



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
MAESTRÍA EN CIENCIAS (NEUROBIOLOGÍA)
INSTITUTO DE NEUROBIOLOGÍA

**CAMBIOS EN LAS OSCILACIONES BASALES DEL
ELECTRORRETINOGRAMA Y EN IMÁGENES DE TOMOGRAFÍA DE
COHERENCIA ÓPTICA Y FONDO DE OJO COMO MÉTODO DE PREDICCIÓN
DEL EDEMA MACULAR DIABÉTICO**

TESIS
QUE OPTA POR EL GRADO EN
MAESTRÍA EN CIENCIAS (NEUROBIOLOGÍA)

PRESENTA:
JORGE ARMANDO HUGHES CANO

TUTOR PRINCIPAL:
DRA. STÉPHANIE COLETTE THÉBAULT
INSTITUTO DE NEUROBIOLOGÍA (INB UNAM)

COMITÉ TUTOR:
DR. ULISES OLIVARES PINTO
ESCUELA DE ESTUDIOS SUPERIORES (ENES JURQUILLA UNAM)
DR. HUGO MERCHANT NANCY
INSTITUTO DE NEUROBIOLOGÍA (INB UNAM)

SANTIAGO DE QUERÉTARO, JURQUILLA, 12 DE SEPTIEMBRE DE 2023



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Agradecemos a la Beca CONACYT CVU:1146197

Agradecemos a los donativos CONACYT CF-2019-1759 y PAPIIT IN 205420 y a los institutos con quienes tuvimos el gusto de colaborar: Instituto Mexicano de Oftalmología (IMO), Asociación Para Evitar la Ceguera (APEC) e Instituto de la Retina del Bajío (INDEREB).

Agradecemos a todos los voluntarios que han participado en el IMO, APEC e INDEREB.

Agradecemos a Luis Fernando Hernández Zimbrón, Mariana López, Nayeli Lucia Piñeiro Ramos y Erick Javier Ramírez Villa por su participación en la toma de datos en los participantes del APEC y del IMO, así como a los médicos, técnicos y administradores del IMO, APEC e INDEREB, particularmente al Dr. Marlon García Roa, el Dr. Ellery López Star y el Dr. Van Charles Lansingh.

Dedico este trabajo a:

Índice	
Índice de figuras	6
Índice de tablas	9
Resumen	10
Abstract	12
Lista de abreviaturas	14
Objeto y razones del estudio	15
Introducción	15
Justificación	18
Pregunta general	18
Preguntas específicas	18
Marco teórico	19
Anatomía	19
Edema macular diabético	21
Diagnóstico del edema macular diabético	22
Bioseñales	23
Procesamiento de bioseñales	23
Electrorretinograma	25
Inteligencia artificial	27
Redes neuronales	28
Red neuronal de aprendizaje profundo	30
Redes neuronales convolucionales	30
Aprendizaje transferido	32
Otros algoritmos de clasificación	33
Aplicación de redes neuronales en oftalmología	34
Hipótesis	37
Objetivos de la investigación	37
Objetivo general	37
Objetivos específicos	37
Metodología	38

Pacientes	38
Toma de ERG	39
Fondo de ojo	39
OCT	39
Modelo	40
Procesamiento de datos	40
Desarrollo y estructura del modelo	42
Modelos híbridos	42
Análisis del desempeño de los modelos	44
Resultados	46
Entrenamiento y pruebas	46
Validación	46
Modelos de OCT	46
Modelos de fondo de ojo	48
Modelos de ERG	49
Modelos híbridos	50
Discusión	55
Conclusiones	68
Bibliografía	69

Índice de figuras

- Figura 1.** La onda bifásica de un ERG típico. Imagen obtenida de Creel, D. (29) 25
- Figura 2. (A)** Ejemplo de toma de muestra de ERG basal (con una luminosidad de 400 lux), **(B)** Trazo de ERG basal en sujeto control y sujeto con EMD y **(C)** Gráfica de promedio de potencia relativa de sujetos con EMD (en verde) y sujetos control (en rosa)..... 27
- Figura 3:** Arquitectura de una red neuronal obtenida de Hernández, E. (24). En verde se muestran las capas de entrada de la red neuronal, en amarillo las capas ocultas y en azul las capas de salida, las cuales son las responsables de hacer la clasificación final. 28
- Figura 4:** Ejemplo de la operación de convolución. A la izquierda se ilustra una convolución bidimensional sobre una matriz de entrada. A la derecha se muestra la idea de la aplicación de esta operación sobre una imagen con tres canales de color. Imagen obtenida de David, A. & Arroyo, E. (39)..... 31
- Figura 5:** Esquema genérico de una red neuronal convolucional. Se esquematizan todas las capas que componen una red neuronal convolucional, desde la entrada, las capas de convolución, las capas de submuestreo, terminando con capas densas o *fully connected*. Imagen obtenida de Contreras, S. & De la rosa, F. (40). 32
- Figura 6.** El esquema muestra cómo funciona el aprendizaje transferido. En el panel de arriba vemos cómo una CNN es entrenada con millones de imágenes del conjunto de datos ImageNet, para clasificar 1000 clases diferentes de imágenes. Mediante el aprendizaje transferido, se aprovecha la información que obtuvo el primer modelo, y añadiendo un nuevo clasificador al final del nuevo modelo, se utiliza esta información y así, entrenando el modelo nuevo con imágenes de fondo de ojo con cuatro patologías diferentes, se puede hacer una clasificación del nuevo tipo de imágenes. Imagen obtenida de Pin, K y col. (42) 35
- Figura 7.** A) Ejemplo de una convolución convencional. B) ejemplo de una convolución atrosa o *atrous convolution*. Como muestra el diagrama, mediante el

uso de una convolución atrosa, se puede abarcar una extensión mayor de la imagen utilizando un kernel más pequeño. Imagen obtenida de Garg, R. y col. (46) 36

Figura 8. Diagrama que muestra el funcionamiento de los modelos híbridos. Se seleccionaron los mejores modelos durante la validación para cada dato de entrada. Para OCT fue OpticNet-71 y para fondo de ojo y ERG utilizando escalogramas, fue DenseNet-201. Se realizaron predicciones con los modelos antes mencionados para cada tipo de entrada (imágenes OCT, fondo de ojo o escalogramas, n = 164 casos en total) y el resultado se multiplicó por una cantidad escalar n_i que representa el porcentaje de ajuste para cada modelo. A continuación, se sumaron los tres valores resultantes para obtener una nueva matriz de predicción, utilizando las predicciones realizadas por cada modelo. El valor máximo se obtiene para cada fila de la matriz, siendo la salida 1 con DME y 2 sin DME..... 43

Figura 9. Desempeño de todos los modelos de OCT. Se muestran las curvas ROC y las matrices de confusión de **(A)** Modelo OpticNet-71 validado con n = 1229 (n= 212 con EMD y n = 1017 sin EMD), **(B)** Modelo DenseNet-201 validado con n = 1229 (n= 212 con EMD y n = 1017 sin EMD), **(C)** Modelo OpticNet-71 validado con n = 164 (n= 42 con EMD y n = 122 sin EMD) y **(D)** Modelo DenseNet-201 validado con n = 164 (n= 42 con EMD y n = 122 sin EMD). **(E)** Resumen de las métricas obtenidas por todos los modelos. AUC, indica el área bajo la curva ROC (por sus siglas en inglés), n indica el número total de datos para la validación. 48

Figura 10. Rendimiento de los modelos de fondo de ojo durante la validación. Curvas ROC, matrices de confusión y métricas obtenidas por el modelo **(A)** ResNet-50, **(B)** MobileNet-V2 y **(C)** DenseNet-201. La validación de todos los modelos incluye un total de 164 imágenes de fondo de ojo (n = 42 imágenes con EMD, y n = 122 imágenes sin EMD). 50

Figura 11. Desempeño de todos los modelos basados en ERG basal. Se muestran las curvas ROC y las matrices de confusión de **(A)** Modelo ResNet-50 alimentado con espectrogramas, **(B)** DenseNet-201 alimentado con espectrogramas, **(C)** ResNet-50 alimentado con escalogramas y **(D)** DenseNet-201 alimentado con escalogramas. Los datos de validación incluyen n = 42 casos con EMD y n = 122

casos sin EMD. **(E)** Resumen de todas las métricas obtenidas durante la validación. AUC, área bajo la curva ROC. 51

Figura 12. Optimización de los modelos híbridos. Los mapas de calor muestran las métricas de rendimiento de los modelos híbridos dobles y triples según la proporción asignada a cada modelo utilizado. **(A)** OCT e imágenes de fondo de ojo, **(B)** OCT e imágenes de escalogramas provenientes de la señal de ERG basal, **(C)** Fondo de ojo e imágenes de escalogramas y **(D)** los tres tipos diferentes de datos. En los ejes X están los porcentajes asignados para cada modelo, estos ejes son complementarios ya que todas las combinaciones suman 100 %. Para el modelo híbrido triple (D), el porcentaje permanece en porciones de 10 % para OCT, y el complemento está dado por los otros dos modelos. Esta metodología busca los valores óptimos de n_1 , n_2 y n_3 como se presentó en la Figura 9. 53

Figura 13. Comparación de todos los modelos utilizados para el diagnóstico predictivo de EMD. El mapa de calor muestra las métricas de rendimiento obtenidas para cada modelo probado, incluidos los modelos individuales o basados en una sola entrada (OCT, fondo de ojo e imágenes espectrales derivadas de ERG basal), así como los modelos híbridos más representativos. $n = 42$ para EMD y $n = 122$ para muestras sin EMD. 54

Índice de tablas

Tabla 1. Distribución y composición de los conjuntos de datos utilizados en este trabajo. Es importante remarcar que las imágenes que fueron utilizadas para el entrenamiento y las pruebas no fueron utilizadas para la validación del desempeño de los modelos. Además, la única base de datos en donde se aplicó aumento de datos fue en los escalogramas y espectrogramas utilizados para entrenamiento y pruebas, pero únicamente en la clase con EMD. 41

Tabla 2. Tabla que resume las métricas obtenidas durante el entrenamiento y las pruebas con los modelos entrenados para cada tipo de dato (OCT, fondo de ojo, escalogramas y espectrogramas). Los modelos de fondo de ojo y OCT obtienen muy buenas métricas durante el entrenamiento y las pruebas. En comparación con estos últimos, los modelos de imágenes espectrales de ERG no obtienen los mismos resultados, indicando un posible sobreajuste de los datos de entrenamiento por la posible poca cantidad de datos 47

Tabla 3. Tabla comparativa entre las validaciones de las clases de EMD y las de RD. En términos generales, las métricas obtenidas en la validación con los mismos modelos, entrenados durante las mismas épocas son mejores en la validación de las clases de RD. Todas las métricas son superiores a 0.60 y las Kappa de Cohen son mejores en esta validación. 57

Resumen

Introducción. El edema macular diabético (EMD), definido como una acumulación de líquido en la porción central de la retina, es la causa más frecuente de nuevos casos de ceguera en adultos económicamente activos con diabetes tipo 2, en todo el mundo. Actualmente, la prueba de tomografía de coherencia óptica (OCT) es la más útil para diagnosticar el EMD. Sin embargo, dicha prueba es costosa y no está disponible en todas las clínicas del país, por lo que es relevante establecer un método de diagnóstico predictivo que sea eficaz, económico y accesible para pacientes con EMD.

Objetivo general. Comprobar si los cambios en las señales del electroretinograma (ERG) basal con o sin imágenes de OCT y fondo de ojo son informativos para crear o mejorar un modelo de diagnóstico predictivo de EMD.

Métodos. Se utilizó el espectro de frecuencia de los ERG basales para generar escalogramas y espectrogramas mediante las transformadas de Wavelet y Fourier, respectivamente. Utilizando el método de aprendizaje transferido, se entrenaron las redes neuronales convolucionales (CNNs) ResNet-50, MobileNet-V2 y DenseNet-201, para la clasificación de imágenes de los escalogramas y fondo de ojo. Para las imágenes de OCT, se utilizaron las CNNs OpticNet-71 y DenseNet-201. Los conjuntos de datos se dividieron aleatoriamente en conjuntos de entrenamiento (80 %) y de prueba (20 %) para permitir la validación interna de todos los modelos. Se seleccionaron los mejores modelos para cada dato de entrada, OpticNet-71 para OCT y DenseNet-201 para fondo de ojo y escalogramas, con el fin de generar modelos combinados mediante la asignación de pesos en las matrices de probabilidad obtenidas de las predicciones de cada uno de los modelos seleccionados. Se validaron los modelos con datos ajenos a los utilizados en el entrenamiento calculando precisión, sensibilidad, especificidad, exactitud, área bajo la curva ROC, puntuación F1 y Kappa de Cohen.

Resultados. Se reportan las métricas obtenidas por los mejores modelos durante la validación (n=164), los cuales fueron: con escalogramas (precisión: 34 %, sensibilidad: 40 %, especificidad: 73 %, exactitud: 65%, Kappa: 0.13, AUC: 57 %, puntuación F1: 37 %); con fondo de ojo (precisión: 82 %, sensibilidad: 88 %,

especificidad: 93 %, exactitud: 92%, Kappa: 0.80, AUC: 91 %, puntuación F1: 85 %); con OCT (precisión: 85 %, sensibilidad: 93 %, especificidad: 94 %, exactitud: 94%, Kappa: 0.84, AUC: 88 %, puntuación F1: 89 %); fondo de ojo en combinación con ERG (precisión: 84 %, sensibilidad: 88 %, especificidad: 94 %, exactitud: 93 %, Kappa: 0.81, AUC: 91 %, puntuación F1: 86 %); y OCT en combinación con ERG: (precisión: 88 %, sensibilidad: 90 %, especificidad: 96 %, exactitud: 94 %, Kappa: 0.85, AUC: 93 %, puntuación F1: 89 %).

Conclusión. Estos resultados muestran que, aunque el ERG no es confiable por sí solo para hacer un diagnóstico de EMD, mejora las métricas de los modelos que usan fondo de ojo y OCT, exceptuando la sensibilidad de este último.

Abstract

Introduction. Diabetic Macular Edema (DME) is defined as the accumulation of liquid in the central portion of the retina, and it is the first cause of new blindness cases in the economically active population diagnosed with Type 2 Diabetes, globally. Nowadays, Optical Coherence Tomography (OCT) is considered the most useful imaging technique to diagnose DME. However, OCT is expensive, and it is not available in all clinics in Mexico; thus, it is relevant to seek a diagnostic method that is effective, economical, and accessible to patients with DME.

Objective. To test whether the variations in the basal Electroretinogram (ERG), either combined with OCT images or as a stand-alone, are informative enough to implement or improve a diagnostic prediction model for DME.

Methods. The power spectra of basal ERGs were used to generate scalograms and spectrograms via Wavelet and Fourier transforms, respectively. Using transfer learning, the following Convolutional Neural Networks (CNN) were trained as classifiers for scalogram images and eye fundus images: ResNet-50, MobileNet-V2 y DenseNet-201. Similarly, OCT images were used as input for OpticNet-71 and DenseNet-201 CNNs. Input data were randomly split into training- and test-sets with a proportion of 80 % to 20 %, respectively, to gauge the performance of each model. The top performers for each type of input were selected, OpticNet-71 for OCT and DenseNet-201 for eye fundus and scalograms, to generate a combined model by assigning different weights for each of the selected models. Finally, model validation was performed using a dataset alien to the training phase of the models; precision, sensitivity, specificity, accuracy, ROC curve, F1-score and Cohen's Kappa were calculated and leveraged to choose the best weight combination.

Results. Metrics of the best models are: precision: 34 %, sensitivity: 40 %, specificity: 73 %, accuracy: 65%, Cohen's Kappa: 0.13, AUC: 57 %, and F1 score: 37 % for ERG scalograms-; precision: 82 %, sensitivity: 88 %, specificity: 93 %, accuracy: 92%, Cohen's Kappa: 0.80, AUC: 91 %, and F1 score: 85 % for the eye fundus-; precision: 85 %, sensitivity: 93 %, specificity: 94 %, accuracy: 94%, Cohen's Kappa: 0.84, AUC: 88 %, and F1 score: 89 % for OCT-; precision: 84 %, sensitivity: 88 %, specificity: 94 %, accuracy: 93 %, Cohen's Kappa: 0.81, AUC: 91 %, and F1

score: 86 % for eyes fundus in combination with ERG; and precision: 88 %, sensitivity: 90 %, specificity: 96 %, accuracy: 94 %, Cohen's Kappa: 0.85, AUC: 93 %, and F1 score: 89 % for OCT combined with ERG-based models.

Conclusion: These results show that, although ERG model is not trustworthy (by itself) *per se* to diagnose DME, it improves all the metrics obtained by the eye fundus and OCT models, with the exception of sensitivity for the OCT model.

Lista de abreviaturas

Abreviatura	Significado
APEC	Asociación Para Evitar la Ceguera
AUC	Área bajo la curva
BRB	Barrera hematorretiniana
Cx	Conexina
EMD	Edema macular diabético
ENSANUT	Encuesta Nacional de Salud y Nutrición
EEG	Electroencefalograma
ERG	Electrorretinograma
FFT	Transformada Rápida de Fourier
FDA	Food and Drug Administration
GPU	Graphics Processing Unit
IA	Inteligencia artificial
INDEREB	Instituto de la Retina del Bajío
IMO	Instituto Mexicano de Oftalmología
MLP	Perceptrón Multicapa
NVC	Neovascularización coroidea
OCT	Tomografía de coherencia óptica
PSD	Densidad espectral de potencial
RD	Retinopatía diabética
ROC	Receiver-Operating Characteristic
UNAM	Universidad Nacional Autónoma de México
UNV	Unidad neurovascular

Objeto y razones del estudio

Introducción

La diabetes tipo 2 es un problema de salud pública mundial. A nivel global, se estima que los casos de diabetes alcanzarán los 592 millones en 2035, lo que afectará al 8.8 % de la población mundial (1).

En México, la diabetes tipo 2 y sus complicaciones representan 50 % del gasto del presupuesto sexenal destinado a salud, lo que significa una cifra de 7.7 mil millones de dólares americanos, por lo que en 2016 fue declarada emergencia epidemiológica (1). La nefropatía diabética es la primera causa de insuficiencia renal, el pie diabético es la primera causa de amputación no traumática, y las complicaciones visuales secundarias a diabetes son la primera causa de ceguera no congénita en el país (1).

Según el último reporte de la encuesta nacional de salud (*ENSANUT 2018*), los diagnósticos de diabetes en México en personas mayores de 20 años aumentaron de 6.4 millones en 2012 a 8.6 millones en 2018 (2). Además, las guías de la *American Diabetes Association* reportan que la retinopatía diabética (RD) y el edema macular diabético (EMD) son las causas más frecuentes de nuevos casos de ceguera en adultos de 20 a 74 años en todo el mundo (3).

El EM puede ocurrir como resultado de muchas patologías retinianas, pero la causa más frecuente es asociada a la diabetes (4). Éste es definido como una acumulación de líquido en la porción central de la retina, (5) con o sin la presencia de otros signos de microangiopatía en el fondo de ojo, y por ello tiene una clasificación independiente de la RD (5). Alrededor del 20 % de los pacientes con diabetes tipo 1 y 25 % de los pacientes con diabetes tipo 2 desarrollan EM después de 10 años de evolución de su enfermedad (3).

Hasta antes de la angiografía por tomografía de coherencia óptica (OCTA por sus siglas en inglés), el EM se dividía en focal y difuso, con diferencias en torno a los hallazgos en la angiografía con fluoresceína (6). La clasificación se vuelve importante para establecer un tratamiento adecuado, por lo que el *Early Treatment Diabetic Retinopathy Study* introdujo el término de EMD clínicamente significativo, cuyas características son: engrosamiento retiniano a 500 micras o menos del centro

de la fovea, exudados duros a 500 micras o menos del centro de la fovea y engrosamiento retiniano al menos en un disco de la fovea, localizado a menos de un diámetro papilar del centro de la fovea (7). Estos pacientes se podrían beneficiar de varios tratamientos; el más clásico es el tratamiento con láser (5). También se pueden aplicar antiangiogénicos y/o corticoides (5), considerando siempre el buen control metabólico como la base de cualquier tratamiento en el paciente diabético. Hasta el momento, únicamente el control metabólico y la fotocoagulación con láser han demostrado de forma concluyente su eficacia en el tratamiento del EMD mediante estudios prospectivos aleatorizados (6). También se utilizan tratamientos intravítreos con medicamentos antiangiogénicos o con implantes de esteroides, pero no todos los pacientes responden, además de que el tratamiento con esteroides se ha relacionado con la formación de cataratas y el desarrollo de hipertensión intraocular (8). Por último, podemos mencionar la vitrectomía, para la cual la mayor parte de los estudios efectuados para comprobar su validez, incluyen un número escaso de pacientes y arrojan resultados contradictorios (6); mientras que algunos encuentran mejoría tanto anatómica como funcional, otros describen sólo mejoría en el aspecto anatómico de la mácula, pero sin cambios en la agudeza visual (9). De ahí la importancia de detectar oportunamente el EMD.

La prueba de OCT es la más útil para diagnosticar y clasificar el EMD. Los cortes tomográficos pueden detectar de forma objetiva y con precisión la presencia de líquido en las capas de la retina y permiten rastrear la existencia de zonas de engrosamiento retiniano (9). Estas medidas proporcionan datos objetivos para valorar la evolución de la patología y la respuesta al tratamiento de estos pacientes (9).

El problema con los métodos diagnósticos actuales, como la OCT, es que son costosos, además de que no es un examen disponible en todas las clínicas del país. Los demás métodos, como la angiografía con fluoresceína, son exámenes altamente invasivos. Y el que aún es el más utilizado, el examen biomicroscópico con lámpara de hendidura, también conocido como fondo de ojo, tiene el problema de que la información que proporciona es exclusivamente cualitativa, subjetiva y

poco sensible, puesto que se necesitan grandes variaciones en el grosor de la retina para que puedan ser apreciadas (6).

Por todo lo anterior, se hace relevante establecer un método de diagnóstico predictivo que sea eficaz, económico y accesible para pacientes que presenten EMD.

Justificación

A pesar de que existen tratamientos para el EMD, todos son sintomáticos y la prevención en términos de complicaciones de la diabetes es la mejor opción. En ese sentido, los métodos diagnósticos actuales son útiles, pero sería bueno contar con un método de diagnóstico predictivo que use, eventualmente, otra fuente diferente a las imágenes de OCT y fondo de ojo. Se propone el desarrollo de un método de diagnóstico predictivo del EMD basado en técnicas de inteligencia artificial, utilizando la clasificación de señales de electroretinograma (ERG) basal, combinada o no con imágenes de OCT y fondo de ojo.

Pregunta general

¿Los cambios en las señales de ERG basal pueden, solos o combinados con imágenes de fondo de ojo y OCT, servir como método de diagnóstico predictivo para el EMD?

Preguntas específicas

¿Se puede generar un modelo de inteligencia artificial del tipo red neuronal convolucional que utilice imágenes de los cambios en el dominio de las frecuencias de las señales de ERG basal para diagnosticar EMD?

¿Se puede mejorar el desempeño de los clasificadores basados en imágenes de fondo de ojo y OCT para predecir el EMD combinándolos con las señales antes descritas?

Marco teórico

Anatomía

La retina es el órgano encargado de procesar la información lumínica (10). Se encuentra en el fondo del globo ocular y se extiende en sentido anterior casi hasta el cuerpo ciliar. Se organiza en capas que contienen a las células encargadas de llevar a cabo la fototransducción y el procesamiento de la información lumínica relativa al contraste, color, direccionalidad y velocidad (10). La capa nuclear externa contiene bastones y conos, que son los fotorreceptores. La capa nuclear interna incluye a las células bipolares, horizontales y amacrinas. La capa ganglionar contiene diversos tipos de células ganglionares, cuyos axones forman el nervio óptico (10). La retina cuenta con células no neuronales que tienen funciones de soporte y que participan en la regulación electrolítica y metabólica de las neuronas; también contribuyen en la actividad inmunológica, la regulación de la angiogénesis y el flujo sanguíneo (11). Estas células son los astrocitos, la microglía, el epitelio pigmentado y las células de Müller (10). Otras células importantes en la retina son las células endoteliales; estas forman los capilares intrarretinianos (10). Por último, podemos mencionar a los pericitos, que son células contráctiles que se encuentran alrededor de las células endoteliales de los capilares intrarretinianos y participan en la barrera hematorretiniana (BRB por sus siglas en inglés) (10).

La mácula es una estructura altamente especializada de la retina y es responsable de la agudeza visual fina (4). Se localiza en la región temporal a la cabeza del nervio óptico y puede reconocerse mediante la visualización del fondo de ojo por su color amarillo, debido a la acumulación de pigmentos como luteína y zeaxantina (4). El centro de la mácula es normalmente avascular, rodeado por capilares dispuestos circularmente que delimitan la zona de la fovea, que es el centro de la mácula y se caracteriza por su gran cantidad de conos, que son los fotorreceptores responsables de la visión a detalle (4).

Las células endoteliales en la retina están conectadas por uniones estrechas (zónula occludens), uniones adherentes (zónula adherens) y por uniones comunicantes (11). Las uniones estrechas endoteliales son estructuras complejas formadas por múltiples proteínas transmembranales, de andamiaje y de

señalización (4). Las proteínas transmembranales más importantes son ocludina y claudina- 1, 2 y 5, pero también existen otras como las moléculas de adhesión específicas de células endoteliales (VCAM), moléculas de adhesión de unión (JAMs) y proteínas 1 y 2 de uniones estrechas (ZO-1, ZO- 2) (11). Las uniones adherentes están formadas por caderinas, y proteínas cateninas, como la beta-catenina, principalmente. Las uniones comunicantes están formadas por dos conexones o hemicanales, cada uno de ellos compuesto a su vez por seis monómeros de proteínas llamadas conexinas (Cx) (12). En la actualidad se han reportado 21 tipos diferentes de Cx (13). En la retina de ratón, cuatro Cx diferentes han sido detectadas a nivel proteico: la Cx36, Cx37, Cx43 y la Cx45; la Cx43 principalmente entre astrocitos, la Cx37 entre células endoteliales y las demás Cx36, Cx45 y Cx57 entre neuronas (14). En el humano, la Cx36 y la Cx45, así como el gen de la Cx59 y Cx62, han sido reportados en retina (15).

En condiciones fisiológicas, diferentes mecanismos mantienen a la retina en un estado transparente y relativamente deshidratado (4). La BRB juega un papel importante en el control de la entrada de líquidos a la retina (4). Su componente interno está formado por las uniones estrechas entre las células endoteliales de los vasos retinianos, así como los pericitos y pies de los astrocitos y células de Müller que, además de recubrir parte de las células endoteliales, interactúan de manera activa con otras células de la retina (4).

En el desarrollo de vasos retinianos, la formación de uniones estrechas entre las células endoteliales depende de sus interacciones celulares con pericitos y astrocitos (11), destacando el papel fundamental de la unidad neurovascular en los primeros pasos de la formación de la BRB interna.

El término “unidad neurovascular” (UNV) se refiere al acoplamiento funcional y a la interdependencia de las neuronas, la glía y la vasculatura altamente especializada en el sistema nervioso central (16). En el contexto de la retina, todas las células componentes de la UNV están en comunicación íntima y mantienen la integridad de la BRB, mientras regulan dinámicamente el flujo sanguíneo en respuesta a las demandas metabólicas (16).

Edema macular diabético

El EMD generalmente es definido como el engrosamiento de la retina, con presencia de exudados dentro de un disco de diámetro desde el centro de la mácula, y se caracteriza por la presencia de edema en la zona macular, ocasionando su deterioro (17).

El EM es consecuencia de una desregulación entre el líquido que entra y el líquido que sale de la retina (4). En condiciones de hiperglucemia, sólo uno de estos mecanismos prevalece, y es el que respecta al incremento de la permeabilidad de la BRB interna (18). Alteraciones en las uniones intercelulares por regulación a la baja, cambio en el estado de fosforilación o pérdida de proteínas de anclaje pueden ocasionar esta desregulación (4).

Algunos estudios sugieren que también los astrocitos pueden contribuir a la pérdida de la homeostasis de la BRB interna (12): la proximidad de estos a la vasculatura retiniana y a las células ganglionares, dos tipos de células que se sabe que sufren apoptosis durante la diabetes (19), los convierte en un factor crítico en la regulación de la UNV (12). Experimentos como los de Ly y *col.* (20), mediante la inducción de diabetes en ratas Sprague-Dawley, muestran que la expresión del ARN mensajero y proteína de Cx26 y Cx43 en astrocitos disminuye después de cuatro semanas de haber inducido la enfermedad; esto ocurre antes de encontrarse con una pérdida significativa de estas células. Al mismo tiempo, se muestra hipoxia en las células ganglionares mediante un aumento en la expresión de factor 1 alfa inducible por hipoxia (HIF – 1a), que coincide con una disminución en la función de las células ganglionares. Después de seis semanas, las ratas presentan gliosis en células de Müller y una reducción en la respuesta de conos y bastones medida mediante electroretinograma (ERG) (20). Esto sugiere que la regulación a la baja de las Cx asociada a la diabetes, compromete la formación de las uniones Gap de los astrocitos con otras células retinianas, lo que dificulta la comunicación entre células y, por lo tanto, influye en la pérdida de la homeostasis en la UNV (11).

Por otro lado, las uniones estrechas también se ven afectadas. La pérdida de la integridad funcional y estructural de las uniones estrechas entre las células endoteliales aumenta la permeabilidad vascular al agua, los solutos y las proteínas

(4). Varias condiciones experimentales (en cultivos celulares o animales de experimentación) que imitan entornos patológicos que favorecen el EM, como la diabetes, pueden alterar las proteínas de unión (4). Existen estudios, como el de Navarra y *col.*, que demuestran que existe una degradación proteolítica de caderina vascular endotelial en condiciones de diabetes (21). Además, se ha demostrado que, en condiciones de hiperglucemia, hay regulación a la baja de claudina 5, ocludina, molécula JAM-A y ZO -1, en células endoteliales de retina humana (4). Todo esto favorece la pérdida de la UNV, lo que conlleva una disfunción en las células neuronales de la retina. Estos cambios se ven reflejados en la manifestación clínica más importante del EMD, la cual es una disminución de la agudeza visual central que puede aparecer desde etapas muy tempranas de la enfermedad (17).

Diagnóstico del edema macular diabético

El diagnóstico de EMD debe emitirse mediante una combinación entre el examen clínico, y la interpretación de imágenes como la OCT o la biomicroscopía con lámpara de hendidura (17).

Hasta antes de la llegada de la OCT, el método de diagnóstico más confiable para el EMD era la biomicroscopía con lámpara de hendidura, en la cual se obtienen imágenes de fondo de ojo que nos permiten observar la presencia de exudados en el área de la retina. Es un examen relativamente sencillo, seguro y no invasivo, además de ser económico (desde \$50 en un negocio minorista o en una cadena de ópticas, realizado por un optometrista hasta \$250 o más en una clínica oftalmológica o en el consultorio privado de un(a) oftalmólogo(a)) que únicamente requiere dilatación pupilar farmacológica para poder obtener una imagen con mejor calidad (17). Este método tiene algunas limitaciones, ya que la información que proporciona es exclusivamente cualitativa, subjetiva y poco sensible, puesto que se necesitan grandes variaciones en el grosor de la retina para que puedan ser apreciadas (6).

Por otro lado, la OCT se ha convertido en la piedra angular del diagnóstico y seguimiento de los pacientes con EMD, debido a que proporciona información acerca del grosor de la retina, ubicación de la zona con edema, y la presencia de

otras anomalías (17). Además, es una tecnología segura, no invasiva y rápida, que proporciona información objetiva, ya que obtiene imágenes tomográficas de la retina a diferentes niveles (17). La limitación de esta tecnología es que, en el contexto de México, no es una herramienta económicamente accesible para toda la población, además de que no se encuentra en todos los hospitales del país.

Bioseñales

El término de bioseñal, se refiere a todos los tipos de señales que pueden ser medidas continuamente en los seres vivos, y hace referencia tanto a bioseñales eléctricas como no eléctricas (22). La obtención de bioseñales permite a investigadores y médicos extraer información sobre el funcionamiento de las diferentes estructuras, como órganos o tejidos, para poder emitir un diagnóstico (23). Entre los ejemplos de bioseñales más conocidas se tiene a los electroencefalogramas (EEG), electrocardiogramas y ERG (22).

Procesamiento de bioseñales

Las bioseñales se dividen en diversas categorías: biopotenciales, impedancias, imágenes, señales mecánicas, acústicas, biomagnéticas y señales bioquímicas.

Las bioseñales suelen ser una fuente de información muy apropiada para el diagnóstico de cualquier tipo de enfermedad, el reto es saber cuál es el tratamiento que se le dará a dicha señal para aprovechar de mejor manera la información obtenida (23).

Algunos tipos de células, como las células nerviosas, presentan la característica de producir potenciales bioeléctricos como resultado de la actividad electroquímica de sus membranas (23). La medida de esta actividad proporciona información sobre su funcionamiento.

Los factores más importantes que caracterizan las bioseñales eléctricas, desde el punto de vista de la instrumentación, son su amplitud y frecuencia (23).

La señal cruda de una bioseñal eléctrica es una serie de tiempo. Estas señales en crudo se utilizan raramente ya que, además de la gran cantidad de datos que representan, suelen contener ruido electromagnético (23).

Para las bioseñales eléctricas, existen diferentes tipos de procesamiento que trasladan la señal pura a dominios diferentes, es decir, una señal que permanecía en un dominio de tiempo y voltaje ahora puede ser trasladada al dominio de la frecuencia o tiempo-frecuencia (23).

Para hacer este cambio de dominio existen diversas herramientas matemáticas. Una de las más utilizadas en el estado del arte es la Transformada Discreta de Fourier (FDT). Se trata de una herramienta muy potente que permite representar una señal periódica mediante sumas de Fourier. La desventaja de esta es que consume una gran cantidad de recursos computacionales, por lo que la transformada convencional puede ser sustituida por la Transformada Rápida de Fourier (FFT).

La FFT es un algoritmo que permite aproximar la transformada discreta de Fourier, con un menor costo computacional (24). El algoritmo estima la magnitud y fase de los componentes sinusoidales periódicos que permiten reconstruir una señal como la combinación lineal de dichos componentes, lo cual permite describir a una señal en el dominio de frecuencia (24).

Por otro lado, el espectrograma es el resultado de calcular el espectro de una señal por ventanas de tiempo (25). El resultado es una gráfica tridimensional que representa la energía de las frecuencias de la señal según van variando en el tiempo (25). El espectrograma es una proyección en dos dimensiones de una sucesión de Transformadas de Fourier, donde la energía y el contenido de frecuencias de la señal van variando a lo largo del tiempo (25).

Otra herramienta matemática para trabajar con el dominio de la frecuencia es la Transformada de Wavelet. Las wavelets son funciones matemáticas que se pueden utilizar para filtrar series temporales de datos y analizar la variación de su contenido espectral, ofreciendo una representación tiempo – frecuencia más precisa para señales no periódicas que las obtenidas mediante el análisis tradicional de Fourier (26). Una wavelet es una señal de duración limitada cuyo valor medio es cero. Así, la Transformada de Wavelet es una operación matemática que descompone la señal que se va a analizar en versiones trasladadas (en el tiempo) y escaladas de una wavelet seleccionada, mejor conocida como wavelet madre (26).

Datos en el dominio de la frecuencia han sido utilizado previamente, en formato de imagen, para el análisis de bioseñales. En el 2017, Türk y Özerdem, usaron una base de datos de señales de EEG para analizar la detección de epilepsia mediante imágenes de las señales. El esquema propuesto por los autores consiste en una transformación de la señal del EEG del dominio del tiempo al dominio de la frecuencia, utilizando la Transformada de Wavelet. A partir de ésta, se genera una imagen de 662x536 pixeles, la cual es analizada posteriormente mediante una red neuronal convolucional (27). De manera similar, en 2021, Villazana y Montilla, proponen un clasificador multiclase para clasificar señales de EEG en tres clases (Normal, Preictal e Ictal), el cual se basa en la imagen de la transformación tiempo-frecuencia de la señal temporal y el uso de una red neuronal convolucional para clasificar la señal en una de estas tres clases (28).

Electrorretinograma

El ERG es una bioseñal, la cual clásicamente mide la respuesta eléctrica de la retina a un estímulo fótico mediante una onda con forma bifásica (**Figura 1**) (29).

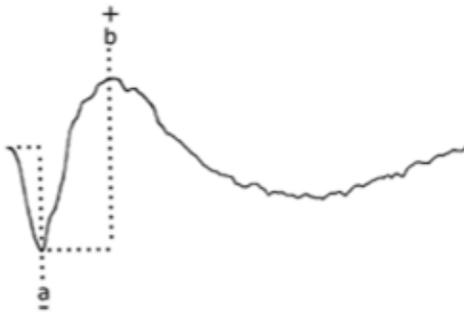


Figura 1. La onda bifásica de un ERG típico. Imagen obtenida de Creel, D. (29)

El ERG es reflejo del estado funcional de capas externas, medias e internas de la retina y su alteración suele deberse al daño funcional o histológico de dichas estructuras (29). La onda A es originada en la capa de los fotorreceptores y corresponde a la primera onda negativa del registro. La onda B es originada por la actividad sináptica entre fotorreceptores y células bipolares, así como por las células ganglionares y la glía de Müller, y es representada por la gran onda positiva (29). Es consecutiva de la onda A. Existen otras ondas del ERG como la onda C

generada en el epitelio pigmentario de la retina y los potenciales oscilatorios, entre otras (29).

Los valores del ERG pueden cambiar según la metodología de registro y parámetros del equipo que se use, por ejemplo: adaptación a la luz u oscuridad, intensidad y características temporales del flash de luz, electrodo, parámetros de registro, entre otros (29). Dado que el ERG mide el estado funcional de la retina, es de gran utilidad para el diagnóstico diferencial de las enfermedades de la retina. En particular, los trastornos metabólicos pueden reducir las amplitudes de las ondas electroretinográficas (29). Debido a esto, las alteraciones tempranas en el EMD pueden ser detectadas por el ERG (30).

Hong – He y *col.*, estudiaron la correlación entre la morfología y la función de 55 ojos en pacientes diabéticos diagnosticados con EM mediante el uso de un ERG multifocal, en comparación con ojos sanos de un grupo control. Encontraron que la amplitud de las ondas negativa (N1) y positiva (P1) disminuyó y, por otro lado, aumentó su latencia en pacientes con diabetes (31). Cabe mencionar que para la formación de N1 en el ERG multifocal, contribuyen las mismas células que contribuyen para la formación de la onda A del ERG, y que P1 incluye las contribuciones de la onda-b y potenciales oscilatorios (29).

Neda y *col.*, también estudiaron la correlación entre la morfología y la función de ojos en pacientes con EMD diagnosticado. Ellos usaron ERG de campo completo para medir la funcionalidad de las capas retinianas y OCT para estudiar la morfología de los pacientes con EMD. Encontraron que la amplitud de las ondas N1 y P1 decrecían en comparación con el grupo control; también encontraron latencia aumentada en las mismas ondas en comparación con el grupo control (32). Todo esto nos indica cómo el ERG muestra cambios significativos en la función de la retina en pacientes con EMD (23, 24). Estos cambios pueden ser medidos e interpretados para poder dar un diagnóstico de la enfermedad mediante un estudio cuantitativo y objetivo. Cabe añadir que, en la actualidad, los dispositivos para medir ERG son completamente no invasivos y económicos, haciendo accesibles los estudios para la población (33).

Es en ese contexto que se plantea la primera parte de este proyecto, que versa sobre el uso de las señales espontáneas registradas con el ERG no invasivo. En el grupo de trabajo se ha encontrado que el ERG permite captar las oscilaciones espontáneas del potencial de campo de la retina y, de manera notable, que esa señal se modifica desde etapas tempranas de la RD (34). Por lo que, cuando se alimenta un algoritmo de aprendizaje supervisado con el espectro de potencia de las oscilaciones espontáneas del ERG, se puede predecir el estado sano versus enfermo, refiriéndose únicamente a factores de riesgo para la diabetes tipo 2 y la RD asociada (34). Aquí queremos determinar si esa serie de tiempo sirve para predecir EMD, ya que tenemos evidencias preliminares de que el espectro de potencia de las oscilaciones espontáneas del ERG es diferente entre pacientes con EMD y el grupo control sin EMD (**Figura 2**).

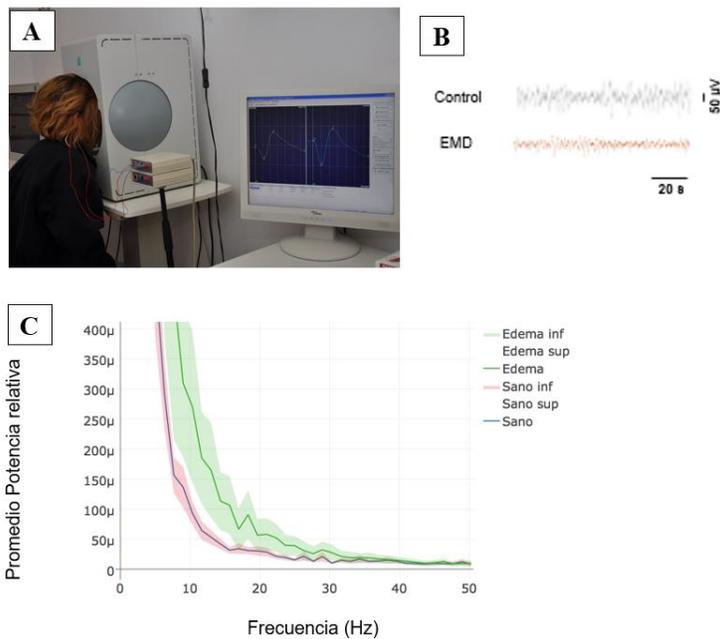


Figura 2. (A) Ejemplo de toma de muestra de ERG basal (con una luminosidad de 400 lux), (B) Trazo de ERG basal en sujeto control y sujeto con EMD y (C) Gráfica de promedio de potencia relativa de sujetos con EMD (en verde) y sujetos control (en rosa).

Inteligencia artificial

La inteligencia artificial (IA) es usada en oftalmología para el análisis de datos, segmentación de imágenes, diagnóstico automatizado, y, en algunos casos, para predicciones (35). La IA consiste, básicamente, en el proceso de automatizar las

tareas intelectuales que son realizadas por los seres humanos (36). Es un campo que abarca el aprendizaje automático y el aprendizaje profundo. Este último está diseñado con una estructura que intenta replicar la forma de las neuronas humanas, denominada red neuronal (37).

Redes neuronales

Para entender el funcionamiento de las redes neuronales, resulta fundamental conocer cómo se organizan y conectan sus componentes individuales: las neuronas (37). La estructura de una red neuronal se compone principalmente del número de neuronas y del número de capas (24). En términos generales, se puede decir que los datos se ingresan a través de las neuronas de la capa de entrada y viajan a lo largo de una o varias capas intermedias, también llamadas capas ocultas, hasta que se llega a la capa de salida, la cual comunica la decisión final del programa (**Figura 3**).

Las redes neuronales artificiales alimentadas hacia adelante (*feedforward*) son la topología más común de redes neuronales; son conocidas como perceptrón multicapa (MLP por sus siglas en inglés) y están formadas por tres partes: capa de entrada, capas ocultas y capa de salida (24).

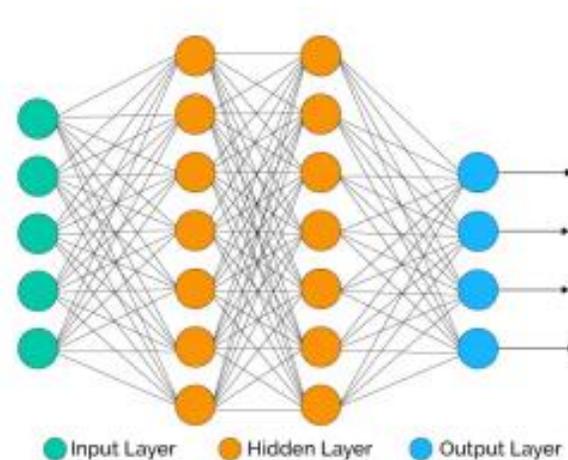


Figura 3: Arquitectura de una red neuronal obtenida de Hernández, E. (24). En verde se muestran las capas de entrada de la red neuronal, en amarillo las capas ocultas y en azul las capas de salida, las cuales son las responsables de hacer la clasificación final.

Es importante mencionar que, en este tipo de redes, ninguna de las conexiones retrocederá a través del sistema, a diferencia de las redes neuronales recurrentes (37). Estas últimas funcionan de forma cíclica, ya que las salidas de la red retroalimentan a las neuronas de entrada. De esta forma, la salida de la red depende no solo de las entradas de ésta, sino también de las salidas anteriores de la propia red (**Figura 4**). El uso de ciclos puede tener un efecto importante en los mecanismos de aprendizaje y su funcionamiento. El resultado es un sistema mucho más dinámico (37,38).

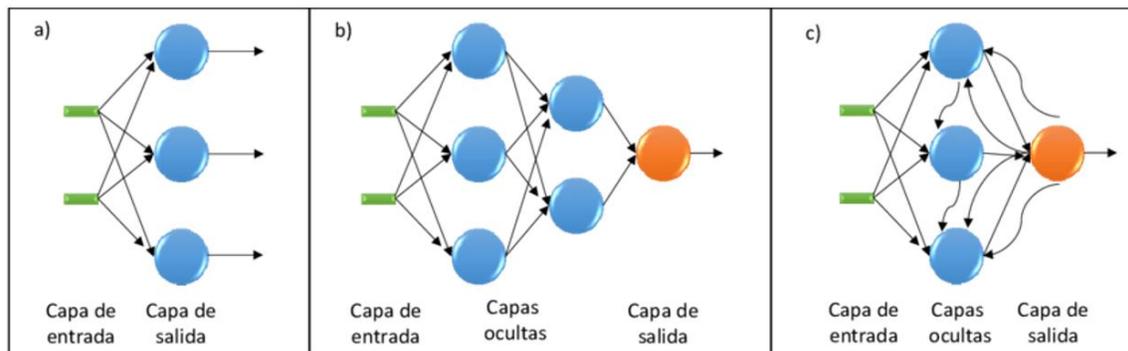


Figura 4: a) Red Neuronal monocapa, b) Red Neuronal tipo perceptrón multicapa y c) Red Neuronal multicapa recurrente. El perceptrón multicapa es la Red Neuronal Artificial más conocida y con un mayor número de aplicaciones, y es importante mencionar que, en este tipo de redes, ninguna conexión retrocederá a través del sistema. Por otro lado, las redes neuronales recurrentes funcionan de forma cíclica, pues las salidas de la red retroalimentan a las neuronas de entrada. Imagen obtenida de Manjarrez, L. (38)

Las redes neuronales tienen la capacidad de ir modificando su comportamiento para ajustarlo a los datos que se le van proporcionando. Esta es su propiedad más llamativa, ya que les confiere cierta capacidad de aprendizaje (24). Existen dos tipos fundamentales de algoritmos de aprendizaje, basados en el grado de interacción humana con el programa: el aprendizaje supervisado y el aprendizaje no supervisado (37).

En el aprendizaje supervisado, el programa conoce la respuesta que está tratando de alcanzar y simplemente necesita resolver los diferentes pasos que necesita

seguir para llegar a ella (24). En el aprendizaje no supervisado, por el contrario, los resultados de los problemas aún no se conocen y el programa utiliza operaciones lógicas para hacer una clasificación (24). El enfoque principal es localizar patrones e identificar propiedades en los datos que le permita clasificarlos.

Red neuronal de aprendizaje profundo

El aprendizaje profundo (*deep learning*) es un subtipo de aprendizaje automático en el cual no es tan necesaria la interacción humana, ya que la máquina tiene la capacidad de alimentarse con datos sin procesar y descubrir automáticamente los patrones o características necesarias para las tareas de detección o clasificación (37). Este tipo de aprendizaje utiliza redes neuronales compuestas de múltiples capas en la que cada capa recibe información procesada proveniente de la capa anterior, la cual analiza y transmite a la capa siguiente. Lo anterior crea conocimiento jerarquizado en el que cada capa recibe información cada vez más abstracta proveniente de la capa anterior (37).

Redes neuronales convolucionales

Las redes neuronales convolucionales son modelos que se utilizan con mayor frecuencia para el procesamiento de imágenes (36). Están diseñadas de tal forma que intentan imitar la estructura de la corteza visual animal. La operación de convolución consiste en superponer un filtro sobre una imagen e ir desplazando el filtro sobre la misma hasta haber recorrido todas las regiones de la imagen (39). Por cada desplazamiento del filtro sobre la imagen se calcula el valor resultante de la suma de todas las multiplicaciones de cada píxel por el valor correspondiente a la posición del filtro que superpone a ese píxel. De esta forma se construye el mapa de características: cada vez que se calcula un nuevo resultado, se almacena en la siguiente posición de la matriz resultado (39) (**Figura 5**).

La entrada a una capa convolucional suele ser una imagen de tamaño $m \times n \times c$, donde m y n son la anchura y altura de la imagen y c , la profundidad o el número de canales (39). Las capas convolucionales suelen constar de más de un filtro

(también denominado *kernel*), de esta forma, cada filtro genera mediante la operación de convolución un nuevo mapa de características bidimensional (39). El total de mapas de características generados a la salida de la capa será igual al número de filtros de la capa, por lo que podemos ver la salida en su conjunto como una nueva imagen cuyo número de canales es igual al número de filtros de la capa. Esta nueva imagen pasará a ser la entrada de la siguiente capa (39).

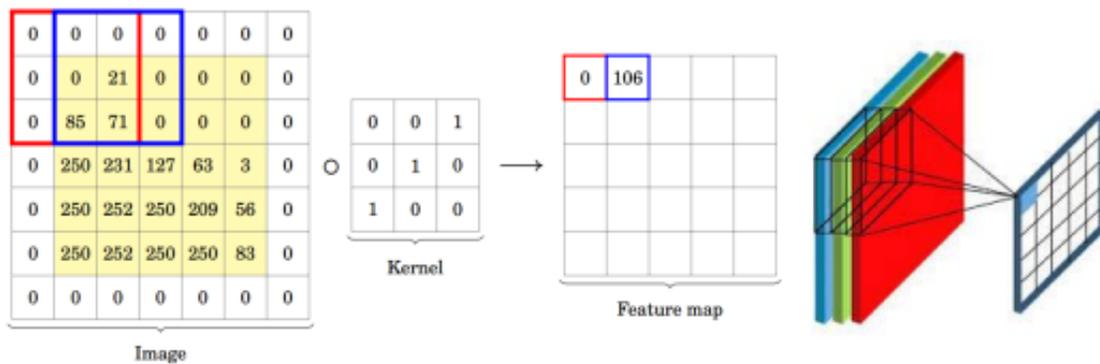


Figura 5: Ejemplo de la operación de convolución. A la izquierda se ilustra una convolución bidimensional sobre una matriz de entrada. A la derecha se muestra la idea de la aplicación de esta operación sobre una imagen con tres canales de color. Imagen obtenida de David, A. & Arroyo, E. (39).

Adicionalmente, se suele introducir una capa de *pooling* tras la convolución, que sirve para reducir el tamaño de las imágenes realizando una operación de agregación sobre los valores de una región contigua (39). Al final, al querer realizarse una tarea de clasificación, se añade una capa para realizar esta tarea (**Figura 6**).

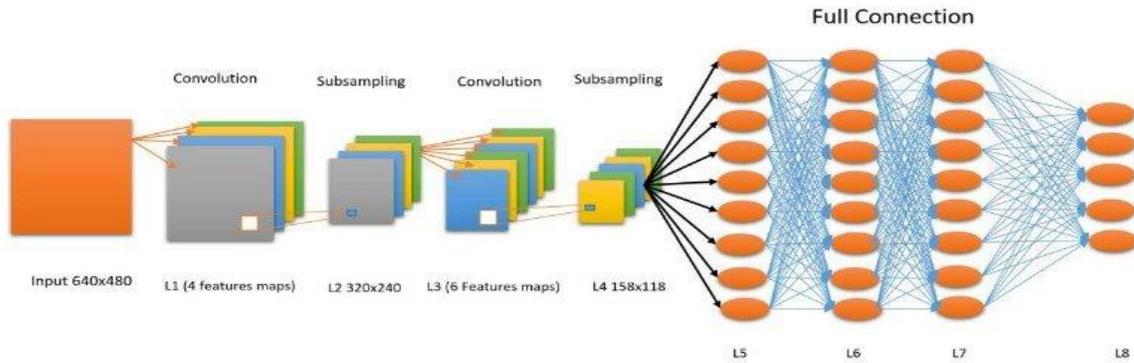


Figura 6: Esquema genérico de una red neuronal convolucional. Se esquematizan todas las capas que componen una red neuronal convolucional, desde la entrada, las capas de convolución, las capas de submuestreo, terminando con capas densas o *fully connected*. Imagen obtenida de Contreras, S. & De la rosa, F. (40).

Aprendizaje transferido

La transferencia de aprendizaje o *transfer learning*, es una técnica que forma parte del aprendizaje profundo y consiste en utilizar una red pre-entrenada, es decir, reusar la arquitectura y los pesos de un modelo entrenado con grandes volúmenes de datos de entrada y aplicarlos a escenarios distintos con otros sets de datos, buscando realizar clasificaciones más rápidas y utilizando menor carga computacional (41).

Un modelo pre-entrenado es un modelo que fue entrenado con un gran conjunto de datos de referencia para resolver un problema similar al que se quiere abordar. Debido al costo computacional que representa el entrenamiento de los modelos, así como en la complejidad al momento de elegir la arquitectura óptima, el *transfer learning* de modelos conocidos y precisos se ha convertido en una práctica común (por ejemplo, ResNet-50 o MobilNet V2) (41).

En la arquitectura de estos modelos, hay una transición donde se pasa de extraer los rasgos más generales a los más específicos. Como resultado, las capas inferiores de una CNN se refieren a características generales, mientras que la parte clasificadora y algunas de las capas superiores de la base convolucional, se refieren a características especializadas (41). La idea principal es mantener la base convolucional en su forma original, y luego usar sus salidas para alimentar un clasificador propio (41,42) (**Figura 7**).

Otros algoritmos de clasificación

Existen otros algoritmos que pueden ser utilizados para la clasificación de imágenes, como los bosques aleatorios, las máquinas de soporte vectorial (SVM) o el algoritmo de k vecinos más cercanos (KNN) (43).

Los métodos de clasificación basados en árboles de decisión son ampliamente utilizados (44). Estos métodos involucran la creación de particiones en el espacio de características de los datos a clasificar (45). De esta manera, se construye un árbol que divide el espacio de características en múltiples regiones, donde cada región tiene un modelo de clasificación asociado (45). Este enfoque permite una representación flexible y jerárquica de los datos, donde las decisiones de clasificación se toman basándose en las características relevantes en cada región específica del espacio de características (45).

Por otro lado, una SVM es un clasificador lineal que parte de la idea de encontrar un parámetro en un hiperplano (45), mediante esta metodología se pretende la extracción de características relevantes, la representación de imágenes como vectores de características, la búsqueda de un hiperplano óptimo que separe las clases en el espacio de características y la clasificación de nuevas imágenes en función de su posición relativa al hiperplano (45).

Otro algoritmo utilizado para este fin es KNN, es relativamente simple y se basa en el principio de que objetos similares tienden a estar cerca en el espacio de características (46). En términos generales, el algoritmo de KNN para la clasificación de imágenes implica calcular la distancia entre una imagen de prueba y las imágenes del conjunto de entrenamiento, seleccionar los k vecinos más cercanos, realizar una votación entre ellos y asignar la clase con mayor número de votos a la imagen de prueba (46).

Aunque estos algoritmos han mostrado ser eficaces para la clasificación de imágenes (43,44,47,48), las CNN han mostrado mejor rendimiento debido a su capacidad de aprender características relevantes, capturar representaciones jerárquicas, manejar imágenes de alta dimensionalidad y tener una mayor capacidad de generalización (48,49). Estas ventajas han llevado a un gran avance

en la clasificación de imágenes y han hecho que las CNN sean el algoritmo preferido en este campo.

Bajo esta idea, el grupo de Wang, P. y col. (48), hicieron una comparación entre el algoritmo SVM y una CNN para la clasificación de imágenes, utilizando el conjunto de datos público *mnist dataset*, midiendo el tiempo de entrenamiento de los modelos, que fue de 27.6 minutos para SVM y 23.2 minutos para la CNN. Además, midieron la exactitud de ambos modelos, que fue de 0.88 y 0.98, respectivamente. Este es sólo un ejemplo de cómo las CNN pueden desempeñarse de mejor manera al momento de clasificar imágenes, en comparación con los modelos tradicionales de aprendizaje automático.

Aplicación de redes neuronales en oftalmología

La asociación de tecnologías de aprendizaje profundo y OCT ha demostrado ser confiable para la detección de enfermedades de la retina y la mejora del rendimiento diagnóstico de las enfermedades del segmento posterior del ojo (35).

En este ámbito, en el 2018, Kermany y col. desarrollaron un algoritmo de aprendizaje por transferencia, efectivo para procesar imágenes médicas y proporcionar un diagnóstico preciso y oportuno de la patología clave en cada imagen. Para este estudio, utilizaron 108 312 imágenes de retina obtenidas por OCT. Utilizaron 37 206 imágenes con neovascularización coroidea (NVC), 11 349 con EMD, 8 617 con drusas y 51 140 imágenes de retinas normales. El propósito del estudio fue que el algoritmo pudiera clasificar las imágenes correspondientes a NVC y EMD como referencias urgentes; a las imágenes con drusas, que son depósitos lipídicos en la mácula (50), como referencias de rutina; y las imágenes normales, como observación. En una comparación multiclase lograron una precisión de 96.6 %, con una sensibilidad del 97.8 %, una especificidad del 97.4 % y un error ponderado del 6.6 % (50).

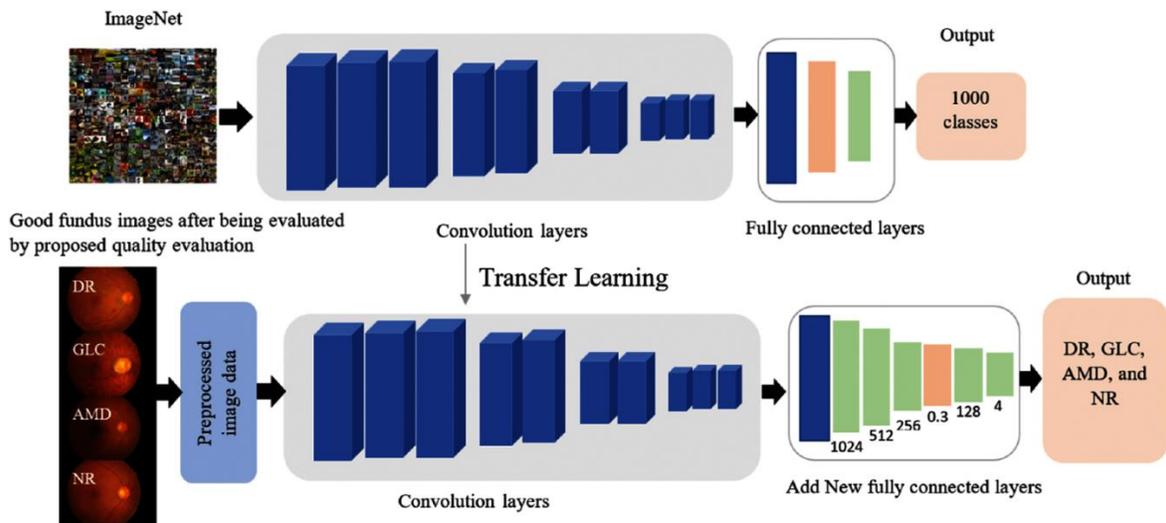


Figura 7. El esquema muestra cómo funciona el aprendizaje transferido. En el panel de arriba vemos cómo una CNN es entrenada con millones de imágenes del conjunto de datos ImageNet, para clasificar 1000 clases diferentes de imágenes. Mediante el aprendizaje transferido, se aprovecha la información que obtuvo el primer modelo, y añadiendo un nuevo clasificador al final del nuevo modelo se utiliza esta información y así entrenando el modelo nuevo con imágenes de fondo de ojo con cuatro patologías diferentes, se puede hacer una clasificación del nuevo tipo de imágenes. Imagen obtenida de Pin, K y col. (42)

Otro estudio importante, pero referido a las imágenes de fondo de ojo, es el trabajo realizado por Varadarajan y col., en el cual propusieron un modelo de IA como una solución potencial para detectar de manera confiable la presencia de EMD, el cual entrenaron mediante imágenes de fondo de ojo (6 039 imágenes de 4 035 pacientes), pero utilizando diagnósticos de EMD realizados por médicos expertos, basados en imágenes de OCT. Después de generalizarlo mediante dos conjuntos de validación, el primero de 1 033 imágenes y el segundo de 990 imágenes, se obtuvo una sensibilidad de 85 % y una especificidad de 80 % para el modelo, mientras que la sensibilidad de los especialistas estuvo entre 82 y 85 %, y su especificidad entre 45 % y 50 %. Los autores concluyen que su modelo puede predecir las características de EMD de la OCT, usando imágenes de fondo de ojo, con mayor especificidad que sus especialistas (51).

En el año 2019, Kamram y col. propusieron una red neuronal convolucional especializada en la identificación de enfermedades de la retina, con la cual obtienen muy buenos resultados. Su arquitectura utiliza una herramienta llamada *Atrous convolution* (Figura 8), la cual es una técnica que permite controlar las resoluciones

con las que se computan las imágenes de la red neuronal, es decir, permite calcular las propiedades de la imagen a diferentes resoluciones mediante un radio r que determina el grado de dilatación (52,53).

Su modelo utiliza imágenes de OCT de mácula, y para sus pruebas, utilizan el set de datos de *Kermany y col.* (54), con el cual obtienen una precisión de 99.8 %, sensibilidad del 99.8 % y especificidad del 99.93 %, obteniendo mejores resultados en comparación con otros algoritmos con la misma finalidad.

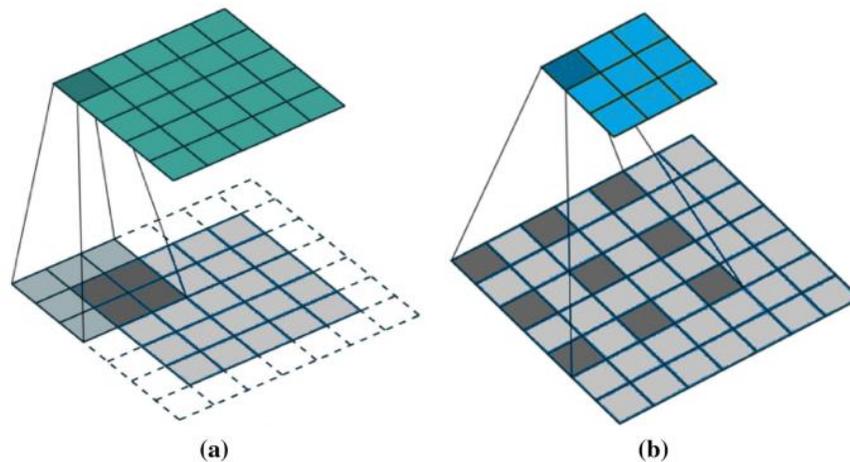


Figura 8. A) Ejemplo de una convolución convencional. B) ejemplo de una convolución atrosa o *atrous convolution*. Como muestra el diagrama, mediante el uso de una convolución atrosa puedes abarcar una extensión mayor de la imagen utilizando un kernel más pequeño. Imagen obtenida de Garg, R. y col. (53)

Estos importantes estudios sustentan el uso de métodos de IA para categorizar y diagnosticar, de manera rápida y con métodos accesibles, al EMD. Además, como lo muestran estos estudios, el uso de análisis de imágenes mediante IA puede presentar una alta precisión en el diagnóstico de EMD, la cual puede ser aumentada utilizando otras variables que se vean afectadas durante la enfermedad. Los estudios de Neda y *col.*, y Hong- He y *col.*, citados previamente, nos muestran que la actividad eléctrica, medida mediante ERG, se ve afectada en pacientes con EMD. Es en este contexto que proponemos la segunda parte de este estudio, que versa sobre el uso de un algoritmo de aprendizaje profundo, de tipo red neuronal convolucional, utilizando imágenes en el dominio de la frecuencia del ERG no evocado, para aumentar la precisión de un modelo predictor de EMD, el cual funciona con imágenes de OCT y fondo de ojo.

Hipótesis

H1: Los cambios observados en el ERG basal pueden servir de base para desarrollar un modelo de diagnóstico predictivo eficiente para el EMD mediante IA.

H2: Los cambios observados en el ERG basal mejoran el desempeño del modelo de diagnóstico predictivo para EMD basado en las imágenes de OCT y/o fondo de ojo.

Objetivos de la investigación

Objetivo general

Comprobar si los cambios en las señales del ERG basal con o sin imágenes de OCT y/o fondo de ojo son informativos para crear o mejorar un modelo de diagnóstico predictivo para el EMD.

Objetivos específicos

1. Comprobar si el algoritmo de aprendizaje profundo, de tipo red neuronal convolucional, es capaz de predecir la presencia de EMD a través del análisis de los datos del ERG basal.
2. Crear un modelo híbrido mediante imágenes de espectros de potencia provenientes del ERG basal combinados con imágenes de OCT y/o fondo de ojo y valorar su desempeño para realizar el diagnóstico predictivo de EMD.

Metodología

El diseño de esta investigación es de tipo observacional analítico.

Todos los experimentos fueron aprobados por el Comité de Bioética del Instituto de Neurobiología (protocolo #74) de la UNAM e IMO (clave NOM-02-ZOO-1999) de acuerdo con las normas y reglamentos de la Sociedad de Neurociencia: Políticas en el uso de Animales y Humanos en la investigación en neurociencia.

Se obtuvo la aprobación del comité de ética de participantes humanos del IMO (referencia: CEI/029-1/2015), del Comité Nacional de Ética (referencia: CONBIOÉTICA-09-CEI-006-20170306), y del comité de investigación de APEC (17 CI 09 003 142). Todos los sujetos dieron su consentimiento informado por escrito. Todos los procedimientos se llevaron a cabo de acuerdo con los principios de la Declaración de Helsinki.

Pacientes

Se cuenta con una base de datos con ERG basal, fondo de ojo y OCT de sujetos metabólicamente sanos, así como pacientes con EMD, del Instituto Mexicano de Oftalmología (IMO) de Querétaro, de la Asociación Para Evitar la Ceguera (APEC) y del Instituto de la Retina del Bajío (INDEREB). Durante la realización de este proyecto, se reclutaron pacientes con EMD en el IMO, con los siguientes criterios de inclusión:

- Hombres y mujeres con edad entre 18 a 80 años
- Pacientes con diagnóstico de EMD con o sin RD
- Pacientes sanos que no cumplan con los criterios de exclusión

Y de exclusión:

- Opacidad en el cristalino
- Menos de 1 mes de haber recibido un tratamiento con fármacos inhibidores de angiogénesis o anticoagulantes
- Pacientes con defectos en la refracción (miopía, hipermetropía)
- Embarazo
- Incapacidad para consentir el estudio
- Oclusión de vena retiniana, cirugía intraocular previa

Toma de ERG

Se tomó ERG no evocado mediante un electroretinógrafo modelo RETIMAX (CSO) o RETeval (LKC Technologies). Previo aseo de la zona del contorno del ojo y de la frente, se colocaron los electrodos de registro y referencia, así como los electrodos activos en los párpados inferiores, para la posterior toma de muestra. La muestra consistió en un registro durante cinco minutos de ERG espontáneo que se tomó en condiciones fotópicas (400 lux).

Fondo de ojo

Para la toma de imágenes de fondo de ojo, el procedimiento consiste en dilatar la pupila del paciente mediante 1 gota de T-P Ofteno (TROPICAMIDA-CLORHIDRATO DE FENILEFRINA, 8 mg/50 mg/ ml), cuidando que el paciente no presente hipersensibilidad a los componentes del midriático (fenilefrina, alfa simpaticomimético y tropicamida, anticolinérgico). Posteriormente, se coloca al paciente frente a la cámara (cámara ZEISS, FM/NA, 60-degree images en el IMO y Visucam® 500 en el APEC) y se le explica la manera en la cual debe colocar su frente y mentón; además, se le indica que debe observar el punto de fijación (punto rojo). Posteriormente, el examinador enfoca las imágenes que sean requeridas y las guarda. La prueba dura unos cuantos minutos.

OCT

Para la toma de imágenes de OCT, la técnica consiste en colocar al paciente frente al tomógrafo (OPTOVUE RTV-1000 en el IMO; Spectralis® Heidelberg Engineering en el APEC), se le indica la parte en la cual va a colocar su frente y mentón y se le indica que debe observar el punto de fijación (punto rojo). Posteriormente, el examinador enfoca las imágenes que sean necesarias y las guarda. Es un procedimiento seguro, por lo cual no representa ningún riesgo para el paciente; además, no requiere dilatación pupilar. La prueba dura unos cuantos minutos.

Modelo

Procesamiento de datos

Para el análisis espectral de los ERGs, los registros de cinco minutos se dividieron en subregistros de un minuto, con la finalidad de aumentar el número de muestras. Estos subregistros se filtraron entre 0.3 y 40 Hz, con una resolución de frecuencia de 0.05 Hz de acuerdo a nuestro estudio previo (34).

En una etapa exploratoria, se optó por probar la capacidad predictiva de las señales procesadas tanto con la transformada rápida de Fourier como con la transformada de Wavelet. Fueron analizados usando scripts personalizados de Python.

Por un lado, se generaron espectrogramas mediante la transformada rápida de Fourier. Los espectrogramas fueron generados a partir de espectros de potencia con la herramienta "Multitaper Spectrogram Code", tal como se describe en (55).

Por otro lado, se aplicó la transformada de wavelet en fragmentos de un minuto, las resoluciones de tiempo y espectral fueron de 0.01 s y 0.05 Hz, respectivamente. Los datos de la transformada de wavelet se representaron como escalogramas. Los escalogramas fueron generados utilizando las librerías Scipy versión 1.7.1 y Matplotlib versión 3.5.1 del paquete Python 3.8.8.

Obtuvimos 1,353 registros de un minuto de ERG basal, con los cuales generamos la misma cantidad tanto de escalogramas como de espectrogramas. Para llevar a cabo la validación de los modelos, la cual se describe más adelante, separamos 164 escalogramas y espectrogramas (con EMD = 42, sin EMD = 122). Las imágenes restantes clasificadas como con EMD ($n = 141$) fueron sometidas a la técnica de aumento de datos modificando la rotación horizontal, el desplazamiento horizontal, el contraste, la variación del brillo y añadiendo ruido en sal y pimienta, tal como se describe en (56). Para el entrenamiento y las pruebas de los modelos, el número final después del aumento de datos, tanto de espectrogramas como de escalogramas fue, con EMD = 846 y sin EMD = 1,048.

Para las imágenes de fondo de ojo, además de las recabadas por nosotros, utilizamos imágenes obtenidas de tres bases de datos públicas diferentes, incluyendo MESSIDOR (57), Pachade et al. (58) y Giancardo et al. (59). De estas bases de datos, utilizamos las imágenes sin EMD ($n = 151$ de MESSIDOR, $n = 180$

de Pachade, $n = 115$ de Giancardo y $n = 495$ de nuestra propia base de datos). Para las imágenes con EMD, utilizamos 1053 imágenes, todas provenientes de nuestra base de datos.

Usamos tres diferentes fuentes de imágenes de OCT. El dataset de Kermany et al. (54) contiene 11,598 imágenes con EMD y 26,565 sin EMD. Únicamente las imágenes contenidas en esta base de datos fueron utilizadas para el entrenamiento y las pruebas de los modelos. La base de datos de Srinivasan et al. (60), de la cual tomamos las imágenes centradas en la mácula, 36 con EMD y 53 sin EMD, y nuestra base de datos compuesta por 176 imágenes con EMD y 964 sin EMD, fueron utilizadas para validar los modelos. Se resume el uso de las diferentes imágenes en la **Tabla 1**.

	FONDO DE OJO			OCT			ESCALOGRAMAS Y ESPECTROGRAMAS		
	Base de datos	EMD	SIN EMD	Base de datos	EMD	SIN EMD	Base de datos	EMD	SIN EMD
Entrenamiento y Pruebas	Presente Estudio	1011	373	Kermany	11598	26565	Presente Estudio	141	1048
	Messidor	0	151	-	-	-	Presente estudio con aumento de datos	846	1048
	Pachade	0	180	-	-	-	-	-	-
	Giancardo	0	115	-	-	-	-	-	-
Validación	Presente Estudio	42	122	Srinivasan	36	53	Presente Estudio	42	122
	-	-	-	Presente Estudio	176	964	-	-	-

Tabla 1. Distribución y composición de los conjuntos de datos utilizados en este trabajo. Es importante remarcar que las imágenes que fueron utilizadas para el entrenamiento y las pruebas no fueron utilizadas para la validación del desempeño de los modelos. Además, la única base de datos en donde se aplicó aumento de datos fue en los escalogramas y espectrogramas utilizados para entrenamiento y pruebas, pero únicamente en la clase con EMD.

Desarrollo y estructura del modelo

Se aplicó un algoritmo de clasificación de tipo Red Neuronal Convolutiva (CNN) en un conjunto de datos humanos mediante el uso de los paquetes Tensorflow versión 2.7.0 y keras versión 2.7.0 de Python 3.8.8. Los detalles del clasificador se pueden encontrar en libros de texto generales de aprendizaje automático. Para las imágenes de espectrogramas, escalogramas y fondo de ojo, se utilizó la técnica de *Aprendizaje transferido* como se describe en (61). Para dicha finalidad, se utilizaron las arquitecturas predefinidas ResNet-50 (62), MobileNet-V2 (63) y DenseNet-201 (64), respetando los hiperparámetros por defecto para cada modelo. Además, a todos los modelos se les aplicó la técnica de *Fine-tuning*, como se describe en (65), añadiendo dos capas LSTM y cuatro capas densas al final del modelo, para estas capas se utilizó la función de activación *ReLU*, como se describe en (66), y capas de *dropout* de 0.2, pues durante la experimentación, fue el valor que mejores resultados generó durante el entrenamiento. Para los tres modelos utilizamos un optimizador Adam, dicha función ajusta los valores de la tasa de aprendizaje automáticamente durante cada época (67), con una tasa de aprendizaje inicial de 0.001, *batch size* de 32 elementos, dicho valor fue seleccionado con base en la literatura previa (66), y función de *Binary Cross Entropy* para el cálculo de la pérdida. En el caso de las imágenes de OCT, se utilizó la CNN OpticNet-71 como se describe en (68), la cual fue entrenada durante veinte épocas sin usar aprendizaje transferido. Los conjuntos de datos finales se dividieron aleatoriamente en conjuntos de entrenamiento (80 % de las observaciones) y de prueba (20 %) para permitir la validación interna de todos los modelos. Se reportan únicamente los modelos con las mejores métricas de desempeño en la etapa de validación, tal como se precisa más adelante.

Modelos híbridos

Para hacer los modelos híbridos, seleccionamos los mejores modelos para cada tipo de variable de entrada. Para OCT fue OpticNet-71, para fondo de ojo fue DenseNet-201 al que se le aplicó *Fine-tuning*, y para las imágenes espectrales de ERG, los mejores resultados se obtuvieron con DenseNet-201 utilizando

escalogramas (**Figura 9**).

Al hacer una predicción con un modelo entrenado, el resultado es una matriz de probabilidades, tal como se describe en la documentación de *Tensorflow* (69). Para realizar los modelos híbridos, tomamos esa matriz de probabilidades para cada tipo de variable de entrada y la multiplicamos por un escalar entre 0 y 1, el cual representa el porcentaje de importancia o de peso que le damos al modelo. Estos valores deben ser exhaustivos, es decir, si el peso del modelo de OCT es 0.5, y el del modelo de fondo de ojo es 0.3, el de escalogramas debe ser necesariamente 0.2, pues con los tres, la suma es igual a 1, o el 100 %. Posteriormente, sumamos las matrices resultantes para cada modelo, obteniendo una nueva matriz de probabilidades, de la cual se obtuvo la predicción final (**Figura 9**).

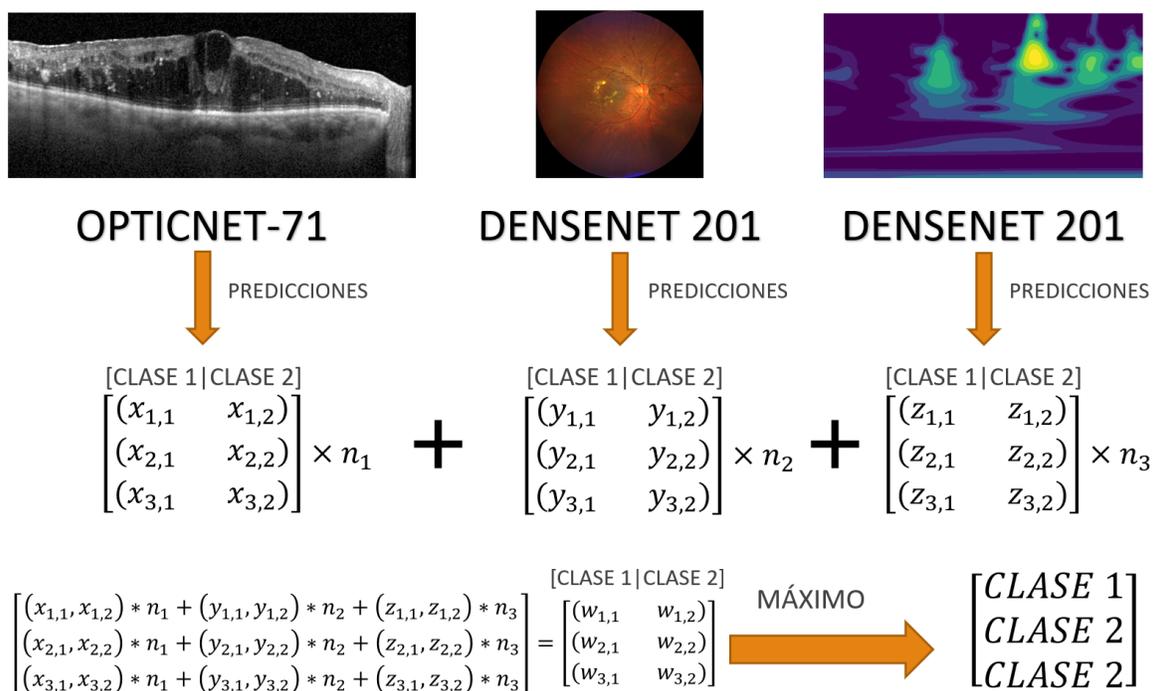


Figura 9. Diagrama que muestra el funcionamiento de los modelos híbridos. Se seleccionaron los mejores modelos durante la validación, para cada dato de entrada. Para OCT fue OpticNet-71 y para fondo de ojo y ERG utilizando escalogramas, DenseNet-201. Se realizaron predicciones con los modelos antes mencionados para cada tipo de entrada (imágenes OCT, fondo de ojo o escalogramas, n = 164 casos en total) y el resultado se multiplicó por una cantidad escalar n_i que representa el porcentaje de ajuste para cada modelo. A continuación, se sumaron los tres valores resultantes para obtener una nueva matriz de predicción, utilizando las predicciones realizadas por cada modelo. El valor máximo se obtiene para cada fila de la matriz, siendo la salida 1 con DME y 2 sin DME.

Para obtener los mejores valores para dichos pesos, se utilizó un script personalizado en Python 3.8, que probara todos los posibles valores para dichos pesos y se seleccionaron aquellos en los cuales las métricas de validación fueran las más altas para el modelo. Dichas métricas se describen a continuación.

Análisis del desempeño de los modelos

Calculamos las diferentes métricas clásicas de rendimiento de la prueba. Las curvas “Receiver-Operating Characteristic” o ROC, se utilizan ampliamente para evaluar la eficacia de los clasificadores binarios y las pruebas de diagnóstico en medicina, machine learning, y otros muchos campos, cuando las clases se distribuyen por igual (70). Construimos la curva ROC trazando la tasa de verdaderos positivos (la razón entre el número de clasificaciones de verdadero positivo, entre el total de positivos o, en este caso, la razón entre el número de casos de enfermedad identificados correctamente y el número total de casos de enfermedad) de nuestro clasificador binario contra la tasa de falsos positivos para todos los posibles umbrales del clasificador. De esta manera, una curva ROC representa el desempeño de una prueba en todos los umbrales posibles. Un método de clasificación binaria que no es mejor que la asignación aleatoria se trazaría cerca de la línea 1:1, mientras que una prueba que es más eficaz que la asignación aleatoria se trazaría por encima de la línea 1:1. El área bajo la curva de la curva ROC (AUC-ROC) cuantifica el rendimiento general de una prueba; una prueba que no sea mejor que la asignación aleatoria correspondería a $AUC-ROC = 0.5$, mientras que las pruebas más precisas tendrían valores de AUC-ROC cercanos a 1. AUC-ROC se utiliza ampliamente para interpretar el rendimiento de un modelo individual, pero también para hacer una comparación entre modelos (70). La precisión, definida como la proporción de predicciones correctas (la relación entre el número de clasificaciones verdaderas y el número total de predicciones), se calcula como una indicación del rendimiento del programa. La sensibilidad corresponde a la tasa positiva verdadera definida anteriormente, y la especificidad o tasa negativa verdadera corresponde a la probabilidad del modelo de realizar una predicción negativa para todo el grupo de observaciones negativas

(70). Se calcula como la relación entre el número de clasificaciones negativas verdaderas y el número total de negativos, o en este caso, la relación entre el número de casos de control / casos sanos correctamente identificados y el número total de casos de control. El valor predictivo positivo (VPP) o precisión, es la proporción de predicciones positivas correctas para el número total de observaciones positivas, mientras que el valor predictivo negativo es la proporción de predicciones negativas correctas para el número total de observaciones negativas (70). También se calculó la Kappa de Cohen (definida como $[(p_o - p_e) / (1 - p_e)]$, donde p_o es la concordancia observada, y p_e es la concordancia esperada (71)) que indica cuánto mejor se está desempeñando el clasificador dado comparado con un clasificador que simplemente adivina al azar de acuerdo con la frecuencia de cada clase. Además, para evaluar el desempeño de modelos que tienen clases desbalanceadas, calculamos la puntuación F1 que se define como un promedio ponderal entre la sensibilidad y la precisión, y es una mejor métrica para evaluar clases desbalanceadas pues normalmente, en estos casos, una de las métricas mencionadas es mucho más elevada que la otra, por lo que su promedio nos indica de mejor manera el rendimiento del modelo (72).

Para evaluar el historial del entrenamiento de los modelos utilizamos la función de pérdida. La función de pérdida en una red neuronal cuantifica la diferencia entre el valor esperado y el valor obtenido por el modelo de aprendizaje automático (73). Las funciones de pérdida basadas en entropía cruzada se usan comúnmente en escenarios de clasificación. La entropía cruzada es una medida de la diferencia entre dos distribuciones de probabilidad (73). La entropía cruzada binaria es apropiada en la configuración de clasificación binaria para obtener uno de dos posibles resultados. La pérdida se calcula de acuerdo con la siguiente fórmula, donde y_i representa el resultado esperado e \hat{y}_i representa el resultado producido por nuestro modelo (73).

$$P\acute{e}rdida = -(y_i \ln(\hat{y}_i) + (1 - y_i) \ln(1 - \hat{y}_i))$$

Resultados

Los resultados son reportados en dos partes. La primera parte es respecto a las fases del entrenamiento y las pruebas de los modelos, y la segunda parte es acerca de la validación y presentación de resultados obtenidos con los mejores modelos probados.

Entrenamiento y pruebas

Para evaluar la efectividad de los modelos de nuestro proyecto, se calcularon las métricas de exactitud y pérdida, como se describe en la metodología. Ambos parámetros se calculan mediante la función fit de Keras, en Python 3.8, con objeto de obtener los mejores modelos para una posterior validación. Los modelos se entrenaron por el número de épocas que maximiza la exactitud y reduce la pérdida. Los resultados se resumen en la **Tabla 2**.

Todos los modelos tienen métricas de exactitud por arriba de 0.85, tanto en entrenamiento como en pruebas. Además, en ninguno de los modelos la pérdida supera el 0.37, ni en entrenamiento ni en pruebas. Obtener estas métricas es importante, pues indica el riesgo que tienen los modelos de presentar sobreajuste de los datos durante el entrenamiento (74). Interesantemente, se puede observar que los modelos con las exactitudes más bajas y pérdidas más altas durante el entrenamiento, son los modelos de imágenes espectrales de ERG, tanto para escalogramas como para espectrogramas, indicando un riesgo de sobreajuste más elevado.

Validación

Modelos de OCT

Los modelos de OCT se validaron con dos bases de datos diferentes, tal como se resume en la metodología (**Tabla 1**). Para la validación interna con $n = 1229$, el modelo OpticNet-71 obtuvo mejores resultados (precisión: 0.93, sensibilidad: 0.93, especificidad: 0.99, puntuación F1: 0.93, exactitud: 0.98, AUC: 0.96 y Kappa de Cohen: 0.91) en comparación con DenseNet-201 (precisión: 0.52, sensibilidad:

0.91, especificidad: 0.83, puntuación F1: 0.72, exactitud: 0.84, AUC: 0.87 y Kappa de Cohen: 0.57).

Para la validación externa con $n = 164$, OpticNet-71 obtuvo, en general, mejores resultados (precisión: 0.85, sensibilidad: 0.93, especificidad: 0.94, puntuación F1: 0.89, exactitud: 0.984 AUC: 0.94 y Kappa de Cohen: 0.84) con todas las métricas iguales o por arriba de 0.84, y aunque la precisión de DenseNet-201 fue superior en esta validación, todas las demás métricas fueron menores en comparación con el primer modelo mencionado (precisión: 0.63, sensibilidad: 0.95, especificidad: 0.80, puntuación F1: 0.75, exactitud: 0.84, AUC: 0.88 y Kappa de Cohen: 0.64). Los resultados de ambas validaciones con OCT se resumen en la **Figura 10**.

			Entrenamiento		Prueba	
Modelo	Tipo de dato de entrada	# de épocas	Exactitud	Pérdida	Exactitud	Pérdida
OpticNet-71	OCT	20	0.98	0.05	0.99	0.01
DenseNet-201 Fine-Tuning	OCT	100	0.97	0.06	0.94	0.18
ResNet-50	Fondo de Ojo	50	1.00	0.01	0.99	0.03
MobileNet-V2 Fine-Tuning	Fondo de Ojo	100	0.96	0.10	0.99	0.03
DenseNet-201 Fine-Tuning	Fondo de Ojo	100	0.96	0.10	0.96	0.05
ResNet-50 (FFT)	Espectrograma	50	0.94	0.18	0.93	0.16
DenseNet-201 Fine-Tuning	Espectrograma	200	0.84	0.37	0.94	0.18
ResNet-50	Escalograma	50	0.98	0.10	0.96	0.16
DenseNet-201 Fine-Tuning	Escalograma	200	0.88	0.27	0.95	0.15

Tabla 2. Tabla que resume las métricas obtenidas durante el entrenamiento y las pruebas con los modelos entrenados para cada tipo de dato (OCT, fondo de ojo, escalogramas y espectrogramas). Los modelos de fondo de ojo y OCT obtienen muy buenas métricas durante el entrenamiento y las pruebas. En comparación con estos últimos, los modelos de imágenes espectrales de ERG no obtienen los mismos resultados, indicando un posible sobreajuste de los datos de entrenamiento por la posible poca cantidad de datos.

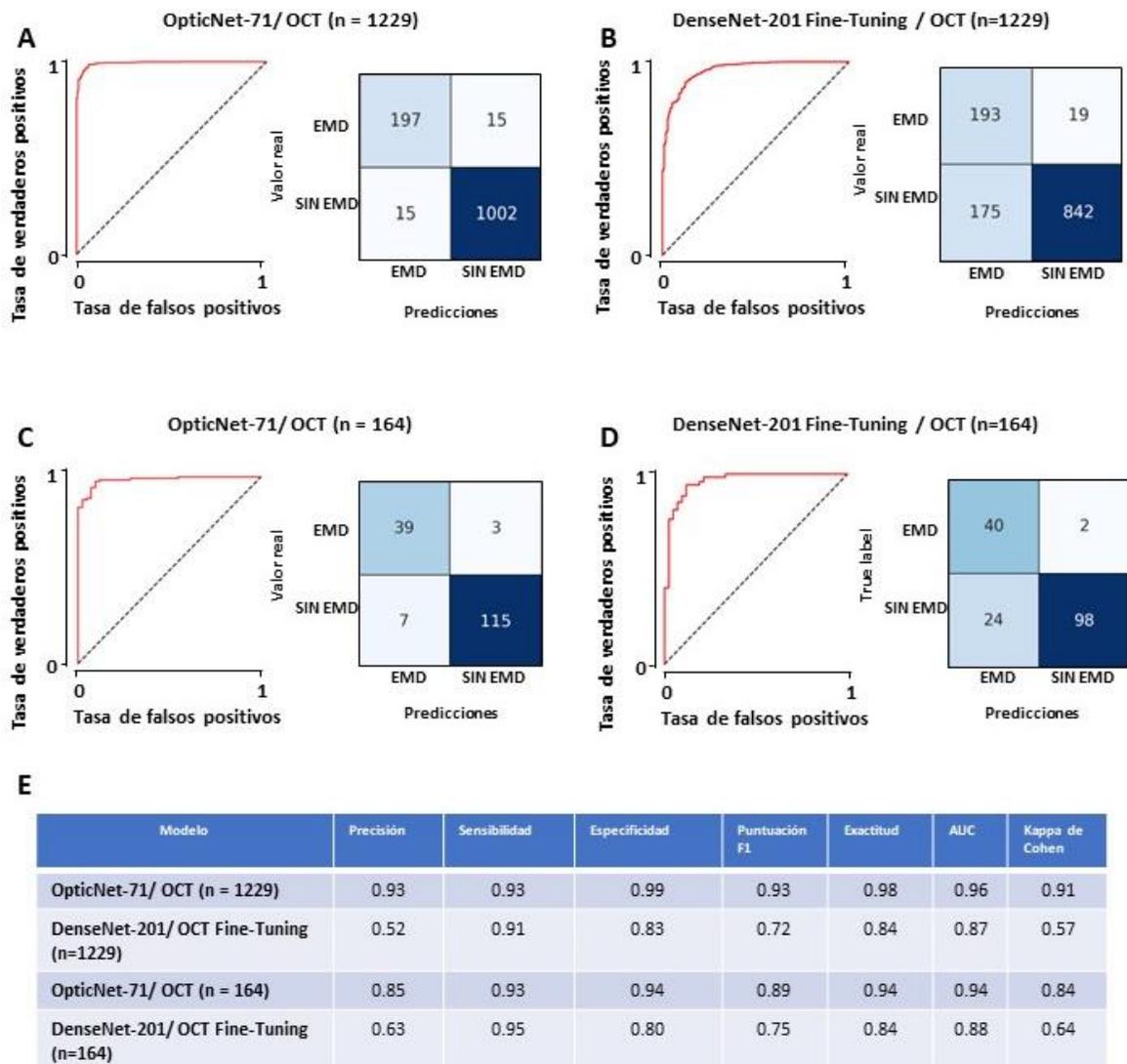


Figura 10. Desempeño de todos los modelos de OCT. Se muestran las curvas ROC y las matrices de confusión de **(A)** Modelo OpticNet-71 validado con n = 1229 (n= 212 con EMD y n = 1017 sin EMD), **(B)** Modelo DenseNet-201 validado con n = 1229 (n= 212 con EMD y n = 1017 sin EMD), **(C)** Modelo OpticNet-71 validado con n = 164 (n= 42 con EMD y n = 122 sin EMD) y **(D)** Modelo DenseNet-201 validado con n = 164 (n= 42 con EMD y n = 122 sin EMD). **(E)** Resumen de las métricas obtenidas por todos los modelos. AUC, indica el área bajo la curva ROC (por sus siglas en inglés), n indica el número total de datos para la validación.

Modelos de fondo de ojo

Para la validación de los modelos de fondo de ojo, reportamos los resultados de los tres modelos que obtuvieron los mejores resultados durante la validación, para la cual se utilizó una n = 164 (con EMD: 42, sin EMD: 122). Los resultados obtenidos fueron, para ResNet-50: precisión: 0.63, sensibilidad: 0.86, especificidad: 0.83,

puntuación F1: 0.73, exactitud: 0.83, AUC: 0.84 y Kappa de Cohen: 0.11; para MobileNet-V2: precisión: 0.89, sensibilidad: 0.60, especificidad: 0.98, puntuación F1: 0.71, exactitud: 0.88, AUC: 0.78 y Kappa de Cohen: 0.64; y para DenseNet-201: precisión: 0.82, sensibilidad: 0.88, especificidad: 0.93, puntuación F1: 0.85, exactitud: 0.92, AUC: 0.91 y Kappa de Cohen: 0.80.

Interesantemente, el modelo ResNet-50 tiene una alta sensibilidad y el modelo MobileNet-V2 tiene una muy alta especificidad, pero es DenseNet-201 el que obtiene métricas más robustas, especialmente en las métricas que sirven mejor para evaluar las clases desbalanceadas, que son la puntuación F1 y Kappa de Cohen (72), obteniendo valores más altos en ambas, en comparación con los primeros dos modelos mencionados. Estos resultados se resumen en la **Figura 11**.

Modelos de ERG

Para la validación de los modelos de imágenes espectrales de ERG basal (espectrogramas y escalogramas), se utilizó una $n = 164$ (DME: 42, No DME: 122) como se resume en la **Tabla 1**. Los resultados de la validación fueron, *Resnet-50 con espectrogramas*: precisión: 0.33, sensibilidad: 0.28, especificidad: 0.80, puntuación F1: 0.31, exactitud: 0.67, AUC: 0.58 y Kappa de Cohen: 0.09; *DesneNet-201 con espectrogramas*: precisión: 0.27, sensibilidad: 0.26, especificidad: 0.75, puntuación F1: 0.26, exactitud: 0.63, AUC: 0.56 y Kappa de Cohen: 0.02; *ResNet-50 con escalogramas*: precisión: 0.38, sensibilidad: 0.36, especificidad: 0.80, puntuación F1: 0.37, exactitud: 0.68, AUC: 0.57 y Kappa de Cohen: 0.15; y *DenseNet-201 con escalogramas*: precisión: 0.34, sensibilidad: 0.40, especificidad: 0.73, puntuación F1: 0.37, exactitud: 0.65, AUC: 0.57 y Kappa de Cohen: 0.13. Estos resultados se resumen en la **Figura 12**.

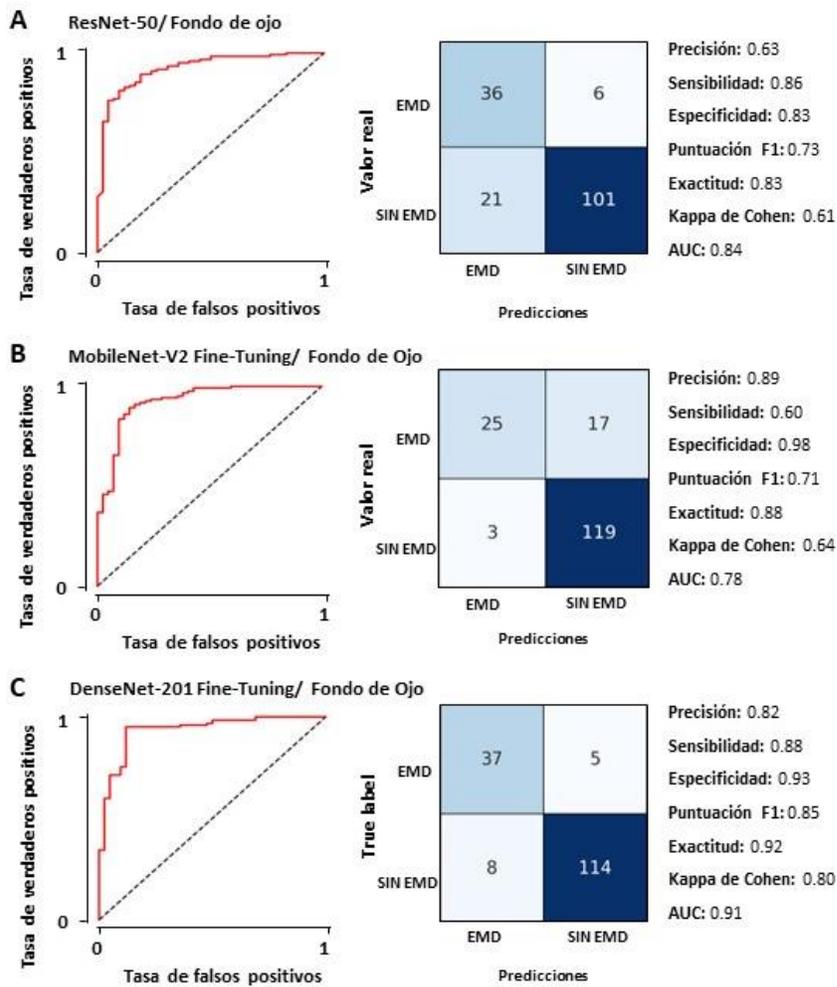


Figura 11. Rendimiento de los modelos de fondo de ojo durante la validación. Curvas ROC, matrices de confusión y métricas obtenidas por el modelo **(A)** ResNet-50, **(B)** MobileNet-V2 y **(C)** DenseNet-201. La validación de todos los modelos incluye un total de 164 imágenes de fondo de ojo ($n = 42$ imágenes con EMD, y $n = 122$ imágenes sin EMD).

Modelos híbridos

Para los modelos híbridos, se seleccionaron los modelos que tuvieran las métricas más elevadas para cada tipo de dato durante la validación. Para OCT, se seleccionó OpticNet-71, para fondo de ojo fue DenseNet-201, y para ERG se eligió DenseNet-201 entrenado con escalogramas.

Para el modelo híbrido de OCT combinado con fondo de ojo, el modelo alcanza sus mejores métricas cuando el valor de n_1 (descrito en la **Figura 9**) toma un valor de 0.60, y el valor de n_2 toma un valor de 0.40, tal como se describe en el panel A de la **Figura 13** (precisión: 0.90, sensibilidad: 0.90, especificidad: 0.97, puntuación F1: 0.90, exactitud: 0.95, AUC: 0.94 y Kappa de Cohen: 0.87). Es interesante ver cómo éste fue el modelo con las métricas más altas durante la validación con $n = 164$,

superando incluso las métricas durante la validación del antes mejor modelo OpticNet-71 con OCT.

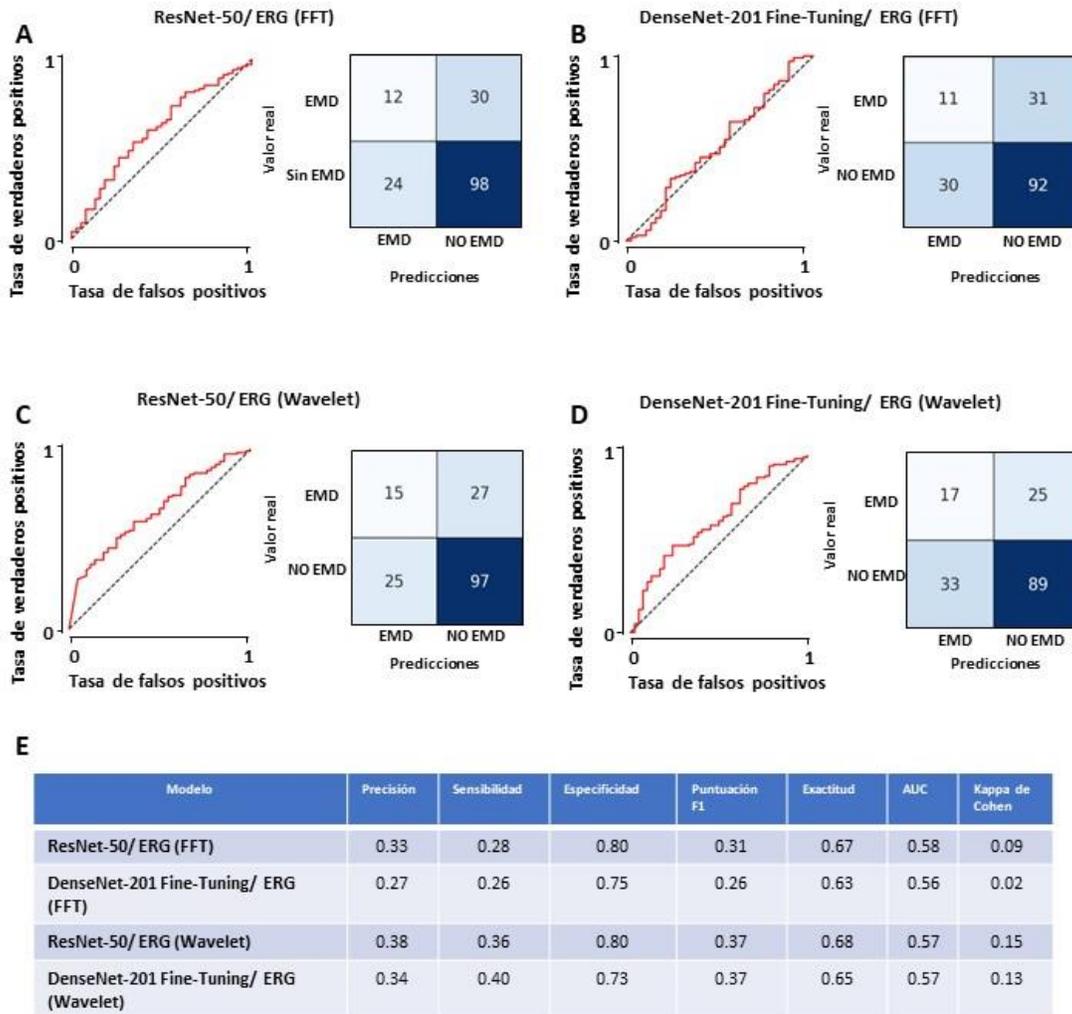


Figura 12. Desempeño de todos los modelos basados en ERG basal. Se muestran las curvas ROC y las matrices de confusión de **(A)** Modelo ResNet-50 alimentado con espectrogramas, **(B)** DenseNet-201 alimentado con espectrogramas, **(C)** ResNet-50 alimentado con escalogramas y **(D)** DenseNet-201 alimentado con escalogramas. Los datos de validación incluyen n = 42 casos con EMD y n = 122 casos sin EMD. **(E)** Resumen de todas las métricas obtenidas durante la validación. AUC, área bajo la curva ROC.

Ninguno de los modelos muestra métricas por arriba de 0.80, siendo en todos los casos la especificidad del modelo la métrica más alta reportada. Es importante destacar que, de los cuatro modelos validados, el que más alta sensibilidad

presenta es DenseNet-201 con escalogramas; es por esta razón que se utilizó para los modelos híbridos.

Para el modelo híbrido de OCT combinado con escalogramas, el modelo se optimiza cuando el valor de n_i para OCT es 0.60 y para ERG es de 0.40, como se muestra en el panel B de la **Figura 13**, (precisión: 0.88, sensibilidad: 0.90, especificidad: 0.96, puntuación F1: 0.89, exactitud: 0.95, AUC: 0.93 y Kappa de Cohen: 0.85). Aunque por sí solo el modelo de OCT es el mejor modelo que utiliza un solo tipo de dato, cuando se combina con ERG con los valores de n_i descritos anteriormente, se obtienen mejores métricas de especificidad, precisión, Kappa de Cohen y exactitud.

Para el modelo híbrido de fondo de ojo combinado con ERG, el modelo se optimiza en una franja muy amplia, entre la que se encuentran los valores de $n_1 = 0.70$ para fondo de ojo y $n_2 = 0.30$ para ERG, (precisión: 0.84, sensibilidad: 0.88, especificidad: 0.94, puntuación F1: 0.86, exactitud: 0.93, AUC: 0.91 y Kappa de Cohen: 0.81). Es interesante observar que, como se describió anteriormente, el modelo DenseNet-201 de fondo de ojo tiene métricas por arriba de 0.80, pero cuando se combina con el modelo de ERG con los valores de n_i descritos, todas las métricas, con excepción de la sensibilidad, aumentan su valor, sin disminuir esta última métrica mencionada.

Por último, para el modelo híbrido que combina todos los datos, se observan diferentes franjas de optimización; la que más llama notoria se presenta cuando $n_1 = 0.60$, $n_2 = 0.40$ y $n_3 = 0$, modelo que se describió previamente. Sin embargo, existen otras franjas de optimización, por ejemplo, cuando $n_1 = 0.60$, $n_2 = 0.28$ y $n_3 = 0.12$, (precisión: 0.88, sensibilidad: 0.90, especificidad: 0.96, puntuación F1: 0.89, exactitud: 0.95, AUC: 0.93 y Kappa de Cohen: 0.85). Todos los resultados, de los modelos individuales y los modelos híbridos se resumen en la **Figura 14**.

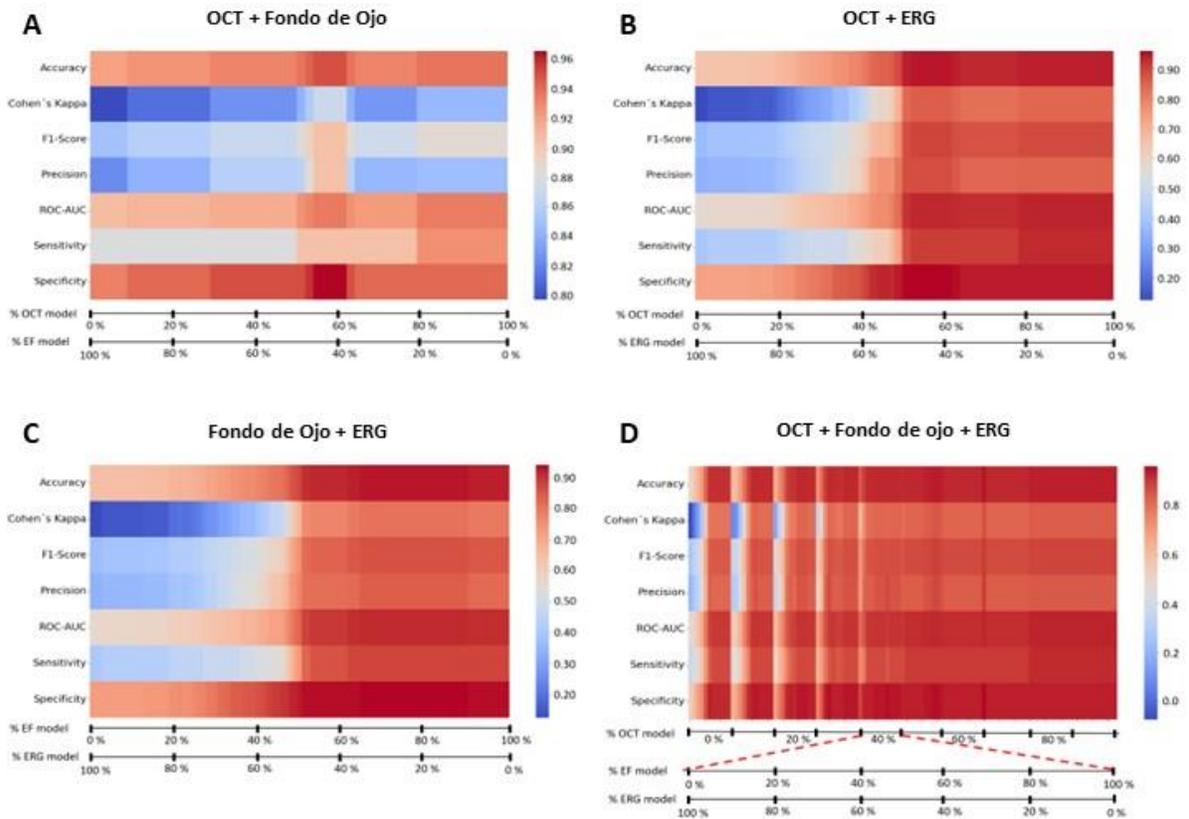


Figura 13. Optimización de los modelos híbridos. Los mapas de calor muestran las métricas de rendimiento de los modelos híbridos dobles y triples según la proporción asignada a cada modelo utilizado. **(A)** OCT e imágenes de fondo de ojo, **(B)** OCT e imágenes de escalogramas provenientes de la señal de ERG basal, **(C)** Fondo de ojo e imágenes de escalogramas y **(D)** los tres tipos diferentes de datos. En los ejes X están los porcentajes asignados para cada modelo, estos ejes son complementarios ya que todas las combinaciones suman 100 %. Para el modelo híbrido triple (D), el porcentaje permanece en porciones de 10 % para OCT, y el complemento está dado por los otros dos modelos. Esta metodología busca los valores óptimos de n_1 , n_2 y n_3 como se presentó en la **Figura 9**.

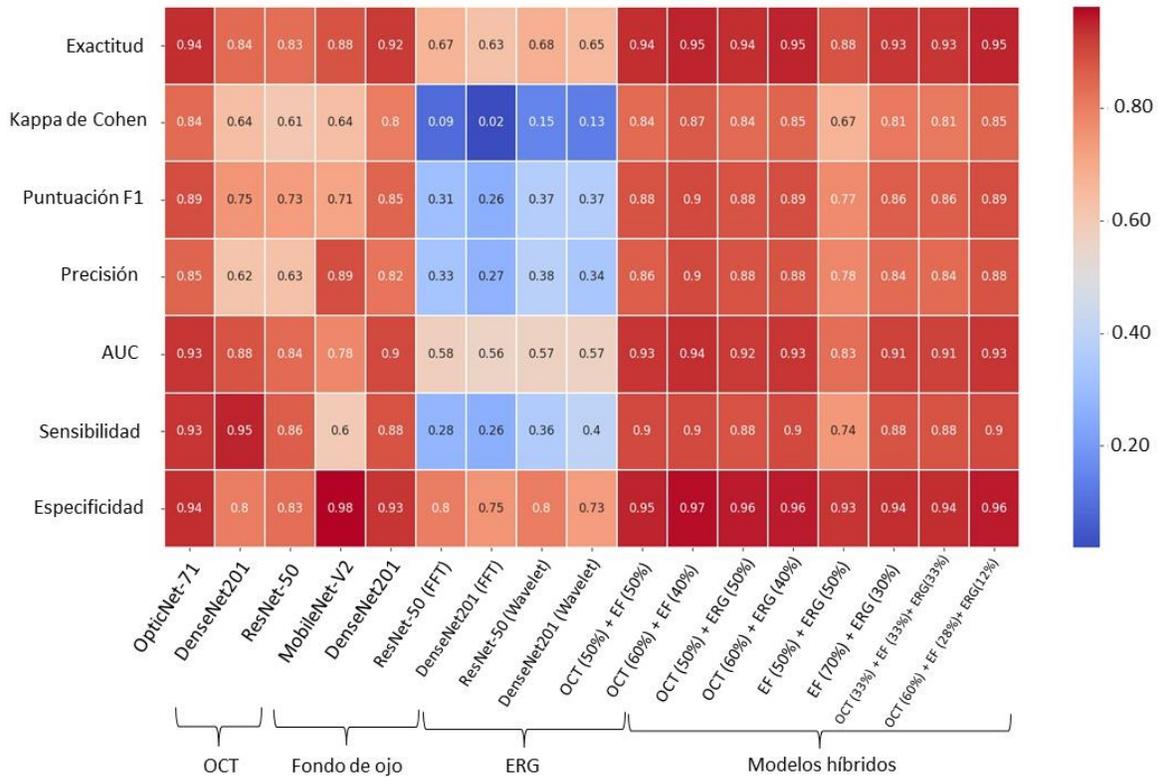


Figura 14. Comparación de todos los modelos utilizados para el diagnóstico predictivo de EMD. El mapa de calor muestra las métricas de rendimiento obtenidas para cada modelo probado, incluidos los modelos individuales o basados en una sola entrada (OCT, fondo de ojo e imágenes espectrales derivadas de ERG basal), así como los modelos híbridos más representativos. n = 42 para EMD y n = 122 para muestras sin EMD.

Discusión

A la fecha, el EMD sigue siendo la principal causa de ceguera prevenible en la población en edad económicamente activa en México y en el mundo (3). Los programas de detección oportuna son extremadamente importantes en el tratamiento de esta complicación de la RD. La introducción de nuevas modalidades de imagen y los avances tecnológicos han facilitado tanto la detección temprana como el seguimiento de pacientes con EMD, en particular la angiografía por OCT y la inteligencia artificial. No obstante, el OCT no es un examen accesible para todos en nuestro país por su costo elevado y la falta de equipos.

Este trabajo está situado en este contexto y en torno a dos preguntas específicas de investigación. La primera cuestiona lo siguiente: ¿Se puede generar un modelo de inteligencia artificial de tipo red neuronal convolucional que utilice imágenes de los cambios en el espectro de la potencia de las señales oscilatorias de ERG basal para diagnosticar EMD? Los resultados muestran que, utilizando esta metodología, no es posible establecer un diagnóstico confiable entre pacientes con EMD de aquellos que no lo tienen, pues las métricas obtenidas por los modelos con imágenes espectrales no son robustas ya que se obtuvieron valores por debajo de 0.80 para todas las métricas calculadas. La explicación para la obtención de estos resultados se puede abordar desde dos perspectivas: la primera es debido al estado de la enfermedad y la segunda debido a la naturaleza de los datos y características de los modelos utilizados.

Para abordar la primera perspectiva del problema, hay que recordar que el EMD puede desarrollarse asociado a diferentes grados de retinopatía diabética (RD), que pueden ir desde una RD no proliferativa moderada hasta una RD proliferativa avanzada (9). Además, se ha observado que la RD, clásicamente definida como una microangiopatía que afecta a los vasos retinianos (75), produce cambios que afectan a las neuronas de la retina, como la apoptosis celular, sin presencia necesaria de cambios vasculares (76). La apoptosis neuronal se inicia en las células ganglionares de la capa interna de la retina, pero puede afectar también a otras células nerviosas retinianas, como son las células bipolares, amacrinas y los fotorreceptores (76). Este daño neuronal en la RD se puede demostrar en las

pruebas funcionales como el ERG (77). Por ejemplo, hay evidencia de que el ERG multifocal puede reflejar anomalías en la UNV, traducidas en retrasos en la latencia de este estudio, entre sujetos con y sin RD (78).

En los datos que utilizamos para el entrenamiento, las pruebas y la validación de los modelos, hay registros de ERG de pacientes con diagnóstico de RD en algún grado, para los dos grupos definidos, a saber, sin EMD y con EMD. Debido a que existen cambios neuronales desde la etapa de RD, antes de desarrollar EMD, consideramos que una explicación probable para los resultados obtenidos es que las imágenes espectrales están mostrando cambios similares en ambos grupos, ocasionando la confusión de los modelos. Para verificar que esta explicación tiene sentido, las imágenes de escalogramas de Wavelet que fueron utilizadas para el entrenamiento y las pruebas durante los experimentos de este proyecto, se reclasificaron en imágenes con algún grado de RD ($n = 293$) e imágenes sin RD ($n = 896$). Para los escalogramas con RD, se aplicó aumento de datos variando los mismos parámetros descritos previamente en la metodología de este proyecto. Con el aumento de datos, se obtuvieron 879 escalogramas con RD y 896 sin RD. Los datos se dividieron aleatoriamente en entrenamiento (80 %) y pruebas (20 %) y se entrenaron los modelos ResNet-50 y DenseNet-201, con el fin de comparar los resultados entre la clasificación binaria hecha de pacientes con EMD y sin EMD, de la hecha con imágenes con RD y sin RD. Los modelos fueron entrenados durante la misma cantidad de épocas que los modelos descritos previamente para la realización de este proyecto.

Para la validación se utilizaron las mismas 164 imágenes utilizadas para la validación explicada previamente, pero estas imágenes fueron recategorizadas en aquellas con RD ($n = 77$) y sin RD ($n = 87$). Se calcularon las mismas métricas descritas en la metodología, y los resultados se muestran en la **Tabla 3**.

MÉTRICAS	MODELOS			
	Con EMD vs Sin EMD		Con RD vs Sin RD	
	ResNet-50	DenseNet-201	ResNet-50	DenseNet-201
Precisión	0.38	0.34	0.67	0.68
Exactitud	0.68	0.65	0.65	0.67
Sensibilidad	0.36	0.40	0.66	0.70
Especificidad	0.80	0.73	0.63	0.64
Puntuación –F1	0.37	0.37	0.67	0.69
Kappa de Cohen	0.15	0.13	0.30	0.34

Tabla 3. Tabla comparativa entre las validaciones de las clases de EMD y las de RD. En términos generales, las métricas obtenidas en la validación con los mismos modelos, entrenados durante las mismas épocas son mejores en la validación de las clases de RD. Todas las métricas son superiores a 0.60 y las Kappa de Cohen son mejores en esta validación.

Los resultados de este experimento muestran que las métricas obtenidas durante la validación hecha con las clases de RD son mejores utilizando los mismos modelos, todas las métricas superan el valor de 0.60, a diferencia de los valores obtenidos en la validación de las clases de EMD. También vemos que los valores de las Kappa de Cohen son mejores en la validación de las clases de RD. Estos resultados apoyan nuestro argumento en torno a la perspectiva del estado de la enfermedad. Aun así, los modelos entrenados para las clases con RD y sin RD, aunque son mejores, no son del todo robustos para hacer una clasificación de suficiente calidad para uso clínico. Pareciera que hay otras variables que están influyendo en la clasificación, es aquí en donde abordaremos la segunda perspectiva, la cual versa sobre la naturaleza de los datos y los modelos.

Previamente, se han reportado trabajos similares en los que utilizan señales electrofisiológicas, como el EEG, para generar imágenes espectrales y entrenar CNNs (27,79), pero en ningún trabajo previo se describe el uso de imágenes espectrales obtenidas de ERG basal de humanos. Eltvik, A. (79) describe un procedimiento similar al que se utilizó en este proyecto, a saber, que mediante la transformada rápida de Fourier, generó espectrogramas y mediante transformada de Wavelet generó escalogramas a partir de EEG y las utilizó para entrenar CNNs,

y así generar, posteriormente, una clasificación que consistía en decidir en qué señal él había añadido datos ajenos a la señal original (79). Sus resultados no superan el 75 % de las métricas obtenidas con espectrogramas y 60 % con escalogramas.

Existen dos razones por las que los resultados no son favorables, tanto en nuestro trabajo como en lo reportado por Eltvik, A. (79). El primer motivo es que la calidad de los datos es baja. Esto puede suceder debido a que la aplicación de las transformadas de Wavelet o de Fourier sobre los registros de la señal electrofisiológica, produce representaciones tiempo-frecuencia que no expresan la información necesaria o suficiente (79), aunque la frecuencia de muestreo es de 2 kHz. Al convertir la señal original en una imagen espectral, algunos datos podrían perderse. En algunos experimentos, los espectrogramas han mostrado baja resolución espacial, lo que puede perjudicar el desempeño de los modelos (80).

El segundo motivo es que la cantidad de datos es muy baja. Las CNNs generalmente utilizan una gran cantidad de imágenes para el entrenamiento, como las redes ResNet-50 o MobileNet, que utilizan el banco de datos de ImageNet, el cual contiene cerca de 14 millones de imágenes (79). Una CNN basada en imágenes espectrales es eficaz en muchos dominios de aplicación; sin embargo, la eficacia de este método depende de conjuntos de datos grandes y exhaustivos para el entrenamiento de dicha CNN (81). Para el entrenamiento y las pruebas, nosotros contamos con poco más de 1,000 imágenes, sin aumento de datos, por lo que es una limitación del estudio contar con tan pocos datos de entrenamiento.

Para contender con el problema de falta de datos, utilizamos la técnica de aprendizaje transferido, la cual se describe en (61). Pero existen trabajos como el de Ruffini, G. y *col.* (82), en el que, mediante la señal de EEG, hacen una clasificación de imágenes de espectrogramas entre pacientes con trastorno de la conducta de movimientos oculares rápidos (RBD por sus siglas en inglés) y controles sanos. Ellos utilizan una CNN simple de 5 capas convolucionales, entrenada con sus datos, y mencionan que no utilizan la técnica de aprendizaje transferido pues, el entrenamiento previo de una CNN con imágenes diferentes a los espectrogramas, para este tipo de tareas, inicializaría los pesos para detectar

características de objetos que no están presentes en los espectrogramas (82), lo cual representaría la última razón del por qué nuestra metodología no fue eficiente para generar un diagnóstico de EMD mediante imágenes espectrales.

La siguiente parte de la discusión versa sobre darle respuesta a la segunda pregunta específica de investigación, la cual cuestiona lo siguiente: ¿Se puede mejorar el desempeño de los clasificadores basados en imágenes de fondo de ojo y/o OCT para predecir el EMD combinándolos con las señales antes descritas?

Con los resultados reportados, podemos decir que sí, se puede mejorar el desempeño, ya que las métricas de especificidad, precisión, Kappa de Cohen y exactitud del mejor modelo de OCT durante la validación con 164 sujetos, mejoraron cuando se les combinó con el mejor modelo de ERG, aunque la sensibilidad disminuyó. No así en el mejor modelo de fondo de ojo, en donde todas las métricas aumentaron cuando se combinó con el mejor modelo de ERG, con excepción de la sensibilidad, aunque esta no disminuyó.

Creemos que esta mejoría en los modelos se puede deber a dos factores: la naturaleza de los datos y la capacidad de los modelos combinados.

Para abordar la primera posible explicación, cabe recordar que las imágenes para diagnóstico se pueden dividir en dos tipos: técnicas de imagenología estructural y funcional (83). Las imágenes estructurales ofrecen una visión estática de la estructura estudiada y son de mucha utilidad a la hora de encontrar la localización de alguna lesión, así como los efectos de una enfermedad (83). Son el tipo de estudio más empleado en el ámbito de la salud; algunos ejemplos de este tipo de imágenes en nuestro contexto son la OCT y las imágenes de fondo de ojo.

Por otro lado, tenemos las pruebas funcionales. Estas se emplean frecuentemente en la investigación, pues proporcionan una visión del funcionamiento de la estructura que queremos estudiar (83). Concretamente, los estudios funcionales consisten en la exposición del sujeto a un determinado estímulo y la observación de su rendimiento, así como del estudio de la actividad basal de dicha estructura. Todo esto permite obtener una medida fisiológica del órgano o estructura que estemos estudiando (83). Estos estudios tienen un enorme potencial en el campo

de la ciencia, pues aportan información directa del funcionamiento del tejido estudiado. En nuestro contexto, los escalogramas y espectrogramas obtenidos mediante ERG basal son parte de este ejemplo, ya que son imágenes funcionales de la retina. Cabe resaltar que, en este trabajo, no fueron consideradas imágenes funcionales derivadas del ERG evocado por flash de luz (84), que es la modalidad clásicamente usada para diagnóstico de enfermedades oculares, sino las imágenes provenientes del ERG espontáneo, ya que se ha reportado recientemente su uso para detectar factores de riesgo de la diabetes tipo 1 (32). Se prefirió esta modalidad dado que prescinde del periodo de adaptación a la luz que se suele requerir en varios protocolos de ERG evocado para diagnóstico de retinopatías asociadas a la diabetes (84), su adquisición es rápida (5-10 minutos) y con algunos aparatos de ERG, no ocupa dilatación pupilar (85).

Contar con información, tanto estructural como funcional de la retina, mejora la capacidad del modelo, pues el modelo cuenta con un panorama más integral para hacer la clasificación.

Para abordar el segundo punto de esta parte de la discusión, es necesario saber que, en el campo del aprendizaje automático, los métodos combinados utilizan múltiples algoritmos de aprendizaje para obtener un rendimiento predictivo que mejore el que podría obtenerse por medio de cualquiera de los algoritmos de aprendizaje individuales que lo constituyen (86).

Hay muchas maneras de construir modelos híbridos o combinados (87). Estudios previos han utilizado esta metodología, aunque en áreas diferentes al contexto de este proyecto. Por ejemplo, Akhtar y col. (88) construyeron un modelo híbrido que es altamente eficaz para el análisis de sentimientos en lenguas de escasos recursos. Utilizaron CNNs en combinación con máquinas de vectores de soporte (SVM por sus siglas en inglés), para la clasificación de sentimientos en texto. El modelo se probó en cuatro conjuntos de datos, y en todas las validaciones el modelo combinado obtuvo mejores resultados que los modelos por separado (88). Otro estudio en este contexto es presentado por Dang y col. (87), quienes usan tres tipos de modelos, SVM, LSTM y CNN, haciendo varias combinaciones entre ellos para el análisis de sentimientos a través de grupos de datos públicos. Concluyen

que los modelos híbridos incrementan la exactitud para el análisis de sentimientos comparados con los modelos individuales en todas las bases de datos utilizadas. Una aproximación más cercana a este trabajo se presentó por Xue y *col.* (89), quienes utilizaron un modelo combinado de CNN y SVM para el estudio de extracción de características y clasificación de tipos morfológicos microvasculares para la ayuda a la detección del cáncer de esófago. Sus resultados mostraron que la combinación de los modelos CNN y SVM supera a los modelos de características tradicionales (89).

Como vemos en estos estudios, en general, los modelos híbridos proveen mejores resultados que los modelos individuales; además, se explica que, empíricamente, se ha comprobado que cuando existe una diversidad significativa entre los modelos individuales, las combinaciones tienden a obtener mejores resultados (86,87). Esta diversidad se puede lograr mediante los datos de entrada que toman los modelos (88). Dang y *col.* (87) explican que sus buenos resultados pueden deberse a estas diferencias en los tipos de datos, pues mientras un algoritmo obtiene información de la duración de las publicaciones, otro obtiene información del contenido (87).

Con esta información, en nuestro contexto, podemos decir que nuestros modelos híbridos fueron mejores puesto que, en general, los modelos combinados en el estado del arte han mostrado mejores resultados que los modelos singulares, debido a la combinación de características y a la variabilidad presente en los datos; es decir, el uso de diferentes tipos de imágenes de retina, tanto funcionales como estructurales, aunado a las características individuales de las CNN utilizadas, aumentan la efectividad de nuestros modelos combinados.

Por lo planteado anteriormente surge una pregunta en cuanto a la comparación del modelo híbrido que utiliza OCT (60 %) en combinación con ERG (40 %) contra del modelo que utiliza OCT (60 %), fondo de ojo (28 %) y ERG (12 %), ya que, en este segundo modelo se están utilizando los tres tipos de datos, es decir, el modelo cuenta con más información para realizar una predicción, entonces ¿por qué ambos modelos obtienen los mismos valores en todas las métricas calculadas?

Una posible respuesta es que, la certidumbre del modelo de OCT es la mayor de entre los tres modelos (**Figura 13**), que, como se mencionó previamente, se puede

inferir por las métricas obtenidas en el entrenamiento y las pruebas, y por lo mismo, sabemos que la certidumbre del modelo de ERG es la más baja, siendo la del modelo de fondo de ojo un punto medio entre las dos. Sabiendo esto, podemos pensar que la certidumbre del modelo de fondo de ojo no es tan grande como para tener un efecto positivo al hacer la predicción, y es superada por la baja certidumbre del modelo de ERG. Para verificar esto calculamos el valor absoluto de la diferencia entre cada una de las probabilidades por cada una de las imágenes de la validación ($n = 164$), así obtenemos un valor que nos indica la distancia que existe entre la probabilidad de que esa imagen pertenezca a la clase con EMD y la probabilidad de que pertenezca a la otra clase. Posteriormente, calculamos las medias aritméticas para esos valores por cada modelo. Para los modelos con más certidumbre, esperamos que ese valor se aproxime a 1, por otro lado, para los modelos con baja certidumbre este valor será próximo a 0.5. Obtuvimos para el modelo de OCT una media de 0.951, para el modelo de fondo de ojo una media de 0.94 y para el modelo de ERG una media de 0.74. Es decir, la baja certidumbre del modelo de ERG anula el posible beneficio del modelo de fondo de ojo, y la alta certidumbre del modelo de OCT no se ve beneficiada por el de fondo de ojo, pues este último tiene valores más bajos.

Otro resultado interesante a discutir es con respecto a las métricas obtenidas por los modelos durante el entrenamiento y las pruebas. En términos generales, las métricas de exactitud y pérdida nos sirven para evaluar el rendimiento del modelo durante estas dos fases (74). A grandes rasgos, los resultados fueron bastante buenos para los modelos de OCT y fondo de ojo. Para los modelos de OCT los buenos resultados se deben, en gran parte, a la cantidad de imágenes con las que se contó para el entrenamiento, pues como se reporta en muchos trabajos, una manera de mejorar las métricas durante el entrenamiento es aumentar el número de muestras (74,87). Por otro lado, para fondo de ojo, todas las métricas de exactitud durante el entrenamiento y las pruebas fueron mayores a 0.95, lo que habla de un buen rendimiento de los modelos, y también se obtienen buenos resultados en la pérdida, nunca superando el 0.1. Aunque no se contó con la misma cantidad de datos como con las imágenes de OCT, las imágenes de fondo de ojo

eran de buena calidad, además que el rendimiento se vio beneficiado por el aprendizaje transferido de los modelos utilizados, como se reporta en trabajos previos de la misma naturaleza (42,63,65).

En comparación con OCT y fondo de ojo, para las imágenes espectrales basadas en ERG basal, no se obtuvieron resultados igual de buenos. Son estos modelos los que obtienen pérdidas más altas en comparación a todos los demás modelos y recordando la definición de la función de pérdida, esta calcula la diferencia entre el valor esperado y el valor obtenido de la predicción (73); es decir, los modelos de ERG son los que tienen peor desempeño para llevar a cabo una clasificación. Por último, es importante el concepto de sobreajuste, también conocido como *overfitting*, que, en términos de inteligencia artificial, se refiere al efecto de sobreentrenar a un algoritmo sobre un conjunto específico de datos (74). Una de las maneras de ver si nuestro modelo está sobreajustado es revisando las métricas de exactitud durante el entrenamiento y las pruebas (74). En condiciones normales, el valor obtenido durante las pruebas no debería superar el obtenido durante el entrenamiento (74). De no ser así, podemos pensar en un sobreajuste del modelo. Interesantemente, vemos que para el modelo DenseNet-201 con imágenes espectrales tenemos un probable sobreajuste, pues la exactitud de las pruebas supera a la del entrenamiento.

Este posible sobreajuste se debe, probablemente, a la poca cantidad de datos con las que se contó para el entrenamiento de los modelos (56), ya que, sin aumento de datos, únicamente había 141 imágenes espectrales con EMD.

Por otro lado, y como se mencionó previamente, un resultado interesante que se obtuvo en este trabajo fue que, aunque los modelos híbridos en general se desempeñaron mejor, los modelos singulares que utilizan como entrada OCT fueron los modelos más sensibles. Creemos que esto se puede deber a dos posibles factores. El primero es que la base de datos de OCT que tuvimos a nuestra disposición es un conjunto de datos muy numeroso. Únicamente la base de datos de *Kermany y col.* (54) contiene más de treinta mil imágenes de OCT de las clases que nos interesan. Recordando la definición de sensibilidad, esta es la capacidad de una prueba de reconocer a los verdaderos positivos (72). Es decir, su utilidad

radica en saber qué sujetos están realmente enfermos, de aquellos que la prueba clasifica como positivos. Debido a que, de la clase con EMD contamos con 11,598 imágenes de OCT para el entrenamiento, a diferencia de las otras clases de datos como fondo de ojo, en donde sólo contamos con 1,011, y las imágenes espectrales de ERG basal en donde, incluso, tuvimos que recurrir al aumento de datos, y cuya cantidad final fue de 846 imágenes; esto podría explicar que la métrica de sensibilidad fuera mayor en los modelos de OCT, pues los modelos podían extraer información de una mayor cantidad de sujetos enfermos. Al momento de combinar las matrices de probabilidad de los modelos, siguiendo el método descrito en la metodología, las bajas sensibilidades de los modelos de ERG y de fondo de ojo tendieron a disminuir las altas sensibilidades de los modelos de OCT.

Otra posible explicación a este resultado es el tipo de imagen empleada. A diferencia de las imágenes de fondo de ojo, que sólo muestran la cara interna de la parte posterior del globo ocular, lo que nos permite ver ciertas estructuras como la mácula o el disco óptico (10) y nos provee de alguna información cualitativa, el estudio de OCT muestra información cuantitativa acerca del grosor de las capas retinianas, así como la presencia o ausencia de fluido subretiniano o intrarretiniano (50), lo que se traduce en mayor cantidad de información acerca del tejido estudiado, y esto permite reconocer a un sujeto enfermo de un sujeto que no lo está con mayor facilidad, aumentando así la sensibilidad de la prueba.

Uno de los resultados más interesantes de este trabajo, es con respecto a las métricas obtenidas por el mejor modelo de OCT, por sí solo, en comparación con las métricas obtenidas por el modelo combinado de fondo de ojo (70 %) más escalogramas de la señal de ERG basal (30 %). Esto debido a que el mejor modelo de OCT, que fue OpticNet-71, fue el que mejor desempeño individual tuvo durante la validación con 164 sujetos, pero las métricas del modelo combinado de fondo de ojo más ERG son comparables. Ambos presentan Kappa de Cohen mayor a 0.8 y puntuación F1 mayor a 0.85, que como se mencionó previamente, son las mejores métricas para evaluar modelos con clases desbalanceadas (72). Quizá la única métrica donde el modelo de OCT es claramente superior, es en la sensibilidad (0.93 vs 0.88), pero este punto ya fue discutido previamente.

Estos resultados se vuelven especialmente relevantes sabiendo que el costo de un estudio de OCT de mácula, en México, oscila entre 910 y 2,500 pesos mexicanos (90,91); por otra parte, tomar un estudio de fondo de ojo cuesta en promedio 400 pesos mexicanos (92). Aunque el ERG basal no es un estudio cuyo uso se haya generalizado todavía en la clínica, consideramos que, entre los dos estudios, fondo de ojo más ERG basal, el costo no debería sobrepasar los 900 pesos mexicanos, lo cual, en el contexto de un país en vías de desarrollo en el que el salario mínimo es de 172.87 pesos mexicanos (93), se vuelve relevante. Además, el precio de un tomógrafo para obtener imágenes de OCT varía ampliamente desde 35,000 hasta poco más de 100,000 dólares americanos (94). Por otro lado, el costo de una cámara no midriática para la toma de imágenes de fondo de ojo oscila entre 10,000 y 20,000 dólares americanos (95) y el electrorretinógrafo RETeval LKC, que fue uno de los aparatos utilizados para este proyecto, cuesta aproximadamente 4,800 dólares americanos (96) y sabemos de otros prototipos en víspera de ser certificados (Retinaut, Francia) que podrían resultar igual o hasta más económicos. Si consideramos la compra de la cámara no midriática más costosa en conjunto con el electrorretinógrafo, obtenemos un costo aproximado de 25 mil dólares americanos, lo cual representa una ventaja en comparación con el costo del aparato más económico para la toma de OCT.

Posterior a este análisis, podemos decir que sí, resulta más económico, en términos generales, realizar un estudio de fondo de ojo más ERG basal, en comparación con el estudio de OCT.

Hablando de las características computacionales para utilizar estos modelos, el mayor costo computacional se requiere durante la fase del entrenamiento y las pruebas; es ahí donde se requieren características específicas para poder entrenar CNNs con una gran cantidad de imágenes. En términos generales, podemos decir que, para la fase de entrenamiento y pruebas, es necesaria una GPU (Unidad de procesamiento de gráficos, por sus siglas en inglés), que es un procesador formado por muchos núcleos pequeños y especializados (97), la gran variedad de núcleos ofrece un desempeño masivo cuando se realiza una tarea compleja, lo que lo hace muy útil para aplicaciones de aprendizaje profundo (97). Estas computadoras

requieren ciertas características para poder entrenar los modelos como una tarjeta gráfica potente o una memoria RAM de al menos 32 GB (97). Pero este proyecto versa sobre el uso de modelos ya entrenados, y una de las aproximaciones interesantes de este proyecto, es que la utilización de los modelos se puede hacer incluso en dispositivos móviles. Por ejemplo, Fragua, A. (98) presenta un trabajo en donde muestra que CNNs, como VVG16 y ResNet-50, que fueron entrenadas mediante el conjunto de datos ImageNet con una GPU, posteriormente pueden utilizarse en dispositivos móviles, con recursos mínimos y alta eficiencia (98). También se sabe que existen CNNs específicamente diseñadas para ser utilizadas en dispositivos móviles, como lo son MobileNet o EfficientNet (98). Con esta información podemos decir que, ya que hicimos el entrenamiento de los modelos, estos son viables para utilizarse en dispositivos con mínimos recursos, tal como un dispositivo móvil o una computadora convencional.

Así, pasamos al último punto de la discusión. El cuidado oftalmológico es inaccesible para mucha gente, y esto se debe a una variedad de factores, entre los cuales se encuentran la disponibilidad de los especialistas, insuficiente infraestructura (costo del equipo) y el transporte hacia las clínicas (99). Actualmente, se busca que los métodos de diagnóstico basados en IA puedan resolver el problema de la falta de especialistas en zonas con brechas en la atención médica (99). La presencia de algoritmos que puedan hacer diagnósticos confiables del EMD resulta en la solución al problema de falta de especialistas en oftalmología, pues un médico general u otro profesional de salud (optometrista o personal de enfermería, por ejemplo) con un adecuado entrenamiento podría utilizar la herramienta de diagnóstico para referir a un centro especializado únicamente a los pacientes que lo necesiten, reduciendo costos en transporte y en atención médica especializada (100).

En julio del año 2015, se publicó una revisión que estudiaba la carga económica de la diabetes en México (100). En ella describen que la carga más elevada se debe a cuatro factores importantes: medicamentos, atención ambulatoria, transporte del paciente y complicaciones.

Para el año 2013, el costo total de transporte del paciente con complicaciones de la diabetes, lo que incluía envíos de clínicas rurales a urbanas, alcanzó un costo total de 921.3 millones de pesos mexicanos (100). Estos costos pueden disminuir con la tecnología que se propone en este proyecto, pues únicamente serían enviados a un nivel mayor de atención los pacientes que estén clasificados adecuadamente como enfermos, en lugar de envío por sospechas clínicas, como suele realizarse (7).

De esta manera, habiendo presentado el análisis de costo económico y computacional, aunado a las ventajas que traería consigo la aplicación de la herramienta presentada en esta investigación, como lo son la reducción en costos de transporte y amortiguamiento en la falta de especialistas en salud visual en el contexto de un país en vías de desarrollo como México, es que se concluye su viabilidad, relevancia y aplicabilidad.

Conclusiones

Mediante el uso de imágenes espectrales de las señales de ERG basal, no es confiable realizar un diagnóstico del EMD, pues las métricas obtenidas de todos los modelos son muy bajas. En contraste, sí es posible mejorar el rendimiento de los modelos que utilizan OCT e imágenes de fondo de ojo mediante los cambios en la señal de ERG basales. Como se demostró en este trabajo, las métricas del mejor modelo de fondo de ojo, por sí solo, aumentan cuando se combinan con el modelo que utiliza espectrogramas provenientes del ERG basal, sin disminuir ninguna de sus métricas. Por otro lado, para OCT, ocurre algo similar a expensas de la sensibilidad, la cual se reduce (0.93 vs. 0.90) cuando se combina el mejor modelo de OCT con el modelo que utiliza escalogramas.

El modelo más robusto fue el modelo híbrido de OCT ($n1 = 0.60$) y fondo de ojo ($n2 = 0.40$), obteniendo la Kappa de Cohen más alta de todos los modelos (0.87) y todas las métricas restantes con valores igual o mayor a 0.90.

Aunque los modelos de OCT fueron los que obtuvieron valores de sensibilidad más altos, fueron los modelos híbridos los que obtuvieron valores más robustos en general.

Por último, cabe mencionar que, aunque los modelos de ERG no son confiables por sí solos para realizar un diagnóstico de EMD, son buenos coadyuvantes para los modelos de OCT y de fondo de ojo, pues son capaces de mejorar las métricas obtenidas por estos últimos. Esto revierte particular importancia al momento de proponer la combinación Fondo de ojo y ERG basal para el diagnóstico del EMD, como una alternativa concreta al examen de OCT, tanto a nivel del costo individual del estudio, como al momento del análisis económico de los aparatos de ERG y fondo de ojo, que resultan ser más económicos que un aparato para la toma de OCT.

Bibliografía

1. col BAA, la Cruz-Góngora D V, Basto-Abreu A, Barrientos-Gutiérrez T, Rojas-Martínez R, Aguilar-Salinas CA, et al. Prevalencia de diabetes y descontrol glucémico en México: resultados de la Ensanut 2016. [cited 2021 Aug 24]; Available from: <https://doi.org/10.21149/10752>
2. Encuesta Nacional de Salud y Nutrición 2018 Presentación de resultados.
3. Association AD. 11. Microvascular Complications and Foot Care: Standards of Medical Care in Diabetes—2021. *Diabetes Care* [Internet]. 2021 Jan 1 [cited 2021 Aug 24];44(Supplement 1):S151–67. Available from: https://care.diabetesjournals.org/content/44/Supplement_1/S151
4. A D, A M, A M, L K, M N, A S, et al. Mechanisms of macular edema: Beyond the surface. *Prog Retin Eye Res* [Internet]. 2018 Mar 1 [cited 2021 Sep 9];63:20–68. Available from: <https://pubmed.ncbi.nlm.nih.gov/29126927/>
5. U SE, J GA, F B, K B, U C, BS G, et al. Guidelines for the Management of Diabetic Macular Edema by the European Society of Retina Specialists (EURETINA). *Ophthalmologica* [Internet]. 2017 May 1 [cited 2021 Sep 8];237(4):185–222. Available from: <https://pubmed.ncbi.nlm.nih.gov/28423385/>
6. Roberto Carmona-Moxica L, Hernández-Núñez F. Práctica clínico-quirúrgica Revisión bibliográfica: edema macular diabético, repercusiones y tratamiento Literature review: Diabetic macular edema. Repercussions and treatment. *Rev Med Inst Mex Seguro Soc*. 2015;53(5):600–7.
7. GUÍA DE PRÁCTICA CLÍNICA GPC Diagnóstico y tratamiento de. [cited 2021 Sep 9]; Available from: <http://www.cenetec.salud.gob.mx/interior/catalogoMaestroGPC.html>
8. Ejecutivo R. Índice.
9. J. Andonegui, L. Jiménez. Edema macular diabético [Internet]. 2008 [cited 2021 Sep 9]. Available from: https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1137-66272008000600004
10. Ganong WF, Barrett KE. *Ganong fisiología medica*. McGraw-Hill; 2016.
11. Roy S, Kim D, Lim R. Cell-cell communication in diabetic retinopathy. *Vision Res* [Internet]. 2017 Oct 1 [cited 2021 Aug 27];139:115. Available from: </pmc/articles/PMC5723213/>
12. Roy S, Jiang JX, Li AF, Kim D. Connexin Channel and its Role in Diabetic Retinopathy. *Prog Retin Eye Res* [Internet]. 2017 Nov 1 [cited 2021 Aug 27];61:35. Available from: </pmc/articles/PMC5653466/>
13. González-Casanova J, Schmachtenberg O, Martínez AD, Sanchez HA, Harcha PA, Rojas-Gomez D. An Update on Connexin Gap Junction and Hemichannels in Diabetic Retinopathy. *Int J Mol Sci* [Internet]. 2021 Mar 2 [cited 2021 Aug 27];22(6):1–19. Available from: </pmc/articles/PMC8004116/>
14. Connexin expression in the retina. *Brain Research Reviews*, 32(1), 138–145 | 10.1016/s0165-0173(99)00074-0 [Internet]. [cited 2021 Sep 1]. Available from: </pubmed.ncbi.nlm.nih.gov/10751663/>
15. Söhl G, Joussen A, Kociok N, Willecke K. Expression of connexin genes in the human retina. *BMC Ophthalmology* 2010 10:1 [Internet]. 2010 Oct 27 [cited 2021 Sep 1];10(1):1–11. Available from: <https://bmcophthalmol.biomedcentral.com/articles/10.1186/1471-2415-10-27>

16. Simó R, Stitt AW, Gardner TW. Neurodegeneration in diabetic retinopathy: does it really matter? *Diabetologia* [Internet]. 2018 Sep 1 [cited 2021 Sep 3];61(9):1902. Available from: [/pmc/articles/PMC6096638/](#)
17. Rodríguez-Guanare ML. Edema macular diabético: definición, epidemiología, factores de riesgo, patogénesis, diagnóstico y clasificación. 2015 [cited 2022 Feb 12]; Available from: [www.revistaalad.com](#)
18. M C, PS C, J O, DE C, P H. Manipulating ocular endothelial tight junctions: Applications in treatment of retinal disease pathology and ocular hypertension. *Prog Retin Eye Res* [Internet]. 2018 Jan 1 [cited 2021 Sep 9];62:120–33. Available from: <https://pubmed.ncbi.nlm.nih.gov/28951125/>
19. Szarka G, Balogh M, Tengölics ÁJ, Ganczer A, Völgyi B, Kovács-Öller T. The role of gap junctions in cell death and neuromodulation in the retina. *Neural Regen Res* [Internet]. 2021 Oct 1 [cited 2021 Aug 27];16(10):1911. Available from: [/pmc/articles/PMC8343308/](#)
20. Ly A, Yee P, Vessey KA, Phipps JA, Jobling AI, Fletcher EL. Early Inner Retinal Astrocyte Dysfunction during Diabetes and Development of Hypoxia, Retinal Stress, and Neuronal Functional Loss. *Invest Ophthalmol Vis Sci*. 2011 Dec 1;52(13):9316–26.
21. Navaratna D, McGuire PG, Menicucci G, Das A. Proteolytic Degradation of VE-Cadherin Alters the Blood-Retinal Barrier in Diabetes. *Diabetes* [Internet]. 2007 Sep 1 [cited 2021 Sep 9];56(9):2380–7. Available from: <https://diabetes.diabetesjournals.org/content/56/9/2380>
22. Procesamiento y Clasificación de Bioseñales con Inteligencia Computacional [Internet]. [cited 2022 Feb 6]. Available from: http://smia.mx/comia/2017/index.php?option=com_content&view=article&id=12&Itemid=127
23. Análisis de bioseñales: Enfoque técnico de la adquisición, procesamiento y sus aplicaciones [Internet]. [cited 2022 Feb 6]. Available from: https://www.researchgate.net/publication/327560974_Analisis_de_biosenales_Enfoque_tecnico_de_la_adquisicion_procesamiento_y_sus_aplicaciones
24. Hernández E. Clasificación de señales EEG basada en representaciones bidimensionales y redes neuronales convolucionales [Internet]. 2020 [cited 2022 Jan 10]. Available from: <https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/2001/1/Tesis%20MC%20EEG%20DL%20v6%2027-01-21.pdf>
25. Blanco D. Análisis EEG de presencia de dolor en voluntarios. 2017;
26. Santa María F, Kouro Rodrigo Musalem M E SR. Técnicas Modernas en Automática Tutorial introductorio a la Teoría de Wavelet.
27. Türk Ö, Sirac M. Epilepsy Detection by Using Scalogram Based Convolutional Neural Network from EEG Signals. 2017;
28. Villazana S, Montilla G, Eblen A, Maldonado C, Villazana S, Montilla G, et al. Detección de señales EEG epilépticas utilizando redes convolucionales basada en la transformada synchrosqueezing acolchada. *Computación y Sistemas* [Internet]. 2021 [cited 2022 Feb 6];25(2):269–86. Available from: http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-55462021000200269&lng=es&nrm=iso&tlng=es

29. Creel DJ. Electroretinograms. In: Handbook of Clinical Neurology. Elsevier B.V.; 2019. p. 481–93.
30. Luu CD, Szental JA, Lee SY, Lavanya R, Wong TY. Correlation between Retinal Oscillatory Potentials and Retinal Vascular Caliber in Type 2 Diabetes. *Invest Ophthalmol Vis Sci*. 2010 Jan 1;51(1):482–6.
31. HH X, JL C, HY C, HJ L. Correlation between optical coherence tomography, multifocal electroretinogram findings and visual acuity in diabetic macular edema. *Int J Ophthalmol* [Internet]. 2020 Oct 18 [cited 2021 Sep 9];13(10):1592–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/33078110/>
32. NM T, H RE, E J, A M, H T, A A, et al. Multifocal Electroretinogram in Diabetic Macular Edema; Correlation with Visual Acuity and Optical Coherence Tomography. *J Ophthalmic Vis Res* [Internet]. 2015 Apr 1 [cited 2021 Sep 9];10(2):165–71. Available from: <https://pubmed.ncbi.nlm.nih.gov/26425320/>
33. Kato K, Kondo M, Sugimoto M, Ikesugi K, Matsubara H. Effect of pupil size on flicker ERGs recorded with RETeval system: New mydriasis-free full-field ERG system. *Invest Ophthalmol Vis Sci*. 2015;56(6):3684–90.
34. Ramsés Noguez Imm, Julio Muñoz-Benitez. Spontaneous electroretinogram signals to screen people with early risk factors for diabetic retinopathy. [cited 2022 Nov 11]; Available from: <https://www.medrxiv.org/content/10.1101/2022.06.26.22276881v1>
35. Moraru AD, Costin D, Moraru RL, Branisteanu DC. Artificial intelligence and deep learning in ophthalmology - present and future (Review). *Exp Ther Med* [Internet]. 2020 Aug 12 [cited 2021 Sep 23];20(4):3469. Available from: </pmc/articles/PMC7465350/>
36. Carrión CB. REDES CONVOLUCIONALES.
37. Zarranz J, Abreu R, Fernández M, Lara J. INTELIGENCIA ARTIFICIAL EN RETINA. MONOGRAFÍA. 2020.
38. 6. a) Red Neuronal monocapa, b) Red Neuronal multicapa de propagación... | Download Scientific Diagram [Internet]. [cited 2022 Nov 11]. Available from: https://www.researchgate.net/figure/Figura-III6-a-Red-Neuronal-monocapa-b-Red-Neuronal-multicapa-de-propagacion-hacia_fig5_315762548
39. David A., Arroyo E. Visualizando neuronas en Redes neuronales convolucionales.
40. Contreras S, de la Rosa F. Aplicación de Deep Learning en Robótica Móvil para Exploración y Reconocimiento de Objetos basados en Imágenes. *Systems and Computing Engineering Department* . 2016;
41. Pérez-Aguilar DRRRCPL. Transfer learning en la clasificación binaria de imágenes térmicas. *Ingenius Revista de Ciencia y Tecnología*, núm 26, 2021. 2021;
42. Pin K, Ho Chang J, Nam Y. Comparative Study of Transfer Learning Models for Retinal Disease Diagnosis from Fundus Images. *Computers, Materials & Continua*. 2022;70(3):5821–34.
43. Laptev N, Laptev V, Gerget O, Kolpashchikov D. Integrating Traditional Machine Learning and Neural Networks for Image Processing. 2021;
44. Cánovas F, Alonso F, Gomariz F. MODIFICACIÓN DEL ALGORITMO RANDOM FOREST PARA SU EMPLEO EN CLASIFICACIÓN DE IMÁGENES DE TELEDETECCIÓN. 2016 Jul;
45. Movilla V. Identificación automática de tejido metastásico en imágenes de nódulos linfáticos usando Machine Learning. [Los Andes]: Universidad de Los Andes; 2019.

46. K-Nearest Neighbors para Clasificación de Imágenes - DataSmarts Español [Internet]. [cited 2023 Jun 25]. Available from: <https://datasmarts.net/es/k-nearest-neighbors-para-clasificacion-de-imagenes/>
47. Karypidids E, Mouslech S, Skoulariki K, Gazis A. Comparison Analysis of Traditional Machine Learning and Deep Learning Techniques for Data and Image Classification. 2022 [cited 2023 Jun 25]; Available from: https://www.researchgate.net/publication/359918228_Comparison_Analysis_of_Traditional_Machine_Learning_and_Deep_Learning_Techniques_for_Data_and_Image_Classification
48. Wang P, Fan E, Wang P. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. *Pattern Recognit Lett.* 2021 Jan;141:61–7.
49. Bhanarkar P. *Machine Learning And Deep Learning: Image Classification.* Ilkogretim. 2021;
50. T S, SM W, H B, F E, A S, AM P, et al. Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning. *Ophthalmology* [Internet]. 2018 Apr 1 [cited 2021 Sep 23];125(4):549–58. Available from: <https://pubmed.ncbi.nlm.nih.gov/29224926/>
51. Varadarajan A v., Bavishi P, Ruamviboonsuk P, Chotcomwongse P, Venugopalan S, Narayanaswamy A, et al. Predicting optical coherence tomography-derived diabetic macular edema grades from fundus photographs using deep learning. *Nat Commun* [Internet]. 2020 Dec 1 [cited 2021 Sep 23];11(1). Available from: </pmc/articles/PMC6949287/>
52. Liang-Chieh Chen YZGPFS and HA. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Google Inc. 2018;
53. Garg R, Kumar A, Bansal N, Prateek M, Kumar S. Semantic segmentation of PolSAR image data using advanced deep learning model. *Sci Rep.* 2021 Dec 28;11(1):15365.
54. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell.* 2018 Feb 22;172(5):1122-1131.e9.
55. Prerau MJ, Brown RE, Bianchi MT, Ellenbogen JM, Purdon PL. Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis. *Physiology.* 2017 Jan;32(1):60–92.
56. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data.* 2019 Dec 6;6(1):60.
57. Decencièrè E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, et al. FEEDBACK ON A PUBLICLY DISTRIBUTED IMAGE DATABASE: THE MESSIDOR DATABASE. *Image Analysis & Stereology.* 2014 Aug 26;33(3):231.
58. Pachade S, Porwal P, Thulkar D, Kokare M, Deshmukh G, Sahasrabuddhe V, et al. Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research. *Data* 2021, Vol 6, Page 14 [Internet]. 2021 Feb 3 [cited 2022 May 1];6(2):14. Available from: <https://www.mdpi.com/2306-5729/6/2/14/htm>
59. Giancardo L, Meriaudeau F, Karnowski TP, Li Y, Garg S, Tobin KW, et al. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. *Med Image Anal.* 2012 Jan;16(1):216–26.

60. Srinivasan PP, Kim LA, Mettu PS, Cousins SW, Comer GM, Izatt JA, et al. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed Opt Express*. 2014 Oct 1;5(10):3568.
61. Zhuang F, Qi Z, Duan K, Xi D, Zhu Y, Zhu H, et al. A Comprehensive Survey on Transfer Learning. 2019 Nov 6;
62. Rivera M. CIMAT. 2019. La Red Residual (Residual Network, ResNet).
63. VAIBHAV KUMAR. MobileNet vs ResNet50 – Two CNN Transfer Learning Light Frameworks. 2020.
64. TensorFlow - tf.keras.applications.DenseNet201 - Instancia de la arquitectura Densenet201. Main aliases tf.keras.applications.den - Español [Internet]. [cited 2023 Jun 25]. Available from: <https://runebook.dev/es/docs/tensorflow/keras/applications/densenet201>
65. Patel R, Chaware A. Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy. In: 2020 International Conference for Emerging Technology (INCET). IEEE; 2020. p. 1–4.
66. Goodfellow I, Bengio Y, Courville A. Optimization for Training DeepModels. In.
67. Adam [Internet]. [cited 2023 Jun 25]. Available from: <https://keras.io/api/optimizers/adam/>
68. Sharif Amit Kamran* SSASS and AT. Optic-Net: A Novel Convolutional Neural Network for Diagnosis of Retinal Diseases from Optical Tomography Images. University of Nevada. 2019;
69. tf.keras.Model | TensorFlow v2.10.0 [Internet]. [cited 2022 Oct 23]. Available from: https://www.tensorflow.org/api_docs/python/tf/keras/Model
70. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med* [Internet]. 2013 [cited 2021 Sep 23];4(2):627. Available from: </pmc/articles/PMC3755824/>
71. España P, Gordillo T, Jesús J, Rodríguez P, Hugo V. Revista de Investigación Educativa. *Investigación Educativa* [Internet]. 2009 [cited 2021 Sep 23];27(1):89–103. Available from: <http://www.redalyc.org/articulo.oa?id=283322804006>
72. Sokolova M, Japkowicz N, Szpakowicz S. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In 2006. p. 1015–21.
73. (PDF) Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review [Internet]. [cited 2022 May 16]. Available from: https://www.researchgate.net/publication/352703047_Accuracy_Assessment_in_Convolutional_Neural_Network-Based_Deep_Learning_Remote_Sensing_Studies-Part_1_Literature_Review
74. Rebecca Roelofs*. A Meta-Analysis of Overfitting in Machine Learning. 33rd Conference on Neural Information Processing Systems. 2019;
75. Muñoz de Escalona-Rojas JE, Quereda-Castañeda A, García-García O. Actualización de la retinopatía diabética para médicos de atención primaria: hacia una mejora de la medicina telemática. *SEMERGEN - Medicina de Familia*. 2016 Apr;42(3):172–6.
76. Ciprés M, Satue M, Melchor I, Gil-Arribas L, Vilades E, Garcia-Martin E. Neurodegeneración retiniana en pacientes diabéticos tipo 2 sin retinopatía diabética. *Arch Soc Esp Oftalmol*. 2022 Apr;97(4):205–18.

77. G. Tenorio, v. Ramírez. Retinopatía diabética; Conceptos actuales. *Revista Médica del Hospital General de México*. 2010;
78. McAnany JJ, Persidina OS, Park JC. Clinical electroretinography in diabetic retinopathy: a review. *Surv Ophthalmol*. 2022 May;67(3):712–22.
79. Audun Eltvik. Deep Learning for the Classification of EEG Time-Frequency Representations. [Norwegian]: Norwegian University of Science and Technology; 2018.
80. Ramos R, Olvera A, Olmos I. Analysis of EEG Signal Processing Techniques based on Spectrograms. Benemérita Universidad Autónoma de Puebla. 2017;
81. Bynum J, Lattanzi D. Combining convolutional neural networks with unsupervised learning for acoustic monitoring of robotic manufacturing facilities. *Advances in Mechanical Engineering*. 2021 Apr 12;13(4):168781402110090.
82. Ruffini G, Ibañez D, Castellano M, Dubreuil-Vall L, Soria-Frisch A, Postuma R, et al. Deep Learning With EEG Spectrograms in Rapid Eye Movement Behavior Disorder. *Front Neurol*. 2019 Jul 30;10.
83. Las técnicas de neuroimagen [Internet]. [cited 2022 Nov 5]. Available from: <https://www.centroacadia.es/tecnicas-neuroimagen/>
84. Robson AG, Nilsson J, Li S, Jalali S, Fulton AB, Tormene AP, et al. ISCEV guide to visual electrodiagnostic procedures. *Documenta Ophthalmologica*. 2018 Feb 3;136(1):1–26.
85. Kato K, Kondo M, Sugimoto M, Ikesugi K, Matsubara H. Effect of Pupil Size on Flicker ERGs Recorded With RET *eval* System: New Mydriasis-Free Full-Field ERG System. *Investigative Ophthalmology & Visual Science*. 2015 Jun 5;56(6):3684.
86. Sancho Fernando. Métodos combinados de aprendizaje [Internet]. 2018 [cited 2022 Nov 5]. Available from: <http://www.cs.us.es/~fsancho/?e=106>
87. Dang CN, Moreno-García MN, De la Prieta F. Hybrid Deep Learning Models for Sentiment Analysis. *Complexity*. 2021 Aug 12;2021:1–16.
88. Md Shad Akhtar, Ayush Kumar. A Hybrid Deep Learning Architecture for Sentiment Analysis.
89. Xue DX, Zhang R, Feng H, Wang YL. CNN-SVM for Microvascular Morphological Type Recognition with Data Augmentation. *J Med Biol Eng*. 2016 Dec 10;36(6):755–64.
90. Estudio de OCT Macular – Visión Digna [Internet]. [cited 2022 Nov 11]. Available from: <https://visiondigna.com/products/oct-macular>
91. 25 Hospitales Angeles en México, ¡ubica tu hospital! [Internet]. [cited 2022 Nov 11]. Available from: <https://www.hospitalesangeles.com/saludyvida/mujeres/articulo.php?id=3351>
92. Estudio de fondo de ojo [Internet]. [cited 2022 Nov 11]. Available from: <https://diabesmart.com.mx/product/retinografia-estudio-de-fondo-de-ojo/>
93. Entra en vigor incremento al salario mínimo del 22% | Secretaría del Trabajo y Previsión Social | Gobierno | gob.mx [Internet]. [cited 2022 Nov 11]. Available from: <https://www.gob.mx/stps/prensa/entra-en-vigor-incremento-al-salario-minimo-del-22?idiom=es>
94. OCT machine prices, comparisons and general must-knows - Medilex [Internet]. [cited 2022 Nov 11]. Available from: <https://www.medilexonline.com/oct-machine-prices/>

95. Topcon Medical TRC-50X Fundus Camera — reLink Medical [Internet]. [cited 2022 Nov 11]. Available from: <https://relinkonline.com/es/products/topcon-medical-trc-50x-fundus-camera-inv-8000228685>
96. Al-Otaibi H, Al-Otaibi MD, Khandekar R, Souru C, Al-Abdullah AA, Al-Dhibi H, et al. Validity, Usefulness and Cost of RET *eval* System for Diabetic Retinopathy Screening. *Transl Vis Sci Technol*. 2017 Jul 17;6(3):3.
97. CPU vs GPU: ¿Cuál es la diferencia? [Internet]. [cited 2022 Nov 11]. Available from: <https://www.intel.la/content/www/xl/es/products/docs/processors/cpu-vs-gpu.html>
98. Superior EP, Fragua Baeza Á, Ángel M, García G. Ejecución de redes neuronales en móviles Android con aceleración hardware mediante Keras y Tensorflow Lite. 2021 [cited 2022 Nov 11]; Available from: <https://repositorio.uam.es/handle/10486/698211>
99. Cheng D, Babij R, Cabrera D, Yuan M, Port A, Mckenney AS, et al. Effective Low-Cost Ophthalmological Screening With a Novel iPhone Fundus Camera at Community Centers. *Cureus*. 2022 Aug 17;
100. Barraza-Lloréns M, Guajardo-Barrón V, Picó J, García R, Hernández C, Mora F, et al. Carga Económica de la Diabetes Mellitus en México, 2013. *Funsalud*. 2015;