



UNIVERSIDAD AUTÓNOMA DEL ESTADO DE MÉXICO



CENTRO UNIVERSITARIO UAEM TEXCOCO

**SISTEMA DE RECONOCIMIENTO DE ROSTROS
UTILIZANDO REDES NEURONALES ARTIFICIALES**

TESIS

QUE PARA OBTENER EL GRADO DE
MAESTRO EN CIENCIAS DE LA COMPUTACIÓN

P R E S E N T A

JOSÉ JAIR VÁZQUEZ PALMA

TUTOR ACADEMICO:

Dra. CRISTINA JUÁREZ LANDIN

TUTORES ADJUNTOS:

Dr. ADRIAN TRUEBA ESPINOZA

M. en E. ANABELEM SOBERANES MARTIN.

TEXCOCO, ESTADO DE MEXICO

JULIO DE 2011



DICTÁMEN DE AUTORIZACIÓN DE OBTENCIÓN DE GRADO DE MAESTRÍA

Texcoco, Méx. , a 01 JULIO DEL 2011



TÍTULO DEL PROYECTO:

SISTEMA DE RECONOCIMIENTO DE ROSTROS UTILIZANDO REDES NEURONALES ARTIFICIALES



TESISTA:

JOSE JAIR VÁZQUEZ PALMA



DICTAMEN:

NO. DE REVISIÓN: Segunda



- RECHAZADO
- SUJETO A MODIFICACIONES
- ACEPTADO, CONDICIONADO
- ACEPTADO

COPIA

OBSERVACIONES GENERALES:

Aceptado el trabajo para la impresión

Aceptada la aplicación de software

Aceptado para la defenza de grado



<p>Tutor Adjunto</p>  <p>Dr. Adrian Trueba Espinoza.</p>	<p>Tutor Académico</p>  <p>Dra. Cristina Juárez Landin.</p>	<p>Tutor Adjunto</p>  <p>M. en E. Anabelem Soberanes Martin.</p>
---------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------



Agradecimientos

A lo largo de este camino han estado muchas personas que con su ayuda han hecho posible atravesarlo exitosamente.

En primer lugar a la Universidad Autónoma del Estado de México y en especial al Centro Universitario UAEM Texcoco por brindarme la oportunidad de realizar mis estudios de licenciatura y maestría.

A COMECYT y a la UAEMex por otorgarme el soporte económico sin el cual hubiera sido imposible continuar mis estudios.

A mis padres por brindarme todo su apoyo, amor y cariño en todo momento, contando con sus consejos para poder continuar y no desvanecer en el proceso de este trabajo; a mi hermana por haberme brindado todo su apoyo.

A mi tutor académico la Dra. Cristina Juárez Landin por todos sus consejos, por su confianza y por haberme aceptado como su estudiante y por el apoyo otorgado para la realización de esta tesis. A la M. en E. Anabelem Soberanes Martín y al Dr. Adrian Trueba Espinoza por su ayuda y amabilidad en todo momento. A los tres por contribuir enormemente a que yo tuviera una mejor formación profesional y por las enormes mejoras del proyecto, resultado de los ratos de trabajo que tuvimos.

A todos mis profesores por sus enseñanzas durante mi instancia como alumno de maestría y al personal del Centro Universitario UAEM Texcoco en especial al departamento del Posgrado que amablemente me ayudó cada vez que lo necesitaba.

A todos mis amigos, compañeros que estuvieron en todo momento brindándome su apoyo y consejos para la mejora continua en la realización de este trabajo de investigación.

RESUMEN

En este trabajo se presenta un método para el reconocimiento de rostros, robusto a cambios de iluminación, postura y expresiones faciales. El método propuesto se encuentra dividido en dos etapas fundamentales; la primera que consiste en la extracción de características faciales por medio de la transformada wavelet discreta (TWD) y la segunda en la clasificación de patrones mediante la red neuronal perceptrón multicapa (MLP) a partir de los vectores característicos extraídos.

La TWD es utilizada con la finalidad de construir un subespacio característico de bajas dimensiones expandido por medio de funciones wavelets a partir del espacio original de la imagen. La representación wavelet describe las características faciales de forma más robusta y simple que la representación original de la imagen. Las variaciones entre vectores característicos debido a los cambios de identidad son mayores que las variaciones intraclase producidas por cambios de iluminación, postura y expresión facial.

Para la etapa de clasificación de patrones, se emplea la MLP por su capacidad para resolver problemas de clasificación no lineal y generalizar a partir de un número limitado de ejemplos. La topología de la MLP consta de 42 nodos en su capa oculta, 30 nodos en sus dos capas ocultas y 15 nodos en la capa de salida los cuales corresponden a las 15 personas a reconocer.

La combinación de las redes neuronales y la TWD permite eliminar o al menos reducir algunos inconvenientes que presentan varios métodos propuestos de reconocimiento de rostros. Las pruebas así como la validación de los resultados usando diferentes filtros wavelet y niveles de descomposición son presentadas en este trabajo así como la comparación con otro método el cual utiliza el histograma para procesar imágenes junto con la MLP.

Abstract

This work shows a method for the recognition of faces, robust to changes of illumination, position and expressions. The proposed method is divided in two fundamental stages; first that consists from the extracted facial feature vectors characteristics by using the discrete wavelet transform (TWD) and second pattern matching from the extracted facial feature vectors based on multilayer perceptron neural network.

The TWD is used in order to construct a subspace characteristic of low dimensions expanded by means of functions wavelets from the original space of the image. The filters wavelet describes to the face characteristics of more robust and simple form that the original representation of the image. The variations between characteristic vectors due to the changes of identity are enlarged that the variations intra-subjet produced by changes of illumination, position and face expression.

For pattern classification stage, the MLP is used for its ability to solve problems of classification nonlinear and generalize from a limited number of examples. Topology of the MLP has 42 hidden layer nodes, 30 nodes in two hidden layers and 15 nodes in the output layer which correspond to the 15 people to recognize.

The combination of neural networks and wavelets allows us to eliminate or at least to reduce the drawbacks of many face recognition methods. The tests as well as the validation of the results using different filters wavelet and levels from decomposition are presented/displayed in this paper as well as the comparison with another method which uses the histogram to process images along with the MLP.

INDICE GENERAL

Índice de figuras	1
Índice de tablas	3
Capítulo 1: Introducción	4
1.1. Antecedentes.....	6
1.2. Planteamiento del problema.....	9
1.3. Objetivos.....	10
1.3.1 Objetivo general.....	10
1.3.1 Objetivos específicos.....	10
1.5. Delimitación o alcances de la investigación.....	11
1.6. Hipótesis.....	12
1.7. Justificación.....	13
1.8. Fundamentación inicial.....	15
Capítulo 2: Transformada wavelet	24
2.1. Historia de los wavelets.....	24
2.2. Principios fundamentales de los wavelets.....	25
2.3. Transformada wavelet continua CWT.....	27
2.4. Transformada wavelet discreta TWD.....	28
2.5. Las funciones wavelet.....	31
2.6. Familias wavelet.....	31
2.6.1. Familia haar.....	32
2.6.2. Familia daubechies.....	32
2.6.3. Familia coiflets.....	33
2.6.4. Familia biortogonal.....	33
2.6.5. Discrete meyer.....	35
2.7. Aplicación de wavelets en imágenes.....	35

Capítulo 3: Redes neuronales artificiales	39
3.1. Modelo de neurona artificial.....	40
3.2. Topología de la RNA.....	42
3.2.1. Aprendizaje de las RNA.....	43
3.2.2. Red neuronal perceptrón multicapa.....	45
3.2.3. Funcionamiento de la MLP.....	47
3.2.4. Algoritmo backpropagation.....	50
Capítulo 4: Sistema propuesto	53
4.1. Método 1: Basado en Histograma.....	53
4.1.1. Implementación.....	54
4.1.2. Resultados del método propuesto.....	56
4.2. Método 2: Basado en la transformada wavelet discreta.....	58
4.2.1. Adquisición de datos.....	59
4.2.2. Extracción de características.....	61
4.2.3. Vectores característicos.....	63
4.2.4. Entrenamiento de la red neuronal artificial.....	66
4.3. Diseño del sistema.....	71
4.3.1. Arquitectura del sistema.....	71
Capítulo 5: Validación de resultados	75
5.1. Sensibilidad.....	76
5.2. Especificidad.....	76
Conclusiones	79
Trabajos a futuro	81
Bibliografía	82
Anexo 1. Base de datos UAEM Texcoco/Valle de Chalco.....	88
Anexo 2. Base de datos ORL.....	89
Anexo 3. Familias wavelet con sus respectivos filtros evaluados.....	94
Anexo 3. Ventana de vista de datos.....	100

INDICE DE FIGURAS

Figura 2.1. Árbol de descomposición por wavelets en tres niveles.	29
Figura 2.2. Esquema de filtrado para calcular la TWD de una imagen.	30
Figura 2.3. Descomposición de la imagen de un rostro en matrices de aproximación y detalles.	31
Figura 2.4. Wavelet de Haar.	32
Figura 2.5. Familia de Wavelets Daubechies.	33
Figura 2.6. Familia de Wavelets Coiflets.	33
Figura 2.7. Familia de Wavelets Biortogonales.	34
Figura 2.8. Esquema de extracción de características mediante la TWD.	35
Figura 2.9. Subimágenes resultantes del proceso de descomposición.	36
Figura 2.10. Ejemplo de descomposición de una imagen empleando Wavelets de Daubechies.	37
Figura 3.1. Modelo de neurona de McCulloch y Pitts.	40
Figura 3.2. Ejemplos de funciones de activación.	41
Figura 3.3. Arquitectura de una red MLP con propagación hacia adelante.	45
Figura 3.4. Arquitectura de una red neuronal de retropropagacion basado en un Perceptron Multicapa.	48
Figura 4.1: Flujo de Información, primer método propuesto.	53
Figura 4.2: Ejemplo de los patrones de entrenamiento para la primera persona.	55
Figura 4.3: Ejemplo de los patrones de prueba para la décima persona.	55
Figura 4.4: Diferencia entre las técnicas utilizadas para la obtención del vector de entrada para la RNA.	56
Figura 4.5: Imágenes mal clasificadas para la primera técnica.	57
Figura 4.6: Imágenes mal clasificadas para la segunda técnica.	57
Figura 4.7. Diagrama del sistema de reconocimiento de rostros.	59
Figura 4.8. Ejemplo de Base de datos ORL.	60
Figura 4.9. Esquema de filtrado para calcular la TWD unidimensional.	61

Figura 4.10. Esquema de la descomposición wavelet.	62
Figura 4.11. Niveles del proceso de descomposición de una imagen facial en pixeles.	63
Figura 4.12. Esquema de extracción de características mediante la TWD.	63
Figura 4.13. Variación entre los vectores característicos de una persona.	64
Figura 4.14. Variación entre los vectores característicos de dos personas.	65
Figura 4.15. Topología general de la red neuronal utilizada.	67
Figura 4.16. Ejemplo de los patrones de entrenamiento.	68
Figura 4.17. Ejemplo de los patrones de prueba.	68
Figura 4.18. Resultados obtenidos aplicando diferentes filtros wavelets. Prueba abierta.	70
Figura 4.19: Esquema general de arquitectura.	72
Figura 4.20: Esquema general de la arquitectura del procesamiento de la RNA.	73
Figura 4.21: Interfaz de usuario de preprocesamiento.	74

INDICE DE TABLAS

Tabla 3.1. Regiones de decisión con diferentes números de capas y de neuronas con una MLP.	46
Tabla 4.1: Parámetros de la Red Neuronal Artificial, primer proceso.	56
Tabla 4.2: Parámetros de la Red Neuronal Artificial, segundo proceso.	57
Tabla 4.3: Bases de datos de rostros estándar disponibles vía internet.	59
Tabla 4.4. Familias wavelet con sus respectivos filtros evaluados.	68
Tabla 4.5. Resultados obtenidos aplicando diferentes filtros wavelets.	69
Tabla 5.1: Tabla de contingencia o matriz de confusión.	75
Tabla 5.2: Tabla de decisiones para la clasificación.	77
Tabla 5.3: Matriz de confusión con los resultados obtenidos usando filtros wavelet.	77

CAPITULO 1

INTRODUCCION

A pesar que para el sistema visual humano reconocer rostros es una tarea rutinaria y sencilla, el diseño de un modelo computacional de reconocimiento de rostros representa un problema de grandes desafíos. Reportes de evaluaciones y estudios sobre reconocimiento de rostros indican que el rendimiento de muchos algoritmos es deteriorado debido a cambios de iluminación, expresión facial, entre otros factores [Chellappa, 1995; Phillips, 2000; Zhao, 2003].

La identificación personal es una función vital en muchas actividades de la sociedad moderna. La necesidad de contar con mecanismos de identificación más confiables y prácticos ha generado un enorme interés en el desarrollo de métodos de identificación biométricos. Los sistemas biométricos utilizan características fisiológicas (iris, huella dactilar, rostro, entre otros) y del comportamiento de los individuos (ej. cadencia de tecleo, voz, etc.) para identificarlos. La biometría es un campo tecnológico que consiste en identificar personas a través de características físicas únicas. Desde la percepción del ser humano, la forma más sencilla de reconocer personas es a través del rostro. El rostro es un elemento del cuerpo humano ampliamente estudiado por distintas disciplinas que van desde la psicología y sociología al procesamiento de imágenes y visión por computadora, entre otras.

El reconocimiento de rostros por medios computacionales ha recibido especial atención debido principalmente a la demanda de seguridad en sitios públicos como estadios, aeropuertos, etc. y al gran potencial comercial de sus aplicaciones. Además, las muestras de los rostros a diferencia de otros atributos biométricos pueden ser obtenidas en forma sencilla y sin la necesidad de solicitar al usuario de su cooperación e incluso en muchos casos sin que tenga conocimiento de estar siendo identificado.

Gracias al desarrollo de computadoras personales de alto rendimiento y al gran auge de las tecnologías digitales, la implementación comercial masiva de sistemas de reconocimiento de rostros confiables y eficientes ofrece prometedoras expectativas. Las aplicaciones del reconocimiento de rostros incluyen el control de acceso, vigilancia, interacción hombre-máquina, etc.

Las aplicaciones de reconocimiento de rostros implican diferentes problemas debido a la difícil situación, dentro de los parámetros a considerar, problemas como la detección de la cara, la oclusión, etc. que deben ser considerados [Oropeza, 2004]. Este trabajo no pretende resolver todos estos problemas, se centran en escenarios donde las imágenes son la cara de frente, en este caso los parámetros como la variación de la iluminación, rotación, expresiones faciales, etc.

En el modelo computacional propuesto nos enfocaremos a los sistemas en donde se requiere encontrar la identidad de un individuo utilizando una base de datos, este tipo de sistemas no necesitan hacer un reconocimiento en tiempo real.

De manera muy general se puede describir dicho modelo como sigue:

- a) Adquisición de bases de datos de imágenes de rostro.
- b) Elección una mejor base de datos de imágenes de rostro de acuerdo a sus características (condiciones de iluminación, expresiones faciales y número de imágenes de rostros que integran la bases de datos).
- c) Diseño, desarrollo e implantación de un algoritmo para la extracción de características faciales con imágenes de rostro que presenten diferentes condiciones de iluminación, escala, traslación, inclinación, desenfoque, ruido, expresiones faciales, vestuario y edad.
- d) Diseño e implantación de la red neuronal perceptron multicapa que permita la clasificación y verificación de cada uno de los individuos para el Reconocimiento de Rostros.

1.1 ANTECEDENTES

En el campo de la Inteligencia Artificial existen algoritmos empleados en el Reconocimiento de Rostros que permiten la extracción de características empleando diversos métodos, entre los que destacan por su uso más común: los Eigenfaces, los Fisherfaces, Redes Neuronales Artificiales (RNA), métodos de Aproximación Probabilística, Redes Bayesianas, métodos de Análisis de Discriminación Lineal (LDA) y métodos de Análisis de Componentes Principales (PCA).

Las Redes Neuronales Artificiales las más usadas en los últimos años con un considerable interés en los campos de la computación y de la inteligencia artificial [INAOE, 2005]. El elemento clave es una estructura computacional compuesta de diversos elementos interconectados (neuronas) trabajando en paralelo. Esta estructura paralela permite realizar muchas operaciones simultáneamente. Actualmente las redes neuronales han probado ya su valía para resolver problemas complejos que a primera vista parecen intratables y son difíciles de formular usando técnicas de cómputo convencionales.

Ejemplos de tales problemas se pueden encontrar en un vasto número en áreas, tales como el reconocimiento de patrones (el reconocimiento de rostros, del habla, etc.), la predicción y el pronóstico de series temporales, el control de procesos, el procesamiento de señales, entre otros.

Una de las tecnologías emergentes que muestra grandes posibilidades de desarrollo en los próximos años, es la Biometría, la cual tiene como propósito la identificación de personas por medio de sus diversas características físicas (voz, huellas digitales, características faciales, etc.).

Los sistemas de reconocimiento de rostros tienen el objetivo de detectar posibles sospechosos, a partir de diversas bases de datos con fotografías de personas consideradas terroristas o delincuentes.

También, tienen otros múltiples usos de identificación en la vida cotidiana: acceso a cajeros automáticos, a estacionamientos, acceso de personal a oficinas, reconocimiento de personas desaparecidas o fallecidas que no presenten identificación, acceso a computadoras (en lugar de utilizar contraseñas), entre otros [Villegas, 2005].

De manera general, los sistemas de reconocimiento de rostros se pueden clasificar en dos categorías:

1. Cuando se requiere encontrar la identidad de un individuo utilizando una base de datos. Este tipo de sistemas no necesitan hacer un reconocimiento en tiempo real.
2. Cuando se requiere identificar a una persona en tiempo real ó cuando se desea permitir el acceso únicamente a un grupo de personas [Oropeza, 2004].

En el modelo computacional propuesto está enfocado a los sistemas de la primera categoría, es decir, los sistemas que emplean una base de datos de imágenes faciales.

Las características que deben satisfacer los sistemas de reconocimiento de rostros son las siguientes:

1. Universalidad, que significa que cada persona debe tener esas características.
2. Unicidad, que significa que dos personas no deben ser la misma en términos de las características.

3. Permanencia, que indica que las características deben ser invariantes con el tiempo.
4. Colectibilidad, que indica que las características pueden ser medibles cuantitativamente.

Actualmente, existen técnicas biométricas que son ampliamente usadas; estas técnicas se basan en: rostro, huella dactilar, geometría de la mano, venas de la mano, iris, patrón retinal, firma, voz, y termograma facial [García, 2006].

Entre las técnicas basadas en rostro utilizando RNA se encuentran sistemas de reconocimiento que utilizan interfaz visual independiente [Villa, 2005], o los basados en línea [García, 2008], así como los que además de utilizar RNA se valen de otros métodos como los algoritmos genéticos [Anam, 2009]; aquellos otros basados en huella dactilar para el acceso a recintos [García, 2006], los que utilizan la firma como seguridad e identificación del personal [Báez, 2006].

En el campo de la medicina para la pre-detección de enfermedades como Creutzfeldt-Jakob, Alzheimer [Rodríguez, 2005], a nivel celular el análisis a los hidratos de carbono en proteínas y otras complejas [Mailk, 2007], la estimación de material en el revestimiento transparente sobre los niveles de colesterol en la sangre [Hamilton, 2009] así como la elaboración de nuevos modelos de cirugía asistida por computadora a partir del reconocimiento de patrones que permiten realizar un análisis profundo y detallado para diagnósticos más precisos y así determinar un tratamiento correcto [ULSA, 2005].

1.2 PLANTEAMIENTO DEL PROBLEMA

Son pocos los sistemas avanzados de identificación física de personas. Teniendo normalmente un alto grado de inseguridad sobre el control de acceso a instalaciones, así como, un inadecuado registro de asistencia ya que el personal checa sus entradas y salidas sin que se tenga la certeza de que el empleado sea quien realice dicha operación, eludiendo los lineamientos y/o reglamentos de la empresa, generando pérdidas económicas por pagos extras a empleados quienes realizan tareas ajenas a su área laboral en las distintas áreas de empresa u organizaciones.

Ahora bien, estudios no tan recientes demuestran que hay ciertas características propias de un individuo como lo son la personalidad en el sentido psicológico, también las hay aquellas que representan singularidades hablando en el sentido estrictamente físico, tales como: huellas dactilares, los ojos, las manos, la complexión, las líneas de las manos, la voz, y la cara, ésta última las más importante en esta investigación.

A pesar de que existen clasificaciones que las distinguen en diferentes tipos, es sabido que el rostro sigue un patrón distinto para cada ser humano y por ende las hace únicas.

En este trabajo se propone el desarrollo de un sistema de reconocimiento de rostros por medio de la extracción de características faciales utilizando redes neuronales artificiales.

1.3 OBJETIVO GENERAL

Desarrollar un sistema computacional para el reconocimiento de rostros, robusto a cambios de iluminación, postura y expresión facial usando técnicas de extracción de características y de Redes Neuronales Artificiales.

1.4 OBJETIVOS ESPECIFICOS

- Construir una Bases de datos de imágenes faciales que cumplan con parámetros apropiados para ser utilizadas en el sistema propuesto.
- Filtrar imágenes digitales empleando el método de la transformada Wavelet (TW) para la obtención de patrones.
- Construir una base de conocimiento para realizar la fase de entrenamiento con la red propuesta a partir de diversas fotografías de personas.
- Implementar un algoritmo de RNA para la autenticación mediante rasgos faciales que proporcione una alta eficiencia.
- Realizar diversas pruebas con el sistema propuesto variando el número de fotografías por persona, para configurar la base de conocimiento.

1.5 DELIMITACIÓN O ALCANCES DE LA INVESTIGACIÓN

En este trabajo de investigación se propone el desarrollo de un sistema de reconocimiento de rostros, basado en redes neuronales artificiales específicamente la red Perceptron Multicapa.

Para efectos de este problema la imagen facial digitalizada deberá ser una imagen de fuente fidedigna, la cual podría ser obtenida de bases de datos, algunas bases de datos de imágenes faciales a considerarse son:

- CVG: (Computational Vision Group), [CVG, 2005].
- The AR Face Database [AR Face Database, 1998].
- CMU: (CMU Image Data Base), [CMU, 2000].

Por lo tanto es importante el desarrollo de herramientas computacionales en el apoyo de seguridad, que empresas gubernamentales y/o privadas realizan al identificar al personal para un control y acceso a las instalaciones. En este sentido, se propone el desarrollo de un sistema de reconocimiento de patrones que representan el rostro de una persona.

El procesamiento de la imagen se efectuara mediante un método del filtrado para extraer las características para el reconocimiento de patrones. El desarrollo del algoritmo de la RNA realizara a partir de las características extraídas en el método de filtrado de la imagen y el entrenamiento de la misma se hará con el algoritmo de backpropagation, donde se efectuará la clasificación de las imágenes.

El sistema se examinara en forma general identificando mejorías y restricciones para realizar las modificaciones en relación a los resultados obtenidos.

1.6 HIPÓTESIS

La implementación de un sistema de reconocimiento de patrones de rostros utilizando el método de filtrado de la transformada Wavelet a imágenes digitales y una red neuronal artificial; podrán ser técnicas adecuadas de procesamiento de imágenes faciales para permitir reconocer a personas con acceso a recintos que manejan seguridad, esperando obtener una herramienta de software para la identificación de personas.

Se demostrará que cuando se utiliza la transformada de Wavelet conjuntamente aplicando la RNA Backpropagation para la clasificación, se espera obtener un reconocimiento facial a partir de un 90% de certidumbre para imágenes de rostro con cambios de iluminación, expresiones faciales y ángulos de orientación del rostro. Para el procesamiento de imágenes se utilizarán herramientas de Matlab, así como para la clasificación de imágenes.

1.7 JUSTIFICACIÓN

Cubrir la mayoría de las necesidades de seguridad e identificación de la persona, se propone el desarrollo un sistema que sea capaz de reconocer personas a partir de sus características faciales y compararlas con una base de datos para finalmente identificar si se trata de una persona registrada o no. Para esto, es necesario que el sistema sea lo más confiable posible para no ocasionar problemas en una situación real.

El tema de reconocimiento de rostros por medios informáticos ha cobrado amplio interés en vista de su importancia en el contexto actual en medidas de seguridad, siendo esta la base fundamental de la realización de este trabajo. La tarea de reconocimiento de rostros la podemos encontrar en sitios públicos como estadios, aeropuertos, bancos, etc. y al sobresaliente potencial comercial de sus aplicaciones donde la seguridad es primordial, lugares donde se requiera identificar de manera automática y con un grado de confiabilidad a las personas que tienen acceso a edificios o áreas restringidas, resultando en este caso atractivo para ser implementado.

Las muestras de los rostros a diferencia de otros atributos biométricos pueden ser obtenidas en forma sencilla y sin la necesidad de solicitar al usuario de su cooperación e incluso en muchos casos sin que tenga conocimiento de estar siendo identificado.

Por tales características, el reconocimiento de rostros se encuentra entre los sistemas biométricos más aceptados por los usuarios. Algunas aplicaciones comerciales han podido ofrecer soluciones apenas adecuadas para la identificación personal por medio de rostros [García, 2008].

Sin embargo, gracias al desarrollo de computadoras personales de alto rendimiento y al tremendo auge de las tecnologías digitales, la implementación comercial masiva de sistemas de reconocimiento de rostros confiables y eficientes ofrece prometedoras expectativas. Las aplicaciones del reconocimiento de rostros incluyen el control de acceso, vigilancia, interacción hombre-máquina, [Fuentes, 2008].

El usar un sistema de control de acceso que utilice algún tipo de tarjeta tiene ciertas desventajas. Por esta razón, la realización de esta propuesta es desarrollar un sistema de control de acceso, que se pueda implementar en cualquier recinto (casa, edificio, etc.), manejando un sistema biométrico el cual utiliza como patrón de identificación los rasgos faciales de cada una de las personas que tengan acceso a éste.

El procesamiento debe ser computacionalmente lo más eficiente posible apuntándose a obtener un producto final que sea de fácil operatividad por parte del usuario.

1.8 FUNDAMENTACIÓN INICIAL

Con el rápido desarrollo de las RNA durante las dos últimas décadas en campos de la inteligencia artificial tales como el reconocimiento de patrones, aparecieron un extenso número de problemas complejos donde no era conveniente una representación explícita del conocimiento y no se disponía de un procedimiento de razonamiento lógico para resolverlo, por esto, las aproximaciones algorítmicas y las estructuras computacionales estándar no eran apropiadas para resolver estos problemas [Oropeza, 2004].

Las RNA fueron introducidas como estructuras computacionales alternativas, creadas con el propósito de reproducir las funciones del cerebro humano. Los modelos computacionales conocidos como redes neuronales están formadas por un número de procesadores o neuronas, dispuestos en varias capas e interconectadas entre sí mediante conexiones con pesos. Estos procesadores realizan cálculos simples basados en la información que reciben de los procesadores vecinos. Las redes neuronales no siguen reglas programadas rígidamente como lo hacen las computadoras digitales más convencionales: más bien usan un proceso de aprendizaje por analogía donde los pesos de las conexiones son ajustados automáticamente para reproducir un conjunto de patrones representativos del problema a aprender. Este aprendizaje también está inspirado en la forma de aprender que tiene lugar en las neuronas, cambiando la efectividad de las sinapsis, de tal manera que la influencia de una neurona en otra varía. Es importante mencionar que las arquitecturas computacionales habituales de redes neuronales son extremadamente simplificadas cuando se analizan desde un punto de vista neurofisiológico; sin embargo, estos modelos tan simples permiten resolver muchos problemas interesantes [Fuentes, 2008].

La identificación de los patrones de una imagen se desarrolla mediante el método de filtrado. El objetivo del filtrado proporcionar una imagen preprocesada que luego pueda ser tratada con facilidad por el resto de las etapas.

Existen numerosos métodos de filtrado como los filtros adaptativos, la transformada de Fourier, la transformada wavelet, entre otros; los cuales resuelven problemas de tratamiento de imágenes.

La aplicación de las RNA ha causado impacto en la resolución de problemas de reconocimiento, simulación, seguridad, diagnósticos médicos. Existen evidencias de la aplicación de estas en algunos trabajos de investigación como:

- Reconocimiento de rostros [Oropeza, 2004] donde se utilizan Redes Neuronales y Filtros Bidimensionales Gabor (FBG); los FBG tienen la tarea de extraer las características necesarias para el reconocimiento contemplando cambios de iluminación, translación, inclinación, ruido, expresiones fáciles en imágenes; la clasificación y verificación para el Reconocimiento de Rostros se propone se realice empleando Redes Neuronales de Retropropagación (RNR).
- Reconocimiento de rostros usando Algoritmos Genéticos y la Red Neuronal Backpropagation [Anam, 2009]; El sistema consta de tres pasos. Al principio algunos pre-procesamientos se aplican sobre la imagen de entrada. En segundo lugar frente a la extracción características, que se tomará como la entrada de la Red Neuronal y algoritmos genéticos. En la tercera etapa se encuentra la clasificación que se lleva a cabo mediante el uso de de la Backpropagation. Mediante el uso del Algoritmo Genético y la Red Neuronal Backpropagation la eficiencia máxima es del 91.3% para el Sistema de reconocimiento de rostro.

- Sistema de Reconocimiento de Rostros basado en el Análisis de Componentes Principales (PCA) [Martínez, 2008]. Técnica que permite reducir la dimensionalidad de un conjunto de datos, el uso de esta técnica permite caracterizar de manera adecuada la información contenida en la imagen de un rostro. Sin embargo, se ha podido comprobar que debido esta caracterización, PCA es sensible a factores de ruido como la iluminación y cambios en la escala. Los resultados muestran que PCA tiene un alto desempeño frente a cambios en la expresión facial y cambios no radicales en la orientación del rostro, mientras que a factores de ruido como la iluminación, enfoque y escala, PCA muestra un bajo desempeño con porcentajes menores al 70% de eficiencia.
- Reconocimiento de rostros utilizando histogramas secuenciales de imagen [Hernández, 2009] el cual aplica técnicas de segmentación de imágenes para poder aislar el rostro de la fotografía capturada y convertirlo a su respectivo en escala de grises, una vez realizado lo anterior, se procede a obtener el histograma de la imagen, el cual es comparado mediante un método estadístico conocido como distancias de Hamming, para poder hacer uso de histogramas secuenciales de imagen, es necesario comparar una imagen con una secuencia de diversas imágenes previamente almacenadas. Con respecto a las comparaciones de fotografías de frente entre diferentes personas, se pudo observar que la eficiencia se incrementó a 82.75%, con lo cual es posible concluir que se obtuvo una eficiencia bastante aceptable comparada con algoritmos más eficientes y de mayor robustez. El uso de la técnica de histogramas secuenciales de imagen tiene muchas limitantes, las cuales hacen que sea necesario combinar con el uso de esta técnica con unas de un mayor nivel de complejidad, como por ejemplo el uso de redes neuronales, para poder así llegar a un nivel de eficiencia excelente, en el que el margen de error sea mínimo o casi despreciable, y así poder desarrollar aplicaciones cada vez más confiables y seguras.

- El reconocimiento y verificación de hablantes por medio de sistemas computacionales [Hernández, 2004] está en pleno auge debido a los diferentes usos y ventajas que se pueden obtener de ello, principalmente cuando se trata de restringir y/o controlar el acceso físico o informático/lógico de individuos. El reconocimiento de hablantes se refiere a que un sistema sea capaz de identificar que hablante es el que está actuando, de un grupo de hablantes enlistados en una base de datos. En cuanto a la verificación de hablantes, ésta se refiere a que un sistema solamente verifica que la persona que está actuando es quien dice ser, entonces su salida sólo puede ser aceptada o rechazada. Con el desarrollo de la Biométrica, se ha ampliado el campo de estudio y creación de sistemas que puedan realizar tanto reconocimiento como verificación de hablantes, la mayoría de éstos están basados en las Redes Neuronales Artificiales.
- Análisis, Reconocimiento y Síntesis de Voz Esofágica [Caeiros, 2007], para el mejoramiento de la calidad de voz de personas especialmente los hablantes con voz esofágica, consiste en extraer y clasificar regiones de baja inteligibilidad de la voz para su posterior sustitución por segmentos de voz de un hablante normal. Este planteamiento lleva a identificar en el mejoramiento de la calidad de voz de pacientes laringectomizados para resolver problemas la extracción de características o parámetros de la señal de voz en zonas de baja inteligibilidad. La voz esofágica se produce insuflando aire desde la nariz y boca hasta el esófago ahí lo retienen y luego lo expulsan produciendo vibraciones que originan la voz. Las ventajas de este método son que no se requiere ninguna intervención quirúrgica ni la necesidad de un dispositivo que se deba portar cada vez que el paciente quiera comunicarse. Su desventaja principal radica en una larga rehabilitación que le permita al paciente comunicarse por medio de voz esofágica ya que no es algo que hacemos de manera natural.

- Diagnóstico y Monitoreo de los patrones de una turbina de gas mediante RNA [Gutiérrez, 2007], donde se diagnostican las fallas de una turbina de gas para el bombeo de gas natural, así como también su monitoreo empleando redes neuronales artificiales, específicamente la red neuronal con múltiples capas entrenada por medio del algoritmo de retropropagación.
- Control neuronal de un generador de inducción para Generación eólica [González, 2005], el impacto para la mejora ecológica que representa el empleo del viento como fuente de energía, surge la necesidad de hacer económicamente más atractiva esta opción energética para los gobiernos e industriales de todo el mundo. Surgiendo la necesidad de aumentar la eficiencia de los sistemas de conversión de energía. Una forma de aumentar la eficiencia en estos sistemas, es mediante el desarrollo de mejores técnicas de control, como lo han demostrado ser las técnicas inteligentes basadas en redes neuronales. Este trabajo presenta la elaboración de un control neuronal para controlar un generador de inducción empleado en plantas eoloelectricas.
- Diseño y Desarrollo de un Simulador para una Planta Virtual, por medio de Redes Neuronales [López, 2008], donde se describe el diseño y programación de una herramienta de software que permita simular el comportamiento dinámico de una planta de proceso, mediante redes neuronales.
- Construcción de una red neuronal artificial para la modelación de un reactor de Hidrodesulfuración de naftas [Paz, 2004]. Uno de los procesos que ha venido cobrando mayor importancia dentro de los complejos industriales de Petróleos Mexicanos (PEMEX) es el de Hidrodesulfuración. La presencia de azufre en las diversas corrientes provoca problemas tales como: corrosión de equipo, deterioro en la calidad del producto terminado, envenenamiento de catalizadores, y contaminación ambiental entre otros.

Actualmente una de las principales preocupaciones a nivel mundial es el creciente índice de contaminación en todos los ámbitos de la naturaleza: aire, agua, y tierra; la cual está directamente vinculada con las actividades industriales tendientes a satisfacer las necesidades de la sociedad moderna. Es esta creciente importancia de la Hidrodesulfuración motivo a la necesidad de contar con una herramienta de simulación por medio de RNA que permita predecir el comportamiento de sus variables.

- Sistema de Reconocimiento de Olores mediante Redes Neuronales Artificiales [Yukse, 2007]. El objetivo de este estudio es desarrollar una nueva red neuronal difusa; la cual sirve como modelo para los clasificadores en tiempo real de sistema de reconocimiento de olor. En este tipo de Red Neuronal, las neuronas de entrada son activaciones que se obtienen a través de la agrupación de los datos de entrada, de modo que el sistema neuronal podrían hacer frente a las estadísticas de los errores de medición directa. Entonces, el rendimiento de la red se compara con la otra red llamada Perceptrón Multicapa (MLP), que también se ha utilizado para lo mismo en el sistema de reconocimiento de olor.
- Red Neuronal de arquitectura paramétrica en reconocimiento de rostros [Salinas, 2000]. Aborda el tema de reconocimiento de rostros haciendo uso directo de redes neuronales de retropropagación, esto es sin preproceso de extracción de características. Este enfoque es poco usual, hay razones teóricas acerca de la desproporción entre información muestral y número de ejemplares que indican que en un espacio de tan poca densidad es imposible obtener resultados satisfactorios. Sin embargo, este trabajo plantea un diseño novedoso de arquitectura de red que logra excelentes resultados. En efecto, la red implementada en Matlab, entrenada con 5 tomas distintas pertenecientes a 40 individuos en aproximadamente 6 minutos. La prueba, con otro set de 5 fotos de los mismos 40 individuos, entrega una precisión de 100% para la identificación.

- Reconocimiento de rostros utilizando secuencias de histogramas como tramas espacio-temporales [Fernández, 2001]. Este artículo describe, un modelo de reconocimiento de rostros humanos, utilizando redes espacio-temporales (STN). Esta red neuronal incorpora una capa de entrada que se encarga de transformar la imagen de entrada al sistema en una secuencia de vectores normalizados que serán aplicados a una red espacio-temporal. Finalmente, la red tiene una capa de salida que utiliza unidades outstar de Grossberg. Presenta el funcionamiento de las redes de contrapropagación (CPN) para el reconocimiento facial; el procedimiento que se sigue es convertir la imagen en un vector de números y pasarlas a la capa de entrada de la red. Como vectores de entrada se probaron vectores de niveles de gris de imágenes y vectores de histogramas por filas de las imágenes.
- Reconocimiento de Rostros usando Datos No Etiquetados [Martínez, 2003]. Este trabajo presenta un método para reconocimiento de rostros utilizando un conjunto de entrenamiento pequeño. El método se basa en expandir el conjunto de entrenamiento original usando datos no etiquetados previamente. Inicialmente, aplicaron una técnica para reducir la dimensionalidad del espacio de atributos, después realizaron un proceso iterativo, clasificando todos los datos no etiquetados con un ensamble de clasificadores construido a partir del conjunto de entrenamiento actual y agregando al conjunto de entrenamiento los ejemplos que han sido clasificados correctamente con un alto nivel de confianza, de acuerdo al ensamble. Se realizaron experimentos usando ensambles basados en el algoritmo de k vecinos más cercanos, redes neuronales artificiales, y regresión lineal localmente ponderada. Los resultados experimentales demostraron que el uso de datos no etiquetados mejora la clasificación en todos los casos. Los mejores resultados, con un porcentaje de clasificación correcta de 92.07.

- Síntesis de expresiones faciales en 3D y su aplicación a sistemas de reconocimiento facial [Ramírez, 2009]. Consiste en un método eficiente para la síntesis de expresiones faciales en 3D y su aplicación a los sistemas de reconocimiento de caras independiente de la expresión (FRS). El modelo propuesto se divide en filtrado de ruido, usado para eliminar los picos (prominentes) presentes en las imágenes de profundidad; detección de puntos de control en la base de datos de caras en 3D; Registro del modelo base con cada una de las imágenes muestra con cara neutral en la base de datos de entrenamiento, para la generación de expresiones faciales sintéticas y su posterior inclusión en la base de datos misma para incrementar el conjunto de entrenamiento. Los resultados muestran que el método propuesto representa fielmente la imagen neutral de la base de datos y además, la adición de expresiones faciales sintéticas para el reconocimiento de caras efectivamente incrementa la tasa de reconocimiento sin requerir algoritmos complejos para el reconocimiento de caras en 3D.
- Reconocimiento de expresiones faciales utilizando análisis de componentes principales kernel [Orozco, 2008]. El trabajo presenta una metodología para el reconocimiento de expresiones faciales con análisis de componentes principales kernel, la base de datos como herramienta de prueba. El método utiliza una función kernel que mapea los datos del espacio característico original a uno de mayor dimensionalidad, de esta forma un problema de origen no lineal se traslada a uno lineal y puede resolverse linealmente, además los métodos basados en kernel pueden reducir el número de parámetros usados para la clasificación, este método es comparado con el análisis de componentes principales y es puesto a discusión donde los porcentajes de acierto encontrados con la base de datos son mayor al 90%.

La transformada de Wavelet es una técnica utilizada en el procesamiento de señales, siendo dicha señal para este caso, una imagen. La cual definiremos como una señal bidimensional, es decir que está conformada por puntos llamados píxeles que se encuentran ordenados y distribuidos a lo largo y ancho de la imagen [Urbano, 2004]. Es una técnica relativamente nueva, para analizar señales de origen biológico. Por ejemplo, se tienen las señales electroencefalográficas (EEG) a las que se les puede analizar utilizando la Transformada Wavelet. Un ejemplo de lo que se puede hacer con la TW es la detección de la enfermedad de Alzheimer; esta enfermedad se diagnostica mediante la sintomatología que presentan quienes se sospecha que la tienen.

Se puede diagnosticar esta enfermedad analizando los EEG de los pacientes mediante el análisis de las técnicas utilizando la TW. La Transformada Discreta Wavelet es una técnica efectiva para procesar una señal [Medina, 2007].

Las redes neuronales con Wavelets son un algoritmo de reconocimiento de rostros que son capaces de clasificar de forma positiva ó negativa a las personas. Las Redes Neuronales de Retropropagación combinados con Wavelets tienen un grado de certidumbre apto para imágenes de rostro con cambios de iluminación, expresiones faciales y ángulos de orientación del rostro [Fuentes, 2008].

CAPITULO 2

TRANSFORMADA WAVELET

La Transformada Wavelet (TW), es una técnica utilizada en el procesamiento de señales e imágenes, la cual permite descomponer una señal o imagen en versiones escaladas y trasladadas de una wavelet madre así como para reducir el tamaño de las imágenes y obtener vectores característicos más robustos a cambios de iluminación, expresión facial y postura. La TW da una representación de señales en el plano de tiempo-frecuencia. Algunas de sus aplicaciones son las siguientes: compresión de señales e imágenes, análisis del electroencefalograma, compresión de imágenes de huellas dactilares, detección de cambios en estructuras y señales, detección de ondas sísmicas, reflectometría óptica, etc. Como se observa el rango de aplicaciones es tan amplio que bien merece la pena conocer sobre esta novedosa herramienta [Sánchez, 2008].

2.1 HISTORIA DE LAS WAVELETS

El análisis con Wavelets es un método relativamente nuevo, aunque sus fundamentos matemáticos datan de tiempo atrás con el trabajo de Joseph Fourier en el siglo XIX. Fourier puso los fundamentos, con su teoría del análisis en frecuencia que demostró ser de enorme importancia y muy influyente ahora en el análisis con Wavelets.

La atención de los investigadores cambió gradualmente del análisis basado en frecuencia, al análisis basado en escala, cuando empezó a quedar claro que el acercamiento que mida las fluctuaciones medias a diferentes escalas podría evidenciar menos sensibilidad al ruido. La primera mención registrada de lo que ahora llamamos "Wavelet" fue hecha en 1909, en una tesis por Alfred Haar. El concepto de Wavelet en su forma teórica presente fue propuesto por Jean Moriet y el equipo del centro de física teórica Marseille trabajando bajo las ordenes de Alex Grossmann en Francia [Medina, 2007].

Los métodos de análisis con Wavelet han sido desarrollados principalmente por Y. Meyer y sus colegas que han asegurado la diseminación de los métodos. Los principales algoritmos datan de tiempo atrás del trabajo de Stephane Mallat en 1988. Desde entonces, investigaciones sobre Wavelets se han realizado internacionalmente. Dicha investigación es principalmente activa en los Estados Unidos, donde es encabezada por el trabajo de científicos como Ingrid Daubechies, Ronald Coifman, y Víctor Wickerhauser [Sánchez, 2008]. A continuación, se presentan los principios fundamentales de la transformada Wavelet.

2.2 PRINCIPIOS FUNDAMENTALES DE LA TRANSFORMADA WAVELET

Es importante conocer los conceptos básicos de la teoría Wavelet; cualquier transformada que se aplica a una señal se hace con la finalidad de obtener información de ella, información que no está disponible en el dominio del tiempo.

Cuando se grafica una señal en el dominio del tiempo, se obtiene una representación de la amplitud de la señal; ésta no es una buena representación para el procesamiento de una señal. La información que interesa se encuentra oculta en la frecuencia. El espectro en frecuencia muestra cuáles son las frecuencias que existen en la señal.

La forma en la que se puede encontrar la frecuencia contenida en una señal es mediante la Transformada de Fourier (TF). Es decir, al obtener la TF de una señal en el dominio del tiempo, se consigue la representación de la señal en la frecuencia. Éstas gráficas dicen cuáles son las frecuencias que existen en la señal. Para cada frecuencia se tiene un valor en la amplitud [Medina, 2007].

La transformada de Fourier de una función está definida como (2.1):

$$F(t) = \int_{-\infty}^{\infty} f(t)e^{-jt} dt \quad (2.1)$$

y $f(t)$ es llamada la inversa de de la transformada de Fourier de $F(t)$, que está definida como (2.2):

$$f(t) = \frac{1}{2} \int_{-\infty}^{\infty} F(t)e^{jt} dt \quad (2.2)$$

Las ecuaciones (2.1) y (2.2) son regularmente llamadas como el par de transformadas de Fourier, y son usadas para trazar funciones en el intervalo de menos infinito hasta el infinito en el dominio del tiempo o la frecuencia en una función continua en el dominio inverso. La llave apropiada para la transformada de Fourier esta en la habilidad de examinar una función o forma de onda desde la perspectiva de los dos, dominio de tiempo y frecuencia. Una función dada puede tener 2 modos equivalentes de representaciones: una es en el dominio del tiempo y es llamada $f(t)$, y la otra es en el dominio de la frecuencia y es llamada $F(t)$. La ecuación (2.1) transforma la función del tiempo en el espectro de frecuencia, y (2.1) sintetiza el espectro de la frecuencia para recuperar la función del tiempo [Reyes, 2009].

Existen numerosas clases de Transformadas (Fourier, Hartley, De Laplace, Gabor, Hankel, Hilbert, Wigner, Radon, Wavelet) y cada una de ellas tiene su propia área de aplicación, con sus ventajas y desventajas.

La TF es una transformada reversible, es decir, tiene inversa; se puede ir de la señal original a la señal procesada. Pero, sólo una de ellas está disponible en un tiempo dado. En otras palabras, cuando se tiene la señal en el dominio del tiempo, no está disponible la información en la frecuencia; y cuando se tiene la señal transformada, no está disponible la información en el tiempo.

La TF nos da la información en la frecuencia de la señal, pero no nos dice cuándo en el tiempo los componentes de esa frecuencia existen.

Obviamente esta información no es necesaria cuando se trata con señales estacionarias. Una señal estacionaria es aquella cuyo contenido en la frecuencia no cambia en el tiempo, los componentes en la frecuencia existen en todo el tiempo [Medina, 2007].

2.3 TRANSFORMADA WAVELET CONTINUA CWT

La Transformada Continua Wavelet (CWT) se desarrolló como una alternativa para resolver el problema de la resolución de la TF [Medina, 2007].

La transformada wavelet continua se muestra en la ecuación 2.3.

$$C(a, b) = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt \quad (2.3)$$

donde C es un conjunto de coeficientes wavelet y $\psi_{a,b}$ representa un conjunto de versiones escaladas y trasladadas de una función wavelet denominada wavelet madre. Estas funciones están definidas por la ecuación 2.4.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2.4)$$

donde a y b son parámetros de escala y posición respectivamente. El parámetro de escala a está relacionado con la frecuencia de la señal; valores pequeños de a corresponden a frecuencias altas (wavelets comprimidas) mientras que valores grandes de a corresponden a frecuencias bajas (wavelets expandidas). Se denomina continua a la TWC debido a que opera en un conjunto continuo de escalas y posiciones, esto es, cada escala y posición posible.

2.4 TRANSFORMADA WAVELET DISCRETA TWD

Los fundamentos de la DWT se remontan a 1976 cuando Croiser, Esteban y Galand planearon una técnica para descomponer señales discretas en el tiempo. Crochiere, Weber y Flanagan hicieron un trabajo similar para codificar señales en ese mismo año. Ellos denominaron su trabajo como *subband coding*. En 1983, Burt definió una técnica muy similar al *subband coding* y lo denominó *pyramidal coding* el cual también es conocido como análisis multiresolución. Después, en 1989 Vatterli y Le Gall hicieron algunas mejoras al *subband coding*, eliminando algunas redundancias existentes en el *pyramidal coding*.

Calcular coeficientes wavelets para todas las posibles escalas requiere de una gran cantidad de cálculos y genera una cantidad enorme de información.

Si únicamente elegimos un subconjunto de escalas y posiciones, por ejemplo, muestras tomadas en potencias de 2 (escalas y posiciones diádicas), entonces nuestro análisis será mucho más eficiente, esto se realiza mediante la transformada wavelet discreta (TWD) [Fuentes, 2008]. Matemáticamente la TWD se define como se muestra en la ecuación 2.5:

$$C_{m,n}(a,b) = a_0^{-m/2} \int_{-\infty}^{\infty} f(t) \psi(a_0^{-mt} - nb_0) dt, \quad (2.5)$$

donde m y n toman valores enteros, $a_0 > 1$ y $b_0 > 0$.

El proceso de descomposición mediante la TWD puede ser iterativo, esto se logra filtrando los coeficientes de aproximación en forma recursiva. A esto se le denomina árbol de descomposición. En cada proceso de filtrado se genera un nivel de descomposición. En la Fig. 2.1 se ilustra el proceso de descomposición en tres niveles, en cada nivel se producen los respectivos coeficientes de aproximación y detalle. Conforme mayor sea el nivel de descomposición, el número de coeficientes es menor y su espectro corresponde a bandas de frecuencia más bajas.

En teoría, la descomposición puede continuar indefinidamente, sin embargo, en la práctica es indispensable elegir un número de niveles de descomposición adecuado a nuestras necesidades.

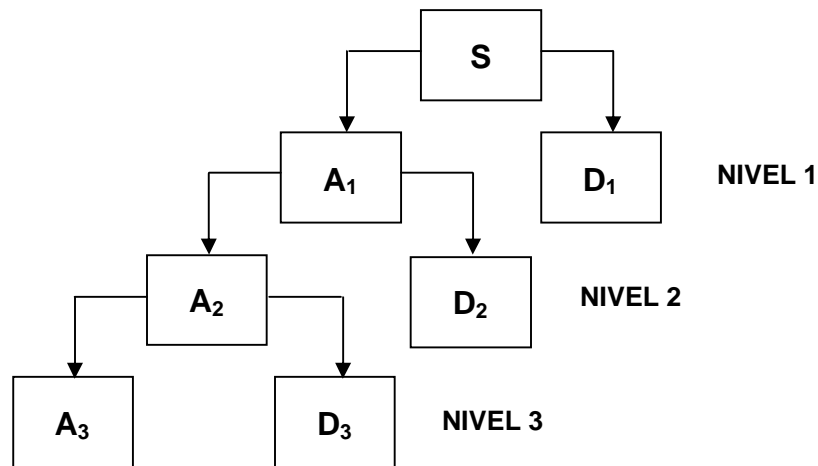


Figura 2.1. Árbol de descomposición por wavelets en tres niveles.

Para el caso bidimensional, gracias a la teoría de variables separables, la TWD puede calcularse aplicando la TWD unidimensional a las filas y posteriormente a las columnas. En la Fig. 2.2 se ilustra el esquema de filtrado para implementar la TWD de una imagen. La aplicación de la TWD a una imagen produce cuatro matrices de coeficientes wavelet, una de aproximación y tres de detalles (horizontal, vertical y diagonal).

La matriz de coeficientes de aproximación contiene los componentes en bandas de baja frecuencia de la imagen original mientras que las matrices de detalles los componentes en bandas de alta frecuencia.

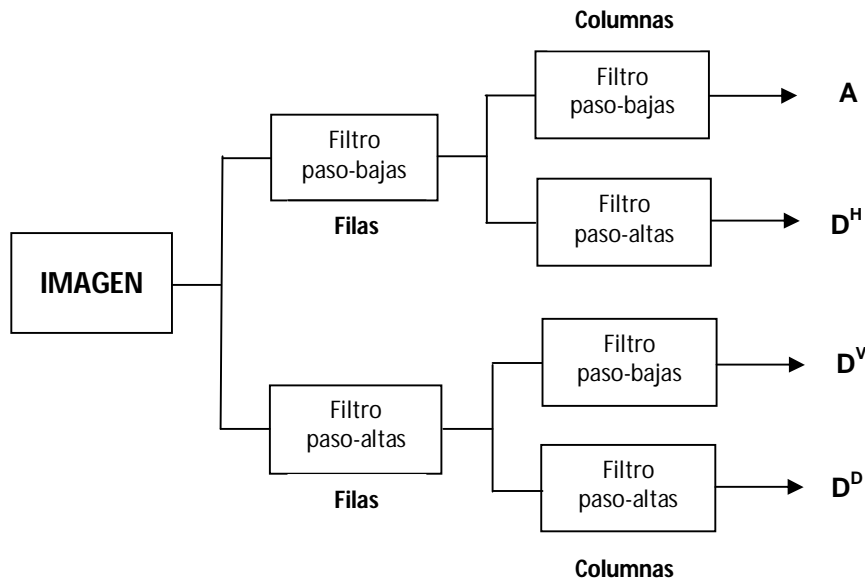


Figura 2.2. Esquema de filtrado para calcular la TWD de una imagen.

La fase de descomposición está determinada por las siguientes expresiones:

$$a_{j+1}(p) = \sum_{n=-\infty}^{+\infty} h(n - 2p)s_j(n), \quad (2.6)$$

donde a_j es el coeficiente de aproximaciones de nivel "j", d_j corresponde al coeficiente de detalles de nivel j . Los coeficientes del filtro están denotados por h y g .

$$d_{j+1}(p) = \sum_{n=-\infty}^{+\infty} g(n - 2p)s_j(n). \quad (2.7)$$

En la Fig. 2.3 se presenta la descomposición de la imagen de un rostro en sus matrices de coeficientes de aproximación (A), detalles horizontales (D^H), detalles verticales (D^V) y detalles diagonales (D^D).

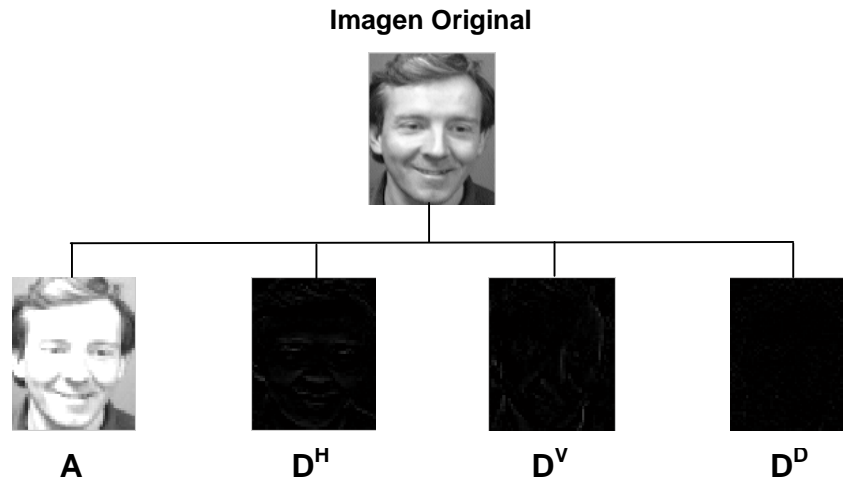


Figura 2.3. Descomposición de la imagen de un rostro en matrices de aproximación y detalles.

2.5 LAS FUNCIONES WAVELETS

Las funciones Wavelets se pueden clasificar de acuerdo a la manera en que procesan las señales, siendo dicha señal para este caso, una imagen. La cual definiremos como una señal bidimensional es decir, que está conformada por puntos llamados píxeles que se encuentran ordenados y distribuidos a lo largo y ancho; partiendo de este conocimiento, se puede realizar su procesamiento de dos maneras: En el primer caso se puede tomar cada dimensión de la imagen (largo y ancho) de manera independiente y se procesa cada dimensión como si fuera una señal unidimensional esta consideración le da origen a los Wavelets de función de variables separables como lo son los Wavelets de Haar y Daubechies. La segunda manera de procesar la imagen, es tomando la señal en forma bidimensional, es decir que no se puede descomponer en señales correspondientes a cada dimensión, y es exactamente aquí donde surgen los Wavelets con funciones de variables no separables como lo son QMF (*Quadratic Media Filter*) y Gabor entre otros [Urbano, 2004].

2.6 FAMILIAS WAVELET

La elección de la familia wavelet más adecuada para la descomposición de un rostro juega un papel muy importante en la obtención de una representación wavelet de los rostros más robusta.

Existen muchas familias de funciones wavelet que han mostrado su gran utilidad en diversas aplicaciones.

A continuación se describen brevemente las características de las familias wavelet propuestas para realizar la descomposición de los rostros [Fuentes, 2008]. Las familias de Wavelets que han demostrado ser más útiles para el procesamiento de señales son las siguientes [Medina, 2007].

2.6.1 Familia Haar

Cualquier discusión sobre Wavelets debe de empezar con la Wavelet Haar, por ser la más simple y sencilla, esta Wavelet es discontinua y se parece a una función “paso”. Es la primera de todas las wavelets.

Es una wavelet continua que se asemeja a una función escalón unitario. Esta función es equivalente a la Wavelet Daubechies db1 [Sánchez, 2008].

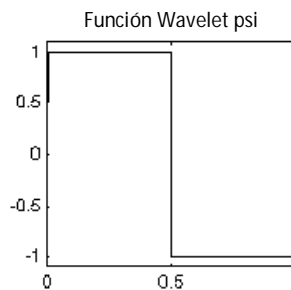


Figura 2.4. Wavelet de Haar.

2.6.2 Familia Daubechies

Ingrid Daubechies, una de las personalidades más importantes en el mundo de la investigación de Wavelets, inventó lo que se conoce como Wavelets ortonormal sólidamente apoyadas, haciendo con esto posible el análisis discreto con Wavelets. El nombre de las Wavelets de la familia Daubechies se escribe dbN, donde N es el orden y db corresponde al apellido de la Wavelet. La Wavelet db1, como se mencionó anteriormente es igual a la Wavelet de Haar [Sánchez, 2008].

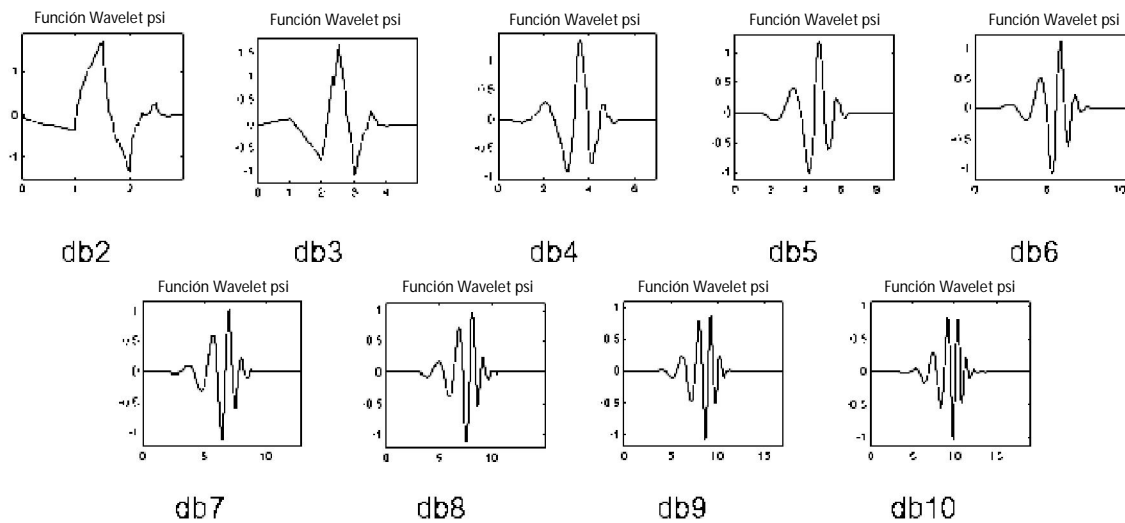


Figura 2.5. Familia de Wavelets Daubechies.

2.6.3 Familia Coiflets

Estas Wavelets fueron desarrolladas por I. Daubechies a petición de R. Coifman. Son Wavelets de soporte compacto con $2N$ momentos iguales a 0 mientras que su función de escala tiene $2N - 1$ momentos iguales a 0, ambas funciones tienen un soporte de magnitud $6N - 1$. Los nombres de la familia Coiflets se escriben como *coifN*, donde N es el orden y *coif* es el sobrenombre [Sánchez, 2008].

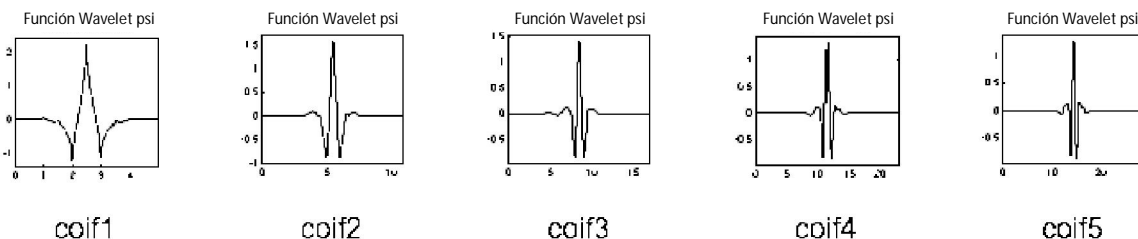


Figura 2.6. Familia de Wavelets Coiflets.

2.6.4 Familia Biortogonal

Esta familia de Wavelets tiene la propiedad de poseer fase lineal, que es necesaria para la reconstrucción de señales e imágenes. Estas utilizan dos Wavelets diferentes, una para descomposición y otra para reconstrucción en lugar de usar solo una para los dos procesos como las otras familias [Sánchez, 2008].

Existen dos variantes:

- 1) *biorthogonal*, las cuales se escriben como *biorNr.Nd*
- 2) *reverse biorthogonal*, escritas como *rbioNd.Nr*. *Nr* y *Nd* que son ordenes para reconstrucción y descomposición respectivamente. Interesantes propiedades se derivan de esto.

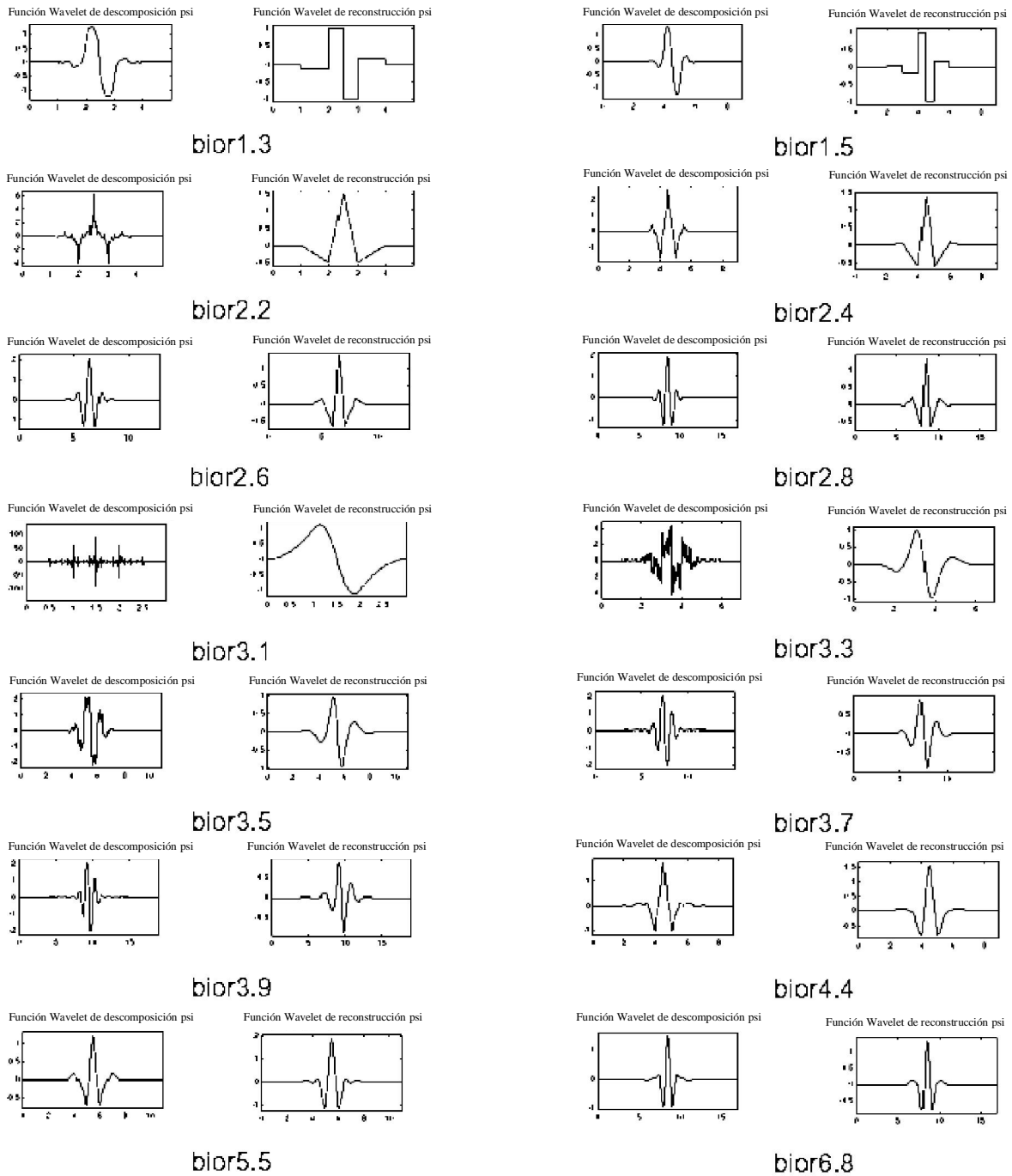


Figura 2.7. Familia de Wavelets Biortogonales.

2.6.5 Discrete Meyer

Esta otra familia de las Wavelet se encuentra definida en el dominio de la frecuencia. Es la versión discreta de la wavelet de Meyer, su alias es *dmey*.

2.7 APLICACIÓN DE WAVELETS EN IMÁGENES

La naturaleza del análisis Wavelet, descomponiendo señales en distintas bandas frecuenciales, es especialmente adecuado cuando tratamos con imágenes. Una característica de las Wavelets es que se adaptan al nivel de detalle de la señal, gracias a que los coeficientes de frecuencias mas altas cubren un menor espacio de la señal, lo que representa otro factor a favor de su uso en compresión de imagen. La imagen procesada se pasa por un filtro pasa bajas y por un filtro pasa altas simultáneamente. Dependiendo del Wavelet a utilizar se usan diferentes tipos de filtros pasa bajas y filtros pasa altas. Un ejemplo de Wavelets utilizados para el reconocimiento de rostros son los Wavelets de Daubechies, este se encuentra constituido por un Wavelet llamado madre y es la forma que tomarán los filtros pasa bajas y pasa altas.

Una vez que se realiza el filtrado, se conservan solamente los resultados del filtro pasa bajas para volver a filtrar la imagen con ambos filtros. El número de veces que se repite esta operación es lo que se conoce como el nivel de descomposición del Wavelet.

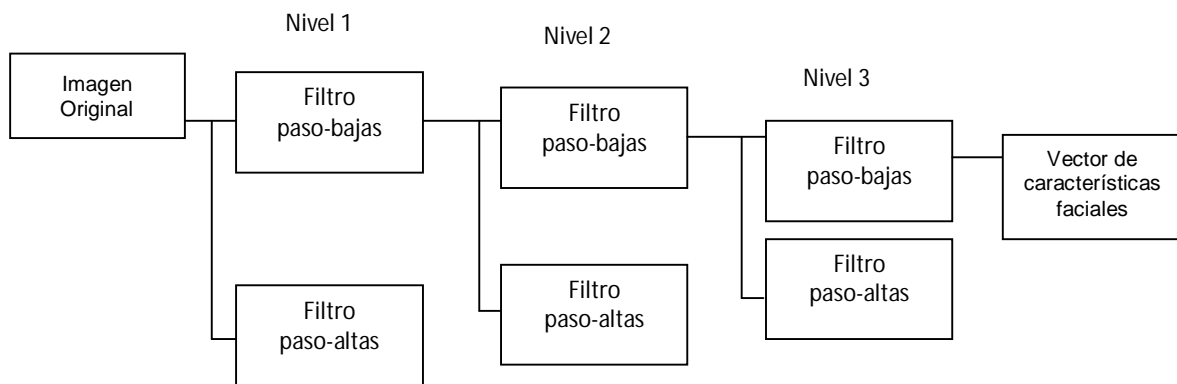


Figura 2.8. Esquema de extracción de características mediante la TWD.

En la Fig. 2.8 se muestra la descomposición en 3 niveles de una imagen, el vector característico es derivado de los componentes de baja frecuencia del tercer nivel de descomposición. La descomposición recursiva de la imagen de una imagen conduce a la reducción de las dimensiones en un cuarto de su tamaño por cada nivel de descomposición [Fuentes, 2008].

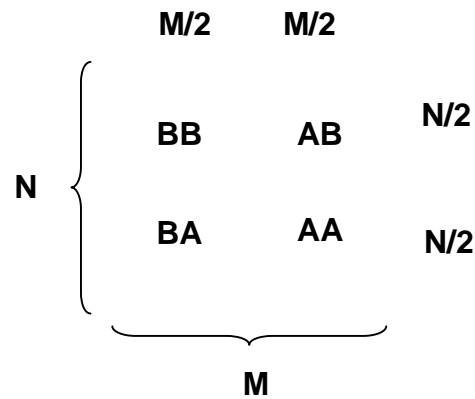


Figura 2.9. Subimágenes resultantes del proceso de descomposición.

Las cuatro subimágenes resultantes se suelen representar juntas formando una imagen de las mismas dimensiones que la imagen original, de manera que quede como se indica en la figura 2.9. También suele usarse una notación más habitual donde H indica la parte de altas frecuencias y L la de bajas frecuencias. Estas letras corresponden a las iniciales en inglés de *high* (*alta*) y *low* (*baja*). Pese a esto, merecía la pena señalar esta notación ya que es muy común en la bibliografía (posiblemente por el hecho de que los coeficientes de las funciones Wavelet siempre se suelen dar como h_k para la función de escala, que se corresponde por tanto con la parte de bajas frecuencias de la señal).

Cuando se desea realizar varios niveles de descomposición, se suele hacer volviendo a tomar la parte de referencia de la señal (la imagen escalada), para aplicar de nuevo este proceso y así de forma sucesiva, hasta llegar a un cierto nivel de descomposición [Sánchez, 2008].

Suponiendo que las dimensiones de esta imagen cuadrada son 512x512 píxeles, y los algoritmos de codificación de imagen Wavelet más eficientes suelen descomponer las imágenes de este tamaño aplicando entre cuatro y seis niveles, de manera que la subimagen de referencia final alcanza tamaños de 32, 16 u 8 píxeles por lado.

En la Fig. 2.10 se ilustra como la imagen se divide en cuatro donde el cuadrante superior izquierdo representa el resultado del filtro pasa bajas. Las frecuencias bajas, según la teoría de los Wavelets, contienen información principal de la imagen y representan la tendencia de la misma. Los tres cuadrantes restantes representan las fluctuaciones.



Figura 2.10. Ejemplo de descomposición de una imagen empleando Wavelets de Daubechies.

Posteriormente, se divide la nueva tendencia de la misma forma que la imagen original. La sumatoria de todos los coeficientes, tanto de la tendencia como de las fluctuaciones, permite recuperar la imagen original. Una vez descompuesta la imagen, la resultante es una matriz de dimensiones $M \times N$. Los Wavelets son casi simétricos y fueron propuestos por *Daubechies*.

Con este tipo de Wavelet se filtra la imagen varias veces para obtener las características principales, es decir la tendencia y las fluctuaciones. La imagen se divide en cuatro secciones, en la esquina superior izquierda queda representada la tendencia y es quien conserva las características principales de la imagen.

La esquina superior derecha representa el detalle horizontal de la misma imagen. En la parte inferior del lado izquierdo se representa el detalle vertical y finalmente en la parte inferior derecha queda representado el detalle diagonal.

En general, el número de transformaciones a realizar depende de varios factores, como son:

- El nivel de compresión deseado, en principio cuanto mayor nivel de compresión se desee, mas niveles de descomposición se llevan a cabo.
- El tamaño de la imagen original, lógicamente las imágenes mas grandes necesitarán más niveles de descomposición.
- El tamaño de los filtros es otro factor importante, que limita el nivel de descomposición de la imagen. Si los filtros son más grandes que la señal a la que se aplican no obtendremos grandes ventajas, ya que estaremos obligados a realizar extensiones artificiales de la señal, que serán más grandes que la propia señal [Sánchez, 2008].

CAPITULO 3

REDES NEURONALES ARTIFICIALES

Las redes neuronales artificiales (RNA) son un paradigma matemático o computacional inspirado en las redes neuronales biológicas.

Las redes neuronales intentan simular, al menos parcialmente, la estructura y funciones del cerebro y sistema nervioso de criaturas vivientes. Las redes neuronales también tienen su analogía con la naturaleza. En una red neuronal, la naturaleza se refiere a características como estructura física (o estructura física inicial en el caso de uno de los pocos modelos de red que modifican su estructura durante el entrenamiento), modelo de aprendizaje usado durante el entrenamiento, y el conjunto inicial de pesos de conexión. Una red neuronal con una naturaleza efectiva inherentemente facilita en gran medida la tarea de desarrollar un sistema que pueda resolver eficientemente un problema particular [Guzmán, 1997].

Una de las características atractivas de las redes neuronales es su capacidad para auto-adaptarse a condiciones ambientales especiales cambiando sus fuerzas de conexión o pesos o cambiando su estructura. Las redes neuronales son consideradas a veces como modelos grotescos del cerebro humano. Este punto de vista es exagerado, debido a que el cerebro humano no es totalmente conocido y su comportamiento es muy complejo. Por esto, es más razonable comparar capacidades de redes neuronales con sistemas nerviosos más simples de animales primitivos, tales como insectos, que tienen la habilidad de auto-adaptarse a ambientes más complejos [Guzmán, 1997].

Un gran número de investigadores de distintas disciplinas han empleado RNA para resolver una gran variedad de problemas que incluyen reconocimiento de patrones, predicción, optimización, memoria asociativa y control.

Además, las RNA han jugado un papel importante en distintas aplicaciones, en particular han mostrado un rendimiento muy convincente en problemas de reconocimiento de patrones [Fuentes, 2008].

3.1 MODELO DE NEURONA ARTIFICIAL

El primer modelo de neurona artificial fue la unidad de umbral binario propuesta por McCulloch y Pitts, la cual se ilustra en la Fig. 3.1 [Báez, 2006].

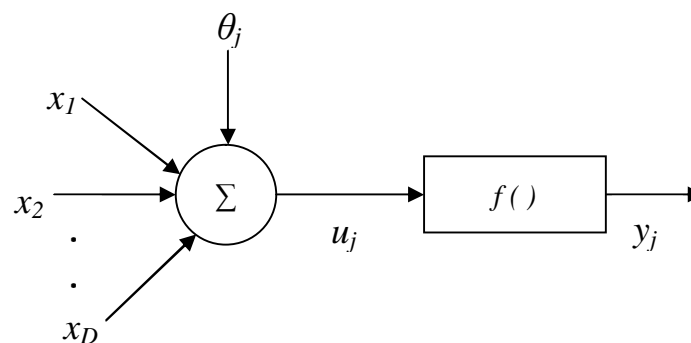


Figura 3.1. Modelo de neurona de McCulloch y Pitts.

Este modelo de neurona se caracteriza por la descripción funcional de la *red de conexiones* y la *activación de neuronas* [Báez, 2006]. Cada neurona recibe valores de entrada x_i , los cuales son propagados desde otras neuronas en la red a través de una red de conexiones unidireccionales. Cada neurona tiene asociada un *peso sináptico* denotado por w_{ij} , el cual determina el efecto de la j -ésima neurona sobre la i -ésima neurona; si el peso es positivo, la conexión se denomina excitatoria; si es negativo, se denomina inhibitoria. Las entradas de la i -ésima neurona son acumuladas junto con una entrada externa θ_i (comúnmente denominada *bias*) para obtener el valor neto u_i . Posteriormente, el valor neto u_i es transformado por medio de una *función de activación* no lineal f para obtener y_i . Las funciones de activación más comúnmente utilizadas son las funciones escalón, rampa, sigmoideal y gaussiana.

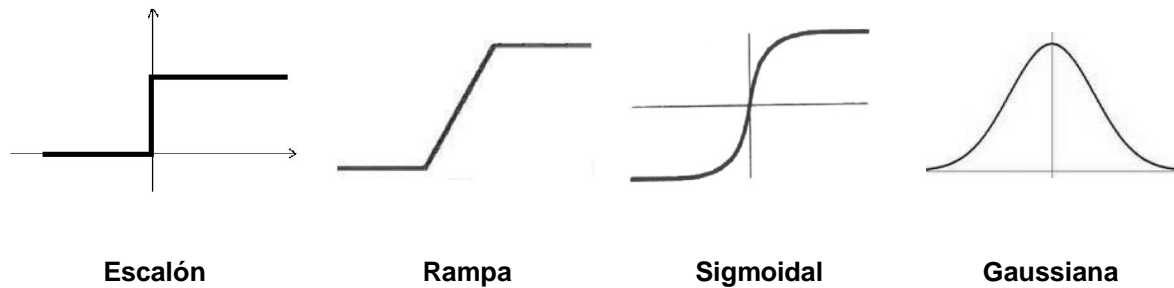


Figura 3.2. Ejemplos de funciones de activación.

La función sigmoideal es por mucho la que se utiliza con mayor frecuencia y esta definida por la ecuación 3.1:

$$f(u_i) = \frac{1}{1 + e^{-(u_j - \theta_j)/\sigma}} \quad (3.1)$$

donde $u_j = \sum_{i=1}^N w_{ji} x_i$. Las conexiones de las RNA están representadas matemáticamente por una *función base* $u(w_j, x)$, donde w_j es el vector de pesos y x el vector de entrada. Esta función base tiene dos formas típicas:

- **Funciones de Base Lineal (LBF)**: son funciones definidas por un hiperplano (funciones base de primer orden). El valor neto de la *j-ésima* neurona es una combinación lineal de las entradas, es decir como lo muestra la ecuación 3.2:

$$u_j = u(w_j, x) = \sum_{i=1}^N w_{ji} \quad (3.2)$$

donde w_{ji} es el peso de conexión de la *j-ésima* neurona y su *i-ésima* entrada, y N es el número de entradas a la *j-ésima* neurona.

- **Funciones de Base Radial (RBF):** son funciones definidas por una hiperesfera que involucra una función base de segundo orden (no lineal). El valor neto u_j de la j -ésima neurona es una función de la distancia entre un patrón de referencia w_j y la entrada x , como se muestra en la ecuación 3.3:

$$u_j = u(w_j, x) = \sqrt{\sum_{i=1}^N (x_i - w_{ji})^2} \quad (3.3)$$

3.2 TOPOLOGÍA DE LAS RNA

Las RNA pueden ser vistas como un grafo dirigido y ponderado, donde los nodos representan neuronas y las aristas dirigidas (con pesos) son las conexiones entre las salidas y entradas de las neuronas.

La forma en la cual las neuronas de una red neuronal se encuentran estructuradas está íntimamente ligada con el algoritmo de aprendizaje utilizado para entrenar la red. Basado en el patrón de conexión (topología o arquitectura) es posible agrupar las redes neuronales en dos categorías:

1. **Hacia adelante:** la señal fluye estrictamente de la entrada a la salida sin existir ciclos ni conexiones entre neuronas de la misma capa.
2. **Recurrentes:** ocurren ciclos debido a que existen conexiones de retroalimentación.

En términos generales, las redes neuronales de propagación hacia adelante son estáticas, esto es, producen un sólo conjunto de valores de salida a partir de una entrada en lugar de una secuencia de estos. Puede decirse que las redes hacia adelante son carentes de memoria en el sentido que su respuesta es independiente a estados anteriores de la red.

Las redes recurrentes por su parte son sistemas dinámicos ya que cada vez que un nuevo patrón de entrada es presentado, las salidas de cada neurona son calculadas. Debido a su retroalimentación, las entradas de cada neurona son modificadas, lo cual produce que la red entre a un nuevo estado.

3.2.1 Aprendizaje de las RNA

Una de las propiedades más importantes de las RNA es su capacidad de aprender de su medio para mejorar su rendimiento. El proceso de aprendizaje de una RNA puede ser visto como el problema de adaptar la arquitectura de la red y los pesos de conexión para que puedan realizar eficientemente una tarea específica. Los paradigmas de aprendizaje se pueden clasificar en dos grandes categorías:

- **Con supervisión.** Este tipo de aprendizaje incorpora una referencia externa o salida deseada. Se da un conjunto de pares de ejemplos (x, y) , $x \in X$ y $y \in Y$ y el objetivo es encontrar una función f en las clases de funciones permitidas que corresponda a los ejemplos. En otras palabras, se infiere el mapeo implícito de los datos. El paradigma de aprendizaje supervisado puede ser visto como el aprendizaje por medio de un maestro en forma de una función que provee continua retroalimentación de acuerdo a la respuesta actual.

Ejemplos de redes que utilizan este paradigma son: el perceptrón simple, ADALINE, perceptrón multicapa y la memoria asociativa bidireccional [Rodríguez, 2005].

- **Sin supervisión.** El paradigma de aprendizaje sin supervisión (o autoorganizado) no necesita de una referencia o conjunto de salidas deseadas. La adaptación de los pesos en el aprendizaje sin supervisión depende de la tarea que se desea realizar y de las propiedades implícitas de nuestro modelo, sus parámetros y variables observadas.

Los problemas que pueden ser resueltos por este paradigma de aprendizaje son en general de estimación. Sus aplicaciones incluyen clustering, estimación de distribuciones estadísticas, compresión y filtrado. Entre las redes que utilizan esta categoría de aprendizaje se encuentran las memorias asociativas, las redes de Hopfield, la máquina de Boltzmann y la máquina de Cauchy, las redes de aprendizaje competitivo, las redes de Kohonen o mapas autoorganizados y las redes de resonancia adaptativa [Rodríguez, 2005].

- **Aprendizaje híbrido.** Es una mezcla de las anteriores. Unas capas tienen un aprendizaje supervisado y otras capas un aprendizaje de tipo no supervisado [Rodríguez, 2005].
- **Aprendizaje reforzado.** Es un aprendizaje con características de tipo supervisado y auto-organizado. No se proporciona una salida deseada, pero si se le indica a la red en cierta medida el error que comete, aunque es un error global [Rodríguez, 2005].

Las redes neuronales de aprendizaje supervisado representan la corriente dominante en el desarrollo de redes neuronales artificiales para autenticación biométrica. El aprendizaje supervisado involucra dos fases: fase de reconocimiento y fase de entrenamiento. El objetivo de la fase de reconocimiento es determinar a cual clase pertenece un patrón. Los valores de salida son funciones de los valores de entrada y los pesos de la red. Por otra parte, en la fase de aprendizaje los pesos son adaptados de tal manera que con mayor probabilidad los patrones (aprendidos y/o no aprendidos) sean clasificados correctamente.

3.2.2 Red Neuronal Perceptrón Multicapa

La red neuronal perceptrón multicapa (MLP, por sus siglas en inglés) es una red hacia adelante con aprendizaje supervisado que representa uno de los modelos neuronales más populares. Existen dos tipos básicos de MLP: LBF y RBF.

El perceptron multicapa emplea en unidad computacional cualquier función de umbral o la función sigmoideal. El perceptron multicapa puede formar fronteras complejas de decisión arbitrariamente y representar cualquier función booleana. El desarrollo del algoritmo de aprendizaje de retro-propagación para determinar los pesos en una red perceptron multicapa ha hecho a esta red la más popular dentro de los investigadores y usuarios de redes neuronales [Sánchez, 2004].

El Perceptron Multicapa es una ampliación y generalización del Perceptron de dos capas. Conserva la estructura del Perceptron simple, pero posee la ventaja de tener múltiples capas ocultas como se muestra en la Fig. 3.3. Esto le permite establecer regiones de decisión mucho más complejas, aprender mayor número de clases y solucionar problemas que el Perceptron simple, sin capa ocultas, no puede resolver [Caceres, 2002].

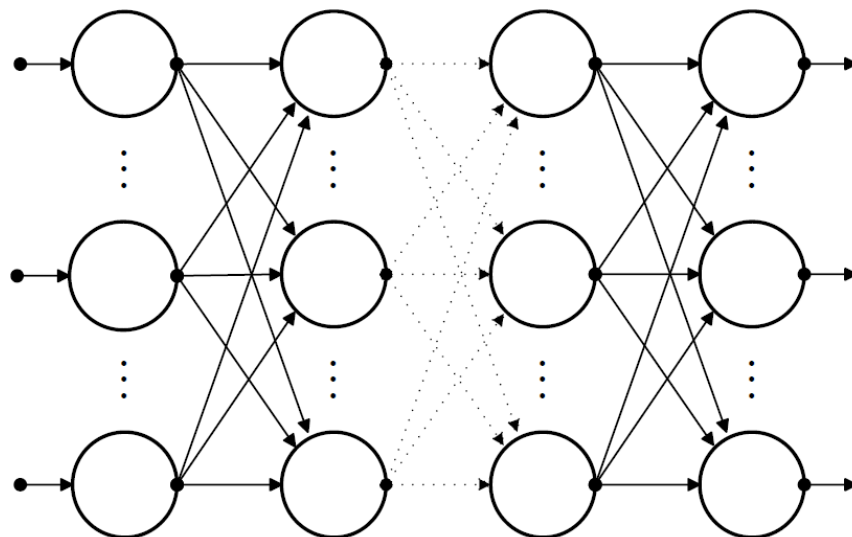


Figura 3.3. Arquitectura de una red MLP con propagación hacia adelante.

Las capacidades del Perceptron Multicapa con una única neurona de salida se pueden observar en la Tabla 3.1 [Lippman, 1987]. Se ve en general, que no se requieren más de cuatro capas para resolver un problema dado, ya que es suficiente para generar regiones de decisión arbitrariamente complejas.

En ciertos problemas, puede mejorarse el aprendizaje mediante la variación de la cantidad de neuronas ocultas por capas y/o con la extensión de la función de activación [Lippman, 1987].

Además, un número excesivo de neuronas en alguna capa oculta, puede generar ruido en el aprendizaje y un número adecuado, aunque sea redundante, permite que la red posea mayor tolerancia a fallos. No existe aún, un método analítico para determinar la cantidad óptima de neuronas en las capas ocultas, por ello, se requiere de un análisis estadístico del comportamiento del sistema con diferentes cantidades de neuronas y que permita determinar la cantidad óptima para un cierto problema dado [Lippman, 1987].

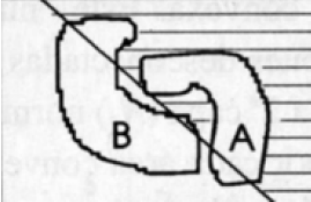
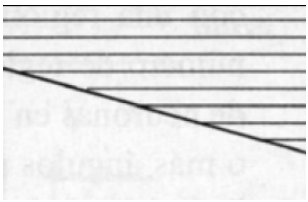
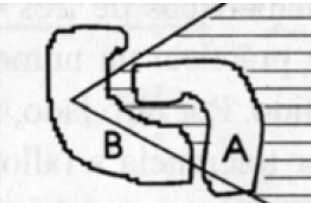
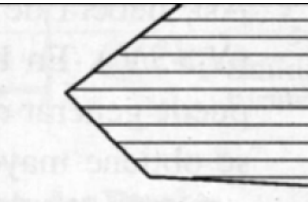
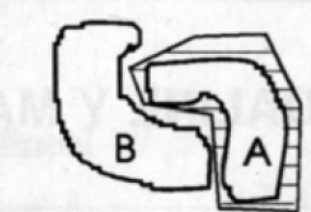
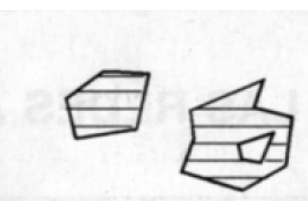
ESTRUCTURA	REGIONES DE DECISION	CLASES CON REGIONES MEZCLADAS	FORMAS DE REGIONES MAS GENERALES
2 CAPAS	MEDIO PLANO LIMITADO POR UN HIPERPLANO		
3CAPAS	REGIONES CERRADAS O CONVEXAS		
4 CAPAS	ARBITRARIA COMPLEJIDAD LIMITADA POR EL NUMERO DE NEURONAS		

Tabla 3.1. Regiones de decisión con diferentes números de capas y de neuronas con una MLP.
Fuente: (Lippman, 1987).

3.2.3 Funcionamiento de La MLP.

El funcionamiento del Perceptron Multicapa es idéntico al Perceptron pero generalizado para múltiples conexiones. Una de las formulaciones más usadas para realizar el entrenamiento de una MLP es la formulación basada en aproximación, la cual puede ser vista como una aproximación/regresión para un conjunto de entrenamiento dado, denotado como la ecuación 3.4 lo muestra:

$$[X, D] = \{[x^{(1)}, d^{(1)}], [x^{(2)}, d^{(2)}], \dots, [x^{(N)}, d^{(N)}]\} \quad (3.4)$$

donde N es el número de pares de entrenamiento, $x^{(1)}, \dots, x^{(N)}$ son los patrones de entrada y $d^{(1)}, \dots, d^{(N)}$ son las salidas deseadas. El objetivo del entrenamiento de la red es encontrar pesos sinápticos óptimos de tal manera que el error entre la salida deseada y la salida actual de la red sea mínimo.

En la Fig. 3.4 se ilustra la arquitectura de una red neuronal multicapa. Las neuronas de la capa de entrada están representadas por $x_i = a_i(0)$ y las de la capa de salida por $y_i = a_i(L)$, donde L es el número de capas. La función de activación denotada por $f(n)$ es comúnmente una función sigmoideal. Para entrenar redes neuronales de aproximación, como lo muestra la ecuación 3.5 los pesos

$$\{w_{ji}(l); i = 1, \dots, M_l; l = 1, \dots, L\} \quad (3.5)$$

donde M_l es el número de nodos (incluyendo el *bias*) en la l -ésima capa y L es el número de capas, se adaptan para minimizar el error de mínimos cuadrados entre la salida deseada y la salida actual de la red neuronal, esto es, como lo indica la ecuación 3.6 :

$$E = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^K [d_k^{(n)} - a_k^{(n)}(L)]^2 \quad (3.6)$$

donde $a_k^{(n)}(L)$ es la salida de la neurona k en la capa de salida, N es el número de patrones de entrada y $K = M_L$ es la dimensión del espacio de salida. Los pesos sinápticos de una red neuronal perceptrón multicapa pueden ser adaptados mediante el algoritmo de aprendizaje de retropropagación de errores o backpropagation (BP).

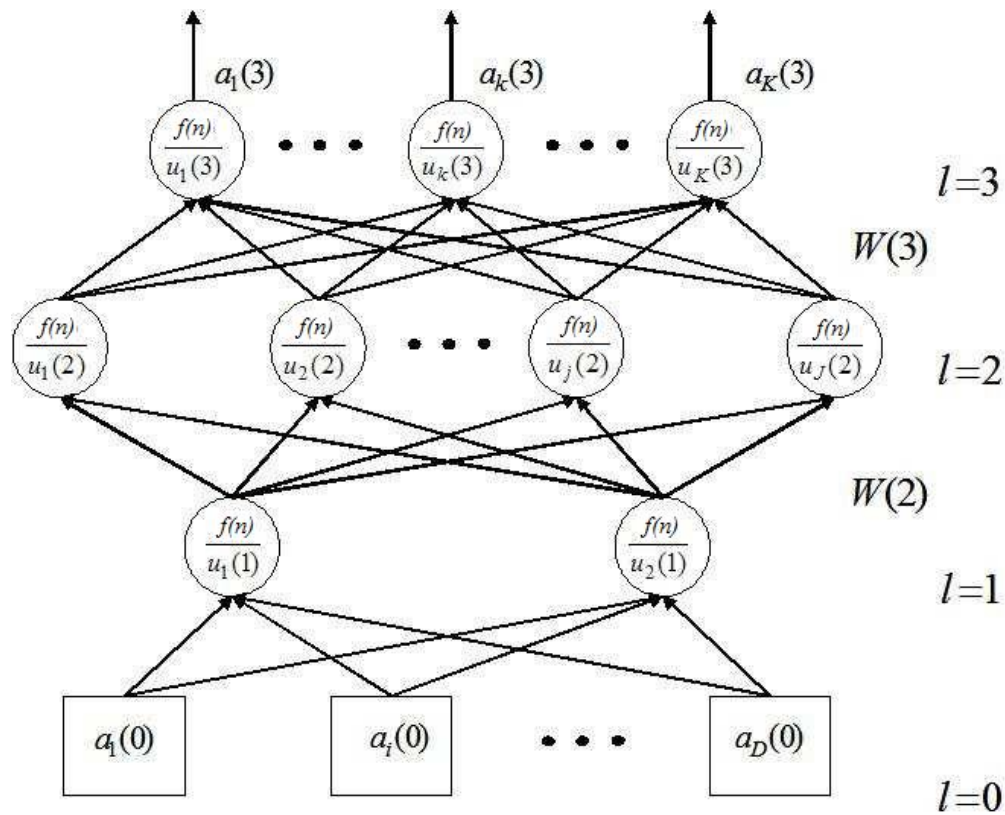


Figura 3.4. Arquitectura de una red neuronal de retropropagación basado en un Perceptrón Multicapa.

Una red neuronal MLP está caracterizada por las siguientes las ecuaciones 3.7 y 3.8 respectivamente:

$$u_j(l) = \sum_{i=1}^{M_{l+1}} w_{ji}(l) a_i(l-1) + \theta(l) = \sum_{i=0}^{M_{l-1}} w_{ji}(l) a_i(l-1) \quad (3.7)$$

$$a_j(l) = f(u_j(l)) \quad 1 \leq j \leq M_1 \text{ y } 1 \leq l \leq L. \quad (3.8)$$

Note que el termino *bias* $\theta(l)$ ha sido integrado a la suma tal que $w_{j_0}(l) = \theta(l)$ y $a_0(l) = 1 \forall l = 0, \dots, L - 1$.

El método de aprendizaje está basado en dos fases de propagación:

- **Hacia adelante.** En la fase de propagación hacia adelante los patrones de entrada son propagados por la red a través de sus múltiples capas hasta generar la salida. Obtenido el vector de salida, este es comparado con el deseado, mediante los errores cuadráticos cometidos.
- **Hacia atrás.** En la fase de propagación del error hacia atrás, debido a que se desconoce la salida deseada para las neuronas pertenecientes a capas ocultas, el error es propagado desde la capa de salida hacia las capas ocultas modificando los pesos asociados a dichas conexiones hasta llegar a la capa de entrada. Así, se habrán modificado todos los pesos de la red, ajustándose al patrón que recientemente le fue presentado. Si se realiza nuevamente una propagación hacia adelante con el mismo patrón, la red habrá disminuido su error en la salida, en comparación con el error cometido en la propagación anterior [Caceres, 2002].

Así sucesivamente, a la red se le podrán ir presentando nuevos patrones, para que vaya autoajustando sus pesos con el fin de aprenderlos y de cometer un error que sea mínimo.

3.2.4 Algoritmo Backpropagation.

El algoritmo más prominente para entrenar redes multicapa es el llamado algoritmo de retropropagación de errores o backpropagation (BP) [McClelland, 2001]. El objetivo es entrenar los pesos $w_{ij}(l)$ para minimizar el error cuadrático medio (MSE). La regla de aprendizaje está determinada por la ecuación 3.9:

$$w_{ij}^{(n+1)}(l) = w_{ij}^{(n)}(l) + \Delta w_{ij}^{(n)}(l) \quad (3.9)$$

donde $n = 1, 2, \dots, N$ es el n -ésimo patrón de entrenamiento. La derivación del algoritmo backpropagation está motivada por la regla de la cadena:

$$\Delta w_{ij}^{(n)}(l) = -\eta \frac{\partial E}{\partial w_{ij}^{(n)}(l)}, \quad (3.10)$$

$$= -\eta \frac{\partial E}{\partial a_j^{(n)}(l)} \cdot \frac{\partial a_j^{(n)}(l)}{\partial w_{ij}^{(n)}(l)}, \quad (3.11)$$

$$= -\eta \delta_j^{(n)}(l) f'(u_j^{(n)}(l)) \frac{\partial u_j^{(n)}(l)}{\partial w_{ij}^{(n)}(l)}, \quad (3.12)$$

donde η es el factor de convergencia (o tamaño de paso) y el error $\delta_j^{(n)}(l)$ está definido por la ecuación 3.13, como:

$$\delta_j^{(n)}(l) = - \frac{\partial E}{\partial a_j^{(n)}(l)}. \quad (3.13)$$

La señal de error $\delta_j^{(n)}(l)$ puede ser obtenida en forma recursiva por retropropagación:

- **Capa Superior.** El valor inicial (de la capa superior) $\delta_j^{(n)}(L)$ puede obtenerse fácilmente utilizando la ecuación 3.14 y 3.15 como sigue:

$$\delta_j^{(n)}(L) = - \frac{\vartheta E}{\vartheta a_j^{(n)}(L)} \quad (3.14)$$

$$= d_j^{(n)} - a_j^{(n)}(L). \quad (3.15)$$

- **Formula Recursiva.** La formula recursiva general de retropropagación para la señal de error $\delta_j^{(n)}(l)$ puede ser derivada utilizando las ecuaciones 3.16 y 3.17 de la siguiente manera:

$$\delta_j^{(n)}(l) = - \frac{\vartheta E}{\vartheta a_j^{(n)}(l)}, \quad (3.16)$$

$$= - \sum_{k=1}^{M_{l+1}} \frac{\vartheta E}{\vartheta u_k^{(n)}(l+1)} \frac{\vartheta u_k^{(u)}(l+1)}{\vartheta a_j^{(n)}(l)}, \quad (3.17)$$

en la secuencia de $l = L - 1, \dots, 1$ tenemos que:

$$\delta_j^{(n)}(l) = \sum_{k=1}^{M_{l+1}} \vartheta_k^{(n)}(l+1) f'(u_k^{(n)}(l+1)) \frac{\vartheta u_k^{(n)}(l+1)}{\vartheta a_j^{(n)}(l)}. \quad (3.18)$$

La formula recursiva es la llave para el aprendizaje por retropropagación de errores ya que permite calcular la señal de error una capa inferior $\delta_j^{(n)}(l)$ como una combinación lineal de las señales de error de la capa superior $\vartheta_k^{(n)}(l+1)$.

De esta manera, las señales de error $\delta_j^{(n)}(.)$ son propagadas hacia atrás a través de todas las capas de arriba hacia abajo. Esto también implica que las influencias de una capa superior a una capa inferior (y viceversa) pueden ser afectadas únicamente a través de capas intermedias.

Sustituyendo la ecuación 3.8 en la 3.12 y 3.18 se obtiene la regla de backpropagation como se muestra en la ecuación 3.19:

$$w_{ji}^{(n+1)}(l) = w_{ji}^{(n)}(l) + \eta \delta_j^{(n)}(l) f' \left(u_j^{(n)}(l) \right) a_i^{(n)}(l-1) \quad l = 1, \dots, L, \quad (3.19)$$

donde la señal de error está definida como lo muestra la ecuación 3.20:

$$\delta_j^{(n)}(l) = \sum_{i=1}^{M_{i+1}} \delta_k^{(n)}(l+1) f' \left(u_k^{(n)}(l+1) \right) w_{kj}^{(n)}(l+1) \quad l = L-1, \dots, 1, \quad (3.20)$$

con la condición inicial dada por la ecuación 3.15.

Podemos deducir de la ecuación 3.12 que para llevar a cabo el algoritmo BP es necesaria una función de activación diferenciable.

CAPITULO 4

SISTEMA PROPUESTO

En este capítulo se proponen dos métodos de procesamiento para el reconocimiento de rostros; ambos se basan en la red neuronal perceptron multicapa (MLP). El primer método utiliza el histograma para procesar las imágenes antes de aplicar la red neuronal; el segundo método se basa en aplicar la transformada wavelet discreta precedida de la red neuronal. Ambos métodos presentaron resultados adecuados para el reconocimiento de rostros, considerando que las características de la base de datos faciales pueden ser diversas. Las bases de datos utilizadas contienen imágenes con variación de ángulos en el rostro, expresiones faciales e incluso el uso de accesorios (lentes). En esta capítulo se presentan los proceso y los resultados obtenidos en cada método implementado.

4.1 METODO 1: BASADO EN HISTOGRAMA

En este método se describe el proceso implementado como lo muestra la figura 4.1. El flujo de información consiste en partir de la imagen original, donde el rostro es enfocado para eliminar información de fondo, se obtiene el histograma y el vector resultante es ingresado a la MLP.

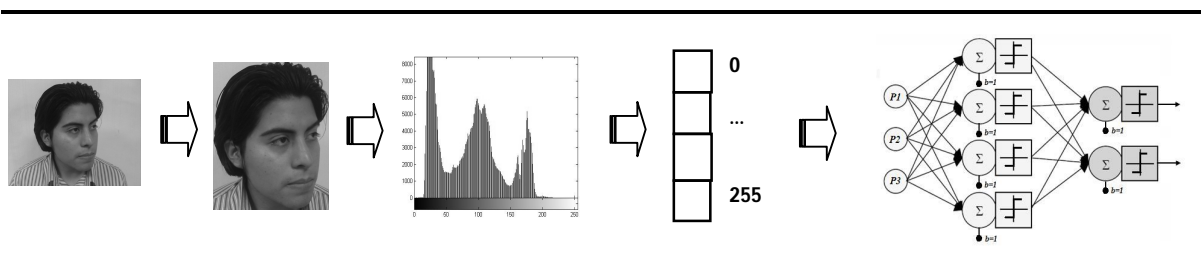


Figura 4.1. Flujo de Información, primer método propuesto.

A continuación se describe con mayor detalle cada una de las etapas propuestas:

- a) **Adquisición de la fotografía** las imágenes fueron tomadas con una cámara digital contemplando fondo blanco con mirada hacia la cámara y expresión normal o seria para reducir la complejidad del problema. La base de datos facial consta de 15 personas; cada una de ellas con 9 ángulos. (Ver anexo 1).
- b) **Preprocesamiento** contempla las imágenes en escala de grises procediendo a extraer de la imagen original solo la parte del rostro evitando con esto errores por cambio de vestuario. Se obtiene el histograma de la imagen para evitar problemas con imágenes de diferente tamaño; ya que aunque cambie el tamaño en la foto, el tamaño del histograma se conserva de 256 datos, correspondientes a los valores de intensidad de la imagen del rostro y este entra como vector a la MLP.
- c) **Procesamiento** los datos obtenidos del histograma son ingresados a la MLP la cual se encarga de normalizar los datos en un rango de 0 a 1, estos datos se propagan por todas las neuronas de la red y en la última etapa la capa de la salida tomara el valor más alto para indicar cual neurona es la activada, correspondiendo al número de clase que la red reconoce logrando identificar el rostro de la persona. La función de activación utilizada es sigmoïdal, en los resultados se describe con detalle las características de la MLP.

4.1.1 Implementación

Primero definiremos los rostros que aprenderá la MLP, es decir el conjunto de entrenamiento, como se observa en la figura 4.2, la base de datos está formada por los histogramas de 15 personas después de haber extraído de la foto la imagen de la cabeza.

Cada clase consta de 5 patrones y cada foto tiene cierto ángulo en el rostro de la persona, 30°, 60°, 90°, 120° y 150° en giro horizontal con cabeza orientada hacia el centro, dando un total de 75 patrones.

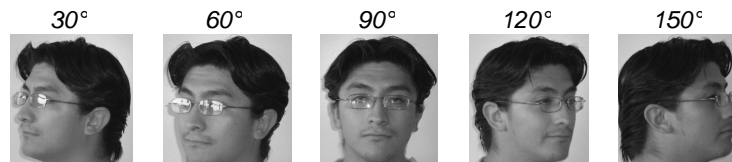


Figura 4.2. Ejemplo de los patrones de entrenamiento para la primera persona.

El conjunto de prueba consta de 15 histogramas de personas pero con fotos tomadas a diferente ángulo, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135° y 150° en giro horizontal y cabeza hacia el centro en posición vertical dando un total de 135 patrones, como se observa en la figura 4.3. Las bases de datos para el conjunto de entrenamiento y prueba se guardaron con formato de texto plano y son utilizados por la MLP.

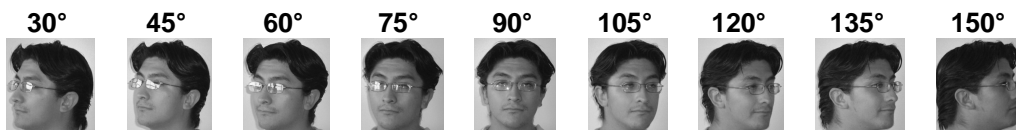


Figura 4.3. Ejemplo de los patrones de prueba para la décima persona.

Para este método se implementaron dos procesos para obtener el vector de entrada de la MLP los cuales consisten en:

- a) Utilizar una imagen correspondiente al rostro de la persona que al obtener su histograma se llega a un vector con 256 datos. Ver figura 4.4.
- b) Se obtiene de la foto, la imagen del rostro de la persona; se segmenta en tres imágenes rectangulares del mismo ancho y diferente largo que corresponden a los ojos con cejas, nariz y boca con barbilla, de cada imagen se obtiene un histograma y se integra en un solo vector con $256 \times 3 = 768$ datos. Ver figura 4.4.

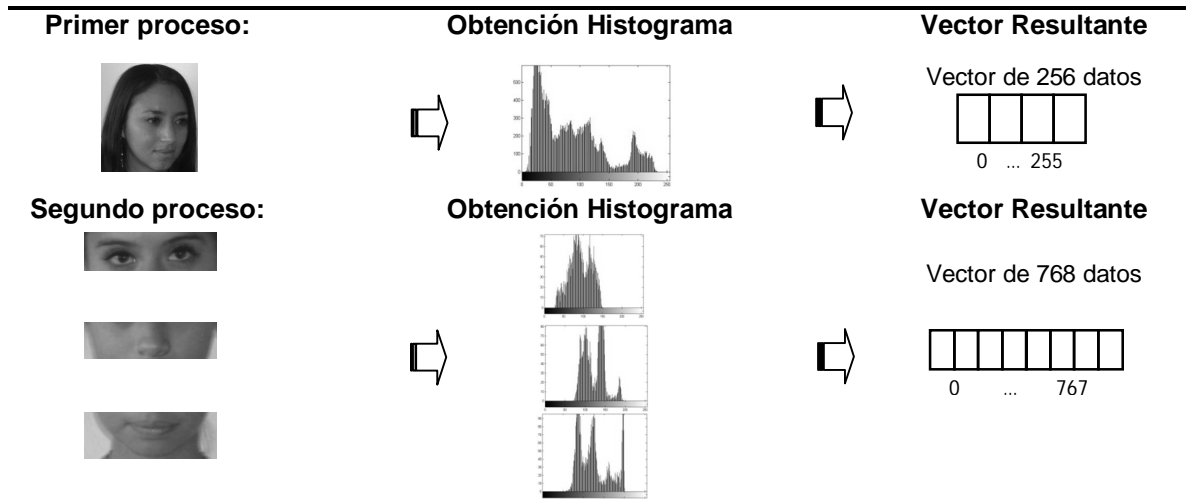


Figura 4.4. Diferencia entre los procesos utilizados para la obtención del vector de entrada para la RNA.

La MLP utiliza un vector de entrada de 256 datos para el uso del primer proceso y 768 datos para el segundo, que corresponden a la longitud del histograma para la imagen en escala de gris, dos capas ocultas con 30 neuronas cada una que utilizan la función de transferencia “*logsig*” donde $f(n) = 1 / [1 + e^{-(n)}]$ y en la capa de salida se tiene 15 neuronas que corresponden al número de personas o clases a identificar.

4.1.2 Resultados del método propuesto

Para el primer proceso se usó un vector de entrada con 256 datos correspondientes al histograma de las fotos de la cabeza de las personas. El mejor resultado fue de 96.296296 % en la clasificación, se necesitaron 30 000 épocas y 30 neuronas en cada capa oculta, con una tasa de aprendizaje de 0.05.

Epocas	Alpha	Capa entrada	Capas Ocultas 1 y 2	Capa Salida	% de Clasificación
10000	0.05	256 neuronas	30 neuronas	15 neuronas	94.8148%
30000	0.05	256 neuronas	30 neuronas	15 neuronas	96.2962%

Tabla 4.1. Parámetros de la Red Neuronal Artificial, primer proceso.



Figura 4.5. Imágenes mal clasificadas para el primer proceso.

Haciendo uso del segundo proceso, se usó un vector de entrada con 768 datos correspondientes a los 3 histogramas, ojos con cejas, nariz y boca con barbilla, de las imágenes de la cara de las personas. El mejor resultado fue de 93.3333 % en la clasificación, se necesitaron 10 000 épocas y 30 neuronas en cada capa oculta, con una tasa de aprendizaje de 0.05.

Epocas	Alpha	Capa entrada	Capas Ocultas 1 y 2	Capa Salida	% de Clasificación
10000	0.05	768 neuronas	30 neuronas	15 neuronas	93.3333%

Tabla 4.2. Parámetros de la Red Neuronal Artificial, segundo proceso.

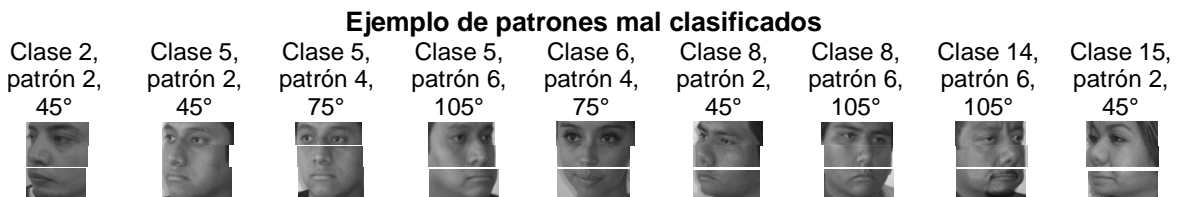


Figura 4.6. Imágenes mal clasificadas para el segundo proceso.

Se demostró que con cinco patrones para el aprendizaje de la red neuronal cuyas fotos tienen los ángulos 30°, 60°, 90°, 120° y 150°, esta es capaz de reconocer los rostros aprendidos y aquellos con ángulo intermedio, en total 9 ángulos del rostro por persona 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135° y 150° aunque el giro de la cabeza sea solo horizontal y no vertical se obtuvieron buenos resultados tomando en cuenta que para las pruebas no se aplicó ningún filtro para mejorar la iluminación y que las fotos tenían diferente tamaño.

Con base a la experimentación, se observó que el uso del histograma benefició el desempeño de la red neuronal, pues en la entrada se configuró con 256 datos para el primer proceso y 768 para el segundo, esto redujo el número de neuronas a 30 para cada una de las 2 capas ocultas lo que permitió obtener un 96.296296% (130 de 135) en la clasificación para la primera técnica utilizando 30000 épocas para el entrenamiento y 93.3333%(126 de 135) en la clasificación para la segunda técnica para el que solo se necesitaron 10000 épocas. Se realizaron diversos experimentos de prueba y error, fijando por último dichos valores.

Aunque se obtuvieron buenos resultados usando el primer método, este presenta problemas cuando hay un cambio de iluminación, expresiones faciales, orientación del rostro e incluso haciendo uso de anteojos; lo cual el uso del histograma no es útil para solucionar dichos problemas, por lo tanto es importante hacer uso de un filtro que ayude a la extracción de características, permitiendo realizar adecuadamente la tarea de reconocimiento de rostros.

4.2 METODO 2: BASADO EN LA TRANSFORMADA WAVELET DISCRETA

La TWD es empleada como un método de extracción de características faciales donde se comprueba la estabilidad de los vectores característicos extraídos en presencia de variaciones debido a cambios de iluminación, postura y expresión facial así como la sensibilidad de estos vectores a cambios de identidad.

El sistema propuesto para el reconocimiento de rostros basado en wavelets y redes neuronales consta de cuatro etapas; la figura 4.7 muestra el diagrama en bloques de un sistema de reconocimiento de rostros basado en redes neuronales.

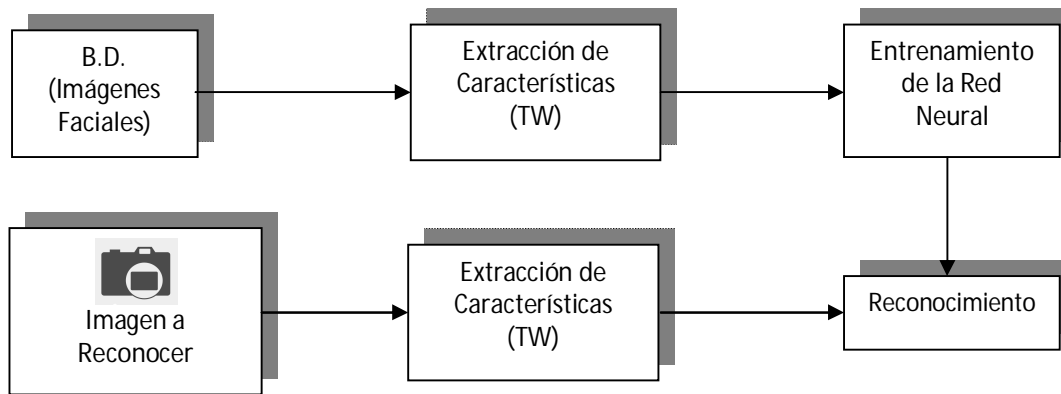


Figura 4.7. Diagrama del sistema de reconocimiento de rostros.

4.2.1 Adquisición de datos

Debido al gran número de técnicas y teorías de reconocimiento de rostros, es clara la necesidad de contar con bases de datos estándar y parámetros que permitan realizar experimentos en investigaciones en el área de visión artificial.

Por otra parte, construcción de una base de datos de alta calidad es una tarea muy laboriosa y que lleva tiempo de realización. Además, la utilización de bases de datos con características específicas y distintos formatos dificulta la evaluación comparativa de los algoritmos. Es por ello que se han creado diversas bases de datos estándar (algunas de ellas se encuentran disponibles vía internet). En la Tabla 4.3 se presentan algunas de las bases de datos disponibles vía internet con sus respectivas direcciones web.

Nombre	Dirección web
The Color FERET Database	http://www.itl.nist.gov/iad/humanid/colorferet/home.html
The Yale Face Database	http://cvc.yale.edu/projects/yalefaces/yalefaces.html
CAS-PEAL Face Database	http://www.jdl.ac.cn/peal/index.html
The Yale Face Database B	http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html
PIE Database	http://www.ri.cmu.edu/projects/project_418.html
The ORL Database of Faces	http://www.cl.cam.ac.uk/research/dtg/attarchive/
The AR Face Database	http://cobweb.ecn.purdue.edu/~aleix/aleix face DB.html
VALID Database	http://ee.ucd.ie/validdb/
The XM2VTS Database	http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/
Caltech Faces	http://www.vision.caltech.edu/html-files/archive.html

Tabla 4.3. Bases de datos de rostros estándar disponibles vía internet.

Los experimentos fueron realizados empleando la base de datos de rostros de Olivetti Research Laboratory [ORL, 2002]. Esta base de datos consiste de 10 imágenes de 40 individuos, las imágenes se encuentran en formato pgm con una resolución de 96x112 pixeles en 256 niveles de gris. (Ver anexo 2).

La base de datos cuenta con imágenes de rostros con variaciones de luminosidad, diferentes expresiones faciales, orientación del rostro, tomas con anteojos y sin anteojos, lo que permite probar la robustez del sistema contra estas variaciones. Todas las imágenes fueron tomadas contra un fondo homogéneo oscuro con el individuo en posición frontal. Además, la base de datos de rostros de ORL es una base de datos estándar utilizada por muchos investigadores, lo que permite realizar comparaciones con diferentes algoritmos propuestos.

En la Fig. 4.8 se muestran algunas imágenes de ejemplo tomadas de la base de datos de rostros ORL. Los ejemplos consisten de 10 tomas diferentes de 2 individuos, las tomas incluyen cambios de expresión facial, iluminación, accesorios (por ej., anteojos) y postura.



Figura 4.8. Ejemplo de Base de datos ORL.

4.2.2 Extracción de características

La información representada por las frecuencias espaciales son muy importantes para el reconocimiento de rostros, la información contenida en una baja y alta frecuencia tienen diferentes roles en la descripción de la imagen. La baja frecuencia constituye la descripción de la imagen global y la alta frecuencia los detalles de la imagen. Los efectos producidos por las diferentes expresiones faciales son reducidos por los componentes de altas frecuencias, los componentes de baja frecuencia son suficientes para la tarea de reconocimiento.

La descomposición de una señal unidimensional se lleva a cabo de la siguiente manera: en primer lugar, la señal se pasa a través de un filtro paso-bajas y otro paso-altas simultáneamente, y posteriormente se realiza la operación de decimación (\downarrow) a la salida de cada filtro para evitar el traslape y eliminar los datos que no son relevantes.

La Transformada Wavelet Discreta permite descomponer una señal o imagen en aproximaciones y detalles. El esquema de filtrado propuesto para implementar la TWD unidimensional se muestra en la Fig. 4.9.

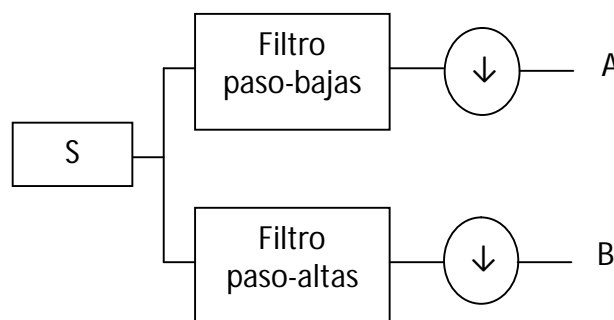


Figura 4.9. Esquema de filtrado para calcular la TWD unidimensional.

La TWD permite descomponer una señal o imagen en aproximaciones y detalles. Los *coeficientes de aproximación* (A) se obtienen de la salida del filtro paso-bajas, mientras que los *coeficientes de detalle* (D) son proporcionados por el filtro paso-altas.

El proceso de descomposición mediante la TWD puede ser iterativo, esto se logra filtrando los coeficientes de aproximación en forma recursiva. A esto se le denomina árbol de descomposición.

Las aproximaciones son componentes de baja frecuencia de la señal y los detalles son los componentes de alta frecuencia. En cada proceso de filtrado se genera un nivel de descomposición. La figura 4.10 muestra el nivel 4 del proceso de descomposición de una imagen de dos dimensiones.

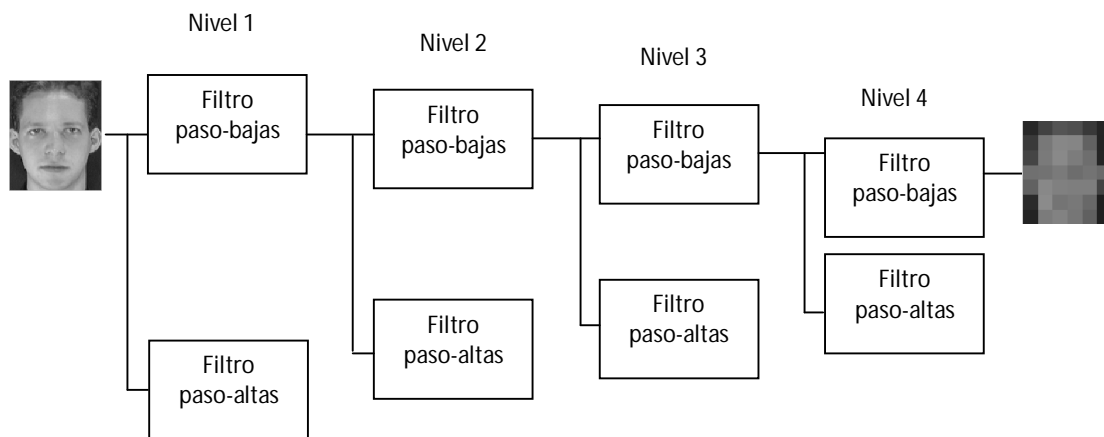


Figura 4.10. Esquema de la descomposición wavelet.

La aplicación de la TWD a una imagen produce cuatro matrices de coeficientes wavelet, una de aproximación y tres de detalles (horizontal, vertical y diagonal). La matriz de coeficientes de aproximación contiene los componentes en bandas de baja frecuencia (filtro paso-bajas) de la imagen original mientras que las matrices de detalles los componentes en bandas de alta frecuencia (filtro paso-altas).

La TWD descompone una imagen de dos dimensiones, utilizando un banco de filtros, obteniendo una descomposición por cada nivel. La Figura 4.11 muestra un enfoque de una descomposición wavelet de nivel 4 de una imagen de dos dimensiones donde se distingue claramente la reducción de las dimensiones en cada nivel, obteniendo de esta manera patrones más simples representados en pixeles. Cada nivel de descomposición reduce el tamaño de la imagen.

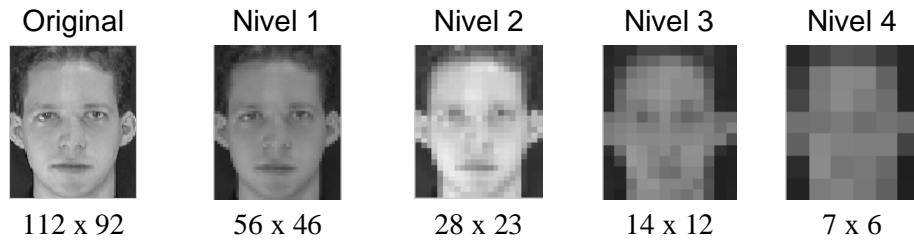


Figura 4.11. Niveles del proceso de descomposición de una imagen facial en píxeles.

4.2.3 Vectores característicos

Los componentes en bandas de baja frecuencia contribuyen a la descripción global de la imagen, la cual creemos contiene la información más relevante para la diferenciación entre clases. Por tal motivo, la extracción de características se lleva a cabo mediante la descomposición recursiva del rostro, de donde los coeficientes en las bandas de más baja frecuencia son extraídos para formar los vectores característicos.

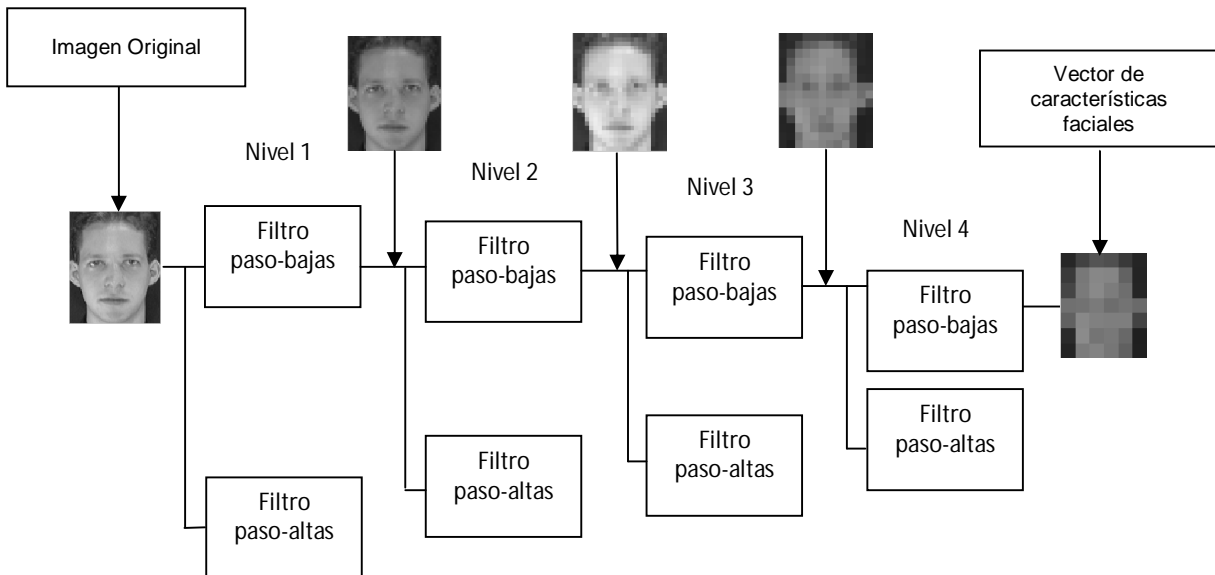


Figura 4.12. Esquema de extracción de características mediante la TWD.

En la Fig. 4.12 se muestra la descomposición en 4 niveles de un rostro, el vector característico es derivado de los componentes de baja frecuencia del cuarto nivel de descomposición.

La descomposición recursiva de la imagen de un rostro conduce a la reducción de las dimensiones en un cuarto de su tamaño por cada nivel de descomposición. Al descomponer recursivamente un rostro empleando la TWD, la banda de frecuencia que corresponde a los detalles de la imagen se representa con menor cantidad de información, además de que la dimensión de la imagen reduce (como se muestra en la fig. 4.12). Sin embargo, después de un determinado número de descomposiciones, es posible que los vectores característicos de distintas clases comiencen a aproximarse entre sí debido a que la representación wavelet del rostro tiende a ser más general. Para determinar el nivel de descomposición mínimo con el cual se obtienen la representación general de los rostros, se realizaron pruebas con distintos niveles de descomposición.

Los vectores característicos son obtenidos de la descomposición de las 10 imágenes de la persona como se muestra en la figura 4.13 donde se aplicó un nivel de descomposición del cuarto nivel utilizando la función Daubechies obteniendo un vector con 42 coeficientes diferentes.

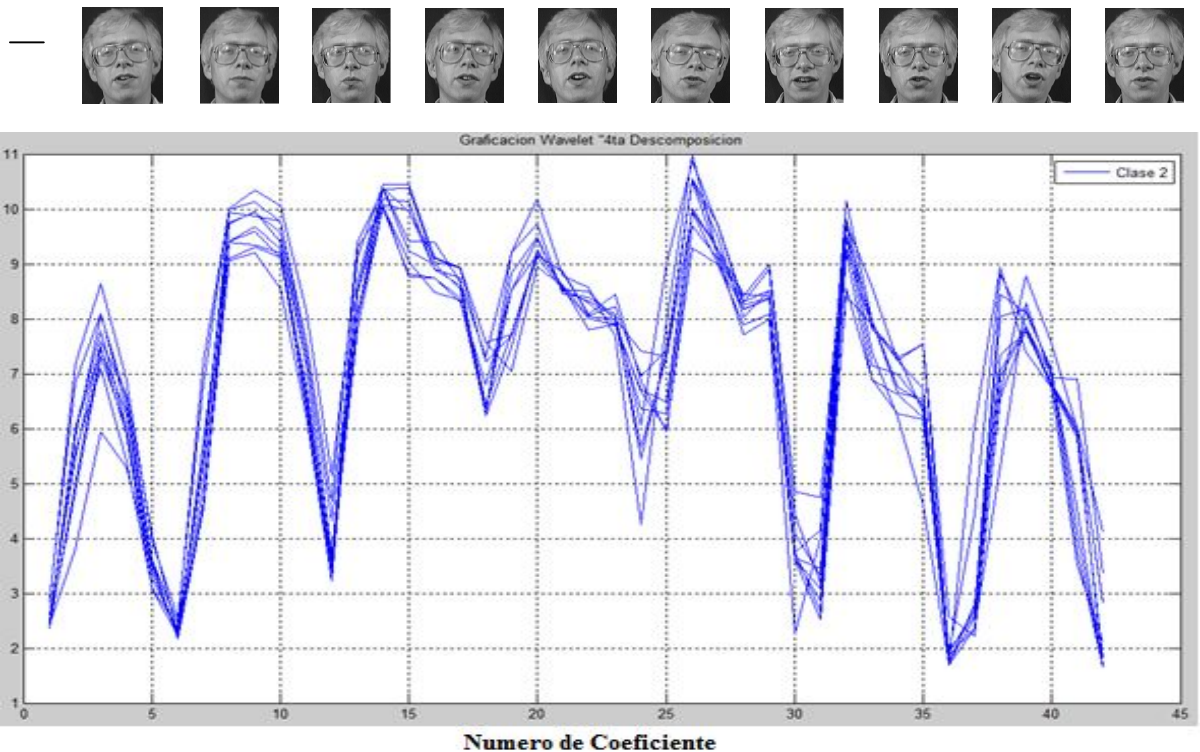


Figura 4.13. Variación entre los vectores característicos de una persona.

Los vectores característicos de 10 imágenes de dos personas con variaciones en la apariencia facial se presentan en la Fig. 4.14; se puede ver que las variaciones entre vectores característicos interclase son mayores que las variaciones intraclase por cambios de iluminación, expresión facial y postura.

Este ejemplo prueba la capacidad de la TWD para extraer características faciales adecuadas para la diferenciación entre individuos.

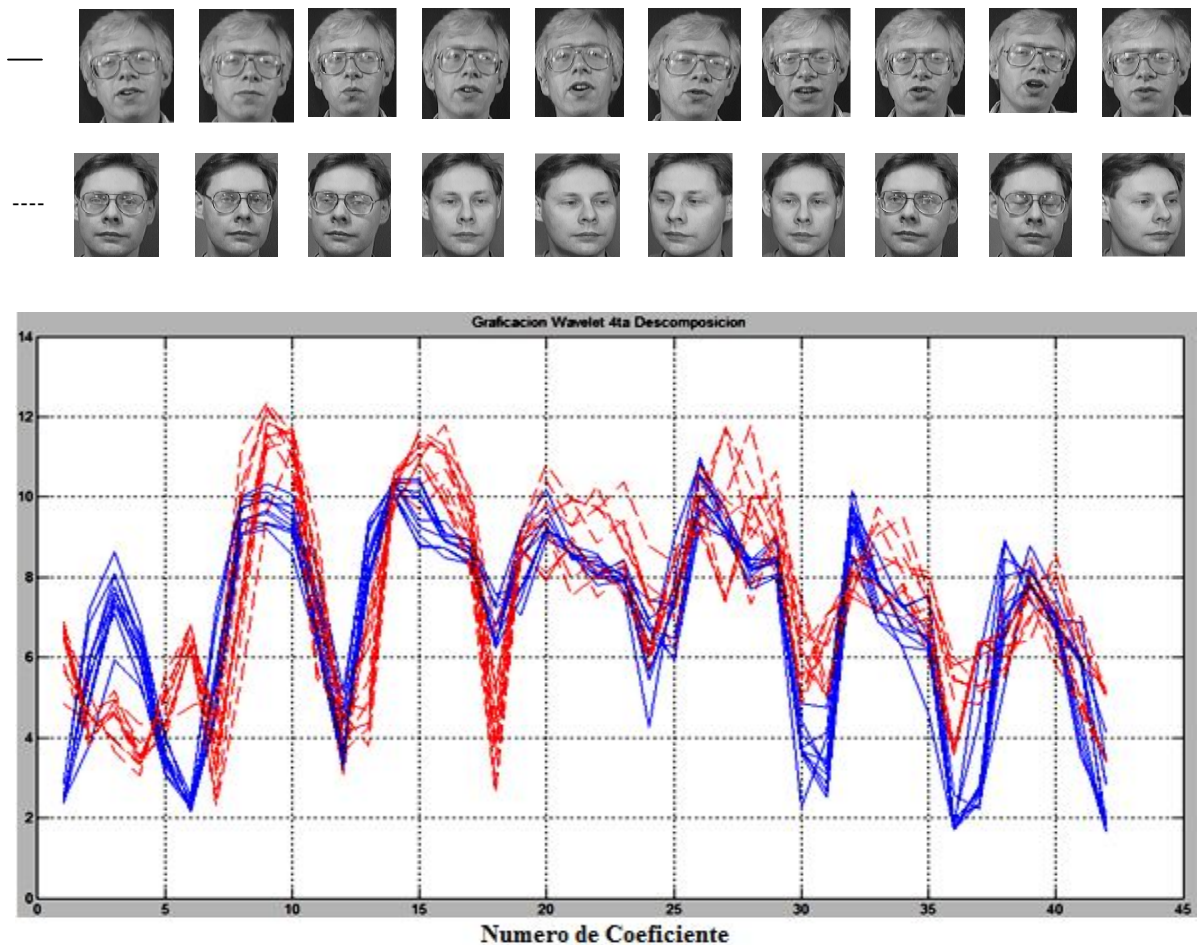


Figura 4.14. Variación entre los vectores característicos de dos personas.

En la figura 4.14 se puede observar que las imágenes de cada persona tiene una tendencia con un patrón similar que varía ligeramente entre sí mismo, pero que tiene características muy diferentes a la de los patrones de las demás personas, identificado en la representación de los coeficientes correspondientes.

Las matrices correspondientes a las imágenes de rostros tienen una dimensión $M \times N$, dichas matrices se utilizan de forma alineada representando un vector que es la entrada de la RNA. La etapa de entrenamiento y prueba se demuestra en los resultados experimentales, donde se puede observar el desempeño de la red neuronal. A continuación se describen las opciones utilizadas para convertir una matriz de $M \times N$ elementos en un vector de “L” elementos.

1. Se toman las columnas y se colocan una debajo de la otra.
2. Se toma una fila y se le aplica una transpuesta y se coloca una debajo de la transpuesta anterior.

Realizando un análisis de las matrices, y para la implementación del sistemas se tomara la primera opción para construir los vectores de entrenamiento de la red neuronal. De tal forma que si se tiene una matriz con dimensiones de $M \times N$ elementos se obtiene un vector de “L” elementos, el cual sirve para alimentar a la red neuronal. Sin embargo, los valores que tiene este vector son muy grandes para entrenar a la red, es más recomendable entrenar a la red neuronal con vectores de valores pequeños.

4.2.4 Entrenamiento de la Red Neuronal Artificial

Para la clasificación de los vectores característicos extraídos por la TWD se propone una red neuronal perceptrón con dos capas ocultas, una de entrada y salida. Las neuronas en la capa de entrada están alimentadas por los vectores característicos obtenidos mediante la aplicación de la TWD a las imágenes de los rostros. El número de neuronas de capa oculta fueron definidos mediante diversas pruebas. La red neuronal artificial multicapa hace uso de neuronas interconectadas entre capas, donde la salida de la capa actual se vuelve la entrada de la siguiente capa; siendo una capa un conjunto de neuronas interconectadas con el mismo número de entradas y mismo tipo de función de transferencia.

El proceso de aprendizaje consta de dos fases, una propagación hacia adelante y una propagación hacia atrás que se itera hasta obtener los pesos que reduzcan el “*error cuadrático medio*”. Para la etapa de aprendizaje, se utilizó la red neuronal perceptrón multicapa, la capa de entrada utiliza 42 nodos, dos capas ocultas con 30 neuronas cada una que utilizan la función de transferencia “*logsig*”, la tasa de aprendizaje de 0.05 y en la capa de salida se tiene 15 neuronas que corresponden al número de personas o clases a identificar como se muestra en la figura 4.15, utilizando el algoritmo de backpropagation para la formación de la red neuronal.

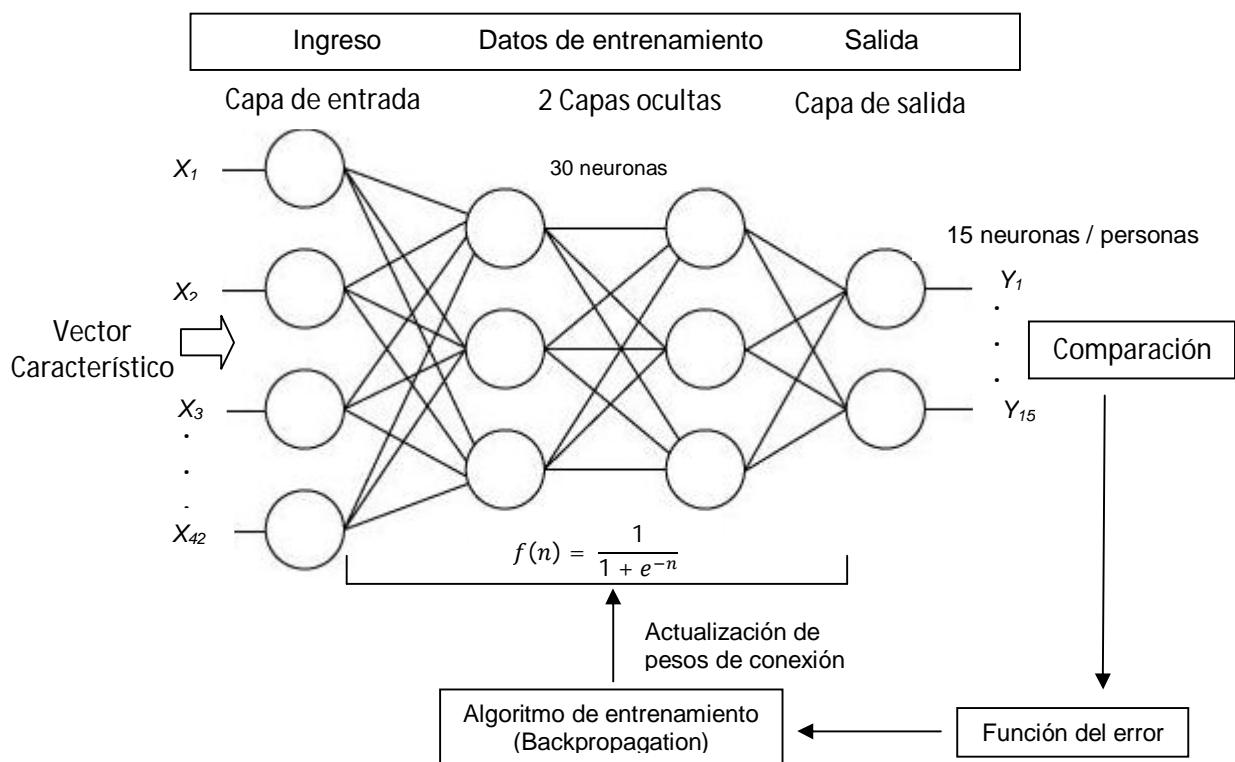


Figura 4.15. Topología general de la red neuronal utilizada.

La prueba se realizó mediante el uso de la base de datos facial ORL, de las cuales se tomaron 15 aleatoriamente como prueba; las imágenes fueron tomadas en diferentes momentos, variando la iluminación, expresiones faciales (ojos abiertos/ cerrados, sonriendo / no sonriendo) y detalles fáciles (lentes / sin lentes).

El conjunto de entrenamiento, se eligieron las primeras 5 imágenes de cada persona como se observa en la figura 4.16, donde se puede identificar claramente los cambios en expresiones faciales, movimientos de rotación de la cara, cambios de iluminación, detalles faciales.



Figura 4.16. Ejemplo de los patrones de entrenamiento.

Para el conjunto prueba se eligieron las otras 5 imágenes restantes por persona como se muestra en la figura 4.17, dando un total de la utilización de las 10 imágenes por persona con las cuales consta la Base de Datos.



Figura 4.17. Ejemplo de los patrones de prueba.

La tabla 4.4 muestra los diferentes filtros wavelet que se evaluaron (de las familias daubechies, haar, symlets, coiflets, biorthogonal, reverse biorthogonal y dmeyer) para la obtención de vectores característicos. (Ver anexo 3).

<i>Familia Wavelet</i>	<i>Filtro wavelet</i>
Wavelet Daubechies	<i>db1, db2, db3, db4, db5, db6, db7, db8, db9, db10</i>
Wavelet Haar	<i>haar</i>
Wavelet Symlets	<i>sym1, sym2, sym3, sym4, sym5, sym6, sym7, sym8, sym12</i>
Wavelet Coiflets	<i>coif1, coif2, coif3, coif4, coif5</i>
Wavelet Biorthogonal	<i>bior1.1, bior1.3, bior1.5, bior2.2, bior2.4, bior3.1, bior3.5</i>
Wavelet Reverse Biorthogonal	<i>rbior1.1, rbior1.6, rbior6.8</i>
Wavelet Dmeyer	<i>dmey</i>

Tabla 4.4. Familias wavelet con sus respectivos filtros evaluados.

La tabla 4.5 muestra los resultados de la aplicación de la red neuronal a varios de los vectores característicos usando diferentes filtros wavelets, donde se pueden observar los porcentajes de reconocimiento para cada filtro wavelet con una descomposición de nivel 4, evaluando el desempeño de cada filtro se usaron los mismos datos utilizados en el entrenamiento (prueba cerrada), así como aquellos no utilizados en la etapa de entrenamiento (prueba abierta).

Los datos presentados corresponden a los resultados en los cuales se obtuvieron los porcentajes de reconocimiento más altos entre las familias wavelet.

<i>Filtro Wavelet</i>	<i>Vector de entrada</i>	<i>Prueba Cerrada</i>	<i>Prueba Abierta</i>	<i>Épocas</i>	<i>Tasa de aprendizaje</i>
db1	42	100%	94.66%	10000	0.05
		100%	96.66%	30000	
haar	42	100%	93.33%	10000	0.05
		100%	94.66%	30000	
sym1	42	100%	94.66%	10000	0.05
		100%	96.66%	30000	
coif1	42	100%	76%	10000	0.05
		100%	82.66%	30000	
bior1.1	42	100%	93.33%	10000	0.05
		100%	93.33%	30000	
rbior1.1	42	100%	94.66%	10000	0.05
		100%	94.66%	30000	
dmey	42	86.6666%	46.66%	10000	0.05
		86.6666%	42.66%	30000	

Tabla 4.5. Resultados obtenidos aplicando diferentes filtros wavelets.

En el grafico 4.1 se puede observar mejor el comportamiento de los porcentajes obtenidos en las pruebas abiertas, variando el número de épocas. Cada filtro wavelet genera subespacios característicos distintos, algunos de estos subespacios pueden proporcionar una representación más robusta que otros, por tal motivo, se realizaron pruebas con diferentes filtros wavelet para determinar el que mejor caracterice un rostro para su reconocimiento.

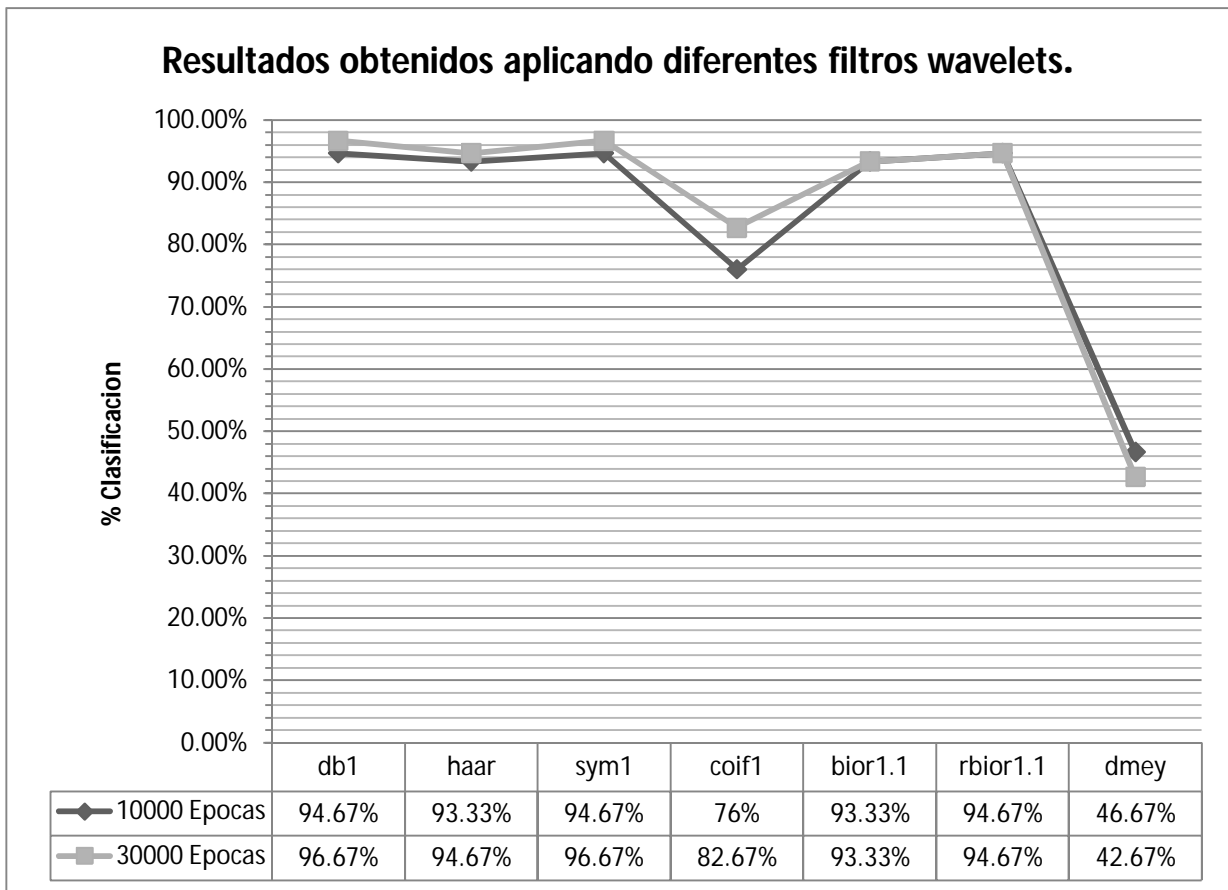


Figura 4.18. Resultados obtenidos aplicando diferentes filtros wavelets. Prueba abierta.

Todos experimentos reportados fueron realizados bajo las siguientes condiciones:

- La extracción de características faciales se realizo por medio de la TWD.
- La etapa de reconocimiento fue implementada por una red neuronal perceptrón multicapa con dos capas ocultas, cada capa oculta con 30 neuronas. El número de neuronas en la capa oculta fue determinado mediante diversas pruebas.
- Se realizaron diversos experimentos de prueba y error, fijando los valores para la MLP anteriormente mencionados.

- Para el entrenamiento de la red se empleó el algoritmo de retropropagación de errores (backpropagation) y se utilizaron únicamente las primeras cinco imágenes de cada individuo.
- En las evaluaciones se emplearon las mismas imágenes usadas en el entrenamiento (evaluación cerrada) además de imágenes no utilizadas durante el entrenamiento (evaluación abierta).

4.3 DISEÑO DEL SISTEMA

Como parte del desarrollo de la investigación y para facilitar la implementación del reconocimiento de rostros, se propuso el diseño un sistema para automatizar dicha tarea; cabe mencionar que el sistema cuenta con parámetros ya establecidos, es decir, la topología de la red neuronal y los niveles de descomposición para la imagen, todo esto de acuerdo con los resultados obtenidos y mostrados en un subcapítulo anterior; todo con la finalidad de eficientar el reconocimiento de tal manera que el resultado arrojado por el sistema sea lo más rápido posible para el usuario final.

4.3.1 Arquitectura del Sistema

Una vez obtenidos los requerimientos se procedió a la fase de diseño de la herramienta computacional, la cual fue desarrollada bajo 2 lenguajes: JAVA, utilizando el estándar NetBeansIDE 6.7.1 y Matlab R2008a. El esquema general está descrito en la figura 4.19.

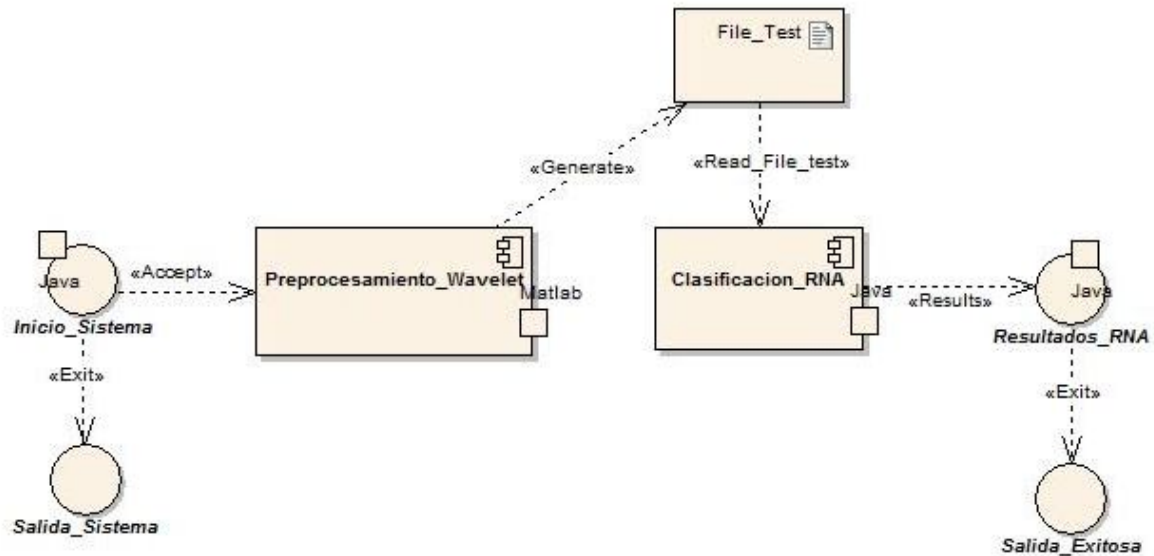


Figura 4.19. Esquema general de arquitectura.

En un nivel más generalizado, se describen cada unas de las etapas de la arquitectura propuesta:

- **Inicio de Sistema:** Interfaz de inicio del sistema, el cual el usuario podrá optar por ingresar a la pantalla principal o realizar la salida del sistema.
- **Preprocesamiento:** Interfaz para el usuario donde podrá elegir una imagen a reconocer, dicha imagen puede ser cargada directamente del disco duro local, así mismo en dicha pantalla mostrara la descomposición wavelet de la imagen facial; generando un archivo prueba con extensión txt, el cual es depositado en el disco duro local. Esta etapa también funge como un proceso para la obtención de la Transformada Wavelet.
- **Archivo prueba:** Documento generado en la etapa de preprocesamiento el cual contiene los patrones de la imagen a reconocer, el documento será utilizado para la etapa de clasificación.

- Clasificación:** Proceso interno del sistema el cual se encarga de leer el documento generado en la etapa de preprocesamiento; dicho proceso será capaz de reconocer a la persona que fue seleccionada en proceso de preprocesamiento. La figura 4.20 muestra el flujo de información el cual modela el funcionamiento general de la red neuronal artificial.

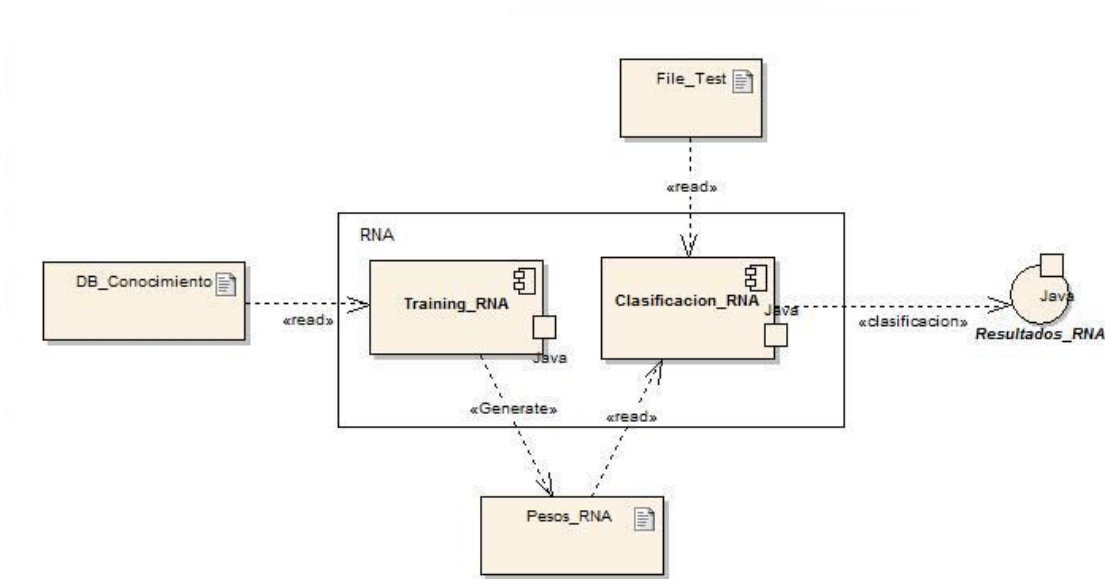


Figura 4.20. Esquema general de la arquitectura del procesamiento de la RNA.

- Resultados:** Interfaz para el usuario donde podrá observar el resultado generado por el clasificador.

La figura 4.21 presenta solo una vista del sistema propuesto, la cual pertenece a la vista de preprocesamiento y que fue descrita de manera general anteriormente. Cabe mencionar que cada que dicha vista es una de las principales dentro del sistema de reconocimiento de rostros y el usuario podrá manipular para poder llevar a cabo la tarea de reconocimiento.

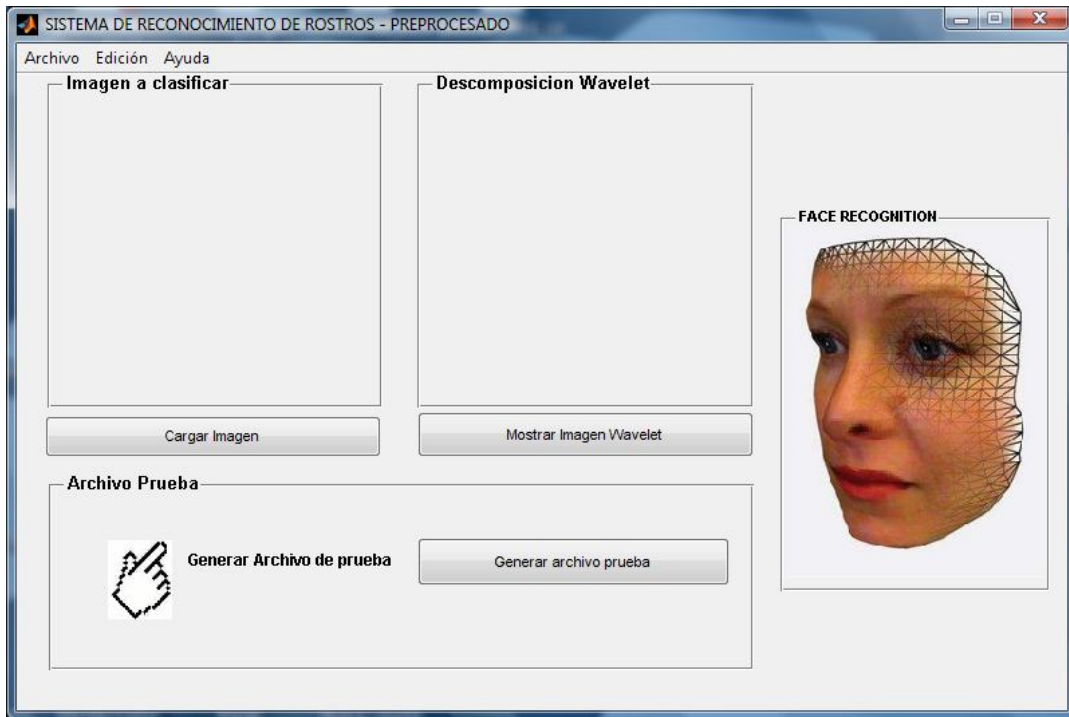


Figura 4.21. Interfaz de usuario de preprocesamiento.

Para una descripción más detallada de las pantallas así como el flujo de información, podrá consultarse la ventana de vista de datos que se encuentra en el anexo 4.

CAPITULO 5

VALIDACION DE RESULTADOS

En el análisis discriminante, el reconocimiento de patrones, inteligencia artificial, estadística de aprendizaje y clasificación se utiliza una matriz para representar los errores en la asignación de categorías [Mandeville, 2009].

Esta matriz, que fue introducida por Massy y se denomina matriz o tabla de confusión, expone los resultados de un procedimiento de clasificación o predicción. Si hay k categorías posibles, es una matriz $k \times k$ donde las filas denotan la verdadera categoría y las columnas indican la categoría a la que se asigna. Se evalúa el comportamiento de tales sistemas por medio de los datos en la matriz. Aquellas entradas en la diagonal de la matriz son correctas. Las entradas fuera de la diagonal son errores de clasificación y representan “confusión”. Con la matriz es fácil ver si el sistema confunde dos categorías [Mandeville, 2009].

		Verdadera clasificación dada por el usuario	
		Positivo	Negativo
Clasificación dada por el sistema de reconocimiento	Positivo	Verdaderos Positivos (TP)	Falsos Positivos (FP)
	Negativo	Falsos Negativos (FN)	Verdaderos Negativo (TN)
	Totales	Desempeño global del sistema en porcentaje de aciertos	Número total de aciertos

Tabla 5.1. Tabla de contingencia o matriz de confusión.

Verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP) y falsos negativos (FN) son las cuatro posibilidades que pueden resultar de una predicción cuando hay dos categorías “sí” y “no”.

Un elemento es considerado **verdadero positivo (TP)** cuando dicho elemento es en realidad positivo para el usuario y es también clasificado como positivo por el sistema de reconocimiento. Si tal elemento es positivo para el usuario pero es clasificado como negativo por el sistema de reconocimiento, es considerado un **falso negativo (FN)**.

Un elemento es considerado **verdadero negativo (TN)** cuando dicho elemento es en realidad negativo para el usuario y es también clasificado como negativo por el sistema de reconocimiento. Si tal elemento es negativo para el usuario pero es clasificado como positivo por el sistema de reconocimiento, es considerado un **falso positivo (FP)** [Salazar, 2006].

5.1. SENSIBILIDAD

Es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como positivo respecto a la condición que estudia la prueba, razón por la que también es denominada fracción de verdaderos positivos (FVP). La ecuación 5.1 representa esta medida [Juárez, 2008].

$$sensibilidad = \frac{Verdaderos\ Positivos}{Falsos\ Negativos + Verdaderos\ Positivos} \quad (5.1)$$

5.2 ESPECIFICIDAD

Es la probabilidad de clasificar correctamente a un individuo cuyo estado real sea el definido como negativo. Un método con especificidad ideal es aquel que no tiene falsos positivos. Es sinónimo de fracción de verdaderos negativos (FVN).

La ecuación 5.2 representa esta medida [Juárez, 2008].

$$\text{especificidad} = \frac{\text{Verdaderos Negativos}}{\text{Verdaderos Negativos} + \text{Falsos Positivos}} \quad (5.2)$$

Para la validación de resultados la categorización de las clases es: “+1” si el patrón pertenece a la clase y “-1” si el patrón no pertenece a la clase; donde C_1 corresponde a la clase de casos con personas a reconocer y C_2 corresponde a la clase de casos a no reconocer por el sistema, como se muestra en la tabla 5.2.

	Etiqueta Salida RNA	=	Etiqueta Real
TP	(+1) C_1	=	C_2 (+1)
TN	(- 1) C_1	=	C_2 (- 1)
FP	(+1) C_1	=	C_2 (- 1)
FN	(- 1) C_1	=	C_2 (+1)

Tabla 5.2. Tabla de decisiones para la clasificación.

La tabla 5.3 presenta la matriz de confusión con los resultados obtenidos para cada una de las pruebas realizadas, dichas pruebas se realizaron con los filtros wavelet que presentaron mayor porcentaje de clasificación; donde “S” hace referencia al concepto de sensibilidad y “E” al concepto de especificidad.

	Wavelet db1	S	E	Wavelet sym1	S	E
P1	73 1	0.97333	0.98666	71 6	0.9466	0.92
	2 74			4 69		
	97.66% 147 de 150			95.33% 143 de 150		
P2	71 6	0.9466	0.92	71 6	0.9466	0.92
	4 69			4 69		
	94.66% 142 de 150			95.33% 143 de 150		
P3	71 8	0.9466	0.91111	71 12	0.9466	0.92
	4 82			4 138		
	94.54% 156 de 165			94.66% 213 de 225		

Tabla 5.3. Matriz de confusión con los resultados obtenidos usando filtros wavelet.

La sensibilidad en nuestro caso, es el porcentaje de personas reconocidas que son clasificados correctamente; un clasificador con sensibilidad cercana a 100% se puede hablar de un clasificador ideal. Por otro lado la especificidad es el porcentaje de personas que no serán reconocidas y es clasificado como negativo por el clasificador.

La tabla 5.3 muestra los resultados de sensibilidad y especificidad obtenidos de cada unas de las pruebas realizadas, donde los porcentajes presentados son adecuados y concluir que las pruebas son satisfactorias para identificar a las personas reconocidas y las no reconocidas.

Para cada unas de las pruebas se selecciona aleatoriamente un conjunto diferente de patrones de entrenamiento.

CONCLUSIONES

La utilización del histograma para el reconocimiento de rostros resulta eficiente cuando el tamaño de las imágenes en píxeles es variado, devolviendo un vector de 256 datos, el uso de esta técnica presenta problemas cuando hay un cambios de iluminación, expresiones faciales, orientación del rostro; por lo tanto es importante hacer uso de un filtro que ayude a la extracción de características.

Los vectores característicos extraídos por la TWD son adecuados para el reconocimiento de rostros, cada vector característico consta de 42 datos los cuales son leídos por la MLP. Los resultados demuestran que con cinco patrones para el aprendizaje de la red neuronal es capaz de reconocer rostros diferentes a los utilizados en la etapa de aprendizaje adquiriendo porcentajes de 96.66%, tomando en cuenta que se utilizó como filtro la transformada wavelet discreta para la extracción de características permitiendo mejorar el reconocimiento de rostros.

La transformada wavelet discreta representa una forma adecuada para la extracción de rasgos faciales, cuando se combina con redes neuronales de retropropagación. Los efectos de las diferentes expresiones faciales, rotación, orientación, etc. se redujeron mediante la descomposición Wavelet, que permite obtener vectores característicos similares cuando se presentan variaciones en imágenes faciales para la misma cara. Se ha demostrado que este método puede ser utilizado con éxito para el problema de clasificación que utiliza imágenes frontales. Los resultados de la evaluación de los distintos filtros Wavelet junto con la red neuronal como clasificador muestran que los filtros Daubechies de orden 1 y Symlets de orden 1 presentan mejores resultados para el reconocimiento de rostros con 94.66 y 96.66% de clasificación respectivamente.

Por su parte, el nivel de descomposición 4 se obtiene la representación más adecuada para la diferenciación entre rostros. El sistema muestra gran eficiencia para la identificación de rostros en condiciones controladas, representando una opción atractiva para algunas aplicaciones.

En este trabajo se realizó la compatibilidad del uso de dos lenguajes de programación; JAVA donde se implementó la red neuronal perceptrón multicapa y Matlab para el procesamiento digital de las imágenes como el cálculo del histograma, implementación de la transformada Wavelet (considerando diversos tipos de familias) y diseño de la interfaz gráfica.

Se generó una interfaz gráfica ejecutable independiente de los lenguajes de programación mencionados previamente. La interacción entre ambos lenguajes de programación permitió generar una aplicación que se puede identificar como sistema, que realiza las etapas de lectura de datos, procesamiento mediante TW y clasificación de imágenes de rostros.

TRABAJOS A FUTURO

Entre los trabajos futuros pueden señalarse la evaluación del sistema en modo verificación, la experimentación y adaptación con otros tipos de bases de datos faciales; la etapa de adquisición de la imagen podría tomarse de un sistema donde la captura automática del rostro sea en tiempo real; permitiendo así un sistema totalmente completo de tal forma que al tomar la imagen facial de una persona, el sistema pueda determinar si puede o no ingresar a las instalaciones dentro de un recinto. La evaluación de la TDW en detección de rostros en imágenes a color, la extensión a imágenes en tres dimensiones implementado algún otro tipo de clasificador diferente al presentado en este trabajo, permitiendo mejorar los porcentajes de clasificación para la tarea del reconocimiento facial.

La utilización correcta de la transformada Wavelet, en especial de los filtros Daubechies y Symlets, agilizan trabajos futuros en la misma línea de investigación, ya que, de acuerdo a los resultados presentados, se logra asimilar que estos filtros son mejores para la clasificación de rostros.

BIBLIOGRAFIA

- [Anam, 2009] Anam S., Shohidul I. M., (2009). Reconocimiento de rostros usando Algoritmos Genéticos y la Red Neuronal Backpropagation. *Multiconferencia Internacional de Ingenieros y Científicos de la Computación 2009. IMECS.*
- [AR FACE DATABASE, 1998] Base de Datos Faciales creado por Aleix Martínez y Robert Benavente en el Centro de Visión por Computador; Consultada el 13 de Junio de 2009, en:
http://cobweb.ecn.purdue.edu/~aleix/aleix_face_DB.html
- [Báez, 2006] Báez Cervantes A.E. (2006). *Identificación de firmas por medio de redes neuronales* Tesis de Maestría en Ciencias con Especialidad en Sistemas Digitales, Centro de Investigación y Desarrollo de Tecnología Digital IPN.
- [Caceres, 2002] Caceres B.F.G. (2002). *Reconocimiento de rostros humanos basado en una red neuronal de retropropagacion.* Tesis de Licenciatura. Universidad Mayor, Facultad de Ingeniería, Chile.
- [Caeiros, 2007] Mantilla Caeiros A.V. (2007). *Análisis, Reconocimiento y Síntesis de Voz Esofáfrica* Tesis de Doctorado en Comunicaciones y Electrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.
- [Chellappa, 1995] Chellappa R., C. L.Wilson, y S. Sirohey. (1995). *Human and machine recognition of faces: A survey.* En Proceedings of the IEEE, tomo 83, páginas 705–741.
- [CMU, 2000] CMU Image Data Base 2000. Consultada el 13 de Junio de 2009, en: <http://vasc.ri.cmu.edu/idb/html/face/index.html>
- [CVG, 2005] Computational Vision Group 2005
Consultada el 12 de Junio de 2009, en:
<http://www.vision.caltech.edu/html-files/archive.html>

- [Fernández, 2001] Fernández C. A.; Gómez J. G. y Fernández A. M. (2001). *“Reconocimiento de rostros utilizando secuencias de histogramas como tramas espacio-temporales”*. Publication of the Department of Computer Science. Universidad de Castilla-La Mancha, Toledo, España.
- [Fuentes, 2008] Fuentes Pineda G. (2008). *Reconocimiento de Rostros Mediante Wavelets y Redes Neuronales*. Tesis de Maestría en Ciencias de Ingeniería en Microelectrónica, Escuela Superior de Ingeniería y Mecánica Eléctrica, IPN.
- [García, 2006] García Ortega V. H. (2006). *Sistema de Reconocimiento de huellas dactilares para el control de acceso a recintos* Tesis de Maestría en Ingeniería de Cómputo con Especialidad en Sistemas Digitales, Centro de Investigación en Computo IPN.
- [García, 2008] García Ricárdez G.A. (2008). *Reconocimiento de rostros para control de acceso de persona Implementado con tecnología .net* Tesis de Maestría en Comunicaciones y Electrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.
- [González, 2005] González García I. (2005). *Control Neuronal de un Generador de Inducción para Generación Eólica* Tesis de Maestría en Ciencias con Especialidad en Ingeniería Eléctrica, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.
- [Gutiérrez, 2007] Gutiérrez Mojica E. (2007). *Diagnostico y Monitoreo de los Patrones de una Turbina de Gas Mediante Redes Neuronales Artificiales* Tesis de Maestría en Ciencias de Ingeniería en Microelectrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.
- [Guzmán, 1996] Guzmán Alacala M.A. *Sistema de reconocimiento de patrones mediante redes neuronales*. Proyecto de Investigación. Universidad Autónoma Metropolitana, Unidad Iztapalapa.
- [Guzmán, 1997] Guzmán Montoya E. (1997). *Implementación del algoritmo de retropropagación para entrenamiento de redes neuronales*. Tesis de Licenciatura en Computación, Universidad Autónoma Metropolitana, Unidad Iztapalapa.

- [Hamilton, 2009] Hamilton I. (2009). Método para el análisis espectroscópico de datos con redes neuronales artificiales. *Escuela de Ingeniería y Ciencias Matemáticas, Universidad de Northampton Square, London, Reino Unido.*
- [Hernández, 2004] Hernández López M. (2004). *Sistema para reconocimiento de hablantes dependiente e independiente del texto.* Tesis de maestría en Ciencias de Ingeniería en Microelectrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica, IPN.
- [Hernández, 2009] Hernández Camargo J. A. y Uribe Andrés M. *Reconocimiento de rostros utilizando histogramas secuenciales de imagen.* Tesis de licenciatura en Ingeniería en Comunicaciones y electrónica, Escuela Superior de de Ingeniería Mecánica y Eléctrica, IPN.
- [INAOE, 2005] Instituto Nacional de Astrofísica, Óptica y Electrónica “Reconocimiento de Patrones” 2005, Consultada el 14 de Junio de 2009 en: <http://ccc.inaoep.mx/~ariel/recpat.pdf>
- [Juárez, 2008] Juárez L.C. (2008). “*Desarrollo de técnicas de reconocimiento y procesamiento utilizando un sistema CAD para aplicaciones en medicina y otras áreas*” Tesis Doctoral en Ciencias de Comunicaciones y Electrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica, IPN. México.
- [Juárez, 2005] Juárez Segura C.D. (2005). *Eliminación de Ruido en Señales de Radar por Medio de Wavelets.* Tesis de Licenciatura en Ingeniero en Robótica Industrial, Escuela Superior de Ingeniería Mecánica y Eléctrica, UPA IPN.
- [Lippman, 1987] Lippman, Richard (1987). “An Introduction to Computing with Neural Nets”. IEEE ASSP Magazine. Vol. 3. No. 4. pp. 4-22.
- [López, 2008] López Padilla A.G. (2008). *Diseño y Desarrollo de un Simulador para una Planta Virtual, por medio de Redes Neuronales* Tesis de Licenciatura en Ingeniería en Control y Automatización, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.

- [Mailk, 2007] Malik A. (2007). Secuencia y características estructurales de la unión de los carbohidratos en las proteínas y evaluación mediante una red neuronal. *Departamento de Ciencias Biológicas, Universidad Jamia Millia Islamia, Nueva Delhi, India.*
- [Martínez, 2003] Martínez C. y Fuentes O. (2003). "Reconocimiento de Rostros usando Datos No Etiquetados". *Computación y Sistemas*. Vol. 7 No. 2 pp. 123 – 129.
- [Martínez, 2008] Martínez Molina M. (2008). *Desarrollo de una aplicación básica de reconocimiento de rostros mediante el análisis de componentes principales (PCA)*. Tesis Licenciatura. Escuela Superior de Ingeniería Mecánica Eléctrica, IPN.
- [McClelland, 2001] J. L. McClelland (2001). "Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises". Vol. 1. Press/Bradford Books, Cambridge, MA, EUA.
- [Medina, 2007] Medina Choperena O. (2007). *Utilización de la transformada Wavelet en el procesamiento de imágenes* Tesis de Licenciatura en Ingeniería Biomédica, Universidad Autonomía Metropolitana, Unidad Iztapalapa UAM.
- [ORL, 2002] ORL (2002) database of faces. AT&T Laboratories Cambridge; <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>. (Julio de 2010).
- [Oropeza, 2004] Oropeza Ortega J.C. (2004). *Reconocimiento de Rostros* Tesis de Maestría en Ciencias de Ingeniería en Microelectronica, Escuela Superior de Ingeniería Mecánica y Eléctrica sección de Estudios de Posgrado e Investigación unidad Culhuacan IPN.
- [Orozco, 2008] Orozco G. A.; Álvarez M. y Fetecua V. Juan (2008); "Reconocimiento de expresiones faciales utilizando análisis de componentes principales kernel (KPCA)", *Scientia Et Technica*, Vol. XIV, Núm. 38, junio 2008, pp. 43-48. Universidad Tecnológica de Pereira Colombia.

- [Paz, 2004] Paz Paredes J.I. (2004). *Construcción de una Red Neuronal Artificial para la Modelación de un Reactor de Hidrodesulfuración de Naftas* Tesis de Maestría en Ciencias con Especialidad en Ingeniería Química, Escuela Superior de Ingeniería Química e Industrias Extractivas IPN.
- [Phillips, 2000] Phillips P.J., H. Moon, S. A. Rizvi, y P. J. Rauss (2000). *The FERET evaluation methodology for face-recognition algorithms*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(10):1090–1104.
- [Ramírez, 2009] Ramírez V. L. y Hasimoto B. R. (2009). “*3D-Facial Expression Synthesis and its Application to Face Recognition Systems*”. Journal of Applied Research and Technology, Vol. 7, Núm. 3, Diciembre, 2009, pp. 323-339.
- [Reyes, 2009] Reyes de Luna R.D. (2009). *Aplicación de la transformada wavelet en el análisis de calidad de la energía*. Tesis de Licenciatura en Ingeniero Electricista, Escuela Superior de Ingeniería Mecánica y Eléctrica, Unidad Profesional Adolfo López Mateos IPN.
- [Rodríguez, 2005] Rodríguez Quintanar R.M. (2005). *Reconocimiento de enfermedades(creutzfeldt-jakob, alzheimer) mediante dimensión fractal y ondaletas usando una red neuronal artificial* Tesis de Maestría en Microelectrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica IPN.
- [Salazar, 2006] Salazar G. E.J y Ortega L. O. (2006). “*Sistema de búsqueda personalizada y recomendación de documentación científica*”. Revista Iberoamericana de Inteligencia Artificial, vol. 10, número 030, pp. 25-42
- [Salinas, 2000] Salinas, R.; Larraguibel L. (2000). “*Red neuronal de arquitectura paramétrica en reconocimiento de rostros*”. Electronic Spreading of Science Knowledge. Vol 17.
- [Sánchez, 2004] Sánchez Pérez G. (2004). *Optimización de estructuras de redes neuronales y aplicaciones*. Tesis Doctoral en Comunicaciones y Electrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica, Sección de Estudios de Posgrado e Investigación unidad Culhuacan IPN.

- [Sánchez, 2008] Sánchez Ramírez J.L. (2008). *Investigación y Optimización de Compresión Basada en Wavelet con aplicación a imágenes Médicas*. Tesis de Doctorado en Ciencias y Comunicaciones y Electrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica, IPN.
- [Sankar, 2004] Sankar k. Pal, Pabitra Mitra. (2004). *Patter Recognition Algorithms for Data Mining*. New York: Washington, D.C.
- [ULSA, 2005] Universidad La Salle, Dirección de Investigación 2005, Consultada el 13 de Junio de 2009, en:
<http://www.ci.ulsa.mx/~egr/docencia/martescib/Presentaciones/ReconocimientoPatrones.pdf>
- [Urbano, 2004] Urbano Sánchez L.E. (2004). *Comprensión de imágenes usando Wavelets y la estimación de sus matrices de detalle*. Tesis de Maestría en Ciencias en Microelectrónica, Escuela Superior de Ingeniería Mecánica y Eléctrica, IPN.
- [Villa, 2005] Villa P.S.M. (2005). Sistema de Reconocimiento de Rostros. *Revista de la Universidad Peruana de Ciencias Aplicadas*, Perú, (10) 1-5.
- [Villegas, 2005] Villegas Quezada C. (2005). *Reconocimiento De Rostros Utilizando Análisis de Componentes Principales: Limitaciones del Algoritmo* Tesis de Maestría en Sistemas y Planeación, Universidad Iberoamericana.
- [Yukse, 2007] Yuksek K., Karlık B., (2007). Sistema de Reconocimiento de Olores mediante Redes Neuronales Artificiales. *Departamento de Ingeniería en Computación, Facultad de Ingeniería, Universidad de Fatih, Estambul, Turquía*.
- [Zhao, 2003] Zhao W., Chellappa R., P. J. Phillips, y A. Rosenfeld (2003). *Face recognition: A literature survey*. *ACM Comput. Surv.*, 35(4):399–458.

ANEXO 1. BASE DE DATOS UAEM TEXCOCO/VALLE DE CHALCO



ANEXO 2. Base de datos ORL.

La base de datos de ORL (ORL, 2002) cuenta con un conjunto de imágenes de rostros tomadas entre Abril de 1992 y Abril de 1994. La base de datos fue elaborada por AT&T Laboratories Cambridge en colaboración con el grupo de habla, visión y robótica del departamento de ingeniería de la Universidad de Cambridge.

La base de datos contiene 10 diferentes imágenes de 40 individuos distintos. Las imágenes de cada individuo fueron tomadas en diferentes sesiones. Las tomas incluyen variaciones de iluminación, expresión facial y el uso de anteojos. Todas las imágenes fueron tomadas contra un fondo homogéneo oscuro con el individuo en posición frontal.

Los archivos se encuentran en formato PGM, con una resolución de 92x112 pixeles con 256 niveles de gris por pixel. Las imágenes se encuentran organizadas en 40 directorios (un directorio por individuo), los cuales tienen nombres de la forma sX, donde $X = 1, \dots, 40$ indica el número de individuo. En cada uno de los 40 directorios hay 10 imágenes, las cuales tienen nombres de la forma Y.pgm, donde $Y = 1, \dots, 10$ es el número de imagen.

La base de datos ORL se puede descargar de forma gratuita de la página de internet de AT&T Laboratories Cambridge (ORL, 2002) (archivo comprimido, disponible en formatos tar y ZIP).

A continuación se anexan las imágenes que componen la base de datos de ORL.

ANEXO 2. Base de datos ORL.



ANEXO 2. Base de datos ORL.



ANEXO 2. Base de datos ORL.



ANEXO 2. Base de datos ORL.



ANEXO 3

T. Wavelet	Patrón Entrenamiento	Patrón Prueba	Épocas	No. de Capas Ocultas	No. Neuronas Capas Ocultas	No. Neuronas Capa Salida	Clasificación
db1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	96.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	96.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	94.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	96.6666%
	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	94.6666%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	96.6666%
db2	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	94.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	95.3300%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	90.6660%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	90.6666%
	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	88.0000%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	88.0000%
db3	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	92.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	88.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	81.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	80.0000%
db4	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	84.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	84.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	65.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	66.0000%
db5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	84.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	82.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	62.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	72.0000%
db6	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	98.6666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	86.6666%

db6	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	85.33333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	69.33333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	73.33333%
db7	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	88.66666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	88.66666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	81.33333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	77.33333%
db8	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	98.66666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	90.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	88.00000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	73.33333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	76.00000%
db9	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	93.33333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	94.66666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	89.33333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	88.00000%
db10	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	98.66666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	98.66666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	88.66666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	90.00000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	78.66666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	76.00000%
haar	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	94.66666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	94.66666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	93.33333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	94.66666%
	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	93.33333%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	94.66666%
sym1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	96.00000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	96.00000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	94.66666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	96.66666%

sym1	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	94.6666%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	96.6666%
sym2	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	94.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	95.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	86.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	89.3333%
sym3	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	86.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	92.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	81.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	86.6666%
sym4	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	98.6666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	98.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	90.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	92.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	82.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	84.0000%
sym5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	87.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	90.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	73.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	80.0000%
sym6	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	86.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	86.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	72.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	72.0000%
sym7	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	86.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	88.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	80.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	81.3333%
sym8	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	85.3333%

sym8	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	86.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	68.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	72.0000%
sym10	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	87.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	87.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	73.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	69.3333%
sym12	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	98.6666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	98.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	89.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	86.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	81.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	84.0000%
coif1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	92.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	91.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	76.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	82.6666%
coif2	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	83.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	84.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	60.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	69.3333%
coif3	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	92.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	91.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	77.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	78.6666%
coif4	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	88.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	91.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	81.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	81.3333%
coif5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%

coif5	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	86.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	89.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	74.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	80.0000%
bior1.1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	97.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	97.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	94.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	94.6666%
	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	93.3333%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	93.3333%
bior1.3	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	97.3333%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	97.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	92.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	92.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	85.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	84.0000%
bior1.5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	85.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	86.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	69.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	76.0000%
bior2.2	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	92.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	92.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	84.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	84.0000%
bior2.4	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	90.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	89.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	80.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	72.0000%
bior3.1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	80.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	82.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	70.6666%

bior3.1	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	70.6666%
bior3.5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	88.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	90.0000%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	76.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	70.6666%
rbior1.1	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	97.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	97.3333%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	94.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	96.0000%
	6,7,8,9,10	1,2,3,4,5	10000	2	30	15	94.6666%
	6,7,8,9,10	1,2,3,4,5	30000	2	30	15	94.6666%
rbior1.5	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	83.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	84.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	73.3333%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	70.6666%
rbior6.8	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	100.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	85.3333%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	88.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	68.0000%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	78.6666%
dmey	1,2,3,4,5	1,2,3,4,5	10000	2	30	15	86.6666%
	1,2,3,4,5	1,2,3,4,5	30000	2	30	15	86.6666%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	10000	2	30	15	60.0000%
	1,2,3,4,5	1,2,3,4,5,6,7,8,9,10	30000	2	30	15	64.6666%
	1,2,3,4,5	6,7,8,9,10	10000	2	30	15	46.6666%
	1,2,3,4,5	6,7,8,9,10	30000	2	30	15	42.6666%

ANEXO 4. VENTANA DE VISTA DE DATOS.

A continuación se presentan las vistas del diseño del sistema; la finalidad es buscar entender el funcionamiento para el usuario, las tareas que se necesitan llevar a cabo para facilitar la tarea de reconocimiento facial.

Cada ventana contiene elementos los cuales serán descritos detalladamente dentro de la ventana de datos.

1 Ventana de vista de datos: Pantalla inicio del sistema.

Pantalla de inicio del sistema el cual permite la inicialización o salida del sistema.



Descripción: Etiquetas.

Hora: Leyenda que indica la hora en la cual fue iniciado el sistema.

Fecha: Leyenda que indica la fecha actual en la cual fue iniciado el sistema.

Versión: Leyenda que indica la versión del sistema.

ANEXO 4. Ventana de vista de datos.

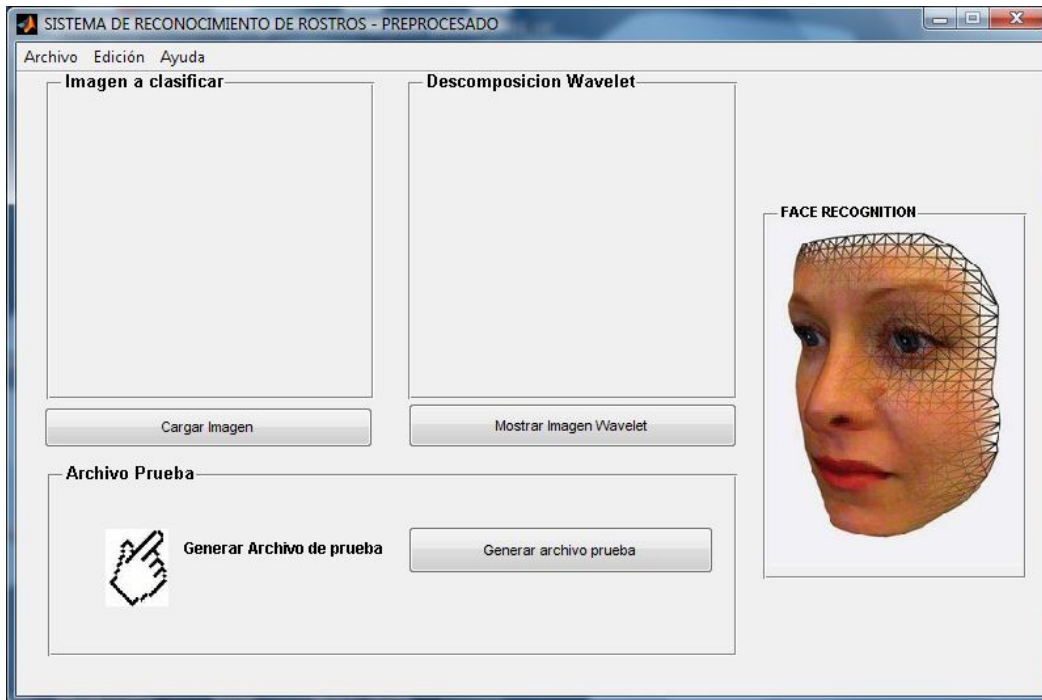
Descripción: Botones.

Inicio: Botón el cual da inicio a la clasificación del sistema; de inmediato mostrando la venta de **Pre-procesado**.

Salida: Botón que permite la salida normal del sistema.

2 Ventana de vista de datos: Pantalla Pre-procesado.

Pantalla en la cual el usuario elegirá el rostro a reconocer, al mismo tiempo se podrá observar la descomposición Wavelet de dicha imagen, para finalizar con la generación del archivo prueba, el cual contiene la información necesaria para el reconocimiento de rostros.



Descripción: Frames.

Imagen a clasificar: Muestra al usuario la imagen facial a reconocer.

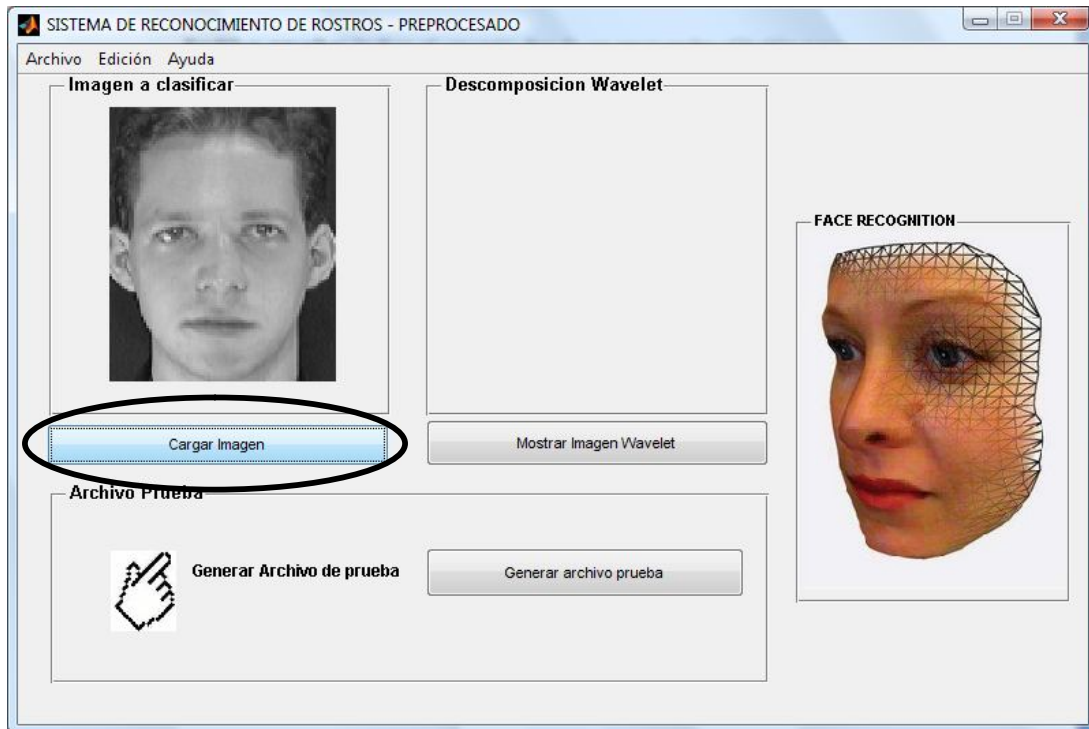
Descomposición Wavelet: Muestra al usuario la imagen facial, aplicándole la descomposición Wavelet.

Archivo prueba: Indica al usuario donde se encuentra el botón para generar el archivo prueba.

ANEXO 4. Ventana de vista de datos.

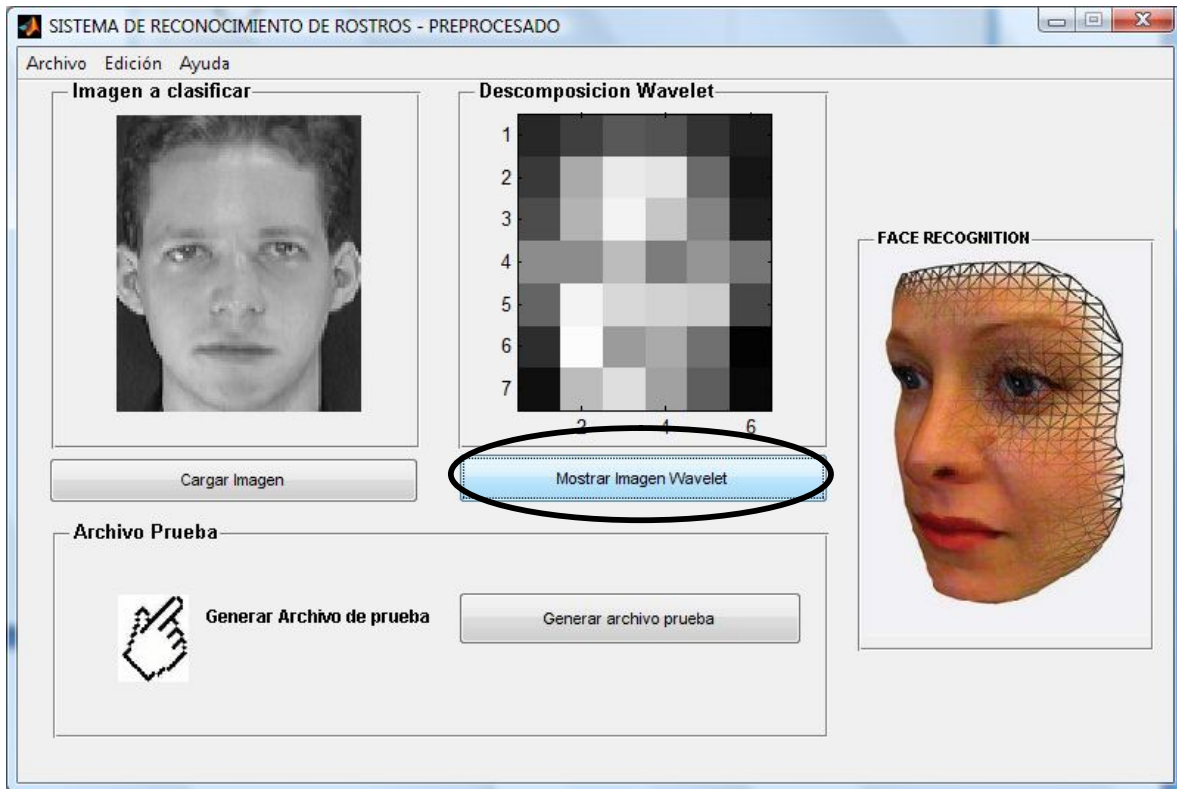
Descripción: Botones.

Cargar Imagen: Componente donde el usuario podrá elegir la imagen facial a identificar, pudiendo elegir de entre un conjunto de rostros; dichas imágenes se encuentran organizadas por ficheros.



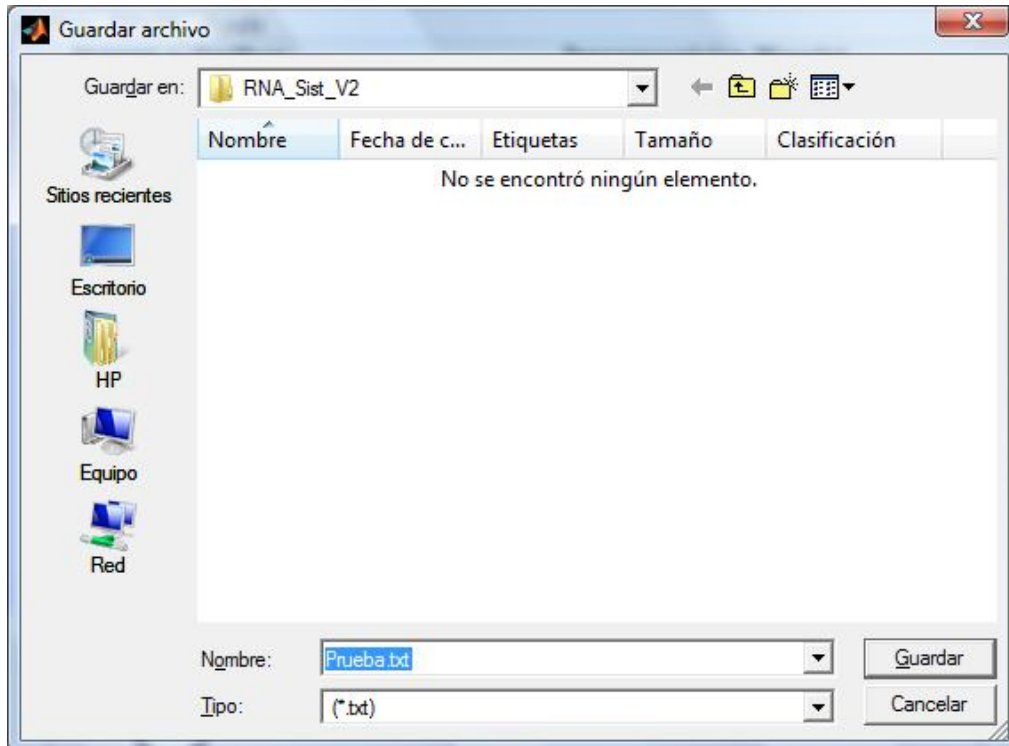
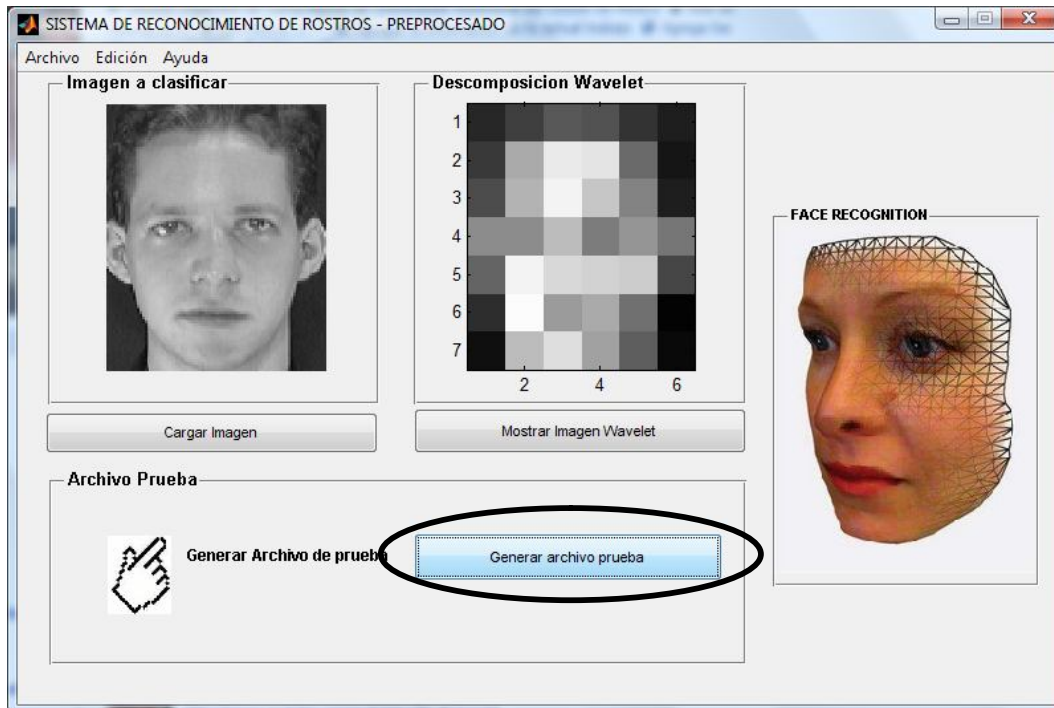
Mostrar Imagen Wavelet: Componente donde el usuario podrá generar la imagen facial descompuesta para su posterior reconocimiento. Este paso no debe de realizarse sin antes no haber seleccionado una imagen a reconocer; como se mostro en la pantalla anterior.

ANEXO 4. Ventana de vista de datos.



Generar archivo prueba: Componente donde el usuario podrá generar el archivo prueba, el cual contiene toda la información necesaria para el reconocimiento; el archivo prueba es leído por la red neuronal artificial, la cual se ejecuta de inmediato solo cuando el archivo prueba es generado; de alguna otra manera la red neuronal no realizara ninguna tarea.

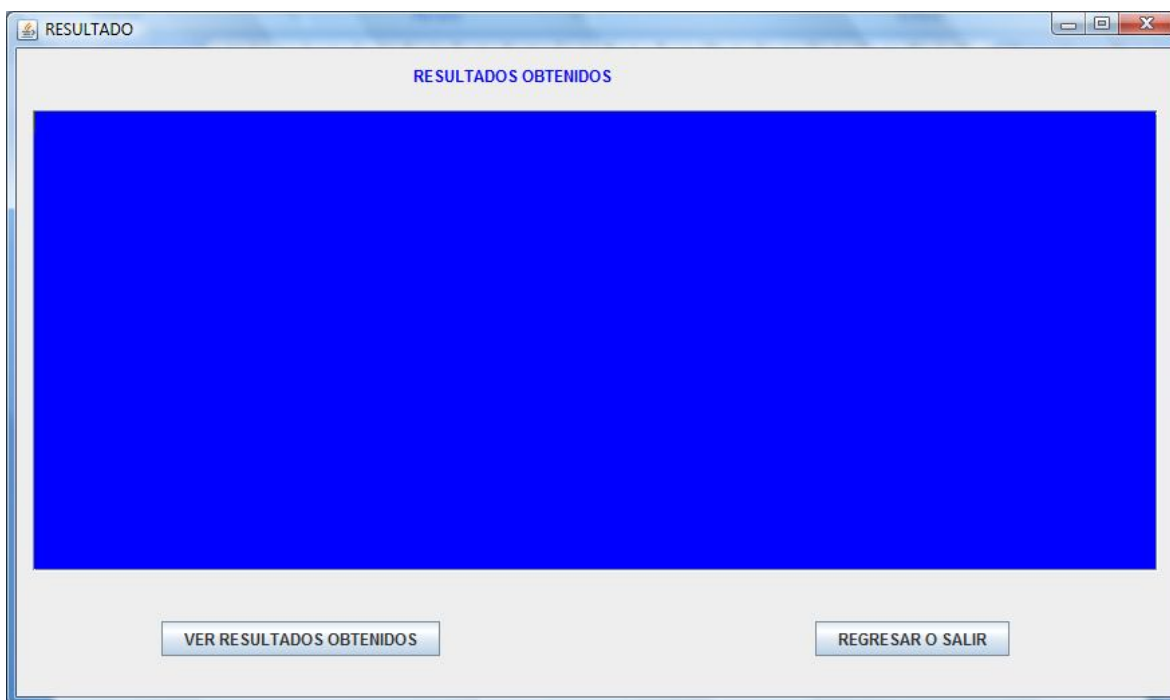
ANEXO 4. Ventana de vista de datos.



ANEXO 4. Ventana de vista de datos.

3 Ventana de vista de datos: Pantalla de resultados RNA.

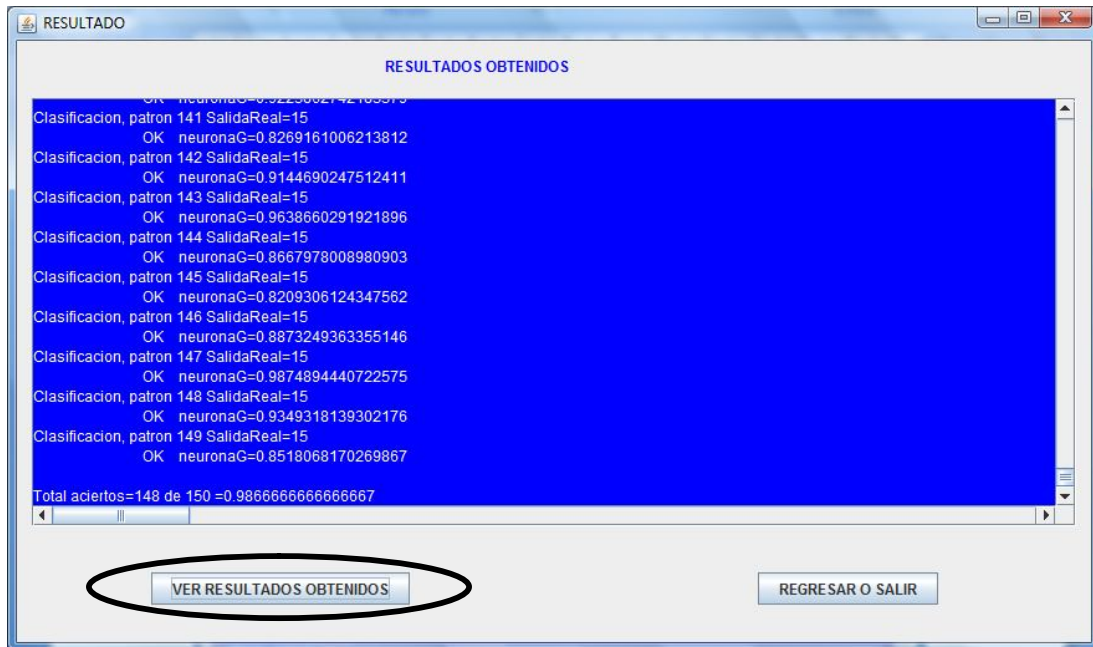
Pantalla en la cual el usuario podrá observar los resultados obtenidos en la clasificación por la RNA, la pantalla de resultados solo se desplegara si el archivo prueba es generado satisfactoriamente, por el contrario la pantalla no mostrara ningún resultado.



Descripción: Botones.

Ver resultados obtenidos: Componente donde el usuario podrá observar los resultados obtenidos por la RNA, los cuales serán interpretados por el usuario.

ANEXO 4. Ventana de vista de datos.



Regresar o Salir: Componente donde el usuario podrá regresar a la pantalla inicial del sistema; y realizar una salida satisfactoria del sistema.

