

DIAGNOSIS 1.0: SOFTWARE PARA LA EXTRACCIÓN DE PARÁMETROS DE CÉLULAS DEL CÉRVIX UTERINO A TRAVÉS DEL PROCESAMIENTO DIGITAL DE IMÁGENES

LUIS EDUARDO CASTRILLÓN
Ingeniero de Sistemas
Universidad Industrial de Santander

LUZ STELLA TRISTANCHO
Ingeniera de Sistemas
Universidad Industrial de Santander

ERNESTO GARCIA AYALA
Director Departamento de Patología
Universidad Industrial de Santander

ALFONSO MENDOZA CASTELLANOS
Grupo de Bioingeniería
Profesor Titular, Escuela de Ingeniería de Sistemas
amendoza@uis.edu.co

RESUMEN

El presente trabajo apoya el diagnóstico preliminar del cáncer del cervix uterino. Determina el estado (normal, displásico o canceroso) en células individuales de muestras citológicas utilizando un sistema de visión artificial. El reconocimiento del estado en las células del cervix uterino se realiza en dos pasos básicos, a saber : Inicialmente se debe clasificar la célula objeto de análisis, como superficial, intermedia o parabasal, tomando como base su aspecto (color, forma, tamaño, etc.). Posteriormente, se detecta cuales criterios de malignidad o premalignidad se encuentran presentes, con el fin de hacer el diagnóstico correspondiente en la muestra.

PALABRAS CLAVE: Procesamiento de imágenes, morfología: erosión, dilatación, detección de bordes, Células: superficial, intermedia y parabasal.

INTRODUCCIÓN

El cáncer del cuello uterino continua siendo la segunda causa de muerte por neoplasias tumorales malignas en mujeres de países en vía de desarrollo como el nuestro. En las áreas más industrializadas del mundo dicha tasa ha disminuido; pero en contraste con el descenso en la mortalidad, la frecuencia con la que se diagnostican los procesos cancerosos y precancerosos es muy elevada[21]. Esta últimas lesiones pueden ser eliminadas a su vez mediante un tratamiento adecuado y oportuno. En gran parte estos buenos resultados dependen de la eficacia de

la citología cervico uterina descrita inicialmente por Papanicolaou en 1942 la cual permite hacer el diagnóstico de las lesiones cervicales precancerosas en una buena proporción de casos. Sin embargo, la sensibilidad de la citología es del 85% lo cual nos habla de un 15% de falsos negativos[21], con errores de detección e interpretación en la lectura de dichos extendidos.

El Objetivo fundamental de este trabajo fue que a través de la extracción de parámetros de células de cervix uterino mediante el procesamiento digital de imágenes se puede mejorar el diagnóstico de dichas entidades.

MATERIALES Y MÉTODOS

Durante el desarrollo del trabajo, se tomó como base el procedimiento para diagnosticar el cáncer, actualmente utilizado en el Departamento de Morfopatología de la Facultad de Salud de la Universidad Industrial de Santander, con el fin de establecer un delineamiento del trabajo experimental que sirviera para definir los criterios de base para el diseño e implementación de un prototipo.

En primera instancia la citología cervico uterina es tomada en el Servicio de Ginecología o de consulta externa del Hospital Universitario Ramón González Valencia y enviada al Departamento de Patología. Después de ser coloreada el patólogo o la citóloga procede a clasificar la célula que esta analizando. Para ello se tiene en cuenta en primera instancia el color que la célula exhibe como reacción a la coloración utilizada en la preparación de la placa, verifican su forma (redondeada, ovalada, poliédrica o con dobleces citoplasmáticos), tamaño de la célula (núcleo y citoplasma), de lo anterior y salvo contadas excepciones se puede concluir si la célula es en primera instancia de origen epitelial y si así fuere si corresponde a una célula escamosa o a una célula endocervical (glandular). Una vez se define que se trata de una célula epitelial escamosa se procede a diagnosticar de cual de los estratos del epitelio escamoso estratificado proviene; es decir, si se trata de una célula superficial, intermedia, parabasal o basal.

Una vez se define de cual tipo de célula se trata, se procede a través de la observación directa de las muestras a determinar su estado, es decir, si se trata de una célula normal, con alteraciones preneoplásicas (diplosia o neoplasia intraepitelial cervical) o cancerosa; de acuerdo a los criterios morfológicos citopatológicos de displasia y malignidad. Estos incluyen la anisonucleosis (diferentes formas y tamaños), pérdida de la relación normal entre el tamaño del núcleo y el citoplasma a expensas del tamaño del núcleo, hiper cromatismo, irregularidades de la membrana nuclear y presencia del nucleolo. En condiciones normales para hacer este diagnóstico el citopatólogo debe evaluar la totalidad de las células presentes en el extendido.

En la actualidad el Departamento de Patología no cuenta con una infraestructura que le permita medir el tamaño de los núcleos y del citoplasma de las células, en ningún momento los especialistas calculan o hallan dichos valores con el fin de emitir su diagnóstico, sino que este se basa sobre imágenes mentales que se han fijado a través de su experiencia con la manipulación de dichas muestras.

Tratando de seguir el mismo procedimiento para apoyar el reconocimiento de la displasia o el cáncer, se pensó en la factibilidad de utilizar los valores de las áreas de los núcleos y citoplasmas y su relación, para determinar la clase de célula y su estado, ante la imposibilidad, en primera instancia, de determinar el color (imágenes en tonos de gris) y las formas angulares debido a las limitaciones impuestas por el hardware, y la naturaleza de las muestras que se procesan (las células se presentan traslapadas y arrolladas sobre si mismas, lo que dificulta la obtención de contornos.).

Inicialmente se desarrollaron tres algoritmos para obtener dichas mediciones, ellos son:

- Rectángulo mínimo: que consiste en tomar el valor del área del menor rectángulo que encierra el objeto de interés (núcleo, citoplasma), ello se hizo encontrando las coordenadas de la mencionada zona, usando la demarcación hecha por el usuario a través del botón izquierdo del ratón.
- Área del Polígono: este procedimiento se basa en una serie de puntos del contorno de la zona a medir, dichos puntos se encuentran a través de las pulsaciones que el usuario hace del botón derecho del ratón; se almacenan en un archivo y posteriormente se utiliza el algoritmo matemático para calcular el área de un polígono irregular.
- Conteo de Pixeles: Como su nombre lo indica, a través de este procedimiento, se recorría la imagen seleccionada pixel a pixel y se cuentan los pixeles, de esta manera se obtiene el área del objeto deseado.

Luego de un análisis detallado y de las pruebas sobre las imágenes de los algoritmos anteriores, se detectaron fallas en la precisión de ellos así:

El algoritmo del rectángulo mínimo si bien presentaba un buen grado de aproximación en el cálculo de las áreas de los núcleos de las células normales, para el caso de los núcleos de células anormales y el de los citoplasmas de células de forma irregular (alargadas, en forma de estrella, etc.) los porcentajes de error fueron elevados, alrededor del 50%, lo cual motivó su eliminación.

Para el método del conteo de pixeles, se descubrió que algunas células presentan zonas claras o de color blanco al interior de las áreas de estudio (núcleo, citoplasma), por lo que el valor arrojado por el procedimiento era incorrecto, se implementó una variación al procedimiento inicial de la siguiente manera:

Primero se hace un recorrido de la imagen con el fin de ubicar los puntos del contorno, los cuales se van grabando en un archivo, de igual manera se almacenan los valores máximos y mínimos en las cuatro direcciones.

Se conserva el recorrido pixel a pixel por el área de interés (marcada por los máximos y mínimos previamente almacenados) y el conteo de los píxeles negros; al momento de encontrar un pixel blanco, se verifica si hay un pixel del contorno por encima, por debajo, a la derecha y a la izquierda de dicho punto, en cuyo caso se contabiliza como elemento integrante del área.

Después de dichas modificaciones, se encontró que el método arrojaba el área real de la imagen, funcionó perfectamente para los núcleos, pero al momento de hacer las pruebas sobre los citoplasmas, se descubrió que el procedimiento era muy lento, debido a las múltiples lecturas que debía realizar al archivo (sobre todo en las células más grandes).

Por último se optó por el método del polígono, aumentando el número de puntos, se decidió recorrer el objeto por su contorno y almacenar dichos puntos en un archivo, posteriormente, se toman dichos puntos como los vértices de un polígono irregular, con los cuales se calcula el área.

La caracterización celular se consigue al momento de calcular los valores de las áreas del núcleo y del citoplasma. Una vez lograda dicha caracterización, se debe proceder a reconocer la presencia de los criterios de malignidad[20], el presente trabajo se ocupa del reconocimiento de dos de los once criterios de malignidad existentes, son ellos:

- Agrandamiento del núcleo
- Aumento de la relación nuclear/citoplasmática

UBICACIÓN DE LA INVESTIGACIÓN EN EL PROBLEMA

Debido a las limitaciones impuestas por el hardware disponible, tales como: cámara y tarjeta digitalizadora de video monocromática[15], microscopios[12] ópticos, etc, el sistema se encarga de hacer un diagnóstico preliminar basado en el criterio del agrandamiento del núcleo, estableciendo la clase de célula y su estado a partir de los datos de su caracterización y a través de su comparación con los valores obtenidos experimentalmente.

ARQUITECTURA DEL SISTEMA

A. Recursos utilizados

La Figura 1. Muestra la arquitectura básica necesaria para la ejecución del sistema software.

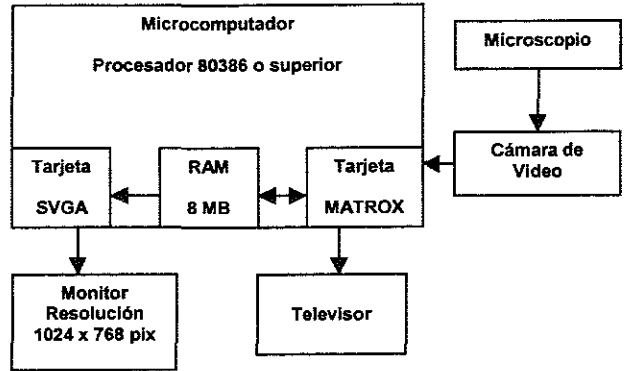


Figura 1. Plataforma Hardware utilizada por Diagnosis 1.0

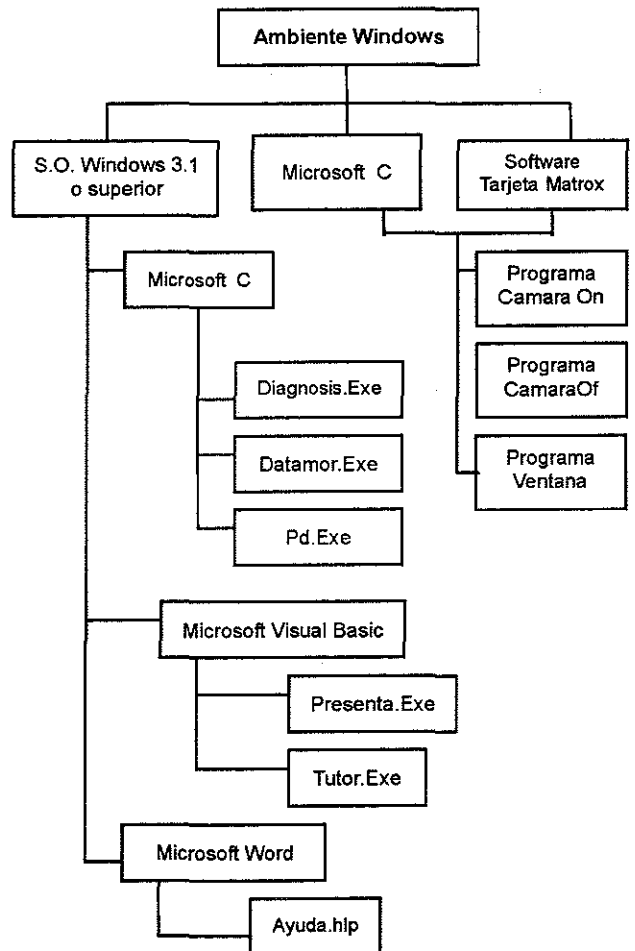


Figura 2. Arquitectura del Software Desarrollado

Se utilizaron un microscopio NIKON SE_DB, una cámara de video HAVELIN CMOS JE3362, una tarjeta de procesamiento de imágenes matrox, un computador PC, un monitor y un televisor a color.

B. Arquitectura del software desarrollado

La Figura 2, muestra la arquitectura del software diseñada para el desarrollo del sistema.

C. Diseño del software

Dadas las condiciones del proyecto, se diseñaron 3 módulos, ver Figura 3 .

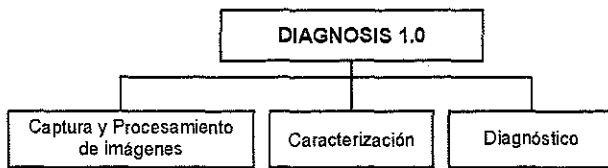


Figura 3. Estructura general del programa Diagnosis 1.0.

Las imágenes obtenidas mediante el módulo de captura y procesamiento, se almacenan en archivos binarios a 256 tonos de gris.

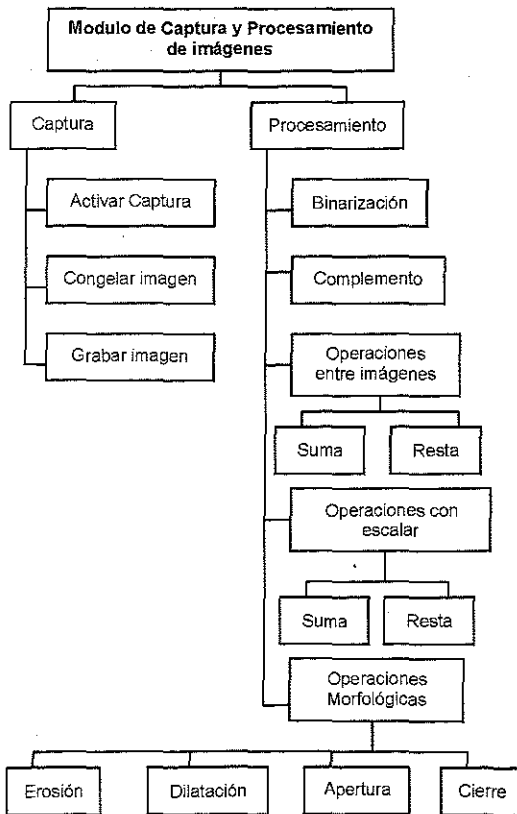


Figura 4. Estructura general de las rutinas disponibles en el módulo de captura y procesamiento de imágenes.

Con el fin de mejorar estas imágenes antes de hacer la extracción de sus características, se desarrollaron los siguientes algoritmos básicos de procesamiento para su mejoramiento[1-3,5-7,9-11,11,13,14,16-19]. Ver figura 4.

• Binarización



Figura 5. Se puede apreciar una imagen en tonos de gris que se ha binarizado con dos umbrales diferentes, el primero corresponde a 190 y el segundo a 115.

• Complemento

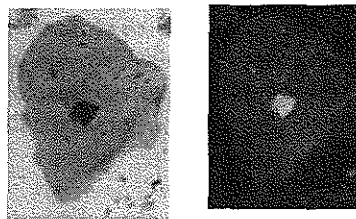


Figura 6. Imagen en tonos de gris y su complemento.

• Suma y resta de imágenes (con escalar)

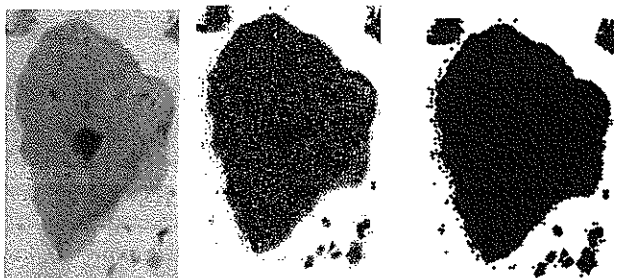


Figura 7. La ilustración permite apreciar una imagen en tonos de gris, a la cual se le ha practicado la operación con escalar, en el primer caso sumándole el escalar 5 y en el segundo caso restándole el mismo escalar a la imagen original.

• Morfología matemática: Erosión, Dilatación, Apertura y Cierre

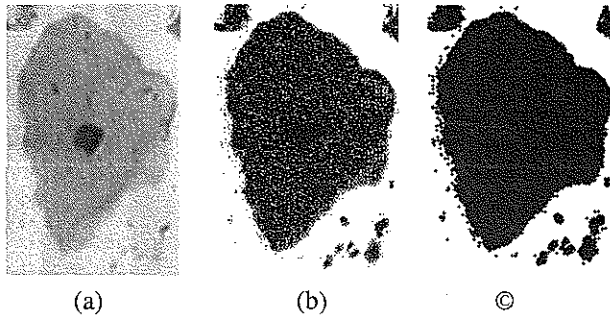


Figura 8. Incidencia de las operaciones morfológicas sobre las imágenes. (a) Imagen en tonos de gris. (b) Imagen binarizada con umbral de 190. (c) Imagen dilatada con elemento estructurante circular de 3x3.

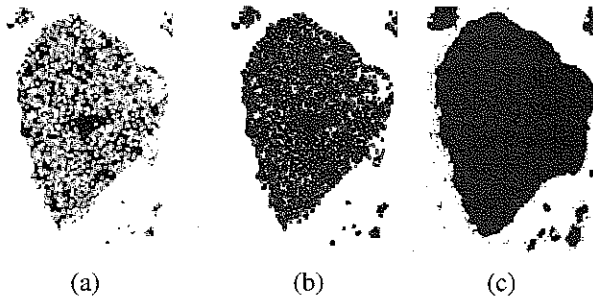


Figura 9. Efecto de las operaciones morfológicas sobre las imágenes, utilizando elemento estructurante circular de 3 x 3. (a) Erosión. (b) Apertura. (c) Cierre.

- Detector de bordes
- Cálculo de área utilizando el método del polígono.

El módulo de caracterización y diagnóstico, tratando de seguir el mismo procedimiento para apoyar el reconocimiento del cáncer, se pensó en la factibilidad de utilizar los valores de las áreas de los núcleos y citoplasmas y su relación, para definir la clase de célula y su estado, ante la imposibilidad de determinar el color (imágenes en tonos de gris) y las formas angulares debido a las limitaciones impuestas por el hardware, y la naturaleza de las muestras que se procesan (las células se presentan traslapadas y arrolladas sobre si mismas, lo que dificulta la obtención de contornos). En la actualidad se están utilizando procedimientos de redes neuronales en el reconocimiento de formas celulares en casos como los que aquí presentan conflicto.

Para la caracterización celular, se pensó inicialmente implementar dos métodos uno manual y otro automático. Mientras en el caso manual el usuario debería realizar las

operaciones de procesamiento de imagen, para cada una de las partes de la célula necesarias para la caracterización (núcleo - citoplasma), en el caso automático, el sistema se encargaría de utilizar la información producto de la investigación con el fin de hacer transparentes las operaciones de mejoramiento de la imagen, sobre todo teniendo en cuenta a los usuarios del área médica. Sin embargo, no fue posible determinar un único umbral de binarización para las células de la misma clase, ni utilizar un elemento estructurante para la aplicación de los filtros morfológicos. Por lo anterior, no fue posible implementar el método automático para la caracterización y diagnóstico de las células, se espera su utilización en el reconocimiento de elementos u objetos que posean una estructura similar a la celular, es decir, un área interior (tono mas oscuro de la imagen) y un área exterior (de tono mas claro). Por lo tanto para la extracción de los valores de área del citoplasma, se debe utilizar el procedimiento de área del polígono, como se mencionó anteriormente.

Una vez se ha hecho la caracterización celular se procede a su clasificación, ella se realizó comparando los valores que experimentalmente se determinaron para cada una de las clases de células (intermedias, parabasales y superficiales), con los valores de los patrones, de igual forma se obtiene el estado celular basados en la relación nuclear - citoplasmática; sólo se utiliza este criterio de malignidad debido a que la implementación de los otros no es posible por las limitaciones en hardware.

ANÁLISIS DE RESULTADOS

Durante el análisis de resultados, se realizaron pruebas sobre diferentes casos comprobados por los médicos, pero no utilizados en la obtención de los patrones, con el fin de determinar la habilidad de diagnóstico del sistema y la facilidad de manejo por parte de dos tipos de usuario.

Como producto de la investigación, se obtuvo la caracterización de las células del cervix uterino humano en términos de las áreas nucleares, citoplasmáticas y sus relaciones, de la siguiente manera:

- Rangos de áreas para los núcleos de las células normales según su clase:
 - ❖ Célula Superficial 171 - 284 pixeles
 - ❖ Célula Intermedia 363 - 573 pixeles
 - ❖ Célula Parabasal 642 - 825 pixeles

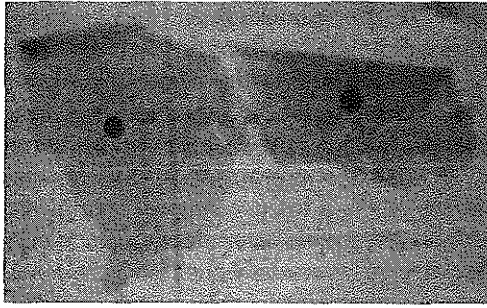


Figura 10. La imagen muestra Células superficiales normales.

- Rangos de áreas para los citoplasmas de las células en estado normal:

❖ Célula superficial	29779 - 38900 pixeles
❖ Célula Intermedia	19406 - 28991 pixeles
❖ Célula Parabasal	5051 - 13457 pixeles

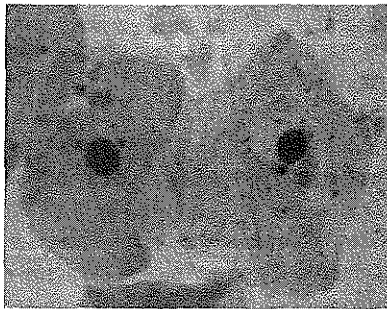


Figura 11. La ilustración presenta un par de Células Intermedias normales.

- Rangos de valores para la relación nuclear/citoplasmática para las células en estado normal o displásico:

❖ Célula superficial	0 - 0.012
❖ Células Intermedias	0.012 - 0.04
❖ Células Parabasales	0.04 - 0.17

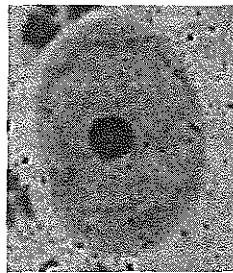


Figura 12. Gráfico de una Célula parabasal normal.

Se logró determinar que las células en estado canceroso exhiben una relación nuclear/citoplasmática superior a 0.17.

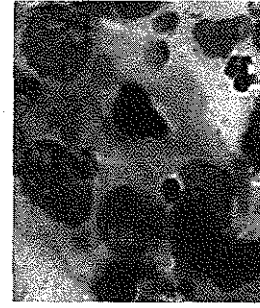


Figura 13. Conjunto de células cancerosas, nótese la ausencia del citoplasma.

La unidad de medida utilizada en todas las operaciones realizadas es el pixel, de igual forma, las imágenes utilizadas para calcular los valores anteriormente mencionados se tomaron sobre un microscopio óptico con un objetivo de 40 aumentos.

Cuando el sistema fue utilizado por personas expertas en el manejo de los computadores y conocedores del área de visiónica y tratamiento de imágenes, se logró una eficiencia del 100% en el diagnóstico emitido por el sistema. Para los usuarios poco expertos en el uso del computador, se observó que en el 10% de los casos el diagnóstico fue erróneo, lo cual puede ser superado con la práctica.

CONCLUSIONES Y RECOMENDACIONES

El sistema Diagnósis se constituye en un primer intento para caracterizar las células del cervix uterino humano, en nuestro grupo de trabajo, con el fin de determinar el grado de avance del cáncer basado en el criterio del agrandamiento nuclear, utilizando los recursos disponibles en el grupo de ingeniería biomédica de la UIS.

Diagnósis aglutina las operaciones básicas de la morfología matemática, constituyéndose así, en una herramienta obligada para las personas que quieran conocer los efectos de los filtros morfológicos sobre una imagen.

Es importante continuar el presente proyecto con el fin de tener en cuenta los once criterios que se utilizan para el diagnóstico del cáncer, las limitaciones del hardware sólo permitieron utilizar el criterio del agrandamiento del núcleo.

Se sugiere investigar la forma de realizar la caracterización muestral automática eliminando las limitaciones impuestas por las diversas tinturas usadas en el proceso de preparación de las placas.

REFERENCIAS

- [1] ADAMS, Programación avanzada de gráficos en C para Windows, Madrid; McGraw Hill, 1.993.
- [2] ANIL Jain K, Fundamentals of digital image processing, New Jersey; Prentice Hall, 1989.
- [3] DOUGHERTY Edward R. Mathematical morphology in image processing, New York; Marcel Dekker, 1992.
- [4] GRAHAM Ruth. The Citologic diagnosis of cancer. Buffalo N.Y.; The Heyman Laboratory, W.B. Sanders Company, 1.972.
- [5] GIARDINA C.R.; DOUGHERTY E.R. Image processing continuos to discrete. Vol 1, Geometric, transform and statistical methods, New Jersey; Prentice Hall, 1987.
- [6] GIARDINA C.R.; DOUGHERTY E.R. Morphological methods in image and signal processing. New Jersey; Prentice Hall, 1987. P.91-182.
- [7] GIARDINA C.R.; DOUGHERTY E.R. Matrix structured image processing, New Jersey; Prentice Hall, 1987. p. 98-130.
- [8] GONZÁLEZ, R.C.; FU, K.S y LEE C.S.. Robótica : control, detección, visión e inteligencia, Madrid; Mc Graw Hill, 1988.
- [9] GONZÁLEZ R.C. Digital image processing, Massachusetts; Addison-Wesley Publishing Company, 1987.
- [10] GROOVER, Mikell P.; WEISS Mitchell; NAGEL Roger N.; ODREY Nicholas G. Robótica industrial, Madrid; Mc Graw Hill, 1989. p. 177-204.
- [11] HARALICK, R. M.; STERNBERG S. R.; ZHUANG X. Image analysis using mathematical morphology. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAMI-9 (Julio 1987) No. 4, p. 532-551.
- [12] NIKON COPORATION. Manual de usuario del microscopio Alphaphot SE YS2, 1991.
- [13] MARAGOS, P Tutorial on advances in morphological image processing and analysis. Optical Engineering, Vol 26 (julio 1987) No. 7, p. 623-632.
- [14] MARAGOS P. A representation theory for morphological image and signal processing. IEEE Transaction on pattern anal machine intelligence, Vol 11 (Junio 1989) No. 6, p. 586-599.
- [15] MATROX. Pip-ez : Software library for the PIP video digitizer. User manual MATROX Electronic systems limited, 1987.
- [16] MICROSOFT CORPORATION. Microsoft C runtime library. Reference for MS-DOS operating system, 1984.
- [17] MICROSOFT CORPORATION. Microsoft C lenguaje reference, 1984.
- [18] PAVLIDIS T. Algorithms for graphics and image processing. Springer-Verlag, 1982.
- [19] SERRA Jean. Image analysis and mathematical morphology, New York; Academic Press, 1983.
- [20] TAKAHASHI L. Atlas Médico Ilustrado, New Jersey; Prentice Hall, 1983.
- [21] ROBBINS. Patología Estructural y Funcional. 5ª Edición Iiteramericana. W. B. Saunders Company. Cap. 23 p. 1155 - 61.