Published in final edited form as:
Cytometry A. 2014 March ; 85(3): 242–255. doi:10.1002/cyto.a.22432.

Histopathological image analysis for centroblasts classification through dimensionality reduction approaches

Evgenios N. Kornaropoulos[1,*], **M Khalid Khan Niazi**[2], **Gerard Lozanski**[3,†], and **Metin N. Gurcan**[2,†,‡]

[1]Informatics and Telematics Institute - Centre for Research and Technology Hellas (ITI - CERTH), Thessaloniki, Greece

[2]Ohio State University, Department of Biomedical Informatics, Columbus, Ohio, United States of America (USA)

[3]Ohio State University, Department of Pathology, Columbus, Ohio, United States of America (USA)

Abstract

We present two novel automated image analysis methods to differentiate centroblast (CB) cells from non-centroblast (Non-CB) cells in digital images of H&E-stained tissues of follicular lymphoma. CB cells are often confused by similar looking cells within the tissue, therefore a system to help their classification is necessary. Our methods extract the discriminatory features of cells by approximating the intrinsic dimensionality from the subspace spanned by CB and Non-CB cells. In the first method, discriminatory features are approximated with the help of Singular Value Decomposition (SVD), whereas in the second method they are extracted using Laplacian Eigenmaps. Five hundred high-power field images were extracted from 17 slides, which are then used to compose a database of 213 CB and 234 Non-CB region of interest images. The recall, precision and overall accuracy rates of the developed methods were measured and compared with existing classification methods. Moreover, the reproducibility of both classification methods was also examined. The average values of the overall accuracy were 99.22% ± 0.75% and 99.07% ± 1.53% for COB and CLEM, respectively. The experimental results demonstrate that both proposed methods provide better classification accuracy of CB/Non-CB in comparison to the state of the art methods.

Keywords

Follicular lymphoma; dimensionality reduction; intrinsic dimensionality; SVD; LDA; Laplacian Eigenmaps

[*]Corresponding author: E. N. Kornaropoulos (ekornaro@gmail.com), Department of Computer Vision, Informatics and Telematics Institute - Centre for Research and Technology Hellas (ITI - CERTH), 1st km Thermi - Panorama, 57001, Thessaloniki, Greece.
[†]Gerard Lozanski and Metin N. Gurcan are both senior authors and contributed equally to this paper.
[‡]Senior Member, IEEE

I. Introduction

Follicular lymphoma (FL) is the second most common lymphoid malignancy in the western world [1]. FL is a disease with highly variable clinical course and is currently only curable with stem cell transplant. In a subset of patients, FL behaves as chronic indolent disease that may never require therapy, whereas in another subset it acts aggressively and if not treated with chemo-immunotherapy will result in patient's rapid death. Moreover, in a subset of patients with an initial indolent presentation, FL may transform into the aggressive type where symptomatic patients must be treated. The reason why chemo-immunotherapy does not apply in indolent cases of the disease is due to the lack of evidence in how beneficial it will be. Therefore, in order to avoid unnecessary toxicity, treatment is reserved for symptomatic patients and for those with the aggressive form of the disease.

Oncologist's clinical decisions at the time of initiation and type of therapy are guided by risk stratification. Risk stratification for each individual patient is based on a combination of clinical and laboratory findings, including morphological characteristics of tumor tissue.

The World Health Organization (WHO) adopted the method proposed by Risa et al., which is the most established and recommended one for morphological risk stratification of FL [39]. In this method, a pathologist stratifies FL cases into histological grades based on the number of large cancer cells, called centroblasts (CB). CBs are counted manually in standard microscopic high power fields (HPFs) from Hematoxilin and Eosin (H&E) stained tissue biopsies of FL. Due to the time constraints, CBs are only counted in ten random HPFs, sampled from sections of malignant follicles [2]. Using this method, FL cases are categorized into three histological grades according to the average CB count per HPF; grade I (0–5 CB/HPF), grade II (6–15 CB/HPF) and grade III (>15 CB /HPF). Grade III cases are further sub-classified into grade IIIa and IIIb. Follicles of grade IIIa contain CBs along with small cells called centrocytes, while follicles of grade IIIb contain pure populations of CBs without admixed centrocytes [1]. Grades I and II are considered as low risk cases and may not require treatment unless the patient is symptomatic. Grades IIIa and IIIb though are evaluated as high risk and indicate an aggressive type of disease that requires immediate life-saving chemotherapy.

Manual histological grading of FL is highly subjective and it requires considerable effort, along with an extensive training. Furthermore, the analysis of even one HPF under a light microscope by the pathologist is a time consuming process. This laborious and time demanding nature of the method is the reason why only ten HPF are analyzed per case. This limitation makes the method highly vulnerable to sampling bias in cases of tumors with high tissue heterogeneity. Moreover, due to the subjective nature of this method, some prognostic clues are not easily observed by all pathologists. Indeed, inter- and intra-pathologist variability has been reported, which normally ranges from 61% to 73% [4, 5]. All these issues directly affect the clinical decisions about the timing and the type of therapy. Therefore, increasing the reliability and reproducibility of the process of histological grading is of great importance.

Computer-aided diagnosis (CAD) is increasingly used in medicine to aid with detection, diagnosis and prognosis of diseases both in radiology and pathology [6, 7, 42, 43, 44]. The quantitative analysis of histological tissues is the subject of several earlier studies and it has been successfully applied in the diagnosis of prostate cancer [19], neuroblastoma (NB) [20, 21], and breast cancer [22]. To better understand the different stages of CAD and some of its applications, a comprehensive review can be found in [23].

CAD has also been reported to be beneficial in classifying tissue subtypes associated with various grades of FL [8-15, 45]. The classification is often achieved by extracting features from H&E stained images of FL. These features can distinguish the cells from each other, and thereby classify them into different categories. Several morphological and topological features [13], or their combination with texture features [8, 12], as well as with graph-based features [14] have been introduced. The most discriminating subset of these feature vectors are often identified in a lower dimensional space, created by employing Principal Component Analysis (PCA) that captures the main modes of variations in the data. Using this approach, classification accuracies ranging from 75% to 85% are reported [8,12,13], but this level of performance may not be sufficient enough for certain clinical applications. Selection of the best and clinically meaningful feature sets, as well as the design of classifiers, is an active area of research [8, 9, 12, 13].

In the current study, our aim is to classify the cells in FL images in a similar manner as perceived by the pathologists when they review tissue of FL under a microscope. For this purpose, we set out to develop a tool called *Kyttaro* (meaning cell in Greek) that acts as a content-based image retrieval system. This system brings the most relevant cell images from its library of cell images, which are already classified into CB or Non-CB categories.

In clinical practice, pathologists identify several features of CB, such as size, circularity, coarse texture, multiple nucleoli, vesicular chromatin and accentuated nuclear membrane. Moreover, pathologists also take into account the structures around the cell, while making a decision. However, not every pathologist uses these features; part of the knowledge is implicit. Therefore, we concluded that we should consider the whole image of a cell with its surroundings as a feature vector. In that way, we incorporate all the features mentioned by the pathologist. Furthermore, redundant features are removed by linear and non-linear dimensionality reduction methods.

The section to follow provides detailed information about the database used in the current study. Section III describes the proposed classification methods along with a preprocessing step necessary to suppress noise from the images. The training process of each proposed classifier, as well as its comparative analysis with the state of the art methods are presented in Section IV. This is followed by a comprehensive discussion in Section V. Finally, the conclusions are given in Section VI.

II. Image Database

Tissue biopsies of FL stained with H&E, from 17 different patients were scanned using a high-resolution whole slide scanner (Aperio - Image Scope). Three board-certified hematopathologists selected 500 HPF images of follicular lymphoma out of the scanned

tissue biopsies. These 500 images are then examined by two expert pathologists by using a remote viewing and annotation tool, developed in our lab, to mark CB cells on the HPF images. Using these markings, a set of images of CB cells was created. Each image contains the CB cell at its center and is of size 71 × 71 pixels (Figure 1a). Similarly, a second set of same size images of cells that were not marked by any pathologist as CB was generated. These images are called Non-CB cells and typically include centrocytes, histoicytes, dendric cells (Figure 1b). All together, the database is composed of 213 CB and 234 Non-CB images. These cases were selected from the archives of The Ohio State University with Institutional Review Board (IRB) approval (Protocol 2007C0069, renewed May 13, 2013).

III. Method

In this section, we describe the process of noise removal in the cell images, as well as the two proposed methods of cells classification in FL images. While the first method extracts discriminative features by utilizing linear dimensionality reduction, the second one uses a non-linear dimensionality reduction to extract the discriminative features. The test image is first projected into a low-dimensional space (discriminating feature space). Then the class label of the image is determined by a distance function. The image retrieval system of *Kyttaro* tool will be based on the most efficient of the two classification methods.

A. Noise removal

Microscopic images show variation within them or between them due to the conditions under which they were acquired. Tissue cutting, processing and staining during slide preparation are some of the steps that cause these variations, making it difficult to perform consistent quantitative analysis on these images [40]. Therefore, all the images in our database were first converted to grayscale and then standardized to partially compensate for these differences. The new image after standardization is a centered, scaled version of the grayscale image of a cell. Moreover, to reduce some salt-and-pepper type of noise while preserving the inherent texture characteristics, we applied median filtering with a kernel size of 5-by-5, to the standardized gray-scale images (see Figure 2).

B. Classification using orthogonal bases (COB)

The design of this classifier is inspired by the theoretical framework provided by Lars Elden [38]. According to this framework, the CB/Non-CB classification problem can be formulated as a minimization problem, where the objective is to minimize the least-square error between a given image of CB/Non-CB and its low rank approximation. We assume that we have n images of CB and the same number of images of Non-CB cells. Let any (i.e. CB or Non-CB) cell image be a vector in a m -dimensional vector space, where m is equal to the number of pixels in each image. Here, the vector entries are the intensity values at each pixel. Within the m -dimensional vector space of all the images, there lie two different hypothetical subspaces, one for the CB cells and the other for the Non-CB cells. We will now present a detailed methodology to determine each hypothetical subspace which will enable us to differentiate between CB and Non-CB cell images.

Let $A^i \in \mathbb{R}^{m \times n}$ be the matrix

$$A^i = \begin{bmatrix} p^i_{11} & \cdots & p^i_{n1} \\ \vdots & \ddots & \vdots \\ p^i_{1m} & \cdots & p^i_{nm} \end{bmatrix}, i \in \{CB, Non-CB\}, \quad (1)$$

where columns are formed by concatenating *n m* -dimensional vectors of CB or Non-CB cells in the training set. So, essentially each column corresponds to a unique CB or Non-CB image from the set of *n* images. Unlike other popular methods, we created two different matrices, one for CB and another one for Non-CB.

To remove the redundancy from each of these rectangular matrices, we can write A^i as:

$$A^i = U^i \begin{pmatrix} \sum^i \\ 0 \end{pmatrix} V^{i,T}, i \in \{CB, Non-CB\}, \quad (2)$$

where $U^i \in \mathbb{R}^{m \times m}$ and $V^i \in \mathbb{R}^{n \times n}$ are the orthogonal matrices, whose columns represent the left and right orthonormal eigenvectors for A^i, respectively. Here, Σ^i corresponds to the eigenvalues of A^i. Further, suppose that A^i has rank *r*. Then, matrices A^i can be written as:

$$A^i = \begin{bmatrix} u^i_1 \dots u^i_r \end{bmatrix} \begin{bmatrix} \sigma^i_1 & & \\ & \ddots & \\ & & \sigma^i_r \end{bmatrix} \begin{bmatrix} \nu^{i,T}_1 \\ \vdots \\ \nu^{i,T}_r \end{bmatrix} + \begin{bmatrix} u^i_{r+1} \dots u^i_m \end{bmatrix} [0] \begin{bmatrix} \nu^{i,T}_{r+1} \\ \vdots \\ \nu^{i,T}_n \end{bmatrix}, i \in \{CB, Non-CB\}, \quad (3)$$

where u^i_k is the k^{th} *m* - dimensional orthonormal eigenvector of U^i, $\nu^{i,T}_k$ is the k^{th} *n* - dimensional orthonormal eigenvector of V^i and σ^i_k is the k^{th} eigenvalue of A^i. Since *m*»*n*, in case A^i has a full column rank, *r* would be equal to *n*. In that case *r* should be substituted by *n* in equation (3) and the following equations.

Let:

$$X = \begin{bmatrix} u^i_1 \cdots u^i_r \end{bmatrix} \begin{bmatrix} \sigma^i_1 & & \\ & \ddots & \\ & & \sigma^i_r \end{bmatrix} = \begin{bmatrix} \sigma^i_1 u^i_1 & \dots & \sigma^i_r u^i_r \end{bmatrix}, i \in \{CB, Non-CB\}$$
$$Y = \begin{bmatrix} \nu^{i,T}_1 \\ \vdots \\ \nu^{i,T}_r \end{bmatrix} \quad (4)$$

Then $A^i = XY$ can be written as an outer product expansion as:

$$A^i = \sum_{j=1}^{r} \sigma^i_j u^i_j \nu^{i,T}_j, i \in \{CB, Non-CB\} \quad (5)$$

Each of the terms $\sigma_j u_j \nu_j^T$ is an outer product of vectors u_j and v_j, weighted by the eigenvalue σ_j.

Equation (5) shows us that any new image of a cell z can be expressed in terms of the orthogonal basis $\left\{u_1^i \cdots u_r^i\right\}$ as:

$$A^i z = \sum_{j=1}^{r} \sigma_j^i u_j^i \nu_j^{i,T} z, i \in \{CB, Non - CB\} \quad (6)$$

Since $\nu_j^{i,T} z$ is a scalar, we can rearrange the order of the factors to:

$$A^i z = \sum_{j=1}^{r} (\nu_j^{i,T} z \sigma_j^i) u_j^i, i \in \{CB, Non - CB\} \quad (7)$$

Now z is expressed as a linear combination of left singular vectors $\left\{u_1^i \cdots u_r^i\right\}$. The left singular eigenvectors are an orthogonal basis for the column space of A^i so for the "image space of CB" or the "image space of Non-CB." We will refer to these eigenvectors as "singular images." Each coefficient in equation (7) is a product of two factors, $\nu_j^{i,T} z$ and σ_j. The factor, $\nu_j^{i,T} z = \nu_j^i \bullet z$, where • denoted the inner product of the two vectors, is j^{th} component of z relative to the orthonormal basis $\left\{\nu_1^i \cdots \nu_r^i\right\}$. Under the action of A_i, each ν component of z becomes a u component after scaling by the proper σ. Hence, the coefficients in equation (7) express the coordinates of the image z in terms of the singular images $\left\{u_1^i \cdots u_r^i\right\}$.

The ordering of the singular images is based on the value of their corresponding eigenvalues. The eigenvalues of u_i are proportional to the values of the covariance matrix among the images, which represent the second dimension of A_i. Therefore the eigenvector with the highest eigenvalue points to the highest variance among the images. We assume that the discriminating features of CB/Non-CB would be revealed in the directions pointing towards the highest variance among the images. We therefore order the singular images on a descending order of their corresponding eigenvalues.

In Figure 3 we present the grayscale version of the first three singular images in the "image space of CB" and the first three singular images in the "image space of Non-CB." Reviewing these images, one can note that the first singular image may represent the size feature of the cells because CB cells are typically bigger than Non-CB cells in size. However, no immediate associations can be made with the features represented by the following singular images.

The singular images make up the bases of the subspace of CB and Non-CB images. Any unknown cell image can be classified as CB or Non-CB by calculating representations of the image in terms of CB or Non-CB bases. The cell class whose bases approximate the unknown image better in terms of least square error is selected as the class of the image.

To determine which of the two bases of cell images better represent an unknown cell image *z*, the norm of the residual vector is calculated in the least squares sense:

$$\rho(z)=\|z-\sum_{h=1}^{k}\alpha_h\, u_h\|_2, \quad (8)$$

where $\rho(z)$ denotes the norm of the residual vector $\left(z-\sum_{h=1}^{k}\alpha_h\, u_h\right)$ in each basis (it can be $\rho(z)^{CB}$ or $\rho(z)^{Non-CB}$ when we have u_h^{CB} and u_h^{Non-CB} corresponding to CB and Non-CB singular basis vector images).

This problem can be restated in the following form:

$$\alpha_{opt} \rightarrow \min_{\alpha}(\|z-U_k\,\alpha\|_2), \quad (9)$$

where $U_k = (u_1 u_2 \ldots u_k)$ and a is the linear combination of weights that need to be assigned to each singular image u_j. As mentioned before, the singular images, which consist of the columns of U_k, are orthonormal, therefore $U_k^T = U_k^{-1}$. The solution of this problem is now given by replacing the factor α in equation (9) by $\alpha = U_k^T z$. The nice thing about this formulation is that we do not need to compute the inverse while solving the least square problem, which is generally the case. This makes it applicable to data residing in very high dimensional space. In the end, we are calculating the norm of the residual vector as follows:

$$\rho(z)=\|(I-U_k\,U_k^T)z\|_2 \quad (10)$$

The classification criterion is defined like this:

Classification algorithm of method 1:

Training: *For the training set of images of known CB and Non-CB cells compute the SVD of each of the set of cells of the two kinds.*

Classification: *If* $\rho(z)^{CB} \leq \rho(z)^{Non-CB}$ *then* z *belongs to CB class otherwise* z *belongs to Non-CB class.*

The reason for assigning the unknown image to *CB* class in cases of a tie is because we want to be conservative in the decision and leave the final decision to the pathologist.

C. Classification based on Laplacian Eigenmaps (CLEM)

A cell can be considered as a collection of numbers, each specifying light intensity. But this collection also specifies the Cartesian coordinates of a point with respect to a set of axes [47]. Therefore, any cell image can be identified with a point in an abstract *image space*.

The second classifier is based on preserving the similarities among the images of CB/Non-CB cells by using Laplacian Eigenmaps (LEM). LEM is a non-linear dimensionality reduction method which approximates the lower-dimensional manifold embedded in the

abstract image space, while preserving the intrinsic spatial relationship among cell images [31, 48]. In that way, it overcomes the limitation of linear dimensionality methods in which the low dimensional space is approximated by projecting the data onto the eigenvectors. The preservation of spatial relationship in LEM makes it relatively insensitive to outliers and noise.

In order to achieve spatial relation preservation, an undirected, weighted graph $G = <V, E>$ of cells is constructed. This is the most crucial step of the LEM algorithm. The cell images constitute the nodes, or vertices (V) of the graph and the approximate similarities among the cells are represented by the edges (E) of the graph. Weights are assigned to the edges based on the similarity values among the images of the cells, resulting in a weight matrix W. Two metrics were considered, a dissimilarity measure (the Euclidean distance metric) and a similarity measure (the linear cross-correlation coefficient metric), both preferred for their simplicity.

Given k number of images of CB and Non-CB cells $x_1,..,x_k$ in the abstract image space $\mathbb{R}^{m^2}$, we construct a weighted graph with k nodes, one for each image, and the set of edges connecting neighboring images to each other. An edge is put between nodes i and j if x_i and x_j are "similar." To define similarity, typically either the ε-neighborhoods or the p-nearest neighbors approach is used. In ε-neighborhood approach, the similarity is defined by a threshold ε without taking into account the number of the neighbors that one node might have. In the p-nearest neighbors approach, on the other hand the p most similar images are chosen as the most similar neighbors. The former approach is more geometrically motivated but may lead to several connected images, in case of a bad choice of a low threshold ε. Although the p nearest neighbors is less geometrically motivated, it is simpler in implementation. Additionally, in the p-nearest neighbors approach, an isolated node might end up having neighbors that have in fact very low similarity with it. In order to define similarity, we use a combination of the ε- neighborhoods approach and the p-nearest neighbors approach. In our method the i_{th} node is connected to the j_{th} node if it is in the group of its p (p=10) closest neighbors (and vice versa), but also satisfies the following criterion:

$$dist\,(|x_i - x_j|) < \varepsilon \quad (11)$$

where *dist* is either the Euclidean distance or the cross-correlation coefficient metric, $\varepsilon = \alpha m_d$ with m_d being the mean score of the Euclidean distances or the cross-correlation coefficients, depending on the metric used, among all the training images. In our case, the cross-correlation coefficient was chosen empirically as the metric to reflect the similarities among the cell images, as it provided better classification results. In the end, our approach is a hybrid of ε-neighborhoods and the p-nearest neighbors approaches. The value of α was determined experimentally as 1.15 with the assumption that the threshold should be close to the value m_d. Therefore, we searched the optimal value of threshold in the range of [0.05m_d 1.5m_d] in incremental steps of 0.05.

Moreover, the weights are assigned by a heat kernel:

$$W_{ij}=e^{-\frac{\|x_i-x_j\|_2^2}{t}} \quad (12)$$

By using the heat kernel approach while assigning weights to the edges of a graph, we end up in a similarity preserving character of classification that emphasizes the natural clusters in the images. Equation 14 holds only for small *t* so as the heat kernel to become increasingly localized and be characterized by Dirac's δ-function [30]. Therefore, this approach is limited by the necessity to choose a value for the parameter *t*. In our work, the value for *t* was chosen empirically as 5.

After constructing the graph, the next step of the method is the creating of the manifold on which the cell images reside. The manifold is created by connecting image-nodes based on their similarity and it spans a lower dimensional space. The problem of defining the manifold can be considered as generalized eigenvector problem. It can then be stated as a problem of finding the eigenvectors that satisfy the following equation:

$$L\nu=\lambda D\nu, \quad (13)$$

where ν are the generalized eigenvectors, λ are the generalized eigenvalues, *D* is the diagonal weight matrix with its entries being column (or row since *W* is symmetric) sums of *W* ($D_{ii} = \Sigma_j W_{ji}$) and $L = D - W$ is the Laplacian matrix or the Laplacian operator of the graph. The use of the Laplacian operator of the graph in order to compute the eigenvectors is the key to achieve the optimal embedding of the manifold. As proved in [30] the Laplacian operator of the graph is an approximation of the Laplace-Beltrami operator defined on the manifold. The use of the Laplacian operator of the graph results in the optimal embedding of the manifold; therefore, its intrinsic geometric structure is reflected.

The ordering of the eigenvectors is based on the criterion of mapping the weighted connected graph *G* to a line so that two nodes ν_i, ν_j, which are connected to each other, stay as close as possible after the mapping. Let $\mathrm{v} = (\nu_1, \nu_2, \ldots, \nu_n)^T$ be the map from the graph to the real line. Then, the optimal eigenvectors are computed based on the minimization of the following objective function:

$$\nu_{opt} \rightarrow \min_{\nu}\left(\sum_{ij}(\nu_i-\nu_j)^2 W_{ij}\right) \quad (14)$$

As proven in [30], for any v,

$$\frac{1}{2}\sum_{i,j}(\nu_i-\nu_j)^2 W_{ij}=\mathrm{v}^T L\mathrm{v} \quad (15)$$

Therefore, the vector ν_{opt} that minimizes equation (14) is given by the minimum eigenvalue solution to the generalized eigenvalue problem of equation (13). In the end, the optimal eigenvector is the one that satisfies the following equation:

$$\nu_{opt}=arg\ min_{\substack{\nu^T D\nu=1 \\ \nu^T D1=0}} \nu^T L\nu, \quad (16)$$

where $\nu^T D\nu = 1$ is a constraint to remove an arbitrary scaling factor in the embedding and $\nu^T D1 = 0$ is a second constraint applied to prevent the collapse of all vertices of *G* onto the real number 1. This might happen in case of a connected graph with the value of 1 assigned to each vertex. It is easy to see that 1 is an eigenvector with eigenvalue 0 and in case of a connected graph, is the only eigenvector for $\lambda = 1$. Therefore, this constraint of orthogonally should be added. The solution ν_{opt} is finally given by the eigenvector with the smallest non-zero eigenvalue.

Having computed the eigenvectors that consist the low dimensional space spanned by the manifold, the classification is achieved by computing the Euclidean distance among the embedded cell images on the manifold (Figure 4). We use the *p*-nearest neighbors concept ($p = 3$) to achieve a more accurate estimation of the "closeness" between two embedded images. Then we create two distance variables, D_{CB} for the Euclidean distance of any tested image from CB cells and $D_{Non–CB}$ corresponding to the closest images of Non-CB cells from the training set. The classification criterion can be then summarized as:

Classification algorithm of method 2:

Training: *For the training set of images of known CB and Non-CB cells, create the manifold in which the images lie by computing the optimum generalized eigenvectors of the Laplacian equation.*

Classification: *For the* j^{th} *image, if* $D_{j,\ CB} < D_{j,\ Non–CB}$ *then image j belongs to CB class otherwise image j belongs to Non-CB class., where* $D_{i,\ CB}$ *and* $D_{i,\ Non–CB}$ *the Euclidean distance of image j from the closest to it images of CB cells and Non-CB cells from the training set respectively.*

D. Classification approaches implemented for comparison

For comparison purposes, in addition to our two classification methods, eight more (both supervised and unsupervised) classification methods were implemented. Two of these also use dimensionality reduction in order to reveal the most discriminating features of images of CB and Non-CB cells. The difference among all the implemented classification methods is the space on which classification is performed, which is either the image or a low-dimensional feature space and the final decision for the classification is based on the similarity criteria.

D.1. Unsupervised methods

1. Euclidean distance (ED): The simplest way to classify the cells is using distance (dissimilarity) or similarity metrics. The most commonly used dissimilarity metric is the Euclidean distance between the corresponding pixels in two images [31]. The classification in that case is the same approach as of the *p* - nearest neighbors. We experimentally determined the value of $p = 3$. Therefore, each test image is assigned a class label based on the label of its three nearest neighbors.

2 – 3. Correlation coefficient (CC) and mutual information (MI): As for other similarity metrics, the cross-correlation coefficient, as well as the mutual information (MI) [33] between two images of cells were examined. Again *p* nearest neighbors approach was used, experimentally setting *p* equal to 17 and 5 for cross-correlation coefficient and MI methods, respectively. As MI is based on entropy, which is a measure of histogram dispersion, this metric is intensity-based, in contrast to the Euclidean distance and cross-correlation coefficient, which are both color-based measures.

4. k-means: This is one of the most commonly used clustering methods and falls into the general category of variance based clustering [35, 36]. Given the set of $N = n_{CB} + n_{Non–CB}$ images of cells, in real m -dimensional space $\mathbb{R}^m$ (m = 5041), and an integer k that defines the number of classes considered for classification (i.e. in our case k = 2 corresponding to CB and Non-CB cells), the problem is to determine a set of k points in $\mathbb{R}^m$, called centers, so as to minimize the mean squared distance from each image to its nearest center [34].

5. Fuzzy C-means(FCM): FCM is based on minimization of the following objective function:

$$J_\gamma = \sum_{i=1}^{N}\sum_{j=1}^{C} u_{ij}^{\gamma} \left\| x_i - c_j \right\|_2^2, \quad (17)$$

where γ is weighting exponent for u_{ij} which controls the fuzziness of resulting clusters and can be any real number greater than 1, u_{ij} is the degree of membership of x_i image in the cluster j, x_i is the i^{th} of m -dimensional images (N images in total), c_j is the m -dimensional center of the cluster (C classes in total) and ‖*‖ is any norm expressing the similarity between any measured data and the center, which in our case was the Euclidean distance among the images. Each test image is assigned a class label based on the membership function [37].

D.2. Supervised methods

1. Classification by Linearly Projection in Low Dimensional Space (CLPLDS): SVD was used again in another, simpler way for dimensionality reduction and cell classification. Using all the training set of images of both CB and Non-CB cells, we calculate the eigenvectors that span the orthogonal eigenspace of all the images of both kinds of cells, and we then project the unknown image on this low dimensionality eigenspace. Again, the eigenvectors were ordered in a sequence of the descending eigenvalues. The projection of the image onto the subspace spanned by only a few eigenvectors would well characterize the cells into the CB and Non-CB. Classification was achieved by calculating the dissimilarity among the projected images, using the Euclidean distance. The smaller the score of the Euclidean distance among the projected unknown image and the projected image of CB from the training set, the more chances the unknown image to be a CB (and vice versa when comparing to a Non-CB from the training set). p -nearest neighbors (p = 3) approach was once again used in order to estimate the closeness of each unknown image to each class of cells.

2. Classification using Linear Discriminant Analysis (CLDA): Since LDA is a data classification rather than a feature extraction method (as SVD), it was used along with SVD in order to maximize the ratio of between-class variance to the within-class variance in the training data set, guaranteeing in that way maximal separability. The projected training and testing images in the eigenspace spanned with the eigenvectors calculated by SVD were used as input to the LDA algorithm. The measure of closeness is the log of the probability density function of the testing image, multiplied by the class' prior probabilities. In the same time, the probability density function of the testing image is based on the number of the neighboring to it images in each class. LDA then assigns labels (CB, Non-CB) to each of the testing images of cells, classifying them to one of the two classes. Using LDA, the distribution of the features extracted through SVD is better differentiable.

3. Classification by optimizing on the Least Squares objective function with a penalty on the L1-norm of the parameters (Lasso): Bridge regression, a special family of penalized regressions of a penalty function $\Sigma|\beta_j|^\gamma$ with $\gamma \geq 1$ is considered in this work as a classifier for CB and Non-CB images. Consider the linear regression problem $y = X\beta + \varepsilon$, where y is a m-dimensional test image, X is a m-by- n matrix formed by concatenating the training CB or Non-CB cell images, β is the n -dimensional vector of parameters and ε a m-dimensional vector of independent and identically distributed random errors. Our goal is to calculate the β parameters that minimize the squared error between $X\beta$ and then use the calculated β parameters to estimate this squared error. The class, in which the squared error is smaller, is the one that better represents the test cell image. We implemented the approach mentioned in [3] to solve bridge regression for $\gamma \geq 1$. Particularly they developed a simple algorithm for the lasso by studying the structure of the bridge estimators. The non-differentiability of the objective function in lasso was handled with "Shooting" algorithm which cycles through the parameters in order. The shrinkage parameter γ and the tuning parameter λ are selected via generalized cross-validation (GCV).

4. Quadratic discriminant Analysis (QDA) based Color Features Classification: Finally, the performance of a recently proposed quantitative methodology was also examined here [12]. This method was designed specifically for classifying CB and Non-CB cells in FL. The method was based on training and testing of a quadratic discriminant analysis (QDA) classifier. The novel aspects of that method were the identification of the CB object with prior information, and the introduction of the principal component analysis (PCA) in the spectral domain to extract color texture features. Both geometric and texture features were used to achieve the classification. The average classification rate of that classifier was equal to 82.56%. We will refer to this method as Color Features Classification (CFC).

IV. Experimental Results

In this section, we first describe the procedure to train both COB and CLEM classifiers and subsequently we present the results of their performance in comparison with the corresponding results of the state of the art classification methods.

A. Training the classifiers

In order to train the classifiers, we divided the images of CB and Non-CB into training and test sets using an 80% (training) - 20% (test) ratio. Due to the higher dimensional feature vectors with limited sample size, the reliability of the proposed methods in extracting the most discriminant features can be questioned. Therefore in order to provide a better estimation of the performance of the tested methods, the pre-mentioned division was repeated K times to yield the “Hold-out K-folds” cross-validation approach [41]. This is considered as the outer loop of the procedure.

The CB and Non-CB images allocated for training were further subdivided into 90%-10% portions, resulting in an inner loop. The division is again repeated K times. During each iteration of the inner loop, the optimal number of eigenvectors and the optimal manifold of embedded training images of CB and Non-CB classes were extracted based on the optimal number of generalized eigenvectors while training COB and CLEM respectively. The optimal number of eigenvectors during training of COB and the optimal manifold during training of CLEM are chosen based on classification result of the 10% sub-testing images. Moreover, in case of training COB, all the training images (i.e. 80%) are used at the outer loop, whereas in case of training CLEM, only the 90% of the training images are used at the outer loop.

Empirically, K is set to 10 in both approaches. An overall flow chart of the proposed classification scheme can be viewed in Figure 5.

B. Demonstration of the results

In order to evaluate the performance of various approaches described in the methods section, we used the two commonly used metrics: precision and recall. Furthermore, the overall accuracy was computed to evaluate the classifier.

Precision is defined as the ratio of the number of correctly classified cells to the total number of classified cells:

$$precision = \frac{t_p}{t_p + f_p}, \quad (18)$$

where t_p (true-positive) corresponds to the sum of the classified CB or Non-CB cells and f_p (false-positive) corresponds to the sum of the misclassified CB or Non-CB cells.

Recall is defined as the ratio of the number of classified cells to the number of cells expected to be classified:

$$recall = \frac{t_p}{t_p + f_n}, \quad (19)$$

where t_p corresponds to the sum of the classified CB and Non-CB cells and f_n (false-negative) corresponds to the sum of the non-classified CB and Non-CB cells.

Overall accuracy (OA) is defined as the ratio of the number of correctly classified CB and Non-CB cells to the total number of classified cells

$$OA=\frac{t_{p,CB}+t_{p,Non-CB}}{n_{CB}+n_{Non-CB}}, \quad (20)$$

where $t_{p,\,CB}$ corresponds to the number of cells correctly classified as CB, $t_{p,\,Non-CB}$ corresponds to the number of cells correctly classified as Non-CB, n_{CB} is the total number of classified CB cells and n_{Non-CB} is the total number of classified Non-CB cells.

Table I demonstrates the results from the COB when no noise removal is applied. Table II on the other hand demonstrates the results of classification using COB when only standardization was applied, whereas in Table II, the corresponding results after applying median filtering are demonstrated. Tables IV through VI show the corresponding results for the CLEM approach. The results in all six tables correspond to the average performance of each iteration during which the images were divided into training and testing sets with 80% - 20% ratio.

In Table VII the classification results of the other methods are presented for comparison purposes. In case of the methods CLPLDS, CLDA, Lasso, CFC, COB and CLEM, which are all supervised methods, the best classification results achieved during ten testing iterations are presented. It should be noted that the particular pre-processing steps, which improves the performance, are part our proposed methods. Therefore, no preprocessing was applied to the images in case of all the implemented methods. Furthermore, it is interesting to see how the classification depends on the number of terms in the CB-basis or Non-CB-basis, in COB, CLDA and CLPLDS. In Figure 6 we show how recall changes depending on the number of terms (up to a hundred) in the two bases.

Finally, we examined the consistency of COB and CLEM. Consistency in our case is apprehended as the degree of agreement when applying a classification method in different images of a cell, created based on different points inside its body (see Figure 7a). We examined the consistency of our methods in 80 different cells (40 CB and 40 Non-CB). Knowing the optimal center-point of each cell, marked by the pathologist, we tested the methods in images of the cell created based on the closest to this, marking points. Closeness is defined based on a radius around the original center of the cell (see Figure 7b). Taking into account points which lie up to a radius of 7 pixels away from the original center (in Figure 7b, $D_{max} = 7$), we ended up with 224 different images of each cell (8 created based on the points that lie 1 pixel away, 16 based on pixels lie 2 pixels away, 24 based on point that lie 3 pixels away etc).

The results in the consistency are shown in Figure 8. These results were obtained after applying all the necessary pre-processing steps (noise removal and standardization of the images), as these led to the best results during training COB and CLEM (see Tables III and VI). Out of the 40 tested original CB images, 78% and 88% were classified correctly in case of COB and CLEM respectively. The classification results are reasonable considering that 40 images represent quite a small sample (19%) of the total amount of cell images used to derive the results in all the presented Tables.

On the other hand, the classification results in case of Non-CB were slightly higher (90% and 85% for COB and CLEM respectively). Furthermore, one can notice that both in the cases of CB, as well as Non-CB the classification accuracy is always above the threshold of 50%, even for 7 pixels away from the original center of the cell. On the contrary, it remains consistent (almost equal to the accuracy in the original image) for up to 2 pixels away from the center of the cell. Finally, the significant drop in accuracy in the Figure 8 of Non-CB cells in large distances from the cell's center (radius of 6 and 7) can be reasonably attributed to the small size of a Non-CB. Due to that, there is a high chance that 6 or 7 pixels away from the original center of the cell are not part of its body anymore.

Based on the results in Figure 8, we thought of examining the accuracy of our methods, when taken into account the classification results of images of the cell created by the closest (up to 7 radius distance) to original center pixels. In Figure 9 we see that, even considering the results of points lying 3 pixels away from the center, accuracy still remains high and close to the one of the original cell image, both for when applying COB or CLEM. Moreover, this holds for both CB and Non-CB cell.

V. Discussion

The results in Tables I-VI indicate the effectiveness and reproducibility of our proposed methods for the CB and Non-CB cell classification. They are effective because of high recall and precision rates. They are reproducible because changing the training and test sets result in similar performances. We observed that the subspaces spanned by PCA of each cell class did not change drastically for different training sets, as this would result in low precision and recall results in some iterations. Therefore, High Dimensional, Low Sample Size (HDLSS) did not have a negative impact for this particular problem. These tables also demonstrate the usefulness of the steps to remove the noise existed in the images (i.e. standardization and median filtering). Preprocessing the input image to COB and CLEM with the median filtering increased the average overall accuracy rate by 11.78% and 18.37% while reducing the variance by 3.44% and 3.03%, respectively. These results justify our initial hypothesis that FL images suffer from noise due to inconsistencies in the preparation phase. Therefore, noise removal is necessary. Although we used median filtering for this study, other non-linear filters with similar characteristics such as anisotropic diffusion filter (ADF) can also be considered. ADF has the advantage of preserving the boundaries of the cells in the image, however, this kind of filter is computationally expensive.

Time efficiency is very important in this application as the pathologist has to examine huge images for every individual patient. Time efficiency is another quality of both COB and CLEM, since only 750 ms and 900 ms are required for the classification of an unknown cell respectively. It should be noted that *Kyttaro* was developed in Matlab ® (Mathworks, Natick, MA) and it was not optimized for efficiency.

The results in Table VII indicate that our methods outperform some of the commonly used or previously developed methods. Those methods use only features extracted from the image while our methods utilize the whole images in the decision. Moreover, the results show the importance of dimensionality reduction to reveal the most discriminative features. As

expected, the methods that used this kind of approach (CLPLDS, CLDA, COB and CLEM) showed the four best results among all the methods.

Furthermore, the results from Figure 8 indicate the consistency of our two proposed methods with some tolerance to the selection of the cell location by the pathologist. These results show that the pathologists don't need to point to the exact same location to receive the same results. Not every pathologist would identify the same exact pixel in the HPF FL image as the center of that cell. In fact, as discussed later, inter-variability among pathologists is quite common. In addition to that, the shape of the cells is not uniform. Therefore based on the results in Figure 9, we concluded that it is better to take into account not only the marking of each pathologist, but also the points around it in a radius to 3 pixels away, without risking on the accuracy of the result. Experimental results indicate that CLEM method is very susceptible to the values of its parameters. Therefore, the low recall results for Non-CB cells may be due to non-optimal selection of its parameters. Optimization of the parameters of different methods is beyond the scope of this study. In our future work, we will optimize the parameters of our system and validate them on a larger dataset using some optimization techniques [46].

In order to demonstrate the usefulness of the system, Belkacem-Boussaid et al., used the input of two experienced pathologists to complete two different experiments on the test set data [12]. In the first experiment the pathologists graded the images without the assistance of our proposed CAD algorithms, whereas in the second experiment, they displayed the results of the designed CAD on the same data set and asked them if either they agreed with the results or not. In case there is a disagreement of judgment between the pathologist and the CAD algorithms, the pathologist were free to change the grading. It was also noted that these experiments were administered on different days for each pathologist. These experiments revealed inter- and intra-reader variability between the grading of the two pathologists. They noticed that the pathologists are not in concurrent in all cases. Inter- and intra-reader variability errors are introduced during their subjective reading. This study revealed that the intra- and inter-reader variability can be more than 25%.

VI. Conclusions

In this paper, we developed a new quantitative methodology to classify cells in a FL image into one of two categories of cells (CB versus Non-CB) using linear and non-linear dimensionality reduction.

The proposed methods could be proven to be a useful diagnostic tool that facilitates pathologists in accurate and reproducible FL grading. Moreover, *Kyttaro* could be beneficial to inexperienced pathologists by retrieving the most similar cells to the ones they have chosen to examine. The proposed methods also incorporate the features mentioned by an expert pathologist while performing classification. The performance of the developed methods was compared with some commonly used and previously developed methods. The current overall classification rates (99.22% ± 0.75% and 99.07% ± 1.53% obtained by COB and CLEM classifier, respectively) are very encouraging considering inter and intra-reader variability of the pathologists which is over 25%.

As a future work, the presented methods will be further refined, so that the selection of the cells in *Kyttaro* will be done automatically, instead of a manual selection by a user. This will be done by integrating *Kyttaro* with previously developed algorithms for the detection of cells from FL images [8, 10-14]. This integration needs to be carried out in such a way to take advantage of the common features of both systems to improve the classification accuracy and efficiency. Moreover, we will further develop an overall automated system for the quantitative analysis of the pathological images of FL, by incorporating the classification methods proposed in this study. Our vision is to translate these efficient classification methods to clinical use. Towards this vision, computer-assisted systems will be regularly used in clinical practice for cancer grading and prognosis. Such a system would have whole-slide FL images as input and the analysis of the images will be done in different scales, starting from a coarse level, in which follicles need to be segmented. Then, moving to a finer scale, HPF images, created around these follicle regions, will be analyzed after automatically detecting the cells in them. After the identification of the cells, the region around each cell would be encoded by one of our proposed methods in order to assess whether the cell belongs to a CB class or a Non-CB class. Depending on the number of CB cells computed in the image, a grade for whole image would be assigned. We are in the process of setting up a clinical trial to assess the effectiveness of this system.

Acknowledgments

The project described was supported in part by Award Number R01CA134451 (PI: Gurcan) from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute, or the National Institutes of Health. This project was also supported by the European Union project MIRACLE.

References

1. Swerdlow, SH.; Campo, E.; Harris, NL.; Jaffe, ES.; Pileri, SA.; Stein, H.; Thiele, J.; Vardiman, JW. WHO Classification of Tumours of Haemaopoietic and Lymphoid Tissues. Fourth Edition. Lyon: World Health Organization; 2008. p. 439
2. Jaffe ES, Harris NL, Stein H, Vardiman JW. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. Annals of Oncology. 2001; 3(3):490–491.
3. Fu WJ. Penalized Regressions: The Bridge Versus the Lasso. Journal of Computational and Graphical Statistics. 1998; 7(3):397–416.
4. Metter GE, Nathwani BN, Burke JS, Winberg CD, Mann RB, Barcos M, Kjeldsberg CR, Whitcomb CC, Dixon DO, Miller TP. Morphological sub-classification of follicular lymphoma: Variability of diagnoses among hematopathologists, a collaborative study between the repository center and pathology panel for lymphoma clinical studies. Journal of Clinical Oncology. 1985; 3:25–38. [PubMed: 3965631]
5. Dick F, VanLier S, Banks P, Frizzera G, Witrak G, Gibson R, Everett G, Schuman L, Isacson P, O'Conor G, Cantor K, Blattner W, Blair Aaron. Use of the working formulation for non-Hodgkin's lymphoma in epidemiological studies: Agreement between reported diagnoses and a panel of experienced pathologists. Journal of National Cancer Institute. 1987; 78:1137–1144.
6. Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D' Orsi C, Berns EA, Cutter G, Hendrick RE, Barlow WE, Elmore JG. Influence of computer-aided detection on performance of screening mammography. The New England Journal of Medicine. 2007; 356(14):1399–1409. [PubMed: 17409321]

7. Wu N, Gamsu G, Czum J, Held B, Thakur R, Nicola G. Detection of small pulmonary nodules using direct digital radiography and picture archiving and communication systems. J Thorac Imaging. 2006; 21(1):27–31. [PubMed: 16538152]
8. Sertel O, Kong J, Catalyurek UV, Lozanski G, Saltz J, Gurcan MN. Histopathological image analysis using model-based intermediate representations and color texture: Follicular lymphoma grading. The Journal of Signal Processing Systems. 2009; 55:169–183.
9. Cooper L, Sertel O, Kong J, Lozanski G, Huang K, Gurcan MN. Feature-Based Registration of Histopathology Images with Different Stains: An Application for Computerized Follicular Lymphoma Prognosis. Computer Methods and Programs in Biomedicine. 2009; 96(3):182–192. [PubMed: 19487043]
10. Sertel O, Kong J, Lozanski G, Catalyurek UV, Saltz J, Joel H, Gurcan MN. Computerized microscopic image analysis of follicular lymphoma. Proceedings of the SPIE Medical Imaging. 2008; 6915 691535.
11. Samsi SS, Krishnamurthy AK, Groseclose M, Caprioli RM, Lozanski G, Gurcan MN. Imaging Mass Spectrometry Analysis for Follicular Lymphoma Grading. Proceedings of IEEE Engineering Medicine & Biology Society. 2009; 2009:6969–6972.
12. Belkacem-Boussaid K, Sertel O, Lozanski G, Ahana'aah A, Gurcan MN. Extraction of color features in the spectral domain to recognize centroblasts in histopathology. Proceedings of IEEE Engineering Medicine & Biology Society. 2009; 2009:3685–3688.
13. Sertel O, Lozanski G, Shana'ah A, Gurcan MN. Computer-Aided Detection of Centroblasts for Follicular Lymphoma Grading Using Adaptive Likelihood-Based Cell Segmentation. IEEE Transaction on Biomedical Engineering. 2010; 57(10):2613–2616.
14. Oztan B, Kong H, Gurcan MN, Yener B. Follicular Lymphoma Grading using Cell-Graphs and Multi-Scale Feature Analysis. Proceedings of the SPIE Medical Imaging. 2012; 8315:9. 831516.
15. Samsi S, Lozanski G, Shana'ah A, Krishanmurthy A, Gurcan MN. Detection of Follicles From IHC-Stained Slides of Follicular Lymphoma Using Iterative Watershed. IEEE Transaction on Biomedical Engineering. 2010; 57(10):2609–2612.
16. Weind KL, Maier CF, Rutt BK, Moussa M. Invasive carcinomas and fibroadenomas of the breast: Comparison of microvessel distributions–implications for imaging modalities. Radiology. 1998; 208:477–4783. [PubMed: 9680579]
17. Bartels PH, Thompson D, Bibbo M, Weber JE. Bayesian belief networks in quantitative histopathology. Analytical And Quantitative Cytology and Histology. 1992; 14:459–473. [PubMed: 1292445]
18. Hamilton PW, Anderson N, Bartels PH, Thompson D. Expert system support using Bayesian belief networks in the diagnosis of fine needle aspiration biopsy specimens of the breast. Journal of Clinical. Pathology. 1994; 47:329–336. [PubMed: 8027370]
19. Tabeshet A, Teverovskiy M, Pang HY, Kumar VP, Verbel D, Kotsianti A, Saidi O. Multifeature prostate cancer diagnosis and gleason grading of histological images. IEEE Transactions on Medical Imaging. 2007; 26(10):1366–1377. [PubMed: 17948727]
20. Kong J, Sertel O, Shimada H, Boyer KL, Saltz JH, Gurcan MN. Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation. Pattern Recognition. 2009; 42:1080–1092.
21. Sertel O, Kong J, Catalyurek UV, Saltz JH, Gurcan MN. Computer-aided prognosis of neuroblastoma on whole-slide images: Classification of stromal development. Pattern Recognition. 2009; 42(6):1093–1103. [PubMed: 20161324]
22. Fatakdawala H, Xu J, Basavanhally A, Bhanot G, Ganesan S, Feldman M, Tomaszewski JE, Madabhushi A. Expectation- maximization-driven geodesic active contour with overlap resolution: Application to lymphocyte segmentation on breast cancer histology. IEEE Transactions on Biomedical Engineering. 2010; 57(7):1676–1689. [PubMed: 20172780]
23. Gurcan MN, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B. Histopathological Image Analysis: A review. IEEE Reviews in Biomedical Engineering. 2009; 2:147–171. [PubMed: 20671804]
24. Gonzalez, RC.; Wintz, P. Digital Image Processing. 2. Reading: Addison-Wesley; 1987. p. 954
25. Rangayyan, RM. Biomedical Image Analysis. Boca Raton: CRC Press; 2005.

26. Tuceryan, M.; Jain, A. Texture Analysis. In: Chen, CH.; Pau, LF.; Wang, PSP., editors. The Handbook of Pattern Recognition and Computer Vision. 2. Singapore: World Scientific Publishing Co.; 1998. p. 207-248.
27. Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE. 1979; 67(5):786–804.
28. Cosatto E, Miller M, Graf HP, Meyer JS. Grading nuclear pleomorphism on histological micrographs. Proceedings of ICPR. 2008; 7:1–4.
29. Chu J, Moon I, Mun M. A Real-Time EMG Pattern Recognition System Based on Linear-Nonlinear Feature Projection for a Multifunction Myoelectric Hand. IEEE Transactions on Biomedical Engineering. 2006; 53(11):2232–2239. [PubMed: 17073328]
30. Belkin M, Niyogi P. Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering. NIPS. 2001:585–591.
31. Sundaresan A, Chellappa R. Model-Driven Segmentation of Articulating Humans in Laplacian Eigenspace. IEEE Transactions on Pattern Analysis and Machine intelligence. 2008; 30(10):1771–1785. [PubMed: 18703830]
32. D' Agostino M, Dardanoni V. What's so special about Euclidean distance? A characterization with applications to mobility and spatial voting. Soc Choice Welf. 2009; 33:211–233.
33. Roshni VS, Revanthy K. Using Mutual Information and Cross Correlation as metrics for registration of images. Journal of Theoritical and Applied Information Technology. 2008; 4(6): 474–481.
34. Kanungo T, Mount D, Netanyahu N, Piatko P. An Efficient k-Means Clustering Algorithm: Analysis and Implementation. IEEE Transactions on Pattern Analysis and Machine Iintelligence. 2002; 24(7):881–892.
35. Inaba M, Imai H, Katoh N. Experimental Results of a Randomized Clustering Algorithm. Proc 12th Ann ACM Symposium On Computational Geometry. 1996:401–402.
36. Inaba M, Katoh N, Imai H. Applications of Weighted Voronoi Diagrams and Randomization to Variance-Based-clustering. Proceedings of 10th Annual ACM Symposium on Computational Geometry. 1994:332–339.
37. Ergüt E, Yardimci Y, Mumcuoglu E, Konu O. Analysis of microarray images using FCM and K-means clustering algorithm. Proceedings of International Conference on Signal Processing. 2003:116–121.
38. Elden, L. Matrix Methods in Data Mining and Pattern Recognition, Fundamentals of Algorithms. Philadelphia: Society for Industrial and Applied Mathematics; 2007. p. 224
39. Mann RB, Berard CW. Criteria for the cytologic subclassification of follicular lymphomas: A Proposed alternative method. Hematological Oncology. 1983; 1(2):187–192. [PubMed: 6376315]
40. Parker, A.; Bain, B.; Gatter, S.; Jack, A.; Matutes, E.; Rooney, N.; Ross, F.; Wilkins, B.; Wotherspoon, A.; Ramsay, A. Best Practice in Lymphoma Diagnosis and Reporting. London: British Committee for Standards in Haematology, Royal College of Pathologists; 2012. p. 154
41. Arlot, S.; Celisse, A. A survey of cross-validation procedures for model selection. In: Yang, Y., editor. Statistics Survey. Vol. 4. American Statistical Association; 2010. p. 40-79.
42. Gurcan MN, Sahiner B, Petrick N, Chan HP, Kazerooni EA, Cascade PN, Hadjiiski LM. Lung nodule detection on thoracic computed tomography images: Preliminary evaluation of a computer-aided diagnosis system. Medical Physics. 2002; 29(11):2552–2558. [PubMed: 12462722]
43. Gurcan MN, Pan T, Sharma A, Kurc T, Oster S, Langella S, Hastings S, Siddiqui K, Siegel EL, Saltz J. GridIMAGE: A Novel use of grid computing to support interactive human and computer-assisted detection decision support. Journal of Digital Imaging. 2007; 20(2):160–171. [PubMed: 17318701]
44. Patterson E, Rayo M, Gill C, Gurcan MN. Barriers and facilitators to adoption of soft copy interpretation from the user perspective: Lessons learned from filmless radiology for slideless pathology. Journal of Pathology Informatics. 2011; 2(1):160–171.
45. Belkacem-Boussaid K, Pennell M, Lozanski G, Shana'ah A, Gurcan MN. Computer-aided classification of centroblast cells in follicular lymphoma. Analytical and Quantitative Cytology and Histology. 2010; 32(5):254–260. [PubMed: 21509147]

46. Gurcan MN, Sahiner B, Chan H-P, Hadjiiski LM, Petrick N. Selection of an Optimal Neural Network Architecture for Computer-aided Detection of Microcalcifications – Comparison of Automated Optimization Techniques. Medical Physics. 2001; 28(9):1937–1948. [PubMed: 11585225]

47. Seung HS, Lee DD. Manifold ways of perception. Science. 2000; 290(5500):2268–2269. [PubMed: 11188725]

48. Niazi MKK, Ibrahim MT, Nilsson MF, Sköld AC, Guan L, Nyström I. Robust Signal Generation and Analysis of Rat Embryonic Heart Rate in Vitro Using Laplacian Eigenmaps and Empirical Mode Decomposition. Computer Analysis of Images and Patterns Lecture Notes in Computer Science. 2011; 6855:523–530.

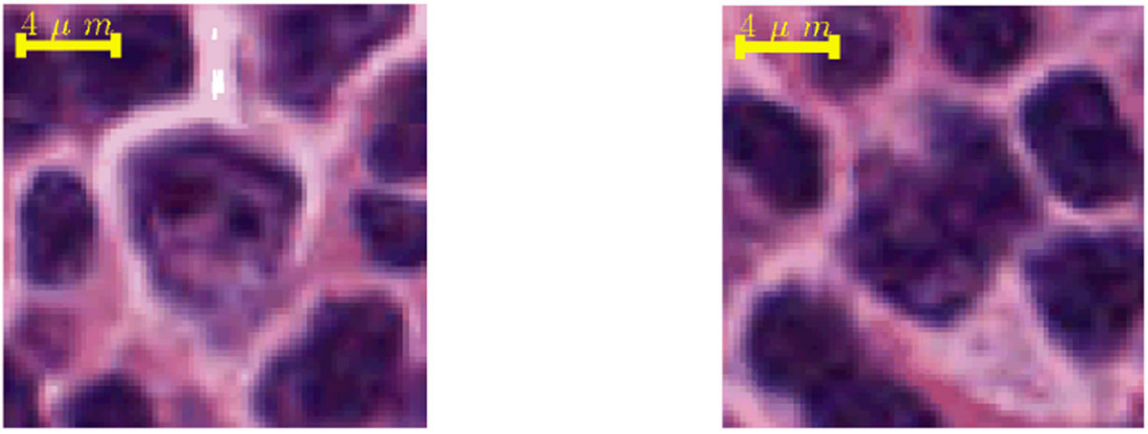

Figure 1.

Images of a CB cell (left image) and Non-CB cells (right image). The scanner's resolution at 40X magnification is 0.25 μm/pixel, therefore the yellow lines indicate a physical length of 4 μm in the tissue.

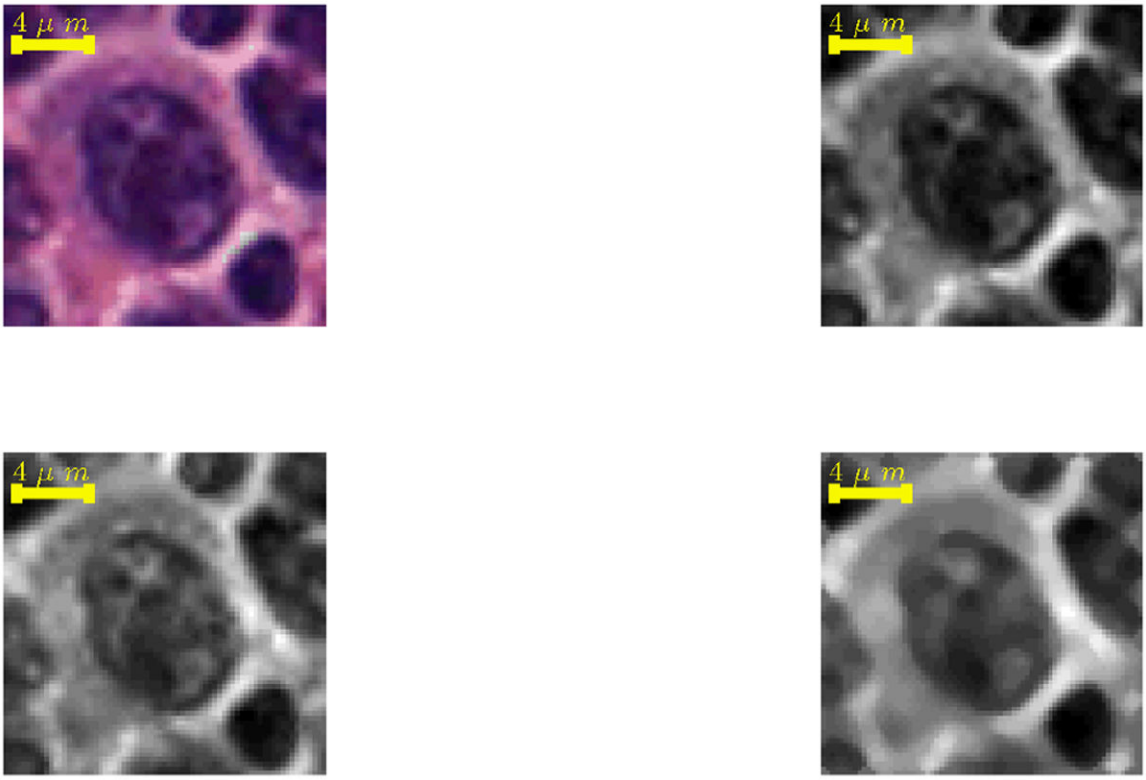

Figure 2.
RGB image of a CB cell (top left image), its intensity representation in gray scale (top right image), the gray scale image after applying standardization (bottom left image) and the filtered with median filter the same gray scale standardized image (bottom right).

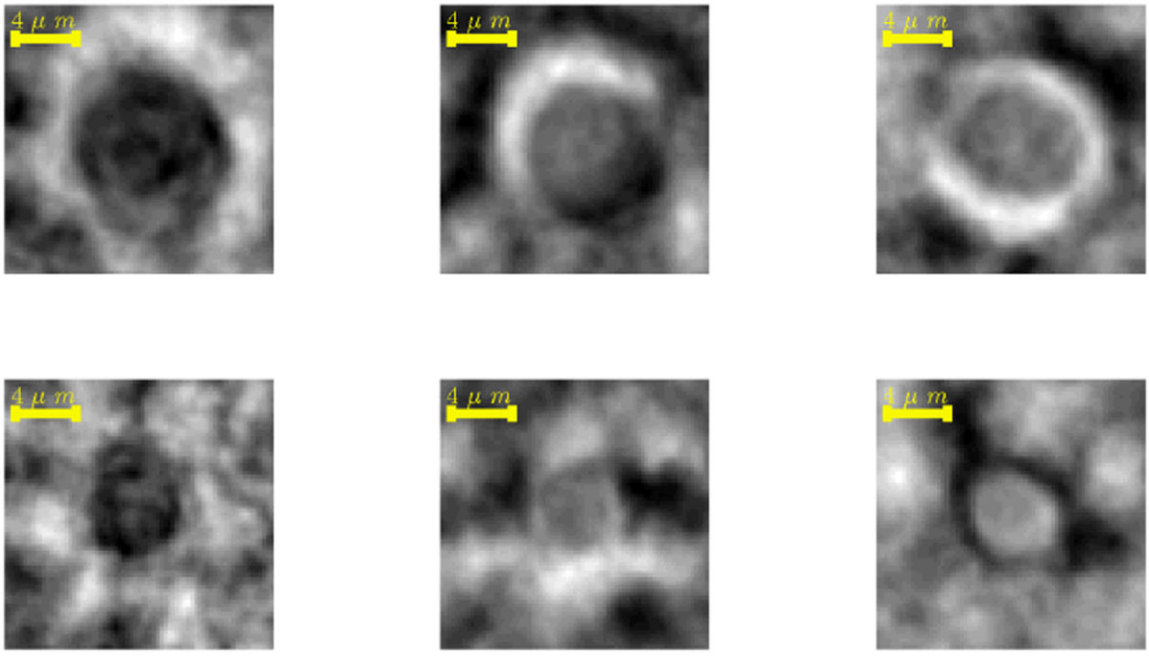

Figure 3.

Grayscale version of the first three singular images in the “image space of CB” (top three images) and the first three singular images in the “image space of Non-CB.” (bottom three images).

Figure 4.

The cell images (CB image cells shown in a hue of blue, whereas Non-CB image cells shown in a hue of red) that lie on the three dimensional manifold (it is consisted of the three first eigenvectors) and the tested image (shown in black) which resides on the top left of the manifold. This specific tested image cell would be classified as a CB image cell since it resides closer to CB image cells.

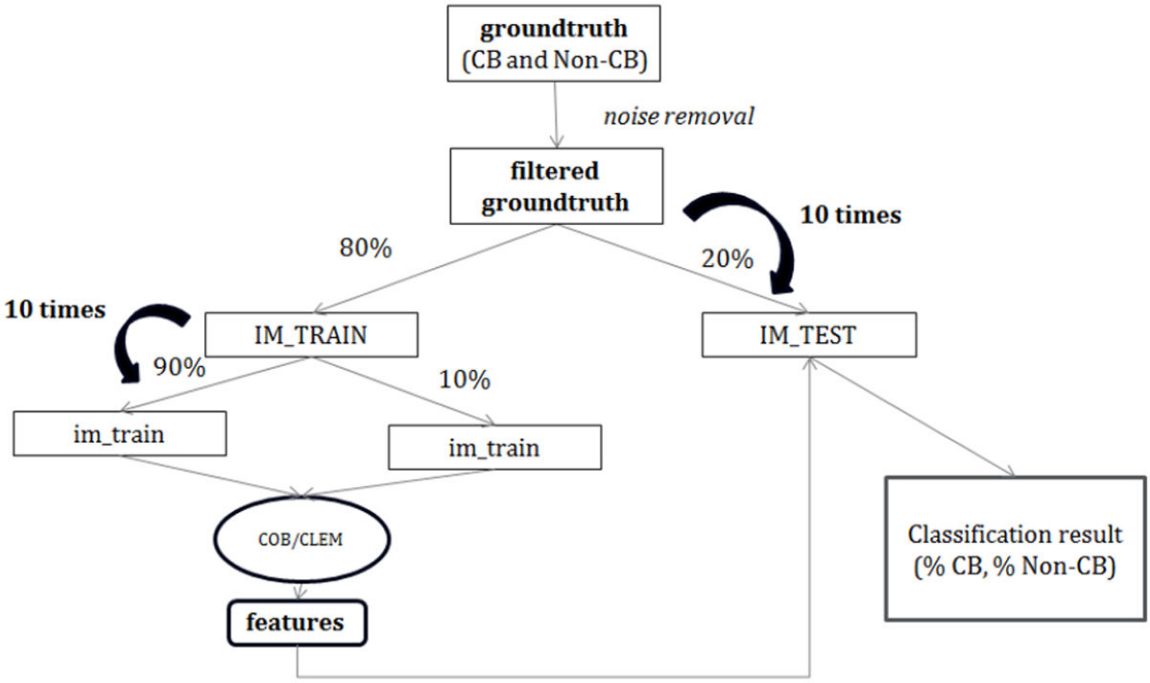

Figure 5.

Flow chart of the proposed scheme for classification of tested image cells into CB and Non-CB classes. The classification method used was either COB or CLEM, both extracting discriminatory features used in the classification of the tested image cells.

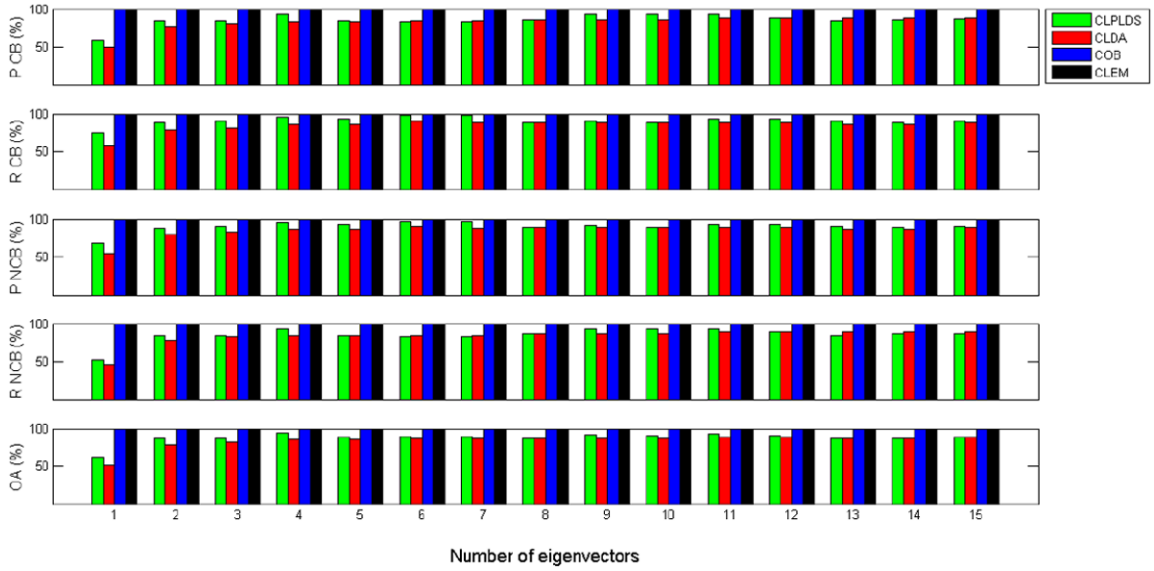

Figure 6.

Results of CB precision (top), CB recall (second from top), Non-CB precision (third from top), Non-CB recall (fourth from top), as well as the overall accuracy results (bottom) relative to the number of eigenvectors (or singular images in case of COB) used in the four supervised classification methods that use dimensionality reduction. Here we plot the results of only the first 15 eigenvector, since the rest do not add any discriminative information of the data. Results of methods that use linear dimensionality reduction (CLPLRDS, CLDA and COB) are shown in green, red, and blue respectively, whereas for CLEM (non-linear dimenionality reduction) results are in black.

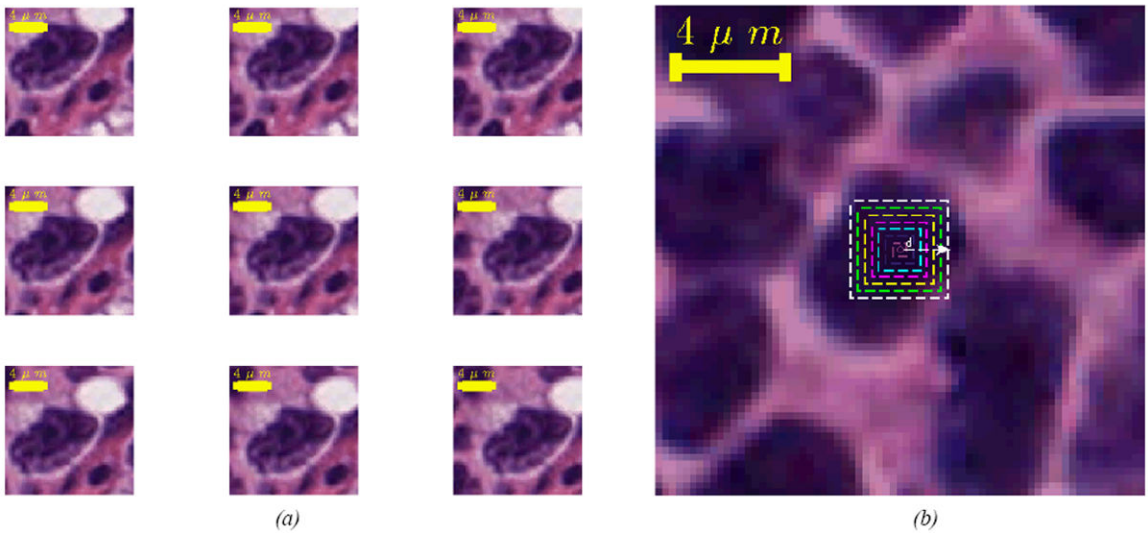

Figure 7.

(a) Images of a CB cell created based on different points inside its body. The middle image is the image created based on the pathologist's marking. (b) Image of a cell used on the examination of COB and CLEM's consistency. New images of the cell were created based on pixels that lie on the colorful rectangulars shown in the figure (created based on a distance d from the cell's center, shown with a white arrow).

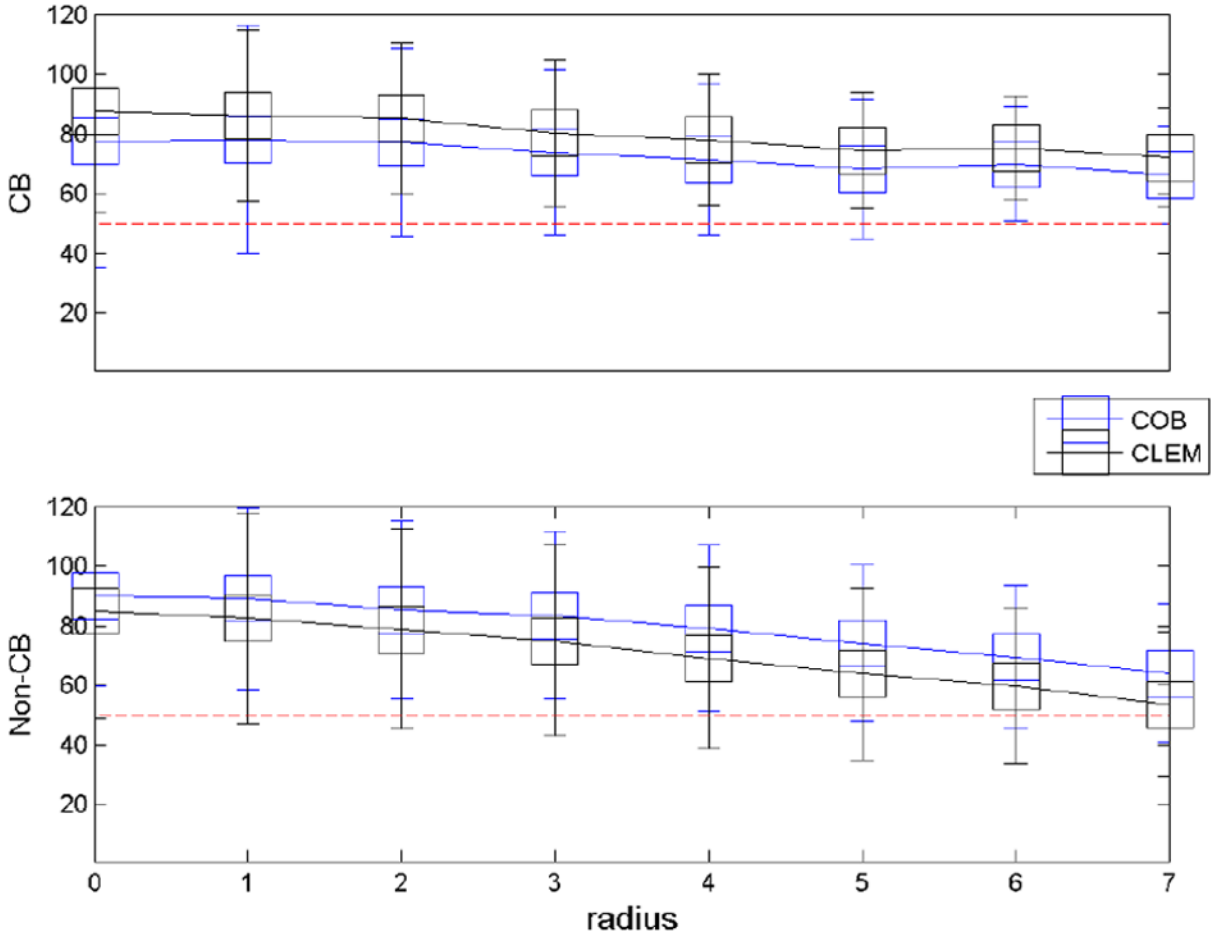

Figure 8.

Average results in consistency of COB (blue) and CLEM (blck) for 40 CB cells (top) and 40 Non-CB (middle). Results in consistency are nothing more than the classification accuracy of the new images of the cells, created based on pixels that lie radius pixels away from the original center of the cell (0 point in x-axis). The bars represent the average results and their range represents the standard deviation from this mean value.

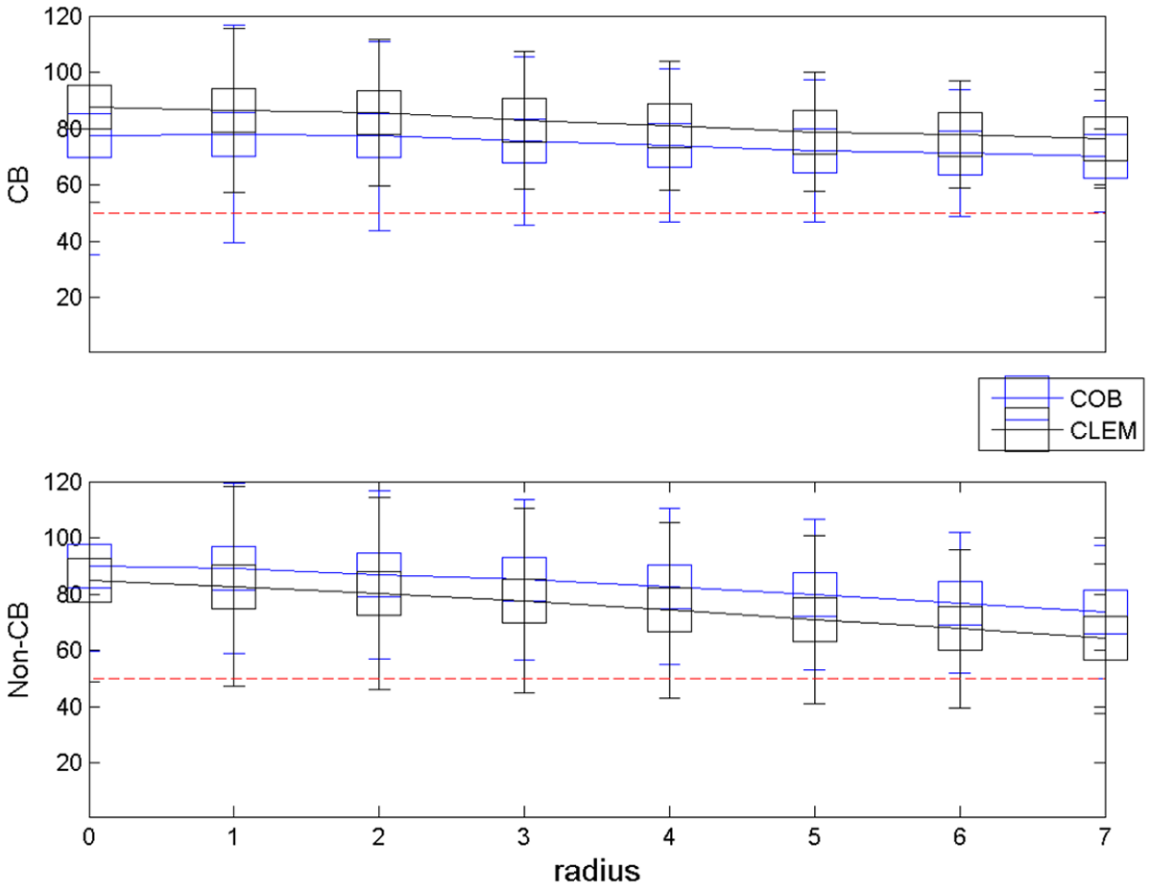

Figure 9.

Results of classification's accuracy in COB (blue) and CLEM (black) for 40 CB (top plot) and Non-CB (bottom plot), when taking into account the results of the cell's images created based on the close to the cells' center pixels in various radius away from it.

Table I

Testing performance of COB without standardization and noise removal

Set	CB		NCB		OA
	P	R	P	R	
1	79.07	89.07	80.85	80.85	**80.00**
2	84.09	86.05	86.96	85.11	**85.56**
3	95.12	90.70	91.84	95.75	**93.33**
4	92.50	86.05	88.00	93.62	**90.00**
5	84.21	74.42	78.85	87.23	**81.11**
6	90.48	88.37	89.58	91.49	**90.00**
7	94.60	81.40	84.91	95.75	**88.89**
8	90.24	86.05	87.76	91.49	**88.89**
9	92.50	86.05	88.00	93.62	**90.00**
10	80.39	95.35	94.87	78.72	**86.67**
mean	88.32	85.35	87.16	89.36	**87.44**
std	5.89	5.91	4.75	6.10	**4.19**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of COB classifier in case of standardizing the cell images. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Table II

Testing performance of COB after standardization

Set	CB		NCB		OA
	P	R	P	R	
1	95.12	90.70	91.84	95.75	**93.33**
2	84.09	86.05	84.96	85.10	**85.56**
3	91.11	95.35	95.56	91.49	**93.33**
4	88.64	90.70	91.30	89.36	**90.00**
5	97.37	86.05	88.46	97.87	**92.22**
6	86.67	90.70	91.11	87.23	**88.89**
7	86.67	90.70	91.11	87.23	**88.89**
8	95.23	93.02	93.75	95.75	**94.44**
9	94.87	86.05	88.24	95.75	**91.11**
10	90.00	83.72	86.00	91.49	**87.78**
mean	90.98	89.30	90.43	91.70	**90.56**
std	4.50	3.67	3.01	4.42	**2.83**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of COB classifier in case of standardizing the cell images. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Table III

Testing performance of COB after standardization and noise removal

Set	CB		NCB		OA
	P	R	P	R	
1	100.00	95.35	95.92	100.00	**97.78**
2	100.00	97.67	97.92	100.00	**98.89**
3	100.00	97.67	97.92	100.00	**98.89**
4	100.00	100.00	100.00	100.00	**100.00**
5	100.00	100.00	100.00	100.00	**100.00**
6	100.00	100.00	100.00	100.00	**100.00**
7	100.00	97.67	97.92	100.00	**98.89**
8	100.00	100.00	100.00	100.00	**100.00**
9	100.00	97.67	97.92	100.00	**98.89**
10	100.00	97.67	97.92	100.00	**98.89**
mean	100.00	98.37	98.55	100.00	**99.22**
std	0.00	1.57	1.39	0.00	**0.75**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of COB classifier in case of standardizing the cell images and applying median filtering on top. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Table IV

Testing performance of CLEM without standardization and noise removal

Set	CB		NCB		OA
	P	R	P	R	
1	78.26	83.72	82.50	76.74	**80.23**
2	75.00	76.74	76.19	74.42	**75.58**
3	82.35	97.67	97.14	79.07	**88.37**
4	87.18	79.07	80.85	88.37	**83.72**
5	78.72	86.05	84.62	76.74	**81.40**
6	77.78	97.67	96.88	72.09	**84.89**
7	81.83	83.72	83.33	81.40	**82.56**
8	88.23	69.77	75.00	90.70	**80.23**
9	81.08	69.77	73.47	83.72	**76.74**
10	71.74	76.74	75.00	69.77	**73.26**
mean	80.22	82.09	82.50	79.30	**80.70**
std	5.07	9.87	8.58	6.80	**4.56**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of COB classifier in case of standardizing the cell images. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Cytometry A. Author manuscript; available in PMC 2015 March 01.

Table V

Testing performance of CLEM after standardization

Set	CB		NCB		OA
	P	R	P	R	
1	87.50	81.40	82.61	88.37	**84.89**
2	87.50	97.68	97.37	86.05	**91.86**
3	82.61	88.37	87.50	81.40	**84.89**
4	85.37	81.40	82.22	86.05	**83.72**
5	87.50	81.40	82.61	88.37	**84.89**
6	90.48	88.37	88.64	90.70	**89.54**
7	81.40	81.40	81.40	81.40	**81.40**
8	100.00	76.74	81.13	100.00	**88.37**
9	82.93	79.07	80.00	83.72	**81.40**
10	85.00	79.07	80.44	86.05	**82.56**
mean	87.03	83.49	84.39	87.21	**85.35**
std	5.33	6.25	5.39	5.40	**3.52**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of CLEM classifier in case of standardizing the cell images. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Table VI

Testing performance of CLEM after standardization and noise removal

Set	CB		NCB		OA
	P	R	P	R	
1	100.00	100.00	100.00	100.00	**100.00**
2	95.56	100.00	100.00	95.35	**97.67**
3	97.73	100.00	100.00	97.67	**98.84**
4	100.00	100.00	100.00	100.00	**100.00**
5	97.56	93.02	93.33	97.67	**95.35**
6	100.00	100.00	100.00	100.00	**100.00**
7	100.00	97.67	97.73	100.00	**98.84**
8	100.00	100.00	100.00	100.00	**100.00**
9	100.00	100.00	100.00	100.00	**100.00**
10	100.00	100.00	100.00	100.00	**100.00**
mean	99.08	99.07	99.11	99.07	**99.07**
std	1.58	2.25	2.15	1.63	**1.53**

Results from the 80% - 20 % testing procedure followed for the evaluation of the performance of CLEM classifier in case of standardizing the cell images and applying median filtering on top. Ten iterations were performed. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy.

Table VII

Classification performance of previous and our proposed methods

Classifier	Performance (%)				
	CB		*NCB*		
	P	***R***	***P***	***R***	***OA***
ED	85.5	60.5	71.5	90.5	**76**
CC	86	87	88	87	**87**
MI	83	23	58	95.5	**61**
k-means	53	60	59	52.5	**56**
FCM	41	41	46	46	**43.5**
CLPLDS	88.5	91	91	89	**90**
CLDA	91	95.5	95.5	91.5	**93**
Lasso	93	60.5	72.5	96	**79**
CFC	87	95	95	87	**91**
COB	*100*	*100*	*100*	*100*	***100***
CLEM	*100*	*100*	*100*	*100*	***100***

Classification results from the various methods examined in this study. Each classifier uses a different method. ED holds for Euclidean distance, CC holds for cross-correlation, MI holds for Mutual Information, FCM holds for Fuzzy C-means, CLPLDS holds for Classification by Linear Projection in Low Dimensional Space, CLDA holds for Classification using Linear Discriminant Analysis, CFC holds for Color Feature-based Classifier, COB for Classification using Orthogonal Bases and CLEM holds for Classification based on Laplacian Eigenmaps. P corresponds to precision, R corresponds to recall and OA corresponds to overall accuracy