



Evaluación de modelos convolucionales en la segmentación de vasos sanguíneos cerebrales

Oliver Jonathan Quintana-Quintana¹, Marco Antonio Aceves-Fernández¹,
Iván Villanueva-Naquid², Saúl Tovar-Arriaga¹ ✉

¹Universidad Autónoma de Querétaro, Facultad de Ingeniería, Querétaro, México,

²Universidad Autónoma de San Luis Potosí, Facultad de Medicina, San Luis Potosí, México.
✉ saul.tovar@uaq.mx

Resumen

La segmentación de vasos sanguíneos cerebrales es una tarea de gran importancia en distintas aplicaciones dentro del área médica; los tratamientos vasculares requieren de una correcta identificación de las diversas malformaciones arteriovenosas, y en general, la identificación de arterias, venas, y capilares son un punto de gran interés para un correcto diagnóstico de padecimientos cerebrovasculares, como tumores y aneurismas, así como en la planificación de intervenciones quirúrgicas para extracción de tumores, estimulación cerebral profunda, entre otras, procedimientos en los cuales es crucial la identificación de los vasos sanguíneos, con el propósito de etiquetar y evitar regiones de riesgo, de tal forma que se minimiza el riesgo de ocasionar severos daños en el paciente al eliminar dichas estructuras de la planeación de la intervención. Así mismo, el proceso de segmentación puede llegar a consumir demasiado tiempo para el experto, debido a que la segmentación manual requiere del etiquetado imagen a imagen dentro de un volumen dado. En este trabajo se evalúan distintos modelos convolucionales, basados en arquitecturas encoder-decoder, que han sido adoptados recientemente en diversas aplicaciones dentro de la visión artificial por computadora, enfocando su uso a la tarea de segmentación de vasos sanguíneos en imágenes de microscopía por excitación de fotones, con el fin de proveer una herramienta de asistencia a especialistas en la realización de dicho procedimiento. Junto con los modelos evaluados, se hace énfasis en el uso de técnicas de regularización para evitar el sobre ajuste de los mismos, así como lograr una mejor generalización en los mapas de segmentación generados, mostrando la manera en que mejoran el proceso de entrenamiento de un modelo convolucional.

Palabras clave: Redes neuronales convolucionales, segmentación de vasos sanguíneos, microscopía, aprendizaje profundo.

Abstract

Within the medical area, the segmentation of cerebral blood vessels is a key procedure in a variety of applications. Vascular treatments require a proper identification of malformations in blood vessels and capillaries, where this task is of huge interest for medical diagnosis of cerebrovascular diseases, such as the identification of tumors and aneurysms, as well as in trajectory planning for surgical interventions in deep brain stimulation and tumor extraction, where the correct identification of blood vessels is key in order to avoid dangerous zones within the brain and reduce the damage done to the patient, while those structures are dismissed from the intervention planning. Segmentation is a very time-consuming task for an expert to perform due to the high amount of individual samples that need to be manually-annotated within a medical imaging volume. In this work, three different convolutional models are evaluated, which are based on encoder-decoder architectures widely adopted for several computer vision applications, while focusing vasculature segmentation task in cerebral multiphoton



microscopy images, providing an assistance tool for physicians while performing such task. Along with the evaluated models, regularization techniques are implemented to prevent such models to over fit, as well as improve the generalization for the segmentation maps obtained, proving how these methods are powerful tools while training convolutional models.

Keywords: Convolutional neural networks, blood vessel segmentation, multiphoton microscopy, deep learning.

1. Introducción

Los algoritmos de aprendizaje automático pertenecen a un área especializada dentro de la Inteligencia Artificial, los cuales dotan de “inteligencia” a las computadoras encontrando patrones entre los datos. Se ha propuesto una gran cantidad de algoritmos que intentan simular la inteligencia humana. Dentro del campo de la visión artificial, ha surgido una arquitectura de red neuronal especializada, las redes neuronales convolucionales (del inglés *Convolutional Neural Network (CNN)*), que intentan emular el procesamiento neuronal que realiza el sistema visual de los seres vivos. Las CNN son consideradas una de las mejores técnicas para el procesamiento de imágenes, ya sea para el reconocimiento y/o detección de objetos, segmentación, entre otras [1]. Se han propuesto diversos métodos para la extracción de información sobre regiones cerebrales de interés, como lo son los vasos sanguíneos, a partir de la segmentación de imágenes médicas, de los cuales se describen algunos a continuación.

Debido a la naturaleza de las técnicas de adquisición de imágenes de angiografía de resonancia magnética cerebrales, los volúmenes de estas son representados por voxeles para conformar estructuras 3D de la cabeza del paciente, comúnmente representados por imágenes bidimensionales apiladas, donde voxeles adyacentes comparten ciertas propiedades debido a la cercanía de la misma, dependiendo también de la resolución del resonador para cada voxel. Esta propiedad es utilizada para extraer características relevantes del ROI (del inglés *Region Of Interest*) del volumen en el método propuesto por [2], siendo estas los vasos sanguíneos cerebrales, a partir de un autoencoder convolucional tridimensional cuya principal aportación es la reducción del ruido de las imágenes de MRA mediante dicha arquitectura para una extracción de vasos sanguíneos más simple, cuyo resultado es un mapa de segmentación de estas estructuras. Este método reporta un índice Dice Score de 73.71, contando con un conjunto de datos de 49 volúmenes de Angiografías de Resonancia Magnética (MRA).

Las arquitecturas encoder-decoder han mostrado un incremento en el rendimiento de los modelos convolucionales. El modelo “U-Net” propuesto por Ronneberger et al. [3] ha establecido un punto de inflexión en la segmentación de imágenes para el área médica, demostrando que una red neuronal convolucional profunda puede ser entrenada en conjuntos de datos relativamente pequeños, logrando buenos resultados en comparación con otras arquitecturas convolucionales. Esto es de relevancia debido a la dificultad de acceder a conjuntos de datos etiquetados por especialistas, los cuales suelen contener un número reducido de elementos.

Considerando lo anterior, diversas arquitecturas convolucionales basadas en la “U-Net” se han propuesto para diversas aplicaciones, dentro y fuera del área médica. Partiendo de esto, Livne et al. [4] proponen su modelo basado en dicha arquitectura, realizando una reducción en su número de parámetros para realizar un entrenamiento más rápido sobre el conjunto de imágenes de angiografías cerebrales para la segmentación de vasos sanguíneos en pacientes con padecimientos cerebrovasculares, obteniendo resultados remarcables de hasta 92.1 en el índice de Dice en sus experimentos.

Teikari et al. [5] desarrollaron un modelo híbrido 2D-3D de red neuronal convolucional para la segmentación de vasos sanguíneos en imágenes de microscopia de fotones, utilizando únicamente un CPU como acelerador del framework, mostrando ciertos modelos de redes convolucionales,



optimizados de manera apropiada, pueden no requerir de hardware más especializado, como GPUs, para la etapa de entrenamiento.

En el trabajo presentado por Tetteh *et al.* [6], depende de transferencia de aprendizaje entre modelos y conjuntos de datos para realizar tres tareas; la localización de la línea central y puntos de bifurcación en vasos sanguíneos, así como la segmentación de los mismos. Para esto, se realiza un pre-entrenamiento con un conjunto de datos sintético para el modelo 3D convolucional propuesto, realizando un entrenamiento posterior y validación en angiografías de resonancia magnética de ratas y humanos.

Utilizando aumento de datos, Zhang *et al.* [7] aplican una transformación de reflejo a cada elemento del conjunto de datos para generar más ejemplos, esto dada la desventaja de contar con conjuntos de datos reducidos para el entrenamiento de modelos de aprendizaje profundo. De manera adicional, los autores del trabajo agregan ruido Gaussiano a los ejemplos para provocar una invarianza a intensidades absolutas del modelo durante el entrenamiento del mismo. Así mismo, proponen una función de pérdida híbrida para lidiar con el desbalance de clases inherente en los mapas de segmentación en imágenes médicas con una sola clase. Como resultado, logran un índice Dice de 75.6 en segmentación de vasos sanguíneos en imágenes de susceptibilidad ponderada (SWIs).

Considerando la variabilidad que conlleva el flujo sanguíneo al realizar la adquisición de imágenes cerebrales, Kandil *et al.* [8] dividen sus volúmenes de MRA en dos partes, por encima y debajo del Circulo de Willis (CoW), entrenando con ambas partes su modelo 3D CNN, obteniendo un índice Dice de 84.37 como resultado de la segmentación.

Zhao *et al.* [9] desarrollan una arquitectura que extrae estructuras candidatas a ser vaso sanguíneo de imágenes de MRA como parte de un pre-procesamiento. Una red neuronal totalmente conectada actúa como clasificador de las regiones extraídas previamente, dando como resultado la probabilidad de cada una de estas de ser vasos sanguíneos.

En el presente trabajo, distintos modelos convolucionales basados en arquitecturas encoder-decoder, conocidos por sus aplicaciones en diversas tareas de segmentación y generación de imágenes, son evaluados con el fin de realizar una comparación entre ellos en la segmentación de vasos sanguíneos cerebrales en imágenes de microscopía por excitación de fotones. Así mismo, se evalúan distintas técnicas de pre procesamiento de las imágenes cerebrales, además de evaluar métodos de regularización y ajuste de hiper parámetros de los modelos convolucionales seleccionados, con el fin de proponer una mejora en cuanto a los resultados del entrenamiento y la inferencia, en comparación con no utilizar estas técnicas y el cómo afectan a la predicción de la segmentación realizada por dichos modelos, con el fin de proponer un método para la asistencia a especialistas en la detección de vasos sanguíneos cerebrales.

2. Marco Teórico

2.1 Redes Neuronales Convolucionales

Dentro del aprendizaje profundo, las redes neuronales convolucionales (CNN) han jugado un papel importante, ya que son referente en cuanto a las aplicaciones que ha logrado el estudio de la anatomía humana aplicada a proyectos de aprendizaje automático. Las redes neuronales convolucionales son una clase de redes neuronales profundas, diseñadas para trabajar con datos de entrada en forma de cuadrícula en dos dimensiones. Comúnmente son utilizadas con imágenes, dentro del campo de la visión artificial, aunque son útiles también con otro tipo de datos, como texto y en general cualquier tipo de datos espaciales y/o temporales. Una red neuronal convolucional es simplemente una red neuronal común que en al menos una de sus capas realiza la operación matricial



de convolucional. En el caso de la segmentación, mediante esta operación de convolución se realiza el filtrado de la información de interés.

2.1.1 Arquitectura

Comúnmente, una red neuronal convolucional contiene tres tipos de operaciones básicas: convolución, *pooling* y ReLU. Cada una de estas capas acepta datos (imágenes) de tipo volumétrico. Una red neuronal convolucional tiene las siguientes propiedades: una entrada de dimensiones $W_1 \times H_1 \times D_1$, donde W_q es el ancho, H_q es la altura y D_q es la profundidad, todos correspondientes a la imagen de entrada. Los hiper parámetros que definen el comportamiento de la CNN son: número de filtros K , longitud del filtro F , el stride S y el padding P [1] [10]. De esta forma, el resultado es una mapa correspondiente a la inferencia de la CNN, dado por $W_2 = \frac{W_1 - F + 2P}{S + 1}$, $H_2 = \frac{H_1 - F + 2P}{S + 1}$ y $D_2 = K$.

2.1.2 Convolución

La convolución de dos funciones está dada por la integral del producto de ambas funciones, después de desplazar una de ellas cierta distancia t , denotada también con el símbolo "" [10], dado por:

$$s(t) = \int x(a)w(t - a)da = (x * w)(t) \quad (1)$$

Para realizar la convolución en una imagen, se realiza el producto punto entre el kernel y una sección de la imagen con dimensiones iguales y posteriormente se suman los resultados. De manera más formal, se considera el kernel como una función y se le conoce como función kernel K , la cual es de dimensiones $F_q * F_q * D_q$. Como resultado a lo anterior se obtiene un valor V en cierta posición con coordenadas x, y sobre una imagen I [10], como lo muestra la ecuación 2.

$$V(x, y) = (I * K)(x, y) = \sum_m \sum_n I(x + m, y + n)K(m, n) \quad (2)$$

2.1.3 Funciones de Activación

Uno de los elementos más importantes en la selección de hiperparámetros dentro de cualquier tipo de red neuronal es la función de activación que usará cada neurona, tanto para entrenamiento como para predicción. En la implementación de modelos profundos o multicapa, es importante el utilizar funciones de activación no lineales, como Sigmoide o ReLU.

Función de activación ReLU

La función de activación ReLU (Rectified Linear Unit) activa una neurona únicamente si el valor de entrada x se encuentra por encima de cierto umbral. Suponiendo un umbral de cero, cuando el valor de entrada es negativo, la función regresa cero y cuando es mayor a cero, simplemente regresa el mismo valor. Por lo tanto, no afecta a la dimensión de las capas previas, siendo únicamente una operación de activación uno a uno [11]. A manera general, se resume que:

$$ReLU(x) = \begin{cases} x & \text{si } x > 0 \\ 0 & \text{en caso contrario} \end{cases} \quad (3)$$

Función de activación Sigmoide

Previo a la adopción de ReLU como función de activación no lineal, era popular el uso de la función sigmoide. La desventaja de esta es el limitado rango de valores que puede dar como salida,

siendo este de 0 a 1, mientras ReLU puede ir desde 0 hasta infinito. Esto es de relevancia en la implementación del algoritmo de propagación hacia atrás para el entrenamiento de las redes neuronales, ya que el gradiente puede convertirse en cero con relativa facilidad, provocando el efecto de desvanecimiento del gradiente [11]. La función sigmoide está definida por:

$$\text{Sigmoide}(x) = \frac{e^{-x}}{1 + e^{-x}} \quad (4)$$

2.1.4 Entrenamiento de Modelos Convolucionales

Para que un modelo logre el objetivo para el que es propuesto, es necesario alimentar al mismo con datos para el entrenamiento. Los modelos basados en redes neuronales implementan el método del descenso del gradiente, basado en la propagación hacia atrás del error en función de la pérdida con el fin de minimizarla y así ajustarse a la distribución de datos de entrada y su correspondiente etiqueta de salida.

El algoritmo de *backpropagation* se compone de dos fases, la propagación hacia adelante, para el cálculo de la pérdida o error, y hacia atrás para actualizar los parámetros del modelo. La propagación hacia adelante se asemeja al proceso de inferencia, donde un vector de ejemplos se propaga por cada una de las capas realizando las operaciones dadas por los parámetros del modelo en el momento dado, obteniendo la predicción del modelo bajo la evaluación de la función de pérdida a minimizar. En la propagación hacia atrás, se calcula el gradiente de la función de pérdida con respecto a los parámetros en cada capa del modelo, partiendo de calcular la derivada de la función de pérdida con respecto de la predicción [1].

U-Net

Esta arquitectura, figura 1, consiste en una sección de encoding, que captura el contexto y extrae las características de la imagen de entrada. Posteriormente, una sección de decoding reconstruye la imagen a partir de las características extraídas, obteniendo así la imagen segmentada. Esta arquitectura ha demostrado obtener un mejor desempeño al ser combinada con el aumento de datos, por lo que es posible obtener buenos resultados con pocas imágenes para entrenamiento [3].

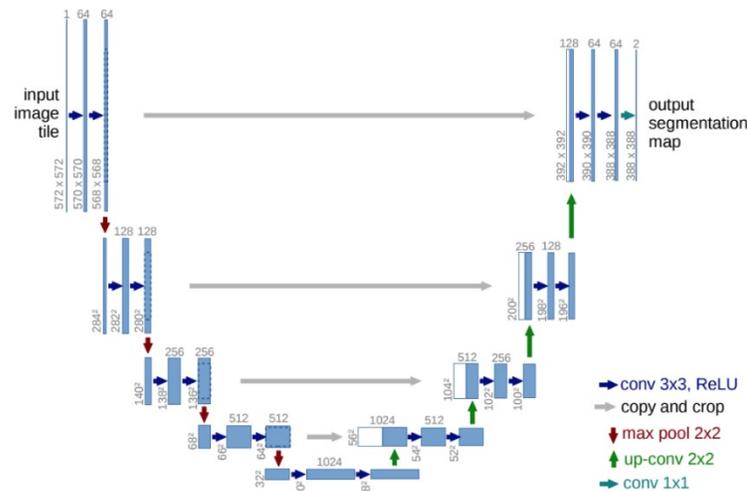


Figura 1. Arquitectura de la U-Net [3].

su respectiva anotación de segmentación objetivo de los vasos sanguíneos. La figura 4 representa una muestra del mismo.

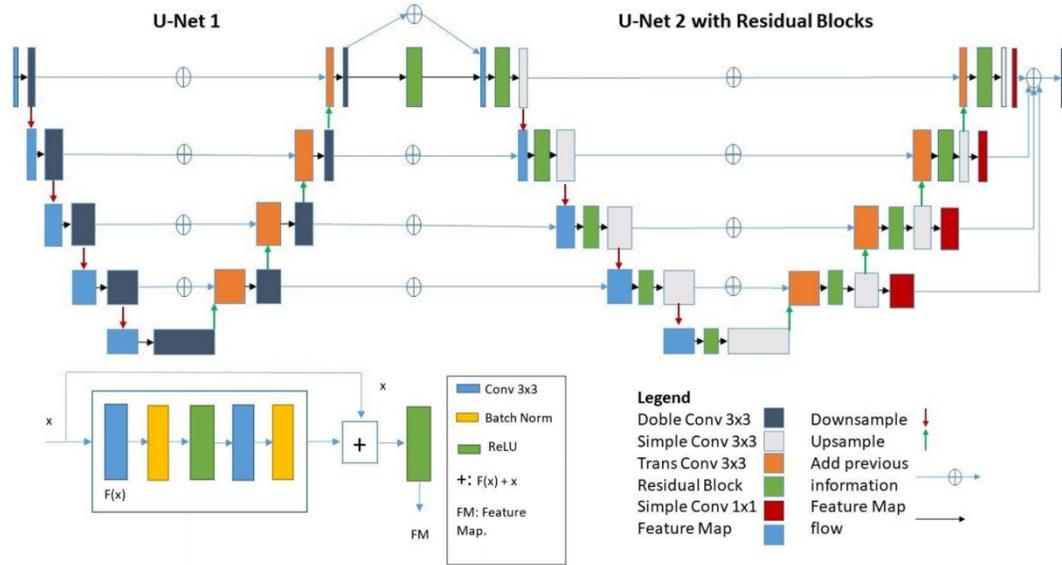


Figura 3. Arquitectura de la U-Net doble [13].

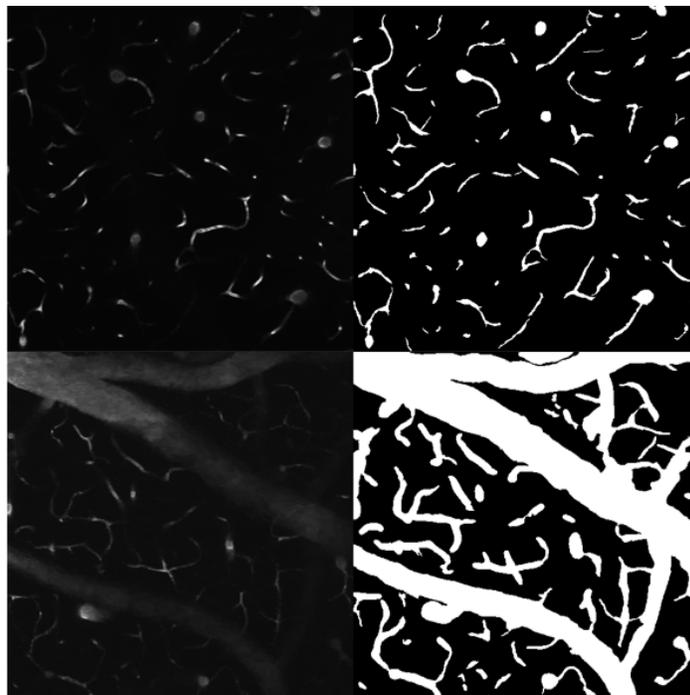


Figura 4. Conjunto de datos de microscopía de excitación de fotones con su respectivo mapa de segmentación de vasos sanguíneos [5].

3.2 Pre procesamiento del conjunto de datos

La ecualización de histograma adaptativo limitada por contraste (CLAHE) [14] ha producido buenos resultados en el procesamiento de imágenes médicas. Este método se formula en la división de imágenes en varias regiones no traslapadas entre sí de el mismo o casi el mismo tamaño. El objetivo de esto es resaltar áreas de interés en la imagen que se encuentran opacadas o en un rango de intensidad de píxeles muy bajo para hacerlas más notorias y que sean más fáciles de identificar por el algoritmo de segmentación. El cálculo del histograma para cada región es relativamente simple. Para cada píxel en escala de grises en la imagen, se contabiliza su intensidad. El conjunto de esto es conocido como el histograma de esa región. Esta función es una aproximación a la función de densidad en escala de grises. La ecualización por histograma es obtenida usando un estimado de la función de distribución acumulativa (CDF).

Si el número de píxeles en escala de grises en cada región son respectivamente M y N , y si $h_{i,j}(n)$ para $n = 0, 1, \dots, N - 1$, es el histograma de la región i, j , entonces una estimación de la CDF en una imagen en escala de grises es:

$$f_{i,j} = \frac{N - 1}{M} \sum_{k=0}^n h_{i,j}(k); \quad n = 1, 2, 3, \dots, N - 1 \quad (5)$$

Una muestra de la ecualización realizada a cada imagen dentro de los volúmenes del conjunto de datos es mostrada en la figura 5, donde se compara la imagen original y la ecualizada, así como los histogramas obtenidos en cada caso.

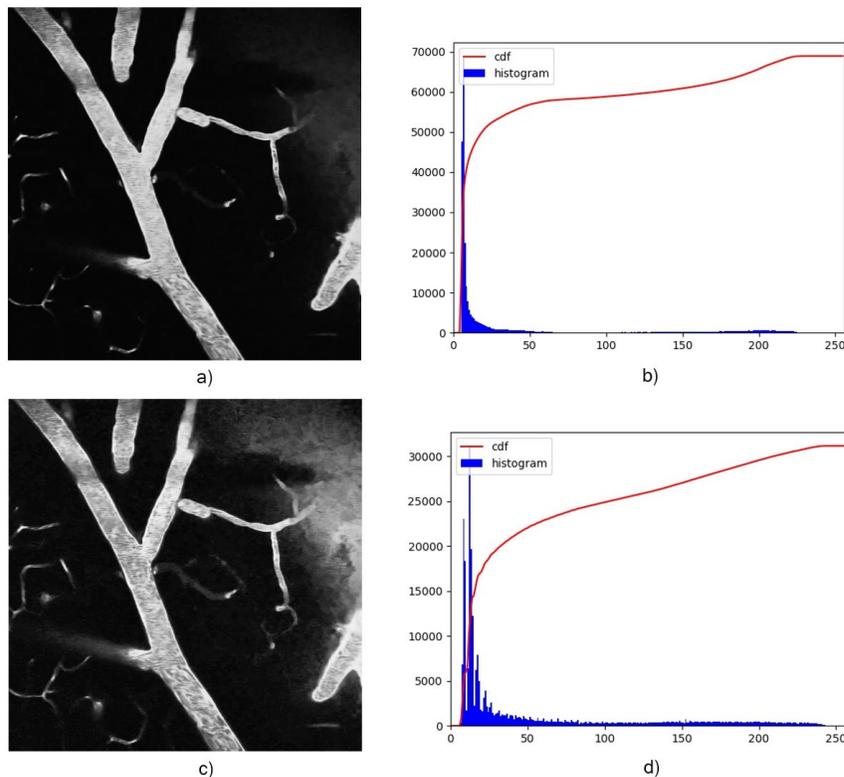


Figura 5. Ecualización por histograma CLAHE. a) Imagen original, b) histograma de la imagen original. c) Imagen ecualizada, d) histograma después de ecualizar la imagen original.

Debido a la limitante en cuanto a cantidad de datos en imágenes médicas, se propone el realizar aumento de datos mediante la librería Keras, la cual permite realizar ciertas transformaciones a las imágenes y así conseguir un mayor volumen de las mismas, sin recurrir a generación de imágenes sintéticas y asegurando que son válidas y pertenecen a la misma distribución de los datos. Las transformaciones aplicadas al conjunto de datos son de acercamiento y reflejo. En la ilustración siguiente se muestra la transformación por acercamiento, técnica mediante la cual se puede obtener cualquier número de imágenes, en este caso se generan 3 a partir de una.

En la figura 6 se aplica la transformación de espejo a las imágenes de la izquierda, obteniendo imágenes totalmente distintas para el entrenamiento de una CNN junto con su etiqueta. De esta manera se puede duplicar el tamaño de cualquier conjunto de datos.

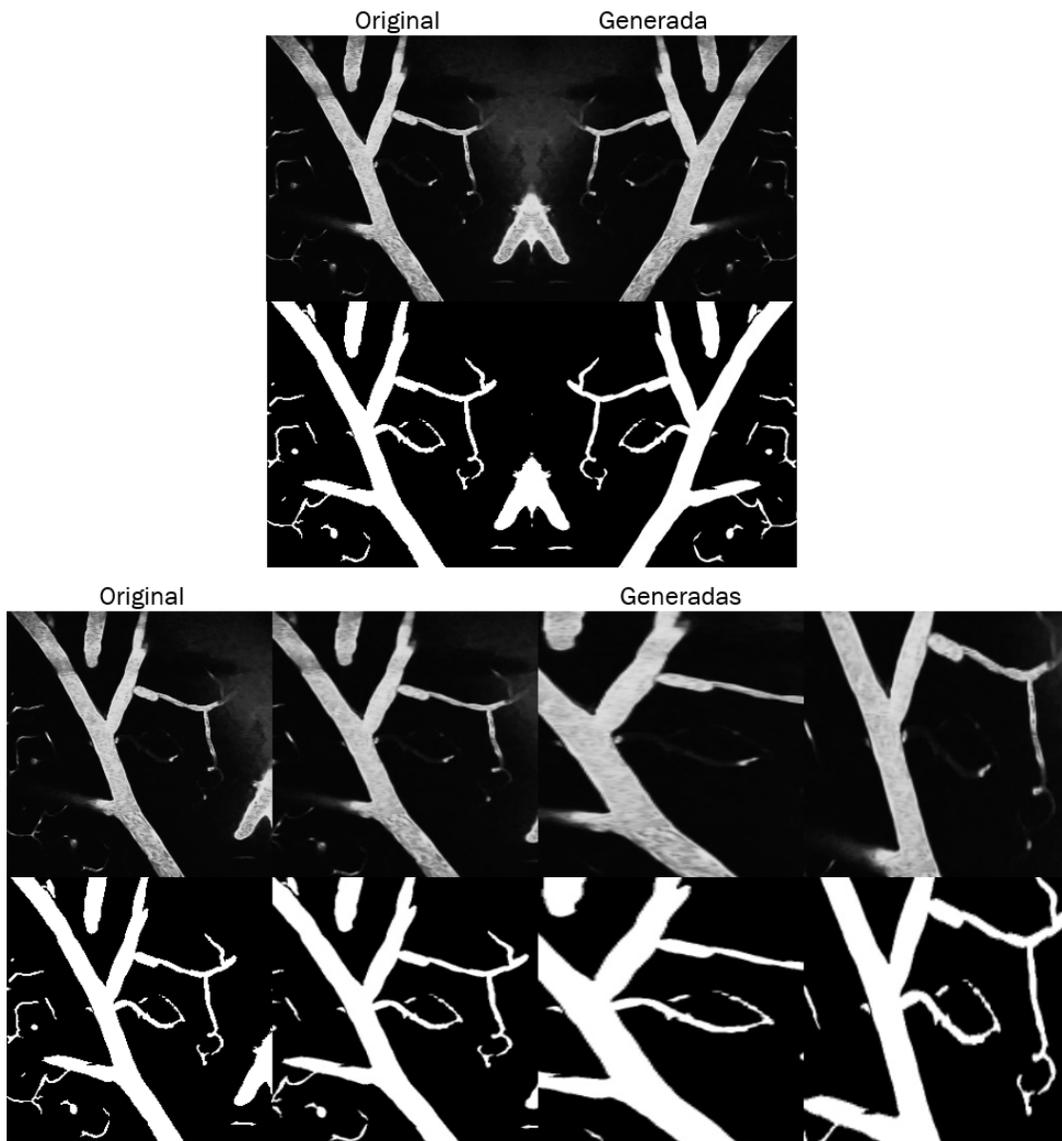


Figura 6. Métodos de aumento de datos. Transformación por reflejo (arriba). Transformación por deformaciones elásticas y acercamiento aleatorio.



3.3 Diseño del Experimento

El entrenamiento de los modelos convolucionales se realizó mediante la herramienta de cómputo en la nube Google Colab, utilizando una GPU Nvidia Tesla T4.

El aumento de datos es realizado como se menciona previamente, obteniendo 4 muestras adicionales a cada una de las imágenes contenidas dentro del conjunto de datos de microscopía, 3 de estas a través de transformaciones elásticas y acercamientos aleatorios, y la última mediante transformación por reflejo.

Para todos los modelos, se realiza una selección de hiperparámetros bastante parecida. Como técnica de regularización para prevenir el sobre ajuste, se implementa dropout [15] con una probabilidad activación de 0.8 en cada capa. El optimizador es Adam [16], con una tasa de aprendizaje de 0.001, $\beta_1 = 0.9$, $\beta_2 = 0.999$ y una tasa de decaimiento exponencial por época sobre la tasa de aprendizaje inicial, mientras que la función de pérdida es logarítmica para clases binarias.

4. Resultados

La implementación de regularización, como aumento de datos, y el pre-procesamiento realizado a las imágenes previo al proceso de entrenamiento de las redes convolucionales conlleva a un mejor resultado en general, donde es posible mejorar el ajuste de cada una sin necesidad de modificar el modelo. Esto es particularmente útil en el caso de contar con un número pequeño de muestras, así como el inherente sobre ajuste de estos modelos, donde es posible lograr un valor elevado de precisión sobre el conjunto de entrenamiento, mientras que en el de validación se obtienen resultados considerablemente más distantes de lo alcanzado en el entrenamiento.

En la figura 7 se muestran las gráficas comparando el proceso de entrenamiento del modelo convolucional de la U-Net doble, donde es evidente el beneficio de la implementación de los métodos descritos. El modelo se ajusta de una mejor manera en ambos casos, entrenamiento y validación, así como logrando una convergencia más suave a lo largo de las épocas por las que fue entrenado. Una muestra de los mapas de segmentación objetivo y las predicciones realizadas a partir de las imágenes de microscopía de entrada, se visualiza en la figura 8, para el caso del mejor de los tres modelos evaluados, la U-Net doble.

Para poder llevar a cabo la comparación entre los tres modelos, el entrenamiento se realiza con las mismas imágenes, así como en la etapa de validación, donde también el tamaño del lote (“batch”) se mantiene constante, con 10 elementos y el mismo pre-procesamiento es realizado al conjunto de datos. En la tabla 1 se muestran los resultados obtenidos.

Tabla 1. Comparación de los modelos convolucionales evaluados.

Modelo	Precisión Entrenamiento	Precisión Validación	Pérdida Entrenamiento	Pérdida Validación	Índice Dice Entrenamiento	Índice Dice Validación
Unet [3]	0.955397	0.930856	0.063940	0.142077	0.850170	0.830051
ResUnet [12]	0.952764	0.937274	0.073630	0.141739	0.857684	0.831315
Unet – ResUnet [13]	0.956317	0.955417	0.063446	0.078425	0.857961	0.841701

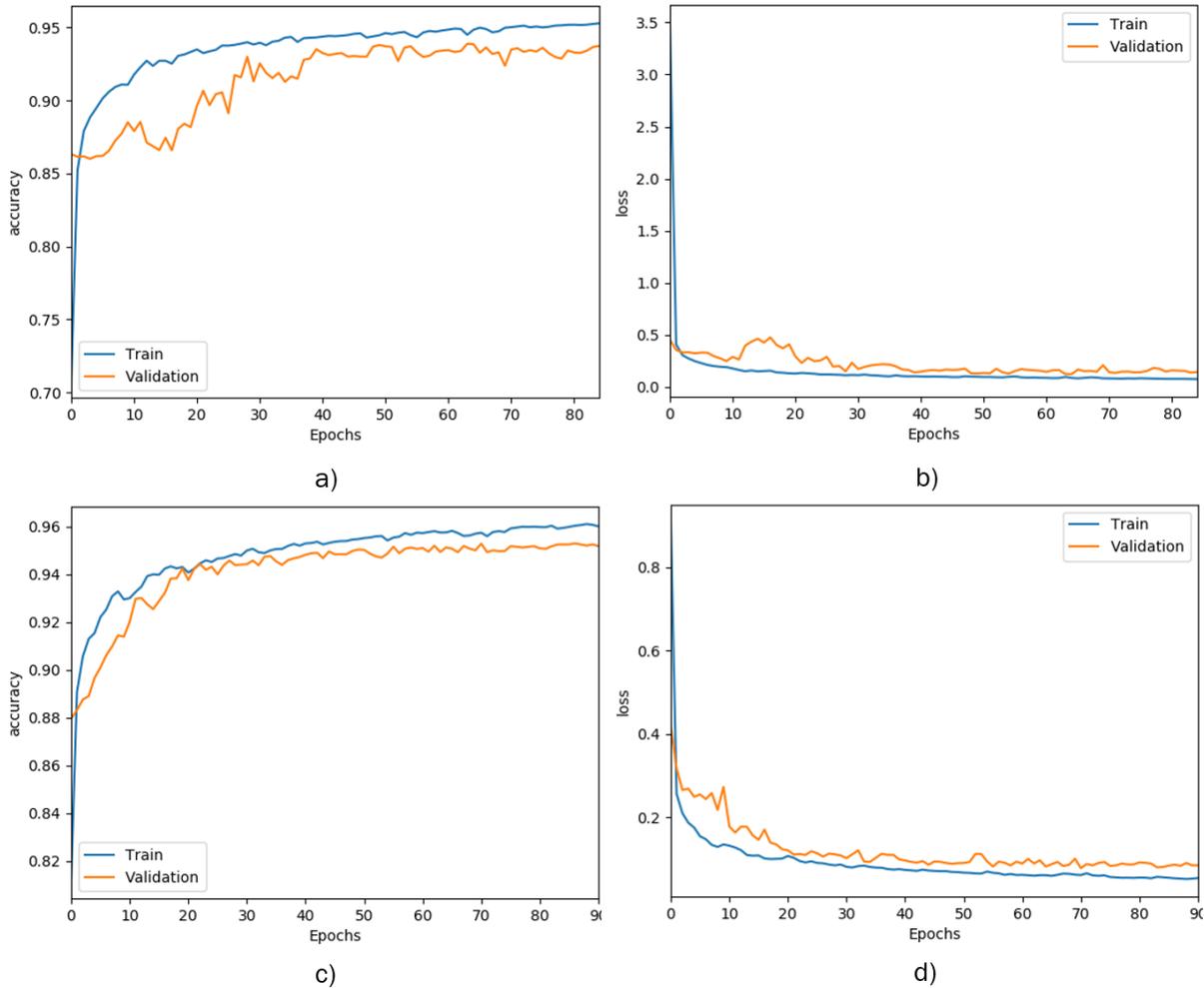


Figura 7. Monitoreo del proceso de entrenamiento de la U-Net doble [13] sobre el conjunto de datos de microscopía con y sin regularización (pre procesamiento de imágenes, aumento de datos, dropout y decaimiento de tasa de aprendizaje). a) Precisión vs épocas en entrenamiento sin regularización. b) Pérdida vs épocas en entrenamiento sin regularización. c) Precisión vs épocas en entrenamiento con regularización. d) Pérdida vs épocas en validación con regularización.

5. Conclusiones

En este trabajo, se muestra la comparativa de diversos modelos convolucionales, enfocados a la segmentación de vasos sanguíneos en imágenes de microscopía por excitación de fotones, donde la implementación de diversos métodos de regularización conlleva a mejores resultados en las predicciones, como se mostró en la figura 7. El aumento de datos es una técnica de gran relevancia al trabajar con imágenes médicas, ya que es difícil encontrar conjuntos de datos públicos con grandes cantidades de muestras, debido al intenso trabajo requerido por parte de expertos en realizar el etiquetado de los mismos. Estas técnicas ayudan a disminuir los efectos negativos al contar con pequeños conjuntos de datos, donde junto con técnicas como el *dropout*, ayudan a evitar el sobre ajuste de dichos modelos. En general se demuestra que las predicciones de estos modelos convolucionales se acercan al trabajo realizado por los especialistas del área médica, requiriendo de menor tiempo para generar los mapas de segmentación y, por lo tanto, acelerando el proceso de diagnóstico para los pacientes que lo requieran. Así mismo, modelos más profundos conllevan a mejores resultados, como se demuestra en la tabla 1, ya que el modelo donde se emplean las dos redes concatenadas obtiene

considerablemente mejores predicciones, incluso con la desventaja de contar con pocas muestras disponibles para el entrenamiento.

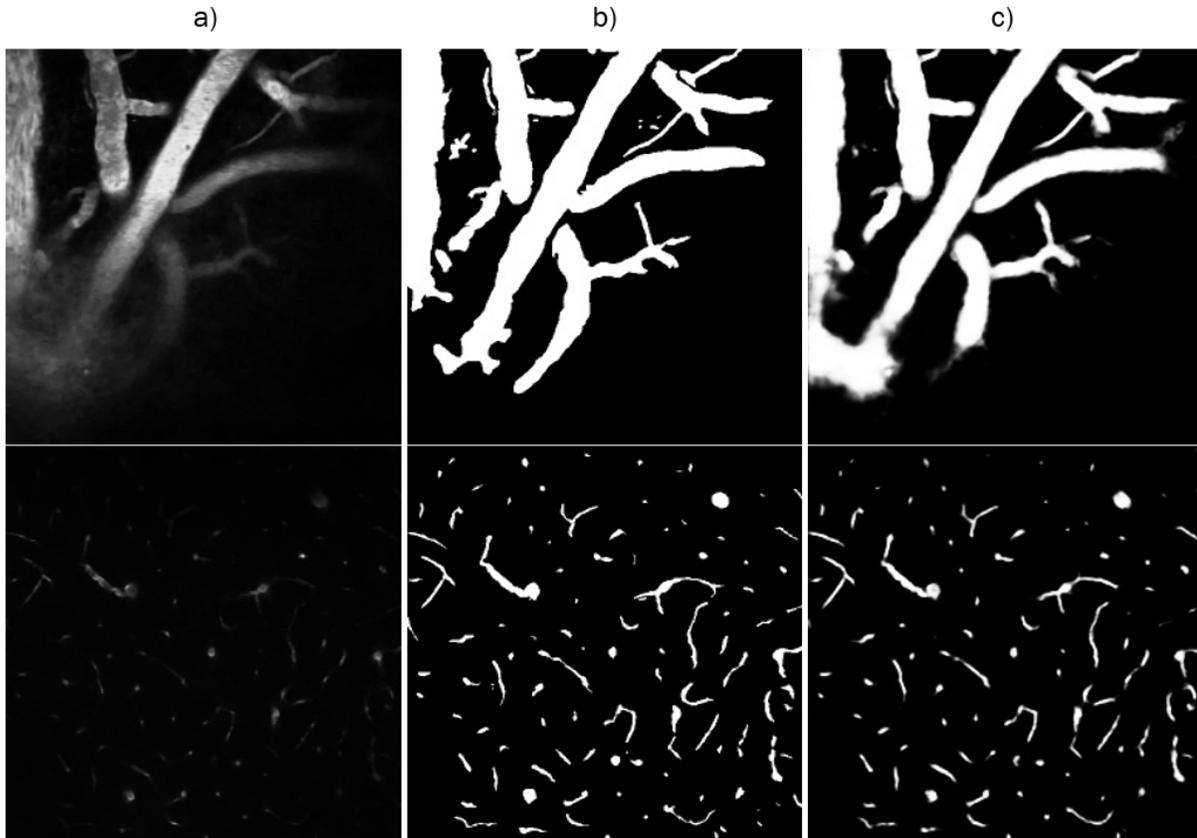


Figura 8. Resultados obtenidos con el modelo de la U-Net doble. a) Imagen de microscopía. b) Segmentación objetivo. c) Predicción.

Referencias

- [1] C. Aggarwal, *Neural Networks and Deep Learning*, Springer, 2018.
- [2] L. Chen, Y. Xie, J. Sun, N. Balu, M. Mossa-Basha, K. Pimentel, T. S. Hatsukami, J. N. Hwang y C. Yuan, «3D intracranial artery segmentation using a convolutional autoencoder,» *Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2017*, pp. 714-717, 2017.
- [3] O. Ronneberger, P. Fischer y T. Brox, «U-net: Convolutional networks for biomedical image segmentation,» *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pp. 234-241, 2015.
- [4] M. Livne, J. Rieger, O. U. Aydin, A. A. Taha, E. M. Akay, T. Kossen, J. Sobesky, J. D. Kelleher, K. Hildebrand, D. Frey y V. I. Madai, «A U-net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease,» *Frontiers in Neuroscience*, vol. 13, pp. 1-13, 2019.
- [5] P. Teikari, M. Santos, C. Poon y K. Hynynen, «Deep Learning Convolutional Networks for Multiphoton Microscopy Vasculature Segmentation,» pp. 1-23, 2016.



- [6] G. Tetteh, V. Efremov, N. D. Forkert, M. Schneider, J. Kirschke, B. Weber, C. Zimmer, M. Piraud y B. H. Menze, «DeepVesselNet: Vessel Segmentation, Centerline Prediction, and Bifurcation Detection in 3-D Angiographic Volumes,» *2018*, pp. 1-13.
- [7] X. Zhang, Y. Zhang y Q. Hu, «Deep learning based vein segmentation from susceptibility-weighted images,» *Computing*, vol. 101, n° 5, pp. 637-652, 2019.
- [8] H. Kandil, A. Soliman, F. Taher, A. Mahmoud, A. Elmaghraby y A. El-Baz, «Using 3-D CNNs and Local Blood Flow Information to Segment Cerebral Vasculature,» *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*, pp. 701-705, 2019.
- [9] F. Zhao, Y. Chen, Y. Hou y X. He, «Segmentation of blood vessels using rule-based and machine-learning-based methods : a review,» *Multimedia Systems*, 2017.
- [10] I. Goodfellow, Y. Bengio y A. Courville, *Deep Learning*, Massachusetts, USA: MIT Press, 2016.
- [11] E. Charniak, *Introduction to Deep Learning*, 2019.
- [12] K. He, X. Zhang, S. Ren y J. Sun, «Deep residual learning for image recognition,» *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016, pp. 770-778, 2016.
- [13] G. Alfonso Francia, C. Pedraza, M. Aceves y S. Tovar-Arriaga, «Chaining a U-Net With a Residual U-Net for Retinal Blood Vessels Segmentation,» *IEEE Access*, vol. 8, pp. 38493-38500, 2020.
- [14] G. Yadav, S. Maheshwari y A. Agarwal, «Contrast limited adaptive histogram equalization based enhancement for real time video system,» *Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2014*, pp. 2392-2397, 2014.
- [15] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever y R. Salakhutdinov, «Dropout: A Simple Way to Prevent Neural Networks from Overfitting,» *Journal of Machine Learning Research*, pp. 1929-1958, 2014.
- [16] D. P. Kingma y J. L. Ba, «Adam: A method for stochastic optimization,» *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1-15, 2015.

Autores

Oliver Jonathan Quintana-Quintana. Ingeniero en robótica por la Universidad Politécnica de Guanajuato. Estudiante de la Maestría en Ciencias en Inteligencia Artificial en la Universidad Autónoma de Querétaro. Su investigación es relacionada con el procesamiento de imágenes, redes neuronales, aprendizaje profundo y machine learning.

Marco Antonio Aceves-Fernández. Es Ingeniero en Telemática por la Universidad de Colima en el año 2000, obtuvo su Maestría y su Doctorado en el área de Sistemas Inteligentes en la University of Liverpool, Reino Unido, éste último en el año 2005. Ha sido reconocido como miembro del Sistema Nacional de Investigadores (SNI) por parte del CONACyT de manera ininterrumpida desde el 2009. Es miembro Senior del IEEE y Presidente honorario de la Asociación Mexicana de Software Embebido. Sus intereses incluyen Sistemas Inteligentes y Embebidos.

Iván Villanueva-Naquid. Ingeniero en Computación por la Universidad Autónoma de San Luis Potosí, obtuvo su grado de Maestría en Ingeniería de Software Distribuido en la Universidad Autónoma de Querétaro y el grado de Doctor en Ciencias de la Computación en la Universidad Autónoma de San Luis Potosí. Actualmente es investigador en el departamento de Innovación Educativa en la Facultad de Medicina en la Universidad Autónoma de San Luis Potosí. Entre sus intereses de investigación se incluyen: análisis y diagnóstico clínico mediante imágenes médicas y empleo de realidad virtual/aumentada para sistemas de entrenamiento médico y asistencia en quirófano.

Saúl Tovar-Arriaga. Obtuvo su grado de Licenciatura en Ingeniería en Electrónica en el Instituto Tecnológico de Querétaro, su Maestría en Ciencias en Mecatrónica en la Universidad de Siegen,



Alemania, y su Doctorado en Ciencias Biomédicas en la Universidad de Erlangen-Nuremberg, Alemania. Actualmente es profesor de tiempo completo e investigador en la Universidad Autónoma de Querétaro. Sus intereses de investigación incluyen robótica médica, diagnóstico automatizado por imagen y visión por computadora.