

## REVIEW ARTICLE

WILEY


 International Journal of  
Laboratory Hematology

# Optimizing morphology through blood cell image analysis

 A. Merino<sup>1</sup> | L. Puigví<sup>2</sup> | L. Boldú<sup>1</sup> | S. Alférez<sup>2</sup> | J. Rodellar<sup>2</sup>

<sup>1</sup>Biomedical Diagnostic Centre, Hospital Clínic, University of Barcelona, Barcelona, Spain

<sup>2</sup>Department of Mathematics, Barcelona Est Engineering School, Technical University of Catalonia, Barcelona, Spain

**Correspondence**

Anna Merino, Biomedical Diagnostic Centre, Hospital Clínic, University of Barcelona, Barcelona, Spain.  
Email: amerino@clinic.cat

**Funding information**

Directory of Science, Technology and Innovation of the Ministry of Economy and Competitiveness of Spain, Grant/Award Number: DPI2015-64493-R

**Abstract**

**Introduction:** Morphological review of the peripheral blood smear is still a crucial diagnostic aid as it provides relevant information related to the diagnosis and is important for selection of additional techniques. Nevertheless, the distinctive cytological characteristics of the blood cells are subjective and influenced by the reviewer's interpretation and, because of that, translating subjective morphological examination into objective parameters is a challenge.

**Methods:** The use of digital microscopy systems has been extended in the clinical laboratories. As automatic analyzers have some limitations for abnormal or neoplastic cell detection, it is interesting to identify quantitative features through digital image analysis for morphological characteristics of different cells.

**Result:** Three main classes of features are used as follows: geometric, color, and texture. Geometric parameters (nucleus/cytoplasmic ratio, cellular area, nucleus perimeter, cytoplasmic profile, RBC proximity, and others) are familiar to pathologists, as they are related to the visual cell patterns. Different color spaces can be used to investigate the rich amount of information that color may offer to describe abnormal lymphoid or blast cells. Texture is related to spatial patterns of color or intensities, which can be visually detected and quantitatively represented using statistical tools.

**Conclusion:** This study reviews current and new quantitative features, which can contribute to optimize morphology through blood cell digital image processing techniques.

**KEYWORDS**

blood, leukemia, lymphoma, morphology

## 1 | INTRODUCTION

Although immunological, cytogenetic, and molecular tests are being increasingly used, the morphologic analysis of blood cells is still a crucial diagnostic aid. The information provided by the morphological analysis of the blood cells is relevant for the selection of additional techniques and follow-up of the patients with malignant blood diseases, including lymphoid neoplasms and leukemias.<sup>1,2</sup> In the WHO classification system, tumor cell morphology, along with the immunophenotype and genetic changes, remains essential in defining disease entities.<sup>3</sup>

In spite of the significant diagnostic role of peripheral blood (PB) morphology, there are no objective values for defining cytological variables. The distinctive cytological characteristics of the blood cells are subjective and influenced by the reviewer's interpretation.<sup>1</sup>

In this respect, digital microscopy and computerized image analysis may provide more accurate and objective morphological evaluation, transforming cytological qualitative parameters into quantitative values. Image analysis is the extraction of meaningful information from digital images through automated or computerized analysis.

An early study about image analysis of small cell lymphoma of the thyroid gland, and the comparison of nuclear parameters of lymphocytes, demonstrated that the nuclear area was the optimum descriptor discriminating between small neoplastic lymphocytes in thyroidal lymphomas and reactive lymphocytes in Hashimoto's thyroiditis.<sup>4</sup>

Image analysis was first developed using histological sections, and the measurements were performed over tissue parameters such as length, area, and cell count, being useful to supply information about the trabecular organization in the marrow space.<sup>5</sup> As blood cells are

seen as individual entities in the smear, their cytological characteristics may be the targets in the definition of features to provide objective numerical scales and quantify them using image analysis.

Traditionally, the morphologic analysis of the PB smear has been performed using the manual microscope method. The close collaboration between cytologists, mathematicians, and engineers over the last few years has eased the development of automatic tools for digital image processing of normal blood cells. Some equipment is capable to preclassify cells in different categories using neural networks, extracting a large number of measurements and parameters, which describe relevant cell morphological characteristics.<sup>6</sup> The use of digital microscopy systems in the clinical laboratories has been extended, and automated analyzers using digital images of leukocytes, erythrocytes, and thrombocytes have been integrated into the daily work.<sup>7,8</sup> Nevertheless, these automatic analyzers have some limitations related to abnormal or neoplastic cell detection.<sup>9</sup> New hematological analyzers are currently under development based on automatic morphology computer imaging of PB smears.<sup>10</sup> The topic of morphological analysis has received much attention with the increasing demands in both bioinformatics and biomedical applications. In the last 20 years, about 1,000 publications have reported the use of morphological cell analysis in biomedical research.<sup>11</sup>

The goal of this study was to review how image analysis helps to define quantitative features able to characterize cytological variables, which are currently used by the pathologist. This review focuses on abnormal lymphoid cells and blasts, as these malignant cells exhibit subtle differences in morphologic characteristics, which makes their visual identification a difficult task. The paper is organized as follows. First, the paper discusses the most important qualitative morphological characteristics to differentiate a variety of abnormal lymphocytes and blast cells. Second, different types of image-based quantitative features are reviewed in 2 directions: their definitions and the current literature discussion. Third, the practical use of quantitative features is evaluated on the automatic recognition of reactive or malignant cell recognition in PB. Finally, some future perspectives are outlined toward objective, morphology-based PB cell assessment through image analysis.

## 2 | QUALITATIVE MORPHOLOGICAL ISSUES IN LYMPHOID AND BLAST CELLS

Morphological evaluation of leukemia and lymphoma cells is essential for their diagnosis and classification. PB film morphology is important to distinguish between reactive lymphocytes, abnormal lymphoid cells, and lymphoid or myeloid blasts in the first morphological evaluation when the patient is explored by the first time in the hospital.

External quality control programs have highlighted difficulties to differentiate between leukemic and non-neoplastic cells, the abnormal lymphoid cells being the most difficult to identify by the participants.<sup>12,13</sup> With respect to lymphoid neoplasms, few studies have been published regarding feature analysis in abnormal lymphoid cells. Chronic lymphocytic leukemia (CLL) cells are typically small

lymphocytes with clumped chromatin and scant cytoplasm. Hairy cells (HCL) are larger than normal lymphocytes with an oval or indented nucleus containing homogeneous or spongy chromatin (less clumped than that of a normal lymphocyte), and they have hyaline cytoplasm with circumferential hairy projections. Splenic marginal zone lymphoma (SMZL) cells usually show short polar villi. B-lineage prolymphocytic leukemia (B-PLL) cells are larger than CLL, show more variation in size and the nuclear chromatin is moderately condensed exhibiting large nucleoli. Follicular lymphoma (FL) cells show notches or deep narrow clefts in some nuclei. In mantle cell lymphoma (MCL), cells are more pleomorphic, have more dispersed chromatin with inconspicuous nucleoli and in the blastoid variant, most cells resemble blast cells. In plasma cell leukemia, the neoplastic cells may resemble mature plasma cells or they may be highly abnormal showing an immature chromatin pattern and nucleoli. T-lineage prolymphocytes are smaller and more pleomorphic than B-lineage prolymphocytes and nuclei are irregular, showing a little apparent nucleolus. Sézary cells may be either small or large, and the chromatin is characteristic due to its convoluted or cerebriform appearance.

With respect to blast cell morphology, one of the major roles of the blood smear analysis is the discrimination between myeloid and lymphoid origin, especially for acute promyelocytic leukemia, in which a blood film is very helpful in the rapid diagnosis and treatment.<sup>14</sup> The majority of lymphoblasts lack granules and have a high nucleus-cytoplasmic ratio, a diffuse chromatin pattern and usually prominent nucleoli. Myeloid blast cells may present a few azurophilic granules and a lower nucleus-cytoplasmic ratio.

## 3 | QUANTITATIVE MORPHOLOGICAL FEATURES BASED ON IMAGE ANALYSIS

The main steps in image analysis are the following: (i) image acquisition and preprocessing, (ii) segmentation,<sup>15-19</sup> and (iii) extraction of features (quantitative descriptors),<sup>18,20,21</sup> which make possible the further classification of the different cell types.

The segmentation step aims to obtain the cells separated from the other objects in the image and divided into regions of interest (ROI). This step is critical to be successful in the feature extraction procedure.<sup>15</sup> In PB cells, 3 main ROI are typically obtained as follows: nucleus, whole cell, and peripheral zone around the cell.<sup>17-19</sup> A fourth ROI is obtained for the cytoplasm by the difference between the sets of pixels that belong to the whole cell and the nucleus region.

From the segmented ROI, the next step is to identify a set of quantitative descriptors, which are usually grouped in the following 3 categories: (i) geometric, (ii) color, and (iii) texture.

### 3.1 | Categories of quantitative features

#### 3.1.1 | Geometric features

In general, geometric features are easy to interpret because they are the most intuitive and closely related to the visual patterns that

are familiar to pathologists. Most common geometric descriptors include parameters such as area, perimeter, circularity, diameter, eccentricity, elongation, roundness, convexity, nuclear size, cell size, nucleus-cytoplasmic ratio, nuclear eccentricity, and others.<sup>1,18,20,21</sup> The *nucleus eccentricity* is calculated as the distance between the cell center and the nucleus center.<sup>20</sup> New geometry-based features may be proposed to quantify more complex characteristics. For instance, a novel descriptor was proposed in<sup>17</sup> to describe the *cytoplasmic profile* or *hairiness* in villous lymphocytes.

Another example of a new geometric feature is the so-called *RBC proximity*, which is an indirect measure of the amount of cell cytoplasm that adheres to red blood cells (RBC). It is well known by pathologists that cytoplasm of reactive lymphocytes (RL) tends to adhere to neighboring (RBC). When a RL touches a RBC, there is a coincidence between the RL cytoplasm perimeter and the edge of the peripheral zone around the RL. Based on this observation, the *RBC proximity* is defined as the proportion of the RL perimeter (in pixels) that intersects the peripheral zone around the cell. The practical use of this feature is described for the first time in this study when it will be used for automatic recognition purposes (see Figure 4).

Geometric features have been previously used for quantitative characterization of lymphocytes in thyroid tissues,<sup>4</sup> such as *nuclear diameter*, *area*, or *irregularity*, useful for the discrimination between abnormal and reactive lymphocytes. In addition, the effect of albumin on CLL lymphoid cell morphology on PB was analyzed using image analysis.<sup>22</sup> It was demonstrated that the addition of albumin to the blood smears resulted into changes in cytological features, decreasing not only cell, but also nuclear areas, while increasing the *nucleus-cytoplasm ratio*. In a later publication, image analysis was used to obtain geometric features to quantify some cytological parameters of the malignant lymphocytes in CLL, MCL, and B-PLL.<sup>1</sup> In that work, some quantitative descriptors, such as *cellular area* and *diameter*, *nuclear area*, and *density*, were found to morphologically discriminate among these 3 groups of abnormal lymphoid cells.

### 3.1.2 | Color features

A digital PB cell image is composed of a finite number of pixels with particular location and color values. Color is a physical property very common in the visual characterization of blood cells. To explore the rich amount of color information in the malignant blood cells, several color spaces are used to obtain quantitative features. A color space is a way of describing a color by a number (3 or 4) of components.<sup>15</sup> In the RGB space, each color is decomposed into 3 bands corresponding to the basic colors: red (R), green (G), and blue (B). CMYK is another space in which the bands are cyan (C), magenta (M), yellow (Y), and black (B). HSV space is defined by hue (H), saturation (S), and brightness (V). Laboratory or Luv spaces use lightness (L) and chromaticity (a, b, u, v). Laboratory color space is characterized by its approximation to human perception.

From each of the color components belonging to a color space, one grayscale image results, which has intensities varying from black (0) to white (255). To obtain quantitative features from these

grayscale images, it is very useful to represent the histogram. The histogram is a 2-axis plot in which the horizontal one is divided into intervals representing intensity values from 0 to 255. The vertical axis gives the number of pixels corresponding to the different intensity intervals. From each histogram, 6 classical *first order statistical features* are usually calculated as follows: *mean*, *standard deviation* (SD), *skewness*, *kurtosis*, *energy* (uniformity), and *entropy* (variability).<sup>19-23</sup>

As an example of statistical features calculated from the histogram, Figure 1 (see<sup>23</sup>) shows 2 original images corresponding to Sézary cells acquired by the CellaVision DM96 (DM96). The 2-axis plot shows the histograms obtained from the magenta component in the grayscale images, in which the pixel count levels were obtained using the scientific software MATLAB (produced by MathWorks, MA, USA). Comparing the blue histogram from image 1 with the red one from image 2, it is possible to note that the first image shows many pixels with different intensity levels, while the second one has the pixels mainly in a short intensity range. This fact is interpreted as image 1 displays more variability and contrast in the nucleus with respect to the cytoplasm, while image 2 shows bright dots practically in the whole cell.

*Entropy* describes the variability and *energy* describes the uniformity of the pixel intensity levels. Accordingly, the histogram of image 1 (in blue) shows more variability of the pixel intensity levels and higher information variety, so that the *entropy* of the image is higher. In contrast, the low uniformity of the pixel intensity levels is translated into lower *energy*. Conversely, the histogram corresponding to image 2 (in red) shows less dispersion and high uniformity in the pixel intensity values, which means low *entropy* and high *energy* in the image.

Color features are involved in the quantification of cytoplasmic basophilia and cytoplasmic granulation, both very important morphological characteristics in the classification of lymphoid cells.<sup>24</sup>

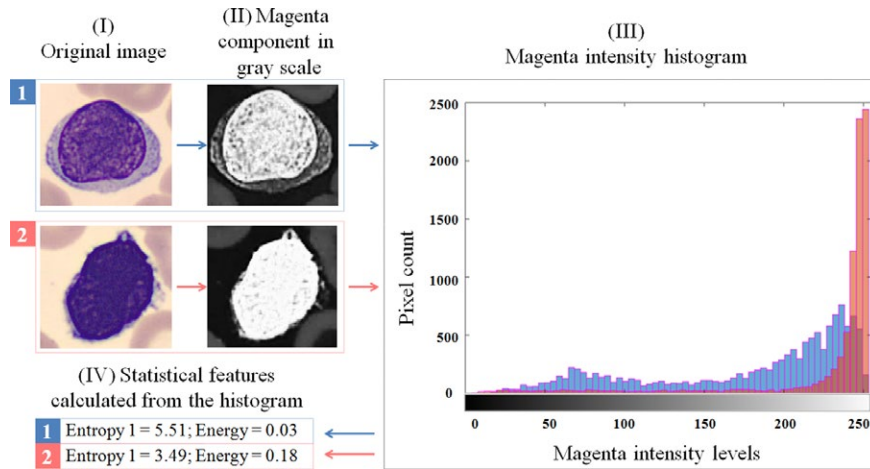
### 3.1.3 | Texture features

In contrast to other morphological parameters, such as *cell size* or *nucleus-cytoplasmic ratio*, the interpretation of the nuclear texture by visual observation is sometimes difficult. The chromatin density is related to the fixation and staining, as both procedures have a high influence on the variations in this parameter. Chromatin distribution reflects DNA nucleus organization containing important cellular diagnostic and prognostic information.

Texture is related to spatial patterns of color or intensities, which can be visually detected. In digital image analysis, texture is quantitatively defined by uniformity, density, pixel tone, and their spatial relationships, among others. Such analysis is not an easy task and is usually performed following 2 main approaches: (i) the so-called *gray level co-occurrence matrix* (GLCM) and (ii) the *granulometry*.

#### Gray level co-occurrence matrix

For a grayscale digital image, the GLCM is defined as the probability that pairs of neighboring pixels have similar intensities.<sup>25</sup> Based on



**FIGURE 1** Two original images corresponding to Sézary cells acquired with the DM96. The 2-axis plot shows the histograms obtained from the magenta component in the grayscale images, in which the pixel count in the magenta intensity levels corresponding to images 1 and 2 was obtained using MATLAB. Comparing the blue histogram from image 1 with the red one from image 2, it is possible to note that the first shows many pixels with different intensity levels, while the second has the pixels mainly in a short intensity range. In fact, image 1 displays more variability and contrast in the nucleus with respect to the cytoplasm, while image 2 shows bright dots practically in the whole cell. More variability of the pixel intensity levels implies higher information variety, so that the *entropy* of the image is higher. In contrast, the low uniformity of the pixel intensity levels is translated into lower *energy*. Conversely, the histogram corresponding to image 2 (in red) shows less dispersion and high uniformity in the pixel intensity values, which means low *entropy* and high *energy* of the image

the GLCM of a digital image, many measurements (second order statistical features) can be calculated providing more information about the texture of the ROI, such as correlation, homogeneity, maximum probability, and others.<sup>26,27</sup> By the way of example, Figure 2 shows 2 original images corresponding to a lymphoblast (left) and a myeloblast (right) acquired with the DM96, as well as their corresponding magenta components as grayscale images. The intensity levels in these images rank from 1 (black) to 8 (white) and, therefore, the GLCM are matrices with 8 rows and 8 columns, which are shown in the figure. As an example, Figure 2 illustrates the meaning and value of the *maximum probability* feature. First, the most repeated pair of pixel intensities is identified. Then, the ratio of the repetitions over the total possible pairs is defined as the *maximum probability* feature. For the lymphoblast case, the pair (8, 8), which is the brightest, has 8,178 repetitions, which means a ratio of .8682. This is in accordance with the visual observation, as the grayscale lymphoblast magenta component appears practically white. On the other hand, for the myeloblast, again (8, 8) is the most probable, but the feature value is .1332, much lower than in the previous case. This is because there are large numbers of pairs of pixels with different gray intensities as numerically shown in the GLCM and visually observed in the grayscale image.

As its introduction by Haralick, the GLCM has been widely used as a texture measurement in medical imaging, such as ultrasound images for solid neoplasms<sup>28</sup> and in bone marrow images to distinguish 4 types of erythrocyte precursor cells stages.<sup>29</sup> Nevertheless, few studies involving GLCM have been performed using PB images and mainly for differentiating among normal leukocytes<sup>30</sup> and blast lymphoid cells.<sup>31</sup> In other paper, the authors calculated 5 textural attributes based on the GLCM (*energy*,

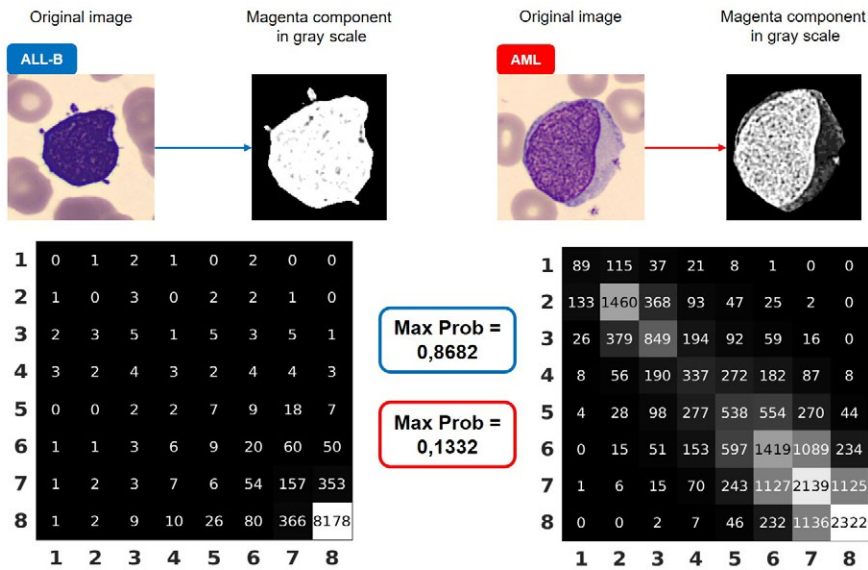
*entropy*, among others) to differentiate between normal leukocytes (5 subtypes) and CLL cells.<sup>32</sup>

### Granulometry

Granulometry is a texture feature that measures the particle size distribution in an image by mathematical morphology operations, such as dilation, erosion, opening, and closing.<sup>26</sup> These authors combined *color granulometries* and *color histograms* to achieve leukocyte classification obtaining *granulometric curves*. The morphology operations are visualized by means of the so-called *granulometric and pseudogranulometric curves*, which represent the proportion of the black and bright particles in the image. Over both curves, it is possible to obtain texture features by calculating parameters such as *mean*, *standard deviation*, *skewness* (measure of asymmetry), and *kurtosis* (measure of uniformity). Reference<sup>23</sup> presented a study in which the granulometry was applied for the morphological differentiation of abnormal lymphoid cells.

A scoring system was proposed<sup>33</sup> to distinguish B-cell disorders based on the following quantitative features: nuclear shape, cellular area and shape, nucleus-cytoplasmic ratio, nuclear red/blue ratio, cytoplasmic green/blue ratio, and nucleolus detection. The study included 87 PB smears of B- and T-cell disorders and 6 healthy donors. For each PB smear, 30 images were analyzed demonstrating the usefulness of the scoring system, but it needed to be confirmed with a larger sampling.

Automatic recognition of 4 abnormal lymphocyte types (HCL, CLL, MCL, B-PLL) using 113 features (geometric and color/texture) was described in.<sup>17</sup> The same authors presented the most relevant features for the automatic classification of normal, HCL, CLL, MCL, and B-PLL.<sup>18</sup> In a further work, FL and RL were added to the previous



**FIGURE 2** The original images correspond to 2 blast cells from lymphoid and myeloid lineage, respectively, and they were acquired with the DM96. In their corresponding magenta component in gray scale, it can be observed that lymphoid blast displays mostly magenta component, which is represented by bright spots in the grayscale component. The intensity levels in the image rank from 1 (black) to 8 (white). The most probable pair of pixels for lymphoid blast is the brightest, where all intensity values are grouped. It resulted in a higher average of the *maximum probability*. In contrast, myeloid blast shows a large number of different pairs of gray pixels' intensities, resulting in a lower maximum value and a lower value for this descriptor

study.<sup>19</sup> In addition, quantitative features were extracted by image analysis to discriminate RL from blast cells and between myeloblasts and lymphoblasts.<sup>34</sup> A number of 20 quantitative features, which were relevant for the discrimination among 12 lymphoid cell types, were also analyzed.<sup>23</sup>

#### 4 | REACTIVE AND MALIGNANT CELL RECOGNITION USING QUANTITATIVE MORPHOLOGY

Quantitative features facilitate the last step of the image analysis process, which is the automatic cell classification. The reader may find more detailed discussions about image processing and machine learning tools in the morphological analysis of blood cells.<sup>35</sup>

This section describes a classification experiment, whose purpose is to assess the usefulness of the quantitative morphological information provided by a number of 140 features (describing size, shape, color and texture from cytoplasm, nucleus and complete cell). The main goal is to check if they are useful for the automatic discrimination between reactive lymphocytes (non-neoplastic) and neoplastic cells circulating in PB. Figure 3 shows the 20 most relevant features in the classification. A machine learning classifier described in<sup>19</sup> was trained with the 140 features for the automatic recognition of RL, blast cells (BL), and abnormal lymphoid cell images, captured by the DM96.

For the classification experiment, 418 PB smears stained with May Grünwald-Giemsa are used from 199 patients with the following confirmed diagnosis: (i) viral infection, (ii) acute leukemia, and (iii) B or T lymphoid neoplasm. For each smear, the pathologist selects a set with between 10 and 90 cell images belonging to the cell group corresponding to the confirmed diagnosis. The final distribution is as follows: RL in 41 smears, BL in 233 (190 myeloid and 43 lymphoid origins), and AL in 144 smears.

The set of images from each individual smear is given to the classifier, which performs a blind automatic recognition whose outcomes are the images belonging to each class. The analysis of a single blood smear by the classifier may be considered as a binary test with *correct* or *incorrect* results. The test is considered to be *correct* if the percentage of images correctly recognized in the true confirmed category is above a prescribed threshold. If none of the 3 groups is above the threshold, then smear is considered in the unknown category.

Figure 3 (left part) shows the confusion matrix, which summarizes the test results with 50% threshold. This 50% threshold, or majority criterion, was chosen because using this value, the highest percentage of smears was correctly classified.

All the 41 smears containing RL are classified in the correct class (100% accuracy). With respect to blast cells images, 97% of the smears obtained from patients with acute leukemia are correctly identified, while the remaining 3% is classified as unknown by the system. Finally, 96% of the smears containing abnormal lymphoid cells are recognized in the right class by the automatic classifier. Global accuracy was 97.7%. These results confirm that the set of quantitative features selected is useful for the automatic recognition between reactive cells and neoplastic (abnormal lymphocyte or blast cells) cells in PB.

It is interesting to mention that 10 of the 20 most relevant features for the classification into reactive or neoplastic PB cells were parameters obtained from the *granulometric* and *pseudo-granulometric curves* (see Figure 4). In contrast, in a previous study reported<sup>23</sup> for the discrimination among 12 lymphoid cell groups, only 2 among the most relevant features were granulometric. The high number of relevant granulometric features contributing to the discrimination between reactive and neoplastic PB cells may be due to the differences in the large cytoplasm present in RL and some myeloblasts with respect to the low cytoplasm present in most of the abnormal lymphoid cells. This is translated into very different

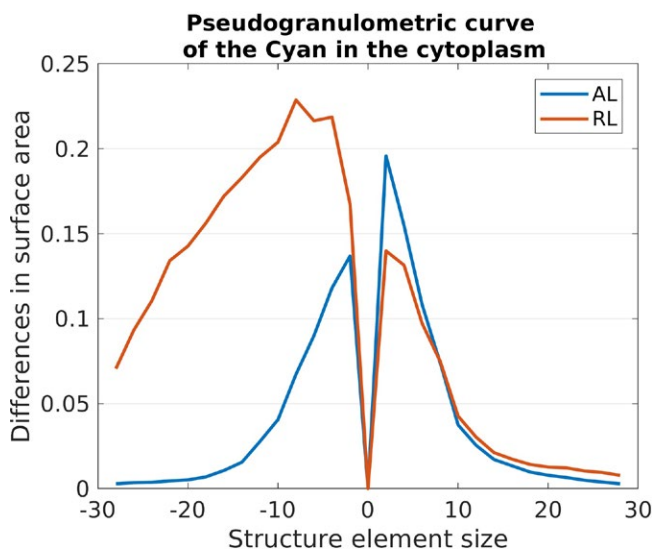
		Predicted class			
		RL	ALC	BL	UNKNOWN
True class	RL	100	0	0	0
	ALC	0	96	3	1
	BL	0	0	97	3

Feature	Color component	Color space	ROI	Feature type
1. Kurtosis	Blue	RGB	C	Stat 1 <sup>st</sup> order
2. Perimeter	-	-	N	Geometric
3. Equivalent diameter	-	-	C	Geometric
4. Skewness of the granulometric curve	Lightness	Lab	N	Granulometric
5. IMC1	Yellow	CMYK	N	Stat 2 <sup>nd</sup> order
6. SD of the pseudogranulometric curve	Cyan	CMYK	Cyt	Granulometric
7. Skewness of the histogram	Blue	RGB	C	Stat 1 <sup>st</sup> order
8. Skewness of the pseudogranulometric curve	v	Luv	C	Granulometric
9. Mean of the granulometric curve	Black	CMYK	Cyt	Granulometric
10. Skewness of the pseudogranulometric curve	v	Luv	N	Granulometric
11. Skewness of the granulometric curve	Black	CMYK	Cyt	Granulometric
12. Correlation	Saturation	HSV	N	Stat 2 <sup>nd</sup> order
13. Cluster prominence	Magenta	CMYK	C	Stat 2 <sup>nd</sup> order
14. SD of the histogram	Blue	RGB	C	Stat 1 <sup>st</sup> order
15. Mean of the pseudogranulometric curve	Magenta	CMYK	Cyt	Granulometric
16. Kurtosis of the pseudogranulometric curve	Green	RGB	N	Granulometric
17. SD of the granulometric curve	v	Luv	C	Granulometric
18. RBC proximity	-	-	-	Geometric
19. IMC1	Saturation	HSV	N	Stat 2 <sup>nd</sup> order
20. Skewness of the pseudogranulometric curve	Hue	HSV	C	Granulometric

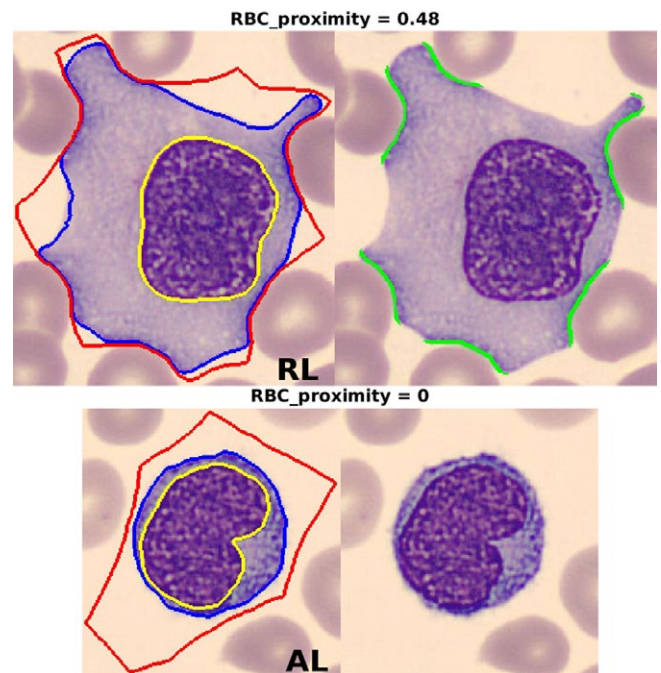
IMC1: information measure of correlation 1; ROI: regions of interest; RBC: red blood cells; SD: standard deviation; Stat: statistical; C: cell; N: nucleus; Cyt: cytoplasm

**FIGURE 3** List of the 20 most relevant features in the classification. Color component, color space, region of interest, and feature type are shown. Left part shows the confusion matrix, which summarizes the predicted classes with respect to the true classes, considering a threshold  $\geq 50\%$  of the set images correctly recognized in the true category. All the 41 smears containing RL were classified in the correct class (100% accuracy). The 97% of the smears obtained from patients with acute leukemia were correctly identified, while the remaining 3% were classified as “unknown” by the system. Finally, 96% of the smears containing abnormal lymphoid cells were recognized in the right class by the automatic classifier. RL, reactive lymphocytes



**FIGURE 4** Different pseudogranulometric curves of the cytoplasm (in the Cyan component of the CMYK color space) obtained for a RL (in red) and for a SMZL (in blue). RL, reactive lymphocytes; SMZL, splenic marginal zone lymphoma

cell texture and discriminative values of the features based on the granulometric curves among these groups. Figure 4 shows an example of the different values for the pseudogranulometric curves of the cytoplasm (in the Cyan component of the CMYK color space) obtained for a RL (in red, mean = 0.1005) and an abnormal lymphocyte (villous lymphocyte in SMZL, in blue, mean = 0.0412) images, respectively. The meaning of the values for both curves is that there are differences in the distribution of the dark and bright granules on the cytoplasm in RL, with respect to abnormal lymphocytes and blast cells.



**FIGURE 5** RBC proximity values in pixels, defined as the proportion of the reactive lymphocyte (RL) perimeter that intersects the peripheral zone around the cell, corresponding to 2 images acquired with the DM96: RL (0.40) and splenic marginal zone lymphoma lymphocyte or SMZL (0). SMZL, splenic marginal zone lymphoma

As observed in Figure 5, when a RL touches a RBC, there is a coincidence between the RL cytoplasm perimeter and the edge of the peripheral zone around the RL. As we explained in the section corresponding to geometric features, the RBC proximity is defined as the proportion of the RL perimeter (in pixels) that

intersects the peripheral zone around the cell. This feature has shown a good discriminative capability between reactive and neoplastic cells as it is shown in the Figure 5, which displays *RBC proximity* values in pixels, corresponding to 2 images acquired with the DM96: RL (0.40) and splenic marginal zone lymphoma lymphocyte or SMZL (0).

## 5 | CONCLUDING REMARKS

Although morphology is used widely in the diagnosis of leukemias and lymphomas, visual morphologic differentiation among subtypes requires much experience and skills. The finding of quantitative features, for qualitative morphological characteristics by image analysis, may provide valuable information on clinically important properties of blood cells. These features could be incorporated to the new hematological analyzers, which are currently under development based on automatic morphology computer imaging of PB smears, to provide additional flags for blast or abnormal lymphoid cell detection. Red cell parameters and reticulocyte indices play an essential role in differential diagnosis of anemia and in its treatment.<sup>36</sup> Similar to red cell indices provided by automated hematology analyzers that indicated changes in red cell size, other specific lymphocyte or blast measures can be introduced into the new automated analyzers to optimize morphology through blood cell image analysis.

Work is in progress to describe the following new descriptors (i) to measure CLL cells typically nuclear clumped chromatin the *Cyan correlation of the nucleus* (statistic parameter based in the GLCM), (ii) to measure the cerebriform chromatin, which is characteristic in Sézary lymphoid cells, the *SD of the granulometric curve of the cyan of the nucleus*, and (iii) to detect LGL granules, the *skewness of the u component histogram of the cytoplasm*.

In addition, a morphologic-genetic correlation has been recognized for certain acute myeloid leukemia subtypes. As an example, there is a strong association between acute myelomonocytic and acute monocytic and myeloid leukemia with nucleophosmin (*NPM1*) mutation.<sup>37</sup> Some other correlations involving morphology and cytogenetic changes in abnormal lymphoid cells may show up using image analysis.

In the future, the expanded use of current and new quantitative features for nuclear, cytoplasmic, and cell abnormalities may lead toward an objective morphological assessment of blood cells. Crucial challenges for the laboratories in the implementation of this technology will include the standardization of the quality of PB smear staining to ensure appropriate, high-quality smears and to minimize the difference in images acquired from different sources. In addition, it would be a requirement for the laboratory professionals to increase their knowledge about image analysis and related topics, and to be involved in multidisciplinary collaborations with engineers and mathematicians, among others.

## ACKNOWLEDGEMENTS

This work is part of a research project funded by the Directory of Science, Technology and Innovation of the Ministry of Economy and Competitiveness of Spain, with reference DPI2015-64493-R.

## REFERENCES

- Jahanmehr SA, Rogers M, Zheng J, Lai R, Wang C. Quantitation of cytological parameters of malignant lymphocytes using computerized image analysis. *Int J Lab Hematol*. 2008;30:278-285.
- Bain BJ. *Blood Cells: A Practical Guide*, 3rd edn. London, UK: Blackwell Science Ltd; 2002.
- Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumors of haematopoietic and lymphoid tissues. In: Lloyd R, Osamura R, Klöppel G, Rosai J, eds. *World Health Organization Classification of Tumours*, 4th edn. Lyon, France: IARC Press; 2017:245-246.
- Eldar S, Sabo E, Cohen A, et al. Computer-assisted image analysis of small cell lymphoma of the thyroid gland. Comparison of the nuclear parameters of small lymphocytes in lymphomas and Hashimoto's thyroiditis. *Comput Med Imaging Graph*. 1998;22:479-488.
- Dalle Carbonaire L, Valenti MT, Bertoldo F, et al. Bone microarchitecture evaluated by histomorphometry. *Micron*. 2005;36:609-616.
- Ceelle H, Dinkelaar RB, van Gelder W. Examination of peripheral blood films using automated microscopy; evaluation of Diffmaster Octavia and Cellavision DM96. *J Clin Pathol*. 2007;60:72-79.
- Briggs C, Longair I, Slavik M, et al. Can automated blood film analysis replace the manual differential? An evaluation of the CellaVision DM96 automated image analysis system *Int J Lab Hematol*. 2009;31:48-60.
- Huisjes R, van Solinge WW, Levin MD, van Wijk R, Riedel JA. Digital microscopy as a screening tool for the diagnosis of hereditary hemolytic anemia. *Int J Lab Hematol*. 2018;40:159-168.
- Merino A, Brugués R, Garcia R, Kinder M, Bedini JL, Escolar G. Comparative study of peripheral blood morphology by conventional microscopy and Cellavision DM96 in hematological and non hematological diseases. *Int J Lab Hematol*. 2011;33(Suppl. 1):112.
- Bruegel M, George T, Feng B, et al. Multi centre evaluation of the sensitivity and specificity of the cobas m 511 integrated hematology analyzer. *ISLH Abstr Proc*. 2017;39(Supplement S2):21.
- Chen S, Zhao M, Wu G, Yao C, Zhang J. Recent advances in morphological cell image analysis. *Comput Math Methods Med*. 2012;2012:101536.
- Gutiérrez G, Merino A, Domingo A, Jou JM, Reverter JC. EQAS for peripheral blood morphology in Spain: a 6-year experience. *Int J Lab Hematol*. 2007;30:460-466.
- Chabot-Richards DS, Foucar K. Does morphology matter in 2017? An approach to morphologic clues in non-neoplastic blood and bone marrow disorders. *Int J Lab Hematol*. 2017;39(Suppl 1):23-30.
- Bain BJ. Diagnosis from the blood smear. *N Engl J Med*. 2005;353:498-507.
- González RC, Woods RE. *Digital Image Processing*. Upper Saddle River, NJ: Prentice Hall; 2008.
- Sadeghian F, Seman Z, Ramli AR, Abdul Kahar BH, Saripan MI. A framework for white blood cell segmentation in microscopic blood images using digital image processing. *Biol Proced Online*. 2009;11:196-206.
- Alfárez S, Merino A, Mujica LE, Ruiz M, Bigorra L, Rodellar J. Automatic classification of atypical lymphoid B cells using digital blood image processing. *Int J Lab Hematol*. 2014;36:472-480.
- Alfárez S, Merino A, Bigorra L, Mujica L, Ruiz M, Rodellar J. Automatic recognition of atypical lymphoid cells from peripheral blood by digital image analysis. *Am J Clin Pathol*. 2015;143:168-176.

19. Alférez S, Merino A, Bigorra L, Rodellar J. Characterization and automatic screening of reactive and abnormal neoplastic B lymphoid cells from peripheral blood. *Int J Lab Hematol*. 2016;38:209-219.
20. Angulo J, Klossa J, Flandrin G. Ontology-based lymphocyte population description using mathematical morphology on color blood images. *Cell Mol Biol (Noisy-le-grand)*. 2006;52:2-15.
21. Mohammed EA, Mohamed MM, Naugler C, Far BH. Application of support vector machine and k-means clustering algorithms for robust chronic lymphocytic leukemia color cell segmentation. Proceedings of the 15th IEEE International Conference on e-Health Networking, Application and Services, 2013.
22. Lunning MA, Zenger VE, Dreyfuss R, et al. Albumin enhanced morphometric image analysis in CLL. *Cytometry B Clin Cytom*. 2004;57B:7-14.
23. Puigví L, Merino A, Alférez S, Acevedo A, Rodellar J. New quantitative features for the morphological differentiation of abnormal lymphoid cell images from peripheral blood. *J Clin Pathol*. 2017;70:108-148.
24. Angulo J. A mathematical morphology approach to cell shape analysis. In: Bonilla LL, Moscoso M, Platero G, Vega JM, eds. *Progress in Industrial Mathematics at ECMI 2006*. Berlin, Heidelberg: Springer; 2008:2-6.
25. Haralick RM, Shanmugan K, Dinstein I. Textural features for image classification. *IEEE Trans Syst Man Cybern*. 1973;3:610-621.
26. Albrechtsen F. *Statistical Texture Measures Computed from Gray Level Co-occurrence Matrices*. Image Processing Laboratory, Department of Informatics, University of Oslo Web site; <http://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/g lcm.pdf>. Published 1995. Updated November 5, 2008. Accessed September 7, 2013.
27. Alférez S. Methodology for automatic classification of atypical lymphoid cells from peripheral blood cell images. PhD Thesis. Technical University of Catalonia, 2015.
28. Yang X, Tridandapani S, Beitler J, et al. Ultrasound GLCM texture analysis of radiation-induced parotid-gland injury in head-and-neck cancer radiotherapy: an in vivo study of late toxicity. *Med Phys*. 2012;39:5732-5739.
29. Kono K, Hayata R, Murakami S, et al. Quantitative distinction of the morphological characteristic of erythrocyte precursor cells with texture analysis using gray level co-occurrence matrix. *J Clin Lab Anal* 2018;32:1-6.
30. Yamamoto M, Kono K, Kuroki M, et al. Texture analysis for quantitative distinction of the morphological characteristics of white blood cells in peripheral blood. *ISLH 2017 Abstr Proc*. 2017;39(Supplement S2):126.
31. Patel N, Mishra A. Automated leukaemia detection using microscopic images. *Procedia Comput Sci*. 2015;58:635-642.
32. Sabino D, Costa L, Gil Rizzatti E, Zago MA, Gil Rizzatti E, Zago MA. A texture approach to leukocyte recognition. *Real-Time Imaging*. 2004;10:205-216.
33. Benattar L, Flandrin G. Morphometric and colorimetric analysis of peripheral blood smears lymphocytes in B-cell disorders: proposal for a scoring system. *Leuk Lymphoma*. 2001;42:29-40.
34. Bigorra L, Merino A, Alférez S, Rodellar J. Feature analysis and automatic identification of leukemic lineage blast cells and reactive lymphoid cells from peripheral blood cell images. *J Clin Lab Anal*. 2017;31:e22024.
35. Rodellar J, Alférez S, Acevedo A, Molina A, Merino A. Image processing and pattern recognition in the morphological analysis of blood cells. *Int J Lab Hematol*. 2018;40:??-??.
36. Buttarello M. Laboratory diagnosis of anemia: are old and new red cell parameters useful in classification and treatment, how? *Int J Lab Hematol*. 2016;38(Suppl 1):123-132.
37. Fallini B, Nicoletti I, Martelli MF, Mecucci C. Acute myeloid leukemia carrying cytoplasmic/mutated nucleophosmin (NPMc+ AML): biologic and clinical features. *Blood*. 2007;109:875-885.

**How to cite this article:** Merino A, Puigví L, Boldú L, Alférez S, Rodellar J. Optimizing morphology through blood cell image analysis. *Int J Lab Hem*. 2018;40(Suppl. 1):54-61. <https://doi.org/10.1111/ijlh.12832>