Segmentation of Bone Marrow Cell Images for Morphological Classification of Acute Leukemia

Carolina Reta, Leopoldo Altamirano, Jesus A. Gonzalez, Raquel Diaz, and Jose S. Guichard

National Institute for Astrophysics, Optics, and Electronics
Luis Enrique Erro No. 1, Puebla, Mexico 72840
E mails: {creta, robles, jagonzalez, raqueld, jguichard}@inaoep.mx

Abstract

This paper presents a two phase methodology to analyze the morphology of abnormal leukocytes images for the classification of acute leukemia subtypes using image processing and data mining techniques. In the first phase we propose a segmentation algorithm that uses color and texture information in order to extract leukocytes and their respective nucleus and cytoplasm from bone marrow images with heterogeneous staining. As usual, these images show a high cell population, we suggest using conical shapes to separate overlapped blood elements. In the second phase we perform feature extraction to the regions segmented and use these attributes to classify the cells into leukemia subtypes. In our experiments we achieved an average accuracy of 95% in the evaluation of the segmentation process. An overall accuracy of 92% was reached in the supervised classification of acute leukemia types, 84% in lymphoblastic subtypes, and 92% in myeloblastic subtypes.

Keywords: cell segmentation, cell separation, leukemia classification, image analysis, data mining.

1. Introduction

Leukemia is a cancer that begins in the bone marrow. It is caused by an excessive production of immature leucocytes that replace normal blood cells (leukocytes, red blood cells, and platelets). It causes the body to be exposed to many diseases with no possibility to fight them for lack of defenses.

Without treatment, this cancer is the cause of many deaths. In Mexico, according to statistics reported by (NIEGI 2006), leukemia was the fifth and sixth cause of death in men (7%) and women (5.8%) with cancer, and it was the first cause of death in children with cancer between 1-4 and 5-14 years old, with 48.5% and 52.2% of deceases, respectively.

Leukemia is curable if it is detected and treated in at early stage. Its detection starts with a complete blood count. If there are abnormalities in this count, a study of morphological bone marrow smear analysis is done to confirm the presence of leukemic cells. In this study, a pathologist observes some cells under a light microscopy looking for abnormalities presented in the nucleus or cytoplasm of the cells in order to classify the abnormal cells in their particular types and subtypes of leukemia. This classification is very important as it determines which treatment is given. This study has an error rate between 30% and 40% depending on the pathologist experience and the difficulty to distinguish leukemia types and subtypes (Morales 2006). A flow cytometry test is highly accurate to classify leukemias but it is very expensive and not all the hospitals have the equipment to perform it.

The classification of leukemia types and subtypes facilitate the physicians' work in deciding what treatment is the best for a given cell type (*lymphocytic* or *myelogenous)* and disease progress (*acute or chronic).*

In this work we are especially interested in automatically determining the type and subtype of acute leukemias by analyzing information contained in digital images. There are 3 subtypes of acute lymphocytic leukemia (ALL): L1-L3, and 8 subtypes of acute myelogenous leukemia (AML): M0-M7.

Few solutions have been proposed to the classification of leukemia cells. (Morales 2006) and (Galindo 2008) extracted geometric, statistical, and texture features from whole cells and classified them into types and subtypes of acute leukemia, respectively, obtaining promising results. In this paper we propose a leukemia cells classification method by means of a more detailed description of the cells. This requires processing the image pixels to segment the cells and separate their respective nucleus and cytoplasm in order to extract from them features that help to improve the classification among leukemia subtypes.

Segmenting the nucleus and cytoplasm of leukocytes from bone marrow images is a very difficult task, as the images show heterogeneous staining and high-cell population. Some segmentation techniques such as thresholding, edge detection, pixel clustering, and growing regions have been combined to extract the nucleus and cytoplasm of leukocytes (Kim et al. 2001, Won et al. 2004, Theera-Umpon 2005, Colantonio et al. 2007, and Dorini et. al 2007). These techniques could be applied as the images showed uniform backgrounds and high contrast that appropriately defined the objects of interest. Conversely, in our work there are images with low contrast among cell elements and a variety of colors and textures that make cellular elements difficult to distinguish. For this reason, we propose a segmentation algorithm based on color and texture pixels features that can work in bone marrow images showing heterogeneous staining.

Regarding the problem of overlapped blood cells few algorithms have been proposed. These algorithms split cells either by joining concave points using separating lines (Won et al. 2004 and Wang 2007) or by eroding and growing regions retaining the shape (Hengena et al. 2002 and Dorini et al. 2007). In this paper we also propose a cell separation algorithm that keeps the original shape of the blood cell and uses information of this shape to split the overlapped regions by drawing a conical curve.

This paper is organized as follows. Section 2 describes the proposed method for segmenting and separating leukocytes, and identifying their respective nucleus and cytoplasm. Section 3 lists the descriptive features of the cells that are used to identify acute leukemia types and subtypes. Section 4 shows the segmentation and classification results. Finally, conclusions are presented in Section 5.

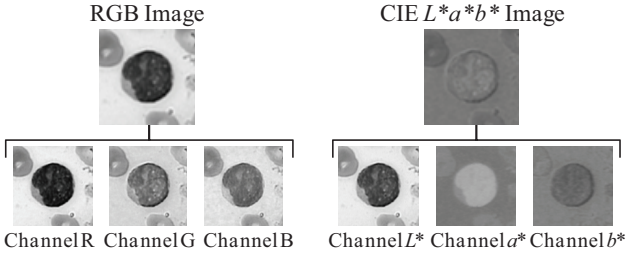
2. Cell Segmentation Model

In this section we propose a segmentation model in order to separate leukocytes into their nucleus and cytoplasm. This model includes three main steps: 1) segmentation of cellular elements, 2) identification of nucleus and cytoplasm, and 3) separation of overlapped blood cells. Good results in this process are very important as we will extract descriptive features to these elements, analyze them, and obtain useful information to classify acute leukemias.

2.1. Segmentation of cellular elements

With the help of an expert, we analyzed the features of samples of bone marrow cell images with different staining in order to design an algorithm that can segment appropriately leukocytes and their respective nucleus. In this analysis we observed some color and texture features that we used to distinguish between blood cells such as: 1) red blood cells get shades of orange and rose, and leukocytes exhibit tonalities of purple in their nucleus, and blue and rose tones in the cytoplasm of lymphocytes and myelocytes, respectively, 2) the color intensities acquired by the nucleus are darker than the cytoplasm, and 3) the texture of the nucleus and cytoplasm of leukocytes, red blood cells, and background are different among them.

2.1.1. Color features. According to these features, we decided to use the CIE *L*a*b** color space because it highlights the visual differences among colors, and it provides accuracy and a perceptual approach in the color difference calculation. Since our image collection is in the RGB color space, a transformation from RGB to the CIE *L*a*b** space was done using the formulas presented in (Paschos 2001). Figures 1(a) and 1(b) show the representation of a bone marrow cell image in the RGB and CIE *L*a*b** color space, respectively.



(a) RGB blood smear (b) CIE *L*a*b** blood smear

Figure 1. *Blood smears in the RGB and CIE L*a*b* color spaces.*

As we can see in figure 1(b), the CIE *L*a*b** color model provides an adequate nucleus color representation by means of its luminosity channel *L**, because it allows identifying objects with different light reflection. We can also notice that channel b* provides a suitable color representation of cells, because it highlights elements with purple and blue tonalities. Since channels *L** and *b** contain valuable information about the nucleus and the cells respectively, they were used to create two similar groups in their channel intensity values. The first group was established by applying the k-means clustering algorithm with k=2 and k=3, and choosing (from the 5 generated clusters) the cluster that better represented the nucleus or cell information. All the pixels that do not correspond to the first group fall into the second one. We finally calculated the mean and standard deviation of the channel intensity values for each group in order to incorporate them as color features in the segmentation model.

2.1.2. Texture model. We use the Wold decomposition model (Francos 1993), as a texture analysis, to separate texture into its structural and stochastic components. We chose this model because blood cells images present heterogeneous textures, hence both periodical and random textures can be found in such images. Additional motivations for choosing this model were its similarity relation with the human visual perception system, and its invariant properties to translation, rotation, and scale.

The Wold decomposition model interprets the image texture by means of the sum of three mutually orthogonal components: a harmonic field, a generalized evanescent

field, and a stochastic field (Francos 1993). The perceptual characteristics of these fields can be described as: periodicity, directionality, and randomness, respectively, according to the three most important human perception dimensions identified by (Rao 1993). Figure 2 shows a diagram of the orthogonal components of the Wold decomposition model for a selected channel.

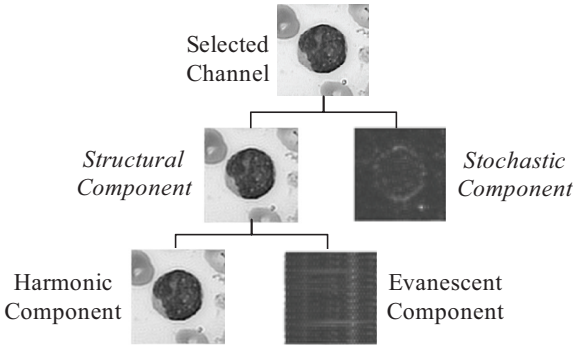


Figure 2. *Wold decomposition texture model.*

In order to parameterize the harmonic field we used the method proposed by (Francos 1993). We first solved the sinusoidals using the discrete Fourier transform (DFT); next we located harmonic peaks by identifying the largest isolated peaks in the harmonic frequencies. We established a value of 10 as the amplitude threshold (experimentally), which is sufficient to find the peaks that were considered harmonic components in the cell images. Finally, the harmonic field parameterization was done by evaluating the amplitude and phase values of the DFT from the frequencies identified as peaks.

In the parameterization of the generalized evanescent field we used the algorithm proposed by (Liu and Picard 1999). In this algorithm, the DFT without harmonic components is used to find four evanescent lines using a Hough transform. The parameterization is carried out by evaluating the amplitude and phase values of the DFT from the frequencies of the evanescent lines identified.

The texture structural component is the sum of the harmonic and generalized evanescent fields. The stochastic component parameterization is done by evaluating the amplitude and phase values of the residual DFT once the structural component is removed.

2.1.3. Color and texture segmentation. In this work we propose a segmentation method that uses contextual color and texture information to classify pixels corresponding to cell elements in bone marrow images with heterogeneous staining. For this, we first represent a channel using a binary Markov Random Field model which consists of a label field and three observation fields (channel intensity, structural texture, and stochastic texture fields). After that, according to the Bayes theorem, we represented the segmentation problem as a Maximum a Posteriori (MAP) estimation of the label field. Theoretical details can be consulted in (Li 2000).

Based on the Iterated Conditional Model algorithm (Kato 1994), we solved the MAP estimation using the procedure shown in table 1.

Table 1. *MAP estimation for cell segmentation.*

1. Initialize the label field f^0 by using eq. 1 and set $i = 0$.

$$f_s^0 = \arg\min_{k=1,k=2}\left(\ln(\sqrt{2\pi}\sigma_k) + \frac{(x_s-\mu_k)^2}{2\sigma^2 k} \right) \quad (1)$$

where x_s is the intensity value of the selected channel, μ_k and σ_k are the mean and standard deviation of the group $k = \{1,2\}$ created in section 2.1.1.

2. For each labeling configuration of f_s that differs in at least one neighbor configuration $f_{N(s)}$, calculate the energy U_s for each class $k = \{1,2\}$ using eq. 2.

$$U_s(x_s \mid f_s = k) = \ln(\sqrt{2\pi}\sigma_k) + \frac{(x_s-\mu_k)^2}{2\sigma^2 k} + w_s + \sum_{r\in f_{N(s)}} v_{sr} \quad (2)$$

Where w_s and $v_{sr} = |v_s - v_r|$ are the spatial variation parameters of the stochastic W and structural V texture components, respectively, defined over second order cliques.

3. Update the labeling configuration f_s^i by selecting the configurations with minimum energy of $f_{N(s)}$ using eq. 3:

$$f_s^{i+1} = \arg\min_{k=\{1,2\}} U_s(k) \quad (3)$$

4. Go to step 2 with $i = i+1$, until $i = 10$ or $f^i = f^{i+1}$.

2.2. Identification of nucleus and cytoplasm

The goal of this step is identifying leukocytes through the recognition of their nucleus and cytoplasm. From the regions obtained in the segmentation process, we analyzed their shape, color, and spatial relation with respect to other regions to determine whether an analyzed region is a nucleus or a leukocyte.

The features that were used to recognize cellular elements are: *circularity* to measure the perimeter complexity of a circular object (circularity= perimeter^2/(4π·area)), *eccentricity* to find out how much the object deviates from being circular (eccentricity= dist(center, focus)), *color* to determine if a region is darker than other, and *containment proportion* to establish whether a region contains or is contained by another region. Using these features and a priori knowledge about the cellular elements structure we designed the rule-based classifiers presented in tables 2 and 3 to label nucleus and cells. We selected a subset of 20 regions with regular shapes (nucleus and cells) and 20 regions with irregular shapes (overlapped regions) and we generated classification rules (using Weka) that discriminate between these types of forms. These rules gave us an idea of the threshold values to use to determine if a region is likely to be a cellular element. Then, we added the color relation and containment proportion rules so that we could match the cells with their respective nucleus. We established a containment proportion threshold of 95% because some pixels are missed in the cell segmentation or in the separation of overlapped regions.

Table 2. *Decision rules to identify nucleus.*

if *color*(region$_1$) < *color*(region$_2$) and *containment proportion* (region$_2$,region$_1$) ≥ 95%
  then if *eccentricity*(region$_1$) ≤ 0.5 and *circularity*(region$_1$) ≤ 1.5
    then region$_1$ is a nucleus.
    else region$_1$ likely is an overlapped nucleus.
  else if *eccentricity*(region$_1$) > 0.5 and *circularity*(region$_1$) > 1.5
    then region$_1$ is not a region of interest.
    if *eccentricity*(region$_2$) > 0.5 and *circularity*(region$_2$) > 1.5
      then region$_2$ is not a region of interest.

Table 3. *Decision rules to identify cells.*

if *color*(region$_1$) >*color*(region$_2$) and *containment proportion* (region$_1$,region$_2$) ≥ 95%
  then if *eccentricity*(region$_1$) ≤ 0.5 and *circularity*(region$_1$) ≤ 1.5
    then region$_1$ is a cell.
    else region$_1$ likely is an overlapped cell.
  else if *eccentricity*(region$_1$) > 0.5 and *circularity*(region$_1$) > 1.5
    then region$_1$ is not a region of interest.
    if *eccentricity*(region$_2$) > 0.5 and *circularity*(region$_2$) > 1.5
      then region$_2$ is not a region of interest.

2.3. Separation of overlapped blood cells

Figure 3 summarizes the process of cell separation once the overlapped region is identified. In order to split the overlapped regions we obtain the edges of the region and its centroid and we provide some concave points as points of separation. Then we transform edges from a cartesian to a polar space, and we interpolate discontinuous points using a linear interpolation. This allows completing cell borders with a conical shape once we come back to the cartesian space. Finally, we join some edges discontinuities by applying morphological operations.

Figure 3. *Cell separation procedure.*

3. Classification of Acute Leukemia Cells

The suitable recognition of leukemia cells requires the definition of good descriptive features that facilitate their classification. In this phase we extract geometric, statistical, texture, and size ratio features from regions obtained in the segmentation process (nucleus, cytoplasm, and whole cell) and we analyze these features to identify types and subtypes of acute leukemia. It is important to mention that we do not normalize the images to extract these features, since the size and color of the cells are important characteristics to distinguish among subtypes of leukemia. Table 4 shows these features. All the geometric features mentioned in table 4 were extracted from each nucleus and cell. Since many of the cytoplasm features are included in the cell, we only calculated its area. Due to the morphological analysis performed by the expert, these features are not relevant to the classification of cells. In case of the nucleus, cytoplasm, and cells, we extracted statistical and texture features from the channels of the RGB image and the gray scale image. We obtained another set of features applying principal components analysis to the images corresponding to the nucleus and the cell. We used as features the firsts 10 eigen values for the channels of the RGB image and gray image, which represent at least the 80% of the data variability in each region. The analysis of these features to classify acute leukemia cells was done using different training and testing sets, attributes selection, and classification algorithms available in (Weka 2009). These tests are described in section 4.3.

Table 4. *Representative features for the cell description.*

Geometrical features
area, perimeter, circularity, weight, height, elongation, major axis length, minor axis length, eccentricity, extension, equivalent diameter, Euler number, convex area, and solidity.
Color features
sum, mode, mean, standard deviation, and variance.
Texture features
homogeneity, contrast, correlation, energy, and entropy.
Size ratio
$\frac{\text{area}_{\text{nucleus}}}{\text{area}_{\text{cytoplasm}}}$, $\frac{\text{area}_{\text{nucleus}}}{\text{area}_{\text{cell}}}$, and $\frac{\text{perimeter}_{\text{nucleus}}}{\text{perimeter}_{\text{cell}}}$.

4.1. Data sets description

In this work we used the MSSI cells image collection that contains 633 bone marrow leukemia cells images with different color staining. These images were digitalized by (Morales et al. 2005) using a digital camera connected to a Carl Zeiss optical microscope with a 100x objective. Hence, all images have the same resolution. Table 5 shows the number of examples of each type and subtype of acute leukemia included in the collection.

Table 5. *Samples of acute leukemia types and subtypes.*

Type / Subtype	**ALL**	L1	L2	**AML**	M2	M3	M5
Number of images	**295**	102	135	**338**	95	47	56

4.2. Cell Segmentation Results

The proposed method showed good qualitative segmentation results allowing the extraction of the 633 leukocytes and their respective nucleus and cytoplasm. In order to measure the accuracy of the segmentation algorithm in a quantitative way, we tested it on a subset of the original images collection, which contains 20 leukemia cells images with a size of 256 x 256 pixels. It is important to notice that this test set includes different samples of images with color variations in their staining; furthermore, there are leukocytes overlapping with other blood cells.

Table 6 shows the results of the evaluation of the algorithm. This evaluation was obtained by comparing our automatic cells segmentation algorithm with cells manually segmented by an expert. The metrics used for evaluating

the segmentation were: Precision = TP/P, FP Rate = FP/P, and FN Rate = FN/N, where TP and FP are the number of pixels correctly and incorrectly classified as cellular elements respectively, P is the number of pixels classified as cellular elements, FN is the number of pixels incorrectly classified as background, and N is the number of pixels classified as background.

Table 6. *Evaluation of the cells segmentation algorithm.*

	Precision	FP Rate	FN Rate
nucleus	95.87%	4.13%	2.33%
cell	95.75%	3.16%	3.83%

Experimental results show that the proposed methodology allows the extraction of the leukocytes and their respective nucleus from cells with good accuracy when experimenting with real bone marrow leukemia cells images. Figure 4 shows some examples with the results of the cell segmentation and identification.

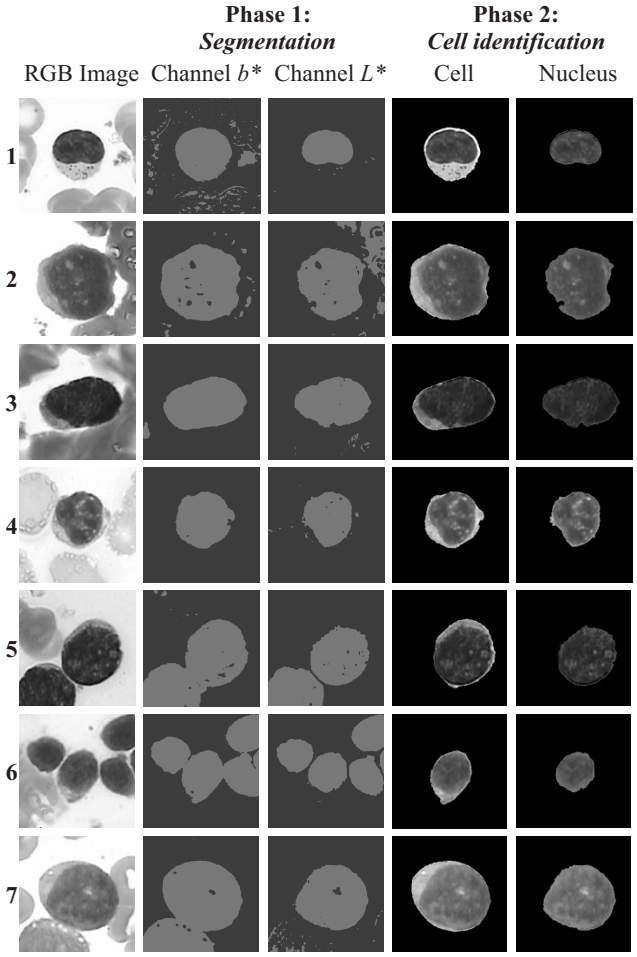


Figure 4. *Examples of the results of the segmentation and identification of cells in images with different staining and cell population.*

As we can see in figure 4, the leukemia cell images that we want to classify show different colors and textures, and leukocytes are overlapped with other blood cells. The images in previous cell segmentation algorithms do not consider these conditions, so these algorithms were not able to work in images with variation in color staining and high-cell population. This is the reason that we didn't compare our results with previous works.

4.3. Cell Classification Results

Experiments were performed by analyzing the nucleus, cell, and cytoplasm features as we propose in this work, and also by using only features of whole cells as proposed in (Galindo 2008).

In each experiment we established different classification tasks to distinguish among types of acute leukemia (ALL and AML), subtypes of ALL (L1 and L2), and subtypes of AML (M2, M3, and M5). In the case of AML subtypes we analyzed the behavior of a subtype with respect to the others performing binary and multiclass classifications. For each classification problem, we performed experiments using different types of features such as geometric, statistical, texture, size ratio, and eigen values. The classification was carried out using instance based classifiers, decision trees, regression functions as well as metaclassifiers available in (Weka 2009). Some of these are: k-Nearest Neighbor (IBk), Random Forest (RF), Simple Logistic (SL), SMO, and Random Committee (RC), which were chosen because they were able to obtain the best results for leukemia subtypes classification in the work of (Galindo 2008).

The evaluation of the classification models was done using 10 cross-validations. The criteria for evaluating classifiers were: overall percentage of correct classifications, true positive rate (TPR), true negative rate (TNR), and area under the ROC curve. Table 7 shows the best classifier for each experiment. These results were achieved using geometric, color, texture, and size ratio features as attributes. The eigen values did not represent useful information to distinguish among leukemia subtypes.

Table 7. *Results of the best classifier for each type or subtype of acute leukemia cells.*

Classification Problem	Clas-sifier	Fea-tures	Prec. %	TPR	TNR	AUC
ALLvsAML	SMO	N&C	92.20	0.920	0.924	0.921
ALLvsAML	SL	C	81.32	0.822	0.803	0.899
L1vsL2	IBk	N&C	84.40	0.835	0.853	0.907
L1vsL2	SL	C	76.78	0.667	0.845	0.814
M2vs(M3&M5)	RC.RF	N&C	92.45	0.883	0.962	0.959
M2vs(M3&M5)	RC.RF	C	74.24	0.706	0.778	0.805
M3vs(M2&M5)	IBk	N&C	91.89	0.805	0.955	0.880
M3vs(M2&M5)	RC.RF	C	80.79	0.390	0.940	0.788
M5vs(M2&M3)	IBk	N&C	91.89	0.870	0.938	0.955
M5vs(M2&M3)	RF	C	84.37	0.730	0.890	0.866
M2vsM3vsM5	RC.RF	N&C	88.39	0.904	0.894	0.945
M2vsM3vsM5	RC.RF	C	66.63	0.801	0.612	0.784

Comparing the results obtained in the classification of leukemia cells using nucleus and cytoplasm features with the results obtained using only cell features, it can be clearly seen that the cell is better described by using features from its cellular elements.

5. Conclusions

In this paper we proposed a novel method that uses color and texture information of images pixels in order to segment leukocytes and their respective nucleus and cytoplasm from bone marrow leukemia cells images.

We used the CIE *L*a*b** color space and the 2D-Wold decomposition texture model as they allow us to analyze blood cells images in a similar way to the human visual perception system. We modeled the color and texture information by using a Markov Random Field in order to obtain regions of cell elements. We then extracted some color, shape, and containment proportion information, and we used a cell identification algorithm to recognize cells and their respective nucleus and cytoplasm as well as overlapped regions. When there were cells or nucleus overlapping other cell elements, we separated them by using a cell separation algorithm based on linear interpolation in the polar space to provide a conical shape.

It is important to note that our method can be applied to images that show heterogeneous color and texture staining and high-cell population, a desirable property when working with bone marrow smears. Experimental results show that our method (with this type of images) achieves a segmentation accuracy of 95% (average) when it is compared with a manual segmentation performed by an expert.

With respect to the cell classification process, we can demonstrate that the use of descriptive features of the nucleus and cytoplasm of the cells improved their representation, allowing the classification of acute leukemia types and subtypes to increase significantly its accuracy (from 7% to 22 %) compared with that obtained when we only use descriptive features of the cell.

Our future work involves the design of a decision algorithm that combines different leukemia cell classifiers in order to provide an automatic diagnosis for a patient by analyzing the information of all the available samples of the patient's cells.

6. Acknowledgments

We thank Dr. Jose E. Alonso for providing support in the manual segmentation and labeling of types and subtypes of the leukemia cell images used in this work.

7. References

Colantonio, S., Gurevich, I., and Salvetti, O., 2007. Automatic Fuzzy-Neural based Segmentation of Microscopic Cell Images. LNCS 4826: 115–127.

Dorini, L., Minetto, R., and Leite, N., 2007. White Blood Cell Segmentation using Morphological Operators and Scale-Space Analysis. IEEE, XX Brazilian Symposium on Computer Graphics and Image Processing. 3118: 294-304.

Francos, J., 1993. An Unified Texture Model based on a 2-D Wold-Like Decomposition. IEEE Transactions on Signal Processing,

Galindo, M., 2008. Obtaining Features of Subtypes of Leukemia in Digital Blood Cell Images for their Classification. Master Thesis. INAOE, Mexico.

Hengena, H., Spoor, S., and Pandi, M., 2002. Analysis of Blood and Bone Marrow Smears using Digital Image Processing Techniques. Progress in Biomedical Optics and Imaging 3: 624-635.

Kato, Z., 1994. Multigrid Markovian Modelization in Early Vision with Application to Satellite Image Segmentation. PhD Dissertation. INRIA, France.

Kim, K., Jeon, J., Choi, W., Kim, P., and Ho, W., 2001. Automatic Cell Classification in Human's Peripheral Blood Images based on Morphological Image Processing. Australian Joint Conference on Artificial Intelligence 2256: 225-236.

Li, S. 2000., Modeling Image Analysis Problems using Markov Random Fields. Handbook of Statistics, Elsevier Science 20: 1-43.

Liu, F., and Picard, R., 1999. A Spectral 2D Wold Decomposition Algorithm for Homogeneous Random Fields. IEEE Proceedings of the International Conference on Acoustics, Speech and Signal Processing.

Morales, B., Olmos, I., Gonzalez, J., Altamirano, L., Alonso, J., and Lobato, R., 2005: Bone marrow smears digitalization. Laboratory of Specialties of the Mexican Social Security Institute at Puebla, Mexico

Morales, B., 2006. Feature Extraction in Bone Marrow Cell Images for the Classification of Acute Leukemia. Master Thesis. INAOE, Mexico.

NIEGI - National Institute of Statistics, Geography, and Informatics 2006. Statistics with respect to the World Day against Cancer. http://www.inegi.org.mx.

Paschos, G., 2001. Perceptually uniform color spaces for color texture. IEEE Transactions on Image Processing.

Rao, A., Lohse, G., 1993. Towards a Texture Naming System: Identifying Relevant Dimensions of Texture. IEEE Conference on Vizualization.

Theera-Umpon, N., 2005. White Blood Cell Segmentation and Classification in Microscopic Bone Marrow Images. LNCS 3614: 787–796.

Wang, W., and Song, H., 2007. Cell Cluster Image Segmentation on Form Analysis. IEEE Third International Conference on Natural Computation.

Weka - Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I., 2009. The WEKA Data Mining Software: An Update; SIGKDD Explorations, 11: Issue 1.

Won, C., Nam, J., and Choe, Y., 2004. Segmenting Cell Images: A Deterministic Relaxation Approach. LNCS 3117: 281–291.