



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIA E INGENIERÍA DE LA COMPUTACIÓN

**“SISTEMA DE CLASIFICACION DE IMAGENES DE
MICROTORNILLOS BASADO EN UNA RED
NEURONAL ARTIFICIAL PARA ASISTIR PROCESOS
DE CORTE EN MICROMAQUINAS”**

T E S I S

QUE PARA OBTENER EL GRADO DE:

**MAESTRO EN CIENCIAS
(COMPUTACIÓN)**

P R E S E N T A:

ANABEL MARTIN GONZALEZ

DIRECTOR DE TESIS: “DRA. TETYANA BAYDYK”

México, D.F.

Noviembre 2006



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Le doy gracias a Dios por permitirme concretar, de manera exitosa y gratificante, una importante etapa de mi vida profesional, apoyándome y cuidándome durante este largo camino de estudios.

A mis padres, Guillermo y Rita, por todo su amor, comprensión, paciencia, cariño y sacrificio. Gracias por brindarme todo su apoyo, oraciones y ánimos para salir adelante en la vida y con este proyecto. Sin ustedes este trabajo no hubiera sido posible. Gracias hoy y siempre. Los quiero con todo mi corazón.

A mis hermanos, Carlos y Guillermo, que siempre los llevo en mi mente y corazón.

De manera muy especial, al Dr. Carlos J. Castilla Centeno, quien con su ejemplo he aprendido que la persistencia y la disciplina son las características esenciales que forjan a un hombre extraordinario en el área profesional. Gracias por todo su apoyo y consejos.

Mi más sincero agradecimiento y admiración a mi asesora de tesis, Dra. Tetyana Baydyk, quien gracias a su invaluable apoyo y orientación he podido finalizar este proyecto. Gracias por abrirme las puertas de su laboratorio y corazón. Fue un placer trabajar con usted.

Un especial agradecimiento a las personas que invirtieron su tiempo para la revisión de este trabajo, mis sinodales: Dra. María Elena Martínez Pérez, Dr. Boris Escalante Ramírez, Dr. Ernest Kussul y Dr. Ernesto Bribiesca Correa. Les agradezco grandemente todos sus comentarios y aportaciones realizadas.

A mis amigos Karen y Carlos, quienes me acompañaron durante este proceso de formación llenando de cariño y apoyo mi espíritu y corazón, y de quienes siempre obtuve palabras de ánimo para seguir adelante en momentos difíciles. Mil gracias.

A mis amigas Diana, Mercedes, Beba y Cinthia quienes siguen siendo un apoyo fundamental en mi vida personal. Las quiero con todo mi corazón.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por todo el apoyo brindado para la realización de mis estudios en el Posgrado en Ciencia e Ingeniería de la Computación.

Al Centro de Ciencias Aplicadas y Desarrollo Tecnológico (CCADET) por su apoyo para el desarrollo de este proyecto.

Al Posgrado en Ciencia e Ingeniería de la Computación, a su coordinador, Dr. Boris Escalante Ramírez y profesores investigadores que laboran en él, les agradezco profundamente el haber compartido sus experiencias e invaluable conocimientos que con mucho esfuerzo han adquirido a lo largo de los años.

Finalmente, a la Universidad Nacional Autónoma de México (UNAM) por abrirme sus puertas y brindarme la oportunidad de crecer profesionalmente. Gracias.

Índice General

Agradecimientos	II
Resumen	III
Abreviaturas	IV
Indice de Figuras	IX
Indice de Tablas	XI
1. Introducción	1
1.1. Micromáquinas	2
1.2. Procesos de Corte en Micromáquinas	3
1.3. Objetivo del Trabajo	5
1.4. Organización de la Tesis	5
2. Microtecnologías	7
2.1. Sistemas MicroElectroMecánicos (MEMS)	7
2.2. Micro y Nano Tecnología (MNT)	9
2.3. Tecnología de MicroMáquinas (MMT)	12
2.4. Tecnología de MicroEquipo (MET)	15
3. Conceptos de Redes Neuronales Artificiales	19
3.1. La Neurona Biológica	19
3.2. La Neurona Artificial	21
3.2.1. Tipos de Funciones de Activación	23

3.3.	Redes Neuronales Artificiales	26
3.3.1.	Propiedades	27
3.3.2.	Componentes	28
3.3.3.	Arquitecturas de Redes Neuronales	28
3.3.4.	Proceso de Aprendizaje o Entrenamiento	29
3.3.5.	Algunos Campos de Aplicación	33
3.4.	Modelo del Perceptrón	34
3.4.1.	Arquitectura	35
3.4.2.	Aprendizaje	36
4.	Descripción de la Metodología	37
4.1.	Descripción y Clasificación de Microtornillos	37
4.2.	Descripción de las Imágenes Digitales de Microtornillos	38
4.3.	Creación del Conjunto de Imágenes de Entrenamiento y Prueba	38
4.4.	Desarrollo e Implementación de la Red Neuronal Artificial	40
4.5.	Entrenamiento y Prueba de la Red Neuronal Artificial	40
4.6.	Implementación del Sistema	41
5.	Clasificador Neuronal: LIRA Grayscale	42
5.1.	Arquitectura	42
5.2.	Conexiones	44
5.3.	Activación de Neuronas	45
5.4.	Proceso de Entrenamiento	45
5.4.1.	Fase Inicial	47
5.4.2.	Método de Selección del Ganador	47
5.4.3.	Adaptación de Pesos	47
5.5.	Optimización del Proceso de Entrenamiento	48
5.5.1.	Codificación de Imágenes	48
5.5.2.	Transformaciones Geométricas de Imágenes	49
5.6.	Proceso de Prueba	49
5.6.1.	Selección del Ganador en el Proceso de Prueba	50

6. Resultados	51
6.1. Resultados con Base de Datos de 40 Imágenes	51
6.1.1. Captura de Imágenes Digitales	51
6.1.2. Experimentos Preliminares	52
6.1.3. Proceso de Entrenamiento	56
6.1.4. Proceso de Prueba	56
6.2. Resultados con Base de Datos de 440 Imágenes	58
6.2.1. Captura de Imágenes	58
6.2.2. Experimentos Preliminares	59
6.2.3. Proceso de Entrenamiento	63
6.2.4. Proceso de Prueba	64
6.3. Velocidad de Clasificación	66
7. Discusión	67
7.1. Tamaño de la Muestra de Entrenamiento	67
7.2. Transformaciones Geométricas de Imagen	68
7.3. Entrenamiento Secuencial y Aleatorio	68
7.4. Imágenes Mal Clasificadas	68
7.5. Restricción del Sistema	70
8. Conclusiones	71
8.1. Trabajo a Futuro	72
Bibliografía	73

Resumen

El presente trabajo de investigación propone el desarrollo de un sistema de clasificación de imágenes de microtornillos con base en una red neuronal artificial que permite reconocer la forma de la micropieza, producida por una micromáquina, con el fin de generar la información necesaria para controlar y optimizar el desempeño del proceso de manufactura. El clasificador neuronal empleado para la tarea de reconocimiento es denominado Limited Receptive Area Grayscale (LIRA Grayscale) cuya arquitectura esta basada en el modelo del perceptrón de Rosenblatt. El sistema de visión fue probado fuera de línea con imágenes de 40 microtornillos de 3 *mm* de diámetro obteniendo un porcentaje de reconocimiento de 98.90% en el mejor de los casos. Este trabajo está motivado bajo la idea de asistir al proceso de manufactura de microtornillos para obtener micropiezas producidas con un alto grado de calidad.

Capítulo 1

Introducción

El reconocimiento de la forma de objetos siempre ha sido un problema relevante en el campo de visión computacional [1]. En el área de la micromecánica, el reconocimiento de forma es una tarea importante relacionada con varios procedimientos empleados en microfábricas. Por mencionar algunos ejemplos de estos tipos de procesos, se encuentran el ensamble de microdispositivos [2], la manufactura de micropiezas [3, 4], el reconocimiento y ubicación de micropiezas [5], entre otros.

En el Laboratorio de Micromecánica y Mecatrónica¹ se propuso una metodología para la producción de microdispositivos mecánicos denominada Tecnología de MicroEquipo (MET). El concepto fundamental de esta propuesta es la creación de generaciones de microequipo, es decir, el desarrollo de máquinas y dispositivos mecánicos (primera generación) que produzcan máquinas y dispositivos más pequeños (segunda generación), los cuales generarán máquinas aún más pequeñas (tercera generación) y así sucesivamente.

El problema principal en el desarrollo de microequipo se encuentra relacionado con el diseño de máquinas que posean un control automatizado de sus procesos para obtener sistemas mecánicos completamente autónomos. Por lo tanto, es necesario el desarrollo de sistemas de visión computacional que permitan obtener la información del estado de los procesos de las máquinas y, con esto, generar una retroalimentación de las mismas para que puedan tomar las decisiones apropiadas en el desempeño de sus tareas.

¹ Este laboratorio pertenece al CCADET de la Universidad Nacional Autónoma de México.

Uno de los modelos en los que se pueden sustentar los sistemas de visión computacional para resolver este tipo de problemas es la red neuronal artificial, debido a que ésta representa una importante herramienta en áreas como el procesamiento de imágenes y reconocimiento de formas; y ha sido objeto de gran investigación [6] y [7]. En [6], [8], [9] y [10] se presentan reportes de aplicaciones de redes neuronales para el problema de reconocimiento de patrones, algunas de ellas poseen la característica de ser invariante al desplazamiento y rotación. Lo que la mayoría de estos métodos realiza es entrenar una red neuronal con patrones en todas las orientaciones y posiciones esperadas, de tal manera que la red se vuelva insensible a orientaciones específicas y cambios de posición.

1.1. Micromáquinas

Una micromáquina es una máquina de dimensiones en la escala de micrómetros² integrada por diminutos elementos funcionales que le permiten desempeñar delicadas y complicadas tareas.

Las micromáquinas tienen varios usos potenciales en diferentes áreas industriales, médicas, biológicas, etc. Las mayores ventajas que presentan las micromáquinas son las siguientes:

- Los diversos campos de aplicación
- El bajo costo de los microdispositivos manufacturados, siempre que sean producidos en grandes cantidades
- Las reducidas dimensiones (escala de micrómetros)
- Su capacidad de dar solución efectiva a varios problemas

Las micromáquinas también tienen desventajas. Dado que son muy pequeñas, son difíciles de manipular y, en ocasiones, puede ser complicado trabajar con ellas.

Las micromáquinas son los componentes principales que integran una microfábrica, la cual se refiere a un sistema miniatura de producción capaz de fabricar piezas a

² Un micrómetro (μ) es una millonésima parte de un metro ($1 \mu = 1 \times 10^{-6} m$).

una escala menor en relación con el tamaño del propio equipo. Este tipo de micro-sistema proporciona ahorros considerables en términos de energía, espacio y costos a comparación de los sistemas de manufactura de tamaño convencional.

Uno de los principales problemas en el diseño de microfábricas es la automatización con base en sistemas de visión computacional. Existen diferentes métodos para construir un sistema de visión computacional para diversos propósitos en micromecánica (ver [11], [12] y [13]).

1.2. Procesos de Corte en Micromáquinas

Uno de los componentes micromecánicos producidos por micromáquinas (ej. micro-torno) desarrollados bajo la tecnología MET es el microtornillo. Estos son componentes necesarios para construir mecanismos miniatura de máquinas. El problema relacionado con el proceso de maquinado de estas micropiezas es que en algunas de ellas se pueden generar defectos en su forma. Estos errores están relacionados con la posición de la herramienta de corte del microtornillo.

En la Figura 1.1 se muestran las diferentes posiciones de la herramienta de corte estudiadas en este trabajo. Se puede notar los cambios de forma que experimenta la materia prima (microtornillo) dependiendo de la ubicación del cortador empleado.

Debido a las reducidas dimensiones que poseerán los objetos elaborados por las micromáquinas (desde milímetros hasta micrómetros) será imposible utilizar instrumentos ordinarios (a gran escala) para el control de los procesos de las mismas.

Una de las formas en las que se ha tratado de lograr el control automático de esta tarea de corte ha sido con la implementación de un sistema de visión computacional basado en una red neuronal artificial denominada Permutation Coding Neural Classifier (PCNC) [3], el cual fue evaluado y mostró resultados aceptables (un porcentaje de reconocimiento de 92.5% con diferentes posiciones de la herramienta de corte), aunque con la posibilidad de poder ser mejorados.

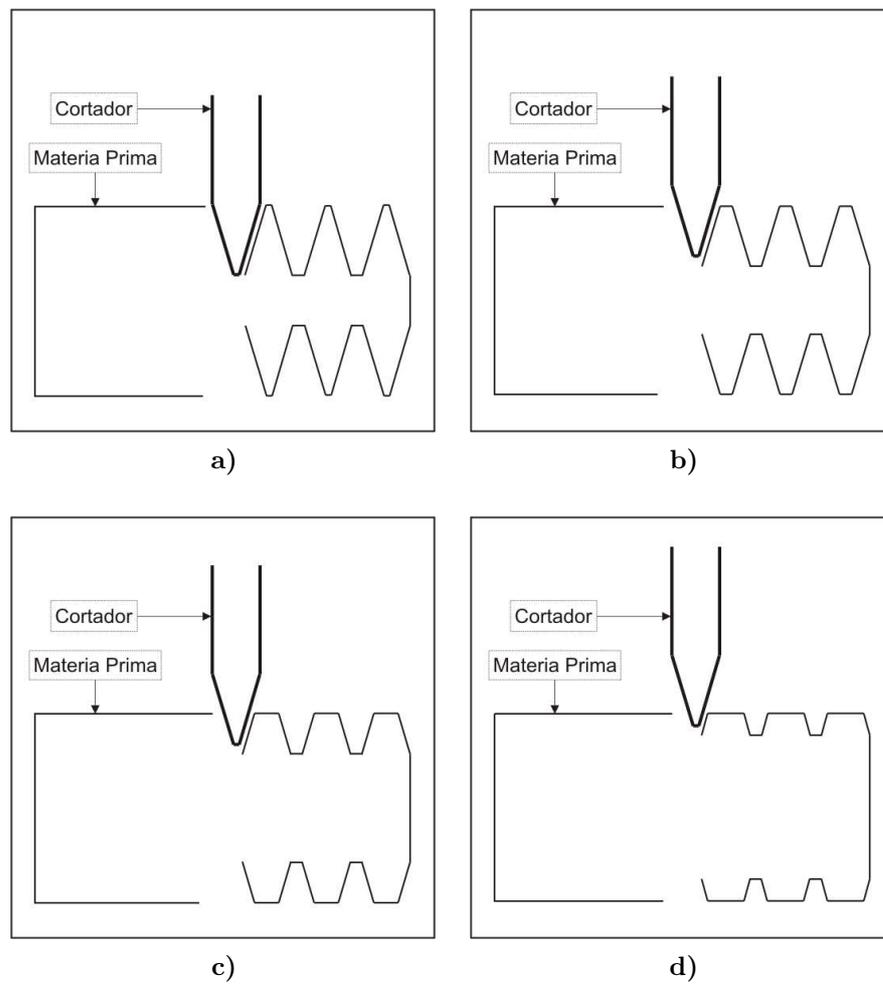


Figura 1.1: Posiciones de la herramienta de corte de microtornillos en relación a la pieza de trabajo. a) Cortador ubicado 0.1 mm por debajo de la posición correcta, b) Cortador en posición correcta, c) Cortador ubicado 0.1 mm por encima de la posición correcta y d) Cortador ubicado 0.2 mm por encima de la posición correcta.

1.3. Objetivo del Trabajo

En el presente trabajo se propone el desarrollo de un sistema de visión computacional, basado en una red neuronal artificial denominada Limited Receptive Area Grayscale (LIRA Grayscale), para asistir al proceso de fabricación, previamente mencionado, de microtornillos manufacturados por micromáquinas.

El objetivo del sistema a desarrollar será el reconocimiento y clasificación de la forma de los microtornillos producidos, brindando de esta manera, la información requerida para la corrección, de ser necesaria, del proceso de fabricación de los mismos, dependiendo de los errores de manufactura detectados. De esta forma, se podrán obtener micropiezas elaboradas con un alto grado de calidad. Se espera que el sistema propuesto mejore los resultados obtenidos en los trabajos previos (consultar [3]).

1.4. Organización de la Tesis

El presente documento se encuentra organizado en un conjunto de ocho capítulos en total, siendo éste el primero de ellos.

En el capítulo 2 se realiza una breve descripción de las microtecnologías que se han ido desarrollando en diferentes partes del mundo, mostrando los diversos enfoques y características que las identifican.

En el capítulo 3 se presentan algunas definiciones y propiedades básicas sobre redes neuronales artificiales que servirán de preámbulo para el entendimiento de los conceptos a los que se hace referencia a lo largo de este documento. Asimismo, se describe, de manera breve, el modelo de red neuronal artificial empleado como base para el diseño y construcción del clasificador propuesto en este trabajo.

Los pasos que constituyen el desarrollo de esta tesis, así como su respectiva implementación se encuentran explicados en el capítulo 4.

A continuación, en el capítulo 5 se presenta a detalle el modelo de red neuronal empleado para resolver el problema propuesto en el capítulo 1, incluyendo su arquitectura, reglas de aprendizaje y procedimientos de entrenamiento y validación. Inmediatamente

después se muestran los resultados obtenidos con la metodología implementada (capítulo 6).

Posteriormente, se realiza una serie de observaciones acerca de los resultados obtenidos en el trabajo desarrollado (capítulo 7), y finalmente, en el capítulo 8 se resume la contribución de este trabajo y se proponen algunas extensiones posibles para el sistema desarrollado.

Capítulo 2

Microtecnologías

En distintas áreas de la ciencia y la tecnología el desarrollo de piezas, dispositivos y sistemas miniatura ha ido evolucionando de manera continua en los últimos años. Esta tendencia es debido a que la producción de microdispositivos ha presentado diversas ventajas, como son: la reducción de espacios, la disminución en el consumo de energía, la minimización de los costos de producción, entre otras.

Actualmente, las tecnologías de desarrollo de microdispositivos se encuentran presentes en diferentes partes del mundo. Ejemplo de éstas son: MicroElectroMechanical Systems (MEMS) en Estados Unidos; MicroSystems Technology (MST) y Micro and Nano Technology (MNT) en Europa; MicroMachine Technology (MMT) en Japón y MicroEquipment Technology (MET) en México.

A continuación se presentará una descripción de cada una de las microtecnologías mencionadas.

2.1. Sistemas MicroElectroMecánicos (MEMS)

Se denomina MicroElectroMechanical Systems (MEMS) o sistemas microelectromecánicos a una tecnología que explora la infraestructura existente de la microelectrónica para crear sistemas complejos a partir del desarrollo de componentes mecánicos cuyas dimensiones, generalmente, varían desde un micrómetro (una millonésima parte de un metro) hasta un milímetro (una milésima parte de un metro) [14].

La investigación sobre MEMS comenzó a fines de la década de los 60, pero los primeros dispositivos MEMS comerciales aparecieron recién a principios de los 90, utilizados en la industria automotriz como detectores de colisión para la implementación de bolsas de aire. Posteriormente en la misma década, otras industrias comenzaron a reconocer el potencial de esta tecnología para reducir los costos y las dimensiones de los sistemas.

Esta tecnología, también conocida en Europa como Tecnología de MicroSistemas (MST), está comúnmente basada en circuitos integrados de silicio¹. Los componentes micromecánicos son fabricados grabando, selectivamente, partes de una placa de silicio y/o añadiendo capas estructurales para formar los dispositivos mecánicos y electromecánicos.

Además del silicio, los dispositivos MEMS pueden ser creados utilizando metales como el oro, níquel, aluminio, cromo, titanio, plata, entre otros.

Las aplicaciones más importantes de los microdispositivos desarrollados bajo tecnología MEMS son:

- Biotecnología: Biochips para la detección de agentes químicos y biológicos peligrosos, microistemas para la detección y selección de fármacos.
- Comunicaciones: Interruptores mecánicos y ópticos.
- Sistemas de seguridad: Microcandados.
- Acelerómetros. Estos microdispositivos miden la aceleración y son empleados para detectar colisiones automovilísticas. Un procesador analiza la magnitud de la aceleración y decide si desplegar o no las bolsas de aire del vehículo.
- Sensores de presión, de temperatura, químicos y de vibración.
- Microactuadores.

Entre otras aplicaciones de MEMS se encuentran: giroscópios, reflectores de luz, control de vehículos, marcapasos y juegos. La tecnología es usada también para ha-

¹ El silicio es un material semiconductor fundamental para el desarrollo de la electrónica de estado sólido actual.

cer cabezas de impresoras de inyección de tinta y microactuadores para cabezas de lectura/escritura.

En la Figura 2.1 se pueden apreciar algunos ejemplos de microdispositivos de tecnología MEMS.

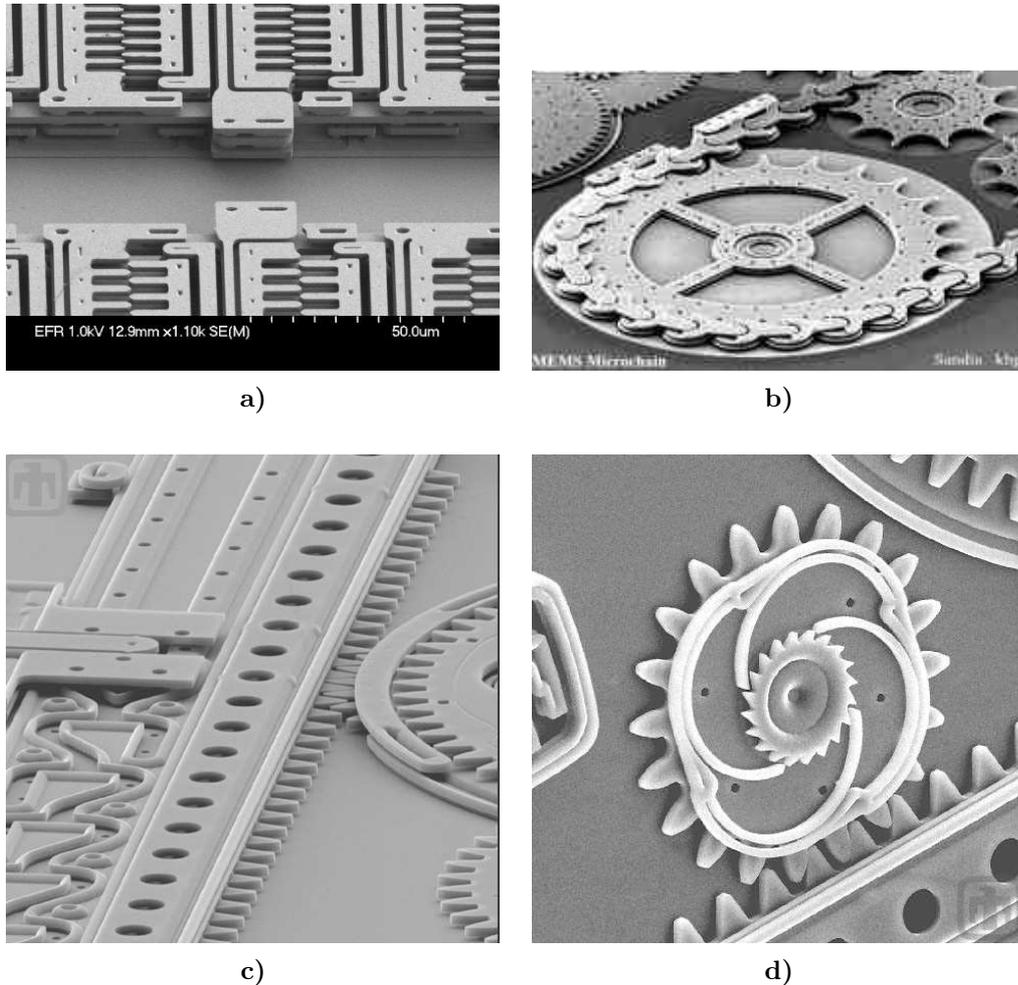


Figura 2.1: Microdispositivos MEMS. a) Microactuador, b) Microengrane con microcadena, c) Microcandado y d) Microengranes. (Sandia National Laboratories, <http://mems.sandia.gov/scripts/index.asp>).

2.2. Micro y Nano Tecnología (MNT)

La miniaturización hacia micro y nano niveles no sólo conduce a productos más pequeños y reducción de costos, sino también permite la adquisición de dispositivos con nuevas funcionalidades que no se pueden obtener en un macro nivel. Estas fun-

cionalidades son consecuencia de los efectos físicos y químicos inherentes a las pequeñas dimensiones.

El nanómetro (nm) es la unidad de longitud que equivale a una milmillonésima parte de un metro ($1\ nm = 1 \times 10^{-9}\ m$), aproximadamente el diámetro de tres o cuatro átomos. Para tener una idea de esta unidad, el cabello humano tiene en promedio 50,000 nanómetros de ancho.

La nanotecnología es un conjunto de técnicas que se utilizan para la creación de materiales funcionales, dispositivos y sistemas a través de la manipulación y control de la materia en la escala nanométrica (1–100 nanómetros), esto es, en la escala de átomos y moléculas. Asimismo, la nanotecnología pretende explotar nuevos fenómenos y sus características (físicas, químicas, biológicas) en dimensiones nanoscópicas.

Las nanofábricas estarán compuestas de nanomáquinas. A diferencia de los millones de millones de átomos que constituyen una micromáquina, las nanomáquinas constarán de millones de átomos. Estos nanodispositivos estarán controlados por computadoras. En la Figura 2.2 se puede observar una impresión artística de una nanofábrica de sobremesa (cortesía de John Burch, Lizard FIRE Studios [15]).

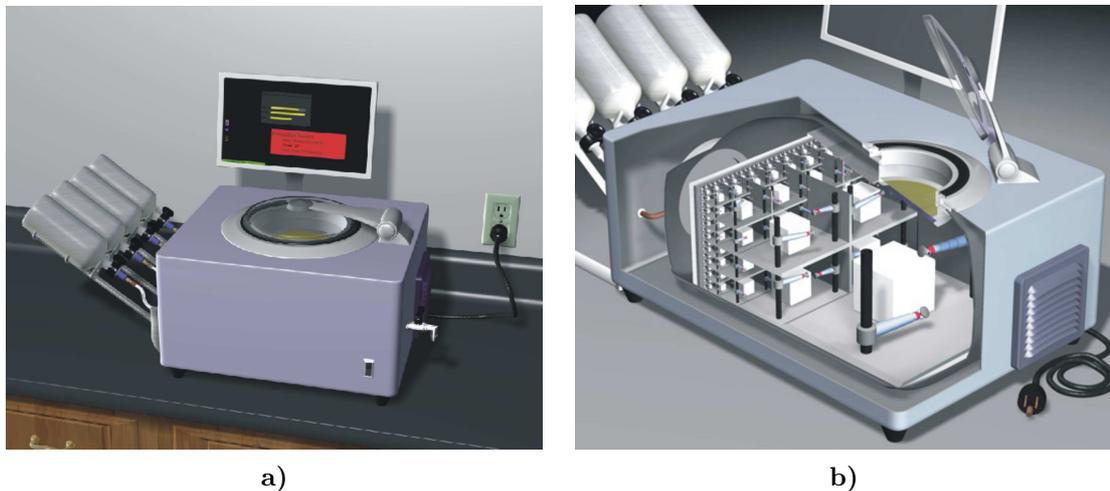


Figura 2.2: Nanofábrica de sobremesa. a) Vista exterior y b) Vista interior.

Una de las aplicaciones más importantes de la nanotecnología será la nanorobótica médica o nanomedicina. La capacidad de diseñar y construir una gran cantidad de nanorobots médicos hará posible la eliminación rápida de enfermedades y la recu-

peración confiable y relativamente sin dolor de traumas físicos. Los nanorobots médicos podrán hacer posible también la corrección de defectos genéticos, y ayudarán a asegurar un tiempo de vida saludable. Asimismo, los nanorobots médicos se podrán utilizar para aumentar capacidades humanas naturales.

Entre otras aplicaciones de la nanotecnología se encuentran: medicina molecular y biotecnología; diagnóstico de enfermedades; sistemas de administración de fármacos; almacenamiento, producción y conversión de energía; desarrollo de materiales avanzados; dispositivos electrónicos más pequeños, rápidos y baratos; tratamiento de aguas; remediación de la contaminación atmosférica; detección y control de plagas; aplicaciones relacionadas con las comunicaciones, la exploración del espacio, la informática, entre otras.

La característica fundamental de la nanotecnología es que constituye un ensamblaje interdisciplinario de varios campos de las ciencias naturales que están altamente especializados. Los físicos juegan un rol importante no sólo en la construcción del microscopio a emplear para investigar los fenómenos, sino también sobre todas las leyes de la mecánica cuántica que intervienen. Alcanzar la estructura del material deseado y las configuraciones de ciertos átomos hacen jugar a la química un papel importante. En medicina, el desarrollo específico de nanopartículas promete ayudar al tratamiento de diversas enfermedades. Aquí, la ciencia ha alcanzado un punto en el que las fronteras que separan las diferentes disciplinas han empezado a diluirse, y es precisamente por esa razón por la que la nanotecnología también se refiere a ser una tecnología convergente.

La mayor parte de los países europeos tienen programas nacionales de nanotecnología. En el 2004, la inversión total en nanotecnología realizada en Europa fue igual o levemente por encima a la generada por Estados Unidos de Norteamérica [16].

En Alemania, la nanotecnología está enfocada en nanoelectrónica, nanomateriales, nano-óptica, microsistemas, nanobiotecnología y manufactura. Nueve centros de nanotecnología se han fundado en este país: Nanomaterials (Karlsruhe), Ultraprecision surface engineering (Brunswick) y nano coatings (Dresde), Nano-optics (Berlín), Nanobio-technology (Múnich y Kaiserslautern), Nanochemistry (Sarrebuck), Hanse Nanotec

(Hamburgo) y CeNtech (Münster).

En el 2003, el Reino Unido lanzó Micro and Nanotechnology Initiative para crear una red de instalaciones destinadas para la micro y nanotecnología. El Reino Unido tiene actualmente 1500 trabajadores en el área MNT. Los centros reconocidos de nanotecnología están en las universidades de Oxford, Cambridge, Newcastle, Durham y Glasgow. Un centro de producción especial de nanomateriales está presente en Farnborough.

En Francia, la estructura de investigación para la nanotecnología se basa alrededor de un grupo de cinco centros de excelencia situados en Grenoble, Toulouse, Marcoussis, Orsay y Lille. Sus áreas de prioridad son:

- Micro y nanoelectrónica
- Opto-electrónica
- Microsistemas y ensamble
- Biotecnología e instrumentación

En Suiza, se han desarrollado los detalles del control de precisión de movimiento y los principios de micromanipulación para futuras microfábricas [17].

2.3. Tecnología de MicroMáquinas (MMT)

La Tecnología de MicroMáquinas (MMT), a diferencia de las tecnologías MEMS y MST, está más enfocada hacia la microingeniería y las micromáquinas para la producción de dispositivos mecánicos, tales como micromotores, microengranes, microbombas y microrobots para aplicaciones médicas e industriales.

En Japón se inició el primer proyecto para la creación de una microfábrica [18]. Un grupo de investigación de micromáquinas del Laboratorio de Ingeniería Mecánica desarrolló una microfábrica de sobremesa [19, 20], el tipo más básico de microfábrica (ver Figura 2.3, cortesía de [19]). Asimismo, realizaron una estimación del efecto de miniaturización de los sistemas de producción. Esta aproximación estableció que al

reducir 1/10 el tamaño de las máquinas de producción, el total de consumo de energía en la fábrica decrementaría aproximadamente 1/100 al de una fábrica convencional.

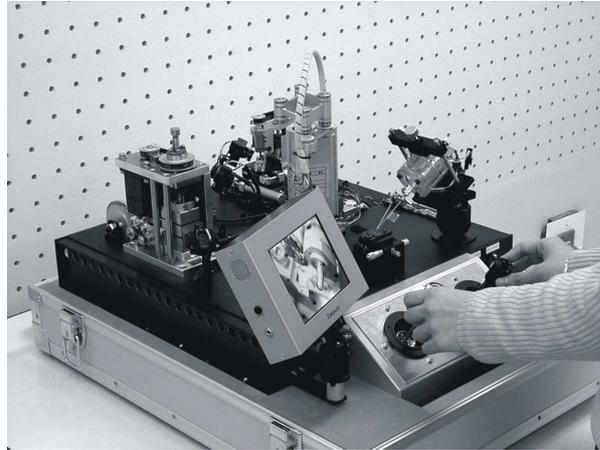


Figura 2.3: Microfábrica japonesa.

La microfábrica desarrollada puede ser dividida en dos secciones básicamente; el área de maquinado y el área de ensamble. Cada una de estas áreas está compuesta de micromáquinas: tres micromáquinas en maquinado (un torno, una fresadora y una prensa) y dos micromáquinas en ensamble (un brazo de transferencia y una mano de dos dedos) (ver Figura 2.4, cortesía de [19]). Estas micromáquinas poseen sistemas de control numérico para su funcionamiento.

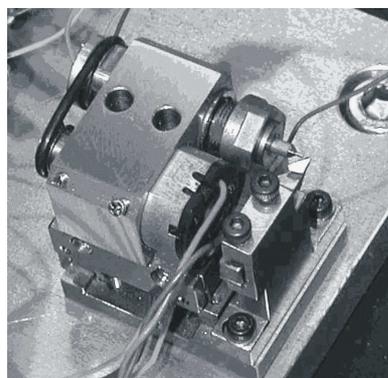
El microtorno tiene 32.0 *mm* de largo, 25.0 *mm* de ancho y 30.5 *mm* de alto, y pesa 100 *g* (Figura 2.4(a)).

Las dimensiones de la microfresadora son 119 *mm* de largo, 119 *mm* de ancho y 102 *mm* de alto (Figura 2.4(b)). Esta micromáquina se encuentra equipada con una cámara CCD miniatura de 7 *mm* de diámetro que permite monitorear el área de corte de las micropiezas de trabajo.

La microprensa mide 111 *mm* de largo, 66 *mm* de ancho y 170 *mm* de alto (Figura 2.4(c)).

La micromano de dos dedos tiene un cuerpo cilíndrico que mide 48 *mm* de diámetro y 65 *mm* de largo con dos agujas de vidrio como dedos de 50 *mm* de largo (Figura 2.4(d)).

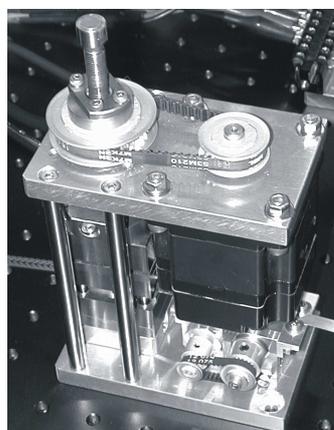
El microbrazo de transferencia tiene 200 *mm* de altura. Este brazo puede trabajar



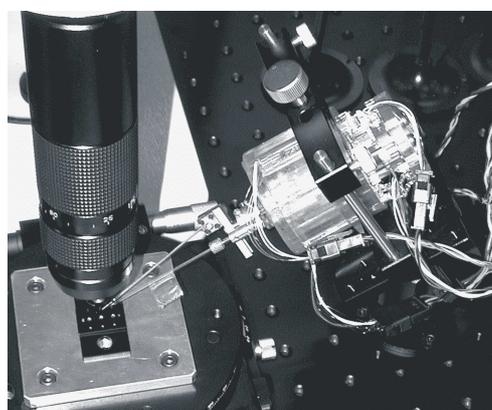
a)



b)



c)



d)



e)

Figura 2.4: Micromáquinas de la microfábrica japonesa. a) Microtorno, b) Microfresadora, c) Microprensa, d) Micromanipulador y e) Microbrazo de transferencia.

en un círculo de 200 *mm* de diámetro (Figura 2.4(e)).

Los componentes de esta microfábrica de sobremesa están integrados en una caja portable con dimensiones externas de $625 \times 490 \times 380$ *mm* y un peso de 34 *kg*. Las micromáquinas son controladas por un operador utilizando dos palancas de múltiples grados de libertad y tres cámaras CCD miniatura, montadas una en cada micromáquina, que despliegan la imagen de una sección de la misma en un monitor LCD de 5.8 pulgadas. Esta microfábrica permite producir y ensamblar piezas miniatura.

2.4. Tecnología de MicroEquipo (MET)

En México, el Laboratorio de Micromecánica y Mecatrónica (LMM) propuso la Tecnología de MicroEquipo para la producción de microdispositivos mecánicos a bajo costo [21]. La idea principal de este método sugiere la creación de generaciones sucesivas de microequipo, donde las dimensiones del microequipo de cada generación son menores que las generaciones predecesoras (Figura 2.5). Cada generación produce las micromáquinas de la siguiente generación. Esta metodología permite emplear componentes de bajo costo para la producción del microequipo y desarrollar microfábricas capaces de producir microdispositivos de bajo costo [22]. Actualmente, algunos prototipos de micromáquinas de primera generación se han desarrollado y evaluado [22] (ver Figuras 2.6 y 2.8, cortesía de [22]).

El primer prototipo de micromáquina desarrollado bajo MET de primera generación tiene las funciones de fresado, torneado, barrenado y pulido. Las dimensiones de este prototipo ($140 \times 153 \times 83$ *mm*) corresponden a las de las micromáquinas japonesas previamente mencionadas.

En la Figura 2.7 (cortesía de [22]) se muestran ejemplos de piezas fabricadas por el primer prototipo de micromáquina de primera generación de la tecnología MET.

El segundo prototipo de micromáquina de primera generación posee las dimensiones de $130 \times 160 \times 85$ *mm* (Figura 2.8). Este segundo diseño realiza las mismas funciones que el primer prototipo de micromáquina desarrollado.

En la Figura 2.9 (cortesía de [22]) se muestran ejemplos de piezas fabricadas por el

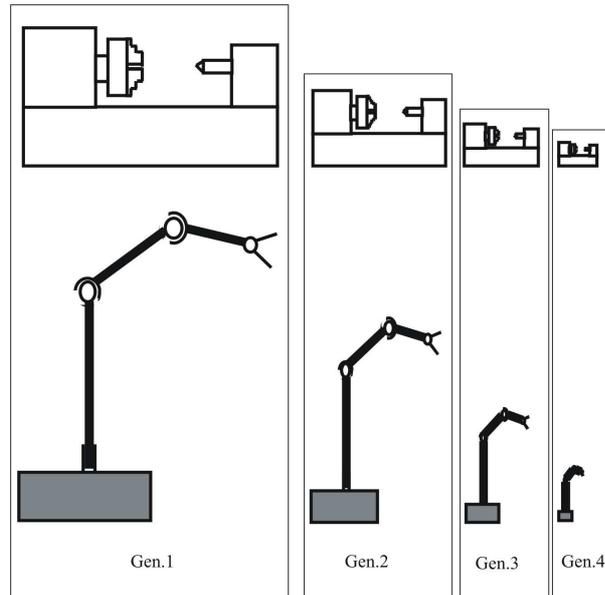
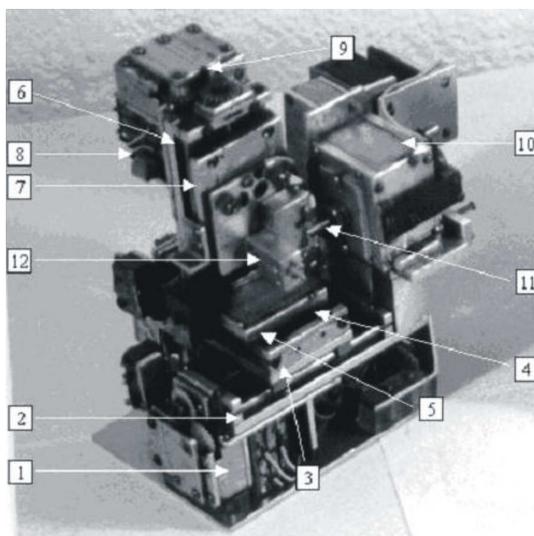
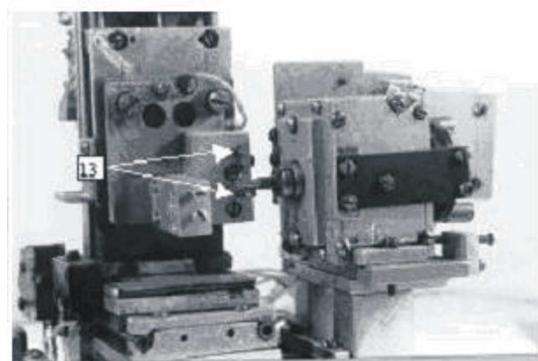


Figura 2.5: Generaciones sucesivas de microequipo de tecnología MET.



a)



b)

Figura 2.6: Micromáquina de tecnología MET. a) y b) Primer prototipo de micromáquina de primera generación.

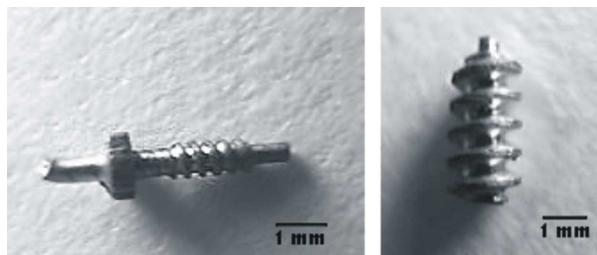


Figura 2.7: Microtornillos manufacturadas con el primer prototipo de micromáquina de primera generación de tecnología MET.

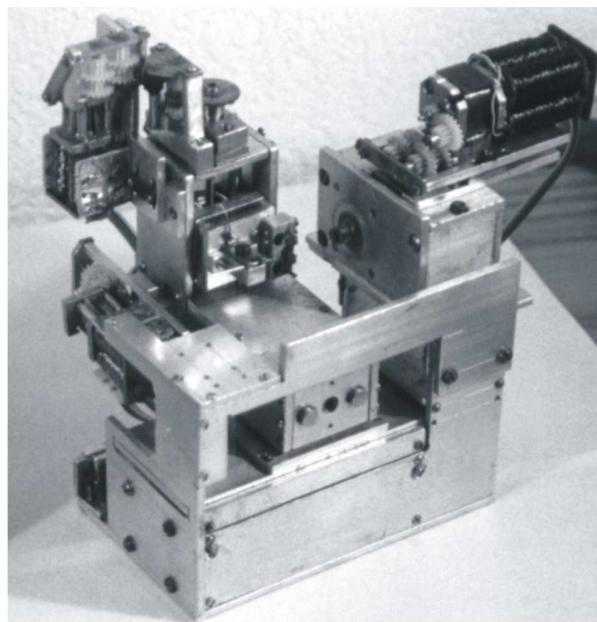


Figura 2.8: Segundo prototipo de micromáquina de primera generación de tecnología MET.

segundo prototipo de micromáquina de primera generación de la tecnología MET. En la Figura 2.9(a) se presenta una microflecha de $50 \mu m$ de diámetro y en la Figura 2.9(b) se pueden apreciar las piezas de un microfiltro y el dispositivo en su etapa final.

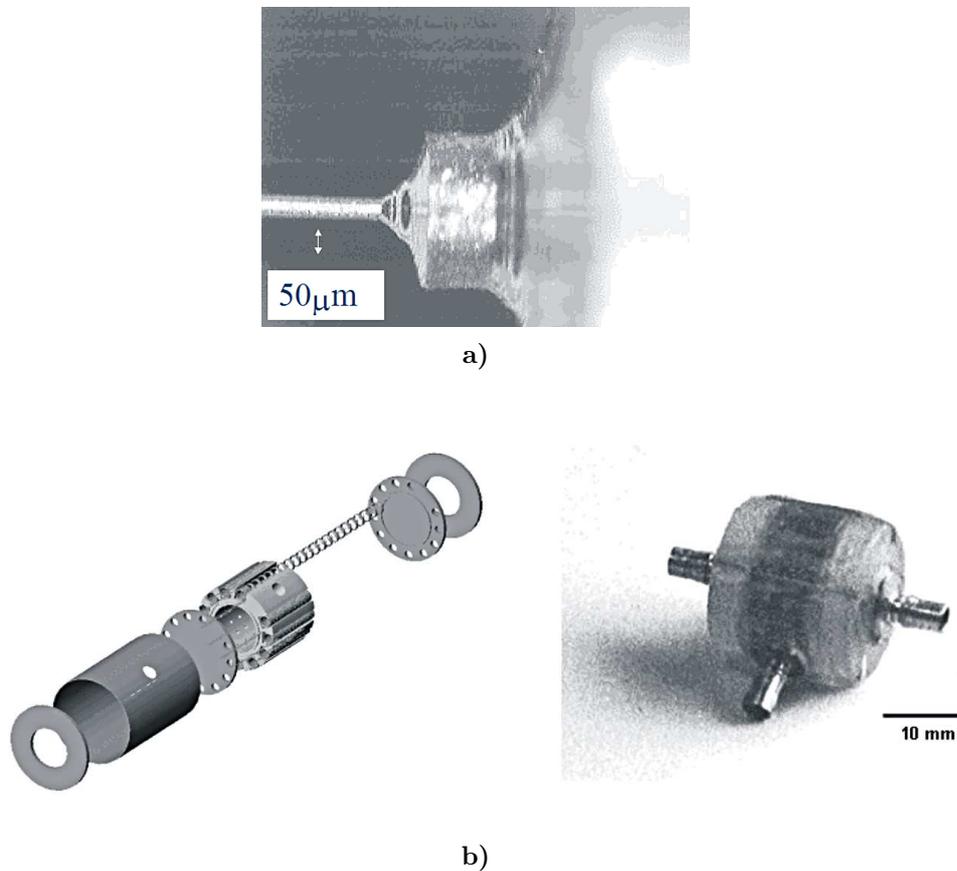


Figura 2.9: Piezas manufacturadas con el segundo prototipo de micromáquina de primera generación de tecnología MET. a) Microflecha de $50 \mu m$ y b) Ensamble de piezas de un microfiltro y prototipo terminado.

Los dos prototipos de micromáquinas de tecnología MET son micromáquinas CNC, es decir, poseen un control numérico computarizado. En [22] se puede encontrar una descripción de los métodos de control CNC desarrollados para este tipo de micromáquinas.

Capítulo 3

Conceptos de Redes Neuronales Artificiales

La ciencia de las redes neuronales artificiales está basada en la creación de modelos matemáticos y computacionales inspirados en la arquitectura y funcionamiento del cerebro humano.

En este capítulo se describirá el concepto de redes neuronales artificiales, los conceptos biológicos básicos que las caracterizan y uno de los primeros modelos de redes neuronales desarrollado, el perceptrón.

3.1. La Neurona Biológica

La neurona biológica es una célula¹ nerviosa que constituye el componente básico del sistema nervioso, incluyendo al cerebro.

Las neuronas forman la arquitectura básica del cerebro. Son elementos de procesamiento de señales que, de manera individual, poseen un poder reducido de procesamiento, pero como grandes conjuntos de neuronas interconectadas son capaces de desempeñar tareas complejas. Asimismo, son células excitables especializadas para la recepción de estímulos y la conducción de impulsos nerviosos.

Se conoce mucho acerca del comportamiento de las neuronas biológicas pero poco sobre su estructura de conexiones, la cual forma lo que se denomina red neuronal, que

¹ Una célula es la unidad o componente básica de los sistemas biológicos.

da origen al comportamiento inteligente del cerebro.

El tiempo requerido por una neurona para reaccionar está en el orden de milisegundos, lo cual es lento en comparación con las computadoras digitales que están en el orden de nanosegundos. Sin embargo, las grandes interconexiones de redes de neuronas son capaces de producir un comportamiento inteligente muy por encima de las capacidades de los sistemas artificiales actuales.

Una neurona está compuesta de tres partes principales [23]: las dendritas, el cuerpo celular (conocido también como soma) y el axón. Ver Figura 3.1 (obtenida de www.cienciateca.com).

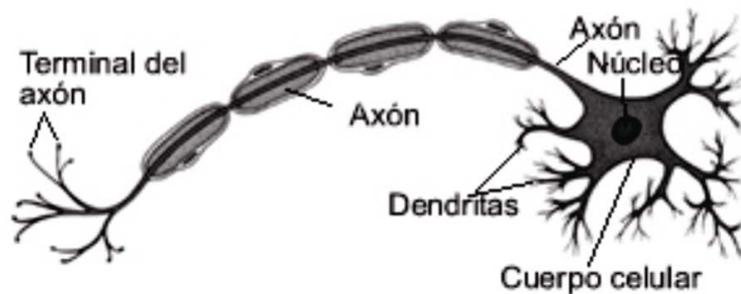


Figura 3.1: Estructura de una neurona biológica.

1. Las dendritas, de número y estructura variable según el tipo de neurona, son usadas para detectar y recibir las señales de las neuronas adyacentes, llevando estas señales o impulsos nerviosos hacia el cuerpo celular.
2. El cuerpo celular o soma reacciona a los impulsos nerviosos que recibe de las dendritas, produciendo su propia señal de respuesta, misma que es conducida hacia el axón para su transmisión.
3. El axón es el responsable de transmitir la señal generada en el cuerpo celular de la neurona hacia una siguiente célula a través de la estimulación de las dendritas de esta última.

El cerebro humano posee al menos cien mil millones de neuronas. Cada neurona recibe señales de otras células, integra las señales y genera una salida, la cual es enviada a otras neuronas, o en algunos casos, acciona órganos como músculos o glándulas.

Una simple neurona puede recibir cientos o miles de impulsos nerviosos y puede transmitir señales de salida a un número similar de neuronas. Una neurona es un dispositivo electroquímico complejo que contiene un potencial interno llamado potencial de membrana². Cuando el potencial de membrana excede un umbral, la neurona puede propagar un impulso nervioso por largas distancias, a través de su axón, a otras neuronas [24].

El punto de conexión entre las terminales del axón de una neurona emisora y la dendrita de una neurona receptora se denomina sinapsis.

Cuando se establece la sinapsis se produce una transformación del impulso nervioso de salida de la neurona emisora, mediante la liberación de unas sustancias llamadas neurotransmisores. El efecto de los neurotransmisores sobre la señal recibida por la neurona receptora puede ser excitatorio o inhibitorio, esto es, los neurotransmisores excitatorios incrementan el impulso nervioso y los neurotransmisores inhibitorios lo decrementan.

3.2. La Neurona Artificial

Una neurona artificial es la unidad fundamental de procesamiento de información. En la Figura 3.2 se muestra un modelo de neurona artificial.

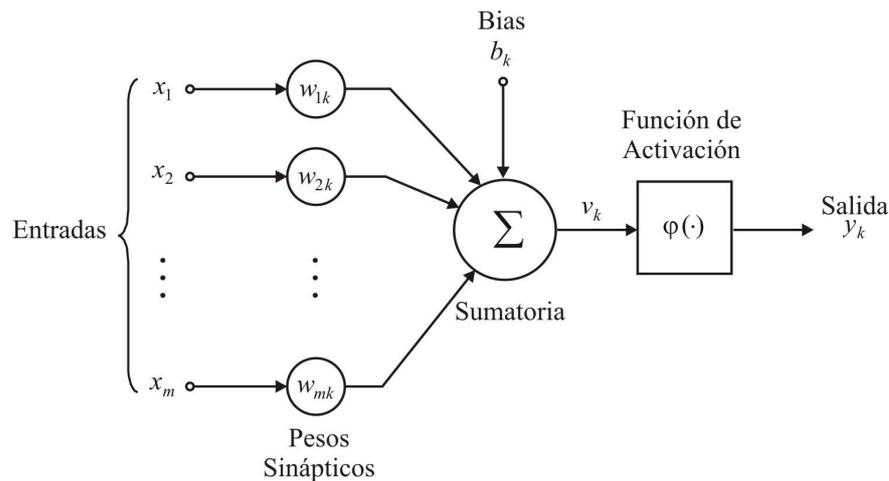


Figura 3.2: Modelo de una neurona artificial.

² El rango del potencial de membrana varía entre $+30 \text{ mV}$ y -70 mV .

Los elementos principales del modelo neuronal ilustrado son los siguientes [25]:

1. Conjunto de sinapsis o conexiones sinápticas. Cada una de estas sinapsis está caracterizada por un peso sináptico. Las conexiones sinápticas se encargan de ponderar la señal de entrada multiplicándola por su correspondiente peso sináptico. Por ejemplo, en la Figura 3.2, la señal x_j en la entrada de la sinapsis j , donde $j = 1, \dots, m$, conectada a la neurona k será multiplicada por el peso sináptico w_{jk} . A diferencia de las sinapsis cerebrales, el peso sináptico de una neurona artificial puede contener valores tanto negativos como positivos. La sinapsis con peso sináptico $w \geq 0$ es llamada conexión excitadora, y con peso sináptico $w < 0$ conexión inhibitoria.
2. Sumador. Se encarga de calcular la sumatoria ponderada de las señales de entrada recibidas, es decir, la suma de las entradas multiplicadas por su correspondiente peso sináptico. Estas operaciones definen una combinación lineal.
3. Función de activación. Permite limitar el rango de amplitud de la señal de salida de la neurona hacia valores finitos. Generalmente, el rango de amplitud normalizado de la salida de una neurona está en el intervalo cerrado $[0, 1]$ o alternativamente $[-1, 1]$.

El modelo neuronal de la Figura 3.2 incluye también una señal externa denominada sesgo (del Inglés *bias*)³, denotada por b_k . El bias b_k tiene el efecto de incrementar o decrementar la entrada de la función de activación, dependiendo si es positivo o negativo, respectivamente.

En términos matemáticos, se puede describir la neurona k con el siguiente par de ecuaciones:

$$u_k = \sum_{j=1}^m w_{jk} \cdot x_j \quad (3.1)$$

y

$$y_k = \varphi(u_k + b_k) \quad (3.2)$$

donde x_1, x_2, \dots, x_m son las señales de entrada; $w_{1k}, w_{2k}, \dots, w_{mk}$ son los pesos sinápticos de la neurona k ; u_k es la combinación lineal de salida generada por las señales de

³ A lo largo del texto se empleará el término *bias* en lugar de sesgo.

entrada; $\varphi(\cdot)$ es la función de activación; y y_k es la señal de salida de la neurona. El uso del bias b_k tiene un efecto sobre la combinación lineal de salida u_k en el modelo de la Figura 3.2, definido por,

$$v_k = u_k + b_k \tag{3.3}$$

Dependiendo de si el bias b_k es positivo o negativo, la relación entre el potencial de activación o campo local inducido v_k de la neurona k y la combinación lineal de salida u_k es modificada como se muestra en la Figura 3.3.

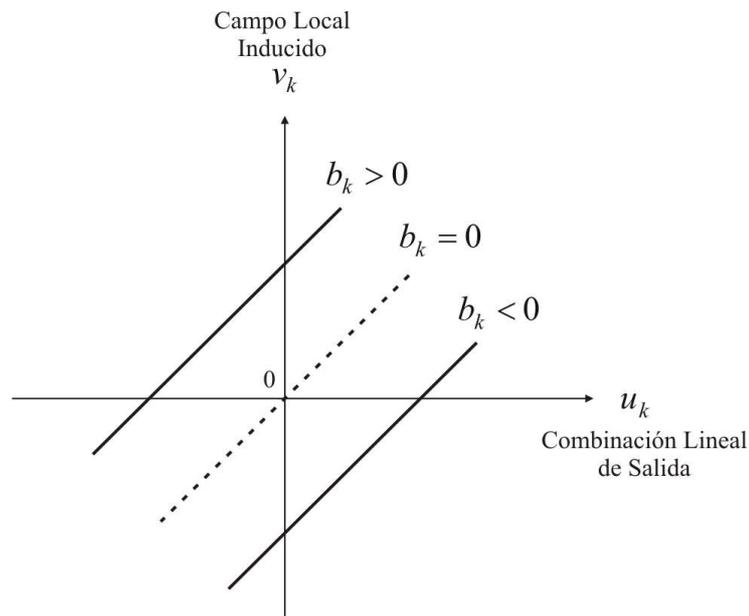


Figura 3.3: Transformación producida por la presencia de un bias b_k .

3.2.1. Tipos de Funciones de Activación

La función de activación, denotada por $\varphi(v)$, define la salida de la neurona mediante la transformación del potencial de activación o campo local inducido v .

A continuación, se describirán algunos tipos de funciones de activación.

- **Función de Umbral o Escalón.** En este tipo de función la salida de una neurona se activa sólo cuando el potencial de activación es mayor o igual que cierto valor de umbral t que representa la mínima entrada total ponderada necesaria para provocar la activación de la neurona. Un ejemplo de función de escalón se ilustra

en la Figura 3.4(a), y se define como,

$$\varphi(v) = \begin{cases} 1, & v \geq t \\ 0, & v < t \end{cases} \quad (3.4)$$

La neurona artificial que emplea esta función de activación es conocida como el modelo de neurona McCulloch y Pitts [26], siendo éste el primer modelo formal de una neurona artificial.

- **Función Lineal a Tramos.** Es una aproximación de una función no lineal. La función lineal a tramos graficada en la Figura 3.4(b) se define como,

$$\varphi(v) = \begin{cases} 1, & v \geq t \\ v, & -t < v < t \\ 0, & v \leq -t \end{cases} \quad (3.5)$$

- **Función Sigmoidal.** La función sigmoidal es la función de activación comúnmente empleada en la construcción de redes neuronales artificiales. Se define como una función estrictamente creciente, suave y de comportamiento asintótico. La función sigmoidal mostrada en la Figura 3.4(c) se define como,

$$\varphi(v) = \frac{1}{1 + e^{-v}} \quad (3.6)$$

- **Función Lineal.** Es también conocida como función identidad, con un rango de salida $(-\infty, +\infty)$, se define como,

$$\varphi(v) = v \quad (3.7)$$

Las funciones de activación definidas en las Ecuaciones 3.4, 3.5 y 3.6 varían en el rango de 0 a +1. En algunas ocasiones es deseable tener funciones de activación cuyo rango varíe entre -1 y +1. Ejemplos de funciones de activación que satisfacen dichos rangos son:

- **Caso de función de umbral:** Se emplea la función comúnmente conocida como función de signo, definida por,

$$\varphi(v) = \begin{cases} 1, & v > 0 \\ 0, & v = 0 \\ -1, & v < 0 \end{cases} \quad (3.8)$$

- **Caso de función sigmoidal:** Se emplea la función tangente hiperbólica, definida por,

$$\varphi(v) = \tanh(v) \quad (3.9)$$

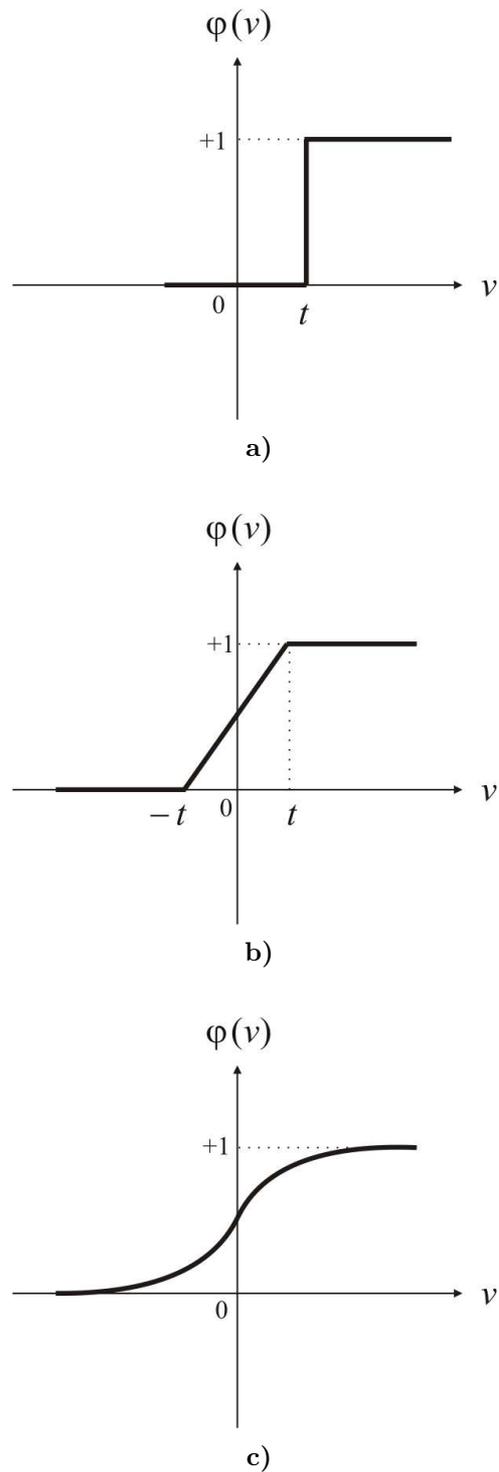


Figura 3.4: Funciones de Activación. a) Escalón o Umbral, b) Lineal a Tramos y c) Sigmoidal.

3.3. Redes Neuronales Artificiales

Las redes neuronales biológicas están constituidas por un gran número de neuronas, conectadas en forma masiva. Estas redes conforman el sistema nervioso y el cerebro. Se estima que el cerebro puede contener 10^{11} neuronas y 10^{15} interconexiones, aproximadamente.

Las redes neuronales artificiales son modelos simplificados de las redes neuronales biológicas. Tratan de imitar la estructura y funcionamiento del cerebro humano a través de modelos matemáticos. Las tres cualidades básicas del cerebro humano que forman los fundamentos de la mayoría de los modelos de redes neuronales son: procesamiento paralelo, memoria distribuida y adaptabilidad.

Una red neuronal artificial (RNA) puede ser definida como una estructura distribuida, de procesamiento paralelo, formada de neuronas artificiales (llamadas también elementos de procesamiento), interconectadas por un gran número de conexiones (sinapsis), las cuales son usadas para almacenar conocimiento que estará disponible para poder ser utilizado.

A continuación se presenta una definición de red neuronal vista como una máquina adaptativa [25].

Una red neuronal es un procesador distribuido masivamente paralelo constituido por unidades de procesamiento simple, que es naturalmente propensa a almacenar conocimiento y dejarlo disponible para su uso. Se asemeja al cerebro en dos aspectos:

1. El conocimiento es adquirido por la red, de su entorno, a través de procesos de aprendizaje.
2. Las ponderaciones de las conexiones, conocidas como pesos sinápticos, son utilizadas para almacenar el conocimiento adquirido.

3.3.1. Propiedades

Algunas de las propiedades más importantes de las redes neuronales son las siguientes⁴:

- **No linealidad:** Una red neuronal hecha de interconexiones de neuronas no lineales es una RNA no lineal. La no linealidad es una propiedad importante, particularmente si el mecanismo físico responsable de generar la señal de entrada (ej. señal de voz) es inherentemente no lineal.
- **Aprendizaje adaptivo:** Las redes neuronales artificiales tienen la propiedad de aprender mediante la modificación de sus pesos sinápticos con el fin de adaptarse a las nuevas condiciones del entorno. Esta es, posiblemente, la característica más importante de las redes neuronales.
- **Auto-organización:** Una red neuronal puede crear su propia organización o representación de la información que recibe mediante un proceso de aprendizaje. Autoorganización significa generalización, de esta forma una red puede responder a datos o situaciones que no ha experimentado antes, pero que puede inferir sobre la base de su entrenamiento. Esta característica es muy útil sobre todo cuando la información de entrada es poco clara o se encuentra incompleta.
- **Tolerancia a fallos:** Debido a que una red neuronal distribuye y almacena la información aprendida de forma redundante a través de sus conexiones sinápticas, ésta puede seguir respondiendo aceptablemente aún en el caso de que algunas sinapsis resulten dañadas. Esta característica evita la pérdida de información.
- **Operación en tiempo real:** Una red neuronal artificial posee una estructura paralela, actualizando todas sus instancias de manera simultánea. Por lo tanto, si ésta es implementada con hardware especialmente diseñado para el procesamiento paralelo, se pueden obtener respuestas en tiempo real.
- **Fácil inserción en la tecnología existente:** Es relativamente sencillo obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilita la integración modular en los sistemas existentes.

⁴ Algunas fueron citadas en [25] y en [27].

3.3.2. Componentes

Una red neuronal artificial está compuesta de neuronas artificiales, las cuales se encuentran conectadas por medio de las sinapsis o conexiones sinápticas.

En general las neuronas suelen agruparse en unidades estructurales denominadas capas neuronales (algunas veces llamadas campos de procesamiento). Las capas neuronales se encuentran conectadas entre sí en una arquitectura, donde el conjunto de una o más capas constituye una red neuronal.

Los tres tipos de capas definidos en una red neuronal son:

- Capa de entrada: También denominada sensorial, está compuesta por neuronas que reciben datos o señales procedentes del entorno.
- Capa de salida: Se compone de neuronas que proporcionan la respuesta de la red neuronal.
- Capa oculta: No tiene una conexión directa con el entorno, es decir, sus señales de entrada y salida permanecen dentro de la red. El tipo de capa oculta proporciona a la red neuronal grados de libertad que le permiten representar, con mayor precisión, determinadas características del entorno que trata de modelar. Una red neuronal puede contener más de una capa oculta.

Asimismo, de acuerdo al tipo de capa al que pertenezcan, se puede hacer distinción entre tres tipos de neuronas: neuronas de entrada, las cuales reciben datos del exterior de la red neuronal; neuronas de salida, las cuales envían datos fuera de la red neuronal; y neuronas ocultas cuyas señales de entrada y salida se mantienen adentro de la red neuronal.

3.3.3. Arquitecturas de Redes Neuronales

Se denomina arquitectura a la organización y disposición de las neuronas que integran una red neuronal.

En general, se pueden identificar dos clases fundamentales de arquitecturas de redes neuronales [28]:

1. Redes de Propagación hacia Delante. En este tipo de red todas las señales van desde la capa de entrada hasta la capa de salida sin existir ciclos, ni conexiones entre neuronas de la misma capa, esto es, presentan conexiones únicamente hacia delante (*feedforward*). En relación al número de capas, las redes de propagación hacia delante se dividen en:
 - Red Monocapa. Las redes neuronales bajo esta arquitectura están compuestas de una capa de nodos de entrada y una capa de neuronas de salida (ver Figura 3.5). El término monocapa hace referencia únicamente a la capa de neuronas de salida debido a que en los elementos de esta capa sí se realiza algún tipo de cálculo computacional, a diferencia de los nodos de la capa de entrada. Ejemplos de redes monocapa son el Perceptrón [29] y Adaline [30].
 - Red Multicapa. Las redes neuronales bajo esta arquitectura se distinguen por la presencia de una o más capas ocultas (ver Figura 3.6). Un ejemplo de red multicapa es el perceptrón multicapa [31].
2. Redes Recurrentes. Este tipo de red, a diferencia de la red de propagación hacia delante, contiene al menos una conexión de propagación hacia atrás (*feedback*), es decir, conexiones que van de las salidas de las neuronas a las entradas de neuronas de la misma capa o de capas anteriores (ver Figura 3.7). Por otra parte, las conexiones de propagación hacia atrás involucran el uso de elementos de retardo (denotados por z^{-1}) en sus conexiones, generando un comportamiento dinámico no lineal, suponiendo que la red neuronal contenga unidades no lineales. Ejemplos de redes recurrentes han sido presentados por Anderson [32], Kohonen [33] y Hopfield [34].

3.3.4. Proceso de Aprendizaje o Entrenamiento

Una red neuronal tiene que ser configurada de tal manera que la aplicación de un conjunto de entradas produzca el conjunto de salidas deseado.

El aprendizaje o entrenamiento de una red neuronal artificial es un proceso mediante el cual se modifican los pesos de las conexiones sinápticas, con base en algún paradigma

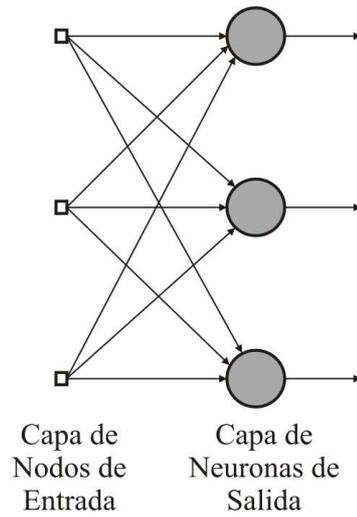


Figura 3.5: Red neuronal monocapa.

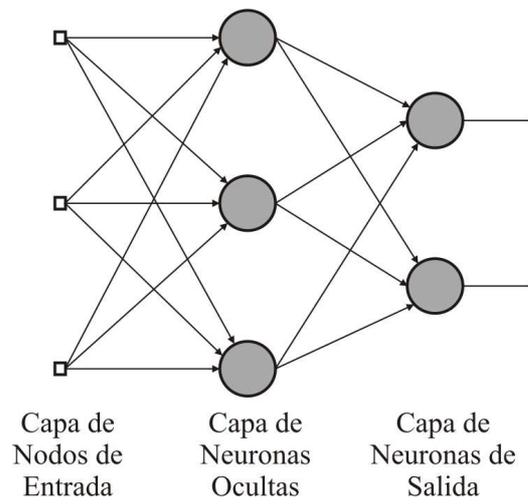


Figura 3.6: Red neuronal multicapa.

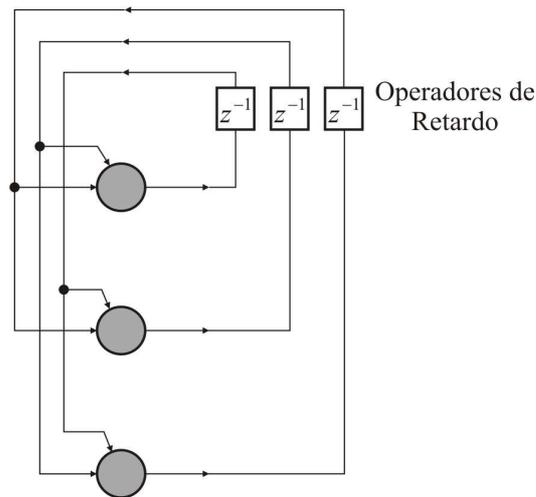


Figura 3.7: Red neuronal recurrente.

y regla de aprendizaje, para que la salida de la red se vaya adaptando de manera paulatina al funcionamiento que se considera correcto.

Existen varios métodos para ajustar el conjunto de pesos sinápticos de las conexiones.

Paradigmas de Aprendizaje

Los paradigmas de aprendizaje se pueden categorizar en dos tipos principales:

1. Aprendizaje supervisado. En esta clase de paradigma, para cada patrón de entrada presentado a la red debe existir la correspondiente respuesta correcta, la cual se empleará para realizar una comparación con la respuesta de salida de la red y con base en dicha comparación realizar los ajustes de los parámetros de la red.
2. Aprendizaje no supervisado. A diferencia del paradigma de aprendizaje supervisado, en este tipo de aprendizaje no se especifica a la red la salida correcta a cada patrón de entrada, sino que el sistema debe encontrar las relaciones existentes en los mismos mediante reglas de aprendizaje.

Reglas de Aprendizaje

Ambos paradigmas de aprendizaje mencionados implican un ajuste de los pesos de las conexiones entre neuronas de acuerdo a alguna regla de modificación. A continuación, se presentarán algunas reglas de aprendizaje.

- Aprendizaje por corrección de error: Este tipo de aprendizaje busca minimizar una función de error mediante el empleo de gradiente descendente. Esta regla de aprendizaje utiliza la diferencia entre la salida real y_k de la red y la salida deseada d_k para el ajuste de los pesos, esto es,

$$\Delta w_{jk} = \gamma y_j (d_k - y_k) \quad (3.10)$$

donde, w_{jk} es el peso sináptico que une a la neurona j con la neurona de salida k , y_j es la salida de la neurona j , y_k es la salida de la neurona k (salida real), d_k es la salida deseada y γ es una constante positiva de proporcionalidad que representa la tasa de aprendizaje ($0 < \gamma < 1$). La tasa de aprendizaje se refiere a un parámetro que determina la velocidad de aprendizaje de una red neuronal, en otras palabras, controla el tamaño del cambio de los pesos sinápticos en cada iteración del entrenamiento. Esta regla de aprendizaje es referida también como regla delta o regla de Widrow-Hoff en honor a sus creadores [30].

Este es el principio usado por el algoritmo *BackPropagation*, muy utilizado para el entrenamiento de redes multicapa como el perceptrón multicapa [31].

- Aprendizaje Hebbiano: Sugerido por Hebb en su libro clásico *The Organization of Behavior* (1949) [35]. La idea básica es que si dos neuronas j y k están simultáneamente activas, la conexión que las une debe ser reforzada, en el caso contrario, deberá debilitarse o eliminarse. Si la neurona k recibe su entrada de la neurona j , la versión simple del aprendizaje Hebbiano prescribe modificar el peso w_{jk} con la siguiente ecuación:

$$\Delta w_{jk} = \gamma y_j y_k \quad (3.11)$$

donde, w_{jk} es el peso sináptico que une a la neurona j con la neurona de salida k , γ es la tasa de aprendizaje, y_j es la salida de la neurona j y y_k es la salida de

la neurona k . Como ejemplo, esta regla fue utilizada en el Modelo de Hopfield [34].

- **Aprendizaje competitivo:** En el aprendizaje competitivo las neuronas de salida de una red neuronal compiten entre ellas para convertirse en ganadoras (activas o disparadas). Mientras que en una red neuronal basada en aprendizaje Hebbiano varias neuronas de salida pueden estar activas simultáneamente, en el aprendizaje competitivo solamente una neurona de salida puede resultar activa a la vez. Esta propiedad hace que este tipo de aprendizaje sea bueno para descubrir características que puedan ser utilizadas para clasificar un conjunto de patrones de entrada. El aprendizaje competitivo es usado en mapas de Kohonen [33] y en la red ART [36].
- **Aprendizaje de Boltzmann:** Es una regla de aprendizaje estocástico obtenido a partir de los principios de la teoría de la información [37] y de la termodinámica. Una red neuronal basada en la regla de aprendizaje de Boltzmann es denominada máquina de Boltzmann [38]. El objetivo del aprendizaje de Boltzmann es ajustar los pesos de las conexiones de tal forma que el estado de las neuronas visibles⁵ satisfaga una distribución de probabilidad deseada en particular.

3.3.5. Algunos Campos de Aplicación

Gracias a la habilidad de aprender de un conjunto de datos de entrada (previamente clasificados o no), las redes neuronales tienen diversos campos de aplicación, como son el modelado, el análisis de series de tiempo, la clasificación y el reconocimiento de patrones, el procesamiento de señales e imágenes, aplicaciones de control, etc.

Las características inherentes de las redes neuronales artificiales las hacen bastante apropiadas para resolver problemas donde la información es imprecisa o no bien conocida. De igual manera se pueden utilizar cuando no existen modelos matemáticos precisos o algoritmos de complejidad razonable. Por ejemplo la red de Kohonen ha sido aplicada con éxito al clásico problema del viajero, un problema para el que no se conoce solución algorítmica de complejidad polinómica.

⁵ Las neuronas visibles interactúan con el ambiente, las ocultas operan libremente.

Varios modelos de redes neuronales artificiales diferentes han sido desarrollados a través de los años. Uno de los primeros y más básico modelo de neuronas es el Perceptrón presentado a continuación.

3.4. Modelo del Perceptrón

El primer modelo de red neuronal artificial fue el Perceptrón, propuesto por el psicólogo Frank Rosenblatt en el año de 1958 [29], al tratar de ilustrar algunas propiedades fundamentales de los sistemas neuronales biológicos. Básicamente, el modelo del Perceptrón emplea el modelo de neurona de McCulloch y Pitts [26] aplicando las ideas de aprendizaje de Hebb [35].

Rosenblatt creía que la conectividad existente en las redes neuronales biológicas contiene un elevado porcentaje de aleatoriedad, es por ello que su modelo está definido en términos de teoría de probabilidad, en lugar de lógica simbólica, empleada por McCulloch y Pitts [26].

En 1959, Rosenblatt probó el teorema sobre el aprendizaje del Perceptron (*perceptron convergence theorem* [39]). En consecuencia, a principios de los años 60, los perceptrones generaron gran interés, no sólo por su capacidad de generalizar a partir de su aprendizaje, sino también por el hecho de comportarse correctamente, aún cuando su estructura presentaba un cierto grado de aleatoriedad.

En 1969 Marvin Minsky y Seymour Papert realizaron la publicación del libro *Perceptrons* [40], el cual para muchos significó el final de las redes neuronales. En este trabajo se presentaba un análisis detallado del Perceptrón, en términos de sus capacidades y limitaciones, en especial sobre las restricciones existentes para los problemas que una red de este tipo puede resolver. La mayor desventaja mostrada para el Perceptrón fue su incapacidad para solucionar problemas que no sean linealmente separables.

En [29] los perceptrones se encuentran descritos en un número de variaciones, etiquetados con un nombre arbitrario. Este trabajo de investigación se basará únicamente en el modelo de red neuronal definida como α -Perceptrón elemental.

3.4.1. Arquitectura

En su forma más simple, el modelo del Perceptrón de Rosenblatt está estructurado por tres áreas o capas neuronales: área sensorial, área asociativa y área de respuesta. Este modelo se muestra en la Figura 3.8.

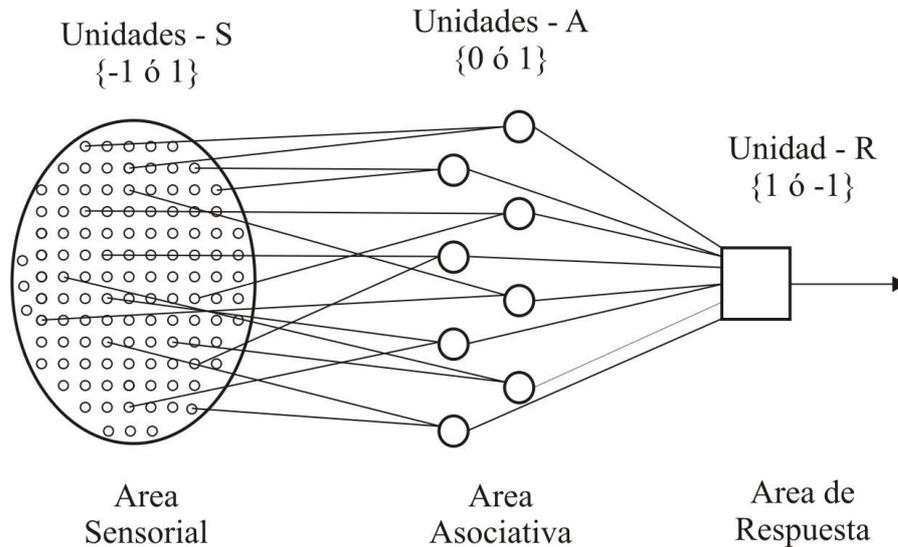


Figura 3.8: Arquitectura del Perceptrón.

El área sensorial, también llamada retina, está constituida por unidades sensoriales (unidades-S) clasificadas en dos tipos: excitatorias e inhibitorias. El tipo de unidades excitatorias tiene un valor de salida 1 y las inhibitorias un valor de salida -1 . Una unidad sensorial será excitatoria si su señal de entrada excede un valor de umbral determinado, en caso contrario, será inhibitoria.

El área asociativa está compuesta de unidades asociativas (unidades-A) con dos estados de salida, 0 ó 1. Cada unidad asociativa está conectada a un número de unidades sensoriales a través de conexiones distribuidas de manera aleatoria. Si la suma algebraica de la salida de unidades excitatorias e inhibitorias conectadas a una unidad-A es mayor o igual a un determinado umbral, entonces la unidad-A se activará y tendrá un valor de salida 1, en caso contrario, se mantendrá inactiva con un valor de salida 0.

El área de respuesta consta de una sola unidad de salida (unidad-R), la cual se encuentra conectada con todas las salidas de las unidades-A en el área asociativa. Este elemento corresponde al modelo de neurona de McCulloch y Pitts, por lo tanto, las

conexiones poseen pesos sinápticos y, se establece que, si la suma ponderada de sus entradas es mayor o igual a un cierto umbral, la salida de esta unidad será 1, en caso contrario, será -1 . Este valor de salida representa la respuesta de la red neuronal.

Con esta arquitectura, este modelo de Perceptrón únicamente puede reconocer dos clases, donde cada clase estará asociada a uno de los dos valores de salida de la red neuronal.

3.4.2. Aprendizaje

El paradigma de aprendizaje del Perceptrón es del tipo supervisado, empleando una regla de aprendizaje por corrección de error. Este proceso de aprendizaje ajusta solamente las conexiones sinápticas que inciden en la única unidad de respuesta presente en el modelo descrito.

El procedimiento de aprendizaje es el siguiente: Se aplicaba una entrada a la retina (área sensorial), y el estímulo se propagaba a través de las capas o áreas neuronales hasta obtener un valor de salida de la unidad de respuesta. Si el valor de salida de la unidad-R coincide con el valor de la respuesta correcta al estímulo de entrada, se incrementan los valores de los pesos sinápticos de la red neuronal, en el caso contrario, se disminuyen los valores de los pesos sinápticos.

El incremento o disminución de los valores de los pesos sinápticos en el proceso de aprendizaje descrito se realiza a través de una sola regla de aprendizaje por corrección de error, definida de la siguiente forma,

$$w_i(t+1) = w_i(t) + \gamma[d(t) - y(t)]a_i(t) \quad (3.12)$$

donde, $w_i(t+1)$ y $w_i(t)$ es el peso sináptico que une a la unidad asociativa $a_i(t)$ con la unidad de respuesta en el paso $(t+1)$ y paso t , respectivamente, γ es la tasa de aprendizaje, $d(t)$ es la salida deseada en el paso t y, $y(t)$ es la salida real de la unidad de respuesta.

Este proceso se repetirá para un número de pasos $t = 1, 2, \dots$

Capítulo 4

Descripción de la Metodología

En este capítulo se describirán los pasos empleados para el desarrollo de este trabajo de investigación. Se iniciará con una descripción de las micropiezas utilizadas para la captura de imágenes, mismas que formarán la base de datos para el entrenamiento y prueba del sistema de visión computacional. Posteriormente, se realizará una breve descripción del clasificador neuronal empleado. Por último, se hará mención de la plataforma y el software para la implementación del trabajo.

4.1. Descripción y Clasificación de Microtornillos

Para el presente trabajo se emplearon 40 microtornillos de 3 *mm* de diámetro, previamente manufacturados con un torno CNC de la compañía Boxford, clasificados en cuatro grupos diferentes, donde cada uno de estos grupos corresponde a piezas producidas con una posición determinada de la herramienta de corte empleada.

Tres de los cuatro grupos están conformados de microtornillos defectuosos, esto es, tres errores de fabricación diferentes relacionados a la ubicación del cortador. Cada uno de los cuatro grupos consta de un total de 10 microtornillos. A continuación se hará una breve descripción de los grupos de microtornillos:

- **Grupo 0:** Microtornillos manufacturados con una distancia, entre el cortador y el eje de la pieza, 0.1 *mm* por debajo de la posición correcta.
- **Grupo 1:** Microtornillos manufacturados con la posición correcta del cortador.

- **Grupo 2:** Microtornillos manufacturados con una distancia, entre el cortador y el eje de la pieza, 0.1 *mm* por encima de la posición correcta.
- **Grupo 3:** Microtornillos manufacturados con una distancia, entre el cortador y el eje de la pieza, 0.2 *mm* por encima de la posición correcta.

Cada uno de estos grupos de microtornillos permitirá al sistema de visión computacional tener varias muestras de la forma resultante de la micropieza, después de su manufactura, dependiendo de cuatro distintas ubicaciones de la herramienta de corte.

4.2. Descripción de las Imágenes Digitales de Microtornillos

La adquisición de las imágenes de los microtornillos fue realizada a través de una computadora PC acoplada a una cámara digital SAMSUNG MPC-M10 ANYCAM Couple, la cual se encuentra montada sobre un microscopio óptico IROSCOPE NZ-14T con un adaptador fabricado por el propio laboratorio LMM. Un software de imagen comercial (ANYCAM Couple) permitió la transferencia de las imágenes digitales, capturadas con la cámara, a la PC. En la Figura 4.1 se puede apreciar el sistema de captura de imágenes.

Las imágenes capturadas de los microtornillos serán empleadas en el entrenamiento y prueba del sistema.

4.3. Creación del Conjunto de Imágenes de Entrenamiento y Prueba

Del conjunto de total de imágenes de microtornillos, se crearán dos subconjuntos de imágenes, uno para entrenamiento y otro para prueba del sistema.

Los conjuntos de imágenes de entrenamiento y de prueba estarán constituidos por un porcentaje del total de imágenes capturadas. Por ejemplo, sea 440 el número total de imágenes, donde por cada uno de los 40 microtornillos existen 11 imágenes capturadas, si el 80 % de las imágenes será empleado para entrenar y el 20 % de imágenes será para

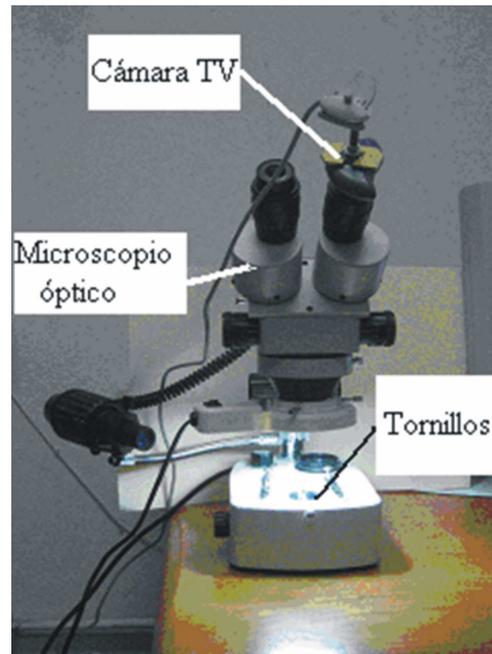


Figura 4.1: Dispositivos para la captura de las imágenes de microtornillos.

probar, entonces, las 88 imágenes de prueba se elegirán aleatoriamente de acuerdo a lo siguiente:

1. Se seleccionan aleatoriamente dos microtornillos de los diez posibles por cada una de las cuatro clases existentes, dando un total de ocho índices de microtornillos escogidos.
2. Para cada uno de los ocho microtornillos seleccionados, se tomarán las 11 imágenes correspondientes a cada uno, formando así, las 88 imágenes del conjunto de prueba.
3. Una vez seleccionadas las imágenes de prueba, las 352 imágenes de los 32 microtornillos restantes serán designadas al conjunto de entrenamiento.

Este proceso de división asegura la independencia entre los dos conjuntos de imágenes de microtornillos, ya que las micropiezas presentadas en el proceso de entrenamiento y en el proceso de prueba son diferentes.

4.4. Desarrollo e Implementación de la Red Neuronal Artificial

Para resolver el problema de clasificación de las imágenes de microtornillos se ha elegido emplear la red neuronal artificial denominada Limited Receptive Area Grayscale (LIRA Grayscale), siendo ésta una extensión del clasificador neuronal Limited Receptive Area (LIRA) [41], diseñado para imágenes binarias¹ únicamente y que ha tenido muy buenos resultados en diversas tareas de reconocimiento, mencionadas posteriormente. Ambos clasificadores están basados en los principios del modelo del perceptrón de Rosenblatt.

La red neuronal LIRA Grayscale será adaptada e implementada en el sistema de visión para la tarea de reconocimiento y clasificación de las imágenes de microtornillos manufacturados por micromáquinas.

La entrada para esta red neuronal será la imagen del microtornillo y la salida será la clase a la que pertenece la imagen de la micropieza (ver Figura 4.2).



Figura 4.2: Diagrama de bloque de entradas y salidas de la LIRA Grayscale.

En el capítulo 5 se describirá ampliamente la arquitectura y procesos de aprendizaje de esta red neuronal artificial.

4.5. Entrenamiento y Prueba de la Red Neuronal Artificial

Una vez construido el clasificador neuronal LIRA Grayscale junto con los conjuntos de imágenes de entrenamiento y prueba, se dará inicio a una etapa pre-experimental, en la cual se llevarán a cabo una serie de experimentos preliminares con los que se

¹ Una imagen binaria es aquella cuyos valores de píxel son cero o uno.

encontrarán los valores óptimos para los parámetros del sistema que no serán modificables durante la etapa de aprendizaje de la red neuronal, como son: número de neuronas asociativas, dimensiones de ventana, la excitación adicional de la neurona ganadora, entre otros. Más adelante, se describirán cada una de estas constantes y su función dentro del sistema.

Al término de la fase pre-experimental se dará comienzo a la fase de entrenamiento o aprendizaje del sistema, el cual será de manera supervisada. En esta etapa, se realizarán los ajustes de los parámetros internos de la red neuronal propuesta, de manera que, el clasificador reconozca de correctamente todas las imágenes de microtornillos que le sean presentadas durante este proceso. El conjunto de imágenes de esta fase será distinto al utilizado para la etapa pre-experimental, es decir, se repetirán los pasos descritos anteriormente para crear nuevos conjuntos de entrenamiento y de prueba.

Posteriormente, se realizará el proceso de validación del clasificador neuronal, con el correspondiente conjunto de imágenes de prueba para obtener el porcentaje de reconocimiento que el sistema ha adquirido a través del proceso de entrenamiento. Asimismo, se podrán realizar modificaciones y correcciones al sistema de acuerdo a los resultados obtenidos.

4.6. Implementación del Sistema

El sistema de clasificación de imágenes fue desarrollado a través de C++ Builder 6 ya que es un lenguaje de programación visual, orientado a objetos, para desarrollo rápido de aplicaciones. Permite crear programas eficientes para Microsoft Windows XP con un mínimo de código manual. Provee todas la herramientas necesarias para desarrollar y probar aplicaciones, incluyendo librerías con componentes que permiten un fácil manejo de imágenes.

La plataforma de implementación del sistema fue Windows XP Professional.

En el capítulo 6 se mostrarán los resultados generados a través de la aplicación de la metodología presentada en este capítulo.

Capítulo 5

Clasificador Neuronal: LIRA Grayscale

El clasificador neuronal Limited Receptive Area (LIRA) está basado en los principios del perceptrón de Rosenblatt [29, 39]. Este clasificador ha sido probado en tareas de ensamble de microdispositivos y reconocimiento de dígitos escritos a mano mostrando, buenos resultados [2, 41].

En este trabajo se ha adaptado LIRA para resolver el problema de clasificación de imágenes en niveles de gris¹ de microtornillos después del proceso de manufactura. La red neuronal artificial resultante se ha denominado Limited Receptive Area Grayscale (LIRA Grayscale) [41].

A continuación se describirá la estructura general de este clasificador para imágenes en niveles de gris.

5.1. Arquitectura

La red neuronal artificial LIRA Grayscale consiste de cuatro capas: capa de entrada (S), capa intermedia o de grupos (I), capa asociativa (A) y capa de salida (R). La arquitectura de este clasificador neuronal se muestra en la Figura 5.1.

Las neuronas de la capa de entrada (S) corresponden al valor de un píxel de la imagen. Así, los valores de salida de estas neuronas están en el rango $[0, 255]$, donde 0

¹ Una imagen en niveles de gris es aquella cuyos valores de píxel varían en el rango de 0 a 255.

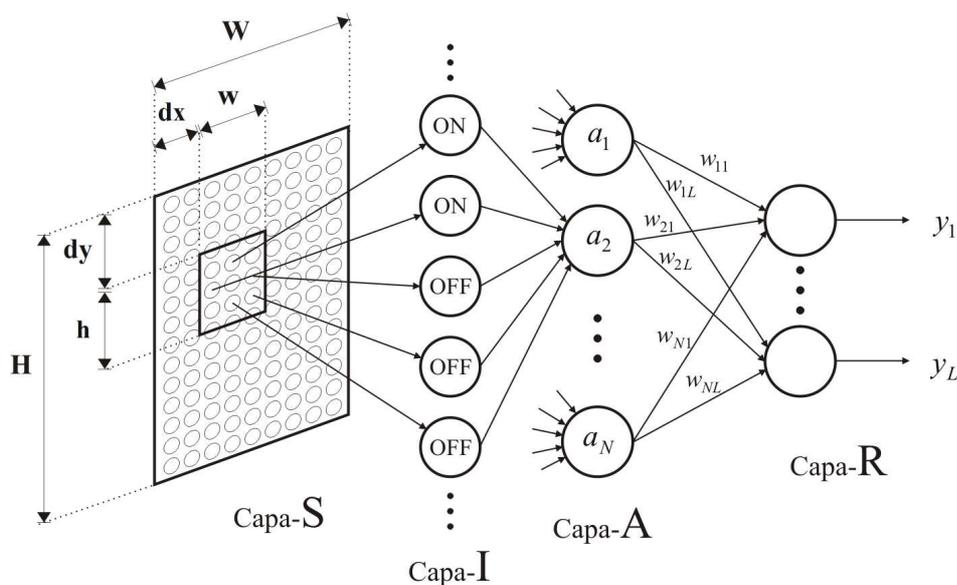


Figura 5.1: Arquitectura de LIRA Grayscale.

indica el valor de brillo mínimo o nulo (negro) y 255 se relaciona con el valor de brillo máximo de la imagen (blanco). Esta capa contiene $W \times H$ neuronas, donde W es el ancho y H es el alto de la imagen a ser clasificada. Dichas neuronas se encuentran organizadas de manera similar a los píxeles en la imagen.

La capa S se encuentra conectada con la capa A , a través de la capa intermedia I , por medio de un procedimiento aleatorio que se describirá posteriormente. Las conexiones que se establezcan serán del tipo no entrenables².

La capa de grupos I contiene N grupos de neuronas, donde N corresponde al número total de neuronas de la capa asociativa A . Existen dos tipos de neuronas en la capa I : neuronas ON y neuronas OFF. De acuerdo a los mecanismos cerebrales de la visión, presentados en [42], las neuronas ON reaccionan ante la presencia de un estímulo y las neuronas OFF se activan ante la ausencia del mismo. Estas neuronas tienen dos estados como valores de salida $\{0, 1\}$, donde un valor de salida 1 indica el estado activo de la neurona.

La capa asociativa A , en conjunto con las capas S e I , cumple la función de extractor de características de una imagen con dos estados de salida $\{0, 1\}$. La capa A se encuentra

² Una conexión no entrenable es aquella que no puede modificar el valor de su peso sináptico durante el proceso de entrenamiento.

completamente conectada con la capa R , donde los pesos en estas conexiones podrán ser modificados durante el proceso de entrenamiento.

La capa de salida R consiste de neuronas con funciones de activación de tipo lineal. El número de neuronas en esta capa corresponde al número de clases a reconocer. Los valores de excitación de las neuronas de la capa R son empleados para la obtención de la respuesta del sistema.

5.2. Conexiones

El procedimiento para conectar la capa de entrada S con la capa asociativa A a través de la capa de grupos I es el siguiente. Sea N el número total de neuronas asociativas. Para cada neurona asociativa a_k , donde $k = 1, \dots, N$, se selecciona aleatoriamente un área rectangular en la capa S (definida como ventana) de $h \times w$ neuronas (ver Figura 5.1). Para obtener las coordenadas de la esquina superior izquierda de la ventana como punto de referencia, se seleccionan valores aleatorios para los parámetros dx y dy del intervalo $[0, W - w]$ y $[0, H - h]$, respectivamente. Los valores de h y w se elijen de manera experimental de los rangos $[0, H]$ y $[0, W]$, respectivamente.

Dentro de la ventana resultante se escogen aleatoriamente m neuronas o puntos, los cuales se dividirán al azar en dos conjuntos, uno de p puntos positivos y otro de n puntos negativos, donde $p + n = m$. Se denomina punto positivo a la neurona de la capa S conectada con una neurona ON de la capa I , y punto negativo a la neurona de la capa S conectada con una neurona OFF de la capa I .

Cada uno de los p puntos positivos seleccionados se conectará a una neurona ON de la capa I que no contenga conexiones previas a otras neuronas de la capa S ; y cada uno de los n puntos negativos se enlazarán a una neurona OFF que, igualmente, no tenga enlaces previos con otras neuronas de la capa S . Este proceso formará un nuevo grupo k de neuronas ON y OFF en la capa I . Asimismo, cada neurona del grupo tendrá un umbral T_{mk} seleccionado de manera aleatoria del intervalo $[0, 255]$, y los subíndices m y k representan el índice de neurona dentro del grupo y el índice del grupo, respectivamente. Este grupo de neuronas obtenido se conectará a la neurona a_k , como se muestra en la Figura 5.1.

El proceso de conexión de las capas S , I y A se realizará una sola vez antes de dar inicio al entrenamiento del clasificador neuronal LIRA Grayscale, manteniendo fijos los valores de todos los parámetros mencionados anteriormente, con excepción de los valores de los pesos sinápticos entrenables ubicados en las conexiones entre la capa A y la capa de salida R .

5.3. Activación de Neuronas

Sea x_{ij} una neurona de entrada de la capa S . La salida de una neurona ON será igual a 1 (estado activo) si su valor de entrada x_{ij} es mayor o igual a un umbral T_{pk} ; y será igual a 0 en otro caso, esto es:

$$\varphi_{on}(x_{ij}) = \begin{cases} 1, & x_{ij} \geq T_{pk} \\ 0, & x_{ij} < T_{pk} \end{cases} \quad (5.1)$$

La salida de una neurona OFF será igual a 1 (estado activo) si su valor de entrada x_{ij} es menor o igual a un umbral T_{nk} ; y será igual a 0 en otro caso, esto es:

$$\varphi_{off}(x_{ij}) = \begin{cases} 1, & x_{ij} \leq T_{nk} \\ 0, & x_{ij} > T_{nk} \end{cases} \quad (5.2)$$

Una neurona asociativa tendrá salida igual a 1 (estado activo) si todas las neuronas ON y neuronas OFF del grupo conectado a ella están en estado activo, de cualquier otra manera su salida será 0. Cada neurona asociativa actúa como una característica dentro de la imagen, cuya salida indica si dicha característica está presente o ausente en la imagen.

5.4. Proceso de Entrenamiento

La red neuronal artificial que se propone emplea un procedimiento de entrenamiento supervisado que implementa un método de selección del ganador.

El entrenamiento supervisado es un paradigma de aprendizaje en el cual existe el conocimiento, de antemano, de la respuesta correcta del sistema en cuestión para un vector de entrada específico. En caso de que el sistema arroje una respuesta de salida

distinta a la correcta (ej. una clasificación errónea) para una muestra dada, se procede a realizar una corrección en los parámetros del modelo con base en el error generado.

La regla de selección del ganador consiste en elegir a la neurona de salida con mayor valor de excitación, siendo el índice de ésta el que indique la clase a la que pertenece el vector de entrada presentado a la red neuronal.

Si la excitación de la neurona ganadora es mucho mayor que la excitación de la neurona competidora más cercana, es decir, la neurona cuya salida tiene el segundo valor más alto, se puede asumir que la neurona ganadora domina no de manera aleatoria. Pero si la diferencia entre las excitaciones es pequeña, cabe la probabilidad de que la neurona ganadora obtuvo la máxima salida debido a factores aleatorios y no se podrá tener la seguridad que en el futuro, bajo las mismas condiciones de entrada, esta neurona gane de nuevo.

Debido a lo anterior, el método de selección del ganador propuesto en este trabajo consiste en aplicar a la regla simple la siguiente modificación: sea y_g la salida de la neurona ganadora y y_c la salida de la neurona competidora. Si

$$\frac{y_g - y_c}{y_g} < T_e \quad (5.3)$$

entonces, la neurona competidora es considerada ahora como ganadora, donde T_e es una constante llamada excitación adicional de la neurona ganadora. Esta modificación fue introducida para hacer más robusto y estable el proceso de entrenamiento, eliminando las posibilidades de que la neurona ganadora haya sido por cuestiones de azar.

La Ecuación 5.3 puede verse también como:

$$y_g(1 - T_e) < y_c \quad (5.4)$$

Se realizaron algunos experimentos específicos para investigar la influencia de esta modificación en el proceso de reconocimiento. Los correspondientes resultados se describirán más adelante.

Antes de dar inicio al proceso de entrenamiento los pesos sinápticos de las conexiones entre la capa A y la capa R son inicializados a cero.

El proceso de entrenamiento consiste de tres fases descritas a continuación.

5.4.1. Fase Inicial

El procedimiento de entrenamiento inicia con la presentación de la imagen a la capa de entrada S de la red neuronal. Las características de la imagen son extraídas y codificadas a través de la capa I y la capa A , respectivamente. Las salidas de las neuronas de la capa R son calculadas mediante la Ecuación 5.5:

$$y_i = \sum_{k=1}^N w_{ki} \cdot a_k \quad (5.5)$$

donde y_i es la salida (excitación) de la i -ésima neurona de la capa R , a_k es la salida (0 ó 1) de la k -ésima neurona de la capa A , w_{ki} es el peso de la conexión entre la k -ésima neurona de la capa A y la i -ésima neurona de la capa R .

5.4.2. Método de Selección del Ganador

Una vez calculada la salida de todas las neuronas de la capa R , la salida y_r , correspondiente a la neurona de la clase real, es modificada por el factor $(1 - T_e)$, esto es, se recalcula mediante la Ecuación 5.6:

$$y_r = y_r \cdot (1 - T_e) \quad (5.6)$$

Después de haber realizado lo anterior, la neurona con excitación máxima es seleccionada como la neurona ganadora g .

5.4.3. Adaptación de Pesos

Una vez que se ha obtenido la neurona ganadora (clase reconocida por la red neuronal), si la neurona r correspondiente a la clase real es igual a la neurona g correspondiente a la clase ganadora, esto es, $r = g$ entonces no se modifica ningún peso en las conexiones; pero si $r \neq g$, entonces:

$$\forall k, w_{kr}(t+1) = w_{kr}(t) + a_k \quad (5.7)$$

$$\forall k, w_{kg}(t+1) = w_{kg}(t) - a_k \quad (5.8)$$

$$\text{Si } (w_{kg}(t+1) < 0) \rightarrow w_{kg}(t+1) = 0 \quad (5.9)$$

donde, $w_{ki}(t)$ es el peso de la conexión entre la k -ésima neurona de la capa A y la i -ésima neurona de la capa R antes del refuerzo, $w_{ki}(t + 1)$ es el peso de la misma conexión después del refuerzo, a_k es el valor de salida de la k -ésima neurona de la capa A .

Este proceso será repetido hasta alcanzar un criterio de convergencia, en este caso, hasta llegar al número máximo de épocas de entrenamiento³ o que el número de errores de reconocimiento sea igual a cero.

5.5. Optimización del Proceso de Entrenamiento

Se realizaron algunas mejoras al proceso de entrenamiento para disminuir el tiempo de cómputo del procedimiento mismo y para incrementar el rango de reconocimiento del clasificador neuronal.

5.5.1. Codificación de Imágenes

Las características extraídas y codificadas en cada imagen a través de las primeras capas permanecen constantes para cada época de entrenamiento. Por este motivo se optó por realizar el proceso de extracción y codificación de características solamente una vez para todas las imágenes, almacenando en disco la lista de los índices de las neuronas asociativas activas por cada imagen. Posteriormente, durante la fase de entrenamiento, se emplearon dichas listas, correspondientes a las neuronas activas de la capa A , en lugar de las imágenes mismas, obteniendo con esto un decremento del tiempo necesario para realizar el proceso de entrenamiento.

Como ejemplo se tiene que, si la salida de la capa asociativa para una imagen determinada es de la forma: 0001000001000010, donde cero indica neurona desactivada y uno indica neurona activada, la correspondiente lista de índices de neuronas asociativas activas para dicha imagen sería: 3, 9, 14.

Esta forma de codificación permite un almacenamiento compacto de la información contenida en la imagen, si se toma en cuenta las miles de neuronas asociativas que po-

³ Una época o ciclo de entrenamiento consiste en la presentación de todas las muestras de entrenamiento a la red una vez.

drían constituir la capa asociativa. Asimismo, genera un cálculo rápido de la actividad neuronal para la capa de salida.

5.5.2. Transformaciones Geométricas de Imágenes

En la referencia [43] se muestra que el desempeño de los sistemas de reconocimiento puede ser mejorado por medio de la implementación de transformaciones geométricas (desplazamientos de posición y rotaciones) sobre la imagen de entrada durante el proceso de entrenamiento. En este trabajo, las imágenes empleadas en los experimentos presentan diferentes transformaciones geométricas, como son: desplazamientos horizontales, desplazamientos verticales y combinaciones de éstos (ver Figura 5.2).

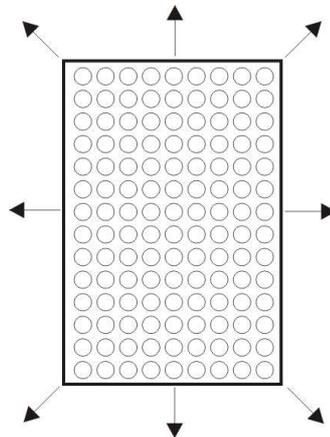


Figura 5.2: Transformaciones geométricas de imagen.

5.6. Proceso de Prueba

Al término del entrenamiento del clasificador neuronal LIRA Grayscale, una fase de prueba es llevada a cabo, en la cual un conjunto nuevo de imágenes de microtornillos es presentado a la red neuronal artificial obteniendo, de esta manera, un error de reconocimiento sobre el nuevo grupo de imágenes para establecer el grado de aprendizaje del sistema sobre la tarea en cuestión.

De acuerdo a lo explicado en el capítulo 4, el conjunto de imágenes de microtornillos empleados en esta etapa de evaluación nunca antes fue visto por el sistema, asegurando con esto la validez de los resultados obtenidos.

5.6.1. Selección del Ganador en el Proceso de Prueba

A diferencia del proceso de entrenamiento, en el proceso de prueba se emplea simplemente la regla de selección del ganador para obtener la salida del clasificador neuronal, esto es, se elegirá como clase reconocida el índice de la neurona de salida de la capa R que posea el mayor valor de excitación.

Capítulo 6

Resultados

En esta sección se expondrán los resultados obtenidos con la metodología propuesta en este trabajo. Asimismo, se plantearán y discutirán algunos puntos en relación a los resultados mostrados en este capítulo, al igual que algunas restricciones presentes en el sistema expuesto.

Para el desarrollo y evaluación del sistema de clasificación propuesto se crearon dos bases de datos de imágenes de microtornillos de 3 *mm* de diámetro, descritos en el capítulo 4. Los resultados correspondientes a cada base de datos se presentan a continuación.

6.1. Resultados con Base de Datos de 40 Imágenes

6.1.1. Captura de Imágenes Digitales

Durante la primera etapa de investigación se desarrolló una base de datos para la cual se capturó una imagen por cada uno de los 40 microtornillos manufacturados, dando un total de 40 imágenes en formato BMP con una resolución de 320×280 píxeles.

La imagen extraída en cada micropieza corresponde a la región de sus cuerdas, siendo ésta la base para caracterizar la forma del microobjeto (ver Figura 6.1).

Dado que cada uno de los 40 microtornillos pertenecen a una de cuatro clases diferentes, de igual manera, las imágenes fueron agrupadas en el mismo número de clases, donde cada una está constituida por diez imágenes de microtornillos distintos. En la

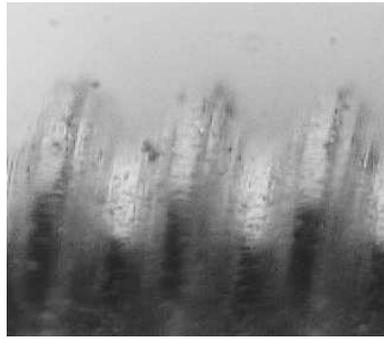


Figura 6.1: Ejemplo de una imagen digital capturada de un microtornillo.

Figura 6.2 se pueden apreciar ejemplos de las imágenes de la base de datos de 40 imágenes.

La implementación de las transformaciones geométricas de imagen, descritas en el capítulo 5, fue a través de software, generando un incremento en el número de imágenes a presentar en el proceso de entrenamiento. Dichas transformaciones (desplazamientos de la imagen original) fueron dos horizontales, dos verticales, y cuatro horizontales y verticales combinadas, de un píxel de distancia, dando un total de ocho transformaciones geométricas por cada imagen. De esta manera, se contó con nueve imágenes (ocho transformaciones más la original) por cada microtornillo de entrenamiento. A las imágenes correspondientes al proceso de prueba no se les aplicó transformación alguna.

6.1.2. Experimentos Preliminares

Previo a la realización de los experimentos de evaluación del sistema, se llevó a cabo una serie de experimentos preliminares. Con base a los resultados obtenidos en esta fase se seleccionó el conjunto de valores de parámetros no modificables de la red neuronal artificial LIRA Grayscale, que maximizaron el desempeño de la misma.

Para esta serie de experimentos se dividió el conjunto de 40 imágenes, de manera aleatoria, con el 80 % para el entrenamiento (32 imágenes) y el 20 % para la prueba (8 imágenes)¹, manteniendo esta relación de imágenes durante toda esta fase experimental.

¹ Para detalles de división de conjuntos de entrenamiento y prueba ver capítulo 4.

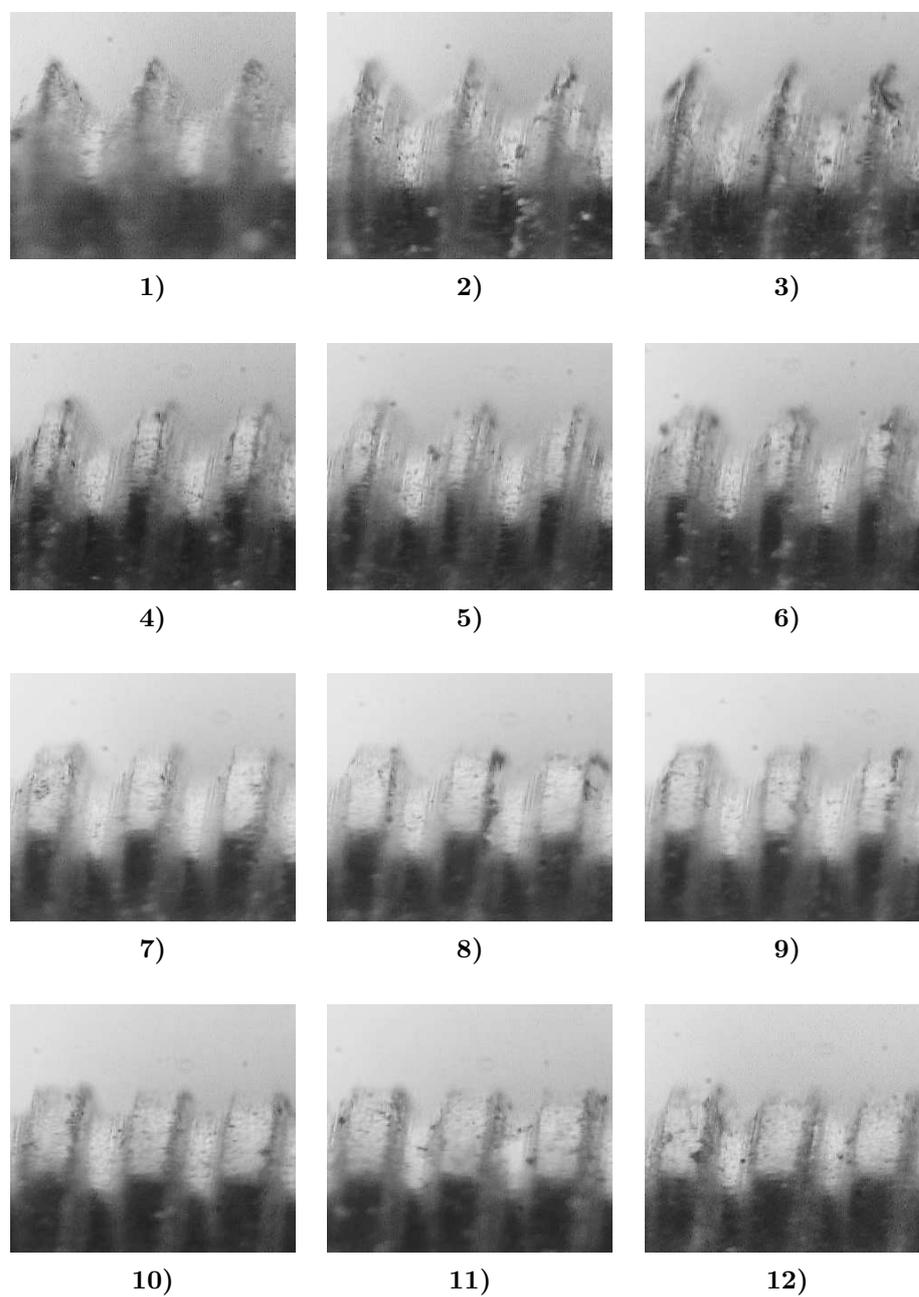


Figura 6.2: Ejemplos de imágenes de microtornillos. 1, 2 y 3) Imágenes de microtornillos de clase 0; 4, 5 y 6) Imágenes de microtornillos de clase 1; 7, 8 y 9) Imágenes de microtornillos de clase 2; y 10, 11 y 12) Imágenes de microtornillos de clase 3.

Para cada experimento se realizaron 20 corridas con el fin de obtener resultados estadísticamente confiables. Así, el número total de imágenes a reconocer fue calculado como el número de imágenes en el conjunto de prueba multiplicado por el valor de 20. En cada una de las 20 corridas se generaban nuevas conexiones entre la capa de entrada S y la capa asociativa A para obtener eventos independientes (para detalles de conexiones ver capítulo 5).

En las Tablas 6.1, 6.2 y 6.3 se puede observar la dependencia de los rangos de reconocimiento en relación al número de neuronas en la capa asociativa y el valor de excitación adicional (T_e) de la neurona ganadora. Estos resultados se obtuvieron con los parámetros fijos de 3 puntos positivos, 4 puntos negativos y una ventana $h \times w$ de dimensiones 10×10 .

Los valores de T_e por arriba de 0.3 generan un efecto oscilatorio del porcentaje de reconocimiento en el proceso de entrenamiento, evitando que el sistema logre su convergencia.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	23 / 160	85.63
0.1	21 / 160	86.88
0.2	18 / 160	88.75
0.3	16 / 160	90.00

Tabla 6.1: Porcentajes de reconocimiento con 64000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	20 / 160	87.50
0.1	18 / 160	88.75
0.2	15 / 160	90.63
0.3	13 / 160	91.88

Tabla 6.2: Porcentajes de reconocimiento con 128000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Dado que con un número de neuronas asociativas $N = 256000$ y una excitación adi-

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	19 / 160	88.13
0.1	15 / 160	90.63
0.2	13 / 160	91.88
0.3	11 / 160	93.13

Tabla 6.3: Porcentajes de reconocimiento con 256000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

cional $T_e = 0.3$ para la neurona ganadora se obtuvo el rango máximo de reconocimiento (93.13%) para estos experimentos, se determinó elegir estos valores como óptimos para dichos parámetros.

Para identificar las dimensiones de ventana y el número de puntos positivos y negativos que maximicen el desempeño del clasificador se realizaron experimentos cuyos resultados se muestran en las Tablas 6.4, 6.5 y 6.6.

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
4 / 5	26 / 160	83.75	156 / 102
3 / 5	23 / 160	85.63	366 / 243
2 / 5	21 / 160	86.88	952 / 586
3 / 4	15 / 160	90.63	613 / 493
2 / 4	16 / 160	90.00	1585 / 1021

Tabla 6.4: Porcentajes de reconocimiento con una ventana de dimensiones 5×5 y diferentes valores de puntos positivos y negativos.

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
4 / 5	24 / 160	85.00	146 / 100
3 / 5	22 / 160	86.25	387 / 306
2 / 5	16 / 160	90.00	926 / 514
3 / 4	7 / 160	95.63	576 / 472
2 / 4	9 / 160	94.38	1522 / 938

Tabla 6.5: Porcentajes de reconocimiento con una ventana de dimensiones 10×10 y diferentes valores de puntos positivos y negativos.

Dado que los mejores resultados fueron obtenidos empleando una ventana de di-

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
4 / 5	25 / 160	84.38	137 / 89
3 / 5	22 / 160	86.25	358 / 240
2 / 5	19 / 160	88.13	879 / 529
3 / 4	14 / 160	91.25	615 / 426
2 / 4	15 / 160	90.63	1586 / 958

Tabla 6.6: Porcentajes de reconocimiento con una ventana de dimensiones 15×15 y diferentes valores de puntos positivos y negativos.

mensiones 10×10 , 3 puntos positivos y 4 puntos negativos, estos serán elegidos como valores óptimos para los parámetros en cuestión.

En resumen, los valores de parámetros no modificables elegidos como resultados de esta fase de experimentos preliminares son los siguientes: número de neuronas asociativas $N = 256000$, una excitación adicional de la neurona ganadora $T_e = 0.3$, una ventana $h \times w$ de dimensiones 10×10 , 3 puntos positivos y 4 puntos negativos.

6.1.3. Proceso de Entrenamiento

El proceso de entrenamiento realizado con el conjunto correspondiente de imágenes de microtornillos fue de tipo secuencial, es decir, para todas las épocas o ciclos de entrenamiento las muestras fueron presentadas en un mismo orden.

En la Figura 6.3 se presenta un ejemplo de curva de aprendizaje del clasificador LIRA Grayscale con un entrenamiento de tipo secuencial.

Con el empleo de una presentación secuencial de muestras en la fase de entrenamiento, se requirió un promedio de 70 épocas para alcanzar la condición de paro de cero errores en el número de imágenes a reconocer, es decir, un 100% como porcentaje de reconocimiento del sistema.

6.1.4. Proceso de Prueba

Para obtener una estimación del desempeño del sistema a través del empleo de esta base de datos se realizaron tres experimentos de evaluación, donde para cada uno de

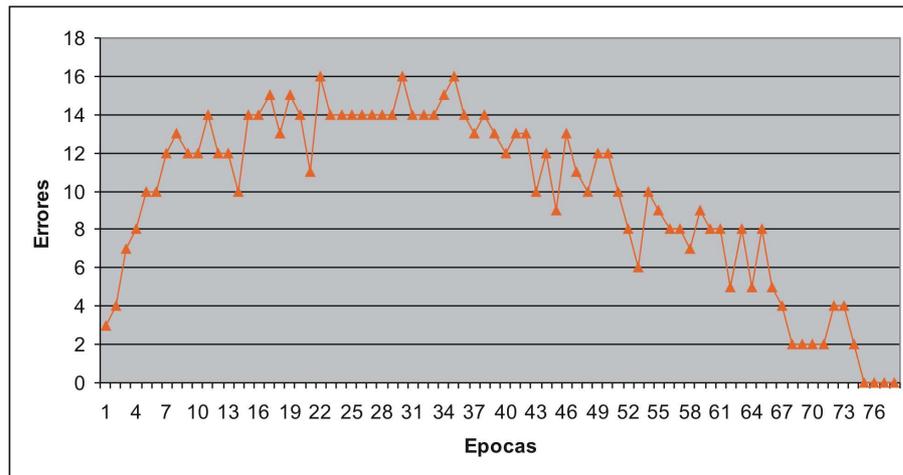


Figura 6.3: Curva de aprendizaje de entrenamiento con presentación secuencial de muestras para LIRA Grayscale.

éstos se utilizó un número diferente de imágenes de entrenamiento y prueba. Cada experimento fue corrido 20 veces para conservar la confiabilidad estadística. De igual manera, en cada experimento se varía el número de transformaciones geométricas de imagen aplicado al conjunto de imágenes de entrenamiento.

En el primer experimento se eligieron seis imágenes por clase para el proceso de entrenamiento y las cuatro restantes, de un total de diez imágenes por clase, para el proceso de prueba. Los rangos de reconocimiento obtenidos en este experimento se presentan en la Tabla 6.7.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	24 / 16	30 / 320	90.63
4	120 / 16	28 / 320	91.25
8	216 / 16	21 / 320	93.44

Tabla 6.7: Porcentajes de reconocimiento con 60 % de imágenes para entrenar y 40 % de imágenes para probar del conjunto total de 40 imágenes.

El segundo experimento consistió de siete imágenes por clase para el proceso de entrenamiento y tres para el proceso de prueba. Los resultados obtenidos se muestran en la Tabla 6.8.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	28 / 12	22 / 240	90.45
4	140 / 12	19 / 240	92.11
8	252 / 12	15 / 240	93.77

Tabla 6.8: Porcentajes de reconocimiento con 70 % de imágenes para entrenar y 30 % de imágenes para probar del conjunto total de 40 imágenes.

En el último experimento se seleccionaron ocho imágenes para entrenamiento y las dos restantes para el proceso de prueba. Los porcentajes de reconocimiento de este experimento se presentan en la Tabla 6.9.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	32 / 8	14 / 160	91.25
4	160 / 8	11 / 160	93.13
8	288 / 8	5 / 160	96.88

Tabla 6.9: Porcentajes de reconocimiento con 80 % de imágenes para entrenar y 20 % de imágenes para probar del conjunto total de 40 imágenes.

En esta fase, el porcentaje de reconocimiento obtenido en el mejor de los casos fue igual a 96.88 % (véase [44]).

6.2. Resultados con Base de Datos de 440 Imágenes

6.2.1. Captura de Imágenes

Con el fin de eliminar las transformaciones geométricas por software empleadas en la base de datos de 40 imágenes y mantener un número mayor de muestras de entrenamiento y prueba del sistema se construyó una segunda base de datos, para la cual se capturaron 11 imágenes por cada uno de los 40 microtornillos fabricados, dando un total de 440 imágenes en formato BMP con una resolución de 320×280 píxeles.

Cada una de las once imágenes correspondientes a un microtornillo presenta una transformación diferente de la posición de la pieza con respecto a un centro de gravedad

aproximado de la imagen, dando como resultado las imágenes siguientes:

- Imagen centrada
- Imágenes con desplazamiento horizontal (dos imágenes)
- Imágenes con desplazamiento vertical (dos imágenes)
- Imágenes con desplazamientos horizontales y verticales combinado (cuatro imágenes)
- Imagen con rotación de +3 grados
- Imagen con rotación de -3 grados

Cada desplazamiento fue realizado con dos píxeles de distancia. Las imágenes obtenidas son similares a las presentadas en la Figura 6.2.

6.2.2. Experimentos Preliminares

Los experimentos preliminares de esta base de datos fueron similares a los realizados en la base de datos de 40 imágenes, previamente expuestos, con la diferencia que en estos experimentos preliminares se empleó la cifra de 512000 neuronas asociativas y distintos valores de puntos positivos y negativos.

Para esta serie de experimentos se dividió el conjunto de 440 imágenes, de manera aleatoria, con el 80 % para el entrenamiento (352 imágenes) y el 20 % para la prueba (88 imágenes), manteniendo esta relación de imágenes durante toda esta fase experimental.

Para cada experimento se realizaron 30 corridas para obtener resultados estadísticamente confiables. Por tanto, el número total de imágenes a reconocer fue calculado como el número de imágenes en el conjunto de prueba multiplicado por el valor de 30. En cada una de las 30 corridas se generaban nuevas conexiones entre la capa de entrada S y la capa asociativa A para obtener eventos independientes.

En las Tablas 6.10, 6.11, 6.12 y 6.13 se puede observar la dependencia de los rangos de reconocimiento en relación al número de neuronas en la capa asociativa y el valor de excitación adicional (T_e) de la neurona ganadora.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	226 / 2640	91.44
0.1	180 / 2640	93.18
0.2	155 / 2640	94.13
0.3	128 / 2640	95.15

Tabla 6.10: Porcentajes de reconocimiento con 64000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	221 / 2640	91.63
0.1	160 / 2640	93.94
0.2	104 / 2640	96.06
0.3	88 / 2640	96.67

Tabla 6.11: Porcentajes de reconocimiento con 128000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	213 / 2640	91.93
0.1	112 / 2640	95.76
0.2	95 / 2640	96.40
0.3	78 / 2640	97.05

Tabla 6.12: Porcentajes de reconocimiento con 256000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Excitación Adicional (T_e)	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0.0	204 / 2640	92.27
0.1	98 / 2640	96.29
0.2	71 / 2640	97.31
0.3	67 / 2640	97.46

Tabla 6.13: Porcentajes de reconocimiento con 512000 neuronas asociativas contra diferentes valores de excitación adicional de la neurona ganadora.

Estos resultados se obtuvieron con los parámetros fijos de 3 puntos positivos, 4 puntos negativos y una ventana $h \times w$ de dimensiones 10×10 .

Los valores de T_e por arriba de 0.3 no permiten la convergencia del sistema y por lo tanto su aprendizaje.

En la Figura 6.4 se pueden apreciar ejemplos del área de estudio cubierta en una imagen con un número diferente de neuronas en la capa asociativa empleando un total de siete puntos distribuidos en una ventana de dimensiones 10×10 .

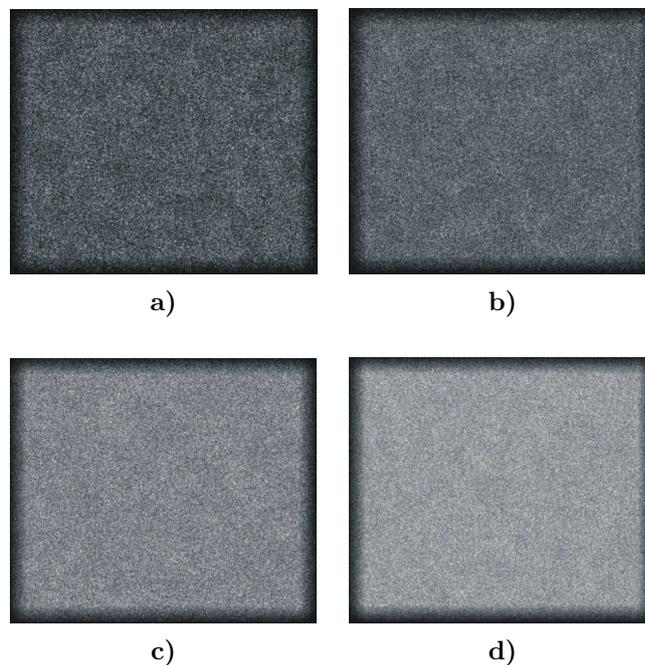


Figura 6.4: Ejemplos de áreas de cobertura en una imagen con: a) 64000 neuronas asociativas, b) 128000 neuronas asociativas, c) 256000 neuronas asociativas y d) 512000 neuronas asociativas.

Dado que con un número de neuronas asociativas $N = 512000$ y una excitación adicional $T_e = 0.3$ para la neurona ganadora se obtuvo el rango máximo de reconocimiento (97.46 %) para estos experimentos, se determinó elegir estos valores como óptimos para dichos parámetros.

Los resultados de experimentos para identificar las dimensiones de ventana y el número de puntos positivos y negativos que maximizan el desempeño del clasificador se muestran en las Tablas 6.14, 6.15 y 6.16.

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
5 / 4	20 / 2640	99.24	654 / 316
5 / 3	7 / 2640	99.74	1682 / 1037
5 / 2	20 / 2640	99.24	5409 / 4038
4 / 3	1 / 2640	99.96	2779 / 1762
4 / 2	1 / 2640	99.96	7858 / 5839

Tabla 6.14: Porcentajes de reconocimiento con una ventana de dimensiones 10×10 y diferentes valores de puntos positivos y negativos.

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
5 / 4	5 / 2640	99.81	681 / 336
5 / 3	6 / 2640	99.77	1732 / 1088
5 / 2	10 / 2640	99.62	5530 / 4197
4 / 3	1 / 2640	99.96	2873 / 1824
4 / 2	1 / 2640	99.96	8040 / 6091

Tabla 6.15: Porcentajes de reconocimiento con una ventana de dimensiones 20×20 y diferentes valores de puntos positivos y negativos.

P. Positivos / P. Negativos	Errores de Reconocimiento / Total Imágenes a Reconocer	Porcentaje de Reconocimiento (%)	Neuronas Activas Max / Min
5 / 4	10 / 2640	99.62	691 / 342
5 / 3	6 / 2640	99.77	1767 / 1149
5 / 2	16 / 2640	99.39	5676 / 4326
4 / 3	3 / 2640	99.88	2966 / 1899
4 / 2	1 / 2640	99.96	8265 / 6344

Tabla 6.16: Porcentajes de reconocimiento con una ventana de dimensiones 30×30 y diferentes valores de puntos positivos y negativos.

Dado que el mejor porcentaje de reconocimiento (99.96 %) fue obtenido en distintos casos, se introdujo el siguiente criterio. Además de buscar los valores que maximicen el rango de reconocimiento del clasificador, un criterio adicional para la selección del número de puntos positivos y negativos conectados a un grupo de neuronas ON y OFF en la capa intermedia I relacionado a una neurona asociativa es elegir estos valores de tal manera que el número de neuronas asociativas activas K esté limitado por la siguiente expresión:

$$K = c \cdot \sqrt{N} \quad (6.1)$$

donde, K es el número de neuronas asociativas activas para una imagen determinada, N es el número de neuronas asociativas y c es una constante seleccionada de manera experimental del rango $[1, 5]$.

Para este caso, eligiendo un valor de $c = 5$, el valor de K debe ser a lo más $5 \cdot \sqrt{512000} = 5 \cdot 715.54 = 3577.71$, esto es, $K \approx 3578$ como máximo valor. Por lo tanto, el número de puntos positivos y negativos que cumplen con el criterio y además maximizan el desempeño del clasificador son 4 y 3, respectivamente.

Dado que la ventana de 10×10 y 20×20 generan resultados similares al emplear 4 puntos positivos y 3 puntos negativos, es indistinta la selección de una u otra ventana, por lo tanto, se elegirá la ventana de 20×20 , ya que en toda la columna en general muestra mejores resultados que la ventana de 10×10 .

En resumen, los valores de parámetros no modificables elegidos como resultados de esta fase de experimentos preliminares son los siguientes: número de neuronas asociativas $N = 512000$, una excitación adicional de la neurona ganadora $T_e = 0.3$, una ventana $h \times w$ de dimensiones 20×20 , 4 puntos positivos y 3 puntos negativos.

6.2.3. Proceso de Entrenamiento

El porcentaje de reconocimiento durante el proceso de entrenamiento fue siempre del 100 %, ya que el criterio de paro para este proceso fue que el número de errores de reconocimiento sea igual a cero.

El proceso de entrenamiento realizado con el conjunto correspondiente de imágenes de microtornillos fue de tipo aleatorio, es decir, en cada época de entrenamiento las

imágenes o muestras fueron presentadas en un orden distinto.

En la Figura 6.5 se muestra un ejemplo de curva de aprendizaje de la red neuronal LIRA Grayscale con un entrenamiento donde las muestras se presentan de manera aleatoria en cada época.

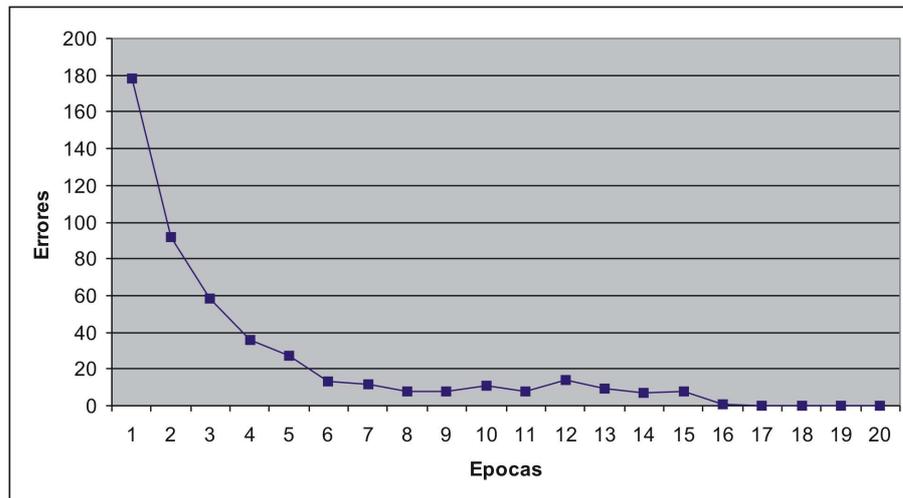


Figura 6.5: Curva de aprendizaje de entrenamiento con presentación aleatoria de muestras para LIRA Grayscale.

Utilizando este tipo de presentación de muestras en el entrenamiento, el promedio de épocas necesarias para conseguir el aprendizaje del sistema fue de 17.

6.2.4. Proceso de Prueba

Para estimar el desempeño del sistema se realizaron cuatro experimentos principales de evaluación. Cada experimento tiene un número diferente de imágenes para los conjuntos de entrenamiento y prueba, respectivamente. Asimismo, se introducen transformaciones geométricas de imagen por software, variando el número de ellas en cada uno de estos cuatro experimentos, con el fin de analizar el efecto que tienen sobre el porcentaje de reconocimiento con este nuevo conjunto de imágenes. Para mantener la confiabilidad estadística de los resultados cada experimento fue corrido 30 veces.

En el primer experimento se eligió el 50% del conjunto total de imágenes para el proceso de entrenamiento y el 50% para el proceso de prueba. Los resultados de este caso se presentan en la Tabla 6.17.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	220 / 220	135 / 6600	97.96
4	1100 / 220	117 / 6600	98.23
8	1980 / 220	115 / 6600	98.26

Tabla 6.17: Porcentajes de reconocimiento con 50 % de imágenes para entrenar y 50 % de imágenes para probar del conjunto total de 440 imágenes.

El segundo experimento utilizó el 60 % del conjunto total de imágenes para entrenar y el restante 40 % para el proceso de prueba. Los resultados de este caso se presentan en la Tabla 6.18.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	264 / 176	88 / 5280	98.33
4	1320 / 176	86 / 5280	98.37
8	2376 / 176	84 / 5280	98.41

Tabla 6.18: Porcentajes de reconocimiento con 60 % de imágenes para entrenar y 40 % de imágenes para probar del conjunto total de 440 imágenes.

Para el tercer caso se empleó el 70 % del conjunto total de imágenes para el entrenamiento y el restante 30 % para probar el sistema. Los rangos de reconocimiento para este experimento se muestran en la Tabla 6.19.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	308 / 132	57 / 3960	98.56
4	1540 / 132	56 / 3960	98.59
8	2772 / 132	52 / 3960	98.69

Tabla 6.19: Porcentajes de reconocimiento con 70 % de imágenes para entrenar y 30 % de imágenes para probar del conjunto total de 440 imágenes.

Por último, en el cuarto experimento utilizó el 80 % del conjunto total de imágenes para el proceso de entrenamiento y el 20 % para el proceso de prueba. Los resultados de este experimento se presentan en la Tabla 6.20.

Número de Transformaciones de Imagen	Imágenes Entrenamiento / Imágenes Prueba	Errores de Reconocimiento / Total de Imágenes a Reconocer	Porcentaje de Reconocimiento (%)
0	352 / 88	34 / 2640	98.71
4	1760 / 88	33 / 2640	98.75
8	3168 / 88	29 / 2640	98.9

Tabla 6.20: Porcentajes de reconocimiento con 80 % de imágenes para entrenar y 20 % de imágenes para probar del conjunto total de 440 imágenes.

En esta fase, el porcentaje de reconocimiento obtenido en el mejor de los casos fue igual a 98.90 %.

6.3. Velocidad de Clasificación

Además del desempeño de clasificación, la velocidad de clasificación es una propiedad importante de un sistema. Todos los experimentos fueron realizados en un procesador Intel Pentium M de 1.6 *GHz* con 512 *Mb* de memoria RAM, consumiendo un tiempo total de 1 hora aproximadamente para llevar a cabo 30 corridas de un experimento con 440 imágenes, donde 352 imágenes son para entrenar y 88 imágenes para probar. Los tiempos para cada operación se desglosan de la siguiente manera: 57 *s* para la creación de los conjuntos de entrenamiento y prueba, 3 *s* para crear las conexiones entre capas, 50 *s* para la codificación de las imágenes, 2 *s* para entrenar, y 12 *s* para probar las imágenes.

El tiempo promedio calculado para clasificar una sola imagen de microtornillo por el sistema desarrollado, después de ser entrenado, es de 0.1364 segundos (136.4 milisegundos), aproximadamente.

Capítulo 7

Discusión

A continuación se plantearán y discutirán algunos puntos en relación a los resultados mostrados en el capítulo 6, así como algunas restricciones presentes en el sistema desarrollado.

7.1. Tamaño de la Muestra de Entrenamiento

Con base a los resultados obtenidos se puede apreciar que el tamaño de la muestra de entrenamiento es un parámetro importante, ya que, a medida que el número de imágenes de entrenamiento crece se observa un incremento en el porcentaje de reconocimiento. Esto es debido a que cada imagen en el conjunto de entrenamiento posee diferentes características que alimentan el aprendizaje del sistema y mientras más características del objetivo a reconocer se le presenten mayor aprendizaje será adquirido por el mismo.

En el caso de la base de datos de 40 imágenes, por ejemplo, en la Tabla 6.7 con 120 imágenes de entrenamiento se obtuvo 91.25 % de reconocimiento y en la Tabla 6.9 con 160 imágenes de entrenamiento se obtuvo 93.13 %, esto es, 1.88 % más de exactitud en el reconocimiento.

En el caso de la base de datos de 440 imágenes, sucede una situación similar, en la Tabla 6.17 con 220 imágenes de entrenamiento se obtuvo 97.96 % de reconocimiento y en la Tabla 6.20 con 352 imágenes de entrenamiento se obtuvo 98.71 %, esto es, 0.75 % mayor exactitud en reconocimiento.

7.2. Transformaciones Geométricas de Imagen

Otra cuestión que ayuda a incrementar el rango de reconocimiento es la aplicación de transformaciones geométricas de imagen. Dado que una imagen transformada actúa como una imagen independiente para el clasificador neuronal, se pueden crear imágenes adicionales a partir de una sola, dándole al sistema de reconocimiento una fuente mayor de aprendizaje. Por ejemplo, en las Tablas 6.8 y 6.18, respectivamente, a medida que el número de transformaciones de imagen aumenta el rango de reconocimiento incrementa.

Aunque las transformaciones de imagen implementadas por software arrojan resultados aceptables, y son fáciles de generar, siempre será mejor emplear transformaciones capturadas de manera directa del objeto y no introducidas de manera artificial, evitando así la inclusión de ruido en las imágenes transformadas, el cual podría afectar al clasificador disminuyendo su rango de reconocimiento.

7.3. Entrenamiento Secuencial y Aleatorio

De acuerdo a las gráficas mostradas en las Figuras 6.3 y 6.5 se puede observar que el tiempo requerido para el proceso de reconocimiento del sistema decremanta ante el empleo de un entrenamiento aleatorio en lugar de uno secuencial. A manera de comparación, en la Figura 7.1 se puede observar la diferencia gráfica de convergencia en ambos tipos de entrenamiento.

7.4. Imágenes Mal Clasificadas

Un punto importante de discusión es el conjunto de imágenes no reconocidas por la red neuronal LIRA Grayscale en esta tarea de clasificación. En la Figura 7.2 se presentan las imágenes de microtornillos que generaron los errores de reconocimiento durante las fases de prueba.

El problema inherente en las imágenes de las Figuras 7.2(a), 7.2(c) y 7.2(d) es la falta de claridad y definición del objeto en la imagen, ésto dificulta la clasificación correcta

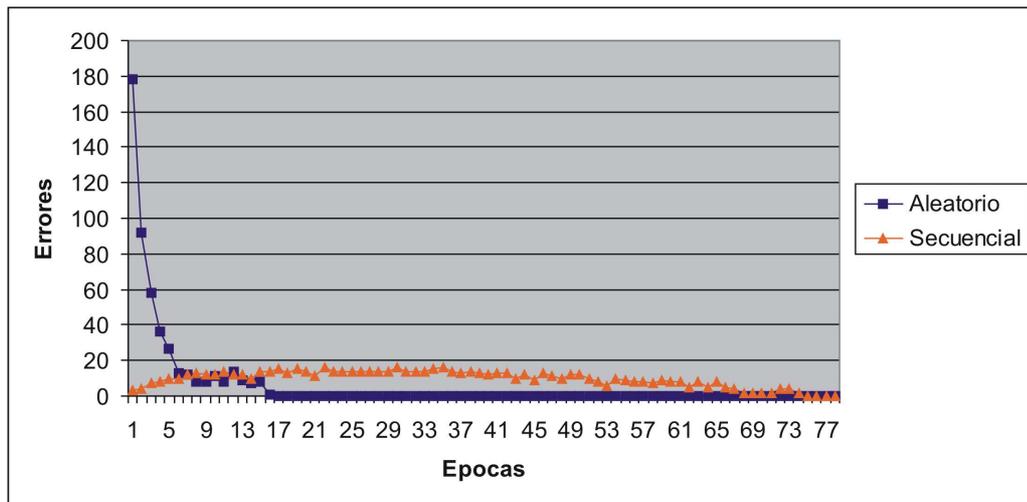


Figura 7.1: Curvas de aprendizaje de entrenamiento con presentación aleatoria y secuencial de muestras.

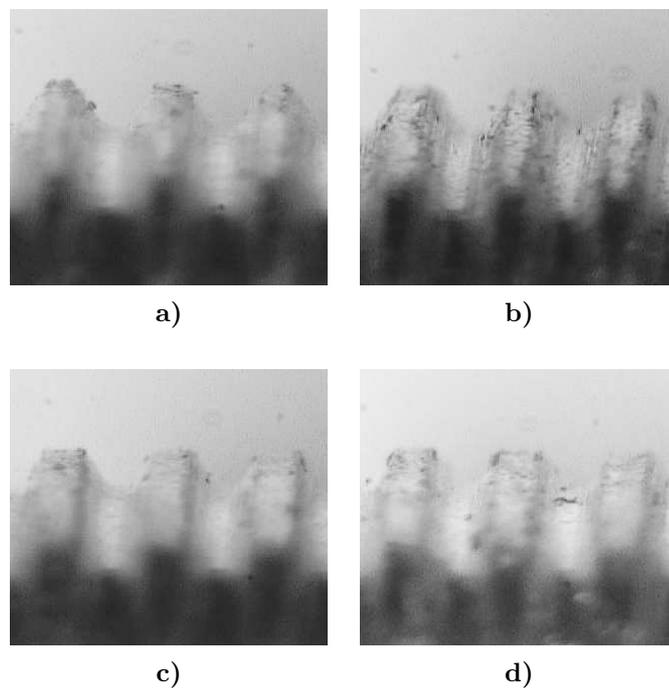


Figura 7.2: Ejemplos de imágenes de microtornillos mal clasificadas. a) Microtornillo de clase 1 reconocido como clase 2, b) Microtornillo de clase 2 reconocido como clase 1, c y d) Microtornillos de clase 2 reconocidos como clase 3.

de la pieza. Este tipo de error en la imagen fue adquirido al momento de capturar la misma.

Cabe hacer notar que en algunos casos, el sistema puede manejar el problema de imágenes borrosas, realizando una clasificación correcta de las mismas, como es el caso de las imágenes mostradas en la Figura 7.3.

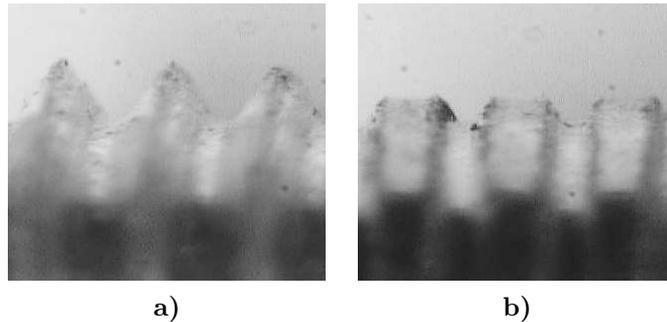


Figura 7.3: Ejemplos de imágenes de microtornillos bien clasificadas a pesar del ruido presente en las imágenes. a) Microtornillo de clase 0 y b) Microtornillo de clase 3.

Un caso particular es el presentado por el microtornillo de la Figura 7.2(b), el cual fue manufacturado y clasificado dentro de los microtornillos de clase 2, sin embargo, al observar detenidamente la forma de la pieza se puede notar que el tamaño de la cuerda queda en un punto intermedio entre el tamaño de cuerda de la clase 1 y el de la clase 2, con mayor tendencia a parecerse a las de clase 1. Es muy probable que por esta razón el sistema la reconozca como imagen de clase 1.

7.5. Restricción del Sistema

La restricción presente en el sistema de clasificación expuesto radica en la falta de heurísticas para estimar el valor de los parámetros no entrenables del clasificador neuronal LIRA Grayscale, obligando a la realización de múltiples experimentos preliminares para elegir los valores que maximicen el desempeño del sistema.

Capítulo 8

Conclusiones

El sistema basado en el clasificador neuronal LIRA Grayscale para el reconocimiento y clasificación de imágenes de microtornillos después del proceso de manufactura fue desarrollado.

La red neuronal LIRA Grayscale fue entrenada y probada con dos bases de datos constituidas por imágenes de 40 microtornillos, de 3 *mm* de diámetro, de cuatro clases distintas, donde cada clase está relacionada con una determinada posición de la herramienta de corte.

La presentación de muestras de entrenamiento en orden aleatorio permitió obtener un menor tiempo para alcanzar la convergencia del sistema en la fase de entrenamiento a comparación de una presentación de muestras secuencial.

Dado que el tamaño de muestra de entrenamiento es un parámetro influyente en el porcentaje de reconocimiento del sistema, la relación para dividir los conjuntos de entrenamiento y prueba que obtuvo mejores resultados fue: 80 % de muestras para entrenamiento y 20 % de muestras para prueba.

La inclusión de transformaciones geométricas de imagen incrementa el rango de reconocimiento del sistema.

El porcentaje de reconocimiento para la fase de entrenamiento fue de 100 % y para las fases de prueba, en el mejor de los casos, fue de:

- 96.88 % para la base de datos de 40 imágenes.

- 98.90 % para la base de datos de 440 imágenes.

En general, se obtuvieron resultados muy aceptables, superando los obtenidos por el clasificador neuronal PCNC (92.5 %). Sin embargo, no se debe eliminar la posibilidad de intentar otros métodos para poder hacer más comparaciones y optimizar los métodos desarrollados en este trabajo para mejorar el porcentaje de reconocimiento en esta tarea de clasificación.

Finalmente, se quiere hacer notar que la mayor contribución del clasificador LIRA Grayscale es su universalidad. No fue diseñado para una tarea particular de reconocimiento de forma, ya que puede ser adaptado e implementado en diferentes problemas de reconocimiento de imágenes.

8.1. Trabajo a Futuro

Dado que el sistema desarrollado fue implementado únicamente fuera de línea, el siguiente objetivo en este trabajo será la conversión a un sistema de clasificación que opere en línea y poder ser integrado con las micromáquinas de tecnología MET.

También será necesario un incremento en la precisión de reconocimiento del sistema. Esto podrá lograrse aumentando el número de clases de microtornillos fabricando micropiezas con una mayor variedad de posiciones de la herramienta de corte. De igual manera, esta optimización generará una mayor precisión al realizar la reposición del cortador en caso de requerirlo.

Otro objetivo a futuro será lograr que el sistema pueda identificar y clasificar el defecto en el microtornillo antes del término de su proceso de manufactura, evitando con esto una pérdida total de la micropieza.

Un último punto a considerar sería el desarrollo de heurísticas para determinar los mejores valores de los parámetros no modificables de la red neuronal LIRA Grayscale, evitando con ello un número considerable de experimentos previos al entrenamiento del sistema.

Bibliografía

- [1] R.C. Gonzalez and R.E. Woods. *Digital image processing*. Prentice Hall, New Jersey, second edition, 2002.
- [2] T. Baidyk, E. Kussul, O. Makeyev, A. Caballero, L. Ruiz, G. Carrera, and G. Velasco. Flat image recognition in the process of microdevice assembly. *Pattern Recognition Letters*, 25(1):107–118, January 2004.
- [3] T. Baidyk and E. Kussul. Neural network based vision system for micro workpieces manufacturing. *WSEAS Transactions on Systems*, 3(2):483–488, April 2004.
- [4] T. Baidyk, E. Kussul, and O. Makeyev. Texture recognition with random subspace neural classifier. *WSEAS Transactions on Circuits and Systems*, 4(4):319–324, April 2004.
- [5] G.K. Toledo, E. Kussul, and T. Baidyk. Neural classidier LIRA for recognition of micro work pieces and their positions in the processes of microassembly and micromanufacturing. *Proceedings of the 7 All-Ukrainian International Conference*, pages 17–20, October 2004.
- [6] R.P. Lippmann. An introduction to computing with neural nets. *IEEE ASSP Mag.*, 4:4–22, April 1987.
- [7] Y.H. Pao. *Adaptive pattern recognition and neural networks*. Reading, Addison-Wesley, 1989.
- [8] K. Fukushima. A neural network for visual pattern recognition. *IEEE Computer*, 21(3):65–75, March 1988.
- [9] B. Widrow, R.G. Winter, and R.A. Baxter. Layered neural nets for pattern recognition. *IEEE Trans. ASSP-36*, (7):1109–1118, July 1988.
- [10] M. Mayorga and L. Ludeman. Neural nets for determination of texture and its orientation. *IEEE ICASSP*, pages 2689–2692, May 1991.
- [11] Q.M.J. Wu, M.F.R. Lee, and C.W. de Silva. Intelligent 3–D sensing in automated manufacturing processes. *IEEE/ASME International Conference on Advanced Intelligent Mechatronics Proceedings*, pages 366–370, 2001.
- [12] S.J. Lee, K. Kim, D.H. Kim, J.O. Park, and G.T. Park. Recognizing and tracking of 3–D shaped micro parts using multiple visions for micromanipulation. *IEEE International Symposium on Micromechatronics and Human Science*, pages 203–210, 2001.

- [13] J.Y. Kim and H.S. Cho. A vision based error-corrective algorithm for flexible parts assembly. *Proceedings of the IEEE International Symposium on Assembly and Task Planning*, pages 205–210, 1999.
- [14] W.S. Trimmer. *Micromechanics and MEMS. Classical and seminal papers to 1990*. IEEE Press, New York, 1997.
- [15] <http://www.lizardfire.com>.
- [16] F. Simonis and S. Schilthuizen. *Nanotechnology: Innovation opportunities for tomorrow's defence*. TNO Science & Industry, March 2006.
- [17] H. Bleuler, R. Clavel, J.M. Breguet, and E. Pernette. Issues in precision motion control and microhandling. *Proceedings of the 2000 IEEE International Conference on Robotics and Automation*, pages 959–964, 2000.
- [18] Some micro machine activities in Japan. *Report ATIP96.021*, 1996.
- [19] N. Ooyama, S. Kokaji, M. Tanaka, K. Ashida, N. Mishima, H. Maekawa, T. Tanikawa, and K. Kaneko. Desktop machining microfactory. *Proceedings of the 2-nd International Workshop on Microfactories*, pages 13–16, October 2000.
- [20] Y. Okazaki and T. Kitahara. Micro-lathe equipped with closed-loop numerical control. *Proceedings of the 2-nd International Workshop on Microfactories*, pages 87–90, October 2000.
- [21] E. Kussul, D. Rachkovskij, T. Baidyk, and S. Talayev. Micromechanical engineering: A basis for the low cost manufacturing of mechanical microdevices using microequipment. *Journal of Micromechanics and Microengineering*, 6(4):410–425, December 1996.
- [22] E. Kussul, T. Baidyk, L. Ruiz, A. Caballero, G. Velasco, and L. Kasatkina. Development of micromachine tool prototypes for microfactories. *Journal of Micromechanics and Microengineering*, 12:795–812, 2002.
- [23] G.J. Tortora and S.R. Grabowski. *Principios de anatomía y fisiología*. Oxford University Press, ninth edition, 2002.
- [24] J.A. Anderson and E. Rosenfeld. *Neurocomputing: Foundations of research*. MIT Press, 1988.
- [25] S. Haykin. *Neural Networks: A comprehensive foundation*. Prentice Hall, second edition, 1999.
- [26] W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.
- [27] J.R. Hilera and V.J. Martínez. *Redes neuronales artificiales: fundamentos, modelos y aplicaciones*. RA-MA, 1995.
- [28] B.J.A. Kröse and P.P. van der Smagt. *Introduction to neural networks*. The University of Amsterdam, fifth edition, 1997.

- [29] F. Rosenblatt. The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65:386–408, 1958.
- [30] B. Widrow and M.E. Hoff. Adaptive switching circuits. *1960 IRE WESCON Convention Record*, pages 96–104, 1960.
- [31] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by backpropagation errors. *Nature*, 323:533–536, 1986.
- [32] J.A. Anderson. *Neural models with cognitive implications*. D. LaBerge and S.J. Samuels (Eds.), Basic Processes in Reading: Perception and Comprehension, New Jersey, 1977.
- [33] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43:59–69, 1982.
- [34] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79:2554–2558, 1982.
- [35] D.O. Hebb. *The Organization of Behavior*. Wiley, New York, 1949.
- [36] S. Grossberg. How does a brain built a cognitive code? *Psychological Review*, 87:1–51, 1980.
- [37] C.E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.
- [38] D.H. Ackley, G.E. Hinton, and T.J. Sejnowski. A learning algorithm for Boltzmann machines. *Cognitive Science*, 9:147–169, 1985.
- [39] F. Rosenblatt. *Principles of Neurodynamics*. Spartan Books, Washington, DC, 1962.
- [40] M. Minsky and S. Papert. *Perceptrons: An introduction to computational geometry*. MA: MIT Press, Cambridge, 1969.
- [41] E. Kussul and T. Baidyk. Improved method of handwritten digit recognition tested on MNIST database. *Image and Vision Computing*, 22(12):971–981, October 2004.
- [42] D.H. Hubel and T.N. Wiesel. Brain mechanisms of vision. *Scientific American*, 241:150–162, 1979.
- [43] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2344, November 1998.
- [44] A. Martín and T. Baidyk. Neural classifier for microscrew shape recognition in micromechanics. *CLEI*, August 2006.