

Estimación del porcentaje de linfoblastos para diagnóstico precoz de leucemia haciendo uso de técnicas de visión artificial

Ruben Darío Angel Correa

Universidad Cooperativa de Colombia, Bogotá D. C., Colombia, rangelcorrea@gmail.com

ABSTRACT

The first part of this document talks about blood tissue components as white or red corpuscles and their normal averages. In the second part is made an analysis of different kinds of leukemia so as their characteristics and percents of white corpuscles (blástos cell in a leukemia) which play a determinant role when we are going to diagnostic the state of this disease. After this and in the third part, is made a work about the different images processing techniques and their applications on the artificial vision systems. Then, we will procedure to do a segmentation of images in order to separate blástos cell from other parts of tissue blood. On the last part we will do a count about how many cells of each one blástos cells and red corpuscles there are in a sample of blood tissue. With these amounts we can do an appreciation about the relation of the blástos cell with the red corpuscles. When we have the percent of both kinds of cells we can do a premature prediction of the disease state which is our objective in this work. It would be a treatment images application in medicine.

Keywords: leukemia, images, artificial, vision, blástos.

1. INTRODUCTION

Cuando se trata del análisis de muestras de sangre, son muchas las características que el analista debe entrar a considerar y muchos los elementos u objetos que en esta muestra pueden aparecer. Afortunadamente las rutinas implementadas permiten de manera rápida y eficaz obtener conclusiones en lapsos de tiempo relativamente cortos. Sin embargo este hecho no quita el que dichas rutinas puedan volverse tediosas y repetitivas. Por ejemplo cuando se trata de contar glóbulos rojos, esto puede ser bastante complicado sobre todo cuando la resolución es baja. En el presente trabajo se presenta una técnica que podría ser bastante útil a la hora de contar los objetos presentes en un frotis sanguíneo, en el cual, partiendo de las imágenes disponibles y de manera automática, permite diferenciar unos objetos de otros y determinar en qué relación porcentual se encuentran. El trabajo asume que las muestras han sido previamente tratadas con lo cual ya se tendrá una previa separación de componentes por medio de mecanismos de tinción.

2. LAS LEUCEMIAS

La leucemia es un tipo de cáncer que afecta a los glóbulos blancos de la sangre. Normalmente los glóbulos blancos defienden al cuerpo de las infecciones y lo protegen de las enfermedades. Pero en la leucemia, los glóbulos blancos se vuelven cancerosos y se multiplican en circunstancias en las que no deberían, produciéndose en consecuencia una elevada cantidad de glóbulos blancos anormales que después interfiere con el funcionamiento de los órganos. *La leucemia se define como la proliferación neoplásica de células*

hematopoyéticas en una estirpe celular con posterior proliferación y expansión, cuya acumulación se acompaña de una disminución del tejido hematopoyético normal en médula ósea y posterior invasión de sangre periférica y otros tejidos. [1]. En las leucemias agudas la población celular predominante está formada por células inmaduras (blástos), y en las crónicas la celularidad presenta un mayor estadio madurativo. Su diagnóstico se realiza mediante el examen de sangre periférica y de médula ósea. En cuanto a la etiopatogenia se han descrito ciertas cromosomopatías, radiaciones ionizantes, fármacos mielotóxicos, virus y ciertos factores genéticos. La hematopoyesis es el proceso por el cual se forman las células de la sangre en la médula ósea a partir de una célula madre pluripotencial o *stem cell*. A partir de dicha célula se originaran las diferentes series: eritrocitos, leucocitos y plaquetas.

Existen diferentes tipos de leucemia como los son: leucemias mieloides agudas, leucemias linfoides agudas, leucemia mieloide crónica y leucemia linfática crónica la cual es la más frecuente en los pacientes adultos de los países occidentales. Es una enfermedad propia del anciano con predominancia del hombre sobre la mujer. La leucemia linfoide crónica que es la que en este trabajo se va a analizar incluye un cuadro clínico cuyos síntomas más frecuentes son adenopatías generalizadas con astenia y pérdida de peso. No es infrecuente una historia previa de infecciones recurrentes, causadas por la hipogammaglobulinemia que suelen presentar estos enfermos. Lo más característico es la presencia de una linfocitosis absoluta que varía de 5 a $500 \times 10^9/l$, pero en la mayoría de los casos es superior a $20 \times 10^9/l$. En el 20% de los pacientes se halla anemia o plaquetopenia. Un 10 a 25% de los casos desarrollan anemia hemolítica autoinmune que en ocasiones se pone de manifiesto al iniciar el tratamiento. La plaquetopenia autoinmune es poco frecuente (2%), y la aplasia pura de la serie roja y neutropenia autoinmune son todavía más raras. La figura 1 ilustra un resumen sobre el cuadro clínico de este tipo de leucemias. [2]

Estadio	Hallazgos clínicos	Supervivencia (meses)
0	Linfocitos	<120
I	Linfocitos y adenopatías	95
II	Linfocitos y hepato y/o esplenomegalia	75
III	Linfocitos y anemia (Hb<11 gr/dl)	30
IV	Linfocitos y trombocitemia (Plts < $100 \times 10^9/l$)	30

Figura 1: cuadro clínico de la leucemia linfoide crónica.

3. DETECCIÓN DE LINFOBLASTOS A PARTIR DE LA IMAGEN DEL TEJIDO SANGUÍNEO.

En el procesamiento digital de imágenes existen diferentes técnicas cuya aplicación metodológica nos permitirán en última instancia realizar una segmentación y posterior separación de los diferentes objetos presentes en la imagen. Dentro de estas técnicas se encuentran pasos procedimentales tales como: eliminación del ruido, detección de bordes, binarización, extracción de características y por último el reconocimiento del objeto. Para la aplicación de estas técnicas se supone un previo conocimiento de la imagen y cuáles han de ser los objetos a separar. En lo que a nuestro trabajo nos incumbe, la imagen indicada en la figura 2 se corresponde con una muestra de tejido sanguíneo a la cual se le ha realizado un proceso de tinción con el fin de revelar la presencia de las células cancerosas linfoblastos o más sencillamente blástos.

Una vez se tiene la imagen, el primer paso a aplicar será la mejora de la imagen, lo cual consiste en mejorar aquellas características tales como el brillo, contraste y nivel de ruido. Por ejemplo, para mejorar el contraste se puede aplicar la técnica basada en el histograma denominada como ecualización uniforme [3], la cual pretende que ante cualquier histograma este se modifique de tal manera que en la imagen resultante los niveles de gris se repartan de forma equitativa en todo el margen dinámico. Siguiendo con la metodología indicada, se tendría la mejora en las condiciones de ruido y la posterior binarización.

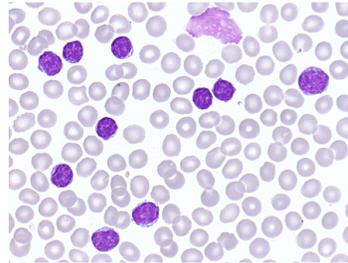


Figura 2: Imagen indicando la presencia de células blástos.

Aplicando estas técnicas a la imagen de la figura 2, se obtiene la imagen indicada en la figura 3 en la cual la elección de un umbral de binarización adecuado nos permite realizar la separación de las células blásto. Obsérvese que la imagen de la figura 3 b) se encuentra substraída en la figura 3 a).

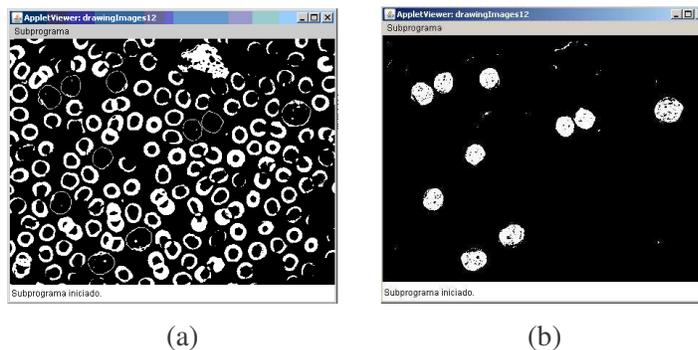


Figura 3: Imagen filtrada y binarizada. a) sin células blásto b) células blásto segmentadas.

Una vez se han separado las células blastos de los demás componentes del tejido sanguíneo, el paso siguiente será realizar el conteo de los objetos contenidos en las imágenes particulares. Para lograr este objetivo se ha implementado la técnica denominada como “*ventaneo*” la cual consiste en generar una ventana con dimensiones acordes con las dimensiones del objeto que se desea detectar. Luego de esto se debe barrer toda la imagen línea por línea y columna por columna e ir *cortando* aquellos objetos que cumplan con las características del objeto que se desea detectar. En la figura 4 se indica el algoritmo utilizado para este fin. El algoritmo es igualmente válido si se van a detectar células blastos o glóbulos rojos o eritrocitos. La diferencia de cual ha de ser el camino a tomar la define las características del objeto a detectar como lo son el área y la relación de aspecto en este caso. Para los objetos considerados: glóbulos rojos y células blástos estas características son diferentes y así lo serán las ventanas implementadas para cada caso.

Para la implementación de las ventanas móviles se hizo uso de las funciones de procesamiento de imágenes existentes en el lenguaje Java[4] con algunas variaciones que permitieron la adecuación al trabajo en particular. Por ejemplo, mediante la aplicación de estas funciones se hizo posible el conteo de pixeles contenidos ya sea en una célula blasto o en un eritrocito y cuyas salidas de programa se encuentran indicadas en la figura 5. El último paso para la determinación e identificación total de los objetos capturados en el anterior paso consiste en el análisis de su forma con el fin de descartar aquellos que no pertenezcan al conjunto que se está separando en este caso las células blástos.

El análisis de forma consiste en hallar la relación de aspecto la cual viene a ser el cociente entre la máxima longitud horizontal sobre la máxima longitud vertical.

$$Ra = \frac{\text{máxima longitud horizontal}}{\text{máxima longitud vertical}}$$

Tanto para los blástos como para los glóbulos rojos se encontró que este valor debe estar en un rango entre 0.8 y 1.2 y no depende del tamaño de los mismos.

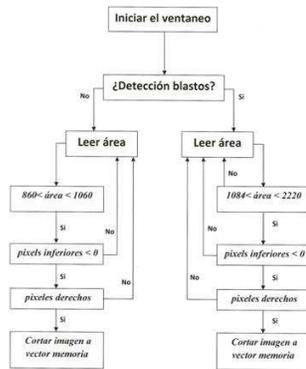


Figura 4. Algoritmo utilizado para la detección de objetos.

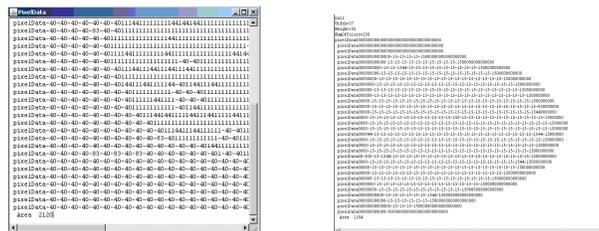


Figura 5. Conteo de pixeles. a) células blásto b) eritrocitos.

4. ANÁLISIS DE RESULTADOS.

De nueve imágenes procesadas 2 presentaron error respecto al número de objetos clasificados lo cual representa un error del 22%. El error de clasificación en las imágenes individuales está por debajo del 5%, esto es debido a la buena segmentación que se puede realizar con este tipo de células. Por otro lado, teniendo en cuenta las imágenes segmentadas de los glóbulos rojos, de nueve imágenes procesadas nueve presentaron error en cuanto al número de

objetos clasificados lo cual da un error del 100%, pero el error de clasificación cometido en cada imagen está por debajo del 20 por ciento, esto es, si se reconocen 35 objetos siendo realmente cuarenta, el error de clasificación será del 12,5%, lo cual se puede considerar como un criterio de buena clasificación. En ésta etapa no es necesario reconocer individualmente los glóbulos rojos sino tener en cuenta todos los objetos presentes en la imagen los cuales también hacen parte del tejido sanguíneo. Esto quiere decir, que siempre se cometerá algún tipo de error en la clasificación de glóbulos rojos pues siempre se habrá de clasificar otro tipo de componentes del tejido sanguíneo junto con los glóbulos rojos. Al fin y al cabo lo que se necesita es hacer una estimación del porcentaje de estos y no un conteo exacto de su número real.

En cuanto al tamaño de la muestra, la clasificación opera de igual modo independientemente del número de éstas, siempre y cuando se mantengan las condiciones impuestas en cuanto a iluminación y resolución. En el algoritmo no se ha previsto un cambio en estas condiciones, por lo tanto todas las imágenes a ser estudiadas deben cumplir con estas restricciones. Esto no se considera que deba ser una mejora posterior al algoritmo porque realmente el funcionamiento de este no debe ser generalizado, siempre se deberán cumplir restricciones en cuanto a las condiciones de la muestra.

5. CONCLUSIONES

- **Respecto de las imágenes.** Con el fin de obtener los mejores resultados, es necesario poder contar con imágenes que hayan sido obtenidas con las mismas condiciones de iluminación y preferiblemente con la misma resolución (x100) ya que en el proyecto no se realizan correcciones de este tipo.

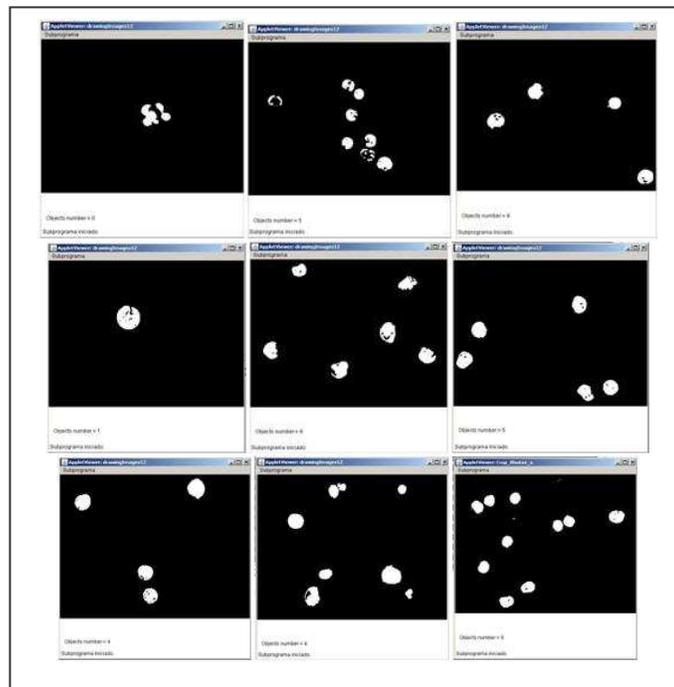


Figura 6. Resultados obtenidos en la clasificación de blásts.

- **Respecto a la clasificación de blásts.** En esta parte se obtuvieron buenos resultados lo cual es consecuencia de las buenas imágenes utilizadas. Se puede mejorar la técnica incluyendo uno o más descriptores de forma como por ejemplo la ausencia de huecos en la imagen o su compacidad.

- **Respecto a la clasificación de glóbulos rojos.** Si se habla de clasificación de glóbulos rojos como tales, los resultados serían muy deficientes debido a la gran cantidad de imágenes solapadas; sin embargo si se habla de objetos presentes en la imagen segmentada, los resultados son mejores debido a que se pueden tomar áreas grandes que pueden ser consideradas como dos o más objetos dependiendo de su magnitud, esto debido a que los glóbulos rojos presentan áreas aproximadamente regulares.
- **Respecto al proyecto en general.** La técnica del ventaneo es perfectamente válida y puede ser mejorada incluyendo los demás descriptores de forma indicados anteriormente. Así como se aplicó en este caso a la detección de células cancerígenas, puede ser aplicada a otro tipo de objetos presentes en la sangre o en otro tipo de sustancias, el algoritmo variaría levemente a la hora de tomar decisiones para el corte de objetos de la imagen.

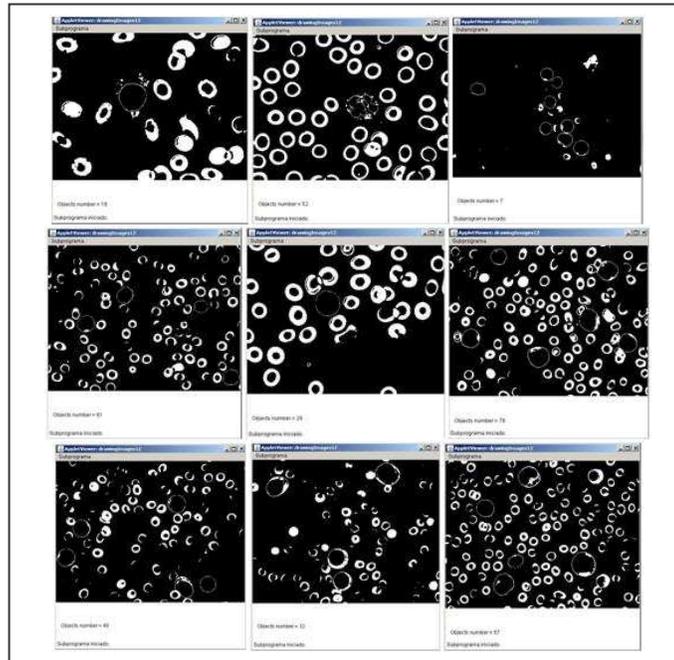


Figura 7. Resultados obtenidos en la clasificación de glóbulos rojos.

6. REFERENCIAS.

- [1] Rosell, Marco y Renau. “Leucemias”. Servicio de hematología H. U. Dr Peset. Valencia. 2008
 [2] Van der Meer, Wim. “Atlas of blood cell differentiation”. Ed Elsevier.2001
 [3] Forero, Manuel Guillermo. “Introducción al tratamiento digital de imágenes”. Silueta ediciones. 2001
 [4] Java documentation. Image processing. <http://www.javasun.download>.

Authorization and Disclaimer

Authors authorize LACCEI to publish the paper in the conference proceedings. Neither LACCEI nor the editors are responsible either for the content or for the implications of what is expressed in the paper.