Methodology for the Construction of a Biometric Facial Recognition System Based on a Neural Classifier

M. Maglianesi and G. Stegmayer, *Member, IEEE*

***Abstract*— This work focuses on the development of a methodology that can be useful in the construction and selection of classifiers based on artificial neural networks biometric systems. The general hypothesis is that it is possible to find an acceptable neural model as a classifier for face recognition (FR) according to data provided during training, according to automated procedures. This could be very useful to assist in the construction process of FR systems, such as for example, it could facilitate the construction of a FR system optimized for a client database at a bank or a government organization**

***Keywords*— Face recognition, neural network classifier, eigenfaces.**

I. Introduccion

LOS SISTEMAS biométricos consisten en la aplicación de técnicas matemáticas y estadísticas sobre los rasgos anatómicos o de conducta de una persona con propósitos de verificación de identidad o identificación. Los sistemas biométricos se basan en el reconocimiento de patrones [1]. Un patrón biométrico para reconocimiento facial (RF) está compuesto por un conjunto de características físicas que definen un vector asociado unívocamente a una persona. Utilizando esta información, es posible verificar e identificar un vector de características dentro de una base de datos con información de otras personas [2].

Un sistema de RF consta básicamente de tres bloques [3]: Procesamiento de las imágenes, Extracción de características y Clasificación (ver Fig. 1)). En el primer paso se realiza, sobre la imagen de entrada, el pre-procesamiento necesario para resaltar los detalles de interés (por ejemplo, reducción de ruido introducido por el dispositivo de captura). En la segunda etapa se extraen las características a utilizar en el reconocimiento, obteniendo un vector patrón para cada imagen, lo cual permite reducir la cantidad de información necesaria para el almacenamiento del rostro. Por último se efectúa la clasificación, comparando el nuevo patrón con los patrones de usuarios registrados almacenados en una la base de rostros autorizados.

Independientemente de la arquitectura con la que se encuentre desarrollado un sistema biométrico, estas tres etapas se ejecutarán en dos momentos de tiempo diferentes. En un primer momento, para introducir un vector de características al sistema para su posterior utilización. Si el sistema se encuentra desarrollado con técnicas de inteligencia computacional (como por ejemplo, redes neuronales (RNs)), se habla de entrenamiento del modelo neuronal y la base de datos de características que utilizará el sistema biométrico representará un estado interno del modelo.

En el momento de uso del sistema, se solicita al sistema biométrico la búsqueda de un patrón que coincida con el presentado. Se entiende que el sistema biométrico ya conoce un conjunto de individuos y se le presenta una o varias imágenes de un individuo para repetir nuevamente el proceso, con el propósito de encontrar un individuo cuyo patrón de características sea lo suficientemente cercano al patrón de características obtenido en tiempo real. Debido al hecho de que un sensor biométrico no capturará exactamente los mismos datos dos veces, el matching (o emparejamiento) de características biométricas es una tarea compleja desde el punto de vista computacional. Recientemente, enfoques basados en RNs han demostrado ser adecuados para este tipo de problemas [4]. En general, el RF puede ser tratado como un problema de reconocimiento de patrones [5], problema que puede resolverse típicamente con modelos basados en neuronas artificiales de tipo perceptrón multicapa (MLP) [6]. Incluso una arquitectura en cascada o arreglo de clasificadores neuronales puede alcanzar mayor precisión en la tarea de RF [7]. Un modelo MLP puede entrenarse con el algoritmo de retroprogacación de errores estándar para clasificar un vector de características extraída de un rostro [8] [9].

En este trabajo se presenta una propuesta de metodología de construcción de un sistema de RF basado en un clasificador neuronal para las imágenes. La metodología se orienta a dar soporte al proceso de construcción y selección de un modelo clasificador neuronal, por lo cual la propuesta se centra en la última etapa de un sistema de RF. En la siguiente sección se presentan los conceptos básicos relacionados a un sistema facial. La sección 3 explica métricas de calidad de éstos sistemas. La sección 4 muestra en detalle la metodología propuesta. La sección 5 presenta la aplicación de la metodología a un caso de estudio con una base de rostros de acceso público. Finalmente, las conclusiones forman la sección 6.

II. Sistema de RF

En el primer paso de un sistema de RF se realizan, sobre la imagen de entrada, los procesos de procesamiento de imagen necesarios para resaltar los detalles de interés. En una segunda etapa se extraen las características a utilizar en el reconocimiento, formando uno o varios vectores patrón. Esta etapa consiste en obtener un patrón único de características de la imagen asociada a una persona. Posteriormente dicho patrón de característica será almacenado en una base de datos de características agrupadas por individuo.

M. Maglianesi , UTN-FRSF, Argentina, maglianesi@gmail.com
G. Stegmayer, Center for Research & Development of Information Systems (CIDISI), CONICET, Argentina, georgina.stegmayer@ieee.org

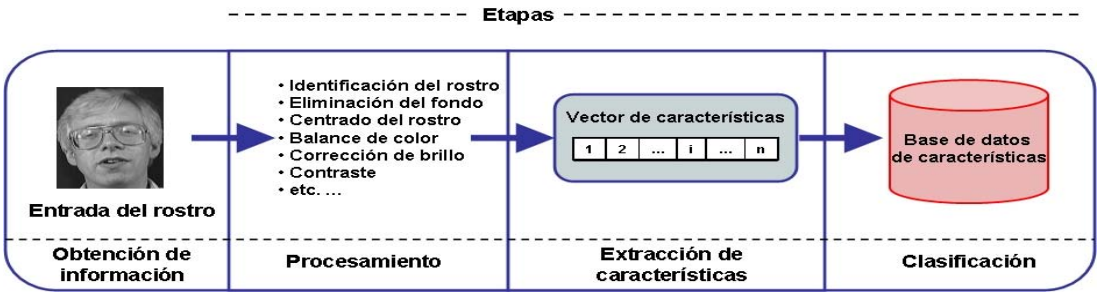

Figura 1. Proceso de Identificación de un Sistema RF.

Existen dos enfoques predominantes en el tratamiento del problema de la extracción de características. El enfoque geométrico (o basado en rasgos) se centra en las relaciones existentes entre los diversos componentes de un rostro, como por ejemplo la distancia entre las dos pupilas, entre las cejas, etc. El enfoque fotométrico no requiere la identificación de los diversos elementos de una cara dado que toma el rostro completo para aplicar una serie de transformaciones cuyo objetivo es reducir su dimensionalidad y al mismo tiempo obtener un vector de características únicas para el sujeto.

El análisis de las características faciales se realiza por medio de operaciones matemáticas recursivas para encontrar los componentes únicos o propios de la muestra, concepto que a veces se refiere con el término alemán eigen. Así la imagen se descompone en un conjunto de áreas de luz y sombras dentro de un determinado patrón. El resultado es entonces que una cara resulta ser la combinación de dichas áreas únicas de esa cara [10]. Éste es uno de los enfoques más ampliamente utilizado para el problema de extracción de características de una imagen. Este método reduce la dimensionalidad de las imágenes de entrada y conserva aquellas que tengan la mayor cantidad de información. Los vectores seleccionados definen un espacio de proyección conocido como eigenspace sobre las imágenes de entrenamiento. La representación de una imagen en el eigenspace es llamada eigenface. Se conoce comúnmente como métodos de las eigenfaces al trabajo de Turk y Pentland [11], que utiliza los eigenvectores principales de la matriz de covarianza de una imagen para su representación. Este enfoques extrae características en un subespacio derivado de las imágenes de entrenamiento. Las eigenfaces describen la forma de toda una cara en lugar de las estructuras locales como la nariz, ojos, línea de mandíbula y pómulos.

La etapa de clasificación consiste en la búsqueda de un patrón clasificado y almacenado previamente en una base de datos que se corresponda con el patrón de la persona que está intentando autenticar contra un sistema biométrico. La Fig. 2 muestra el esquema genérico de un clasificador dentro de un sistema de RF, al cual se le introduce una imagen almacenada en una BD de Características faciales durante la etapa de entrenamiento y una imagen utilizada durante la fase de test o validación. La algorítmica del clasificador determinará si las dos imágenes que se le presentaron corresponde o no a la misma persona. El objetivo del clasificador es diseñar una función que clasifique una instancia dentro de muchas clases predefinidas [12]. En general, cuando se requiere identificar un rostro, se transforma la imagen del mismo al espacio de proyección y se evalúan las diferencias entre los pesos de entrada y los pesos pertenecientes a las imágenes del entrenamiento. Si la diferencia entre estos es menor a cierto umbral, el rostro es identificado.

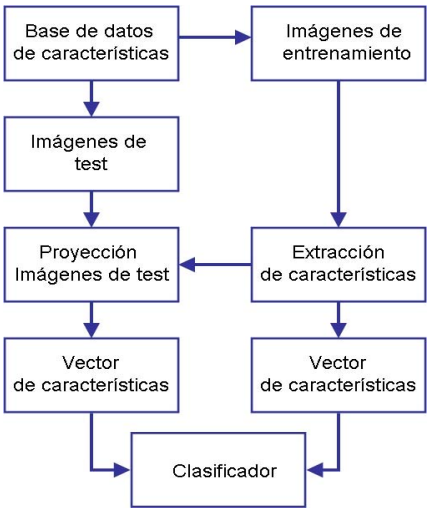

Figura 2. Clasificador en un sistema de RF.

Persiguiendo un incremento en el rendimiento en el proceso de clasificación, algunos autores proponen combinar varios clasificadores para inducir un clasificador de nivel más alto [13][7]. Independientemente de que técnica se haya empleado para la construcción del clasificador, el proceso de identificación es básicamente el mismo. Cuando se requiere identificar un rostro se transforma la imagen del mismo al espacio de proyección generado durante el entrenamiento del sistema y se evalúan las diferencias entre los pesos de entrada y los pesos pertenecientes a las imágenes del entrenamiento (ver Fig. 2). Si la diferencia entre estos es menor a cierto umbral, el

rostro es identificado [2]. Con independencia si la decisión del sistema biométrico se encuentra embebido en el clasificador o resulta en un módulo distinto, por lo general la algorítmica del mismo recae en la distancia euclidea existente entre la imagen de entrada y la más cercana en el sub espacio. Dicha distancia se transforma en el umbral del sistema biométrico, límite entre aceptación o rechazo de un individuo.

III. Métricas de calidad.

Se definen las siguientes cualidades ideales en un sistema biométrico:

Robustez: La robustez significa que la características de los individuos que son utilizadas por un sistema biométrico no cambian significativamente a través del tiempo. La robustez de un sistema biométrico es medida por lo que se conoce como tasa de "falso rechazo" o FRR (False Rejection Rate) e indica la probabilidad de que un dispositivo rechace una persona autorizada. El FRR se define como:

$$FRR = \frac{Cant.AutorizadosRechazados}{Cant.TotalAccesosAutorizados} \times 100\,\%. \quad (1)$$

En general, mientras más bajo son los valores del FRR, más alta es la precisión del sistema biométrico.

Carácter distintivo: El carácter distintivo muestra una gran variación sobre la población. Se mide por la tasa de falsa aceptación o FAR (False Acceptance Rate). El FAR, también conocido como "error de tipo II", indica la probabilidad de que una muestra –conocida o desconocida por el sistema biométrico– sea erróneamente identificada como perteneciente a otra persona. Esta tasa muestra el porcentaje de número de veces que el sistema produce una falsa aceptación. Es decir cuando un individuo es identificado como usuario de manera incorrecta. Este valor debe ser lo suficientemente bajo como para que no se impida el ingreso a los usuarios, pero no tanto que permita el ingreso de personal no autorizado. El FAR se define como:

$$FAR = \frac{Cant.\ ImpostoresAceptados}{Cant.TotalAccesosImpostores} \times 100\,\%. \quad (2)$$

En general, mientras más bajo son los valores del FAR, más alta es la precisión del sistema biométrico.

Rendimiento general de un sistema de RF: En la actualidad no existe ningún sistema biométrico que maximice las características mencionadas anteriormente de forma tal que reconozca a todos los individuos en un 100% (FRR), rechace a todos los individuos desconocidos por el sistema (FAR) y a la vez sea de alta disponibilidad, accesible y aceptable. Por lo que en el proceso de construcción de un sistema biométrico se ponderan estas características priorizando algunas de ellas sobre las otras.

Mientras que el FRR y el FAR representa medidas básicas de error de un sistema biométrico, existen otras métricas que las combinan para obtener una medida general del rendimiento del sistema. La más relevante es la "tasa global de éxito" o TSR (Total Success Rate):

$$TSR = 100\,\% - \left(1 - \frac{FAR + FRR}{CantidadTotalAccesos}\right). \quad (3)$$

IV. Metodología propuesta.

Un sistema biométrico puede ser diseñado con dos objetivos excluyentes entre si: i) Que las muestras presentadas son de un individuo conocido por el sistema; ii) Que las muestras presentadas son de un individuo no conocido por el sistema. Los sistemas biométricos construidos en torno al primer objetivo se los suele llamar sistemas de "Identificación Positiva" mientras que aquellos construidos en torno al segundo objetivo se los llama sistemas de "Identificación Negativa". Todos los sistemas biométricos son de un tipo o de otro. Los sistemas de identificación positivo sirven generalmente para prevenir coincidencia de múltiples personas en una sola identidad mientras que los sistemas de identificación negativos sirven para prevenir coincidencia de múltiples identidades de una sola persona.

En este trabajo se propone una metodología para la construcción de un sistema de reconocimiento facial, el cual use como clasificador un modelo neuronal. La metodología está enfocada en dar soporte al proceso de selección de un modelo clasificador adecuado, por lo cual no se tiene en cuenta la etapa de preprocesamiento de las imágenes, sino que se presupone que ésta ya fue realizada y se cuenta con una base de datos (BD) de rostros de sujetos autorizados para ingresar a un sistema. Los pasos de la metodología propuesta para construcción de un sistema de RF basado en un clasificador neuronal son los siguientes:

A. Preparación de conjuntos de entrenamiento y prueba del clasificador

En esta etapa se seleccionan *k*-particiones de la BD de rostros para poder aplicar la técnica de validación cruzada *(k*-fold cross-validation). Ésta es una técnica del área de machine learning para evaluar qué tan bien generaliza un modelo predictivo sobre un conjunto de datos independientes. Una vuelta de validación cruzada involucra la partición de un conjunto de datos en subconjuntos complementarios, realizando el análisis (en este caso, el entrenamiento del clasificador) sobre un subconjunto denominado de entrenamiento, y validándolo sobre el otro subconjunto (denominado de validación o prueba). Para reducir la variabilidad, se realizan n rondas de validación cruzada usando *k* diferentes particiones, y los resultados finales de validación son calculados como el promedio sobre todas las rondas de prueba [6].

Las imágenes pueden variar en condiciones de iluminación, contraste, ángulo del rostro, etc.. En la práctica es habitual que el subconjunto de entrenamiento involucre la mayoría de las imágenes de los individuos y el subconjunto de test o validación un número menor de muestras de los mismos individuos utilizados durante la fase de entrenamiento. Esto con el propósito de presentarle al clasificador en la fase de test una muestra "diferente" de

un individuo "conocido". Por lo que la partición que se realiza en cada ronda de validación cruzada i debe considerar la división de los subconjuntos $train_i$ y $test_i$ de forma tal que ambos involucren los mismos individuos aunque el número de muestras en $train_i$ será mayor que el contenido en $test_i$, para $i=[1..k]$.

La medida de rendimiento que se obtiene es la tasa de falso rechazo o FRR definida en la sección 3. El FRR es calculado por cada individuo presentado durante el test y luego se promedian los mismos para obtener un único índice de falso rechazo para la partición *i*. El rendimiento general del clasificador con relación al falso rechazo es calculado a partir del promedio de cada uno de los FRR obtenidos en cada ronda de validación. También es práctica habitual en la construcción de un clasificador neuronal separar algunas muestras completas de individuos para que no formen parte de los subconjuntos train y test. Estas muestras son presentadas al clasificador en la etapa de test como ''intrusos" al sistema y se evalúa el comportamiento del clasificador. En este proceso se obtiene como medida de rendimiento la tasa de falsa aceptación o FAR descripta en la sección 3. El FAR es calculado por cada individuo presentado durante el test y luego se promedian los mismos para obtener un único índice de falsa aceptación para la partición i. El rendimiento general del clasificador con relación a la aceptación de intrusos es calculado a partir del promedio de cada uno de los FAR obtenidos en cada ronda de validación.

B. Extracción de características con eigenfaces

La presente metodología asume que el sistema biométrico que se está diseñando utilizará el método de las Eigenfaces. El método PCA permite reducir la dimensionalidad de las imágenes de entrada conservando aquellas que tengan la mayor cantidad de información. Para ello, se vale de la matriz de covarianza obtenida a partir de cada imagen de ejemplo para construir un espacio de características denominado Eigenspace.

En este paso, sobre los conjuntos de entrenamiento resultantes de la etapa anterior, se calculan las eigenfaces correspondientes al 75%, 80% y 85% de varianza total de las imágenes a ser representada. Estos porcentajes corresponden a lo usualmente seleccionado para este tipo de sistemas [4] y la cantidad de eigenfaces que se obtienen depende del tamaño de la BD de rostros sobre la cual se aplica el método. Una mayor cantidad de eigenfaces implicará seguramente mejores resultados de clasificación pero tendrá un costo de almacenamiento asociado más alto. De todos modos, en comparación al almacenamiento de las imágenes originales, se reduce drásticamente la cantidad de información a almacenar dado que finalmente se guardan sólo las eigenfaces del espacio de representación de características, en lugar de las imágenes originales.

El análisis de componentes principales o transformación Karhunen-Loëve es una técnica estándar usada en reconocimiento estadístico de patrones para reducción de dimensionalidad. Ésta aplica una transformación lineal ortogonal que decorrelaciona variables, reteniendo las que más contribuyen a la máxima varianza contenida en el patrón (las componentes principales) y descartando el resto [14]. Una imagen en escala de grises de tamaño W en ancho y H en alto puede ser considerada también como un vector de una dimensión $N=W\times H$. Considerando un grupo de imágenes con la misma configuración, cada uno de ellos corresponde a un punto en un espacio N-dimensional. De esta manera, si nos ocupamos de imágenes similares de caras, éstas se ubicarán en una pequeña región del espacio. Aquí, el objetivo del PCA es seleccionar un subespacio de menor dimensión que mejor representa las imágenes originales. Como veremos, los nuevos vectores están dadas por los vectores propios de la matriz de covarianza correspondientes a las imágenes originales (eigenfaces).

Para calcular las eigenfaces, primero se necesita un conjunto M de imágenes Γ_1, Γ_2,..., Γ_m, idealmente de un conjunto de sujetos conocidos en la BD. Cada Γ_i es un vector N-dimensional conteniendo los N pixels de cada i-ésima imagen. El objetivo es calcular los eigenvectores u_i de la matriz de covarianza Γ_i. Definiendo a la imagen media como

$$\vec{\Psi} = \frac{1}{M}\sum_{i=1}^{M}\vec{\Gamma}_i \tag{4}$$

y la diferencia entre cada imagen y la imagen media como $\vec{\Phi}_i = \vec{\Gamma}_i - \vec{\Psi}$, la matriz de covarianza está dada por

$$C = \frac{1}{M}\sum_{i=1}^{M}\vec{\Phi}_i\vec{\Phi}_i^T \tag{5}$$

resultando en una matriz de tamaño $N\times N$, aún para imágenes de tamaño pequeño.

El cálculo de los eigenvectores de (5) es una tarea computacionalmente intensiva. El método descripto en [11] es el más comúnmente usado, el cual define

$$A = \left[\vec{\Phi}_1\vec{\Phi}_2\cdots\vec{\Phi}_M\right] \tag{6}$$

como una matriz de tamaño $N\times M$ (N filas, una por cada pixel de cada imagen, y M columnas, una por cada sujeto).

De este modo, la matriz de covarianza AA^T tiene tamaño $N\times N$ pero A^TA es una matriz de tamaño $M\times M$. Luego, definiendo

$$L = A^TA \tag{7}$$

donde

$$L_{ij} = \vec{\Phi}_i^T\vec{\Phi}_j \tag{8}$$

y calculando los M eigenvectores $\vec{v_i}$ de L, las eigenfaces se calculan como

$$\vec{u}_i = \sum_{i=1}^{M}\vec{v}_{ij}\vec{\Phi}_j, \qquad i = 1,\ldots,M. \tag{9}$$

Además, los autovectores pueden ser rankeados a través de su autovalor asociado, y las primeras M' más altas eigenfaces pueden ser seleccionadas de acuerdo a la proporción deseada de varianza total de las imágenes a representar. Generalmente, el número de imágenes de entrenamiento será menor que la cantidad de pixels en las imágenes ($M<<N$), tal que el número de operaciones se reduce significativamente.

Cualquier imagen de entrada puede ser proyectada en un espacio M'-dimensional (el espacio de las caras) haciendo

$$\vec{\Omega} = \vec{u}^T(\vec{\Gamma} - \vec{\Psi}). \tag{10}$$

De este modo, $\vec{\Omega}$ forma el vector de características M'−dimensional que representa la imagen y que puede ser

usado como entrada de un clasificador con tamaño de entrada fija, sin importar el número de pixels *N* de la imagen original. Si bien PCA permite reducir en gran medida el espacio de características, la aplicación de una técnica denominada Scree Test [15] permite reducir aun más la dimensionalidad del espacio centrándose únicamente en aquellas eigenfaces más representativas. Scree Test se aplica para determinar el número de componentes principales (c) de la matriz de covarianza de train o test según el porcentaje de variabilidad de los datos que se desea representar [15]. El término Scree se define como analogía a los escombros en el fondo de un acantilado, es decir, las componentes principales usadas son el acantilado y el resto los escombros.

La aplicación de esta técnica permite obtener en forma ordenada las distintas componentes en función a la variabilidad de los datos que contiene. Por lo que la primera componente principal representará el vector de datos de mayor variabilidad. Los vectores de datos son obtenidos a través de diversas operaciones que se aplican sobre las matrices de train y test según corresponda. Independientemente de los datos que se encuentran modelados con la técnica PCA, es habitual que valores cercanos al 80% de la varianza total del espacio de características sea suficiente para su identificación. Una mayor cantidad de eigenfaces implicará seguramente mejores resultados de clasificación pero tendrá un costo de almacenamiento asociado más alto y mayores tiempos de cómputo en todos los procesos involucrados (entrenamiento, prueba y producción). Por lo que en este paso se pretende explorar el número de características mínimas que permitan obtener un desempeño aceptable del clasificador conforme a los objetivos de diseño establecidos.

C. Construcción del clasificador

Se construyen y entrenan dos posibles arquitecturas para el clasificador neuronal. Se han elegido las topologías más usuales (según la literatura actual) para este tipo de sistema [16] [17] [8]. La arquitectura I) consiste en un perceptrón multicapa (MLP), con una neurona de salida para cada sujeto (s) a validar en el sistema (ver Fig. 3). La arquitectura II) es un arreglo de s perceptrones multicapa (s-MLPs), una para cada sujeto s, en el cual cada modelo posee una única neurona de salida [7] (ver Fig. 4).

La capa de entrada de cada clasificador neuronal MLP es un conjunto E de neuronas de entrada, donde cada entrada es un punto del espacio de eigenfaces (obtenido en la etapa de extracción de características del sistema), y hay O neuronas en la capa oculta de cada modelo. Cada neurona de salida del modelo clasificador debe tomar un valor de 1 si reconoce al sujeto presentado como entrada del modelo, o 0 en caso contrario. Una vez entrenado el clasificador con el método de validación cruzada, cuando una imagen tiene que ser clasificada, es proyectada en el espacio de las eigenfaces y usada como entrada del clasificador (ecuación 10). Esto implica presentarla al modelo MLP y a las s redes neuronales que conforman el modelo s-MLPs. En ambos casos, a la salida máxima obtenida se le asigna una etiqueta de clase.

Por ejemplo, si se presenta al sistema la imagen del sujeto i (i∈[1□s]), se espera un valor de salida cercano a 1 en la neurona i del modelo MLP y un valor cercano a 1 en la red neuronal i del modelo s-MLPs. Para ambas arquitecturas, los valores de sus parámetros son inicializados en forma aleatoria entre [-1,+1] al inicio de la etapa de entrenamiento y se utiliza el algoritmo estándar de retropropagación de errores, en la variante de Levenberg-Marquardt [6]. En el caso del arreglo de MLP, se entrena cada modelo neuronal con las imágenes que corresponden a su clase y todas las otras imágenes de los sujetos que no corresponden a su clase.

Una vez establecidos los criterios de partición de datos que se utilizarán para el entrenamiento y validación, transformadas las imágenes a vectores, conformando los conjuntos train y test y determinada la cantidad de componentes principales a utilizar, se construyen uno o varios clasificadores neuronales con distintas topologías y se seleccionan las variables sobre las que operarán cada uno de ellos. Luego, se entrena cada modelo por cada valor que pueda asumir cada una de las variables del mismo. Una vez entrenado cada uno de los clasificadores posibles, cuando una imagen Γ perteneciente a un conjunto k tiene que ser clasificada, es proyectada en el espacio de las eigenfaces y usada como entrada del clasificador. Esto implica presentarla al modelo y evaluar la salida que produce el mismo. Para cada clasificador que se entrena se almacenan todas sus variables asociadas y se obtienen múltiples medidas de rendimiento que también son almacenadas para su análisis posterior.

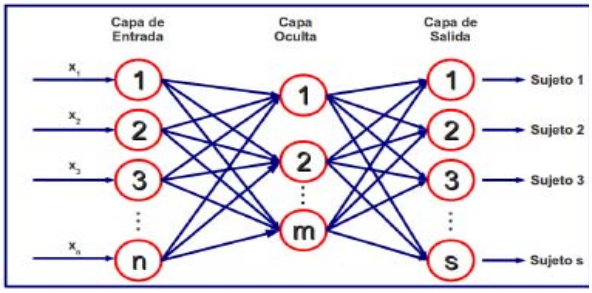

Figura 3. Topología I - Perceptrón multicapa (MLP).

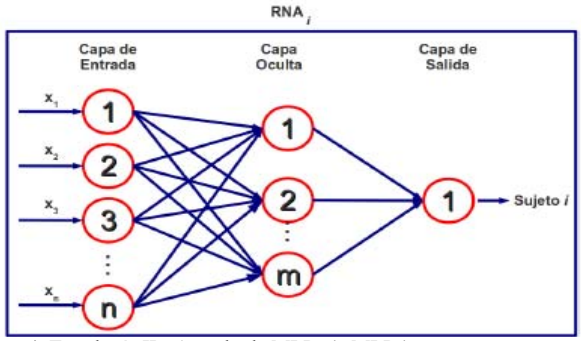

Figura 4. Topología II - Arreglo de MLPs (s-MLPs).

La determinación de la cantidad de neuronas ocultas representa una de las variables sobre las que se basa la metodología propuesta pudiendo adoptar la forma de un

rango discreto de valores sobre los cuales probar y evaluar el rendimiento del clasificador resultante. La interpretación final de la salida del clasificador depende de su topología. En el caso de la topología I se analiza la neurona si con mayor valor de activación como la salida relevante del clasificador y como posible identificación del sujeto. En el caso de la topología II se analiza la red con mayor valor de activación como relevante, representando una posible identificación del sujeto.

Independientemente de la topología empleada, el clasificador devuelve inicialmente un sujeto si como candidato de identificación a través de un valor que representa el máximo valor de activación del clasificador.

Esto es así incluso cuando se le es presentado un intruso al sistema. Esto hace necesario aplicar a la salida del clasificador un valor de umbral que limite la aceptación de personas no autorizadas. El umbral consiste en establecer un valor mínimo para que la máxima salida entregada por el clasificador sea identificada como el sujeto si. Si la máxima salida no supera el valor de umbral establecido el clasificador rechazará al sujeto presentado, sea este una persona autorizada o un intruso del sistema. Asimismo, si el valor de umbral es muy bajo el clasificador será proclive a aceptar intrusos y si el umbral es muy alto a rechazar personas autorizadas. El valor umbral a aplicar debe ser evaluado por cada clasificador y considerando los objetivos de diseño del sistema de RF.

C.1. Cálculo del FRR

La TABLA I muestra la información inherente al proceso de clasificación de 3 sujetos (s), 2 muestras (m) por cada uno de ellos para el clasificador descripto en la topología I (MLP) diseñado para 12 sujetos. El ejemplo de clasificación se produce en la etapa de test del clasificador en la cual se conoce a-priori a que sujeto corresponde la imagen suministrada. Las columnas de datos representan los valores máximos de activación de cada una de las s neuronas de la capa de salida una vez introducida la imagen al clasificador.

TABLA I
EJEMPLO DE CLASIFICACIÓN DE 3 SUJETOS, 2 MUESTRAS POR SUJETOS EN EL CLASIFICADOR DE LA TOPOLOGÍA I DISEÑADO PARA 12 SUJETOS

s	s_1		s_2		s_3	
	m_1	m_2	m_1	m_2	m_1	m_2
1	0.4641	0.0009	0.0021	0.0000	0.0010	0.0004
2	0.0000	0.0020	0.1520	0.6392	0.0000	0.0000
3	0.0000	0.0000	0.0000	0.0000	0.4497	0.7280
...	...	...	...	...	...	...
12	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Cada fila representa un sujeto. Para facilitar la lectura denotaremos $s_{(s,m)}$ donde el primer subíndice corresponde al número de sujeto y el segundo al número de muestra. Para cada muestra presentada el clasificador el proceso de identificación consiste en evaluar la neurona con mayor valor de activación (resaltada en la tabla) y cotejar si la misma se corresponde con la del sujeto con el cual se está probando. Para el ejemplo ilustrado en la tabla de referencia, las muestras $s_{(1,1)}$, $s_{(2,1)}$, $s_{(2,2)}$, $s_{(3,1)}$, $s_{(3,2)}$, $s_{(4,1)}$, $s_{(4,2)}$ y $s_{(5,1)}$ son clasificadas correctamente. La muestra $s_{(1,2)}$ es clasificada erróneamente como perteneciente a s_2 y la muestra $s_{(5,2)}$ como perteneciente al sujeto $s_{(1,1)}$. Si aplicamos al ejemplo citado la ecuación 1, el cálculo del FRR se aplicaría como: $FRR= 2/10 \times 100\%=20\%$, lo cual expresa que el 20% de los sujetos presentados al clasificador son erróneamente clasificados y por lo tanto rechazados por el mismo. Como se dijo en su oportunidad el valor óptimo para este índice es 0.

C.2. Cálculo del FAR

La TABLA II ejemplifica la presentación de 4 intrusos a un clasificador de tipo MLP diseñado para 12 sujetos. Las columnas de datos representan los valores de cada una de las s neuronas de la capa de salida una vez introducida la imagen del intruso al clasificador. Cada fila representa un sujeto tal y cada columna a un intruso.

Para cada muestra presentada al clasificador el proceso de validación de intrusos consiste en evaluar la neurona con mayor valor de activación (resaltada en la tabla) y cotejar si la misma supera cierto umbral de activación. En el ejemplo de referencia, si no se aplica un valor umbral, el clasificador aceptará erróneamente los 4 intrusos presentados como si se trataran de los sujetos s_8, s_5, s_{11} y s_2. Si aplicamos al ejemplo el cálculo del FAR (ecuación 2), se obtendría $FAR= 4/4 \times 100\%=100\%$, lo cual expresa que el 100% de los intrusos presentados al clasificador son erróneamente aceptados por el mismo. Como se dijo en su oportunidad el valor óptimo para este índice es 0.

TABLA II
EJEMPLO DE PRESENTACIÓN DE 4 INTRUSOS AL CLASIFICADOR.

s	I_1	I_2	I_3	I_4
1	0.0436836	0.1771888	0.0008657	-0.0507513
2	-0.0204900	0.0528491	0.0091070	**0.1674717**
3	-0.0223433	0.0590199	0.0566782	-0.0563901
4	0.1182653	-0.0362740	0.2311245	0.0203761
5	-0.0451147	**0.3217476**	-0.1480765	0.0465514
6	0.0042542	-0.0659941	0.0317070	-0.0292746
7	-0.0490030	0.0453813	-0.0449852	-0.0901531
8	**0.349627**	0.0095236	0.0672050	-0.0394631
9	0.1338885	0.0355615	0.0667065	0.0685316
10	-0.0719714	0.0574900	0.0958392	0.0482997
11	0.0554922	-0.0096233	**0.7490629**	-0.0674618
12	0.0030577	0.0985803	0.1530010	0.1052294

En general, el clasificador no rechazará la totalidad de los intrusos sin variar el umbral de activación de cada neurona de salida. Considerando que uno de los objetivos de diseño del sistema de reconocimiento facial podría ser justamente el rechazo total de intrusos, esta metodología propone para cada clasificador calcular los índices de rendimiento repetidas veces variando incrementalmente los umbrales de activación de cada neurona de salida a fin de hallar el mínimo valor de umbral para un rechazo total. La TABLA III ejemplifica el proceso de exploración de umbrales hasta alcanzar un rechazo total de intrusos.

TABLA III
EXPLORACIÓN DE UMBRALES PARA ALCANZAR UN RECHAZO TOTAL

Umbral	*Aceptados*	*Intrusos Totales*	*FAR*
0,0	4	4	100 %
0,1	4	4	100 %
0,2	3	4	75 %
0,3	3	4	75 %
0,4	1	4	25 %
0,5	1	4	25 %
0,6	1	4	25 %
0,7	1	4	25 %
0,8	0	4	0 %
0,9	0	4	0 %
1,0	0	4	0 %

Aplicando un valor de umbral de 0.1, el ejemplo representado en la TABLA II sigue aceptando a todos los intrusos dado que los valores de activación son todos superiores a 0.1. Con valores de umbral de 0.2 o 0.3, I_4 será correctamente catalogado como intruso ya que el máximo valor de activación que produce es de 0.1675. Para valores de umbral de 0.4 a 0.7, los intrusos se corresponden con I_1, I_2 *e* I_4. Al aplicar valores de umbral de 0.8 en adelante todos los intrusos son rechazados.

Por cada valor umbral que se explora se debe volver a calcular el FRR correspondiente del clasificador. El 20% obtenido como índice de falso rechazo a partir de los datos ilustrados en la TABLA I ha sido obtenido sin aplicar ningún valor de umbral. La TABLA IV muestra como aumenta el FRR a medida que se incrementan los umbrales.

TABLA IV
AUMENTO DEL FRR A MEDIDA QUE SE EXPLORAN LOS UMBRALES

Umbral	*Rechazados*	*Sujetos Presentados*	*FRR*
0,0	2	10	20 %
0,1	2	10	20 %
0,2	3	10	30 %
0,3	4	10	40 %
0,4	4	10	40 %
0,5	7	10	70 %
0,6	7	10	70 %
0,7	8	10	80 %
0,8	10	10	100 %
0,9	10	10	100 %
1,0	10	10	100 %

Con un valor de umbral de 0.2 la tasa de falso rechazo se ve afectada dado que $s_{(2,1)}$ es inicialmente clasificado correctamente pero su máximo valor de activación es de 0.1520. Similarmente ocurre a medida que se incrementan a umbrales mayores y sujetos que inicialmente fueron correctamente clasificados comienzan a ser rechazados. En el ejemplo de referencia cuando el umbral alcanza el valor 0.8 rechazará la totalidad de todas las muestras y el valor de FRR sera 100%.

C.3. Cálculo del TSR

Al igual que el FRR y el FAR, el mejor rendimiento del clasificador se logra con TSR=0 (ecuación 3). La TABLA V muestra el valor que adquiere con los diferentes umbrales el ejemplo de las subsecciones anteriores.

TABLA V
TSR RESULTANTE A MEDIDA QUE SE EXPLORAN LOS UMBRALES

Umbral	*TSR*
0,0	42.86 %
0,1	42.86 %
0,2	42.86 %
0,3	50.00 %
0,4	35.71 %
0,5	57.14 %
0,6	57.14 %
0,7	64.29 %
0,8	71.43 %
0,9	71.43 %
1,0	71.43 %

C.4. Selección del clasificador neuronal

Inicialmente, y como paso previo al proceso de selección, se deben promediar los resultados obtenidos de cada uno de los indicadores obtenidos en las diferentes rondas de validación cruzada a fin de disponer de un único valor por cada indicador y para cada clasificador. Una vez creados y entrenados un conjunto de clasificadores neuronales y calculados sus diferentes indicadores de rendimiento, se procede a un proceso de selección conforme a los objetivos de diseño que persiga el sistema de reconocimiento facial. Se proponen tres alternativas para seleccionar el mejor modelo clasificador:

a Priorizando el Reconocimiento

(a) Por cada modelo, seleccionar el valor de umbral más alto que no modifique el FRR del clasificador sin aplicar umbrales. De esta forma se obtiene la mejor tasa de falso rechazo y se mejora todo lo posible la tasa de falsa aceptación.

(b) Realizar un ranking de los modelos ordenados por mejor TSR.

(c) Seleccionar el modelo que haya obtenido mejor puntuación en el ranking.

La TABLA VI muestra la información de uno de los modelos generados en el Caso de Estudio a partir del cual se puede apreciar que utilizar el valor umbral=0.2 no va en detrimento del FRR pero mejora la tasa de falsa aceptación.

b Priorizando el Rechazo Total de Intrusos

(a) Analizar la información de umbrales para las neuronas de salida asociada a cada topología de clasificador. Aplicar los umbrales hasta obtener una tasa de rechazo de intrusos del 100%.

(b) Realizar un ranking de los modelos ordenados por mejor FRR.

(c) Seleccionar el modelo que haya obtenido mejor puntuación en el ranking.

c Priorizando el mejor rendimiento

(a) Explorar los diferentes valores del TSR para cada clasificador aplicando diferentes umbrales a cada uno de ellos y seleccionar el de mejor rendimiento.

(b) Realizar un ranking de los modelos ordenados por mejor TSR.

(c) Seleccionar el modelo que haya obtenido mejor puntuación en el ranking.

TABLA VI
EJEMPLO DE SELECCIÓN DEL MEJOR FRR DE UN CLASIFICADOR OPTIMIZANDO A LA VEZ EL FAR DEL MISMO

Umbral	*FRR*	*FAR*	*TSR*
0,0	23.61 %	100 %	50.89 %
0,1	23.61 %	72.50 %	41.07 %
0,2	23.61 %	67.50 %	39.29 %
0,3	25.00 %	57.50 %	36.61 %
0,4	26.39 %	55.00 %	36.61 %
0,5	26.39 %	52.50 %	35.71 %
0,6	27.78 %	50.50 %	35.71 %
0,7	27.78 %	50.00 %	35.71 %
0,8	30.56 %	45.00 %	35.71 %
0,9	31.94 %	32.50 %	31.25 %
1,0	100.00 %	0.00 %	64.29 %

Observando los valores de rendimiento expuestos en la TABLA VI puede apreciarse cómo el TSR mejora gradualmente entre los umbrales 0 y 0.9. En el ejemplo expuesto, esto ocurre principalmente por la mejora que se produce en el indicador del FAR que resulta ser significativamente mayor que la degradación del índice del FRR. Una vez que se encuentran en el umbral 0.9 el FRR y el FAR se produce el mejor rendimiento del clasificador. Luego, con el último incremento del umbral, si bien se logra un rechazo total de intrusos (objetivo primordial del criterio b) de selección), la degradación del FRR es tan alta (rechaza a todos los sujetos autorizados) que provoca que el índice de rendimiento global desmejore considerablemente.

V. CASO DE ESTUDIO

Este capítulo ejemplifica cada uno de los pasos de la metodología propuesta para la construcción de un sistema de reconocimiento facial de rostros basado en redes neuronales. Para ello se seleccionó una base de datos de rostros de dominio público para entrenar y validar los clasificadores neuronales y se definieron 2 topologías diferentes de redes neuronales para la ejecución de las pruebas. A continuación se describe la base de datos de rostros utilizada y luego se describen paso a paso las tareas que se llevaron a cabo en el marco de la metodología propuesta.

A. Base de datos de rostros

Existen multitud de librerías de imágenes de caras humanas que se utilizan para testear el funcionamiento de los algoritmos de reconocimiento de rostros. En el presente trabajo se ha optado por utilizar la base de rostros del Olivetti Research Laboratory, más conocida como ORL, la cual está conformada por 40 personas, 10 imágenes por persona, conformando un total de 400 imágenes. Las fotografías fueron tomadas entre los años 1992 y 1994. Las tomas de algunas personas se realizaron en momentos diferentes variando condiciones de iluminación, las expresiones faciales (ojos abiertos/cerrados, sonriente y no sonriente) y los detalles faciales (con lentes y sin lentes). Todas las imágenes están tomadas contra un fondo oscuro homogéneo con las personas en posición vertical y frontal (con tolerancia de algunos movimientos laterales). La Fig. 5 muestra algunos de las personas contenidas en la base de datos . La resolución de las imágenes es de 92x112 pixeles y con una profundidad de color de 8 bits (escala de 256 grises por pixel).

Figura 5. Ejemplos de la BD ORL.

B. Aplicación de la metodología

Al conjunto de imágenes de la base de rostros ORL se le aplicó la técnica de validación cruzada "k-fold cross validation". Atento a la cantidad de sujetos diferentes y muestras por sujetos que conforma la base de rostros, se decidió emplear un valor de k=3, conformando así 3 rondas de validación para cada clasificador neuronal. De los 40 sujetos disponibles, 4 se reservaron para las pruebas de detección de intrusos. Los 36 sujetos restantes fueron utilizados para el proceso de entrenamiento y test. Considerando las 10 muestras por sujetos, el tamaño del conjunto de intrusos fue de 40 imágenes y el de entrenamiento y test de 360 imágenes. Sobre éste último conjunto, el 80% se utilizó para entrenar y el 20% para testear por lo que el tamaño de los conjuntos train fue de 288 imágenes y el tamaño de los conjuntos test de 72 imágenes. Sobre los conjuntos de entrenamiento resultantes de la etapa anterior, se calcularon las componentes principales correspondientes al 75%, 80% y 85% de varianza total de las imágenes a ser representadas. Estos porcentajes corresponden a lo usualmente seleccionado para este tipo de sistemas.

Una vez aplicado PCA + Scree Test y reducido de esta forma el espacio de características, se ejecutaron varios ciclos anidados que fueron variando de a una vez las diferentes variables a aplicar y por cada ciclo, se procedió a construir un clasificador neuronal por cada topología en estudio. La topología I) consiste en un perceptrón multicapa (MLP), con una neurona de salida para cada sujeto (s) a validar en el sistema. La topología II) es un arreglo de s perceptrones multicapa (s-MLPs), una para cada sujeto s, en el cual cada modelo posee una única neurona de salida.

Como paso previo al proceso de selección, se promediaron los distintos indicadores de rendimiento de las redes creadas en el paso anterior en función de las 3 rondas de validación cruzada que se realizaron. Por lo tanto se obtuvieron 162 modelos de los cuales 81 corresponden a la Topología 1 y otros tantos a la Topología 2. Luego, se siguieron las 3 alternativas propuestas por la metodologías. A continuación se describen los procedimientos y resultados obtenidos por cada una de ellas.

a Priorizando el Reconocimiento

La TABLA VII muestra los primeros 5 modelos obtenidos por cada topología luego de aplicar el ranking descendente. Los resultados se muestran agrupados por cada una de las topologías a fin de comparar el rendimiento de cada una de ellas. Se pueda apreciar que el mejor rendimiento corresponden a los modelos 98 y 104 pertenecientes a la Topología 2 con un valor de 2.31%. Esto representa que como máximo, 2 individuos de los 72 autorizados por el sistema no serán reconocidos por el sistema. Observése que el modelo 78, también perteneciente a la Topología 2, está tercero en el ranking obtenido pero el rendimiento global del sistema (medido por el TSR) es mejor que el de sus antecesores. El modelo 78 también rechazará como máximo 2 sujetos de los 72 autorizados en el sistema pero el índice de aceptación de intrusos mejora al punto de aceptar un máximo de 25 intrusos (60.83% x 40) en lugar de los 31 (76.67% x 40) del modelo 98.

TABLA VII
PRIORIZANDO EL RECONOCIMIENTO

Topología 1				Topología 2			
Modelo	FRR	FAR	TSR	Modelo	FRR	FAR	TSR
95	2.78 %	64.17 %	24.70 %	**98**	**2.31 %**	**76.67 %**	**28.87 %**
129	4.17 %	100.00 %	38.39 %	**104**	**2.31 %**	**63.33 %**	**24.11 %**
145	4.17 %	99.17 %	38.10 %	**78**	**2.78 %**	**60.83 %**	**23.51 %**
147	4.63 %	100.00 %	38.69 %	96	3.24 %	60.00 %	23.51 %
23	5.09 %	100.00 %	38.99 %	158	3.24 %	63.33 %	24.70 %

b Priorizando el Rechazo Total de Intrusos

Para esta alternativa se descartaron todos aquellos modelos que no alcanzaron un rechazo total de intrusos y se confeccionó un ranking descendente por el TSR de cada uno de ellos. La TABLA VIII muestra los primeros 5 modelos del ranking obtenido por cada una de las topologías en estudio. Se puede apreciar que el modelo 158, perteneciente a la topología 2, es el que lidera el ranking de un rechazo total de intrusos y con un rechazo máximo de 11 individuos de los 72 autorizados por el sistema (15.28% x 72). De los datos expuestos se puede apreciar también que la topología 2 obtiene en general mejores resultados cuando se prioriza el rechazo total de intrusos.

c Priorizando el mejor rendimiento

En esta variante se ordenaron en forma descendente cada uno de los modelos por el mejor índice global de rendimiento (TSR) obtenido. Los primeros modelos que lideran el ranking conformado puede apreciarse en la TABLA VII. Los modelos 114, 96 y 134 son los que han logrado el mejor rendimiento del clasificador existiendo entre ellos pequeñas diferencias entre el índice de reconocimiento y el de rechazo de intrusos. Para el modelo 114 los valores obtenidos indican que el clasificador rechazará 5.6 individuos autorizados y permitirá el acceso de 3 intrusos al sistema. El modelo 96 y 134 ofrece resultados similares, es decir rechazará entre 5 y 6 individuos autorizados y permitirá el acceso de 3 o 4 intrusos al sistema.

TABLA VIII
PRIORIZANDO EL RECHAZO TOTAL

Topología 1				Topología 2			
Modelo	FRR	Umbral	TSR	Modelo	FRR	Umbral	TSR
149	42.59 %	0.78	27.38 %	**158**	**15.28 %**	**0.58**	**9.82 %**
131	50.93 %	0.97	32.74 %	134	15.74 %	0.57	10.12 %
145	50.93 %	0.58	32.74 %	104	18.52 %	0.65	11.90 %
73	62.96 %	0.42	40.48 %	162	18.98 %	0.83	12.20 %
75	63.89 %	0.38	41.07 %	140	25.46 %	0.62	16.37 %

TABLA IX
PRIORIZANDO EL RENDIMIENTO

Topología 1				Topología 2			
Modelo	FRR	FAR	TSR	Modelo	FRR	FAR	TSR
149	10.19 %	6.67 %	8.93 %	**114**	**7.87 %**	**7.50 %**	**7.74 %**
95	11.11 %	10.00 %	10.71 %	**96**	**6.94 %**	**9.17 %**	**7.74 %**
41	10.19 %	15.00 %	11.90 %	**134**	**7.41 %**	**8.33 %**	**7.74 %**
131	12.96 %	10.83 %	12.20 %	78	7.41 %	9.17 %	8.04 %
53	14.35 %	14.17 %	14.29 %	158	12.04 %	3.33 %	8.93 %

VI. CONCLUSIONES.

Este trabajo ha presentado una metodología para la construcción y selección de clasificadores basados en redes neuronales artificiales para sistemas biométricos de Reconocimiento Facial. Se aplicó la metodología a un caso de estudio que involcura una base de datos de rostros disponible públicamente, verificando la hipótesis de que es factible poder encontrar una red neuronal que actúe aceptablemente como un clasificador para el RF, conforme a los requerimientos del sistema y a los datos proporcionados durante su entrenamiento, siguiendo procedimientos automáticos.

REFERENCIAS

[1] P. S. y. J. A. K. Prabhakar S., "Biometric recognition: Security and privacy concerns," IEEE Security and Privacy, 2003.
[2] P. R. y. P. B. Carrasco M. A., "Reconocimiento biométrico de audio y rostro: Un sistema viable de identificación," Departamento de Ciencia de la Computación, Pontificia Universidad Católica de Chile, 2006.
[3] O. Muller, Verificación Biométrica Automática de Identidad Mediante Reconocimiento Facial, Tesis de grado. Santa Fe: FICH-UNL, 2007.
[4] X. Qinghan, "Biometrics-technology, application, challenge, and computational intelligence solutions," Technology Review IEEE, Mayo 2007.
[5] =2plus 43minus 4W. Zhao, R. Chellappa, P. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," ACM Comput. Surv., vol. 35, no. 4, pp. 399–458, 2003. [Online]. Available: http://portal.acm.org/citation.cfm? id=954339.954342 =0pt
[6] S. Haykin, Neural Networks: A Comprehensive Fondation. New York: Prentice Hall, 2002.
[7] D. Capello, C. Martinez, D. Milone, and G. Stegmayer, "Array of multilayer perceptrons with no-class resampling training for face recognition," Revista Iberoamericana de Inteligencia Artificial, vol. 13, no. 44, pp. 5-13, 2009
[8] A. Eleyan and H. Demirel, "Pca and lda based neural networks for human face recognition," in Face Recognition, K. Delac and M. Grgic, Eds., Vienna, Austria: I-Tech Education and Publishing, 2007, pp. 93–106.
[9] A. Khashman, "A modified backpropagation learning algorithm with added emotional coefficients," IEEE Transactions on Neural Networks, vol. 19, no. 11, pp. 1896–1909, 2008.
[10] M. O. M. Caballero A. M., "Técnicas biométricas de identificación personal," Departamento de Ingeniería y Arquitectura Telemáticas (DIATEL), Universidad Politécnica de Madrid, 2004.
[11] P. A. P. Turk M. A., "Face recognition using eigenfaces," Proc. IEEE, pp. 586–591, 1991.

[12] R. P. D. Jain, A. K. and J. Mao, "Statistical pattern recognition: A review," IEEE Trans. Pattern on Analysis and Machine Intelligence, vol. 22, no. 1, 2000.
[13] S. Dzeroski and B. Zenki, "Is combining classifiers better than selecting the best one," in ICML, 2000, pp. 123–130.
[14] K. Fukunaga, Introduction to Statistical Pattern Recognition, second edition, Academic Press, 1990.
[15] J. Jackson, A users guide to principal components. New York: John Wiley and Sons, 1991.
[16] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," in Proc. of the IEEE Computer Vision and Pattern Recognition, 1996, pp. 203–208.
[17] A. Khashman, "Face Recognition Using Neural Networks and Pattern Averaging," in Lecture Notes in Computer Science, vol. 3972, no. 1, pp. 98–103, 2006.

Martín Maglianesi received a bachelor's degree in Information Systems at the Universidad Católica de Santa Fe, Argentina, in 2005 and the Master of Engineering in Information Systems in the Universidad Tecnologica Nacional - Regional Santa Fe, SantaFe, Argentina, in 2011.

Georgina Stegmayer received the Engineering degree in Information Systems Engineering from Universidad Tecnologica Nacional - Regional Santa Fe, SantaFe, Argentina, in 2000, and the PhD in electronic devices from Politecnicodi Torino, Torino, Italy, in 2006. She is Adjunct Researcher at CONICET (Argentina) since 2007. Her current research interest are applications of neural networks to modeling and data mining problems.