

**UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE CIENCIAS PURAS Y NATURALES
CARRERA DE INFORMATICA**



TESIS DE GRADO

**“CUANTIFICACIÓN Y DETECCIÓN DE ALTERACIONES
MORFOLÓGICAS DE GLOBULOS ROJOS EN IMÁGENES
DIGITALES MICROSCÓPICAS”**

PARA OPTAR AL TÍTULO DE LICENCIATURA EN INFORMATICA
MENCION: INGENIERIA DE SISTEMAS INFORMATICOS

POSTULANTE: BRIGIDA TARQUI QUISPE

TUTOR METODOLÓGICO: M. Sc. MIGUEL COTAÑA MIER

ASESOR: LIC. MARCELO ARUQUIPA CHAMBI

LA PAZ – BOLIVIA

2013

DEDICATORIA

Dedico este presente trabajo a mis Padres Eugenia y Marcelo porque nunca perdieron la fe en mí y siempre me apoyaron y colaboran en todo momento.

A mis hermanos por su apoyo, confianza y colaboración en tantos años de buenos y malos momentos.

AGRADECIMIENTOS

A Dios por ser quien ha estado a mi lado en todo momento dándome las fuerzas necesarias para continuar luchando día tras día y seguir adelante.

A Lic. M. Sc. Miguel Cotaña Mier, quien gentilmente dedico su tiempo al seguimiento de este trabajo brindándome su apoyo y conocimiento para la elaboración y culminación de este trabajo.

A Lic. Marcelo Aruquipa Chambi, quien fue mi revisor, agradecerle pos su paciencia y tiempo, que me otorgo en el transcurso del desarrollo de este trabajo y que con su conocimiento profesional realizo el seguimiento llevando a la culminación de este trabajo.

Por último agradezco a todos mis amigos quienes me acompañaron en el transcurso de estos años, en especial a Gloria, Edith, Judith y Maritsa por ser amigos de verdad, por todo el apoyo y comprensión que incondicionalmente me brindan.

RESUMEN

Debido al avance de tecnologías y herramientas como procesamiento de imágenes y el reconocimiento de patrones mediante redes neuronales aplicadas en áreas como la medicina, biología, industria, agricultura, en otras, ha cobrado gran importancia.

El uso de estas herramientas es esencial en el análisis de imágenes microscópicas para la generación de diagnósticos, en hematología el estudio de muestras sanguíneas pueden servir para diagnosticar algún tipo de enfermedad hematológica.

Este trabajo propone utilizar técnicas de procesamiento de imágenes y reconocimiento de patrones permita la cuantificación y detección de alteraciones morfológicas de glóbulos rojos en imágenes microscópicas de muestras de sangre.

Para este trabajo se desarrollo un algoritmo de identificación y cuantificación de las alteraciones morfológicas de glóbulos rojos mediante descriptores elementales de forma.

Se diseño e implemento el modelo de red neuronal de base radial el cual clasifica 5 tipos de alteraciones morfológicas según la variación de forma mediante 7 descriptores geométricos, se utilizo la metodología en cascada para el desarrollo del software y Matlab como lenguaje de programación.

INDICE

CAPITULO I MARCO REFERENCIAL

| | |
|--|---|
| 1.1 INTRODUCCIÓN..... | 1 |
| 1.2 ANTECEDENTES..... | 2 |
| 1.3. PLANTEAMIENTO DEL PROBLEMA..... | 5 |
| 1.3.1 Formulación del Problema..... | 5 |
| 1.4 OBJETIVOS..... | 5 |
| 1.4.1 Objetivo General..... | 5 |
| 1.4.2 Objetivos Específicos..... | 6 |
| 1.5 HIPÓTESIS..... | 6 |
| 1.5.1 Identificación de Variables..... | 6 |
| 1.6 JUSTIFICACIÓN..... | 7 |
| 1.6.1 Justificación Técnica..... | 7 |
| 1.6.2 Justificación Social..... | 7 |
| 1.6.3 Justificación Económica..... | 7 |
| 1.7 LIMITES Y ALCANCES..... | 7 |
| 1.7.1 Limites..... | 7 |
| 1.7.2 Alcances..... | 8 |
| 1.8 APORTES..... | 8 |
| 1.9 DISEÑO METODOLÓGICO..... | 9 |

CAPITULO II MARCO TEORICO

| | |
|-------------------------------------|----|
| 2.1 HEMATOLOGÍA..... | 10 |
| 2.1.1 HEMOGRAMA..... | 10 |
| 2.1.2 LA SANGRE..... | 10 |
| 2.1.2.1 Elementos de la Sangre..... | 10 |

| | | |
|-----------|---|----|
| 2.1.3 | GLÓBULOS ROJOS..... | 11 |
| 2.1.3.1 | Alteraciones Morfológicas de los Glóbulos Rojos..... | 12 |
| 2.1.3.2 | Tipos de Alteraciones Morfológicas de los Glóbulos Rojos..... | 12 |
| 2.1.3.2.1 | Alteraciones de Tamaño de los eritrocitos..... | 12 |
| 2.1.3.2.2 | Alteraciones de Forma de los eritrocitos..... | 12 |
| 2.1.3.2.3 | Alteraciones de Coloración de los eritrocitos..... | 13 |
| 2.1.3.4 | FROTIS SANGUINEO | 16 |
| 2.1.3.4.1 | Tinción del Frotis Sanguíneo..... | 16 |
| 2.1.3.4.2 | Tipos de Tinción del Frotis Sanguíneo..... | 17 |
| 2.2 | PROCESAMIENTO DIGITAL DE IMÁGENES..... | 18 |
| 2.2.1 | REPRESENTACION DE UNA IMAGEN DIGITAL..... | 19 |
| 2.2.1.1 | Imagen en niveles de gris..... | 19 |
| 2.2.1.2 | Imágenes binarias..... | 19 |
| 2.2.2 | REPRESENTACIÓN DEL COLOR..... | 20 |
| 2.2.2.1 | Espacio de Color RGB..... | 20 |
| 2.2.3 | SEGMENTACIÓN DE LA IMAGEN..... | 21 |
| 2.2.3.1 | Binarizacion..... | 21 |
| 2.2.4 | EXTRACCIÓN DE CARACTERÍSTICAS..... | 22 |
| 2.2.5 | DESCRIPTORES ELEMENTALES DE FORMA..... | 23 |
| 2.2.5.1 | Coefficiente de Circularidad..... | 24 |
| 2.2.5.2 | Coefficiente Elíptico..... | 24 |
| 2.3. | RECONOCIMIENTO DE PATRONES..... | 24 |
| 2.3.1. | ALGORITMOS DE CLASIFICACIÓN..... | 25 |
| 2.3.2 | REDES NEURONALES..... | 25 |
| 2.3.2.1 | RED NEURONAL BIOLÓGICA..... | 26 |
| 2.3.2.2 | RED NEURONAL ARTIFICIAL..... | 27 |
| 2.3.2.2.1 | Elementos de una red neuronal artificial..... | 28 |
| 2.3.2.2.2 | Arquitectura de las redes neuronales artificiales..... | 29 |
| 2.3.2.2.3 | Aprendizaje de una red neuronal artificial..... | 29 |
| 2.3.2.2.4 | Tipos de Aprendizaje..... | 30 |

| | |
|---|----|
| 2.3.2.2.5 Tipos y Modelos de redes neuronales artificiales..... | 30 |
| 2.3.3 REDES DE FUNCIONES DE BASE RADIAL..... | 33 |
| 2.3.3.1 ARQUITECTURA DE LAS REDES RBF | 33 |
| 2.3.3.2 MODELO GENÉRICO DE UNA RED RBF..... | 34 |
| 2.3.3.3 APRENDIZAJE EN REDES DE FUNCIONES DE BASE RADIAL | 37 |
| 2.4 METODOLOGIA DE DESARROLLO DEL SOFTWARE..... | 39 |
| 2.4.1 METODOLOGIA DE DESARROLLO EN CASCADA..... | 39 |
| 2.4.2 DESCRIPCIÓN DE ACTIVIDADES..... | 40 |
| 2.4.2.1 Análisis de los requisitos del software..... | 40 |
| 2.4.2.2 Diseño del sistema..... | 40 |
| 2.4.2.3 Implementación o Codificación..... | 40 |
| 2.4.2.4 Prueba de sistema..... | 40 |
| 2.4.2.5 Mantenimiento..... | 41 |
| 2.5 LENGUAJE DE MODELADO UNIFICADO UML..... | 41 |
| 2.6 PRUEBA DE HIPÓTESIS..... | 41 |
| 2.6.1 Prueba de t Student..... | 43 |
| 2.7 LENGUAJE DE PROGRAMACIÓN..... | 43 |
| 2.7.1 Matlab..... | 43 |

CAPITULO III MARCO APLICATIVO

| | |
|---|----|
| 3.1 MODELO GENERAL DEL PROTOTIPO..... | 46 |
| 3.1.1 Descripción General..... | 46 |
| 3.1.1.1 Adquisición de Imágenes..... | 46 |
| 3.1.1.2 Procesamiento de Imagen..... | 47 |
| 3.1.1.3 Identificación y Cuantificación de alteraciones morfológicas de glóbulos rojos mediante descriptores elementales de forma..... | 48 |
| 3.1.1.4 Extracción de Características..... | 50 |
| 3.1.1.5 Normalizar..... | 50 |
| 3.1.1.6 Reconocimiento mediante Redes Neuronales..... | 51 |

| | | |
|-----------|--|----|
| 3.1.1.6.1 | Diseño de la Red Neuronal..... | 51 |
| 3.1.1.6.2 | Arquitectura de la Red Neuronal..... | 51 |
| 3.1.1.6.3 | Entrenamiento de la Red Neuronal..... | 53 |
| 3.2 | DESARROLLO DEL PROTOTIPO..... | 55 |
| 3.2.1 | ANALISIS DE REQUERIMIENTOS..... | 55 |
| 3.2.1.1 | Diagrama de casos de uso..... | 55 |
| 3.2.1.2 | Diagrama de casos de uso en formato expandido..... | 56 |
| 3.2.2 | DISEÑO DEL PROTOTIPO..... | 60 |
| 3.2.2.1 | Diagrama de Secuencia..... | 60 |
| 3.2.2.2 | Diagrama de Colaboración..... | 64 |
| 3.2.2.3 | Diagrama de Clases..... | 66 |
| 3.2.2.4 | Diseño de Interfaces..... | 67 |
| 3.2.3 | IMPLEMENTACION Y PRUEBAS..... | 69 |
| 3.2.3.1 | Implementación del prototipo..... | 68 |
| 3.3.3.1.1 | Implementación de algoritmos basado en descriptores Elementales de forma..... | 70 |
| 3.2.3.2 | Pruebas del prototipo..... | 71 |
| 3.3 | PRUEBA DE HIPÓTESIS..... | 72 |

CAPITULO IV CONCLUSIONES Y RECOMENDACIONES

| | | |
|-----|----------------------|----|
| 4.1 | CONCLUSIONES..... | 74 |
| 4.2 | RECOMENDACIONES..... | 75 |
| | BIBLIOGRAFÍA | 76 |
| | ANEXOS | 78 |

INDICE DE FIGURAS Y TABLAS

| | |
|--|----|
| Figura 2.1: Imagen de una muestra de Globulos Rojos..... | 11 |
| Figura 2.2: Imágenes con diferentes tipos de tincion..... | 16 |
| Figura 2.3: Etapas del procesamiento digital de imágenes..... | 18 |
| Figura 2.4: Imagen digital en un arreglo de mxn pixeles..... | 19 |
| Figura 2.6: Representacion de los componentes RGB del color..... | 20 |
| Figura 2.7: Imagen de extendido sanguineo de globulos rojos..... | 21 |
| Figura 2.9: Etapas de reconocimiento de patrones..... | 24 |
| Figura 2.10: Neuronas Biologicas..... | 27 |
| Figura 2.11: Neuronas Artificiales y sus elementos..... | 27 |
| Figura 2.11: Funciones de Activacion..... | 29 |
| Figura 2.12: Arquitectura de las redes neuronales artificiales..... | 29 |
| Figura 2.13: Arquitectura de las redes neuronales RBF..... | 33 |
| Figura 2.14: Funcion gaussiana de una red neuronal RBF..... | 35 |
| Figura 2.15: Funcion sigmoidea de una red neuronal RBF..... | 36 |
| Figura 2.16: Ciclo de vida del modelo en cascada..... | 39 |
| Figura 3.1: Modelo General..... | 46 |
| Figura 3.2: Imagen de globulos rojos con eliptocitos..... | 47 |
| Figura 3.3: Aplicación de algoritmos de Procesamiento de imágenes..... | 48 |
| Figura 3.4: Algoritmo de identificación y conteo de Alteraciones Morfológicas de los Glóbulos Rojos | 49 |
| Figura 3.7: Arquitectura de la Red Neuronal de Base Radial | 52 |
| Figura 3.8: Diagrama de casos de uso..... | 54 |
| Figura 3.9: Diagrama de Secuencia..... | 60 |
| Figura 3.21: Diagrama de Colaboración..... | 62 |

| | |
|--|----|
| Figura 3.22: Diagrama de Clases..... | 65 |
| Figura 3.23: Interfaz gráfica del procesamiento de imágenes..... | 66 |
| Figura 3.24: Interfaz grafica del entrenamiento de la red neuronal..... | 67 |
| Figura 3.25: Interfaz gráfica del reconocimiento..... | 68 |
| Figura 3.26: Algoritmo basado en descriptores elementales de forma..... | 69 |
| Figura 3.27: Distribución de t_0 y $t_{\alpha/2}$ en el grafico para la toma de decisión..... | 71 |
| Tabla 2.1: Alteraciones Morfológicas de los Glóbulos Rojos..... | 13 |
| Tabla 3.1: Vector de Características..... | 46 |
| Tabla 3.2: Vector de características Normalizado..... | 46 |
| Tabla 3.3: Vector de entrada..... | 47 |
| Tabla 3.4: Vector de salida..... | 52 |
| Tabla 3.5: Datos de Entrenamientos..... | 52 |
| Tabla 3.6: Parametros de la red neuronal de base radia..... | 52 |
| Tabla 3.7: Descripción de caso de uso: Ingresar al sistema..... | 55 |
| Tabla 3.8: Descripción de caso de uso: Abrir imagen..... | 56 |
| Tabla 3.9: Descripción de caso de uso: Identificar y contar células..... | 56 |
| Tabla 3.10: Descripción de caso de uso: Extraer Características..... | 57 |
| Tabla 3.11: Descripción de caso de uso: Entrenar Red Neuronal..... | 58 |
| Tabla 3.12: Descripción de caso de uso: Reconocer células..... | 60 |
| Tabla 3.13: Resultados del prototipo..... | 71 |

CAPITULO I

MARCO REFERENCIAL

1.1 INTRODUCCION

Las Alteraciones Morfológicas de los Glóbulos Rojos o eritrocitos en la sangre son provocadas por algunas enfermedades, como las Anemias que afecta a entre 300 y 500 millones de personas en el mundo anualmente, según datos de OMS. Existen alrededor de 200 millones de portadores con algún tipo de anemia.

Actualmente en nuestro país, el diagnóstico de anomalías morfológicas de los eritrocitos es un proceso realizado de manera tradicional o automática. Para detectar y contar estas alteraciones de los glóbulos rojos se lo realiza mediante un examen o análisis de sangre. Este examen es realizado mediante la tinción de láminas de extendido sanguíneo o frotis sanguíneo, es un mecanismo científico que consiste de una gota de sangre en la superficie de un portaobjetos con el fin de analizarla, bajo un microscopio por un hematólogo, en ella pueden identificarse claramente dos tipos de células contra un fondo uniforme. La primera clase que corresponde a los glóbulos rojos (GR) normales consta de células de dimensiones y formas muy similares, mientras que la segunda clase corresponde a los glóbulos rojos deformados o con variaciones, esta clase se conforma de células de mayor tamaño y de formas irregulares [Lanz 98].

Existe gran cantidad de analizadores o contadores automáticos para laboratorios clínicos. Estos equipos utilizan diferentes tecnologías:

Citometría de flujo, fluorescencia, impedancia, espectrofotometría, etc. También diversas técnicas de dilución con uno o más reactivos para el conteo de células sanguíneas. A pesar que estos equipos son altamente eficientes tienen algunos inconvenientes como: recalibración de sus sensores ópticos, Controladores de fluido necesarios para controlar la tasa de flujo de fluido a través de sus sensores, gran volumen de reactivos y diluyentes requeridos para realizar las pruebas de sangre, y a su vez crean problemas como la eliminación de residuos, aumentos en el costo de la prueba, gran consumo de energía, etc. Estos analizadores automáticos se limitan a detectar células normales y, a lo más, indican la presencia de células anormales, inmaduras o desconocidas.

Procesamiento Digital de Imágenes y Reconocimiento de patrones están tomando un gran impulso e importancia en el campo de la Inteligencia Artificial (IA), en la línea de desarrollar sistemas que permitan automatizar el análisis de imágenes.

En el presente proyecto de Tesis, se desarrollara un prototipo de cuantificación y detección de alteraciones morfológicas de glóbulos rojos mediante técnicas de procesamiento digital de imágenes y reconocimiento de patrones que existe en el campo de la Inteligencia Artificial.

1.2 ANTECEDENTES

Los primeros intentos para contar el número de células sanguíneas fueron realizados en el siglo XVII, el investigador Van Leeuwenheok utilizando un rudimentario microscopio y la sangre de un pollo, Van Leeuwenheok contó el número de eritrocitos presentes en una muestra, depositada en un capilar graduado de vidrio.

La técnica general del conteo de células empleando un microscopio, en un contenedor transparente, se fue desarrollando gradualmente, a través de mejoras tanto al microscopio como al contenedor, hasta llegar al diseño de "cámaras de conteo", esto después de dos siglos de investigación continua.

Es en el siglo XIX cuando se desarrollan las técnicas de dilución de la sangre, que permiten el diseño de cámaras más precisas.

Los trabajos relacionados con el tema de investigación son los siguientes:

“Segmentación de las Imágenes de la sangre usando operadores Morfológicos”.

Elaborado por Carlos Platero, Ingeniería de Sistemas, Escuela Universitaria de Ingeniería Técnica Industrial (UPM), año 2004.

Este trabajo describe una parte de un sistema de proceso palúdico de imagen, para detectar, clasificar y contar los parásitos de la malaria en imágenes dentro de las diapositivas manchadas en la sangre para evaluar el parasitemia de la sangre. Un requisito importante del sistema es un método eficiente para dividir imágenes de la célula en segmentos.

“Segmentación y Clasificación de las Células de sangre blanca”

Elaborado por Damián Almarales Núñez, Investigación en Robótica y Sistemas Inteligentes, Universidad del Norte, Colombia, año 2000.

En este trabajo aborda el problema del análisis blanco en la forma de la célula de sangre basado en las características morfológicas de su contorno y núcleos externos.

“Desarrollo de un Distema de Análisis Automático de imágenes de extendidos sanguíneos”.

Elaborado por Roger J. Bello, Computación Gráfica y Geometría Aplicada, Universidad Central de Venezuela, año 2004.

Este trabajo utiliza reconocimiento de imágenes para detectar glóbulos rojos infectados con la malaria, estos glóbulos rojos al ser infectados cambian de forma lo que hace posible su reconocimiento. Como resultado de este trabajo ellos lograron hacer el diagnóstico de la malaria de las personas con muestras de diferentes laboratorios con un gran porcentaje de certeza.

Existen trabajos de tesis realizadas en la Carrera de Informática de la Universidad Mayor de San Andrés, los cuales son los siguientes:

“Reconocimiento de Células Sanguíneas Leucocitarias mediante Redes Neuronales” [Bustos m. 2009].

En este trabajo de tesis se utiliza el modelo de red neuronal Backpropagation para el reconocimiento de leucocitos, a partir de una imagen de muestra sanguínea, este modelo de red neuronal está diseñada para el reconocimiento de 5 tipos de leucocitos utiliza como parámetros de entrada descriptores de características, además utiliza el método de umbralización técnica de tratamiento digital de imágenes.

“Reconocimiento de Cromosomas mediante Redes Neuronales” [Morales m. 2000].

En este trabajo de tesis se aplica el modelo de red neuronal Backpropagation para el reconocimiento de cromosomas en imágenes, este modelo de red neuronal reconoce 7 grupos de cromosomas por bandeado G, clasifica según la posición del centrómero y tamaño, utiliza además el método de umbralización técnica de tratamiento digital de imágenes ampliamente utilizada en la segmentación de imágenes.

“Reconocimiento de Patrones en imágenes digitales de Cromosomas”[Fernández f. 2008].

La red neuronal que se utiliza en esta Tesis es el modelo de Hopfield para el reconocimiento de patrones de cromosomas de Bandeamiento G en imágenes digitales, este modelo de red neuronal reconoce y clasifica 7 grupos de cromosomas según la posición del centrómero y tamaño.

“Reconocimiento de Cromosomas Humanos en imágenes basado en una Red Neuronal de Retropropagación” [Quelca r. 2011].

En este trabajo de tesis se utiliza técnicas de tratamiento digitales de imágenes como el algoritmo de detección de bordes de Prewit y umbralización para detectar los cromosomas en muestras de imágenes para su posterior clasificación por el modelo de red neuronal Retropropagación,

el cual clasifica 5 tipos de cromosomas en imágenes digitales por Bandeamiento G.

“reconocimiento de patrones de impresiones nasales en canes usando redes neuronales” [chacon g. 2010].

En este trabajo de tesis se aplica el método de binarización técnica de tratamiento digitales de imágenes y se utiliza red neuronal Kohonen que requiere que sus los valores de las entradas de la red oscilen entre -1 y 1, esta red neuronal está diseñada para reconocer las impresiones nasales en canes en imágenes digitales.

1.3. PLANTEAMIENTO DEL PROBLEMA

La cuantificación¹ y detección de las alteraciones morfológicas de glóbulos rojos en muestras de sangre son procesos de un examen morfológico de Frotis, Para llevar a cabo este examen, el hematólogo necesita observar la muestra y realizar manualmente estos procesos, esta manera general de realizar este examen puede tardar varios horas e incluso días y pueden cometerse errores en el análisis de las muestras sanguíneas.

1.3.1 FORMULACION DEL PROBLEMA

El problema de investigación que se plantea es el siguiente:

¿Las técnicas de procesamiento digital de imágenes y redes neuronales aplicadas a imágenes digitales de muestras de sangre permitirán la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos?

1.4 OBJETIVOS

1.4.1 Objetivo General

Desarrollar un prototipo que, mediante técnicas de procesamiento digital de imágenes y redes neuronales permita cuantificar y detectar alteraciones

morfológicas de los glóbulos rojos en imágenes digitales de muestras de sangre.

1.4.2 Objetivos Específicos

- Desarrollar un algoritmo que permita identificar y cuantificar las alteraciones morfológicas de glóbulos rojos en imágenes digitales de muestras de sangre.
- Diseñar e implementar un modelo de red neuronal para el reconocimiento de las alteraciones morfológicas de glóbulos rojos en imágenes digitales de muestras de sangre.
- Construir una interfaz gráfica y amigable que permita la visualización de las imágenes y los resultados.

1.5. HIPÓTESIS

La hipótesis que se plantea surge del análisis del árbol de problemas y objetivos.

“La Aplicación de técnicas de Procesamiento digital de Imágenes y reconocimiento de patrones optimizara la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos en imágenes digitales de muestras de sangre.”

1.5.1 Identificación de Variables

a) Variable Independiente

Procesamiento Digital de imágenes, reconocimiento de patrones, imagen digital de una muestra de sangre.

b) Variable Dependiente

Cuantificación y Detección de las alteraciones morfológicas de glóbulos rojos

1.6 JUSTIFICACIÓN

1.6.1 Justificación Técnica

El presente proyecto de tesis es justificable técnicamente ya que se cuenta con equipos de computación necesarios para la desarrollo del prototipo, así también con recursos de software.

1.6.2 Justificación Social

El prototipo a desarrollarse se justifica socialmente, constituye un aporte al personal médico en el área de hematología de laboratorios clínicos, hospitales, centros de salud, etc. Ya que será una herramienta de apoyo al diagnóstico médico para el análisis de la morfología de los glóbulos rojos, ayudara a facilitar las tareas repetitivas y rutinarias como es el conteo y la clasificación de los distintos tipos de anormalidades de los glóbulos rojos, ayudará a mejorar el control de los resultados de las pruebas que se genera en el laboratorio, aportando confiabilidad y calidad en los mismos. Ya que estos resultados permitirán el diagnóstico acertado por parte del médico, lo cual beneficiará al paciente.

1.6.3 Justificación Económica

El prototipo a desarrollarse es económicamente factible, pues los beneficios serán mayores que los gastos debido a que reducirá los procesos manuales de conteo y de detección de alteraciones morfológicas de glóbulos rojos,

reducirá el tiempo de análisis de cada muestra de sangre, aumentara la productividad del hematólogo.

1.7 LIMITES Y ALCANCES

1.7.1 LIMITES

- El prototipo se limitara a cuantificar y detectar de las alteraciones morfológicas de glóbulos rojos según la variación de forma (poiquilocitosis).
- Las imágenes digitales microscópicas de las muestras de sangre estarán en formato JPEG.
- El prototipo se limitara a cuantificar y detectar las alteraciones morfológicas de glóbulos rojos en imágenes digitales microscópicas de extendido sanguíneo con cualquier tipo de tinción Romanowsky.
- El prototipo no utilizara un gestor de base de datos para el almacenamiento de las imágenes.

1.7.2 ALCANCES

- El prototipo reconocerá 5 tipos de alteración morfológica de los glóbulos rojos según a la variación de forma de estas células (poiquilocitosis) en imágenes microscópicas.
- El algoritmo de identificación y cuantificación de las alteraciones morfológicas de glóbulos rojos se basara en descriptores de forma para determinar si un eritrocito o glóbulo rojo es normal o patológico.
- El prototipo no contempla a imágenes digitales de muestras de sangre con solapamiento de glóbulos rojos.

1.8. APORTES

El desarrollo de un prototipo para cuantificar y detectar las alteraciones morfológicas de glóbulos rojos constituye un gran aporte al sistema de salud especialmente a los laboratorios clínicos en el diagnóstico, tratamiento y control de enfermedades de la sangre como la anemia, en nuestro país, además ayudara a dar seguridad de los resultados obtenidos.

1.9 DISEÑO METODOLOGÍCO

1.9.1 METODO DE INVESTIGACION

Se utilizara el método de investigación científica hipotético-deductiva, siguiendo los lineamientos por [Mario Bunge, 1997] de la siguiente manera:

- a) **Observación.** Se realizara la obtención de información necesaria para la detección y cuantificación de las alteraciones morfológicas de los glóbulos rojos en imágenes digitales, así como información sobre el procesamiento de imágenes y de redes neuronales.
- b) **Planteo de la hipótesis.** Es posible identificar las tipos de alteraciones morfológicas de los glóbulos rojos mediante el reconocimiento de patrones a través de una red neuronal artificial.
- c) **Diseño de aplicación.** Se procederá al diseño de una red neuronal para el reconocimiento de patrones en imágenes digitales de glóbulos rojos.
- d) **Casos de prueba.** Se realizaran casos de prueba en el prototipo de software para evaluar la calidad de software.
- e) **Conclusiones.** Se realizara las conclusiones una vez terminada la investigación relacionada con las alteraciones morfológicas de los glóbulos rojos.

1.9.2 TIPO DE INVESTIGACION

El tipo de investigación será teórica y aplicada, de tipo teórica debido a que se basara en fundamentos de redes neuronales para el reconocimiento de patrones de las distintas alteraciones morfológicas de los glóbulos rojos y de tipo aplicada porque se aplicara o utilizara conocimientos para la construcción de un prototipo de software.

CAPITULO II MARCO TEORICO

2.1 HEMATOLOGIA

La hematología es el estudio de la sangre, de los tejidos formadores de sangre, los órganos hematopoyéticos² y los desórdenes asociados a ellos. Los hematólogos mediante un análisis de hematología en laboratorios clínicos, diagnostican y tratan desordenes de la sangre, como la leucemia, la anemia y la hemofilia, incluyendo los ganglios linfáticos, la medula ósea y el bazo [Theml et.al., 2004].

2.1.1 HEMOGRAMA

El Hemograma es un examen cuantitativo y cualitativo de las diferentes poblaciones celulares en una muestra de sangre. El examen nos da información de parámetros hematológicos como: Recuento total del número de células (eritrocitos, leucocitos y trombocitos) en millones*mm³, Recuento diferencial de glóbulos blancos (WCB) en millones*mm³, Calculo de concentración de la hemoglobina (HGB) en g/ dl, índices eritrocitarios (VCM, HCM, RDW, VPM y PRD.), índices plaquetarios, etc.

El Hemograma también es llamado Biometría Hemática, Citometria o Citología Hemática o Conteo Sanguíneo Completo (CSC).

2.1.2 LA SANGRE

La sangre es una forma especializada de tejido conectivo que consta de elementos formes, o células sanguíneas, y una sustancia intercelular líquida, el plasma sanguíneo, la sangre constituye alrededor del 8% del peso corporal y son de dos tipos principales: rojas (eritrocitos) y blancas (leucocitos). [Dihigo & Menendez, 1987].

2.1.2.1 ELEMENTOS DE LA SANGRE

Los elementos formes de sangre son células sanguíneas, los cuales son: glóbulos rojos (eritrocitos o hematíes) son alrededor de 4.5 a 5.5 millones por mm^3 , los glóbulos

Blancos (leucocitos) hay un promedio de 5.000 a 9.000 leucocitos por mm^3 en la sangre humana normal, y las plaquetas entre 200.000 a 300.000, denominadas trombocitos [Moore et. al., 1970].

2.1.2.2. GLOBULOS ROJOS

Los glóbulos rojos, también llamados eritrocitos o hematíes son células formes cuantitativamente más numerosas de la sangre. Los eritrocitos tiene la forma de un disco bicóncavo para adaptarse a una mayor superficie de intercambio de oxígeno por dióxido de carbono en los tejidos, miden alrededor entre 7 y 9 micras de diámetro en promedio, y 1.9 micras en su parte más gruesa en frotis secos, son de color rojizo debido a la hemoglobina, no tienen núcleo y carecen de mitocondrias, tiempo de vida es alrededor de 120 días aproximadamente, oscilan en el cuerpo humano alrededor de 4.5 a 5.5 millones por mm^3 , sin embargo esta cantidad varía según el sexo, la edad, el peso y el lugar geográfico en que habita la

persona. Los eritrocitos constituyen el 95% a 99% del total del volumen células sanguíneo [Gigola G,2005].

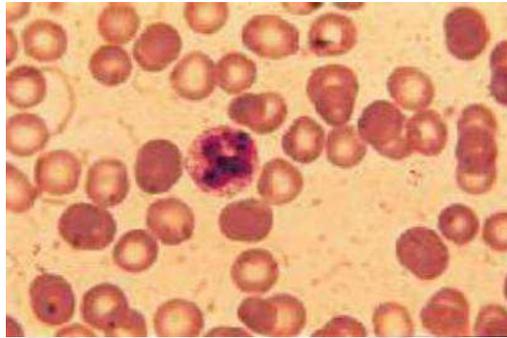


Figura 2.1: Imagen de una muestra de sangre de globulos rojos

Fuente: [Atlas de Hematología,2005]

2.1.2.3. ALTERACIONES MORFOLOGICAS DE LOS GLOBULOS ROJOS

Las alteraciones morfológicas son anomalías o deformaciones morfológicas de los eritrocitos o glóbulos rojos se presentan en tres categorías: Anomalías en el tamaño, anomalías en la forma y anomalías de coloración. Estas anomalías son detectadas en el extendido sanguíneo o frotis coloreado con un método de tinción hematológico, observado bajo un microscopio a través de este examen se puede diagnosticar diferentes tipos de anemias [García D, 2005].

2.1.2.4 TIPOS DE ALTERACIONES MORFOLOGICAS DE LOS GLOBULOS ROJOS

Las Alteraciones morfológicas de los Glóbulos rojos se clasificación según la variación de tamaño, forma y coloración [García D, 2005].

2.1.2.4.1 ALTERACIONES DE TAMAÑO DE LOS ERITROCITOS

La variación del tamaño de los eritrocitos se llama Anisocitosis. El tamaño del diámetro de la célula roja oscila entre 6.2-8.2 micras con un tamaño promedio de 7.2 micras, según al tamaño se clasifican en los siguientes:

- Microcitos
- Macrocitos

2.1.2.4.2 ALTERACIONES DE LA FORMA DE LOS ERITROCITOS

La variación o deformación en la forma de los eritrocitos se llama *Poiquilocitosis*. Las alteraciones morfológicas de los glóbulos rojos según la forma son los siguientes:

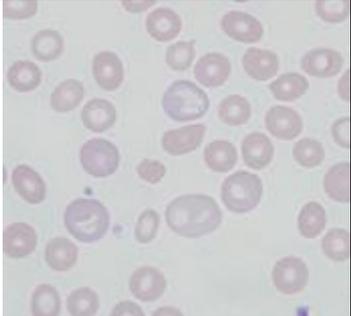
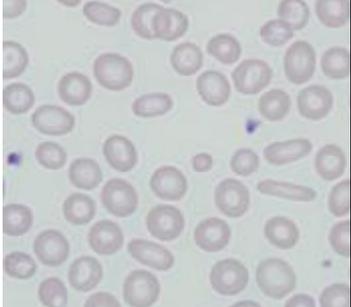
- Eliptocitos
- Esferocitos
- Equinocitos
- Dianocitos
- Drepanocitos
- Estomatocitos

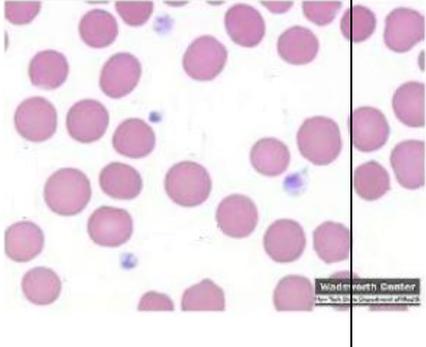
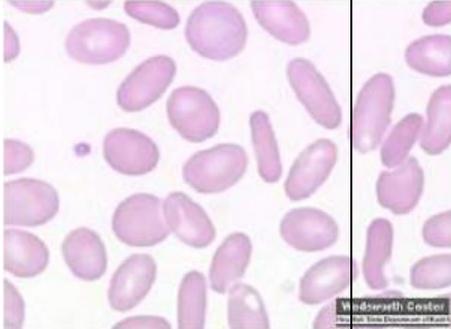
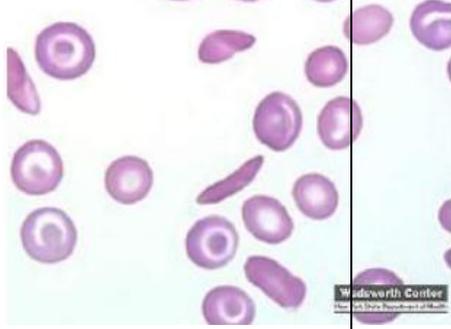
2.1.2.4.3 LAS ALTERACIONES DEL COLOR DE LOS ERITROCITOS

Reflejan las anomalías en el contenido hemoglobínico de los glóbulos rojos o eritrocitos, esta alteración es llamada Anisocromia, son las siguientes:

- Hipocromica
- Policromatofilia

Las alteraciones morfológicas de los glóbulos rojos o eritrocitos se muestran a continuación en la Tabla 2.1.

| Tipo de Alteración Morfológica | Nombre | Características | Imagen de una muestra digital |
|--------------------------------|------------|--|---|
| Anisocitosis | Macroцитos | <p>Descripción: El macrocito es un eritrocito morfológicamente normal con un diámetro mayor a 9,0 μm</p> <p>Sinónimos: megalocito.</p> |  |
| | Microцитos | <p>Descripción: El microcito es un eritrocito morfológicamente normal pero anormalmente pequeño, con un diámetro inferior a 7 μm.</p> |  |

| | | | |
|-----------------|----------------------------|--|---|
| Poiquilocitosis | Esferocitos | <p>Descripción: Células rojas de forma esférica que no presentan el halo de luz central característico de los eritrocitos normales.</p> <p>Sinónimos: microesferocito, microesferocito.</p> |  |
| | Elíptocitos | <p>Descripción: Células de forma ovalada o elíptica.</p> <p>Sinónimos: eliptocito, célula en lápiz, célula en cigarro.</p> |  |
| | Estomatocitos | <p>Descripción: Células que adoptan una configuración de boca de pez.</p> <p>Sinónimos: célula en boca de pescado.</p> |  |
| | Dianocitos | <p>Descripción: Representan eritrocitos con forma de diana o tienen forma de lágrima.</p> <p>Sinónimos: célula diana, célula en forma de tiro al blanco (<i>target cell</i>), célula en sombrero mejicano, codocito.</p> |  |
| | Drepanocitos o Falciformes | <p>Descripción: Células en forma de hoz o han tomado forma de media luna.</p> <p>Sinónimos: célula falciforme, célula en hoz, célula en media luna, meniscocito.</p> |  |

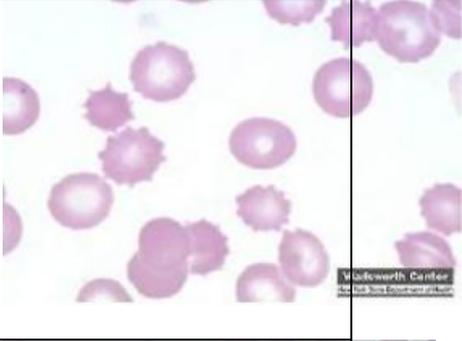
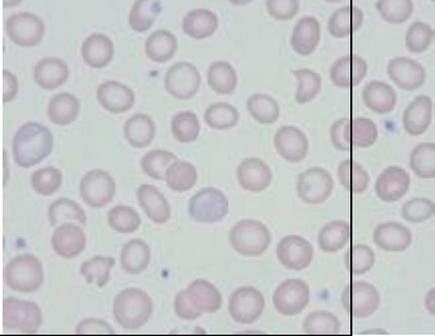
| | | | |
|-------------|------------------|---|--|
| | Equinocitos | <p>Descripción: Célula dentada, semejante a un erizo de mar conocido como crenocito.</p> <p>Sinónimos: crenocito, célula crenada, <i>burr cell</i>.</p> |  |
| Anisocromia | Hipocromia | <p>Descripción: Los eritrocitos se tornan casi invisibles debido a su poca tinción en la hemoglobina.</p> |  |
| | Policromatofilia | <p>Descripción: Los hematíes con tonalidad gris azulada, esto es signo de inmadurez celular.</p> |  |

Tabla 2.1: Alteraciones Morfológicas de los Glóbulos Rojos

Fuente: [Atlas de Hematología, 2005]

2.1.2.5 FROTIS SANGUINEO O EXTENDIDO DE SANGRE PERIFERICA

El Frotis sanguíneo o extendido de sangre periférica son laminas de muestras de sangre sobre un portaobjetos, tratadas con tintes o colorantes hematológicos y a partir del cual se observan las características morfológicas de las células sanguíneas (leucocitos eritrocitos y plaquetas) bajo el microscopio. Este estudio ayuda a diagnosticar muchas enfermedades de la sangre, como la leucemia, la malaria o la anemia drepanocítica.

2.1.2.5.1 TINCION DEL FROTIS SANGUINEO

La tinción del frotis sanguíneo consiste en teñir las muestras de sangre con un colorante hematológico. La tinción del frotis sanguíneo con colorante de

tipo Romanowsky, es la más frecuentemente empleada en el diagnóstico hematológico.

Una tinción o coloración de tipo Romanowsky, está constituido fundamentalmente por la mezcla de eosina (colorante ácido) y azul de metileno (colorante básico), además de sus derivados por oxidación del azul de metileno que se conoce con el nombre de azures (A, B, C), así como de tiacinas, eosina Y o eosina B. La acción combinada de estos colorantes produce el efecto Romanowsky que da una coloración púrpura a los núcleos de los leucocitos y a los gránulos neutrofílicos y da color rosado a los eritrocitos.

La Tinción del frotis en muestras sanguíneas se ve en la figura 2.2.

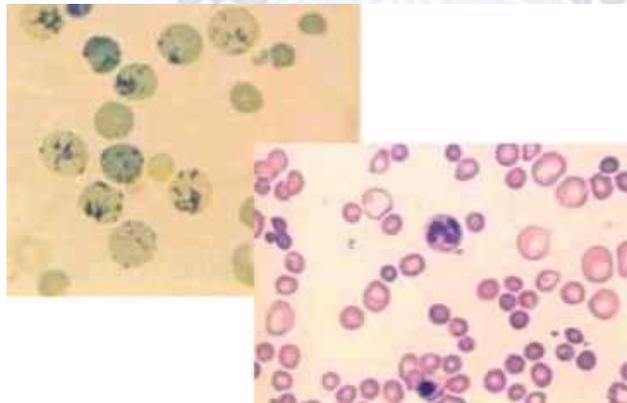


Figura 2.2: Imágenes con diferentes tipos de tinción

Fuente: [Atlas de Hematología, 2005]

2.1.2.5.1.1 TIPOS DE TINCION DEL FROTIS SANGUINEO

a. TINCIÓN DE WRIGHT

La tinción de Wright es una tinción de tipo Romanowsky usada en hematología para facilitar la diferenciación de los tipos de células de la sangre, conocida como policromática debido a que produce varios colores en las estructuras celulares. La tinción de Wright es una solución de alcohol metílico y eosina junto con el azul de metileno y azur B. Esta tinción provoca que los eritrocitos se observen de un color rojo, naranja o gris rosáceo debido a su contenido en

hemoglobina y en su zona central pálida debido a su forma de disco bicóncavo.

b. TINCIÓN GIEMSA

La tinción de Giemsa es un método habitual para el examen de frotis sanguíneos, cortes histológicos y otro tipo de muestras biológicas. Este colorante está constituido por una mezcla de azul de metileno, eosina y varios azules A, B en dilución acuosa. Los efectos de esta tinción, es que los eritrocitos se tiñen de color rosa asalmonado o color violeta-azulado intenso.

c. TINCIÓN DE MAY-GRÜNWARD

La Tinción de May-Grünwald es una modificación de la coloración de Wright. Esta tinción está constituido por el eosinato de azul de metileno de Jenner disuelto en alcohol metílico, destaca bien los gránulos de leucocitos y es por tanto especialmente útil en su recuento diferencial. Los eritrocitos se tiñen mal, adquiriendo coloración rojo pálido o violeta azulado.

d. TINCIÓN DE PANÓPTICA

Es el resultado de combinar la tinción de Giemsa con la de May-Grünwald, con el fin último de combinar las ventajas de dichos colorantes, de ahí que, esta tinción resalta de manera especial las granulaciones leucocitarias y mejora la coloración de los hematíes. Los efectos que tiene los eritrocitos es que se tiñen de color rosado-narajado o violeta-azulado.

2.2 PROCESAMIENTO DIGITAL DE IMÁGENES

El Procesamiento de imágenes digitales es el estudio de teorías, modelos y algoritmos que permite la manipulación de las imágenes en un ordenador con el fin de obtener información objetiva del mundo real, de manera automática a partir de una imagen de la escena captada a través de un

dispositivo electrónico o digital. De acuerdo a [González & Woods, 2002] y [Young et. al., 1998] existen tres tipos de procesamiento computarizado: bajo, medio y alto nivel de procesamiento.

Con los avances en las computadoras en cuanto a velocidad, procesamiento y almacenamiento, ha sido posible realizar grandes progresos en el procesamiento de imágenes, de esta forma, de una imagen se puede obtener: reconocimiento y la extracción de objetos individuales, la determinación del color y la clasificación de dichos objetos, entre otras características [Alcañiz M. ,1999].

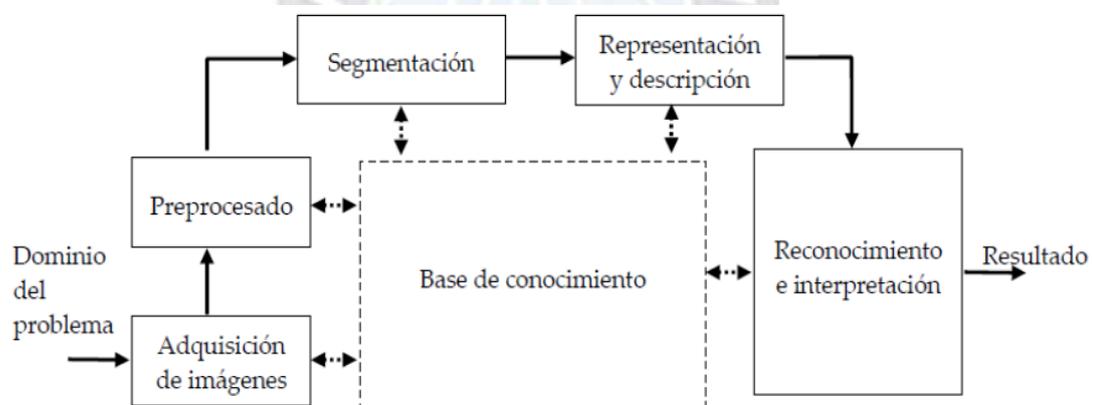


Figura 2.3: Etapas del procesamiento digital de imágenes

Fuente: [Alcañiz M. ,1999].

2.2.1 REPRESENTACION DE UNA IMAGEN DIGITAL

Se define una imagen como una función bidimensional de intensidad de luz $f(x, y)$, donde x e y representan coordenadas espaciales. La amplitud de f en un punto cualquiera (x, y) , es proporcional al brillo (nivel de gris) de la imagen en ese punto. Cuando x , y y la amplitud de valores de f son finitos y además las cantidades discretas, entonces a la imagen se le llama imagen digital. Una imagen digital está compuesta por un número finito de

elementos, cada elemento tiene un valor y una localidad específica, lo que se conoce como pixel [Añil, 1989] [Lindley, 1992].

La imagen digital puede considerarse como una matriz cuya fila y columna indican un punto en la imagen, el valor correspondiente a ese elemento (pixel) indica el nivel de gris en ese punto. En la Figura se indican los ejes de la representación digital de una imagen.

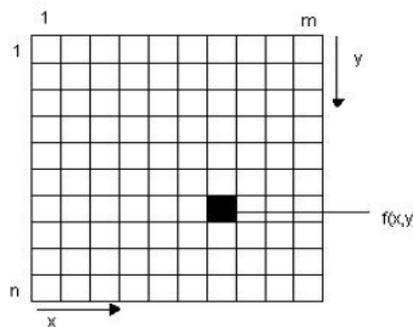


Figura 2.4: Imagen digital en un arreglo de $m \times n$ píxeles.

Fuente: [Añil, 1989].

2.2.1.1 IMAGEN EN NIVELES DE GRIS

Una imagen bidimensional es una región en el plano en el cual un brillo en cada punto es definido como una función $f(x, y)$ donde x e y son las coordenadas de dicho punto en el plano. Los elementos que constituyen una imagen digital se les conoce como píxeles, si cada uno de los píxeles de una imagen tiene asociado un nivel de brillo para su representación, entonces se le llama imagen en niveles de gris.

2.2.1.2 IMÁGENES BINARIAS

Una imagen binaria tiene solo dos niveles de gris (0,255), el nivel más claro se representa por el valor de 255 (blanco) y el más oscuro por el negro (0).

2.2.2 REPRESENTACIÓN DEL COLOR

Un espacio de color es un método por el que podemos especificar, crear y visualizar color. Un color es una o varias coordenadas o atributos. Estas coordenadas no nos dicen cual es el color, sino que nos señalan cual es su posición dentro de un espacio de color [Wik02].

2.2.2.1 ESPACIO DE COLOR RGB

RGB es conocido como un espacio de color aditivo ya que, todos los colores dentro de este espacio son entendidos como la suma de las componentes rojo(R), verde (G) y azul (B). Estos tres colores fueron elegidos porque corresponden aproximadamente a los tres tipos de conos sensitivos al color en el ojo humano [Wik02].

Este modelo es frecuentemente visualizado a través de un cubo unitario. Cada componente de color es asignada a uno de los tres ejes ortogonales de coordenadas del espacio tridimensional [Bou95].

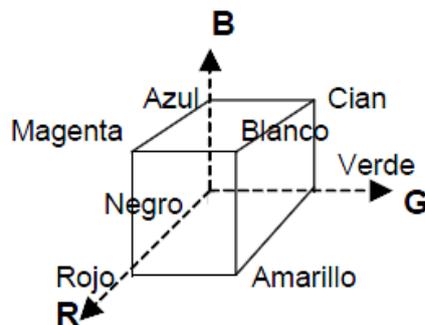


Figura 2.5: Representación de los componentes de color RGB.
Fuente: [Añil. ,1989].

2.2.3 SEGMENTACION DE LA IMAGEN

La Segmentación consiste en subdividir una imagen en sus respectivas regiones u objetos, sin que ningún pixel pertenezca a mas de una región u objeto. Esta técnica de descomposición de imágenes en partes con

significado, es decir consiste en la identificación de las regiones u objetos que se distinguen por los cambios de contraste en la escala de grises o las texturas con lo que facilitan la identificación de regiones con áreas homogéneas. [Muoz, 2007]

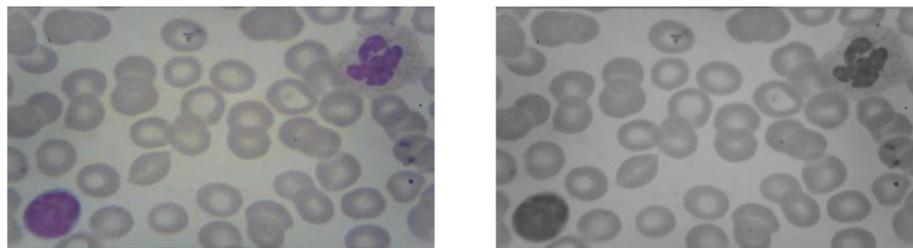
2.2.3.1 BINARIZACION

El método de umbralización o binarización consiste en poder los pixeles que sobrepasen en umbral u dado, en el valor más alto dentro del rango de niveles, generalmente blanco, y los que sean iguales o estén por debajo de dicho umbral en el valor más pequeño en el mismo rango de niveles, generalmente negro.

Sea X una matriz que contiene todos los niveles de la imagen donde $X_{i,j}$ representa el nivel de pixel en la columna i y la fila j; el nivel, generalmente, se encuentra en el rango (0,255). La imagen binarizada es una matriz X^* cuya ij-esima componente está dada por la ecuación 2.1:

$$X_{i,j}^1 = \begin{cases} 255; & \text{si } X_{i,j} > u \\ 0; & \text{si } X_{i,j} \leq u \end{cases} \quad (2.1)$$

El valor del umbral es usualmente seleccionado en forma experimental de acuerdo a la aplicación.



a) Imagen de glóbulos rojos b) Imagen de glóbulos rojos binarizada

Figura 2.6: Imagen de extendido sanguíneo de glóbulos rojos a) y b).
Fuente: [Martínez, 2000].

2.2.4 EXTRACCION DE CARACTERISTICAS

Para reconocer un objeto es necesario tener una *descripción* de los mismos (descriptor o modelo del objeto). Los descriptores deben de ser

independientes del tamaño, localización u orientación del objeto, y deben ser suficientes para discriminar objetos entre sí. Los descriptores se basan en la evaluación de alguna *característica* del objeto [NA02].

2.2.4.1 Descriptores Geométricos

Son descriptores de contornos y regiones de los objetos, son aquellos que cuantifican las formas geométricas de los objetos.

- **Área**

En una imagen digital, el área de un objeto está dada por el número de píxeles en blanco que representan al mismo.

$$A = \sum_{x=0} \sum_{y=0} f(x, y) \quad (2.2)$$

- **Perímetro**

Es un parámetro geométrico que al igual que el área, puede ser calculado a partir del número de píxeles en blanco de la imagen binaria que contiene el borde del objeto.

$$P = \sum_{x=0} \sum_{y=0} f(x, y); \forall f(x, y) = 0 \quad (2.3)$$

- **Diametro Eje Mayor**

Longitud del eje mayor del objeto, viene definido por la distancia entre los píxeles del contorno del objeto en el eje x.

$$K = \sqrt{\frac{2(\mu_{20} + \mu_{02} - \Delta)}{\mu_{11}}} \quad (2.4)$$

- **Diametro Eje Menor**

Longitud del eje mayor del objeto, viene definido por la distancia entre los pixeles del contorno del objeto en el eje x.

$$L = \sqrt{\frac{2(\mu_{20} + \mu_{02} + \Delta)}{\mu_{11}}} \quad (2.5)$$

- **Excentricidad**

Es el cociente entre la longitud o diametro del eje mayor del objeto y su diametro del eje menor.

$$e = 2\left(\frac{\sqrt{\left(\frac{L}{2}\right)^2 - \left(\frac{K}{2}\right)^2}}{L}\right) \quad (2.6)$$

- **Circularidad**

Es un parametro que no depende del tamaño de la region y vienen dado por los coeficientes entre el area y el perimetro.

$$Circularidad = \frac{4\pi Area}{perimetro^2} \quad (2.7)$$

- **Elongacion**

Razon entre la longitud o diametro del eje mayor del objeto y su diametro del eje menor del contorno del objeto.

$$Elipticidad = \frac{Diametroe\grave{M}ayor}{Diametroe\grave{M}enor} \quad (2.8)$$

2.2.5 DESCRIPTORES ELEMENTALES DE FORMA

Los descriptores elementales de forma son parámetros que se basan en la circularidad de los objetos es decir en las características geométricas, los valores numéricos de los parámetros tienen un rango entre [0,1], los cuales son 1 si el objeto es circular y 0 si el objeto es una línea.

2.2.5.1 Coeficiente de Circularidad

El Coeficiente de circularidad es un parámetro geométrico, empleado en la comparación de objetos, desde diferentes distancias y no depende del tamaño del objeto. El Coeficiente de circularidad o compacidad sirve para identificar figuras geométricas, que dependen de la relación entre el área y el perímetro, la ecuación de este parámetro es el mismo de la ecuación (2.7).

2.2.5.1 Coeficiente de Elíptico

El Coeficiente Elíptico o elongación se determina por medio de la razón entre la longitud mayor y la longitud menor de la región u objeto, con este valor numérico se puede determinar cuan redondo es un objeto, ya que si la longitud mayor y la longitud menor de una región circular son iguales, entonces el valor de la elongación se aproxima a 1, la ecuación de este parámetro es el mismo de la ecuación (2.8).

2.3. RECONOCIMIENTO DE PATRONES

El Reconocimiento de patrones tiene una gran importancia en el área de Inteligencia Artificial debido a sus grandes aplicaciones implementadas en aéreas como la industria, medicina, seguridad, entre muchas otras, para resolver problemas como en el reconocimiento de voz, clasificación de documentos, reconocimiento de imágenes, recuperación de datos, etc.

A continuación se muestra el esquema en la figura de un sistema de reconocimiento de patrones.

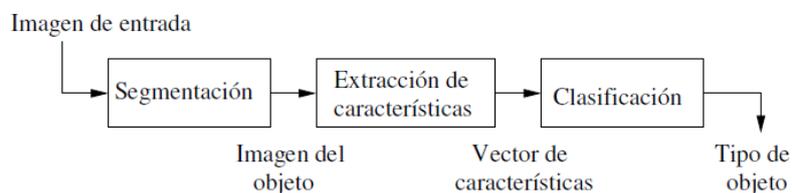


Figura 2.7: Etapas del reconocimiento de patrones

Fuente: [Alcañiz M. ,1999].

2.3.1. ALGORITMOS DE CLASIFICACION

En la clasificación de datos se desarrolla una descripción o modelo para cada una de las clases presentes, un clasificador no reconoce objetos, reconoce al patrón que define al objeto. Existen muchos métodos de clasificación como aquellos basados en arboles de decisión TDIDT como el ID3 y el c4.5, los métodos estadísticos, las redes neuronales, los conjunto finitos, entre otros. [Servente, 2002]

A continuación se describen brevemente aquellos Algoritmos de clasificación:

- **Algoritmos estadísticos:** Los algoritmos estadísticos han sido utilizados por los analistas para detectar patrones inusuales en los datos y explicar dichos patrones mediante la utilización de modelos estadísticos, como por ejemplo los modelos lineales.
- **Redes Neuronales:** Las redes neuronales imitan la capacidad de la mente humana para encontrar patrones. Han sido aplicadas en aplicaciones que trabajan sobre la clasificación de los datos.
- **Algoritmos genéticos:** Técnicas de optimización que utilizan procesos como el entrecruzamiento genético, la mutación y la selección natural en un diseño basado en los conceptos de la evolución natural.
- **Método de K vecino más cercano:** es una técnica que clasifica cada registro de un conjunto de datos en base a la combinación de las clases de los k registros más similares.
- **Regla de Inducción:** La extracción de reglas si-entonces a partir de datos de importancia estadística.
- **Clasificador basado en instancias o ejemplos:** Una manera de clasificar un caso es a partir de un caso similar cuya clase es conocida, y predecir que el caso pertenecerá a esa misma clase.

2.3.2 REDES NEURONALES

Las Redes Neuronales Artificiales (RNA) son sistemas implementados de manera tal que funcionen como las neuronas biológicas de los seres vivos. El cerebro humano continuamente recibe señales de entrada de muchas

fuentes y las procesa a manera de crear una apropiada respuesta de salida. Nuestros cerebros cuentan con millones de neuronas que se interconectan para elaborar “Redes Neuronales”. Las RNAs fueron originalmente una simulación abstracta de los sistemas nerviosos biológicos, formados por un conjunto de unidades llamadas “neuronas” o “nodos” conectadas unas con otras. Estas conexiones tienen una gran semejanza con las dendritas y los axones en los sistemas nerviosos biológicos.

Las Redes Neuronales Artificiales (RNA o ANN), son una aproximación de la inteligencia artificial fundamentada en el cerebro de los mamífero, mas explícitamente en las neuronas biológicas.

2.3.2.1 RED NEURONAL BIOLÓGICA

Una neurona (o célula nerviosa) es una célula biológica especial que procesa información (Figura 3.1). Está compuesta de un cuerpo celular, o *soma*, y dos tipos de ramas que llegan al exterior: el *axón* y las *dendritas*. El cuerpo celular tiene un núcleo que contiene información acerca de rasgos hereditarios y un plasma que sostiene el equipo molecular para producir material necesario para la neurona. Una neurona recibe señales (impulsos) de otras neuronas a través de sus dendritas (receptores) y transmite señales generadas por su cuerpo celular a lo largo del axón (transmisor), el cual eventualmente se divide en ramales y subramales. En los extremos de estos ramales están las *sinapsis*. Una sinapsis es una estructura elemental y una unidad funcional entre dos neuronas (un ramal axón de una neurona y una dendrita de otra). Cuando el impulso alcanza el extremo de la sinapsis, ciertos químicos llamados neurotransmisores son liberados. Los neurotransmisores se difunden a través del espacio sináptico, para aumentar o inhibir, dependiendo del tipo de la sinapsis, la tendencia de la neurona receptora a emitir impulsos eléctricos. Esta dependencia de la historia actúa como una memoria, la cual es posiblemente responsable de la memoria humana [Dermuth & Beale, 1996].

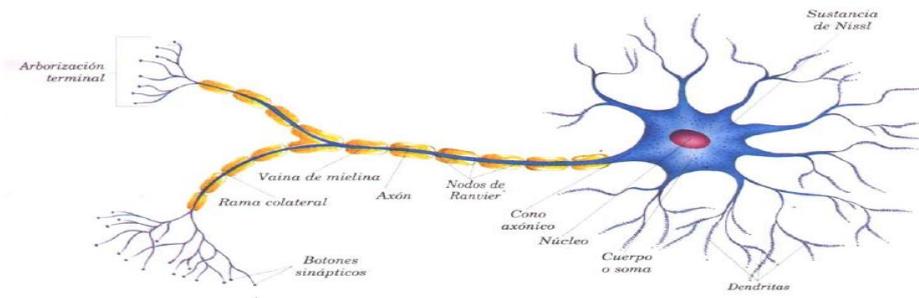


Figura 2.8: Neurona Biologica

Fuente: [Shapiro, 2000].

La corteza cerebral contiene aproximadamente 10^{15} millones de neuronas interconectadas. Las neuronas se comunican a través de un corto tren de pulsos, normalmente de milisegundos de duración, a una velocidad de conmutación mucho más lenta que en circuitos electrónicos. Las neuronas artificiales usadas para construir RNAs son mucho menos complejas que las encontradas en el cerebro.

2.3.2.2 RED NEURONAL ARTIFICIAL

La Neurona Artificial es la base estructural de las Redes Neuronales Artificiales. Una neurona artificial es una Unidad elemental de proceso que realiza cálculos a partir de un vector de entrada procedente del exterior o de otras neuronas.

En la figura 2.11 puede ver la representación de una Red Neuronal Artificial.

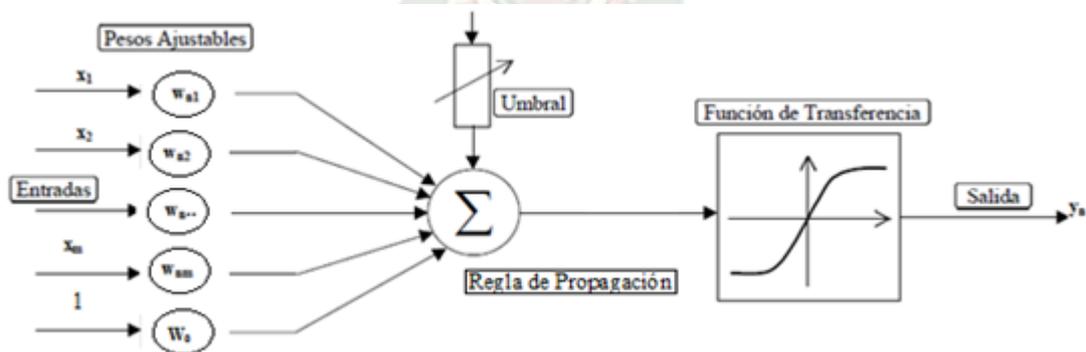


Figura 2.9: Neurona Artificial y sus elementos

Fuente: Modificado [Shapiro, 2000].

2.3.2.2.1 ELEMENTOS DE UNA RED NEURONAL ARTIFICIAL

Una neurona artificial es un procesador elemental, en el sentido de que procesa un vector de entrada $\mathbf{x}=\{ x_1,x_2,\dots,x_m \}$ y produce una respuesta o salida única Y_n . Los elementos clave de una neurona artificial los podemos ver en la figura anterior y son los siguientes:

- Las **entradas** x que reciben los datos de otras neuronas a las que está conectada.
- **Salida:** Es el resultado en función del estado de activación, la ecuación de la salida es el siguiente:

$$Y_n = F(Un + bn)$$

Donde F es la función de activación

U_n es la combinación lineal de las entradas por los pesos sinápticos

b_n es el umbra, toma valor entre 0 y 1.

- Los **pesos sinápticos** w_{nj} . En una neurona artificial a las entradas que vienen de otras neuronas se les asigna un peso. El parámetro w_0 se suele conocer como *bias* o sesgo, y puede ser considerado como otro peso de una entrada extra fijada permanentemente a +1.
- **Regla de propagación.** Con esas entradas y los pesos sinápticos, se realiza una suma ponderada. A la combinación de entradas se le denomina Net.

$$Un = \sum x_i * w_{nj}$$

donde: x_i *datos de entrada*

w_{nj} *pesos sinápticos*

- **Función de activación o transferencia.** El valor obtenido con la regla de propagación, se filtra a través de una función conocida como función de activación y es la que nos da la salida de la neurona, toma valores en el intervalo $[0,1]$ o $[-1,1]$. En la Figura 2.10 se muestran las funciones de activación más usuales.

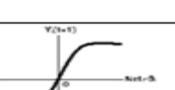
| FUNCION | DESCRIPCION | GRAFICA |
|--------------|--|---|
| LINEAL | $Y_i(t+1) = net_i - \theta_i$ |  |
| ESCALON | $Y_i(t+1) = \begin{cases} +1 & [net_i > \theta] \\ y(t) \sin net_i = \theta \\ -1 & [net_i < \theta] \end{cases}$ |  |
| LINEAL MIXTA | $Y_i(t+1) = \begin{cases} b & [net_i \leq b + \theta] \\ net_i = \theta_i b + \theta_i & [net_i < b + \theta] \\ b & [net_i \geq b] \end{cases}$ |  |
| HIPERBOLICA | $Y_i(t+1) = \frac{(1 - e^{-(net_i - \theta)})}{(1 + e^{-(net_i - \theta)})}$ |  |
| SIGMOIDAL | $Y_i(t+1) = \frac{1}{(1 + e^{-(net_i - \theta)})}$ |  |

Figura 2.10: Funciones de Activacion

Fuente: [Shapiro, 2000].

2.3.2.3 Arquitectura de las Redes Neuronales Artificiales

Se denomina arquitectura de una red neuronal a la estructura que conecta las neuronas entre sí para formar una red neuronal. Las arquitecturas de la red neuronal son de diversos tipos, de acuerdo a los requerimientos para la solución de determinados problemas [Del Pezzo, 2003].

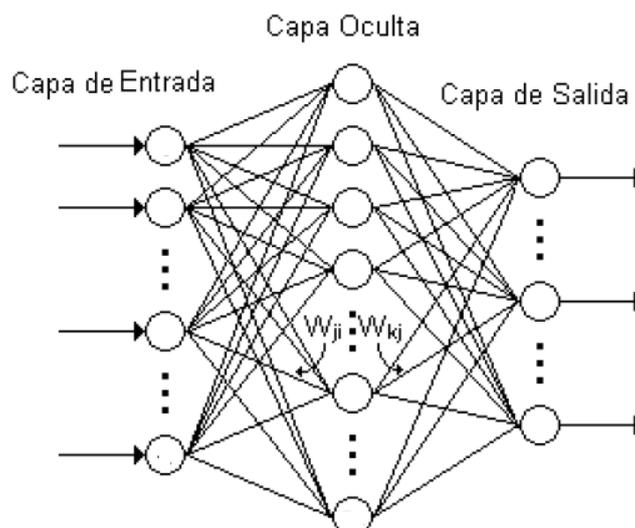


Figura 2.11: Arquitectura de las redes neuronales artificiales

Fuente: Modificado [Shapiro, 2000].

La Arquitectura de las redes Neuronales artificiales se muestra en la figura, la distribución de las neuronas en una red se realiza formando niveles o capas de neuronas de la siguiente manera:

- *Capa de entrada*: Contiene n-neuronas que representan n-valores o datos de entrada de la red que se reciben provenientes del exterior.
- *Capas ocultas*: Las capas ocultas y el número de neuronas dependen del tipo de red que se emplee. Este tipo de capas realiza gran parte del procesamiento.
- *Capa de salida*: Contiene n-neuronas que representan n-valores o datos de salida de la red. Esta capa proporciona la respuesta de la red neuronal. Normalmente también realiza parte del procesamiento.

2.3.2.4 APRENDIZAJE DE UNA RED NEURONAL ARTIFICIAL

Las redes neuronales artificiales tienen la capacidad de “aprender de su ambiente” a través de un proceso iterativo de ajustes de sus pesos. Este proceso recibe el nombre de *regla de aprendizaje* o simplemente aprendizaje. Para realizar el aprendizaje de una RNA, primero se alimenta a la red con los datos de entrenamiento. Cada observación de entrenamiento se propaga a través de la red para producir una respuesta en la capa de salida, esta respuesta se compara con la salida real. Luego se obtiene una diferencia que se usa para calcular una función de pérdida del error de la respuesta de la red con un algoritmo apropiado permitiendo modificar los valores de los pesos con el fin de reducir el error. Estas correcciones deben realizarse varias veces, para todo el conjunto de observaciones de entrenamiento.

2.3.2.4.1 TIPOS DE APRENDIZAJE

Existen dos tipos de aprendizaje en las RNA:

1. **Aprendizaje no Supervisado**: Este tipo de aprendizaje solo se tienen los datos no existe un agente externo, la red modifica los pesos a

partir de información interna, determinando características como rasgos significativos, regularidades o redundancias cuando .En este caso la red debe de auto enseñarse.

2. **Aprendizaje Supervisado:** Este tipo de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida de la red y en el caso de que ésta no coincida con la salida correcta se procederá a modificar los pesos de las conexiones, con el fin de conseguir que la salida obtenida se aproxime a la correcta.
3. **Aprendizaje Supervisado Por refuerzo.** En este aprendizaje se tienen los datos y los resultados deseados, los pesos se modifican hasta que para todos los ejemplos de entrenamiento, la salida sea lo más parecida a la deseada.

2.3.2.5 TIPOS Y MODELOS DE REDES NEURONALES ARTIFICIALES

Existen diferentes tipos y modelos de redes neuronales, las cuales se pueden clasificar de acuerdo a sus características propias de las redes, según [Charytoniuk, 1995]:

- a) **Según su arquitectura**, se clasifican en:
 - **Redes Neuronales Monocapas.** Se corresponde con la red neuronal más sencilla ya que tiene una capa de neuronas que proyectan las entradas a una capa de neuronas de salida donde se realizan diferentes cálculos. El modelo de red conocido de este tipo es la red Perceptron simple, Adaline y Madaline
 - **Redes Neuronales Multicapas.** Es una generalización de la anterior red existiendo un conjunto de capas intermedias entre la entrada y la salida. Los modelos de red Perceptron Multicapa, Red de Base radial, Hopfield.
- b) **Según el tipo de conexiones**, se clasifican en :
 - **Redes neuronales recurrentes.** Estas redes se caracterizan por la existencia de lazos retroalimentación, los modelos de redes son:

Redes competitivas, Hopfields, Teoría de resonancia adaptativa (ART), Elman, Máquina de Boltzmann.

- **Redes neuronales no recurrentes.** Estas se redes se subdividen en:
 - Redes con conexiones hacia adelante (feedforward). En este tipo de red las señales se propagan hacia adelante (de izquierda a derecha), los modelos de red de este tipo son: Perceptrón Simple, Adaline, Madalena, Perceptrón multicapa, Redes de base radial.
 - Redes con conexiones hacia adelante y hacia atrás (feedforward/feedback), las capas de esta red están conectadas hacia adelante/hacia atrás, la información circula durante el funcionamiento de la red en ambas direcciones. Los modelos de red son: Teoría de resonancia adaptativa (ART), Mapas auto-organizativos de Kohonen (SOM).
- c) **Según su aprendizaje**, se clasifican en :
 - Redes neuronales con Aprendizaje supervisado, este tipo de redes requiere de influencia externa (entrenado) para ajustar los pesos de las conexiones entre neuronas y así obtener las salidas deseadas.
 - Redes neuronales con Aprendizaje no supervisado, no se necesita de influencia externa, este tipo de red trabaja con datos internos para obtener las salidas.
- d) **Según su aplicación**, se clasifican en :
 - **Redes neuronales de predicción.** Las redes de predicción tienen por objetivo modelar una función no lineal para la predicción de la aplicación del problema
 - **Redes neuronales de clasificación**, son redes que clasifican a base de las entradas a una clase en específico.

2.3.3 REDES NURONALES DE FUNCION DE BASE RADIAL

Las **Redes Neuronales de función de base radial** o **RBF** (**Radial Basis Functions por sus siglas en ingles**) propuestas por M.J.D. Powell, D.S Broomhead y D. Lowe a mediados de los años 80, es una red multicapa de tipo FeedForward y aprendizaje híbrido, que implementa una capa oculta un algoritmo de aprendizaje no supervisado y en su capa de salida un aprendizaje supervisado; esta red busca la aproximación global de la función implementando aproximaciones locales mediante funciones de base radial [Caicedo B, Eduardo;Lopez S,Jesus,2009].

2.3.3.1 Arquitectura de las redes RBF

La arquitectura genérica de una red **RBF** consta de tres capas de neuronas con conexiones de tipo *feedforward* como puede ver en la Figura 2.12.

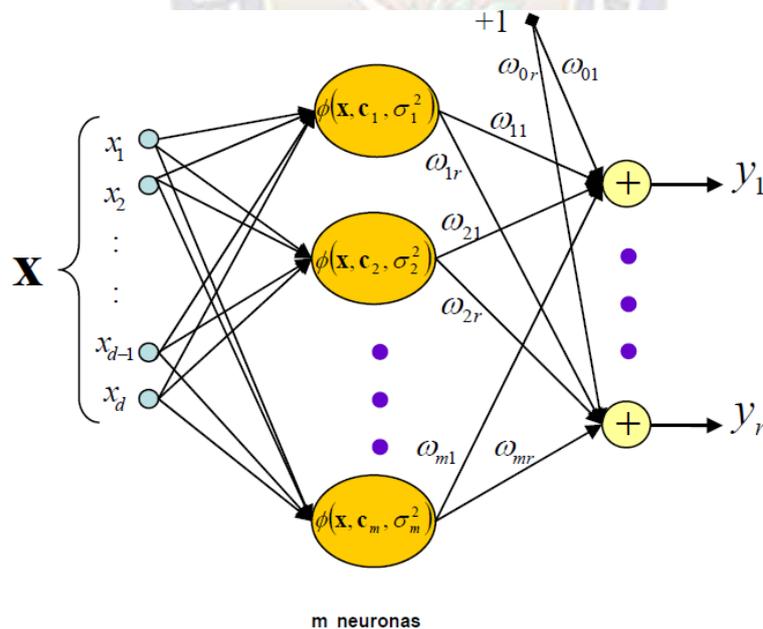


Figura 2.12: Arquitectura de las redes neuronales RBF

Fuente: [Caicedo B, Eduardo;Lopez S,Jesus,2009].

Capa Entrada. La primera capa está formada por las neuronas de entrada, representando cada una de estas neuronas cada una de las características

que componen el vector o patrón de entrada $x=\{x_1,x_2,\dots,x_d\}$, es decir, la neurona i -ésima de la capa de entrada está asociada a la característica i -ésima x_i del patrón de entrada x . Esta primera capa no realiza ningún procesado de los datos de entrada, pasándolos tal cual a la siguiente capa; es por esta razón por la que algunos autores consideran a estas redes como de dos capas, sólo considerando aquellas capas que realizan algún tipo de procesado de los datos. La capa de entrada simplemente realiza una función de simple transmisión de información.

Capa Oculta. La siguiente es la capa oculta, que al contrario que en otros tipos de redes, es única. Cada una de las neuronas de la capa oculta está asociada a un vector sináptico llamado comúnmente *centroide*. Al contrario de lo que suele ser habitual en otras redes, en vez de calcular la suma ponderada de las entradas (producto escalar del vector de entrada y de pesos sinápticos) y aplicarla una cierta función de activación, calcula la distancia euclídea entre el vector de las entradas $x=\{x_1,x_2,\dots,x_d\}$ y el centroide de cada una de las neuronas ocultas, y a esa distancia se le aplica una *función de base radial* que realiza una transformación no lineal, y cuya salida pasara a la siguiente capa. La función más utilizada es la gaussiana, aunque existen otras funciones que se pueden utilizar.

Capa Salida. Finalmente, las neuronas de la capa de salida, realizan una combinación lineal de las funciones de base radial, es decir, las salidas proporcionadas por las neuronas de la capa oculta son multiplicadas por el peso de su conexión y sumadas todas ellas, para a continuación el resultado aplicarlo a una función de activación como por ejemplo la lineal (generalmente) o la sigmoide.

2.3.3.2 Modelo genérico de una red RBF

Para una **red neuronal de base radial** con d neuronas en la capa de entrada, m neuronas en la capa oculta, neuronas en la capa de salida y salida lineal, la formula que describe la salida de la neurona k -ésima de la capa oculta,, viene dada por la siguiente expresión:

$$Y_k(x) = \sum_{j=1}^m \omega_{kj} \bullet \phi_j(x) + \omega_{k0} \quad k = 1 \dots r \quad (2.9)$$

donde ω_{kj} representa el **peso** asociado a la conexión que une la neurona j -ésima de la capa oculta con la neurona k -ésima de la capa de salida, ω_{k0} es el **umbral** asociado a la neurona de salida k -ésima, x es la entrada de la red y $(\phi_j)_{j=1 \dots m}$ son las **funciones de base radial o de activación** de cada uno de las neuronas de la capa oculta.

Las **funciones de base radial** se expresan normalmente como traslaciones de una función prototipo de la siguiente forma:

$$\phi_j(x) = \phi\left(\frac{\|x - c_j\|}{d_j}\right) \quad (2.10)$$

donde c_j es el centro asociado a la neurona j -ésima de la capa oculta, d_j es la desviación, anchura o factor de escala para la distancia euclídea entre el vector de entrada x y el centroide $\|x - c_j\|$.

Como función prototipo se suele emplear una gaussiana de la forma:

$$\phi(r) = e^{-\frac{r^2}{2\sigma^2}} \quad (2.11)$$

donde r es la distancia entre el centroide y el vector de entrada (ver Figura 2.14)

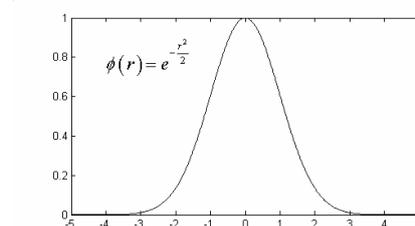


Figura 2.13: Funcion gaussiana de una redes neuronales RBF

Fuente: [Caicedo B, Eduardo;Lopez S,Jesus,2009].

La mayoría de estas funciones tienen similar dependencia radial, caracterizándose porque alcanzan un máximo en torno al origen ($r=0$) y decrecen hacia cero cuando. Por tanto, estas funciones se caracterizan

porque la respuesta es *localizada*. Se dice que además son *aproximadores universales* ya que pueden aproximar arbitrariamente bien cualquier función continua y acotada [Poggio & Girosi, 1990].

También se han usado funciones de base radial con una dependencia distinta, como por ejemplo la multicuadrática, cuya expresión se encuentra en la tabla anterior, pero en el que la raíz está en el numerador. En este caso la respuesta no es localizada.

El **parámetro de normalización** σ es un factor de escala que nos da la anchura de la función. Viene a ser como el radio de influencia de la neurona en el espacio de las entradas. En algunos casos se pueden escoger diferentes valores de σ para cada función de cada neurona, con lo que estas funciones $\phi_j(r)$ dejan de tener simetría radial y adquieren formas elipsoidales.

Como ya se ha comentado, la salida de la red es generalmente lineal aunque también existen otras posibilidades como la función de activación de forma sigmoidea (Figura 2.14)

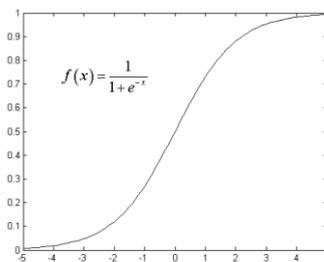


Figura 2.14: Funcion sigmoidea de una redes neuronales RBF

Fuente: [Caicedo B, Eduardo;Lopez S,Jesus,2009].

2.3.3.3 APRENDIZAJE DE LAS REDES NEURONALES DE BASE RADIAL

El aprendizaje de las redes RBF se divide en dos métodos, un Aprendizaje Híbrido o Aprendizaje totalmente supervisado.

2.3.3.3.1 Aprendizaje Híbrido

El aprendizaje Híbrido se divide en dos etapas:

a) Aprendizaje No Supervisado

En el Aprendizaje no supervisado se determina los parámetros de la capa oculta, la determinación de centros se realiza mediante algoritmos k-medidas o mapas de Kohonen; mientras que, las desviaciones se deben calcular de manera que cada neurona de la capa oculta se active en la región del espacio de entradas de manera que el solapamiento de las zonas de activación sea lo más ligero posible, para el cálculo de estas desviaciones existe una serie de aproximaciones que se pueden considerar, uno de estos criterios viene dado por la ecuación:

$$d = \frac{D_{\max}}{\sqrt{2 * m_1}} \quad (2.12)$$

Donde D_{\max} es la máxima distancia entre los centros m_1 es el número de centros.

b) Aprendizaje Supervisado

En el aprendizaje supervisado, se determinan los pesos de la capa de salida, usando una técnica de minimización de la función de error.

2.3.3.3.2 Aprendizaje Totalmente Supervisado

Este tipo de aprendizaje no conserva las propiedades o características locales de las redes de base radial, en este caso todos los parámetros de la red se determinan de manera completamente supervisada con el objetivo de minimizar el error cuadrático medio. Este tipo de aprendizaje no tiene en cuenta que las desviaciones tomen valores tales que provoquen solapamiento de las activaciones de las neuronas de la capa oculta, para el cálculo de los parámetros se usa las siguientes ecuaciones:

- Pesos

$$w_{ik}(n) = w_{ik}(n-1) - \alpha_1 \frac{\delta e(n)}{\delta w_{ik}} \quad (2.13)$$

- Centros

$$c_{ij}(n) = c_{ij}(n-1) - \alpha_2 \frac{\delta e(n)}{\delta c_{ij}} \quad (2.14)$$

- Desviaciones

$$d_i(n) = d_i(n-1) - \alpha_3 \frac{\delta e(n)}{\delta d_i} \quad (2.15)$$

- Umbrales

$$u_k(n) = u_k(n-1) - \alpha_1 \frac{\delta e(n)}{\delta u_k} \quad (2.16)$$

2.4 METODOLOGÍA DE DESARROLLO DEL SOFTWARE

La Metodología de desarrollo del software es un conjunto de procedimientos, técnicas, herramientas y un soporte documental que ayuda a los desarrolladores a realizar nuevo software. Una Metodología puede seguir uno o varios modelos de ciclo de vida, es decir, el ciclo de vida indica que es lo que hay que obtener a lo largo del desarrollo del proyecto pero no como hacerlo [Pressman, 98].

2.4.1 METODOLOGÍA DE DESARROLLO EN CASCADA

El Modelo de Cascada [Bennington 1956, Modificado por Royce en 1970]. El modelo de cascada también conocido como **modelo lineal secuencial**, se basa en un enfoque sistemático y secuencial para el desarrollo del software que comienza en un nivel de sistemas y progresa con el análisis, diseño, implementación o codificación, pruebas y mantenimiento [Pressman, 98].

Este modelo en cascada es una Metodología de desarrollo del software, que permite el modelar el sistema en los siguientes etapas: Análisis de Requerimientos, diseño del prototipo, implementación o codificación, pruebas y mantenimiento. La siguiente figura ilustra el modelo en cascada para la ingeniería de software.

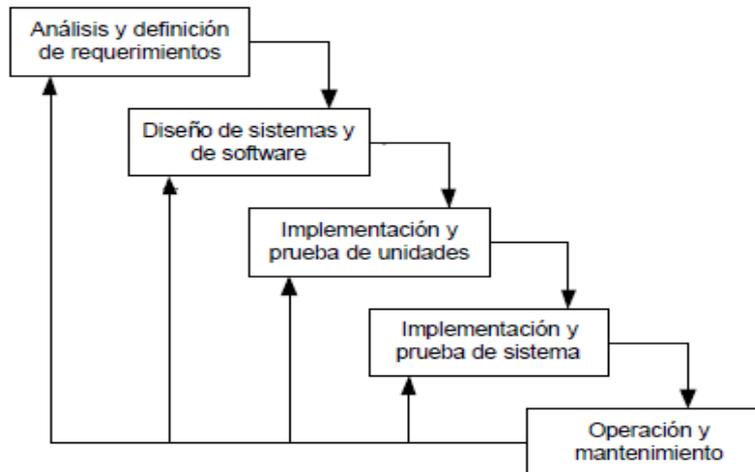


Figura 2.16: Ciclo de vida del modelo en cascada

Fuente: modificado [Pressman, 98].

2.4.2 DESCRIPCIÓN DE ACTIVIDADES

2.4.2.1 Análisis de los requisitos del software

El análisis de requisitos es fundamental que a través de una colección de requerimientos funcionales y no funcionales, el desarrollador o desarrolladores del software comprendan completamente la naturaleza de los programas que se deben construir para desarrollar la aplicación, la función requerida, comportamiento, rendimiento e interconexión [Pressman, 98].

2.4.2.2 Diseño del sistema

El diseño del software, en general, la actividad del diseño se refiere al establecimiento de las estructuras de datos, la arquitectura general del software, representaciones de interfaz y algoritmos. El proceso de diseño traduce los requisitos en un representación de software [Pressman, 98].

2.4.2.3 Implementación o Codificación

Es la fase de programación. Aquí se desarrolla el código fuente, el diseño debe traducirse en una forma legible para la maquina, haciendo uso de prototipos así como pruebas y ensayos para corregir errores. El paso de

codificación realiza esta tarea. Si el diseño se realiza de una manera detallada la codificación puede realizarse mecánicamente [Pressman, 98].

2.4.2.4 Prueba de sistema

Una vez que se ha generado el código comienza la prueba del software o sistema que se ha desarrollado. De acuerdo con Pressman, El proceso de pruebas se centra en la lógica interna del software, y en las funciones externas, realizando pruebas que aseguren que la entrada definida produce los resultados que realmente se requieren. Se comprueba que funciona correctamente antes de ser puesto en explotación [Pressman, 98].

2.4.2.5 Mantenimiento

El software indudablemente sufrirá cambios, y habrá que hacer algunas modificaciones a su funcionalidad, es de suma importancia que el software de calidad pueda adaptarse con fines de acoplarse a los cambios de su entorno externo [Pressman, 98].

2.5 LENGUAJE DE MODELADO UNIFICADO UML

El lenguaje de Modelado, UML propuesto por Booch, Rumbaugh y Jacobson “es un lenguaje de modelado visual, que permite especificar, visualizar, construir y documentar los componentes que forman parte de un sistema de software orientada a objetos” [Pressman, 98].

UML no es un método de desarrollo de software. Es un estándar para representar los diagramas del análisis al diseño y de este al código, UML al no ser un método de desarrollo es independiente del ciclo de desarrollo que se vaya seguir, puede encajar en métodos como el modelo en cascada, el modelo en espiral, incluso en los métodos ágiles de desarrollo de software.

UML ofrece diagramas con los cuales se pueden modelar sistemas:

- Diagrama de Casos de Uso para modelar los procesos
- Diagrama de Secuencia para modelar el paso de mensajes entre objetos.

- Diagrama de Colaboración para modelar interacciones entre objetos
- Diagrama de Estado para modelar el comportamiento de los objetos en el sistema.
- Diagrama de Clases para modelar la estructura estática de las clases en el sistema.
- Diagrama de objetos para modelar la estructura estática de los objetos en el sistema.
- Diagrama de Componentes para modelar los componentes.

2.6 PRUEBA DE HIPÓTESIS

Una hipótesis estadística es una suposición hecha con respecto a la función de distribución de una variable aleatoria. Para establecer la verdad o falsedad de una hipótesis estadística con certeza total, será necesario examinar toda la población. En la mayoría de las situaciones reales no es posible o práctico efectuar este examen, y el camino más aconsejable es tomar una muestra aleatoria de la población y en base a ella, decidir si la hipótesis es verdadera o falsa.

En la prueba de una hipótesis estadística, es costumbre declarar la hipótesis como verdadera si la probabilidad calculada excede el valor tabular llamado el nivel de significación y se declara falsa si la probabilidad calculada es menor que el valor tabular.

La prueba a realizar dependerá del tamaño de las muestras, de la homogeneidad de las varianzas y de la dependencia o no de las variables.

Si las muestras a probar involucran a más de 30 observaciones, se aplicará la prueba de Z, si las muestras a evaluar involucran un número de observaciones menor o igual que 30 se emplea la prueba de t de student. La fórmula de cálculo depende de si las varianzas son homogéneas o heterogéneas, si el número de observaciones es igual o diferente, o si son variables dependientes.

Para determinar la homogeneidad de las varianzas se toma la varianza mayor y se divide por la menor, este resultado es un estimado de la F de Fisher. Luego se busca en la tabla de F usando como numerador los grados de libertad $(n-1)$ de la varianza mayor y como denominador $(n-1)$ de la varianza menor para encontrar la F de Fisher tabular. Si la F estimada es menor que la F tabular se declara que las varianzas son homogéneas. Si por el contrario, se declaran las varianzas heterogéneas. Cuando son variables dependientes (el valor de una depende del valor de la otra), se emplea la técnica de pruebas pareadas.

Como en general estas pruebas se aplican a dos muestras, se denominarán a y b para referirse a ellas, así entenderemos por:

- n_a al número de elementos de la muestra a;
- n_b al número de elementos de la muestra b;
- \bar{x}_b al promedio de la muestra b;
- s^2_a la varianza de la muestra a;
- Y así sucesivamente.

Entonces se pueden distinguir 6 casos a saber:

1. Caso de muestras grandes ($n > 30$)
2. Caso de $n_a = n_b$ y $s^2_a = s^2_b$
3. Caso de $n_a = n_b$ y $s^2_a <> s^2_b$
4. Caso de $n_a <> n_b$ y $s^2_a = s^2_b$
5. Caso de $n_a <> n_b$ y $s^2_a <> s^2_b$
6. Caso de variables dependientes

2.6.1 Prueba de t student

Técnicamente se puede describir la prueba t de Student como aquella que se utiliza en un modelo en el que una variable explicativa (var. independiente) dicotómica intenta explicar una variable respuesta (var. dependiente) dicotómica. Es decir en la situación: dicotómica explica dicotómica. La prueba t de Student como todos los estadísticos de contraste se basa en el cálculo de estadísticos descriptivos previos: el número de observaciones, la media y la desviación típica en cada grupo. A través de estos estadísticos previos se calcula el estadístico de contraste

experimental. Con la ayuda de unas tablas se obtiene a partir de dicho estadístico el p-valor. Si $p < 0,05$ se concluye que hay diferencia entre los dos tratamientos. Las hipótesis o asunciones para poder aplicar la t de Student son que en cada grupo la variable estudiada siga una distribución Normal y que la dispersión en ambos grupos sea homogénea (hipótesis de homocedasticidad = igualdad de varianzas). Si no se verifica que se cumplen estas asunciones los resultados de la prueba t de Student no tienen ninguna validez. Por otra parte no es obligatorio que los tamaños de los grupos sean iguales, ni tampoco es necesario conocer la dispersión de los dos grupos. La distribución t es continua, tiene forma de campana y es simétrica respecto al cero como la distribución z.

La distribución t es dispersa y es más plana en centro que la distribución z, pero se acerca a ella cuando el tamaño de muestra crece.

El estadístico de prueba para el caso de una muestra está dado por:

$$t = \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

Para una prueba de dos colas con la distribución t, se rechaza la hipótesis nula cuando el valor del estadístico de prueba es mayor que

$$t_{n-1, \alpha/2}$$

o si es menor que

$$t_{n-1, \alpha/2}$$

Para una prueba de cola izquierda con la distribución t, se rechaza la hipótesis nula cuando el valor del estadístico de prueba es menor que

$$t_{n-1, \alpha/2}$$

2.7 LENGUAJES PROGRAMACIÓN

Para el desarrollo del prototipo se empleara el lenguaje de programación Matlab, por su facilidad de manejo de imágenes y redes neuronales.

La herramienta de software Matlab es un lenguaje técnico para aplicaciones computacionales que permite integrar computación, visualización y programación en un ambiente fácil de usar y en donde las soluciones son expresadas en notación matemática.

2.7.1 MATLAB

MATLAB (abreviatura de *MATrix LABoratory*, "laboratorio de matrices") es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M).

Entre sus prestaciones básicas se hallan: la manipulación de matrices, la representación de datos y funciones, la implementación de algoritmos, la creación de interfaces de usuario (GUI) y la comunicación con programas en otros lenguajes y con otros dispositivos hardware. El paquete MATLAB dispone de dos herramientas adicionales que expanden sus prestaciones, a saber, Simulink (plataforma de simulación multidominio) y GUIDE (editor de interfaces de usuario – GUI). Además, se pueden ampliar las capacidades de MATLAB con programas de apoyo especializados, denominados *toolboxes*. Entre los *toolboxes* disponibles cabe destacar el Image Processing Toolbox y Network Neural Toolbox, por el interés que tiene para este proyecto.

CAPITULO III

MARCO APLICATIVO

3.1 MODELO GENERAL DEL PROTOTIPO

El Modelo para la cuantificación y detección de alteraciones morfológicas de glóbulos rojos en imágenes microscópicas digitales está basada técnicas de procesamiento digital de imágenes y el reconocimiento de patrones, en la figura 3.1 se observa las etapas del modelo.

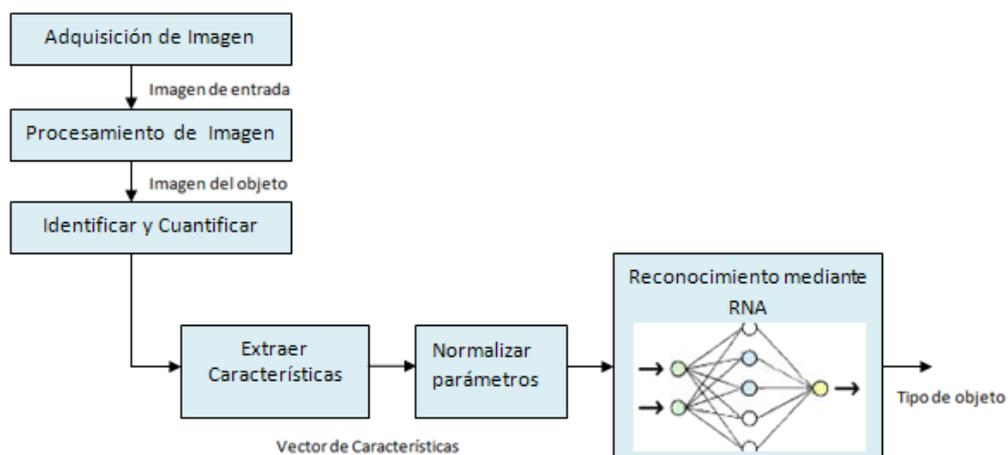


Figura 3.1: Modelo general

Fuente: Elaboración Propia

3.1.1 DESCRIPCION GENERAL

3.1.1.1 ADQUISICION DE IMÁGENES

La Adquisición de imágenes se realizó mediante la Web de la página <http://www.iqb.es/hematologia/atlas/eritrocitos.htm>, debido a que la información y el acceso a estas son restringidos en nuestro medio, se seleccionaron 25 imágenes fueron almacenadas en una base de imágenes en formato JPEG por razones de espacio de memoria. De las 25 imágenes se obtuvieron 25 imágenes individuales de diferentes alteraciones

morfológicas de Glóbulos Rojos que también fueron almacenadas en una base de imágenes en formato JPEG. En la figura 3.2 se puede observar una imagen digital de la base de imágenes.

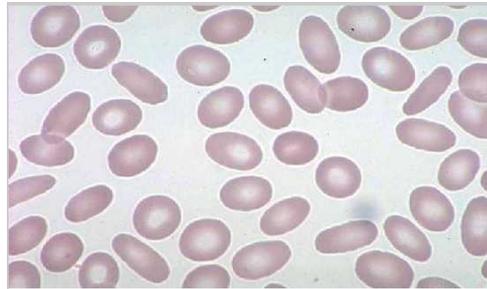
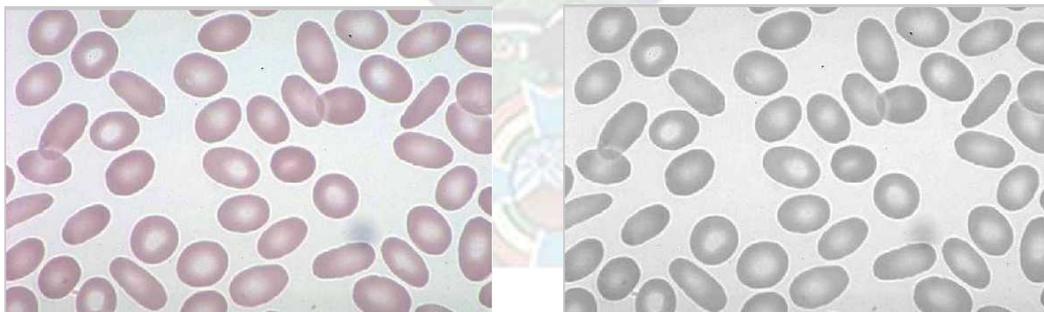


Figura 3.2: Imagen de globulos rojos con eliptocitos
Fuente: [Atlas de hematologia]

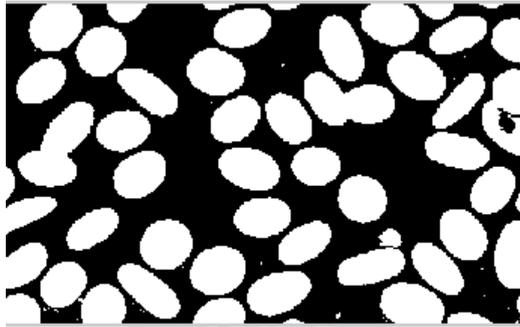
3.1.1.2 PROCESAMIENTO DE IMAGEN

Se realizo un análisis de las imágenes de las alteraciones morfológicas de los glóbulos rojos y se aplico los siguientes algoritmos para el procesamiento de estas imágenes conversión a escala de grises el cual permite separar los tres colores de capa pixel, obteniendo su valor correspondiente a rojo, verde y azul, la binarización se aplico a la imagen a escala de grises para segmentar los glóbulos rojos de la imagen, las cuales facilitaran los procesos de las siguientes etapas como la identificación y conteo, extracción de características y reconocimiento, en la Figura 3.2 se puede observar los resultados de la aplicación de los algoritmos.



a) Imagen original

b) Imagen a escala de grises



c) Imagen binaria

Figura 3.3: a), b) y c) Aplicación de algoritmos de Procesamiento de imágenes
Fuente: Elaboración Propia.

3.1.1.3. IDENTIFICACION Y CUANTIFICACION DE ALTERACIONES MORFOLOGICAS DE GLOBULOS ROJOS MEDIANTE DESCRIPTORES ELEMENTALES DE FORMA

La Identificación y cuantificación de las alteraciones morfológicas de los glóbulos rojos en imágenes se realizara mediante un algoritmo se basa en dos descriptores elementales de forma.

El algoritmo se basa en dos descriptores elementales de forma como el coeficiente de circularidad y el coeficiente elíptico se basan en la circularidad de los objetos estos descriptores son valores numéricos definidos entre $[0,1]$.

Geoméricamente, un glóbulo rojo sano o normal se puede describir como un objeto circular o redondo debido a que no presenta ni núcleo ni mitocondrias [Domínguez, 1996], si el objeto es circular o redondo entonces el CE (Coeficiente Elíptico) y CC (Coeficiente de Circularidad) tienen valor de 1, y si un objeto tiene forma irregular entonces su valor es <1 .

Realizando un estudio en imágenes con alteraciones morfológicas de glóbulos rojos según la variación de forma, se puede determinar si un glóbulo rojo esta deformado o no, empleando diferentes valores numéricos para los coeficientes de circularidad y elíptico de los cuales se hallaron los siguientes valores numéricos:

- Si CC (Coeficiente de Circularidad) < 0.8 y $CE > 0.5$ el glóbulo rojo es deformado.
- Si $CC > 0.8$ y el $CE < 0.5$ el glóbulo rojo es considerado Normal.

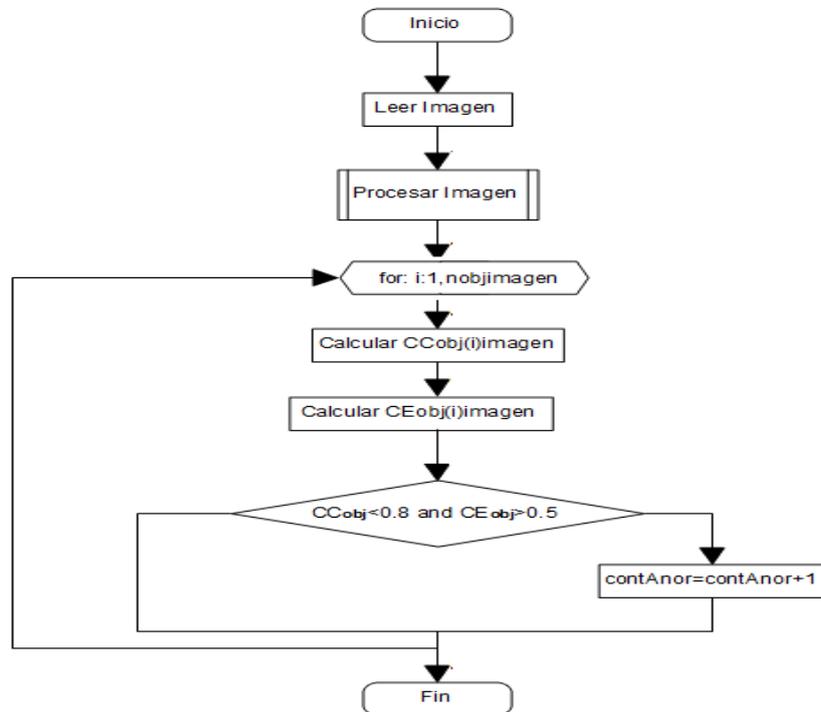


Figura 3.4: Algoritmo de identificación y conteo de Alteraciones Morfológicas de los Glóbulos Rojos según la variación de forma
Fuente: Elaboración Propia

En la figura 3.4 se puede observar la aplicación del algoritmo basado en descriptores elementales de forma en una imagen.

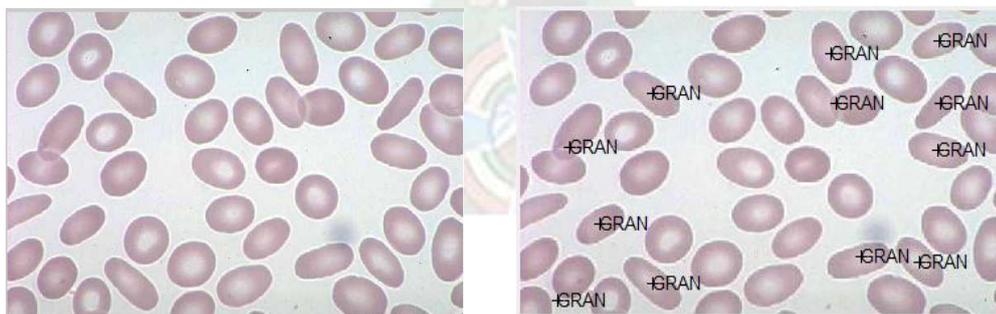


Figura 3.5: a) Imagen de glóbulos rojos con eliptocitos b) identificación y conteo de alteraciones morfológicas de glóbulos rojos superpuesta en la imagen
Fuente: Elaboración propia

3.1.1.4 EXTRACCION DE CARACTERISTICAS

En esta etapa se realizo la extracción de características de las alteraciones morfológicas de los glóbulos rojos utilizando descriptores geométricos como: Área, Perímetro, Excentricidad, Diámetro del eje mayor, Diámetro del eje menor, Circularidad y Elongación. Se realizo la extracción de características de 25 imágenes individuales de 5 tipos de alteraciones morfológicas de los glóbulos rojos para generar vectores de características como datos para los parámetros de entrada para el reconocimiento de patrones mediante una red neuronal. En la Figura 3.6 se puede observar la extracción de características de la imagen

| Img_dian02_01.jpg | Área | Perímetro | Excentricidad | Eje mayor | Eje menor | Circularidad | Elongación |
|---|------|-----------|---------------|-----------|-----------|--------------|------------|
|  | 7419 | 131.10 | 0.541 | 56.251 | 49.2198 | 0.718 | 1.400 |

Figura 3.6: Extracción de características de una imagen

Fuente: Elaboracion Propia

En Anexos C se encuentra la tabla 3.1 del vector de características generado por las 25 imágenes.

3.1.1.5. NORMALIZAR

Se deben normalizar los datos de los parámetros de entrada y salida según la función de activación utilizada en red neuronal de base radial, estos valores pueden ser [0,1]. Se normalizar los datos de acuerdo a la siguiente ecuación:

$$D_n = \frac{D_i - D_{\min}}{D_{\max} - D_{\min}} \quad (3.1)$$

Donde:

D_n = Dato Normalizado.

D_{\min} = Valor mínimo del conjunto de datos.

D_{\max} = Valor máximo del conjunto de datos.

D_i = Valor a normalizar.

En Anexos D se encuentra la tabla 3.2 el vector de características con los datos normalizados.

3.1.1.6 RECONOCIMIENTO MEDIANTE REDES NEURONALES

Para reconocer el tipo de alteración morfológica de los glóbulos rojos se utilizara las redes neuronales de base radial.

3.1.1.6.1 Diseño de la Red Neuronal

El modelo de red que se va utilizar es una red neuronal de base radial (RBF), esta red consta de 3 capas (ver Capitulo II).

3.1.1.6.2 Arquitectura de la Red Neuronal

La arquitectura de la red propuesta muestra las siguientes características:

Tipo de Red Neuronal: Red Neuronal con base radial es una red FeedForward (conexiones hacia adelante) de tipo multicapa.

Numero de neuronas en la capa de entrada: 7, de acuerdo a las variables de entrada.

$$X = [X_1, X_2, X_3, X_4, X_5, X_6, X_7]$$

| Variables de Entradas | Descripción |
|-----------------------|---------------|
| X_1 | Área |
| X_2 | Perímetro |
| X_3 | Excentricidad |
| X_4 | Eje Mayor |
| X_5 | Eje Menor |
| X_6 | Circularidad |
| X_7 | Elongación |

Tabla 3.3: Variables de entrada
Fuente: Elaboracion Propia.

Número de capas Ocultas: 1, de acuerdo a la topología de la red, explicada en el capítulo II.

Número de neuronas de la capa Oculta: será determinado después de la implementación de la red.

Número de neuronas de la capa salida: 5, de acuerdo a las variables de salida, para reconocer 5 tipos de alteración morfológica de glóbulos rojos según la variación de forma, , por lo cual se ha utilizado el valor de 0.9999 para indicar que una neurona de la capa de salida debe activarse ante un tipo de alteración morfológica de los glóbulos rojos reconocida y el valor de 0.0001 para indicar que no corresponde al tipo de alteración morfológica de los glóbulos rojos indicado por tal neurona de salida.

$$Y = [Y_1, Y_2, Y_3, Y_4, Y_5]$$

| Variables de Salida | Descripción |
|---------------------|-------------|
| Y_1 | Acantocito |
| Y_2 | Eliptocito |
| Y_3 | Equinocito |
| Y_4 | Dianocito |
| Y_5 | Drepanocito |

Tabla 3.4: Variables de salida
Fuente: Elaboracion Propia.

Función de transferencia: La función de transferencia usada es la gaussiana.

En la Figura 3.7 se representa el modelo de red neuronal de base radial que se utilizara para la clasificación o reconocimiento de las alteraciones morfológicas de glóbulos rojos.

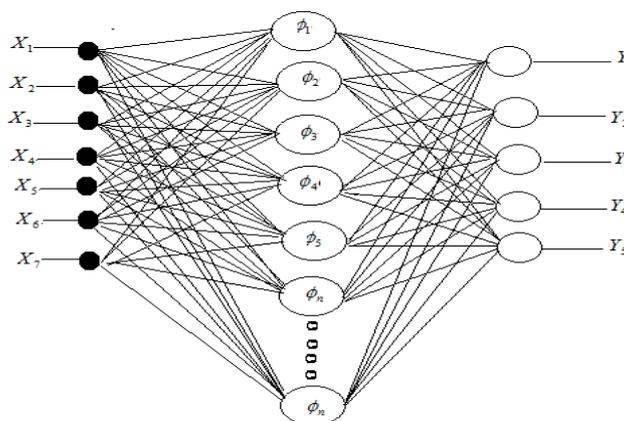


Figura 3.7: Arquitectura de la Red Neuronal de Base Radial

Fuente: Elaboración Propia

3.1.1.6.3 Entrenamiento de la Red Neuronal

Para el entrenamiento de la red neuronal de base radial se utilizó el software de Matlab 7.6, que tiene disponible el aprendizaje totalmente supervisado, estas redes neuronales de base radial pueden ser implementadas por medio de la función: newrb.

La función newrb plantea una forma más eficiente para crear una red de base radial, la función crea la red de forma iterativa, creando neuronas en cada iteración.

net=newrb(P,T,GOAL, SPREAD)

P: Vector del patrón de entrada.

T: Vector de las salidas deseadas.

GOAL: Es la precisión con que se quiere se realice el proceso de entrenamiento, no se debe colocar 0 pues la red tiende a memorizar los ejemplos de entrenamiento y la generalización que realice es muy mala, para entrenar la red neuronal el parámetro goal será definido después de varias pruebas y errores.

SPREAD: (Ajuste de precisión) es el radio con el que se quiere que tengan las funciones de base radial, este parámetro también será definido después de la implementación de la red neuronal.

Para el entrenamiento de la RFB se utilizará los datos de la tabla 3.5 (Anexos E) y los parámetros de la tabla 3.6.

| Parámetros | Valor |
|------------|-------|
| GOAL | 0.001 |
| SPREAD | 10.5 |

Tabla 3.6: Parámetros de la red de base radial
Fuente: Elaboración Propia.

Los valores de los parámetros fueron definidos después de realizar varias pruebas, en la Figura 3.7 se observa el entrenamiento de la red neuronal de

base radial, se muestra que fueron necesarias 550 neuronas en la capa oculta para poder alcanzar una respuesta adecuada por parte de la red.

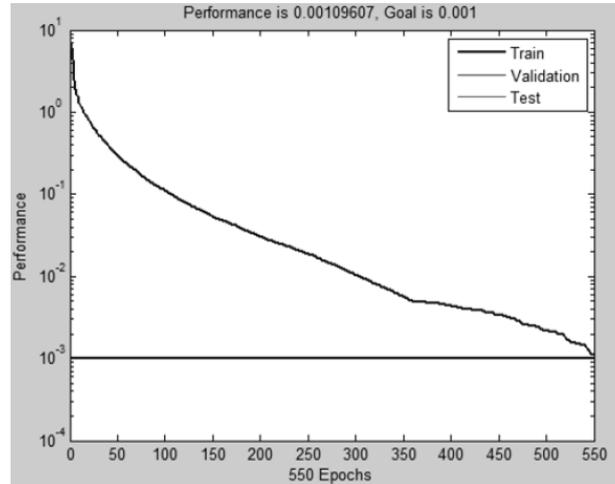


Figura 3.8: Entrenamiento de Red Neuronal de Base Radial
Fuente: Elaboración Propia

En la Figura se observa el nivel de correlación que existe entre la red de base radial creada y los patrones de entrenamiento para la red.

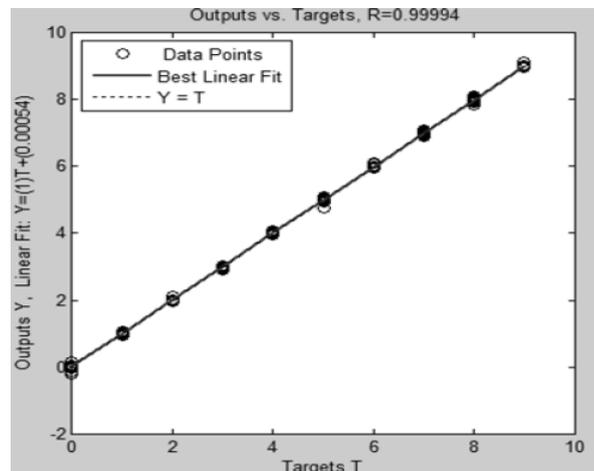


Figura 3.9: Correlación de la red y los patrones de salida T
Fuente: Elaboración Propia

3.2 DESARROLLO DEL PROTOTIPO

La Metodología de desarrollo de software del prototipo de cuantificación y detección de alteraciones morfológicas de glóbulos rojos en imágenes

microscópicas es el modelo en cascada. Dentro de la etapa de análisis y diseño se hace uso del Lenguaje de Modelado Unificado (UML).

3.2.1 ANALISIS DE REQUERIMIENTOS

En esta etapa se realiza la recopilación de los requisitos, y se procede a modelar la aplicación mediante casos de uso que representaran la funcionalidad del sistema y su interacción con los agentes externos.

3.2.1.1 DIAGRAMA DE CASOS DE USO

Los diagramas de caso de uso presentado a continuación refleja las funcionalidades en el prototipo.

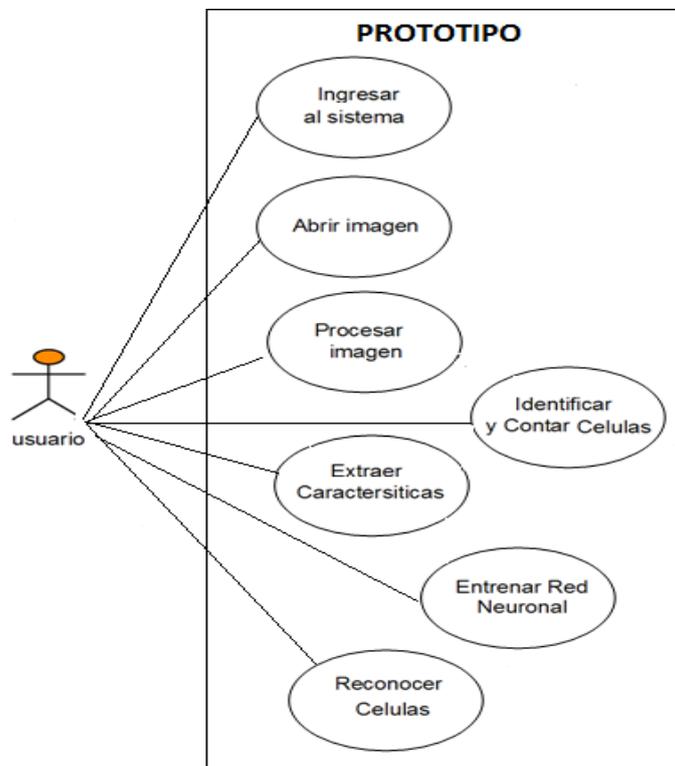


Figura 3.10: Diagrama de casos de uso de Alto Nivel
Fuente: Elaboración Propia

3.2.1.2 DIAGRAMA DE CASOS DE USO EN FORMATO EXPANDIDO

Caso de uso: Ingresar al sistema

| | |
|---------------------|----------------------------|
| CASO DE USO: | Ingresar al sistema |
|---------------------|----------------------------|

| | |
|-------------------------------------|---|
| ACTOR: | Usuario |
| PROPÓSITO: | Ingresar al sistema |
| DESCRIPCIÓN: | El usuario hace un clic en ingresar y se muestra la pantalla de inicio del sistema. |
| PRECONDICIÓN: | |
| POSTCONDICION: | El usuario ha ingresado al sistema |
| CURSO NORMAL DE EVENTOS | |
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| | 1. El sistema muestra la pantalla de ingreso al sistema. |
| 2. El usuario hace clic en ingresar | 3. El sistema muestra la pantalla de inicio. |

Tabla 3.7: Descripción de caso de uso: Ingresar al sistema

Caso de uso: Abrir imagen

| | |
|---|--|
| CASO DE USO: | Abrir imagen |
| ACTOR: | Usuario |
| PROPÓSITO: | Permitir al usuario la posibilidad de elegir la imagen que desea analizar, previamente digitalizada en formato JPEG. |
| DESCRIPCIÓN: | Al darle clic en “Abrir imagen”, el usuario podrá elegir la imagen de muestra sanguínea, previamente digitalizada en formato JPEG. |
| PRECONDICIÓN: | Debe haber un archivo de imágenes en formato JPEG. |
| POSTCONDICION: | El usuario ha cargado la imagen de muestra sanguínea seleccionada. |
| CURSO NORMAL DE EVENTOS | |
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| 1. El usuario hace clic el “Abrir imagen”. | 2. El sistema le presenta el directorio donde se encuentra el banco de imágenes. |
| 3. El usuario selecciona una de las imágenes presentadas por el sistema y oprime la opción “Aceptar”. | 4. El sistema carga la imagen en el espacio respectivo. |

Tabla 3.8: Descripción de caso de uso: Abrir imagen

Caso de uso: Identificar y Contar células

| | |
|---------------------|-------------------------------------|
| CASO DE USO: | Identificar y Contar células |
|---------------------|-------------------------------------|

| | |
|---|---|
| ACTOR: | Usuario |
| PROPÓSITO: | Permitir al usuario detectar y contar células con cualquier tipo de deformaciones en una imagen de muestra sanguínea, previamente segmentada. |
| DESCRIPCIÓN: | Al darle clic en “Identificar y Contar células”, el sistema aplicara el algoritmo de detección y conteo alteraciones morfológicas de Glóbulos rojos basado en descriptores de forma. |
| PRECONDICIÓN: | El usuario debe haber segmentado la imagen de muestra sanguínea en formato JPEG. |
| POSTCONDICION: | En la imagen de muestra sanguínea, se indicara los glóbulos rojos con deformaciones identificados y contados.. |
| CURSO NORMAL DE EVENTOS | |
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| 1. El usuario elige la opción “Identificar y Contar células”. | 2. El sistema aplicara el algoritmo de identificación y conteo alteraciones morfológicas de Glóbulos rojos basado en descriptores de forma a una imagen de muestra sanguínea, previamente segmentada. |
| | 3. El sistema mostrara en pantalla la imagen de muestra sanguínea con los glóbulos rojos identificados y contados. |

Tabla 3.9: Descripción de caso de uso: Identificar y contar células

Caso de uso: Extraer Características

| | |
|-----------------------|--|
| CASO DE USO: | Extraer Características |
| ACTOR: | Usuario |
| PROPÓSITO: | Permitir al usuario extraer características de un glóbulo rojo con cualquier tipo de deformaciones en una imagen de muestra sanguínea, previamente segmentada. |
| DESCRIPCIÓN: | Al darle clic en “Extraer Características”, el sistema extraerá las características de la imagen de un glóbulo rojos con cualquier tipo de alteración morfológica, previamente segmentada. |
| PRECONDICIÓN: | El usuario debe haber segmentado la imagen de muestra sanguínea en formato JPEG. |
| POSTCONDICION: | Obtener las características de las glóbulos rojos deformados, para generan un vector de características |

| | |
|--|---|
| | o patrón. |
| CURSO NORMAL DE EVENTOS | |
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| 1. El usuario elige la opción “Extraer Características”. | 2. El sistema extraerá las características de los glóbulos rojos deformados y los almacenara en un vector de características. |
| | |

Caso de uso: Entrenar Red Neuronal

| | |
|--|---|
| CASO DE USO: | Entrenar Red Neuronal |
| ACTOR: | Usuario |
| PROPÓSITO: | Dar a la Red Neuronal la capacidad de aprender a reconocer o clasificar un tipo de alteración morfológica de glóbulos rojos. |
| DESCRIPCIÓN: | El usuario entrenara la Red Neuronal configurando los parámetros de aprendizaje de la red, tanto la taza de entrenamiento, como la modificación de los pesos. |
| PRECONDICIÓN: | La Red Neuronal debe estar creada y diseñada para clasificar alteraciones morfológicas de glóbulos rojos. |
| POSTCONDICION: | Red Neuronal entrenada. |
| CURSO NORMAL DE EVENTOS | |
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| 1. El usuario elige la opción “Entrenar Red Neuronal”. | 2. El sistema mostrara una pantalla para la configuración de la red neuronal |
| 3. El usuario configura los parámetros de aprendizaje de la red. | 4. El sistema entrena la red según los parámetros introducidos por el usuario. |

Tabla 3.11: Descripción de caso de uso: Entrenar Red Neuronal

Caso de uso: Reconocer célula

| | |
|-----------------------|---|
| CASO DE USO: | Reconocer célula |
| ACTOR: | Usuario |
| PROPÓSITO: | Reconocer o clasificar el tipo de alteración morfológica de glóbulos rojos según la variación de forma en una imagen. |
| DESCRIPCIÓN: | El usuario |
| PRECONDICIÓN: | Prototipo con Red Neuronal entrenada. |
| POSTCONDICION: | El sistema muestra la salida de red |

| CURSO NORMAL DE EVENTOS | |
|---|---|
| ACCIONES DE ACTOR | RESPUESTA DEL SISTEMA |
| 1. El usuario elige la opción "Reconocimiento". | 2. El sistema muestra la pantalla de reconocimiento. |
| 3. El usuario elige la imagen que desea reconocer y el archivo de datos del patrón de entrada de la imagen. | 4. El sistema carga la imagen seleccionada en pantalla y el archivo de datos del patrón de entrada de la imagen |
| 5. El usuario presiona reconocer | 6. El sistema reconoce la imagen seleccionada y muestra las salidas de la red. |

Tabla 3.12: Descripción de caso de uso: Reconocer células

3.2.2 DISEÑO DEL PROTOTIPO

En base a lo descrito en el análisis de requerimientos se desarrollara un modelo para el diseño del prototipo del sistema de cuantificación y detección de alteraciones morfológicas de glóbulos rojos en imágenes microscópicas de muestras de sangre.

3.2.2.1 DIAGRAMA DE SECUENCIA

Los diagramas de secuencia destacan el paso del mensajes tal como se desarrolla a lo largo del tiempo lo cual permite modelar el comportamiento dinámico de los casos de uso.

3.2.2.1.1 Ingresar al sistema

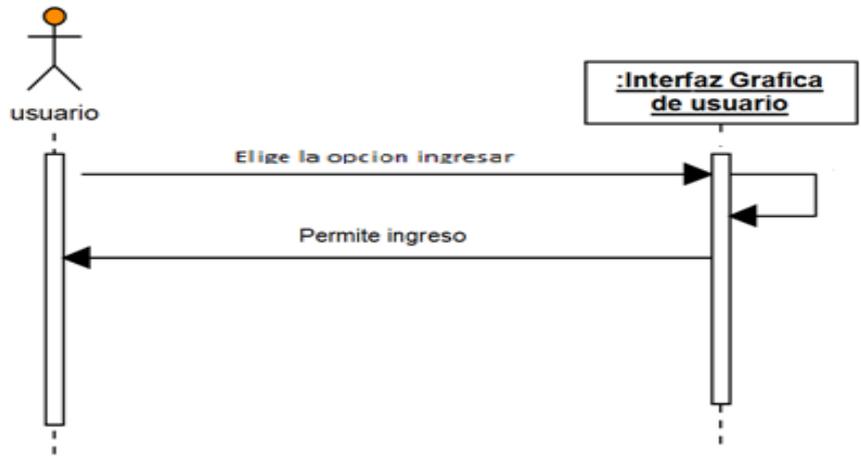


Figura 3.11: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.1.2 Abrir Imagen

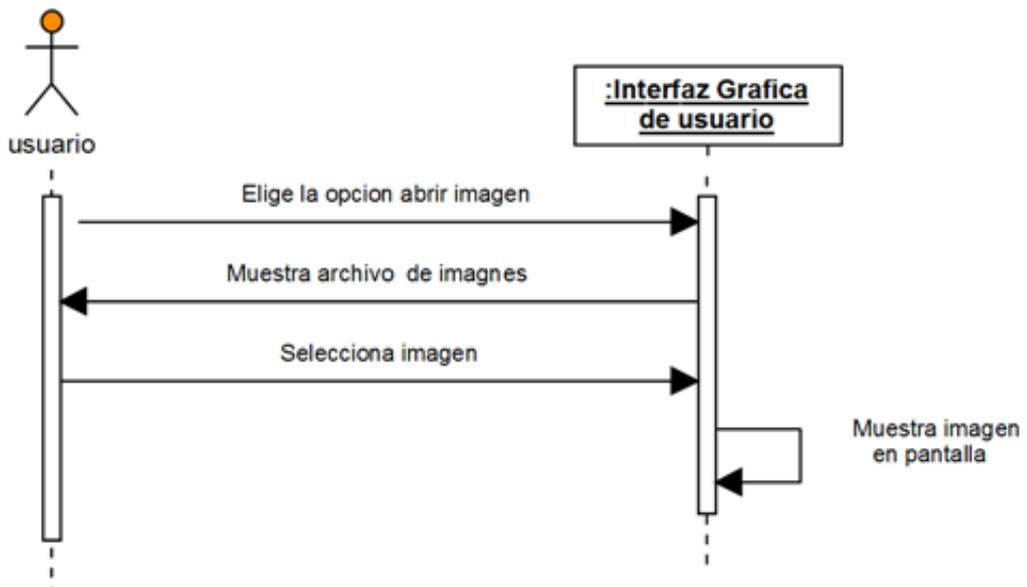


Figura 3.12: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.1.3 Procesar imagen

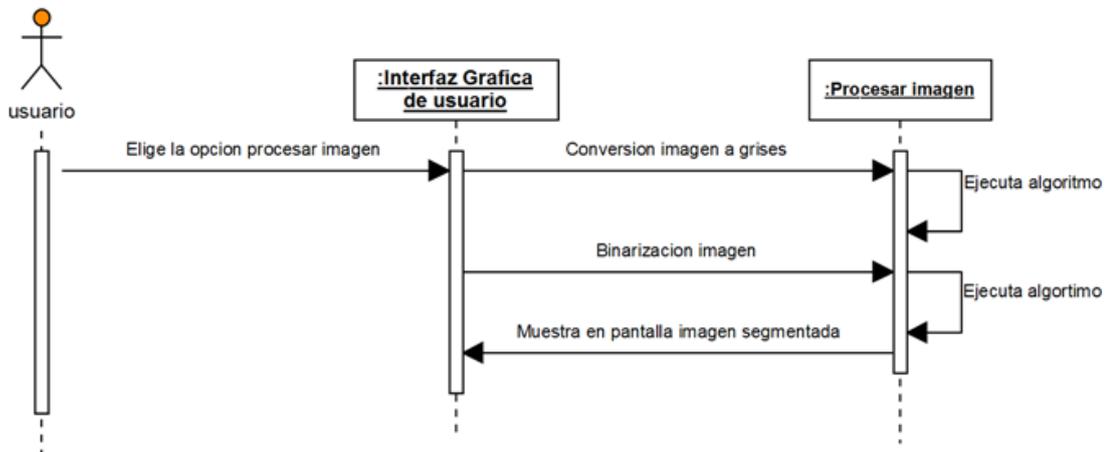


Figura 3.13: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.1.4 Identificar y Contar células

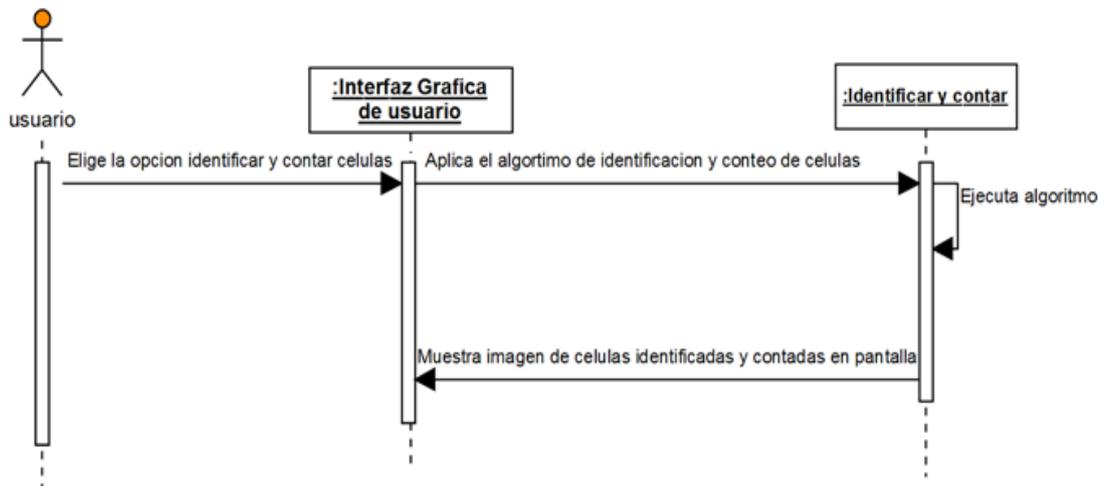


Figura 3.14: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.1.5 Extraer Características

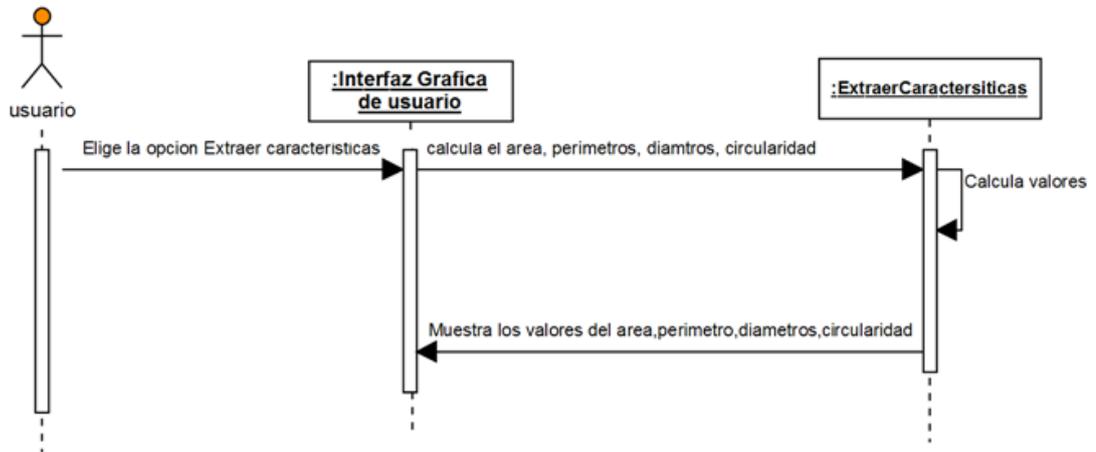


Figura 3.15: Diagrama de Secuencia: Extraer Características
Fuente: Elaboración

3.2.2.1.6 Entrenar Red Neuronal

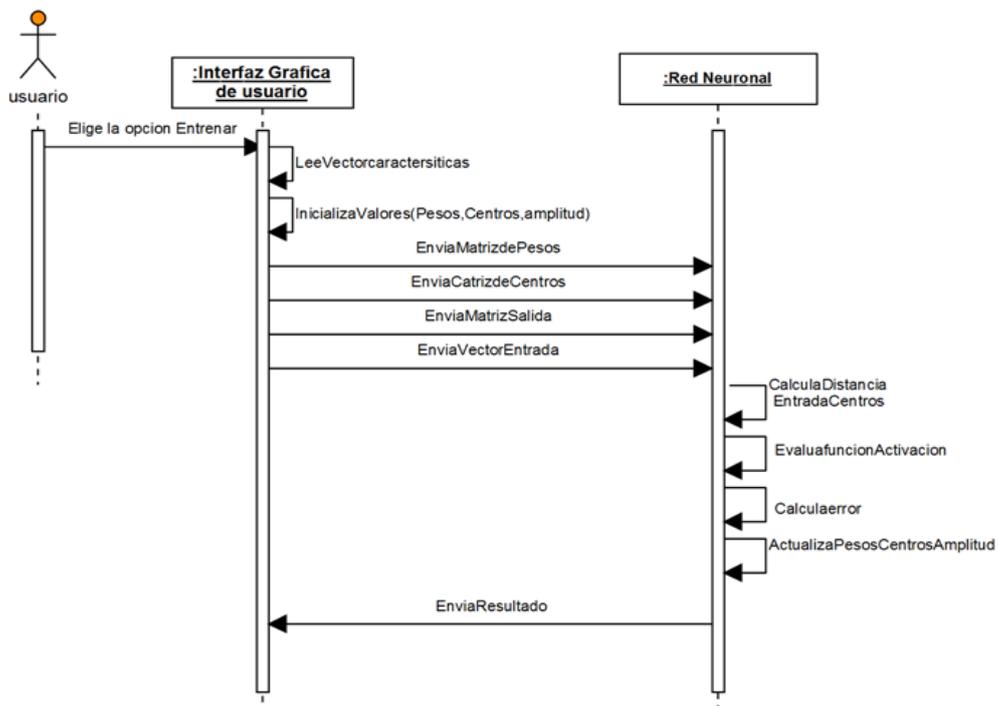


Figura 3.16: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.1.7 Reconocer Célula

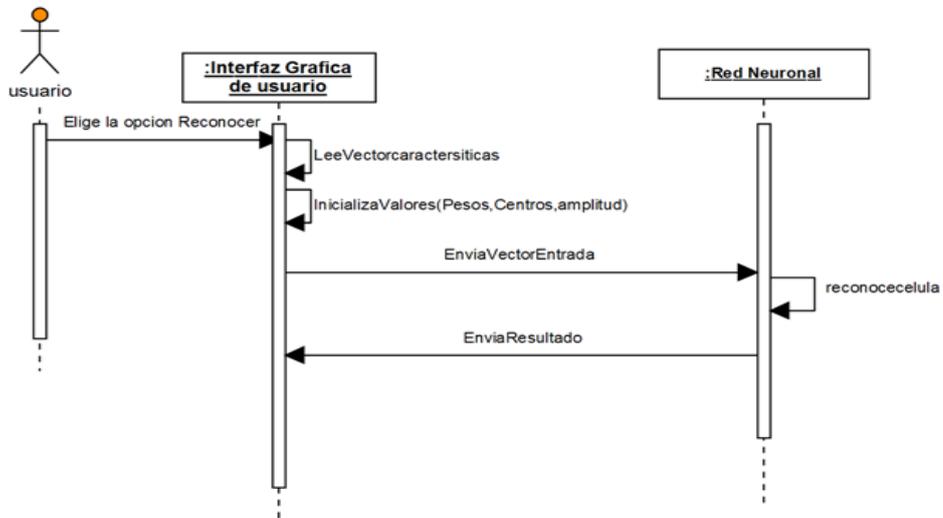


Figura 3.17: Diagrama de Secuencia
Fuente: Elaboración Propia

3.2.2.2 DIAGRAMA DE COLABORACION

3.3.2.2.1 Abrir imagen

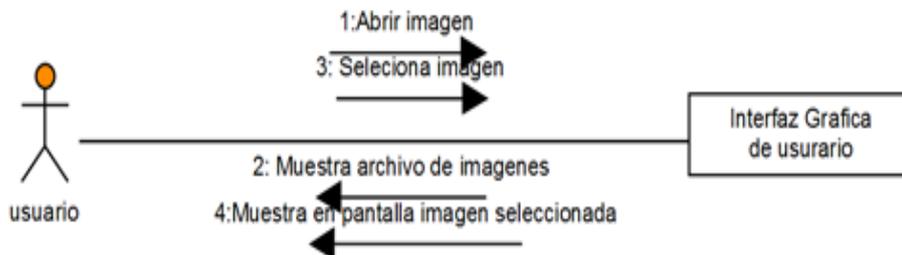


Figura 3.18: Diagrama de Colaboración
Fuente: Elaboración Propia

3.3.2.2 Identificar y Contar células

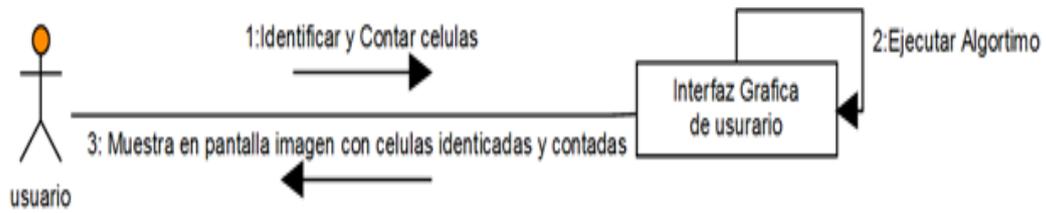


Figura 3.19: Diagrama de Colaboración
Fuente: Elaboración Propia

3.3.2.2.3 Entrenar Red Neuronal

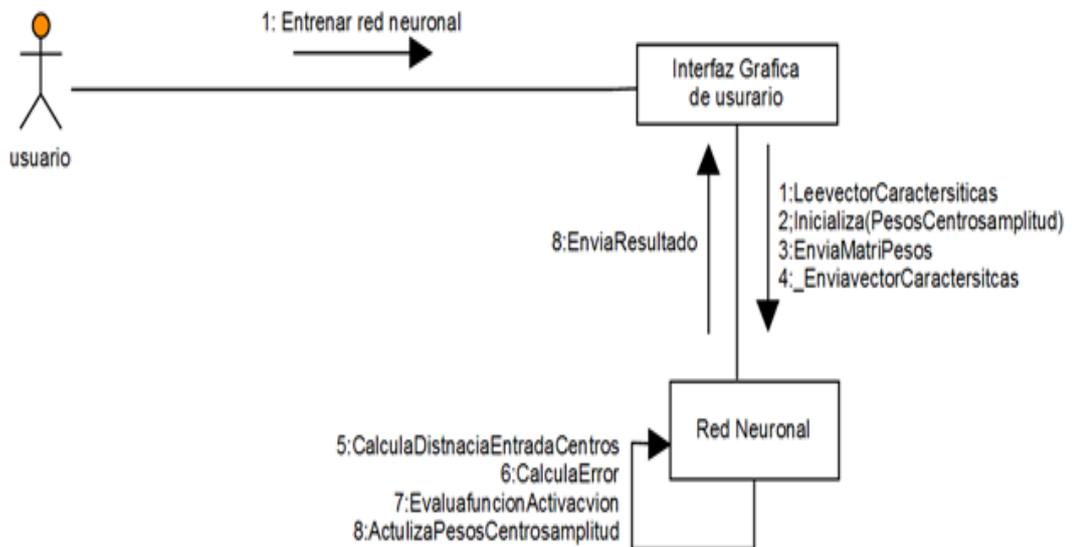


Figura 3.20: Diagrama de Colaboración
Fuente: Elaboración Propia

3.2.2.3 DIAGRAMA DE CLASES

En la figura 3.21, se muestra el diagrama de clases del prototipo, las clases y las relaciones entre ellas.

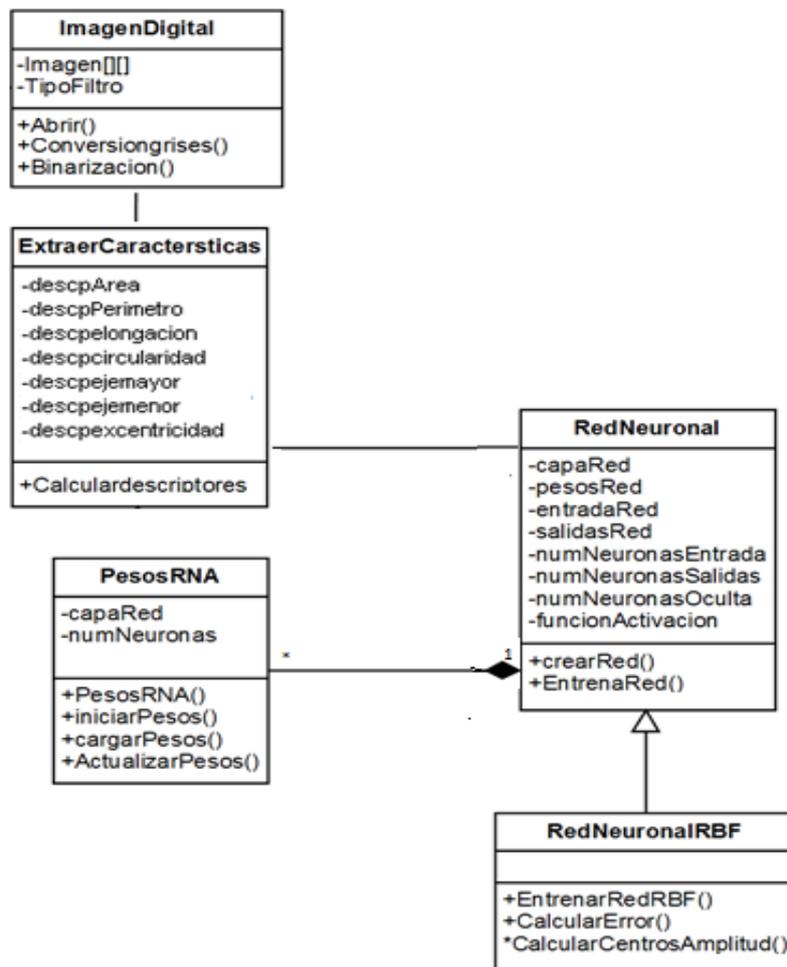


Figura 3.21: Diagrama de Clases del prototipo
Fuente: Elaboración

3.2.2.4 DISEÑO DE INTERFACES

El prototipo presenta una interfaz grafica diseñada en función al análisis y diseño ya presentados.

a. PANTALLA DE PROCESAMIENTO DE IMAGEN

En esta ventana de procesamiento de imágenes una vez cargada la imagen, se realizara la identificación y cuantificación de las alteraciones morfológicas de glóbulos rojos y se seleccionara una subimagen.

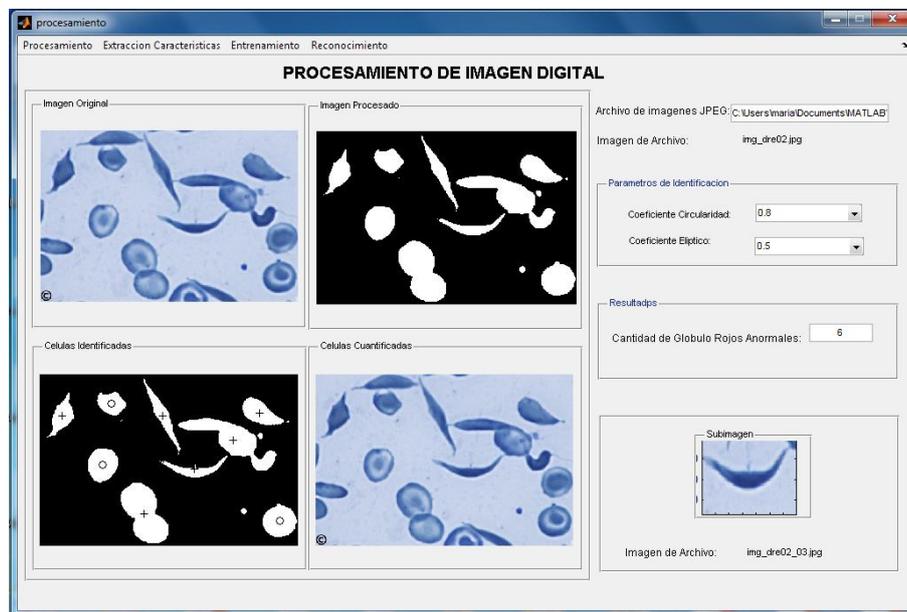


Figura 3.22: Interfaz gráfica del procesamiento de imágenes

b. PANTALLA DE EXTRACCIÓN DE CARACTERSTICAS

En esta ventana se cargara la imagen seleccionada se extraerá características, se normalizara los datos y guardar en un archivo de texto TXT que contiene los patrones de entrada de la red neuronal.

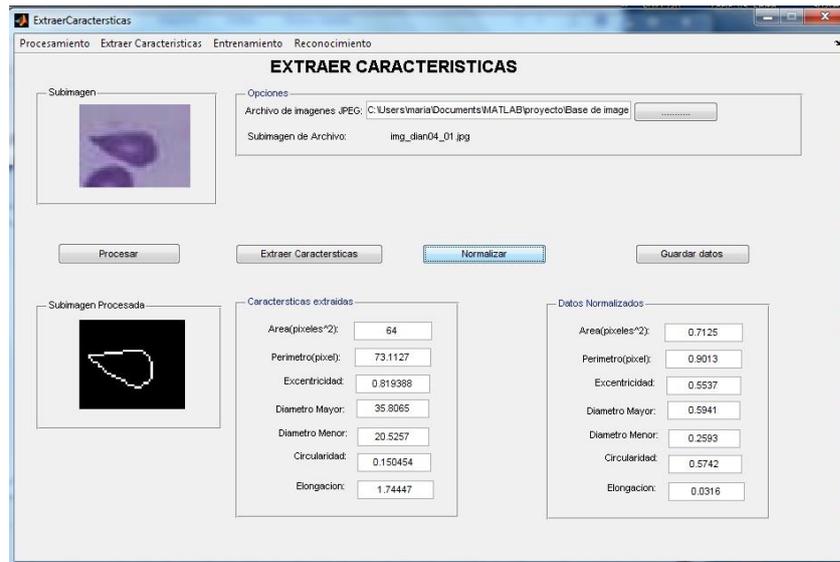


Figura 3.23: Interfaz gráfica de la extracción de características

c. PANTALLA DE ENTRENAMIENTO DE LA RED NEURONAL

En esta ventana se configurara la red neuronal y se cargara los archivos de texto TXT de entrada y salida para entrenar la red neuronal de base radial

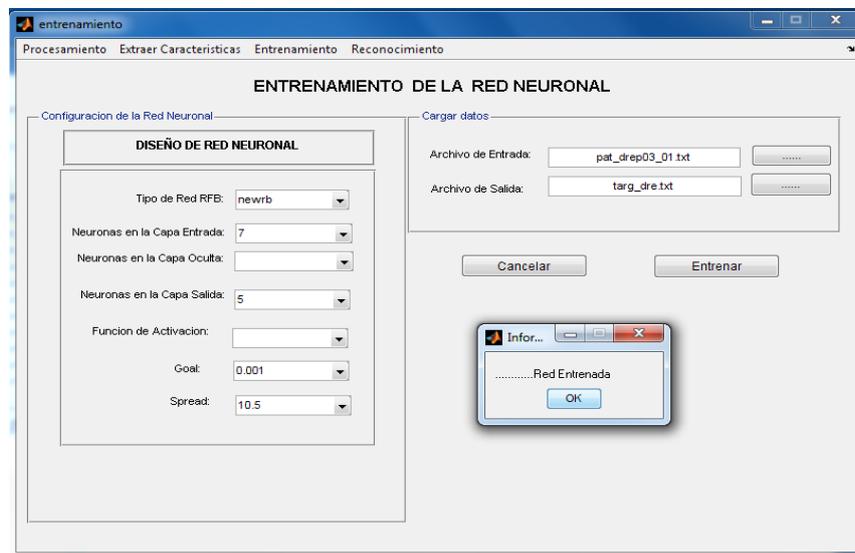


Figura 3.24: Interfaz gráfica del entrenamiento de la red neuronal

d. PANTALLA DE RECONOCIMIENTO

En esta ventana se reconocerá el tipo de alteración morfológica de los glóbulos rojos y cargara la imagen a reconocer y el archivo de entrada.

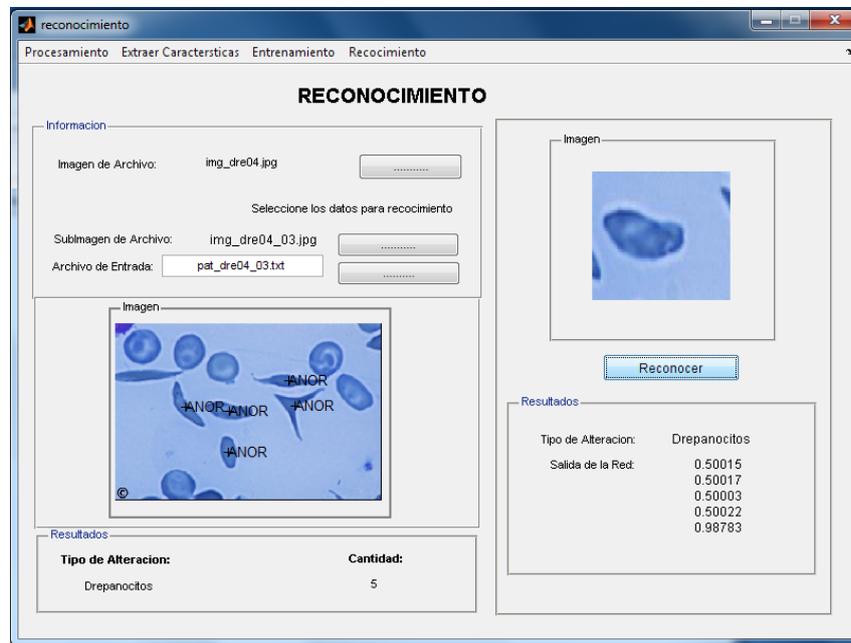


Figura 3.25: Interfaz gráfica del reconocimiento

3.2.3 IMPLEMENTACION Y PRUEBAS

3.2.3.1 Implementacion del prototipo

La implementacion del prototipo fue desarrollada en base al de diseño del prototipo propuesto en la etapa anterior. La herramienta de software utilizada para llevar a cabo la implementación de este proyecto es Matlab 7.6, con librerías de Image Processing Toolbox y Neural Network Toolbox.

3.2.3.1. 1. Implementación del algoritmo basado en descriptores de elementales de forma

En la figura 3.36 se muestra la implementación en Matlab 7.6. del algoritmo basado en descriptores de elementales de forma.

```
%% Algoritmo basado en descriptores elementales de forma
im=imread('eliptocitos5.jpg'); % imagen de prueba
figure; imshow(im);
grises=rgb2gray(im);
umb=graythresh(grises);
bin=im2bw(grises,umb);

imc=imcomplement(bin);
imgr=imfill(imc,'holes');
iml=imclearborder(imgr);
figure; imshow(iml)
labels = bwlabel(iml);
features = regionprops(labels,'Area','Perimeter','Centroid',
'MajorAxisLength', 'MajorAxisLength ');

%%Sobreponer el resultado de la identificación y conteo de glóbulos rojos.
imshow(im);hold on;
figure; imshow(label2rgb(labels));hold on;
for i=1:numel(features)
if(4*pi*features(i).Area/(features(i).Perimeter)^2 > .8)
    &&( (features(i).MajorAxisLength/(features(i).MajorAxisLength)<.5 )
plot(features(i).Centroid(1),features(i).Centroid(2),'ko');
else
plot(features(i).Centroid(1),features(i).Centroid(2),'k+');
text(X,Y,'GRAN');
ci=ci+1;
end
end
hold off;
title(strcat('Objetos Identificados: ',num2str(cantidadObjetos)))
```

Figura 3.26: Algoritmo basado en descriptores elementales de forma
Fuente: Elaboración Propia

3.2.3.2 Pruebas del prototipo

Para las pruebas del prototipo, se utilizaron 15 imágenes, los resultados que se presentan en la tabla 3.12 a continuación.

| Nº | Imagen de Archivo | Conteo de GR con Alteración | Tipo de Alteración Identificado | Salida de la red neuronal | %Efectividad del prototipo |
|----|-------------------|-----------------------------|---------------------------------|---------------------------|----------------------------|
| 1 | img_equi03.jpg | 12 | Equinocitos | 0.97943 | 95 |
| 2 | img_dre05.jpg | 16 | Drepanocitos | 0.97950 | 95 |
| 3 | img_acan01.jpg | 6 | Acantocitos | 0.98869 | 97 |
| 4 | img_equi04.jpg | 5 | Equinocitos | 0.9945 | 98 |
| 5 | img_elip02.jpg | 16 | Eliptocitos | 0.99300 | 98 |
| 6 | img_dian02.jpg | 18 | Dianocitos | 0.96904 | 95 |
| 7 | img_dian03.jpg | 21 | Dianocitos | 0.99838 | 98 |
| 8 | img_acan02.jpg | 13 | Acantocitos | 0.99562 | 98 |
| 9 | img_dre03.jpg | 9 | Drepanocitos | 0.99731 | 98 |
| 10 | img_elip03.jpg | 11 | Eliptocitos | 0.99300 | 98 |
| 11 | img_acan04.jpg | 7 | Acantocitos | 0.97965 | 95 |
| 12 | img_equi02.jpg | 22 | Equinocitos | 0.96942 | 95 |
| 13 | img_elip05.jpg | 32 | Eliptocitos | 0.98723 | 97 |
| 14 | img_dre04.jpg | 11 | Drepanocitos | 0.98943 | 97 |
| 15 | img_dian04.jpg | 14 | Dianocitos | 0.98851 | 96 |

Tabla 3.13 Resultados del prototipo

3.3 PRUEBA DE HIPOTESIS

Para probar la hipótesis planteada:

“Aplicando Procesamiento digital de Imágenes y reconocimiento de patrones a una imagen digital de una muestra de sangre optimizara la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos”.

Aplicaremos la distribución t-student , a una muestra de datos de la tabla 3.1 de los resultados de prototipo.

Para comprobar la hipótesis planteada se tomara en cuenta las siguientes hipótesis de trabajo:

H_0 = La Aplicación de Procesamiento digital de Imágenes y reconocimiento de patrones a una imagen digital optimizara la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos en un % ≤ 85

H_1 = La Aplicación de Procesamiento digital de Imágenes y reconocimiento de patrones a una imagen digital optimizara la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos en un % > 85 .

Calculando la distribución t-student en la ecuación 3.1:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \quad (3.1)$$

$$t = \frac{96,9 - 85}{\frac{2.061154}{\sqrt{15}}} = 1,0695$$

Una distribución t-student , tiene una distribución normal con n-1 grados de libertad, usando un nivel de significancia de $\alpha=0.05$ significa que el nivel de confianza es 95%.

$$t_{\frac{\alpha}{2}} = t_{0.025} = 1.96 \text{ (Detablas)}$$

Utilizando el criterio de decisión y comparando el valor de los estadísticos t_0 y $t_{\alpha/2}$, se observa que el valor del estadístico de la prueba se encuentra en la región de aceptación, por lo tanto se rechaza la hipótesis nula H_0

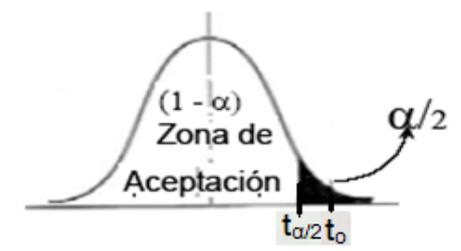


Figura 3.27: Distribución de t_0 y $t_{\alpha/2}$ en el grafico para la toma de decisión

Cuando rechazamos la hipótesis nula se considera que la prueba es potente, si aceptamos la hipótesis nula el criterio de decisión es débil, ya que generalmente se busca rechazar H_0 .

Por lo tanto la Aplicación de Procesamiento digital de Imágenes y reconocimiento de patrones a una imagen digital optimizara la cuantificación y detección de las alteraciones morfológicas de glóbulos rojos en un $\% >85$.

CAPITULO IV

CONCLUSIONES Y RECOMENDACIONES

4.1 CONCLUSIONES

En este trabajo de tesis se desarrollado un prototipo para la cuantificación y detección de alteraciones morfológicas de glóbulos rojos en imágenes microscópicas digitales que cumple con los objetivos planteados.

La aplicación de de técnicas de procesamiento de imágenes como: conversión a escala de grises, binarizacion, fueron aplicadas a las imágenes digitales.

Se ha propuesto un algoritmo basado en descriptores elementales como el coeficiente de circularidad (CC) y el Coeficiente de Elíptico (CE) para la identificación y conteo automático de los glóbulos rojo según la variación de forma, los valores numéricos del CC y CE son 0.8 y 0.5 respectivamente.

Se utilizo el modelo de red neuronal de base radial implementada en el prototipo para la clasificación, este utiliza 7 características extraídas de una imagen digital para el reconocimiento de las 5 tipos de alteraciones morfológicas según la variación de forma (acantocitos, eliptocitos, equinocitos, dianocitos , drepanocitos) con un porcentaje de certeza de 96,9%.

El desarrollo del prototipo mediante el software Matlab 7.6 (MatlabR2008a) fue el adecuado, ya que esta herramienta facilita el uso de sus librerías como Image Processing toolbox y Neural Network toolbox para el manipulación y procesamiento de imágenes, además de la creación, entrenamiento de la red neuronal de base radial.

4.2 RECOMENDACIONES

Para una mayor eficiencia del prototipo, se recomienda utilizar imágenes que sean de una misma resolución, y que las imágenes de los glóbulos rojos de la imagen no sean solapadas.

Aplicar otras técnicas de procesamiento de imágenes para la segmentación de las alteraciones morfológicas de glóbulos rojos.

Aplicar otros valores numéricos del CC y CE para la identificación y conteo automático de los glóbulos rojo según la variación de forma.

Para una mayor eficiencia del prototipo, se recomienda utilizar más parámetros de entrada a red la red neuronal de base radial, para obtener un mayor porcentaje de eficiencia.

Utilizar otros modelos de redes neuronales como backpropagacion, para el reconocimiento de las alteraciones morfológicas de glóbulos rojos para obtener un mayor porcentaje de eficiencia.

BIBLIOGRAFIA

HERNANDES, Fernandez & Baptista, "**Metodologías de la investigación**", 3ra. Edición, 2003, Ed. McGraw, 706p.

PAJARES G., CRUZ J. M. "**Visión por Computador**". México, Editorial AlfaOmega, 2002. pp. 291-294.

PLATT W., "**Atlas de Hematología en color**". Barcelona, Editorial JIMS, 1972. pp. 111-114, 185-190.2005.

RODRÍGUEZ J., CORREA C., "**Caracterización geométrica de los glóbulos rojos. Diferenciación de muestras normales y patológicas**". Colombia. 2008.

CUEVAS, E., "**Visión por Computador utilizando Matlab y el Toolbox de Procesamiento Digital de Imágenes**", 2007.

HUMBERTO S., SOTELO R, "**Nuevo método para contar objetos en imágenes**". México. 1999.

JACOBSON, I. "**Aplicación UML en el Proceso Unificado**" Presentacion. Rational Software. Editorial McGraw Hill. pp.400.

PRESSMAN, R. "**Ingeniería del software**". 5ta edición, McGraw Hill, España. 2002.

FRAGA, L. "**Procesamiento de imágenes digitales**", Mayo 2001.

MEDINA, R. **"Bases del Procesamiento de imágenes Médicas"**.
Universidad de los Andes. Facultad de ingeniería, Grupo de Ingeniería
Biomédica de la ULA. Venezuela.

MARIN DIAZARAQUE, J. **"Introducción a las Redes Neuronales
aplicadas"**, 2002.

HILERA, J.R., MARTINEZ, V.J. **"Redes Neuronales Artificiales"**. Addison-
Wesley. Iberoamérica

FUENTES ELECTRONICAS

ACOSTA, M.I & ALFONZO Z. **"Tutorial sobre redes neuronales
artificiales aplicadas en ingeniería eléctrica"**, disponible:
<http://ohm.utp.edu.co/neuronales/Navigator.html>. Consultado: 21/08/20013

DAZA, S. **"Redes neuronal artificial. Fundamentos, Modelos y
aplicaciones"**, disponible: [http://www.ilustrados.com/publicaciones/
EpyVZEUFAAboUbISx.php](http://www.ilustrados.com/publicaciones/EpyVZEUFAAboUbISx.php). consultado: 21/08/20013

DIAZ, M. **"Redes neuronales artificiales"**, disponible:
<http://www.cs.us.es/~delia/sia/html98-99/pag-alumnos/web9/indice.html>.
Consultado 14/08/20013

GONZALES, M. **"Librería de redes neuronales para usar con Matlab"**.
http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf.
Consultado: 30/08/20013

LOPEZ, M. **"Redes neuronales artificiales."**, disponible:
[http://www.iiia.csic.es/~mario/rna/tutorial/ index.html](http://www.iiia.csic.es/~mario/rna/tutorial/index.html). Consultado:15/10/2013.

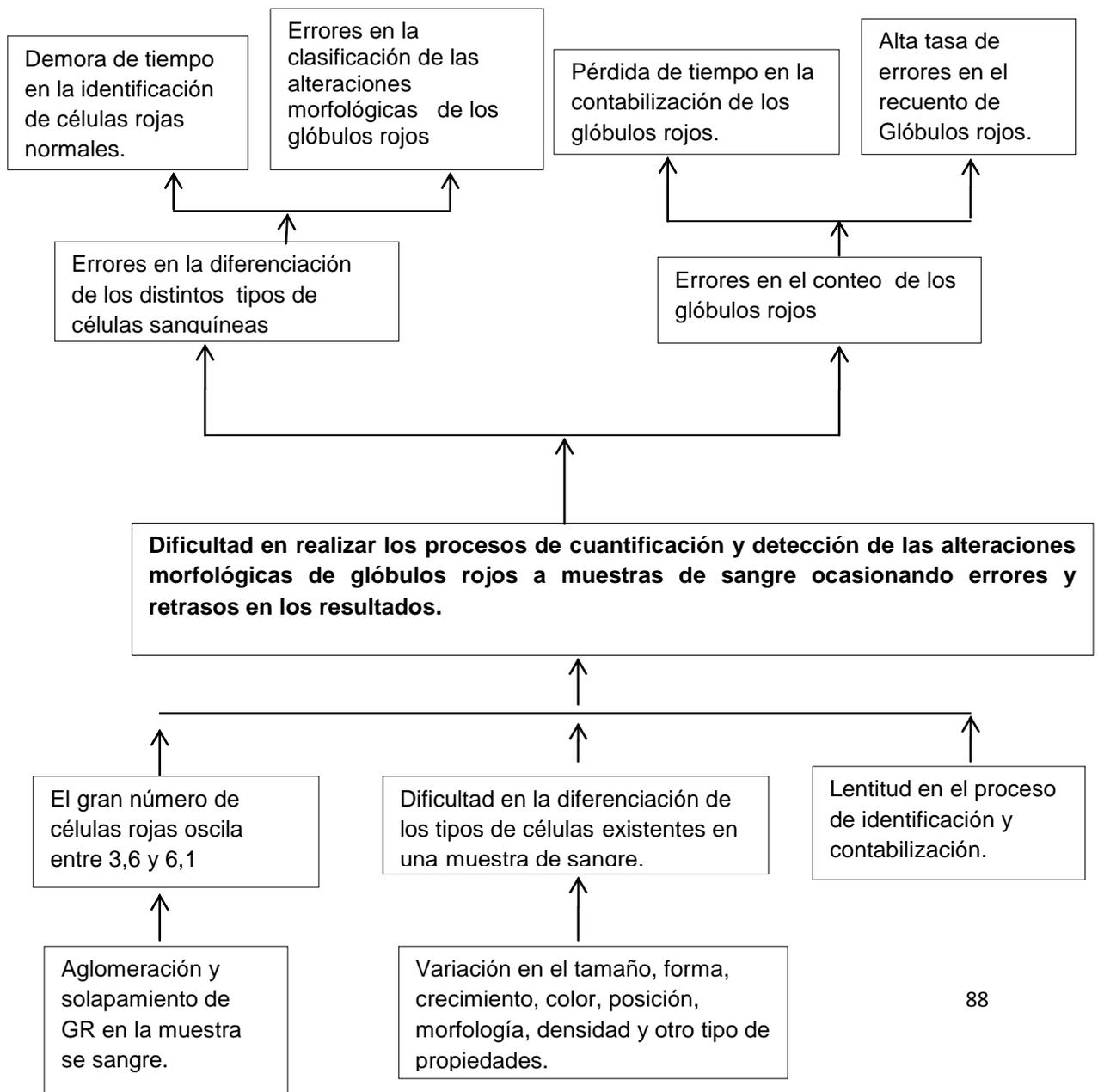
HILMAN, D, **"Manual de Hematología"**, disponible:
<http://www.cc.uoa.gr/health/pathology/aoh/images.htm>. Consultado:17/11/2013

FREGGAIRO. E. "**hematología de Eritrocitos**", disponible:
http://orbita.starmedia.com/forobiop/a_h_eritro.html. Consultado: 22/11/2013

OSORIO, G. "**Atlas de hematología**", disponible:
<http://www.iqb.es/hematologia/atlas/eritrocitos.htm>. Consultado: 25/11/2013

ANEXOS

ANEXO A: ARBOL DE PROBLEMAS



ANEXO B: ARBOL DE OBJETIVOS

