

Clasificación de A/V en imágenes de retina usando arquitectura inspirada en U-Net con ConvLSTM bidireccionales

M.G. Hernández-Ramírez, Saul Tovar-Arriaga,
Jesús Carlos Pedraza-Ortega, Efrén Gorrostieta-Hurtado

Universidad Autónoma de Querétaro,
Facultad de ingeniería,
México

{mghernandez1705, efren.gorrostieta}@gmail.com,
saul.tovar@uaq.mx, caryoko@yahoo.com,

Resumen. La retina es la única forma no invasiva de tener acceso al sistema cardiovascular, permite estudiar cambios en las estructuras que se encuentran en esta zona, como microaneurismas, hemorragias y los cambios que presentan los vasos retinianos, todos estos cambios se pueden utilizar para diagnosticar una amplia gama de enfermedades. Una manera accesible de estudiar estos cambios es por medio de imágenes de fondo de ojo. El objetivo de este trabajo es clasificar los vasos retinianos en arteria y vena. Actualmente las redes basadas en aprendizaje profundo han logrado los mejores resultados en la segmentación y clasificación de imágenes médicas. Una de las redes más usadas por sus buenos resultados para esta tarea es U-Net. En este trabajo, proponemos una modificación de U-Net con bloques ConvLSTM bidireccionales para la clasificación de A/V en imágenes de fondo de ojo. En lugar de una concatenación en los bloques de U-Net, empleamos ConvLSTM bidireccionales para combinar los mapas de características extraídos, además de solo usar dos bloques de la U-Net, esto hace el entrenamiento menos tardado. El modelo propuesto se evalúa en la base de datos DRIVE logrando un rendimiento competente con los métodos existentes.

Palabras clave: Clasificación de vasos retinianos, redes neuronales convolucionales, U-Net, ConvLSTM.

A/V Classification in Retina Images Using U-Net Inspired Architecture with Bidirectional ConvLSTM

Abstract. The retina is the only non-invasive way to the cardiovascular system, it permits to study changes in the structures found in this area, such as microaneurysms, hemorrhages and changes in the retinal vessels. These alterations in the retina can be used to diagnose a wide range of

diseases. An accessible way to study these changes is through fundus images. This paper focuses on the classification of retinal vessels in arteries and veins. Networks based on deep learning have, in recent years, achieved the best results in segmentation and classification of medical images. One of the more commonly used networks, because of its high performance, is U-Net. In this work, a modified U-Net, with bidirectional ConvLSTM blocks is proposed for A / V classification in fundus images. Instead of concatenating across U-Net blocks, two-way ConvLSTM are used to combine the extracted feature maps. In addition, using only two U-Net blocks makes training less time consuming. The proposed model was evaluated using the DRIVE database, achieving competent performance with existing methods.

Keywords: Classification of retinal vessels, convolutional neural networks, U-Net, ConvLSTM.

1. Introducción

Algunas características del fondo del ojo proporcionan información sobre el estado de salud del ser humano, como problemas oculares y problemas relacionados a enfermedades crónicas [1], estas enfermedades provocan cambios en la estructura y función del ojo. Una de las enfermedades que está asociada a los cambios en el fondo de ojo es la hipertensión [8]. Uno de los cambios que provoca esta enfermedad es el estrechamiento de las arterias de menor tamaño, este cambio en la vasculatura puede evaluarse mediante las relaciones del diámetro de arteria y vena (AVR) [2], también se ven afectados los nervios retinianos, corioideo y óptico, estos experimentan una serie de cambios fisiopatológicos en respuesta a la presión arterial elevada, teniendo como resultado una serie de signos clínicos llamados retinopatía hipertensiva [3].

Una de las principales anomalías en la retinopatía hipertensiva es el estrechamiento de vasos retinianos y el cruce arteriovenoso patológico [4]. Estos cambios están relacionados no solo con los niveles actuales de presión arterial del paciente, sino también con los niveles medidos en el pasado, lo que sugiere que estos signos son marcadores persistentes de daño hipertensivo crónico [7]. El estrechamiento de vasos sanguíneos también se puede usar para predecir el desarrollo de la hipertensión en individuos normotensos. Por lo tanto, el estrechamiento podría ser un marcador temprano de hipertensión [9].

Algunas otras enfermedades también afectan la circulación sanguínea y como consecuencia pueden causar un AVR anormal como [5]. Así mismo, el AVR se puede utilizar para predecir el riesgo de muerte por enfermedad coronaria y cerebrovascular [10]. Estos cambios vasculares a menudo preceden al inicio de los síntomas y, si se detectan, pueden permitir tratamientos preventivos que pueden reducir el riesgo de pérdida de visión. Para abordar estos problemas se estudia la vasculatura de la retina mediante la segmentación, clasificación y medición de los vasos.

La retina es la única forma no invasiva para poder visualizar el sistema cardiovascular, es una ventana que permite estudiar los cambios de las carac-

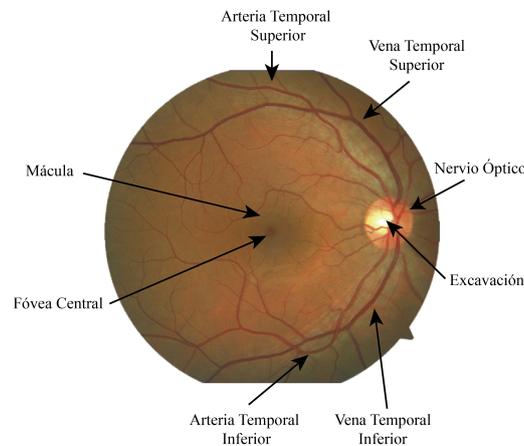


Fig. 1. Imagen sana de fondo de ojo, *Mexican Advanced Imaging Laboratory for Ocular Research (MAILOR)*.

terísticas en esta estructura. Esto es posible utilizando una variedad de métodos de imágenes, que incluyen fotografía de fondo de ojo, angiografía con fluoresceína y tomografía de coherencia óptica. La fotografía de fondo de ojo es la manera preferida para obtener la información, tiene un menor costo que las anteriores y es fácil de usar. En la Fig. 1 se muestra una imagen de fondo de ojo en condición sana donde se pueden ver las partes más importantes. En la retina, todas las arterias y venas emanan del nervio óptico.

Por lo tanto, todas las arterias están conectadas entre sí, y también todas las venas, se identifican las venas por un color más oscuro al de las arterias porque transportan sangre desoxigenada al corazón [6]. Las venas normalmente tienen un calibre mayor al de las arterias, teniendo una relación AVR normal de 2:3. La clasificación de vena y arteria es una tarea realizada manualmente por especialista, es una tarea tardada y laboriosa. En una imagen de alta resolución, a menudo se pueden detectar más de cien vasos o segmentos de vasos, muchos de los cuales son ambiguos y requieren una cuidadosa visualización para clasificarlos. Este artículo se centra en la segmentación y clasificación automática de los vasos sanguíneos de imágenes de retina.

En los últimos años, el uso de redes neuronales ha mostrado avances significativos y resultados prometedores en la detección de patrones en imágenes, específicamente las redes neuronales convolucionales (CNN). Las CNN son particularmente apropiadas para la clasificación y segmentación de imágenes [11], [12]. En 2015 Ronneberger *et al.* propusieron una red totalmente convolucional en forma de U [13] para la segmentación de imágenes médicas. La arquitectura U-Net propuesta ha demostrado una mejor capacidad para segmentar estructuras más finas, como la membrana celular, con muy pocas imágenes de entrenamiento, ha tenido éxito y se han propuesto diferentes variaciones.

En este estudio, se modifica la arquitectura de la U-Net y se propone una U-Net con 2 bloques y conectada con ConvLSTM bidireccionales inspirado por [21].

2. Antecedentes

Los métodos de clasificación de los vasos de la retina: arteria y vena, encontrados en la literatura se pueden clasificar en dos clases, en métodos basados en árboles y en características y los más populares en los últimos años, basados en aprendizaje profundo. En la Tabla 1 se muestra un resumen de algunos de los métodos usados en los últimos años.

Tabla 1. Descripción general de trabajos para la clasificación de vena y arteria de imágenes de retina.

Autor	Método	Métrica
Estrada <i>et al.</i> [14] (2015)	Estimación de topología, teoría de grafos	Exactitud = 0.935
Xu <i>et al.</i> [15] (2017)	Clasificación de píxeles, clasificador kNN.	Exactitud = 0.923
Welikala <i>et al.</i> [18] (2017)	CNN	Exactitud = 0.919
Xu <i>et al.</i> [16] (2018)	CNN U-Net	Exactitud = 0.832
Girard <i>et al.</i> [17] (2019)	CNN adaptación U-Net	Exactitud = 0.924

En la primera clase encontramos trabajos como el de Estrada *et al.* [14] donde utilizó un enfoque de estimación de topología basado en grafos, que estima la topología más probable para las etiquetas de arteria y de vena. Buscando posibles árboles etiquetados, usando algoritmos de búsqueda heurísticos.

El algoritmo fue probado con 4 bases de datos: WIDE, AV-INSPIRE, AV-DRIVE, CT-DRIVE. Otro trabajo es el propuesto por Xu *et al.* [15] donde propone una técnica de clasificación arteria y vena donde la imagen de retina es regularizada y normalizada para extraer múltiples características de los vasos sanguíneos. Se extraen diferentes características basadas en el color y textura para cada píxel y se clasifican en arteria y vena usando un clasificador kNN. Esta técnica fue probada en la base de datos DRIVE y logró una exactitud del 0.923.

Actualmente y gracias al poder de computo con el que hoy contamos y las bases de datos disponibles, los métodos más recurridos están basados en aprendizaje profundo. Un beneficio de estos modelos es su capacidad para realizar la extracción automática de características de los datos sin procesar, también llamada aprendizaje de características.

Un método de esta clasificación es el propuesto por Welikala *et al.* [18] donde propuso un método basado en el aprendizaje profundo para la clasificación de arterias y venas en toda la imagen de la retina.

La red propuesta esta formada por 6 capas, de las cuales 3 son capas convolucionales y 3 son capas completamente conectadas, la capa ReLU siguió a cada capa convolucional y las dos primeras capas completamente conectadas. Este método se evaluó en el conjunto de datos del Biobanco del Reino Unido y en la base de datos publica DRIVE.

Xu *et al.* [16] propuso una red neuronal convolucional inspirada en la arquitectura U-Net para la segmentación de venas y arterias en imágenes a color de la retina. La red fue entrenada y probada en el conjunto de datos publica DRIVE. La sensibilidad, especificidad y exactitud es de 0.870, 0.980, 0.832 respectivamente. Las tasas de clasificación errónea registradas fueron de 0.098 para venas y 0.237 para arterias. Girard *et al.* [17] Propone un método de clasificación A-V usando una CNN simple que clasificaba los píxeles en arteria y vena, la red está compuesta de cinco capas convolucionales de codificación y cinco capas convolucionales de decodificación junto con conexiones de derivación. La arquitectura propuesta es una derivación de la U-Net. Realiza un preprocesamiento a la imagen donde corrigen problemas de iluminación no uniforme, obteniendo 3 canales diferentes a los originales (RGB), concatenan los 3 canales nuevos a los 3 originales, teniendo una imagen de 6 canales. El método se probó en dos bases de datos DRIVE y MESSIDOR.

3. Metodología

En esta sección, se describe a detalle el conjunto de datos usados, el preprocesamiento realizado a las imágenes y la arquitectura usada.

Para la implementación de este trabajo se utilizó la base de datos DRIVE, la cual es una base de datos gratuita y la más usada para trabajos de segmentación y clasificación de vasos sanguíneos de retina, así como para la búsqueda y detección de diferentes patologías que pueden identificarse en esta zona.

3.1. Preprocesamiento

Las imágenes de fondo de ojo pueden tener una amplia variación en la pigmentación, debido a problemas con la iluminación no uniforme y a la mala exposición al momento de tomar la fotografía, lo que puede influir el rendimiento de la CNN. Realizar un preprocesamiento mejora la calidad de las imágenes y permitir que la CNN detecte con mayor facilidad las características buscadas.

El primer paso es la normalización de datos, la normalización es usada para escalar un atributo dentro de un rango específico, facilitar y acelerar el entrenamiento de las redes. Las CNN no funcionan bien cuando los atributos numéricos de entrada tienen escalas muy diferentes. En este trabajo se utilizan dos técnicas de normalización, z-score y min-max.

La normalización z-score es útil cuando se tienen valores atípicos en un atributo:

$$x' = \frac{x - \mu_A}{\sigma_A}, \quad (1)$$

donde μ_A y σ_A son la media y desviación estándar.

En la normalización min-max se hace una transformada lineal de los datos, los valores se desplazan y reescalan para que terminen en un rango de 0 a 1:

$$x' = \frac{x - \min(A)}{\max(A) - \min(A)}, \quad (2)$$

donde x' es el valor normalizado del valor original x relacionado con los valores mínimos y máximos de un atributo A .



Fig. 2. Resultado de preprocesamiento de imagen. A la izquierda imagen original; a la derecha imagen preprocesada.

El segundo paso es la corrección de contraste y brillo. Para el primero se aplica una ecualización de histograma adaptativo limitada por contraste (CLAHE). Para mejorar de brillo se utiliza una corrección gamma. El resultado de este paso se puede ver en la Fig. 2.

3.2. Aumento de datos

Las bases de datos para la segmentación y clasificación de vasos retinianos suelen no ser muy grandes debido a su naturaleza, para abordar este problema se hace un aumento de datos para mejorar el entrenamiento de la red neuronal.

Se implementa un aumento de datos mediante la extracción de parches aleatorios de cada imagen, haciendo giros vertical y horizontal aleatoriamente en cada parche. Con la intención de aumentar artificialmente los datos y agregar más variabilidad al conjunto de datos de entrenamiento. En la Fig. 3 podemos ver algunos parches extraídos con sus respectivas etiquetas.

3.3. Arquitectura

En este trabajo se usan bloques ConvLSTM bidireccionales y parte reducida de la estructura U-net como se muestra en la Fig. 4. La estructura original, de

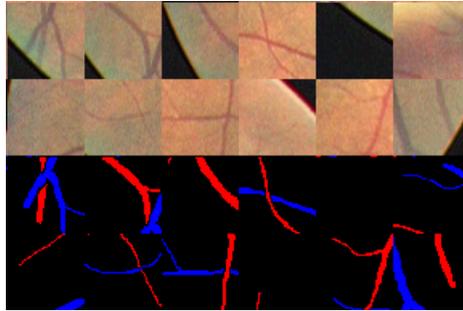


Fig. 3. Ejemplo de extracción de parches. Dos filas superiores imágenes; dos filas inferiores etiquetas.

U-net está compuesta por una parte de codificación, donde se captura tanto la información semántica como la contextual a medida que se duplica el número de filtros. La segunda parte está compuesta por decodificación, donde se restaura la información espacial y se fusionan las características obtenidas en los niveles más bajos.

La estructura propuesta consta de 2 niveles, en los que se aplican una doble convolución con un filtro de 3x3, cada una seguida de una unidad lineal rectificadora (ReLU), una operación de agrupación máxima de 2x2 con paso de 2 para disminuir la resolución. La imagen de entrada es de 3 canales un tamaño de 64x64. En el primer nivel el número de canales utilizados fue de 64 y el tamaño de la imagen 32x32 tras dobles convoluciones. En el segundo nivel el número de canales utilizados fue de 128 y el tamaño de la imagen 16x16. La parte de decodificación sigue la estructura anterior, el número de canales fue el mismo, pero en sentido contrario. También se añaden las características que se encuentran en cada uno de los diferentes niveles, concatenando la salida del nivel de contracción con la entrada de expansión.

Los mapas de características correspondientes a la capa de codificación tienen una resolución más alta, mientras que los mapas de característica correspondientes a la decodificación contienen más información semántica. En lugar de una simple concatenación en esta propuesta se hace la combinación de estos dos mapas de características con ConvLSTM bidireccional inspirado en [21]. Después de cada capa de convolución se aplica una capa de normalización por bloque y la función de activación ReLU.

4. Materiales y experimentación

Para la presente investigación de segmentación de vena y arteria de imágenes de fondo de ojo, se trabaja con la base de datos publicada por Qureshi *et al.* [19] donde un oftalmólogo capacitado etiquetó manualmente los vasos del conjunto de pruebas como arteria, vena o superposición.

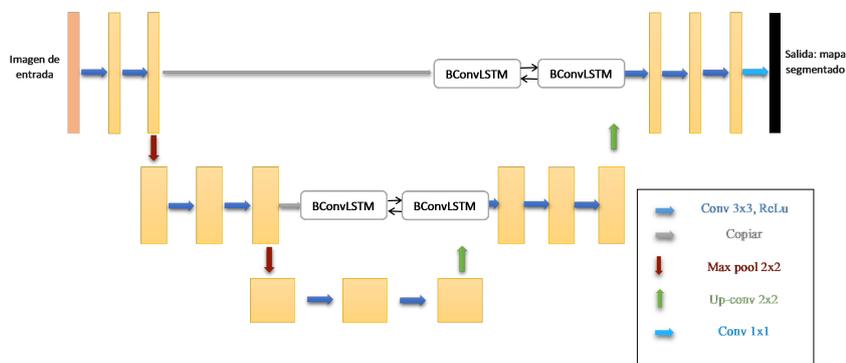


Fig. 4. Arquitectura propuesta.

Para esta clasificación se usó la base de datos publica DRIVE [20] que incluye un conjunto de cuarenta fotografías a color del fondo de ojo, obtenidas de un programa de detección de retinopatía diabética. En este conjunto de datos, siete de las cuarenta imágenes contienen patologías anormales como exudados, hemorragias y epitelio pigmentario. Las imágenes tienen un tamaño de 565x584 píxeles, se adquirieron utilizando una cámara Canon CR5 no midriática 3-CCD con un campo de visión (FOV) de 45 grados. La base de datos DRIVE se divide en dos conjuntos, el conjunto de entrenamiento y el conjunto de prueba, cada uno con veinte imágenes. El conjunto de prueba esta segmentado manualmente por dos observadores, el primer observador es aceptado como verdad fundamental.

Se utilizó optimización ADAM, con una tasa de aprendizaje de 0.001. En total, se realizaron 100 épocas, con tamaño de lote de 16, este es un hiperparámetro importante, es el número de muestras que se propagarán a través de la red. La métrica que se evalúa para decidir si la CNN está o no mejorando, es la función de perdida.

Para la validación se usó validación cruzada, con 5 particiones. De las 40 imágenes de la base de datos, 24 se usaron para entrenar, 6 para validar y 10 para prueba, estos resultados se muestran en la Tabla 2.

También se hicieron pruebas con la configuración predeterminada de la base de datos DRIVE, donde se usa 20 imágenes para entrenamiento, las cuales se dividen 16 para entrenar, 4 para validar y las 20 restantes de prueba.

Toda la implementación de este trabajo se desarrolló en TensorFlow y Keras, ambas librerías se especializan en aprendizaje profundo, y se usa Python como lenguaje de programación.

El equipo utilizado en las pruebas cuenta con un procesador Intel (R) Core (TM) i5-8400 CPU@2.80 GHz, con 16 GB de RAM y una tarjeta gráfica NVIDIA GeForce GTX 1070 con 8 GB de VRAM.

5. Resultados

Las métricas usadas para la evaluación del desempeño del modelo son sensibilidad, especificidad y exactitud. El cálculo de estas métricas se basa en cuatro posibles interpretaciones de los datos:

- Verdadero positivo (TP): la etiqueta de la muestra es positiva y se clasifica como tal.
- Verdadero negativo (TN): la etiqueta de la muestra es negativa y se clasifica como tal.
- Falso positivo (FP): la etiqueta de la muestra es negativa, pero se clasifica como positiva.
- Falso negativo (FN): la etiqueta de la muestra es positiva, pero se clasifica como negativa.

En función de estos valores, se pueden calcular las métricas mencionadas anteriormente de acuerdo a las ecuaciones 3, 4, 5:

- Sensibilidad: También llamada tasa de Verdaderos Positivos. Es la proporción de casos positivos que fueron correctamente identificados:

$$\text{Sensibilidad} = \frac{TP}{TP + FN}. \quad (3)$$

- Especificidad: También llamada Tasa de Verdaderos Negativos. Casos negativos que se clasificaron correctamente:

$$\text{Especificidad} = \frac{TN}{TN + FP}. \quad (4)$$

- Exactitud: Mide cuántas observaciones, tanto positivas como negativas, se clasificaron correctamente. Es la proporción de instancias negativas que se clasifican incorrectamente como positivas:

$$\text{Exactitud} = \frac{TP + TN}{TP + FP + FN + TN}. \quad (5)$$

En la Tabla 2 podemos ver una comparación de trabajos similares basados en el conjunto de datos DRIVE, donde se observa sensibilidad, especificidad y exactitud, métricas descritas anteriormente.

En la Fig. 5 se puede ver un ejemplo de un resultado obtenido del conjunto de datos DRIVE. Los falsos negativos son especialmente vasos retinianos pequeños que no se alcanzan a detectar.

En los vasos mayores a 5 píxeles, los falsos positivos se ubican en los bordes de los vasos. Existe una confusión entre arteria y vena o viceversa en los vasos mas pequeños o en los puntos de intersección entre estos dos.

En la Fig. 6 se proporcionan las curvas *Receiver Operating Characteristic* (ROC) y *Precision-Recall* (PR) de la arquitectura propuesta para los resultados de rendimiento de la segmentación arteria y vena. La curva ROC muestra el

Tabla 2. Comparación de resultados de clasificación Vena/Arteria con estado del arte en base de datos DRIVE.

Método	Sensibilidad	Especificidad	Exactitud
Estrada <i>et al.</i> [14] (2015)	0.9170	0.9170	0.9100
Xu <i>et al.</i> [15] (2017)	-	-	0.9230
Welikala <i>et al.</i> [18] (2017)	0.7810	-	0.9190
Xu <i>et al.</i> [16] (2018)	0.8700	0.9800	0.8320
Girard <i>et al.</i> [17] (2019)	0.9370	0.9290	0.9493
Propuesto	0.8000	0.9780	0.9550

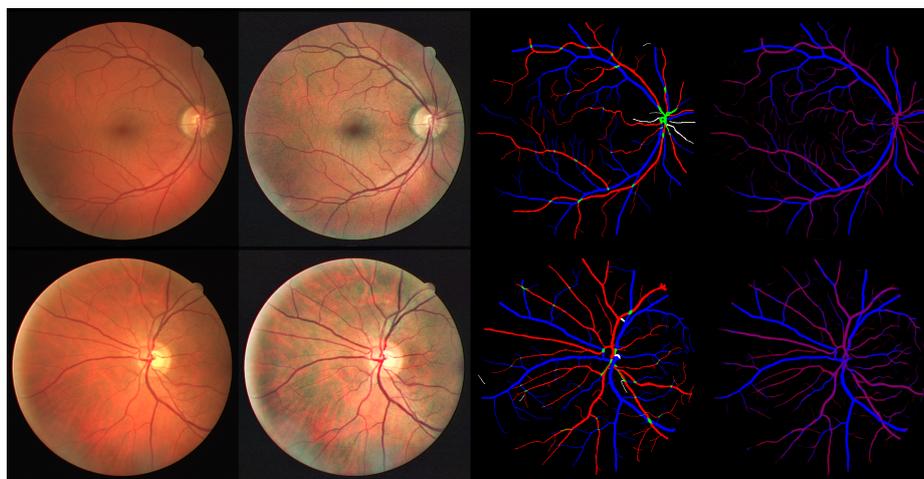


Fig. 5. Resultado de clasificación de arteria y vena de dos imágenes. Iniciando con la primera a la izquierda, imagen original; segunda imagen, preprocesamiento; tercera imagen verdad fundamental; última imagen salida clasificada.

rendimiento de un modelo de clasificación, el área bajo la curva es de 0.9722 para este modelo. La curva traza la sensibilidad contra especificidad, parámetros de tasa de verdaderos positivos y tasa de falsos positivos. La curva *Precision-Recall* tiene un área bajo la curva de 0.9003, esta gráfica permite ver a partir de qué *recall* se tiene una degradación de la precisión y viceversa, esta curva se centra principalmente en el rendimiento de la clase positiva.

6. Conclusiones

El estado patológico del fondo de ojo también puede afectar la clasificación de arteria y vena, específicamente cuando existe una lesión obstruyendo los vasos, cuando las lesiones son pequeñas no se tiene problema y se clasifica correctamente pero cuando las lesiones son grandes y en un color contrastante al resto de la retina los vasos quedan ausentes o clasificados erróneamente. El preprocesamiento de las imágenes de entrada es de suma importancia ya que corrige algunos

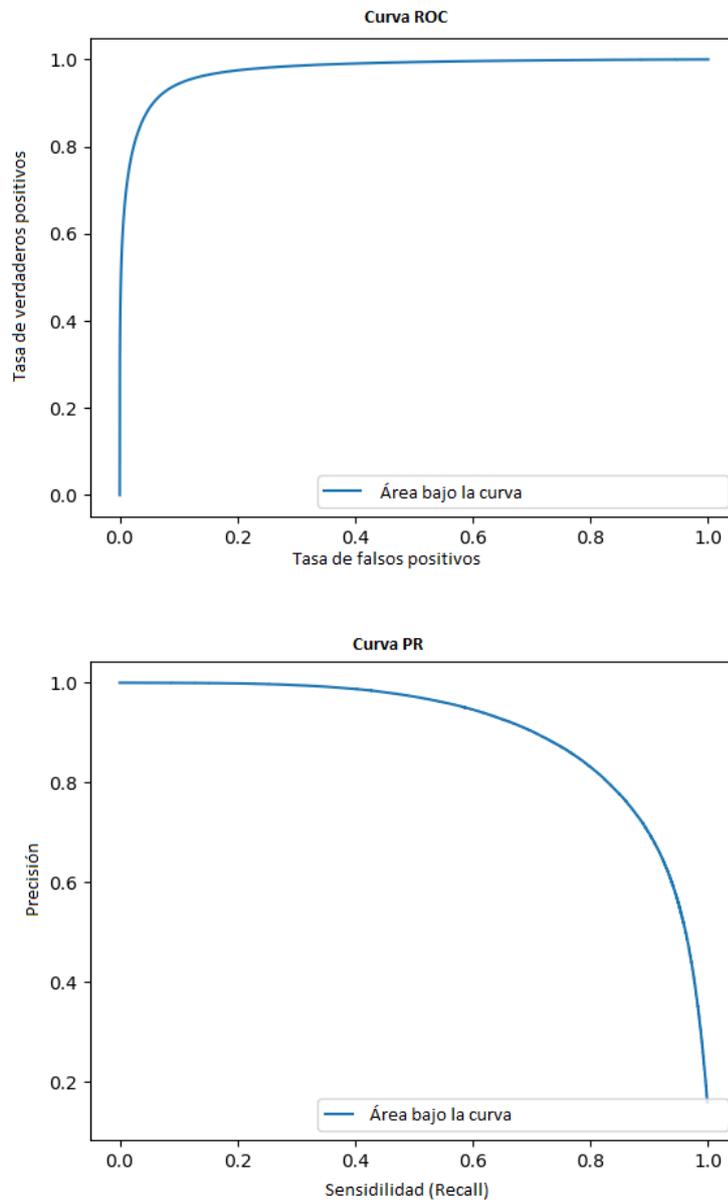


Fig. 6. Curvas de rendimiento de la arquitectura propuesta en la segmentación A/V en DRIVE.

errores mencionados anteriormente, resaltando las características deseadas en la imagen.

Acompañado del preprocesamiento, un aumento de datos es importante en este trabajo, la base de datos DRIVE cuenta con pocas imágenes para el entrenamiento de una CNN, lo cual implementar aumento de datos por extracción de parches y girarlos aleatoriamente ayuda a un mejor entrenamiento.

Tener una cantidad mayor de datos etiquetados mejoraría la clasificación, el modelo aprendería más información de la estructura provocando que el modelo sea robusto para diferentes patologías severas que compliquen la clasificación de los vasos.

El trabajo propuesto acelera el etiquetado manual de arterias y venas, proporcionando un etiquetado automático que requiere correcciones mínimas manualmente por parte de un experto. Los resultados en este estudio son muy similares o superiores a los resultados de los métodos de mayor rendimiento, se simplificó la complejidad de métodos similares con resultados parecidos. Al usar bloques ConvLSTM bidireccionales ayuda a combinar los mapas de características extraídos hace la arquitectura más susceptible.

El trabajo futuro se centrará en encontrar nuevos biomarcadores y mediciones clínicas utilizando la clasificación de arteria y vena para detectar los cambios en los vasos y con ellos poder detectar signos en temprana edad en diversas patologías.

Agradecimiento. Este trabajo ha sido apoyado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) Agradecemos a la Dra. Renata García Franco por el apoyo dado para la realización de esta investigación como al *Mexican Advanced Imaging Laboratory for Ocular Research* (MAILOR).

Referencias

1. Kanski, J. J., Bowling, B.: Clinical ophthalmology: a systematic approach. Elsevier Health Sciences, Location (2011)
2. Cheung, C. Y. L., Ikram, M. K., Sabanayagam, C., Wong, T. Y.: Retinal microvasculature as a model to study the manifestations of hypertension. *Hypertension*, 60(5), pp. 1094–1103 (2012)
3. Wong, T. Y., Ming, C., Cheung, G., Larsen, M., Sharma, S.: Diabetic retinopathy. *Medicine*, 77–85 (2016)
4. Sharma, K., Kanaujia, V., Mishra, P., Agarwal, R., and Tripathi, A. .: Hypertensive retinopathy. *Clinical queries, nephrology*, 2(3), pp. 136–139 (2013)
5. Kondermann, C., Kondermann, D., and Yan, M.: Blood vessel classification into arteries and veins in retinal images. In: *Medical Imaging 2007: Image Processing*, Vol. 6512, p. 651247 (2007)
6. Yu, H., Barriga, S., Agurto, C., Nemeth, S., Bauman, W., Soliz, P.: Automated retinal vessel type classification in color fundus images. In: *Medical Imaging 2013: Computer-Aided Diagnosis*, International Society for Optics and Photonics, Vol. 8670, p. 86700P (2013)
7. Sharrett, A. R., Hubbard, L. D., Cooper, L. S., Sorlie, P. D., Brothers, R. J., Nieto, F. J., ... y Klein, R.: Retinal arteriolar diameters and elevated blood pressure: the Atherosclerosis Risk in Communities Study. *American journal of epidemiology*, 150(3), pp. 263–270 (1999)

8. Wong, T. Y., Klein, R., Klein, B. E., Tielsch, J. M., Hubbard, L., Nieto, F. J.: Retinal microvascular abnormalities and their relationship with hypertension, cardiovascular disease, and mortality. *Survey of ophthalmology*, 46(1), pp. 59–80 (2001)
9. Leung, H., Wang, J. J., Rochtchina, E., Wong, T. Y., Klein, R., Mitchell, P.: Impact of current and past blood pressure on retinal arteriolar diameter in an older population. *Journal of hypertension*, 22(8), pp. 1543–1549 (2004)
10. Wang, J. J., Liew, G., Klein, R., Rochtchina, E., Knudtson, M. D., Klein, B. E., ... and Mitchell, P.: Retinal vessel diameter and cardiovascular mortality, pooled data analysis from two older populations. *European heart journal*, 28(16), pp. 1984–1992 (2007)
11. Krizhevsky, A., Sutskever, I., Hinton, G. E.: Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp. 1097–1105 (2012)
12. Badar, M., Haris, M., Fatima, A.: Application of deep learning for retinal image analysis: A review. *Computer Science Review*, 35, p. 1002035 (2020)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net, Convolutional networks for biomedical image segmentation. In: *International Conference on Medical image computing and computer assisted intervention*. Springer, Cham, pp. 234–241 (2015)
14. Estrada, R., Allingham, M. J., Mettu, P. S., Cousins, S. W., Tomasi, C., Farsiu, S.: Retinal artery-vein classification via topology estimation. *IEEE transactions on medical imaging*, 34(12), pp. 2518–2534 (2015)
15. Xu, X., Ding, W., Abramoff, M. D., Cao, R.: An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image. *Computer methods and programs in biomedicine*, 141, pp. 3–9 (2017)
16. Xu X., Tan T., Xu F.: An improved U-net architecture for simultaneous arteriole and venule segmentation in fundus image. *Medical Image Understanding and Analysis, MIUA 2018, Communications in Computer and Information Science*, Springer, vol. 894, pp. 333–340 (2018)
17. Girard, F., Kavalec, C., Cheriet, F.: Joint segmentation and classification of retinal arteries/veins from fundus images. *Artificial intelligence in medicine*, 94, pp. 96–109 (2019)
18. Welikala, R. A., Foster, P. J., Whincup, P. H., Rudnicka, A. R., Owen, C. G., Strachan, D. P., Barman, S. A.: Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. *Computers in biology and medicine*, 90, pp. 23–32 (2017)
19. Qureshi, T. A., Habib, M., Hunter, A., Al-Diri, B.: A manually-labeled, artery/vein classified benchmark for the DRIVE dataset. In: *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pp. 485–488 (2013)
20. J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken: Ridge-based vessel segmentation in color images of the retina. *IEEE Trans Med Imaging*, (vol. 23, no. 4, pp. 501–509.) (2004)
21. Azad, R., Asadi-Aghbolaghi, M., Fathy, M., Escalera, S.: Bi-Directional ConvLSTM U-Net with Densely Connected Convolutions. In: *Proceedings of the IEEE International Conference on Computer Vision Workshops* (2019)