

# Biometric recognition model using deep convolutional neural networks and computer vision

Christian Ovalle<sup>1</sup> , Daniel Sumire Qqenta<sup>2</sup> , Julián Vilca Sucapuca<sup>3</sup> , Rebeca Sumire Qqenta<sup>4</sup> 

<sup>1</sup>Universidad Privada del Norte, Perú, [denis.ovalle@upn.pe](mailto:denis.ovalle@upn.pe)

<sup>2</sup>Universidad Peruana Unión, Perú, [rebecasumire@upeu.edu.pe](mailto:rebecasumire@upeu.edu.pe)

<sup>3</sup>Universidad Peruana Unión, Perú, [jvilca@upeu.edu.pe](mailto:jvilca@upeu.edu.pe)





<sup>4</sup>Universidad Peruana Unión, Perú, [danielsumire@upeu.edu.pe](mailto:danielsumire@upeu.edu.pe)

*Abstract - Authentication of the person by means of unique features such as finger veins is used in various fields such as security. In this research, a verification method based on convolutional neural networks with the help of computer vision is proposed. Through experimentation, it was possible to create an artificial intelligence model that shows the measurements of loss and precision when verifying the images of a data set. Finally, it is concluded that the loss of biometric recognition, the lower the percentage, the better the model performs. For the modified VGG16 model, 40 epochs were carried out for training, while the Mobilenet was 50 epochs. Additionally, at the end of the execution, the proposed architecture finished in 13 minutes and 45 minutes.*

*Keywords. Biometric recognition, neural networks, computer vision, VGG16 model, verification method*

**Digital Object Identifier:** (only for full papers, inserted by LACCEI).  
**ISSN, ISBN:** (to be inserted by LACCEI).  
**DO NOT REMOVE**

# Modelo de reconocimiento biométrico mediante redes neuronales convolucionales profundas y visión computacional

Christian Ovalle<sup>1</sup> , Daniel Sumire Quenta<sup>2</sup> , Julian Vilca Sucapuca<sup>3</sup> , Rebeca Sumire Quenta<sup>4</sup> 

<sup>1</sup>Universidad Privada del Norte, Perú, [denis.ovalle@upn.pe](mailto:denis.ovalle@upn.pe)

<sup>2</sup>Universidad Peruana Unión, Perú, [danielsumire@upeu.edu.pe](mailto:danielsumire@upeu.edu.pe)

<sup>3</sup>Universidad Peruana Unión, Perú, [jvilca@upeu.edu.pe](mailto:jvilca@upeu.edu.pe)

<sup>4</sup>Universidad Peruana Unión, Perú, [rebecasumire@upeu.edu.pe](mailto:rebecasumire@upeu.edu.pe)

**Resumen**—La autenticación de la persona por medio de características únicas como las venas de los dedos se emplea en varios campos como la seguridad. En esta investigación se propone un método de verificación basado en las redes neuronales convolucionales con la ayuda de la visión computacional. Por medio de la experimentación se logró realizar un modelo de inteligencia artificial que muestre las medidas de pérdida y precisión al verificar las imágenes de un conjunto de datos. Finalmente, se concluye que la pérdida del reconocimiento biométrico mientras menor sea el porcentaje el modelo rinde mejor. Para el modelo VGG16 modified se realizó 40 épocas para el entrenamiento mientras que el Mobilenet fueron 50 épocas. Adicionalmente al término de la ejecución la arquitectura propuesta termino en 13 minutos y 45 minutos.

**Palabras Claves**—Reconocimiento biométrico, redes neuronales, vision computacional, modelo VGG16, método de verificación.

## I. INTRODUCCION

El reconocimiento biométrico vascular (RBV) es una de las variantes biométricas que más favorece al usuario al momento de identificarse. Sin embargo, aún existen factores que retrasan el uso de estos sistemas. Como los son el excesivo costo, el tamaño del equipo (la mayoría de estos sistemas ocupan un espacio considerable), pero los principales factores son la precisión y el tiempo de respuesta, los cuales se ven afectados de una manera directamente proporcional al costo. Puesto que, si se trata de reducir el costo al momento de realizar estos sistemas la precisión disminuye y el tiempo de respuesta aumenta haciendo que el uso de estos sistemas sea restringido [1]. Existen varias investigaciones al respecto en las cuales se proponen distintos métodos para poder reducir el porcentaje de error y de esta manera el RBV tenga una mayor precisión. Actualmente contamos con varios autores que han realizado diversas pruebas con distinta metodologías como la CNN, DCNN, TL, DL, Deep Vein, ROI entre otros [4].

De las investigaciones y pruebas realizadas por los distintos autores se puede destacar que existen metodologías que contienen mejores resultados que otras con respecto a la tasa de reconocimiento [5]. Por ello surge la siguiente pregunta ¿Cómo se diseña un modelo de reconocimiento Biométrico Vascular, mediante redes neuronales convoluciones profundas y visión computacional?

Este trabajo tiene como objetivos: Analizar la data de la muestra de imágenes para reconocer las características de las venas de los dedos; Aplicar la metodología DCNN a la base de datos para aumentar la precisión al emparejar las Imágenes y Evaluar el resultado del proceso de emparejamiento para el reconocimiento biométrico Vascular de los dedos. Por ende, nuestra contribución en este trabajo es diseñar el modelo propuesto, utilizando técnicas de aprendizaje profundo, los cuales incluye las redes neuronales convolucionales. Esto con el fin de mejorar la precisión y acelerar el tiempo de respuesta del RBV, también se evaluará el desempeño del modelo mediante una matriz de confusión para comprobar la efectividad del modelo y confirmar si tiene un mejor rendimiento que otros modelos de reconocimiento biométrico vascular.

## II. ESTADO DEL ARTE

### A. Reconocimiento Biométrico Vascular

El reconocimiento se utiliza para rastrear y recopilar la máxima cantidad posible de información sobre un sistema de destino. Existen diversas técnicas para la recopilación de información en las cuales son el escaneo, escáneres de puertos, rastreadores de paquetes, entre otras. Las modalidades biométricas son más seguras de usar ya que es difícil de adquirir los rasgos de los individuos. Para obtener una imagen biométrica vascular se necesita un sistema de imagen óptica de luz infrarroja que por medio de la mano o dedos del sujeto [3].

Gran parte de los sistemas de reconocimiento biométrico utilizan la bioinformática. Estas herramientas tienen como desventaja la cual es que ocupan un espacio de suma importancia. Una manera de contrarrestar este problema es buscar reducir los costos, aumentar el porcentaje de tasa de reconocimiento y también aumentar la tasa de respuesta [10].

## B. Deep Learning en Informática

Las redes neuronales (RN) son un conjunto de algoritmos de machine learning con el propósito de aprender a resolver tareas complejas. Se basa en aprender ejemplos de patrones que figura de una base de datos de entrenamiento. Las RN al momento de aprendizaje toman su propia estructura y parámetros para entender las señales de entrada [9]. No obstante, hay casos donde los datos de aprendizaje son insuficientemente escasos esto es debido a que a una red neuronal son sensibles a la ubicación espacial, es decir, que objeto de aprendizaje tiene que ser muy idéntico a los datos de la entrada. En una imagen cuando el objeto es cambiado de posición, identifica mal el objeto entre otras categorías generando que la salida del reconocimiento se vea afectada. Por ello, está la red neuronal convolucional, que, a comparación de las tradicionales RN, una arquitectura CNN pueden mejorar el rendimiento de clasificación de imágenes. Debido a esto, el CNN puede extraer características de una imagen de entrada y con ello hacer un reconocimiento y seguimiento de objetos con el fin de tomar una decisión acertada [6].

## C. Transfer Learning

El aprendizaje por transferencia tuvo sus orígenes en la investigación cognitiva, su idea principal es la transferencia de conocimiento por medio de la relación que existe entre las tareas. De esta manera las nuevas tareas tienen un mejor desempeño. El ser humano tiene la capacidad de aprovechar cualquier conocimiento previo para resolver tareas similares. Pan y Yang definieron TL con nociones de dominios y tareas. “El dominio este compuesto por una distribución de probabilidad marginal  $P(X)$  y un espacio de características, en el cual  $x = \{x_1, \dots, x_n\} \in X$ . El dominio específico es  $e = \{X, P(X)\}$ , donde la tarea esta denotada por  $Y, f(\cdot)$ , un espacio de etiqueta, siendo una predicción predictiva objetiva, las tareas se van aprendiendo de la pareja  $X_i, Y_i$  en donde  $X_i \in X, Y_i \in Y$ . El dominio de origen, la tarea de aprendizajes está representada por  $D_S$ , el dominio de destino es  $D_T$  y el aprendizaje de la función predictiva  $F_T(\cdot)$  en  $D_T$ , utiliza el conocimiento proveniente de  $D_S$ ” (Pan SJ, Yang Q. Una encuesta sobre el aprendizaje por transferencia [4].

El objetivo principal del aprendizaje por transferencia (TL) con CNN es mejorar el rendimiento de una nueva tarea. Al tener tareas similares aprendidas, aprovecha el conocimiento de estas.

## III. METODOLOGIA

Este trabajo es de tipo aplicada, esto porque se aplicarán los conocimientos sobre las redes neuronales convolucionales profundas y la visión computacional para realizar el reconocimiento biométrico vascular, de tal manera evaluará el desempeño de la metodología para mejorar la precisión del reconocimiento a comparación de otras arquitecturas y con ello mejorar futuras investigaciones. Nuestra investigación es de tipo aplicada, porque se aplicarán los conocimientos sobre las redes neuronales convolucionales profundas y la visión computacional para el reconocimiento biométrico vascular. [7], se conoce la investigación aplicada como investigación

empírica. Porque su principal característica se basa en la búsqueda del estudio empleando conocimientos previos. Sin embargo, se adquirirán nuevos conocimientos al implementar y sistematizar todo tipo de práctica basada en la investigación. Esta investigación se elaboró en un enfoque cuantitativo. Esto porque según [2] este fenómeno se describe objetivamente, puesto que los datos se pueden extraer con diferentes tipos de mediciones y la recopilación de información de tal manera poder probar y comprobar una hipótesis para luego examinar a detalle la información adquirida. Es necesario especificar los índices donde se cuantificarán las variables y sus respectivas observaciones. Es uno de los enfoques más empleados en la ciencia aplicada, puesto que busca una explicación mediante pruebas de hipótesis o la realización de un modelo. Por ello este enfoque es el que mejor se adapta a los requerimientos de nuestra investigación. De esta manera, nos permitirá determinar el rendimiento de la metodología DCNN en las regiones de interés de las imágenes y con ello mejorar la precisión en la tarea de pareo para que la identificación de las personas sea más rápida y eficiente. Este trabajo de investigación solo posee un variable, el modelo de reconocimiento biométrico vascular mediante redes neuronales convolucionales será la variable responsable de mejorar la precisión de emparejamiento de imágenes para cumplir con el objetivo establecido. Gracias a esta variable se extraerán las características de un conjunto de datos (Imágenes de las venas de los dedos). Además [8]; aseguran que el método de las redes neuronales convolucionales profundas aumenta la robustez de las pruebas y procedimientos del VBR. Se recurrirá a un diseño experimental. Según [5] esta investigación implica manipular una o más variables, las cuales no se han comprobado en circunstancias netamente controladas. Con el objetivo de describir el modo y la causa en que se produjo esa situación. Si la investigación es causada por el investigador, él puede añadir variables específicas y controlar el incremento o reducción de las variables, para observar el efecto que estas producen. También es pre-experimental, porque no se usará un grupo control, tampoco se manipulará más de una variable, solo se medirá una variable (el modelo de reconocimiento biométrico vascular mediante redes neuronales convolucionales) y el grado de control que tendremos en esta investigación será el mínimo. Los pre-experimentos nos ayudan a aproximarnos al fenómeno estudiado. Esto se logra con la administración de un estímulo o tratamiento al grupo experimental con el fin de generar una hipótesis, para realizar la medición de variables y observar los efectos. [2]

### A. Diseño del modelo de reconocimiento biométrico vascular

Para diseñar el modelo nos basamos en una técnica de la inteligencia artificial que es el Deep Learning, ya que usa algoritmos de clasificación de imágenes.

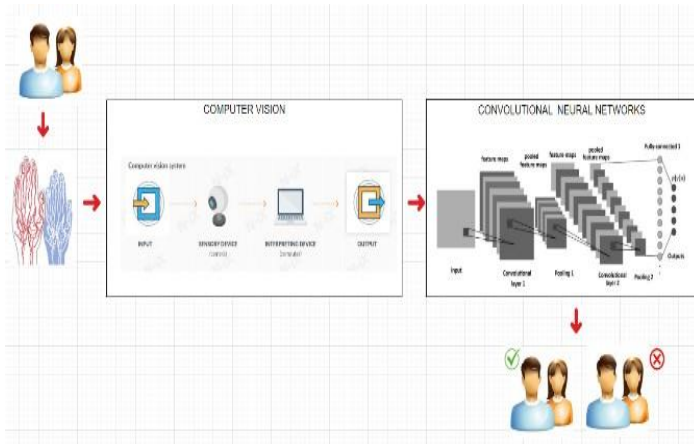


Fig. 1 Modelo Propuesto de reconocimiento biométrico

### B. Desarrollo de la metodología propuesta

Para su desarrollo nos basamos en varios antecedentes que realizan diferentes tipos de modelos de Deep Learning entre ellas nos asentamos en 5 fases importantes para diseñar la metodología. A continuación, se mostrará un diagrama de cómo se va diseñando el modelo.

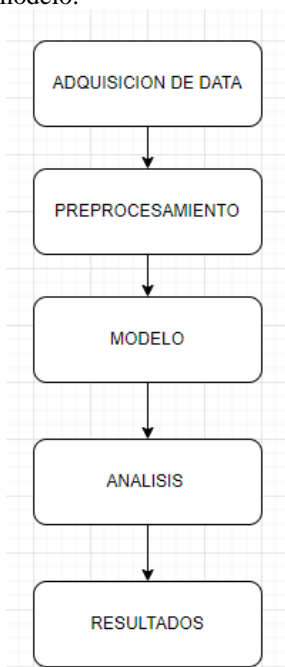


Fig. 2 Fases de la metodología propuesta

### C. Procedimiento

- Adquisición de la data es obtenida de un repositorio llamado “Kagle”. Esta base de datos cuenta con una muestra de 123 individuos en la que está dividida en 83 hombres y 40 mujeres en un promedio de edad entre los 20 y 52 años. Las imágenes tienen una dimensión de 100x300 pixeles y cuentan con alrededor de 5904 imágenes de los dedos.

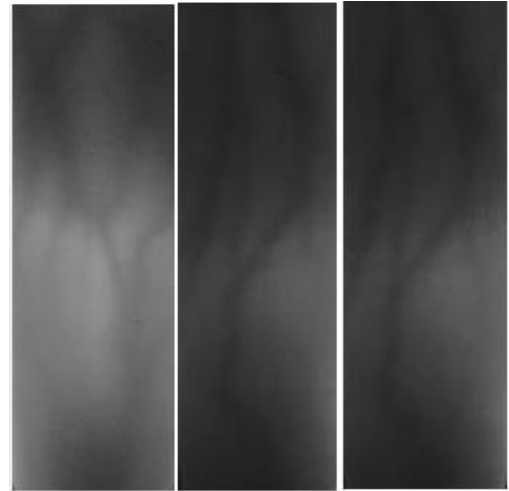


Fig. 3 Imágenes ROI

- Preprocesamiento de imágenes, se aplicará 3 técnicas para mejorar la calidad de la imagen entre ellas está la el mejorar el brillo y para ello se aplica la ecualización adaptiva del histograma para solucionar este problema. La segunda técnica es la eliminación de ruido, ya que tiene el objetivo de mejorar los pixeles de la imagen mediante filtros gaussianos que se encarga de reducir el ruido de la imagen que tiene regiones borrosas o componentes de alta frecuencia. Por último, la técnica de mejora de contraste, la ecuación de corrección gamma mejora ese contraste. Debido al valor gamma de luminosidad que se ingresa las cuales no se pueden detectar si se da una baja intensidad y se calcula de la siguiente forma:

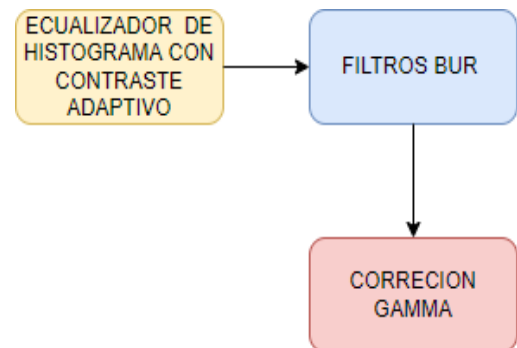


Fig. 4 Técnicas aplicadas

$$\gamma = \frac{dlog(V_{out})}{dlog(V_{in})}$$

Fig. 5 Ecuación de corrección gamma

- Modelo de redes neuronales convolucionales nos basamos en una arquitectura existente de una red neuronal convolucional pre entrenada denominada VGG16. Las razones son las siguientes: Según [1] tiene un resultado entre un 96% y 97% a comparación de otros modelos de CNN. Contiene 13 capas convolucionales y 3 densas y es fácil de comprender e implementar. Agregando lo anterior, se modificará esta arquitectura para lograr un mejor resultado.

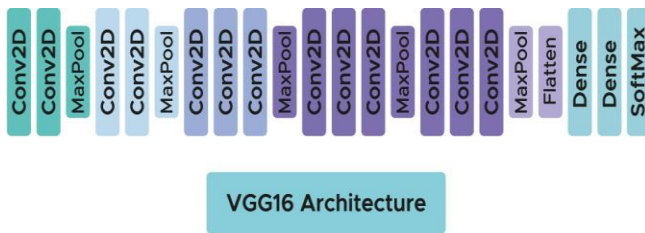


Fig. 6 Mapeo de la arquitectura

- Análisis y resultado nos basamos en una serie de métricas que nos permite evaluar el desempeño del modelo propuesto. Para ello usaremos la matriz de confusión que es un tablero que describe el rendimiento de un modelo de aprendizaje supervisado.

#### D. Experimento

En el entrenamiento se configuró parámetros para que el rendimiento del modelo sea efectivo. Para el DCNN propuesto le aplicamos un data augmentation que simulara más data en el set de entrenamiento y set de validación con el fin de mejorar el performance y eliminar el overfitting.

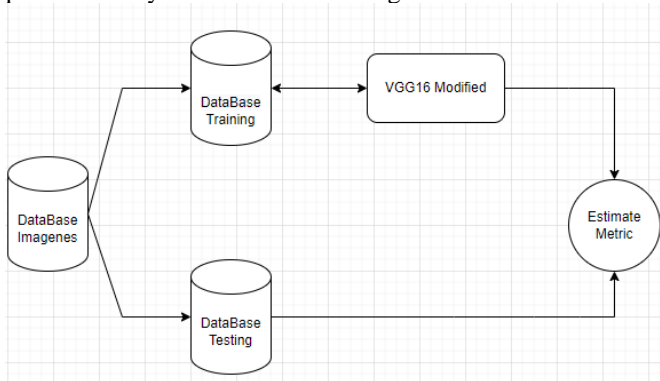


Fig. 7 Modelo de entrenamiento del DCNN

Configurado los parámetros e aumentado los datos, seguimos con el siguiente que es aplicar el VGG16 modified. Para ello, adicionamos una capa densa con n cantidad de clases ya que la capa densa es la más eficaz en cuanto al tratamiento de información. Sin embargo, es la más lenta al momento de procesar datos por la abundante información que gestiona de una capa a otra.

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool1 (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool1 (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool1 (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool1 (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool1 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
output (Dense)	(None, 60)	245820

-----  
 Total params: 134,506,364  
 Trainable params: 245,820  
 Non-trainable params: 134,260,544

Fig. 2 Estructura VGG16 Modified

#### IV. RESULTADOS

En el siguiente párrafo se analizará los resultados obtenidos por las métricas ya establecidas en la metodología:

- Se uso el repositorio de BD-FV-USM para el entrenamiento, validación y prueba.
- Se usaron dos modelos diferentes con los mismos parámetros, pero diferentes épocas.
- Los parámetros elegidos para el proceso de clasificación fueron 40 épocas, donde cada época tomo 30 segundos de la GPU. El optimizador ADAMmejora el rendimiento del modelo y reduce la función de pérdida.

TABLE I  
TABLA DE DATOS

Total	Entrenamiento del modelo		
	Entrenamiento	Validacion	Prueba
1450	1020	300	130

Para entrenar el modelo se separó la data en un 70% de entrenamiento, 20% de validación y 10% de prueba. Es necesario realizar pruebas para verificar si la imagen está prediciendo.

En el resultado de la perdida mientras menor sea el porcentaje el modelo rinde mejor. Para el modelo VGG16 modified se realizó 40 épocas para el entrenamiento mientras que el Mobilenet fueron 50 épocas. Adicionalmente al término de la ejecución la arquitectura propuesta termino en 13 minutos y 45 minutos la otra.

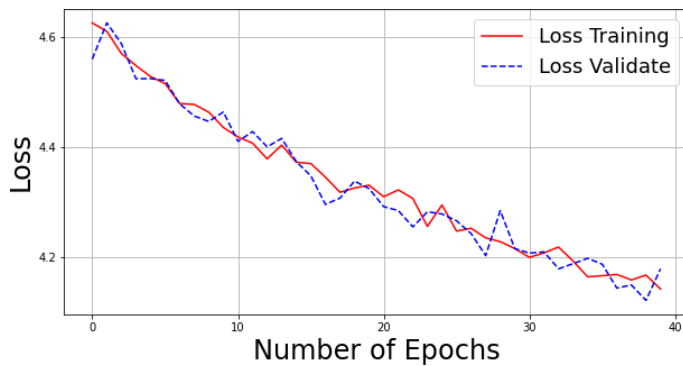


Fig 3. Loss VGG16 Modified

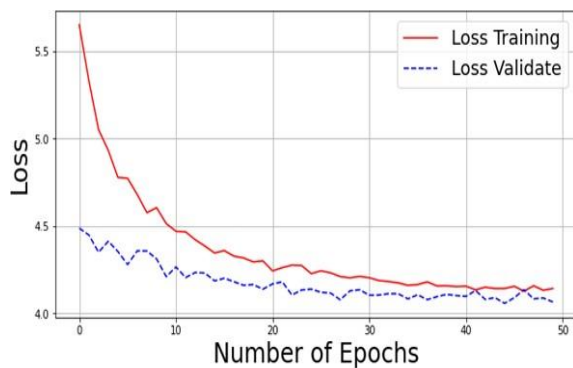


Fig 4. Loss Mobile Net

En resultado de entrenamiento se observa que la precisión aumenta, en la que se puede deducir que la precisión es alta. Por lo cual la precisión del modelo está en un rango favorable que puede ser aplicado sin ningún inconveniente.

TABLE II  
TABLA DE COMPARACIÓN DE LOS MODELOS

MODELO	PRECISION	PERDIDA
VGG16 Modified	0.9705	0.2351
Mobilenet	0.9655	0.2498

Como resultado del entrenamiento, se observó un aumento en la precisión, por lo que se puede concluir que la precisión es alta. Por lo tanto, la precisión del modelo se encuentra en un rango favorable que se puede aplicar sin daño.

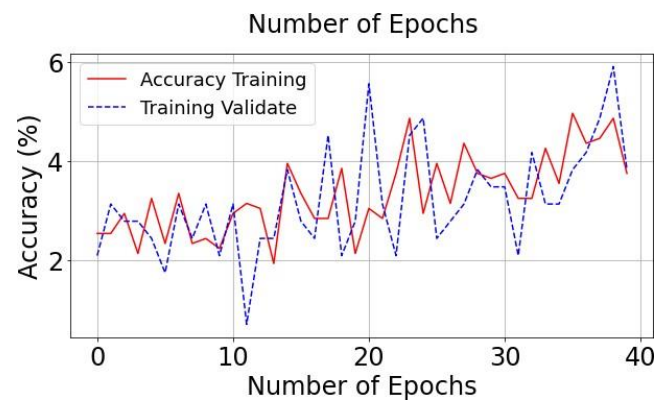


Fig 5. Accuracy VGG16 Modified

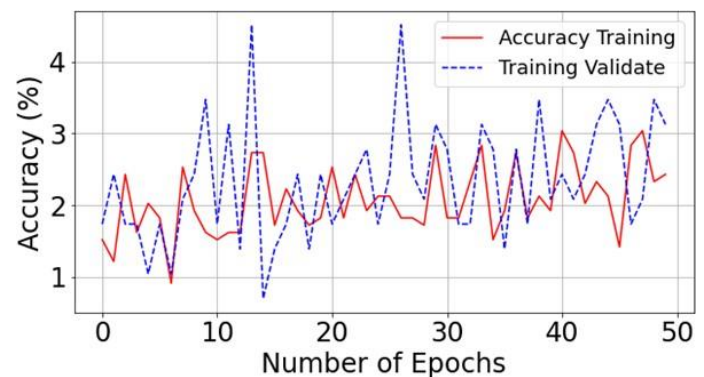


Fig 6. Accuracy Mobile Net

La comparación de los gráficos muestra que el modelo basado en DCNN es mejor que el método implementado por mobilenet, debido a que el entrenamiento es muy efectivo ya que el DCNN implementado supera al otro método en todos los conjuntos de pruebas.

La precisión del conjunto de pruebas aumenta el tamaño del conjunto de entrenamiento del modelo. Dependiendo de cuantas clases se configure en el parámetro la precisión será mucho más estable.

## V. CONCLUSIÓN

En este artículo, presentamos un nuevo modelo para el reconocimiento biométrico vascular de los dedos mediante un DCNN y visión computacional. Realizamos experimentos con dos diferentes arquitecturas de redes neuronales convolucionales y usando una matriz de confusión para evaluar las métricas planteadas en este documento. Como se muestra en los resultados, el método propuesto logra un mejor rendimiento en la precisión y pérdida, lo que demuestra que las redes neuronales convolucionales profundas tengan un buen impacto en coincidir los patrones de las venas de los dedos.

Con los resultados de la figura X se comprobó que no es necesario aumentar más épocas para un mejor rendimiento, generando que el tiempo de respuesta sea muy efectiva al momento de querer implementar este modelo en un sistema de verificación.

## REFERENCIA

- [1] Ahmad Radzi, S., M. Khalil-Hani, and R. Bakhteri. "Finger-Vein Biometric Identification Using Convolutional Neural Network." *Turkish Journal of Electrical Engineering and Computer Sciences* 2016 24(3):1863– 78. doi: 10.3906/elk-1311-43.
- [2] Campbell, D.T. Reforms as experiments. *American Psychologist*, 24, 409-429, 1969
- [3] Debnath, T., Reza, M. M., Rahman, A., Beheshti, A., Band, S. S., & Alinejad-Rokny, H. Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity. *Scientific Reports*, 2022, 12(1) doi:10.1038/s41598-022-11173-0
- [4] Da Costa, J. F. P., & Cabral, M. Statistical methods with applications in data mining: A review of the most recent works. *Mathematics*, 2022, 10(6) doi:10.3390/math10060993
- [5] Franchini, G., Ruggiero, V., Porta, F., & Zanni, L. Neural architecture search via standard machine learning methodologies. *Mathematics in Engineering*, 2023, 5(1) doi:10.3934/MINE.2023012
- [6] Hsiao, C. -, Lin, C. -, Wang, P. -, & Wu, Y. -. Application of convolutional neural network for fingerprint-based prediction of gender, finger position, and height. *Entropy*, 2022, 24(4) doi:10.3390/e24040475
- [7] Liu, W., Li, W., Sun, L., Zhang, L., & Chen, P. Finger vein recognition based on deep learning. Paper presented at the *Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications, ICIEA 2018*, 205-210. doi:10.1109/ICIEA.2017.8282842.
- [8] Sarma, D., & Bhuyan, M. K. Methods, databases and recent advancement of vision-based hand gesture recognition for HCI systems: A review. *SN Computer Science*, 2021, 2(6) doi:10.1007/s42979-021-00827
- [9] Swiderski, B., Osowski, S., Gwardys, G., Kurek, J., Slowinska, M., & Lugowska, I. Random CNN structure: Tool to increase generalization ability in deep learning. *Eurasip Journal on Image and Video Processing*, 2022, 1 doi:10.1186/s13640-022-0
- [10] S. -Y. Jhong et al., "An Automated Biometric Identification System Using CNN-Based Palm Vein Recognition," 2020 International Conference on Advanced Robotics and Intelligent Systems (ARIS), Taipei, Taiwan, 2020, pp. 1-6, doi: 10.1