

Identificación y localización de células gliales en imágenes de tejido cerebral

Karla Mandujano-López¹, Juan Villegas-Cortez², Graciela Román-Alonso¹,
Arturo Zuñiga-López³, Edgar Rangel-López⁴, César Benavides-Alvarez¹

¹ Universidad Autónoma Metropolitana,
Departamento de Ingeniería Eléctrica,
Ciudad de México,
México

² Universidad Autónoma Metropolitana,
Departamento de Sistemas,
Ciudad de México,
México

³ Universidad Autónoma Metropolitana,
Departamento de Electrónica,
Ciudad de México,
México

⁴Instituto Nacional de Neurología y Neurocirugía
Manuel Velasco Suárez,
Ciudad de México,
México

{karlamanlop, cesarbenavides32}@gmail.com, juanvc@azc.uam.mx,
grac@xanum.uam.mx, raledg@hotmail.com

Resumen. En este artículo se propone un sistema de procesamiento digital de imágenes para la identificación automática de células gliales de tejido cerebral. El propósito es apoyar al especialista, que al seleccionar una región de interés (ROI) en una imagen de tejido cerebral, se procede a buscar esa ROI en toda la imagen en cuestión. La metodología propuesta trabaja a nivel textura por lo que primero aplica un preprocesamiento a toda la imagen: se extraen atributos CBIR probados en la literatura, dando como resultado la posibilidad de analizar de forma local a una imagen específica para localizar regiones similares en la misma imagen. Nuestra propuesta hace uso de una discriminación de los patrones conformados usando métrica euclidiana, tras implementar el método K-Nearest Neighbors para una auto-organización de los patrones de textura conformados. Se han realizado pruebas con un total de 312 imágenes de laboratorio, puramente experimentales obtenidas del microscopio en formato TIFF, y se han obtenido resultados prometedores bajo la supervisión del especialista en biología celular.

Palabras clave: Clasificación de células, procesamiento digital de imágenes, CBIR, K-NN, reconocimiento de patrones.

Identification and Localization of Glial Cells of the Brain Tissue

Abstract. This paper proposes an image digital processing system for the automatic identification of glial cells of the brain tissue. The purpose is to support the specialist to find within an image all the areas where exist the region of interest (ROI) previously defined by the specialist. The proposed methodology is developed at the texture level, therefore, it is first applied a preprocessing of the whole image by extracting CBIR attributes found in references. This preprocessing stage results in the possibility of analyzing on a local basis, a specific image to localize similar regions within the same image. Our proposal discriminates the conformed patterns by using an Euclidean metric after implementing a K-Nearest Neighbors method in order to self organize the conformed texture patterns. 312 images were analyzed. These images are completely experimental and were obtained from a microscope in TIFF format. The results are promising and were discussed with the aid of a specialist in cell research.

Keywords: Classifying cells, digital image processing, CBIR, K-NN, pattern recognition.

1. Introducción

Para nosotros como seres humanos, seres inteligentes y dinámicos, el sentido de la vista es el más común que usamos, prácticamente pasa desapercibido que de los cinco sentidos que tenemos éste sea actualmente el más explotado. Un ejemplo de lo refinado que es nuestro sentido de la vista es el “reconocimiento automático” que hacemos de imágenes, ya sea en color y/o forma, logrando la clasificación o identificación de objetos.

Un problema actual es la identificación de objetos, o puntos, en imágenes de tejidos celulares; estas imágenes son adquiridas y analizadas para poder diagnosticar enfermedades en pacientes acorde con un propósito médico, siendo así que actualmente, el diagnóstico para cada paciente requiere del análisis visual de muestras de tejidos por parte de un médico especializado, o personal debidamente capacitado. Siendo ésta una tarea repetitiva y dependiente de la experiencia y propósitos particulares del médico, es que no siempre es posible garantizar un correcto análisis visual de dichas imágenes.

La *glía* es un grupo heterogéneo de células neurales que cumplen funciones diversas en la fisiología del cerebro. Con base en su morfología, fisiología y localización en el cerebro podemos identificar los siguientes tipos principales de glía: microglía, oligodendroglía y astrogía [8]. En particular, el estudio de las células gliales es muy importante porque puede otorgar información sobre si existe o no una enfermedad neurodegenerativa en el paciente.

Para esto, se cuenta con dos tipos de imágenes de campo claro y de fluorescencia, que aunque a primera vista son muy diferentes por los colores, la información que encontramos dentro de ellas es la misma.

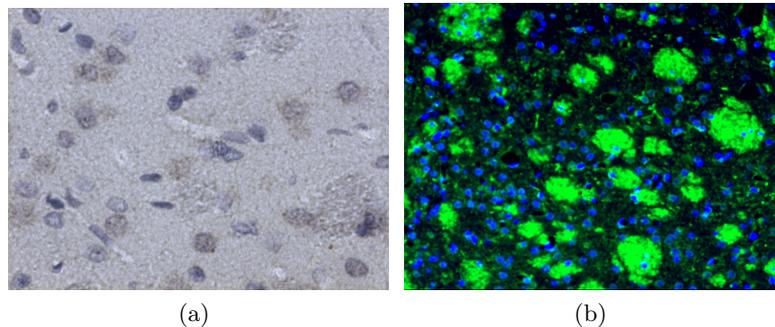


Fig. 1. Ejemplo de imágenes de tejido cerebral a) Campo Claro y b) Fluorescencia.

Hasta ahora lo que se realiza es que la persona especialista observa tejidos a través de un microscopio ver Figura 1, ve su forma y hace un conteo de los objetos que se encuentran dentro de la imagen. Esta tarea suele llevar mucho tiempo por cada imagen y también depende de la experiencia del especialista y de la capacidad visual de la persona.

En este trabajo mostramos el desarrollo e implementación de un sistema de identificación y localización de células gliales, aplicado en una base de datos de un grupo de imágenes digitales de tejido cerebral con muestras experimentales del Instituto Nacional de Neurología en México¹. Presentamos un avance basado en el estudio de textura de la imagen, con la finalidad de caracterizar el patrón de la textura de superficie a partir de rasgos estadísticos, basándonos en rasgos ya probados en otros estudios de reconocimiento y clasificación de imágenes por medio de la técnica CBIR (Content Based Image Retrieval), a saber: la media, la varianza y la homogeneidad en el espacio de color HSI.

La metodología de “reconocimiento de imágenes basada en su contenido propio”, o Content Based Image Retrieval (CBIR), se ha utilizado con éxito para el reconocimiento de imágenes partiendo del contenido de las imágenes mismas, sin tener que apoyarse en el uso de datos externos, e.g. etiquetas de texto que indiquen qué contiene la imagen, tan sólo estudiando y analizando la imagen propia desde diferentes enfoques, como lo es por medio de la textura [1]. Debido a las características de las imágenes del tejido cerebral, consideramos prudente su utilización en este trabajo.

En nuestra propuesta presentamos un alcance de clasificación de regiones de imágenes de tejido cerebral a partir de un análisis local de la imagen.

Para esto se hace una extracción de atributos de textura en un espacio de color HSI buscando tener rasgos independientes del color, para tener un patrón robusto que pueda caracterizar una región específica de la imagen. El procedimiento requiere que previamente el especialista de la salud indique la región

¹ Página oficial: <https://www.gob.mx/salud/innn>, Fecha de consulta: 19 de mayo, 2020.

que sea de su interés localizar dentro de la imagen. Posteriormente hacemos un entrenamiento de ese tipo de regiones a buscar dentro de la imagen, a partir de un algoritmo K-Nearest Neighbors de auto organización de atributos, para poder hacer comparaciones posteriores y así poder determinar el tipo de regiones más parecidas dentro de la misma.

El artículo se organiza de la siguiente forma: en la Sección 2 presentamos el estado del arte de nuestro problema de estudio, en la Sección 3 la metodología se presenta con detalle. La sección 4 muestra la implementación y resultados obtenidos, así como una discusión de los mismos. Finalmente, en la Sección 5 compartimos nuestras conclusiones.

2. Estado del arte

En la última década, con el aumento de la capacidad de almacenamiento en la red o en la nube y la masificación de las redes sociales, la disponibilidad por Internet de un gran volumen de imágenes digitales con diversas características es posible; y el hecho de que tales recursos no estén clasificados ha propiciado el desarrollo de herramientas que traten de paliar este problema. Hablando de cantidades reportadas en el año 2017, se estimó en un billón la cantidad de imágenes digitales acorde a una investigación reportada por el servicio MYLIO², tal que el 85 % de ellas fueron tomadas por un teléfono móvil. Ante esta realidad se tiene la posibilidad de crear nuevos repositorios de bases de datos de imágenes, pero el reto es clasificarlas acorde a su contenido, de forma no-supervisada, y con una alta eficiencia de forma simple y rápida, todo ello orientado al trabajo fundamental de organizar, proteger y compartir.

En particular, en el campo de la ciencia los grandes avances en la tecnología digital, así como en el campo de la microscopía de luz en los últimos años, determinaron que la imagen celular digital había adquirido una importancia creciente en la biología celular [7]. En el diagnóstico médico, el recuento de células sanguíneas juega un papel muy importante, ya que su aumento o disminución hace que se produzcan muchas enfermedades. Existen diferentes técnicas de recuento de células sanguíneas, el método convencional de conteo manual bajo microscopio requiere mucho tiempo y produce resultados inexactos [3], puede tener algunas limitaciones inherentes de los observadores humanos [7].

Por estas limitaciones existen ya nuevos métodos de adquisición y más sofisticados, como los empleados en la detección de alto contenido, generalmente producen una gran cantidad de datos que exige el poder de las computadoras para analizarlos [7], si bien existen soluciones de hardware como el Contador de hematología automatizado, los países en desarrollo no son capaces de disponer de máquinas costosas e inasequibles en todos los laboratorios de los hospitales [3].

Por lo cual al no tener el hardware necesario, podemos hacer uso de el análisis de imágenes por computadora en microscopía celular ya que puede abordar diversas tareas como clasificación y conteo celular, estudios sobre estructuras

² Página oficial: <http://tinyurl.com/y2jo5wal>, Fecha de consulta: 19 de mayo, 2020.

subcelulares o estudios sobre células vivas, entre otras, sin necesidad de grandes máquinas costosas. También se tiene utilidad en campos como patología, biotecnología vegetal, toxicología, desarrollo de fármacos y otros [7].

Para proporcionar una alternativa rentable basada en software y una alternativa eficiente para reconocer y analizar las células del tipo sanguínea [7], es necesario aportar una metodología para hacer un conteo automático de células basado en el procesamiento digital de imágenes. Diferentes etapas deben ser consideradas, desde la adquisición de la imagen, preprocesarla, mejorarla, segmentarla, la extracción de características, aplicación de operaciones morfológicas y la detección de anomalías, se pueden considerar diferentes problemáticas, como la restauración de la imagen, el filtrado de ruido y la reducción del desenfoque. La segmentación de imágenes es también un procesamiento importante aplicado mediante diferentes algoritmos para incluir la separación de los agregados celulares [7].

Para este propósito, se han implementado pocas técnicas de preprocesamiento y posprocesamiento en la imagen de las células sanguíneas para proporcionar una imagen mucho más clara y limpia [3], es aquí donde existe un campo amplio de investigación. Una de las técnicas que ha dado resultados exitosos es la técnica CBIR la cual ha sido probada en identificación y clasificación de imágenes de interés [1], tales como los escenarios naturales y la clasificación de rostros humanos.

Muchos sistemas y herramientas CBIR se han desarrollado para la consulta de imágenes [2], tal como Query By Image Content (QBIC) desarrollado en los años 90s por IBM [4], que permite al usuario consultas dentro de una gran base de datos con base a atributos de la imagen.

CBIR es un sistema que tiene dos etapas principales: la inserción de datos y el procesamiento de consultas. La inserción de imágenes, es un método no supervisado para la extracción de información (tal como color, forma o texturas), de la base de datos para ser guardados y autoagrupados por medio de un algoritmo para su uso futuro. En el procesamiento de consultas, mediante una imagen de muestra, se obtienen sus características, para ser comparadas con métricas específicas hacia los racimos de datos auto agrupados en la etapa previa, y así obtener el valor dentro de la clasificación previamente hecha de la base de datos [6].

Anteriormente la clasificación de las imágenes se llevaba a cabo de forma manual y por medio de metadatos insertados en las imágenes tales como etiquetas de texto, pero, en una base de datos muy grande conlleva un costo enorme para su clasificación [12]. Con el uso de CBIR la identificación y clasificación de células sanguíneas se puede automatizar como se ha aplicado a otro tipo de imágenes (rostros y escenarios naturales), ya que se realiza por medio de características de la misma imagen sin necesidad de metadatos, etiquetas o intervención externa, esta propuesta es presentada en el presente trabajo.

Por definición, las enfermedades neurodegenerativas (Parkinson, Alzheimer, Huntington, Esclerosis lateral amiotrófica y tumores cerebrales entre las principales) comprenden tanto la pérdida progresiva e irreversible de neuronas; así

como la proliferación descontrolada de algunos otros tipos celulares en el caso de cáncer en el Sistema Nervioso. Todo esto conlleva a una pérdida del control óptimo de aquellas funciones reguladas a nivel nervioso. La determinación microscópica de muestras de pacientes con padecimientos neurodegenerativos sigue siendo considerada una herramienta muy útil para el diagnóstico y evolución de estas enfermedades. Bajo técnicas diversas de histoquímica e inmunofluorescencia contra marcadores específicos es posible obtener un estimado en biopsias de estos pacientes sobre la disminución en el número de neuronas o la proliferación descontrolada de oligodendrocitos o astrocitos en casos de tumores cerebrales [9–11].

3. Metodología

Para el problema de la identificación y localización de las células gliales en imágenes digitales se propone el modelo mostrado en la Figura 2, en el que se plantean las siguientes etapas:

Etapas 1

- *Adquisición* de la base de datos de imágenes de tejido cerebral, todas ellas se encuentran en el espacio de color RGB.
- *Preprocesamiento*:
 1. Se aplica a cada una de las imágenes un filtro gaussiano.
 2. Ecuilibración de imagen.
 3. Exportación de las imágenes del espacio de color RGB original, al espacio HSI, con el objetivo de trabajar con más información independiente del color de la imagen y buscando tener información de las texturas de las imágenes.

Etapas 2

- El usuario selecciona una sola imagen (de referencia) de la pila de imágenes previamente procesadas en la etapa 1, y en dicha imagen selecciona una región de interés (ROI) con el cursor.
- A continuación se procede a obtener los atributos, el sistema obtiene los atributos CBIR de la ROI, los cuales los disponemos en un vector de la forma: $V_R = [\mu_H, \sigma_H, H_H, \mu_S, \sigma_S, H_S, \mu_I, \sigma_I, H_I]_{9 \times 1}$, de tamaño 9×1 , ya que se tienen tres capas de la imagen y por cada capa se extraen los atributos: media, desviación estándar y homogeneidad.

Etapas 3

- En la imagen de referencia que el usuario seleccionó anteriormente se genera un emparrillado con dos casos: regular de 100 puntos, o aleatorio de 300 puntos distribuidos uniformemente.
- En cada punto generado se abre una ventana de tamaño 10×10 píxeles.

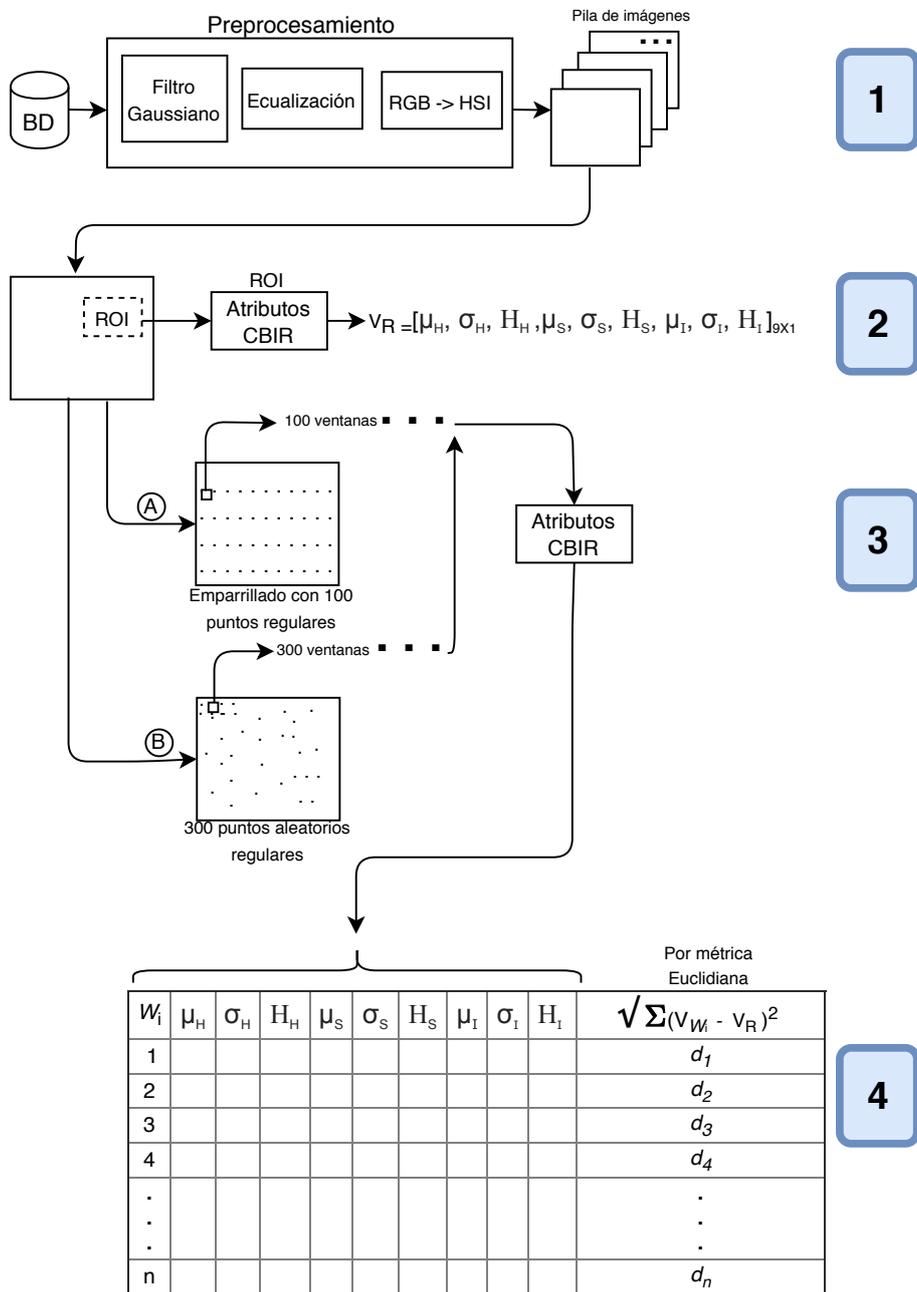


Fig. 2. Metodología.

- A cada una de las (100 o 300) ventanas se extraen sus atributos CBIR teniendo como resultado una tabla de tamaño (9×100 , o 9×300), nueve de los atributos CBIR y (100 o 300) por cada ventana según el caso.

Etapa 4

- Aplicación de un clasificador de distancia K-NN, con $K = 3, 5, 11$ o 31 [5]. A cada vector de la tabla obtenida se le aplica la métrica euclidiana contra el vector V_R y se guardan esas distancias en dicha tabla para después aplicar un ordenamiento ascendente y tomar las 3, 5, 11 o 31 (según considere el especialista) regiones más parecidas a la ROI. Buscamos un número de k-vecinos de cantidad impar, para lograr discriminar mejor acorde a la literatura reportada, y dada la naturaleza de la imagen local analizada, que puede replicarse en muchas otras subáreas de la imagen de estudio, es que se llega a considerar 11 u 31 vecinos cercanos.

4. Experimentos

Se uso una computadora tipo PC con:

- Procesador: Intel(R) Core(TM) i7-8750H CPU @ 2.20GHz (12 CPUs), 2.2GHz.
- Sistema Operativo: MS Windows 10, 64 bits.
- Memoria RAM: 8 GB.
- Lenguaje: Python 3.8.1.
- Librería: Open CV 4.1.2.

Base de datos cuenta con un total de 312 imágenes, 190 de campo claro y 122 de fluorescencia, en formato TIF:

Tipo	Resolución
Campo Claro	1392×1040 pixeles
Fluorescencia	355×402 a 1280×960 pixeles

Las primeras pruebas que se hicieron son probando con dos formas de emparrillado la primera es un emparrillado regular de 100 puntos con subventanas de tamaño 10×10 pixeles y las segundas pruebas fueron con un emparrillado de 300 puntos aleatorios con subventanas de tamaño 10×10 y de 30×30 pixeles.

Se realizaron diferentes pruebas con las imágenes de campo claro y de fluorescencia, lo que se quiere realizar es tener una imagen de células gliales y ya sea seleccionando una ROI de manera manual o una de manera aleatoria buscar ese objeto de interés en la misma imagen y mostrar esos objetos parecidos.

Por lo cual se tienen dos tipos de pruebas una es de manera supervisada y otra de forma no supervisada. En la primera forma supervisada lo que se hace es que el médico especialista selecciona una ROI de forma manual y él da su punto de vista al decir si los objetos localizados son semejantes o no. Y la segunda forma de manera no supervisada, la ROI se selecciona de manera aleatoria en la imagen y se muestran los más parecidos.

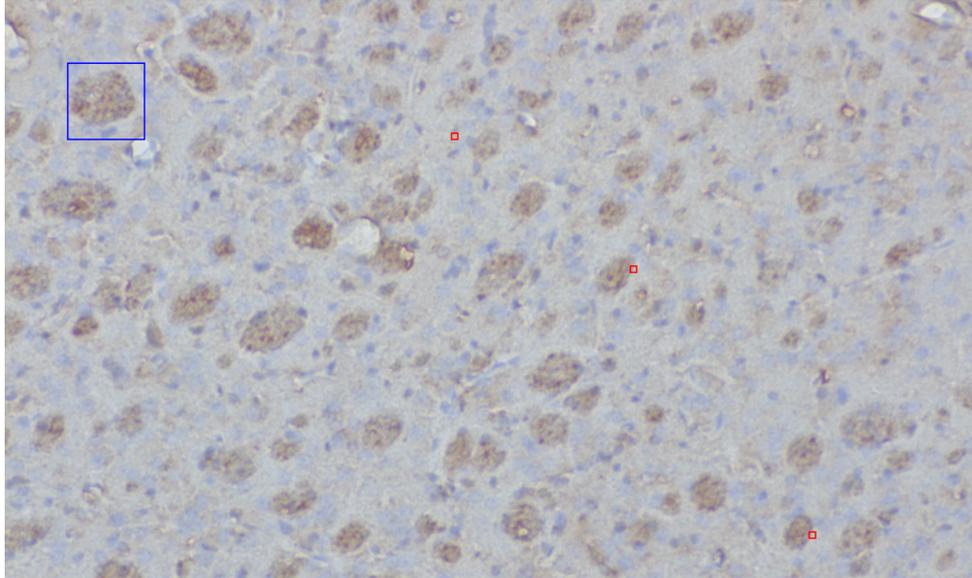


Fig. 3. Se muestra una imagen celular con la ROI seleccionada por el usuario y las tres ROI de tamaño 10x10 más parecidas, considerando un emparrillado regular de la imagen.

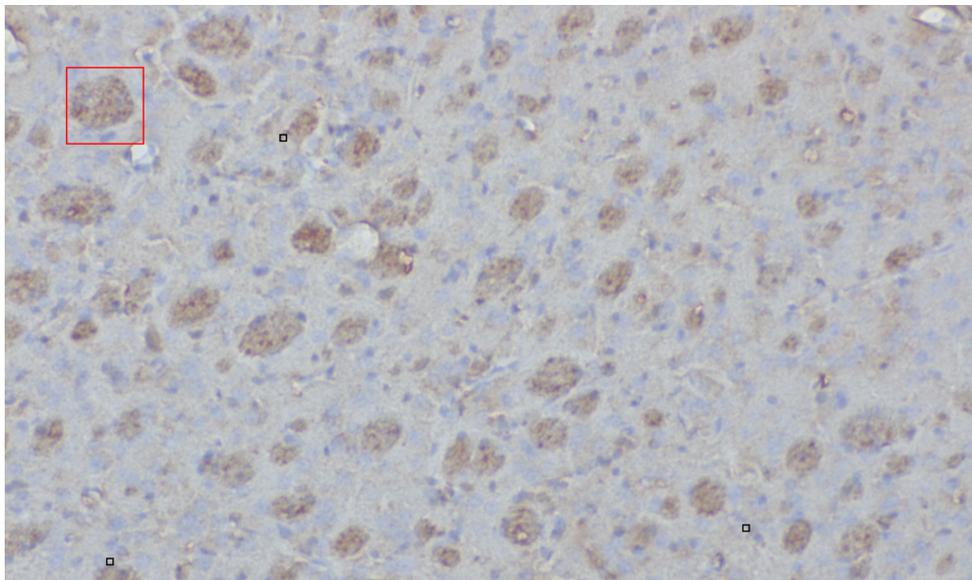


Fig. 4. Se muestra una imagen celular con la ROI seleccionada por el usuario y las tres ROI de tamaño 10×10 más parecidas, considerando un emparrillado aleatorio de la imagen.

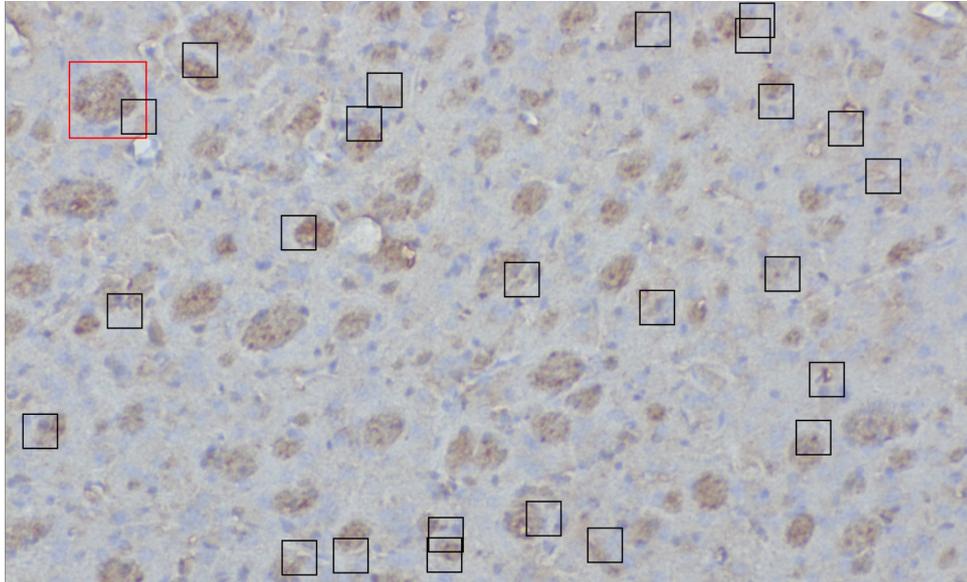


Fig. 5. Se muestra una imagen celular con la ROI seleccionada por el usuario y las tres ROI de tamaño 30×30 más parecidas, considerando un emparrillado aleatorio de la imagen.

4.1. Análisis de resultados

En los primeros experimentos que se realizaron se eligió un emparrillado de 100 puntos regulares en la imagen, se selecciona la ROI y a partir de eso se buscaban las 3 ROI que más se parecían a los atributos CBIR de la selección, como se muestra en la Figura 3.

Las ROI más parecidas se fijaron como ventanas de tamaño 10×10 , pero al observar los resultados estas ventanas eran muy pequeñas para el tamaño de la imagen que aproximadamente es de 1040×1392 .

El segundo experimento fue realizado con un emparrillado aleatorio de 300 puntos y de igual forma se obtuvieron las tres ROI que más se parecían con respecto a las características CBIR obtenidas de la región seleccionada, pero de igual forma las ventanas de tamaño 10×10 eran muy pequeñas ver Figura 4.

Por último, se ajustó el tamaño de las ventanas a un espacio de 30×30 píxeles y así fue más fácil de visualizar los resultados, como se muestra en la Figura 5.

Por el momento se está trabajando en tener una aplicación amigable para el profesional de la salud, y también estamos recabando la mayor cantidad de imágenes probadas en el laboratorio del Dr. Edgar, para validar nuestros resultados con su experiencia.

4.2. Criterio de pruebas semi-supervisadas automáticas

Con la finalidad de poder tener un criterio para analizar la posible efectividad de nuestra metodología, se partió de marcar o generar una ROI aleatoria sobre cada una de las imágenes de prueba, de tal forma que se tiene la suerte de que caiga o no encerrando o delimitando una región verdadera de interés, i.e. una célula tipo glía, tal como se observa en la Figura 6 la ROI seleccionada ya que es aleatoria selecciona más de un objeto y los objetos que se muestran son similares, pero en la Figura 7 la ROI seleccionada contiene muchos objetos, esto debido a que la imagen tiene más objetos o fue realizada con un acercamiento más detallado, por lo cual no podemos decir si los objetos que encuentra son o no son parecidos.

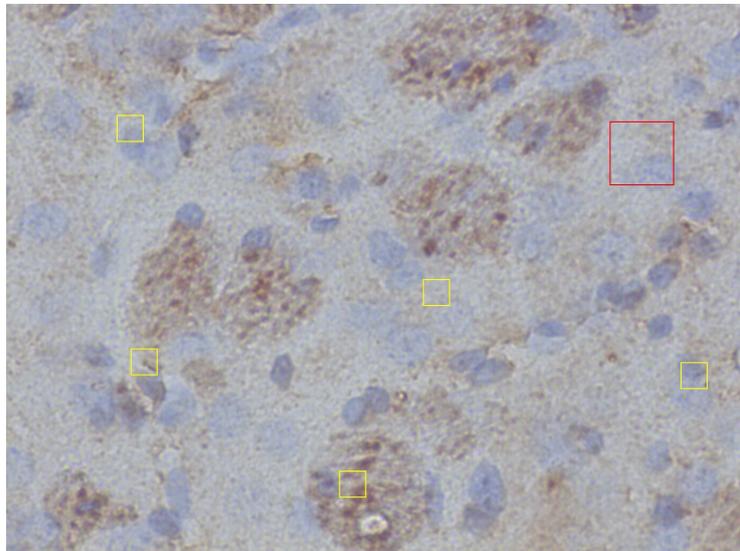


Fig. 6. Se muestra una imagen celular con la ROI seleccionada aleatoriamente y las cinco ROI de tamaño 50×50 más parecidas, considerando un empaquetado aleatorio de la imagen.

4.3. Discusión de resultados

El tamaño de las subventanas depende mucho del tamaño de las imágenes, ya que el tamaño que manejamos de las mismas es muy grande, una ventana de tamaño 10×10 píxeles es muy pequeña para visualizar los resultados. El tamaño de las ventanas para señalar los objetos parecidos al objeto señalado por el usuario en la ROI debería ser de un tamaño aproximado al tamaño del objeto de interés. Se puede ver en las muestras presentadas que por el momento

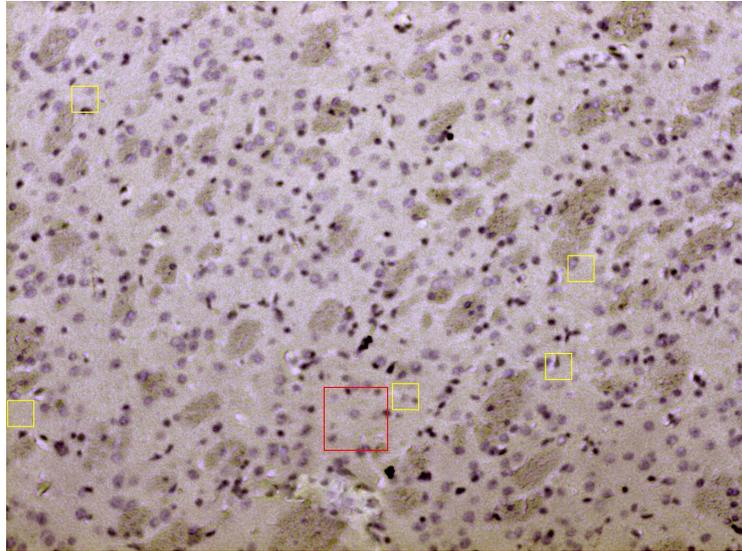


Fig. 7. Se muestra una imagen celular con la ROI seleccionada aleatoriamente y las cinco ROI de tamaño 50×50 más parecidas, considerando un emparrillado aleatorio de la imagen.

la ROI aleatoria para la batería de pruebas no precisamente cae encerrando una célula tipo glia, pero los resultados de las más parecidas aparentemente son satisfactorias para el profesional de la salud según su criterio de supervisión de las pruebas y su resultado.

Si el tamaño de la ROI después de aplicarle el preprocesamiento es de (n, m) se tomará el máximo de ellos por lo tanto el tamaño de las ventanas será de $(\max(n, m) \times \max(n, m))$, esto para que se busquen objetos de las mismas dimensiones.

5. Conclusiones

Hemos presentado una metodología para estudiar las células tipo glia en tejido cerebral a partir de sus atributos de textura, dentro de una imagen obtenida experimentalmente en microscopio, de tal forma que este enfoque no ha sido reportado en la literatura y este es nuestro principal aporte.

A partir de estos primeros resultados vemos que sí es posible poder brindar un apoyo para la localización de este tipo de células dentro de las imágenes sin importar si es de tipo campo claro o fluorescencia, ya que no impacta el color y sus variaciones dado que hacemos uso de la textura y no del color como atributos para conformar el patrón de caracterización de nuestro objeto de estudio. Se observa que con los tres atributos CBIR el algoritmo es capaz de encontrar objetos parecidos al seleccionado por el usuario.

Como trabajo futuro se espera que en la ROI seleccionada por el usuario en ella se haga un preprocesamiento para solo encerrar el contorno deseado con un algoritmo de detección automática y así no se busquen regiones que no sean de nuestro interés, para esto se pretende aplicar un procesamiento morfológico y en base a esto se pretende tener una nueva capa de procesamiento automático de la imagen.

Referencias

1. Benavides, C., Villegas, J., Roman, G., Aviles, C.: Face classification by local texture analysis through cbir and surf points. *IEEE Latin America Transactions* 14(5), 2418–2424 (2016)
2. Benavides Alvarez, C.: Sistema no-supervisado de clasificación de rostros con técnicas basadas en CBIR. Universidad Autónoma Metropolitana, Unidad Iztapalapa (2015)
3. Bhamare, M.G., Patil, D.S.: Automatic Blood Cell Analysis By Using Digital Image Processing: A Preliminary Study. *International Journal of Engineering Research & Technology (IJERT)* 2(9), 3137–3141 (2013)
4. Flickner, M., Sawhney, H., Niblack, W., Ashley, J., Qian Huang, D., Gorkani, M., Hafner, J., Lee, D., Petkovic, D., Steele, D., Yanker, P.: Query by image and video content: the QBIC system. *Computer* 28(9), 23–32 (1995)
5. Fukunaga, K.: *Introduction to statistical pattern recognition* (2nd ed.). Academic Press Professional, Inc., San Diego, CA, USA (1990)
6. Gonzalez, R.C., Woods, R.E.: *Digital image processing*. Prentice Hall (2002)
7. Morales, R.O., Valentín, J., Ginori, L.: *Cell Microscopy Imaging: A Review on Digital Image Processing Applications* (2015)
8. Reyes H., D., Bulavina, L., Pivneva, T.: La glia el pegamento de las ideas. *Ciencia* pp. 12–18 (2014)
9. Saxena, S., Caroni, P.: Selective neuronal vulnerability in neurodegenerative diseases: From stressor thresholds to degeneration. *Neuron* (2011)
10. Subramaniam, S.: Selective neuronal death in neurodegenerative diseases: The ongoing mystery. *Yale J. Biol. Med.* (2019)
11. Verma, D., Gupta, S., Biswas, J., Joshi, N., Singh, A., Gupta, P., Tiwari, S., Sivarama, R.K., Chaturvedi, S., Wahajuddin, M.: New therapeutic activity of metabolic enhancer piracetam in treatment of neuro-degenerative disease: Participation of caspase independent death factors, oxidative stress, inflammatory responses and apoptosis. *Biochim Biophys Acta Mol Basis Dis* (2018)
12. Vogel, J., Schiele, B.: Semantic modeling of natural scenes for content-based image retrieval. *International Journal of Computer Vision* 72, 133–157 (2007)