

## Identificación de casos potenciales de edema macular diabético en imágenes de fondo de ojo utilizando técnicas de segmentación dinámica y clasificadores basados en redes neuronales.

I. G. Rodríguez-Leon<sup>1</sup>, R. Díaz-Domínguez<sup>1</sup>, E. Zañudo-Angulo<sup>1</sup>, G. Sanchez-Sosa<sup>1</sup>,  
R. Fuentes-Aguilar<sup>1</sup>, A. Garcia-Gonzalez<sup>1</sup>, A. Santos-Garcia<sup>1,2</sup>

<sup>1</sup>Departamento de Bioingenierías, Tecnológico de Monterrey, Campus Guadalajara.

<sup>2</sup>Centro de Retina Médica y Quirúrgica, Guadalajara, Jalisco.

**Resumen**— El edema macular diabético (EMD) es uno de los principales padecimientos derivados de la diabetes. Su identificación requiere de la atención brindada por médicos especialistas en retina, mismos que a nivel mundial, no son suficientes para atender a la población propensa a esta enfermedad. El siguiente trabajo presenta un sistema para la identificación de casos potenciales de EMD a través de imágenes obtenidas con un oftalmoscopio comercial. El trabajo consta de tres partes: el pre-procesamiento de las imágenes para la corrección de iluminación, ruido y movimiento generados por la cámara y el usuario. El cálculo de una transformación de cuerpo rígido para realizar un análisis de correlación entre imágenes obtenidas por el oftalmoscopio y las imágenes de un estudio de tomografía de coherencia óptica (OCT) y la implementación de una red neuronal artificial para aproximar el grosor macular usando las imágenes de fondo de ojo, este dato permite la clasificación del potencial daño en la retina.

Para validar el sistema se generó una base de datos de 120 imágenes de fondo de ojo en pacientes sanos de 19 a 24 años. El error promedio en la aproximación generada por la red y el valor real dado por la OCT fue menor al 8%.

**Palabras clave**—edema macular diabético, imagen de fondo de ojo, morfología matemática, red neuronal, transformación de cuerpo rígido

### I. INTRODUCCIÓN

De acuerdo a la Organización Mundial de la Salud (OMS), en el mundo existen alrededor de 347 millones de personas con Diabetes Mellitus (DM); el 97 % de las personas diagnosticadas con DM tipo I y el 77.8% con DM tipo II presentan algún tipo de complicación provocada por esta enfermedad [1].

La retinopatía diabética (RD) es la complicación más común, de la vista, derivada de la diabetes; genera una afectación directa en los vasos sanguíneos de la retina, misma que puede causar ceguera parcial o total. Después del diagnóstico, el 27%, 71% y el 95% de los pacientes diabéticos presentan el primer signo de RD, el edema macular diabético (EMD) en los siguientes, 5, 10 y 20 años respectivamente [2].

El EMD produce daño en la mácula, la porción central de la retina, haciendo que se genere una fuga en el fluido sanguíneo que circula en los vasos de la misma [3].

De acuerdo al Centro Nacional para la Excelencia Tecnológica en Salud (CENETEC) en su Guía para la Práctica Clínica para el diagnóstico de retinopatías [4], se recomienda que éste sea realizado por un sub-especialista en retina, ya que su identificación no es evidente sin

entrenamiento/experiencia, además de requerir un estudio de tomografía de coherencia óptica (OCT), equipo que permite medir las variaciones del grosor de la retina.

En México, existen registrados alrededor de 300 médicos especialistas en retina [5]; lo que implica una falta de especialista para atender a la población latente de presentar RD y particularmente EMD. Es por ello que acercar este tipo de estudios de diagnóstico a un segmento más grande de la población médica permitiría identificar de forma temprana los potenciales casos del padecimiento. En este sentido existen alrededor de 80 mil médicos familiares, internistas y generales en México, a los que se les deben de otorgar las herramientas necesarias para realizar un diagnóstico precoz del padecimiento con un mínimo de entrenamiento.

Se han generado algunas alternativas que buscan acercar el diagnóstico del EMD a través de plataformas basadas en procesamiento de imagen, técnicas de inteligencias artificial y sistemas de telemedicina; como ejemplo tenemos el PeekVision [13], una aplicación desarrollada por la Universidad de Higiene y Medicina Tropical de Londres que permite la visualización de una imagen de fondo de ojo a bajo costo adaptando un teléfono inteligente a un lente para su envío a una central de diagnóstico remoto, ésta al igual que otras aplicaciones requiere del análisis de especialistas en retina.

Considerando esta situación, la investigación en procesamiento de imágenes de fondo de ojo para el diagnóstico ha centrado sus esfuerzos en tres áreas particulares: cuantificación de la calidad de la imagen, localización y segmentación de estructuras retínicas y la identificación de anomalías; existen alrededor de 700 artículos relacionados con estos temas [6], en todos ellos el diagnóstico se da cuando las lesiones son evidentes.

El siguiente trabajo presenta el desarrollo de un sistema para la identificación de casos potenciales de edema macular diabético por el médico de primer contacto utilizando técnicas de procesamiento y segmentación de imágenes, extracción de características y utilizando un clasificador basado en redes neuronales para la aproximación del grosor macular en imágenes bidimensionales obtenidas por un oftalmoscopio.

### II. METODOLOGÍA

La primera parte de la metodología busca identificar de manera semiautomática dos pares de intersecciones vasculares en las imágenes obtenidas por el oftalmoscopio y

la OCT; esto con la finalidad de obtener referencias espaciales que permitan realizar una transformación de cuerpo rígido (TCR) entre ellas. La TCR ajusta en tamaño y ángulo la imagen de oftalmoscopio para dividirla en las nueve áreas, tipo mallado, de estimación de grosores dado por la OCT (**Fig. 1-B**). Sobre cada una de las nueve áreas definidas, se busca extraer características propias de la imagen y generar una red neuronal artificial para la estimación de los grosores de la región macular.

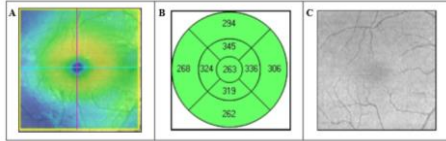


Fig. 1: (A) ROI definida por la OCT, (B) distribución de las 9 áreas para cálculo de grosor en función de la ROI [12] y (C) escalamiento de la imagen de oftalmoscopio

### 1) Base de datos

En el desarrollo de este trabajo participaron 20 sujetos de estudio entre 19 y 24 años de edad (15 hombres y 5 mujeres), clasificados como sanos por médicos especialistas en retina del Centro de Retina Médica y Quirúrgica (CRMQ, Jalisco, México), todos ellos firmaron los permisos correspondientes posteriores a la autorización del comité de bioética.

A cada sujeto se le realizaron tres estudios de fondo de ojo: tomografía de coherencia óptica, imagen de fondo de ojo con retinógrafo y una con oftalmoscopio **Fig. 2**

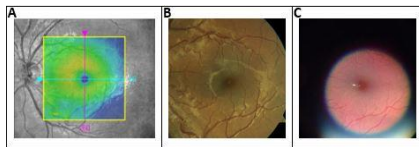


Fig. 2. Imágenes de fondo de ojo obtenidas con: (A) OCT, (B) retinógrafo y (C) oftalmoscopio

Con ello se generó una base de datos de 120 imágenes de fondo de ojo:

- 40 capturadas con un oftalmoscopio marca PanOptics® adaptado con un iPhone 4S®. Las imágenes obtenidas tiene una resolución de 2248x3264 pixeles en formato JPG. (**Fig. 2-C**)
- 40 capturadas de una cámara de fondo de ojo Zeiss FF450 plus con Visupac®. Las imágenes obtenidas tienen una resolución de 2300x2000 pixeles en formato TIFF (para el correcto diagnóstico por parte del especialista) (**Fig. 2-B**)
- 40 imágenes capturadas con una OCT Zeiss Cirrus HD-5000®. (**Fig. 2-A**)

Con la base generada se procedió con las técnicas para la identificación intersecciones vasculares de referencia con la siguiente metodología:

- Pre-procesamiento para el ajuste en la variabilidad en color y ruido en las imágenes obtenidas
- Eliminación dinámica de ruido e identificación de intersecciones vasculares

- Transformación de cuerpo rígido entre las imágenes de los diferentes sistemas.
- Extracción de características
- Clasificador basado en redes neuronales

El procesamiento digital se realizó usando Matlab® 2013B en una computadora portátil de doble núcleo con 4 GB de RAM.

### 2) Pre-procesamiento

Considerando la imagen capturada con el oftalmoscopio como una función bidimensional,  $I_Q(x, y)$  sobre la cual se define manualmente una región de interés ROI- $(I_{Q-ROI}(x, y))$  – que contiene dos intersecciones vasculares (**Fig 3-A**) en ella, éstas servirán de puntos de referencia para comparar las imágenes.

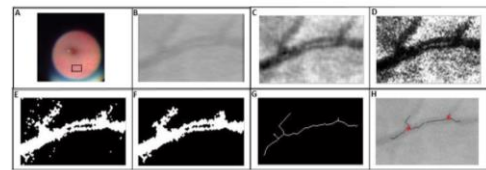


Fig. 3. Pre-procesamiento: (A) Selección de la ROI manual, (B) aplicación del filtro "motion", (C) magnitud del gradiente, (D) Ecuilización, (E) Binarización, (F) Eliminación del ruido, (G) Esqueleto vascular e (H) Identificación de intersecciones

La primera modificación realizada fue la aplicación de un filtro [9] de movimiento (**Fig. 3-B**) para aproximar de manera lineal el movimiento generado por la cámara, esto es:  $I_{Q-M}(x, y) = T(I_{Q-ROI}(x, y))$ ; donde  $I_{Q-M}(x, y)$  es la imagen resultado de aplicar un operador T de vecindad 8, en cada punto  $(x, y)$  de la imagen de entrada  $I_{Q-ROI}(x, y)$ .

Posterior al filtrado, se procede a obtener la magnitud del gradiente en  $I_{Q-M}(x, y)$ , tanto vertical como horizontal (**Fig. 3-C**), utilizando las siguientes máscaras:

$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \dots (1) \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \dots (2)$$

Y calculando la magnitud del gradiente:

$G_{I_{Q-M}} = \sqrt{G_x(I_{Q-M})^2 + G_y(I_{Q-M})^2}$ , donde  $G_y(I_{Q-M})$  y  $G_x(I_{Q-M})$  son las imágenes procesadas aplicando (1) y (2) respectivamente a  $I_{Q-M}$  y  $G_{I_{Q-M}}$  es la imagen resultante. Con la magnitud calculada, se identifican las variaciones locales de iluminación, así mismo permite identificar la dirección de máxima variación, en  $I_{Q-M}(x, y)$ , de las diferencias entre los niveles de intensidad de grises en la imagen [7]. Enseguida se procede a ecualizar la imagen tal que:  $I_{E-Q} = E(G_{I_{Q-M}})$ , donde E es definido como el operador de ecualización utilizado en [8] para corregir la distribución no-uniforme en la iluminación (**Fig. 3-D**) y con la que finalmente se caracteriza la iluminación de los vasos sanguíneos (intensidad en escala de grises entre 30-50) para usarlo como referencia en la binarización por umbral utilizada (**Fig. 3-E**). La imagen pre-procesada es definida como:  $I_{Q-BN}$ .

### 3) Eliminación dinámica de ruido

Con el pre-procesamiento aplicado, se procede a la eliminación de ruido en la imagen para aislar dos intersecciones vasculares en las imágenes de OCT y de oftalmoscopio. Sobre la imagen  $I_{Q-BN}$  se realiza un conteo de objetos, definiendo un objeto como un conjunto de píxeles unidos a través de vecindad 4 u 8 a los cuales se les calculó el tamaño de su área (en píxeles). Con el cálculo de las áreas se procedió a identificar el ruido en las imágenes (**Fig. 3-E**). Analizando las imágenes, se estableció como ruido a todo objeto que tuviera un área menor al 10% del tamaño de la ROI, considerando que el tamaño de la ROI varía en función del usuario.

Los objetos con área mayor al 10% de la ROI, son sometidos a  $n=1, 2, 3, \dots$  operaciones de apertura y cierre, para su unificación, con ello se busca identificar una sola estructura vascular tal como se realiza en [9]; el algoritmo detiene las operaciones cuando se ha registrado un solo objeto en la imagen (**Fig. 3-F**).

Con el objeto identificado se procede a obtener el esqueleto del mismo, ello se logra aplicando operaciones de erosión sobre la imagen bajo dos condiciones: mantener el objeto íntegro (**Fig. 3-G**) y contar con cuatro puntos de terminación, es decir que se presenten dos ramificaciones. Finalmente se aplica una máscara, de 0 a  $360^\circ$ , en la imagen esqueletada para la identificación de intersecciones de la siguiente forma:

$$Y = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \dots (3)$$

Como resultado se obtienen las intersecciones vasculares, mismas que se restauran en la imagen original, conservando su posición y coordenadas (**Fig. 3-H**). El pre-procesamiento y la eliminación dinámica del ruido son aplicados a las imágenes del retinógrafo y la OCT

### 4) Transformación de cuerpo rígido

Una vez realizado el pre-procesamiento y la segmentación dinámica para la identificación de los puntos en las imágenes, tomamos las intersecciones  $P$  y  $Q$  en las imágenes de OCT y oftalmoscopio (**Fig. 4**) para realizar la TCR [10]

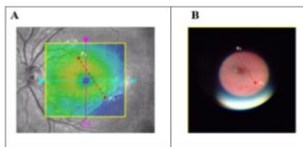


Fig. 4. Intersecciones vasculares de referencia en las imágenes de (A) OCT y (B) del oftalmoscopio, puntos  $p$  &  $q$  respectivamente.

Sea TCR el operador de transformación de cuerpo rígido de un punto  $Q$  a un punto  $P$ , entonces la ecuación:

$$Y = TCR(Q, P) \dots (5)$$

Significa que  $Y$  es la nueva localización del conjunto  $Q$  tal que **los vectores  $y_i$  son los las posiciones más cercanas que pueden alcanzar los puntos  $q_i$**  a través de la transformación de movimiento rígido, dada por una matriz de rotación  $R_{TCP} \in \mathcal{R}^{2 \times 2}$  y 2 vectores de traslación  $\gamma_i \in \mathcal{R}^{2 \times 1}$ , esto es:

$$y_i = R_{TCP} x_i + q_i \quad i = 1, 2 \quad \dots (6)$$

Con lo anterior se realizar un escalamiento y ajuste espacial de la imagen del oftalmoscopio (**Fig. 1-C**) a las dimensiones de la obtenida por la ROI de la OCT (**Fig. 1-A**). Ello se realizar debido a que el estudio de OCT ofrece un mallado (**Fig. 1-B**), dividido en nueve áreas, con los diferentes valores de grosor obtenidos para cada segmento macular. La imagen del oftalmoscopio se escala para generar estas mismas nueve áreas y obtener características en ellas.

### 5) Extracción de características

Una vez realizada la transformación de cuerpo rígido y la segmentación de las nueve áreas en la imagen de oftalmoscopio, se obtienen siete parámetros sobre cada una de las áreas dadas: homogeneidad, contraste, energía, entropía, grosor, promedio de intensidad de color y el promedio de intensidad de color aplicando una transformación de oponencia [8] definida como:

$$L = (L_r + L_g) - 2L_b \dots (7)$$

Donde  $L_r, L_g$  y  $L_b$  son los componentes de color de la imagen de oftalmoscopio en formato RGB. De ello se realiza análisis de correlación respecto del grosor dado en el mallado de la OCT.

### 6) Red Neuronal Artificial

Para realizar la aproximación a partir de la característica seleccionada, se usó un perceptrón multicapa del toolbox de Matlab®. Se utilizó una neurona para la capa de entrada, cinco neuronas en la capa oculta usando una función sigmoidea como función de activación. Para la capa de salida se colocó una neurona con una función de activación tangente-sigmoidea. De la base de datos se utilizó un 75% para el entrenamiento y 25% para las pruebas.

## III. RESULTADOS

Del ajuste espacial entre imágenes con el algoritmo para el escalamiento, TCR, se consume un tiempo menor a los 5 segundos en el empalme y el ajuste espacial de los dos pares de puntos de la imagen del oftalmoscopio y la imagen de la OCT. Respecto de la extracción de características de las nueve áreas de la OCT, se determinó que solo el área central aporta información relevante sobre el grosor macular esto debido a que, por las distintas dimensiones de las imágenes, se pierde información al realizar el escalamiento a la ROI

definida. En ella se realizó análisis de correlación entre las siete características y el grosor central dado, Tabla 1:

Tabla 1

Análisis de correlación entre los siete parámetros y el grosor dado por la OCT

	Oponencia	Promedio	Entropía	Energía	Contraste	Correlación Homogeneidad	Grosor
Oponencia	1	-0.15308	-0.124269	-0.088467	-0.112607	0.0295139	0.140741822
Prom edio	-0.15308	1	-0.192902	-0.828504	-0.512011	-0.233428	0.43620965
Entropía	-0.124269	-0.192902	1	0.3784555	-0.150124	0.219821	0.125974977
Energía	-0.088467	-0.828504	0.3784555	1	0.6007417	0.2437748	-0.547907126
Contraste	-0.112607	-0.512011	-0.150124	0.6007417	1	-0.382738	-0.956323436
Correlación	0.0295139	-0.233428	0.219821	0.2437748	-0.382738	1	0.379577594
omogeneidad	0.1407418	0.4362096	0.125975	-0.547907	-0.956323	0.3795776	1
Grosor	0.7662483	-0.366286	-0.094654	0.167737	0.0636651	0.077898	-0.027991489

El análisis de correlaciones nos permite plantear el uso de un estimador del grosor a partir de la información de la imagen transformada en oponencia, es decir, existe una relación importante, de 0.7662, entre la intensidad de color en la imagen y el grosor dado por la OCT. De esta forma, utilizando un 25% de los datos de la base, se puede aproximar el grosor con un algoritmo de redes neuronales artificiales. Los resultados obtenidos se encuentran en la tabla 2:

Tabla 2  
Comparativo de error obtenido con el clasificador

Aproximación	Real	Error Absoluto	Rango normal 220.5-294.8
243.5374	264	7.751	Sano
243.5949	263	7.3784	Sano
256.4531	252	1.7671	Sano
253.7578	248	2.3217	Sano
253.1218	230	10.053	Sano
259.7102	228	13.908	Sano
253.3323	229	10.6255	Sano
274.5548	237	15.8459	Sano
244.4783	239	2.2922	Sano
273.3455	243	12.4878	Sano

En las 10 muestras mostradas (tabla 2) se observa un error máximo de 15.8459 y un error mínimo de 1.7671 entre el valor real del grosor de la mácula y el valor aproximado por el algoritmo. Es importante mencionar que para verificar si la aproximación del grosor de la RNA a través de la transformación de oponencia era afectada por la calidad de la imagen, la metodología descrita fue aplicada a la imagen en alta definición obtenida por el retinógrafo.

#### IV. DISCUSIÓN

El pre-procesamiento de la imagen permite corregir: la distribución de la iluminación de la imagen, los artefactos de movimiento dados por el usuario, así como el ruido generado por la calidad de la cámara del oftalmoscopio.

Respecto de la TCR, esta tiene como fundamento el algoritmo de iteración del punto más cercano [14] (ICP por sus siglas en inglés), esta modificación permite disminuir los tiempos, a menos de 5 segundos, para empalmar los dos pares de coordenadas de las intersecciones de las referencias vasculares de las dos imágenes: OCT y oftalmoscopio.

Finalmente es debe considerar que la calidad de la imagen obtenida con el oftalmoscopio no afecta a la RNA de manera considerable ya que, como se mencionó, la metodología se aplicó utilizando la retinografía; el análisis

de correlación de la transformación de oponencia dio como resultado un coeficiente de 0.7667, es decir, solo 0.0005 mayor que el obtenido con la imagen del oftalmoscopio.

#### V. CONCLUSIÓN

Para el proceso de validación, se generó una base de datos de 120 personas. Únicamente el 25% de los datos fue empleado para aproximar el grosor. El error promedio entre el grosor real de la mácula dado por la OCT y la aproximación de grosor generada por la red neuronal artificial es del 8.44%. Sin embargo, en todos los casos, el valor obtenido es considerado aceptable ya que se encuentra en el rango etiquetado como sano de acuerdo al OCT (220.5-294.8). Ello se debe a que dichos tomógrafos manejan desviaciones estándares que varían de 10 a 20  $\mu\text{m}$  de grosor [12].

Los métodos actualmente propuestos han centrado su esfuerzos en la identificación de exudados y lesiones rojas en la imágenes de fondo de ojo, manifestaciones detectables con precisión cuando el paciente ya se encuentra en una etapa avanzada de la enfermedad, es por ello la importancia del trabajo presentado, ya que este permite generar una alerta del padecimiento cuando aún no es evidente.

#### VI. BIBLIOGRAFÍA

- [1] OMS. (2014). *Diabetes* [en línea]. Disponible en [www.who.int/mediacentre/factsheets/fs312/es/](http://www.who.int/mediacentre/factsheets/fs312/es/)
- [2] IMSS. (2014). *Detección de Retinopatía Diabética en Primer Nivel de Atención* [en línea]. Disponible en [www.imss.gob.mx](http://www.imss.gob.mx)
- [3] T. A. Ciulla, A. G. Amador, and B. Zinman, "Diabetic retinopathy and diabetic macular edema: Pathophysiology, screening, and novel therapies". *Diabetes Care*, vol. 26, no. 9, pp. 2653–2664, 2003.
- [4] CENETEC. (2014). *Retinopatía Diabética en Primer Nivel de Atención* [en línea]. Disponible en [www.cenetec.salud.gob.mx/](http://www.cenetec.salud.gob.mx/)
- [5] ASRS. (2015). *Asociación Mexicana de Retina* [en línea]. Disponible en <http://www.asrs.org/>
- [6] M. D. Abramoff, M. K. Garvin, and M. Sonka, "Retinal Imaging and Image Analysis," *IEEE Trans. Biomed. Eng.*, vol. 3, pp. 169–208, Dic. 2010.
- [7] T. E. Choe, G. Medioni, I. Cohen, A. C. Walsh, and S. R. Sadda. "2-D Registration and 3-D Shape Inference of the Retinal Fundus from Fluorescein Images," *ScienceDirect. Med Image Anal.* Vol. 12, no.2, pp. 174-190, Oct. 2007.
- [8] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and G. B. van, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [9] M. Omar, A. Hossain, L. Zhang, and H. Shum, "An intelligent mobile-based automatic diagnostic system to identify retinal diseases using mathematical morphological operations", in *Proc. 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, Dacca, Bangladés, pp. 1 – 5.
- [10] S. Rusinkiewicz, and M. Levoy. "Efficient variants of the ICP algorithm". In *Proc. Third International Conference on IEEE in 3-D Digital Imaging and Modeling*, Quebec City, Que.,
- [11] D. J. Calkins, Y. Tsukamoto, and P. Sterling, "Microcircuitry and mosaic of a blue-yellow ganglion cell in the primate retina," *J. Neurosci.*, vol. 18, no. 9, pp. 3373–3385, 1998.
- [12] *Measurements: Comparison between Cirrus HD-OCT and Stratus OCT*. Carl Zeiss Meditec Inc, Dublin, CA, 2007.
- [13] S.A. (2013). PEEK VISION. Recuperado el 20 de junio de 2014 de: <http://www.peekvision.org/>
- [14] A. Santos, R.Díaz. Et all."Reconstrucción antropomórfica de la región abdominal utilizando nubes de puntos de profundidad." Congreso Nacional de Ingeniería Biomédica 2013, Aguascalientes, México.