

## RED NEURONAL CONVOLUCIONAL DE BAJO COSTO PARA EL RECONOCIMIENTO DE IRIS

González Estrada Adrián Alberto, Ramírez Quintana Juan Alberto y Vega Pineda Javier

Tecnológico Nacional de México  
Instituto Tecnológico de Chihuahua  
División de Estudios de Posgrado  
Av. Tecnológico 2909, Chihuahua, Chih., CP 31310, México  
[adriahn@gmail.com](mailto:adriahn@gmail.com), [jaramirez@itchihuahua.edu.mx](mailto:jaramirez@itchihuahua.edu.mx), [jvega@itchihuahua.edu.mx](mailto:jvega@itchihuahua.edu.mx)

### RESUMEN.

Se presenta la arquitectura de una red neuronal convolucional (CNN) de bajo costo basada en el modelo AlexNet y se comparó su rendimiento con el de las arquitecturas CNN AlexNet (sin modificar) y la CNN VGG-16. El reconocimiento de iris se refiere al proceso automático de reconocer individuos en base a sus patrones visuales específicos de su iris. La naturaleza de las características aleatorias y distintivas del iris lo convierte en una señal adecuada para el reconocimiento biométrico. Avances recientes en el área de aprendizaje profundo ("deep learning") indican que las CNN extraen características que representan y generan información importante existente en las imágenes. Por lo cual, se trabajó en el desarrollo de un modelo basado en la CNN AlexNet y el modelo propuesto se probó con 15 clases seleccionadas de la base de datos UTIRIS V1 (imágenes de ojos humanos) para evaluar su desempeño y complejidad computacional. Para este propósito, se introducen imágenes segmentadas y normalizadas sin ruido, logrando buenos resultados comparativos con respecto a las antes mencionadas CNN.

**Palabras Clave:** Reconocimiento de iris, red neuronal convolucional (CNN), aprendizaje profundo.

### ABSTRACT.

An architecture of a low-cost convolutional neural network (CNN) model based on the AlexNet model is presented and its performance was compared with the CNN AlexNet (unmodified) and the CNN VGG-16 architectures. Iris recognition refers to the automatic process of recognizing individuals based on their iris specific visual patterns. The nature of the random and distinctive features of the iris makes it an adequate signal for biometric recognition. Recent advances in the deep learning area indicate that CNNs extract characteristics that can represent and generate important information existing in images. Therefore, we worked on the development of a model based on the CNN AlexNet and the proposed model was tested with 15 selected classes of the UTIRIS V1 database (images of human eyes) to evaluate its performance and computational complexity. For that purpose, we introduce segmented and normalized images without noise, achieving good comparative results with respect to the CNN.

**Keywords:** Iris recognition, convolutional neural networks (CNN), deep learning.

### 1. INTRODUCCIÓN

La identificación automática de una persona se lleva a cabo en función de las características únicas que posee en su cuerpo y se conoce en algunas áreas del conocimiento como identificación por medios biométricos. Se utiliza el reconocimiento biométrico debido a que una contraseña o una tarjeta de identificación puede ser fácilmente extraviada o falsificada, relacionándose en ámbitos como la banca, seguridad o compras por Internet.

El reconocimiento del iris del ojo humano se considera el sistema de identificación biométrico más confiable y preciso disponible en la actualidad. El iris se caracteriza por tener una textura única, estabilidad en el tiempo, accesibilidad y existe la posibilidad de su escaneo no invasivo para la persona. Además, con ventaja de que la variabilidad de patrones del iris entre personas es enorme [1].

El primer paso en el reconocimiento del iris es localizar mediante un proceso de detección de bordes los dos límites que definen el iris. La contribución más significativa en este sentido fue realizada por Daugman quien aplicó un operador integro diferencial que localiza los límites del iris [2].

El segundo paso es la normalización de la imagen del iris ya segmentada. Consiste en la transformación del espacio de coordenadas cartesianas  $(x, y)$  a un sistema de coordenadas polares  $(r, \theta)$  para realizar una comparación de imágenes de manera más directa para el sistema [3]. Utilizando la imagen normalizada en coordenadas cartesianas, se aplican técnicas tradicionales de procesamiento de imágenes para lograr los resultados de la aplicación final.

La extracción de características efectivas de las imágenes es uno de los pasos más importantes para el reconocimiento robusto de objetos y tareas de visión en la clasificación de imágenes. Hoy en día, se presta mucha atención a los algoritmos de aprendizaje de características y redes neuronales convolucionales (CNN). En estas estructuras algorítmicas, la imagen se alimenta directamente como entrada a la CNN, luego el algoritmo extrae las características óptimas para obtener la salida deseada para su clasificación o en este caso para la identificación de la persona.

Se han realizado intentos para apropiarse de los principios del aprendizaje profundo a la tarea de reconocimiento del iris. La limitación de los métodos de aprendizaje profundo al

problema del reconocimiento del iris se debe al hecho de que requiere una gran cantidad de datos de entrenamiento, aún no disponibles para la mayoría de los investigadores del iris. Sin embargo, el aprendizaje profundo también es un proceso costoso desde el punto de vista computacional por lo que es indispensable contar con procesamiento paralelo como pueden ser varias unidades de procesamiento gráfico (GPU, *Graphics Processing Unit*). Esto es un reto para la implementación física de las aplicaciones que utilizan el aprendizaje profundo.

En la literatura relacionada con el reconocimiento de iris no se ha encontrado información sobre porque el aprendizaje profundo debería funcionar para este problema y no se ha realizado un análisis sistemático para determinar la mejor manera de aprovechar los enfoques profundos modernos para diseñar una arquitectura óptima de redes profundas para lograr una alta precisión y baja complejidad computacional.

En trabajos reportados, se presentan CNNs que apilan múltiples capas en su diseño para el reconocimiento del iris. Sin embargo, no necesariamente es lo más óptimo debido a las múltiples opciones para el diseño (número de capas, configuración de cada capa) y puede ser ineficiente (debido a capas redundantes).

Se han utilizado modelos ya diseñados para el reconocimiento de imágenes de redes neuronales preentrenadas bajo la arquitectura AlexNet [4] la cual se tomó como referencia para el presente estudio. También, se han diseñado arquitecturas específicas para el reconocimiento de iris como las arquitecturas DeepIrisNet [5]. La red DeepIrisNet-A se basa en el estándar de capas convolucionales y la red DeepIrisNet-B utiliza apilamiento de capas de inicio, los entrenamientos y características utilizados fueron por medio de análisis experimental. En relación con los datos de entrenamiento, las arquitecturas poco profundas o muy profundas pueden resultar en ajuste insuficiente o excesivo.

La arquitectura DeepIrisNet-A se compone de ocho capas convolucionales (conv1 - conv8), cada una seguida por su capa de normalización por lotes y sus capas de agrupamiento máximo y finalmente por tres capas totalmente conectadas. En la arquitectura DeepIrisNet-B se apilan primero las capas convolucionales (conv1 - conv5), y luego dos capas iniciales (Inception6 e Inception7) y finalmente tres capas totalmente conectadas. La entrada a la arquitectura DeepIris-Net son imágenes del iris de niveles de gris de  $128 \times 128$  píxeles sin procesamiento previo y capturadas por el sensor LG2200, Fig. 1. Las dos redes se entrenan utilizando solo una escala y no se realiza ningún aumento de datos. Los resultados se dividieron en tres pruebas con base en modificaciones realizadas a las imágenes de entrada (segmentación, rotación y tamaño), obteniendo como el mejor desempeño un 97.31% de clasificación correcta.

En este documento se propone un esquema para el reconocimiento de iris en base a la arquitectura CNN AlexNet consistente de 8 capas: 5 capas convolucionales y 3 capas totalmente conectadas. Se modificó la arquitectura AlexNet para adecuarse a las necesidades del reconocimiento de iris en donde se trabaja para la clasificación de  $N$  clases, es decir,

la identificación de  $N$  personas diferentes utilizando aprendizaje profundo con las CNN.

## 2. FUNDAMENTOS DE RECONOCIMIENTO DE IRIS

Se presentan las componentes del ojo humano y el preprocesamiento previo a sus imágenes para la obtención de las imágenes de entrada a las CNN. Las dos etapas iniciales que engloban el proceso son: segmentación y normalización. La mayoría de las investigaciones utilizan la imagen de iris normalizada para la extracción de características [4-6].

### 2.1. Estructura

El iris se considera un órgano interno del cuerpo humano y puede ser observado sin necesidad de realizar una intervención al cuerpo. El iris corresponde a la sección pigmentada o coloreada dentro del ojo entre la pupila y la córnea. Su principal función es controlar la abertura de la pupila para regular la entrada de la luz hacia la retina [7]. En la Figura 1 se indican las principales componentes del ojo siendo la región del iris y la región de la pupila los puntos de interés.

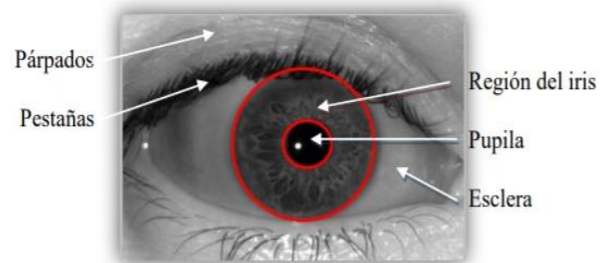


Figura 1. Anatomía del ojo humano.

### 2.2. Segmentación

La segmentación tiene como objetivo encontrar en la imagen del ojo los píxeles que representan al iris. Se logra determinar el iris en la imagen al ubicar su límite interno (pupilar) y su límite exterior (límbico). Existen posibles problemas en la segmentación del iris como reflexión especular, oclusión por párpados y/o pestañas. Se usa una máscara para denotar estos píxeles de no iris en la imagen [8].

### 2.3. Normalización

En el modelo homogéneo “rubber sheet” introducido por Daugman [2], se realiza una transformación de la región del iris de cada punto en un sistema cartesiano  $(x, y)$  a un sistema normalizado de coordenadas polares  $(r, \theta)$  de acuerdo al número de muestras de radio establecidas en un rango de  $[0, 1]$  y en una dirección theta de  $[0, 2\pi]$ , Figura 2 [3]. El reacomodo de píxeles en la normalización se modela como:

$$I(x(r, \theta), y(r, \theta)) \rightarrow I(r, \theta) \quad (1)$$

donde  $I(x, y)$  es la región del iris en coordenadas cartesianas  $(x, y)$  y  $(r, \theta)$  son las coordenadas de los píxeles en el sistema polar ya normalizadas.

Para encontrar los puntos en coordenadas polares se aplica:

$$x(r, \theta) = (1 - r)x_p(\theta) + rx_i(\theta) \quad (2)$$

$$y(r, \theta) = (1 - r)y_p(\theta) + ry_i(\theta) \quad (3)$$

donde  $(x_p, y_p)$  son las coordenadas de los límites de la pupila y  $(x_i, y_i)$  son las coordenadas de los límites del iris, ambas coordenadas en dirección de theta,  $\theta$ .

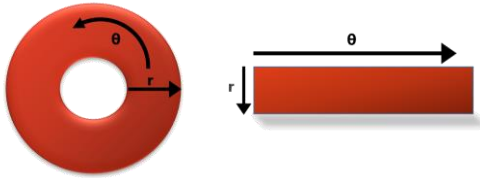


Figura 2. Normalización del iris.

### 3. MATERIALES Y HERRAMIENTAS

El desarrollo de la modificación a la arquitectura AlexNet CNN y la evaluación del algoritmo de reconocimiento de iris se realizó con la herramienta Matlab 2018a y con los paquetes “Image Processing Toolbox” y “Neural Network Toolbox”.

La base de datos de imágenes de iris (ojos) utilizada fue UTIRIS creada por la Univ. de Teherán [9]. Es el primer banco de datos biométricos registrados en dos sesiones distintas de imágenes de longitud de onda visible (VW) e infrarrojo cercano (NIR). UTIRIS está construida con 1540 imágenes de 79 individuos de ambos ojos (derecho e izquierdo). Los individuos se enumeran igual en sesiones NIR y VW, un ejemplo se muestra en la Fig. 3. En la Tabla 1 se muestran las características de la base de datos.

Tabla 1. Configuración para la adquisición de la base de datos.

	VW	NIR
Dimensión de la imagen (número de píxeles)	2048 x 1360	1000 x 776
Formato de compresión	JPEG	BMP
Dispositivo de cámara	Canon EOS 10D	ISG Lightwise LW
Espacio de color	RGB	GRIS
Longitud focal	100	
Ojos rojos	si	
Tiempo de exposición	1/60	

Se seleccionaron 15 clases diferentes (número de personas a identificar) dentro de los 79 individuos, cada conjunto de datos por individuo está compuesto por entre 8 y 10 imágenes incluyendo ambos ojos. Se trabajó con imágenes VW RGB en su canal rojo [10], Figura 4 y el color rojo tiene una longitud de onda visible muy cercana al infrarrojo cercano alrededor de 640 nm. Incluso para imágenes de iris oscuro utilizando el canal rojo se aprecia que se tiene un mayor contraste entre el iris y la pupila. Además, se resaltan los detalles en el iris, permitiendo a su vez una mejor detección de la pupila requerida para el proceso de normalización.

El conjunto de imágenes de las 15 clases se dividió en tres conjuntos, para entrenamiento, validación y pruebas. Primero, del total de las imágenes se seleccionaron cinco imágenes de cada clase de manera aleatoria por cada entrenamiento de la red para formar el conjunto de entrenamiento. El resto de las imágenes se dividió entre los conjuntos de validación y prueba.

Para la selección del conjunto de validación se seleccionaron aleatoriamente dos imágenes del conjunto restante. El resto de las imágenes forman el conjunto de pruebas para probar las CNN y obtener los desempeños. En la Tabla 2 se presentan las cantidades de las imágenes totales para cada conjunto.



Figura 3. Base de datos UTIRIS, imagen VW RGB.



Figura 4. Canal rojo de la imagen RGB de entrada.

Tabla 2. Conjunto de imágenes para el reconocimiento de iris.

Base de Datos	Número de clases	Cantidad de imágenes	Imágenes de entrenamiento	Imágenes de validación	Imágenes de pruebas
UTIRIS	15	138	75	30	33

### 4. RED NEURONAL CONVOLUCIONAL (CNN)

Se presentan los componentes principales de una arquitectura CNN y su funcionamiento. Una CNN es una red neuronal multicapa *feed-forward*, que difiere de las redes neuronales tradicionales totalmente conectadas al combinar varias capas conectadas localmente destinadas al reconocimiento automático de características, seguidas de una serie de capas conectadas totalmente dirigidas a la clasificación [11]. La arquitectura CNN de la Figura 5 comprende varias capas distintas que incluyen conjuntos de capas convolucionales conectadas localmente con un número específico de diferentes filtros en cada capa, capas de submuestreo llamadas capas de agrupación y una o más capas totalmente conectadas.

Las CNN se aplican en problemas de reconocimiento de patrones; desafortunadamente, pueden ser difíciles de entrenar debido a la alta potencia computacional que requieren para su entrenamiento como es el caso del modelo AlexNet [12].

Entre los parámetros que contribuyen significativamente al rendimiento de las CNN están: el tamaño de los filtros para las operaciones de convolución, el número de mapas de características en cada capa de convolución, el tipo de operación de agrupamiento en la capa de submuestreo, la compilación de mapas de convolución, la profundidad de la red y las funciones de activación [13].

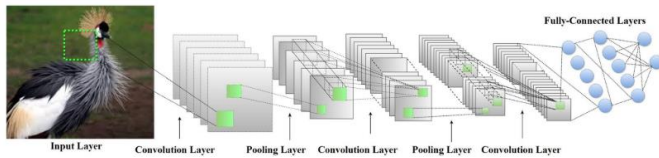


Figura 5. Arquitectura tradicional de una CNN [4].

#### 4.1. Capa convolucional

Esta capa tiene un conjunto de filtros que se componen de pesos generados aleatoriamente y se modifican mediante el algoritmo de *backpropagation*. Cada kernel resultante realiza el proceso de convolución con la imagen de entrada para obtener los mapas de características y el número de mapas de características es igual al número de kernels aplicados en esa capa. La Figura 5 muestra una primera capa de convolución que consta de seis mapas de características apilados y generados a partir de seis kernels que aplican la convolución en toda la imagen de entrada [11].

#### 4.2. Normalización por lotes (*Batch normalization*)

Esta técnica mejora la velocidad y estabilidad de la CNN durante el entrenamiento. También tiene un efecto benéfico en el flujo de gradiente a través de la red esto al reducir la dependencia de los gradientes en la escala de los parámetros o de sus valores iniciales. Lo anterior nos permite utilizar tasas de aprendizaje mucho más altas sin el riesgo de divergencia [14].

#### 4.3. Capa de agrupamiento

Su función es implementar una operación de reducción de muestreo para disminuir el tamaño espacial de las capas convolucionales y producir una forma limitada de invariancia traslacional [4]. Una forma comúnmente utilizada es la agrupación máxima con regiones de tamaño  $2 \times 2$  y un paso de 2. La dimensión de profundidad del volumen de entrada se mantiene sin cambios [11].

#### 4.4. Capa totalmente conectada

La salida de la última capa convolucional o de agrupación se alimenta a una o más capas completamente conectadas como en una red neuronal tradicional. En esas capas, las salidas de todas las neuronas en la última capa de la etapa de extracción de características están completamente conectadas a cada neurona de la capa siguiente y así sucesivamente para las capas posteriores [11]. Esto es, se utilizaron las características extraídas en las capas anteriores para realizar la tarea de clasificación.

#### 4.5. Capa de clasificación Softmax

La red tiene variaciones en la precisión con las que las neuronas aprenden debido a que se inicializan los pesos de forma aleatoria. Consecuentemente, se utiliza una función de pérdida de entropía cruzada denominada Softmax, la cual se emplea para calcular el error normalizado entre la etiqueta del *ground truth* (salida deseada) y la etiqueta de predicción obtenida [15].

La capa Softmax se encuentra entre la capa de salida de la red neuronal y la capa de predicción de los resultados de la red y se define como [16]:

$$\text{Softmax}(y_i) = y'_i = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}} \quad (4)$$

donde  $y_i$  es la  $i$ -ésima salida de la red neuronal y la suma incluye desde la primera clase hasta la  $n$ , esto es, la cantidad de clases en la red neuronal. La salida de esta capa es la probabilidad de que la imagen pertenezca a cada una de las clases para la identificación de la persona.

#### 4.6. El proceso de aprendizaje

Una CNN resuelve el problema de interés entrenando varios coeficientes como son los pesos de los kernels de las capas convolucionales y los pesos de las capas totalmente conectadas. Los parámetros mencionados se inicializan aleatoriamente o con otros métodos más estables. Luego, un proceso iterativo de aprendizaje alterna entre los pasos de *feed-forward* y *backpropagation*. Estas iteraciones operan en pequeños lotes de imágenes de entrada y se actualizan los pesos con cada lote procesado.

El proceso de aprendizaje comprende en minimizar una función de costo,  $J$ , la cual en tareas simples de clasificación mide el error entre las salidas deseadas,  $\bar{y}$  (*targets*), y la salida real de la red,  $y$  [17].

Para minimizar el error de la función de costo existen diversos esquemas de solución. El más utilizado es el método de descenso de gradiente estocástico (SGD, *Stochastic Gradient Descent*), en el cual se incluye la tasa de aprendizaje  $\alpha$  que permite controlar la magnitud del cambio en el ajuste de los pesos. Además, en este trabajo se incluyó la constante de momento, utilizada como una estrategia para acelerar la convergencia de la red hacia el error mínimo. Como se explicó con anterioridad, al evitar los cambios bruscos de la modificación de los pesos sinápticos al mantenerlos en una misma tendencia general lleva a la red neuronal al aprendizaje.

### 5. ARQUITECTURA CNN PROPUESTA

Se describe el modelo propuesto para la clasificación de las imágenes de iris en base al modelo AlexNet [12]. La entrada es la imagen normalizada del iris y la extracción de características tiene 4 capas convolucionales. Para la clasificación, se cuenta con 3 capas totalmente conectadas. La red AlexNet tiene 5 capas convolucionales y 3 totalmente conectadas. Se consideró una gran reducción de parámetros en cada capa convolucional y en la cantidad de neuronas en las capas totalmente conectadas.

En la Fig. 6(a), se muestra la imagen normalizada del iris y en la Fig. 6(b), el segmento seleccionado (de la izquierda) para asegurar que la imagen del entrenamiento no contenga ruido por oclusión de párpados. La sección coloreada de rojo en la Figura 6(a) izquierda es debido a la mala calidad de datos de entrenamiento y daría como resultado un mal entrenamiento en la red neuronal afectando en los resultados finales para la clasificación.

La arquitectura propuesta se muestra en la Fig. 7 y en la Tabla 3 están los detalles de cada componente en la red neuronal indicando el nombre de cada capa, el tipo de capa y el tamaño de cada resultante de aplicar cada operación en cada capa. Los filtros establecidos en la CNN propuesta se

modificaron proporcionalmente a partir de los filtros establecidos originalmente por la arquitectura AlexNet. Para no modificar por completo su funcionamiento al ir reduciendo los filtros, se probó el desempeño de la red estableciendo la cantidad de filtros con los mejores desempeños obtenidos.

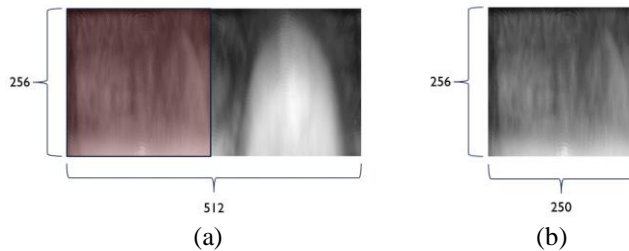


Figura 6.(a) Imagen normalizada. (b) selección de sección de la imagen sin ruido (por oclusión en la imagen derecha).

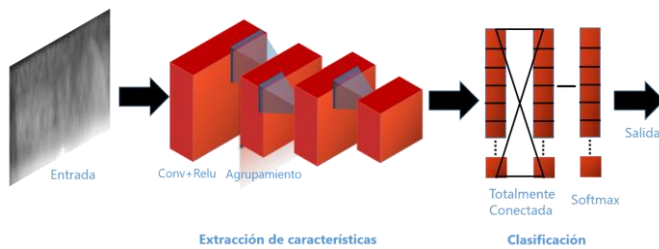


Figura 7. Representación visual de la arquitectura del modelo propuesto.

## 6. RESULTADOS EXPERIMENTALES

En la Tabla 4, se muestran los resultados al entrenar cada una de las tres arquitecturas implementadas: el modelo propuesto, el modelo AlexNet y el modelo VGG-16. Las arquitecturas se entrenaron bajo las mismas condiciones de hardware con un procesador Intel i5 de 7a generación, una GPU GTX 1050 de 4Gb de video dedicados y utilizando la misma versión de Matlab 2018a. Las imágenes de entrada están normalizadas y no cuentan con oclusión de párpados.

Se tomaron los desempeños obtenidos con cada conjunto de imágenes en entrenamiento, validación y prueba para cuantificar el funcionamiento del sistema. Se consideran de mayor importancia los desempeños obtenidos del conjunto de prueba, debido a que es el conjunto que no intervino en el entrenamiento de la red.

Se capturó el tiempo de clasificación del conjunto de pruebas con la red previamente entrenada y como se mencionó anteriormente, se obtuvo el error de desempeño del conjunto de pruebas.

Con los resultados (Tabla 4), se hizo una comparativa de los recursos que se requieren para el entrenamiento de cada una de las tres arquitecturas. Se tomaron en cuenta factores como el número de capas para extracción de características, el espacio de memoria que ocupan, la cantidad de filtros aplicados en todas las capas, entre otros. En la Tabla 5 se muestra los resultados del análisis al calcular un promediado del tiempo de clasificación. La arquitectura propuesta y la AlexNet tienen un promedio igual, teniendo ambas modelos ventajas sobre la VGG-16. En lo referente al número de capas, memoria ocupada y filtros totales la arquitectura propuesta ocupa menos recursos computacionales que

AlexNet y VGG-16. Adicionalmente, el desempeño promedio de la arquitectura propuesta es mejor que en las otras redes.

Tabla 3. Arquitectura del modelo propuesto detallando cada capa.

Nombre	Tipo	Activación
imageinput 256x250x1	Image input	256x250x1
conv_1 24 11x11x1 stride [4 4] padding [0 0 0]	Convolution	62x60x24
relu_1	ReLU	63x60x24
batchnorm_1 Batch Normalization with 24 channels	Batch normalization	62x60x24
maxpool_1 3x3 maxpooling stride [2 2] padding [0 0 0]	Max Pooling	30x29x24
conv_2 64 5x5x24 stride [1 1] padding [2 2 2]	Convolution	30x29x64
relu_2	ReLU	30x29x64
batchnorm_2 Batch Normalization with 64 channels	Batch normalization	30x29x64
conv_3 96 3x3x64 stride [1 1] padding [1 1 1]	Convolution	30x29x96
relu_3	ReLU	30x29x96
batchnorm_3 Batch Normalization with 96 channels	Batch normalization	30x29x96
maxpool_2 3x3 maxpooling stride [2 2] padding [0 0 0]	Max Pooling	14x14x96
conv_4 64 3x3x96 stride [1 1] padding [1 1 1]	Convolution	14x14x64
relu_4	ReLU	14x14x64
batchnorm_4 Batch Normalization with 64 channels	Batch normalization	14x14x64
maxpool_3 3x3 maxpooling stride [2 2] padding [0 0 0]	Max Pooling	6x6x64
fc_1 1024 fully connected layer	Fullly Connected	1x1x1024
relu_5	ReLU	1x1x1024
dropout_1 50% dropout	Dropout	1x1x1024
fc_2 1024 fully connected layer	Fullly Connected	1x1x1024
relu_6	ReLU	1x1x1024
dropout_2 50% dropout	Dropout	1x1x1024
fc_3 15 fully connected layer	Fullly Connected	1x1x15
relu_7	ReLU	1x1x15
softmax	Softmax	1x1x15
classoutput crossentropyex with '002' and 14 other classes	Classification Output	-

En base a los resultados obtenidos al calcular la desviación estándar del error, el modelo propuesto tiene mejor resultado que AlexNet por una diferencia aproximada de 8% por debajo de la media de los desempeños. Se omitieron los resultados obtenidos en el modelo VGG-16, ya que al ser una arquitectura demasiado grande en comparación a la cantidad de imágenes para entrenar la red no alcanza a converger a un

error mínimo, llevando al caso de un sobreentrenamiento de la red.

Tabla 4. Desempeños de la CNN propuesta con los diferentes conjuntos de datos establecidos.

Arquitectura	Desempeño entrenamiento	Desempeño validación	Desempeño prueba	Tiempo clasificación(ms)	Error
Propuesto	100	84.44	93.33	171	6.67
	100	95.55	93.33	218	6.67
	100	91.11	96.66	226	3.34
	100	97.78	93.33	195	6.67
	100	91.11	96.66	239	3.34
	100	80	93.33	188	6.67
	100	86.66	90	205	10
	100	95.56	96.66	177	3.34
	100	93.33	86.66	170	13.34
	100	100	96.66	146	3.34
AlexNet	100	71.11	60	204	40
	100	64.44	40	176	60
	100	66.66	50	191	50
	100	64.44	60	182	40
	100	53.33	56.66	178	43.34
	100	60	70	188	30
	100	80	80	177	20
	100	53.33	56.66	211	43.34
	100	66.67	56.66	214	43.34
	100	77.78	73.33	214	26.67
VGG-16	6.66	6.66	6.66	1060	93.34
	6.66	6.66	6.66	990	93.34
	6.66	6.66	6.66	997	93.34
	6.66	6.66	6.66	1010	93.34
	6.66	6.66	6.66	972	93.34
	6.66	6.66	6.66	1010	93.34
	6.66	6.66	6.66	1020	93.34
	6.66	6.66	6.66	875	93.34
	6.66	6.66	6.66	877	93.34
	6.66	6.66	6.66	937	93.34

Tabla 5 Análisis de los resultados obtenidos.

Arquitectura	Tiempo promedio	Desempeño promedio	Desviación estándar del error	Número de capas	Memoria ocupada	Filtros totales aplicados
Propuesta	193.5	93.66	3.313795407	7	13 MB	248
AlexNet	193.5	60.33	11.59530599	8	136 MB	928
VGG-16	974.8	6.66	1.49796E-14	16	488 MB	4224

Cuando se tienen más neuronas de las necesarias y en cuanto a los pesos que componen los kernels dio como resultado una red que pierde su capacidad de generalizar los resultados debido a que solo aprendió patrones específicos del conjunto de entrenamiento.

## 7. CONCLUSIONES

Se propone un modelo de CNN de bajo consumo de recursos computacionales para el reconocimiento del iris basada en la modificación de la red AlexNet. El modelo propuesto cuenta con un total de siete capas, de las cuales cinco son capas convoluciones y tres capas totalmente conectadas.

La red propuesta tuvo un funcionamiento más adecuado que AlexNet y VGG-16 teniendo un mejor desempeño promedio de un 93%, un poco por debajo de trabajos relacionados que alcanzaron un 98%. Sin embargo, considerando los recursos computacionales utilizados, la red propuesta es más factible para la creación de aplicaciones que puedan ser introducidas a un sistema embebido.

## REFERENCIAS

[1] M. Oravec, "Feature extraction and classification by machine learning methods for biometric recognition of face and iris," *ELMAR 56th Int. Symp.*, September, pp. 1-4, 2014.  
 [2] J. Daugman, "How Iris Recognition Works," *Essent. Guid. to Image Process.*, vol. 41, no. 1, pp. 715-739, 2009.  
 [3] M. Shamsi, P. Saad, and A. Rasouli, "Iris segmentation and normalization approach," *J. Teknol. Mklm.*, vol. 19, no. 2, pp. 88-101, 2008.

[4] M. G. Alaslani and L. A. Elrefaei, "Convolutional Neural Network Based Feature Extraction for IRIS Recognition," *Int. J. Comput. Sci. Inf. Technol.*, vol. 10, no. 2, pp. 65-78, 2018.  
 [5] A. Gangwar and A. Joshi, "Deepirisnet: Deep Iris Representation With Applications In Iris Recognition And Cross-Sensor Iris Recognition," *2016 IEEE Int. Conf. Image Process.*, vol. 2015-November 2016.  
 [6] F.F. Khan et al., "Iris Recognition using Machine Learning from Smartphone Captured Images in Visible Light," *ELMAR, 56th Int. Symp.*, vol. 7, pp. 26-28, September 2017.  
 [7] Z. Othman and A. Satria Prabuwo, "Preliminary study on iris recognition system: Tissues of body organs in iridology," *Proc. IEEE EMBS Conf. Biomed. Eng. Sci. IECBES*, pp. 115-119, December 2010.  
 [8] S. El-Naggar and A. Ross, "Which dataset is this iris image from?," *IEEE Int. Work. Inf. Forensics Secur. WIFS 2015 - Proc.*, November, 2015.  
 [9] M.S. Hosseini, B.N. Arabbi, and H. Soltanian-Zadeh, "Pigment melanin: Pattern for iris recognition," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 792-804, 2010.  
 [10] F. F. Khan, A. Akif, and M. A. Haque, "Iris Recognition using Machine Learning from Smartphone Captured Images in Visible Light," *IEEE Intl. Conf. Telecomm. and Photonics (ICTP)*, pp. 26-28, 2017.  
 [11] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. M. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Anal. Appl.*, vol. 21, no. 3, pp. 783-802, 2018.  
 [12] G. E. H. Alex Krizhevsky, Ilya Sutskever, "ImageNet Classification with Deep Convolutional Neural Networks," *LSVRC-2010*, vol. 12, 2012.  
 [13] O. Oyedotun and A. Khashman, "Iris nevus diagnosis: Convolutional neural network and deep belief network," *Turkish J. Electr. Eng. Comput. Sci.*, vol. 25, no. 2, pp. 1106-1115, 2017.  
 [14] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv Prepr. arXiv1502.03167*, 2015.  
 [15] K. Wang and A. Kumar, "Cross-spectral iris recognition using CNN and supervised discrete hashing," *Pattern Recognit.*, vol. 86, pp. 85-98, 2019.  
 [16] W. Zhang, C. Wang, and P. Xue, "Application of convolution neural network in Iris recognition technology," *4th Int. Conf. Syst. Informatics, ICSAI 2017*, pp. 1169-1174, January 2018.  
 [17] F. Ma, G. Po, S. Sa, and L. Ve, "A deep learning approach for iris sensor model identification," vol. 46, no. 3, pp. 1-8, 2016.