputo para 200 células. Por el contrario, los tiempos de cómputo que ofrecen los métodos de fuerza bruta y Feret resultan inviables.

La segunda célula analizada es la siguiente:

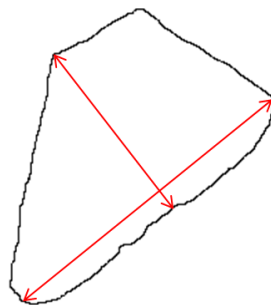


Figura 73: Segunda célula analizada, con sus diámetros en rojo.

La medida manual de sus diámetros ha sido:

$$\text{Diámetro}_{\text{mayor}} = 234.7 \text{ píxeles}$$

$$\text{Diámetro}_{\text{menor}} = 143.5 \text{ píxeles}$$

La siguiente tabla muestra los resultados obtenidos para la segunda célula analizada:

	Medida		% exactitud		Tiempo de cómputo	Tiempo estimado para 200 células	FdM
	Mayor	Menor	Mayor	Menor			
Fuerza Bruta	234.8	143	99.9%	99.7%	19.1 s	63.7 min	0.11
Feret (1º)	234.5	142	99.9%	98.9%	15.1 s	50.3 min	0.13
Feret (5º)	234.3	141	99.8%	98.3%	3.1 s	10.3 min	0.32
Tensiones	233.8	145.2	99.6%	98.8%	0.45 s	1.5 min	4.41

Tabla 12: Resultados para la segunda célula analizada.

Se observa cómo, de nuevo, el bajo tiempo de cómputo y la aceptable exactitud hacen que la FdM del método de las tensiones vuelva a ser la mayor con diferencia. El alto tiempo de cómputo en los otros dos métodos los hacen poco factibles en un software.

Por último, se muestra la tercera célula analizada:

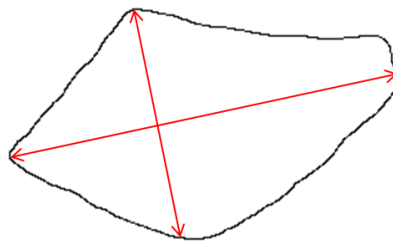


Figura 74: Segunda célula analizada, con sus diámetros en rojo.

La medida manual de sus diámetros ha sido:

$$\text{Diámetro}_{\text{mayor}} = 269.1 \text{ píxeles}$$

$$\text{Diámetro}_{\text{menor}} = 159 \text{ píxeles}$$

La siguiente tabla muestra los resultados obtenidos tras aplicar los tres algoritmos:

	Medida		% exactitud		Tiempo de cómputo	Tiempo estimado para 200 células	FdM
	Mayor	Menor	Mayor	Menor			
Fuerza Bruta	268.3	160	99.7%	99.4%	27.8 s	92.7 min	0.07
Feret (1º)	269.1	160	100%	99.4%	20.2 s	67.3 min	0.1
Feret (5º)	269.1	160	100%	99.4%	4.1 s	13.7 min	0.49
Tensiones	259.7	161	96.6%	98.8%	0.61 s	2 min	3.2

Tabla 13: Resultados para la tercera célula analizada.

Se repite el resultado de los dos casos anteriores. El método de las tensiones proporciona una FdM mucho mayor que los demás, debido a su alto porcentaje de exactitud en el cálculo de los diámetros y por su bajo tiempo de cómputo.

Tras analizar las tres células que se han mostrado, y muchas otras a lo largo de la realización de este trabajo, queda patente que el método de las tensiones principales

proporciona una mejor solución al cálculo del diámetro menor, y es por ello que este método se ha incluido en la interfaz gráfica de la herramienta.

El resto de parámetros necesarios para el patólogo pueden ser calculados a partir de los datos obtenidos en este capítulo. La implementación del cálculo de dichos parámetros se muestra en el siguiente capítulo.

Capítulo 6

Desarrollo de la interfaz gráfica de usuario: MuCSA

En este capítulo se describe el desarrollo de la interfaz gráfica que actúa como herramienta de apoyo al diagnóstico de distrofias musculares. Las siglas MuCSA son el acrónimo de “Muscle Cross Section Analyzer”, o en español, Analizador de la Sección Transversal de Músculo.

En primer lugar, se muestran las posibilidades que ofrece MATLAB a la hora de crear interfaces gráficas. En segundo lugar, se describen las funcionalidades implementadas en la interfaz, las cuales principalmente parten de los algoritmos desarrollados anteriormente, así como de otras características añadidas. En tercer lugar, se describe la utilización de la herramienta a modo de tutorial. En cuarto y último lugar, se ha considerado el análisis del código generado y de los procesos del mismo, para determinar aquellos que son computacionalmente más complejos y con mayor carga temporal, y proponerlos como línea futura para su aceleración.

6.1 Introducción a la creación de interfaces gráficas de usuario con MATLAB

Una interfaz gráfica es el vínculo entre el usuario y un programa computacional, constituida generalmente por un conjunto de comandos, menús, instrumentos o métodos por medio de los cuales el usuario se comunica con el programa durante las operaciones que se desean realizar, facilitando la entrada y salida de datos e información. El nombre en inglés de una interfaz gráfica de usuario es Graphical User Interface, también conocido por su acrónimo GUI.

MATLAB proporciona un lenguaje de alto nivel y herramientas de desarrollo con las que es posible desarrollar y analizar algoritmos, así como crear aplicaciones. MATLAB posibilita en el desarrollo de GUIs la inclusión de listas de selección, menús desplegables, botones de pulsación o de opción, deslizadores, casillas de verificación, entre otros. Resulta muy útil la herramienta GUIDE (Graphical User Interface Development Environment) de MATLAB, ya que permite diseñar y editar GUIs.

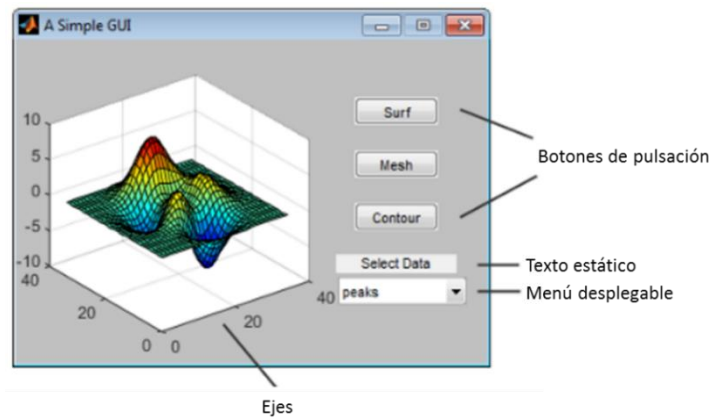


Figura 75: Ejemplo de interfaz creada con GUIDE [34].

Existen dos formas de crear GUIs en MATLAB:

- Mediante GUIDE: escribiendo GUIDE en la línea de comandos de MATLAB, es posible comenzar a crear una figura nueva en el editor de diseño gráfico que se abre. Dicha figura, puede rellenarse con los componentes necesarios. Tras guardar, GUIDE genera un archivo de extensión “.fig” que contiene la ventana o figura de la interfaz, así como un archivo de extensión “.m”, que contiene el código asociado a las funciones de “callback” de los elementos que se han incluido en la interfaz. Estas funciones se programan en función de lo requerido.

Por tanto, el desarrollo de GUIs con GUIDE se realiza en dos etapas:

- o Diseño de los componentes (controles, menús, ejes, etc) que forman el GUI.
- o Programación de la función respuesta (callback) de cada uno de los componentes ante la interacción del usuario.
- Mediante la programación de código: es posible crear directamente un archivo de extensión “.m” que contenga el código que define la ventana, figura o elemento, además de las propiedades de los componentes y las funciones que indican su comportamiento.

Es importante tener en cuenta que, en una GUI, cada objeto está asociado a un identificador (handle) único desde el momento de su creación. A través de este identificador se pueden modificar las características de un objeto gráfico, también llamadas propiedades del objeto. Naturalmente, además se pueden establecer las propiedades de un objeto en el momento de su creación (cambiarlas con respecto a los valores por defecto).

Para la realización de un entorno de apoyo al diagnóstico de distrofias musculares, se ha empleado la herramienta GUIDE junto con figuras creadas mediante programación, que apoyan las funcionalidades del software.

6.2 Desarrollo del entorno: MuCSA

Recordando el procedimiento tradicional para el diagnóstico de distrofias musculares mostrado en la sección 2.4, el patólogo realiza manualmente el siguiente estudio a una muestra, mediante microscopía óptica:

- Cuantificación de fibras Tipo I, Tipo II y la suma total.
- Selección de 100 células aleatorias de cada tipo para medir el diámetro menor de cada una.

Tras realizar el anterior procedimiento, el patólogo continúa calculando los parámetros siguientes:

- Cálculo del porcentaje de fibras Tipo I y Tipo II frente al total.
- Cálculo del diámetro medio de cada tipo y del diámetro medio total.
- Generación de dos histogramas con los diámetros, uno por cada tipo de célula.
- Evaluación de la variabilidad mediante el Coeficiente de Variabilidad (visto en la sección 2.4, con la ecuación (1)).
- Cálculo de los Factores de Atrofia e Hipertrofia de cada tipo y total (vistos en la sección 2.4, con las ecuaciones (2) y (3)).

Se pretende, por tanto, proporcionar al patólogo una herramienta que realice el procedimiento descrito de forma automática, facilitando el diagnóstico. Además, se quiere que esta sea simple, intuitiva y que el especialista no necesite poseer conocimientos técnicos sobre procesamiento digital de imágenes para utilizarla, a diferencia de otras herramientas como las nombradas en el capítulo 3.

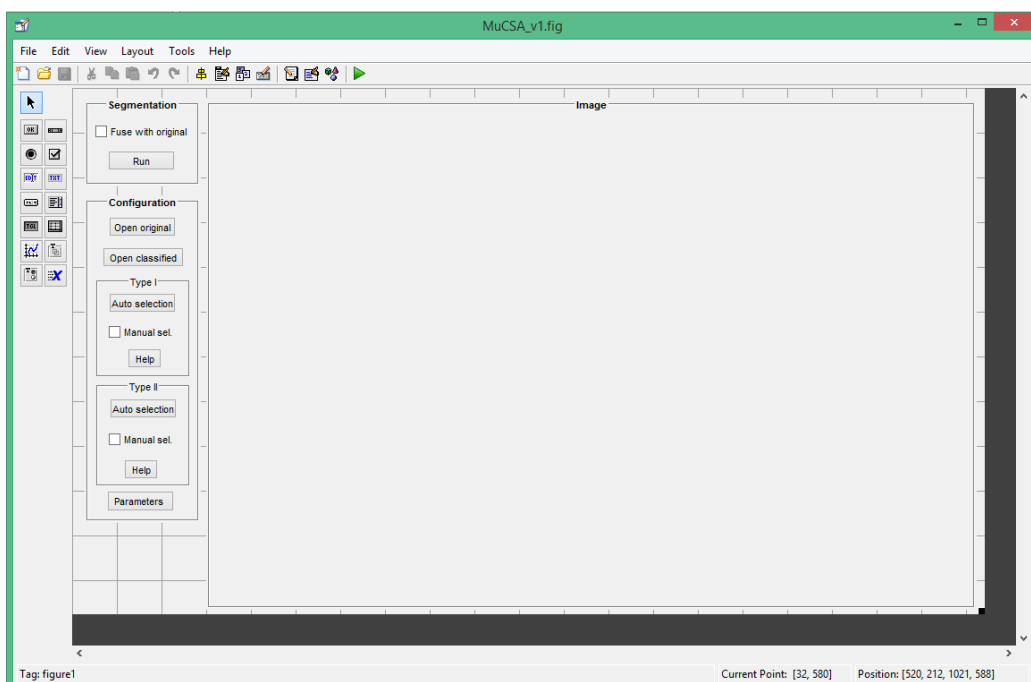


Figura 76: Editor GUIDE, durante la creación de la interfaz gráfica de MuCSA.

Las siguientes subsecciones describen la implementación de las funcionalidades contenidas en el entorno MuCSA, así como el cálculo de los parámetros en los que se apoya el patólogo para realizar un diagnóstico. En la sección 6.3 se exhibe la utilización de MuCSA, a modo tutorial, con una imagen de ejemplo.

6.2.1 Ejecución del entorno y cargar una imagen

El entorno se ejecuta desde MATLAB, escribiendo su nombre en la ventana de comandos. Una vez abierto, el usuario puede cargar una imagen de una muestra de músculo desde el menú habilitado para ello. La extensión de dicha imagen puede ser .jpg, .tiff, .tif, .png y .bmp. Una vez cargada la imagen, el patólogo la observa en la ventana principal de la interfaz y puede proceder al análisis de la misma.

6.2.2 Segmentación de la imagen

Tras cargar la imagen, existe un botón que ejecuta la segmentación de la imagen. Dicha segmentación es llevada a cabo por el algoritmo desarrollado en este trabajo, que es el que mejores resultados ofrece.

Cuando se ha ejecutado la segmentación, en la ventana principal se sustituye la imagen original por la segmentada, quedando en negro los bordes de las células y en blanco el resto. Es posible superponer la imagen segmentada sobre la imagen original en escala de grises en la ventana principal de la herramienta, para que el patólogo realice una inspección visual del resultado de la segmentación si lo desea (ver Figura 77). Asimismo, en cualquier momento se puede volver a visualizar en la ventana principal la imagen segmentada sin fusionar, sin tener que volver a realizar la segmentación, y viceversa.

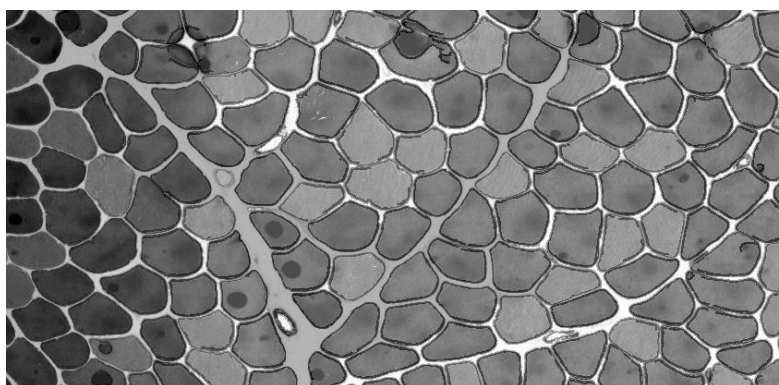


Figura 77: Imagen original en escala de grises fusionada con su segmentada.

La visualización tanto de la fusionada como de la segmentada, indistintamente sobre la ventana principal, es posible gracias a una función desarrollada cuyos argumentos de entrada son la imagen en escala de grises, la segmentada y un “flag” binario cuyo valor 0 implica que no se fusionan ambas imágenes y el valor 1 indica que sí. Este “flag” se activa mediante una casilla de verificación habilitada para ello (ver Figura 78)

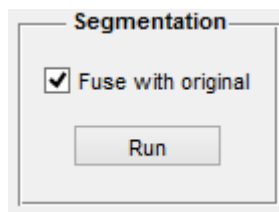


Figura 78: Casilla de verificación para habilitar la superposición de la imagen segmentada sobre la original en escala de grises.

En el momento en el que se ejecuta la segmentación (Run), el software también realiza la cuantificación y clasificación en tipos de las células, tal y como se ha descrito en la sección 5.1. Dicha clasificación no es mostrada en la ventana principal, pero sí es posible abrir una nueva figura con la imagen segmentada y con células clasificadas (como la imagen b) de la Figura 79) tras pulsar un botón habilitado para ello. Además, mediante otro botón, es posible visualizar la imagen original en escala de grises con contraste mejorado, por si el especialista desea comparar visualmente.

6.2.3 Selección de las células cuyos diámetros son calculados

Para seleccionar automáticamente las células Tipo I (claras) a las que realizar el cálculo de los diámetros menores, se aprovecha la clasificación previamente realizada. Se toman 100 células Tipo I de forma aleatoria mediante la comprobación del valor de cada célula en la escala de grises. Como se ha mostrado en la sección 5.1, el valor 150 en la escala de grises se corresponde con las células Tipo I. En cuanto a las Tipo II, se seleccionan 100 aleatorias cuyo valor en la escala de grises, sobre la imagen con las células clasificadas, sea 50. En caso de que haya menos de 100 células de cada tipo en la imagen, se seleccionan todas las células posibles para calcular su diámetro. En la selección de células que se muestra en la Figura 79, se muestran en rosa las seleccionadas del Tipo I y en rojo las seleccionadas del Tipo II. La interfaz incluye una ventana de ayuda que muestra a qué tipo pertenece cada color.

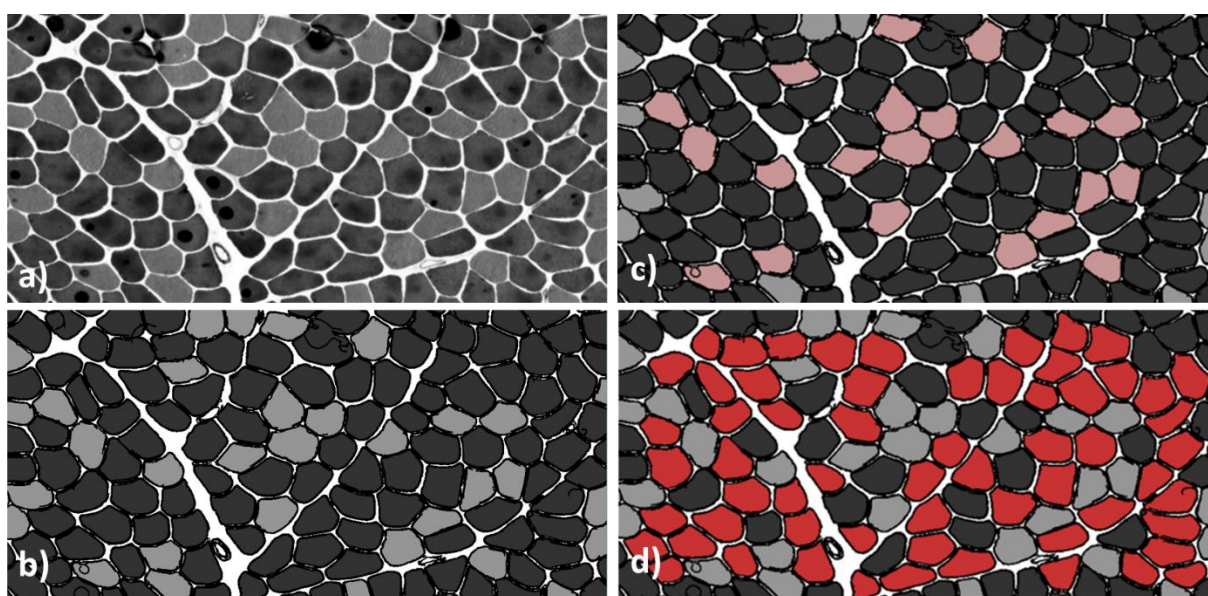


Figura 79: a) Imagen I en escala de grises, b) Imagen I segmentada y con sus células clasificadas en Tipo I (claras) y Tipo II (oscuras), tanto a) como b) se abren en ventanas emergentes tras pulsar los botones pertinentes, c) en rosa las células Tipo I seleccionadas para calcular sus diámetros y d) en rojo las células Tipo II seleccionadas.

En la Figura 79, se observa como en el caso de la selección de células Tipo I, al no haber 100 en la muestra, se ha seleccionado todas. Cabe destacar, que las células localizadas en los márgenes de la imagen no son seleccionadas por el algoritmo, ya que se encuentran incompletas por su posición. Esto se ha conseguido incluyendo una condición en el código, para que, al seleccionar una célula, primero se tenga en cuenta si alguna de las coordenadas de sus píxeles se encuentra en los márgenes de la imagen.

Sin embargo, para la cuantificación de células de cada tipo sí se tienen en cuenta las células de los márgenes, ya que son esenciales en el cálculo del porcentaje de células de cada tipo sobre el total.

En la herramienta, se ha implementado la posibilidad de ofrecer al especialista una interacción con las imágenes con células seleccionadas, para que sea él quien decida si las seleccionadas son de su agrado. Mediante la función “ginput()”, es posible determinar las coordenadas del puntero del ratón sobre la imagen cuando se hace “click” izquierdo. Cuando el especialista clicca sobre un píxel cuyas coordenadas se corresponden con una célula de las que han sido seleccionadas para calcular su diámetro en el procedimiento automático, esta se deselectiona y vuelve a su color correspondiente (50 o 150 en escala de grises), dejando de pertenecer al grupo de células seleccionadas. Asimismo, si clicca sobre un píxel cuyas coordenadas se corresponden con una célula de las que no fueron seleccionadas, esta se selecciona cambiando su color a rosa o rojo, dependiendo del tipo. Esto, da al patólogo la oportunidad de valorar las células visualmente y seleccionar las que sean de su interés o corregir problemas ocasionados en la segmentación de los bordes, como que se haya seleccionado un elemento que en realidad no se corresponda con una célula.

6.2.4 Cálculo de los diámetros

Una vez el especialista está conforme con las células seleccionadas, se procede a calcular los diámetros de todas. Se ha implementado el método de las tensiones principales, cuyos resultados ofrecen una relación exactitud-tiempo de cómputo, determinada en la Figura de Mérito, mejor que los demás métodos.

6.2.5 Obtención y tratamiento de los resultados

MuCSA permite analizar de forma interactiva los resultados obtenidos en forma de parámetros. Cuando el cálculo de los diámetros se ha llevado a cabo, se abre una ventana secundaria, sobre la cual el patólogo puede examinar los diámetros de cada célula de forma independiente.

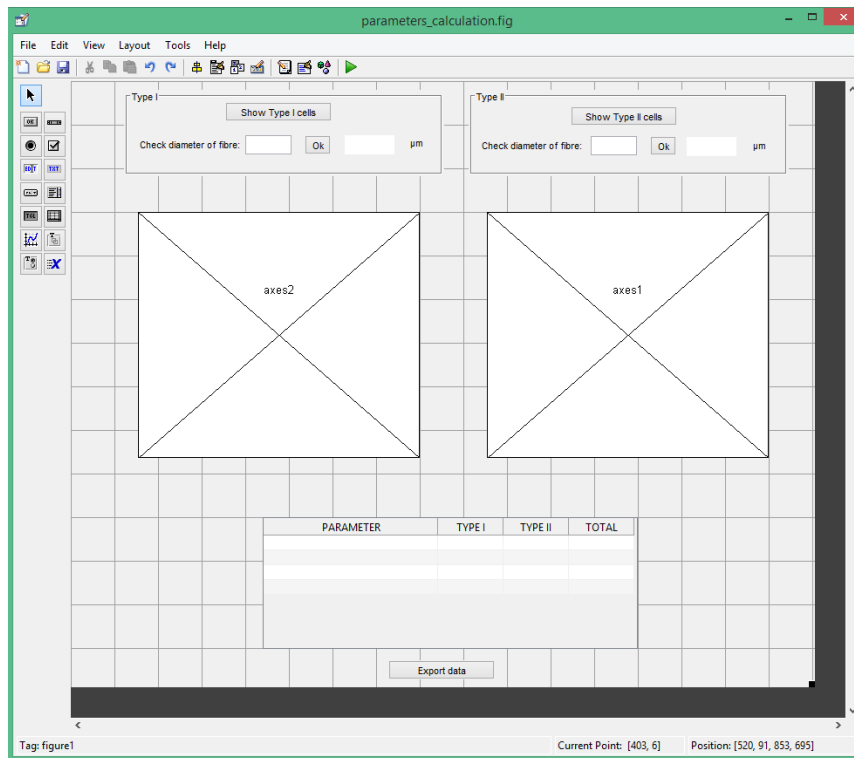


Figura 80: Aspecto de la ventana secundaria que muestra los resultados en el editor GUIDE.

La Figura 80 muestra los elementos que contiene la ventana secundaria, que son:

- Dos paneles, como se observa en la parte superior. Los botones “Show Type I cells” y “Show Type II cells” abre, en ventanas aparte, imágenes que contienen en blanco las células de dicho tipo que han sido seleccionadas para calcular sus diámetros. Estas células, han sido etiquetadas numéricamente y su etiqueta puede verse en su interior (ver Figura 81).

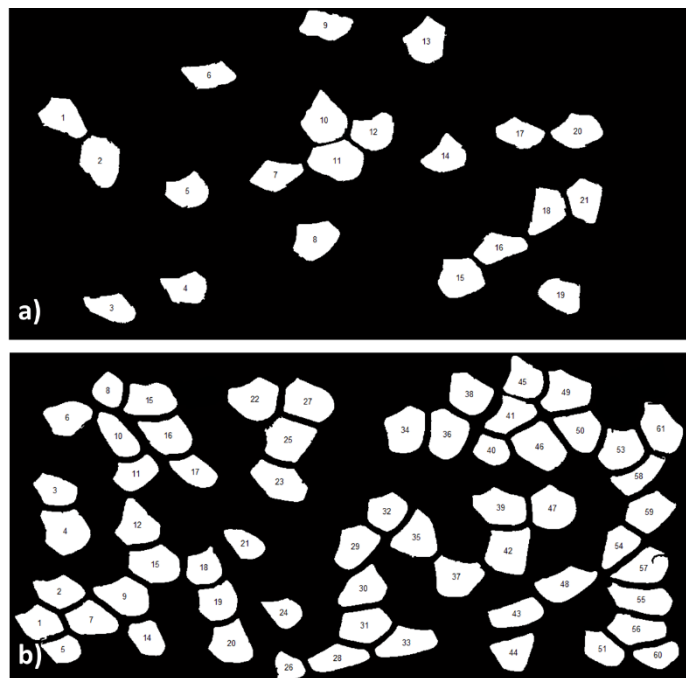


Figura 81: a) células Tipo I etiquetadas y cuyos diámetros han sido calculados y b) las Tipo II.

Los paneles de la parte superior de la Figura 80, permiten al especialista consultar la medida del diámetro de la célula que desee. Basta con escribir el número de la célula (que indica su etiqueta) en la caja de texto editable y pulsar “Ok” (ver Figura 82).

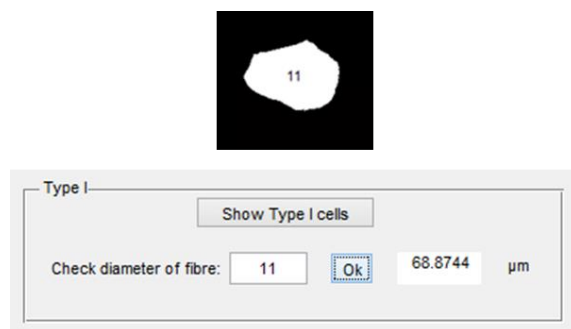


Figura 82: Ejemplo de consulta de un diámetro.

- En medio de la Figura 80, aparecen dos ejes, los cuales contienen los histogramas de los diámetros de las células de cada tipo, como se observa en la Figura 83.

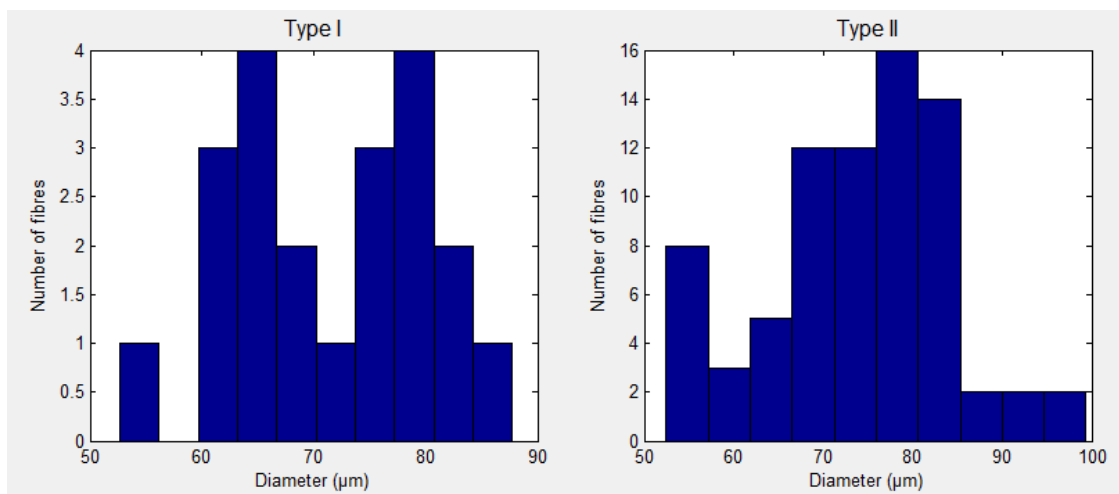


Figura 83: Histogramas de diámetros de las células. A la izquierda las Tipo I y a la derecha las Tipo II.

Para transformar los diámetros calculados de píxeles a μm , que es la unidad requerida en este tipo de procedimientos, se ha pedido al patólogo que mida varios diámetros menores, en μm , de células reales, los cuales se han comparado con sus medidas en píxeles para obtener el factor.

6.2.6 Cálculo de otros parámetros necesarios

El resto de los parámetros se muestran en la tabla que aparece en la parte inferior de la Figura 80 y se corresponden con los descritos en el procedimiento tradicional. En la Figura 84 se observa con mayor detalle.

PARAMETER	TYPE I	TYPE II	TOTAL
Number of fibres	34	121	155
% of fibres to total	21.9355	78.0645	100
Mean of diameters	71.2924	73.8253	73.2769
Standard deviation	9.0087	10.2733	10.2733
Variability coefficient	126.3626	139.1572	136.7829
Atrophy factor	0	0	0
Hypertrophy factor	302.6316	142.8571	268.0412

Figura 84: Tabla con los parámetros resultantes.

Los parámetros obtenidos son, por filas:

- Número de fibras de cada tipo y total en la imagen completa.
- Porcentaje de fibras de cada tipo sobre el total.
- Diámetro medio de cada tipo y total.
- Desviación estándar de cada tipo y total.
- Coeficiente de Variabilidad de cada tipo y total (sección 2.4, ecuación (1)).
- Factor de Atrofia de cada tipo y total (sección 2.4, ecuación (2)).
- Factor de Hipertrofia de cada tipo y total (sección 2.4, ecuación (3)).

6.2.7 Exportación de los resultados a EXCEL

Por último, también se ha implementado la posibilidad de exportar a EXCEL los datos contenidos en la ventana de resultados de la Figura 80. Al exportar los datos, el usuario tiene la posibilidad de guardar el archivo con el nombre que desee. Dicho archivo, de extensión “.xls”, contiene la información en tres hojas. En la primera, se muestra la tabla descrita anteriormente. La segunda contiene la lista de los diámetros de las células Tipo I calculados. Por último, en la tercera, la lista con los diámetros de las células Tipo II.

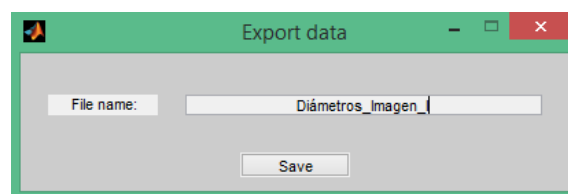


Figura 85: Ventana emergente al pulsar el botón de exportar datos.

Se ha decidido que el formato de salida de los datos sea en EXCEL dado que ofrece un entorno conocido por el patólogo. Además, permite volver a representar los histogramas, realizar otros cálculos que considere necesarios, o simplemente almacenar los datos y revisarlos en otro momento.

6.3 Tutorial de utilización de MuCSA

Esta sección tiene como objetivo mostrar la funcionalidad de la interfaz gráfica, mediante su utilización en un caso real, el de la Imagen I.

En primer lugar, es necesario que el usuario sitúe el ambiente de trabajo de MATLAB en el directorio en el que la interfaz está situada. Para ejecutar MuCSA v1.0, basta con escribir su nombre en la ventana de comandos de MATLAB y pulsar “Enter”. El aspecto que ofrece MuCSA, en su primera versión, es el que se observa en la Figura 86.

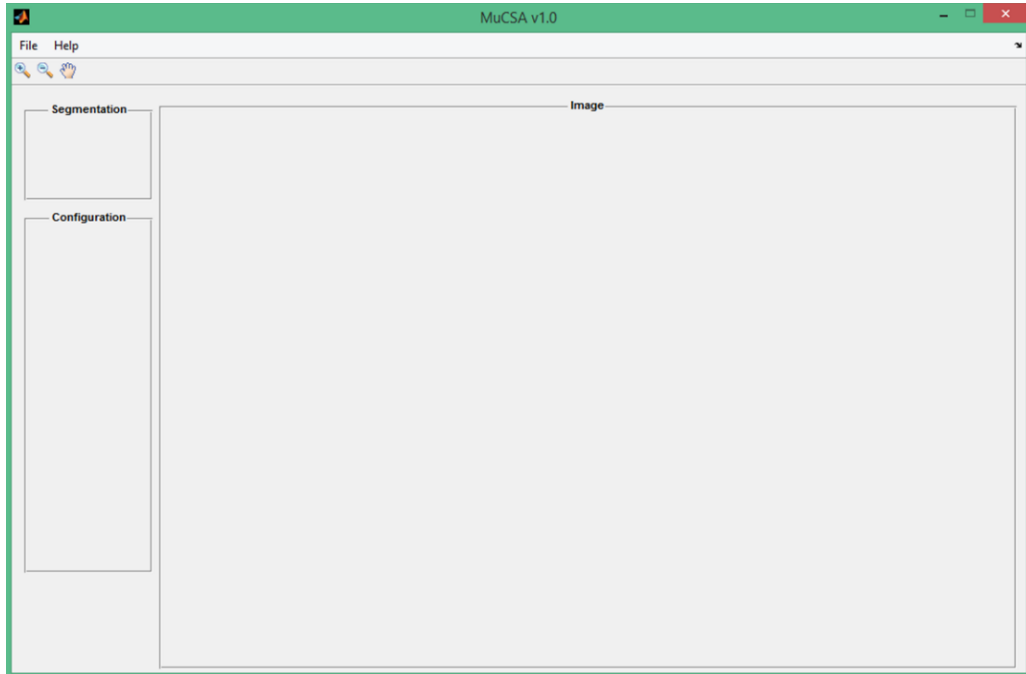


Figura 86: Aspecto de MuCSA v1.0 al ejecutarse desde MATLAB.

Para cargar una nueva imagen y comenzar una nueva sesión de trabajo, es necesario seleccionar el menú “File” y consecutivamente “Load Image”, como se observa en la Figura 87. El menú de ayuda “Help” contiene una versión en “.pdf” de este tutorial.

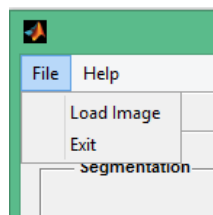


Figura 87: Cargar una nueva imagen.

Tras seleccionar “Load Image”, emerge una ventana en la que el usuario puede seleccionar la imagen en el directorio en el que se encuentre (ver Figura 88).

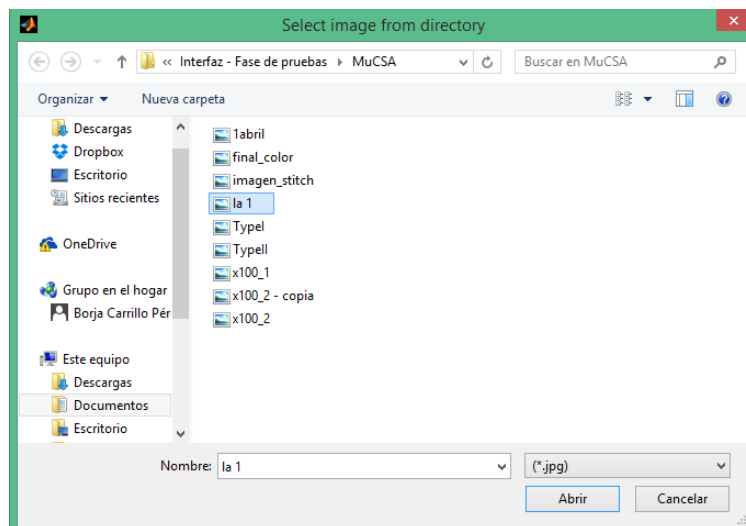


Figura 88: Selección de la imagen desde el directorio en el que se encuentre.

La imagen cargada se muestra en la ventana principal como se observa en la Figura 89. A continuación, en el panel de segmentación, llamado “Segmentation”, aparecen habilitados los elementos de actuación que permiten segmentar la imagen. Para segmentar la imagen, es necesario pulsar el botón “Run”. Es posible marcar la opción “Fuse with original” para fusionar la imagen resultante de la segmentación con la Imagen I en escala de grises (ver Figura 91). También es posible seleccionar (o deseleccionar) la opción cuando el usuario desee, presionando “Run” después, para actualizar la vista en la ventana principal.



Figura 89: Imagen I cargada en la interfaz.

Tras pulsar en “Run”, el resto de opciones en el menú de configuración “Configuration” aparecen habilitadas, como se observa en la Figura 90.

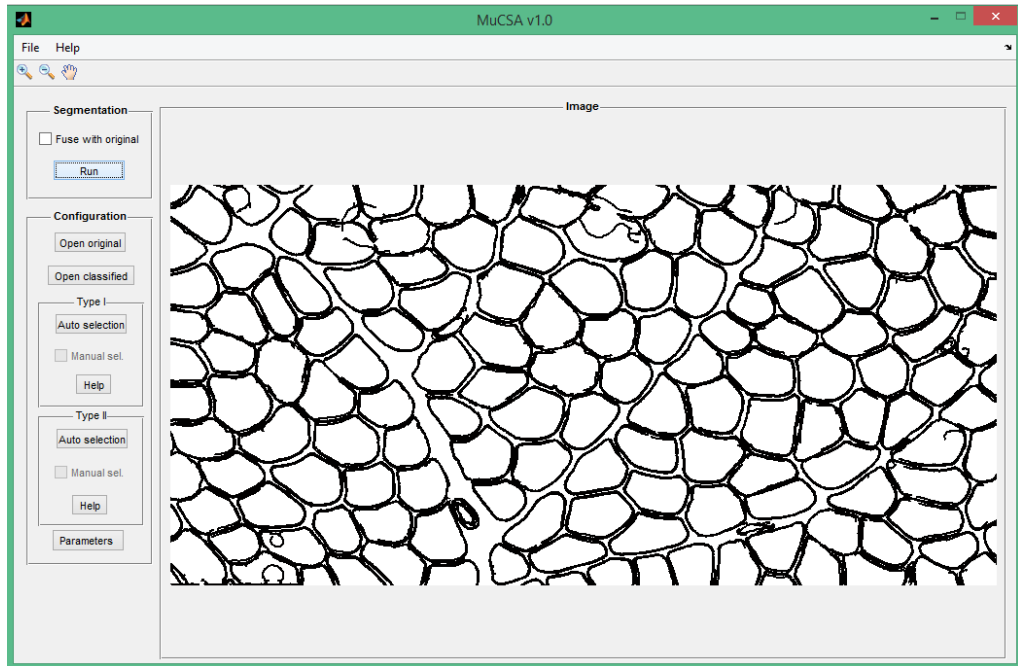


Figura 90: Interfaz con la Imagen I segmentada.

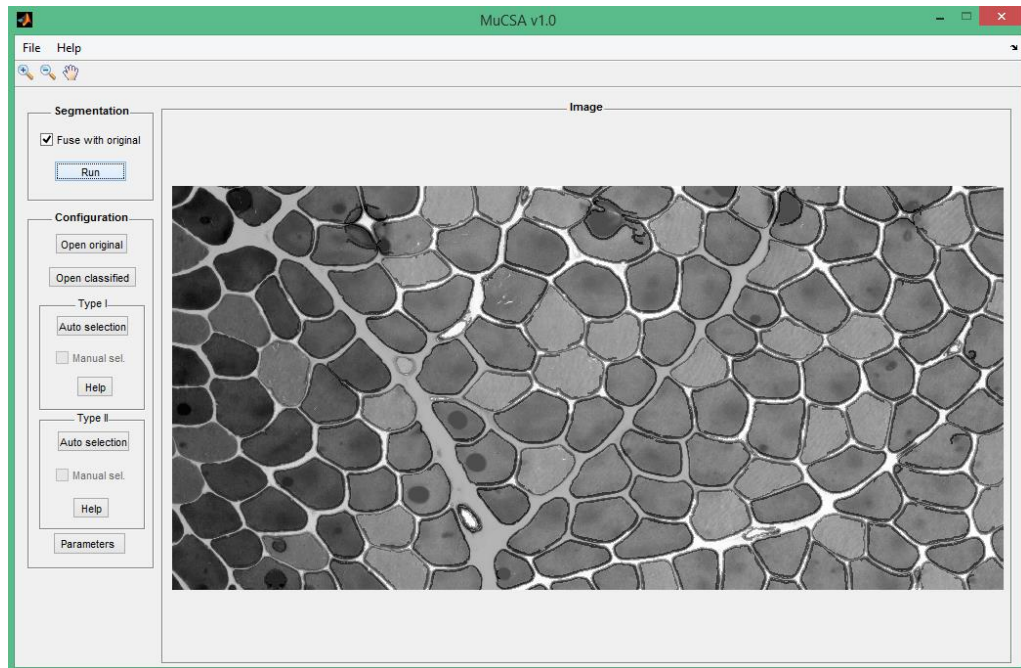


Figura 91: Interfaz con la Imagen I segmentada y fusionada con la original en escala de grises.

Para cualquier imagen que se observe en la ventana principal, es posible realizar una ampliación (“Zoom In”, en el símbolo de la lupa) y navegación (símbolo de la mano).

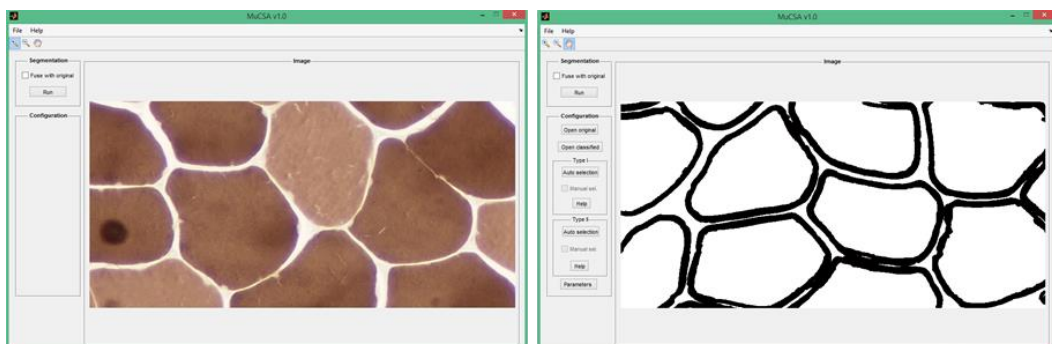


Figura 92: Aplicación de zoom y navegación por las imágenes.

El botón “Open original” muestra en una nueva ventana la imagen original en escala de grises. El botón “Open classified” muestra la imagen segmentada y con la clasificación en tipos ya procesada por el programa. En color gris claro se muestran las células Tipo I y en color oscuro las Tipo II (ver Figura 94).

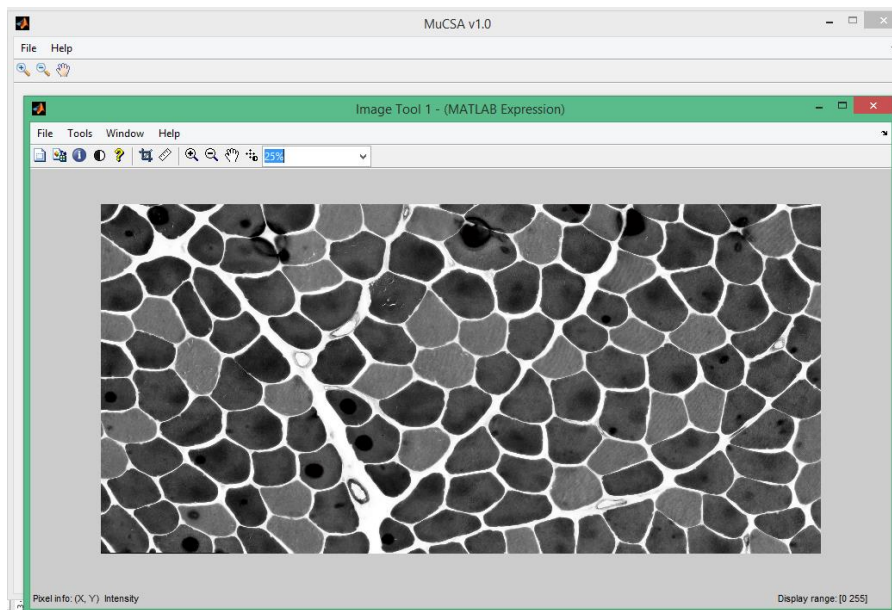


Figura 93: Ventana secundaria con la imagen original en escala de grises, generada tras pulsar “Open original”.

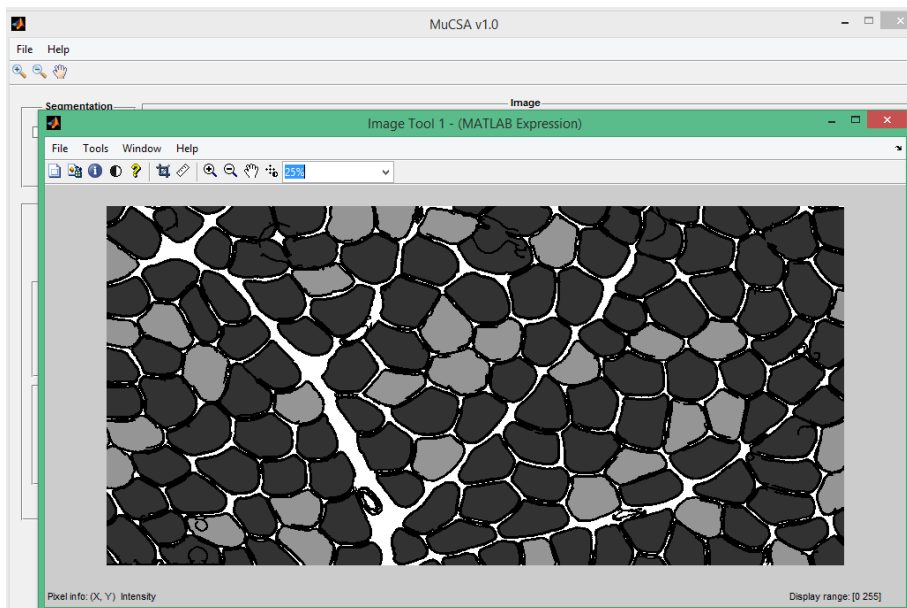


Figura 94: Ventana secundaria con la imagen segmentada y con sus células clasificadas, generada tras pulsar “Open classified”. En gris claro, las células Tipo I. En gris oscuro las Tipo II.

A continuación, el botón de selección automática de células, “Auto selection”, situado en el panel “Type I”, produce la selección automática por parte del programa de las células Tipo I a las cuales se procederá después a calcular sus diámetros menores. En la Figura 95, se observan en rosa las células Tipo I seleccionadas. Al haber menos de 100 células de este tipo en la imagen, se han seleccionado todas, sin embargo, si hubiese más, se hubieran seleccionado automáticamente 100 al azar. Nótese que las células de los márgenes de la imagen no han sido seleccionadas porque no aparecen completas.

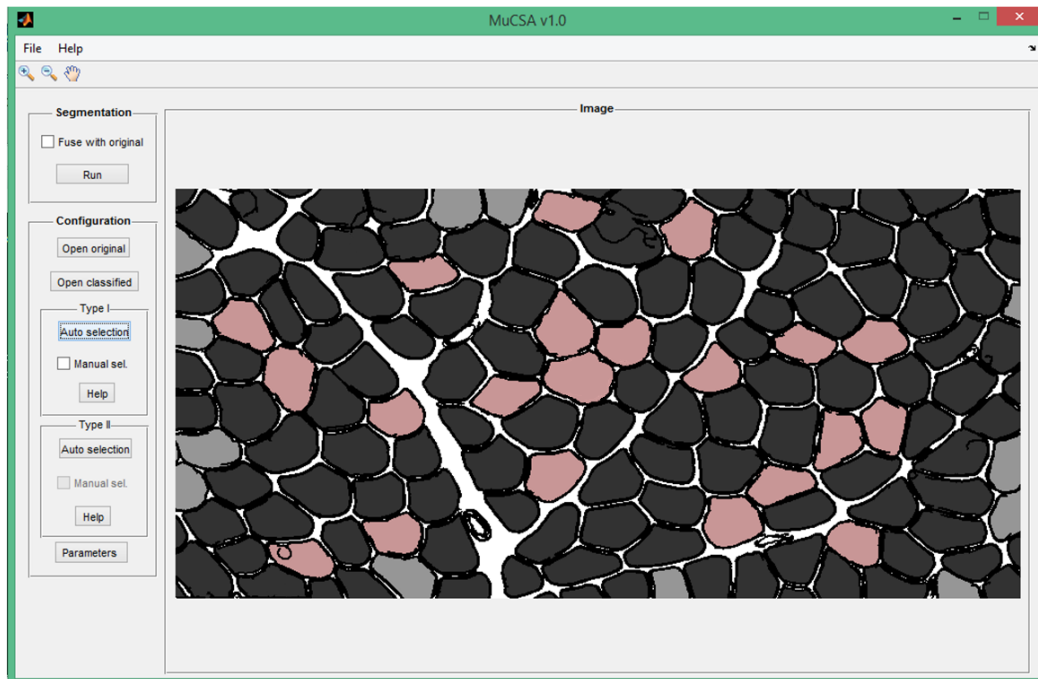


Figura 95: Células Tipo I seleccionadas, en rosa.

Al seleccionar la casilla de verificación “Manual sel.”, un cursor aparece. Haciendo “click” derecho con el ratón, se pueden seleccionar células Tipo I (o Tipo II), en función de lo que el usuario considere necesario, ya sea para subsanar errores o por la necesidad de conocer algún diámetro concreto.

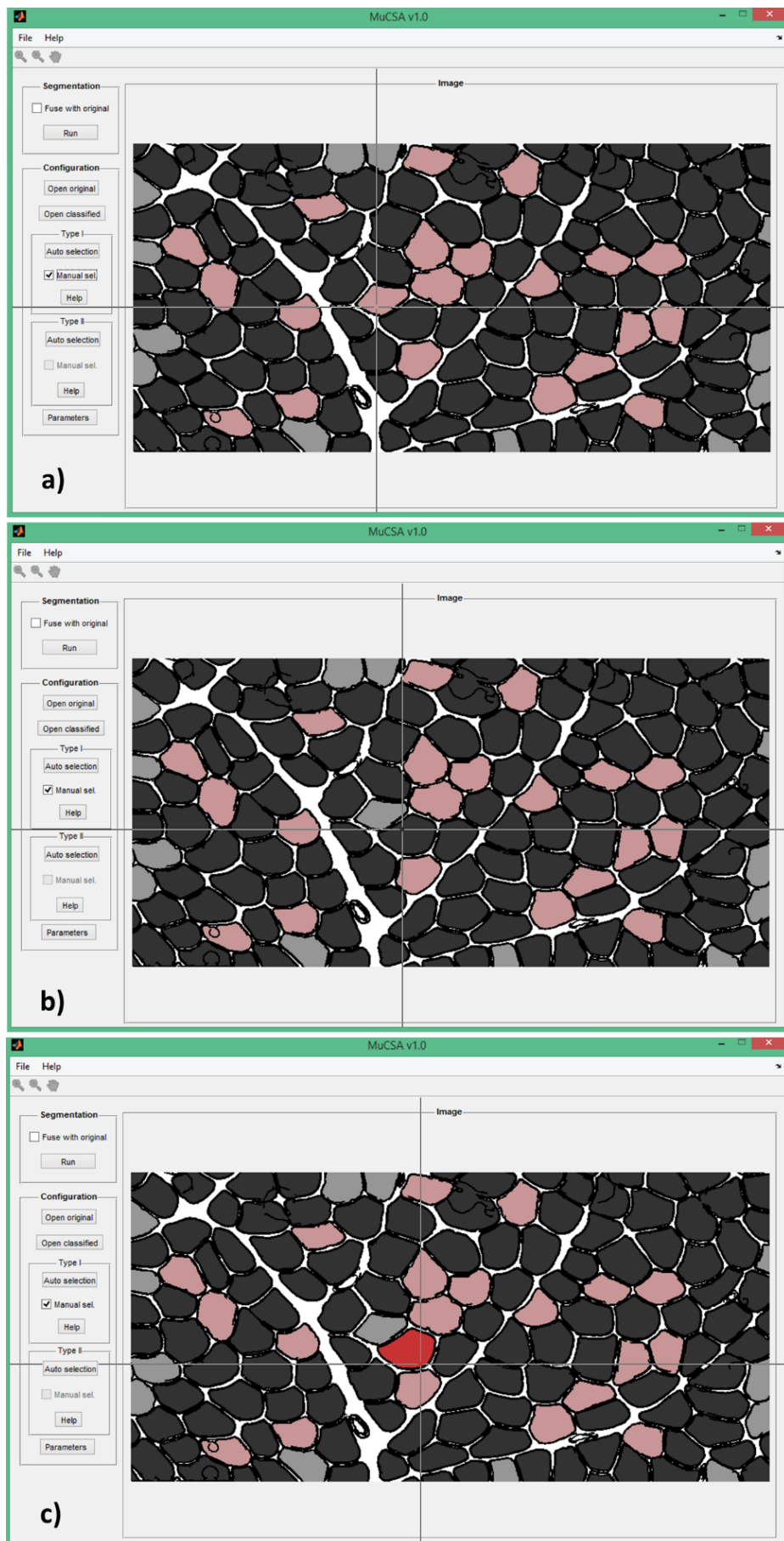


Figura 96: Proceso de selección manual de células Tipo I. a) y b) desección de una célula Tipo I, c) selección de una célula Tipo II como si fuera una Tipo I (por decisión del usuario)

La Figura 96 muestra la desección y selección de células mediante el cursor. Cuando se selecciona una célula nueva, que en programa ha clasificado como Tipo I, esta pasa a ser color rosa. Sin embargo, se puede seleccionar una célula que el programa ha clasificado como Tipo II, en caso de que el usuario haya concluido, mediante inspección visual de la imagen original, que dicha célula es Tipo I. En dicho caso, la célula pasa a ser de color rojo. Este color es diferente al rosa del resto del Tipo I para que el usuario recuerde que el programa había clasificado dicha célula como Tipo II. En el botón de ayuda, “Help”, el usuario puede encontrar la explicación a los colores.

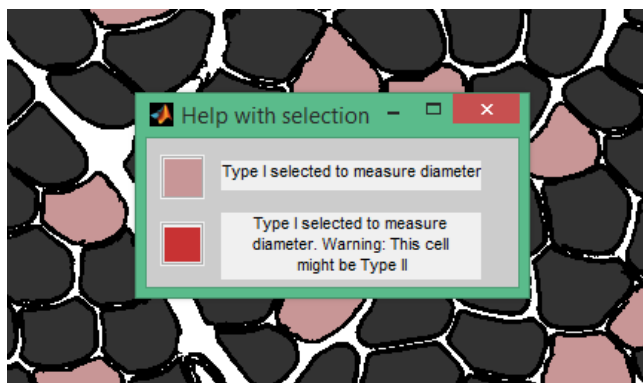


Figura 97: Ventana de ayuda que muestra la explicación a los colores en el caso de la selección de las células Tipo I.

A continuación, se selecciona la opción de selección automática de las células Tipo II. Se sigue el mismo procedimiento que el descrito con anterioridad, con la salvedad de que los colores que muestran las células seleccionadas ahora son al contrario, como se observa en la Figura 99.

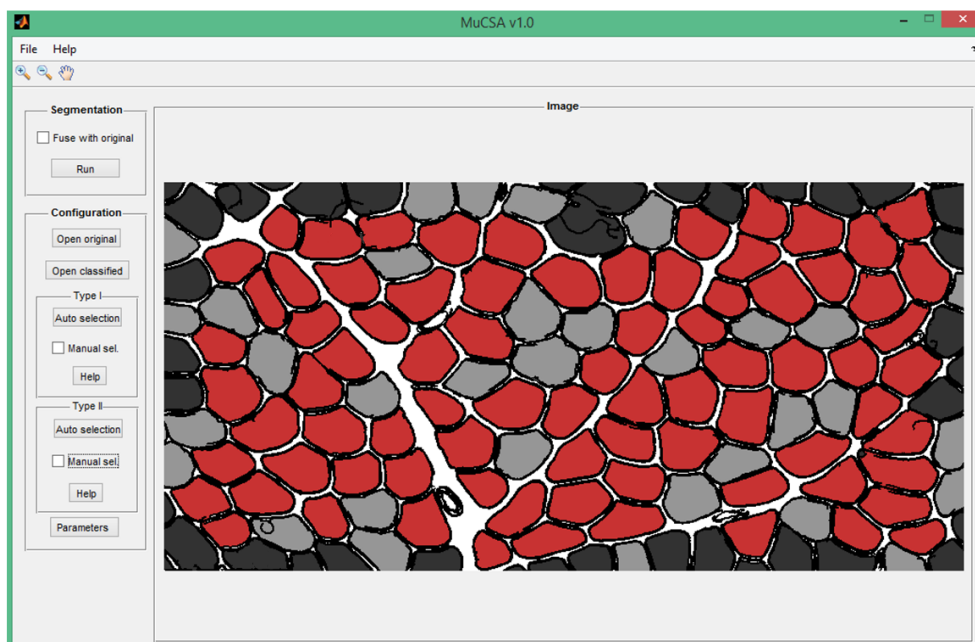


Figura 98: Células Tipo II seleccionadas, en rojo.

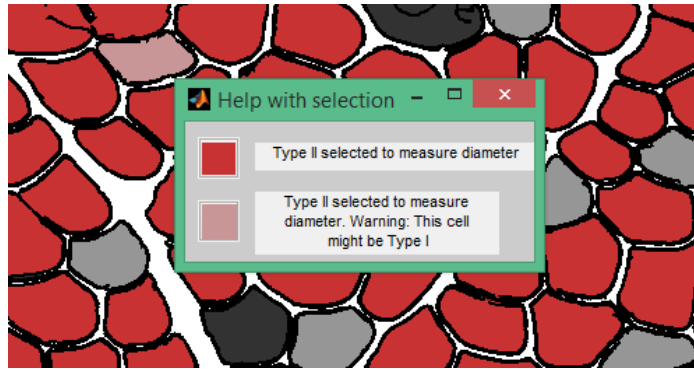


Figura 99: Ventana de ayuda que muestra la explicación a los colores en el caso de la selección de las células Tipo I.

A continuación, se pulsa el botón “Parameters” para que se produzca el cómputo del cálculo de los diámetros y los demás parámetros necesarios. Los resultados se muestran en una ventana secundaria como la que se observa en la Figura 100.

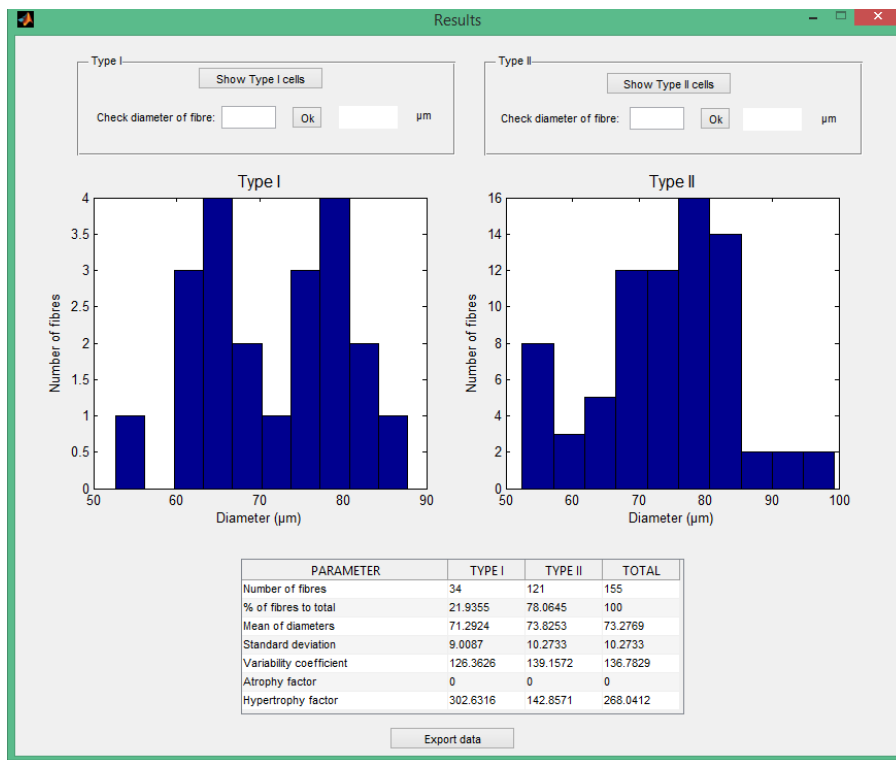


Figura 100: Ventana secundaria con los resultados.

En la parte superior de la Figura 100, se observan dos paneles. Los botones “Show Type I cells” y “Show Type II cells” muestran, en ventanas secundarias, las células Tipo I y Tipo II seleccionadas, de color blanco y el resto en negro. Dichas células están etiquetadas numéricamente y su etiqueta puede observarse en el interior de las mismas (ver Figura 101 y Figura 102).

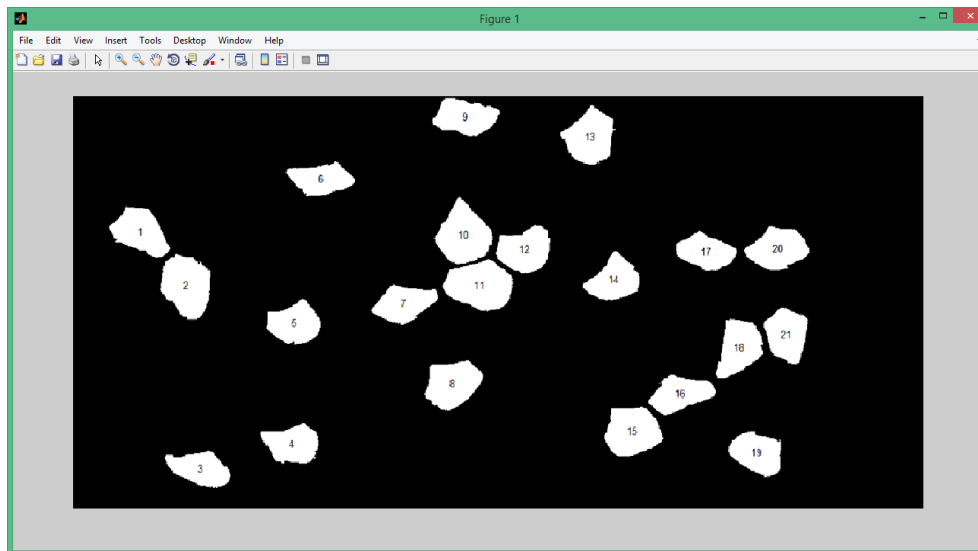


Figura 101: Ventana que emerge tras pulsar “Show Type I cells”. Muestra las células Tipo I, etiquetadas para las cuales han sido calculados los diámetros menores.

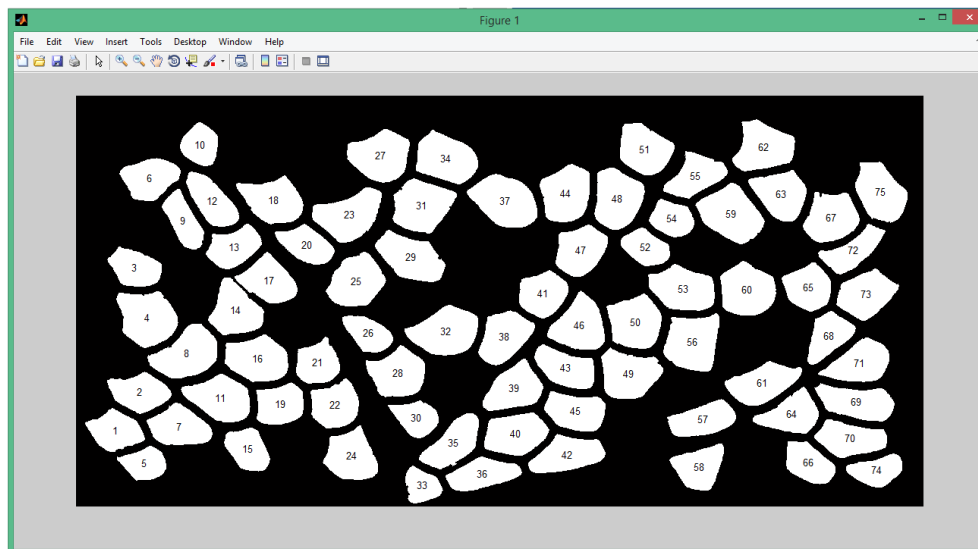


Figura 102: Ventana que emerge tras pulsar “Show Type II cells”. Muestra las células Tipo II, etiquetadas para las cuales han sido calculados los diámetros menores.

Una vez que el usuario conoce cómo puede identificar de forma individual cada célula, puede consultar, si lo desea, el diámetro menor de cualquiera de las células de ambos tipos, en los recuadros habilitados para ello (ver Figura 103).

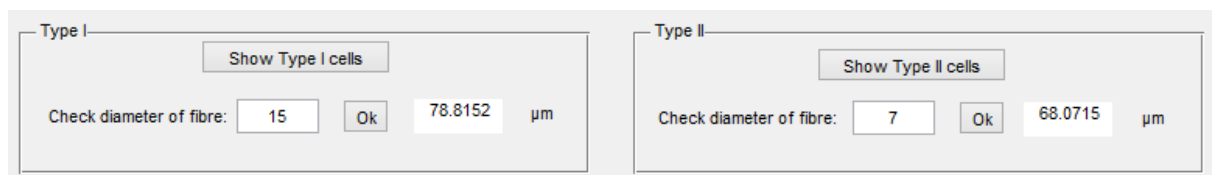


Figura 103: Consulta de diámetros. En este caso el 15 de las células Tipo I y el 7 de las Tipo II.

Los dos histogramas que se observan en medio de la Figura 100, ordenan los diámetros calculados en las células de cada tipo. En el eje de abscisa se muestran valores de diámetros (en μm) y en el eje de ordenada la frecuencia con la que aparecen dichos diámetros.

La tabla de la parte inferior consiste en un resumen de los datos obtenidos a partir de la cuantificación de las células, descritos en la sección 6.2.6.

Por último, es posible exportar todos los datos a EXCEL, en formato “.xls”. Al pulsar el botón “Export”, aparece una ventana emergente en la que el usuario debe incluir el nombre del fichero (ver Figura 104).

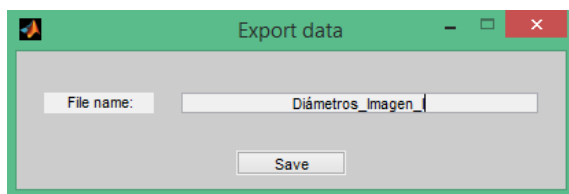


Figura 104: Ventana emergente al pulsar el botón de exportar datos.

El fichero se almacena en el directorio en el que se encuentra el ejecutable de MuCSA, con el nombre que el usuario haya indicado. La información está contenida en tres hojas. En la primera, se muestra la tabla descrita anteriormente. La segunda contiene la lista de los diámetros de las células Tipo I calculados. Por último, en la tercera, la lista con los diámetros de las células Tipo II (ver Figura 105).

	A	B	C	D
1		Type I	Type II	Total
2	Number of fibres	121	34	155
3	% of fibres to total	78,0645161	21,9354839	100
4	Mean of diameters	73,8714904	71,2924256	73,30732
5	Standard deviation	10,334553	9,00869823	10,334553
6	Variability coefficient	139,899073	126,362628	137,382706
7	Atrophy factor	0	0	0
8	Hypertrophy factor	142,857143	306,666667	270,833333

	A	B	C
1	Type I diameters list (µm)		
2	70,7507261		
3	80,8360167		
4	67,1239305		
5	60,3778122		
6	69,2983459		
7	52,6652598		
8	62,6090241		
9	79,286938		
10	60,0663155		
11	87,7544991		
12	78,2189908		
13	75,67983		
14	81,7893775		
15	76,213381		
16	78,8152255		
17	64,3187447		
18	63,2272394		
19	78,4594854		
20	77,1174505		
21	66,6146723		
22	65,9176724		

	A	B	C
1	Type II diameters list (µm)		
2	67,9266756		
3	62,7294804		
4	72,417525		
5	83,5103413		
6	53,0843574		
7	69,1475531		
8	68,0715077		
9	94,6055665		
10	74,9320988		
11	56,9737658		
12	77,1067869		
13	52,4494096		
14	74,4734593		
15	78,4779932		
16	59,4225383		
17	71,8923271		
18	80,9417591		
19	78,8728183		
20	72,8930052		
21	73,766789		
22	71,1918336		
23	79,2990727		
24	95,8596768		
25	80,9670397		
26	72,375159		
27	55,3404947		

Figura 105: Aspecto de las tres hojas de EXCEL, de izquierda a derecha, del fichero “.xls” generado.

6.4 Análisis de rendimiento del software

Se ha analizado la carga temporal de los algoritmos empleados en MuCSA durante su utilización, a fin de determinar aquellos con mayor tiempo de ejecución. El código de los algoritmos ha sido previamente depurado para que el rendimiento sea el mejor

posible. Con este análisis se estudia qué partes del código consumen mayor cantidad de tiempo y podrían ser susceptibles a una mejora.

Para realizar el análisis de rendimiento de la herramienta desarrollada, se ha realizado una investigación de su comportamiento a través de la observación de la información reunida tras su ejecución. El objetivo es averiguar el tiempo dedicado a la ejecución de las diferentes partes del programa, para detectar los puntos donde podría ser llevada a cabo una optimización del rendimiento mediante paralelización de los algoritmos.

Se ha utilizado MATLAB Profiler, que es una herramienta proporcionada por MATLAB, la cual puede ser ejecutada a la vez que el código y que tras terminar devuelve los resultados temporales de las funciones implementadas. Cuando se desea examinar un extracto concreto de código, basta con escribir “profile on” antes del código y “profile viewer” después. Cuando esta última línea se ejecuta, se abre una ventana que contiene un resumen del “profiling” (ver Figura 106). Para los bucles del código que no forman parte de funciones, se ha utilizado el cronómetro de MATLAB, el cual se activa escribiendo “tic” y se desactiva escribiendo “toc”, devolviendo este último el lapso de tiempo.

Primero se han determinado las partes del programa cuyo consumo de tiempo es notablemente mayor a los demás. Después se han analizado dichas partes, tanto con el Profiler como con el cronómetro “tic-toc”, y los resultados obtenidos se muestran en las siguientes subsecciones. Los tiempos han sido calculados con un ordenador cuyo procesador es un Intel Core i7-5500U CPU @ 2.40 GHz, el sistema operativo es Windows 8.1 y la versión de MATLAB es la R2014a.

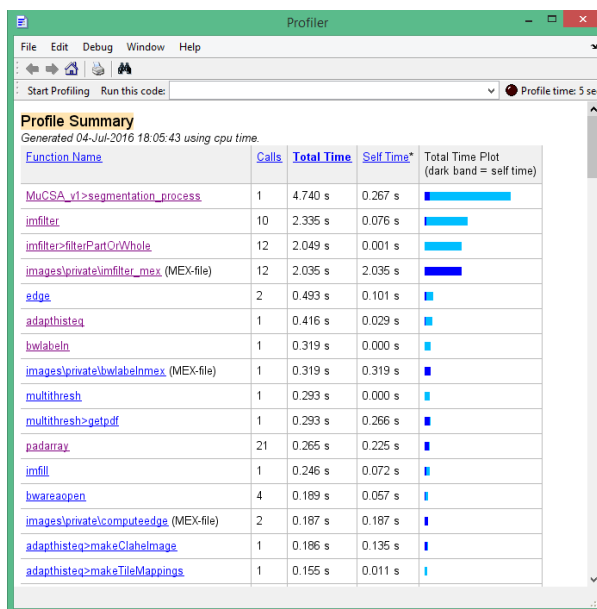


Figura 106: Ejemplo de uso de la herramienta MATLAB Profiler.

6.4.1 Análisis temporal de la segmentación

Para analizar temporalmente la segmentación, se han examinado las cuatro imágenes utilizadas de test en el capítulo 4.

La Tabla 14 muestra el tiempo de cómputo total de segmentación para cada imagen y el porcentaje temporal de las cuatro funciones que más tiempo conllevan, contando todas las veces que estas son llamadas. Estas funciones son:

- `imfilter()`: para el filtrado de las imágenes.
- `edge()`: para detectar bordes en imágenes en función de la intensidad.
- `adapthisteq()`: para mejorar el contraste de la imagen mediante la ecualización del histograma.
- `bwareaopen()`: para eliminar elementos interconectados del tamaño que se imponga.

	Imagen I	Imagen II	Imagen III	Imagen IV
Tiempo total	4.7 s	4.7 s	1.8 s	4.2 s
Función	%Tiempo	%Tiempo	%Tiempo	%Tiempo
imfilter	45.0%	40.9%	43.0%	41.3%
edge	10.4%	8.9%	9.3%	9.3%
adapthisteq	8.8%	10.1%	7.2%	9.0%
bwareaopen	3.9%	10.5%	9.1%	10.8%

Tabla 14: Análisis temporal de la segmentación.

Se observa que el mayor porcentaje de tiempo se emplea en el filtrado de las imágenes, con más del 40% del tiempo que conlleva segmentar por completo una imagen.

6.4.2 Análisis temporal de la clasificación y cuantificación de células

Se ha analizado la parte del código dedicada a la clasificación y cuantificación de células en tipos. Durante dicho análisis, se ha percibido que son dos los bucles del código que ocupan la mayor parte del cómputo temporal de esta parte del programa. Estos son:

- El que atañe a la medida de la luminancia media de cada área aislada por la segmentación, para posteriormente clasificarlas.
- El que, tras haber realizado la clasificación, cuantifica las células que hay de cada tipo.

La Tabla 15 muestra tiempo de cómputo total de la parte de clasificación y cuantificación para las cuatro imágenes de test, así como los porcentajes sobre el tiempo total de los dos bucles comentados.

	Imagen I	Imagen II	Imagen III	Imagen IV
Tiempo total	49.4 s	84.1 s	12.3 s	56.1 s
Función	%Tiempo	%Tiempo	%Tiempo	%Tiempo
Media lumínica	45.4%	45.6%	43.4%	46.2%
Cuantificación	53.4%	53.8%	51.7%	52.76%

Tabla 15: Análisis temporal de la clasificación.

Entre estos dos bucles analizados, se ocupa más del 98% del tiempo de cómputo de la cuantificación y clasificación de las células de la imagen.

6.4.3 Análisis temporal del cálculo de los diámetros

En esta subsección, se estudia la parte del código dedicada al cálculo de los diámetros de las células ya seleccionadas previamente por el programa.

En la Tabla 16 se muestran los tiempos totales de ejecución del cálculo de los diámetros para cada imagen, así como el número de células para las cuales han sido calculados los diámetros y el tiempo medio que ha conllevado calcular cada diámetro.

	Imagen I	Imagen II	Imagen III	Imagen IV
Tiempo total	37.2 s	45.1 s	14.9 s	49.7 s
Células medidas	97	138	107	162
Tiempo medio	0.39 s	0.33 s	0.14 s	0.31 s

Tabla 16: Análisis temporal del cálculo de los diámetros menores.

Capítulo 7

Conclusiones y líneas futuras de trabajo

En este capítulo se exponen las conclusiones a las que se ha llegado con la realización de este Trabajo Fin de Máster, teniendo en cuenta los resultados obtenidos en capítulos anteriores. Además, se muestran las líneas futuras que formarían parte de la continuación de esta investigación.

7.1 Conclusiones

El procedimiento tradicional realizado por los patólogos del Servicio de Anatomía Patológica, implica el reconocimiento y examen de las secciones transversales de músculo esquelético de forma manual, llevando consigo un procedimiento tedioso y largo. En este trabajo se han diseñado e implementado, en MATLAB, algoritmos de procesamiento digital de imágenes para dar una solución automática al procedimiento manual, mejorando de forma clara los tiempos de obtención de los resultados. De forma desglosada, se pueden concluir los siguientes puntos:

- Tras investigar tres posibles soluciones a la segmentación de las imágenes de muestras musculares, queda patente que el tercer método, desarrollado en este trabajo, ofrece mejores resultados con respecto al de K-means y el de Watershed, como indica su Figura de Mérito.
- Se ha determinado que el método de las tensiones principales proporciona una mejor relación en porcentaje de exactitud en el cálculo de los diámetros mayor y menor con el tiempo que conlleva computarlos, con respecto al método de fuerza bruta y el del diámetro de Feret.
- Los algoritmos anteriores han sido unificados en una interfaz gráfica de usuario, llamada MuCSA v1.0, que además permite al patólogo:
 - o Segmentar una imagen de una muestra de músculo, obteniendo el contorno de las fibras de forma automática.
 - o Comprobar el resultado de la segmentación, comparando la imagen segmentada con la original en escala de grises.
 - o Observar la imagen original en escala de grises y con contraste mejorado en una ventana aparte.

- Examinar visualmente la clasificación de células tras segmentar la imagen.
 - Seleccionar de forma aleatoria 100 células de cada tipo, a las cuales se les calculará el diámetro menor posteriormente. En caso de que haya menos de 100 células de cada tipo en la imagen, se seleccionan todas las posibles, menos las que se encuentran en los márgenes, que aparecen incompletas.
 - Corregir manualmente, si lo desea, la selección de las células a calcular los diámetros, o seleccionar nuevas según se considere oportuno.
 - Calcular automáticamente los diámetros menores de las células seleccionadas.
 - Revisar cada célula individualmente.
 - Observar el resultado de los diámetros de cada tipo en los histogramas correspondientes.
 - Comprobar los resultados de los parámetros correspondientes en una tabla resumen. Los parámetros calculados se corresponden fielmente con los descritos en los objetivos.
 - Exportar la tabla anterior y dos listas con los diámetros de cada tipo de célula a un fichero “.xls” de EXCEL.
- Han sido analizados los tiempos de ejecución de los algoritmos para detectar aquellos que conllevan mayor carga temporal, a fin de determinar cuáles de ellos son potencialmente susceptibles de ser mejorados, ya que se ha detectado que sus partes podrían ser ejecutadas independientemente. Estos han resultado ser:
- Las funciones que implican la aplicación de filtros a la imagen, en la segmentación.
 - El bucle dedicado al cálculo de la media lumínica de cada área aislada por la segmentación, para posteriormente clasificar dichas áreas como células Tipo I y II, y el bucle encargado de realizar la cuantificación de células de cada tipo.
 - La función que realiza el cálculo de los diámetros menores de las células.

Con todo lo anterior, se ha dado el primer paso hacia el desarrollo de una herramienta comercial, la cual lleva a cabo el análisis de muestras musculares de forma automática, acelerando el procedimiento entre 20 y 30 veces. Esto quiere decir, que de los 30 a 60 minutos que un patólogo experimentado dedica al análisis de una muestra, sin contar el tratamiento de los datos obtenidos para el posterior cálculo de parámetros necesarios, puede ser reducido a 2 ó 3 minutos. Visto de otra forma, podría acelerarse el trabajo de un año, 365 días, a 10 ó 20 días. Por otro lado, el análisis de rendimiento del software muestra que los con mayor carga temporal son los que conciernen

Por último, se muestra, a modo de tutorial, el funcionamiento de la interfaz gráfica de MuCSA, con la Imagen I como ejemplo. En la última reunión con los especialistas del Servicio de Anatomía Patológica del C.H.U. Insular – Materno Infantil, fue presentada la herramienta MuCSA, cuyo funcionamiento resultó del agrado de los patólogos.

7.2 Líneas futuras

Este trabajo cuenta con una amplia proyección futura, la cual puede estar orientada a un producto comercial en el que un sistema reconfigurable acelere aún más los procedimientos. Las líneas futuras entre las que podría continuar este trabajo, pueden ser las siguientes:

- Aumentar el número de imágenes para conseguir una mayor robustez.
- Proporcionar la herramienta al Servicio de Anatomía Patológica para que realicen el procedimiento tradicional en paralelo a la utilización de la herramienta, para después de un tiempo prudencial, poder tomar los datos obtenidos y generar un modelo estadístico con el que se pueda determinar el grado de confianza del software y mejorarlo si fuera preciso.
- Paralelizar los algoritmos con mayor carga temporal para la optimización del programa. Dichos algoritmos poseen paralelismo, es decir, la posibilidad de procesarlos mediante la realización de varias operaciones por unidad de tiempo, ya que dichas operaciones no dependen unas de otras.

Los algoritmos encargados de la medida de la media lumínica de cada área aislada podrían ser paralelizados, ya que podría calcularse la media lumínica de cada área aislada de forma independiente. Asimismo, la cuantificación de células de cada tipo también, ya que se podría cuantificar cada tipo en dos hilos paralelos. Por último, el algoritmo relacionado con el cálculo de los diámetros menores podría ser implementado con computación paralela, mediante el cálculo de diámetros en grupos de células al mismo tiempo, por ejemplo.

La paralelización podría ser llevada a cabo mediante la implementación de dichas partes del código en sistemas reconfigurables, que permitieran conseguir una buena diagnosis de una forma más eficiente mediante la aceleración de los algoritmos mencionados y con vista a crear un sistema más compacto. Entre estos se encuentran las FPGAs, las GPUs o las CPUs multi-core, debido a su potencial para acelerar en gran medida una amplia variedad de aplicaciones. Por un lado, una de las ventajas de las plataformas como FPGAs y GPUs es la arquitectura paralela que estos dispositivos poseen. Por otro lado, generalmente poseen menores frecuencias de reloj que las CPUs convencionales.

Finalmente, cabe comentar que, con el contenido de este trabajo, se pretende redactar un artículo que será enviado a BIOIMAGING 2017 (4th International Conference on Bioimaging [35]) y a CASEIB 2016 (XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica [36]), con el objetivo de ser presentado.

Referencias

- [1] L. P. Gartner y J. L. Hiatt, Texto Atlas de Histología, 2^a edición, McGraw-Hill, 1995.
- [2] V. Dubowitz, C. A. Sewry, A. Oldfors y R. Lane, Muscle Biopsy: A Practical Approach, London: ELSEVIER, 2013.
- [3] «Muscular Dystrophy,» National Institute of Neurological Disorders and Stroke. Office of Communications and Public Liaison, 2009. [En línea]. Available: http://espanol.ninds.nih.gov/trastornos/distrofia_muscular.htm. [Último acceso: 12 Junio 2016].
- [4] «Federación Española de Enfermedades Raras,» [En línea]. Available: <http://www.enfermedades-raras.org/>. [Último acceso: 12 Junio 2016].
- [5] F. E. d. E. Neuromusculares, Guía de las Enfermedades Neuromusculares: Información y apoyo a las familias, Alcalá la Real (Jaén): Formación Alcalá, 2008.
- [6] C. Ridaura-Sanz, «Biopsia muscular,» *Acta Pediátrica de México*, vol. 29, n^o 6, pp. 347-54, 2008.
- [7] R. C. Gonzalez, R. E. Woods y S. Eddins, Digital Image Processing Using Matlab, Natick, USA: Gatesmark Publishing, 2009.
- [8] Mathworks, «Marker-Controlled Watershed Segmentation,» [En línea]. Available: <http://es.mathworks.com/help/images/examples/marker-controlled-watershed-segmentation.html>. [Último acceso: 12 Junio 2016].
- [9] M. J. Jiménez y R. González, «Procesamiento de imágenes en el dominio espacial,» Grado en Ingeniería Informática, Universidad de Sevilla, 2015.
- [10] N. Otsu, «A Threshold Selection Method from Gray-Level Histograms,» *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, n^o 1, pp. 62-66, 1979.
- [11] R. J. Herrera, H. Velasco, H. A. Roncancio, C. A. Polanco y J. Rosado, «Segmentación de imágenes de microanatomía de madera mediante una clasificación multiescalada basada en operaciones morfológicas,» *Ciencia, Investigación, Academia y Desarrollo*, vol. 7, n^o 2, pp. 64-72, 2002.
- [12] M. Mathworks, «Structural Elements Classes,» [En línea]. Available: <http://es.mathworks.com/help/images/ref/strel-class.html>. [Último acceso: 9 Junio 2016].
- [13] P. Tzekis, A. Papastergiou, A. Hatzigaidas y A. Cheva, «A sophisticated edge detection method for muscle biopsy image analysis,» de *7th WSEAS International Conference on Signal, Speech and Image Processing*, Beijin, 2007.
- [14] M. Megías, «Atlas de histología vegetal y animal. Técnicas histológicas. Protocolos,» Departamento de Biología Funcional y Ciencias de la Salud , Universidad de Vigo, 2016.

- [15] O. Sertel, B. Dogdas, C. Chi Sung y M. Gurcan, «Microscopic image analysis for quantitative characterization of muscle fiber type,» *Computerized Medical Imaging and Graphics*, vol. 35, pp. 616-628, 2011.
- [16] F. Liu, A. Mackey, R. Srikuea, E. K y Y. L, «Automated image segmentation of haematoxylin and eosin stained skeletal muscle cross-section,» *Journal of Microscopy*, vol. 252, nº 3, pp. 275-285, 2013.
- [17] C. Pertl, M. Eblenkamp, A. Pertl, S. Pfeifer, E. Wintermantel, H. Lochmüller, M. Walter, S. Krause y C. Thirion, «A new web-based method for automated analysis of muscle histology,» *BMC Musculoskeletal Disorders*, pp. 2-9, 2013.
- [18] Definiens, «Cognition Network Technology,» [En línea]. Available: <http://www.definiens.com/tissue-phenomics#tissue-phenomics-details>. [Último acceso: 7 Junio 2016].
- [19] S. LifeScience, «S.CORE,» [En línea]. Available: <http://www.sco-lifescience.de/technology.php5>. [Último acceso: 7 Junio 2016].
- [20] A. Sáez, E. Rivas, A. Montero-Sánchez, C. Paradas, B. Acha, A. Pascual, C. Serrano y L. M. Escudero, «Quantifiable diagnosis of muscular dystrophies and neurogenic atrophies through network analysis,» *BMC Medicine*, pp. 1-11, 2013.
- [21] Y.-J. Kim, T. Brox, W. Feiden y J. Weickert, «Fully Automated Segmentation and Morphometrical Analysis of Muscle Fiber Images,» *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, vol. 71, pp. 8-15, 2007.
- [22] W. S. Rasband, *ImageJ*, Maryland, USA: U. S. National Institutes of Health, 1997-2016.
- [23] P. Karen, M. Stevanex, V. Smerdu, E. Cvetko, L. Kubinova y I. Erzen, «Software for muscle fibre type classification and analysis,» *European Journal of Histochemistry*, vol. 53, nº 2, pp. 87-95, 2009.
- [24] ViDiTo, «Ellipse software,» [En línea]. Available: <http://www.ellipse.sk/>. [Último acceso: 2016 Junio 8].
- [25] J. Dubach-Powell, «Quantitative determination of muscle fiber diameter (minimal Feret's diameter) and percentage of centralized nuclei,» *TREAT-NEUROMUSCULAR NETWORK DISEASES*, pp. 1-16, Last reviewed 2014.
- [26] D. I. Systems, «OLYMPUS: Your Vision, Our Future,» [En línea]. Available: <http://www.dis-imaging.gr/OLYMPUS/software.html#>. [Último acceso: 7 Junio 2016].
- [27] L. R. Smith y E. R. Barton, «SMASH - semi-automatic muscle analysis using segmentation of histology: a MATLAB application,» *Skeletal Muscle Journal*, vol. 4, pp. 1-21, 2014.
- [28] R. Szeliski, «Image Alignment and Stitching,» Microsoft Research Coporation, Redmond, WA, 2006.

- [29] M. Sahu y K. Parvathi, «Segmentation of Colour Data Base Image by Implementing K-Means Clustering,» *International Journal of Innovations in Engineering and Technology*, vol. 2, nº 4, pp. 229-34, 2013.
- [30] K. Teknomo, «K-means Clustering Tutorial,» Escuela Colombiana de Ingeniería, [En línea]. Available: <http://people.revoledu.com/kardi/tutorial/kMean>. [Último acceso: 13 Junio 2016].
- [31] J. Tan, «Segmenting cells in a monolayer using fluorescently-tagged cadherin,» 24 Junio 2015. [En línea]. Available: <http://jiongyitan.me/work/2015/06/24/cell-segmentation/>. [Último acceso: 16 Junio 2016].
- [32] J. Piña y V. Bucalá, «Procesamiento de cólidos, capítulo 2: propiedades de una partícula,» Universidad Nacional del Sur, Buenos Aires, 2013.
- [33] J. Santo Domingo Santillana, «Resistencia de Materiales,» Escuela Politécnica Superior de la Universidad de Salamanca, Zamora, 2008.
- [34] Mathworks, «GUIDE,» [En línea]. Available: <http://es.mathworks.com/>. [Último acceso: 28 Junio 2016].
- [35] BIOIMAGING2017, «4th International Conference On BIOIMAGING,» 2017. [En línea]. Available: <http://www.bioimaging.biostec.org/CallForPapers.aspx>.
- [36] CASEIB2016, «XXXIV Congreso Anual de la Sociedad Española de Ingeniería Biomédica,» 2016. [En línea]. Available: <http://caseib16.es/>.

