

**IMPLEMENTACIÓN DE UN MODELO COMPUTACIONAL PARA CLASIFICACIÓN NORMAL –
DISPLÁSICA DE LAS CÉLULAS ESCAMOSAS DE CITOLOGÍAS CÉRVICO UTERINAS**

VÍCTOR EDUARDO MARTÍNEZ ABAUNZA

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2004**

**IMPLEMENTACIÓN DE UN MODELO COMPUTACIONAL PARA CLASIFICACIÓN NORMAL –
DISPLÁSICA DE LAS CÉLULAS ESCAMOSAS DE CITOLOGÍAS CÉRVICO UTERINAS.**

VÍCTOR EDUARDO MARTÍNEZ ABAUNZA

**Trabajo de Grado para Optar al Título de
Ingeniero de Sistemas**

**Director:
Alfonso Mendoza Castellanos
Bachelor of Science DEA.**

**Codirector:
Ernesto García Ayala. MD.
Especialista en Patología**

**Codirectora:
Claudia Janeth Uribe Pérez. MD.
Especialista en Patología**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER
FACULTAD DE INGENIERÍAS FÍSICO MECÁNICAS
ESCUELA DE INGENIERÍA DE SISTEMAS E INFORMÁTICA
BUCARAMANGA
2004**

A Chechis,
La ciencia te debe una explicación

AGRADECIMIENTOS

El autor desea agradecer a las siguientes personas.

A la Esperanza, por ser su hijo

Al Rocío, por refrescar cada día de la vida.

Al profesor Alfonso Mendoza, por todo lo aprendido.

Al doctor Ernesto García, por mantener la esperanza.

A la doctora Claudia Uribe, por creer en este proyecto.

A los amigos del GIIIB, porque se trabaja y se goza.

A los Campesinos de Tona, por defender la vida cuando la única ley es la muerte.

A los Campesinos de Cachirí, por mantener la esperanza en donde reina el miedo.

A quienes no se mencionan, porque de hacerlo escribiría un libro sólo para ellos.

A la familia, por permanecer.

A Dios, por todo lo anterior.

CONTENIDO

	pág.
INTRODUCCIÓN	19
1. PRESENTACIÓN DEL PROYECTO	20
1.1 OBJETIVOS	20
1.1.1 Objetivo General	20
1.1.2 Objetivos Específicos	20
1.2 JUSTIFICACIÓN	20
1.2.1 Definición Del Problema	20
1.2.2. Impacto Esperado	21
1.2.3 Viabilidad	21
1.2.4 Estado del Arte	22
2. MARCO TEÓRICO	23
2.1 CITOLOGÍA CÉRVICO UTERINA O EXTENDIDO DE PAPANICOLAOU	23
2.1.1 Técnica de Obtención de Muestras	23
2.2 MORFOLOGÍA DEL CÉRVIX	23
2.2.1 Anatomía Macroscópica	23

2.2.2 Histología	23
2.3 COMPONENTES CELULARES EN LA CITOLOGÍA CÉRVICO UTERINA	24
2.4 ESTRUCTURAS CELULARES QUE SON PARÁMETROS DE EVALUACIÓN	26
2.5 PROCESOS PATOLÓGICOS QUE AFECTAN LAS CÉLULAS	26
2.5.1 Efectos morfológicos sobre células que resultan en una actividad normal. Euplasia	26
2.5.2 Efectos morfológicos sobre células de procesos asociados con cáncer	26
2.5.3 Efectos morfológicos sobre células de procesos retrogresivos resultando en una disminución de la actividad. Retroplasia.	27
2.5.4 Efectos morfológicos sobre células de procesos progresivos resultando en actividad aumentada. Proplasia.	27
2.5.5 Efectos morfológicos sobre células de procesos cambiantes y mixtos	27
2.6. DIAGNÓSTICO EN CITOLOGÍA	27
2.7 TRATAMIENTO DIGITAL DE IMÁGENES	28
2.7.1 Fundamentos	28
2.7.2 Adquisición	28
2.7.3 Preprocesamiento	29
2.7.4 Segmentación	29
2.7.5 Descripción	31
2.7.6 Reconocimiento	31

3. CONSTRUCCIÓN DEL MODELO COMPUTACIONAL	32
3.1. METODOLOGÍA	32
3. 2 RECOLECCIÓN DE MUESTRAS	34
4. ADQUISICIÓN	35
5. PREPROCESAMIENTO	38
6. SEGMENTACIÓN	41
6.1 UMBRALIZACIÓN	41
6.1.1 Pun-Kapur	41
6.1.2 Iterativo	42
6.1.3 Triangular	44
6.1.4 Matlab	45
6.1.5 Manual	46
6.2 MORFOLOGÍA MATEMÁTICA	46
6.3 EXTRACCIÓN DE BORDES	47
7. DESCRIPCIÓN	49
8. RECONOCIMIENTO	51
8.1 RED NEURONAL	51
8.2 SISTEMA BASADO EN EL CONOCIMIENTO	53

8.2.1 Base de Conocimiento	53
8.2.2. Motor de Inferencia	53
8.2.3 Subsistema de Explicación	54
9. BASE DE DATOS	55
10. MODELO COMPUTACIONAL	57
11. RESULTADOS	59
12. FUNCIONAMIENTO DE LA HERRAMIENTA SOFTWARE	63
13. CONCLUSIONES	81
14. RECOMENDACIONES	82
BIBLIOGRAFÍA	83
ÍNDICE	87
ANEXO A. Requerimientos del sistema	89
ANEXO B. Instalación de la herramienta software	90
ANEXO C. Sistema Bethesda de clasificación	91
ANEXO D. Microscopio Adicional	92

LISTA DE TABLAS

	pág.
Tabla 1. Parámetros de las células basales	59
Tabla 2. Parámetros de las células parabasales	60
Tabla 3. Parámetros de las células intermedias	61
Tabla 4. Parámetros de las células superficiales	62

LISTA DE FIGURAS

	pág.
Figura 1. Lámina de citología cérvico uterina fijada y coloreada	23
Figura 2. Células escamosas superficiales	24
Figura 3. Células escamosas intermedias	25
Figura 4. Células escamosas parabasales	25
Figura 5. Diagrama de Casos de Uso	33
Figura 6. Diagrama del Modelo de análisis	33
Figura 7. Imagen tomada en el Laboratorio de Óptica y Tratamiento de Señales	35
Figura 8. Imagen tomada utilizando una cámara de vídeo Sony Handycam como digitalizador	36
Figura 9. Imagen tomada utilizando una tarjeta de adquisición MPEG TV Station	36
Figura 10. Imagen utilizada para el desarrollo del proyecto	37
Figura 11. Filtrado con Promedio, Circular y Gauss	39
Figura 12. Filtrado con Rotacional, Unsharp y Mediana	40
Figura 13. Resultado de la segmentación utilizando el método de Pun-Kapur	42
Figura 14. Resultado de la segmentación utilizando el método Iterativo	43
Figura 15. Localización de umbrales por el método Triangular	44
Figura 16. Resultado de la segmentación utilizando el método Triangular	45
Figura 17. Resultado de la segmentación utilizando el método de Matlab	45
Figura 18. Resultado de la segmentación utilizando el método Manual	46
Figura 19. Extracción de Bordes	48
Figura 20. Rejillas utilizadas para calcular la relación entre micra-píxel	50
Figura 21. Arquitectura de la red de clasificación	52
Figura 22. Clasificación de la Red Neuronal para 104 formas nucleares	52

Figura 23. Base de Datos	56
Figura 24. Esquema del Modelo Computacional	58
Figura 25. Ejecución de la herramienta en la línea de Comandos Matlab	63
Figura 26. Ventana Principal	64
Figura 27. Menú Archivo	64
Figura 28. Submenú Paciente	65
Figura 29. Ventana de Registro	65
Figura 30. Ventana de Identificación	66
Figura 31. Submenú Abrir	66
Figura 32. Submenú Configuración	67
Figura 33. Ventana de Configuración	67
Figura 34. Cargar Imagen	68
Figura 35. Datos de Adquisición	68
Figura 36. Visualización de la Imagen	69
Figura 37 Selección de la célula	69
Figura 38. Menú de Filtrado	70
Figura 39. Parámetros de filtrado	70
Figura 40. Menú Segmentar	71
Figura 41. Parámetros de segmentación	71
Figura 42. Plano de segmentación	72
Figura 43. Extracción de Bordes	72
Figura 44. Resultado de la segmentación	73
Figura 45. Segmentación Manual	73
Figura 46. Parámetros del núcleo	74
Figura 47. Medidas del núcleo	74
Figura 48. Color del núcleo	75
Figura 49. Descriptores de Fourier	75

Figura 50. Forma del núcleo	76
Figura 51. Clasificación del núcleo	76
Figura 52. Parámetros del citoplasma	77
Figura 53. Medidas del citoplasma	77
Figura 54. Relaciones Núcleo/Citoplasma	78
Figura 55. Guardar datos	78
Figura 56. Clasificación de células normales	79
Figura 57. Clasificación de células displásicas	79
Figura 58. Guardar datos de células clasificada	80
Figura 59. Menú Edición	80
Figura 60. Selección del Origen de Base de Datos (ODBC)	90
Figura 61. Imagen adquirida con el nuevo microscopio que integra las funciones de adquisición	92

LISTA DE ANEXOS

	pág.
ANEXO A. Requerimientos del sistema	89
ANEXO B. Instalación de la herramienta software	90
ANEXO C. Sistema Bethesda de clasificación	91
ANEXO D. Microscopio Adicional	92

GLOSARIO

ADQUISICIÓN. Etapa del procesamiento digital de imágenes que permite obtener una imagen digital de un objeto, a partir de una escena.

BASOFÍLICO. Que presenta alto grado de coloración.

BINARIZACIÓN. Procedimiento para convertir una imagen que se encuentra en el formato de color verdadero, o en escala de grises, a un formato de dos colores (Blanco y Negro)

CÉLULAS ESCAMOSAS. Células encontradas en una muestra de citología cérvico uterina. De acuerdo con su etapa de maduración se clasifican en: basales, parabasales, intermedias y superficiales.

CÉRVIX. Llamado también cuello uterino. Porción del útero que se une con la vagina.

CITOLOGÍA CÉRVICO UTERINA. Prueba de tamizaje realizada a las mujeres, que consiste en obtener una muestra adecuada de la unión exoendocervical y del endocérvix.

CIANOFÍLICO. Que tiene afinidad por los colorantes azules o verdes.

CROMATINA. Sustancia protoplásmica, coloreable, del núcleo de la célula

DESECACIÓN CELULAR. Degeneración celular.

DESCRIPCIÓN. Etapa del procesamiento digital de imágenes que permite describir las características de un objeto.

DISCRETIZACIÓN. Procedimiento para asignar un valor discreto a una función continua.

DISPLASIA. Anomalía en el desarrollo de una célula.

DIGITALIZADOR. Dispositivo electrónico que permite convertir una señal análoga a un formato digital, por medio de la discretización.

DESCRIPTORES DE FOURIER. Coeficientes de una serie de Fourier.

FILTRADO. Etapa del procesamiento digital de imágenes que permite mejorar las características de una imagen digital.

FROTIS CÉRVICO UTERINA. Véase citología cérvico uterina.

HIPERCROMASIA. Abundancia de cromatina en una célula. En imágenes digitales se manifiesta con bajos niveles de color.

HIPERCROMATISMO. Véase hipercromasia.

HISTOGRAMA. Gráfico utilizado para la representación de la cada uno de los niveles de color.

LESIÓN INTRAEPITELIAL ESCAMOSA. Término acuñado por el sistema Bethesda para referirse a un cambio en el tejido del cérvix, a veces se usa como otra palabra para tumor

NEOPLASIA INTRAEPITELIAL CERVICAL. Crecimiento anormal que se origina en el cérvix, causado por una sola célula alterada.

NIVEL DE COLOR. Valor asociado con la cantidad de luz detectada por un dispositivo de adquisición. Depende del número de bits empleado, para 8 bits se encuentra entre el intervalo [0, 255].

ORANGOFÍLICO. Que presenta una coloración anaranjada.

PICNOSIS. Condensación celular en la que el protoplasma se hace más denso, propia de una célula que se muere

PLEOMORFISMO NUCLEAR. Variaciones morfológicas nucleares.

PREPROCESAMIENTO. Véase filtrado.

RADIO NÚCLEO/CITOPLASMA. Relación entre el tamaño del núcleo y el citoplasma.

RECONOCIMIENTO: Etapa del procesamiento digital de imágenes que permite la clasificación de los objetos presentes en una escena.

RED NEURONAL. Conjunto de operadores interconectados en forma de red, que son utilizados para resolver problemas de optimización y clasificación, mediante funciones de activación.

SEGMENTACIÓN. Etapa del procesamiento digital de imágenes que permite la separación de los objetos presentes en una imagen.

SISTEMA BASADO EN EL CONOCIMIENTO. Sistemas informáticos que involucran conocimiento propio de especialistas.

TRANSFORMADA DE FOURIER. Función matemática que permite obtener las componentes de frecuencia de una señal.

UMBRALIZACIÓN. Técnica de segmentación que consiste en binarizar una imagen, tomando como referencia un nivel del histograma.

VPH. Virus de Papiloma Humano, relacionado ampliamente con el cáncer de cérvix.

RESUMEN

Título: IMPLEMENTACIÓN DE UN MODELO COMPUTACIONAL PARA CLASIFICACIÓN NORMAL – DISPLÁSICA DE LAS CÉLULAS ESCAMOSAS DE CITOLOGÍAS CÉRVICO UTERINAS.*

Autor: Víctor Eduardo Martínez Abaunza.**

Palabra Claves: Tratamiento Digital de Imágenes, Morfología Matemática, Redes Neuronales, Sistema Basado en el Conocimiento, Citología Cérvico Uterina, Células Escamosas, Lesión Intraepitelial, Cáncer de Cérvix.

Descripción: El presente proyecto continúa el trabajo realizado por el Grupo de Investigación en Ingeniería Biomédica (GIIB), en la línea de tratamiento de imágenes médicas, orientada a la construcción de aplicaciones para la detección precoz de cáncer de cuello uterino. Esta enfermedad ha venido mostrando las mayores tasas de incidencia en los países en vías de desarrollo, como Colombia, mientras que en la población europea se ha notado un marcado descenso, atribuido principalmente a la realización de campañas preventivas y al uso de nuevas tecnologías de diagnóstico.

El desarrollo de herramientas informáticas que permiten la identificación de células escamosas superficiales ha sido posible, sin embargo el alto costo que representan, hace inaccesible a este tipo de tecnología en una región como Santander. El principal objetivo del GIIB en esta línea de investigación es desarrollar una herramienta que pueda ser utilizada en la región, y que apoye la clasificación de las células, hallazgos y lesiones encontradas en una citología cérvico uterina.

Han sido empleadas diversas técnicas de tratamiento de imágenes, la metodología de Proceso Unificado de Desarrollo de Software, junto con herramientas de inteligencia artificial, que permitieron clasificar las células. Se determinaron los parámetros correspondientes a los criterios de normalidad existentes, que logran la identificación de displasias. Se recomienda la caracterización de cada una de las lesiones intraepiteliales relacionadas en el Sistema Bethesda y el uso de arquitectura en paralelo para la construcción de algoritmos más robustos.

* Trabajo de Grado

** Facultad de Ingenierías Físico Mecánicas. Escuela de Ingeniería de Sistemas. Director: Alfonso Mendoza Castellanos. Codirectores: Ernesto García Ayala, Claudia Janeth Uribe Pérez.

ABSTRACT

Title: COMPUTATIONAL MODEL IMPLEMENTATION FOR ABNORMAL-NORMAL CLASSIFICATION OF THE SQUAMOUS CELLS OF CERVICAL SMEAR CYTOLOGIES.

Author: Víctor Eduardo Martínez Abaunza.^{**}

Keywords: Digital Image Processing, Mathematical Morphology, Neural Networks, Knowledge-Based System, Cervical Smear Cytology, Squamous Cells, Intraepithelial Lesion, Cervical Cancer.

Description: This project continues the work carried out by the Biomedical Engineering Research Group (GIIB), in the line of medical images processing, guided to the construction of applications for the early detection of cervical cancer. For this illness, the highest incidence rates have been reported from the underdeveloped countries, as Colombia, while in the European population a marked descent has been noticed, attributed mainly to the realization of preventive campaigns and the use of new diagnosis technologies.

The development of computer tools has been possible, it allows the identification of superficial squamous cells, however they represent a high cost, and it makes inaccessible this kind of technology in a region like Santander. The main objective for the GIIB in this investigation line is to develop a tool that can be used in the region, and that it supports the classification of the cells, discoveries and lesions found in cervical smear cytology.

Diverse techniques of images processing have been employed, the methodology of Unified Process of Development of Software, together with tools of artificial intelligence that allowed to classify the cells. The parameters corresponding to the existent approaches of normality were determined to achieve the dysphasia identification. The characterization is recommended for each one of the intraepithelial lesions according to the Bethesda System and the use of parallel architecture for the construction of more robust algorithms.

* Grade Work

** Faculty of Engineering Physique-Mechanics. School of Systems Engineering and Informatics. Director: Alfonso Mendoza Castellanos. Co-directors: Ernesto García Ayala, Claudia Janeth Uribe Pérez.

INTRODUCCIÓN

El presente proyecto se constituye en una herramienta que permite la medición y extracción de características de células de citologías cérvico uterinas, para su correspondiente clasificación entre los parámetros de normalidad existentes; esto contribuye a la detección temprana de anormalidades o displasias, que posibilitarían la evolución del cáncer de cuello uterino.

El cáncer de cérvix es el segundo tipo de cáncer más frecuente en la población mundial femenina, diagnosticándose cerca del 80% de los casos en países no desarrollados. Regionalmente se presentaron 1897 casos de displasias leves, 736 displasias moderadas, 169 carcinomas in situ y 262 de cáncer invasor; entre los años 2000 y 2001, de acuerdo con el registro poblacional de cáncer del área metropolitana de Bucaramanga.

La citología cérvico uterina es el examen más utilizado a nivel mundial, como prueba de tamizaje para la detección precoz de esta enfermedad, el propósito es obtener una muestra suficiente de las células que conforman el epitelio escamoso y el glandular, para detectar los cambios que sean causados por una enfermedad, especialmente la neoplasia maligna; sin embargo, en países como Colombia la cobertura de este método para mujeres sexualmente activas es muy baja, entre el 20% y el 30%.

De acuerdo con estudios internacionales, se han obtenido rangos de sensibilidad entre un 11%-99% y de especificidad de 14%-97%. La International Agency of Research on Cancer (IARC) ha reportado un 40% de falsos positivos y falsos negativos, atribuidos en una tercera parte a errores en la lectura e interpretación de las muestras.

En un estado normal las estructuras nucleares clave son: núcleo, nucleolo y cromocentro redondeados, un patrón de cromatina y paracromatina uniforme con un grosor regular del anillo cromatínico asociado internamente a la membrana nuclear. Los cambios debido al cáncer están relacionados a alteraciones de estas estructuras como: membrana nuclear irregular, cromatina condensada en grumos gruesos, paracromatina más fácilmente apreciable, nucleolo más grande, irregular y de forma angulada, presencia de mitosis anormales, multinucleación con núcleos que exhiben pleomorfismo nuclear e incremento en el radio núcleo/citoplasma.

El objetivo de este proyecto es desarrollar un software mediante un modelo computacional que apoye la clasificación de células escamosas en citologías cérvico uterinas dentro de los parámetros de normalidad existentes. La metodología empleada para la construcción de la herramienta ha sido el Proceso Unificado de Desarrollo de Software. Las imágenes han sido digitalizadas mediante una cámara de vídeo acoplada a un microscopio y conectada a un computador.

Con la investigación se busca contribuir en la clasificación de una citología cérvico uterina. Esto permitirá a los patólogos reconocer con mayor precisión posibles anormalidades celulares en las pacientes, y facilitará el diagnóstico precoz de cáncer de cérvix, disminuyendo el porcentaje de falsos positivos o negativos para minimizar el perfil inmunológico de enfermedad y mortalidad.

1. PRESENTACIÓN DEL PROYECTO

1.1. OBJETIVOS

1.1.1 Objetivo General. Desarrollar un software mediante un modelo computacional que apoye la clasificación de células escamosas en citologías cérvico uterinas dentro de los parámetros de normalidad existentes.

1.1.2 Objetivos Específicos.

- Establecer los criterios correspondientes a cada uno de los parámetros a evaluar en las células escamosas normales en cada uno de sus estadios como son: células superficiales, intermedias, parabasales y basales.
- Establecer los criterios cualificables y cuantificables en cada uno de los parámetros de las células escamosas correspondientes a células displásicas.
- Identificar las características de las células normales o, que presenten cambios por displasias.
- Desarrollar los algoritmos necesarios para la identificación de los parámetros de clasificación de células escamosas normales.
- Desarrollar los algoritmos para la identificación los criterios de cada uno de los parámetros de las células escamosas que permitan clasificarlas como células displásicas.
- Construir un software que permita clasificar las células de acuerdo con los parámetros extraídos de una lámina de citología cérvico uterina

1.2 JUSTIFICACIÓN

1.2.1 Definición Del Problema. Entre la población mundial femenina, el cáncer de cuello uterino ocupa el segundo lugar de frecuencia, diagnosticándose cerca del 80% de los casos en los países en vías de desarrollo. Esta enfermedad presenta en Colombia una tasa de incidencia reportada mundialmente de 32,86 casos por 100 mil habitantes, siendo en el área metropolitana de Bucaramanga de 35,7; lo que nos dice acerca de la magnitud del problema en nuestro medido. En el período comprendido entre los años de 1983 y 1991, la tasa de mortalidad en Colombia debido a esta patología fue de 6,5 por 100 mil habitantes. Actualmente, según Profamilia y la Liga contra el Cáncer “es la primera causa de muerte de cáncer en la mujer colombiana y nuestro país posee una de las más altas tasas de esta enfermedad en el mundo”.

Esta patología es detectable precozmente mediante un examen conocido como citología cérvico uterina. En varios países en vías de desarrollo como Colombia la cobertura que tienen las mujeres para acceder a esta prueba es aún muy baja: sólo entre el 20% y el 30% de las mujeres sexualmente activas pueden optar al mencionado examen. En la población europea la localización tumoral ha mostrado un marcado descenso, atribuido en parte a la extensión de la realización periódica de esta prueba de tamizaje.

En diversos estudios reportados a nivel mundial existe una importante diferencia en la agudeza diagnóstica de este examen, mostrando una gran variación en los resultados de sensibilidad y especificidad. Se han obtenido con diseños y tiempos de seguimiento diferentes rangos de sensibilidad entre un 11% - 99%, y de especificidad de 14% - 97%. La Agencia Internacional de Investigación del Cáncer (IARC), analizando la experiencia de 10 programas realizados en Europa y Canadá, estimó tasas de falsos positivos y falsos negativos próximas al 40%. Este alto porcentaje preocupa por la posibilidad de sobre diagnosticar o, en su defecto sospechar lesiones neoplásicas que no existen, siendo peor aún aquellos casos que existiendo la lesión preneoplásica o neoplásica, la sospecha diagnóstica se escapa de la detección. Por esta razón sólo con 2 ó 3 extendidos normales realmente se certifica la ausencia de cáncer y, ante la duda, se hace necesario recurrir a otro tipo de pruebas de mayor costo.

Hay factores que afectan la efectividad del tamizaje y condicionan también una baja eficiencia de estos programas en nuestro medio, como son las limitaciones en la validez de la propia citología cérvico uterina. Existe cierto grado de acuerdo en atribuir una tercera parte de los resultados falsos negativos a errores de lectura e interpretación de las muestras en el laboratorio y las dos terceras partes restantes a fallos en la toma de la muestra (no se logra recoger células anormales existentes).

La calidad de un frotis cérvico uterino depende fundamentalmente de la habilidad y entrenamiento específico del que hace la toma. La formación adecuada y la retroalimentación sobre la calidad de las tomas, con citotecnólogos y patólogos, se han mostrado muy eficaces para mejorar la calidad de los frotis; pero se han encontrado evidencias que, por más adecuado entrenamiento, la subjetividad del observador interfiere en el diagnóstico o, muchas veces, cambios sutiles que se escapan al ojo humano, dando origen a la no adecuada clasificación de la lesión.

1.2.2 Impacto Esperado. Se espera que las instituciones de salud y los laboratorios de la región puedan utilizar la información suministrada por el proyecto para diseñar planes de control y prevención de cáncer de cuello uterino y elaborar tratamientos efectivos de acuerdo al riesgo, consiguiendo de este modo que las instituciones hagan extensivas sus funciones al controlar y mejorar la calidad de vida de los pacientes.

A nivel de investigación, el modelo ayuda a sentar bases para futuros proyectos que relacionen conocimientos de ingeniería y de medicina, con el fin de aportar soluciones a problemáticas de connotación social, primero a nivel regional, y en un futuro a nivel nacional.

1.2.3 Viabilidad. Existen ventajas en el aspecto humano y tecnológico, como son las necesidades actuales del Departamento de Patología para apoyar el desarrollo de herramientas que faciliten la investigación en curso, para dar respuesta a las necesidades académicas y asistenciales propias de su profesión, también cabe resaltar la disposición que se tiene para involucrarse en el desarrollo de proyectos que beneficien el aprendizaje individual y cooperativo.

Algunos de los equipos necesarios para la ejecución del proyecto se encuentran en el laboratorio del Grupo de Investigación de Ingeniería Biomédica y otros, son aportados por los respectivos codirectores del proyecto. Se espera que mediante gestiones con entidades que financian la investigación, como en el caso de Colciencias, se logre la consecución de equipos con mayor capacidad de trabajo.

1.2.4 Estado del Arte. La utilización de tratamiento digital de imágenes ha sido demostrada en la construcción de sistemas de citometría. El Grupo de Investigación en Visiónica (nombre anterior del grupo) desarrolló un proyecto que determinaba el tamaño nuclear y la relación núcleo-citoplasma para las células superficiales, intermedias y parabasales normales de una muestra de citología cérvico uterina. A nivel nacional se encuentran trabajos como el Sistema para la segmentación de células en imágenes al microscopio, del epitelio poliestratificado vaginal; desarrollado en la Universidad Nacional de Colombia (Sede Manizales). Internacionalmente se encuentran productos comerciales de alto costo, diseñados para detectar cáncer de cérvix como el Papnet, el Inpath, el ThinPrep Pap System y el TriPath Imaging. Estos sistemas complejos incluyen los dispositivos, los accesorios para la adquisición de las láminas y su posterior análisis en equipos especializados.

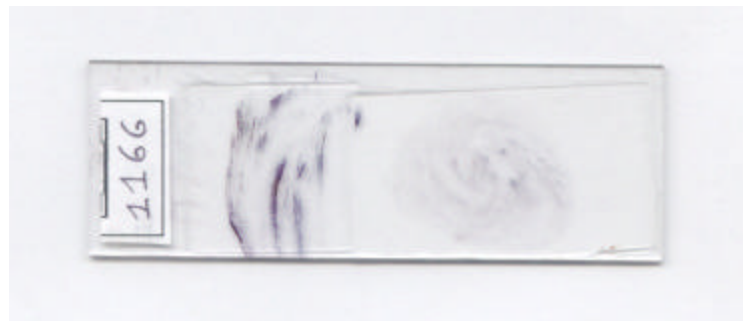
2. MARCO TEÓRICO

2.1 CITOLOGÍA CÉRVICO UTERINA O EXTENDIDO DE PAPANICOLAOU

Es el examen más usado, como prueba de tamizaje para la detección precoz del cáncer de cérvix, con efectividad demostrada mediante estudios epidemiológicos realizados en todo el mundo, disminuyendo la incidencia y mortalidad de cáncer hasta un 60% en las poblaciones sometidas a tamizaje sistemático. Este no es un examen diagnóstico, siendo necesario que los hallazgos anormales encontrados sean confirmados por otros medios más específicos como la Colposcopia y la biopsia cervical.

2.1.1 Técnica de Obtención de Muestras La citología se realiza anualmente a mujeres mayores de 18 años o antes, si tiene vida sexual activa. Se hace un raspado de la superficie mucosa del exocérnix y del fondo de saco vaginal, se introduce en el canal endocervical el cepillo de citología, para obtener adecuada muestra de la unión exoendocervical y del endocérnix. Todas estas muestras se extienden en el área de la lámina portaobjetos correspondiente según criterios de toma de muestra. La muestra debe ser fijada y secada. Finalmente se remite al laboratorio para su procesamiento en donde se colorea. La siguiente figura presenta una lámina de citología.

Figura 1. Lámina de citología cérvico uterina fijada y coloreada.



2.2 MORFOLOGÍA DEL CÉRVIX

2.2.1 Anatomía Macroscópica. El cérvix es la porción del útero que protruye en la vagina, está constituido por dos porciones, el endocérnix y el exocérnix, este último es redondo, pequeño y recubierto por una mucosa rosada, húmeda y brillante. En la parte central existe un orificio denominado orificio externo que da origen al canal endocervical, el cual está recubierto por una mucosa con múltiples plegamientos en su recorrido al interior del útero. En el orificio cervical externo se unen la porción exocervical y endocervical, sitio donde se alberga gran parte de las lesiones preneoplásicas y neoplásicas del cérvix.

2.2.2 Histología. La mucosa del exocérnix tiene un estroma fibromuscular, y un revestimiento de epitelio escamoso húmedo, posee un estrato basal, de células con núcleos alargados,

perpendiculares a la membrana basal y escaso citoplasma; un estrato parabasal, de células poligonales y mayor cantidad de citoplasma; un estrato intermedio, con células de núcleos ovalados y abundante citoplasma y un estrato superficial, de células dispuestas horizontalmente y núcleos pequeños, hiper cromáticos. Estos cambios son de maduración celular.

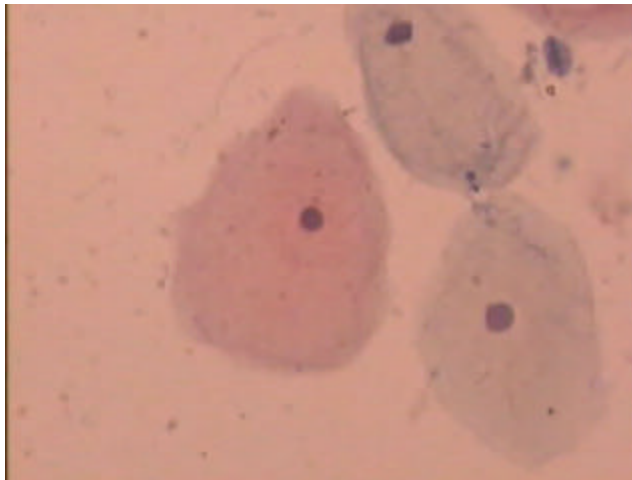
La mucosa del endocérvi x tiene un estroma conjuntivo revestido por epitelio cilíndrico simple, mucosecretor, con pliegues que forman glándulas.

2.3 COMPONENTES CELULARES EN LA CITOLOGÍA CÉRVICO UTERINA

El diagnóstico citológico está basado en el análisis de las características morfológicas de las células exfoliadas del cérvix, que dependerán de la edad de la paciente, fase del ciclo menstrual e influencia de factores exógenos tales como inflamación, terapias, o estados patológicos. Las células epiteliales más numerosas son las células escamosas que se desprenden tanto del exocérvi x como de la pared vaginal. De este tipo de células se pueden identificar tres clases:

- **Células escamosas superficiales.** Corresponden a la capa más externa del epitelio. Miden entre 40–50 μm . Son poliédricas, planas, con citoplasma claro eosinofílico con bordes bien definidos. Pueden contener pequeños gránulos basofílicos de queratohialina. El núcleo es pequeño, redondo, picnótico homogéneo, mide entre 3 y 5 μm . No se ve la red cromatínica. Esta característica permite diferenciarla de una intermedia. Algunos autores sugieren un criterio de núcleo métrico: Todo núcleo con diámetro menor de 6 μm , debería llamarse núcleo picnótico.

Figura 2. Células escamosas superficiales



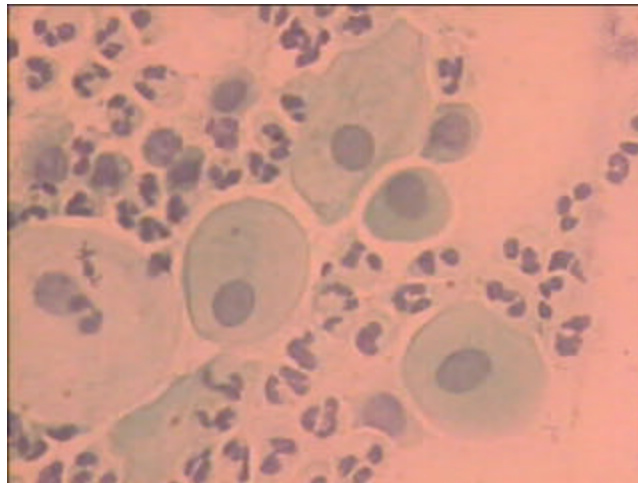
- **Células escamosas intermedias.** Células grandes con un tamaño aproximado de 50 μm poligonales, con citoplasma transparente, usualmente cianofílico. Su núcleo ovalado tiene un diámetro entre 8–10 μm , está centralmente ubicado con un patrón de cromatina fino homogéneo, y su membrana nuclear está claramente definida.

Figura 3. Células escamosas intermedias



- **Células escamosas parabasales.** Son células pequeñas con medidas entre 12–30 μm . Su citoplasma es opaco, basofílico y puede mostrar vacuolización. Tiene un núcleo redondo vesicular, con un radio núcleo/citoplasma alto.

Figura 4. Células escamosas parabasales



Se pueden encontrar otros dos tipos celulares que son: células queratinizadas y células basales. Las primeras son grandes, aplanadas, con citoplasma orangofílico y generalmente anucleadas. Las segundas son uniformes en tamaño, redondas, con citoplasma basofílico y opaco, con núcleo hipercromático y alto radio núcleo/citoplasma. También están representadas las células endocervicales, son cilíndricas altas, con abundante citoplasma claro, ligeramente cianofílico y finamente vacuolado. Se organizan en grupos con apariencia de “panal de abejas”. Los núcleos basalmente localizados son ovales, de tamaño variable, con cromatina fina granular difusamente distribuida. Tienen uno o dos pequeños nucleolos.

2.4 ESTRUCTURAS CELULARES QUE SON PARÁMETROS DE EVALUACIÓN

Para la evaluación tanto de actividad biológica y funcional de las células se deben tener en cuenta los siguientes parámetros:

- Núcleo.
- Nucleolo.
- Cromatina.
- Borde cromatínico y membrana nuclear.
- Red de cromatina.
- Paracromatina.
- Mitosis.
- Multinucleación.
- Radio núcleo/citoplasma.
- Coloración del citoplasma

2.5 PROCESOS PATOLÓGICOS QUE AFECTAN LAS CÉLULAS

2.5.1 Efectos morfológicos sobre células que resultan en una actividad normal. Euplasia. En este estado las estructuras clave vistas al microscopio de luz son: núcleo, nucleolo, cromocentro redondeados, y un patrón de cromatina y paracromatina uniforme con un grosor regular del anillo cromatínico asociado internamente a la membrana nuclear.

2.5.2 Efectos morfológicos sobre células de procesos asociados con cáncer. Algunos cambios morfológicos claves son: membrana nuclear de forma irregular. Hay ángulos hacia adentro y hacia afuera del núcleo. La cromatina se condensa en grumos gruesos que se distribuyen irregularmente, principalmente cerca del anillo cromatínico. La paracromatina es fácilmente apreciable, ocasionada por la mayor cantidad de zonas claras. El nucleolo es grande, irregular y de forma angulada. Con mitosis anormales. Multinucleación con núcleos pleomórficos. Incremento del radio núcleo/citoplasma.

2.5.3 Efectos morfológicos sobre células de procesos retrogresivos resultando en una disminución de la actividad. Retroplasia. Hay edema celular dado por degeneración hidrópica o crenación, cambio en la coloración de tinción citoplasmática, núcleo arrugado, colapsado, disminución de la basofilia nuclear, cromatina agregada en grumos gruesos con apariencia de gelatina.

2.5.4 Efectos morfológicos sobre células de procesos progresivos resultando en actividad aumentada. Proplasia. La membrana nuclear muestra ondulaciones, hiper cromasia nuclear, mayor aclaramiento de la paracromatina distribuida regularmente, nucleolos inmensos y en ocasiones múltiples, con esquinas redondeadas, mitosis normales, multinucleación frecuente con regularidad nuclear, y una relación núcleo/citoplasma aumentada.

2.5.5 Efectos morfológicos sobre células de procesos cambiantes y mixtos. Pueden mezclarse los eventos anteriormente descritos.

2.6. DIAGNÓSTICO EN CITOLOGÍA

Para analizarla se utiliza la clasificación del Sistema Bethesda, esta tiene como complemento informar sobre lo adecuado de la muestra, y un diagnóstico descriptivo, con reconocimiento de patrones morfológicos permitiendo discernir entre cambios celulares benignos por infecciones o reparativos y anormalidades de las células epiteliales. En esta última categoría se debe clarificar si los cambios observados pertenecen a condiciones premalignas o malignas y en caso dudoso se establece la categoría de *SIGNIFICANCIA INDETERMINADA*.

En el diagnóstico descriptivo hay áreas que generan gran controversia particularmente las denominadas células escamosas o glandulares atípicas de significancia indeterminada. (ASCUS – AGUS). Se usa cuando los cambios encontrados exceden a los ocasionados por proceso reactivo benigno, pero son insuficientes para un diagnóstico definitivo de lesión.

La *LESIÓN INTRAEPITELIAL ESCAMOSA (LISL)* se divide en lesión de bajo y alto grado. Hay características morfológicas relacionadas con cada una de estas:

- En el diagnóstico por VPH hay cambios nucleares que incluyen agrandamiento nuclear con hiper cromasia o picnosis, con cromatina y contorno nuclear arrugado, en el citoplasma hay una cavidad perinuclear bien definida con reforzamiento del citoplasma adyacente.
- Las células de la displasia leve pertenecen a la categoría de bajo grado, muestran un citoplasma traslúcido, con bordes angulados bien definidos, recordando células superficiales o intermedias pero con menor cantidad de citoplasma y un núcleo más grande que ocupa menos de un tercio del área total de la célula. La cromatina nuclear es finamente granular, irregularmente distribuida con ligera hiper cromasia.
- En las de alto grado se incluyen la displasia moderada, severa y el carcinoma in situ. En la displasia moderada las células tienen un tamaño variable, la mayoría de ellas son ovales,

pero pueden encontrarse células fusiformes bizarras. Predominan los citoplasmas eosinofílicos con núcleos agrandados de forma irregular. Hay mayor aumento en el radio núcleo/citoplasma. El nucleolo usualmente está ausente. En la displasia severa y carcinoma in situ el tamaño de las células se compara con las células parabasales, mostrando gran núcleo que ocupa dos tercios del área total de la célula con hiper cromasia, cromatina en grumos gruesos distribuidos irregularmente y membrana nuclear con irregularidades.

- Cuando la lesión tumoral sobrepasa la membrana basal en donde se apoya el epitelio escamoso, es considerada como invasora, siendo muy agresiva. Existen parámetros definidos para el diagnóstico de carcinoma invasor como son: menor tamaño de las células escamosas tumorales, comparadas con las células normales con formas redondas, ovales y alargadas conteniendo fibrillas intracitoplasmáticas. Las células se disponen en sincitios con bordes no definidos, con polaridad alterada. El núcleo tiene pleomorfismo, con un alto radio núcleo/citoplasma.

2.7 TRATAMIENTO DIGITAL DE IMÁGENES

2.7.1 Fundamentos. El tratamiento de imágenes surge como un área de la computación con el fin de resolver los problemas de envíos intercontinentales de información gráfica, junto a las dificultades que presentaban las agencias aeroespaciales con la información que recibían de las sondas espaciales y satélites artificiales respecto a la topografía y condiciones ambientales de los planetas; también se han desarrollado significativos avances en aplicaciones industriales, donde la integridad y la seguridad de las personas se ve amenazada. Desde la década de los ochenta se incrementó su desarrollo con la construcción de equipos con mayor capacidad de cómputo y con la implementación de modelos matemáticos que permitían extraer las características de una imagen. Su uso se popularizó en medicina con la aparición de dispositivos que apoyaban el diagnóstico de enfermedades mediante imágenes de rayos X, ecografías, tomografías axiales computarizadas e imágenes microscópicas de alta resolución. En el campo de la microscopía y la citometría, se han desarrollado sistemas que permiten ejecutar este tipo diagnóstico con la utilización de imágenes digitales.

Básicamente el tratamiento de imágenes se ocupa de las siguientes etapas: Adquisición, preprocesamiento, segmentación, descripción y reconocimiento.

2.7.2 Adquisición. Mediante un dispositivo sensible a una determinada banda del espectro electromagnético (rayos X, ultravioleta, visible o infrarrojo) se produce una señal eléctrica de salida proporcional al nivel de energía detectado y mediante otro dispositivo llamado digitalizador se convierte esta señal en un formato digital, obteniéndose de esta manera la imagen.

El dispositivo convertidor análogo digital debe tener en cuenta dos criterios muy importantes al momento de realizar la discretización de las intensidades: *el muestreo*, relacionado con la dimensión espacial e indica el tamaño del elemento de la imagen (píxel), el segundo criterio hace referencia a la amplitud de la señal, es decir el nivel de intensidad del píxel y es conocido como *la cuantificación*.

Para el caso de la microscopía, se ha generalizado el uso de una cámara de vídeo digital acoplada al microscopio.

2.7.3 Preprocesamiento. En el proceso de adquisición la imagen que se genera no es totalmente confiable, dado que se pueden presentar obstáculos; fallos en la iluminación del microscopio, partículas de polvo en los objetivos, insuficiente coloración de las muestras, etc.

El filtrado de una imagen permite corregir algunos de estos problemas y ha sido agrupado conforme a la técnica utilizada. Cada filtro es implementado de acuerdo a la imagen que se desea obtener.

- **Mejoramiento por procesamiento de punto.** Se basa en la mejora de las intensidades de los niveles de grises considerando los píxeles individualmente. Entre los más utilizados están: Negativos de imágenes, Aumento del Contraste, Compresión del Rango Dinámico y Procesamiento del Histograma (Ecuación y Especificación).
- **Filtrado en el Dominio Espacial.** El filtrado espacial de una imagen es implementado mediante pequeñas matrices cuadradas, conocidas como máscaras, núcleos o ventanas, las cuales, aplicando el teorema de la convolución con la imagen a filtrar, produce como salida otra imagen donde ha sido modificado el contenido en escala de grises. A su vez han sido clasificados como Paso-Bajo y Paso-Alto, los primeros permiten homogenizar las superficies de la imagen mediante el difuminado, los segundos son útiles en la extracción de bordes. Entre los filtros espaciales más comunes encontramos: El Filtro Mediana, El Filtro Promedio y los Filtros Diferenciales (Operadores de Roberts, Prewitt y Sobel).
- **Filtrado en el Dominio de la Frecuencia.** Para implementar esta serie de filtros se calcula la Transformada de Fourier de una imagen, con el fin de obtener las componentes frecuenciales de la imagen para multiplicarlas por una función de transferencia de un filtro, seguidamente aplicamos la transformada Inversa de Fourier para producir la imagen filtrada. Estos se encuentran clasificados como Paso-Bajo y Paso-Alto, los primeros atenúan las componentes de alta frecuencia y los segundos las bajas. La función de transferencia más utilizada es conocida como Filtro de Butterworth tanto para frecuencias altas como para frecuencias bajas.

2.7.4 Segmentación. Esta etapa nos permite descomponer en sus partes constituyentes los objetos que se encuentran en la imagen (Núcleo, Citoplasma, etc.). De esta etapa depende el éxito o fracaso del análisis de la imagen. Se utilizan diversos algoritmos con el fin de detectar discontinuidades, líneas y bordes. Entre estos algoritmos se encuentran el Laplaciano de una función bidimensional, la Transformada de Hough y el procesamiento global por medio de teoría de grafos.

En la segmentación de la imagen se utiliza también la técnica de umbralización, la cual consiste en convertir una imagen de escala de grises a una imagen binaria (Blanco y Negro puros) utilizando uno de los niveles de gris como umbral. Se acepta que los objetos son aquellas regiones que dan como resultado el color blanco mientras que la región negra pertenece al fondo de la imagen.

Una herramienta robusta para este proceso es la morfología matemática, que provee un enfoque para el tratamiento digital de imágenes basado en la forma de los objetos. Ésta sirve para simplificar los datos de una imagen preservando sus características esenciales de forma, y ha sido aplicada ampliamente en el tratamiento digital de imágenes para microscopía. Surgió en Europa en la década de los sesenta como un método de teoría de conjuntos para el análisis de imágenes motivado principalmente por problemas en la microscopía cuantitativa. Sus herramientas matemáticas están relacionadas con la geometría integral y la estereología. El lenguaje utilizado es similar al utilizado en la teoría de conjuntos, que en morfología matemática representan las formas, las cuales se manifiestan sobre imágenes binarias imágenes o en escala de grises.

Sea A la imagen segmentada y B el elemento estructurante, las operaciones básicas en morfología matemática son:

- **Erosión.** Definida como:

$$A \odot B = \{x \mid (\hat{B})_x \subseteq A\}$$

Todos los elementos de B reflejado se encuentran contenidos en A

- **Dilatación.** Definida como:

$$A \oplus B = \{x \mid (\hat{B})_x \cap A \neq \emptyset\}$$

Al menos uno de los elementos de B reflejado se encuentra contenido en A

La utilización de estas operaciones permite construcción de las siguientes rutinas:

- **Apertura.** Definida como:

$$A \circ B = A \oplus (A \odot B)$$

Esto es una erosión, seguida de una dilatación. Permite separar objetos traslapados o unidos.

- **Cierre.** Definida como:

$$A \bullet B = A \odot (A \oplus B)$$

Esto es una dilatación, seguida de una erosión. Permite llenar huecos en los objetos.

2.7.5 Descripción. Una vez realizada la segmentación de la imagen se continúa a la etapa de descripción, consistente en describir las características de los objetos presentes en la imagen respecto a sus características externas (contorno) y sus características internas (píxeles que comprenden la región). Esta etapa permite diferenciar cada una de las características del núcleo (tamaño, forma, cromatina) y del citoplasma (tamaño, coloración, relación núcleo/citoplasma). Entre los esquemas de representación para contornos encontramos:

- **Códigos de Cadena.** Utilizados para representar contornos por medio de una sucesión conexa de segmentos de longitud y dirección específica. Se podría generar un código de cadena siguiendo las manecillas del reloj.
- **Aproximaciones Poligonales.** El objetivo es lograr una aproximación poligonal de un contorno con el menor número de lados posibles.
- **Firmas.** Consiste en representar con una función unidimensional un contorno.
- **Descriptores Geométricos.** Se utilizan medidas geométricas como; longitud, diámetro, área, centroide, etc.
- **Descriptores de Fourier.** Son los coeficientes de la Transformada Discreta de Fourier de una serie, donde las coordenadas de los píxeles son utilizados como una serie de números complejos.

Referente a las regiones son utilizados descriptores topológicos que determinan por ejemplo el número de huecos, o la cantidad de componentes conexas. La textura también es utilizada como descriptor de región debido a que proporciona medidas de propiedades como suavizado, rugosidad y regularidad. Se utilizan métodos estadísticos, estructurales y espectrales para proporcionar características de texturas.

2.7.6 Reconocimiento. En esta etapa se asigna una etiqueta a los objetos de acuerdo con sus descriptores permitiendo reconocer e identificar patrones de los objetos para ser comparados con otros patrones y poder así reconocer e identificar el objeto de la imagen que se ha analizado. Mediante clasificadores como los de mínima distancia, de correspondencia por relación y los clasificadores bayesianos, el proceso que se la ha aplicado a la imagen permite establecer criterios de decisión.

Una herramienta que está implementándose en el reconocimiento de objetos es la Red Neuronal; definida como un conjunto de elementos no lineales (neuronas), organizados en una red, y se piensa que se asemejan a la forma en que están interconectadas las neuronas del cerebro. Dada la expectativa generada con el descubrimiento de lo que parecería un modelo de aprendizaje que permitiría un gran desarrollo en diferentes áreas de la computación, hoy se reconoce que en algunos casos este modelo no es conveniente, en el área de tratamiento de imágenes está permitiendo la posibilidad de reconocer patrones mediante técnicas híbridas entre los modelos lineales y las redes neuronales. Este tipo de técnica es común en la construcción de sistemas de citometría.

3. CONSTRUCCIÓN DEL MODELO COMPUTACIONAL

3.1. METODOLOGÍA

La metodología utilizada para el desarrollo de la herramienta fue el Proceso Unificado de Desarrollo de Software. Éste, utiliza el Lenguaje Unificado de Modelado con el cual se preparan todos los componentes de un sistema software. La construcción de la herramienta se hizo por medio de casos de uso, centrándose en la arquitectura, iterativa e incremental. Se conoce como caso de uso a toda acción que proporcione al usuario un resultado importante.

Para el Proceso Unificado es importante definir claramente los requisitos del usuario, es decir, lo que el usuario quiere que el sistema haga; ha sido planteado como la construcción de una herramienta que permita clasificar las células de una citología cérvico uterina. Este caso de uso permitió diseñar la arquitectura del sistema a medida que evolucionaba el proceso. El trabajo se dividió en partes más pequeñas llamadas iteraciones. Estas iteraciones permitieron diferenciar el crecimiento del producto.

Para construir una versión del producto siguiendo esta metodología se realizaron cuatro fases, de la siguiente manera:

- **Inicio.** Las tareas principales realizadas; definir el protocolo y la realización de un estudio en el área de patología clínica.
- **Elaboración.** En esta fase se buscó diseñar los modelos matemáticos para la caracterización de cada una de las clases celulares, se postuló la utilización de la Transformada de Fourier; se diseñaron los algoritmos, las bases de datos y el motor de inferencia, utilizando principalmente redes neuronales y reglas de inferencia que nos permitieron clasificar las células. En esta fase se inició la recolección de las muestras.
- **Construcción.** Una vez realizados los diseños mencionados en la fase anterior, se continuó con la construcción y codificación de los algoritmos, además de tomar la totalidad de las muestras necesarias para el proyecto.
- **Transición.** Se realizó la verificación de la herramienta; y se espera publicar los respectivos resultados en congresos y artículos de revistas especializadas.

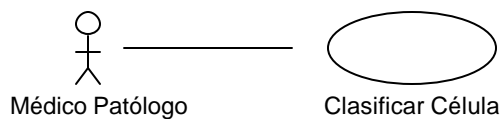
Adicionalmente, en cada una de las iteraciones se realizaron los siguientes flujos de trabajo:

- **Requisitos.** Se buscó describir los requisitos de la iteración, es decir, lo que esta desarrollaría.

- **Análisis.** Su objetivo fue lograr una comprensión más precisa de los requisitos que permitiera estructurar la iteración de una manera más fácil.
- **Diseño.** su propósito fundamental fue formular los modelos que permitieran la evolución de la iteración.
- **Implementación.** su propósito fue la elaboración de la iteración por ejemplo; código fuente, almacenamiento en la base de datos, construcción de interfaces.
- **Prueba.** su objetivo consistió en comprobar los resultados de la implementación mediante pruebas.

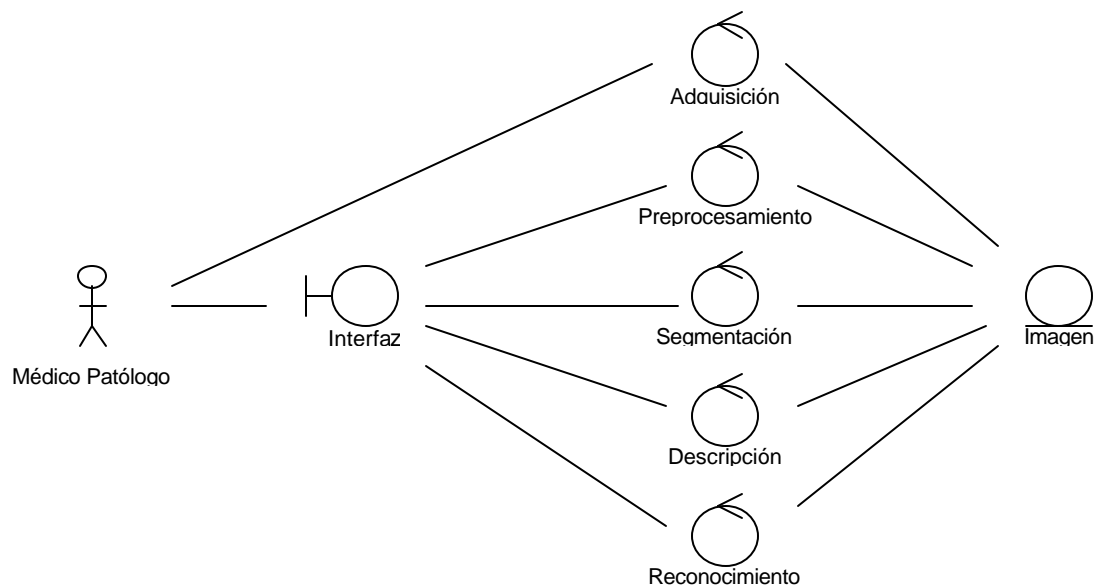
El primer paso fue la definición de los casos de uso o acciones que proporcionan un resultado importante. Como ha sido estipulado en el objetivo general, consistió en la clasificación de las células como normal o displásica, donde el actor es el Médico Patólogo.

Figura 5. Diagrama de Casos de Uso



El caso de uso Clasificar Célula, permite a su vez crear el modelo de análisis, descrito como una estructura de clases y relaciones entre ellas.

Figura 6. Diagrama del Modelo de análisis.



3. 2 RECOLECCIÓN DE MUESTRAS

Las muestras de citologías cérvico uterinas eran recolectadas por los patólogos del proyecto, se tuvieron en cuenta los siguientes criterios de inclusión:

- Citologías con coloración de papanicolaou.
- Muestras consideradas como satisfactorias por los patólogos, cumpliendo los requisitos del sistema de Bethesda.
- Citologías normales.
- Citologías que presentaron cambios reactivos o lesiones preneoplásicas y neoplásicas

Las muestras que no cumplieran estos criterios eran rechazadas. Las células presentes eran clasificadas como basales, parabasales, intermedias o superficiales; luego se digitalizaban cada una de las imágenes con el propósito de identificar los parámetros.

Se adquirieron un total de 225 imágenes celulares que se sometieron a un proceso de selección, con el fin de obtener una muestra que no pudiese presentar ambigüedades en la información. Esto produjo un total de 183 imágenes evaluadas.

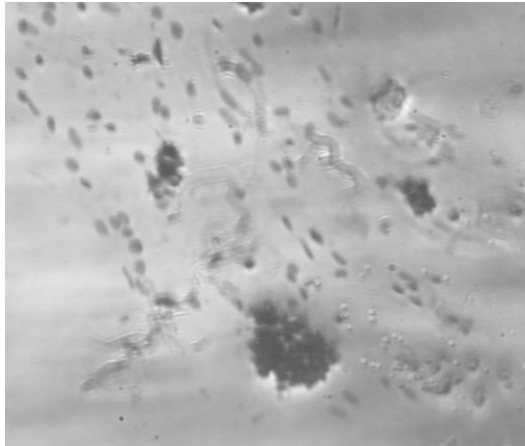
4. ADQUISICIÓN

Con el propósito de obtener una imagen con las mejores condiciones de iluminación, aumento y resolución; se construyeron diversos modelos de adquisición. Se varió cada uno de los componentes necesarios para realizarla: cámaras, dispositivos digitalizadores y equipos de cómputo.

El primer modelo de adquisición, localizado en el Laboratorio de Óptica y Tratamiento de Señales de la Escuela de Física, consistió en una cámara a blanco y negro, un microscopio interferométrico y una tarjeta digitalizadora.

Éste no fue utilizado porque la lámina de citología cérvico uterina debía procesada sin el tratamiento de fijación, y en esas condiciones las células sufren desecación. La figura 7 muestra una imagen adquirida con este modelo.

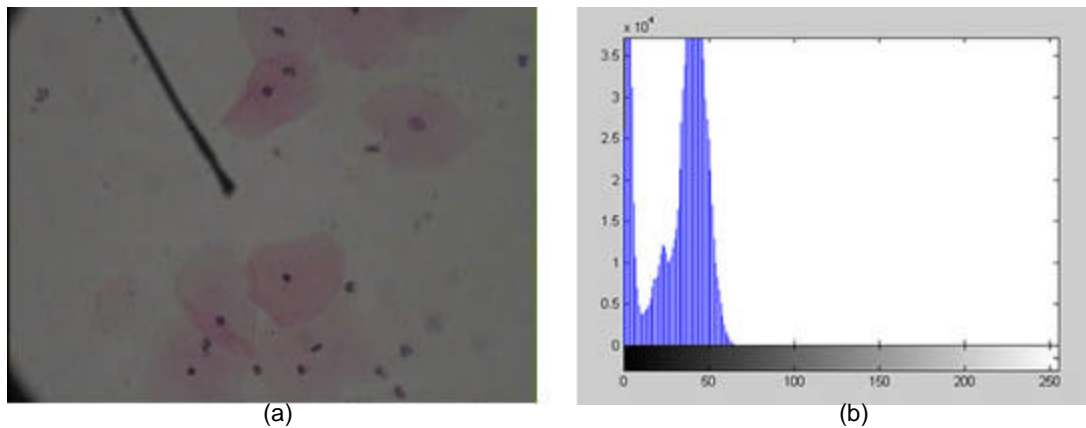
Figura 7. Imagen tomada en el Laboratorio de Óptica y Tratamiento de Señales



El segundo modelo fue elaborado en la Facultad de Medicina de la Universidad Autónoma de Bucaramanga; se utilizó un microscopio Carl Zeiss, con cámara de vídeo Panasonic. Gracias al cambio de microscopio se pudo utilizar el aumento de 40x.

Puesto que no se contaba con un dispositivo digitalizador, se empleó una cámara de vídeo Sony Handycam con conexión USB, ésta permitía almacenar las imágenes en la memoria para descargarlas posteriormente en un computador. Eran guardadas en el formato JPEG (Joint Photographic Experts Group), en una resolución de 320x240 píxeles. La compresión daba como resultado imágenes a color con un tamaño en bytes menor al primer modelo, pero se perdía información respecto a cromaticidad y el histograma de niveles de gris mostraba poca iluminación. La figura 8 corresponde a una imagen adquirida con este segundo modelo.

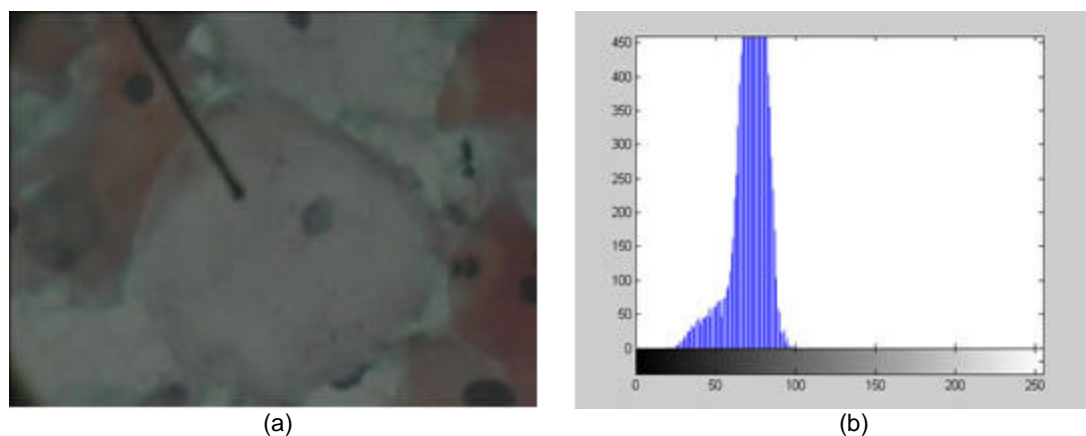
Figura 8. Imagen tomada utilizando una cámara de vídeo Sony Handycam como digitalizador.
a) Imagen celular. b) Histograma de la imagen



Para el tercer modelo se cambió la cámara Sony Handycam por una tarjeta de adquisición de vídeo MPEG TV Station de tipo USB; se conectó a un Computador Portátil Dell Inspiron con Sistema Operativo Windows XP, Las imágenes fueron almacenadas en el formato TIFF (Tagged Image File Format), con una resolución de 640x480 píxeles y un tamaño aproximado a los 901 Kbytes.

Este modelo no fue utilizado, debido a que el histograma de la imagen no presentaba la diferenciación de los objetos y se observaba poca iluminación, adicional a esto la pantalla del equipo visualizaba en algunos casos una imagen borrosa, dificultando su enfoque, y el microscopio poseía un indicador que en ocasiones fraccionaba las células. La figura 9 presenta la imagen resultado de este tercer modelo.

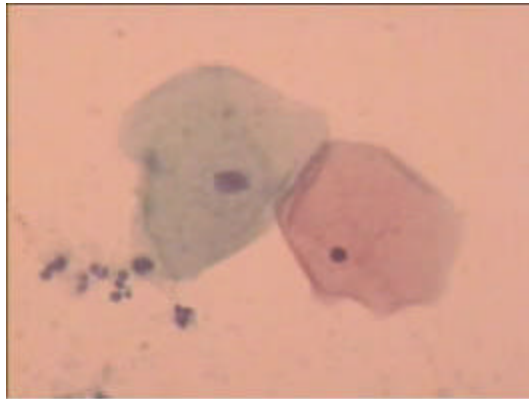
Figura 9. Imagen tomada utilizando una tarjeta de adquisición MPEG TV Station.
a) Imagen celular. b) Histograma de la imagen



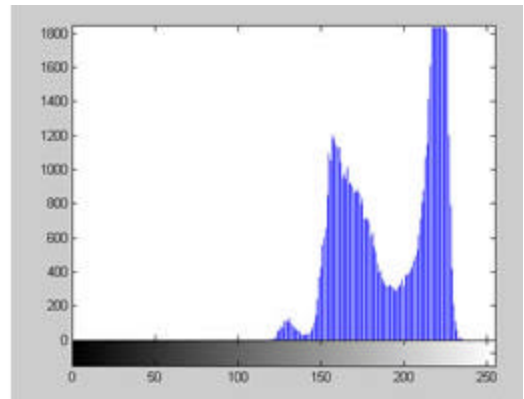
Para el modelo definitivo se cambiaron el microscopio, la cámara de vídeo y se trabajó con un computador portátil que presentara mejor visualización de las imágenes, se manejó un equipo Sony Vaio. Se conservó la tarjeta de adquisición. El tamaño de las células aumentó, y se encontró

en el histograma que la iluminación mejoró considerablemente, mientras que los objetos se diferenciaban fácilmente. Para la adquisición de las imágenes se buscaron células individuales y que mostraran un alto contraste con el fondo. La figura 10 representa el tipo de imágenes que fueron utilizadas para el desarrollo del proyecto.

Figura 10. Imagen utilizada para el desarrollo del proyecto.
a) Imagen celular. b) Histograma de la imagen



(a)



(b)

5. PREPROCESAMIENTO

Una vez adquirida la imagen se procedió a la construcción de los algoritmos para la etapa de preprocesamiento o filtrado, se utilizaron filtros clasificados como espaciales paso bajo. Se recurrió a esta serie de matrices porque su funcionamiento permite suavizar la imagen, y cada una de las superficies, núcleo y citoplasma, tiende a ser homogénea. El filtrado se realiza como la convolución de la matriz del filtro $h(x, y)$ con la imagen original $f(x, y)$, dando como resultado la imagen filtrada $g(x, y)$:

$$g(x, y) = h(x, y) * f(x, y)$$

Los filtros empleados se encuentran implementados en la herramienta Matlab:

- **Promedio.** Este filtro promedia el valor del píxel con sus vecinos, su resultado depende del tamaño de la matriz utilizada. Un filtro de tamaño tres está definido de la siguiente manera:

$$1/9 * \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

- **Promedio Circular.** Utiliza una matriz de forma circular para aplicar el promedio del píxel con sus vecinos, el resultado depende del radio. Un filtro de radio dos está definido de la siguiente manera:

$$\begin{bmatrix} 0 & 0.0170 & 0.0381 & 0.0170 & 0 \\ 0.0170 & 0.0784 & 0.0796 & 0.0784 & 0.0170 \\ 0.0381 & 0.0796 & 0.0796 & 0.0796 & 0.0381 \\ 0.0170 & 0.0784 & 0.0796 & 0.0784 & 0.0170 \\ 0 & 0.0170 & 0.0381 & 0.0170 & 0 \end{bmatrix}$$

- **Rotacional.** El resultado del filtro es la aproximación al movimiento de una cámara, que depende de la longitud y el ángulo. Para 0° y 90° funciona como un filtro promedio de una fila o una columna respectivamente. Si la longitud es tres y el ángulo de 45° se obtiene:

$$\begin{bmatrix} 0 & 0.0754 & 0.1883 \\ 0.0754 & 0.3215 & 0.0754 \\ 0.1883 & 0.0754 & 0 \end{bmatrix}$$

- **Gauss.** Se aplica un filtro Gaussiano rotacionalmente simétrico, el resultado depende del tamaño de la matriz y de la desviación estándar utilizada (sigma). Un filtro de tamaño tres y sigma ½ está definido de la siguiente manera:

$$\begin{bmatrix} 0.0113 & 0.0838 & 0.0113 \\ 0.0838 & 0.6193 & 0.0838 \\ 0.0113 & 0.0838 & 0.0113 \end{bmatrix}$$

- **Unsharp.** Conocido también como High-Boost, utiliza el negativo del operador Laplaciano, el cual es un filtro paso alto. El resultado depende del factor alfa (α) utilizado. Para $\alpha = 0.2$ está definido de la siguiente manera:

$$\begin{bmatrix} -0.1667 & -0.6667 & -0.1667 \\ -0.6667 & 4.3333 & -0.6667 \\ -0.1667 & -0.6667 & -0.1667 \end{bmatrix}$$

- **Mediana.** Este filtro no posee una matriz de coeficientes definida, dado que opera intercambiando el valor del píxel por la mediana de éste con sus vecinos. El funcionamiento depende del número de filas y columnas utilizadas.

La figuras 11 y 12 presentan el resultado de la aplicación de los filtros mencionados para una célula basal.

Figura 11. Filtrado con Promedio, Circular y Gauss.

(a) Imagen Original. (b) Filtrado Promedio (c) Filtrado Circular (d) Filtrado Gauss

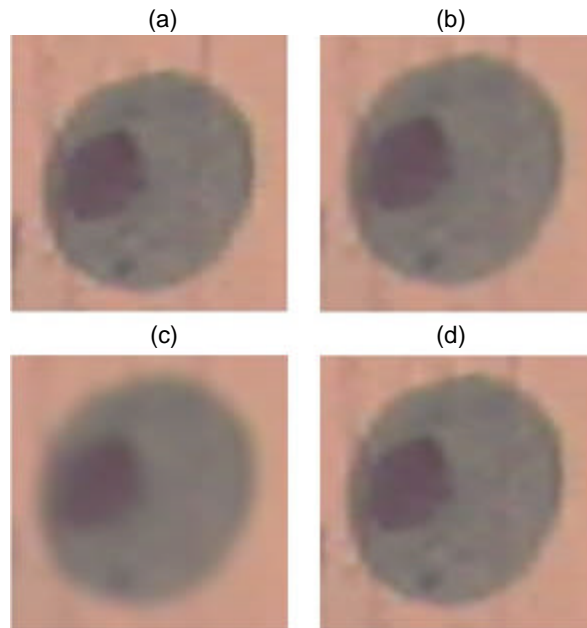
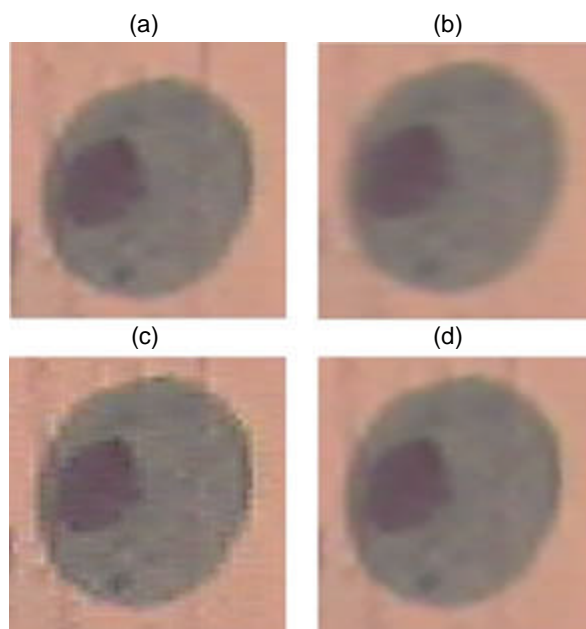


Figura 12. Filtrado con Rotacional, Unsharp y Mediana.

(a) Imagen Original. (e) Filtrado Rotacional. (f) Filtrado Unsharp. (g) Filtrado Mediana.



6. SEGMENTACIÓN

Las etapas anteriores permitieron obtener imágenes de alto contraste (Adquisición), con superficies homogéneas (Preprocesamiento). Debido a estas características, los algoritmos utilizados en la segmentación están basados en la umbralización, permitiendo obtener imágenes binarias de las regiones. El procedimiento es el siguiente:

1. Se detectó el umbral con alguno de los métodos implementados y se efectuó la umbralización sobre uno de los tres planos de color (Rojo, Verde o Azul) o en la escala de grises.
2. Se aplicó morfología matemática.
3. Se extraen los bordes utilizando principalmente filtros espaciales paso alto.

6.1 UMBRALIZACIÓN

6.1.1 Pun-Kapur. El primer método empleado se denomina Pun-Kapur. Éste localiza el umbral utilizando la entropía del histograma. Pun Define dos entropías a posteriori, una para el objeto y otra para el fondo. Kapur hace una extensión de estas involucrando distribuciones de probabilidad.

$$H_b(t) = - \sum_{i=0}^t \frac{p_i}{P_t} \ln \left(\frac{p_i}{P_t} \right)$$

$$H_w(t) = - \sum_{i=t+1}^{l-1} \frac{p_i}{1-P_t} \ln \left(\frac{p_i}{1-P_t} \right)$$

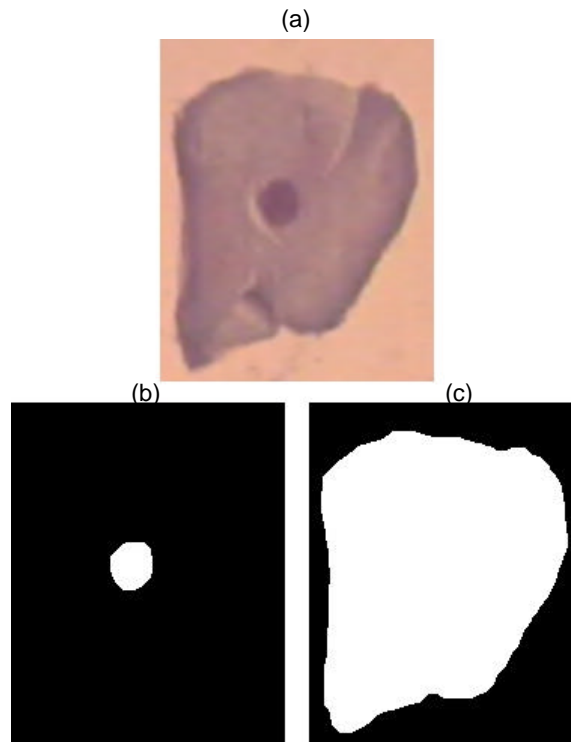
$$H = H_b + H_w$$

Donde H_w es la entropía del objeto, H_b es la entropía del fondo, p_i es la probabilidad de un nivel de gris, P_t es la distribución de probabilidad del histograma, y el umbral t es localizado cuando H es máximo.

Dado que sólo detecta umbrales para histogramas bimodales, se realizó una modificación que permitió trabajar aquellos que presentan tres distribuciones, característicos de las imágenes celulares: Se obtiene el primer umbral, se eliminan los niveles menores a éste y se ejecuta nuevamente el algoritmo.

El método modificado permitió segmentar las células que presentaban una gran diferenciación entre el núcleo, el citoplasma y el fondo. La figura 13 presenta el resultado de aplicar la segmentación empleando este algoritmo.

Figura 13. Resultado de la segmentación utilizando el método de Pun-Kapur.
(a) Imagen Original. (b) Núcleo segmentado. (c) Citoplasma segmentado



6.1.2 Iterativo. El segundo algoritmo implementado se denomina iterativo. Para la localización del umbral se hacen comparaciones sucesivas entre dos umbrales hasta encontrar el adecuado, el método es el siguiente:

- El primer paso consiste en determinar los dos umbrales iniciales; el primero (U_0) se localiza en cero y el segundo (U_1) en la mediana de los niveles del histograma diferentes a cero.

$$U_0 = 0$$

$$U_1 = \text{mediana de los niveles diferentes a cero}$$

- Se calcula la diferencia entre U_1 y U_0 .

$$Dif = U_1 - U_0$$

- Mientras que el valor absoluto de la diferencia sea mayor que uno, se calculan nuevamente los umbrales de la siguiente manera:

$$U_0 = U_1$$

$$U_1 = \frac{\text{Promedio}(\text{Mayores a } U_0) + \text{Promedio}(\text{Menores a } U_0)}{2}$$

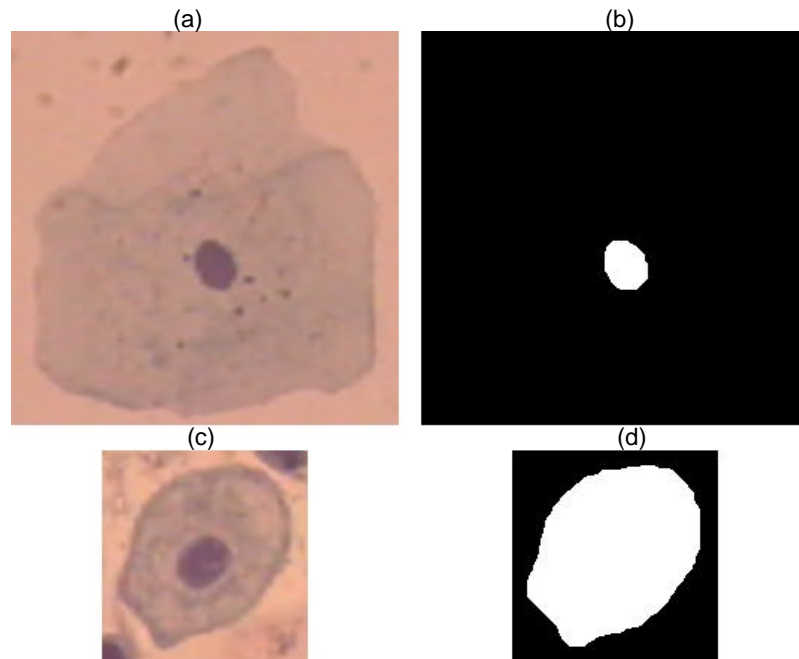
- Se realiza la comparación de la diferencia hasta que sea menor o igual que uno y se obtiene el umbral adecuado.

$$\text{Umbral} = U_1$$

Este método permitió segmentar en algunas células el núcleo y en otras el citoplasma. Como también se trataba de un método para histogramas bimodales, se aplicó la misma variación diseñada para el método de Kapur, sólo que para este caso no se obtuvieron los resultados esperados. La figura 14 presenta el resultado de aplicar la segmentación por medio de este algoritmo.

Figura 14. Resultado de la segmentación utilizando el método Iterativo.

(a) Célula Intermedia. (b) Núcleo segmentado. (c) Célula Parabasal. (d) Citoplasma segmentado

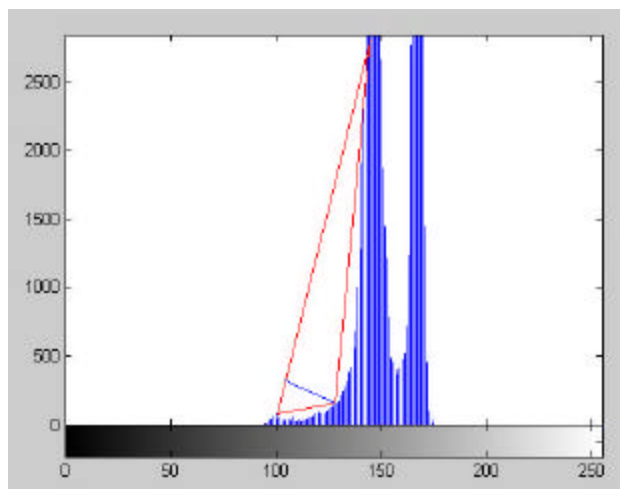


6.1.3 Triangular. El tercer método de detección de umbrales se denomina triangular; este método detecta el umbral por medio de la construcción de triángulos entre los picos máximos (crestas) del histograma. Diseñado de la siguiente manera:

- Se detectan dos crestas sucesivas del histograma.
- Se trazan triángulos con cada uno de los niveles intermedios a éstas.
- El nivel que determine el triángulo con mayor altura es considerado umbral.

La siguiente figura presenta el esquema para la localización de umbrales utilizando el algoritmo mencionado.

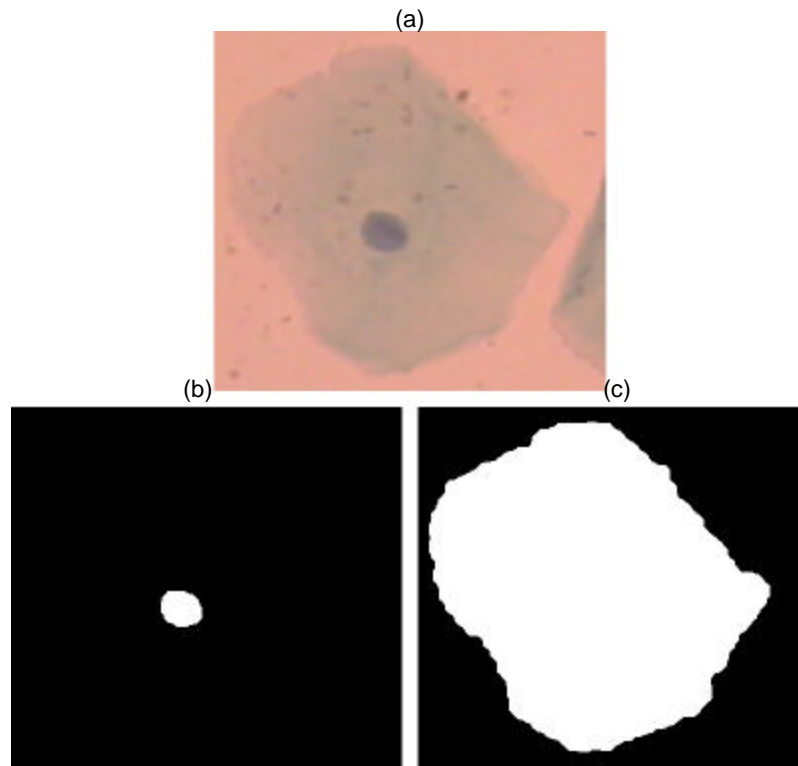
Figura 15. Localización de umbrales por el método Triangular



Con el fin de eliminar falsos picos se le aplicó un promediado a cada punto de histograma y sus vecinos, y en la localización de las crestas se utilizó la derivada.

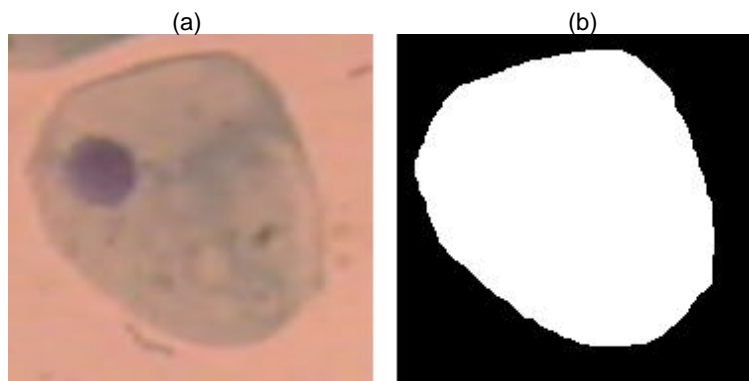
La figura 16 presenta el funcionamiento de este método, que permite segmentar el núcleo y el citoplasma en imágenes que presenten poco contraste.

Figura 16. Resultado de la segmentación utilizando el método Triangular.
(a) Imagen Original. (b) Núcleo segmentado. (c) Citoplasma segmentado



6.1.4 Matlab. El cuarto método empleado para la detección del umbral, consistió en el algoritmo por defecto de Matlab implementado en la función `graythresh`. Esta función permite detectar sólo un umbral para cada imagen, la variación en este procedimiento consistió en igualar los niveles de gris menores al primer umbral localizado y aplicar nuevamente la función con el propósito de determinar el segundo umbral. En algunos casos permitió segmentar el núcleo o el citoplasma. La siguiente figura presenta la segmentación realizada por el algoritmo.

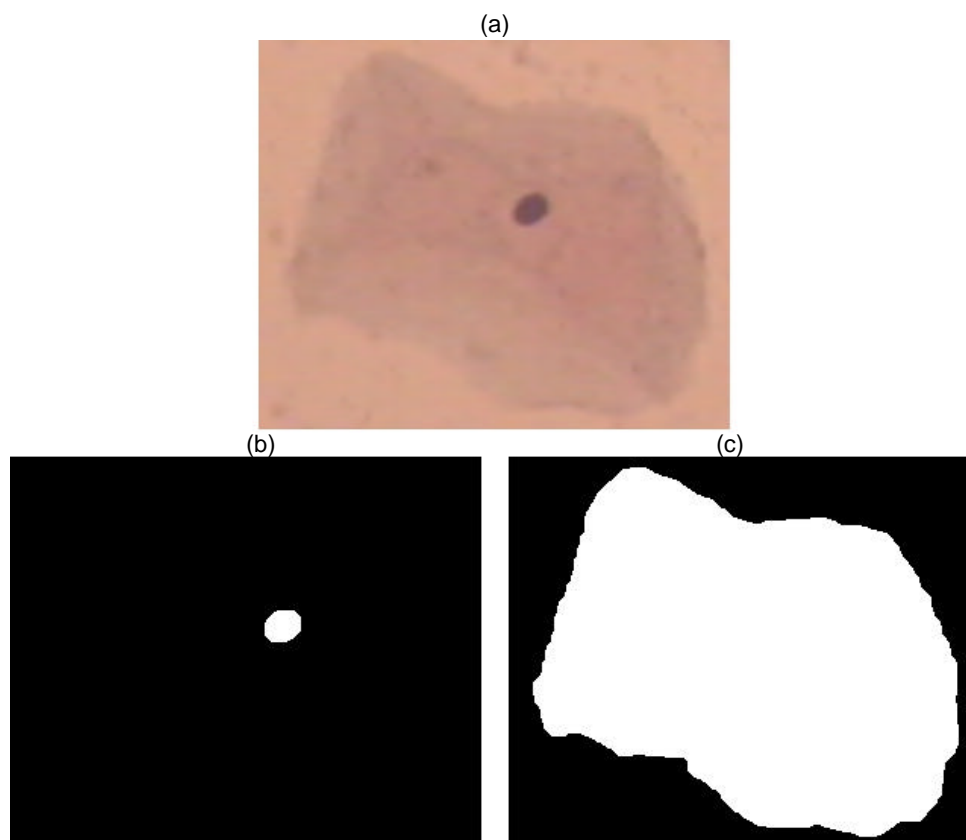
Figura 17. Resultado de la segmentación utilizando el método de Matlab.
(a) Imagen Original. (b) Citoplasma segmentado.



6.1.5 Manual. Dado que los métodos mencionados no segmentaron correctamente la totalidad de la muestra de imágenes empleadas, se utilizaron métodos manuales, determinando el umbral hasta que la célula quedara segmentada. La figura 18 presenta la utilización de este procedimiento.

Figura 18. Resultado de la segmentación utilizando el método Manual.

(a) Imagen Original. (b) Núcleo segmentado, umbral = 127. (c) Citoplasma segmentado, umbral = 158



6.2 MORFOLOGÍA MATEMÁTICA

La Morfología Matemática fue utilizada como un procedimiento interno de la segmentación. Es aplicada inmediatamente después de la umbralización. Ésta permite eliminar los objetos que no corresponden a un núcleo o citoplasma. El procedimiento Morfológico es explicado a continuación:

- Se realizó una apertura con el siguiente elemento estructurante:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Esta apertura permitió separar regiones traslapadas.

- Se calculó el área de los objetos, se tomó el objeto con mayor área y se eliminaron los demás.
- Una vez obtenida la imagen con el objeto de mayor área, se efectuó un cierre con el siguiente elemento estructurante:

$$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \end{pmatrix}$$

Se decidió que los elementos estructurantes fuesen redondos, debido a la forma presentada por las células, y el tamaño se obtuvo experimentalmente.

6.3 EXTRACCIÓN DE BORDES

En la extracción de bordes se experimentaron varios procedimientos, el objetivo consistió en lograr que la imagen resultante presentara un borde de ancho uno. El primer método

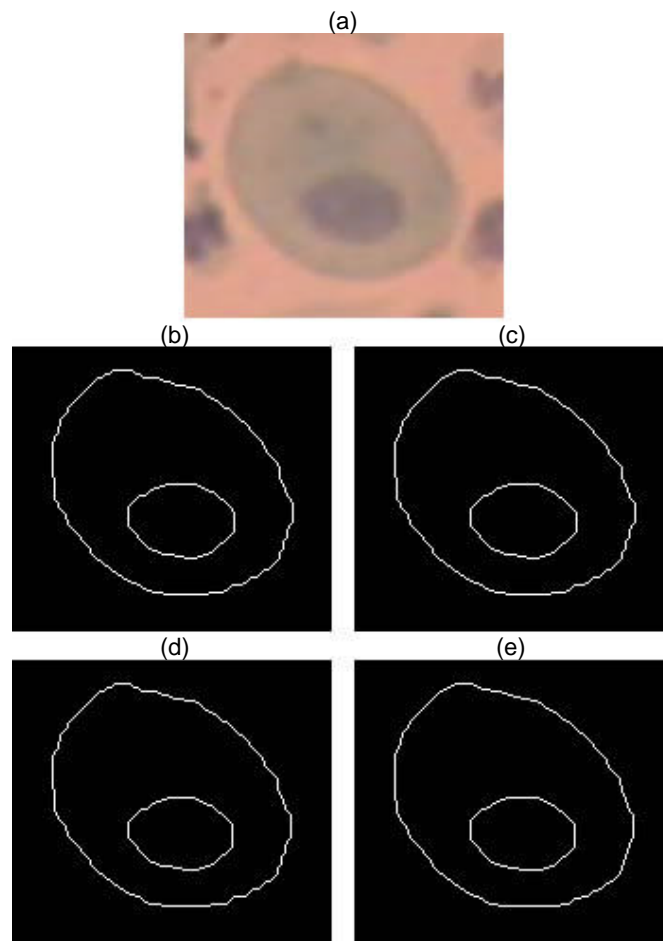
implementado utilizó los operadores de Kirsch, un conjunto de ocho matrices que calculan el gradiente de la imagen, pero no cumplieron la condición dada.

El siguiente paso fue utilizar una función implementada en Matlab para la detección de bordes. Se utilizaron cada uno de los parámetros, dando como resultado la elección de los siguientes métodos: Sobel, Prewitt, Laplaciano del Gaussiano y Canny.

La siguiente figura presenta la extracción de bordes realizada en una célula parabasal.

Figura 19. Extracción de Bordes.

(a) Imagen Original (b) Operadores de Sobel. (c) Operadores de Prewitt
(d) Laplaciano del Gaussiano. (e) Algoritmo Canny



7. DESCRIPCIÓN

El resultado de la segmentación es un grupo de imágenes binarias, y corresponden con las regiones del núcleo, el citoplasma y sus respectivos bordes. Éstas permiten extraer los diversos parámetros de las células, utilizados para la identificación y clasificación de células escamosas de citologías cérvico uterinas.

Los parámetros utilizados para describir el núcleo son:

- **Área:** calculada como el número de píxeles en blanco de la imagen binaria que contiene el núcleo.
- **Perímetro:** calculado como el número de píxeles en blanco de la imagen binaria que contiene el borde del núcleo.
- **Ejes:** calculados como los ejes de la elipse con el segundo momento igual a la región del núcleo.
- **Circularidad:** relación entre el cuadrado del perímetro y el área, esta medida permite conocer la tendencia circular del núcleo. Si un objeto presenta esta forma geométrica el valor es 4π .
- **Niveles de Color:** para cada plano de color (R, G, B) de la imagen adquirida, se obtienen el nivel menor, la moda, el nivel mayor y el promedio de la región que contiene el núcleo.
- **Descriptores de Fourier:** calculados como los coeficientes de la serie de Fourier, de la función del contorno expresada en números complejos. Para la reconstrucción de los contornos se utilizaron 70 descriptores.

Para la descripción del citoplasma se utilizan:

- **Área:** calculada como el número de píxeles en blanco de la imagen binaria que contiene el citoplasma.
- **Perímetro:** calculado como el número de píxeles en blanco de la imagen binaria que contiene el borde del citoplasma.
- **Ejes:** calculados como la longitud de los ejes de la elipse con el segundo momento igual a la región del citoplasma.

- **Relaciones Núcleo/Citoplasma:** llamada también Radio N/C, indica la proporción entre el tamaño del núcleo y el citoplasma; se calcula para el área, el perímetro y los ejes.

Con el fin de determinar las medidas del área, el perímetro y los ejes en las unidades utilizadas para la medición de las células, se calculó una relación entre un píxel y la cantidad de micras contenidas en esta área.

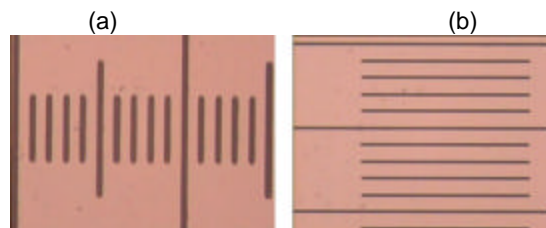
Para obtener esta relación se utilizó un par de rejillas diseñadas para el cálculo de medidas en el microscopio, mediante el siguiente procedimiento utilizando un aumento de 40x:

- Se tomó una imagen de cada rejilla, vertical y horizontal.
- De acuerdo con la fabricación de las rejillas, cada línea estaba separada una de otra por diez micras.
- Se diseñó un algoritmo que permitía obtener líneas con un píxel de ancho.
- Se calculó el número de píxeles que separaba cada línea.
- Se obtuvo experimentalmente que la separación en píxeles era 41, tanto vertical como horizontalmente.
- Con esta medida se obtuvieron los siguientes factores para los demás aumentos, mediante una relación lineal:

$$5x = \left(\frac{80}{41}\right)^2, \quad 10x = \left(\frac{40}{41}\right)^2, \quad 40x = \left(\frac{10}{41}\right)^2, \quad 100x = \left(\frac{4}{41}\right)^2$$

Las unidades de estos factores son micras²/píxeles. La siguiente figura presenta las rejillas utilizadas para el cálculo.

Figura 20. Rejillas utilizadas para calcular la relación entre micra-píxel.
(a) Rejilla para el cálculo horizontal. (b) Rejilla para el cálculo vertical.



8. RECONOCIMIENTO

8.1 RED NEURONAL

La etapa de reconocimiento se inicia con la clasificación del núcleo como regular o irregular. Para que el modelo pudiera tomar este tipo de decisión se planteó su solución por medio de una Red Neuronal. Se utilizaron las arquitecturas implementadas en Matlab.

La idea original fue construir una red de 70 entradas, cada una correspondiente al módulo de un descriptor de Fourier; y una salida, que produciría un valor positivo si la forma era regular, cero o negativo para el caso irregular. Este tipo de red no funcionó y se planteó agregar a la red una capa oculta, se varió el número de neuronas en esta capa, se modificó el número de épocas para el entrenamiento, sin llegar a obtener un resultado preciso. Dados estos resultados se planteó la necesidad de cambiar el número de entradas.

Para diseñar la nueva Red Neuronal se plantearon los siguientes criterios:

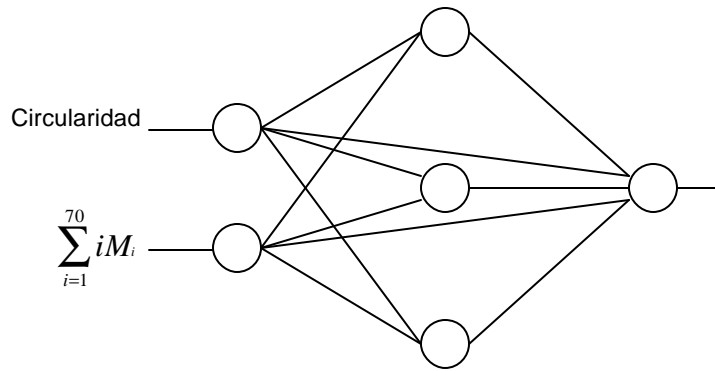
- La primera entrada correspondería a la circularidad.
- Se obtuvo experimentalmente que las formas regulares presentaban los mayores módulos en los últimos descriptores, mientras que las irregulares incrementaban los primeros. Con esta condición se diseñó que la segunda entrada debía ser la sumatoria del producto del índice del descriptor por el módulo.

$$Entrada = \sum_{i=1}^{70} iM_i$$

- Se utilizaría la arquitectura Backpropagation de tipo Cascade-Forward. Capa de entrada, capa oculta y capa de salida. Las formas regulares producirían salidas positivas y las irregulares negativas o cero.
- La capa de entrada con las dos neuronas mencionadas, la capa oculta con tres neuronas y la capa de salida con una neurona.
- Las funciones de transferencia, entrenamiento, aprendizaje, y funcionamiento serían 'Tansig', 'TrainLm', 'LearnGdm' y 'Mse' respectivamente.

La arquitectura Cascade-Forward interconecta las capas sucesivas con todas las capas anteriores. Como se ha utilizado una sola capa oculta significa que la entrada se interconecta a su vez con la salida. La figura 21 presenta el esquema de la red utilizada para la clasificación de la forma nuclear.

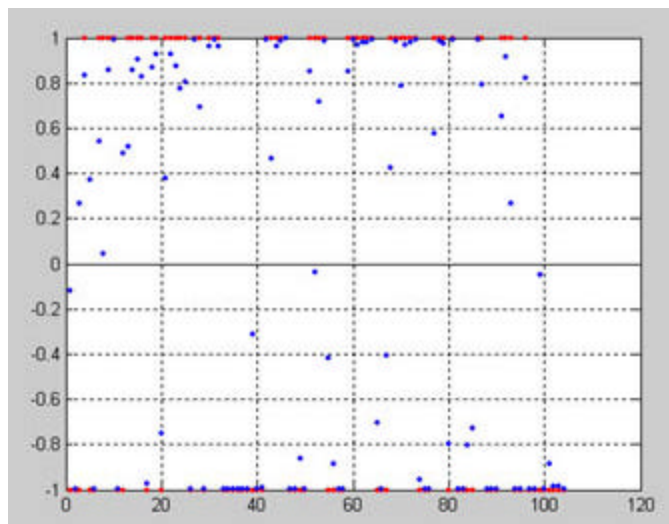
Figura 21. Arquitectura de la red de clasificación.



Una vez implementada la red se procedió a entrenarla utilizando 100 épocas y 104 formas nucleares que incluían igual número de módulos de descriptores de Fourier, circularidades y formas conocidas. Luego se realizó una prueba con los mismos valores utilizados en el entrenamiento y produjo un total de 100 aciertos, 3 falsos negativos y un falso positivo. La figura 22 presenta las salidas de la red (Azules) y las formas conocidas (Rojos).

Se consideran falsos negativos las formas irregulares clasificadas como regulares y falsos positivos las formas regulares clasificadas como irregulares.

Figura 22. Clasificación de la Red Neuronal para 104 formas nucleares.



8.2 SISTEMA BASADO EN EL CONOCIMIENTO

Para la clasificación de cada una de las células dentro de los parámetros de normalidad existentes, se implementó un Sistema Basado en el Conocimiento. Se construyó utilizando la siguiente arquitectura:

- Base de Conocimiento.
- Motor de Inferencia.
- Subsistema de Explicación.

8.2.1 Base de Conocimiento. En su construcción se empleó una muestra de 183 imágenes, de las cuales se extrajeron 286 núcleos y 152 citoplasmas. A cada uno de estos se le calcularon los parámetros mencionados en el capítulo anterior. Después que los valores son almacenados en una base de datos, se evalúan los intervalos para cada clase de célula (basal, parabasal, intermedia, superficial).

Los parámetros almacenados en la base de conocimiento han sido agrupados de la siguiente manera:

- Forma del Núcleo.
- Relaciones Núcleo/Citoplasma. (Área, Perímetro, Eje Mínimo, Eje Máximo).
- Medidas del Núcleo (Área, Perímetro, Eje Mínimo, Eje Máximo).
- Color del Núcleo (Rojo Promedio, Rojo Mínimo, Rojo Moda, Rojo Máximo, Verde Promedio, Verde Mínimo, Verde Moda, Verde Máximo, Azul Promedio, Azul Mínimo, Azul Moda, Azul Máximo).
- Medidas del Citoplasma (Área, Perímetro, Eje Mínimo, Eje Máximo).

8.2.2. Motor de Inferencia. Para el diseño del motor de inferencia se utilizó la regla de Modus Ponens, que aplica el siguiente teorema:

Si $p(X) \rightarrow q(X)$ es verdadero
Si se confirma $p(a)$, entonces se concluye que $q(a)$ es verdadero.

Se evalúan cada uno de los parámetros y se comparan con los intervalos almacenados en la Base de Conocimiento. Las reglas se definieron de la siguiente manera:

- Si Parámetros (Basal) → Célula (Basal).
- Si Parámetros (Parabasal) → Célula (Parabasal).
- Si Parámetros (Intermedia) → Célula (Intermedia).
- Si Parámetros (Superficial) → Célula (Superficial).
- Si Parámetros (Ninguna) → Célula (Displásica).

8.2.3 Subsistema de Explicación. Este se emplea cuando la célula ha sido clasificada como displásica y le permite ver al médico patólogo qué parámetro no se cumplió en cada clase de célula; por ejemplo, cuando el área de un núcleo es demasiado grande respecto al citoplasma.

El subsistema de explicación rastrea el parámetro que no se cumple y da la siguiente respuesta:

- Célula Basal: Eje Mínimo del Núcleo Por Fuera del Rango.
- Célula Parabasal: Color del Núcleo Por Fuera del Rango.
- Célula Intermedia: Relación del Área Por Fuera del Rango.
- Célula Superficial: Relación del Área Por Fuera del Rango.

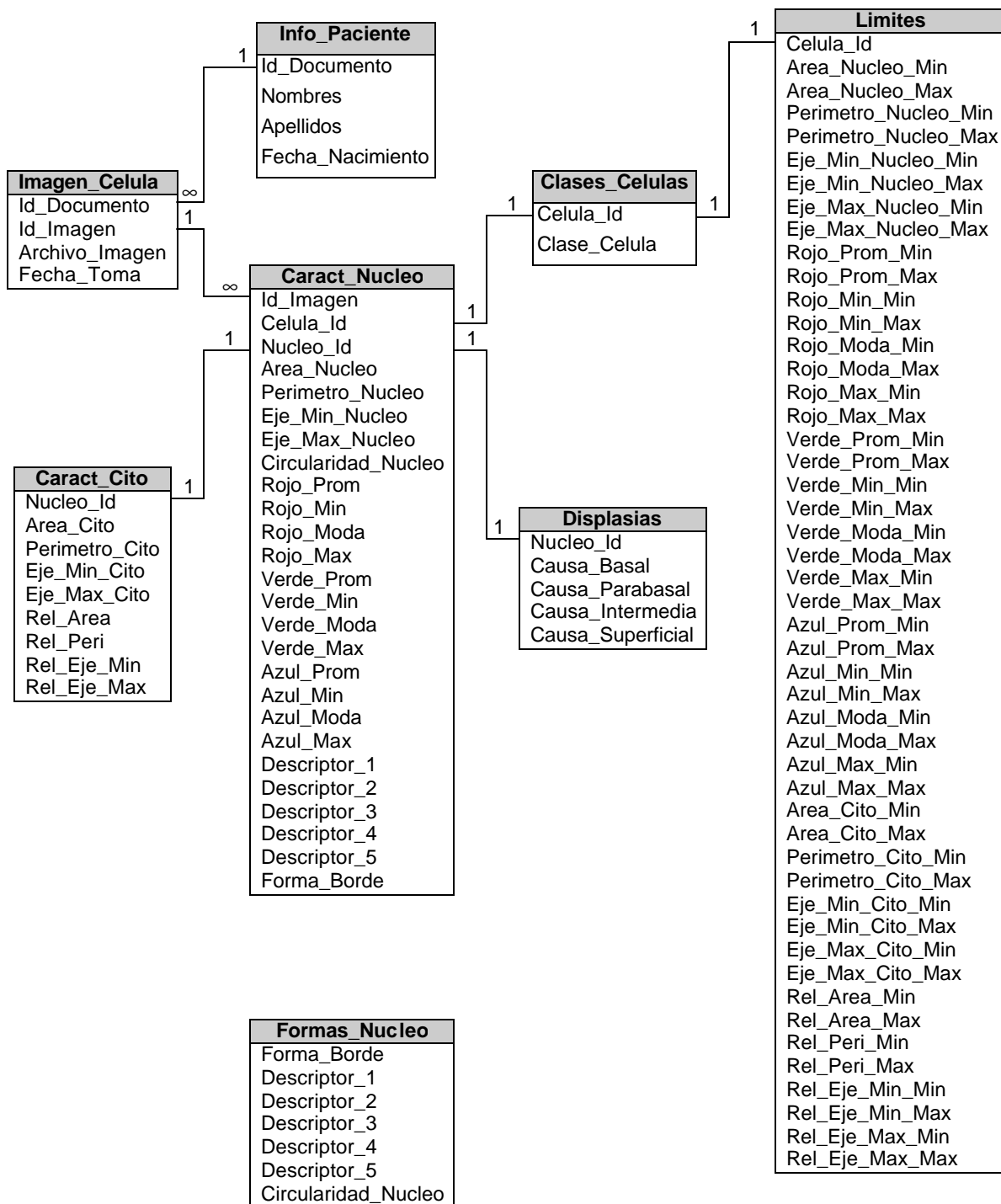
9. BASE DE DATOS

El diseño de la base de datos se hizo de forma relacional, esto permite reducir la redundancia de la información. Las tablas son las siguientes:

- **Info_Paciente:** contiene la información respecto a la paciente que se realiza la citología. La relación con Imagen_Celula permite controlar que una imagen no sea vinculada a diferentes pacientes.
- **Imagen_Celula:** contiene el vínculo al archivo de las imágenes de los pacientes. Relacionada con Caract_Nucleo, esto permite conocer de qué imágenes se han extraído los parámetros.
- **Caract_Nucleo:** contiene los parámetros extraídos de los núcleos. Es la tabla que presenta el mayor número de relaciones; con Clases_Celulas para determinar la clasificación celular, con Caract_Cito para relacionar las mediadas del citoplasma y con Displasias para identificar el resultado obtenido por el Sistema Basado en el Conocimiento.
- **Caract_Cito:** contiene los parámetros de los citoplasmas.
- **Clases_Celulas:** contiene las diferentes clasificaciones celulares.
- **Limites:** tabla del Sistema Basado en el Conocimiento, contiene todos los intervalos de cada una de las clasificaciones celulares. Su relación con Clases_Celulas permite diferenciarlas unas de otras.
- **Formas_Nucleos:** única tabla que no ha sido relacionada. Contiene las formas nucleares que son utilizadas como entradas en el entrenamiento de la red neuronal.

La siguiente figura muestra cada una de las tablas con sus respectivas relaciones.

Figura 23. Base de Datos



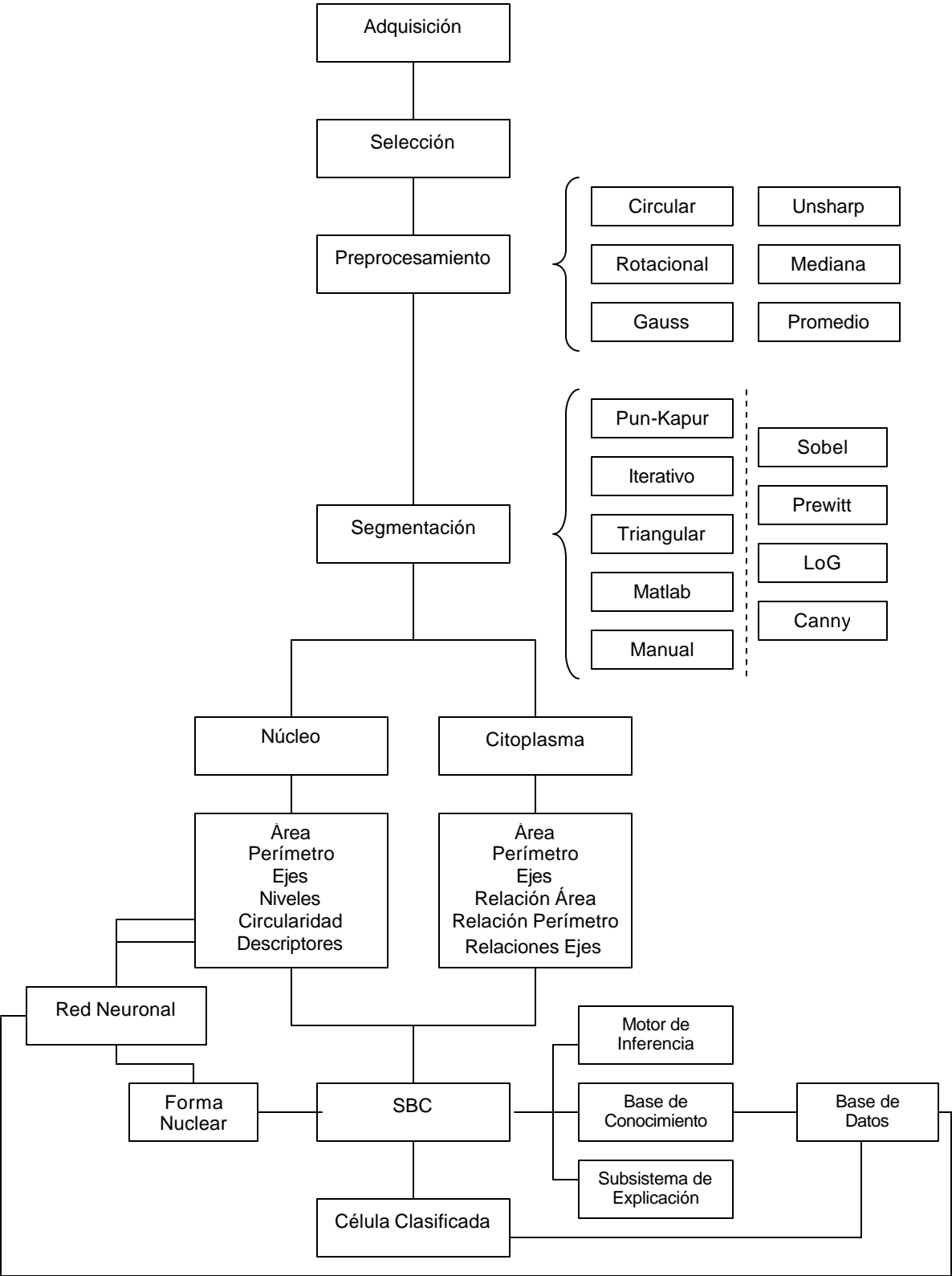
10. MODELO COMPUTACIONAL

El Modelo Computacional permite, por medio de una serie de iteraciones, obtener una célula escamosa clasificada. Es necesario que se sigan cada una de los pasos mencionados a continuación:

- Se realiza la adquisición de la imagen de acuerdo con el modelo explicado en el capítulo cuatro.
- En la imagen adquirida, se selecciona la célula que será clasificada.
- La imagen de la célula es filtrada.
- Se detectan las regiones que corresponden al núcleo y al citoplasma.
- Se calculan los parámetros mencionados en el capítulo siete.
- Se clasifica la forma nuclear, utilizando la red neuronal.
- El Sistema Basado en el Conocimiento clasifica la célula de acuerdo a los parámetros.
- Los datos de la célula son almacenados en la base de datos.

Si alguno de los pasos falla, o no es realizado conforme a este modelo, la célula no es clasificada correctamente. La siguiente figura presenta el esquema del Modelo Computacional.

Figura 24. Esquema del Modelo Computacional.



11. RESULTADOS

Para cada clase de célula normal se calcularon los intervalos de cada uno de los parámetros utilizados en la clasificación. Estos permiten determinar el desempeño del Sistema Basado en el Conocimiento. Cada una de las tablas corresponde a las células basales, parabasales e intermedias. Las unidades que expresan el área se encuentran en micras cuadradas, las longitudes en micras, y los colores el nivel encontrado.

Tabla 1. Parámetros de las células basales.

Parámetro	Valor Mínimo	Valor Máximo
Núcleo:		
Área	20,40	56,51
Perímetro	14,39	24,39
Eje Mínimo	4,62	8,02
Eje Máximo	5,64	9,12
Rojo Promedio	104	138,5
Rojo Mínimo	87	120
Rojo Moda	97	128
Rojo Máximo	115	166
Verde Promedio	83,97	109
Verde Mínimo	63	86
Verde Moda	69	94
Verde Máximo	99	135
Azul Promedio	81,39	109,65
Azul Mínimo	56	98
Azul Moda	82	107
Azul Máximo	107	135
Citoplasma:		
Área	166,38	458,35
Perímetro	35,85	70
Eje Mínimo	12,24	22,46
Eje Máximo	17,58	32,40
Relaciones Núcleo/Citoplasma:		
Área	0,04	0,20
Perímetro	0,21	0,45
Eje Mínimo	0,21	0,50
Eje Máximo	0,19	0,40

Estos parámetros permiten estipular algunas de las características de una célula basal. Núcleo con bajas tonalidades de color (hipercromático), tamaño mediano, al igual que el citoplasma, y una relación N/C entre 1:5 y 1:2.

Tabla 2. Parámetros de las células parabasales.

Parámetro	Valor Mínimo	Valor Máximo
Núcleo:		
Área	26,94	142,59
Perímetro	16,09	38,53
Eje Mínimo	5,72	11,79
Eje Máximo	6,03	15,40
Rojo Promedio	125,5	184,5
Rojo Mínimo	106	167
Rojo Moda	121	171
Rojo Máximo	136	202
Verde Promedio	99	135,5
Verde Mínimo	83	125
Verde Moda	89	134
Verde Máximo	111	152
Azul Promedio	105,58	133,13
Azul Mínimo	85	122
Azul Moda	98	133
Azul Máximo	118	147
Citoplasma:		
Área	264,01	1273,71
Perímetro	38,29	116,09
Eje Mínimo	14,41	38,64
Eje Máximo	22,24	44,75
Relaciones Núcleo/Citoplasma:		
Área	0,04	0,24
Perímetro	0,20	0,54
Eje Mínimo	0,21	0,59
Eje Máximo	0,22	0,44

De acuerdo con esta tabla se encuentra que los niveles de color del núcleo han aumentado respecto a una célula basal, lo que produce tonalidades más claras. Los tamaños, junto con sus relaciones N/C, también se han incrementado levemente.

Tabla 3. Parámetros de las células intermedias.

Parámetro	Valor Mínimo	Valor Máximo
Núcleo:		
Área	27,95	113,08
Perímetro	16,82	33,65
Eje Mínimo	5,03	11,06
Eje Máximo	6,37	13,72
Rojo Promedio	115,82	184
Rojo Mínimo	87	163
Rojo Moda	100	170
Rojo Máximo	130	205
Verde Promedio	67	136
Verde Mínimo	57	121
Verde Moda	61	128
Verde Máximo	77	151
Azul Promedio	73	130
Azul Mínimo	61	115
Azul Moda	73	130
Azul Máximo	85	149
Citoplasma:		
Área	748,06	4788,16
Perímetro	85,85	254,87
Eje Mínimo	25,05	71,95
Eje Máximo	38,49	86,41
Relaciones Núcleo/Citoplasma:		
Área	0,01	0,09
Perímetro	0,08	0,28
Eje Mínimo	0,09	0,37
Eje Máximo	0,08	0,23

En esta tabla se observa que las células intermedias muestran el mayor aumento en el citoplasma, respecto a las anteriores, se presenta una leve disminución de los niveles de coloración y el área nuclear, que produce una menor relación N/C.

Tabla 4. Parámetros de las células superficiales.

Parámetro	Valor Mínimo	Valor Máximo
Núcleo:		
Área	14,	34,97
Perímetro	11,95	18,53
Eje Mínimo	4,35	6,40
Eje Máximo	4,35	7,37
Rojo Promedio	112,51	166,06
Rojo Mínimo	84	151
Rojo Moda	95	156
Rojo Máximo	133	210
Verde Promedio	57,5	110
Verde Mínimo	43	100
Verde Moda	48	101
Verde Máximo	68	141
Azul Promedio	63	124,5
Azul Mínimo	47	111
Azul Moda	56	122
Azul Máximo	75	152
Citoplasma:		
Área	1026,59	3984,77
Perímetro	103,41	220,24
Eje Mínimo	30,01	67,18
Eje Máximo	41,83	91,27
Relaciones Núcleo/Citoplasma:		
Área	0,00	0,04
Perímetro	0,06	0,18
Eje Mínimo	0,07	0,26
Eje Máximo	0,06	0,19

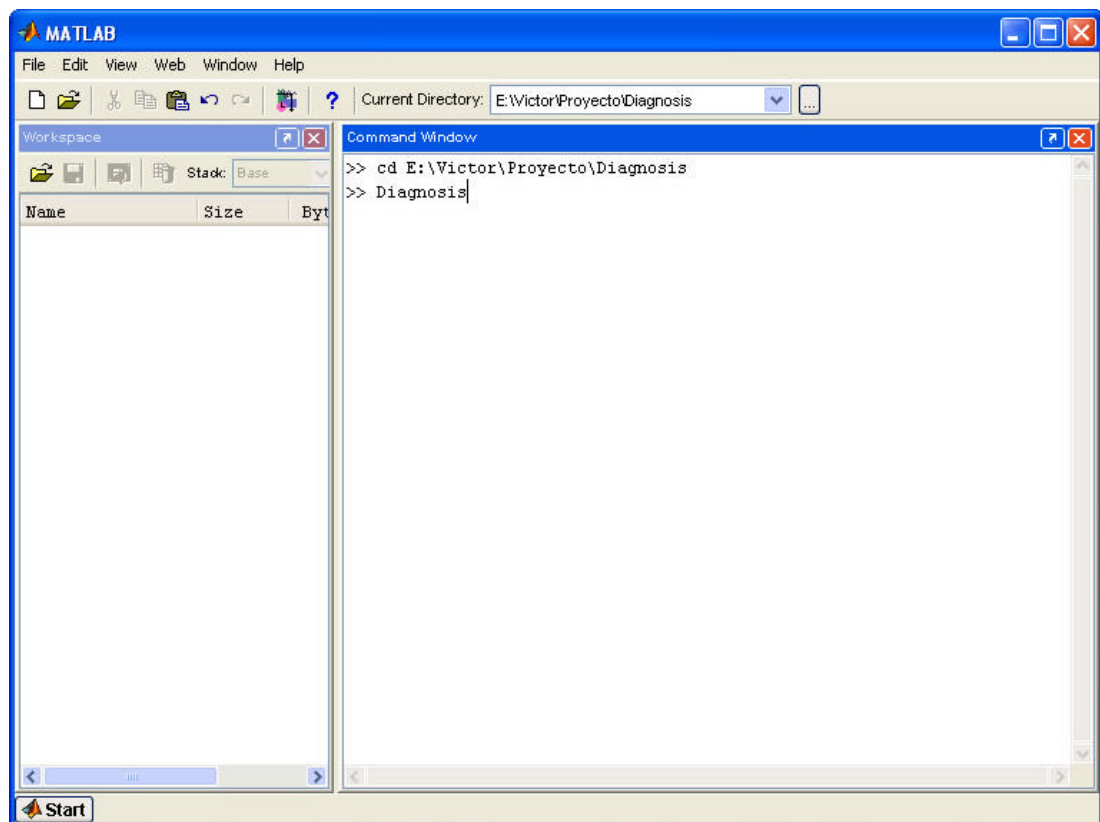
De acuerdo a estos parámetros, las células superficiales presentaron el menor nivel de color, esto es el mayor hiperchromatismo; el tamaño del núcleo se ha reducido considerablemente, se exhibe una leve disminución del citoplasma. La relación N/C es la más pequeña de las células escamosas.

12. FUNCIONAMIENTO DE LA HERRAMIENTA SOFTWARE

Con el propósito de presentar una mayor claridad en el uso de la herramienta software, se mostrará a continuación cada uno de los pasos necesarios para clasificar las células escamosas, de una muestra de citología cérvico uterina:

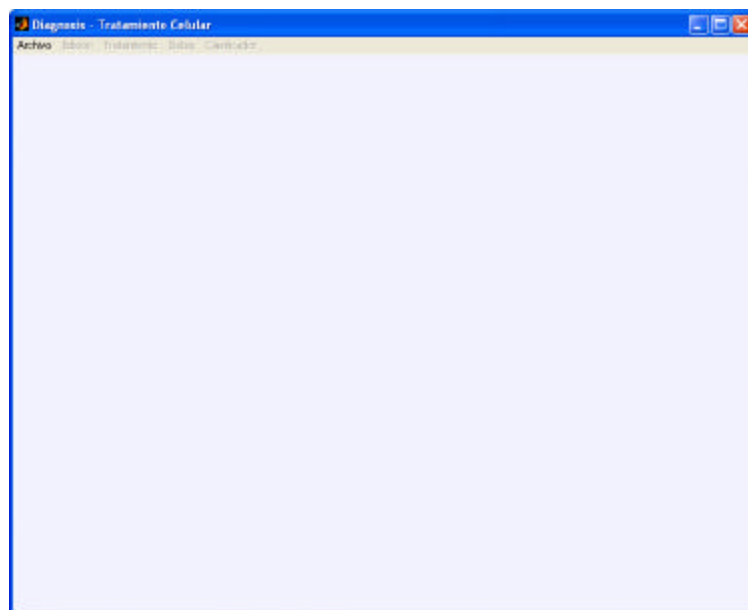
- **Ejecución.** Una vez se ubique la línea de comandos de la plataforma Matlab en el directorio que contiene la herramienta software, se ejecuta la aplicación con el nombre de: *diagnosis*. Como se muestra en la siguiente figura.

Figura 25. Ejecución de la herramienta en la línea de Comandos Matlab.



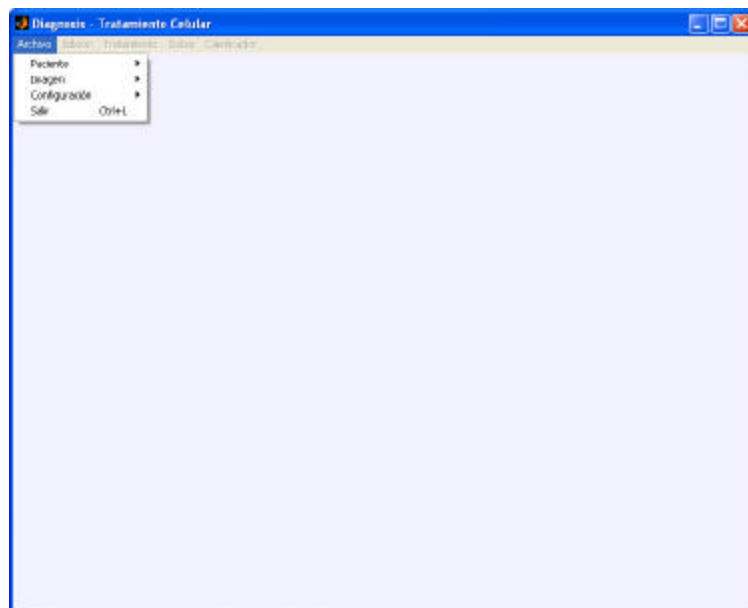
- **Ventana Principal.** Ejecutada la instrucción anterior se despliega la ventana que controla todas las opciones y los algoritmos de la herramienta software. Presentada en la siguiente figura. Esta sólo tiene activado el menú archivo, los restantes se activaran a medida que avanza el proceso de clasificación.

Figura 26. Ventana Principal



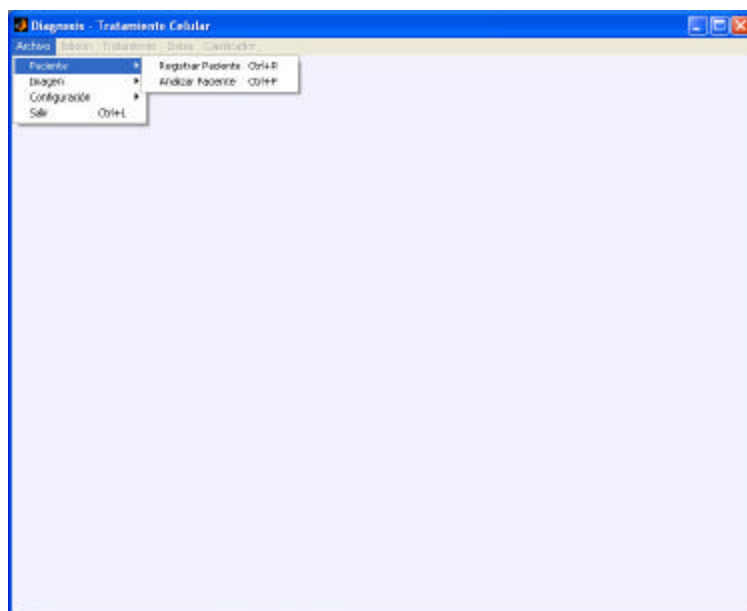
- **Menú Archivo.** Contiene las primeras opciones de la herramienta: paciente, imagen, configuración y salir. La siguiente figura presenta su despliegue.

Figura 27. Menú Archivo.



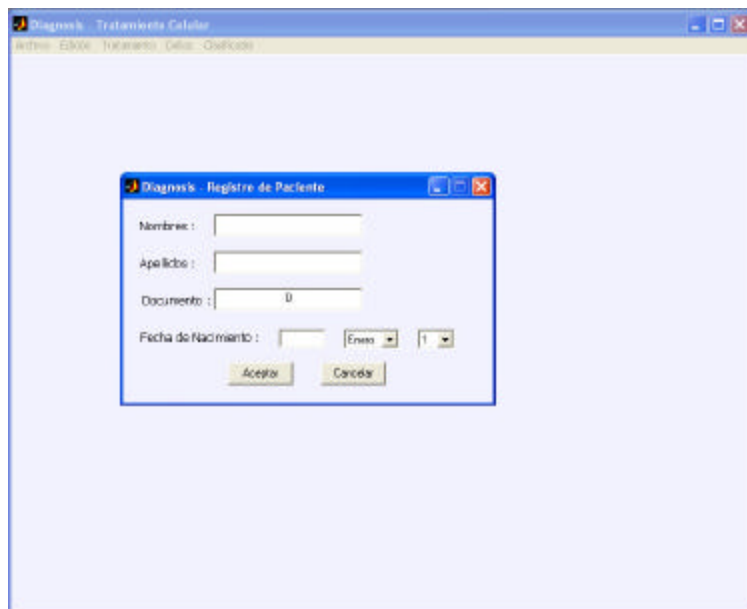
- **Submenú Paciente.** Permite registrar una nueva paciente, o indicarle a la herramienta la paciente que presentó la citología a analizar, de acuerdo con la siguiente figura.

Figura 28. Submenú Paciente.



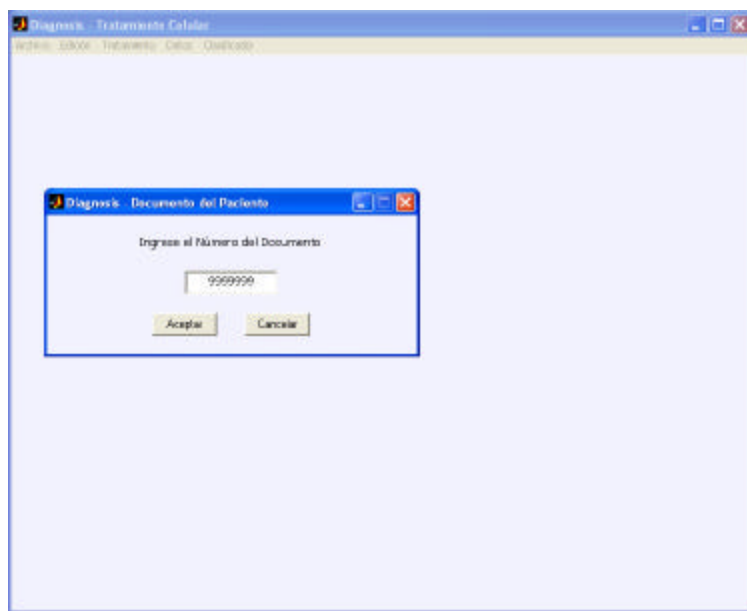
- **Ventana de Registro.** Si es elegida la opción de registrar paciente, se despliega una ventana para agregar el registro a la base de datos. Se ingresa el nombre, los apellidos, el documento de identidad y la fecha de nacimiento. Como se indica en la siguiente figura.

Figura 29. Ventana de Registro.



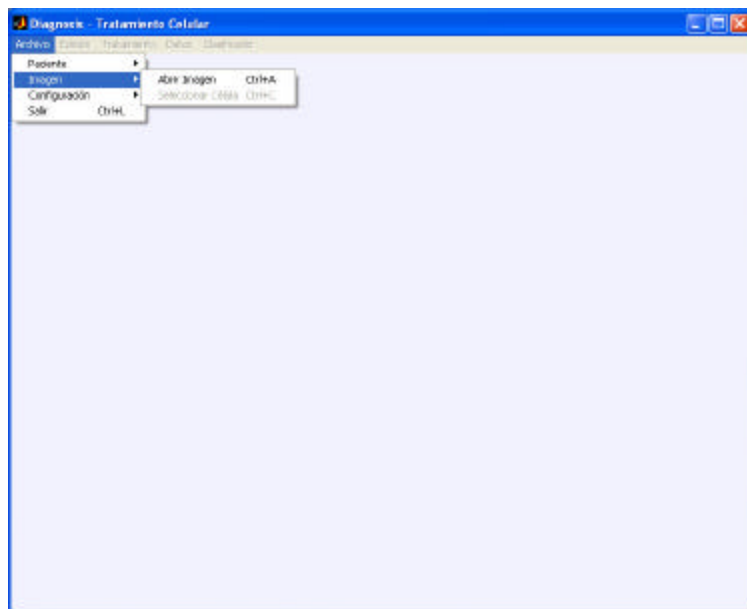
- **Ventana de Identificación.** Si la paciente ha sido registrada deberá ser identificada para realizar la clasificación de las células. Se despliega la ventana mostrada en la figura 30.

Figura 30. Ventana de Identificación



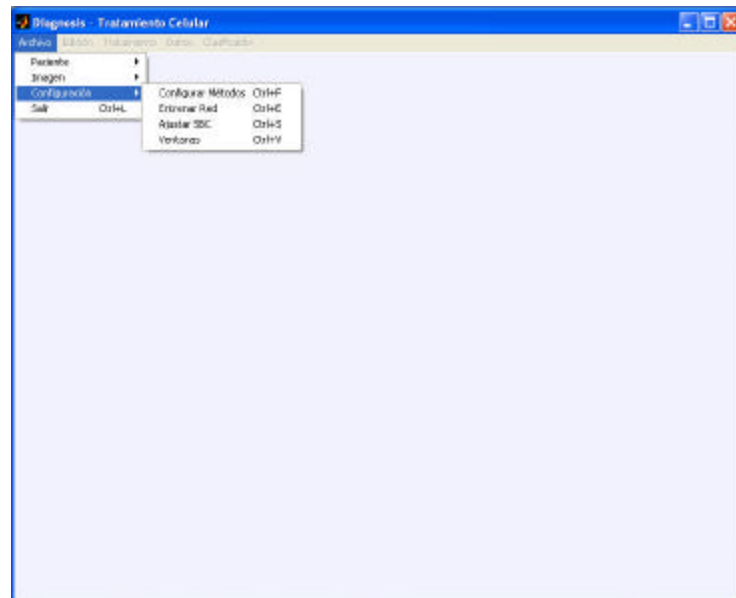
- **Submenú Imagen.** Cuando el paciente ha sido identificado, se habilitan las opciones del submenú imagen: Abrir Imagen, y Seleccionar Célula; presentado a continuación.

Figura 31. Submenú Abrir.



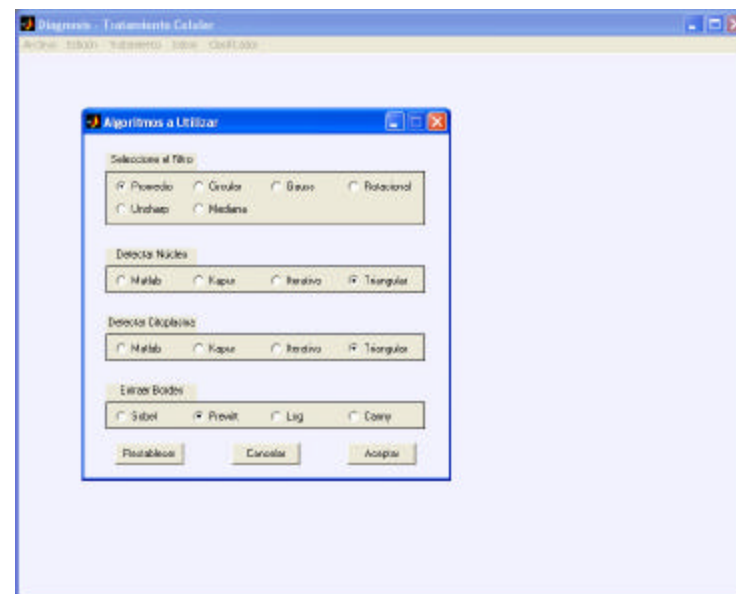
- **Submenú Configuración.** Contiene las opciones para configurar los métodos de segmentación, el entrenamiento de la red neuronal, el ajuste del Sistema Basado en el Conocimiento y el color de las ventanas, de la siguiente manera:

Figura 32. Submenú Configuración.



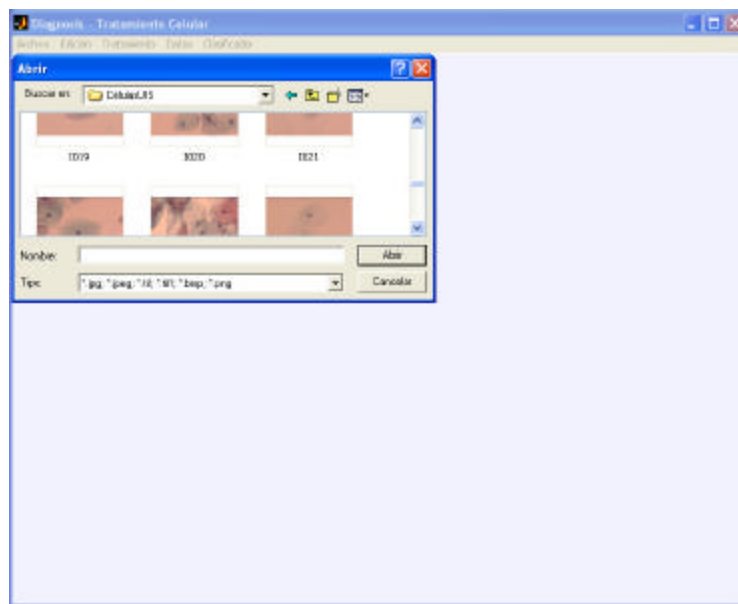
- **Ventana de Configuración.** Contiene las opciones para configurar los métodos de segmentación. Se despliega la ventana que se muestra a continuación.

Figura 33. Ventana de Configuración.



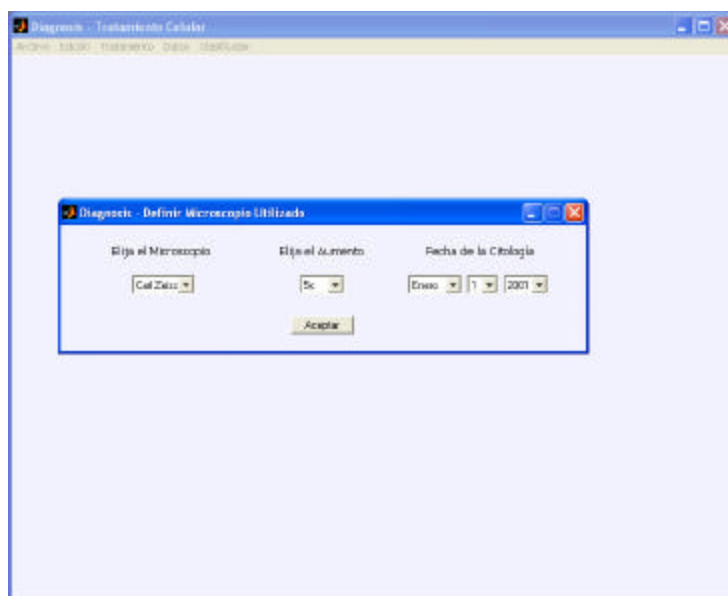
- **Cargar Imagen.** La opción Abrir Imagen permite desplegar una ventana, como se muestra en la figura 34, para localizar la imagen que se desea analizar.

Figura 34. Cargar Imagen.



- **Datos de Adquisición.** Una vez se ha cargado la imagen, se requiere conocer el microscopio y el aumento utilizado para la adquisición, junto con la fecha en que se realizó la citología. Se despliega la ventana presentada en la siguiente figura.

Figura 35. Datos de Adquisición.



- **Visualización de la imagen.** Cuando se ha abierto una imagen, ésta se ubica en la parte superior izquierda de la ventana principal, tal y como se muestra en la figura.

Figura 36. Visualización de la Imagen.



- **Selección.** La célula debe ser seleccionada para la clasificación, esta opción se habilita una vez la imagen ha sido abierta. Como lo presenta la siguiente figura.

Figura 37 Selección de la célula.



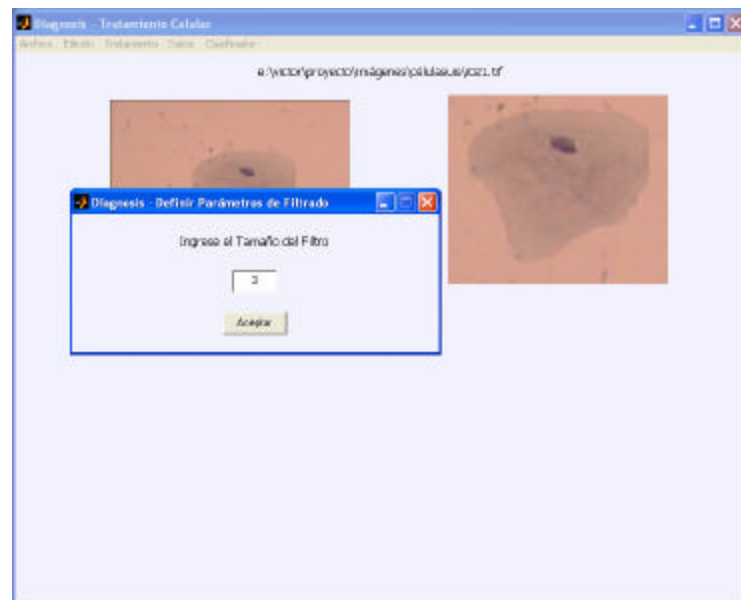
- **Filtrado o Preprocesamiento.** Seleccionada la célula se inicia el proceso que clasificará la célula, el primer paso es filtrar la imagen. La siguiente figura presenta el menú de filtrado.

Figura 38. Menú de Filtrado.



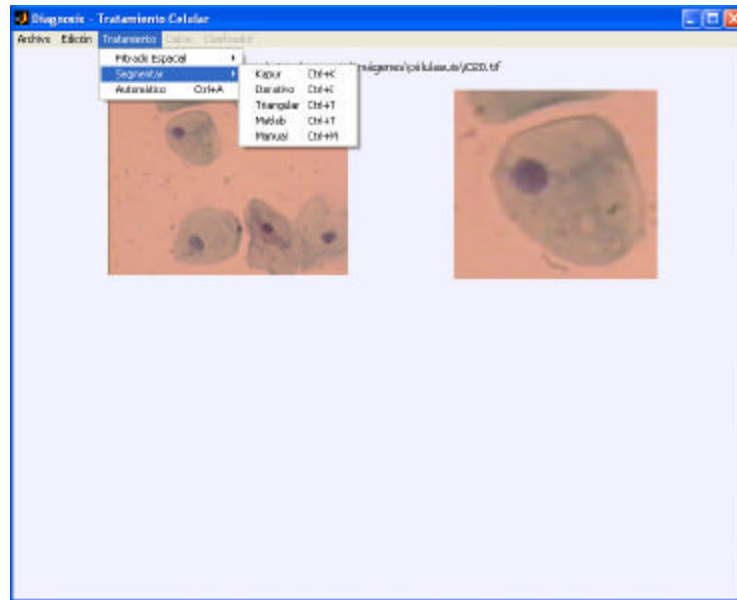
- **Parámetros de filtrado.** Si se selecciona un método de filtrado, se despliega una ventana que permite ingresar los parámetros de ese filtro, como se muestra en la siguiente figura.

Figura 39. Parámetros de filtrado



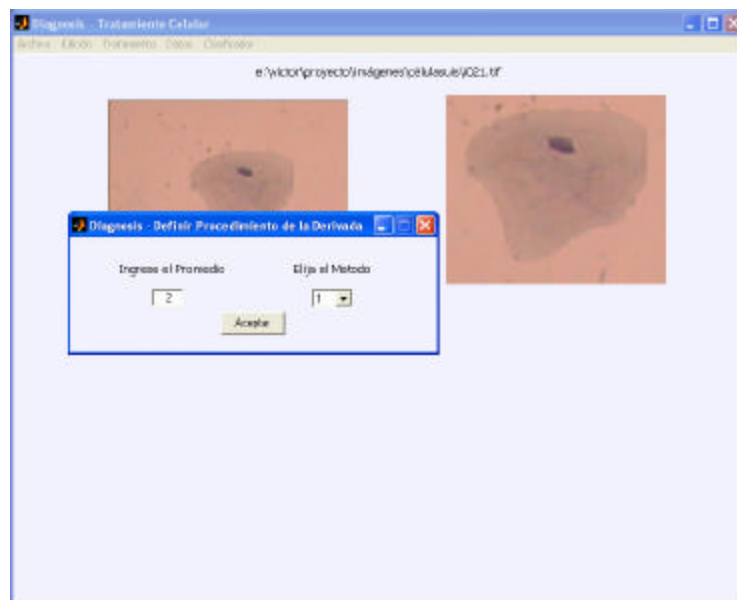
- **Segmentación.** Si la imagen es filtrada se sigue con la segmentación por medio de la elección de uno de los métodos disponibles. La siguiente figura presenta este menú.

Figura 40. Menú Segmentar.



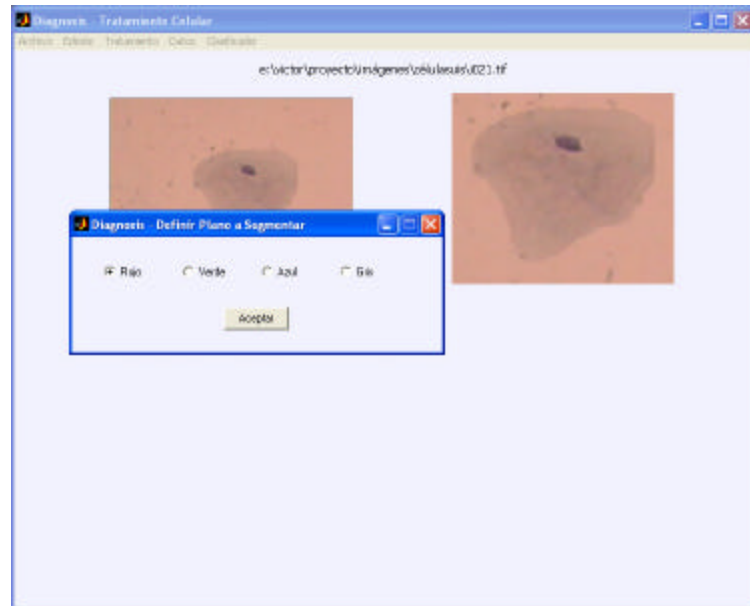
- **Parámetros de segmentación.** Si es elegido el método triangular, se despliega una ventana con el fin de ingresar los datos de este procedimiento, como se muestra en la figura.

Figura 41. Parámetros de segmentación.



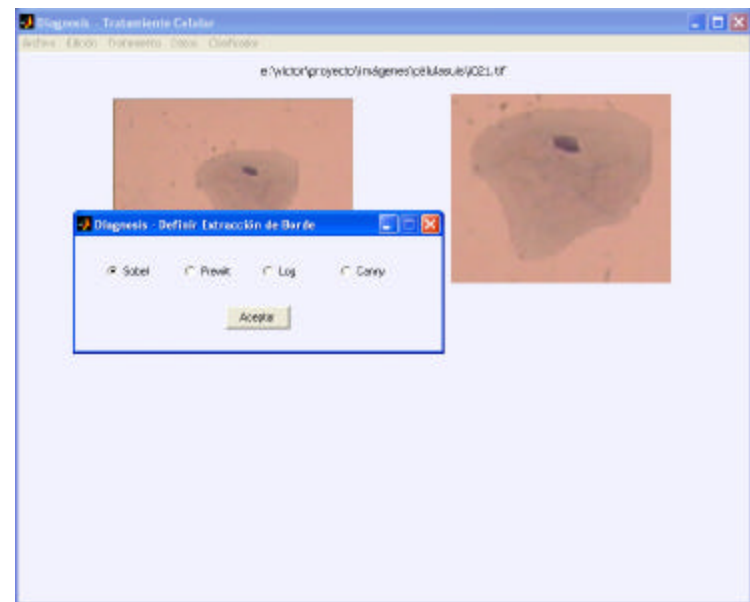
- **Plano de segmentación.** En cada uno de los métodos de segmentación se despliega la ventana que indica el plano a segmentar, presentada en la siguiente figura.

Figura 42. Plano de segmentación.



- **Detección de bordes.** Uno de los resultados de la segmentación es el borde de la célula, la siguiente figura indica la ventana que permite seleccionar el método de extracción.

Figura 43. Extracción de Bordes.



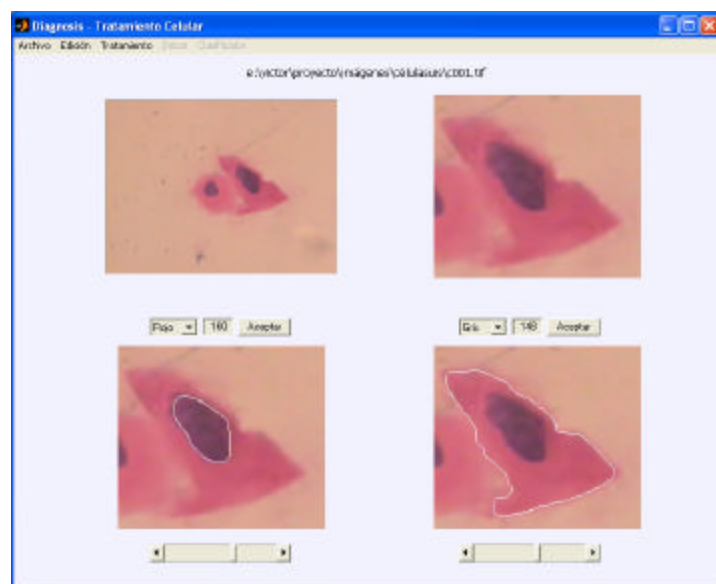
- **Resultado de la segmentación.** Cuando la imagen ha sido segmentada, por los procedimientos paso a paso o por el método automático, el resultado es mostrado en la ventana principal como se revela a continuación. Se visualizan los bordes detectados, los cuales serán presentados a medida que se avanza con los botones de selección.

Figura 44. Resultado de la segmentación.



- **Segmentación manual.** Si la imagen no es segmentada correctamente, por medio de los algoritmos, se elige el método manual; se define el plano y el umbral a utilizar.

Figura 45. Segmentación Manual.



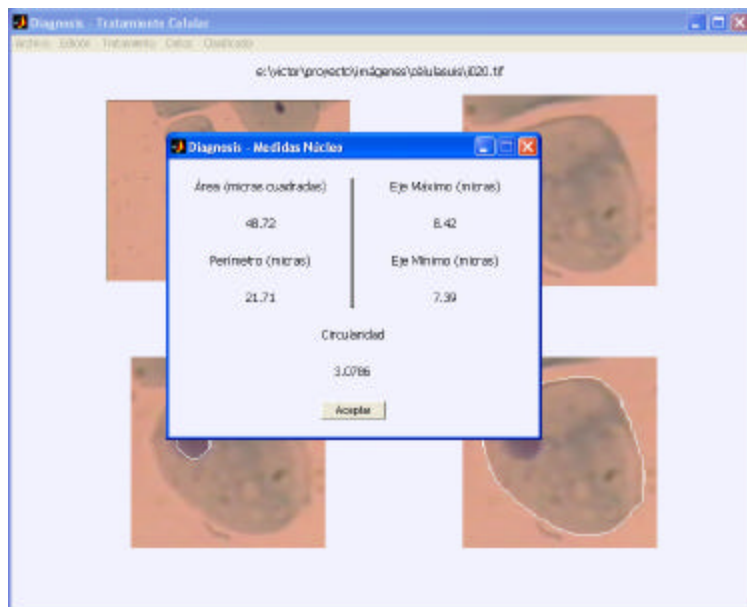
- **Parámetros del núcleo.** Cuando es fijado el núcleo, se activa el menú que permite describir sus características, como se presenta a continuación.

Figura 46. Parámetros del núcleo.



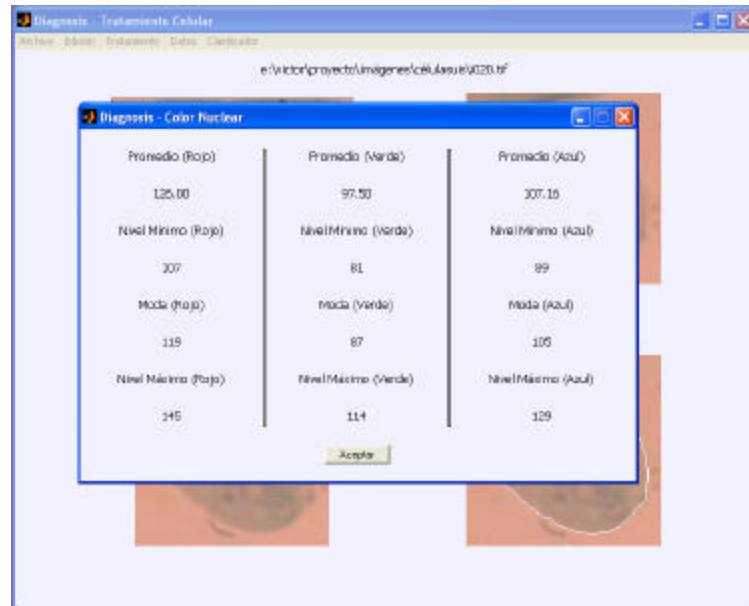
- **Medidas del núcleo.** Esta opción calcula el área, el perímetro, la longitud de los ejes y la circularidad del núcleo. Son visualizados como lo indica la siguiente figura.

Figura 47. Medidas del núcleo.



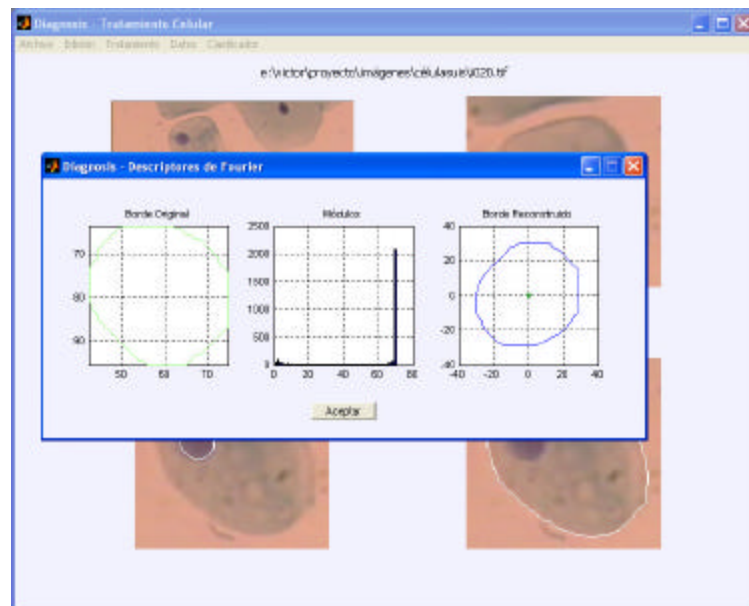
- **Color del núcleo.** Se obtienen los niveles menor, mayor, moda y promedio de cada plano de color, en la región que corresponde al núcleo. Se muestran en la siguiente figura.

Figura 48. Color del núcleo



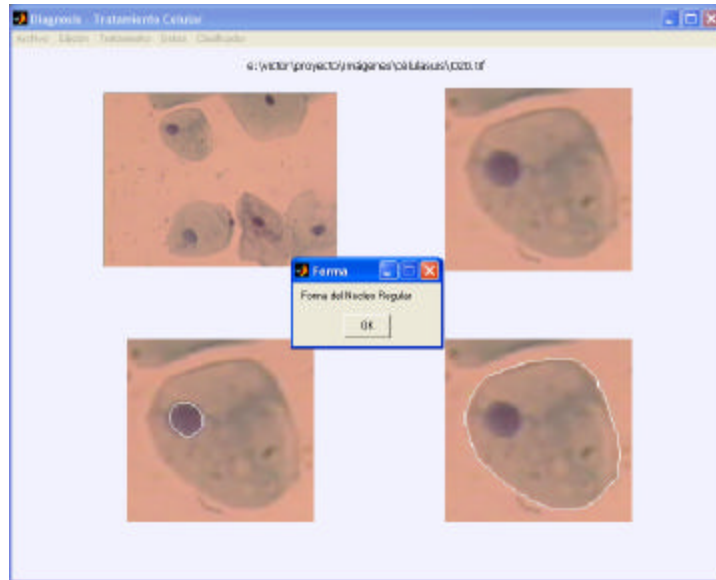
- **Descriptores.** Se obtienen los descriptores de Fourier del borde nuclear. Se presenta el borde original, los módulos y el borde reconstruido, como lo muestra la siguiente figura.

Figura 49. Descriptores de Fourier.



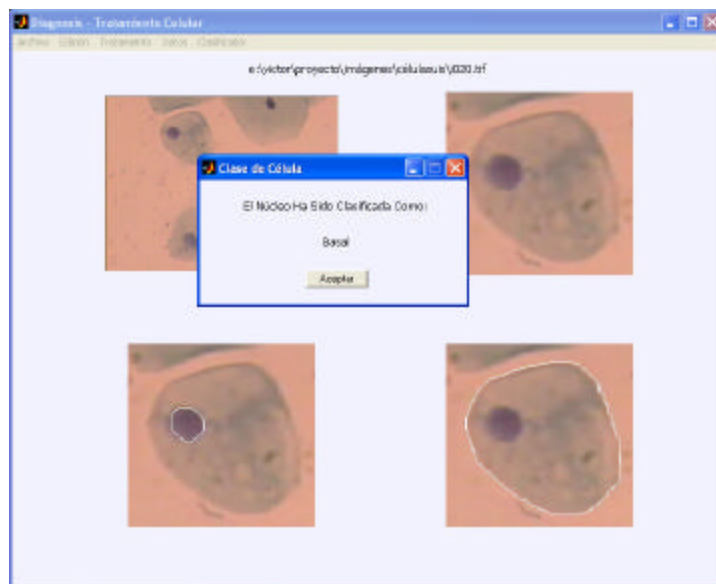
- **Forma del núcleo.** Si se han calculado las medidas y los descriptores se puede determinar si la forma que presenta el núcleo es regular, esto se hace con la red neuronal. Se presenta de la siguiente manera:

Figura 50. Forma del núcleo.



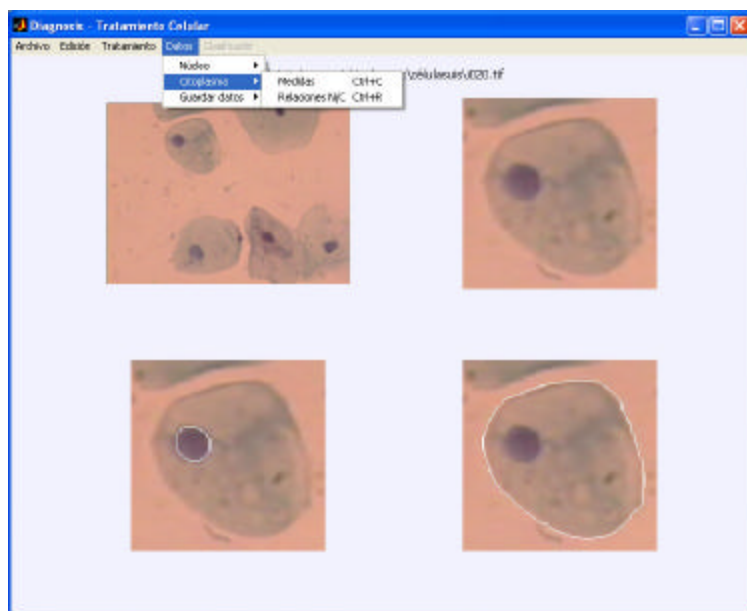
- **Clasificación del núcleo.** Si han sido extraídos todos los parámetros del núcleo, se habilita una opción que permite utilizar el SBC, para la clasificación del núcleo. No se recomienda utilizarla en núcleos ambiguos. La siguiente figura muestra el resultado.

Figura 51. Clasificación del núcleo.



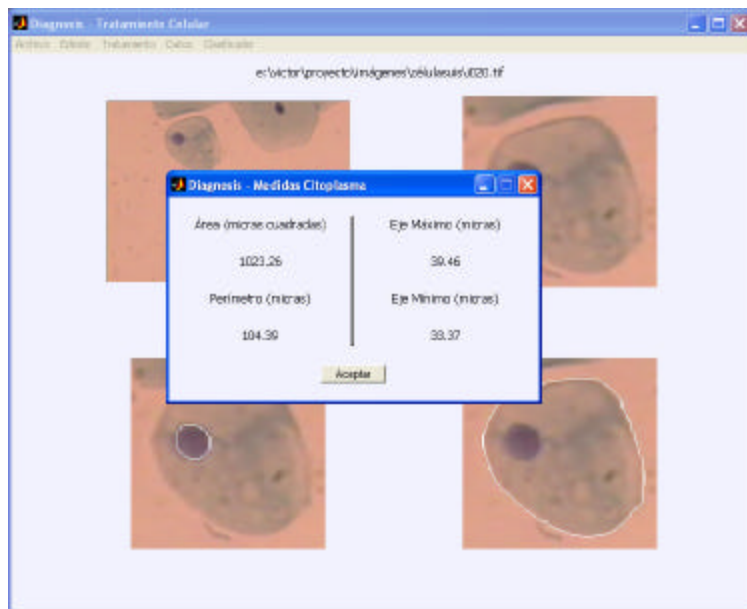
- **Parámetros del citoplasma.** Cuando es fijado el citoplasma, se activa el menú que permite describir sus características, como se presenta a continuación.

Figura 52. Parámetros del citoplasma.



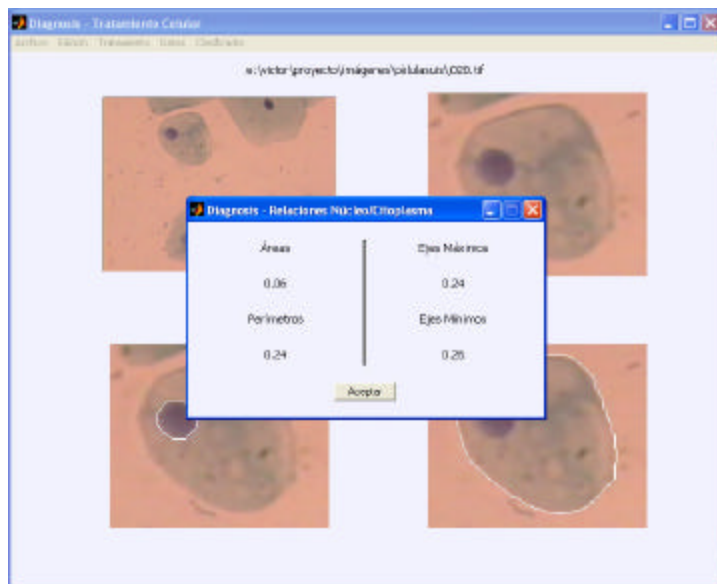
- **Medidas del citoplasma** Esta opción calcula el área, el perímetro, y la longitud de los ejes del citoplasma. Son visualizados como lo indica la siguiente figura.

Figura 53. Medidas del citoplasma.



- **Relaciones Núcleo/Citoplasma.** Cuando se calculan las medidas del núcleo y el citoplasma puede ser halladas las relaciones. Es resultado es mostrado en la siguiente figura.

Figura 54. Relaciones Núcleo/Citoplasma.



- **Guardar datos.** Calculados los parámetros del núcleo o el citoplasma, se habilita una opción que permite guardar en la base de datos la clase de célula que se ha encontrado. Esta opción sólo se recomienda ser utilizada por expertos, se indica en la siguiente figura.

Figura 55. Guardar datos.



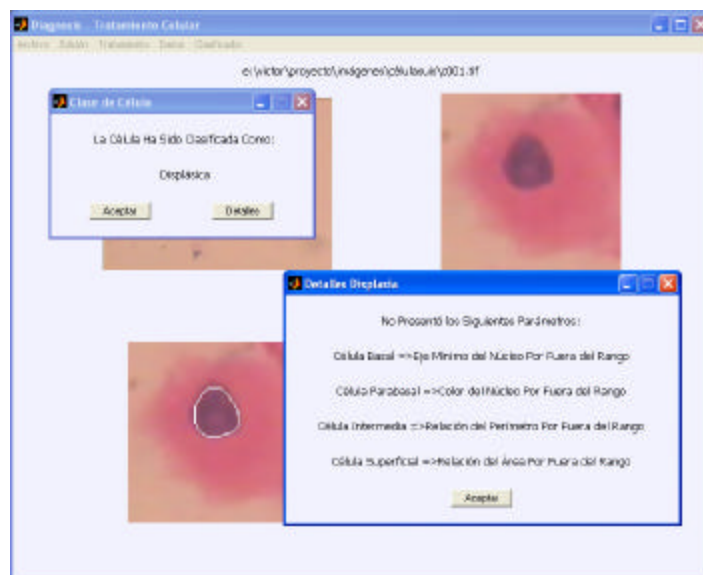
- **Clasificación de la célula.** Si han sido extraídos todos los parámetros del núcleo y el citoplasma, se habilita una opción que permite utilizar el SBC, para la clasificación de la célula. Esta opción es la que permite apoyar el trabajo del experto. La siguiente figura muestra el resultado.

Figura 56. Clasificación de células normales .



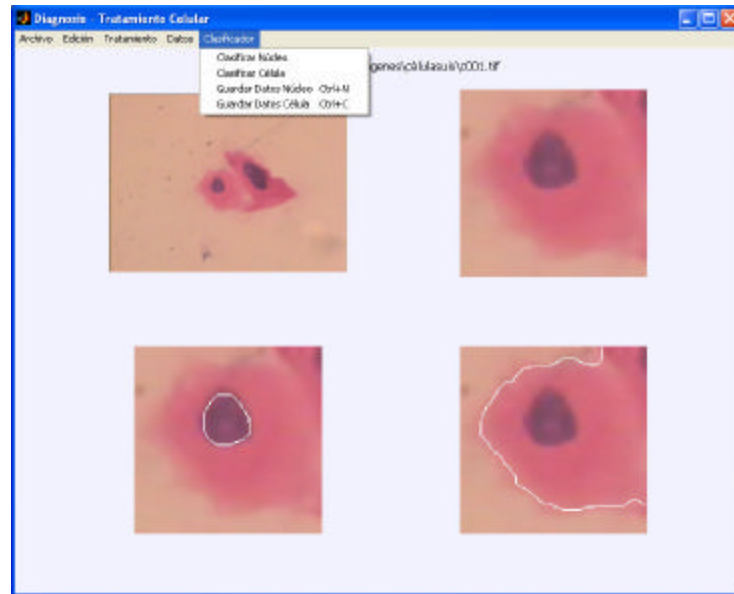
- **Clasificación de células displásicas.** Cuando un células es clasificada de esta manera, es subsistema de explicación permite conocer qué parámetros no se cumplen en este resultado, como lo indica la siguiente figura.

Figura 57. Clasificación de células displásicas .



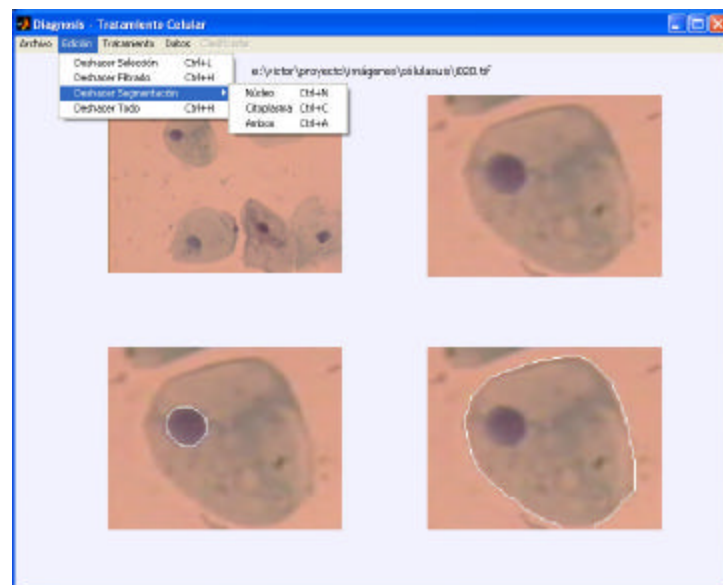
- **Guardar datos de células clasificadas.** Si ha sido utilizado el SBC para la clasificación celular, los resultados pueden ser almacenados, gracias a la activación del menú. La siguiente figura presenta esta opción.

Figura 58. Guardar datos de células clasificadas



- **Menú Edición.** En cualquier instante del proceso los algoritmos pueden ser cancelados, las imágenes borradas y los resultados eliminados, utilizando las opciones del menú Edición; mostrado en la siguiente figura.

Figura 59. Menú Edición.



13. CONCLUSIONES

Se construyó una herramienta software que permite la extracción de características de células escamosas de citologías cérvico uterinas, para apoyar la clasificación entre los parámetros de normalidad existentes.

Se establecieron los criterios correspondientes a cada uno de los parámetros a evaluar en las células superficiales, intermedias, parabasales y basales.

Mediante la utilización de la herramienta software se logra la clasificación de las células escamosas normales y la identificación de los parámetros de las displásicas.

La utilización de las técnicas de tratamiento de imágenes y su complemento con la metodología del Proceso Unificado de Desarrollo de software permitió desarrollar de forma ordenada e iterativa la herramienta software.

Las técnicas de inteligencia artificial empleadas, redes neuronales y sistema basado en el conocimiento, lograron la clasificación de las células a partir de los parámetros extraídos.

Los resultados de las características celulares obtenidos experimentalmente, presentaron coherencia con aquellas estipuladas por los expertos.

14. RECOMENDACIONES

Extraer los parámetros de cada una de las lesiones intraepiteliales estipuladas en el Sistema Bethesda.

Ampliar la Base de Conocimiento con las características de cada una de las lesiones intraepiteliales.

A medida que el conocimiento sea ampliado se hace necesario ajustar el Sistema Basado en el Conocimiento y entrenar la Red Neuronal.

Utilizar arquitectura en paralelo para la construcción de algoritmos de segmentación basados en regiones.

Describir las formas nucleares empleando dimensiones fractales.

BIBLIOGRAFÍA

Abu-Jawdeh GM, Trawinski G, Wang HH. Histocytological study of squamous atypia on Pap smears. Mod Pathol. 1994 Dec; 7(9):920-4.

Abati A, Jaffurs W, Wilder AM. Squamous atypia in the atrophic cervical vaginal smear: a new look at an old problem. Cancer. 1998 Aug 25; 84(4):218-25.

Adad SJ, Souza MA, Etchebehere RM, Saldanha JC, Falco VA, Murta EF. Cyto-histological correlation of 219 patients submitted to surgical treatment due to diagnosis of cervical intraepithelial neoplasia. S Paulo Med J 1999; 117:81-4

ALFARO FERRERES, Luis. GARCÍA ROJO, Marcial. PURAS GIL, Ana María. Manual de Telepatología. Sociedad Española de Anatomía Patológica. Pamplona. 2001. 249 p.

CASTLEMAN. Keneth. R. Digital Image Processing. Prentice Hall. Upper Saddle River, NJ. 1996. 667 p.

CORTÉS GABAUDAN, Francisco. Diccionario Médico Bibliográfico. Universidad de Salamanca. Junio 2004. <http://clasicas.usal.es/dicciomed/>.

CORREDOR MONTAGUT, Martha Vitalia. Principios de Inteligencia Artificial & Sistemas Expertos. Universidad Industrial de Santander. 2000. 88p

Creasman, WT. Disaia, P.J. Oncología ginecológica clínica. Quinta Edición. Mosby Company. USA. 1997.

DEMUTH, Howard. BEALE, Mark. Neural Network Toolbox. User's guide. The Mathworks Inc. 1997. 700 p.

Diccionario General de la Lengua Española Vox. Spes Editorial. 2002. www.diccionarios.com

DOUGHERTY, Edward R. Mathematical Morphology in Image Processing. Marcel Dekker Inc. 1993. 529p.

Fahey MT, Irwing L, Macaskill. Metaanalysis of Pap test accuracy. Am J Epidemiol 1995; 141(7): 680 686.

GIARDINA R, Charles. DOUGHERTY, Edward R. Morphological Methods in Image and Signal Processing. Prentice Hall. 1988. 321 p.

GONZÁLEZ Rafael C, WOODS Richard E. Tratamiento Digital de Imágenes. Addison Wesley / Díaz de Santos.1996. 773p.

Guerrini L, Sama D, Visani M, Cotignoli T, Sintoni C, Maioli P, Lanza G, Schincaglia P. Is it possible to define a better ASCUS class in cervicovaginal screening? A review of 187 cases. Acta Cytol. 2001 Jul-Aug; 45(4): 532-6.

Hakama M, Miller AB, Day NE, eds. Screening for cancer of the uterine cervix. IARC Scientific Publications num. 76. Lyon: International Agency for Research on Cancer, 1986.

Howell LP, Davis RL. Follow-up of Papanicolaou smears diagnosed as atypical squamous cells of undetermined significance. Diagn Cytopathol. 1996 Feb; 14(1):20-4.

HUDSON, Donna L. COHEN Maurice. Neural Networks and Artificial Intelligence for Biomedical Engineering. New York. IEEE Press Series in Biomedical Engineering. 1999. 305 p

HUSSAIN, Zahid. Digital Image Processing. Ellis Horwood. 1991. 406 p.

Image Processing Toolbox. User's guide. The Mathworks Inc. 2003. 831 p.

INSTITUTO NACIONAL DE CANCEROLOGÍA. Guías de práctica clínica en enfermedades neoplásicas. Ministerio de Salud. Bogotá. Colombia 2001.

INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS. Normas Colombianas para la presentación de tesis, trabajos de grado y otros trabajos de investigación. Quinta actualización. Bogotá D.C. ICONTEC. 2002. NTC 1486.

JACOBSON, Ivar. BOOCH, Grady, RUMBAUGH, James. El Proceso Unificado de Desarrollo de Software. Editorial Addison Wesley. 2000.438 p.

JAIN, Anil K. Fundamentals of Digital Image Processing. Prentice Hall. 1989. 569 p.

Janerich D, Hadjimichael O, Schwartz PE, Lowell DM, Meigs W, Merino MJ, Flannery JT, Polednak AP. The screening histories of women with invasive cervical cancer. Connecticut. Am J Public Health 1995; 85: 791-794

LASPRILLA, María del Carmen. TORRES, Yezid. Caracterización y Pre-Clasificación de Zonas de Interés en Fotografía Aérea. Tesis de Postgrado. Bucaramanga 1989. 114p.

BIBBO, Marluce. Comprehensive Cytopathology. Second edition. W.B. Saunders

Pitman MB, Cibas ES, Powers CN, Renshaw AA, Frable WJ. Reducing or eliminating use of the category of atypical squamous cells of undetermined significance decreases the diagnostic accuracy of the Papanicolaou smear. Cancer. 2002 Jun 25; 96(3):128-34.

Pundel JP, Lichtfus C: La pycnose nucleaire des cellules vaginales. Bull Soc R Belge Gynecol Onstet 26:630,1956.

Raab SS, Steiner AL, Hornberger J. The cost-effectiveness of treating women with a cervical vaginal smear diagnosis of atypical squamous cells of undetermined significance. Am J Obstet Gynecol 1998; 179:411-20.

REY SERNA, Adrian Marcela. CASTILLO CASTELBLANCO, Sergio Fernando, TORRES, Yezid. Caracterización y procesamiento digital de imágenes celulares. Tesis de grado. Bucaramanga. 1990. 87 p.

Robert L. Ehrmann, m.d. Bening to malignant progression in cervical squamous epithelium. Iguaku – Shoin Medical Publishers Tokyo 1994

SÁNCHEZ NOGUERA, Jorge Mario. SOTAQUIRÁ GUTIERREZ Miguel Angel. GUALDRÓN GONZÁLEZ, Oscar. VARGAS, Clara Inés. Obtención de Cariotipos Mediante Procesamiento de Imágenes. Tesis de Grado. Bucaramanga. 2001. 78 p.

Sawaya George F, Brown Adalstein, Washington A.E, Garber, A. Current approaches to cervical – Cancer Screening. N Engl J Med Vol 344, Nº 21 Mayo 24 .2001 1603 -1607.

Sigurdsson, K. Quality assurance in cervical cancer screening: the Icelandic experience 1964-1993. Eur J Cancer 1995; 31A(5): 72834.

Solomon D, Davey D, Kurman R, Moriarty A, O'Connor D, Prey M, Raab S, Sherman M, Wilbur D, Wright T Jr, Young N; Forum Group Members; Bethesda 2001 Workshop. The 2001 Bethesda System: Terminology for reporting results of cervical cytology. JAMA 2002; 287:2114-9.

Solomon D, Frable WJ, Vooijs GP, Wilbur DC, Amma NS, Collins RJ, Davey DD, Knight BK, Luff RD, Meisels A, Navin J, Rosenthal DL, Sauer T, Stoler M, Suprun HZ, Yamauchi K. ASCUS and AGUS criteria. International Academy of Cytology Task Force summary. Diagnostic Cytology Towards the 21st Century: An International Expert Conference and Tutorial. Acta Cytol. 1998 Jan-Feb;42(1):16-24

Stoler MH. New Bethesda terminology and evidence-based management guidelines for cervical cytology findings. JAMA 2002; 287:2140-1.

Stoler MH, Schiffman M. Interobserver reproducibility of cervical cytologic and histologic interpretations: realistic estimates from the ASCUS-LSIL Triage Study. JAMA 2001; 285:1500-5.

TRISTANCHO CELIS Luz Stella, CASTRILLÓN DÍAZ, Luis Eduardo, GARCÍA AYALA, Ernesto, MENDOZA CASTELLANOS, Alfonso. SUESCÚN, Pedro. Diseño e implementación de un sistema software de visión artificial para el reconocimiento del grado de avance del cáncer en células de muestras citológicas del cérvix uterino basado en el criterio del agrandamiento del núcleo. Tesis de grado. Bucaramanga. 1995. 120 p.

UNIVERSIDAD AUTÓNOMA DE BUCARAMANGA. Registro Poblacional de Cáncer del Área Metropolitana de Bucaramanga. 2000 – 2001.

ÍNDICE

Adquisición, 28, 35
Área, 49
Base de datos, 55, 56, 58
Binarización, 29, 41
Cámara, 35
Cáncer de cérvix, 19, 20, 26
Células, 34, 69
basales, 25, 59
displásicas, 26, 54, 79
intermedias, 24, 61
parabasales, 25, 60
superficiales, 24, 62
Cérvix, 19, 23
Citología cérvico uterina, 19, 23, 34
Circularidad, 49
Descriptores, 31
de Fourier, 31, 49, 75
Geométricos, 31
Digitalizador, 28
Ejes, 49
Euplasia, 26
Extracción de bordes, 47
canny, 48
kirsch, 48
laplaciano del gaussiano, 48, 72
prewitt, 29, 48
sobel, 29, 48
Filtrado, 29, 38, 70
en el dominio espacial, 29
en el dominio de la frecuencia, 29
gauss, 39
mediana, 29, 39
promedio, 29, 38
promedio circular, 38
rotacional, 38
unsharp, 39
Hipercomatismo, 27, 28, 59
Histograma, 29, 44
Lenguaje Unificado de Modelado, 32
Lesión intraepitelial escamosa, 27
de bajo grado, 27
de alto grado, 27
Microscopio, 35, 36, 68
Modelo Computacional, 57, 58
Morfología Matemática, 30, 46
apertura, 30, 47
cierre, 30, 47
dilatación, 30

- erosión, 30
- Niveles de color, 49, 75
- Neoplasia Intraepitelial Cervical, 27
- Parámetros, 25, 59, 60, 61, 62, 74, 77
- Perímetro, 49
- Pleomorfismo, 26
- Preprocesamiento, 29, 70
- gauss, 39
- mediana, 29, 39
- promedio, 29, 38
- promedio circular, 38
- rotacional, 38
- unsharp, 39
- Proceso Unificado de Desarrollo de Software, 32
 - análisis, 33
 - casos de uso, 32, 33
 - construcción, 32
 - diseño, 33
 - elaboración, 32
 - implementación, 33
 - inicio, 32
 - modelo de análisis, 33
 - prueba, 33
 - requisitos, 32
 - transición, 32
- Proplasia, 27
- Reconocimiento, 51
- Red neuronal, 51
 - backpropagation, 51
 - cascade-forward, 52
- Relación N/C, 25, 26, 27, 50, 78
- Retroplasia, 27
- Segmentación, 29, 41, 71
- Sistema Basado en el Conocimiento, 53, 58
 - base de conocimiento, 53
 - motor de inferencia, 53
 - subsistema de explicación, 54
- Sistema Bethesda, 27, 34
- Tratamiento digital de imágenes, 28
- Trasformada de Fourier, 31
- Umbralización, 29, 41
 - iterativo, 42
 - matlab, 45
 - manual, 46, 73
 - pun-kapur, 41
 - triangular, 44
- VPH, 27

ANEXO A

Requerimientos del Sistema

Para la correcta ejecución de la herramienta software se hace necesario que el equipo de cómputo donde será instalada, cumpla las siguientes especificaciones:

- Procesador: Intel Pentium III, 1.0 Ghertz o superior.
- Memoria RAM: 256 MB o superior.
- Tarjeta de vídeo: 32MB o superior.
- Resolución del monitor: 1024 x 768 píxeles.
- Tarjeta de adquisición de vídeo: Mpeg TV Station / USB.
- Sistema Operativo: Microsoft Windows XP.
- Plataforma: Matlab 6.5.
- Toolbox instalados: Signal Processing, Filter Design, Image Processing, Neural Network, Database.
- Microsoft Office XP: Base de Datos Access.

Si no se cumple alguno de estos requisitos, es posible que la herramienta no pueda ser utilizada, no se reconozcan las funciones, o se produzcan bloqueos en el sistema operativo.

ANEXO B

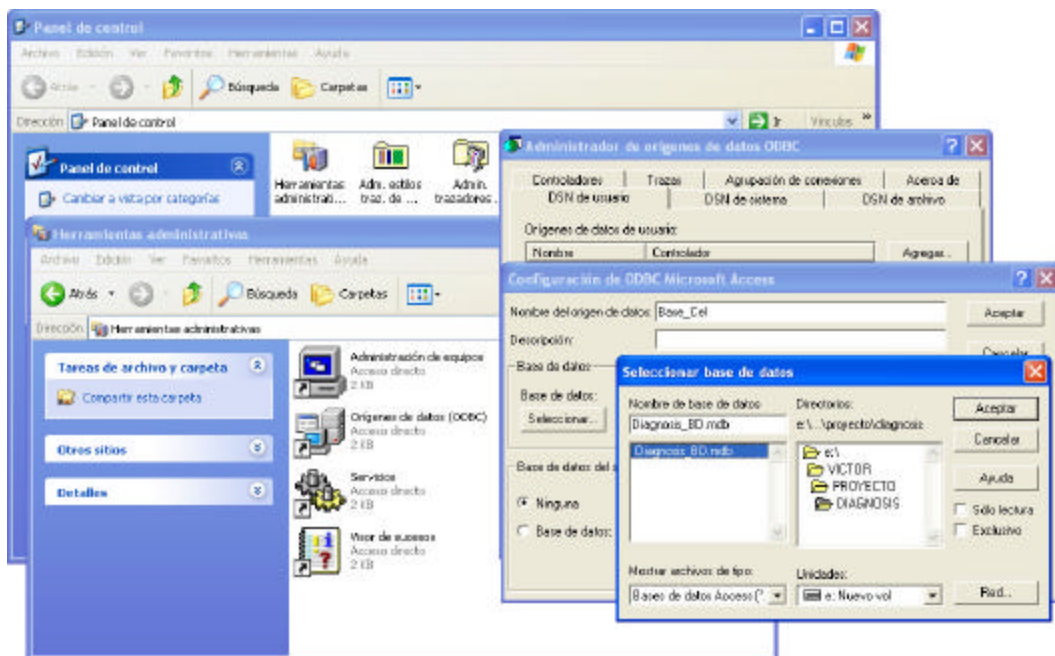
Instalación de la Herramienta Software

La herramienta software se compone de un conjunto de funciones, almacenadas en archivos, que deben ser instalados en un directorio para su ejecución. Los pasos a seguir para este procedimiento son los siguientes:

- Ejecutar el archivo diagnosis.exe, se solicitará la carpeta donde se desea descargar las funciones, que pueden ser almacenadas en un directorio personal.
- Ir a panel de control, herramientas administrativas, orígenes de bases de datos, DSN de usuario y agregar.
- Se elige controlador de Microsoft Access, en el nombre del origen de datos se escribe "Base_Cel"; con el botón seleccionar se localiza, en la carpeta donde han sido almacenadas las funciones, el archivo "Diagnosis_DB.mdb".

Finalizados estos pasos la herramienta podrá ser utilizada conforme al capítulo 12. La siguiente figura presenta las ventanas del sistema, para la selección del origen de base de datos (ODBC).

Figura 60. Selección del Origen de Base de Datos (ODBC).



ANEXO C

Sistema Bethesda de Clasificación

El sistema Bethesda ha sido diseñado para estandarizar cada una de las clases de células, hallazgos y lesiones intraepiteliales, encontradas en una citología cérvico uterina, es necesario completar la siguiente información:

ESPECIMEN ADECUADO

Satisfactorio para evaluación (presencia/ausencia de transformación endocervical)

Insatisfactorio para evaluación (especificar razón)

 Especimen rechazado/ no procesado (especificar razón)

 Especimen procesado y examinado pero insatisfactorio para anormalidad epitelial por (especificar razón)

CATEGORIZACIÓN GENERAL (Opcional)

Negativo para lesión intraepitelial o malignidad

Anormalidad epitelial

Otro

INTERPRETACIÓN/RESULTADOS

Negativa para Lesión intraepitelial o Malignidad

Organismos

 Trichomona Vaginalis

 Organismos morfológicamente consistentes con Candida species

 Cambio en la flora sugestivo de vaginosis bacteriana

 Bacteria morfológicamente consistente con Actinomyces species

 Cambios celulares consistentes con virus Herpes simples

Otros hallazgos no sugestivos de neoplasia

 Cambios celulares reactivos asociados con

 Inflamación

 Radiación

 DIU

 Estado celular glandular posthisterectomía

 Atrofia

Anormalidades de las células epiteliales

Células Escamosas

 Células escamosas atípicas (ASC)

 De significado incierto (ASC-US)

 No se puede excluir lesión intraepitelial de alto grado –LIEA/HSIL (ASC-H)

 Lesión intraepitelial de Bajo grado LIEB/LSIL

 VPH-displasia moderada- NIC 1

 Lesión Intraepitelial de Alto grado LIEA/HSIL

 Displasia moderada y severa, carcinoma in situ; NIC 2 y NIC 3

 Carcinoma escamocelular

Células Glandulares

 Células Glandulares Atípicas (AGS) (especificar endocervical, endometrial, o sin especificar)

 Células Glandulares atípicas, favor neoplástico (especificar)

 Adenocarcinoma Endocervical in situ (AIS)

 Adenocarcinoma

Otra

 Células endometriales en una mujer >40 años

Tomado de Bethesda System 2001

ANEXO D

Microscopio Adicional

La consecución de un nuevo microscopio, que integra los dispositivos necesarios para la adquisición, simplifica este procedimiento.

Con pruebas realizadas en este dispositivo se llegó a la conclusión de utilizar el aumento de 100x; sin embargo, este equipo fue conseguido en la fase de transición, por consiguiente, no se lograron adquirir un número de muestras suficientes para la extracción de parámetros.

En las condiciones actuales, la herramienta no permite analizar el parámetro de la coloración del núcleo, debido a que la iluminación es diferente.

Se recomienda realizar un análisis de la coloración nuclear. La siguiente figura presenta una imagen adquirida con este modelo.

Figura 61. Imagen adquirida con el nuevo microscopio que integra las funciones de adquisición.

