

Análisis de Modelos de Segmentación U-Net, PSP-Net, SegNet y FCN en Imágenes de Fondo de Ojo

1

J. E. Hernández, *Estudiante MCIA, UAQ*, J. Anguiano, *Estudiante MCIA, UAQ*, C.G. Rodríguez, *Estudiante MCIA, UAQ*, J. J. Flores *Estudiante MCIA, UAQ*,
Deep Learning

Resumen—La segmentación de imágenes se ha posicionado como un campo de gran importancia en la Inteligencia Artificial que nos permite brindar a los sistemas informáticos la habilidad de “visualizar” su entorno y “extraer” regiones de interés. Las imágenes de fondo de ojo son una técnica no-invasiva con un alto índice de empleabilidad por oftalmólogos para detectar factores de riesgo asociados al glaucoma, entre los que se encuentra la Atrofia Peripapilar (APP). La tele oftalmología se desarrolla para cerrar la brecha entre la medicina y los sistemas informáticos.

La base de datos en este trabajo consta de 300 imágenes de fondo de ojo con etiquetas de las regiones de APP delimitadas manualmente y verificadas por una especialista en Glaucoma. Estas imágenes fueron extraídas de la base de datos pública ORIGA.

En el presente trabajo se utilizaron cuatro distintos modelos para la segmentación: U-Net, PSP-Net, SegNet y FCN, en donde se obtuvieron resultados positivos por parte del modelo U-Net.

Temas claves—Machine Learning, Segmentación de imágenes, Autocodificadores, U-Net, PSP-Net, SegNet y FCN, Fondo de ojo, Retina.

I. INTRODUCCIÓN

EL glaucoma es una de las principales causas de ceguera irreversible en el mundo, y los modelos actuales de atención médica no pueden cerrar la brecha entre la creciente prevalencia de la neuropatía, y los desafíos tanto del acceso a los servicios de salud especializada como la dificultad de proveer de un diagnóstico oportuno, por lo que la tele oftalmología se desarrolla para cerrar estas brechas.

Actualmente se cuenta con técnicas de inteligencia artificial capaces de analizar imágenes y generar su debida segmentación, la cual tiene un objetivo similar al modelo de detección de objetos visto anteriormente. Ambos necesitan localizar sujetos: persona, animal, vehículo, etc. Sin embargo, en la segmentación de imágenes, la salida de las redes neuronales es mucho más precisa en términos de localización

debido a que se necesita realizar la clasificación de cada píxel de la imagen de entrada.

En este trabajo se implementan cuatro distintas arquitecturas basadas en autocodificadores para la segmentación de Atrofia Peripapilar Alfa y Beta en pacientes con y sin glaucoma, las cuales se entrenaron por medio de una muestra de 300 imágenes de la base de datos pública ORIGA [1], con el fin de realizar una comparación de los resultados obtenidos.



Fig. 1. Ejemplo de imagen de fondo de ojo en un paciente con Glaucoma [1].

II. JUSTIFICACIÓN

El glaucoma alcanzó a nivel mundial un total de 60.5 millones de casos en el 2010, y de acuerdo con estudios de metaanálisis epidemiológicos que estimaron la prevalencia de glaucoma primario de ángulo abierto y cerrado, para los años 2020 y 2040 en pacientes entre los 40 y 80 años, obteniendo proyecciones donde revelaron que en el año 2013 vivían 64.3 millones de personas afectadas por glaucoma en el mundo, para el 2020 se incrementó a 76 millones y estimando finalmente para alcanzar 111.8 millones en el 2040, lo que representa el 3.54% de la población mundial. Los resultados de los estudios expresan un comportamiento de crecimiento desproporcionado. Dado que el glaucoma es una enfermedad que progresa lentamente con daño neuronal irreversible, el diagnóstico precoz y la monitorización sensible de la progresión son fundamentales para el tratamiento del glaucoma [2].

De acuerdo con las guías prácticas del IMSS, en México, el diagnóstico de glaucoma se realiza en hospitales de segundo nivel de atención médica mediante el estudio de perimetría automatizada, con el cual a su vez es posible determinar el grado de la patología; no obstante, este estudio va acompañado de otras pruebas que evalúan la calidad de la visión y rasgos

¹J. E. Hernández estudia la Maestría en Ciencias en Inteligencia Artificial en la Universidad Autónoma de Querétaro (e-mail: zeedhermx@gmail.com)

C.G. Rodríguez estudia la Maestría en Ciencias en Inteligencia Artificial en la Universidad Autónoma de Querétaro (e-mail: gaby96.flores@gmail.com)

J. Anguiano estudia la Maestría en Ciencias en Inteligencia Artificial en la Universidad Autónoma de Querétaro (e-mail: anguiano_javier@outlook.com)

J. J. Flores estudia la Maestría en Ciencias en Inteligencia Artificial en la Universidad Autónoma de Querétaro (e-mail: @outlook.com)

morfológicos en el nervio óptico, como lo es la atrofia peripapilar beta [2]. Es por ello, que es necesario desarrollar estrategias de salud pública para garantizar la detección oportuna y tratamiento temprano con el objetivo de retrasar la pérdida visual.

La detección y segmentación de estructuras morfológicas en imágenes de fondo de ojo son fundamentales para el desarrollo de sistemas automáticos de diagnóstico de patologías oftalmológicas [3].

Las técnicas de DL son herramientas que han demostrado ser útiles y efectivas en aplicaciones médicas, como es el reconocimiento de imágenes, donde destacan la segmentación del disco óptico, clasificación de retinopatía diabética referible, segmentación de vasos, entre otras [4]; generando un pronóstico favorable para el desarrollo de nuevas metodologías que integren técnicas de DL para la detección y segmentación de APP beta en imágenes de fondo de ojo.

Con la adaptación de los principales avances en las tecnologías de diagnóstico que ofrecen una visión más amplia del estado de la retina y de las enfermedades oculares, los médicos y los proveedores de atención médica han adoptado nuevos métodos de medición de la actividad visual en un esfuerzo por mejorar la atención al paciente, acelerar el proceso de gestión y minimizar la cantidad de herramientas necesarias para cada evaluación.

III. MARCO TEÓRICO

A. Glaucoma

El glaucoma representa un grupo diverso de enfermedades que tienen en común el aumento de la excavación del nervio óptico, pérdida del campo visual y en la gran mayoría de los casos, elevación de la presión intraocular [5].

El daño glaucomatoso temprano puede ser difícil de identificar mediante la evaluación funcional de la visión, ya que esta se ve afectada cuando se encuentra en estados avanzados, por ello es por lo que se requiere de una observación minuciosa de las características morfológicas del nervio óptico.

La evaluación clínica de la imagen de fondo de ojo a través de un oftalmoscopio requiere de la interpretación experta de un oftalmólogo altamente capacitado que provea de una evaluación semi-cuantitativa [6].

Entre los factores estructurales que pueden ser observados dentro de las imágenes de fondo de ojo están: Tamaño del disco óptico, tamaño y forma del nervio neuro retiniano, capas de fibras nerviosas retinianas, presencia de atrofia peripapilar y presencia de hemorragias retinianas o del disco óptico[7].

B. Atrofia Peripapilar

Se ha demostrado que la atrofia peripapilar (PPA) presenta una fuerte correlación con el diagnóstico de glaucoma [8]. Esta se puede definir como el adelgazamiento y degeneración del tejido coriorretiniano justo fuera del disco óptico, que se asocia con el desarrollo y la progresión del glaucoma [9], [10]. Existen dos

formas de atrofia peripapilar: alfa y beta. La atrofia alfa, se encuentra presente en la mayoría de los ojos sanos, así como en los ojos glaucomatosos, y se caracteriza por regiones que presentan hipopigmentación, con otras con hiperpigmentación irregular del epitelio pigmentario de la retina (EPR) como se muestra en la Figura 2 señalada con las flechas blancas y la atrofia beta con flechas negras. Para describir la extensión de la atrofia beta, se puede utilizar la relación entre la longitud circunferencial de la atrofia peripapilar a lo largo del borde del disco óptico (entre A y B) y la circunferencia del disco óptico. La atrofia beta es el más importante para el diagnóstico de glaucoma, ya que es causada por la atrofia del EPR y coriocapilar que conduce a que los vasos coroideos y la esclerótica se vuelvan visibles [7].

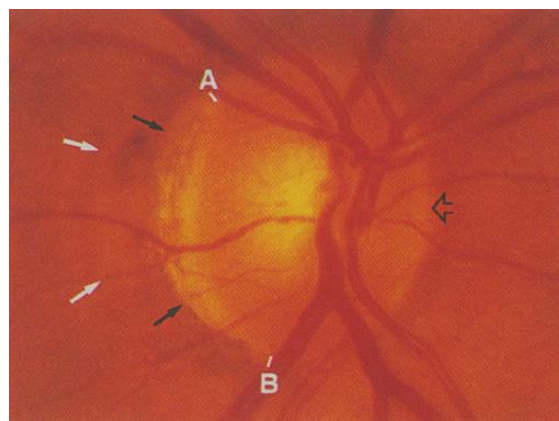


Fig. 2. Acercamiento del nervio óptico de imagen de fondo de ojo con atrofia peripapilar Alfa y Beta [3].

C. Segmentación

La segmentación semántica, considerada como una tarea de alto nivel que encauza a una mayor comprensión de la escena, consiste en clasificar cada píxel de una imagen como perteneciente a una clase particular que permite identificar a todos los elementos que constituyen la escena.

El objetivo de la segmentación semántica es el mismo que el de la clasificación de imágenes tradicional en la teledetección, que generalmente se lleva a cabo mediante la aplicación de técnicas tradicionales de aprendizaje automático, como el bosque aleatorio y el clasificador de máxima verosimilitud.

D. Modelos

Hay muchos algoritmos de segmentación semántica como U-net, Feature Pyramid Network (FPN), PSP-Net, SegNet, etc. A continuación, se presentan los modelos mencionados anteriormente.

Modelo SegNet

SegNet es una arquitectura de red neuronal totalmente convolucional, profunda que realiza una segmentación semántica por píxeles. Su arquitectura consta de una red de codificador, una red de decodificador correspondiente seguida de una de clasificación por píxeles.

La red del codificador es topológicamente idéntica a las trece capas convolucionales de la red VGG16 [11]. El papel de la red del decodificador es mapear los mapas de características del codificador de baja resolución a mapas de características de resolución de entrada completa para la clasificación por píxeles.

La novedad de SegNet radica en la forma en que el decodificador aumenta la muestra de sus mapas de características de entrada de menor resolución. Específicamente, el decodificador utiliza índices de agrupación calculados en el paso de agrupación máxima del codificador correspondiente para realizar un muestreo ascendente no lineal. Esto elimina la necesidad de aprender a aumentar la muestra. Los mapas sobre muestreados son escasos y luego se convolucionan con filtros entrenables para producir mapas de características densos [12].

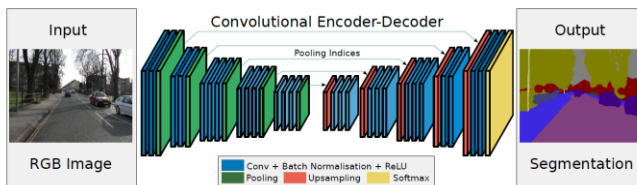


Fig. 3. Arquitectura SegNet [5].

Modelo U-Net

La red U-Net se trata de un modelo de red neuronal totalmente convolucional. Este modelo fue desarrollado originalmente por Olaf Ronneberger, Phillip Fischer y Thomas Brox en 2015 para la segmentación de imágenes médicas.

Su arquitectura contiene dos caminos, la primera ruta, también conocida como como ruta de contracción o bien codificador, en donde por medio de ella es posible capturar el contexto de una imagen, es decir, contiene un conjunto de capas de convolución, así como capas de max pooling, con las cuales se consigue reducir el número de parámetros de la red.

Así bien, la segunda ruta, también conocida como ruta de expansión simétrica o bien decodificador, se utiliza para obtener una localización precisa mediante la convolución transpuesta [13].

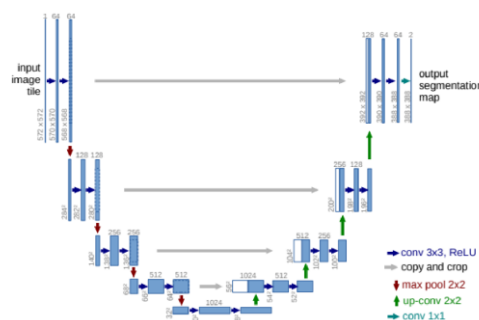


Fig. 4. Arquitectura U-Net

Modelo FCN

Las redes totalmente convolucionales (FCN) han demostrado ser una solución eficaz de extremo a extremo para la segmentación semántica de imágenes. Un FCN produce un mapa de puntuación de un tamaño proporcional a la imagen de entrada, donde cada píxel representa un clasificador de objetos. La idea clave de FCN es utilizar múltiples escalas de mapas de características para generar la predicción final [14].

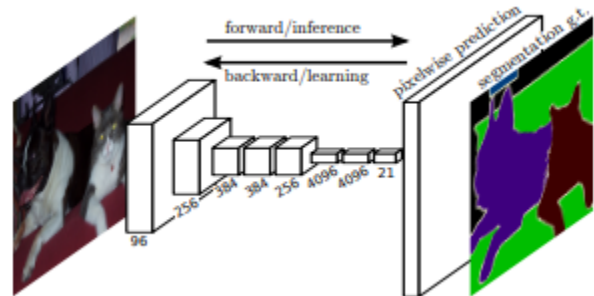


Fig. 5. Arquitectura FCN [14].

Modelo PSP-Net

El modelo PSP-Net (Pyramid Scene Parsing Network), o análisis piramidal de escena, basado en segmentación semántica, es un tema fundamental en la visión artificial. El objetivo es asignar a cada píxel de la imagen una etiqueta de categoría. Así, el análisis de escenas proporciona una comprensión completa de la escena. Este, predice la etiqueta, ubicación así como la forma de cada elemento.

Los modelos PSP son de gran interés para aplicaciones en conducción autónoma, sensado robótico, etc.

El modelo PSP-Net, emplea un módulo de agrupación piramidal, y proporciona una prioridad global contextual efectiva para el análisis de escenas pixel a pixel. El módulo de agrupación piramidal puede recopilar niveles de información, más representativos que la agrupación global. Empleando una pirámide de 4 niveles, el modelo logra cubrir mediante los pooling kernels, la totalidad, la mitad, y pequeñas porciones de la imagen. Estos, se emplean con prioridad global, y entonces son concatenados con el mapa de características original. Finalmente, una capa convolucional genera el mapa final de predicciones. [15]

La arquitectura del modelo PSP-Net tiene en cuenta el contexto global de la imagen para hacer predicciones a nivel local, lo que ofrece un mejor rendimiento en paisajes urbanos o conjuntos de referencia como PASCAL VOC 2012. El modelo era necesario pues los clasificadores de píxeles basados en FCN no pueden capturar el contexto de la imagen completa.

El codificador en el modelo PSP-Net contiene la red nuclear de CNN con convoluciones dilatadas a través del módulo de agrupación piramidal. Una vez que el codificador ha extraído las características de la imagen, el decodificador toma esas características y las convierte en predicciones gracias a sus capas.

El modelo PSP-Net no es un modelo de segmentación completo en sí mismo, es solo un codificador, lo que significa que es solo la mitad de lo que se requiere para segmentación de imágenes. Los decodificadores más comunes que se encuentran en varias implementaciones de PSP-Net son una capa de convolución seguida de un muestreo ascendente bilineal 8x (aunque los resultados no capturan información de alta resolución de la imagen.). [16]

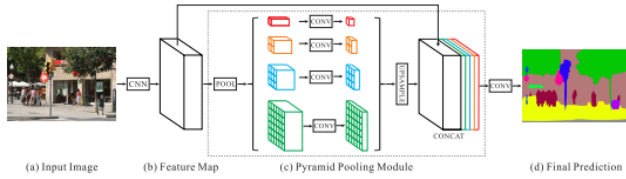


Fig. 6. Arquitectura PSP-Net [15].

IV. DESARROLLO

Para la elaboración de este trabajo, se hizo uso de las redes FCN, U-Net, PSP-Net, Seg-Net. La metodología utilizada para la implementación de los modelos mencionados anteriormente fue muy similar entre sí.

Antes de detallar cada uno de los modelos propuestos, se procede a mencionar la base de datos a utilizar, así como el proceso de acondicionamiento de la base de datos.

A. Adquisición de base de datos

La base de datos está compuesta de 300 imágenes, de las cuales se tomaron 240 para entrenamiento y 60 para prueba con un tamaño de imagen de [3072,2048]. Al cual se le aplicaron técnicas de redimensionamiento para adecuar los datos a cada modelo que se implementó.

B. Etiquetado

Se realizó un etiquetado manual de la atrofia peripapilar alfa y beta en las 300 imágenes, apoyado del etiquetador RoboFlow y CVAT [4], para poder extraer las etiquetas en el formato PNG. No obstante, cada imagen fue analizada y validada por un especialista en Glaucoma.

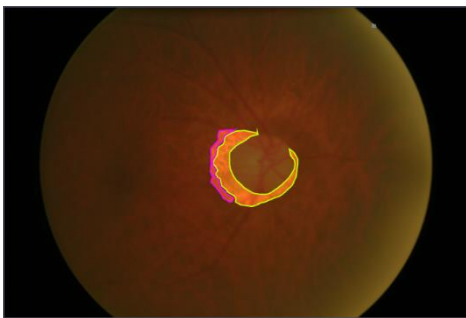


Fig. 7. Ejemplo de etiquetado manual de APP Alfa y Beta en imagen de fondo de ojo, mediante el etiquetador RoboFlow.



Fig. 8. Ejemplo de etiquetado manual de APP Alfa y Beta en imagen de fondo de ojo, mediante el etiquetador CVAT.

TABLA 1
PARÁMETROS ASIGNADOS A CADA MODELO PROPUESTO

Modelos	Núm. de Clases	Tamaño de las imágenes	Loss	Optimizer
VGG U-Net	255	320x640	Categorical crossentropy	Adam
ResNet50 U-Net	255	320x640	categorical crossentropy	Adam
MobileNet U-Net	255	320x640	Categorical crossentropy	Adam
FCN 8 MobileNet	255	320x640	Categorical crossentropy	Adam
VGG PSP-Net	255	320x640	Categorical crossentropy	Adam
ResNet PSP-Net	256	384x576	Categorical crossentropy	Adam
SEGNET	256	416x608	Categorical crossentropy	Adam
VGG_SEG NET	255	416x608	Categorical crossentropy	Adam
ResNet50_SEGNET	255	416x608	categorical crossentropy	Adam
FCN_8_V GG	255	416x608	Categorical crossentropy	Adam
FCN_8	255	416x608	Categorical crossentropy	Adam
ResNet50 FCN	255	416x608	Categorical crossentropy	Adam

C. Resultados

Después de la etapa de entrenamiento fue posible obtener diversos resultados para cada uno de los modelos propuestos, por lo cual a continuación se mostrarán los diferentes parámetros obtenidos de cada uno de los modelos.

TABLA 2
RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO VGG U-NET

Época	Loss	Accuracy
1	1.6928	0.8511
2	0.0320	0.9917
3	0.0277	0.9918
4	0.0248	0.9920
5	0.0226	0.9925

TABLA 3

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

RESNET50 U-NET

Época	Loss	Accuracy
1	0.0262	0.9925
2	0.0123	0.9954
3	0.0082	0.9968
4	0.0068	0.9973
5	0.0074	0.9973

TABLA 4

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

MOBILENET U-NET

Época	Loss	Accuracy
1	1.5008	0.9153
2	0.0143	0.9951
3	0.0124	0.9955
4	0.0073	0.9972
5	0.0067	0.9973

TABLA 5

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

FCN 8 MOBILENET

Época	Loss	Accuracy
1	0.2605	0.9677
2	0.0185	0.9940
3	0.0126	0.9952
4	0.0108	0.9959
5	0.0072	0.9971

TABLA 6

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

VGG PSP-NET

Época	Loss	Accuracy
1	0.1706	0.9778
2	0.0574	0.9905
3	0.0544	0.9905
4	0.0515	0.9906
5	0.0496	0.9906

TABLA 7

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

RESNET PSP-NET

Época	Loss	Accuracy
1	0.1445	0.9778
2	0.0448	0.9913
3	0.0401	0.9919
4	0.0365	0.9923
5	0.0337	0.9924

TABLA 8

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

SEGNET

Época	Loss	Accuracy
1	2.2374	0.8980
2	0.0600	0.9901
3	0.0503	0.9904
4	0.0466	0.9906
5	0.0443	0.9908

TABLA 9

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

VGG_SEGNET

Época	Loss	Accuracy
1	1.8878	0.8661
2	0.0523	0.9903
3	0.0477	0.9906
4	0.0436	0.9911
5	0.0423	0.9913

TABLA 10

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

RESNET50_SEGNET

Época	Loss	Accuracy
1	1.7630	0.9092
2	0.0377	0.9925
3	0.0295	0.9940
4	0.0251	0.9952
5	0.0241	0.9955

TABLA 11

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

FCN_8_VGG

Época	Loss	Accuracy
1	2.5575	0.9588
2	2.3693	0.9739
3	0.1184	0.9902
4	191.6377	0.9583
5	1.8462	0.9833

TABLA 12

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

FCN_8

Época	Loss	Accuracy
1	0.2691	0.9662
2	0.0489	0.9908
3	0.0447	0.9909
4	0.0397	0.9911
5	0.0368	0.9913

TABLA 13

RESULTADOS OBTENIDOS EN LA ETAPA DE ENTRENAMIENTO DEL MODELO

RESNET50 FCN

Época	Loss	Accuracy
1	0.1568	0.9705
2	0.0135	0.9949
3	0.0096	0.9963
4	2.9049	0.9934
5	0.8191	0.9891

D. Análisis de resultados

En esta sección se realizará una comparación de los resultados obtenidos en la implementación de cada uno de los modelos propuestos, por lo cual a continuación se podrá visualizar la siguiente tabla comparativa en donde se encuentran los diferentes parámetros como son la exactitud y la pérdida de cada uno de los modelos en la etapa de entrenamiento.

TABLA 14
COMPARACIÓN DE LOS RESULTADOS OBTENIDOS EN LA ETAPA DE
ENTRENAMIENTO DE LOS MODELOS PROPUESTOS

Modelos	Loss	Accuracy
VGG U-Net	0.0226	0.9925
ResNet50 U-Net	0.0074	0.9973
MobileNet U-Net	0.0067	0.9973
FCN 8 MobileNet	0.0072	0.9971
VGG PSP-Net	0.0496	0.9906
ResNet PSP-Net	0.0337	0.9924
SEGET	0.0443	0.9908
VGG_SEGNET	0.0423	0.9913
ResNet50_SEGNET	0.0241	0.9955
FCN_8_VGG	1.8462	0.9833
FCN_8	0.0368	0.9913
ResNet50 FCN	0.8191	0.9891

En términos de precisión, los modelos U-Net (VGG U-Net, ResNet50 U-Net, MobileNet U-Net) exhibieron los mejores resultados, siendo 0.9925, 0.9973 y 0.9973, respectivamente. Lo cual concuerda con un análisis visual comparativo de las predicciones de cada modelo con respecto a *ground truth*.

Los modelos FCN 8 MobileNet y ResNet50 SEGET presentaron de igual manera buenos resultados de precisión, siendo 0.9971 y 0.9955, respectivamente. Lo cual concuerda con un análisis visual comparativo de las predicciones de cada modelo con respecto a *ground truth*.

A pesar de que el modelo FCN 8 MobileNet presentó uno de los mejores resultados, el modelo FCN 8 VGG, presentó el peor resultado en precisión (0.9833, mostrando una ausencia de segmentación como se muestra en la Fig. 9.), lo que demuestra la importancia de emplear distintas versiones, con el fin de encontrar el modelo que mejor se adapte a la resolución de nuestro problema.

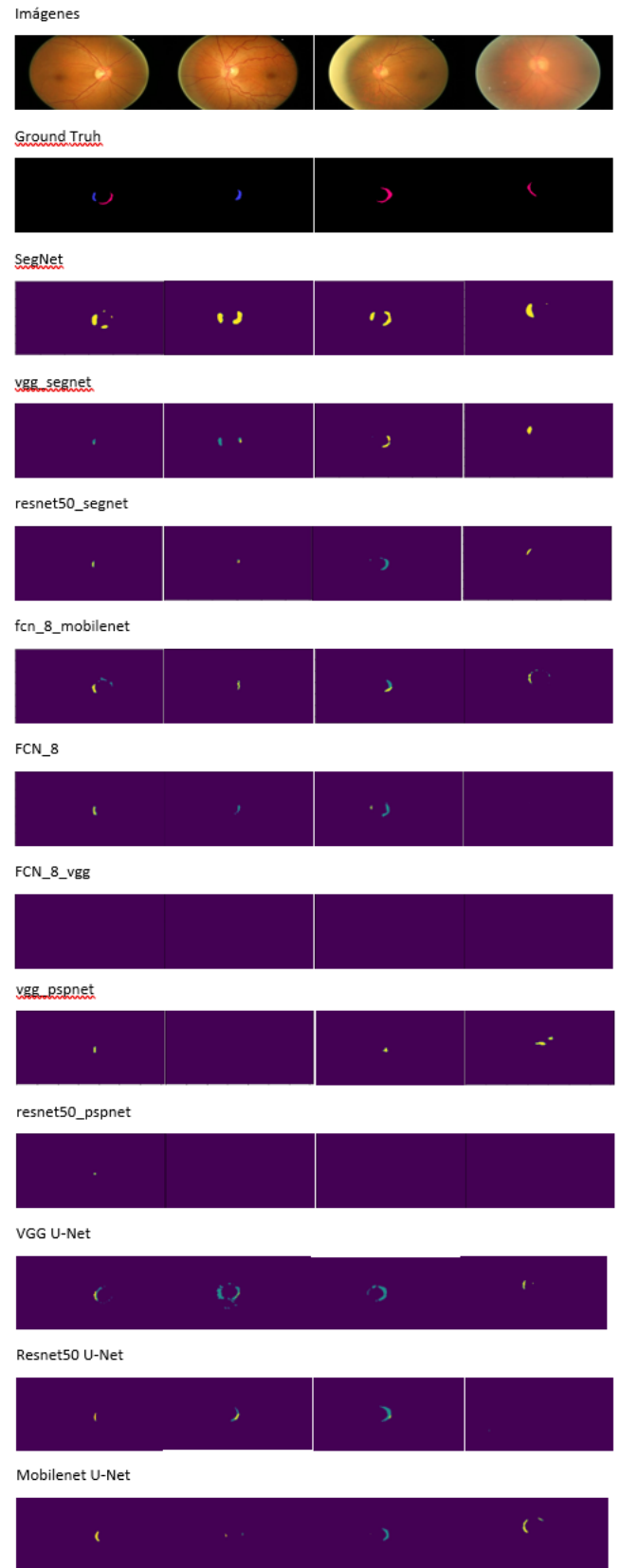


Fig. 9. Resultados de Segmentación de Atrofia Peripapilar Alfa y Beta.

E. Conclusiones

Se realizó un análisis de modelos de segmentación de Atrofia Peripapilar (APP) Alfa y Beta en imágenes de fondo de ojo, utilizando doce modelos diferentes que seguían las arquitecturas de U-Net, PSP-Net, SegNet y FCN.

Los resultados obtenidos demostraron una variabilidad en la segmentación de las imágenes, no obstante, U-Net expresó resultados que se acercan a la estructura morfológica y localización de los biomarcadores Alfa y Beta (Precision de 0.9925, 0.9973 y 0.997 para los modelos VGG U-Net, ResNet50 U-Net, MobileNet U-Net, respectivamente).

Particularmente PSP-Net exhibió resultados desalentadores, dado que en sus predicciones no identificaron ninguna de clases propuestas. Los resultados generan una dirección hacia el análisis de las variantes de U-Net para resolver el problema de la detección de APP Alfa y Beta, sirviendo, así este trabajo como precedente en la búsqueda de identificación de biomarcadores en imágenes de fondo de ojo que permitan brindar un diagnóstico oportuno de Glaucoma.

V. REFERENCIAS

- [1] IEEE Staff and IEEE Staff, *2010 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*.
- [2] D. Y. Tratamiento and E. Recomendaciones, *GUÍA DE PRÁCTICA CLÍNICA GPC DE GLAUCOMA PRIMARIO DE ANGULO ABIERTO*. [Online]. Available: <http://www.cenetec.salud.gob.mx/contenidos/gpc/catalogoMaestroGPC.html>
- [3] J. M. Molina Casado, "Detección y Localización Automática de Estructuras Anatómicas en Imágenes de Retina utilizando Técnicas de Visión Artificial."
- [4] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, "Artificial intelligence in retina," *Progress in Retinal and Eye Research*, vol. 67. Elsevier Ltd, pp. 1–29, Nov. 01, 2018. doi: 10.1016/j.preteyeres.2018.07.004.
- [5] "Oftalmología en la practica de la medicina general by [httpbooksmedicos.org](http://booksmedicos.org) (z-lib.org)".
- [6] M. Fallon, "Protocolo Meta-analisis Dispositivos de Imagen y Retinografía en Glaucoma", doi: 10.13140/RG.2.1.1850.8564.
- [7] M. Fingeret, F. A. Medeiros, R. Susanna, and R. N. Weinreb, "Five rules to evaluate the optic disc and retinal nerve fiber layer for glaucoma."
- [8] "Optic Nerve and Peripapillary Retina Characteristics in Primary Open Angle Glaucoma."
- [9] J. B. Jonas, P. Martus, F. K. Horn, A. Jünemann, M. Korth, and W. M. Budde, "Predictive factors of the optic nerve head for development or progression of glaucomatous visual field loss," *Investigative Ophthalmology and Visual Science*, vol. 45, no. 8, pp. 2613–2618, 2004, doi: 10.1167/iovs.03-1274.
- [10] I. An *et al.*, "Parapapillary Chorioretinal Atrophy in Patients With Ocular Hypertension Evaluation as a Predictive Factor for the Development of Glaucomatous Damage SI NCE," 1997. [Online]. Available: <http://archophth.jamanetwork.com/>
- [11] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," Sep. 2014, [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [12] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation," Nov. 2015, [Online]. Available: <http://arxiv.org/abs/1511.00561>
- [13] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015, [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [15] Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2881-2890).
- [16] <https://developers.arcgis.com/python/guide/how-pspnet-works/>