



Using neural networks for face recognition in controlled environments

Uso de redes neuronales para el reconocimiento de rostros en ambientes controlados

Holman Montiel Ariza¹, Fernando Martínez Santa², Diego Armando Giral Ramírez³

Fecha de recepción: 12 de julio de 2014

Fecha de aceptación: 18 de agosto de 2015

Cómo citar: Montiel Ariza, H., Martínez Santa, F., & Giral, D. (2015). Using neural networks for face recognition in controlled environments. *Revista Tecnura*, 19, 67-77. doi: 10.14483/udistrital.jour.tecnura.2015.SE1.a05

Abstract:

This paper discusses the implementation and validation of neural networks for face recognition in controlled environments, an implementation strategy that consists of 3 stages: acquisition, extraction of facial parameters and validation using neural networks is presented. This paper seeks to validate the operation of two neural networks in particular for face recognition, these are the Perceptron and ART. We seek to analyze and validate through a series of tests the possible operation of this type of neural networks in biometric identification systems and access control.

Keywords: MLP, ART, Neural Network, Facial Parameters, Perceptron

Resumen:

Este artículo presenta la implementación y validación de redes neuronales para el reconocimiento de rostros en entornos controlados, se presenta una estrategia de implementación que consta de 3 etapas: adquisición, extracción de parámetros faciales y validación mediante redes neuronales. Este trabajo busca validar el funcionamiento de dos redes neuronales en particular para el reconocimiento de rostros, estas son el Perceptron y la ART. se busca analizar y validar mediante una serie de pruebas el posible funcionamiento de este tipo de redes neuronales en sistemas biométricos de identificación y control de acceso.

Palabras claves: MLP, ART, Red neuronal, Parámetros Faciales, Perceptrón

¹ Ingeniero en Control Electrónico e Instrumentación, especialización en Telecomunicaciones, master en Seguridad Informática. Docente Asistente Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Contacto: hmontiela@udistrital.edu.co

² Ingeniero en Control Electrónico e Instrumentación, magister en Ingeniería Electrónica y de Computadores. Docente Asistente Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Contacto: fmartinezs@udistrital.edu.co

³ Ingeniero en Electricidad, candidato magister en Ingeniería Eléctrica. Docente tiempo completo Universidad Distrital Francisco José de Caldas. Bogotá, Colombia. Contacto: dagiralr@udistrital.edu.co

INTRODUCCIÓN

La medición biométrica se ha venido estudiando desde tiempo atrás y se considera en la actualidad el método ideal de identificación humana (Tolosa Borja y Giz Bueno, s. f.)

Cada sistema biométrico utiliza una cierta clase de interfaz para recopilar la información sobre la persona que intenta acceder. Un software especializado procesará esa información en un conjunto de datos que se pueden comparar con los modelos de los usuarios que se han introducido previamente al sistema. Si se encuentra un “matching” o concordancia con la base de datos, se confirma la identidad de la persona y se concede el acceso (Tolosa Borja y Giz Bueno, s. f.)

Este artículo pretende validar el funcionamiento de redes neuronales para el reconocimiento de rostros previamente almacenados y bajo un ambiente controlado frente a condiciones lumínicas.

Este proceso se inicia con la adquisición de los rostros por identificar mediante una cámara digital; en este bloque se deben tener en cuenta las restricciones del proyecto, ya que al tomar una fotografía para el reconocimiento de rostros hay muchos parámetros que influyen en el desempeño del algoritmo de extracción de parámetros faciales, puesto que si no se tienen características como iluminación uniforme del rostro, un fondo uniforme o de color semejante al de la piel se tendrán dificultades en el reconocimiento. La relación entre el largo y ancho de las fotos siempre se debe mantener, se proyecta una resolución de trabajo de 1280 x 960 píxeles y mayores; se debe tener en cuenta que a mayor resolución tendremos un procesamiento más lento y si tenemos menos resolución se dificultará la extracción de los parámetros faciales.

Existen varios métodos y algoritmos para la compensación de iluminación, en este caso se implementó *White World (Perfect Reflector)*.

Este algoritmo toma una imagen, determina cuál es su mayor intensidad y define esta intensidad como el “blanco”. Para calcular el punto de

mayor intensidad lo que se hace es buscar el punto de menor distancia al blanco [255,255,255]. Cada uno de los canales de color R, G y B son normalizados con respecto al punto de “blanco” encontrado.

El algoritmo de compensación de iluminación que se utiliza es el propuesto por M. Abdel-Mottaleb y A. K. Jain (Hsu, Abdel-Mottaleb, & Jain, 2002, pp. 696-706); esta técnica usa un “blanco de referencia” para normalizar la apariencia del color.

Se consideran píxeles que estén en 5% del valor superior de la luminancia como “blanco de referencia”, solo si el número de estos píxeles es suficientemente grande (> 100). Las componentes RGB de la imagen se ajustan de manera que el valor medio de esos píxeles del “blanco de referencia” es escalado linealmente a 255; la imagen no se cambia si no se detecta un número suficiente de píxeles de “blanco de referencia” (Gámez Jiménez, 2009).

Extracción de parámetros faciales

Esta etapa se divide en dos procedimientos, el primero identifica en la fotografía los posibles rostros y luego los procesa para identificar los parámetros faciales.

Identificación de rostro

Una vez almacenada la fotografía, el software debe identificar los píxeles que contengan el color piel (Terrillon, Shirazi, Fukamachi, & Akamatsu, 2000, pp. 54-61). Se hace un estudio comparativo de nueve espacios de color distintos para la detección de caras y se concluye que el espacio TSL (tinte-saturación-luminancia) brinda los mejores resultados modelando las densidades de probabilidad de clústeres de piel y no piel, con dos gaussianas o con una mezcla de gaussianas. En este y otros espacios como el YCbCr la luminancia y la croma están en canales separados (Aguerreberre, Capdehourat, Delbracio, y Mateu, 2005). Sin embargo, este es solo el primer paso para la identificación del rostro; se utiliza un clasificador de piel, que define específicamente los límites de la

región correspondiente al color de la piel, a través de unas reglas numéricas. Este método es aplicable en diferentes espacios de color (Benito, 2005). La regla utilizada para una imagen en RGB es:

$$R > 95 \wedge G > 40 \wedge B > 20 \text{ y}$$

$$\max\{R, G, B\} - \min\{R, G, B\} > 15 \text{ y}$$

$$|R - G| > 15 \text{ y } R > G \text{ y } R > B$$

Una vez obtenida la clasificación de los colores de la imagen con base en el color piel, esta área debe pasar por un proceso de erosión y dilatación. Primero se realiza una apertura; aquí la erosión elimina las pequeñas y finas partículas que quedan de la etapa anterior y que no son más que componentes de ruido y tienen muy poca probabilidad de ser una cara. Luego la dilatación recompone los componentes que no son eliminados (Aguerreberre, Capdehourat, Delbracio, y Mateu, 2005) (figura 1).

Selección de candidatos a cara

Para que una región de la imagen se considere un candidato a cara debe cumplir las siguientes condiciones:

Área: las regiones candidatas a cara deben tener un área mínima de 400 píxeles. Este valor permite eliminar regiones que son excesivamente pequeñas como para ser una cara. Este valor determina el tamaño mínimo que debe tener la cara para ser reconocida.

- **Factor de forma:** mide la regularidad de una región (ecuación (1)). Las caras son más o menos circulares, por lo que su factor de forma debería tender a 1, pero al detectar los píxeles de piel es muy probable que también se detecten zonas de cuello, por lo que la forma de la región será más alargada, más parecida a una elipse, por lo que el factor de forma se reduce. El valor de este parámetro debe ser superior a 0.1 (Gámez Jiménez, 2009).

$$ff = 4 * \pi * \frac{\text{área}}{\text{perímetro}^2} \quad (1)$$

- **Relación de aspecto:** es la relación entre el diámetro mayor y el diámetro menor de la región (ecuación (2)). Con este parámetro se tiene una medida cuantitativa de lo regular que es la región. Se ha determinado que el valor mínimo de esta propiedad debe ser 0.3 (Gámez Jiménez, 2009).

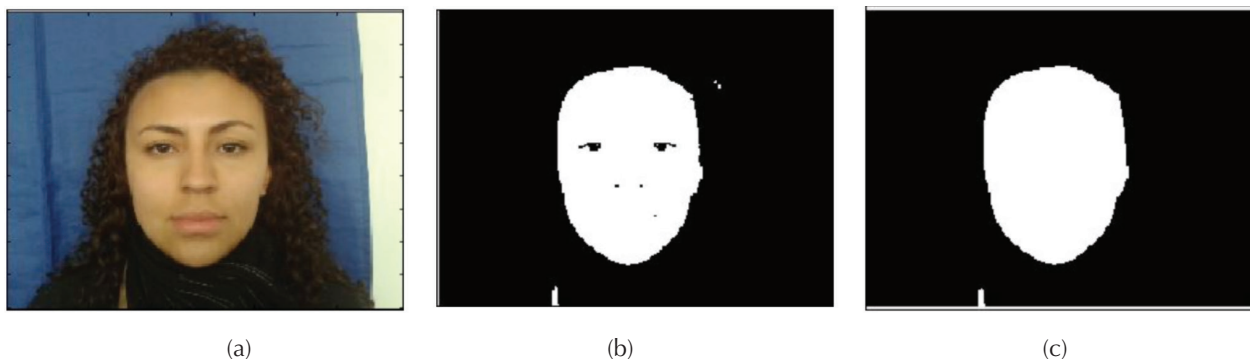


Figura 1. (a) Imagen RGB; (b) Clasificador de color piel; (c) Proceso de erosión y dilatación.

Fuente: elaboración propia

$$ra = \frac{\text{eje mayor}}{\text{eje menor}} \quad (2)$$

- **Solidez:** indica qué porcentaje de la región está dentro del cerco convexo de la misma. El cerco convexo es la mínima región convexa que contiene a la región. Su valor debe ser alto para regiones candidatas a cara, por lo que su valor mínimo se fija arbitrariamente en 0.52 (Gámez Jiménez, 2009).
- **Extensión:** da una medida del área que ocupa una región en el Bounding Box correspondiente. El Bounding Box es el rectángulo de menor tamaño que encierra a la región de interés. También se asume que las caras tendrán un valor alto de extensión (Aguerreberre, Capdehourat, Delbracio, y Mateu, 2005).

Si el área cumple con estas cinco condiciones se debe tomar como un candidato a cara, luego cada uno de los candidatos debe ser procesado para identificar candidatos a ojos y boca.

Identificación de parámetros faciales

Entre las diversas características faciales, las más importantes para el reconocimiento de rostros son los ojos y la boca. La mayoría de los enfoques para la localización de ojos se basa en plantillas; sin embargo, se puede situar directamente los ojos,



Figura 2. Imagen RGB e imagen YCbCr.

Fuente: elaboración propia

boca y contorno facial en función de sus características de mapas derivados tanto de Luma como de Croma de la imagen (Hsu, Abdel-Mottaleb, & Jain, 2002, pp. 696-706) (figura 2).

Identificación de ojos

- Para poder efectuar la identificación de los ojos se deben realizar dos mapas de ojos por separado, uno de los componentes de luminancias y uno de los componentes de los cromas; una vez hechos los dos mapas, estos se combinan para obtener el mapa único de ojos. Basados en la observación se puede afirmar que los ojos tienen un alto componente de croma azul (C_b) y un baj componente de croma rojo (C_r) (figura 3). Con base en este análisis Hsu, *et al.* (Hsu, Abdel-Mottaleb, & Jain, 2002, pp. 696-706) proponen la ecuación (3).

$$\text{Mapa de ojo } C = \frac{1}{3} * \left\{ (C_b^2) + (255 - C_r^2) + \left(\frac{C_b}{C_r} \right) \right\} \quad (3)$$

Donde C_b y C_r deben ser normalizados en un rango de [0 , 255] (figura (3)).

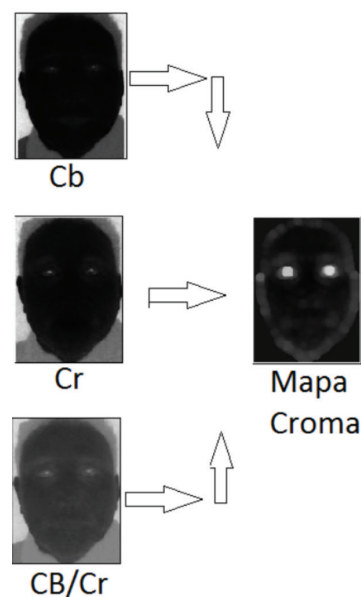


Figura 3. Mapa de cromas.

Fuente: elaboración propia

Para implementar la segunda sentencia se utilizó morfología en escalas de grises. Para esto se utiliza un elemento de estructura hemisférica g_o . Lo que se propone es resaltar las zonas que tienen alto contenido de Y y que a su vez tienen vecinos cercanos con bajo contenido de Y . Para esto se dilata el componente Y (imagen de luma), obteniendo así una imagen que resalta las zonas en donde hay alto contenido de luma-brillos. Por otro lado, se erosionó la Y , resaltando (hacia el 0) las zonas en donde hay poco componente de luma. Si el elemento de estructura es suficientemente grande, las zonas que tienen alto y bajo componente de luma van a aparecer en las dos imágenes resaltadas (en el primer caso, resaltadas positivamente y en el segundo, negativamente). Luego se realizó la división entre la imagen dilatada y la erosionada, obteniendo así el mapa de ojos considerando la luma, la ecuación (4) rige el sistema.

$$\text{Mapa de ojo } L = \frac{Y(x,y) \oplus g_o(x,y)}{Y(x,y) \ominus g_o(x,y) + 1} \quad (4)$$

El tamaño del elemento de estructura utilizado es una medida del mínimo tamaño de ojo por encontrar en la imagen (figura 4); se estima a partir de la ecuación (5).

$$\sigma = \frac{\sqrt{W * H}}{2 * Fg} \quad (5)$$

Donde: W : ancho de la imagen (máscara); H : alto de la imagen (máscara); Fg : cociente entre máximo tamaño de ojo sobre tamaño promedio de cara. Tomamos $Fg = 7$ píxeles, \oplus : dilatación, \ominus : erosión.

Identificación de boca

Para construir el mapa de boca de la imagen se supone que en la boca existe mayor cantidad de croma roja Cr que de croma azul Cb . Hsu *et al.* (Hsu, Abdel-Mottaleb, & Jain, 2002, pp. 696-706) proponen la ecuación (6) y (7) en función de los cromas (figura 5).

$$\text{Mapa de boca} = C_r^2 * \left(C_r^2 - \eta * \frac{C_r}{C_b} \right)^2 \quad (6)$$

Donde:

$$\eta = 0.95 * \frac{\text{promedio}(C_r^2)}{\text{promedio}\left(\frac{C_r}{C_b}\right)} \quad (7)$$

Todos los componentes de la ecuación (6) y (7) deben ser normalizados a $[0, 255]$ (figura 5).

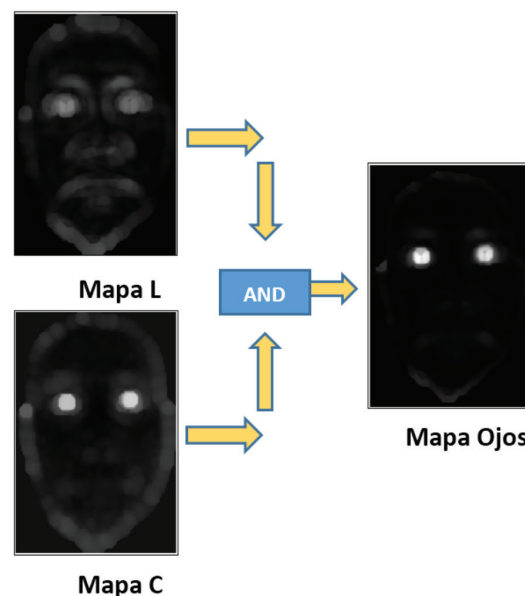


Figura 4. Mapa de ojos.

Fuente: elaboración propia

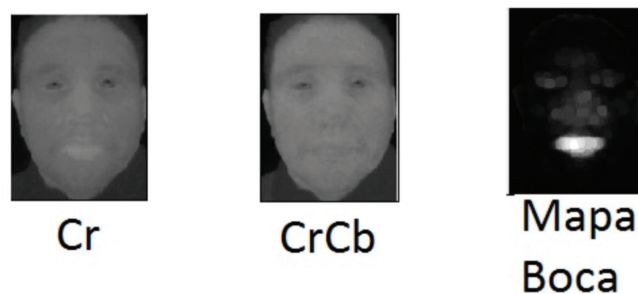


Figura 5. Mapa de boca.

Fuente: elaboración propia

Validación por restricciones geométricas

Por último, se desarrolla una etapa que combina los resultados encontrados en las etapas anteriores con el fin de validar si la región de piel considerada es una cara o no. En el caso en que se considere como cara, se debe indicar la (mejor) ubicación de los ojos y la boca. Se realiza todas las combinaciones posibles entre los candidatos a ojos y boca hallados (*si hay n bocas, m ojos, entonces hay n combinaciones de m tomadas de a dos ternas posibles*). A cada combinación la se le denomina *candidato*.

Para cada candidato que supere un conjunto mínimo de restricciones se calcula un score que contempla qué tan *buen* candidato es.

Con cada par de ojos y cada boca se forma un triángulo, cuyos vértices son los centroides de las regiones asociadas a los ojos y boca en consideración. Primero se aplica una etapa denominada de **mínimas condiciones**, en la cual si un candidato no las cumple es eliminado. Esta consiste en:

- El triángulo formado por los tres centroides (2 ojos, 1 boca), debe ser agudo (no puede tener ninguno de sus ángulos superior a 90°).
- El triángulo debe tener un área mínima de 100 pixeles (esta restricción está asociada al tamaño mínimo de cara por detectar).

Score de simetría y orientación

Para el cálculo de este score se tuvo en cuenta dos aspectos:

- Simetría: qué tan isósceles es el triángulo formado (ángulo q_1).
- Orientación: qué tan inclinado respecto a la vertical se encuentra (ángulo q_2).

Para el cómputo del score se utilizó la ecuación (8).

$$ScOrienta = \prod_{k=1,2} e^{-3 \cdot \text{sen}^2\left(\frac{\theta_k}{2}\right)} \quad (8)$$

Está formado básicamente por gaussianas que dependen del cuadrado del seno del ángulo. El propósito es favorecer a las caras simétricas y a las que se encuentran muy cerca de la vertical.

Score de simetría de tamaño de ojos

Se implementó un score que contempla positivamente el hecho de que las regiones formadas por cada ojo tengan áreas similares. El score es una campana descrita por la ecuación (9).

$$ScOrienta = \prod_{k=1,2} e^{\frac{Area_i - Area_j}{\min\{Area_i, Area_j\}}} \quad (9)$$

Score de coherencia de tamaño de cara

Mediante un breve estudio estadístico se concluye que el área del triángulo formado por los ojos y la boca es 10% del área total de la cara. Si se aproxima el área de la cara al área de la máscara, se puede implementar un score para favorecer a los candidatos que cumplan esto (ecuación (10)).

$$ScOrienta = e^{-\sqrt{Area(Trgulo) - Area(Mask)}} \quad (10)$$

Luego, los tres scores se agrupan en un único score mediante una combinación lineal (ecuación (11)).

$$Sc = 0,5 \cdot ScOrienta + 0,25 \cdot ScAreaOjos + 0,25 \cdot ScTamaño \quad (11)$$

Es necesario tener en cuenta que el score más determinante para la validación de una cara es el de orientación y simetría.

Por último, se devuelve como válido aquel candidato que tenga mayor score y cuyo valor sea superior a 0.6. Si no existen candidatos que tengan score superior a 0.6, se considera que el candidato no es una cara.

Una vez terminado el proceso de validación de los candidatos a cara, se obtienen 12 coordenadas de los puntos característicos del rostro. Para poder hacer el reconocimiento del rostro se deben tener magnitudes normalizadas, por lo cual las coordenadas no servirían debido a que en el momento de cambiar la fotografía o que el rostro estuviera más

lejos o cerca de la cámara, estas cambiarían, por lo cual se toman las distancias euclidianas entre cada punto de referencia (ecuación (12)).

$$d = \sqrt{(x1-x2)^2 + (y1-y2)^2} \quad (12)$$

Esta distancia se obtiene en pixeles; esta medida cambiará dependiendo de la resolución de la cámara y de la distancia que haya entre la cámara y el individuo, por lo cual no serviría como patrón para el reconocimiento; por este motivo esta distancia es normalizada con base en el área del rostro con el fin de que los patrones se vean menos afectados con los cambios de resolución y la distancia entre la cámara y el individuo. En la figura 6 se observan los puntos característicos y las distancias medidas.

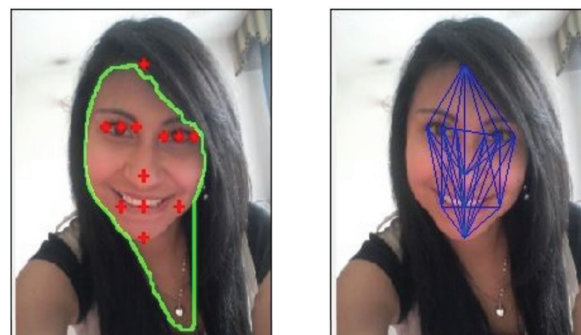


Figura 6. Derecha: puntos característicos del rostro; izquierda: mediciones efectuadas para obtener el patrón

Fuente: elaboración propia

REDES NEURONALES IMPLEMENTADAS

En general, cualquier sistema neuronal puede utilizar distintos paradigmas para el aprendizaje de la red, al igual que distintos algoritmos de entrenamiento. La figura 7 muestra el proceso de aprendizaje dividido en los paradigmas y algoritmos antes mencionados (Mejía Sánchez, 2004).

En las redes neuronales utilizadas se emplea el paradigma de aprendizaje supervisado y el algoritmo de corrección del error, algunas veces conocido como la regla Delta. Hablar de aprendizaje

supervisado (supervised learning) hace referencia al tipo de entrenamiento en el cual se provee al sistema con información de las entradas, al igual que se proveen las salidas esperadas o destinos correspondientes a dichas entradas a fin de que el sistema tenga los destinos como punto de referencia para evaluar su desempeño con base en la diferencia de estos valores y modificar los parámetros libres con base en esta diferencia, como lo muestra la figura 8 (Mejía Sánchez, 2004).

Basados en estos criterios de operación y una vez normalizado un vector de entrada para los datos obtenidos en el proceso de identificación de parámetros faciales se evaluar dos tipos de redes neuronales, las cuales son:



Figura 7. Procesos de aprendizaje.

Fuente: elaboración propia

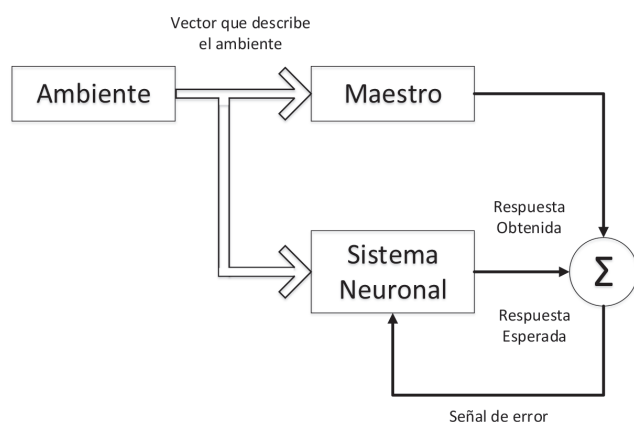


Figura 8. Diagrama de bloques del aprendizaje supervisado.

Fuente: elaboración propia

MLP: El perceptrón multicapa es una red de alimentación hacia adelante (feed-forward) compuesta por una capa de unidades de entrada (sensores), otra capa de unidades de salida y un número determinado de capas intermedias de unidades de proceso, también llamadas capas ocultas porque no tienen conexiones con el exterior (García García, 2013).

ART2: Este modelo se basa en la idea de hacer resonar la información de entrada con los prototipos o categorías que reconoce la red (Sesmero Lorente, 2012). Si entra en resonancia con alguno (es suficientemente similar), la red considera que pertenece a dicha categoría y se realiza una adaptación que incorpora algunas características de los nuevos datos a la categoría existente.

ANÁLISIS Y RESULTADOS DE LAS REDES NEURONALES

MLP-BackPropagation:

Para el desarrollo del proyecto se planteó una red neuronal con dos capas ocultas; sin embargo, estas condiciones pueden ser modificadas para validar tiempos de respuesta e iteraciones. Primero se debe realizar la selección de valores aleatorios

para el número de neuronas de las capas ocultas; el único valor que se tiene predefinido es el número de neuronas de salida, el cual es igual al número de sujetos por identificar. Se efectuaron pruebas a 1.000 y a 10.000 iteraciones, las cuales no fueron favorables ya que con este número de iteración no se logró llegar al error estimado de 0.001. Una aproximación para calcular el número de neuronas de cada una de las capas ocultas depende del número de sujetos por identificar y del número de imágenes con que se cuenta de cada sujeto para el entrenamiento de la neurona. Las ecuaciones (13) y (14) permiten determinar las neuronas.

$$\#NO1 = 2 \cdot (s + 1) \quad (13)$$

$$\#NO2 = (s + f) \quad (14)$$

Donde:

#NO1 = número de neuronas de la capa oculta 1;
 #NO2 = número de neuronas de la capa oculta 2;
 s = número de sujetos por reconocer;
 f = número de fotografías por sujeto.

En la tabla 1 se puede apreciar los datos obtenidos para algunas redes de prueba; allí se puede observar que en las redes seleccionadas se mejoró el entrenamiento en el primer caso cerca de 50%, y en el segundo caso cerca de 33%.

Se realizó el entrenamiento de una red neuronal para tres sujetos, esta red fue entrenada con 5 fotos de cada sujeto y el error resultado del entrenamiento fue de 0.001. Luego se verificaron los resultados con 10 fotos de cada sujeto, dando como resultado la tabla 2.

Como se observa en la tabla 2, la red solo cometió un error en la fotografía número 22, la cual se refería al sujeto número 1; los datos en cero son imágenes en las cuales no existen rostros. Esto nos indica que la red neuronal tuvo una efectividad de 96.66%. Este dato puede ser mejorado, solo se debe reentrenar la red neuronal con más fotografías de cada uno de los sujetos; sin embargo, esto implica un tiempo de entrenamiento mayor.

Tabla 1. Comparación del número de neuronas con respecto al error.

# Neuronas O1	# Neuronas O2	# Neuronas salida	# Iteraciones	Error mse
5	5	2	1000	0.277
5	5	2	10000	0.00147
8	7	2	1000	0.161
8	7	2	8214	0.001
5	5	3	10000	0.0228
5	5	3	1000	0.287
10	8	3	1000	0.0956
10	8	3	10000	0.00646

Fuente: elaboración propia

Tabla 2. Resultados de verificación del entrenamiento, Red MLP.

	S1	S2	S3	S1	S2	S3
1	0.99732671	5.5421E-08	0.04141823	1	0	0
2	0.0222233	0.99638852	0.02252849	0	1	0
3	0.02961242	7.2819E-05	0.96707754	0	0	1
4	0.92098959	0.00497382	0.03042537	1	0	0
5	0.0222233	0.99638852	0.02252849	0	1	0
6	0.02961242	7.2819E-05	0.96707754	0	0	1
7	0.99629076	4.6695E-08	0.07227149	1	0	0
8	0.0222233	0.99638852	0.02252849	0	1	0
9	0.02961242	7.2819E-05	0.96707754	0	0	1
10	0.994998	1.4536E-06	0.00085969	1	0	0
11	0.00058194	0.99996957	0.07923894	0	1	0
12	0.02961242	7.2819E-05	0.96707754	0	0	1
13	0.98716473	0.00580057	0.00176849	1	0	0
14	0.0222233	0.99638852	0.02252849	0	1	0
15	0.02961242	7.2819E-05	0.96707754	0	0	1
16	0.86376236	3.6642E-09	0.65972678	0	0	0
17	0.0222233	0.99638852	0.02252849	0	1	0
18	0.02961242	7.2819E-05	0.96707754	0	0	1
19	0.994998	1.4536E-06	0.00085969	1	0	0
20	0.25029691	8.5261E-06	0.61607902	0	0	0
21	0.02961242	7.2819E-05	0.96707754	0	0	1
22	0.07245527	0.98030011	0.02447715	0	1	0
23	0.0222233	0.99638852	0.02252849	0	1	0
24	0.02961242	7.2819E-05	0.96707754	0	0	1
25	0.994998	1.4536E-06	0.00085969	1	0	0
26	0.10562683	0.92410413	0.01244898	0	1	0
27	0.02961242	7.2819E-05	0.96707754	0	0	1
28	0.994998	1.4536E-06	0.00085969	1	0	0
29	0.89030601	7.7654E-07	0.27551846	0	0	0
30	0.02961242	7.2819E-05	0.96707754	0	0	1

Fuente: elaboración propia

ART-2

El planteamiento de la red Art 2 es muy distinto al de la red MLP, debido a que la red ya no se entrena; esto se debe a que esta red pertenece a la familia de redes no supervisadas. Esta familia tiene las facultades de autoentrenarse a medida que los patrones van entrando; sin embargo, el resultado esperado con esta red no fue el óptimo. Esto se debe a que los datos obtenidos de la extracción de parámetros cambian dependiendo de la foto, inclusive siendo del mismo sujeto. La red ART 2 ajusta sus pesos utilizando la ecuación (15).

$$W_{ji}(t+1) = \frac{e_i^k + w_{ji}(t) * Num_j(t)}{Num_j(t) + 1} \quad (15)$$

La ecuación (15) describe el comportamiento de los patrones pertenecientes a la misma categoría; sin embargo, en el vector patrón proveniente del módulo de extracción, la máxima distancia en la que se encuentran los datos es de 0.0345. En este orden de ideas la red no reconocerá a la perfección los patrones entregados; si se aumentan los parámetros de vigilancia la red empezara a crear más clases de los mismos sujetos. En la tabla 3 se observa los datos de la prueba realizada con los mismos 3 sujetos con los que se entrenó la red MLP, se utiliza un parámetro de vigilancia de 0.06.

Para mejorar los datos obtenidos por esta red se requiere que las fotografías sean tomadas en entornos totalmente controlados, incluyendo iluminación, resolución, distancia, para una aplicación real estas condiciones no son fáciles.

Tabla 3. Resultados de las pruebas ART 2.

Iter.	Sujeto	Semejanza	Neurona	Iter.	Sujeto	Semejanza	Neurona
1	S1	0,22105323	1	23	S1	0,04386593	NN
2	S1	0	1	24	S2	0,00426833	2
3	S2	0,06375721	NN	25	S3	2.95E-03	3
4	S2	0	2	26	S1	0,06710177	NN
5	S3	0,04080242	NN	27	S2	0,00373479	2
6	S3	0	3	28	S3	2.95E-03	3
7	S1	0	1	29	S1	0	5
8	S1	5.42E-04	1	30	S2	0,10104278	NN
9	S2	2.60E-04	2	31	S3	2.86E-03	3
10	S2	4.77E-04	2	32	S1	0,02796869	2
11	S3	0	3	33	S2	0,00432322	2
12	S3	5.20E-04	3	34	S3	3.08E-03	3
13	S1	0,35881126	NN	35	S1	0	5
14	S1	0	4	36	S2	0,05587395	NN
15	S2	8.67E-05	2	37	S3	4.64E-03	3
16	S2	3.47E-04	2	38	S1	0	5
17	S3	0	3	39	S2	0,14013228	NN
18	S3	6.51E-04	3	40	S3	4.77E-03	3
19	S1	0	5	41	S1	0,09682883	NN
20	S1	0	5	42	S2	0,02522165	1
21	S2	0,02987835	2	43	S3	0,06988724	NN
22	S3	1.65E-03	3	44	S2	0,03943498	1

Fuente: elaboración propia

CONCLUSIONES

Este artículo presenta los resultados obtenidos frente a la implementación de dos redes neuronales para el reconocimiento de rostros, determinando que el proceso de adquisición de parámetros faciales se puede lograr obteniendo una selección de puntos específicos o patrones de medición localizados en el rostro de una persona; dicho procedimiento se argumentó matemáticamente y fue probado mediante una herramienta computacional robusta como MATLAB. Frente a la eficacia de las redes neuronales se obtuvieron mejores resultados con el MLP que con la red ART2; este resultado se soporta sobre el hecho de que el perceptrón multicapa realiza un proceso más eficaz de entrenamiento en el que se debe tener una base de datos de fotografías de los sujetos amplia para disminuir el error en el entrenamiento; además, la composición de la red perceptrón puede cambiar con base en el número de sujetos y el número de imágenes que se tienen y, si se requiere ingresar un nuevo sujeto, es necesario hacer de nuevo todo el proceso de entrenamiento y calibración.

Mientras que con respecto al uso de la red AR podemos concluir que depende muy singularmente del parámetro de vigilancia establecido en su proceso de selección; si el vector de entrada establece una similitud enmarcada dentro del parámetro de vigilancia (tolerancia al error) reconocerá efectivamente el rostro de la persona, y a su vez recalculará los pesos de la red neuronal sin necesidad de una etapa previa de entrenamiento. Esta red facilita el proceso de creación de nuevos patrones sin necesidad de entablar etapas previas de entrenamiento.

REFERENCIAS

Aguerreberre, Cecilia, Capdehourat, Germán, Delbracio, Mauricio, y Mateu, Matías (2005). *Detección*

de caras en Imágenes a Color. Uruguay: Universidad de la República.

Benito, Darío de Miguel (2005). *Detección automática del color de la piel en imágenes bidimensionales basado en el análisis de regiones*. España: Universidad Rey Juan Carlos.

Cabello Pardos, Enrique (2004). *Técnicas de reconocimiento facial mediante redes neuronales*. Tesis doctoral. Archivo electrónico. España: Universidad Politécnica de Madrid.

Gámez Jiménez, Carmen Virginia (2009). *Diseño y desarrollo de un sistema de reconocimiento de caras*. España: Universidad Carlos III de Madrid.

García García, Pedro Pablo (2013). *Reconocimiento de imágenes utilizando redes neuronales artificiales*. Tesis, trabajo fin de máster. España: Universidad Complutense de Madrid.

Hsu, R.L., Abdel-Mottaleb, M., & Jain, A. K. (2002). *Face detection in color images*. *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, n.º 5, pp. 696-706.

Mejía Sánchez, Juan Arturo (2004). *Sistema de detección de intrusos en redes de comunicaciones utilizando redes neuronales*. Mexico: Universidad de las Americas Puebla.

Palacios, Sandra María (2010). *Sistema de reconocimiento de rostros*. Universidad Peruana de Ciencias Aplicadas. Archivo electrónico.

Sesmero Lorente, M. Paz (2012). *Diseño, análisis y evaluación de conjuntos de clasificadores basados en redes de neuronas*. España: Universidad Carlos III de Madrid.

Terrillon, J.C., Shirazi, M. N., Fukamachi, H., & Akamatsu, S. (2000). Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. *Fourth IEEE International Conference on Automatic Face and Gesture Recognition Proceedings*, pp. 54-61.

