



BUAP Benemérita Universidad
Autónoma de Puebla

Facultad de Ciencias de la Computación

“Medición de la concentración de iones de calcio en
células de tejido endotelial”

T E S I S

Para obtener el grado de:

Maestría en Ciencias de la Computación

Presenta:

Pedro García Ramírez

Director de Tesis:

Dr. César Bautista Ramos

Asesor de Tesis:

Dr. Rafael Lemuz López

21 de enero de 2021

Contenido

| | |
|--|--------------------------------------|
| RESUMEN | 5 |
| INTRODUCCIÓN | 6 |
| DESCRIPCIÓN DEL PROBLEMA | 6 |
| MOTIVACIÓN | 6 |
| PREGUNTAS DE INVESTIGACIÓN | 6 |
| OBJETIVOS | 7 |
| CONTRIBUCIÓN | 7 |
| CAPÍTULO I CONCEPTOS PREVIOS | 8 |
| 1.1 EL SISTEMA CIRCULATORIO | 9 |
| 1.2 BIOLOGÍA CELULAR ENDOTELIAL | ¡Error! Marcador no definido. |
| 1.3 MICROSCOPIA | 10 |
| 1.4 REVISIÓN DEL ESTADO DEL ARTE | 10 |
| 1.5 MATERIALES Y MÉTODOS | 12 |
| CAPÍTULO 2 SEGMENTACIÓN DE NÚCLEOS CELULARES | 14 |
| 2.1 HISTOGRAMA DE IMAGEN | 15 |
| 2.2 ECUALIZACIÓN DEL HISTOGRAMA | 16 |
| 2.3 SEGMENTACIÓN DE IMÁGENES BASADA EN INTENSIDAD | 17 |
| 2.3.1 SEGMENTACIÓN DE IMAGEN POR AGRUPACIÓN DE K-MEANS | 18 |
| 2.3.2 SEGMENTACIÓN POR FUZZI C-MEANS | 20 |
| 2.3.3 SEGMENTACIÓN BASADA EN MODELADO DE MEZCLAS | 22 |
| 2.3.4 COMPARATIVA | 24 |
| 2.4 SEGMENTACIÓN DE IMAGEN POR MODELADO DE CAMPO ALEATORIO DE MARKOV | 26 |
| 2.4.1 SECUENCIA ALEATORIA DE MARKOV | 27 |
| 2.4.2 CAMPOS ALEATORIOS | 28 |
| 2.4.3 SISTEMAS DE VECINDAD | 28 |
| 2.4.4 DISTRIBUCIÓN DE GIBBS | 28 |
| 2.4.5 CAMPO ALEATORIO DE MARKOV | 29 |
| 2.4.6 MODOS CONDICIONALES ITERADOS (ICM) | 29 |
| 2.5 ALGORITMOS DE RECOCIDO SIMULADO | 31 |
| 2.5.1 ALGORITMOS DE METRÓPOLIS | 31 |
| 2.5.2 MUESTREO DE GIBBS | 33 |

| | |
|---|--------------------------------------|
| 2.5.3 COMPARATIVA..... | 34 |
| CAPITULO 3 MÉTODO DE CARACTERIZACIÓN PROPUESTO | 34 |
| 3.1 PRE PROCESAMIENTO | ¡Error! Marcador no definido. |
| 3.2 MÉTODO DE UMBRALIZACIÓN ADAPTATIVO | 34 |
| 3.2.1 LA UMBRALIZACIÓN | 34 |
| 3.3 ALGORITMO DE PAGERANK | 36 |
| 3.4 APLICACIÓN DEL ALGORITMO A LA CARACTERIZACIÓN DE CÉLULAS..... | 37 |
| 3.5 TRIANGULACIÓN DE DELAUNAY | 39 |
| 3.6 EJEMPLO DELAUNAY Y PAGERANK CON 6 CÉLULAS | 40 |
| CAPITULO 4 RESULTADOS EXPERIMENTALES | 43 |
| 4.1 RESULTADOS DE LOS ALGORITMOS DE SEGMENTACIÓN | 43 |
| 4.2 ANÁLISIS DE LOS DATOS..... | 43 |
| 4.3 K-MEANS..... | 44 |
| 4.4 FUZZY C MEANS..... | 45 |
| 4.5 MODELADO DE MEZCLAS | 45 |
| 4.6 ICM..... | 46 |
| 4.7 RESULTADOS ALGORITMO METRÓPOLIS..... | 46 |
| 4.8 RESULTADOS DEL MUESTREO DE GIBBS | 47 |
| 4.9 RESULTADOS DE LA MEDICIÓN DE LAS CONCENTRACIONES DE CALCIO Y EL ÁREA..... | 49 |
| 4.9.1 GRANULOMETRÍA DIGITAL..... | 50 |
| 4.10 RESULTADOS DE LA CARACTERIZACIÓN CON PAGERANK..... | 51 |
| 4.11 SEGUNDO ANÁLISIS CON NUEVAS CÉLULAS | ¡Error! Marcador no definido. |
| CONCLUSIONES Y TRABAJO FUTURO | 61 |
| REFERENCIAS..... | 63 |

RESUMEN

Esta tesis describe un sistema para medir la concentración de las corrientes iónicas de calcio en células del tejido endotelial analizando secuencias de imágenes. Se propone un método para caracterizar la variabilidad de la concentración de iones de calcio Ca^{2+} y la forma de las células antes y después de infringir una lesión al tejido endotelial. Las corrientes iónicas de sodio, potasio y calcio de las células regulan diversos procesos fisiológicos como el flujo de la actividad eléctrica en los órganos y la cicatrización de tejidos expuestos a lesiones. El método propuesto ubica el área de la lesión, segmenta los núcleos y citoplasma celular en imágenes de fluorescencia, mide la concentración de Ca^{2+} y el área de los núcleos celulares. Finalmente se analiza la concentración de calcio promedio de las células con el algoritmo PageRank de un grupo de células alrededor de la lesión.

En esta tesis se presentan conceptos previos para un mejor entendimiento sobre el problema aquí presentado, en el capítulo 2 se conocerán diferentes tipos de segmentación, así como los conceptos de la granulometría digital.

En el capítulo 3 se presenta el algoritmo de PageRank, la triangulación Delaunay y su aplicación del algoritmo a la caracterización de células.

En el capítulo 4 se muestran los resultados cuantitativos de la medición de las oscilaciones de los niveles de concentración de Ca^{2+} y el área del núcleo en el tiempo antes y después de la lesión y la caracterización con el algoritmo de PageRank, este algoritmo nos ayuda a saber cuáles son las células más importantes alrededor de la lesión.

INTRODUCCIÓN

DESCRIPCIÓN DEL PROBLEMA

Describir cual es el problema que motiva el estudio de la variabilidad en la concentración de calcio en las células epiteliales. PORQUE ES IMPORTANTE HACER ESTA MEDICION

MOTIVACIÓN

El proceso de análisis de la información tradicionalmente requiere que un experto realice diversas tareas repetitivas sobre las imágenes que tiene como resultado un proceso tedioso propenso a errores. Es por esta razón que se desarrolló un sistema de procesamiento y análisis de la información automatizada que utiliza técnicas de procesamiento y análisis de imágenes digitales para caracterizar la variabilidad de la concentración de calcio en células endoteliales antes y después de una lesión.

En específico se propone un algoritmo de detección de núcleos celulares en imágenes microscópicas que detecte regiones de alta luminosidad. Calcular el índice de variabilidad de la concentración de calcio en el núcleo calculando el promedio de las intensidades de los píxeles que representa una célula o conjuntos de ellas.

PREGUNTAS DE INVESTIGACIÓN

¿Cómo cambia la concentración de los iones calcio (Ca^{2+}) en las células del tejido endotelial cuando se infringe una lesión?

Los iones de calcio son utilizados por las células para regular múltiples actividades, el Ca^{2+} es utilizado como una señal intracelular para el control de diferentes procesos, el mecanismo general de acción del Ca^{2+} es el siguiente: cuando las concentraciones del Ca^{2+} citoplasmático se mantienen en niveles bajos, las células permanecen quiescentes, sin embargo, cuando esas concentraciones se elevan dichas células se activan para realizar sus funciones específicas.

¿La concentración de calcio en las células que permanecen después de la lesión es diferente con respecto a la distancia del corte?

El incremento de la concentración de Ca^{2+} se realiza predominantemente en las células que se encuentran más cerca del área de la lesión, una vez cargadas las células más cercanas el Ca^{2+} se sigue regando hacia las células siguientes. Las células con mayor distancia al área del corte no se cargan del Ca^{2+} esparcido en el momento. La distancia de las células que se cargan con Ca^{2+} depende en gran medida del tamaño de la lesión.

OBJETIVOS

En esta tesis se propone desarrollar un algoritmo de procesamiento digital de imágenes para caracterizar la variabilidad del índice de calcio en las células del endotelio vascular con respecto a la distancia de las células a la región de lesión.

ESCRIBIR OBJETIVOS

CONTRIBUCIÓN

Un algoritmo que utiliza un método de segmentación y seguimiento de células en el tiempo para describir el comportamiento de las células a través del tiempo y cómo cambian después de infligir una lesión. También se podrá identificar cuáles son las células más importantes dentro de la imagen con las células que se estén estudiando.

CAPÍTULO I CONCEPTOS PREVIOS

Las corrientes iónicas de sodio, potasio y calcio de las células regulan diversos procesos fisiológicos como el flujo de la actividad eléctrica en los órganos y la cicatrización de tejidos expuestos a lesiones. Las células endoteliales (CE) controlan el tono vascular, la activación y agregación plaquetaria y la angiogénesis. La disfunción endotelial finalmente resulta en trastornos cardiovasculares severos, como hipertensión, aterosclerosis, enfermedad de las arterias coronarias y derrames cerebrales, que pueden causar la muerte del paciente [1].

La restauración del revestimiento endotelial se puede lograr mediante la activación de las células endoteliales vecinas liberadas por inhibición de contacto y por las células progenitoras endoteliales circulantes (EPC). La señalización intracelular de Ca^{2+} es esencial para promover la curación de heridas.

La identificación de núcleos celulares teñidos con fluorescencia es la base de la detección celular, la segmentación y la extracción de características en experimentos de microscopía de alto contenido. La morfología de las células también es uno de los indicadores esenciales de la variación de características o rasgo observable.

La segmentación celular utilizando técnicas de procesamiento digital de imágenes es un componente fundamental para analizar la información que aparece en una secuencia de video, para después hacer la extracción de características en función de la morfología como la delgadez, redondez o su intensidad de Ca^{2+}

Un aumento en la concentración intracelular de Ca^{2+} representa una de las vías de señalización más importantes mediante las cuales las CE ejercen su control sofisticado de la homeostasis cardiovascular, que es el funcionamiento del sistema cardiovascular se mantiene en un nivel de eficiencia funcional adecuado a las condiciones que cada conducta requieres [2].

La biología moderna es una ciencia cada vez más cuantitativa con la que se obtienen conocimientos sobre los procesos biológicos microscópicos y macroscópicos. Se requiere un enfoque abstracto que utilice estadísticas y modelos para inferir patrones y tendencias que no se pueden reconocer de manera intuitiva. Las técnicas de microscopía cada vez más sofisticadas y diversas están produciendo conjuntos de datos más grandes y complejos. Las tecnologías ópticas y la instrumentación proporcionan capacidades inimaginables en la adquisición de imágenes. Por ejemplo, la microscopía robótica ha aumentado la velocidad y la capacidad con que los microscopios recopilan datos de imágenes de experimentos. Las consecuencias de estas tendencias han llevado a la

necesidad de enfoques automatizados de análisis de imágenes computarizados para recopilar datos de los conjuntos de datos basados en imágenes [3].

1.1 EL SISTEMA CIRCULATORIO

En organismos complejos, el sistema circulatorio facilita el transporte de la linfa y la sangre. El sistema que soporta la circulación de la linfa se conoce como el sistema linfático, y el sistema responsable de la circulación de la sangre se conoce como el sistema cardiovascular. El sistema cardiovascular es una red cerrada que comprende: la sangre, el corazón y los vasos sanguíneos. Es extenso, de modo que casi todas las células en la mayoría de los tejidos de un vertebrado están ubicadas dentro de 50 μm a 100 μm de un capilar [4].

Si no se mantiene la integridad física de la red cerrada, distribuida y presurizada que forma el sistema cardiovascular, se produce un sangrado y una pérdida de presión. Estos tienen un serio impacto en la viabilidad de un organismo y, por lo tanto, existen mecanismos para mantener la integridad física del sistema. Dos mecanismos principales que mantienen la integridad física del sistema cardiovascular y sustentan el trabajo para esta tesis son:

1. La capacidad del sistema para reparar agujeros en las paredes de los buques.
2. La capacidad de permitir una salida controlada a través de la pared del vaso donde se requiere la reparación del tejido subyacente.

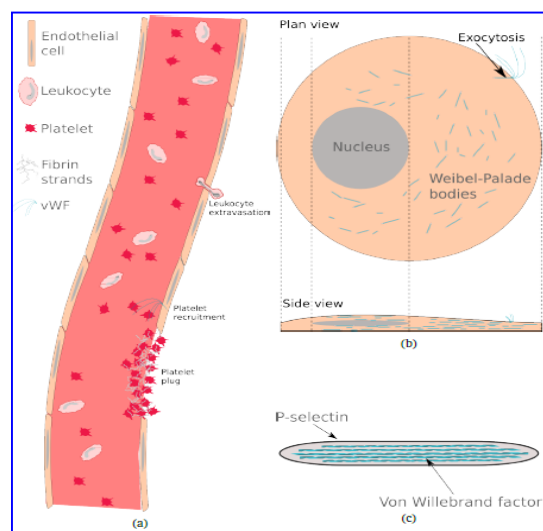


Figura 1. Diagramas para dilucidar la maquinaria para la reparación cardiovascular a escala vascular, celular y de orgánulos.

La figura 1 muestra el revestimiento endotelial de un vaso sanguíneo con plaquetas y leucocitos circulantes, responsables de proporcionar primeros auxilios en el entorno vascular.

1.2 MICROSCOPÍA

Existen tres categorías generales de microscopía: microscopía óptica, microscopía electrónica y microscopía con sonda de barrido. La más antigua es la microscopía de luz, que utiliza una serie de lentes para crear una imagen ampliada de un objeto a través de la difracción, reflexión o refracción de las ondas electromagnéticas en el espectro de luz visible. Los avances en microscopía de luz desde el siglo XVII han llevado a la plétora de instrumentos diversos y sofisticados disponibles en la actualidad [9].

Las alternativas más recientes a la microscopía de luz son la microscopía electrónica y la microscopía con sonda de barrido. Imágenes de microscopía electrónica de una muestra por la interacción de un haz acelerador de electrones con la muestra. Un haz de electrones acelerado da un poder de resolución más alto que el disponible con microscopía de luz visible. Esta resolución más alta se puede lograr porque los electrones excitados pueden tener una longitud de onda más corta que la luz visible. La tercera técnica de microscopía es la microscopía con sonda de barrido, esta familia diversa de métodos escanea la muestra para estudiar las propiedades de la superficie a través de una sonda física. Los microscopios de exploración de túneles (STM) y los microscopios de fuerza atómica (AFM) son dos tipos prominentes de microscopios de sonda de exploración.

La microscopía óptica es una herramienta clave en la biología celular porque: la resolución es muy adecuada para obtener imágenes de las estructuras celulares, la luz es relativamente imperturbable, lo que significa que las células vivas se pueden visualizar a lo largo del tiempo, y se puede usar una variedad de sondas fluorescentes para marcar proteínas y orgánulos [10].

Las imágenes adquiridas mediante métodos de microscopía óptica forman la base para el procesamiento y análisis de imágenes.

1.4 REVISIÓN DEL ESTADO DEL ARTE

En el artículo “Mediciones in vitro de las fuerzas hemodinámicas y sus efectos en la mecánica de células endoteliales a nivel subcelular”. Publicado en Octubre del 2018 [9]; se presentan mediciones de velocimetría de rastreo de micro partículas sobre monocapas de células endoteliales aórticas bovinas cultivadas en micro canales.

El objetivo fue cuantificar las fuerzas de fluidos y la morfología celular a escala subcelular para mono capas sometidas a velocidades de corte constantes [11]. La velocimetría de seguimiento de partículas (PTV) es una técnica de medición de flujo estándar en dinámica de fluidos. Implica el uso de una cámara de video para grabar secuencias de flujo sembradas con pequeñas partículas flotantes neutrales.

En el artículo “Regulation of Blood Flow In the Microcirculation: Role of Conducted Vasodilation.” Publicado en Diciembre del 2011 [12]; se presenta un sistema de procesamiento de imágenes para facilitar la cuantificación de cómo las células endoteliales dentro del primer arco aórtico y proximales al corazón del pez cebra se comportan en respuesta al inicio del latido cardiaco.

En el artículo “Calcium signaling in endothelial cells. Cardiovascular Research.” Publicado en el 2000 [13]; los autores evalúan los efectos del marcador fluorescente PKH26 sobre la proliferación de células endoteliales e identificar las condiciones de etiquetado que produciría la mayor fluorescencia en el tiempo sin afectar negativamente a la viabilidad celular.

En el artículo “Morfología precisa que preserva la segmentación de las células superpuestas en función de los contornos activos.” Publicado en 2016 [14]; presenta que las células endoteliales no activadas expresan actividad anticoagulante, anti adhesiva y vasodilatadora, mientras que la célula endotelial activada expresa actividad pro coagulante, pro adhesiva y vasoconstrictora.

En el artículo “Imagej for microscopy. Biotechniques.” Publicado en 2007 [15]; se presentan las funciones del endotelio, la principal función del endotelio es regular el flujo y la perfusión sistémicos a través de cambios en el diámetro y en el tono vascular. La disfunción endotelial (DE) comprende una pérdida del balance entre los factores vasodilatadores y vasoconstrictores derivados del endotelio, donde el estado vasoconstrictor llega a ser dominante, llevando a cambios patofisiológicos progresivos. Colectivamente, estos cambios endoteliales exhiben características proinflamatorias, prooxidantes, proliferativas, procoagulantes y de adhesión vascular. El proceso de aterosclerosis comienza temprano en la vida y la DE contribuye a la aterogénesis en cada fase de la aterosclerosis. La disfunción endotelial es un marcador de la aterosclerosis y es el precursor más claramente definido de la aterosclerosis.

En el artículo “Assessment of tumour viability in human lung cancer xenografts with texture-based image analysis.” Publicado en 2015 [16]; se presenta un modelo que puede detectar núcleos celulares y su morfología incluso en cultivos celulares de alta confluencia con muchos núcleos celulares superpuestos. Combinan el modelo de contorno activo "gas de círculos cercanos", que favorece las formas circulares pero permite ligeras variaciones a su alrededor, con un nuevo modelo de datos.

Se ha propuesto una gran variedad de algoritmos de detección de células individuales. La mayoría de los métodos de segmentación [17] usan umbrales locales o globales, generalmente basados en el histograma de las intensidades de imagen. Otros métodos utilizan propiedades inherentes de los valores de intensidad de la imagen, como la textura, para detectar celdas con patrones característicos.

Los métodos de aprendizaje automático supervisados [18] y no supervisados [19] han demostrado su utilidad práctica en aplicaciones de detección unicelular: superan ampliamente las técnicas de segmentación clásica al combinar información derivada de imágenes multiparamétricas y superficies de decisión no triviales. Aunque estos métodos solo se enfocan solo en la detección de células y su segmentación.

1.5 MATERIALES Y MÉTODOS

Las imágenes y el procedimiento experimental detallado se describen en [20], aquí presentamos un resumen de la forma en que se obtuvieron las imágenes.

Ratas MaleWistar de 2 a 3 meses de edad fueron provistas por las instalaciones de animales de Benemérita Universidad Autónoma de Puebla, donde fueron mantenidas en condiciones de temperatura ambiental, constante y expuestas a ciclos de luz oscura de 12 hrs., con agua y alimento a voluntad. Todos los procedimientos experimentales en animales se realizaron de acuerdo con los protocolos aprobados por el Comité de Uso y Cuidado de Animales de la Universidad.

El procedimiento para obtener las imágenes de CE a partir de aorta de rata fue desarrollado por Berra-Romani y Cols., 2004 y se describe brevemente a continuación. El día del experimento, las ratas se sacrificaron con una sobredosis de solución intraperitoneal de ketamina-xilazina, 0,2 ml por 100 g de peso; posteriormente se sometieron a una toracotomía anterior para exponer el arco aórtico y el corazón. Se extrajo el corazón para canular y perfundir la aorta con una solución salina fisiológica (PSS). La aorta torácica y abdominal se diseccionaron y se colocaron en una placa de Petri con PSS a temperatura ambiente.

Usando un microscopio estereoscópico (Nikon SMZ-2T), se eliminaron los tejidos conectivo y graso que rodeaban la aorta. Posteriormente, la aorta se cortó en anillos de ~ 5 mm de ancho, se almacenó en PSS a temperatura ambiente (22-24 C) y se usó dentro de las 5 h. Usando unas tijeras de micro disección, los anillos aórticos se cortaron cuidadosamente para abrirse y se tuvo cuidado para evitar cualquier daño al endotelio y obtener tiras aórticas con endotelio intacto.

Las tiras aórticas con endotelio intacto se bañaron en PSS con 16 μmol de Fura-2 / AM durante 60 minutos a temperatura ambiente, se lavaron durante 30 minutos para permitir la desesterificación intracelular de Fura-2 / AM y se fijaron (con la cara luminal

hacia arriba) al fondo de una placa de Petri cubierta por silicona inerte (Silgard® 184 Silicone Elastomer, Dow Corning, MI, EE. UU.) mediante el uso de cuatro pasadores de acero inoxidable de 0,4 mm de diámetro.

Las CE se visualizaron mediante un microscopio Axiolab de epifluorescencia vertical. Para controlar el $[Ca^{2+}]_i$, las CE se excitaron alternativamente a 340 y 380 nm y la luz emitida se detectó a 510 nm. Se acopló un filtro de densidad neutra (densidad óptica = 1,0) al filtro de 380 nm para aproximarse a la intensidad de la luz de 340 nm. Se usó un diafragma redondo para aumentar el contraste [21].

Se usó un software personalizado, que funciona en el entorno LINUX, para controlar la cámara. Se realizaron mediciones de razón y se trazaron en línea cada 3 s. El procedimiento en línea permitió el monitoreo de $[Ca^{2+}]_i$ al mismo tiempo que la lesión por micro pipeta. Las imágenes de fluorescencia obtenidas a 340 nm y 380 nm se almacenaron en el disco duro y el análisis de datos se realizó más tarde como se describe en la sección de análisis de datos.

Los experimentos se realizaron a temperatura ambiente (20–22 C) para limitar las disminuciones dependientes del tiempo en la intensidad de la señal de fluorescencia.

Como se muestra en la Referencia [22] el endotelio aórtico se lesionó bajo control microscópico por medio de una micro pipeta de vidrio con una punta de aproximadamente 30 μ m de diámetro, impulsada por un micro manipulador hidráulico XYZ. Los electrodos se fabricaron utilizando un extractor de micro pipetas de color marrón fuego P-1000 (Sutter Instruments, Novato, CA, EE. UU.). Se tomaron imágenes de EC cargados con Fura-2, junto con un ROI numerado, antes de la lesión, para identificar las células que se enfrentan al sitio de la lesión.

Los datos de imágenes endoteliales de Ca^{2+} se procesaron usando un procedimiento de procesamiento de datos automatizado. El procedimiento de procesamiento de datos se realiza de la siguiente manera, como primer paso, las imágenes de fluorescencia obtenidas a 340 y 380 nm se cargan por separado; luego, se inserta una etapa de calibración para determinar los parámetros utilizados por la segmentación de la imagen.

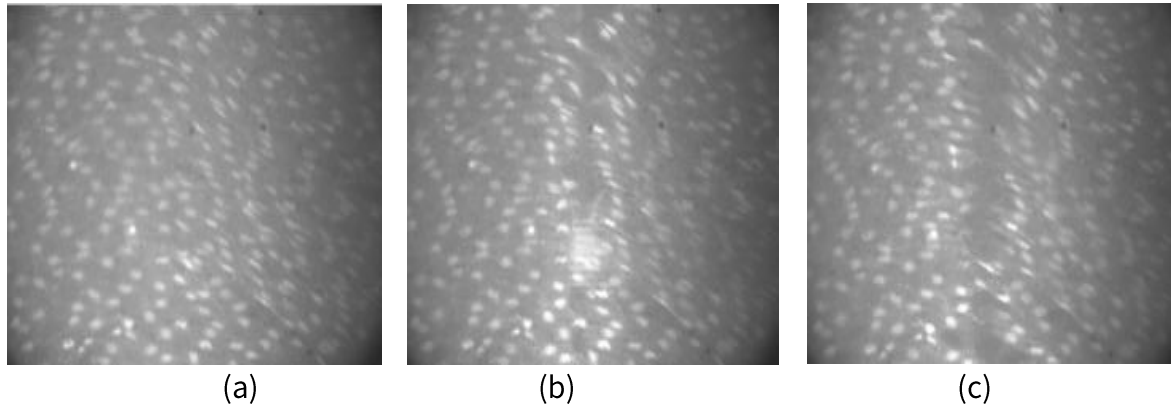


Figura 3. Imágenes obtenidas por [18]. (a) Imagen del experimento antes del estímulo mecánico, corte por un micro pipeta. (b) imagen en el momento del corte. (c) imagen después del corte.

En este primer capítulo se presentó la gran importancia del estudio de las CE para entender y detectar problemas en el sistema cardiovascular. La microscopia es una herramienta muy importante ya que a través de ella se obtienen imágenes microscópicas de CE y con ello poder realizar su estudio. Una vez obtenida las imágenes conocimos conceptos sobre una imagen digital y su histograma, es importante conocer estos conceptos ya que en el siguiente capítulo mostraremos métodos de segmentación y para lograrlo es importante realizar una ecualización del histograma, esta misma se presentara en el siguiente capítulo.

CAPÍTULO 2 SEGMENTACIÓN DE NÚCLEOS CELULARES

En este capítulo se presenta un resumen de tres métodos de segmentación de núcleos celulares, la descripción detallada se puede consultar en [23]; y los resultados con las imágenes de los núcleos celulares obtenidos. En los seres humanos, el sistema visual es experto en percibir e interpretar intuitivamente la información visual para comprender una escena. Tiene la capacidad de filtrar a la perfección información irrelevante como el ruido o las variaciones de iluminación dentro de una escena [24].

A medida que la biología moderna se desplaza hacia experimentos con grandes conjuntos de datos altamente cuantitativos, los enfoques manuales para la interpretación de imágenes se están volviendo anticuados. Con conjuntos de datos experimentales tan grandes, no es factible una inspección visual experta y anotación. Además, la inspección humana no es adecuada para detectar diferencias de población sutiles en decenas o cientos de miles de imágenes [25].

La detección y medición de estructuras celulares en imágenes de microscopía es esencial para la biología celular. La segmentación y la extracción de características también son temas importantes en la visión por computadora y el procesamiento de imágenes. Por lo tanto, el desarrollo de algoritmos para identificar y medir automáticamente las estructuras de interés dentro de las imágenes biológicas proporciona la base para muchas colaboraciones productivas entre los biólogos y los desarrolladores de algoritmos de procesamiento de imágenes computacionales. Todavía no se han establecido métodos de análisis de imágenes de propósito general debido a la diversidad de los métodos de microscopía y adquisición de imágenes que se utilizan en biología celular. Los enfoques de análisis de imágenes siguen adaptándose a la configuración experimental específica, sin embargo, a medida que la inteligencia artificial (AI) y los métodos de aprendizaje automático continúan desarrollándose, se podrían desarrollar métodos más genéricos [26].

2.1 HISTOGRAMA DE IMAGEN

Un histograma de intensidad representa la frecuencia de aparición de intensidades en una imagen. Es la aproximación de la función de distribución de probabilidad que siguen las intensidades.

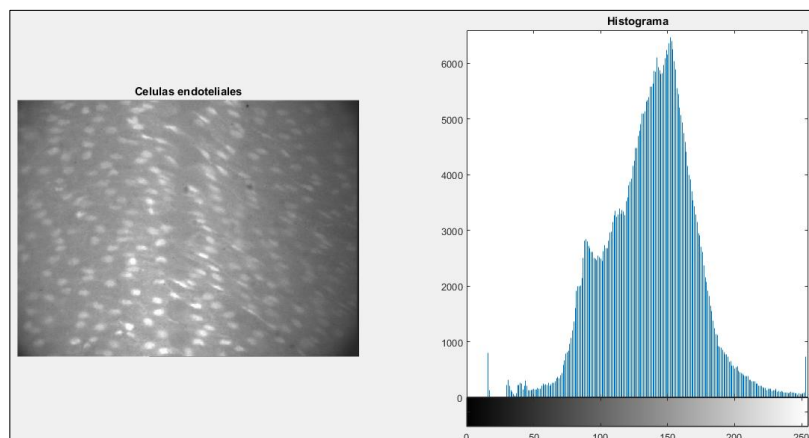


Figura 4. Se muestra del lado izquierdo la imagen de células endoteliales y a la derecha su histograma.

En este trabajo se utiliza MATLAB como herramienta de programación. La mayoría de los programas de edición de imágenes ofrecen la opción de mostrar un histograma de la imagen por ejemplo, la función *imhist* y *histc* son alternativas en la computación de histogramas de imágenes. Para el uso de estas funciones, el lector puede consultar las páginas de ayuda de MATLAB.

2.2 ECUALIZACIÓN DEL HISTOGRAMA

Para mejorar el contraste de una imagen se utiliza la ecualización del histograma.

La ecualización del histograma de una imagen es una transformación que pretende obtener, para una imagen, un histograma con una distribución uniforme. Es decir, que exista el mismo número de píxeles para cada nivel de gris del histograma de una imagen monocromática [27].

Dada una imagen $M \times N$ con nk píxeles para cada nivel r_k la ecualización del histograma consiste en realizar la siguiente transformación sobre los niveles de intensidad de la imagen:

$$S_K = T(r_k) = (L - 1) \sum_{j=0}^k p_{r(r_j)} = \frac{L-1}{MN} \sum_{j=0}^k n_j \quad (1)$$

Lo que resulta en una dispersión del histograma en un rango mayor dentro del intervalo $[0, L-1]$ [28].

Es conocido que si se usa sólo información global no se alcanza un buen realce de contraste, debido a que las técnicas globales causan un efecto de saturación de intensidades. Para solucionar este problema existe una ecualización del histograma localizada o también conocida como adaptativa, la cual considera una ventana local para cada píxel y calcula el valor de la nueva intensidad basado en el histograma local definido para cada ventana. Adicionalmente la ecualización adaptativa del histograma es un operador local uniforme en el sentido que todos los píxeles de una ventana local contribuyen igualmente a la determinación del nuevo valor del píxel central que está siendo considerado.

La ecualización de histograma adaptable de contraste limitado, a diferencia de una ecualización global, opera en pequeñas regiones de datos (mosaicos) en lugar de toda la imagen. El contraste de cada mosaico se mejora para que el histograma de cada región de salida coincida aproximadamente con el histograma especificado. La mejora de contraste puede ser limitada con el fin de evitar amplificar el ruido que podría estar presente en la imagen.

Los métodos de umbral se denominan globales si se calcula un único umbral para toda la imagen. Si la imagen se divide en subbloques y se calcula un umbral para cada subbloque, este método de umbralización se denomina umbralización local [29].

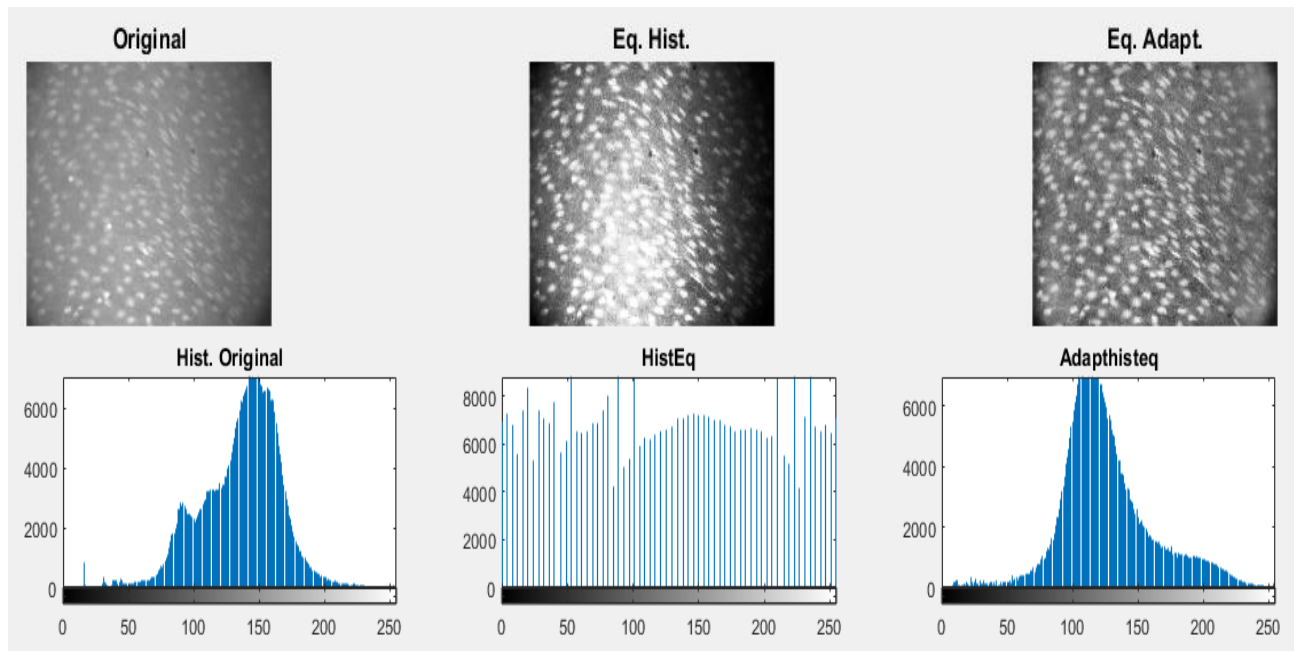


Figura 5. Imagen original, Ecuación de la imagen, ecuación adaptativa. En la segunda fila se muestra una binarización de las imágenes, según su media y su desviación estándar. Última fila su histograma.

En la figura 5 se observa, en la parte superior, la imagen original microscópica de células endoteliales y su transformación al aplicar la ecuación del histograma y la adaptativa. En la parte inferior de la figura se muestra el histograma de la imagen original y después de aplicar las diferentes ecuaciones.

2.3 SEGMENTACIÓN DE IMÁGENES BASADA EN INTENSIDAD

La segmentación de imágenes mediante el uso de umbrales basados en histogramas es probablemente el enfoque más común, ya que es fácil de implementar y requiere menos recursos de CPU para ejecutarse. Estos métodos generalmente emplean la maximización o minimización de una función de criterio basada en el histograma de la imagen. El umbral óptimo es la intensidad de nivel de gris a la que la función de criterio alcanza su extremo (máximo o mínimo). Los métodos de umbral se denominan globales si se calcula un único umbral para toda la imagen. Si la imagen se divide en subbloques y se calcula un umbral para cada subbloque, este método de umbralización se denomina umbralización local.

Uno puede interpolar los valores de umbral para cada píxel colocando el umbral local en el píxel central de cada bloque. Esto se conoce como umbral paramétrico. La principal desventaja de los métodos de umbral basados en histogramas es su desprecio del contexto espacial dentro del cual se produce el valor de intensidad.

En otras palabras, no consideran los valores de intensidad de los píxeles circundantes. Los métodos basados en histogramas bidimensionales consideran, hasta cierto punto, la información contextual. Los métodos de umbral también se utilizan con frecuencia para la segmentación inicial de imágenes antes de la aplicación de un método de segmentación más sofisticado con el fin de reducir el tiempo de convergencia [30]. Una revisión exhaustiva de los métodos de umbral se puede encontrar en Mardia y Hainsworth [31] y Pal y Pal [32].

En este capítulo discutiremos los métodos de umbral basados en histogramas que se usan de manera ubicua en la segmentación de imágenes. También discutiremos el modelado de mezclas y los algoritmos de agrupación K-means y C-means difusos aplicados a la segmentación de imágenes.

2.3.1 SEGMENTACIÓN DE IMAGEN POR AGRUPACIÓN DE K-MEANS

La agrupación de K-means (MacQueen, 1967) es un método comúnmente utilizado para particionar automáticamente un conjunto de datos en k grupos.

La agrupación, también conocida como clasificación no supervisada, busca las agrupaciones naturales en un conjunto de datos multidimensionales empleando una medida de similitud o disimilitud. En la agrupación, las clases son desconocidas y exploradas a partir de los datos mismos. Mientras que el agrupamiento divide los datos en grupos similares, la clasificación asigna una observación a uno de los grupos ya conocidos. La agrupación se puede dividir en dos categorías principales, a saber, agrupación particional y jerárquica.

Podemos ver la segmentación de imágenes como la división de las intensidades observadas en grupos similares. Luego, se puede emplear un método de agrupación para segmentar imágenes en regiones o agrupaciones [33].

El procedimiento sigue una forma simple y fácil de clasificar un conjunto de datos. La idea principal es definir K centroides, la mejor opción es colocarlos lo más lejos posible el uno del otro. El siguiente paso es tomar cada punto que pertenece a un conjunto de datos determinado y asociarlo al centroide más cercano. Cuando no hay ningún punto pendiente, se completa el primer paso y se realiza una agrupación temprana.

En este punto, necesitamos volver a calcular K nuevos centroides como baricentros de los grupos resultantes del paso anterior. Después de tener estos K nuevos centroides, se debe hacer un nuevo enlace entre los mismos puntos de ajuste de datos y el nuevo centroide más cercano. Los K centroides cambian su ubicación paso a paso hasta que no se realicen más cambios.

Finalmente, este algoritmo apunta a minimizar una función objetivo, en este caso una función de error al cuadrado. La función objetivo:

$$J = \sum_{j=1}^k \sum_{i=l}^k || x_i^{(j)} - c_j ||^2 \quad (2)$$

Donde $|| x_i^{(j)} - c_j ||^2$ es una medida de distancia elegida entre un punto de datos $x_i^{(j)}$ y el centro del grupo c_j , es un indicador de la distancia de los n puntos de datos desde sus respectivos centros de grupo.

Aunque se puede demostrar que el procedimiento siempre terminará, el algoritmo K-means no necesariamente encuentra la configuración óptima, correspondiente al mínimo de la función objetivo global [34].

Existen dos estrategias populares de agrupamiento particional: error cuadrado y modelado de mezclas. La suma de las distancias al cuadrado euclidianas entre las muestras en un grupo y el centro del grupo se llama variación dentro del grupo. La suma de las variaciones dentro del grupo en un esquema de agrupación se utiliza como criterio en la agrupación de K-medias. Esta agrupación también se conoce como partición de varianza mínima.

La agrupación de K-means es computacionalmente muy eficiente y proporciona resultados satisfactorios si los grupos son compactos y están bien separados en el espacio de características, pero el número de grupos debe especificarse a priori. El algoritmo K-means puede o no converger al mínimo global. Es sensible a los centros de clúster seleccionados inicialmente. La K determina el número de grupos que se desean formar, lo que significa que el algoritmo se puede ejecutar varias veces para aliviar el último problema.

La aplicación del algoritmo K-means a la segmentación de imágenes es razonablemente sencilla. Primero, escribiremos nuestra imagen 2-D o 3-D en un formato vectorial ordenado por columnas. Esto se puede hacer apilando cada columna después de la anterior. Llamemos a nuestro vector de muestra $M * N \times d$, cada elemento del cual comprende la intensidad de nivel de gris en el caso de una imagen de nivel de gris ($d = 1$) o los valores R, G y B en el caso de una imagen en color ($d = 3$). Tenga en cuenta que el orden de las columnas da como resultado la pérdida de información espacial. Por lo tanto, K significa agrupación sin tener en cuenta la información contextual o espacial [35].

El algoritmo K-means para la segmentación de imágenes incluye los siguientes pasos:

1. Inicialice el número de clases K y centroides μ .
2. Asigne cada píxel o vóxel al grupo cuyo centroide es el más cercano.
3. Después de que se hayan asignado todos los píxeles, vuelva a calcular los centroides.
4. Repita los pasos 2 y 3 hasta que los centroides ya no cambien.

Al calcular los centroides, se pueden emplear diferentes medidas de distancia, aunque la distancia euclidiana es probablemente la más común.

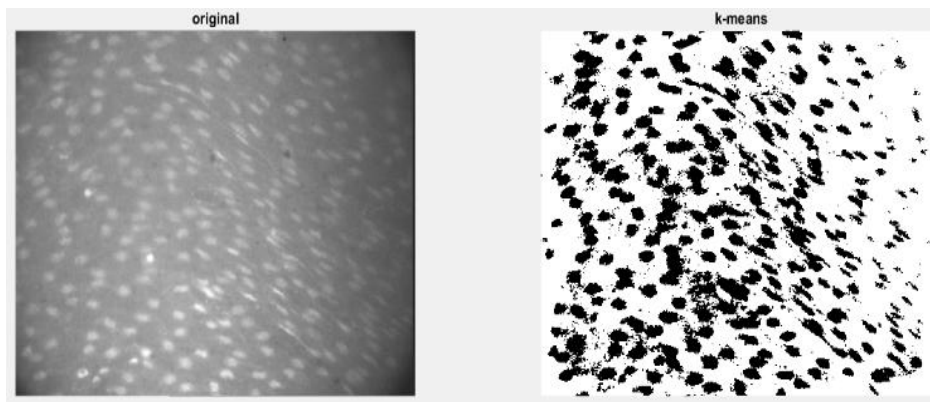


Figura 6. Segmentación utilizando k -means con 2 clusters, y calculando la distancia euclidiana para los centroides. El algoritmo para una imagen tardó 3.314054 segundos.

2.3.2 SEGMENTACIÓN POR FUZZI C-MEANS

El algoritmo de agrupamiento de C-medias difuso (FCM) fue propuesto por Dunn en 1973 y mejorado por Bezdek 1981. A diferencia de la agrupación K-means, en la que cada observación tiene una pertenencia binaria clara, el método FCM propone una pertenencia difusa que asigna un grado de pertenencia para cada clase. El concepto de grado de pertenencia en el agrupamiento difuso es similar a la probabilidad posterior en un entorno de modelado mixto.

Se basa en la minimización de la siguiente función objetivo:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C U_{ij}^m \|X_i - C_j\|^2 \quad , \quad 1 \leq m < \infty \quad (3)$$

Donde:

m es cualquier número real mayor que 1,

U_{ij} es el grado de pertenencia de X_i en el grupo j ,

X_i es el i ésimo de los datos medidos de dimensión d ,

C_j es el centro de dimensión d del grupo, y

$\|*\|$ es cualquier norma que expresa la similitud entre cualquier dato medido y el centro [36].

Al monitorear los puntos de datos que tienen valores de membresía cercanos a las clases existentes, es posible formar nuevos grupos; ésta es la principal ventaja de la agrupación FCM sobre la agrupación K-means [37].

La matriz de membresía U satisface la condición:

$$\sum_{i=1}^c U_{ij} = 1, \forall j = 1, \dots, n \quad (4)$$

Tenga en cuenta que los centroides son la media de los datos ponderados por su membresía y normalizados a la membresía total. Tenga en cuenta también que el grado de membresía es inversamente proporcional a la distancia al centro del grupo.

Seudocódigo FCM

1. Agrupar los datos en k grupos donde k está predefinido.
2. Seleccionar k puntos al azar como centros de agrupación.
3. Asignar objetos a su centro de clúster más cercano de acuerdo con la función de distancia euclidiana.
4. Calcular el centroide o la media de todos los objetos en cada grupo.
5. Repetir los pasos 2, 3 y 4 hasta que los mismos puntos se asignen a cada grupo en rondas consecutivas [38].

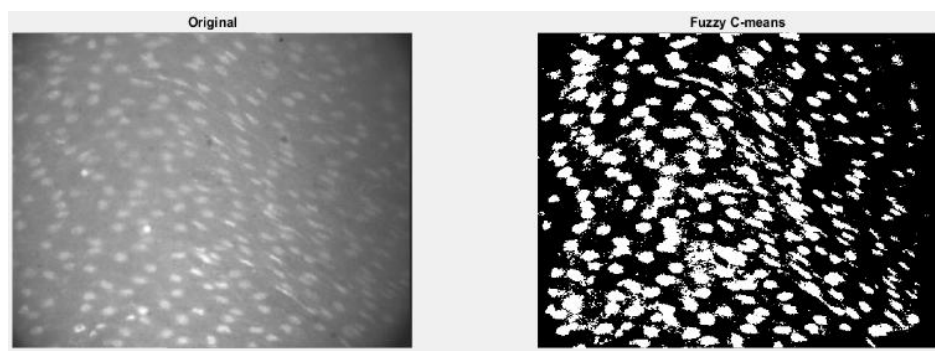


Figura 7. Imagen segmentada por el algoritmo Fuzzy C-means, especificado con 100 iteraciones y 2 clúster el algoritmo tardó 19.948726 segundos para una imagen.

Estos resultados se pueden usar como una imagen inicial para métodos de segmentación más sofisticados que también utilizan información espacial. Una buena segmentación inicial es esencial para que los métodos iterativos sofisticados logren una convergencia global más rápida.

2.3.3 SEGMENTACIÓN BASADA EN MODELADO DE MEZCLAS

En este enfoque, suponemos que los valores de píxeles que forman nuestra imagen provienen de una densidad multimodal. Existen protuberancias en el histograma de la imagen que indican que algunos píxeles alcanzan valores relativamente altos. Los modos en el histograma pueden corresponder a ciertas regiones o características (por ejemplo, un cierto tipo de tejido) en la imagen.

Por lo tanto, si podemos modelar la función de densidad de probabilidad (PDF) de los valores de píxel, entonces podemos usar ese modelo para encontrar las probabilidades de cualquier valor de píxel que pertenezca a diferentes clases o regiones.

El modelado de mezclas surgió como una técnica para la estimación de PDF y encontró aplicaciones significativas en varios problemas biológicos [39].

El primer paso en la aplicación de este enfoque a un problema de segmentación de imágenes sería obtener un histograma de imágenes. El histograma nos da una idea aproximada sobre el PDF de los valores de píxeles. Aunque la estimación de PDF es un problema complicado en sí mismo, el histograma sigue siendo una herramienta útil / práctico de gran valor.

Si la estimación del histograma del PDF no se obtiene con cuidado, algunos modos pueden desaparecer o aparecer falsamente. Afortunadamente, en el caso de los datos de imagen, en su mayoría sabemos cuáles pueden ser los valores de píxeles. Por ejemplo, cuando se trata de una imagen en escala de grises de 256 valores de intensidad, es obvio que deberíamos establecer los centros del contenedor de histograma en [0,255].

Ahora discutiremos cómo podemos modelar o ajustar las densidades normales a esos modos y proponer un modelo de mezcla normal general para el PDF de valores de píxeles.

Sea χ un valor de píxel y sea $f(x)$ sea la probabilidad de obtener ese valor de píxel. También podemos decir que " $f(x)$ es el PDF de x ". Si tenemos modos o clases K en el PDF de valores de píxeles, podemos argumentar que χ puede provenir de cualquiera de esas densidades K con diferentes probabilidades. Debido a que estos eventos son disjuntos, escribimos $f(x)$ como la suma de esas probabilidades de clase [40].

$$\Pr(x \text{ \& } x \text{ provienen de la Clases } i) = \Pr(\text{Clase } i) f_i(x | \text{Clase } i) \quad (5)$$

Donde $f_i(x | \text{Clase } i), i=1,2,\dots,k$ representa clases PDF 's condicionales de x .

En general, suponemos que el valor de píxel χ es de dimensión d y, en consecuencia, $f(x)$ es un PDF multivariado. Es por eso que hemos usado la notación vectorial para χ . Para

imágenes de intensidad, la dimensión de píxeles / observaciones es 1, mientras que para imágenes de color verdadero, los píxeles son 3-D (RGB) y d es 3.

La ecuación (5) simplemente establece que un *PDF* multimodal puede expresarse como una suma ponderada de *PDF*'s condicionales de clase k , siendo las ponderaciones las probabilidades anteriores. Con un pequeño cambio de notación, podemos reescribir la ecuación (5) como:

$$f(x) = \sum_{i=1}^k w_i f_i(x|\theta_i) \quad (6)$$

Se renombro la clase de probabilidad $\Pr(\text{Clase } i)$ como w_i que representa el "peso" o "coeficiente de mezcla" del i -ésimo componente *PDF*, y $f_i(x|\text{Clase } i)$ como $f_i(x|\theta_i)$, porque cada clase condicional *PDF* se identifica con un conjunto de parámetros Θ_i .

La estimación de máxima verosimilitud *ML* de θ se puede encontrar numéricamente usando varios algoritmos de optimización. Un método de optimización popular en particular que se usa ampliamente en el modelado de mezclas es la maximización de expectativas (*EM*).

El *EM* es un método iterativo para optimizar la función de probabilidad cuando falta información. En nuestro caso, la información que falta es la pertenencia a la clase de las observaciones, es decir, los valores de píxeles. El algoritmo *EM* fue formulado por Dempster, Laird y Rubin en su artículo seminal [41] y ahora es una herramienta estándar en estadística. El algoritmo *EM* para el modelado de mezclas se puede resumir como

Seudocódigo modelado de mezclas

1. Inicialización: Comience con una estimación inicial adecuada de los parámetros de la mezcla Θ . Suponiendo que conocemos el número de componentes k , esta inicialización se puede lograr utilizando uno de los algoritmos de agrupamiento no paramétricos disponibles.
2. Paso de expectativa: Calcule los pdf $(x|\theta_i)$, usando el valor estimado actual de θ y estimando la probabilidad posterior de una observación X_i .
3. Paso de maximización: dados los pesos estimados (ω_i) del paso de expectativa, actualizamos el vector de parámetros θ maximizando la función de probabilidad. **QUE ES S**
4. Iterar: (es decir, repetir los pasos de expectativa y maximización) hasta que las actualizaciones / cambios relativos en las estimaciones de los parámetros de la mezcla sean más pequeños. Seleccionamos la tolerancia del parámetro y el recuento máximo de iteraciones.

Si usamos demasiados parámetros del modelo, podemos producir un sobreajuste, lo que significa que podemos estar incorporando componentes que realmente no existen. Por lo tanto, debemos determinar un número óptimo de parámetros del modelo que expliquen adecuadamente nuestro PDF. Un método popular para evaluar o determinar la calidad del ajuste es el Criterio de Información Bayesiano (BIC) propuesto por Schwartz.

$$BIC = -2L(\theta) + M \log(n) \quad [7]$$

Donde $L(\theta)$ es el valor maximizado de la función log-verosimilitud, y M y n son los números de parámetros y píxeles del modelo, respectivamente.

Otra herramienta poderosa para evaluar la bondad de ajuste de un modelo de mezcla es el gráfico cuantil-cuantil (qq). Para verificar que el modelo de mezcla normal se ajusta bien a la distribución subyacente de los datos, generamos datos aleatorios de nuestro modelo de mezcla estimado con cuatro componentes y comparamos sus cuantiles con los cuantiles de los datos de píxeles [42].

Debemos tener en cuenta que EM no es la única opción para estimar los parámetros de un modelo de mezcla. La estimación de los parámetros del modelo de mezcla, basada en la maximización de la función de probabilidad logarítmica, es un problema típico de estimación y optimización de parámetros; como tal, también se podrían utilizar otros métodos de optimización.

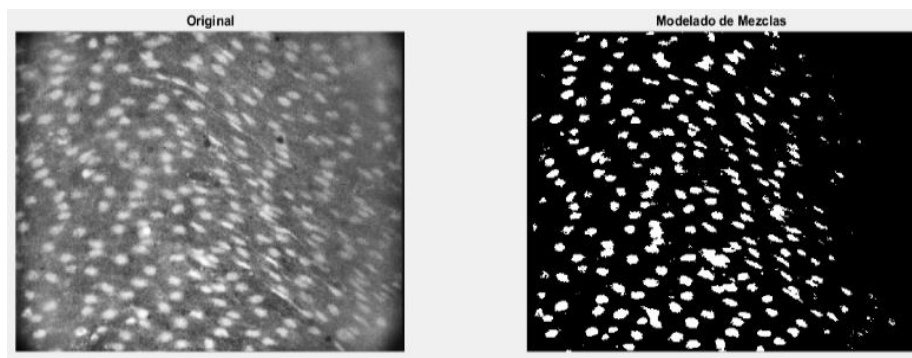


Figura 8. Segmentación de la imagen utilizando el algoritmo de modelado de mezclas, pasando como parámetro el número 2 para los clúster. Tardo en ejecutarse para una imagen 6.955721 segundos.

2.3.4 COMPARATIVA

El método de segmentación basado solo en la umbralización es el método más rápido, y el método más tardado fue el Fuzzy C-means con un parámetro de 100 iteraciones. El tiempo

de ejecución del Fuzzy C-means podría reducirse si se cambia el parámetro de iteraciones, pero a su vez se vuelve menos eficiente en la agrupación de los clúster.

El algoritmo de k-means tarda en promedio la mitad que el de modelado de mezclas, y se pudo observar que cuando se desea segmentar la imagen en 2 diferentes grupos el K-means logró segmentar más células que el modelado de mezclas. A diferencia de cuando se desea tener 3 grupos diferentes, ahí, el algoritmo de modelado de mezclas, segmenta mejor las células.

Podríamos concluir que para nuestro caso el mejor método de segmentación es el modelado de mezclas, aunque este lleve más tiempo de ejecución. Ya que el k-means abarca más células pero no logra una segmentación tan específica como la del modelado de mezclas.

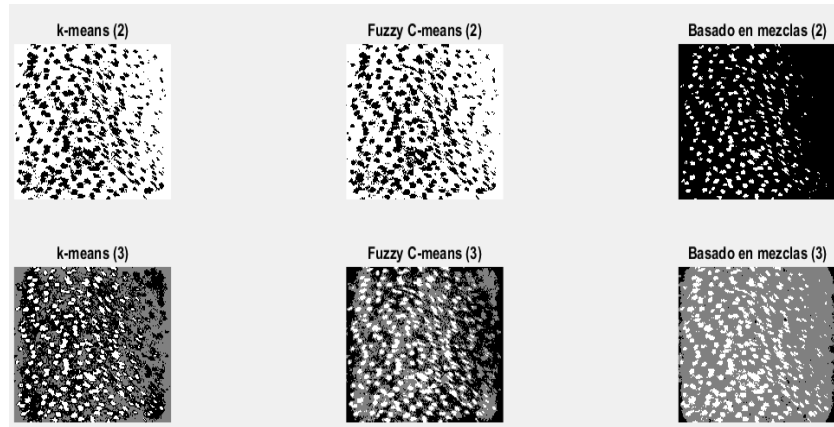


Figura 9. Comparación de los 3 algoritmos de segmentación con 2 y 3 clústeres.

Como se muestra en la figura 9 se realizaron 2 pruebas diferentes con cada uno de los algoritmos de segmentación, con 2 y 3 clusters. Cuando nosotros definimos un agrupamiento de 2, el método de k-mean es el que logra abarcar mayor células endoteliales aunque su segmentación no está bien definida. A diferencia el método basado en mezclas logra una mejor segmentación de las células endoteliales.

2.4 SEGMENTACIÓN DE IMAGEN POR MODELADO DE CAMPO ALEATORIO DE MARKOV

Sabemos que el valor de píxel en una imagen refleja ciertas características del objeto en la imagen. Las imágenes digitales que consisten en píxeles y se definen en una red 2D limitada tienen dos atributos importantes.

Primero, las intensidades de píxeles de un objeto o región casi homogénea seguirán una cierta distribución estadística. Llamaremos a esto la distribución de intensidad condicional de las intensidades de píxeles. Las intensidades de píxeles pueden ser escalares o una cantidad vectorial dependiendo del tipo de imagen.

En segundo lugar, los píxeles que se encuentran juntos o en un vecindario tenderán a tener valores de intensidad similares. La última restricción también se conoce como información contextual.

Los métodos de segmentación basados en la intensidad que discutimos en el tema 2.3 ignoran el último atributo de las imágenes. En esta sección, discutiremos cómo se puede modelar la información contextual mediante el modelado MRF. Este modelo que representa las características locales de la imagen subyacente, cuando se combina con la distribución condicional de las intensidades de píxeles en un marco bayesiano, nos permitirá estimar las intensidades verdaderas de la imagen con mucha más precisión que aquellas basadas solo en la información derivada de Intensidades de imagen.

Los métodos de segmentación basados en el modelado de MRF se han estudiado y utilizado ampliamente en aplicaciones de imágenes médicas y no médicas [43]. Proporcionan una manera conveniente de combinar la intensidad observada y la información contextual en un marco bayesiano. En las siguientes secciones, discutiremos el modelado MRF en detalle y su aplicación a la segmentación de imágenes a través de un marco bayesiano.

El modelado de campo aleatorio de Markov (MRF) de imágenes con el propósito de segmentación o restauración se inspiró en las técnicas computacionales desarrolladas en mecánica estadística. La mecánica estadística es una disciplina en física que estudia las propiedades de los átomos en la materia condensada. Para poder calcular tales propiedades de los átomos como se mencionó anteriormente, se necesitaba un modelo más simple para metales ferromagnéticos. En 1924, Ernst Ising propuso un modelo de interacción simple entre átomos en un metal ferromagnético.

Aunque el modelo de Ising puede parecer simple, fue a través de este modelo simple que se descubrió la existencia de la temperatura de Curie. Desde su publicación, el modelo Ising se ha convertido en uno de los modelos más utilizados para estudiar las propiedades de los

materiales ferromagnéticos por parte de físicos / científicos sucesivos. De hecho, esta es una tarea abrumadora incluso con las CPU rápidas de hoy.

Metropolis en 1953 introdujo un algoritmo para simular tales sistemas [44]. En su algoritmo, a uno de los átomos en la configuración actual se le da un pequeño desplazamiento, y se calcula el cambio en la energía del sistema ΔE .

Si $\Delta E < 0$, se acepta el desplazamiento, en otras palabras, la nueva configuración. Si $\Delta E > 0$, la nueva configuración se acepta con una probabilidad definida por la probabilidad de Boltzman $P(\Delta E) = e^{-\Delta E/KT}$. La decisión probabilística en el último caso se toma dibujando un número aleatorio uniforme del intervalo (0, 1). Luego, se acepta la nueva configuración, si $P(\Delta E)$ es mayor que este número; de lo contrario, se conserva la configuración actual. Esta es la esencia del conocido algoritmo Metropolis, que se puede utilizar para simular diferentes configuraciones de un espacio de estado A.

Este algoritmo es análogo a extraer muestras al azar de una población que representa, en este contexto, nuestro espacio de todas las configuraciones o el espacio de estado A. En lugar de sumar todas las configuraciones, que sabemos que no es factible de manera realista, se podrían extraer muestras del espacio de configuración en números suficientemente grandes y calcular el promedio del conjunto utilizando solo las muestras extraídas, que representan todo el espacio de estado.

El algoritmo Metropolis fue adaptado posteriormente por Kirkpatrick [45] para resolver la minimización de los problemas de tipo NP completo (polinomio no determinista). Esta nueva clase de algoritmos se conoce como métodos de recocido simulado (SA). Analizaremos los algoritmos basados en SA, así como los otros tipos no deterministas, cuando analicemos los algoritmos de recocido simulado que se han utilizado en el contexto de la segmentación de imágenes basada en MRF para encontrar la estimación máxima a priori (MAP).

2.4.1 SECUENCIA ALEATORIA DE MARKOV

Una secuencia aleatoria U se llama Markov de orden P si la probabilidad condicional de U dado todo su pasado es igual a la probabilidad condicional de $u(n)$ dado solo:

$$u(n-1), \dots, u(n-p) \quad (8)$$

En otras palabras, esto significa que si conocemos lo reciente, podemos estimar todo el pasado. Esto también se conoce como propiedad Markoviana de secuencias.

2.4.2 CAMPOS ALEATORIOS

Si cada elemento de una secuencia 2D es una variable aleatoria, la secuencia 2D se llama un campo aleatorio discreto [46]. A la luz de esta definición, podemos ver cada imagen como un campo aleatorio en el que cada píxel se considera como una variable aleatoria. Este concepto de campo aleatorio puede extenderse a 3D, y en el caso de una imagen en 3D, lo consideraremos un campo aleatorio en 3D. Una red rectangular en 2D (es decir, nuestra imagen) de píxeles $M_x \times M_y$. Tenga en cuenta que una red tridimensional (o red cúbica) tendrá voxels $M_x \times M_y \times M_z$ cuyas ubicaciones centrales estarán en la red.

2.4.3 SISTEMAS DE VECINDAD

Para describir o definir el modelo MRF dependiente localmente, primero debemos definir un sistema de vecindad sobre el cual se definirá el MRF. De nuevo, permítanme denotar una red $M_x \times M_y$ que indexa los píxeles en una región objetivo determinada. También sea S el punto de red o píxel, también llamado sitio, y N_s denota la vecindad de S . En este caso, el sistema de vecindad N_s debe cumplir las dos condiciones siguientes:

1. $s \notin N_s$
2. $r \in N_s$ iff $s \in N_r$ (9)

La primera condición establece que un sitio no debe ser vecino de sí mismo. La segunda condición, sin embargo, implica el requisito de simetría, que establece que si r es vecino de s , entonces s es vecino de r .

2.4.4 DISTRIBUCIÓN DE GIBBS

Una distribución de Gibbs con respecto a N_s es una medida de probabilidad dada por:

$$P(X = x) = \frac{1}{Z} \exp\{-\beta U(x)\} \quad (10)$$

Sabemos que no es factible calcular Z en la práctica. Por lo tanto, se dice que Z es computablemente intratable. La función de energía $U(x)$, a la que también nos referiremos como Energía de Gibbs [47], tiene la siguiente forma:

$$U(x) = \sum_{c \in C} V_c(X_c) \quad (11)$$

2.4.5 CAMPO ALEATORIO DE MARKOV

En las secciones anteriores, discutimos todos los ingredientes que necesitamos para definir el MRF. Ahora estamos listos para definir el modelo MRF.

La primera propiedad establece que el proceso de etiquetado asigna una probabilidad positiva a todas las imágenes en el espacio de configuración (sistema de vecindad). La segunda propiedad, sin embargo, es la misma propiedad de Markovian como se mencionó anteriormente en la discusión de las secuencias de Markov. Para expresar la última condición en palabras, la probabilidad de etiquetar un píxel dadas todas las etiquetas en la imagen completa es igual a la probabilidad de la etiqueta dada solo las etiquetas de sus vecinos [48].

En otras palabras, es una distribución condicional que depende de las etiquetas de los píxeles en el vecindario cercano. Es por esta razón que los modelos MRF también se denominan modelos MRF dependientes localmente.

Después de esto, nos gustaría combinar las dos fuentes de información, a saber, distribución de intensidad e información contextual, para llegar a una regla de decisión más poderosa con respecto a la etiqueta verdadera de un píxel o vóxel. Haremos esto por medio del teorema bayesiano, que establece que uno puede combinar la creencia previa junto con la probabilidad de obtener la creencia posterior sobre un evento, es decir:

$$Posterior = \frac{Probabilidad \times Prioridad}{Evidencia} \quad [12]$$

2.4.6 MODOS CONDICIONALES ITERADOS (ICM)

Besag propuso el algoritmo iterativo de modos condicionales (ICM), que es una solución aproximada a la estimación al máximo a posteriori (MAP). El ICM resuelve el problema de minimización actualizando secuencialmente las etiquetas.

El rendimiento del algoritmo ICM depende en gran medida del etiquetado inicial. Si es posible un buen etiquetado inicial, el algoritmo ICM puede converger rápidamente a una solución deseada. Si no es posible un etiquetado inicial razonablemente bueno, los algoritmos estocásticos, que se discutirán más adelante, pueden ser una mejor opción. La idea principal de ICM consiste en actualizar el estado de cada píxel (i, j) en cada iteración teniendo en cuenta dos factores. Primero, la evidencia radiométrica que es la probabilidad de la observación dado el estado oculto

$$l: p(X_{ij}|l, \mu l, \Sigma l). \quad [13]$$

Y segundo, la información contextual que son los estados ocultos de los píxeles vecinos. Juntando la evidencia radiométrica y la información contextual se asigna al píxel (i, j) el estado l que maximiza:

$$g(l) = p(X_{ij}|l, \mu_l, \Sigma_l) \exp \beta U_{ij}(l) \quad [14]$$

En la ecuación (14), $U_{ij}(l)$ es la cantidad de píxeles de la vecindad del píxel (i, j) con estado oculto l . Entonces $U_{ij}(l)$ incorpora información contextual junto con el parámetro de suavidad β . Para $\beta > 0$ vale: mientras más píxeles vecinos tengan el estado oculto l , más grande $U_{ij}(l)$ y por lo tanto más probable que el píxel (i, j) también tome el estado l .

Como el parámetro β es generalmente desconocido, hay que determinarlo de alguna manera. Básicamente hay dos opciones. El usuario fija algún valor para β o se estima el valor de β [49].

Seudocódigo ICM

1. Inicializar mapa de estados
2. Estimar parámetros μ_l y Σ_l para todos $l \in \Lambda$
3. Estimar o definir parámetro β
4. Fijar esquema de visita
5. Actualizar el estado oculto de cada píxel (i, j) según ecuación (3.1)
6. Opcional: re-estimar el parámetro β
7. Iterar el paso 5) y posiblemente el paso 6) hasta convergencia.

Podemos implementar el algoritmo ICM para segmentar imágenes con solo dos regiones. Tenga en cuenta que un método basado en histograma no nos daría un resultado tan bueno como el algoritmo ICM proporcionado. El ICM visita cada píxel (es decir, exploración de trama) en la imagen y lo vuelve a etiquetar.

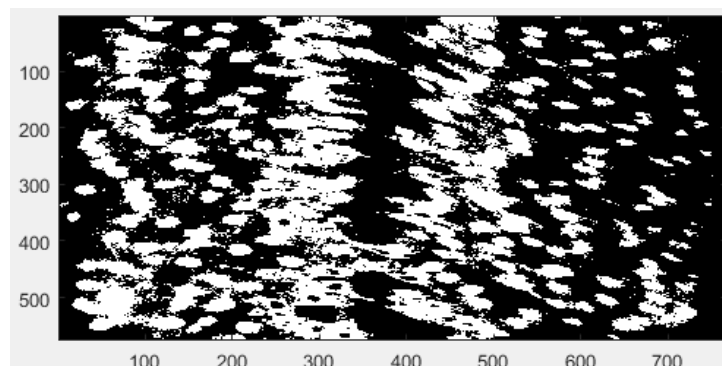


Figura 10. Células endoteliales segmentada por el algoritmo ICM en su versión vectorizada. Tenga en cuenta que en la vectorización, los píxeles no se etiquetan secuencialmente, sino en paralelo.

El impacto de la segmentación inicial se puede explorar mediante el uso de diferentes valores de umbral durante la segmentación inicial, tenga en cuenta que existe una versión vectorizada donde los píxeles no se etiquetan secuencialmente, sino en paralelo.

2.5 ALGORITMOS DE RECORRIDO SIMULADO

Anteriormente, formulamos el problema de segmentación como un problema de optimización. En otras palabras, estamos tratando de encontrar la estimación MAP, que es una maximización a posteriori. En esta sección, discutiremos dos algoritmos basados en recorrido simulado (SA) que resuelven la estimación MAP de una manera similar al proceso de recocido físico. En un proceso de recocido físico, la materia se calienta a una temperatura muy alta y luego se enfría gradual y muy lentamente para alcanzar el estado fundamental. Inspiradas en el recorrido físico, las soluciones basadas en SA introducen una variable de temperatura, similar a la temperatura física en concepto, en nuestras funciones energéticas.

Esta variable nos permitirá comenzar nuestro proceso de optimización desde un estado en el que todas las configuraciones tienen la misma probabilidad, en otras palabras, desde un estado muy caliente. Luego, al disminuir gradualmente la variable de temperatura, llegaremos a la solución global. La principal ventaja de estos enfoques es que el rendimiento de la optimización ya no depende del etiquetado inicial. Hay dos algoritmos bien conocidos que entran en esta categoría. Estos son el algoritmo Metrópolis y la muestra de Gibbs.

2.5.1 ALGORITMOS DE METRÓPOLIS

Lo que hace el algoritmo de Metrópolis es generar un vecino, calcularle su energía y aceptar ese vecino si tiene menor energía o aceptarlo con mayor energía pero con cierta probabilidad que depende de la temperatura (T). Se puede probar que si se realiza este proceso durante muchas transiciones se puede llegar a lo que se conoce como un equilibrio térmico. El equilibrio térmico está caracterizado por la distribución de Boltzmann (K_b).

Dado un estado i con energía E_i , genera un nuevo estado j mediante un mecanismo de perturbación (pequeña distorsión del estado i) y calcula la energía del nuevo estado E_j .

Si $(E_j - E_i) \leq 0$ entonces acepta el estado j como estado nuevo sino, acepta el estado con probabilidad:

$$\exp\left(\frac{E_i - E_j}{k_b T}\right) \quad (15)$$

La distribución da la probabilidad de que el sólido esté en el estado i con energía E_i a temperatura T y está dada por:

$$P_T\{X = i\} = \frac{1}{Z(T)} \exp\left(\frac{-E_i}{k_b T}\right) \quad (16)$$

Donde X es la variable estocástica denotando el estado actual del sólido, y $Z(T)$ es la función de partición definida como:

$$Z(T) = \sum_j \exp\left(\frac{-E_j}{k_b T}\right) \quad (17)$$

El algoritmo Metrópolis se implementó para imágenes con dos regiones [50]. Tenga en cuenta que el algoritmo anterior de hecho está tratando de minimizar. Similar al proceso de recocido físico, necesitamos comenzar con una temperatura alta y enfriarla disminuyendo la temperatura muy lentamente. En la figura 11 se observa como es la segmentación de células con este algoritmo. El programa de enfriamiento más comúnmente utilizado es $T_{k+1} = C * T_k$, donde C adquiere valores en el rango $[0.97, 1)$. Un valor de 0.97 parece dar resultados aceptables.

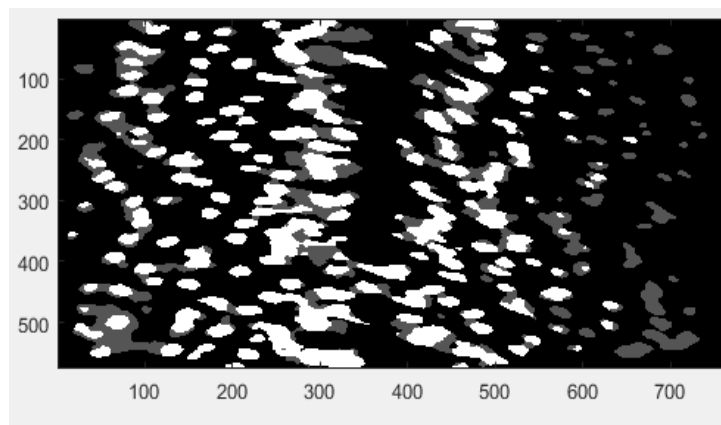


Figura 11. Segmentación de células endoteliales utilizando el algoritmo de metrópolis.

2.5.2 MUESTREO DE GIBBS

El muestreo de Gibbs fue sugerido por primera vez por Geman y Geman [51] para ser utilizado en la restauración o segmentación de imágenes. Similar al algoritmo Metrópolis, el muestreo de Gibbs proporciona una solución estocástica al problema de maximización, a saber, el cálculo de la estimación MAP. Vamos a ilustrar cómo funciona la muestra de Gibbs utilizando el siguiente pseudocódigo.

Pseudocódigo

1. Inicialice la temperatura $T \leftarrow T_0$ (debe ser lo suficientemente alta) y el número de iteración $k \leftarrow 0$.
2. Asigne etiquetas iniciales a los píxeles (se pueden elegir diferentes enfoques aquí). Se pueden elegir etiquetas arbitrarias para todos los píxeles.
3. Por cada píxel en la imagen se extrae una muestra de modo que se acepte x_s con la probabilidad:

$$\frac{\exp(-U(X_s))}{\sum_{r \in N_s} \exp(-U(r))} \quad (18)$$

4. Disminuya la temperatura $T \leftarrow T_{k+1}$ y aumente el número de iteración.

La muestra de Gibbs requiere que tomemos muestras de una distribución discreta conocida. Necesitamos saber primero cómo podemos tomar muestras de una distribución discreta. Esto se puede hacer mediante el uso de un método conocido como el método de transformación inversa [52]. Tengamos en cuenta que nuestro objetivo principal es resolver el problema de estimación MAP utilizando la muestra de Gibbs.

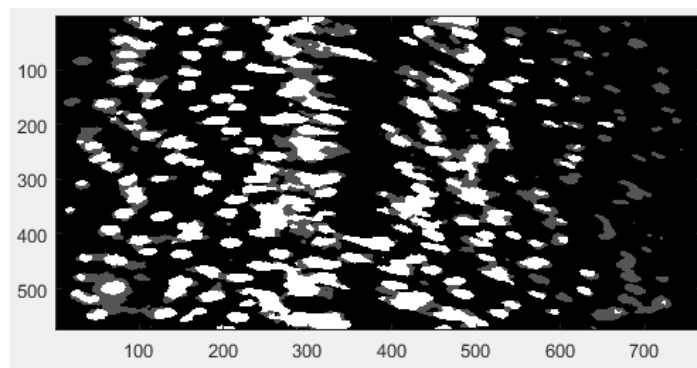


Figura 12. Segmentación de células endoteliales utilizando el muestreo de Gibbs.

2.5.3 COMPARATIVA

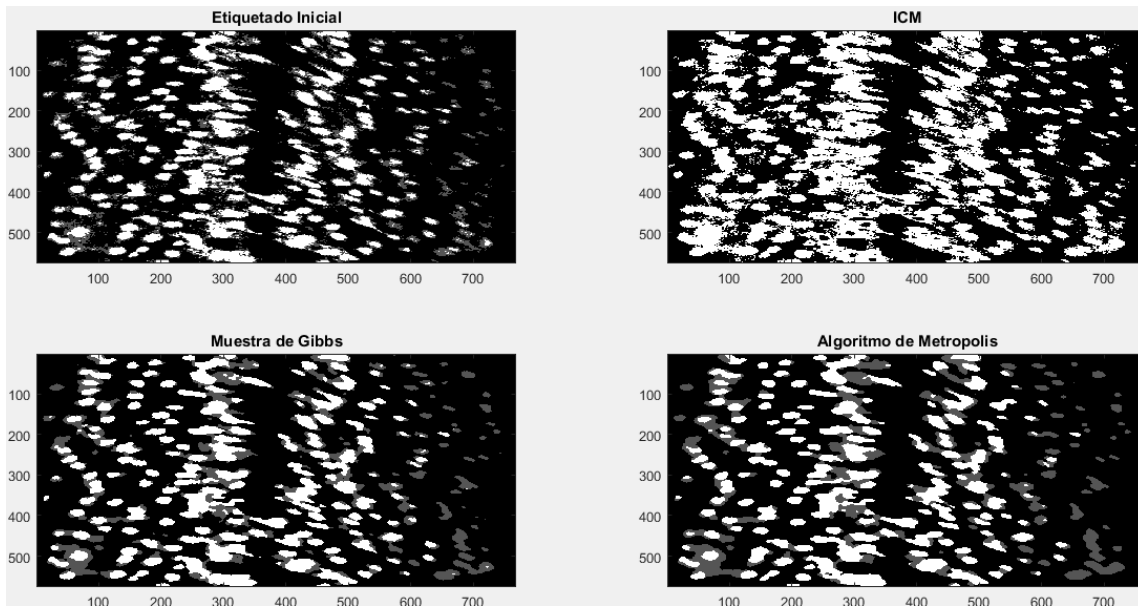


Figura 13. Comparación de los diferentes algoritmos de segmentación vistos en este capítulo.

En la figura 13 se muestran las células endoteliales segmentadas usando ICM y dos algoritmos de recorrido simulado: la muestra de Gibbs y el algoritmo de Metrópolis. Como podemos observar, los algoritmos de recorrido simulado obtienen similares mientras que el algoritmo ICM no logró una buena segmentación de las células.

CAPITULO 3 MÉTODO DE CARACTERIZACIÓN PROPUESTO

3.1 PRE PROCESAMIENTO

Cuando sobre una imagen aparecen sombras o manchas se hacen casi imposible realizar una binarización. Debido a esto es necesario calcular el umbral para diferentes partes de la misma imagen de forma localizada, teniendo en cuenta la vecindad de los píxeles.

3.2 LA UMBRALIZACIÓN

La umbralización es una técnica de segmentación simple y eficiente que permite separar los píxeles de una imagen en escalas de grises, y separarlos en dos categorías a partir de una ventana o borde. El umbral está en función de la posición $p(r, c)$, de la vecindad $N(r, c)$ y de la intensidad $I(r, c)$ actual del pixel.

$$T = T((r, c), N(r, c), I(r, c)) \quad (19)$$

Cuando la imagen que se encuentra bajo análisis no posee un fondo definido y su histograma no es bimodal, se debe encontrar un umbral T para cada pixel de la vecindad o través de una máscara que se debe recorrer sobre toda la imagen.

Dicha máscara puede ser de diferentes formas y tamaños lo que se define en el momento de implementar el algoritmo para la determinación del umbral local. [54]

El método de umbralización adaptativa, logra obtener resultados parecidos cuando segmenta antes y después del momento de la lesión, no se utilizara como método de segmentación el muestreo de Gibbs porque después del momento del corte tiene baja asertividad.

Optamos por utilizar como método de segmentación la umbralización adaptativa, este método obtiene resultados de segmentación más homogéneos durante el procesamiento de la secuencia de video completa.

| | Células Segmentadas | | Células NO Segmentadas | | Tiempo por imagen |
|------------|---------------------|---------|------------------------|---------|-------------------|
| | correctamente | | correctamente | | |
| | Antes | Después | Antes | Después | |
| K MEANS | 111 | 99 | 77 | 92 | 4.7 seg. |
| C MEANS | 114 | 96 | 82 | 90 | 18.7 seg. |
| M. MEZCLAS | 123 | 65 | 97 | 69 | 9.5 seg. |
| ICM | 75 | 130 | 94 | 77 | 150 seg. |
| METRÓPOLIS | 158 | 33 | 97 | 56 | 90 seg. |
| GIBBS | 161 | 84 | 72 | 58 | 74.5 seg. |
| ADAPTATIVA | 138 | 132 | 48 | 54 | 89.15 seg. |

Tabla 1. Comparativa de los 7 métodos de segmentación

Como se muestra en la tabla 1 el método de umbralización adaptativa, logra obtener resultados parecidos cuando segmenta antes y después del momento de la lesión, a diferencia de la segmentación por el muestreo de Gibbs porque después del momento del corte tiene baja asertividad. El método que segmenta en menor tiempo es el Kmeans.

Optamos por utilizar como método de segmentación la umbralización adaptativa, este método obtiene resultados de segmentación más homogéneos durante el procesamiento de la secuencia de video completa.

3.3 ALGORITMO DE PAGERANK

Una explicación extensa del algoritmo PageRank se puede consultar en [55]. Aquí se presenta un resumen. El objetivo del PageRank es conectar todas células y estudiar su comportamiento, no solo de una célula sino de un conjunto de ellas.

PageRank es un algoritmo de clasificación inventado por Larry Page y Serguéi Brin. Mide la autoridad de una página web de forma numérica en base a su importancia y calidad de los enlaces que apuntas hacia ella. Es entonces que la web se convierte en un sistema “democrático” donde las páginas web votan por el contenido de otras páginas web, cuantos más votos tenga una página web mayor popularidad tendrá, sin embargo la página que tenga mayor popularidad no significa que sea la más relevante, ya que cada voto tiene diferente “calidad” y esta calidad es hereditaria [56]. Por lo tanto la importancia se encuentra en la misma página y ésta se puede medir, utilizando el algoritmo de PageRank.

Para entender mejor el funcionamiento de este algoritmo empezaremos considerando que la web es un grafo dirigido muy grande, donde cada vértice es una página web y cada arista es un enlace que va dirigido a otra página web, por lo tanto la asignación de una puntuación a cualquier página web, se deriva de los enlaces realizados a esa página desde otras páginas web. Así, la web se convierte en una “democracia” donde las páginas votan por otras páginas mediante los enlaces a ellas [57].

PageRank confía en la naturaleza democrática de la web utilizando su vasta estructura de enlaces como un indicador del valor de una página en concreto. Google interpreta un enlace de una página A a una página B como un voto de la página A, para la página B. Pero Google mira más allá del volumen de votos, o enlaces que una página recibe; también analiza la página que emite el voto. Los votos emitidos por las páginas consideradas "importantes", es decir con un PageRank elevado, valen más, y ayudan a hacer a otras páginas "importantes". Por lo tanto, el PageRank de una página refleja la importancia de la misma en Internet.

El algoritmo inicial del PageRank lo podemos encontrar en el documento original donde sus creadores presentaron el prototipo de Google: “The Anatomy of a Large-Scale Hypertextual Web Search Engine” [58].

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(i)}{C(i)} \quad [20]$$

Donde:

$PR(A)$ Es el PageRank de la página A.

d Es un factor de amortiguación que tiene un valor entre 0 y 1.

$PR(i)$ Son los valores de PageRank que tienen cada una de las páginas i que enlazan a A .

$C(i)$ Es el número total de enlaces salientes de la página i (sean o no hacia A).

La calidad de la página y el número de posiciones que ascienda se determina por una "votación" entre todas las demás páginas de la World Wide Web acerca del nivel de importancia que tiene esa página. Un hiperenlace a una página cuenta como un voto de apoyo. El PageRank de una página se define recursivamente y depende del número y PageRank de todas las páginas que la enlazan. Una página que está enlazada por muchas páginas con un PageRank alto consigue también un PageRank alto. Si no hay enlaces a una página web, no hay apoyo a esa página específica. El PageRank de la barra de Google va de 0 a 10. Diez es el máximo PageRank posible y son muy pocos los sitios que gozan de esta calificación, 1 es la calificación mínima que recibe un sitio normal, y cero significa que el sitio ha sido penalizado o aún no ha recibido una calificación de PageRank. Parece ser una escala logarítmica. Los detalles exactos de esta escala son desconocidos. En los últimos tiempos Google está tratando de mantener un poco "privado" su PageRank para evitar manipulaciones, pero existen sitios donde se puede comprobar el PageRank [59].

3.4 APLICACIÓN DEL ALGORITMO A LA CARACTERIZACIÓN DE CÉLULAS

Después de procesar la secuencia de video se obtiene la intensidad y el área promedio de células representativas ubicadas en diferentes regiones en la imagen. En la siguiente tabla se muestran las 60 células seleccionadas de las cuales se conoce su área y su intensidad promedio para cada imagen de la secuencia.

| | Centro de célula | | Intensidad Promedio | | Área Promedio | |
|----|------------------|-----|---------------------|------------|---------------|------------|
| | X | y | Antes | Después | Antes | Después |
| 1 | 30 | 80 | 126.610833 | 125.947692 | 126.610833 | 125.850833 |
| 2 | 19 | 156 | 136.567508 | 134.591142 | 235.5 | 219.75 |
| 3 | 33 | 304 | 147.027925 | 144.499475 | 191.916667 | 178.166667 |
| 4 | 19 | 354 | 144.494033 | 141.650517 | 213.166667 | 190.75 |
| 5 | 36 | 440 | 140.224192 | 138.529183 | 220 | 209.25 |
| 6 | 50 | 546 | 130.420508 | 126.903475 | 383.833333 | 399.083333 |
| 7 | 70 | 496 | 156.091617 | 153.417142 | 201.25 | 172.25 |
| 8 | 66 | 389 | 161.114733 | 158.784375 | 204.5 | 205.666667 |
| 9 | 97 | 329 | 168.237508 | 165.751117 | 158.583333 | 153.75 |
| 10 | 67 | 274 | 158.270308 | 155.722342 | 189.083333 | 179 |
| 11 | 44 | 237 | 148.740875 | 145.983633 | 325.083333 | 302.416667 |
| 12 | 83 | 114 | 160.027808 | 158.073417 | 205 | 209.916667 |
| 13 | 79 | 42 | 140.358692 | 140.615775 | 148.75 | 137.333333 |

| | | | | | | |
|----|-----|-----|------------|------------|------------|------------|
| 14 | 154 | 23 | 157.445483 | 159.905167 | 257.583333 | 249.25 |
| 15 | 159 | 100 | 168.047733 | 166.853092 | 225.5 | 234.583333 |
| 16 | 171 | 165 | 161.797942 | 168.093542 | 79.75 | 95.1666667 |
| 17 | 147 | 228 | 172.698658 | 170.839817 | 217.25 | 202.166667 |
| 18 | 175 | 318 | 178.333883 | 178.58515 | 190.166667 | 180.333333 |
| 19 | 162 | 388 | 183.336017 | 184.265725 | 468.166667 | 488.666667 |
| 20 | 150 | 484 | 189.875275 | 185.795958 | 205.583333 | 224.25 |
| 21 | 130 | 542 | 186.907175 | 179.724842 | 295.916667 | 313.75 |
| 22 | 187 | 511 | 201.987183 | 197.068192 | 226.5 | 240.083333 |
| 23 | 211 | 413 | 193.644225 | 202.109158 | 276 | 287.416667 |
| 24 | 234 | 314 | 188.04035 | 207.41 | 121.916667 | 160.583333 |
| 25 | 219 | 243 | 170.768025 | 174.123533 | 86.8333333 | 79.5833333 |
| 26 | 223 | 128 | 164.528108 | 169.9604 | 168 | 167.333333 |
| 27 | 211 | 67 | 165.580392 | 171.211775 | 217.416667 | 203.916667 |
| 28 | 268 | 21 | 154.162133 | 175.023325 | 887 | 1067.83333 |
| 29 | 291 | 85 | 176.115217 | 206.8657 | 120.083333 | 138.75 |
| 30 | 347 | 134 | 179.0243 | 159.540458 | 246 | 243.833333 |
| 31 | 306 | 205 | 175.33075 | 204.022025 | 87.5833333 | 107.666667 |
| 32 | 337 | 287 | 185.248775 | 196.226467 | 182.833333 | 575.25 |
| 33 | 321 | 358 | 196.407383 | 196.239817 | 208.666667 | 210.083333 |
| 34 | 319 | 410 | 199.061275 | 211.363858 | 132.666667 | 177.166667 |
| 35 | 319 | 479 | 210.586683 | 231.29855 | 191.583333 | 333.166667 |
| 36 | 298 | 550 | 198.078317 | 211.779775 | 348.083333 | 582.75 |
| 37 | 400 | 489 | 197.318883 | 199.212308 | 194.583333 | 239.666667 |
| 38 | 452 | 422 | 197.457842 | 206.846983 | 164.916667 | 182.25 |
| 39 | 440 | 377 | 190.166783 | 211.156467 | 132.666667 | 184.166667 |
| 40 | 434 | 315 | 179.340283 | 200.3841 | 200.3841 | 216.833333 |
| 41 | 450 | 246 | 168.541058 | 183.979292 | 156.083333 | 110.833333 |
| 42 | 463 | 160 | 172.131392 | 191.324667 | 166.75 | 248 |
| 43 | 415 | 41 | 173.822142 | 167.609592 | 385.166667 | 321 |
| 44 | 350 | 24 | 139.075783 | 149.532833 | 177 | 394.083333 |
| 45 | 545 | 47 | 137.509375 | 141.28 | 148.666667 | 142.25 |
| 46 | 531 | 126 | 150.168508 | 157.168617 | 215.416667 | 251.166667 |
| 47 | 528 | 217 | 154.663192 | 155.605708 | 299.833333 | 290.333333 |
| 48 | 497 | 329 | 177.684383 | 181.606625 | 509.75 | 486 |
| 49 | 512 | 399 | 172.066783 | 173.738475 | 166.666667 | 185.166667 |
| 50 | 485 | 508 | 178.4873 | 183.818842 | 255.75 | 237.166667 |
| 51 | 574 | 509 | 154.688058 | 147.027692 | 111 | 89 |
| 52 | 601 | 409 | 141.332483 | 136.809275 | 145.75 | 123.75 |
| 53 | 576 | 321 | 154.587725 | 151.064175 | 97.5833333 | 92.8333333 |
| 54 | 616 | 239 | 127.952583 | 126.855525 | 74.25 | 81.5 |

| | | | | | | |
|----|-----|-----|------------|------------|------------|------------|
| 55 | 628 | 150 | 124.741958 | 123.794917 | 85.6666667 | 80.1666667 |
| 56 | 627 | 69 | 119.84385 | 119.450867 | 78.5 | 92.5833333 |
| 57 | 722 | 101 | 93.2682833 | 92.8705583 | 116.083333 | 131.166667 |
| 58 | 694 | 259 | 112.453225 | 110.6881 | 97.75 | 99.9166667 |
| 59 | 692 | 319 | 115.187475 | 113.879742 | 276.25 | 291.25 |
| 60 | 749 | 333 | 93.9814583 | 92.5135833 | 391.333333 | 388.25 |

Tabla 2. Intensidad y área promedio de las células antes y después del corte endotelial.

La secuencia de video se subdividió en 2 subsecuencias una antes del corte y otra después del corte; de la células se utilizaron las intensidades y las áreas promedio para calcular la importancia de cada célula, para poder caracterizar la importancia de cada célula en el último cuadro de la secuencia.

Para adaptar el problema de caracterización de células al método de Page Rank cada célula representa un nodo; si dos células están conectadas con una arista la flecha apunta a la célula con mayor intensidad o área. Para conectar las células se utilizó la triangulación de Delaunay.

3.5 TRIANGULACIÓN DE DELAUNAY

Para un conjunto dado P de puntos discretos en un plano unidos por aristas formando triangulaciones. Las triangulaciones de Delaunay maximizan el ángulo mínimo de todos los ángulos de los triángulos en la triangulación. La triangulación lleva el nombre de Boris Delaunay por su trabajo sobre este tema desde 1934.

Para un conjunto de puntos en la misma línea no hay triangulación de Delaunay. Para cuatro o más puntos en el mismo círculo la triangulación de Delaunay no es única: cada una de las dos posibles triangulaciones que dividen el cuadrángulo en dos triángulos satisface la "condición de Delaunay", esta es, que los círculos de todos los triángulos tienen interiores vacíos.

Al considerar las esferas circunscritas, la noción de triangulación de Delaunay se extiende a tres y mayores dimensiones. Las generalizaciones son posibles para métricas distintas de la distancia euclidiana. Sin embargo, en estos casos no se garantiza que una triangulación de Delaunay exista o sea única [60].

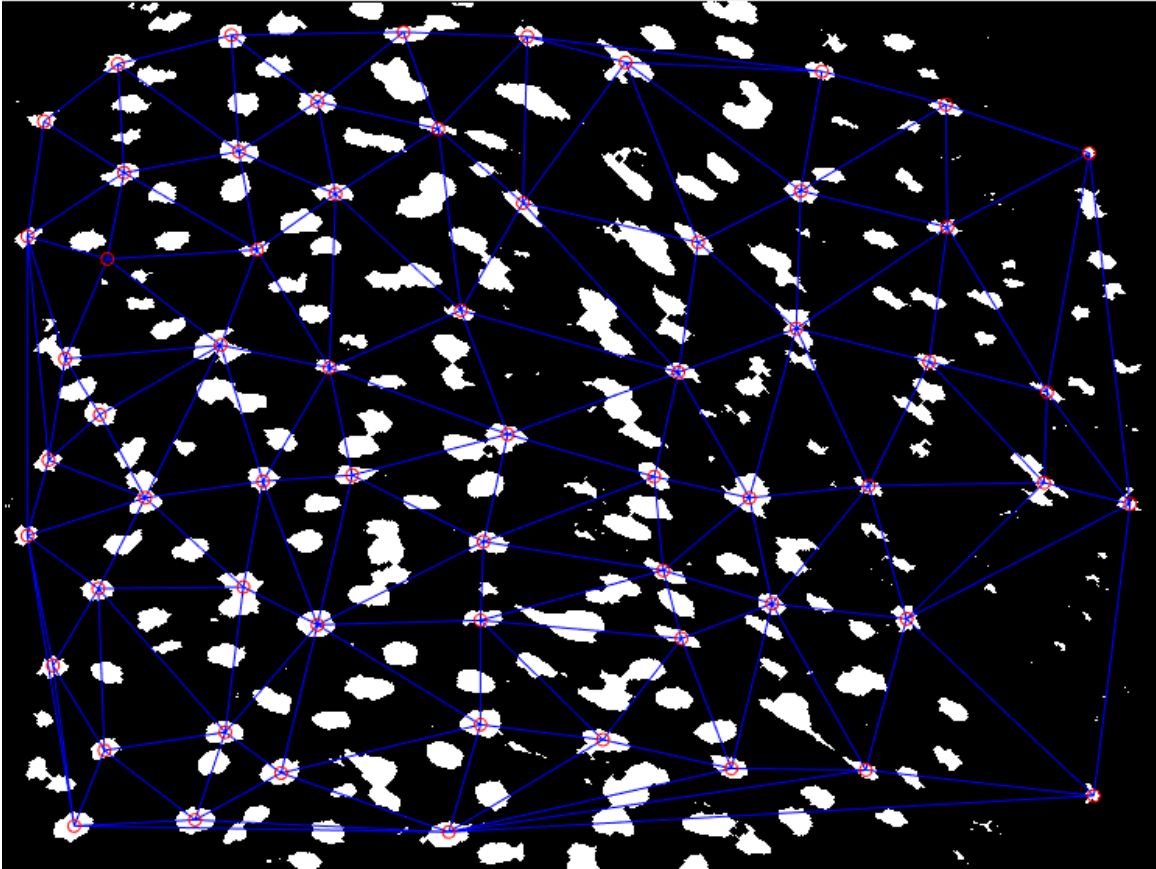


Figura 14. 60 células conectadas entre ellas por la triangulación de Delaunay.

En la figura 14 se observan 60 células que fueron seleccionadas previamente, donde se le pasa al algoritmo de la triangulación Delaunay el centro de cada célula así como su área e intensidad.

3.6 EJEMPLO DELAUNAY Y PAGERANK CON 6 CÉLULAS

Se mostrará un pequeño ejemplo para poder entender un poco mejor lo que se hace con el total de células, que en nuestro problema se escogieron 60 células que se encuentra esparcidas por toda la imagen. En este ejemplo se seleccionaron 6 células.

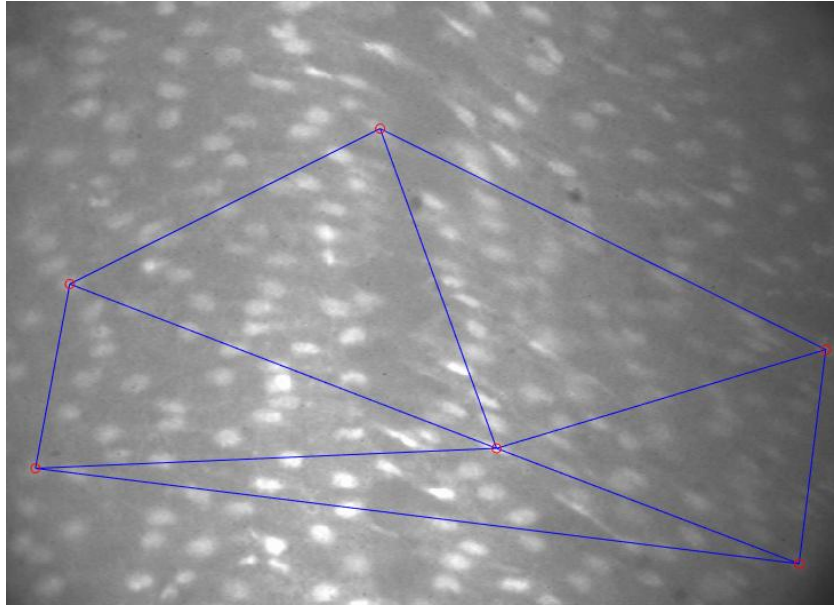


Figura 15. Células representan nodos color rojo y las aristas la conexión entre ellas.

En la figura 15 se muestra donde se encuentran 4 células alejadas del área del corte (dos de cada lado de la imagen). El primer paso es procesar la imagen original para obtener las coordenadas de las células seleccionadas.

El éxito del buscador Google se encuentra en el punto de calcular la relevancia de cada página y lo hace mediante el algoritmo PageRank. Adaptamos este algoritmo a nuestro problema, donde en vez de trabajar con páginas lo haremos con células y poder hacer los cálculos para el PageRank. La importancia de cada célula será medida a través de su intensidad y área promedio como se muestra en la Tabla 2.

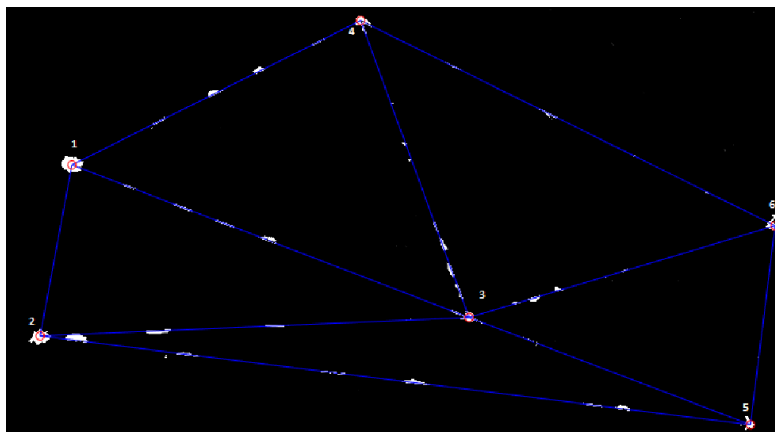


Figura 16. Figura de las células procesadas, obteniendo la intensidad y área de las células seleccionadas.

El algoritmo de PageRank adaptado a nuestro problema regresa la importancia de todas las células. El puntaje de un nodo debe aumentar más si tiene un enlace proveniente de un nodo con mayor intensidad que el de un nodo con menor intensidad.

Entonces declaramos a X_i como la importancia del nodo (i) tal que $1 < i < n$, siendo X_i un número no negativo. La importancia de un nodo i se obtiene mediante la resta de las intensidades de todos los nodos con enlaces hacia él.

La tabla 3 pertenece a las 6 células de la figura 16 y se muestra la importancia de cada una de ellas, donde 'InDegree' se refiere a las células con enlaces hacia ella con menor intensidad. 'OutDegree' son las células a las que se conecta con mayor intensidad.

| Célula | Importancia | InDegree | OutDegree |
|--------|-------------|----------|-----------|
| 1 | 0.094476 | 0 | 3 |
| 2 | 0.12123 | 1 | 2 |
| 3 | 0.25754 | 3 | 0 |
| 4 | 0.19949 | 3 | 2 |
| 5 | 0.094476 | 0 | 3 |
| 6 | 0.23279 | 3 | 0 |

Tabla 3. Resultados del PageRank para el ejemplo de 6 células.

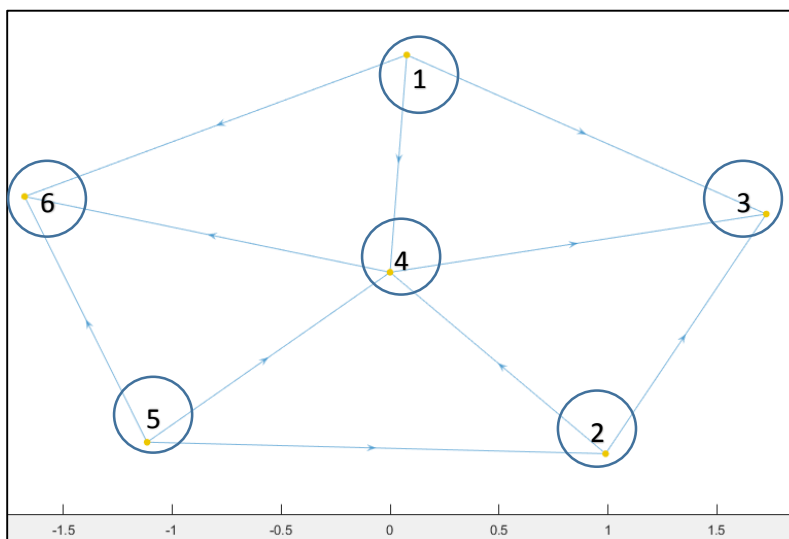


Figura 17. Grafica de los resultados de PageRank.

En la figura 17 se puede observar las células del ejemplo de 6 células ordenadas por el algoritmo de PageRank donde las flechas apuntan hacia las células con mayor

importancia, se observa en la tabla 3 que la célula con el índice 3 es la de mayor importancia.

En la figura 16 la célula con ese índice se encuentra en el centro dentro del área de corte donde el Ca^{+2} se empieza a esparcir y las primeras células en absorberlo son las más cercas al área del corte

CAPITULO 4 RESULTADOS EXPERIMENTALES

4.1 RESULTADOS DE LOS ALGORITMOS DE SEGMENTACIÓN

En este capítulo analizaremos los métodos de segmentación comparándolos antes, durante y después del corte. Para esto se contó a mano el número de células que se encuentran en una imagen después de realizar una ecualización adaptativa. Se verificará que método logro segmentar mejor las células y cuantas lograron segmentar.

4.2 ANÁLISIS DE LOS DATOS

Las imágenes de células endoteliales se procesaron usando un procedimiento de procesamiento de imágenes automatizado. En la figura 18 se presenta el diagrama de bloques del proceso de caracterización de las imágenes. Como primer paso las imágenes obtenidas se ecualizarán para luego ser segmentadas con un algoritmo de agrupación.

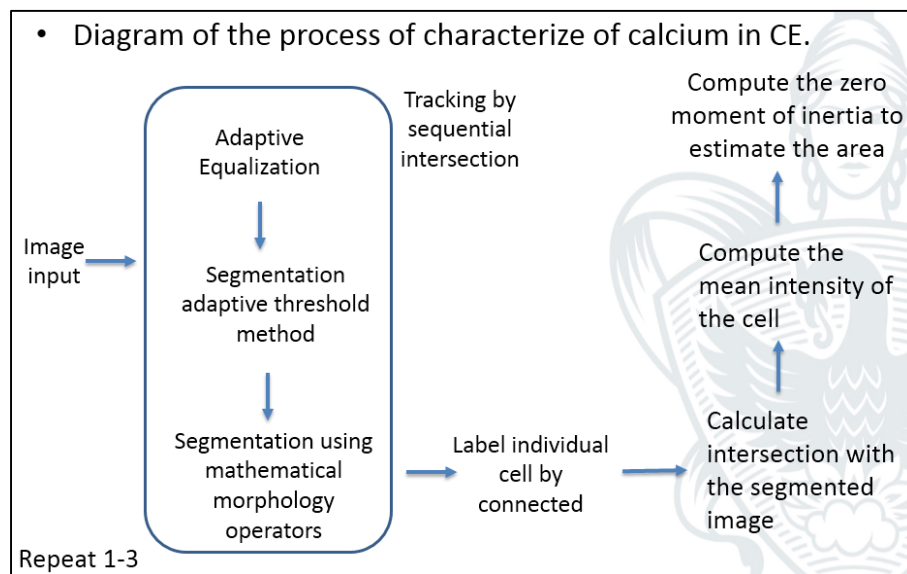


Figura 18. Diagrama del proceso de caracterización de calcio en CE.

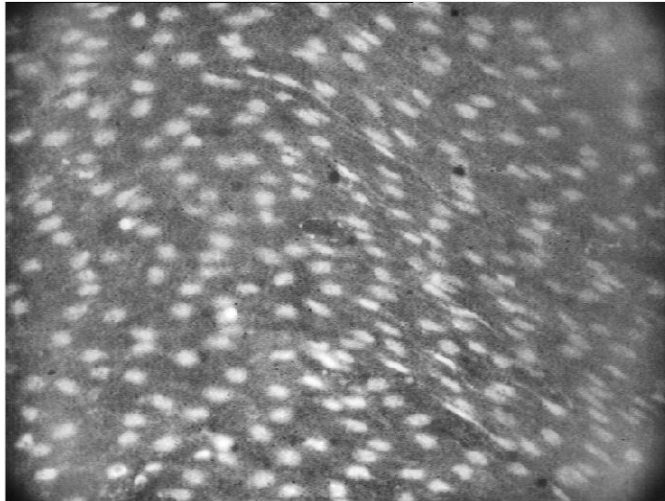


Figura 19. Imagen antes del corte con un filtro de ecualización adaptativa.

En la figura 19 se observa la imagen completa de las células, aquí se realizó un conteo a mano de las células endoteliales visibles antes del corte, para después comparar los resultados con los diferentes métodos de segmentación. En total se contaron 240 células.

4.3 K-MEANS

Se muestra la segmentación de la imagen de CE en tres diferentes momentos, antes del corte, durante el corte y después del corte. El algoritmo de segmentación de k-means tarda en promedio 4.749 segundos por imagen. Se puede observar a simple vista que no logra segmentar las células bien durante y después del corte.



Figura 20. Segmentación de células con el algoritmo de K-means.

Antes del corte conto 111 células, durante y después del corte conto 77 células, muchas de ellas mal segmentadas. Hay células que el algoritmo cuenta como una cuando realmente son 3 células juntas.

4.4 FUZZY C MEANS



Figura 21. Segmentación de células por el algoritmo Fuzzy C-means

Este algoritmo es más tardado que el k-means con un promedio de segmentación por imagen de 18.796 segundos. Este algoritmo segmentó 114 células antes del corte, y después del corte solo segmentó 82 células sin buenos resultados en la segmentación.

4.5 MODELADO DE MEZCLAS



Figura 22. Segmentación de células por el algoritmo de Modelado de Mezclas

Este algoritmo tarda en segmentar una imagen en promedio 9.568 segundos obteniendo buenos resultados en la segmentación de las células. Antes del corte segmentó 123 células, y durante y después del corte logó segmentar 97 células.

Como se puede observar los algoritmos son más precisos antes del corte y esto se debe a que la intensidad de las células aumenta y los algoritmos agrupan algunas de ellas. Las células que se encuentran más alejadas del corte son aquellas que logran una mejor segmentación.

Cabe recordar que estos algoritmos utilizan sólo las intensidades de los píxeles para su agrupamiento, los siguientes algoritmos que vamos a ver toman en cuenta dos aspectos; las intensidades de los píxeles y también los píxeles que se encuentran juntos o en un vecindario con un valor de intensidad similar.

4.6 ICM

El algoritmo ICM tiene dos parámetros de entrada que denotan las imágenes de nivel de gris e inicialmente etiquetadas. Para esto podemos utilizar la imagen segmentada del algoritmo de modelado de mezclas y la imagen con una ecualización adaptativa, lo cual ayudará a que el algoritmo pueda converger y encuentre una solución óptima en menor tiempo.

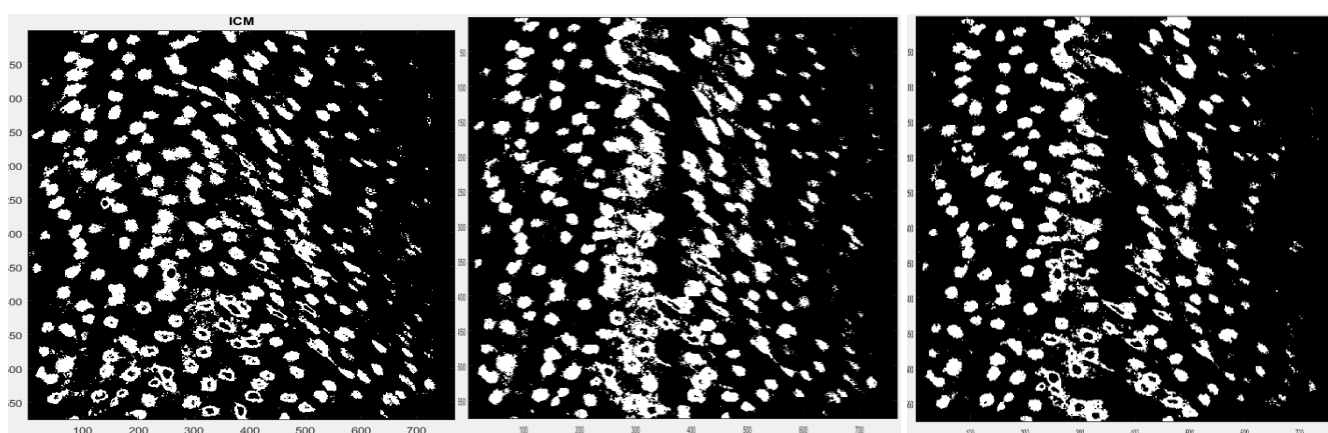


Figura 23. Segmentación de células por el algoritmo de modos condicionales iterativos ICM.

Como se puede observar en la imagen 18 hay células que no están bien segmentadas ya que a simple vista se puede ver que en el centro de algunas células existen huecos, estas aparecen más cerca al área de la lesión. A diferencia de los algoritmos anteriores donde la segmentación cambiaba mucho antes y después de la lesión, el ICM trata de mejorar esto y aunque cambia la segmentación antes y durante la lesión lo hace en menor cantidad que los algoritmos anteriores. Conto 75 células antes del corte y 94 durante y después de la lesión. En el algoritmo ICM se especifican variables como el número de clases, número de iteraciones, en nuestro caso fueron de 4 clases y 30 el número de iteraciones, con esto el algoritmo tarda en promedio 2.5 minutos por imagen.

4.7 RESULTADOS ALGORITMO METRÓPOLIS

Lo que hace el algoritmo de Metrópolis es generar un vecino, calcularle su energía y aceptar ese vecino si tiene menor energía o aceptarlo con mayor energía, pero con cierta probabilidad que depende de la temperatura. Para esto nosotros tenemos que darle tres parámetros al algoritmo, el primero será una variable beta que nosotros tendremos que

especificar, la siguiente será el número de regiones o clases y por último el número de iteraciones.

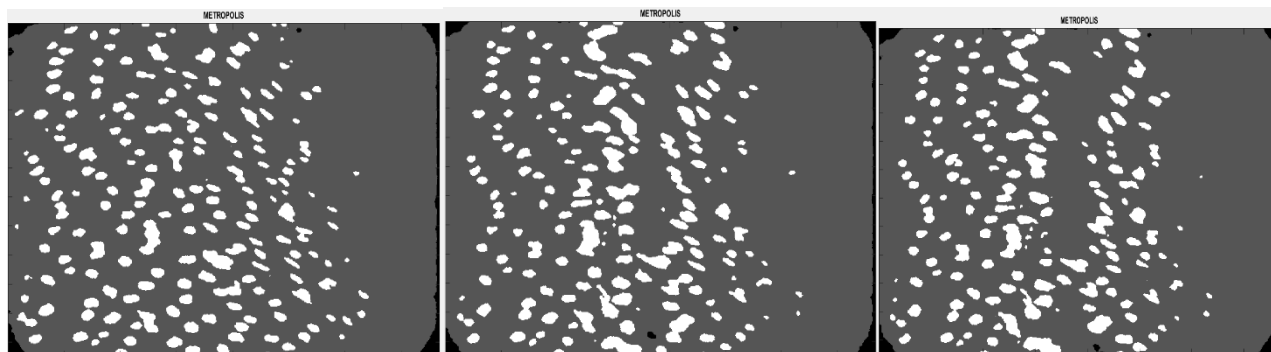


Figura 24. Segmentación de células con el algoritmo Metrópolis.

4.8 RESULTADOS DEL MUESTREO DE GIBBS

Como se puede observar en la figura 24 este algoritmo no logra segmentar todas las células, de hecho, en la parte derecha de la imagen ya sea antes del corte o después del corte no segmenta ninguna célula. Este algoritmo logra segmentar bien las células más cerca al área de la lesión, y al igual que el algoritmo de Metrópolis la segmentación no cambia tanto como los algoritmos de K-means y C-means, que varía mucho la segmentación antes y después del corte. Se contaron 158 células antes de corte y 97 células después del corte. Metrópolis tarda en promedio 1.5 minutos en segmentar una imagen.

El algoritmo de Metrópolis es muy general y es necesario afinar diferentes parámetros para tener una distribución propuesta hasta que el algoritmo funcione correctamente, esto puede llegar a ser muy complicado. Por otra parte, el Muestreo de Gibbs no necesita de una distribución propuesta, esta es su principal ventaja.

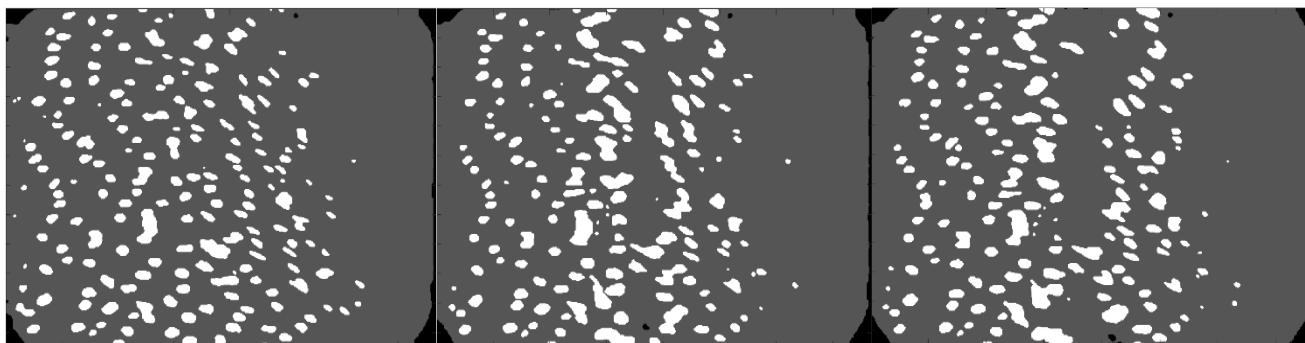


Figura 25. Segmentación de células por el Muestreo de Gibbs.

Los últimos dos algoritmos son basados en recorrido simulado SA, esto quiero decir que manejan conceptos como temperatura mejor explicados en el tema 2.4 de esta tesis. El muestreo de Gibbs tarda aproximadamente 74.547 segundos en segmentar una imagen. En la imagen antes de corte el algoritmo segmento 161 células y 102 después del corte.

4.9 RESULTADOS DEL ALGORITMO DE UMBRALIZACIÓN ADAPTATIVO

ESCRIBIR POR EL ALGORITMO ADAPTATIVO

A continuación, se mostrara una tabla donde se aprecia mejor los resultados de estos algoritmos.

| | Células Segmentadas | | Células NO Segmentadas | | Tiempo por imagen |
|------------|---------------------|---------|------------------------|---------|-------------------|
| | correctamente | | correctamente | | |
| | Antes | Después | Antes | Después | |
| K MEANS | 111 | 99 | 77 | 92 | 4.7 seg. |
| C MEANS | 114 | 96 | 82 | 90 | 18.7 seg. |
| M. MEZCLAS | 123 | 65 | 97 | 69 | 9.5 seg. |
| ICM | 75 | 130 | 94 | 77 | 150 seg. |
| METRÓPOLIS | 158 | 33 | 97 | 56 | 90 seg. |
| GIBBS | 161 | 84 | 72 | 58 | 74.5 seg. |
| ADAPTATIVA | 138 | 132 | 48 | 54 | 89.15 seg. |

Tabla 4. Comparación de algoritmos de segmentación de células endoteliales.

Como se puede observar en la tabla 4 el mismo algoritmo segmenta de forma diferente las células antes y después del corte esto se debe a que la intensidad de las células cambia drásticamente al absorber el Ca^{2+} derramado después de la lesión. Se puede concluir que el mejor algoritmo para segmentar es la umbralización adaptativa ya que es el que segmenta más células correctamente antes y después del corte además de que en su algoritmo no es necesario afinar parámetros específicos.

4.10 RESULTADOS DE LA MEDICIÓN DE LAS CONCENTRACIONES DE CALCIO Y EL ÁREA

Para medir la concentración de calcio primero procesamos la imagen original para su división y etiquetado, segmentamos y binarizamos las células, esto quiere decir que la imagen la convertimos en ceros y unos, donde el uno nos indica que es la célula endotelial, después definimos un elemento estructurante para después hacer un cierre morfológico en la imagen, esto es una dilatación seguida de una erosión utilizando el mismo elemento estructurante para ambas operaciones.

La segunda parte será el etiquetado de las células. Los píxeles con un valor de 0 son el fondo de la imagen, los píxeles con valor de 1 son aquellos que forman un objeto, en nuestro caso formarán una célula, para esto se recorre toda la imagen y se asignan etiquetas preliminares, se vuelve a recorrer la imagen y se asigna la misma etiqueta a los píxeles que estén conectados, con esto se tendrá una única etiqueta para cada célula.

El siguiente paso será calcular el promedio de intensidad y el área de la célula, para esto utilizamos un método que devuelve mediciones para el conjunto de propiedades especificado para cada objeto conectado, esto es propiedades para cada una de las células. Estas propiedades son: el área que es el número real de píxeles de la región, el centro de cada célula y la intensidad promedio. Para esta última propiedad primero se hará la intersección de la célula con la dilatación de la misma célula, con esto eliminaremos cualquier ruido en la célula.

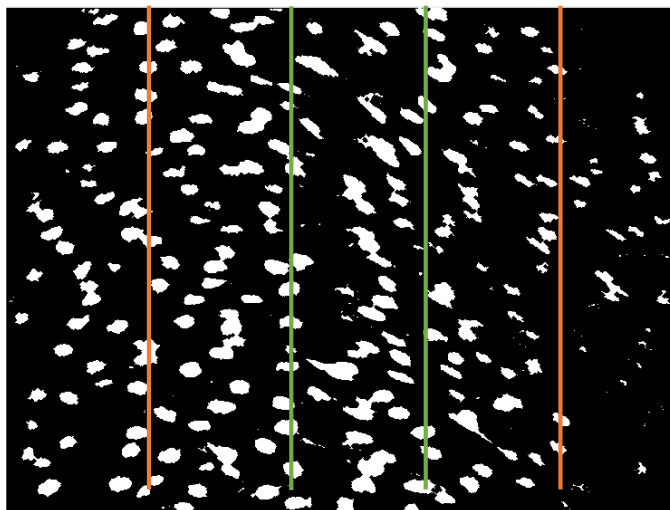


Figura 26. Imagen de estudio dividida por secciones.

Para una mejor comprensión de los datos dividimos la imagen, la primera sección será la más cercana al área de la lesión, la siguiente será a distancia media y la última serán las células más alejadas al área de la lesión, como se muestra en la figura 26.

4.11 GRANULOMETRÍA DIGITAL

La granulometría es el estudio de la distribución estadística de los tamaños de una colección de elementos de un material sólido, fraccionado o líquido multifásico. El análisis granulométrico es el conjunto de operaciones cuyo fin es determinar la distribución del tamaño de los elementos que componen una muestra. Para que la granulometría óptica sea completamente exitosa, se debe tomar una foto precisa, con suficiente iluminación y utilizando la tecnología adecuada, para obtener resultados cuantificados. Si se cumplen estos requisitos, se puede implementar un sistema de análisis de imágenes [61]. Lo más importante en esta área es el estudio del tamaño, la forma y el número. Para esto graficaremos el área de las células a través del tiempo y la intensidad de cada una de ellas.

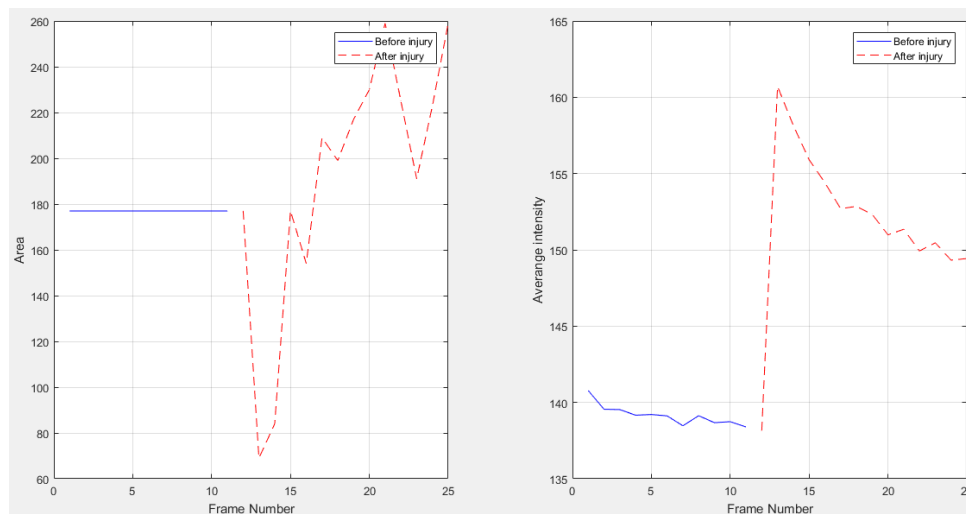


Figura 27. Grafica del lado izquierdo del área y del lado derecho el promedio de intensidad a través del tiempo.

En la figura 27 del lado izquierda se muestra la intensidad promedio de la célula, antes del corte se mantiene la intensidad y después del corte la intensidad aumenta y baja gradualmente, esto se debe al desprendimiento de calcio de las células cortadas. La figura 27 muestra el área y la intensidad de una célula cercana al área de la lesión.

Como se puede observar en la gráfica del área de la célula, esta se mantiene hasta el momento de la lesión después disminuye drásticamente para después aumentar gradualmente con diferentes oscilaciones a través del tiempo.

En la siguiente grafica se muestra los cambios de área y de intensidad en una célula a distancia media del área de la lesión.

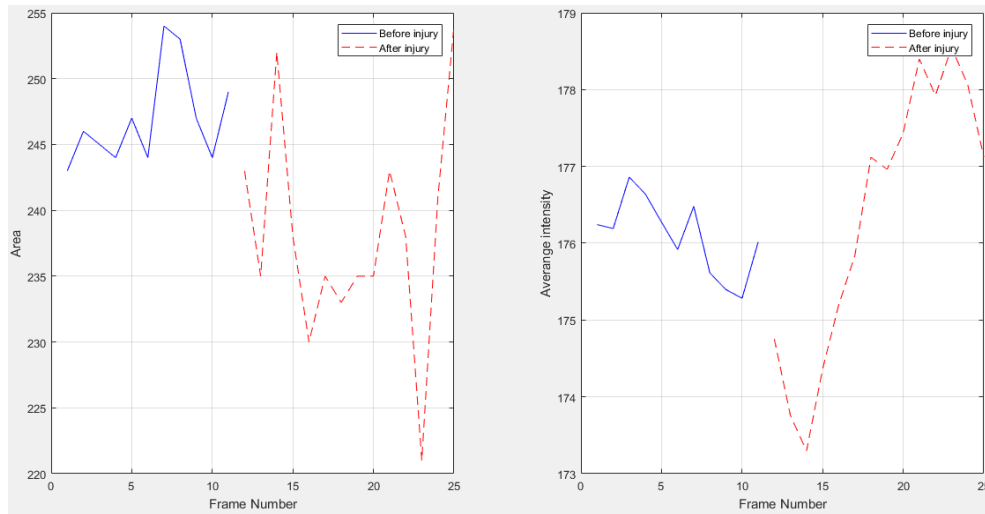


Figura 28. Grafica del lado izquierdo del área y del lado derecho el promedio de intensidad a través del tiempo.

Como se puede observar en la figura 28, tanto el área como la intensidad cambian a través del tiempo, a diferencia de lo que ocurre en una célula que se encuentra cerca del área de la lesión. El área sufre grandes cambios después del corte, tiene oscilaciones hasta alcanzar su área mayor, es de esperarse ya que al absorber el calcio la célula incrementa su área, la intensidad promedio incrementa al igual que una célula cercana a la lesión. Existen picos al final de la intensidad y en el área de la célula y esto es porque el fluoruro con el que se inyectó a las células para identificar el Ca^{2+} se va esparciendo hasta las células más lejanas y hay oscilaciones al absorber el Ca^{2+} de las células eliminadas.

Los datos de la intensidad y el área de cada célula se guardaron en un documento de Excel donde se encuentra la información completa de cada célula a través del tiempo así como sus coordenadas en el plano. A cada célula se le asignó un identificador, estos datos se utilizaron para obtener la importancia de las células y poder identificar como es que cambian a través del tiempo.

4.12 RESULTADOS DE LA CARACTERIZACIÓN CON PAGERANK

En esta sección se muestran los resultados obtenidos con el algoritmo PageRank usando las 60 células seleccionadas. Como en el ejemplo del capítulo 3.5 se aplicó el algoritmo de PageRank para 6 células ahora se aplicará para todas las células seleccionadas que suman un total de 60 como se muestra en la figura 14.

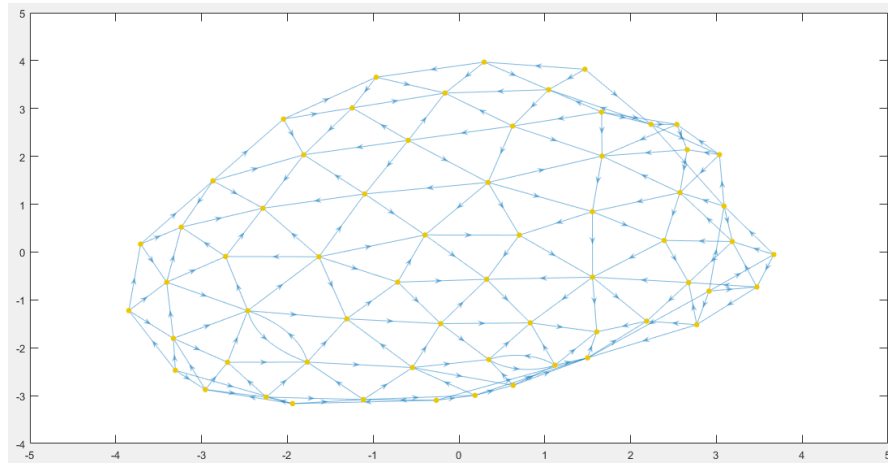
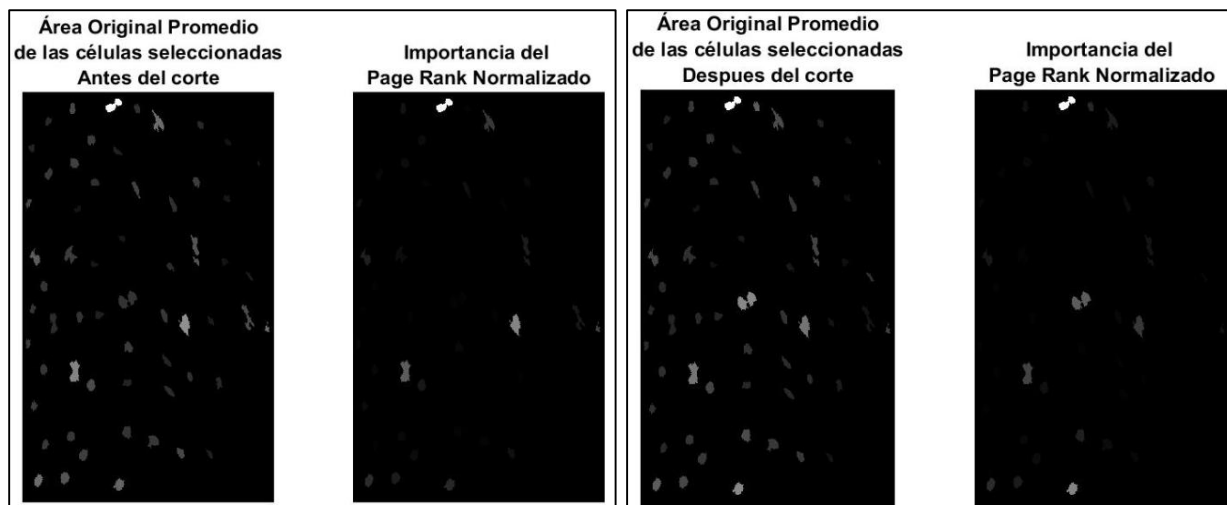


Figura 29. Grafica de PageRank con 60 células seleccionadas.

En la figura 29 se muestra el grafo construido por el algoritmo de PageRank usando las 60 células. Las aristas van dirigidas hacia la célula con mayor intensidad. Se aplicó el algoritmo de PageRank con 4 parámetros diferentes:

- Intensidad promedio antes del corte
- Intensidad promedio después del corte
- Área promedio antes del corte
- Área promedio después del corte



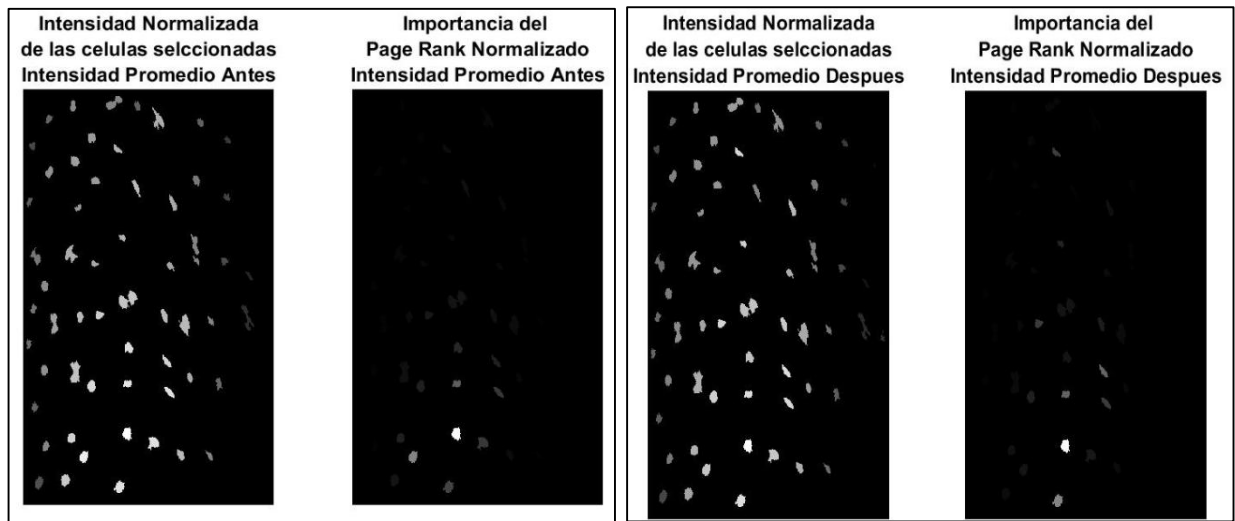


Figura 30. PageRank con diferentes parámetros de importancia.

En el cuadro superior izquierdo de la figura 30 se muestran los resultados obtenidos por el algoritmo PageRank usando los datos de la intensidad normalizada de las células antes del corte. En la imagen de la izquierda se puede observar que hay células con diferente valor en la escala de grises, esto quiere decir que las células con un color más cercano a 255 es decir (blanco), son las que tienen un valor más grande de intensidad. En la imagen de la derecha se muestran células de mayor importancia para el PageRank.

En el cuadro superior derecho se hizo el mismo proceso, pero esta vez con los datos de las intensidades promedio después del corte. Al aplicarle el algoritmo de PageRank se obtiene que las células más importantes cambian, pero se siguen manteniendo en el centro donde fue el área de corte.

En el cuadro inferior izquierdo el proceso se hizo con los parámetros del área promedio de cada célula antes del corte y a un lado se muestra las células con mayor importancia con el algoritmo de PageRank normalizado, a diferencia del PageRank con intensidades donde las células más importantes se muestran en la parte del centro inferior, con los datos del área antes del corte se muestran en el centro superior y en la parte izquierda inferior en su mayoría.

En el cuadro inferior derecho el proceso se realiza con los datos del área promedio de cada célula después del corte y al realizar el proceso de PageRank las células más importantes se encuentran distribuidas en el centro de la imagen donde fue el área de la lesión.

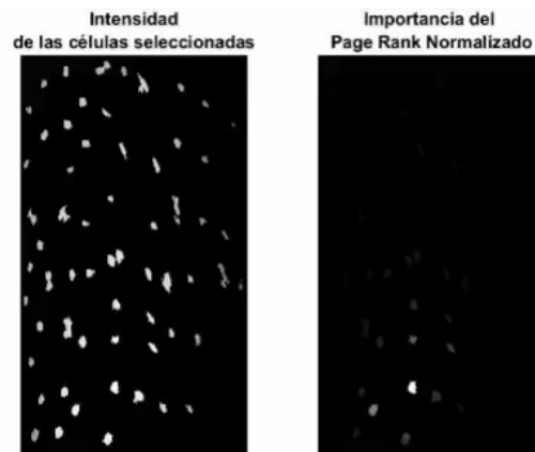


Figura 31. Captura de pantalla del Video 1 adjunto. Intensidad de las células y PageRank de 24 imágenes.

En los ejemplos anteriores se trabajó con intensidades y áreas promedio, en el video 1 adjunto se muestra cómo cambia las intensidades y la importancia del PageRank a través del tiempo, con un total de 24 imágenes donde las 12 primeras son antes del corte y de la 13 a la 24 son después del corte. En este video se puede apreciar como las células varían a través del tiempo en su intensidad y con esto en su importancia para PageRank.

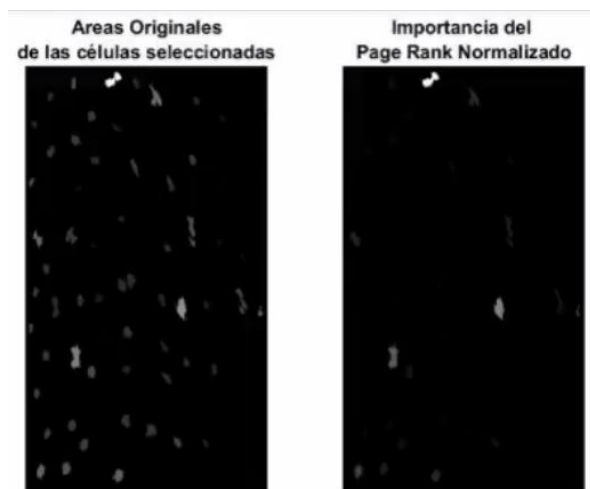


Figura 32. Captura de pantalla del video 2 adjuntos. Áreas originales y PageRank en 24 imágenes.

En el video 2 se hizo el PageRank de las 60 células y como fueron cambiando a través del tiempo con un total de 24 imágenes, a diferencia del video 1, este muestra cómo cambia

el área y las células más importantes para el PageRank con los parámetros del área de cada célula.

La Tabla 5 muestra los resultados del algoritmo de PageRank, para las 60 células, usando parámetros como son la intensidad promedio y el área promedio.

| Indice de célula | ORDEN DE IMPORTANCIA METODO DE PAGERANK | | | | PORCENTAJE DE IMPORTANCIA METODO DE PAGERANK | | | |
|------------------|---|---------|---------------|---------|--|---------|---------------|---------|
| | Intensidad Promedio | | Área Promedio | | Intensidad Promedio | | Área Promedio | |
| | Antes | Después | Antes | Después | Antes | Después | Antes | Después |
| 1 | 1 | 1 | 1 | 1 | 0.0048 | 0.0049 | 0.0067 | 0.0067 |
| 2 | 6 | 6 | 7 | 7 | 0.0048 | 0.0049 | 0.0067 | 0.0067 |
| 3 | 57 | 58 | 9 | 9 | 0.0048 | 0.0049 | 0.0067 | 0.0067 |
| 4 | 58 | 62 | 16 | 25 | 0.0048 | 0.0049 | 0.0067 | 0.0067 |
| 5 | 62 | 5 | 51 | 34 | 0.0048 | 0.0057 | 0.0067 | 0.0067 |
| 6 | 5 | 56 | 54 | 51 | 0.0056 | 0.0062 | 0.0067 | 0.0067 |
| 7 | 56 | 13 | 56 | 55 | 0.0058 | 0.0062 | 0.0067 | 0.0067 |
| 8 | 13 | 30 | 10 | 3 | 0.0061 | 0.0063 | 0.0075 | 0.0075 |
| 9 | 28 | 57 | 25 | 16 | 0.0062 | 0.0064 | 0.0076 | 0.0075 |
| 10 | 45 | 45 | 24 | 24 | 0.0064 | 0.0066 | 0.0077 | 0.0075 |
| 11 | 44 | 14 | 31 | 31 | 0.0066 | 0.0066 | 0.0077 | 0.0075 |
| 12 | 61 | 44 | 34 | 38 | 0.0068 | 0.0067 | 0.0078 | 0.0076 |
| 13 | 26 | 2 | 35 | 54 | 0.0069 | 0.0071 | 0.0078 | 0.0076 |
| 14 | 2 | 61 | 39 | 56 | 0.007 | 0.0072 | 0.0078 | 0.0076 |
| 15 | 7 | 7 | 42 | 29 | 0.0072 | 0.0073 | 0.0078 | 0.008 |
| 16 | 21 | 33 | 53 | 41 | 0.0076 | 0.0077 | 0.0078 | 0.008 |
| 17 | 59 | 21 | 29 | 53 | 0.0077 | 0.0077 | 0.008 | 0.0083 |
| 18 | 25 | 59 | 41 | 13 | 0.0081 | 0.0078 | 0.008 | 0.0086 |
| 19 | 14 | 37 | 45 | 45 | 0.0083 | 0.0078 | 0.0081 | 0.0088 |
| 20 | 31 | 60 | 13 | 22 | 0.0087 | 0.0086 | 0.0085 | 0.009 |
| 21 | 4 | 51 | 8 | 4 | 0.0087 | 0.0087 | 0.0086 | 0.0091 |
| 22 | 51 | 4 | 38 | 10 | 0.0087 | 0.0088 | 0.0089 | 0.0091 |
| 23 | 3 | 3 | 52 | 52 | 0.0087 | 0.0088 | 0.0091 | 0.0092 |
| 24 | 60 | 11 | 55 | 39 | 0.0089 | 0.0097 | 0.0092 | 0.0093 |
| 25 | 41 | 52 | 3 | 59 | 0.0094 | 0.0099 | 0.0096 | 0.0093 |
| 26 | 11 | 55 | 20 | 62 | 0.0095 | 0.0102 | 0.0096 | 0.0098 |
| 27 | 55 | 12 | 62 | 18 | 0.0096 | 0.0103 | 0.0097 | 0.0099 |
| 28 | 16 | 50 | 59 | 30 | 0.0099 | 0.0104 | 0.0097 | 0.0104 |
| 29 | 12 | 8 | 18 | 8 | 0.0102 | 0.011 | 0.0102 | 0.0106 |
| 30 | 52 | 10 | 44 | 42 | 0.0103 | 0.0115 | 0.0103 | 0.0108 |
| 31 | 50 | 54 | 22 | 20 | 0.0106 | 0.0124 | 0.0103 | 0.0108 |
| 32 | 8 | 26 | 26 | 33 | 0.0109 | 0.0124 | 0.0113 | 0.0112 |
| 33 | 46 | 38 | 32 | 26 | 0.011 | 0.0125 | 0.0118 | 0.0114 |
| 34 | 27 | 15 | 12 | 21 | 0.0112 | 0.0129 | 0.0119 | 0.012 |
| 35 | 10 | 27 | 37 | 12 | 0.0113 | 0.013 | 0.0119 | 0.0123 |
| 36 | 54 | 32 | 4 | 57 | 0.0122 | 0.0135 | 0.012 | 0.0124 |
| 37 | 42 | 28 | 21 | 5 | 0.0138 | 0.0137 | 0.0128 | 0.0127 |
| 38 | 32 | 16 | 57 | 27 | 0.014 | 0.0137 | 0.0128 | 0.0129 |
| 39 | 40 | 41 | 27 | 40 | 0.0145 | 0.0141 | 0.0128 | 0.0129 |
| 40 | 49 | 49 | 5 | 60 | 0.0158 | 0.0145 | 0.013 | 0.0143 |
| 41 | 29 | 46 | 60 | 49 | 0.0163 | 0.0151 | 0.0138 | 0.0148 |
| 42 | 37 | 53 | 40 | 46 | 0.0165 | 0.0153 | 0.0139 | 0.0153 |
| 43 | 43 | 43 | 46 | 50 | 0.017 | 0.0156 | 0.014 | 0.0157 |
| 44 | 43 | 43 | 43 | 27 | 0.0176 | 0.0155 | 0.0145 | 0.0162 |
| 46 | 48 | 40 | 30 | 58 | 0.0197 | 0.0163 | 0.015 | 0.0183 |
| 47 | 53 | 47 | 2 | 23 | 0.02 | 0.0173 | 0.018 | 0.0194 |
| 48 | 38 | 9 | 47 | 32 | 0.0205 | 0.0178 | 0.0198 | 0.0202 |
| 49 | 39 | 31 | 33 | 47 | 0.0206 | 0.019 | 0.0214 | 0.0214 |
| 50 | 24 | 19 | 50 | 43 | 0.0208 | 0.02 | 0.0216 | 0.0217 |
| 51 | 20 | 36 | 15 | 15 | 0.0212 | 0.0205 | 0.0219 | 0.0224 |
| 52 | 47 | 20 | 58 | 2 | 0.0215 | 0.0207 | 0.0248 | 0.0234 |
| 53 | 19 | 25 | 14 | 19 | 0.0215 | 0.0214 | 0.0313 | 0.0238 |
| 54 | 15 | 17 | 43 | 35 | 0.0224 | 0.0236 | 0.0322 | 0.0246 |
| 55 | 17 | 18 | 23 | 6 | 0.0229 | 0.0245 | 0.0371 | 0.0257 |
| 56 | 18 | 24 | 61 | 14 | 0.0306 | 0.0301 | 0.0381 | 0.0318 |
| 57 | 22 | 42 | 48 | 44 | 0.0321 | 0.0323 | 0.0392 | 0.032 |
| 58 | 23 | 29 | 36 | 48 | 0.0405 | 0.0433 | 0.0453 | 0.0411 |
| 59 | 30 | 39 | 19 | 11 | 0.0441 | 0.0664 | 0.0458 | 0.0467 |
| 60 | 34 | 35 | 11 | 28 | 0.0639 | 0.0719 | 0.0464 | 0.0658 |

Tabla 5. Resultados de los resultados de PageRank con diferentes parámetros.

Como se puede observar en la tabla 5 del lado derecho se muestra el porcentaje de importancia de cada célula y cómo cambia si se mide la importancia antes o después del corte, también si se mide la intensidad o el área promedio.

5. SEGUNDO ANÁLISIS CON NUEVAS CÉLULAS

Una vez obtenidos los resultados del PageRank de las 60 células seleccionadas, se hará un segundo análisis seleccionando 30 células diferentes cercanas al área de la lesión, a distancia media y las más alejadas como se hizo con las primeras 60 células.

Se obtuvo su intensidad y área promedio, así como la triangulación Delaunay de esas 30 células y se obtuvo su importancia con el algoritmo de PageRank los resultados se muestran en la figura 33 y la tabla de los datos de las siguientes 30 células seleccionadas.

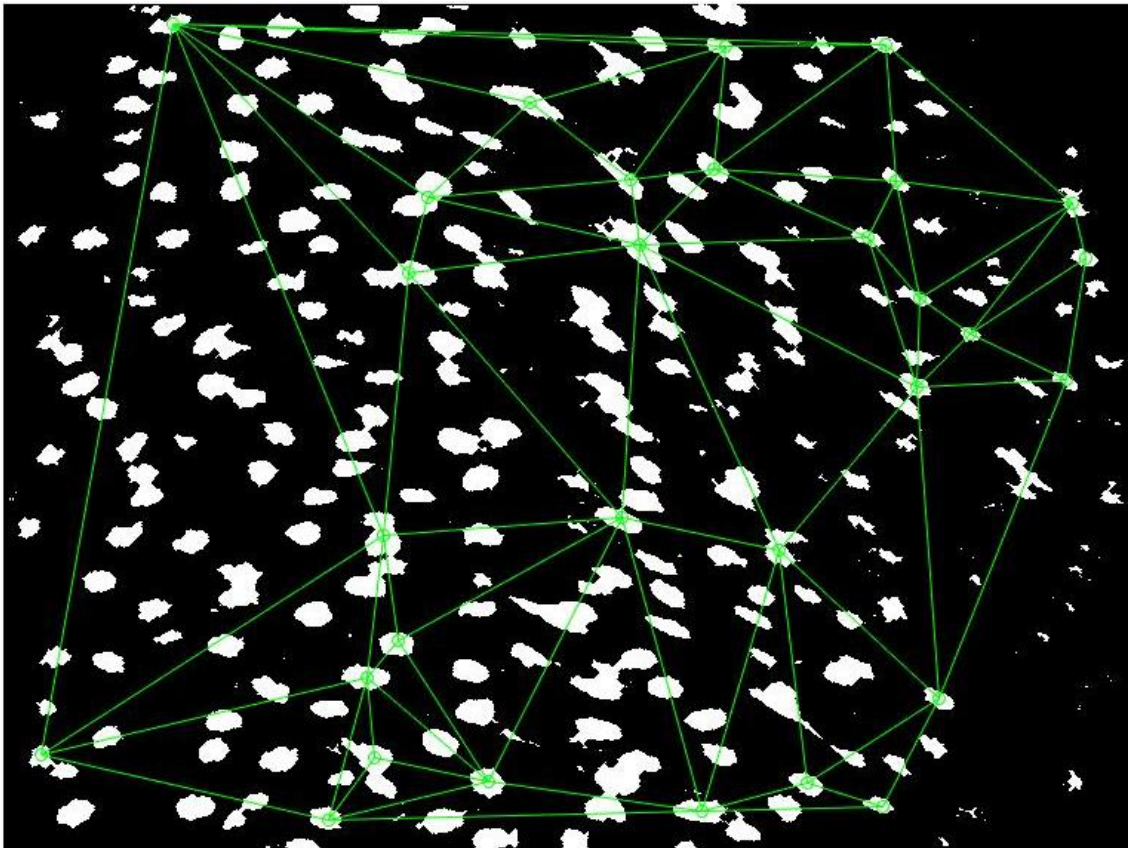


Figura 33. Triangulación Delaunay con 30 nuevas células seleccionadas.

| | Centro de célula | | Intensidad Promedio | | Area Promedio | | Intensidad Promedio | | Area Promedio | | Intensidad Promedio | | Area Promedio | |
|----|------------------|-----|---------------------|---------|---------------|---------|---------------------|---------|---------------|---------|---------------------|---------|---------------|---------|
| | x | y | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues |
| 61 | 115 | 14 | 139 | 138 | 574 | 661 | 4 | 4 | 1 | 1 | 0.0160 | 0.0167 | 0.2045 | 0.2059 |
| 62 | 26 | 507 | 133 | 129 | 477 | 457 | 7 | 7 | 2 | 7 | 0.0136 | 0.0126 | 0.1105 | 0.0335 |
| 63 | 220 | 551 | 189 | 196 | 217 | 248 | 14 | 6 | 4 | 9 | 0.0374 | 0.0426 | 0.0167 | 0.0196 |
| 64 | 251 | 509 | 218 | 229 | 457 | 505 | 6 | 5 | 7 | 12 | 0.2050 | 0.2041 | 0.0776 | 0.0502 |
| 65 | 246 | 455 | 199 | 221 | 263 | 303 | 5 | 13 | 9 | 4 | 0.0500 | 0.0604 | 0.0220 | 0.0255 |
| 66 | 267 | 430 | 200 | 222 | 231 | 252 | 13 | 3 | 10 | 15 | 0.0606 | 0.0774 | 0.0185 | 0.0208 |
| 67 | 257 | 359 | 210 | 225 | 443 | 639 | 3 | 8 | 12 | 10 | 0.1108 | 0.1103 | 0.0605 | 0.1113 |
| 68 | 274 | 182 | 172 | 194 | 303 | 441 | 15 | 9 | 15 | 2 | 0.0256 | 0.0379 | 0.0256 | 0.0303 |
| 69 | 287 | 131 | 172 | 194 | 442 | 574 | 10 | 15 | 14 | 8 | 0.0256 | 0.0379 | 0.0499 | 0.0781 |
| 70 | 356 | 67 | 177 | 181 | 404 | 470 | 12 | 12 | 30 | 30 | 0.0302 | 0.0236 | 0.0426 | 0.0376 |
| 71 | 424 | 119 | 166 | 169 | 232 | 240 | 8 | 14 | 8 | 5 | 0.0207 | 0.0207 | 0.0195 | 0.0185 |
| 72 | 430 | 163 | 174 | 184 | 370 | 544 | 9 | 10 | 19 | 13 | 0.0276 | 0.0275 | 0.0373 | 0.0609 |
| 73 | 418 | 347 | 196 | 218 | 185 | 297 | 19 | 21 | 5 | 14 | 0.0427 | 0.0498 | 0.0159 | 0.0237 |
| 74 | 328 | 525 | 209 | 182 | 306 | 294 | 11 | 11 | 27 | 6 | 0.0778 | 0.0254 | 0.0304 | 0.0221 |
| 75 | 472 | 545 | 179 | 188 | 350 | 482 | 21 | 22 | 11 | 3 | 0.0334 | 0.0301 | 0.0333 | 0.0429 |
| 76 | 543 | 526 | 157 | 149 | 219 | 175 | 16 | 19 | 6 | 11 | 0.0185 | 0.0175 | 0.0175 | 0.0147 |
| 77 | 594 | 542 | 141 | 131 | 161 | 145 | 22 | 16 | 16 | 22 | 0.0167 | 0.0130 | 0.0141 | 0.0131 |
| 78 | 632 | 469 | 137 | 133 | 162 | 167 | 17 | 1 | 3 | 21 | 0.0146 | 0.0140 | 0.0146 | 0.0141 |
| 79 | 524 | 369 | 168 | 167 | 303 | 178 | 1 | 25 | 13 | 27 | 0.0221 | 0.0184 | 0.0256 | 0.0153 |
| 80 | 617 | 259 | 133 | 132 | 151 | 164 | 25 | 24 | 22 | 19 | 0.0136 | 0.0135 | 0.0135 | 0.0136 |
| 81 | 480 | 112 | 160 | 178 | 129 | 213 | 18 | 29 | 18 | 16 | 0.0195 | 0.0220 | 0.0131 | 0.0168 |
| 82 | 488 | 29 | 149 | 168 | 168 | 229 | 24 | 18 | 17 | 18 | 0.0175 | 0.0195 | 0.0153 | 0.0176 |
| 83 | 596 | 28 | 124 | 124 | 114 | 130 | 2 | 20 | 20 | 20 | 0.0118 | 0.0118 | 0.0122 | 0.0127 |
| 84 | 603 | 120 | 134 | 134 | 114 | 119 | 20 | 17 | 21 | 17 | 0.0141 | 0.0153 | 0.0122 | 0.0119 |
| 85 | 584 | 158 | 139 | 138 | 128 | 123 | 29 | 2 | 25 | 23 | 0.0160 | 0.0167 | 0.0126 | 0.0122 |
| 86 | 619 | 199 | 130 | 128 | 105 | 113 | 26 | 26 | 23 | 25 | 0.0122 | 0.0122 | 0.0114 | 0.0115 |
| 87 | 722 | 135 | 98 | 96 | 240 | 209 | 23 | 23 | 24 | 24 | 0.0111 | 0.0111 | 0.0207 | 0.0160 |
| 88 | 731 | 172 | 124 | 124 | 45 | 46 | 28 | 28 | 26 | 26 | 0.0118 | 0.0118 | 0.0108 | 0.0108 |
| 89 | 653 | 223 | 132 | 134 | 104 | 102 | 27 | 27 | 29 | 29 | 0.0126 | 0.0153 | 0.0111 | 0.0112 |
| 90 | 718 | 254 | 94 | 93 | 306 | 318 | 30 | 30 | 28 | 28 | 0.0108 | 0.0108 | 0.0304 | 0.0277 |

Tabla 6. Información de las nuevas 30 células seleccionadas

Los resultados de la importancia fueron comparados con las primeras 60 células seleccionadas y se observó que el resultado es muy similar, las células que se encuentran justo en el área de la lesión son las más importantes, ya que tienen más acumulación de iones de calcio que las demás, son aquellas que muestran un patrón muy específico.

Las células más cercanas al área de la lesión muestran en las gráficas que justo en el momento de infligir una lesión se cargan de los iones de calcio, para después ir disminuyendo a través del tiempo.

Para estar seguros se seleccionaron una última vez 30 células nuevas y fueron estudiadas a través del algoritmo de PageRank y conectadas por la Triangulación Delaunay.

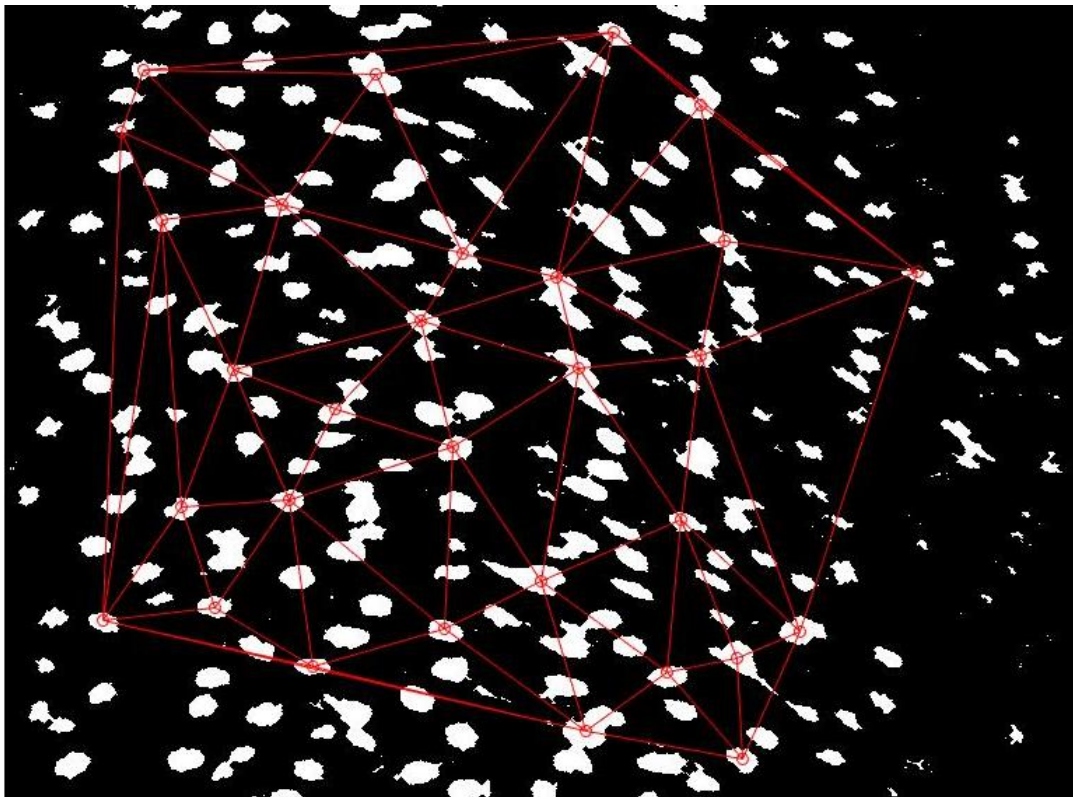


Figura 34. Triangulación Delaunay con las últimas 30 células seleccionadas.

| | | | | | | | ORDEN DE IMPORTANCIA METODO DE PAGERANK | | | | PORCENTAJE DE IMPORTANCIA METODO DE PAGERANK | | | |
|------------------|-----|-----|---------------------|---------|---------------|---------|---|---------|---------------|---------|--|---------|---------------|---------|
| Centro de célula | | | Intensidad Promedio | | Area Promedio | | Intensidad Promedio | | Area Promedio | | Intensidad Promedio | | Area Promedio | |
| | x | y | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues | Antes | Despues |
| 91 | 100 | 50 | 152 | 152 | 167 | 170 | 13 | 14 | 8 | 15 | 0.0118 | 0.0118 | 0.0122 | 0.0122 |
| 92 | 84 | 94 | 156 | 152 | 97 | 65 | 14 | 9 | 13 | 13 | 0.0122 | 0.0118 | 0.0108 | 0.0108 |
| 93 | 113 | 157 | 158 | 157 | 146 | 150 | 15 | 13 | 16 | 8 | 0.0130 | 0.0126 | 0.0119 | 0.0115 |
| 94 | 164 | 264 | 188 | 182 | 223 | 260 | 7 | 7 | 15 | 10 | 0.0379 | 0.0195 | 0.0185 | 0.0176 |
| 95 | 127 | 362 | 171 | 169 | 181 | 189 | 12 | 15 | 20 | 11 | 0.0175 | 0.0167 | 0.0136 | 0.0126 |
| 96 | 151 | 434 | 185 | 184 | 204 | 212 | 9 | 12 | 17 | 20 | 0.0275 | 0.0221 | 0.0153 | 0.0147 |
| 97 | 221 | 476 | 199 | 208 | 179 | 210 | 4 | 8 | 27 | 9 | 0.0603 | 0.0606 | 0.0127 | 0.0141 |
| 98 | 266 | 53 | 172 | 201 | 543 | 541 | 26 | 27 | 11 | 16 | 0.0184 | 0.0374 | 0.2059 | 0.0779 |
| 99 | 298 | 229 | 189 | 215 | 245 | 436 | 10 | 25 | 30 | 27 | 0.0425 | 0.1108 | 0.0237 | 0.0375 |
| 100 | 395 | 198 | 187 | 180 | 222 | 516 | 6 | 26 | 24 | 19 | 0.0301 | 0.0185 | 0.0176 | 0.0607 |
| 101 | 411 | 263 | 183 | 190 | 300 | 514 | 27 | 11 | 19 | 18 | 0.0235 | 0.0254 | 0.0335 | 0.0501 |
| 102 | 321 | 319 | 197 | 202 | 195 | 279 | 11 | 22 | 9 | 17 | 0.0497 | 0.0427 | 0.0141 | 0.0185 |
| 103 | 384 | 415 | 218 | 211 | 513 | 589 | 25 | 6 | 29 | 25 | 0.2040 | 0.0778 | 0.1113 | 0.1110 |
| 104 | 315 | 449 | 213 | 231 | 234 | 281 | 22 | 28 | 14 | 14 | 0.1103 | 0.2050 | 0.0208 | 0.0208 |
| 105 | 416 | 522 | 204 | 206 | 403 | 807 | 28 | 4 | 22 | 24 | 0.0774 | 0.0500 | 0.0609 | 0.2053 |
| 106 | 524 | 470 | 171 | 165 | 462 | 423 | 8 | 10 | 4 | 12 | 0.0175 | 0.0147 | 0.0781 | 0.0334 |
| 107 | 569 | 451 | 157 | 152 | 347 | 324 | 5 | 24 | 10 | 4 | 0.0126 | 0.0118 | 0.0429 | 0.0236 |
| 108 | 515 | 172 | 152 | 164 | 219 | 339 | 16 | 5 | 18 | 30 | 0.0118 | 0.0136 | 0.0168 | 0.0255 |
| 109 | 436 | 23 | 161 | 167 | 277 | 352 | 24 | 29 | 23 | 29 | 0.0141 | 0.0153 | 0.0255 | 0.0276 |
| 110 | 498 | 75 | 148 | 165 | 370 | 511 | 23 | 19 | 6 | 22 | 0.0111 | 0.0147 | 0.0502 | 0.0428 |
| 111 | 653 | 194 | 127 | 126 | 121 | 130 | 29 | 16 | 26 | 6 | 0.0108 | 0.0108 | 0.0112 | 0.0111 |
| 112 | 474 | 480 | 179 | 188 | 232 | 240 | 19 | 20 | 12 | 7 | 0.0206 | 0.0236 | 0.0196 | 0.0153 |
| 113 | 71 | 443 | 162 | 159 | 206 | 192 | 30 | 18 | 5 | 26 | 0.0153 | 0.0131 | 0.0160 | 0.0131 |
| 114 | 199 | 146 | 170 | 174 | 287 | 281 | 3 | 23 | 25 | 23 | 0.0159 | 0.0175 | 0.0277 | 0.0208 |
| 115 | 328 | 181 | 181 | 197 | 180 | 284 | 17 | 3 | 7 | 5 | 0.0220 | 0.0302 | 0.0131 | 0.0221 |
| 116 | 204 | 357 | 188 | 193 | 203 | 196 | 2 | 30 | 1 | 1 | 0.0379 | 0.0276 | 0.0147 | 0.0136 |
| 117 | 237 | 292 | 184 | 199 | 340 | 364 | 1 | 1 | 3 | 28 | 0.0254 | 0.0334 | 0.0376 | 0.0302 |
| 118 | 484 | 371 | 176 | 184 | 141 | 169 | 18 | 2 | 28 | 3 | 0.0195 | 0.0221 | 0.0115 | 0.0118 |
| 119 | 498 | 254 | 162 | 169 | 243 | 243 | 20 | 17 | 21 | 21 | 0.0153 | 0.0167 | 0.0221 | 0.0160 |
| 120 | 528 | 542 | 161 | 155 | 293 | 260 | 21 | 21 | 2 | 2 | 0.0141 | 0.0122 | 0.0303 | 0.0176 |

Tabla 7. Información de las últimas 30 células seleccionadas

En la tabla 6 y 7 muestra la importancia del PageRank con solo su conjunto de células, es decir, la importancia solo es de las nuevas 30 células seleccionadas y de las últimas 30 células seleccionadas y así se compara la importancia con las primeras 60 células seleccionadas.

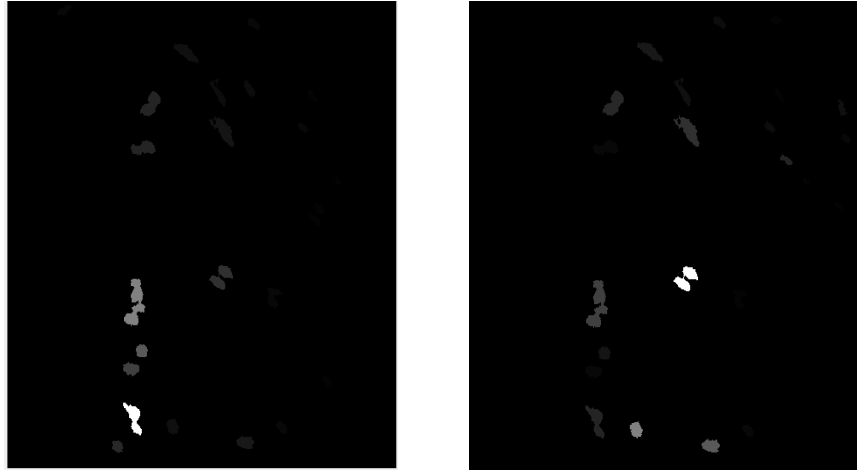


Figura 35. Del lado izquierdo la importancia del PageRank de las primeras 60 células, y del lado derecho la importancia del PageRank de las últimas 60 células seleccionadas.

Los datos obtenidos con las primeras 60 células en el PageRank son muy similares con las siguientes 60 células seleccionadas. Con esto concluimos que las células más importantes siempre se encontrarán justo en el área de la lesión. Esto ayudará a estudios posteriores de cómo medir la concentración de iones de calcio de un grupo de células y de una célula en particular, y así poder centrarse en el estudio de ciertas células en particular.

CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se evaluaron seis métodos de segmentación. Además se implementó un método para obtener el índice de la concentración de calcio en una secuencia de imágenes; este método evalúa una célula en particular a través del tiempo, como cambia su área y su intensidad promedio. Como vimos existen diferentes métodos de segmentación y en algunos casos el tiempo de ejecución es mucho para segmentar una sola imagen, podemos concluir que el mejor método de segmentación es el método adaptivo ya que separa la mayor cantidad de células antes y después del corte.

Además, se desarrollaron una serie de métodos de procesamiento de imágenes para mejorar el resultado cuantitativo de los métodos experimentales existentes en biología de células endoteliales.

Una vez hecha la segmentación se hizo un proceso de etiquetado para poder identificar a cada célula dentro de la imagen. El siguiente punto importante para lograr buenos resultados es la forma en cómo se conectaron las células entre ellas para poder comparar su comportamiento entre ellas mismas, para esto utilizamos la triangulación de Delaunay y como siguiente paso detectar las células con mayor importancia a través del algoritmo de PageRank.

En específico se diseñó un algoritmo de caracterización del índice de concentración de calcio en secuencias de imágenes del tejido endotelial antes y después de una lesión.

Un algoritmo de detección de núcleos celulares en imágenes microscópicas que detecte regiones de alta luminosidad.

Se calcula el índice de variabilidad de la concentración de calcio en el núcleo calculando el promedio de las intensidades de los píxeles que representan una célula o conjuntos de ellas.

Durante el proceso de PageRank se observó que las células con más cambios se encuentran justo en el área de la lesión y dependiendo en qué momento del tiempo se encuentre las células de mayor importancia irán cambiando, concluimos que el mejor parámetro para estudiar las células con el algoritmo de PageRank es la intensidad de cada una de las células.

En el capítulo 4.11 se puede observar cómo es que cambia el área y la intensidad de las células a través del tiempo, esto nos ayudará a entender mejor como es el comportamiento de las células después de infligir una lesión.

Con esto se espera tener la posibilidad de realizar mediciones de función endotelial de forma rápida y segura, y así poder reproducir la aplicación de este método en forma amplia.

REFERENCIAS

- [1] Ayaori, M.; Iwakami, N.; Uto-Kondo, H.; Sato, H.; Sasaki, M.; Komatsu, T.; Iizuka, M.; Takiguchi, S.; Yakushiji, E.; Nakaya, K.; et al. Dipeptidyl peptidase-4 inhibitors attenuate endothelial function as evaluated by flow-mediated vasodilatation in type 2 diabetic patients. *J. Am. Heart Assoc.* 2013, 2, 7–10. [CrossRef] [PubMed]
- [2] Moccia, F.; Tanzi, F.; Munaron, L. Endothelial Remodelling and Intracellular Calcium Machinery. *Curr. Mol. Med.* 2014, 14, 457–480. [CrossRef] [PubMed].
- [3] B. A. S. G. Eskin, C. L. Ives, L. V. McIntire, and L. T. Navarro, “Response of cultured endothelial cells to steady flow,” *Microvascular Research*, vol. 28, no. 1, pp. 87–94, 1983.
- [4] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*. New York: Garland Science, 4 ed., 2002.
- [5] Bagher P, Segal SS (2011) Regulation of Blood Flow In the Microcirculation: Role of Conducted Vasodilation. *Acta Physiol* 202: 271–284.
- [6] Berra-Romani, R.; Raqeeb, A.; Torres-Jácome, J.; Guzman-Silva, A.; Guerra, G.; Tanzi, F.; Moccia, F. The mechanism of injury-induced intracellular calcium concentration oscillations in the endothelium of excised rat aorta. *J. Vasc. Res.* 2011, 49, 65–76. [CrossRef] [PubMed].
- [7] Iyer, R. K., Chui, J., & Radisic, M. (2009). Spatiotemporal Tracking of Cells in Tissue Engineered Cardiac Organoids. *Journal of Tissue Engineering and Regenerative Medicine*, 3(3), 196–207.
- [8] Prahlad G. Menon, Elizabeth R. Rochon, Beth L. Roman, "In-vivo cell tracking to quantify endothelial cell migration during zebrafish angiogenesis", *Proc. SPIE 9784, Medical Imaging 2016: Image Processing*, 97844A (21 March 2016).
- [9] D. L. Spector and R. D. Goldman, *Basic Methods in Microscopy a laboratory manual*. New York: Cold Spring Harbor Laboratory Press, 2006.
- [10] K. Thorn, “A quick guide to light microscopy in cell biology,” *Molecular Biology of the Cell*, vol. 27, no. 2, pp. 219–222, 2016.
- [11] L. M. Lambert, I. I. Pipinos, B. T. Baxter, In vitro measurements of hemodynamic forces and their effects on endothelial cell mechanics at the sub-cellular level. *American Institute of Physics*.
- [12] Bagher P, Segal SS (2011) Regulation of Blood Flow In the Microcirculation: Role of Conducted Vasodilation. *Acta Physiol* 202: 271–284.
- [13] Tran QK, Ohashi K, Watanabe H (2000) Calcium signaling in endothelial cells. *Cardiovascular Research* 48: 13–22.
- [14] Molnar, C., Jermyn, IH, Kato, Z., Rahkama, V., Östling, P., Mikkonen, P., ... y Horvath, P. (2016). Morfología precisa que preserva la segmentación de las células superpuestas en función de los contornos activos. *Informes científicos*, 6, 32412.
- [15] Collins, T. J. *Imagej for microscopy*. *Biotechniques* 43, 25–30 (2007).
- [16] Turkki, R. et al. Assessment of tumour viability in human lung cancer xenografts with texture-based image analysis. *J Clin Pathol* (2015).

- [17] Linder, N. et al. A malaria diagnostic tool based on computer vision screening and visualization of Plasmodium falciparum candidate areas in digitized blood smears. *PloS One* 9(8) (2014).
- [18] D. Tschumperle and R. Deriche, Diffusion PDE's on vector-valued images. *IEEE Signal Processing Magazine* 19:16-25,2002.
- [19] T. Jones, A. Carpenter, D. Sabatini, and P. Golland, "Methods for high-content, high throughput image-based cell screening," *Proceedings of the First MICCAI Workshop on Microscopic Image Analysis with Applications in Biology*, pp. 65–72, 2006.
- [20] Marcial Sanchez-Tecuatl, Ajelet Vargaz-Guadarrama, Juan Manuel Ramirez-Cortes, Pilar Gomez-Gil, Francesco Moccia and Roberto Berra-Romani. Automated Intracellular Calcium Profiles Extraction from Endothelial Cells Using Digital Fluorescence Images. (2 November 2018).
- [21] Berra-Romani, R.; Rinaldi, C.; Raqeeb, A.; Castelli, L.; Magistretti, J.; Taglietti, V.; Tanzi, F. The duration and amplitude of the plateau phase of ATP- and ADP-Evoked Ca²⁺ signals are modulated by ectonucleotidases in in situ endothelial cells of rat aorta. *J. Vasc. Res.* 2004, 41, 166–173. [CrossRef] [PubMed].
- [22] Berra-Romani, R.; Raqeeb, A.; Torres-Jácome, J.; Guzman Silva, A.; Guerra, G.; Tanzi, F.; Moccia, F. The mechanism of injury-induced intracellular calcium concentration oscillations in the endothelium of excised rat aorta. *J. Vasc. Res.* 2011, 49, 65–76. [CrossRef] [PubMed].
- [23] William S. N. Grimes, MRC Laboratory for Molecular Cell Biology, University College London Image processing and analysis methods in quantitative endothelial cell biology, September 2016.
- [24] A. R. Parker, "On the origin of optics," *Optics & Laser Technology*, vol. 43, pp. 323–329, mar 2011.
- [25] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision* (1), New York: Addison-Wesley, 1992.
- [26] T. Jones, A. Carpenter, D. Sabatini, and P. Golland, "Methods for high-content, high throughput image-based cell screening," *Proceedings of the First MICCAI Workshop on Microscopic Image Analysis with Applications in Biology*, pp. 65–72, 2006.
- [27] Iyer, R. K., Chui, J., & Radisic, M. (2009). Spatiotemporal Tracking of Cells in Tissue Engineered Cardiac Organoids. *Journal of Tissue Engineering and Regenerative Medicine*, 3(3), 196–207.
- [28] Maza-Sastre, H., Ochoa-Montiel, R., Sánchez-López, C., Pérez-Corona, C., Carrasco-Aguilar, M. A., & Morales-López, F. E. (2014, November). Identification of Trypanosoma with digital image processing. In *2014 IEEE Central America and Panama Convention (CONCAPAN XXXIV)* (pp. 1-4). IEEE.
- [29] ZEYUN, Yu and CHANDRAJIT, Bajaj. A fast adaptive method for image contrast enhancement. USA, 2001.
- [30] Y. Zhang, M. Brady, and S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20: 45-57,2001.
- [31] K. V. Mardia and T. J. Hainsworth, A spatial thresholding method for image segmentation. *IEEE PAMI*10: 919-927,1988.
- [32] N. R. Pal and S. K. Pal, A review on image segmentation techniques, *Pattern Recogn* 26:1277-94,1993.

- [33] T. Kurita, N. Otsu, and N. Abdelmalek, Maximum likelihood thresholding based on population mixture models. *Pattern Recogn* 25:1231-1240,1992.
- [34] MacQueen (1967): "Algunos métodos para la clasificación y análisis de observaciones multivariadas, Actas del 5° Simposio de Berkeley sobre estadística matemática y probabilidad" Berkeley, University of California Press, 1: 281-297.
- [35] J. B. MacQueen, Some methods for classification and analysis of multivariate observations, presented at Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California, 1967.
- [36] JC Dunn (1973): "Un pariente difuso del proceso ISODATA y su uso en la detección de grupos compactos bien separados", *Journal of Cybernetics* 3: 32-57.
- [37] O. Demirkaya and M. Asyali, Determination of image bimodality thresholds for different intensity distributions. *Signal Process: Image Commun* 19: 507-516,2004.
- [38] N. R. Pal and S. K. Pal, A review on image segmentation techniques, *Pattern Recogn* 26:1277-94, 1993.
- [39] S. S. Reddi, S. F. Rudin, and H. R. Keshavan, An optimal multiple threshold scheme for image segmentation. *IEEE Trans Syst. Man Cybernetics SMC-14*: 661-665,1984.[10] J. Besag, On the statistical analysis of dirty images, / *R Stat Soc B* 48: 259- 302,1986
- [40] G. J. McLachlan, R. W. Bean, and D. Peel, A mixture model-based approach to the clustering of microarray expression data. *Bioinformatics* 18: 413-422, 2002.
- [41] A. Dempster, N. Laird, and D. Rubin, Maximum likelihood from incomplete data via the EM algorithm. / *R. Stat. Soc B*39:1-38, 1977.
- [42] G. Schwartz, Estimating the dimension of a model. *Ann. Stat*, 461-464, 1978.
- [43] Y. Zhang, M. Brady, and S. Smith, Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm, *IEEE Trans Med Imaging* 20: 45-57, 2001
- [44] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, Equation of state calculations by fast computing machines, / *Chem Phys* 21: 1087-1092, 1953.
- [45] S. Kirkpatrick, C. D. G. Jr., and M. P. Vecchi, Optimization by simulated annealing, *Science* 220: 671-680, 1983.
- [46] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.
- [47] S. Geman and D. Geman, Stochastic Relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal Machine Intelligence* 6: 721-741, 1984.
- [48] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, 3rd, R. Kikinis, and H. W. Muller-Gartner, Markov random field segmentation of brain MR images, *IEEE Trans Med Imaging* 16: 878-86,1997.
- [49] J. Besag, On the statistical analysis of dirty images, / *R Stat Soc B* 48: 259- 302, 1986.

- [50] Haario, H., Saksman, E. y Tamminen, J. (2001). Un algoritmo adaptativo de Metrópolis. *Bernoulli*, 7 (2), 223-242.
- [51] S. Geman and D. Geman, Stochastic Relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal Machine Intelligence* 6: 721-741,1984.
- [52] W. L. Martinez and A. R. Martinez, *Computational Statistics Handbook with MATLAB*. Boca Raton, FL: CRC Press, 2001.
- [53] F. Pedregosa, O. Grisel, R. Weiss, A. Passos, and M. Brucher, "Scikit-learn : Machine Learning in Python," vol. 12, pp. 2825–2830, 2011.
- [54] Rafael C. González, Richard Eugene Woods, "Digital Image Processing", Prentice Hall, 2008.
- [55] K. Bryan and T. Leise. The 25, 000, 000, 000 Eigenvector The Linear Algebra Behind Google. *Society for Industrial and Applied Mathematics (SIAM)*. Vol 48, No 3, páginas 569 581, 2006.
- [56] A. Arasu. PageRank Computation and the Structure of the Web: Experiments and Algorithms. *Computer Science Department, Stanford University, CA 94305*.
- [57] Sergio Cruz Vázquez, Abril de 2017. Algoritmo de Jerarquización de Google.
- [58] Page, Lawrence; Brin, Sergey; Motwani, Rajeev y Winograd, Terry (1999). The PageRank citation ranking: Bringing order to the Web.
- [59] Cheng, Alice; Friedman, Eric J. Manipulability of PageRank under Sybil Strategies. *Proceedings of the First Workshop on the Economics of Networked Systems (NetEcon06)*.
- [60] Legrá-Lobaina, A. A., Atanes-Beatón, D. M., & Guilarte-Fuentes, C. (2014). Contribución al método de interpolación lineal con triangulación de Delaunay. *Minería y Geología*, 30(2), 58-72.
- [61] Maerz, NH (1996). Técnicas de muestreo de imágenes y requisitos para el análisis automatizado de imágenes de fragmentación de rocas. *Actas del FRAGBLAST*, 5, 115-120.