

# Identificación de personas a través de las manos mediante redes neuronales artificiales

EDWIN ALEJANDRO PERÉZ D. <sup>(1)</sup>, MIGUEL FELIPE ARÉVALO C. <sup>(2)</sup>,  
DIANA JANETH LANCHEROS C. <sup>(3)</sup>

(1) eperez28@unisalle.edu.co

(2) marevalo30@unisalle.edu.co

(3) dilancheros@unisalle.edu.co

Universidad de La Salle  
Bogotá, Colombia

# Identificación de personas a través de las manos mediante redes neuronales artificiales

---

## RESUMEN

---

*Palabras clave:*

***Red neuronal artificial;  
acondicionamiento y  
tratamiento de imágenes;  
red neuronal multicapa;  
resolución de imágenes;  
momentos invariantes;  
carpograma***

**S**on muchos los campos en donde llegar a tener un muy buen control de calidad, con la mínima intervención de la actividad humana es realmente necesario. En muchas ocasiones los costos de producción al emplear a uno o más técnicos, para un control de calidad, puede llegar a ser reemplazado por la automatización. Hoy en día con el fin de desarrollar sistemas más seguros se hace necesario el uso de los rasgos físicos de las manos de cada una de las personas. El presente artículo muestra el procesamiento de la imagen para poder lograr la identificación de diferentes personas a partir de una red neuronal artificial entrenada para reconocer las manos.

## I. INTRODUCCIÓN

El presente artículo describe el desarrollo de un sistema inteligente (software y hardware) para la identificación de personas a partir de las manos usando redes neuronales artificiales. En el desarrollo del proyecto inicialmente se realiza el reconocimiento de las manos mediante la toma de las imágenes que servirán como muestra, con el objetivo de establecer algunos patrones de reconocimiento. La metodología de desarrollo incluye las siguientes etapas:

La primera fase es la adquisición de imágenes, que involucra el ajuste de la iluminación y de la cámara web, esto con el montaje de un domo para la mejor visualización de los objetos.

La etapa dos, consta de realizar la selección de los filtros que se utilizarán para que, a partir de la imagen tomada, se pueda llegar a una imagen con unas características muy claras, que nos permitan la diferenciación entre las imágenes de muestreo y las imágenes en tiempo real.

La tercera etapa, consiste en el entrenamiento de una red neuronal artificial, a partir de las imágenes filtradas, para que, a partir de una imagen tomada y procesada, sea capaz de identificar la mano de la persona que ingresó al sistema.

La cuarta y última etapa, es la creación de una interfaz gráfica en Matlab en donde se integren todos los algoritmos utilizados previamente, realizando una interfaz amigable con el usuario y siendo así más fácil su operabilidad, esto como parte de las pruebas y la validación del sistema.

El presente artículo menciona en la sección II los antecedentes, en la sección III describe el diseño y el análisis del sistema, y en la sección IV menciona los resultados y la validación para que, a partir de estos, se realicen en la sección V los resultados y las recomendaciones finales del proyecto.

## II. ANTECEDENTES

Realizando una revisión de trabajos relacionados con el tratamiento de imágenes en sistemas de calidad, se encuentran trabajos como el de Drozdowicz *et al.* [1] donde el objetivo principal fue explorar el potencial de diferentes estructuras de redes neuro-

nales para realizar la segmentación de imágenes de resonancia magnética (IRM). La investigación permitió validar procesos de segmentación de imágenes mediante la utilización de RNA [1].

Vargas *et al.* [2] presentaron un sistema óptico de reconocimiento de patrones de imágenes usando redes neuronales artificiales para la identificación del lenguaje de señas empleado por las personas sordomudas y con dificultad en el habla para su comunicación. El proceso que se llevó a cabo para la identificación consistió en pre procesar las imágenes para adaptarlas y mejorar la capacidad de la red para discriminar los patrones, incluyendo filtrado, reducción y eliminación de ruido y detección de bordes. Finalmente concluyen que las RNA constituyen una herramienta óptima para la identificación y reconocimiento de patrones, debido a que reconoce las señas sin movimiento del abecedario de dicho lenguaje [2].

Achinga *et al.* [3] desarrollaron un sistema con algoritmos de extracción de características, clasificación, y verificación de imágenes de huellas dactilares en grandes bases de datos. Se empleó técnicas sofisticadas de procesamiento de imágenes en las diferentes etapas del sistema, se localizó el punto central de la huella a través del índice de Poincare y usando un banco de filtros de Gabor 2D, obteniendo así el FingerCode, se utilizaron algoritmos de k-vecindad, redes neuronales y el algoritmo de Euclides. Como principal conclusión de este trabajo se tiene que los algoritmos de tratamiento de imágenes utilizados permitieron reconocer las huellas dactilares [3].

Como se puede observar el desarrollo de sistemas que involucran visión artificial utilizan algoritmos para el procesamiento de las imágenes y en la mayoría de las ocasiones se emplean redes neuronales para el reconocimiento.

## III. DISEÑO Y ANÁLISIS DEL SISTEMA

El sistema se desarrolló teniendo en cuenta las siguientes fases:

### A. Adquisición de las imágenes.

En esta parte se capturaron 14 fotos, una de cada una de las manos de las personas que se tomaron en

cuenta para la muestra, con el fin de obtener, después del tratamiento y acondicionamiento de cada una de estas, suficiente información que permita más adelante poder entrenar la red neuronal artificial.

La adquisición de las imágenes se realizó utilizando una cámara Logitech Webcam Pro 9000 2MP, con el fin de llegar a tener una buena resolución en las imágenes con una cámara que sea más económica y factible a su uso, dado que posee una resolución de 1600x1200 y una velocidad de envío proporcionada por la conexión USB 2.0.

La Fig. 1 es una parte del resultado que se obtuvo al momento de capturar la fotografía.

**Fig. 1** Fotografía tomada con la cámara de 2 Mega píxeles.



Como se pudo evidenciar el resultado de la imagen, hay algunos problemas de ubicación de la mano, así como unas sombras en la imagen.

### *B. Ajuste de la parte de iluminación y de la cámara web.*

De acuerdo a la calidad de las imágenes que se obtuvieron en la etapa de adquisición, se observa una resolución que resalta con mayor detalle la forma de las manos. Sin embargo, fue necesario realizar un ajuste en el sistema de iluminación para poder obtener una mayor claridad en la imagen, en este caso se ingresó una segunda barra de leds, y su resultado se observa en la Fig. 2.

**Fig. 2.** Fotografía tomada con segunda barra de leds.



La forma de la mano es mucho más clara, con lo cual se puede garantizar que el proceso en el cual se le va a realizar un acondicionamiento va a obtener resultados más precisos.

A pesar de la resolución de la cámara, la calidad de las imágenes se encuentra dentro de un rango aceptable. Por lo tanto, se procedió a capturar catorce fotografías, una por cada mano de las siete personas seleccionadas para la muestra

### *C. Tratamiento de la imagen.*

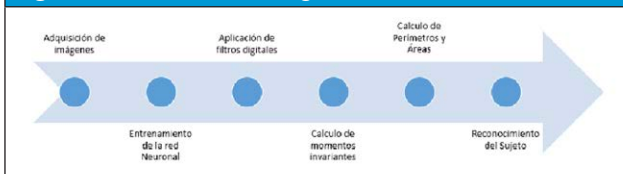
El objetivo final de esta etapa consistió en tomar cada una de las imágenes de la sección anterior y convertirlas en imágenes binarias en las cuales solo se observará la silueta de la mano, como se puede observar en la Fig. 3

**Fig. 3.** Imagen binaria de la Fig. 2



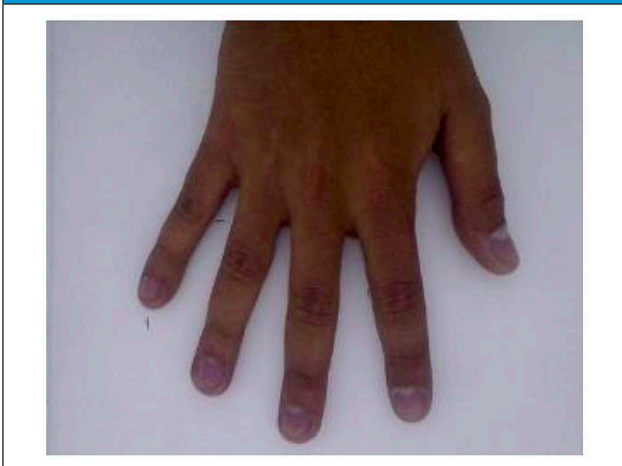
La estructura del algoritmo está compuesta por cuatro partes, la primera de ellas es la lectura de la imagen RGB que se obtuvo en la etapa de adquisición de imágenes mediante la GUI de Matlab; las otras tres partes corresponden al tratamiento que se le aplica a la imagen tomada, más adelante se explicara la función que cumple cada uno de estos. El siguiente diagrama (Fig. 4) representa la serie de pasos que se ejecutó.

**Fig. 4.** Tratamiento de la imagen



*Adquisición de la imagen RGB:* esta es la única parte del algoritmo que no realiza ningún tipo de tratamiento de la imagen, debido a que esta corresponde a la lectura de la fotografía con espacio de color RGB de la etapa anterior (Ver Fig. 5).

**Fig. 5.** Adquisición de imagen RGB.

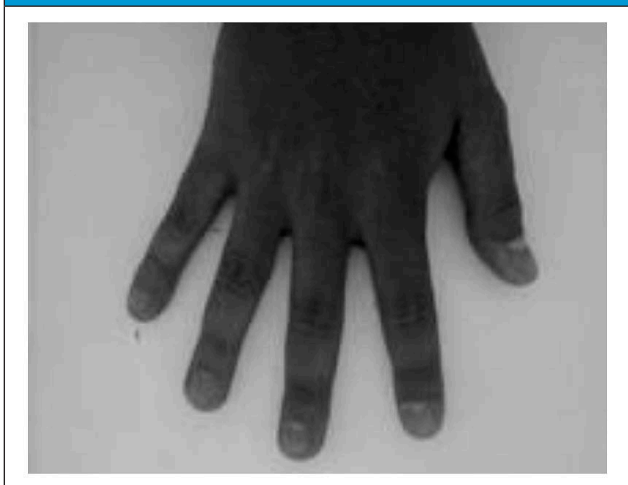


*Convertir a escala de grises:* convierte una imagen RGB a escala de grises, mediante la eliminación de la tonalidad y la saturación de información, mientras conserva la luminancia (Ver Fig. 6).

*Convertir a binario:* cambia una imagen en escala de grises a una imagen binaria. En este caso se sustituye todos los píxeles de la imagen de entrada con luminancia mayor a un nivel que el usuario define por un valor de 1 (blanco) y todos los otros píxeles los reemplaza por un valor de 0 (negro).

El valor de nivel se especifica en el rango  $[0, 1]$  [4] (Ver Fig. 7).

**Fig. 6.** Imagen con escala de grises.



**Fig. 7.** Imagen convertida a binaria.



*Esqueletización:* La esqueletización es una técnica que permite obtener, a partir de la imagen digital dada, otro que contiene la misma información topológica, y gran parte de la información geométrica del objeto inicial [3], esta es una operación morfológica que se le aplica a las imágenes para binarizarlas. (Ver Fig. 8).

Una vez que se obtuvo una imagen como la anterior para cada una de las diferentes muestras (Personas) se procedió a calcular los momentos invariantes (F1, F2, F3, etc.) y el CR (relación entre el perímetro y el área), los cuales se registraron en la siguiente tabla (Tabla N°1).

Fig. 8. Esqueletización de la imagen para binarizarla.



TABLA I. PERÍMETRO, ÁREA Y CR DE LAS IMÁGENES TOMADAS.

Persona	Mano	Área	Perímetro	CR
Diana	Derecha	30491 1	10348	351.1881
	Izquierda	33672 6	11211	373.3605
Edward	Derecha	47599 1	8561	153.975
	Izquierda	33807 8	10683	337.5744
Julián	Derecha	33996 6	9532	267.2591
	Izquierda	35658 7	11238	354.1706
Lucho	Izquierda	29655 9	10103	344.1831
Miguel	Derecha	35927 2	12846	459.3169
	Izquierda	32450 9	12710	497.8108
Pérez	Derecha	33518 6	8245	202.8128
	Izquierda	39001 1	9532	232.9653
Ing. Diana	Derecha	24952 5	7506	225.7891

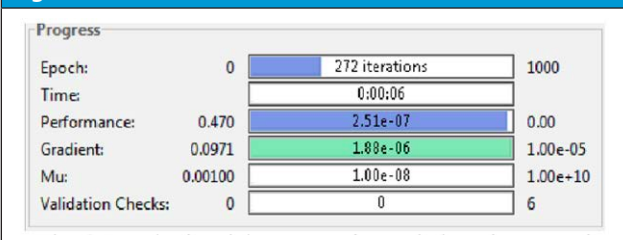
**D. Programación de la red neuronal artificial**

La arquitectura de la red neuronal que se empleó para identificar las manos es de tipo multicapa, en esta se empleó la función de entrenamiento por re-

tro-propagación Levenberg- Marquardt y la función de rendimiento del error cuadrático medio normalizado (mse por sus siglas en ingles). La red neuronal está configurada por ocho datos de entrada, doce capas ocultas, dos capas de salida y dos datos de salida.

Los datos de entrada corresponden a los valores que se registraron en las tablas anteriores (etapa de programación de filtros). Es necesario tener en cuenta que la red neuronal se entrenó empleando el toolbox de MATLAB, “Neuronal Network Training (ntraintool)”, con el cual se obtuvo información del progreso muy importante, un número de iteraciones igual a 270, un tiempo de entrenamiento de 0:00:20, un rendimiento de 2.50e-06 y un gradiente de 1.88e-06. En la Fig. 8 se puede observar estas cifras.

Fig. 8. Resultados del entrenamiento de la red neuronal.



Si el entrenamiento de la red se llevaba a cabo con un mayor número de capas, por ejemplo 50, el proceso tardaba más y el error era muy alto, en este caso 0:00:36 y 0.000796, respectivamente. El consumo de CPU también se elevaba. Las (Fig. 9 y Fig. 10) corroboran esta información.

Fig. 9. Resultados del entrenamiento, empleando 30 capas ocultas.

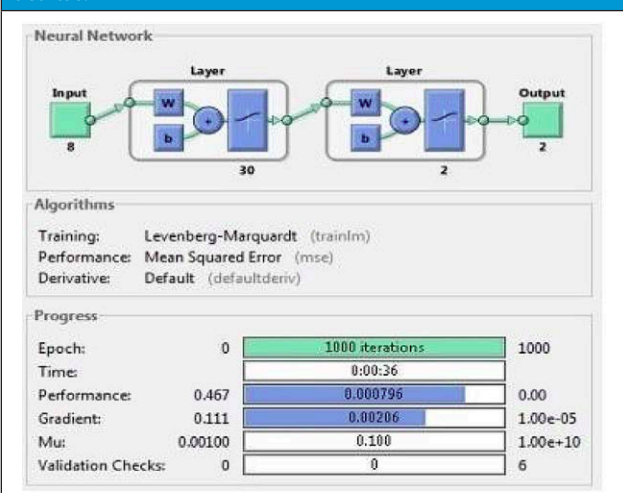
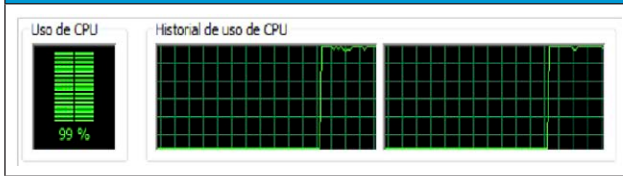


Fig. 10. Uso del CPU



### E. Diseño de la interfaz gráfica

El diseño de la interfaz gráfica se realizó a través de la herramienta GUIDE de Matlab, la cual nos permite realizar un sistema más amigable al momento de que el usuario la utilice, el uso de la interfaz es intuitivo al momento de utilizarse y está integrada para que el usuario pueda realizar de forma sencilla la identificación del objeto. A partir del uso de cada uno de los botones en la interfaz, se indican el tipo de tareas que el sistema debe realizar, se encuentra en la interfaz la opción de "Entrenar red neuronal", la función de activación de la cámara mediante el botón "iniciar Cámara, y a sí mismo, la captura de la imagen, para que finalmente con el botón "procesar" se reconozca la persona a través de su mano, cabe destacar, que con este comando, internamente, se inicia el proceso de tratamiento de la imagen; el resultado se puede visualizar por medio de un mensaje que se despliega en la parte inferior de la interfaz.

Debido a que el entrenamiento de la red neuronal puede llegar a aumentar el uso de CPU, como se pudo evidenciar anteriormente, el usuario solo podrá entrenar la red neuronal o encender la webcam, un proceso a la vez, así que cuando se presione el botón "Aprender" se recomienda no ejecutar otra tarea. En la parte superior de la interfaz saldrá un mensaje como el siguiente "El Sistema Terminó de Aprender" al momento de terminar el entrenamiento de la red.

Otra consideración a tener en cuenta, es el momento en el que se captura una fotografía, ya que una vez que se importe o se capture se tendrá que realizar la identificación de manera obligatoria, lo que quiere decir que estos dos procesos no se podrán realizar de forma simultánea, de no ser así el programa entraría en error. Estas consideraciones hacen del programa un sistema más eficiente y fácil de usar.

## IV. PRUEBAS Y VALIDACION DEL SISTEMA

En esta etapa se realizaron las pruebas correspondientes para verificar si el programa funciona correctamente, para ello se tomaron fotos piloto que posteriormente se ingresaron a la red neuronal, luego de su procesamiento para observar su resultado. Los resultados se muestran a continuación por medio de la imágenes.

Fig. 11. Reconocimiento de una mano ejemplo

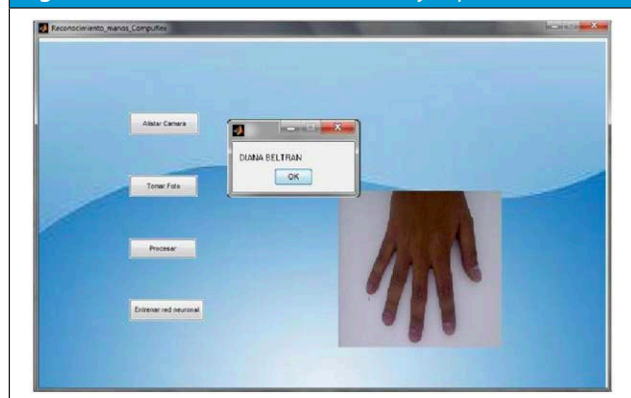


TABLA II. PRUEBAS

Persona	Mano	Número total de pruebas	Pruebas acertadas	Porcentaje de aciertos %
Diana	Derecha	20	20	100
	Izquierda	20	20	100
Edward	Derecha	20	13	65
	Izquierda	20	11	55
Julián	Derecha	10	6	60
	Izquierda	10	7	70
Lucho	Izquierda	10	2	20
Miguel	Derecha	7	4	57.14
	Izquierda	7	3	42.85
Pérez	Derecha	10	6	60
	Izquierda	10	4	40
Ing. Diana	Derecha	3	2	66.67

Al revisar el resultado de la identificación en cada caso, se encontró que, en personas como Diana, el sistema funcionó correctamente, mas sin embargo las pruebas con las demás personas, llevaban a una confusión en la red neuronal y el sistema no respondía correctamente en todos los casos.

### V. CONCLUSIONES

Con base a los resultados que se obtuvieron se pudo demostrar que sí es posible desarrollar un sistema inteligente sencillo que permita identificar los rasgos característicos de las personas; en esta primera prueba se logró la identificación de las personas cuyas características físicas de las manos era mucho más sobresalientes que las de las demás (Tamaño, Forma etc.).

Los elementos físicos como la cámara web y el sistema de iluminación se pueden conseguir muy fácilmente sin llegar a invertir mucho dinero.

Es un sistema muy económico y eficiente, ya que el error en la etapa más crítica, "programación de la red neuronal artificial", no fue muy alto. Este valor permitió identificar sin problemas las manos de aquellas personas que poseían características sobresalientes en sus manos.

El sistema puede llegar a ser un sistema a un nivel mucho más especializado con ayuda de una mejor cámara y con la utilización de una red neuronal mucho más amplia, en este proyecto se ve la limitación de las capas de la red con respecto al reconocimiento de las manos de las personas.

Para realizar un trabajo completamente funcional, se planea la idea de mejorar el sistema, para llegar a identificar las manos de todas las personas que se ingresen, dado que, a pesar de las similitu-

des, dos manos nunca pueden ser iguales, demostrado por cartogramas.

Durante el proceso de entrenamiento y a partir de los resultados que se obtuvieron de la red neuronal artificial se llegó a la conclusión que se debe colocar la cantidad suficiente de capas, con el fin de no sobredimensionar el sistema, porque en ocasiones, el error es mucho más bajo cuando se empleaban muchas capas, a diferencia de cuando se empleaban menos.

### REFERENCIAS

- [1] Bartolomé Drozdowicz, Guillermo Bernasconi, Mariano Reyes, Fernando Saba, and Gonzalo Simón, "Segmentación semiautomática de imágenes de resonancia magnética, basada en redes neuronales artificiales," *Ciencia, Docencia y Tecnología*, vol. XVI, no. 30, pp. 117155, Mayo 2005.
- [2] L. Vargas, L. Barba, and L. Mattos, "Sistema de identificación de lenguaje de señas usando redes neuronales artificiales," *Revista Colombiana de Física*, vol. 42, no. 2, 2010.
- [3] Erik Valdemar Cuevas Jimenez and Daniel Zaldivar Navarro. slideshare. [Online]. <http://www.slideshare.net/umbrela/23371-procesamientodeimagenesconmatlab>
- [4] Omar Sanchez. (2008, Mayo) slideshare. [Online]. <http://www.slideshare.net/omarspp/imagen-morfologicas>
- [5] Susan Lucy Palacios Salcedo, Registro de imágenes multiespectrales aéreas utilizando la transformada de Wavelets. Lima, Perú, 2012.
- [6] [http://master.us.es/masterma1/TfM\\_pdfs%28Julio11%29/Raul\\_Reina\\_Molina.pdf](http://master.us.es/masterma1/TfM_pdfs%28Julio11%29/Raul_Reina_Molina.pdf)
- [7] <http://www.mathworks.com/help/images/ref/bw-morph.html>