



**UNIVERSIDAD INDUSTRIAL DE SANTANDER**  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES  
Perfecta combinación entre Energía e Intelecto



**“SOFTWARE DE ASISTENCIA PARA EL DIAGNOSTICO DE LEUCEMIA  
MEDIANTE PROCESAMIENTO DIGITAL DE IMAGENES MICROSCOPICAS”**

**SANDRA MEDRANO MEDINA  
JARLEUDER ALBERTO MORENO MENESES**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECHANICAS  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES  
BUCARAMANGA  
2012**



**UNIVERSIDAD INDUSTRIAL DE SANTANDER**  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y TELECOMUNICACIONES  
Perfecta combinación entre Energía e Intelecto



**“SOFTWARE DE ASISTENCIA PARA EL DIAGNOSTICO DE LEUCEMIA  
MEDIANTE PROCESAMIENTO DIGITAL DE IMAGENES MICROSCOPICAS”**

**SANDRA MEDRANO MEDINA  
JARLEUDER ALBERTO MORENO MENESES**

**Proyecto de grado para optar al título de Ingeniero Electrónico**

**DIRECTOR:  
MPE JAIME BARRERO**

**COORDIRECTOR:  
MsC (c) CARLOS NIÑO**

**UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECHANICAS  
ESCUELA DE INGENIERÍAS ELÉCTRICA, ELECTRÓNICA Y  
TELECOMUNICACIONES  
BUCARAMANGA  
2012**

## CONTENIDO

	<b>Pág.</b>
INTRODUCCIÓN	14
1. LEUCEMIA Y TOMA DE DATOS	16
1.1. LEUCEMIA	16
1.2. CARACTERÍSTICAS DE LA LEUCEMIA LINFOCITICA AGUDA LLA.	18
2. IMAGEN DE ENTRADA Y PRE PROCESAMIENTO	21
2.1. ADQUISICIÓN DE IMÁGENES	21
2.2. INGRESO DE DATOS	22
2.3. ESPACIO DE COLOR Y FILTRADO	23
3. SEGMENTACIÓN	26
3.1. ETAPAS DEL PROCESO	26
3.2. HISTOGRAMAS Y OTSU	26
3.3. MÁSCARAS Y ÁREAS	29
3.4. SEGMENTACIÓN POR ETIQUETADO Y RECONSTRUCCIÓN DE OBJETOS	30
4. CARACTERIZACIÓN, RESULTADOS Y CONCLUSIONES	31
4.1. CARACTERIZACIÓN	31
4.2. TEXTURA	31
4.3. DESCRIPTOR DE CONTORNO SIGNATURA O FIRMA	33
4.4. FRACTALES	34
4.5. CLASIFICACION USANDO REDES NEURONALES	42

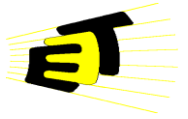
4.6. PROCEDIMIENTO	43
4.7. SEGURIDAD DE LA PRUEBA DIAGNOSTICA	48
4.8. CONCLUSIONES Y RECOMENDACIONES	52
REFERENCIAS:	54
ANEXOS	57

## LISTA DE FIGURAS

	<b>Pág.</b>
Figura 1.1. Diagrama de bloques del procedimiento.	15
Figura 1.2. Evolución de una célula sanguínea. Una célula madre sanguínea pasa por varias etapas para convertirse en un glóbulo rojo, una plaqueta o un glóbulo blanco.	17
Figura 2.1. Diagrama del procedimiento general de adquisición de imágenes.	21
Figura 3.1. Imágenes de las 3 componentes de una imagen en HLS y los histogramas de sus componentes correspondientes filtradas. (a), (b), (c), (d), (e) y (f).	27
Figura 3.2. (a) Histograma Bidimensional L/S. (b) Objeto segmentado.	30
Figura 4.1. Grafica de la signatura de una célula tipo F	34
Figura 4.2. Curva de cuadrados fractales para tres iteraciones de una célula tipo F.	35

## LISTA DE TABLAS

	<b>Pág.</b>
Tabla 1. Existen Dos Tipos Básicos De Leucemia (3)	16
Tabla 2. Concepto de falso positivo en medicina diagnostica (7)	19
Tabla 3. Clasificación de características malignas leucemia linfocítica aguda (6)	20
Tabla 4. Comparación de los métodos de descripción de texturas (22).	37
Tabla 5. Descriptores y matriz de coocurrencia (23)	38
Tabla 6. Proceso de Segmentación características malignas célula tipo A.	47
Tabla 7. Clasificación celular realizada por el modelo computacional.	51
Tabla 8. Numero de Falsos Positivos y Negativos obtenidos con el modelo computacional.	51



## LISTA DE ANEXOS

	<b>Pág.</b>
ANEXO A SOFTWARE	57
ANEXO B. TABLA DE DATOS USADOS SE ENCUENTRA EN DAT.XLS	57

## RESUMEN

**TITULO:** “SOFTWARE DE ASISTENCIA PARA EL DIAGNOSTICO DE LEUCEMIA MEDIANTE PROCESAMIENTO DIGITAL DE IMAGENES MICROSCOPICAS”\*

**AUTORES:** SANDRA MEDRANO  
JARLEUDER MORENO\*\*

**PALABRAS CLAVES:** Segmentacion, color, coocurrencia, descriptores, red neuronal

Se diseña un algoritmo en MatLab, como ayuda para diagnosticar LLA, encontrando características malignas en los glóbulos blancos linfoides que se encuentran en la sangre periférica. Éste debe ser un apoyo (asistente virtual) para el médico especialista. Las muestras médicas fueron donadas por el HUS. Consta de cuatro etapas de procesamiento: Imagen de entrada, Segmentación, Caracterización y Reconocimiento. En la primera etapa se lee la imagen que se selecciona de una base de datos adquirida previamente, donde se encuentran todas las imágenes de prueba. Se convierte la imagen original a una imagen en escala de grises. En la etapa de segmentación se utiliza segmentación multinivel combinada con espacios de color RGB-HLS unido al método de OTSU y descriptores con matriz de Co-ocurrencia. Dentro de la misma etapa de segmentación se utilizan operadores morfológicos descriptores de forma, para lograr caracterizar la célula; un objeto puede ser descubierto fácilmente en una imagen si el objeto tiene el contraste de fondo suficiente. En la etapa de caracterización se usan descriptores de forma y color y se comparan los datos obtenidos con la tabla de clasificación. En la etapa de reconocimiento se tiene una red neuronal artificial backpropagation la cual se encarga de dar un diagnóstico. Finalmente se diseña una interfaz como conexión entre el asistente y el usuario. Con el algoritmo propuesto se obtuvo una sensibilidad y una especificidad de la prueba diagnóstica de 94,4% y 83,3%.

---

\* Proyecto de Grado

\*\* Facultad de Ingenierías fisicomecánicas, Escuela de Ingenierías eléctrica, electrónica y telecomunicaciones  
Director: Jaime Barrera. Codirector: Carlos Niño

## ABSTRACT

**TITLE:** "ASSISTANCE SOFTWARE FOR THE DIAGNOSIS OF LEUKEMIA USING DIGITAL PROCESSING OF IMAGES MANIKINSMODELSMICROSOPIC"

**AUTHORS:** SANDRA MEDRANO  
JARLEUDER MORENO\*\*

**Keywords:** Segmentation, color, co-occurrence, descriptors, neural network

Designing an algorithm in Matlab, to help diagnose LLA, finding malignant features in white blood cells found in the peripheral blood. This should be a support (virtual assistant) to the specialist. Medical samples were donated by the HUS. It consists of four processing steps: Input Image, Segmentation, Characterization and Recognition. The first stage reads the image that is selected from a database, where there are all the test images. It converts the original image to a gray scale image. In the segmentation step, multilevel segmentation is used combined with RGB-HSL color space, attached to OTSU method and Concurrency Matrix descriptors. Within the same stage of segmentation it uses morphological operators. An object can be found easily in an image if the object has an enough background contrast. In the Characterization stage, color and shape descriptors are used and the result is compared with the classification table. The recognition stage uses an artificial neural network back propagation which is responsible for providing a diagnosis. Finally, a user interface is designed as a connection between the wizard and the user. With the proposed algorithm it obtained a sensitivity and specificity of diagnose 94,4% y 83,3%. the wizard and the user. With the proposed algorithm it obtained a sensitivity and specificity of diagnose 94,4% y 83,3%.

---

\* Project of grade

\*\* Faculty of Engineerings physicochemical, school of Engineerings electica, electronics and telecomunicaicones Director: Jaime Barrera. Co-Director: Carlos Niño

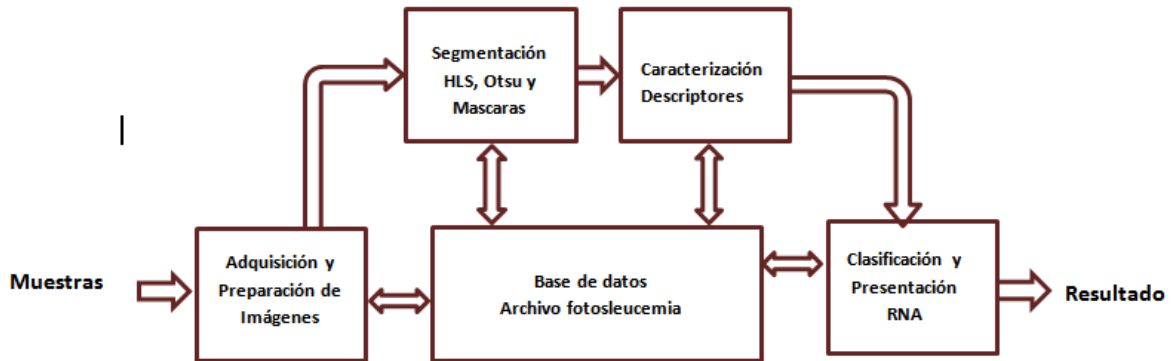
## INTRODUCCIÓN

Uno de los grandes retos a los que se enfrenta la humanidad es tratar de mejorar la salud y la calidad de vida de las personas. De esta manera se ha visto a lo largo de los años, los esfuerzos para idear herramientas que puedan garantizar el bienestar humano.

Cada día crece la preocupación de adquirir alguna enfermedad y entre las más importantes se encuentra el cáncer en todas sus manifestaciones, siendo los más comunes el cáncer de mama, pulmón, estomago entre otros; el cáncer en la sangre, especialmente en los glóbulos blancos es llamado también leucemia linfocítica y una de las más habituales es la leucemia linfocítica aguda que es la patología clínica a trabajar en el siguiente libro.

Se diseña un algoritmo en MatLab, como asistencia para alertar sobre características malignas en los glóbulos blancos que se encuentran en la sangre periférica. Éste debe ser un apoyo (asistente virtual) para el médico especialista.

**Figura 1.1. Diagrama de bloques del procedimiento.**



Fuente: Autores del Proyecto

Esta tesis consta de cuatro etapas de procesamiento ver *Figura 1.1*, Imagen de entrada, Segmentación, Caracterización y Reconocimiento. En la primera etapa se lee la imagen que se selecciona de una base de datos adquirida previamente y donde se encuentran todas las imágenes de prueba, puede ser manual o automáticamente. Se convierte la imagen original a double para poder operar. En la etapa de segmentación se utiliza segmentación multinivel combinada con espacios de color RGB-HLS unido al método de Otsu y umbrales adecuados, además a través de la función regionprops se obtienen las propiedades de la imagen que nos permiten analizar más detalladamente las imágenes segmentadas individualmente con la ayuda de (marcos) o cuadros que toman cada arreglo matricial u objeto de interés y lo guarda en una estructura para luego ser evaluado por separado analizando su comportamiento ante los descriptores; se hace un corte en veinte células por que más de estas en un mismo extendido son un clave indicador de la existencia de la enfermedad . En la etapa de clasificación se usan descriptores de **forma**: (signatura y fractales), **textura**: (energía, entropía, homogeneidad y correlación) y **color**: (matiz de la imagen segmentada y coloreada) y se comparan los datos obtenidos para encontrar diferencias. En la etapa de reconocimiento se entrena una red neuronal artificial la cual se encarga de dar un diagnóstico.

## 1. LEUCEMIA Y TOMA DE DATOS

### 1.1. LEUCEMIA

Según la fundación mundial contra el cáncer el 75% de casos de cáncer en la sangre son debidos a la LLA o leucemia linfocítica aguda siendo esta la más común de las leucemias y se presenta sobre todo en los niños (2), por eso este sistema se considera de gran importancia y un aporte de bajo coste computacional que puede ser usada como herramienta de asistencia virtual a la hora de poder diagnosticar a tiempo y evitar que la enfermedad degenerativa se siga desarrollando.

Una clasificación de la leucemia o cáncer en la sangre es de acuerdo al tipo de célula o glóbulo blanco afectado, ver Tabla 1.

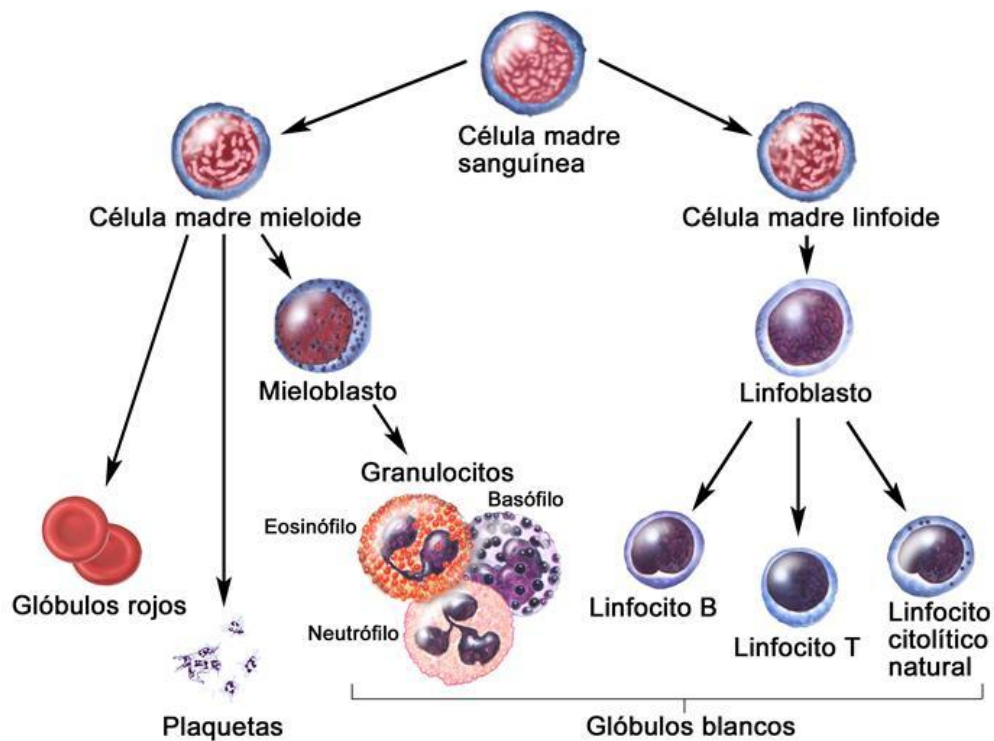
**Tabla 1. Existen Dos Tipos Básicos De Leucemia (3)**

<p><b>LEUCEMIA LINFOBLASTICA</b></p>	<p><b>SE DESARROLLA EN LOS GLOBULOS BLANCOS LINFOIDES: LINFOCITOS. Puede ser Cónica o Aguda. Más común.</b></p>
<p><b>LEUCEMIA MIELOGENA</b></p>	<p><b>SE DESARROLLA EN LOS GLOBULOS BLANCOS MIELOIDES: MONOCITOS Y GRANULOCITOS. Puede ser Cónica o Aguda. Más difícil de contraer.</b></p>

Fuente: Autores del Proyecto

En la Figura 1.2 se puede observar la evolución de una célula sanguínea desde que es una célula madre y su división entre linfoide y mieloide (4). Más adelante y dependiendo de las condiciones con las cuales se encuentren en la sangre periférica se pueden convertir en los dos grupos básicos de leucemia más conocidos: la leucemia linfocítica y la leucemia mielogena que dependiendo de su grado de desarrollo se divide en aguda y crónica (1).

**Figura 1.2. Evolución de una célula sanguínea. Una célula madre sanguínea pasa por varias etapas para convertirse en un glóbulo rojo, una plaqueta o un glóbulo blanco.**



© 2007 Terese Winslow  
 U.S. Govt. has certain rights

Fuente: Autores del Proyecto

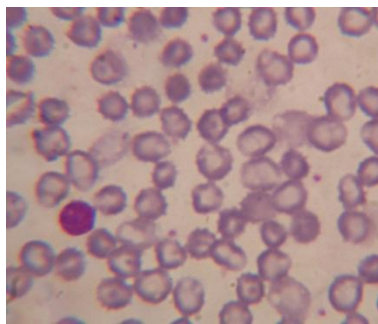
## **1.2. CARACTERÍSTICAS DE LA LEUCEMIA LINFOCITICA AGUDA LLA.**

Uno de los grandes retos de un especialista en análisis de extendidos sanguíneos es poder diagnosticar acertadamente si existe alguna patología evidente en las muestras, para ello se requiere mucho tiempo bajo el microscopio y desgaste visual ya que es necesario cubrir toda la placa con el lente buscando características malignas en los glóbulos blancos si es que existen, la Tabla 3 refleja los diferentes tipos de glóbulos blancos presentes en los linfocitos inmaduros en sangre periférica o también llamada leucemia linfocítica aguda con sus características definidas bajo parámetros evidentes de forma, color y textura que deben ser analizados en su totalidad para lograr obtener un buen resultado de diagnóstico final. Es importante tener en cuenta que solo una de estas características presentes en un mismo extendido sugiere que el paciente presente probablemente un verdadero positivo (5), ver Tabla 3 de sangre contaminada por eso el especialista crea una alerta de cuidado para así a través de otros exámenes como pruebas histoquímicas poder decidir realmente si posee o no este tipo de cáncer e iniciar inmediatamente tratamiento de quimioterapia para lograr tratar a tiempo dicha enfermedad y no permitir que involucre otros órganos en la metástasis de la misma (6).

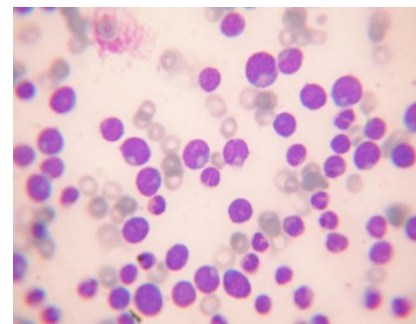
**Tabla 2. Concepto de falso positivo en medicina diagnostica (7)**

Tipos de diagnósticos		Enfermedad	
		Ausente	Presente
Prueba Diagnostica	negativa	<b>Verdadero negativo</b> (diagnóstico negativo enfermedad ausente)	<b>Falso negativo</b> (diagnóstico negativo enfermedad presente)
	Positiva	<b>Falso positivo</b> (diagnóstico positivo enfermedad ausente)	<b>Verdadero positivo</b> (diagnóstico positivo enfermedad presente)

**Figura. 1.3.** Extendido sangre periférica, Tinción Wright, a 100 X. a) Célula Sanguínea Sana. b) Célula Sanguínea con LLA (1).

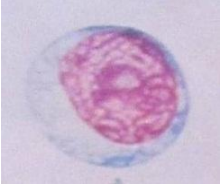



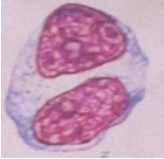
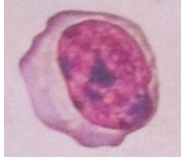
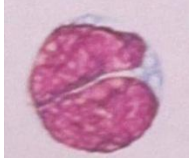


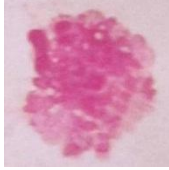


a)



b)

**Tabla 3. Clasificación de características malignas leucemia linfocítica aguda (6)**

 <p><b>TIPO A</b>                  LINFOBLASTO CON NUCLEOLO (célula inmaduro blasto con nucléolo evidente)</p>	 <p><b>TIPO B</b>                  LINFOBLASTO CON VARIOS NUCLEOS EVIDENTES igual al tipo a pero con dos o más nucléolos</p>	 <p><b>TIPO C</b>                  PROLINFOCITO CON NUCLEOLO INDISTINTO(NO LINFOBLASTO) casi no hay citoplasma solo estructura linfocítica sin núcleo evidente</p>	 <p><b>TIPO D</b>                  PROLINFOCITO CON CROMATINA DE ESTRUCTURA INTERMEDIA Y CITOPLASMA DE ASPECTO RIZADO poco frecuentes</p>
 <p><b>TIPO E</b>                  PROLINFOCITO CON NUCLEO DOBLE Y CROMATINA NUCLEAR INMADURA, se puede ver núcleo dividido blastos dobles inmaduros</p>	 <p><b>TIPO F</b> LINFOCITO ATIPICO CON CROMATINA NUCLEAR CONGLOMERADA Y CITOPLASMA PARDO,SIN GRANULOS</p>	 <p><b>TIPO G</b>                  PROLINFOCITO CON NUCLEO PROFUNDAMENTE HENDIDO\ Forma de m hendidura causando posible división no tiene nucléolos evidentes</p>	 <p><b>TIPO H</b> LINFOCITO CON NUCLEO MULTILOBULADO Muy común en extendidos sanguíneos es una célula multividida</p>
 <p><b>TIPO I</b> LINFOCITO CON FRAGMENTO NUCLEAR</p>	 <p><b>TIPO J</b> TIZNON especie de mancha que se puede encontrar en extendidos sanguíneos,</p>		

## 2. IMAGEN DE ENTRADA Y PRE PROCESAMIENTO

### 2.1. ADQUISICIÓN DE IMÁGENES

Las muestras médicas de extendido sanguíneo con leucemia fueron donadas por el Hospital Universitario de Santander HUS. Las laminillas con extendido sanguíneo normal son obtenidas de la siguiente forma: Se toman muestras de sangre del brazo con una jeringuilla estándar desechable y nueva a voluntarios sanos, se coloca una gota de sangre de 2 a 3 mm de diámetro en el extremo de una lámina porta objeto limpia. El tamaño de la gota es importante, demasiado grueso nos da un extendido muy grueso o largo, demasiado pequeño nos da un extendido muy delgado o corto. Con otro portaobjetos limpio se toca la gota de sangre que se encuentra en el primer portaobjeto con un ángulo de 30 a 45 grados y se empuja a lo largo del mismo de forma suave y rápida.

**Figura 2.1. Diagrama del procedimiento general de adquisición de imágenes.**



Fuente: Autores del Proyecto

El procedimiento se hizo de inmediato, pero se tiene máximo de 2 a 3 horas después de la toma de la muestra para hacerlo. Luego se espera un tiempo de secado para las láminas y son expuestas al colorante Wright el cual les da la tonalidad morada o rosada dependiendo el tiempo de exposición, se usa de 1 a 2 minutos, esto se hace con el fin de facilitar la diferenciación de los tipos de células sanguíneas.

Para mantener las condiciones de luz y color constantes, se cubre con una cartulina negra el microscopio, la cámara fotográfica y parte del observador, se usa en el microscopio el nivel 6 de luz.

Teniendo en cuenta un observador con capacidades visuales 20/20, se utiliza el microscopio para efectuar el enfoque de las imágenes primeramente, se verifica la imagen en la cámara (la cual se encuentra en enfoque manual) y con el enfoque fino del microscopio se toma la imagen deseada que es entregada en un formato .jpg guardado en el espacio RGB.

## **2.2. INGRESO DE DATOS**

A continuación se presenta el modo de ingreso de los datos a nuestro sistema computacional, además se explica el método de pre procesamiento de la imagen, los dos tipos de guardado en memoria, el cambio de tamaño para lograr una convergencia más rápida y el cambio de espacio de color a uno más parecido al que podemos percibir los seres humanos.

Como se explicó en la sección anterior el formato de entrega por parte de la cámara en el microscopio es .jpg existen otros formatos como: (.png;.bmp;.tga) entre los más usados. Las imágenes se encuentran en un espacio de color

definido, RGB (red, blue, green), el cual representa muy bien toda la gama de colores que se pueden percibir en una pantalla a través de tres matrices de datos que al combinar sus colores primarios permite visualizar una imagen.

El sistema permite el ingreso de datos de manera manual y automática a través de un banco de imágenes celulares; leucemia1.m ingresa los datos de esta última forma ejecutando un comando simple que re-direcciona a un folder llamado 'fotosleucemia', uno de los bancos de imágenes obtenidos en el capítulo 1 y que puede ser ampliado cada vez que se desee con imágenes de diferentes extendidos o nuevos universos de glóbulos blancos linfocíticos. Pero si se prefiere evaluar la imagen manualmente porque se tiene duda acerca de cierto extendido, entonces cabe la posibilidad de ejecutar leucemia2.m que emplea otro modo de ingreso [filename, pathname] que direcciona a la memoria donde se encuentran guardadas las imágenes que se desean estudiar y la carga dando doble clic sobre la misma.

### **2.3. ESPACIO DE COLOR Y FILTRADO**

Para lograr operar matemáticamente de una manera más sencilla el espacio de color RGB que esta inicialmente en uint8 se convierte a double y se normaliza cada componente espacial transformado, luego es guardado mediante la función size (8), en las variables [a,b,caps] que representan el tamaño de la matriz, la longitud y el número de capas que contienen para luego ser usado en la transformación del espacio de color a hls.

Según Gil (9), en su artículo presenta un método para tratar de reducir la influencia de sombras y brillos que aparecen formando parte de información acromática presente en las imágenes capturadas, para lograrlo se acepta el hecho de hacer

un cambio en el espacio de color a trabajar, Gil lo hace de RGB-HLS el cual se adopta en nuestro trabajo pero con un breve cambio sugerido por (Angulo y Serra), donde se pretende hacer una clara separación de lo cromático/acromático a través de un umbral de la componente de la saturación que se define como la cromaticidad de un color o intensidad del color.

Dentro del método de transformación las cantidades máximas y mínimas asociadas a la saturación se logran mejorar sustancialmente el espacio HLS según Hanbury y Serra (10), Sin embargo, esta semi-norma no puede utilizarse por que la luminancia (no es invariante por proyección sobre el eje acromático). Una media ponderada, en cambio, si es una norma, y siguiendo la recomendación de Hanbury y Serra (10) se usa para la luminancia de este sistema aquella recomendada por el estándar de televisión de alta definición (similar a los monitores). De donde se obtiene la siguiente grupo de ecuaciones del nuevo espacio de color sugerido por (11):

$$\begin{aligned}
 hHLS = & \quad g - b \quad \max R, G, B - \min(R, G, B) \quad \text{si } r = \max r, g, b \quad , \\
 & \quad b - r \quad \max R, G, B - \min(R, G, B) + 2 \quad \text{si } b = \max r, g, b \\
 & \quad r - g \quad \max R, G, B - \min(R, G, B) + 4 \quad \text{si } g = \max r, g, b
 \end{aligned} \tag{1.0}$$

$$IHSV = 0.213r + 0.715g + 0.072b \tag{1.1}$$

$$sHSV = \max R, G, B - \min(R, G, B) \tag{1.2}$$

Los resultados obtenidos de la conversión de color pueden verse reflejados en la tabla de proceso de segmentación de características malignas célula tipo A en el último capítulo, Tabla 6.

Como puede verse se muestran las imágenes en RGB y cada una de las imágenes de matiz, luminancia y saturación, además como técnica de suavizado de superficie se usó un *filtro pasabajas promediado*, convolucionado con cada una de las componentes obtenidas H, L Y S.

La convolucion ocasiono que se creara un marco negro alrededor de la imagen el cual se eliminó de cada componente espacial para tener la imagen pura, tal cual se había obtenido inicialmente cuando se hizo la adquisición pero con un diferente espacio de color ahora solo queda esperar el proceso de segmentación y la máscara total que entregaran el glóbulo blanco segmentado al final del siguiente capítulo.

### **3. SEGMENTACIÓN**

#### **3.1. ETAPAS DEL PROCESO**

En este capítulo se muestra la forma como los Glóbulos blancos que son nuestro objeto de interés, son detectados y extraídos fuera de las imágenes de los extendidos sanguíneos, mediante varios pasos (9) (12):

1. Obtención histogramas de la imagen en el espacio IHLS.
2. Cálculo automático de los umbrales de Otsu para separar las regiones cromáticas en la imagen, independientemente de sus condiciones de luz.
3. Construcción de máscaras binarias para separar las células blancas del fondo.
4. Erosión binaria de la imagen y discriminación por área, para eliminar elementos que no corresponden al objeto de interés y/o separar células unidas.
5. Segmentación mediante etiquetado para separar las células por cuadros.
6. Reconstrucción de los objetos de interés para la extracción de sus características.

#### **3.2. HISTOGRAMAS Y OTSU**

Luego de que la imagen o el conjunto de imágenes a analizar han sido convertidas de RGB a HLS, y pasadas por un filtro pasa bajas para suavizar discontinuidades, son analizadas a través de sus histogramas.

Considerando la definición del histograma de una imagen como la representación gráfica o analítica de la frecuencia relativa de los niveles de gris de la imagen (13). Cada valor posible de pixel de imagen, y en caso de imágenes grises de 8 bits

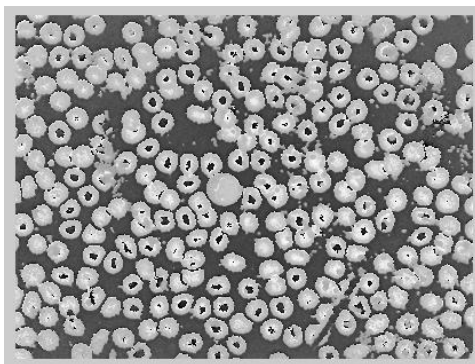
será un vector de 256 componentes, siendo la componente  $i$  el número de píxeles de nivel  $i$  en la imagen, dividido por el número total de píxeles. El histograma es formalmente la función estadística de densidad de probabilidad en forma discreta de los distintos niveles de gris dentro de la imagen. Si se integra el resultado es la función de distribución  $\omega[i]$ . (14).

$$\text{histograma } i = \frac{N_{\text{píxeles de nivel } i}}{N_{\text{total}}} \quad 0 \leq i \leq 255 \quad (3.1) \quad (14)$$

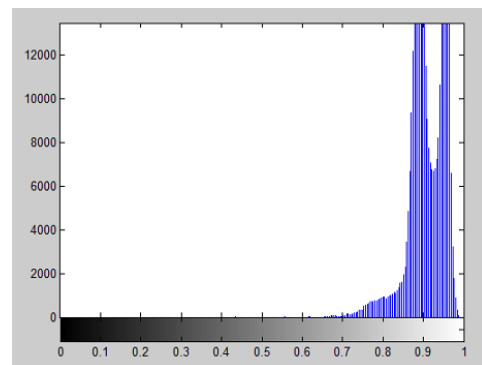
$$\omega_i = \sum_{j=0}^i \text{histograma } j \quad 0 \leq i \leq 255 \quad (3.2) \quad (14)$$

En los histogramas de las tres componentes filtradas (Matiz, Luminancia y Saturación) de la figura 3.1, se distinguen regiones de color bien definidas lo cual nos brinda un índice aproximado que sirve para hallar los umbrales de color de las regiones y que posiblemente los más adecuados sean los que se encuentran en los valles del histograma. Nuestro objetivo ahora es encontrar estos umbrales óptimos de manera que los glóbulos blancos puedan ser separados del fondo.

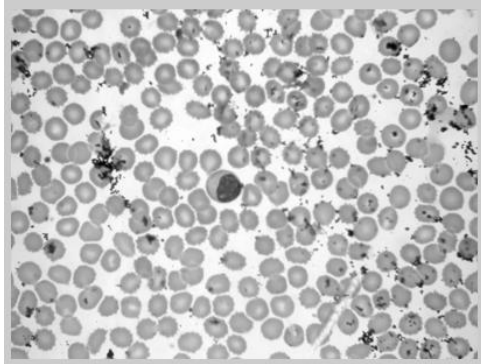
**Figura 3.1. Imágenes de las 3 componentes de una imagen en HLS y los histogramas de sus componentes correspondientes filtradas. (a), (b), (c), (d), (e) y (f).**



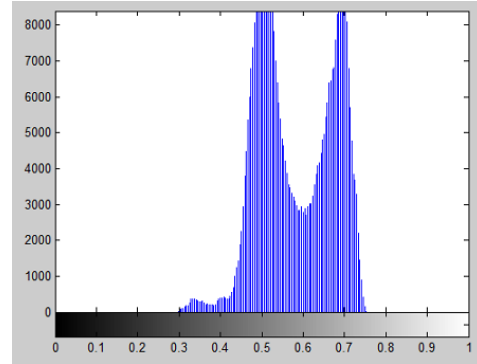
(a) Componente hHLS



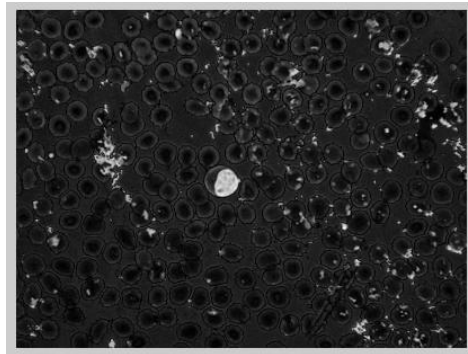
(b) Histograma Componente hHLS filtrada.



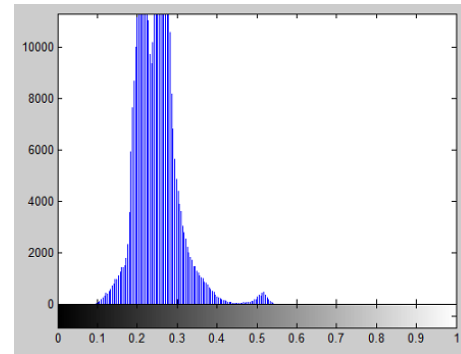
(c) Componente IHLS.



(d) Histograma Componente IHLS filtrada.



(e) Componente sHLS.



(f) Histograma Componente sHLS filtrada.

Fuente: Autores del Proyecto

En la literatura actual ya sea encontrada en la web o en libros paginados, existe una gran variedad de métodos y algoritmos de umbralización. En esta oportunidad se utiliza el método de Otsu como una fase del proceso completo de segmentación, debido a su selección automática de umbrales, a que no necesita información previa de la imagen, su independencia de la luz cuando se requiere binarizar una imagen y finalmente sus buenos resultados con imágenes del mundo real que proponen en los artículos (12) (15).

Se hallan los umbrales de las regiones cromáticas de forma automática aplicando el algoritmo del método Otsu, hallado en la web (16), a las componentes filtradas (Matiz, Luminancia y Saturación) de la imagen de entrada.

`function [IDX,sep, k] = otsu(l,n) (16)`

Y por cada componente del espacio HLS ingresada a otsu, se obtiene una matriz IDX del mismo tamaño de la componente, donde en lugar de tener el valor de intensidad del punto contiene el índice de la región segmentada a la cual pertenece. También se obtiene un valor de separabilidad y un vector con los valores de los umbrales característicos de las regiones.

### 3.3. MÁSCARAS Y ÁREAS

Se dibuja un histograma bidimensional Luminancia Vs Saturación para separar de forma intuitiva las zonas cromáticas donde se encuentran los glóbulos blancos y fijar sus umbrales para la creación de máscaras a partir de las matrices índice (IDX1, IDX2, IDX3) que otsu proporciona. Estas máscaras son imágenes binarias cuyo valor es 1 para las zonas donde se encuentran los glóbulos blancos y 0 para los demás puntos. Figura 3.2.

Las 3 máscaras resultantes `masLuminancia`, `masSaturacion` y `masMatiz` se convolucionan para formar una máscara final:

`masctotal = masLuminancia.*masSaturacion.*masMatiz;`

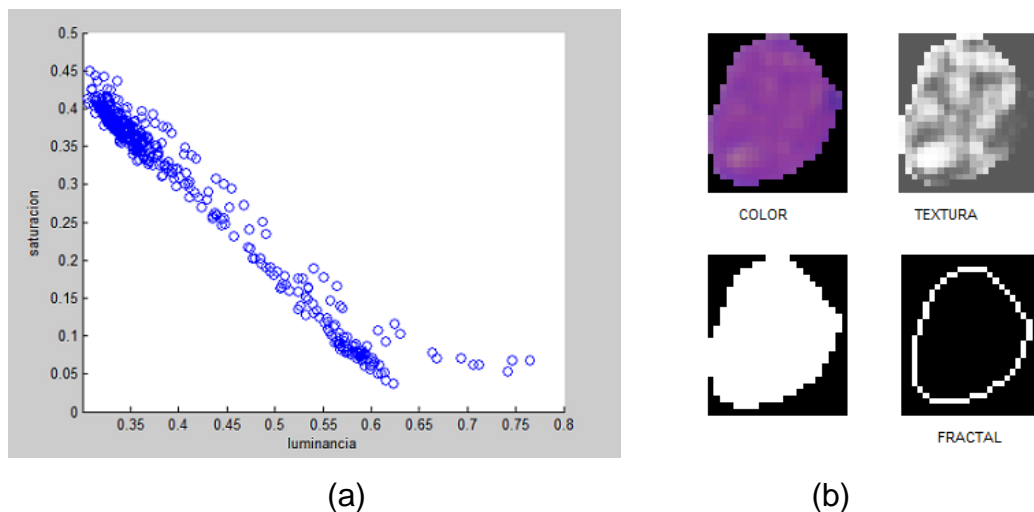
Esta máscara total o final aún presenta objetos que no son de nuestro interés. Se obtienen las propiedades de área y boundingbox con la herramienta de *Matlab*

*regionprops*, para efectuar un último tamizaje por áreas con un umbral que permite el paso de los objetos mayores al 38% del objeto con mayor área que corresponde en su mayoría a los glóbulos blancos.

### 3.4. SEGMENTACIÓN POR ETIQUETADO Y RECONSTRUCCIÓN DE OBJETOS

Con la propiedad de boundingbox se obtiene una matriz que llamamos limcadcuad (límite de cada cuadro) y con ella se mapea con un cuadrado a cada objeto de interés, se separa cada objeto en un cuadro o imagen individual cuyo objetivo es analizar las características de cada célula blanca por separado. Se hace un corte si hay más de 20 objetos ya que esto implica una característica inmediata de Cáncer en la muestra. Y finalmente se reconstruye la imagen con cada célula blanca encontrada. Figura 3.2.

**Figura 3.2. (a) Histograma Bidimensional L/S. (b) Objeto segmentado.**



Fuente: Autores del Proyecto

## **4. CARACTERIZACIÓN, RESULTADOS Y CONCLUSIONES**

### **4.1. CARACTERIZACIÓN**

En el proceso de caracterización se pretende definir una imagen mediante ciertos parámetros que la describen como son: la forma, color, textura, tales parámetros se conocen con el nombre de descriptores y son muy importantes a la hora de brindar información relevante y detallada del objeto en estudio.

### **4.2. TEXTURA**

El uso de la textura de una imagen proviene de la habilidad innata de los humanos de reconocer diferencias texturales (17). En el procesamiento digital de imágenes, la textura es una de las características más importantes en la etapa de caracterización y clasificación de patrones.

Existe un método que facilita el uso de los parámetros texturales ya que se simplifica en el análisis de la matriz de coocurrencia y surge como herramienta importante del histograma de segundo orden, también conocido como matriz de coocurrencia de nivel de gris.

Los estadísticos de segundo orden miden las dependencias espaciales que caracterizan a una textura ver Tabla 4. Mientras los descriptores calculan varias medidas sobre la matriz de co-ocurrencia de niveles de gris en la imagen y no sobre la imagen ver Tabla 5.

La matriz de coocurrencia fue usada para obtener los descriptores de textura se expresa a través de la función `cooccurrence.m` que permite hallar la relación entre pixeles vecinos de semejante valor de gris agrupándolos en una nueva matriz que tiene información relevante acerca de la textura (17).

Los descriptores son hallados a través de la función `cocaract2.m` que calcula las ocho características de cada matriz de con-ocurrencia a las direcciones (0°, 45°, 90°, 135°) y luego se promedian. Se le pasa 'glcm' que contiene las N matrices de co-ocurrencia de una sub-matriz de detalles y devuelve sus ocho características de co-ocurrencia.

En nuestro caso solo obtenemos a la salida cuatro características explicadas en las Tablas 4 y 5, bajo la forma de la función:

`function [energia,entropia,homogen,correlac]=cocaract2(glcm,N) (18)`

glcm es un método estadístico para el análisis de texturas considerando la relación espacial de los niveles de gris de la matriz de coocurrencia y entregando las características de co-ocurrencia

Dentro de los descriptores de textura tenemos la energía, entropía, correlación y homogeneidad, que se pueden ver en la Tabla 5 en esta podemos ver la relación entre la matriz de coocurrencia y los descriptores que usamos en nuestro trabajo, también tenemos otros descriptores de forma y color de los cuales hablaremos enseguida.

### 4.3. DESCRIPTOR DE CONTORNO SIGNATURA O FIRMA

La función signature.m obtenida en la WEB y definida de la siguiente manera (19):

```
function [st, angle, x0, y0] = signature(b, varargin)
```

Busca el centroide de la imagen segmentada y traza una línea hasta el contorno evitando cruces por dos o más superficies, entrega la distancia radial y el ángulo teta siguiendo la periferia de todo el objeto generando los puntos que luego son graficados en la función unidimensional como puede verse en la figura 2.2.

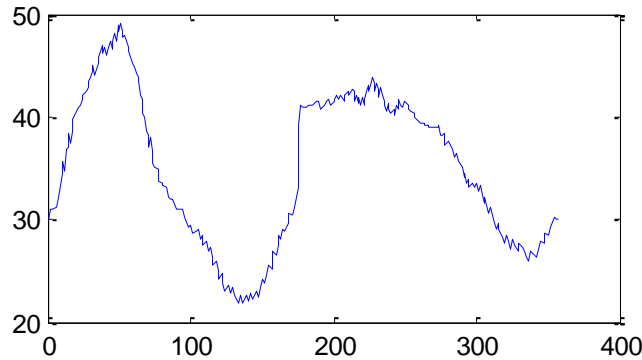
En el caso de que la línea trazada cruce por dos o más puntos del contorno de la imagen se hace un tratamiento de la imagen rellenando los bordes y eliminando esa discontinuidad mediante las siguientes líneas de código:

```
Bj=zeros(abB, abB1);
for kh=1:length(Bb2)
    Bj(round(Bb2(kh,1)),round(Bb2(kh,2)))=1;
end
B2=Bj;
```

Donde Bj es una matriz de ceros con elementos dentro de la matriz es decir se crea una especie de marco para el posterior análisis y se dilatan los bordes de manera que solo cruce por un solo punto.

Con la signatura se puede caracterizar la forma de un objeto; en la Figura 2.2 se ve la signatura de una célula tipo F, básicamente es tomar el contorno de la misma en 2D y llevarlo a una curva en 1D manteniendo la forma para tener una calificación o firma del borde o contorno del cual requerimos información.

**Figura 4.1. Grafica de la signatura de una célula tipo F**



Fuente: Autores del Proyecto

#### 4.4. FRACTALES

Las formas encontradas en la naturaleza no son fácilmente descritas por la geometría tradicional, la geometría fractal provee una descripción y un modelo matemático para definir estas formas complicadas (20).

Un fractal es un ente geométrico el cual en su desarrollo espacial se va reproduciendo cada vez a una escala menor. Una característica esencial consiste en que si observamos muy detalladamente e hiciéramos un zoom en cualquier parte del mismo este reproduce a escala menor la figura total del fractal.

**Construcción de la curva de cuadrados fractal** El punto de partida del método recursivo para dibujar la curva de cuadrados es un cuadrado simple. Debe usarse las cuatro esquinas del cuadro como el centro de 4 cuadros pequeños, cada uno con la mitad del tamaño del cuadrado principal y así iterando cada vez hasta llegar a cuadros más pequeños Figura 2.3.

**Figura 4.2. Curva de cuadrados fractales para tres iteraciones de una célula tipo F.**



a)

b)

c)

Fuente: Autores del Proyecto

Como puede verse, la estrategia más sencilla para hallar un fractal, es tomar una figura y reproducirla en versiones más pequeñas. Sin embargo, se pueden conseguir objetos cada vez mucho más complejos.

La función `fractalHB.m` implementada en matlab recibe como entrada la imagen del contorno de cada una de las células segmentadas Figura 2.4d y entrega a la salida cada fractal D0, D1 Y D2

```
function [frac0, fracD1, fracD2]=fractalHB(Imag)
```

Además esta función cuenta con una función interna `fracpq2.m` que se encarga de dibujar cada uno de los cuadrados fractales como se ve en la figura 2.5.

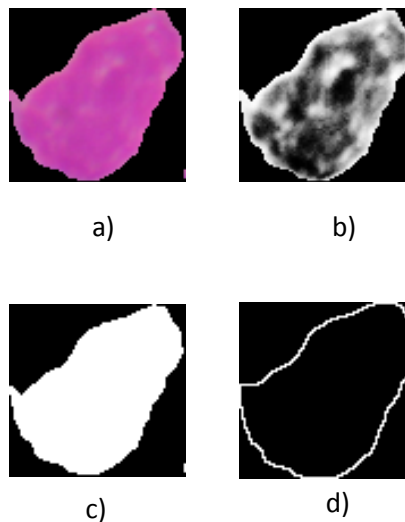
```
function frac=fracpq2(a,b,ac,bc,a1,b1)
```

Cada uno de los tres cuadrados fractales buscan lograr una aproximación del contorno de la figura tomando cada uno de los cruces por los bordes de la imagen como se observa en la figura 2.5, a) cruza dos veces, b) cruza seis veces y por

último en c) se logran apreciar 18 cruces lo cual indica que se lograría una mejor aproximación evidente con el fractal D2.

Las siguiente figuras son un ejemplo de la aplicación de textura color y forma.

**Figura 4.3.** a) Célula tipo F segmentada color, b) Textura, c) Imagen binarizada, d) contorno de la imagen.



Fuente: Autores del Proyecto

Como puede verse en la figura 2.4. a) Presenta la célula segmentada y coloreada, b) una imagen de la textura, c) imagen binarizada, d) imagen del contorno.

A continuación puede encontrarse las curvas de cuadrados fractales de una célula tipo F, hallado a partir del procedimiento explicado en la Tabla 5 descriptor de forma fractal, Figura 2.3.

Para caracterizar el color se hizo uso de la transformación a HSV usando la función RGB2HSV (21) de MatLab y usamos la componente H como se explica en la Tabla 5.

**Tabla 4. Comparación de los métodos de descripción de texturas (22).**

METODOS		VENTAJAS	INCONVENIENTES
DESCRIPTORES ESPECTRALES "FRECUENCIA"	FOURIER	BUENOS RESULTADOS PARA IMAGENES DE TEXTURA HOMOGENEA	SOLO ES UTIL PARA TEXTURAS CON PATRONES PERIODICOS SOLO REALIZA ANALISIS EN FRECUENCIA NO TEMPORAL
	WAVELET	UNICA QUE PERMITE ANALISIS MULTIRRESOLUCION	ELEVADO COSTE COMPUTACIONAL DEBIDO A LOS VARIOS NIVELES DE DESCOMPOSICION
DESCRIPTORES ESTRUCTURALES		RESULTADOS RAPIDOS SI LA TEXTURA ESTA FORMADA POR UN ELEMENTO ESTRUCTURAL RECONOCIBLE POR EL SISTEMA	NO TODAS LAS TEXTURAS ESTAN FORMADAS POR ELEMENTOS REPETITIVOS. NECESITA MEMORIA DONDE ALMACENAR POSIBLES ELEMENTOS ESTRUCTURALES
DESCRIPTORES BASADOS EN MODELOS	CADENAS DE MARKOV	RESULTADOS MUY BUENOS PARA MICROTERTURAS INTERRELACIONA PÍXELES VECINOS	ALTA COMPLEJIDAD COMPUTACIONAL NECESIDAD DE DEFINIR MASCARA ADECUADA
	FRACTALES	BUENOS RESULTADOS PARA	SOLO UTIL PARA TEXTURAS

METODOS		VENTAJAS	INCONVENIENTES
		SUPERFICIES RUGOSAS	RUGOSAS
DESCRIPTORES ESTADISTICOS	1er ORDEN ENERGIA E HISTOGRAMAS	SENCILLOS Y RAPIDOS NECESITAN POCA MEMORIA	NO TIENE EN CUANTA LA INFORMACION DADA POR LA POSICION RELATIVA DE LOS PIXELES
	2ord MATRIZ DE CO-OCURRENCIA	ADECUADA PARA TEXTURAS DE ORIGEN ALEATORIO Y QUE NO RESPONDAN A UN PATRON DETERMINADO Y TIENEN EN CUANTA LA POSICION RELATIVA DE LOS PIXELES ENTRE SI.	ALTA CARGA COMPUTACIONAL

Fuente: Autores del Proyecto

**Tabla 5. Descriptores y matriz de coocurrencia (23)**

Descriptores		CONCEPTO	FORMA DE APLICACIÓN
	<b>SIGNATURA</b>	Una signatura es una representación de un contorno mediante una función real unidimensional que	La implementación se logra dibujando la distancia del centroide del objeto al contorno en función del ángulo. Después de esto se hace

Descriptores		CONCEPTO	FORMA DE APLICACIÓN
<b>DESCRIPTOR ES DE FORMA</b>		sea más sencilla que la función bidimensional que define el contorno.	una representación unidimensional de cada punto sobre el contorno.
	<b>FRACTALES</b>	Un fractal es un objeto cuya estructura se repite a diferentes escalas. Es decir, por mucho que nos acerquemos o alejemos del objeto, observaremos siempre la misma estructura. De hecho, somos incapaces de afirmar a qué distancia nos encontramos del objeto, ya que siempre lo veremos de la misma forma.	Esta implementación se logró a través de la función fractalHB.m donde se toma la mitad del tamaño del cuadrado de la imagen y luego la mitad de esa mitad y los nuevos cuadros son colocados en los vértices y de esta manera hasta encontrar el fractal D2
<b>DESCRIPTOR DE COLOR</b>		Un descriptor de este tipo nos permite como su nombre lo indica representar el color en sus diferentes clases y un gran indicador es el matiz ya que es el	El descriptor de color implementado se hizo tomando la imagen coloreada después del proceso de segmentación y hallando el matiz de esta a través de la función: RGB2HSV (22) que transforma de



Descriptores	CONCEPTO	FORMA DE APLICACIÓN
	alcanzando su máximo cuando todos los elementos de la matriz de coocurrencia son iguales.	
HOMOGENEIDAD	Es lo opuesto al contraste. La homogeneidad es alta cuando la matriz de coocurrencia se concentra a lo largo de la diagonal. Esto ocurre cuando la imagen es localmente homogénea.	$HL = \frac{1}{i} \sum_j \frac{C_{ij}}{1 + (i - j)^2}$ (4.3)
CORRELACION	Este descriptor nos mide la semejanza de la imagen consigo misma desplazada. Algunas propiedades de la correlación son: <ul style="list-style-type: none"> <li>• Un objeto tiene más alta correlación dentro</li> </ul>	$Cr = \frac{1}{\sigma_i \sigma_j} \sum_i \sum_j (u_i - u_j) C_{ij}$ (4.4) Donde:

Descriptores	CONCEPTO	FORMA DE APLICACIÓN
	de él que entre objetos adyacentes. • Píxeles cercanos están más correlacionados entre si que los píxeles más distantes	$u_i = \sum_j C_{ij} u_j$ (4.5)
		$\sigma_i = \sum_j (i - u_i) C_{ij} - \sum_j (j - u_j) C_{ij}$ (4.6)
		(4.7)
		(4.8)

#### 4.5. CLASIFICACION USANDO REDES NEURONALES

La arquitectura de red usada se denomina back propagations multicapa. Esta red se entrena a partir de cuatro células con características similares de leucemia la red pretende separar entre células sanas y enfermas según las características de descriptores similares como energía, entropía, homogeneidad, correlación, signatura y cada uno de los 3 fractales que se redujeron a solo tres entradas a la red (energía, signatura y la varianza de todos los datos) ya que estas eran las más significativas entrando así una matriz de tamaño 3x55 servirán como parámetros de entrada a la red de características malignas facilitando de esta manera la clasificación de la red neuronal; cabe recalcar nuevamente que si se presenta solo uno de los tipos en un mismo extendido de sangre periférica se crea una alerta

(enferma) sobre la posibilidad de poseer características malignas para que así el especialista chequee minuciosamente y con más detalle el paciente en estudio.

La red neuronal utilizo como entrada un tipo de función llamada tansig ya que las entradas son tanto positivas como negativas y la función de salida es una logsig las salidas serán números enteros positivos, el algoritmo de entrenamiento es trainlm con dos capas ocultas de tres y cinco neuronas con las que se obtuvo un modelo más acertado. (24)

#### 4.6. PROCEDIMIENTO

##### 1) ingresan los datos a través de:

```
[filename, pathname] = uigetfile( ...
{'*.jpg;*.png;*.bmp;*.tga','Image Files (*.jpg;*.png;*.bmp;*.tga)', ...
'Select your file');
```

##### 2) se cargan en la variable E y se disminuye su tamaño:

```
E = imread([pathname,filename]);
fact=.25
```

##### 3) Se transforma de espacio de color RGB-HLS:

###### a) Se halla matiz viendo los valores máximos de r,g,b:

$$\left\{ \begin{array}{l} \text{hHLS}(i,k)=(G(i,k)-B(i,k))/(q-W); \quad \text{cuando } r \text{ es maximo} \\ \text{hHLS}(i,k)=((B(i,k)-R(i,k))/(q-W))+2; \quad \text{cuando } g \text{ es maximo} \\ \text{hHLS}(i,k)=((R(i,k)-G(i,k))/(q-W))+4; \quad \text{cuando } b \text{ es maximo} \end{array} \right.$$

**b) Se halla luminancia bajo el estándar de televisión recomendado por hambury y serra:**

$$IHLS(i,k)=0.213*R(i,k)+0.715*G(i,k)+0.072*B(i,k);$$

**c) Se halla la saturación:**

$$sHLS(i,k)=q-W;$$

**4) Se pasa un filtro pasa bajas para suavizar discontinuidades y se cortan los bordes de la imagen ya que el filtro los oscurece:**

$$H1=ones(4); \quad H2=H1./sum(sum(H1)); \quad T1=conv2(hHLS,H2,'same');$$

$$T1=T1(2:end-1,2:end-1);$$

**5) Se aplica otsu :**

$$[IDX1,sep,kT1]=otsu(T1,oum); \quad [IDX2,sep,kT2]=otsu(T2,ou); \dots$$

**6) Conocer los nuk umbrales normalizados y se crea la máscara total célula segmentada:**

$$masctotal=masLuminancia.*masSaturacion.*masMatiz;$$

**7) Se colorea nuevamente con el color original y se pasa a uint8:**

$$T1nuevo=(R(2:end-1,2:end-1)).*masctotal; \dots$$

$$P=uint8(255*(cat(3,T1nuevo,T2nuevo,T3nuevo)));$$

**8) Se hallan las propiedades de los objetos y se eliminan las áreas que no usamos, además se tiene en cuenta los objetos sobre los bordes de la imagen y los que están por encima de estos se eliminan:**

```
s= regionprops(ff,'area','boundingbox');
```

```
limcadcuad(k,:)=round(s(k,1).BoundingBox) s(k,1).Area];
```

```
for kp =1:a
```

```
    if limcadcuad(kp,5) > .38*max(limcadcuad(:,end))
```

```
        pq=limcadcuad(kp,1)+limcadcuad(kp,3);
```

```
        if pq>ppy
```

```
            pq=ppy;
```

```
        end
```

**9) Se saca cada objeto de interés en un cuadro para analizar individualmente sus características, se hace un corte si hay más de 20 objetos.**

```
limcadcuad1(c,:)= limcadcuad(kp,:);
```

```
lmc=P(limcadcuad(kp,2):limcadcuad(kp,2)+limcadcuad(kp,4),limcadcuad(kp,1):pq,:  
);
```

```
lmcB=ff(limcadcuad(kp,2):limcadcuad(kp,2)+limcadcuad(kp,4),limcadcuad(kp,1):pq  
,:)
```

**10) Hallamos el descriptor de color de cada objeto:**

```
H = RGB2HSV(lmc); matiz(c)=mean2(H(:,:,1));
```

**11) Hallamos parámetros y contorno con strel, edge y dilatamos los bordes con imdilate:**

```
se1 = strel('square',7); EB=imdilate(EB,se1); B2=edge(EB,'canny');
```

**12) Encontramos descriptores de forma signature.m y fractalHB.m:**

**Signatura,regularidad o firma**

```
[aB bB]=find(B2==1); [st, angle, x0, y0] = signature([aB bB]); regularidad=std(st);
```

**Cuadrados fractales**

```
D0=calcD(~frac0.*B2,N1,a);DT1=calcD(~(frac0+fracD1).*B2,N1,a);DT2=calcD(~(frac0+fracD1+fracD2).*B2,N1,a);
```

**13) Hallamos descriptor de textura con ayuda de la matriz de coocurrencia :**

```
nuimagtex=cooccurrence(E(:, :, 1),oo, 1, 1);
```

```
[energia,entropia,homogen,correlac]=cocaract2(nuimagtex,1);
```

**14) Llamamos la red neuronal entrenada y entregamos diagnóstico:**

```
load esce1 LEUC
```

```
EC=seronoser(daaat');
```

```
datos=[EC;daaat(2);daaat(6)];
```

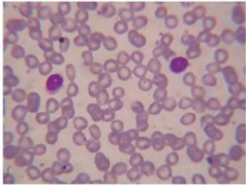
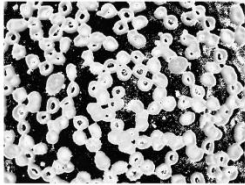
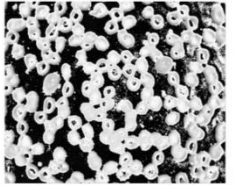
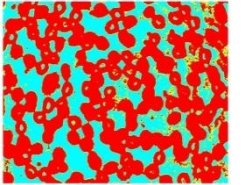
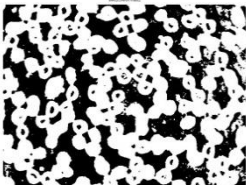
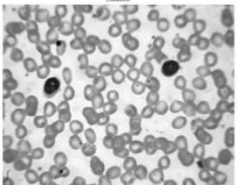
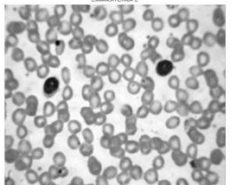
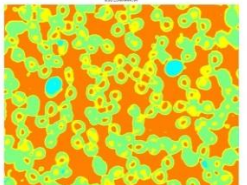

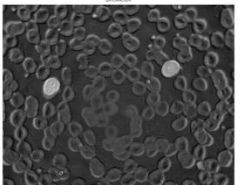
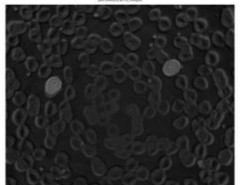
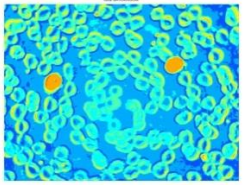
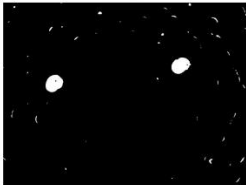



```
EC= sim(LEUC,datos);
```

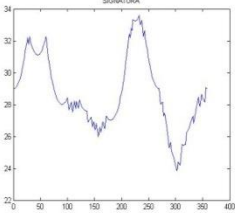
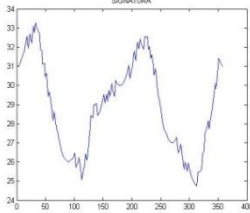
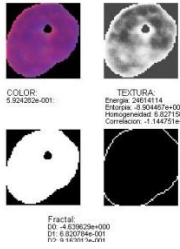
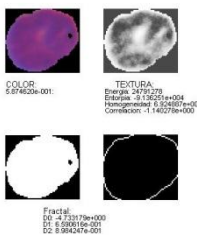
```
if EC>0,6
```

```
'sana'
else
'Enferma'
End
```

Todo el procedimiento anterior se puede observar en la Tabla 6.

**Tabla 6. Proceso de Segmentación características malignas célula tipo A.**

 <p>Imagen en RGB</p>	 <p>Matiz</p>	 <p>Matiz Filtrada</p>	 <p>Otsu del Matiz</p>
 <p>Mascara Matiz</p>	 <p>Luminancia</p>	 <p>Luminancia Filtrada</p>	 <p>Otsu de Luminancia</p>
 <p>Mascara Luminancia</p>	 <p>Saturación</p>	 <p>Saturación Filtrada</p>	 <p>Otsu Saturación</p>
			

Mascara Saturación	Mascara Total; Convolucion De Todas Las Mascaras	Fractal Célula De La Izquierda	Fractal Célula De La Derecha
 <p>Signatura Célula Izquierda</p>	 <p>Signatura Célula Derecha</p>	 <p>Descriptores de textura, forma y color célula izquierda</p>	 <p>Descriptores de textura forma y color célula derecha</p>

#### 4.7. SEGURIDAD DE LA PRUEBA DIAGNOSTICA

Para las pruebas se cuenta con un banco de imágenes de 468 tomas de extendidos sanguíneos de células enfermas y sanas. La prueba de confiabilidad de la herramienta se hizo con 60 imágenes, para la cual se calculó la sensibilidad y especificidad de la medida analizando los errores y aciertos cometidos.

Se revisarán los conceptos que determinan la validez de un test (sensibilidad y especificidad) y su seguridad (valores predictivos positivos y negativos) (25).

##### Sensibilidad

Es la probabilidad de clasificar correctamente a una célula enferma, es decir la probabilidad que para una imagen de células enfermas se obtenga en la prueba

un resultado positivo. La sensibilidad es por lo tanto la capacidad del modelo computacional para detectar células enfermas.

$$\text{Sensibilidad} = \frac{VP}{VP+FN} \quad (4.9)$$

Donde VP corresponde a Verdaderos Positivos y FN a Falsos Negativos.

### **Especificidad**

Es la probabilidad de clasificar correctamente a una célula sana, es decir la probabilidad que para una imagen de células sanas se obtenga un resultado negativo. En otras palabras se puede definir la especificidad como la capacidad del modelo de identificar células sanas.

$$\text{Especificidad} = \frac{VN}{VN+FP} \quad (4.10)$$

Donde VN corresponde a Verdaderos Negativos y FP a Falsos Positivos.

Los conceptos de sensibilidad y especificidad permiten valorar la validez de una prueba diagnóstica, sin embargo, carecen de utilidad en la práctica clínica.

Tanto la sensibilidad como la especificidad proporcionan información acerca de la probabilidad de obtener un resultado concreto (positivo o negativo) en función de la verdadera condición del enfermo con respecto a la enfermedad. Sin embargo, cuando a algún paciente se le realiza alguna prueba, el medico carece de información a priori acerca de su verdadero diagnóstico, y más bien la pregunta se plantea en sentido contrario: ante un resultado positivo (negativo) en la prueba, ¿cuál es la probabilidad de que el paciente este realmente enfermo (sano)? Así pues, resulta obvio que hasta el momento solo se ha abordado el problema en una

sola dirección. Por medio de los valores predictivos se complementa esta información.

### **Valor Predictivo Positivo**

Es la probabilidad de padecer la enfermedad, o de la célula ser catalogada como maligna, si se obtiene un resultado positivo mediante el modelo computacional.

El valor predictivo positivo puede estimarse, por tanto, a partir de la proporción de células con un resultado positivo en la prueba que finalmente resultaron ser malignas:

$$VPP = \frac{VP}{VP+FP} \quad (4.11)$$

Siendo  $VP$  Verdaderos Positivos y  $FP$  Falsos Positivos.

### **Valor Predictivo Negativo**

Es la probabilidad de que una célula con un resultado negativo en el modelo computacional este realmente sana. Se estima dividiendo el número de verdaderos negativos entre el total de células con un resultado negativo a través del modelo propuesto:

$$VPN = \frac{VN}{VN+FN} \quad (4.12)$$

Siendo  $VN$  Verdaderos Negativos y  $FN$  Falsos Negativos.

Se tomaron 84 imágenes de las cuales obtuvimos los siguientes resultados:

**Tabla 7. Clasificación celular realizada por el modelo computacional.**

Clasificación Real	Clasificación del modelo computacional
Negativa	68
Positiva	10

Fuente: Autores del Proyecto

**Tabla 8. Numero de Falsos Positivos y Negativos obtenidos con el modelo computacional.**

Falsos negativos	4
Falsos positivos	2

Fuente: Autores del Proyecto

Lo cual arrojo los siguientes resultados: Sensibilidad= 94,4%, Especificidad= 83,3%, VPP= 97,1%, VPN=71,3 %.

Para aumentar a futuro los valores de especificidad y sensibilidad y disminuir los falsos positivos y negativos y así tener una mayor probabilidad de éxito se recomienda tener en cuenta otros parámetros y características como: la cantidad de glóbulos rojos contra glóbulos blancos presentes en un mismo extendido, la relación núcleo citoplasma de las células y el tamaño de los glóbulos blancos en la sangre periférica para hacer más minucioso el proceso de clasificación y lograr mejores resultados .

#### **4.8. CONCLUSIONES Y RECOMENDACIONES**

Como pudo observarse en la Figura 1.3 la evolución de una célula sanguínea pasa por diferentes etapas de desarrollo antes de llegar a la sangre periférica donde debe cumplir con sus funciones naturales; una célula patógena en cambio se caracteriza por estar de manera prematura en el torrente sanguíneo de forma irregular ya que no ha cumplido su ciclo de desarrollo habitual, es decir es una célula inmadura tratando de cumplir funciones de defensa ante los virus en el caso de los glóbulos blancos.

Una célula inmadura se caracteriza como se vio en la Tabla 2 por tener nucléolos evidentes en el linfoblasto, dos o más divisiones en el núcleo, multilobulaciones, cromatina nuclear conglomerada, hendiduras causantes de posible división, manchas o tizón que demuestra la debilidad de la célula y por eso su sensibilidad ante el colorante, manchas negras en el núcleo lo cual evidencia casos de fagocitosis, al contrario una célula madura se caracteriza por tener una textura homogénea y casi esferoidal lo cual demuestra que cumplió su ciclo de desarrollo normal.

La toma de imágenes permitió tener un contacto inicial con los extendidos sanguíneos los cuales no fueron fáciles de obtener ya que una persona diagnosticada con leucemia es intervenida lo más pronto posible con quimioterapia para empezar a atacar de manera temprana la enfermedad esto ocasiona que los glóbulos blancos empiecen a deformarse dependiendo del tipo de químico utilizado hay mayor o menor trastorno por ende es aconsejable analizar los datos de una muestra no tratada ya que mantienen las características de las células patógenas sin alteraciones.

Los algoritmos empleados y el procesamiento implementado mostraron resultados positivos que se pueden mejorar, de esta manera se cuenta con una herramienta computacional capaz de disminuir el número de falsos positivos y en el mejor de los casos también reducir la cantidad de falsos negativos.

Si bien no se alcanzaron resultados satisfactorios, la continuidad de la investigación debería ampliar su universo de imágenes, abriendo horizontes hacia la detección de sitios primarios para características de células malignas de cualquier tipo en los glóbulos blancos, favoreciendo de esta manera la detección oportuna de cáncer y su pronta intervención que redunde en beneficio del paciente y su círculo familiar.

El campo de la Inteligencia Artificial es un amplio terreno sobre el cual explorar y experimentar, buscando la creación de nuevos y mejores algoritmos que aplicados al campo del tratamiento de imágenes médicas impacte de manera positiva no solo en el dominio académico e institucional sino también en el más importante, la salud de la población más vulnerable quien en últimas resulta ser la más beneficiada con estas técnicas de bajo costo económico y computacional.

## REFERENCIAS:

1. “*La morfología de las células de la sangre humana*”. **Diggs, Sturm, Dorothy y Bell, Ann.** 1971, frotitis de sangre periférica y de médula ósea teñidos con el colorante de Wright, págs. 40-41.
2. Leucemia-linfoblástica infantil. [En línea]  
<http://www.cancer.gov/espanol/pdq/tratamiento/leucemia-linfoblasticainfantil/patient>.
3. **Organization, World Health.** *Classification of tumors of haematopoietic and lymphoid tissues.* s.l. : Lyon IARC Press, 2001.
4. Evolución de una célula sanguínea. [En línea]  
<http://www.cancer.gov/PublishedContent/MediaLinks/448392.html>.
5. **Martínez, Bibiana, y otros, y otros.** *Relevancia de la morfología a través del tiempo en el diagnóstico de la Leucemia Linfocítica Aguda. Grupo de Hematología Especial del Instituto de Cancerología.* Bogotá, Colombia : s.n., 2007.
6. **Pérez, B., Gómez, X. y Patricia, M.Sc.** *Diagnóstico histoquímico de Leucemia Linfoblástica aguda infantil. Instituto SELADIS Facultad de Cs. Farmacéutica y Bioquímica UMSA.* 1998.
7. Medicina-diagnóstica. [En línea]  
[http://es.wikipedia.org/wiki/Falso\\_positivo\\_\(medicina\)](http://es.wikipedia.org/wiki/Falso_positivo_(medicina)).
8. Size. [En línea] <http://www.mathworks.com/help/matlab/ref/size.html>.
9. “*Detección de Objetos por Segmentación Multinivel Combinada de Espacios de Color*”. **Gil, P., Torres, F. y Ortiz, F.G.** España : s.n., 2004. XXV Jornadas de Automática. Ciudad Real.
10. **Hanbury, A. y Serra, J.** “Colour Image Analysis in 3D-polar Coordinates”. [aut. libro] Serra J. Hanbury A. *DAGM.* Magdeburg, Germany. : s.n., 2003.

11. *"Image Colour Segmentation Using Bi-Variate Histograms in Luminance/Saturation/Hue Polar Color Space"*. **Angulo, J. y Serra, J.** 4, Mexico : s.n., 2005, Computación y Sistemas, Vol. 8.
12. Universidad Nacional de Quilmes. Ingeniería en Automatización y Control Industrial. *Visión Artificial. Segmentación por Umbralización. Método otsu*. [En línea] Octubre de 2005. <http://iaci.unq.edu.ar/materias/vision/apuntes.htm>.
13. Histograma. [En línea] <http://es.wikipedia.org/wiki/Histograma>.
14. **Romero, F. J.** Fundamentos y algoritmos de visión Artificial. *Integración de sistema de visión artificial y robot en aplicación tipo Pick&Place*. s.l. : Tesis de Grado, 2006. Vol. III.
15. **Gonzalez, R., Woods, R. y Eddins, S.** *Digital Image Processing: Using MATLAB*. 1. s.l. : Upper Saddle: Pearson Education, 2004.
16. Algoritmo del método de Otsu. [En línea] [www.biomecardio.com/matlab/otsu.m](http://www.biomecardio.com/matlab/otsu.m).
17. *"La matriz de coocurrencia en la clasificación multiespectral."*. **Pressuti, Miriam y Venetsanopoulos, A.N.** 11 a 13 de agosto de 2004, São Leopoldo, RS, Brasil : s.n., 2004. Universidad nacional de la plata, 4<sup>a</sup> Jornada de Educação em Sensoriamento Remoto no Âmbito do Mercosul.
18. [www.mathworks.com](http://www.mathworks.com). [En línea] [www.mathworks.com/matlabcentral/fileexchange/11727-cooccurrence-matrix/content/cooccurrence.m](http://www.mathworks.com/matlabcentral/fileexchange/11727-cooccurrence-matrix/content/cooccurrence.m).
19. Código Signature. [En línea] [codigo.signature.m/fourier.eng.hmc.edu/e161/dipum/signature.m](http://codigo.signature.m/fourier.eng.hmc.edu/e161/dipum/signature.m).
20. Geometriafractal. [En línea] [www.geometriafractal.com/geometriafractal.asp](http://www.geometriafractal.com/geometriafractal.asp).
21. rgb2hsv. [En línea] <http://www.mathworks.com/help/matlab/ref/rgb2hsv.html>.
22. Textura. [En línea] <http://bibing.us.es/proyectos/abreproy/11494/direccion/PROYECTO%252F/CAPITULO3/TEXTURA>.

23. CARACTERIZACION DE TEXTURAS. [En línea]

<http://bibing.us.es/proyectos/abreproy/11494/direccion/PROYECTO%252F/CAPITULO5/CARACTERIZACION DE TEXTURAS>.

24. **Hu, Yu Hen y Hwang, Jenq-Neng.** *Handbook of NEURAL NETWORK SIGNAL PROCESSING*. Boca Raton : CRC Press LLC, 2002.

25. **Pita, F. S. y P´ertegas, D. S.** *Pruebas diagnosticas*. s.l. : Cad Aten Primaria, 2003.

## ANEXOS

### ANEXO A SOFTWARE

Carpeta leucemia contiene: programas en .m leucemia1.m y leucemia2.m , junto a sus respectivas funciones.

#### ANEXO B. Tabla de datos usados se encuentra en dat.xls

##### Algoritmo Leucemia2.m

```

clc
clear all
close all

tic
[filename, pathname] = uigetfile( ...
    {'*.jpg;*.png;*.bmp;*.tga', 'Image Files (*.jpg,*.png,*.bmp,*.tga)'}, ...
    'Select your file');
E = imread([pathname,filename]);
fact=.25;
E=imresize(E,fact);%reducimos el tamaño a un cuarto de la imagen
D=double(E);%convierto a doble para poder operar
R=D(:,:,1)/255;
G=D(:,:,2)/255;
B=D(:,:,3)/255;
[a,b,CAP]=size(D);
imshow(E)
    for i=1:a
        for k=1:b

            q=max([R(i,k),G(i,k),B(i,k)]);%HALLO VALOR MAXIMO DE RGB
            W=min([R(i,k),G(i,k),B(i,k)]);

            if q==W
                hHLS(i,k)=0 ; %elimino los posible NaN DIVISION POR 0
            else

```

```

        if R(i,k)==q %si r=max
        hHLS(i,k)=(G(i,k)-B(i,k))/(q-W);
        hHLS(i,k)=gradohsl(hHLS(i,k));

        end
        if G(i,k)==q%si g=max
        hHLS(i,k)=(B(i,k)-R(i,k))/(q-W)+2;
        hHLS(i,k)=gradohsl(hHLS(i,k));
        end
        if B(i,k)==q%si b=max
        hHLS(i,k)=(R(i,k)-G(i,k))/(q-W)+4;
        hHLS(i,k)=gradohsl(hHLS(i,k));
        end
        end
        %% HALLO lHLS
        lHLS(i,k)=0.213*R(i,k)+0.715*G(i,k)+0.072*B(i,k);
        %HALLO sHLS
        sHLS(i,k)=q-W;
        end
    end
    H1=ones(4);
    H2=H1./sum(sum(H1));
    T1=conv2(hHLS,H2,'same');
    T2=conv2(lHLS,H2,'same');
    T3=conv2(sHLS,H2,'same');
    T1=T1(2:end-1,2:end-1);
    T2=T2(2:end-1,2:end-1);
    T3=T3(2:end-1,2:end-1);
    %%% calculamos umbrales y regions de otsu
    oum=4;
    ou=4;
    oo=8;
    [IDX1,sep,kT1]=otsu(T1,oum);
    [IDX2,sep,kT2]=otsu(T2,ou);
    [IDX3,sep,kT3]=otsu(T3,oo);
    kT1=kT1/255;
    kT2=kT2/255;
    kT3=kT3/255;

    %%BUSCAMOS HISTOGRAMA L/S con la ayuda de scatter para luego fijar
    %%los umbrales donde se encuentran los objetos de interés.
    [xi,yi,value]=impixel(T3,[])
    lc=T2(min(yi):max(yi),min(xi):max(xi));
    sc=T3(min(yi):max(yi),min(xi):max(xi));
    figure,imshow(sc,[])
    figure,scatter(lc(:),sc(:)),xlabel('luminancia'),ylabel('saturacion')

    % CREACIÓN DE MASCARA
    masSaturacion0=(IDX3==7);
    masSaturacion0=(IDX3==8);
    masSaturacion=masSaturacion0+masSaturacion0;
    
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
masLuminancia=(IDX2==1);%1
masLuminancia2=(IDX2==2);%1
[a1 b1]=find(masLuminancia==1);
[a12 b12]=find(masLuminancia2==1);
tamL=length(a1);
tamL2=length(a12);
sum=0;
sum2=0;
for ku=1:tamL
    sum=T2(a1(ku),b1(ku))+sum;
end
for ku1=1:tamL2
    sum2=T2(a12(ku1),b12(ku1))+sum2;
end

if sum2/tamL2 >.6
    masLuminancia=masLuminancia+ masLuminancia2;
end
%hallamos la mascara total
masMatiz=(IDX1==5);%
masMatiz1=(IDX1==4);%
masMatiz2=(IDX1==3);%
masMatiz=masMatiz+masMatiz1+masMatiz2;

masctotal=masLuminancia.*masSaturacion.*masMatiz;
T1nuevo=(R(2:end-1,2:end-1)).*masctotal;
T2nuevo=(G(2:end-1,2:end-1)).*masctotal;
T3nuevo=(B(2:end-1,2:end-1)).*masctotal;
P=uint8(255*(cat(3,T1nuevo,T2nuevo,T3nuevo)));
ff=logical(masctotal);

s= regionprops(ff,'area','boundingbox');
[a,b]=size(s);
for k=1:a
    limcadcua(k,:)= [round(s(k,1).BoundingBox) s(k,1).Area];
end
%return
c=1;
[ppx ppy oi]=size(P);
%ignoramos areas no utiles
for kp =1:a
    if limcadcua(kp,5) > .38*max(limcadcua(:,end))
        pq=limcadcua(kp,1)+limcadcua(kp,3);
        if pq>ppy
            pq=ppy;
        end
        limcadcua1(c,:)= limcadcua(kp,:);
    end
end

```

```

Imc=P(limcadc cuad(kp,2):limcadc cuad(kp,2)+limcadc cuad(kp,4),limcadc cuad(kp,1)
:pq,:);
ImcB=ff(limcadc cuad(kp,2):limcadc cuad(kp,2)+limcadc cuad(kp,4),limcadc cuad(kp,
1):pq,:);
        H = RGB2HSV(Imc);
        matiz(c)=mean2(H(:, :, 1));
        ImccwB(c)={ImcB};
        ImccwC(c)={Imc};
        Imcc=rgb2gray(Imc);
        Imcc=ecualizada(uint8(Imcc));
        Imccw(c)={Imcc};
        c=c+1;
    end
end
%nos preparamos para hallar signatura

oo=2;
dtci=10;
format short
datos=zeros(5,2);
cmn=1;
if c>20
    c=20;
end

for tt=1:c-1

    E=Imccw{1,tt};
    EEE=E;
    [a,b]=size(E);
    E=[zeros(dtci,b); E; zeros(dtci,b)];
    E=[zeros(20+a,dtci) E zeros(20+a,dtci)];
    E1=ImccwC{1,tt};
    EB=ImccwB{1,tt};

    EBB=EB;
    EB=[zeros(10,b); EB; zeros(10,b)];
    EB=[zeros(20+a,10) EB zeros(20+a,10)];
    [afs acs]=size(EB);
    if afs*acs >2500
        sel = strel('square',7);
        EB=imdilate(EB,sel);
    end
    B2=edge(EB, 'canny');
    [abB abB1]=size(B2);
    [B,L,N,A] = BWBOUNDARIES(B2);

    may=0;

```

```

pc=1;
for hjk=1:length(B)-1
    [as bs]=size(B{hjk});
    [ass bss]=size(B{hjk+1});

    if as>ass & as>may;
        may=as;
        pc=hjk;
    end
end

end

Bb2=B{pc};
Bj=zeros(abB, abB1);
for kh=1:length(Bb2)
    Bj(round(Bb2(kh,1)),round(Bb2(kh,2)))=1;
end
B2=Bj;

[aB bB]=find(B2==1);
[st, angle, x0, y0] = signature([aB bB]);
regularidad=std(st);
%usamos matriz de coocurrencia

B2=B2(9:a+11,9:b+10);
nuimagtex=cooccurrence(E(:, :, 1),oo, 1, 1);
[energia,entropia,homogen,correlac]=cocaract2(nuimagtex,1);

%%se hallan los tres fractales
ax=1.2*a;
bx=1;
    if a>30
        bx=5;
    end
    if a>45
        bx=5;
    end
[sa sb]=find(B2==1);
N1=length(sa);
[frac0, fracD1, fracD2]=fractalHB(B2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
D0=calcD(~frac0.*B2,N1,a);
DT1=calcD(~(frac0+fracD1).*B2,N1,a);
DT2=calcD(~(frac0+fracD1+fracD2).*B2,N1,a);

figure
set(gcf, 'color', 'white', 'Position', [350,120,350,420])
    
```

```

title('clasificacion')
subplot(221),imshow(E1)
text(1,ax,sprintf('%s: %d','COLOR'));

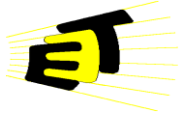
text(1,ax+bx,sprintf('%s: %d',matiz(tt),'FontSize',8);
subplot(222),imshow(EEE)
text(10,ax,sprintf('%s: %d','TEXTURA'),'FontSize',10);
text(2,ax+bx,sprintf('%s: %d','Energia',
energia),'FontSize',8);
text(2,ax+bx*2,sprintf('%s: %d','Entropia',
entropia),'FontSize',8);
text(2,ax+bx*3,sprintf('%s:
%d','Homogeneidad',homogen),'FontSize',8);
text(2,ax+bx*4,sprintf('%s:
%d','Correlacion',correlac),'FontSize',8);
subplot(223),imshow(EBB)
text(30,ax,sprintf('%s: %d','Fractal'),'FontSize',10);
text(30,ax+bx,sprintf('%s: %d','D0',D0),'FontSize',8);
text(30,ax+bx*2,sprintf('%s: %d','D1',DT1),'FontSize',8);
text(30,ax+bx*3,sprintf('%s: %d','D2',DT2),'FontSize',8);
subplot(224),imshow(B2)

qwe=[matiz(tt) energia/1000000 entropia/1000000 homogen/1000
correlac regularidad D0 DT1 DT2];
%AA=mean(qwe);
pause(.5)
cmn=cmn+1;
format short

daaat(1)=qwe(1);
daaat(2)=qwe(2);
daaat(3)=qwe(3);
daaat(4)=qwe(4);
daaat(5)=qwe(5);
daaat(6)=qwe(6);
%daaat(7)=qwe(7)/2;
daaat(7)=qwe(8);
daaat(8)=qwe(9);

% llamamos la red
load escell LEUC
EC=seronoser(daaat');
datos=[EC;daaat(2);daaat(6)];
EC= sim(LEUC,datos);
if EC>.6
'sana'
else
'Enferma'
end

```



end