Lymphoma images analysis using morphological and non-morphological descriptors for classification

Marcelo Zanchetta do Nascimento [a,c,*], Alessandro Santana Martins [b],
Thaína Aparecida Azevedo Tosta [c], Leandro Alves Neves [d]

[a] *UFU - FACOM, av. João Neves de Ávila 2121, Bl.B, Uberlândia-MG 38400-902, Brazil*
[b] *IFTM, r. Belarmino Vilela Junqueira S/N, Ituiutaba-MG 38305-200, Brazil*
[c] *UFABC - CMCC, av. dos Estados 5001, Bl.B, St. André-SP 09210-580, Brazil*
[d] *UNESP - DCCE, r. Cristóvão Colombo 2265, S.J. Rio Preto-SP 15054-000, Brazil*

ARTICLE INFO

ABSTRACT

Mantle cell lymphoma, follicular lymphoma and chronic lymphocytic leukemia are the principle subtypes of the non-Hodgkin lymphomas. The diversity of clinical presentations and the histopathological features have made diagnosis of such lymphomas a complex task for specialists. Computer aided diagnosis systems employ computational vision and image processing techniques, which contribute toward the detection, diagnosis and prognosis of digitised images of histological samples. Studies aimed at improving the understanding of morphological and non-morphological features for classification of lymphoma remain a challenge in this area. This work presents a new approach for the classification of information extracted from morphological and non-morphological features of nuclei of lymphoma images. We extracted data of the RGB model of the image and employed contrast limited adaptive histogram equalisation and 2D order-statistics filter methods to enhance the contrast and remove noise. The regions of interest were identified by the global thresholding method. The flood-fill and watershed techniques were used to remove the small false positive regions. The area, extent, perimeter, convex area, solidity, eccentricity, equivalent diameter, minor axis and major axis measurements were computed for the regions detected in the nuclei. In the feature selection stage, we applied the ANOVA, Ansari-Bradley and Wilcoxon rank sum methods. Finally, we employed the polynomial, support vector machine, random forest and decision tree classifiers in order to evaluate the performance of the proposed approach. The non-morphological features achieved higher AUC and AC values for all cases: the obtained rates were between 95% and 100%. These results are relevant, especially when we consider the difficulties of clinical practice in distinguishing the studied groups. The proposed approach is useful as an automated protocol for the diagnosis of lymphoma histological tissue.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Investigation through pathological anatomy can be performed on tissue samples stained with hematoxylin and eosin (H&E). This task is essential in order to allow for the identification of the type of lesion, the monitoring of disease progression and the choice of treatment for the patient [1]. Specialists currently use this strategy to investigate various cancer types.

Lymphomas are neoplasms that originate in the lymphatic system and represent one of the most common types of cancer found in the World population. There are more than 38 subtypes of lymphomas [1,2] divided into Hodgkin lymphomas (HL) and the non-Hodgkin lymphomas (NHL), which are differentiated by morphological, genetic and clinical information [3]. According to published statistics, between 2,530 and 10,180 new cases were estimated in Brazil, respectively, for HL and NHL between 2018 and 2019 [4]. In the United States, 8500 new cases of HL and 74,680 new cases of NHL are expected for the year 2018 [5]. In worldwide, more than 65 and 385 thousand new cases of HL and NHL, respectively, were estimated in 2012 [4]. NHL is the most common cancer of the lymphatic system with almost 90% of all cases [6].

These statistical data demonstrate the importance of new research that contributes to a more accurate diagnosis of lymphomas belonging to this class of lesion. Mantle cell lymphoma (MCL), follicular lymphoma (FL) and chronic lymphocytic leukemia (CLL) are the main subtypes of the NHL group. The diversity of clinical pre-

sentations and the histopathological features have made diagnosis a complex task for specialists [7]. Factors related to subjectivity and workload of specialists does not contribute to a more precise diagnosis. These problems have led to the development of algorithms in computational systems that can assist the specialists in reaching a more desirable result [8,9].

The computer aided diagnosis system (CAD) employs computational vision and image processing techniques, which contribute toward the detection, diagnosis and prognosis of digitised images of histological samples. These systems allow greater objectivity and less variability in the process of image analysis [10]. The techniques commonly used for tissue analysis are thresholding, region growing, graph cuts, active contour, morphological descriptors and texture features.

In the Literature, the authors [11] presented a lymphoma image from a public domain dataset that can be employed for developing and evaluating biological image analysis methods. There are several studies that proposed the combination of detection methods, feature extraction and classification to the diagnosis of these subtypes of lymphoma. There are those works that investigated the degree of severity of the FL samples [12–14]. However, a limited number of samples of several lesion subtypes was employed to evaluate the performance of the proposed system. The researchers in Orlov et al. [1] presented a classification system for histological images of the MCL, FL and CLL subtypes using the Chebyshev and Fourier wavelets for realising feature extraction and the weighted neighbours distance, radial basis functions, naive Bayes network algorithms were used for the classification of the samples. The system presented relevant performance, but the authors did not investigate the nuclei regions during the feature extraction stage and classification. The authors in [15] developed an approach for classification that consists of dividing the image into 25 blocks for the extraction of a set of 505 visual features from each block. In Nascimento et al. [16] an investigation was also made into the classification of these lesions with descriptors based on the stationary wavelet transform. In addition authors also presented approaches for quantifying and extracting features from the histological images of the lymphoma image dataset [17,18]. In these works, the authors did not employ the detection stage for nuclei before the feature extraction stage. The analysis of detected regions allows for the extraction of morphological features and this information may contribute toward results of higher relevance in the classification of the lesions. Moreover, new studies aimed at improving the understanding of morphological and non-morphological features for classification of lymphoma remain as challenges in this area.

In this paper, we present a new approach for the classification of nuclei in MCL, FL and CLL images with information extracted from morphological and non-morphological features. We extracted data of the Red channel of the RGB model of the image and employed the contrast limited adaptive histogram equalisation and 2D order-statistics filter methods to enhance the contrast and remove the noise. The detected regions were identified by a global thresholding method. The flood-fill and watershed techniques were used to remove the false positive regions and the non-morphological features were extracted from the channels of the RGB colour model and the grayscale image. The area, extent, perimeter, convex area, solidity, eccentricity, equivalent diameter, minor axis and major axis measurements were computed for the detected regions of nuclei. In the feature selection stage, we applied the ANOVA, Ansari–Bradley and Wilcoxon rank techniques. Finally, we employed the polynomial, support vector machine, random forest and decision tree classifiers to evaluate the performance of the proposed approach.

The contributions of this paper can be summarized as follows:

- It presents a novel approach for classification lymphoma subtypes employed to images with variation of contrast and illumination from a public dataset [11]. This condition is relevant for evaluating the robustness of the novel approach as presented in the study by [19];
- It shows the detection of the regions of interest with nuclei from lymphoma subtypes;
- The evaluation of morphological and non-morphological features based on statistical measurements of lymphoma lesions;
- It proposes the use of feature selection algorithms to evaluate the performance of the classifier with information obtained from morphological and non-morphological descriptors;
- It shows the usefulness of our approach when applied to the different lymphoma subtype: MCL, FL and CLL.

The paper is organised as follows. In Section 2, we detail the methodologies used in the various stages of the proposed approach. Section 3 presents the principle results and discussions, introducing the criteria for classification performance evaluation and showing the comparisons with other works. Finally, the conclusions are summarized in Section 4.

2. Methodology

The approach proposed in this work is illustrated in Fig. 1 from which one can verify the presence of steps: (1) detection of nuclei from the histological images; (2) extraction of morphological descriptors; (3) extraction of non-morphological descriptors using statistical metrics of each channel of the RGB colour model and grayscale image; (4) feature selection through statistical techniques for the generated set; (5) classification using the polynomial (PL), random forest (RF), decision tree (DT) and support vector machine (SVM) algorithms. The algorithms were developed using MATLAB® language with the help of the WEKA platform 3.6.6 for classifying the features. These experiments were performed on a notebook Intel Core i5, 8 GB RAM and 1 TB.

2.1. Image database

In this study, we used a public dataset from studies conducted by researchers from National Cancer Institute and National Institute on Aging, in the United States, available for download at [11]. A total of 30 histological slides of lymph nodes stained with H&E were used, containing 10 cases from the three cited types of lymphoma. To better represent the clinical environment, instead of the tightly controlled laboratory environment, the slides were obtained with significant variations in sectioning and staining.

The microscopic images were digitally obtained, using a light microscope (Zeiss Axioscope) with 20× objective and a color digital camera (AXio Cam MR5). Regions from each slide were digitally photographed and stored with no compression in the tif format, RBG colour model, resolution of 1388 × 1040, with quantisation of 24 bits. In total, 375 images were generated, containing 113, 140 and 122 regions of CLL, FL and MCL, respectively.

Fig. 2 shows samples of histological images employed in our experiments to evaluate the proposed approach. CLL sample is depicted in Fig. 2(a), FL sample is shown in Fig. 2(b) and MCL sample is presented in Fig. 2(c). The cases are named according to the codification defined by National Cancer Institute and National Institute on Aging.

2.2. Detection of nuclei

2.2.1. Preprocessing

Firstly, we should improve the quality of the object contrast in the histological image and make it more suitable for the application of a thresholding algorithm.
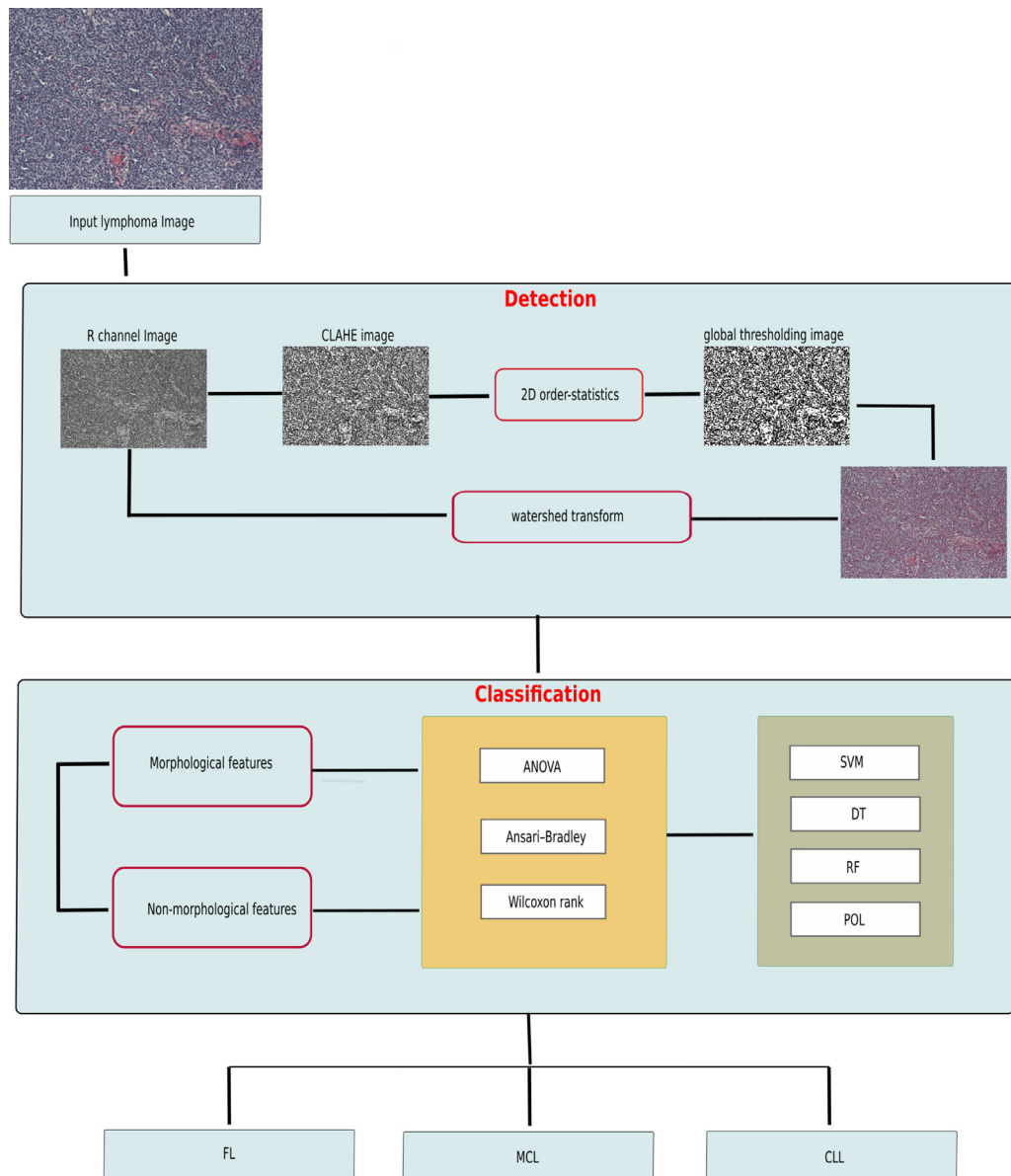
Fig. 1. Block diagram of the main stages of the proposed approach: detection of nuclei from the lymphoma images; the morphological and non-morphological descriptors; the feature selection; and the classification algorithms.
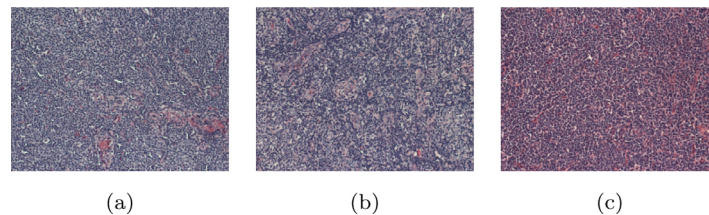

(a)   (b)   (c)

Fig. 2. Samples: (a) sj-03-2810_008; (b) sj-05-588-R1_006; (c) sj-04-4525-R4_001.

Then, we evaluate the histogram and the images of the three colour channels from the RGB model. This experiment allows for the identification of the high contrast between objects and background in an image based on the R colour channel. Therefore, we chose the R colour channel image $g(x, y)$ for applying the methods. The contrast in the R channel is due to the features of the H&E stain used in the tissue pigmentation process and hematoxylin is attracted by nucleic acids in cell nuclei, which are subsequently stained a violet color.

The contrast limited adaptive histogram equalisation (CLAHE) method allows for the rearranging of the range of intensity levels of the image improving the image contrast level. The technique uses an adaptive histogram equalisation function, which limits the contrast in homogeneous areas of the image and avoids enhancement of noise. Further details can be found at [20]. Thus, we applied this technique to the R colour channel in order to increase the differences between intensities of nuclei and the background of the image. In this processing step, we used the Matlab function
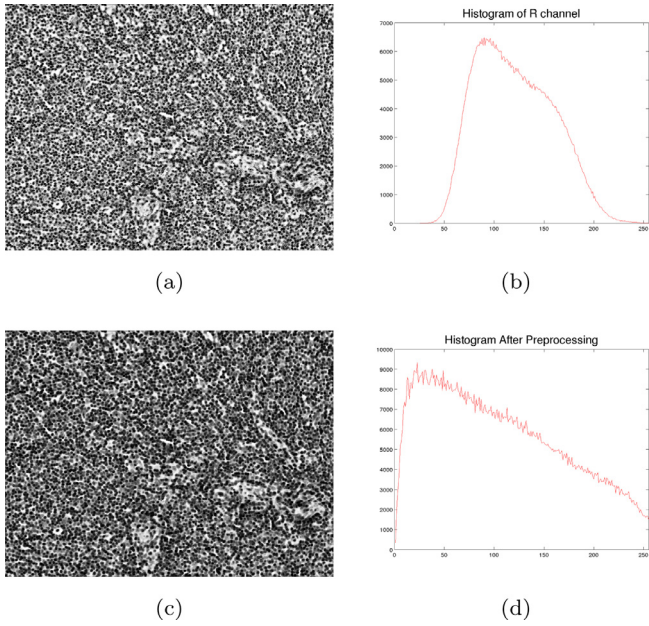
(a)   (b)

(c)   (d)

Fig. 3. Images of the R channel before (a) and after (c) applying the filters. Histogram of the R channel before (b) and after (d) applying the techniques.

`adapthisteq` and the parameters were set in a number of tiles [8,8], clip-limit of 0.01, NBins of 256, distribution uniform and alpha of 0.4. The default values are determined based on suggestion derived from [21].

After applying the CLAHE algorithm to $g(x, y)$, the 2D order-statistics filter was employed to remove discrete points and noise. This filter is a generalised version of the median filter, where a neighbourhood is defined by the non-zero pixels of the domain and each pixel in the image is replaced by the smallest orderth from among its neighbours inside this domain [22]. This process was performed by applying the Matlab function `ordfilt2`, with a 3×3 neighbourhood and order index 1, which represents the minimum value. These parameters allow removing discrete points without to degraded the cell nuclei, which have the larger size than a 3×3 mask.

Fig. 3(a)–(d) show the histogram and the R channel of the RGB model, before and after applying the techniques.

2.2.2. Thresholding

For the thresholding step, we computed a grayscale in order to perform image binarization. A global thresholding was employed for detection of these structures [23]. The threshold value was defined and set at 90. In this stage, different threshold values were tested in order to obtain the best match between the structures and background of the image. The differences between intensities of nuclei and the other structures on the image were more distinct when the threshold value was set at 90. Other threshold values can be applied, but with results that do not benefit the next steps of the proposed method.

However, there are regions that were not detected correctly due the differences in brightness level affecting the structures. Then, we applied a flood-fill operation with 4-connected neighborhoods to select the pixels located within the regions of interest that were not detected using the thresholding method [24]. This step obtained a thresholded image as defined by $f(x, y)$.

Some of the cells that remained connected during the use of global thresholding can be separated by the use of the watershed transform with 8-connected neighborhood [25]. Therefore, we applied the watershed algorithm to the R channel image [26].

Given $M_1, M_2, \ldots, M_R$ as the different regional minima of the R channel image $g(x, y)$, let $C(M_i)$ be the set of points in a catchment basin associated with the regional minimum M_i. The minimum and maximum values of $g(x, y)$ are represented by the notation *min* and *max*, respectively. Then, $T[n]$ represents the set of coordinates (s, t) of the image in which $g(s, t) < n$. Mathematically, we can represent it as

$$T[n] = \{(s, t) | g(s, t) < n\}, \tag{1}$$

where, $T[n]$ is a set of coordinates of the image $g(x, y)$ lying below the plane $g(x, y) = n$. The flooding process starts from the range of $n = min + 1$ to $n = max + 1$. Let $C_n(M_j)$ represent the set of coordinates in the catchment basin associated with the minimum M_j, which are flooded at step n. Then, $C_n(M_i)$ can be viewed as a binary image represented as

$$C_n[M_i] = C(M_i) \cap T[n]. \tag{2}$$

Then, $C[n]$ represents the union of the flooded catchment basins at the stage n:

$$C[n] = \bigcup_{i=1}^{R} C_n(M_i), \tag{3}$$

and $C[max + 1]$ is the union of all the catchment:

$$C[max + 1] = \bigcup_{i=1}^{R} C(M_i). \tag{4}$$

Then, $C[n − 1]$ is a subset of $C[n]$ and $C[n]$ is a subset of $T[n]$. Therefore $C[n − 1]$ can be represented as a subset of $T[n]$.

Next, $C[min + 1]$ is initialised with $T[min + 1]$ to find the watershed lines. The algorithm recursively finds $C[n]$ from $C[n − 1]$. The calculation of $C[n]$ from $C[n − 1]$ is as follows. Let Q be the set of connected components in $T[n]$. Then for each connected component $q \in Q[n]$ there may be three possibilities:

- $q \cap C[n − 1]$ is empty,
- $q \cap C[n − 1]$ contains one connected component of $C[n − 1]$,
- $q \cap C[n − 1]$ contains more than one connected component of $C[n − 1]$.

Further flooding would lead to the merging of these catchment basins. A one pixel thick dam can be constructed by dilating $q \cap C[n − 1]$ with a 3×3 structuring element of 1s and constraining the dilation to q.

We then perform an element-by-element multiplication of the thresholded image and the watershed image, as illustrated in Fig. 4(a) and (b), respectively. In Fig. 4 the image is shown after the application of this operation to where the connected cells were separated. The region marked in red was amplified to show the areas before and after separation of the candidate cell nuclei.

2.2.3. Post-processing

As false positive areas may remain in the image, regions with an area smaller than a fixed value represented by n should be removed. By means of the experiments conducted on a group of known abnormal images, we established $n = 30$ since the cell nuclei have the larger size than this area value. On the other hand, regions with an area value smaller than n were removed because they represent small noises. Moreover, the tests performed show that values greater than $n = 30$ degraded the regions representing cell nuclei and made the nuclei feature computation step unviable.

After identifying the groups of pixels, we applied the connected-components labelling technique to identify groups of pixels on the image. This procedure assigns unique identifiers for each group of pixels on a image [27] and allows for the individual analysis of the structures. In this study, we used the algorithm proposed by [28]. Fig. 5 shows the regions marked in red with the identified nuclei.
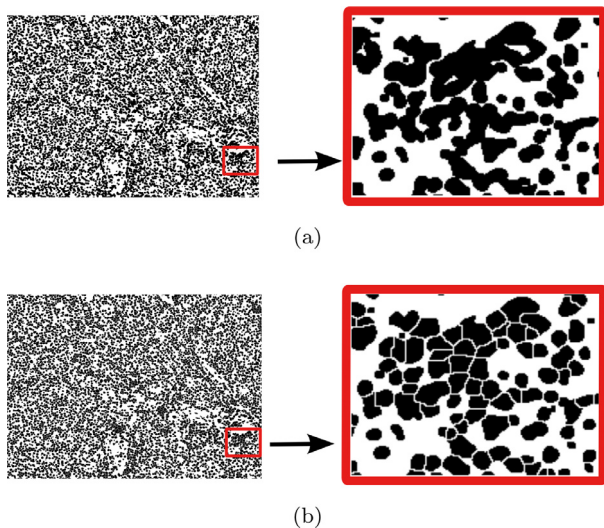
Fig. 4. Image with regions generated by application of these techniques. The marked region in red on the thresholding image was zoomed (a) and the image obtained after the separation of connected cells was zoomed (b). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
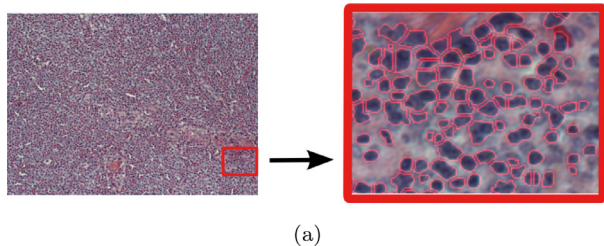

Fig. 5. Nuclei marked in red after the post-processing stage with a value of $n = 30$. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Morphological features computed from the image.

Feature	Name
mf_1	Area
mf_2	Extent
mf_3	Perimeter
mf_4	Convex area
mf_5	Solidity
mf_6	Eccentricity
mf_7	Equivalent diameter
mf_8	Minor axis
mf_9	Major axis

2.3. Nuclei feature computation

In the next step, we computed the morphological and non-morphological features from the detected nuclei regions. For morphological features we computed shape and geometric information, which are defined on Table 1 [29,30].

We calculated the information (see Table 1) for each nuclei present on the detected image. Then, the mean, median, standard deviation and mode metrics were computed for all nuclei of the histological image. The information obtained was stored in the vector.

$$Morphological = [mf_1 \cup \cdots \cup mf_{36\,features}]. \qquad (5)$$

For the non-morphological features, we obtained the information for pixel intensity values, first order statistics, from the different colour channels of the RGB model and grayscale channel. The separation of the colour channels of the histological image resulted in 4 new images with the brightness information of each separation channel of the RGB model and grayscale. Through use of the brightness level information from each colour channel, we computed the mean, median, standard deviation, minimum and maximum of pixels for each nuclei region on the image. Following this, the mean, median, standard deviation and mode metrics were obtained for each colour channel image. Then, this information was stored in a vector with 80 features of which 5 represent the metrics of brightness for each nuclei metrics of each image colour channels (RGB and grayscale).

$$Non\text{-}Morphological = [nmf_1 \cup \cdots \cup nmf_{80\,features}]. \qquad (6)$$

2.4. Feature selection

The performance of the results can be affected when there are redundant or irrelevant features in stage of classification [31]. Then, the feature selection technique removes the redundant features and retains information, which contributes to the performance of the classifier. Moreover, it avoids the so-called curse of dimensionality and over-fitting problems [32,33]. After applying the feature extraction methods, we computed the analysis of variance (ANOVA), Wilcoxon rank sum (WRS) and AnsariBradley (AB) techniques over all extracted features in order to select a relevant set of features. The tests were applied with a confidence interval of ρ=5% where the features with ρ-values smaller than 0.05 and can be used for discrimination of information.

By applying ANOVA a dataset of features can be grouped into statistically similar subgroups, except for isolated features discarded as noise [30]. In this method, the null hypothesis is that the mean for all features is the same. The alternative hypothesis is that at least two of these are different. Then, the samples are used to test and evaluate if each attribute is different from the others. If any two groups are statistically the same, both are discarded since they do not contribute to the classification step. The WRS method is a non-parametric test of the null hypothesis, where two independent populations are the same against an alternative hypothesis and these two distributions differ only with respect to the median. This test assumes that within each population under investigation the observations are independent and identically distributed. It has higher efficiency on non-normal distributions, such as a mixture of normal distributions [34]. The AB technique is a non-parametric hypothesis test that assumes that two independent populations come from same distribution, and not any alternative hypothesis where two distributions differ only in scale [34,35].

2.5. Performance of the proposed approach

The purpose of the classifier is to identify and interpret the information about the lesions based on features extracted from images [36]. In this case, the classification algorithms were employed with two groups separately (CLL-FL, CLL-MCL and FL-MCL).

Some of the main machine learning algorithms were used in this work. We chose J48 (WEKAs own version of C4.5) module, Random Forest module and SMO (WEKAs own version of support vector machine) module for implementation of the classifiers DT, RF, and SVM, respectively, in WEKA [37–39]. The parameters for each machine learning algorithm are summarized in Table 2. Moreover, we also employed PL classification algorithm to distinguish the classes for each image group of lymphoma.

We employed the 10-fold cross-validation method, where 90% of data were used for training and 10% were used for testing. The feature selection algorithms were applied to each training set in order to select the relevant features and these features were evaluated using the test set. The classifiers were trained on k-1 sets

Table 2
The machine learning parameters used for each of the different algorithms in WEKA.

Algorithms	Parameters
DT	weka.classifiers.trees.J48 -C 0.25 -M 2
RF	weka.classifiers.trees.RandomForest -I 100 -k 0 -S 1
SVM	weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V 1 -W 1 -K

and tested on the one holdout set. In these subsections, we provide details of the algorithms.

2.5.1. Decision tree

The DT algorithm is a method of inductive inference, robust to noisy data and capable of learning disjunctive expressions [40]. On the root node, the data splitting process is applied to each internal node based on a particular rule of the tree. This process is repeated until a previously specified stop criterion is reached. Each split is based on an optimal threshold value of a single feature.

There are many ways a decision tree can be structured from a set of features. The algorithms based on heuristics are proposed for the induction of decision trees. The top-down induction of a decision tree is employed as the basis for many decision tree induction algorithms. The main algorithms are ID3 [41], C4.5 [42] and CART [43].

On the WEKA, the C4.5 algorithm is used for generating the decision tree. This algorithm is an extension of the ID3 method. To select the splitting attributes, the concepts of entropy and information gain are used. In this case, for a set of samples (X), entropy is given by

$$Entropy(X) = \sum_{c=1}^{m} -p_c.log_2 p_c, \qquad (7)$$

where p_c is the proportion of examples in X associated with the cth class label and m is the total number of class labels. The entropy metric measures the impurity of a collection of examples relative to their values of the class attribute. The higher entropy values correspond to more uniformly distributed examples and improve the classification process. Using the entropy measure, the information gain of a descriptor is given by

$$Gain(X, D) = Entropy(X) - \sum_{v=1}^{T} \frac{|X_v|}{|X|}.Entropy(X_v), \qquad (8)$$

where X_v is the number of examples in the subset of X for which the descriptor D has the vth value in the domain of D, and $|X|$ is the number of examples in X. Further details for C4.5 can be found in [42].

2.5.2. Random forest

The RF algorithm uses a decision tree as a base classifier to make predictions that are built from different bootstrap samples of the training data, with random features selected in the induction tree process [37].

In the algorithm, two parameters must be determined; these are m and N. m is the number of variables to be used for each node and N is the number of trees to be created.

The method produces a large set of features that is called v. In turn, v is submitted to the algorithm, which starts by taking a random training set $I \subset v$. After acquiring I, the algorithm makes each tree grow randomly from different trunks $I' \subset I$. Each node of the tree is labelled by a number n, and the corresponding $I_n \subset I$ is split into left and right subsets I_ℓ and I_r, respectively. Noteworthy here is that $I_1 = I'$. This split is determined by a threshold t and a function $f(v_i)$, $v_i \in v$ and $i \in I_n$. Once these decision trees are obtained, the algorithm combines these in order to construct a new
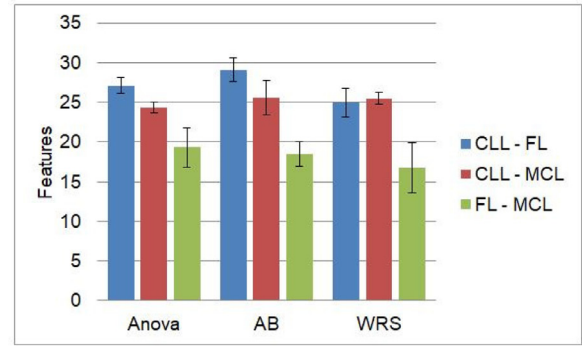

(a)
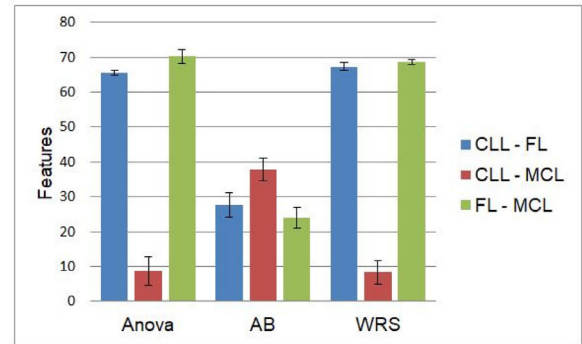
(b)
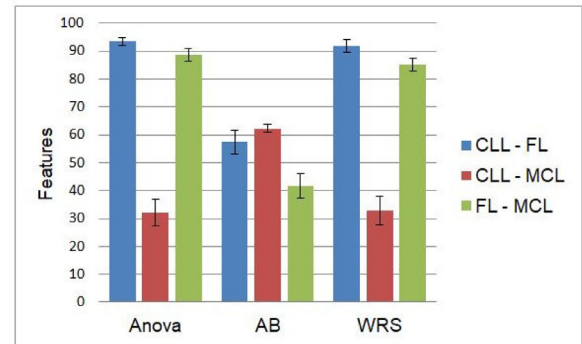
(c)

Fig. 6. Resulted of the average values of features obtained after the application of the selection algorithms for the investigated lesions: (a) morphological, (b) non-morphological and (c) a subspace that combines these two information.

tree called I^*. After construction, the forest is used to predict the estimated labels. Then, the final labelling of the group is performed using a majority vote from the forest.

2.5.3. Support vector machines

The SVM technique is based on a supervised learning algorithm that can be used to analyse data and recognise patterns in statistical classification and regression analysis [44]. Its purpose is to minimise and balance two types of data: training set and test set. Then, given a training set $\{(z_i, y_i) : z_i \in \mathbb{R}^{n_f}, y_i \in \{-1, +1\}\}_{i=1}^{n_s}$ with n_f the number of features and n_s the number of samples. Hence, in order to establish the two classes we can use the dual form to find the Lagrange multipliers $\{\alpha_i\}_{i=1}^{n_s}$ that minimise the objective
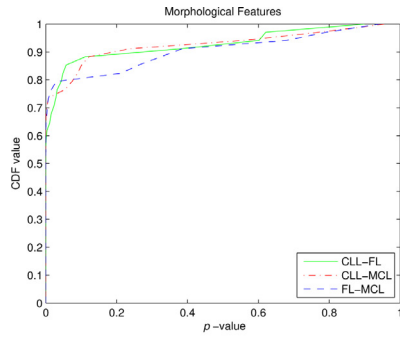
(a)
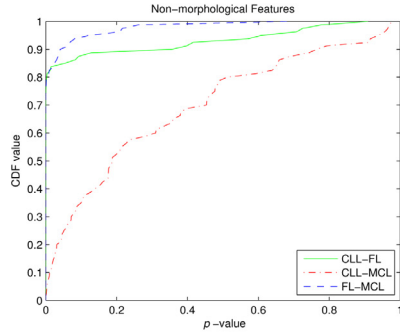
(b)
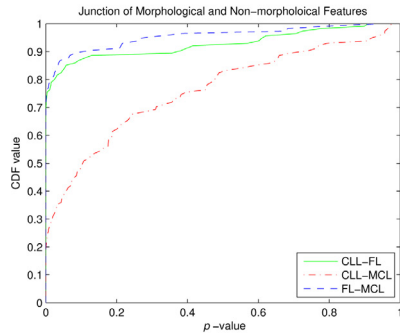
(c)

Fig. 7. Empirical cumulative distribution function of the *p*-values of the feature groups: (a) Morphological, (b) Non-morphological and (c) A subspace that combines these two types of features.

(a)

(b)

(c)

Fig. 8. The AUC values obtained with the classifiers and the ANOVA algorithm for the investigated lesion of lymphoma: (a) CLL-FL; (b) CLL-MCL and (c) FL-MCL.

function

$$Q(\alpha) = \frac{1}{2} - \sum_{i=1}^{n_s} \alpha_i + \sum_{i=1}^{n_s} \sum_{j=1}^{n_s} \alpha_i \alpha_j y_i y_j K(z_i, z_j) \tag{9}$$
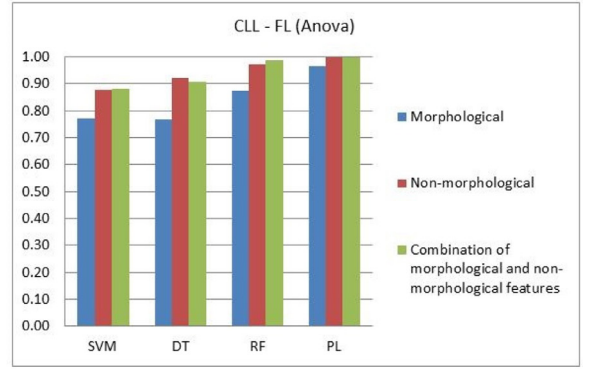
constrained to

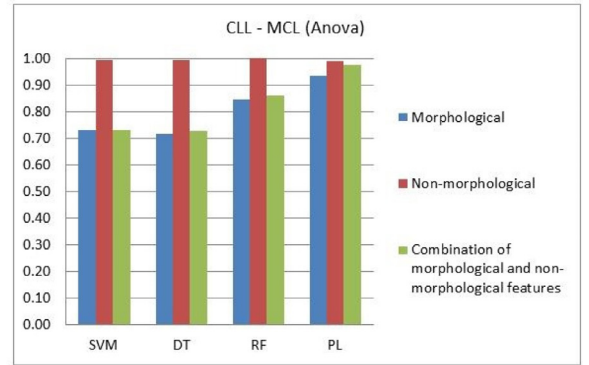$$\sum_{i=1}^{n_s} \alpha_i y_i = 0, \qquad 0 \le \alpha_i \le C, \ \forall i = 1, 2, \ldots, n_s, \tag{10}$$

where C is a positive constant given by the user, and $K(\cdot, \cdot)$ is a partially defined positive kernel function.

Let $\{\alpha_i\}_{i=1}^{n_s}$ be an optimal solution of (9) and (10). The decision function is given by

$$\beta_{SVM}(z) = b + \sum_{i=1}^{n_s} \alpha_i y_i K(z_i, z), \tag{11}$$

where b is some bias value. Here, we have used the polynomial kernel function given by

$$K(x_i, x_j) = (x_i x_j + 1)^2. \tag{12}$$

2.5.4. Polynomial

The PL classifier is a non-linear parametrized method that non-linearly expands a sequence of input vectors into a larger dimension space and identifies these vectors for a desired output sequence. The purpose of this expansion is to improve the separation of the different classes in the expanded vector space [45].

The polynomial function is defined by Eq. (13)

$$g(\mathbf{x}) = \mathbf{a}^T p_n(\mathbf{x}), \tag{13}$$

Table 3
Results obtained (%) with classification algorithms for the lymphoma images of the dataset with the 10-fold cross-validation method with morphological features.

Classification algorithm	Features select	CLL-FL			CLL-MCL			FL-MCL		
		SE	SP	AC	SE	SP	AC	SE	SP	AC
SVM	ANOVA	70	83	76	64	79	71	67	67	66
	WRS	71	81	76	64	80	72	67	67	66
	AB	69	81	75	65	78	71	69	66	67
DT	ANOVA	61	85	73	76	73	74	69	61	65
	WRS	72	77	74	69	67	67	69	72	70
	AB	74	73	73	64	74	69	74	65	69
RF	ANOVA	68	88	77	76	81	78	63	74	68
	WRS	71	83	77	69	82	76	65	79	71
	AB	74	89	81	70	75	72	65	75	70
PL	ANOVA	99	94	96	95	91	94	95	92	94
	WRS	96	97	96	92	95	94	86	96	91
	AB	99	94	96	95	95	95	92	94	93

Table 4
Performance of classifier (%) with the non-morphological features from the histological images.

Classification algorithm	Features select	CLL-FL			CLL-MCL			FL-MCL		
		SE	SP	AC	SE	SP	AC	SE	SP	AC
SVM	ANOVA	91	80	86	99	100	99	99	100	99
	WRS	91	84	87	99	100	99	99	100	99
	AB	82	74	78	99	100	99	99	100	99
DT	ANOVA	87	92	89	99	100	99	99	100	99
	WRS	91	89	90	99	100	99	99	100	99
	AB	87	88	87	99	100	99	99	100	99
RF	ANOVA	91	94	92	99	100	99	94	97	95
	WRS	91	94	92	99	99	99	95	98	96
	AB	86	91	88	93	100	96	100	99	99
PL	ANOVA	100	100	100	99	99	99	100	100	100
	WRS	100	100	100	100	99	99	100	100	100
	AB	100	100	100	100	100	100	100	100	100

Table 5
Results (%) of the investigation of the subspace that combines morphological and non-morphological features for classification of lesions in the lymphoma histological images.

Classification algorithm	Features select	CLL-FL			CLL-MCL			FL-MCL		
		SE	SP	AC	SE	SP	AC	SE	SP	AC
SVM	ANOVA	89	87	87	71	83	77	87	96	91
	WRS	92	87	89	68	79	73	87	94	90
	AB	85	87	85	84	83	83	89	92	90
DT	ANOVA	93	85	88	80	74	77	92	89	90
	WRS	90	91	90	82	74	78	90	84	87
	AB	87	92	89	86	76	80	92	90	91
RF	ANOVA	92	95	98	79	81	80	88	95	91
	WRS	93	95	94	73	82	77	92	96	94
	AB	89	93	91	79	86	82	92	96	94
PL	ANOVA	100	100	100	99	96	98	100	100	100
	WRS	100	100	100	99	97	98	100	100	100
	AB	100	100	100	100	100	100	100	100	100

where a is a coefficient vector of the polynomial basis function, $p_n(\mathbf{x})$ is the polynomial basis function and n represents the order or degree of the polynomial function.

The vector of the polynomial discriminant function $\mathbf{a}$ can be calculated by Eq. (14)

$$\mathbf{a} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T\mathbf{b} = \mathbf{M}^{\dagger}\mathbf{b}, \tag{14}$$

where the matrix $\mathbf{M}^{\dagger}$ of dimension $L \times N$ is shown in Eq. (15)

$$\mathbf{M}^{\dagger} = (\mathbf{M}^T\mathbf{M})^{-1}\mathbf{M}^T. \tag{15}$$

This is known as the pseudo-inverse of $\mathbf{M}$ [46].

If there are only two classes, the objective of the classifier is similar to the decision rule given by Eq. (16)

$$\text{Decide} \begin{cases} \omega_1, & \text{if} \quad g(\mathbf{x}) > 0 \\ \omega_2, & \text{if} \quad g(\mathbf{x}) < 0 \end{cases}. \tag{16}$$

The algorithm is divided into two stages, namely training and test, which are detailed in [47]. In this study, the feature set for 3-dimensions was used to calculate the polynomial basis of the 4th order. At this step, different values of degree were tested to calculate the polynomial basis. The value of degree 4th was chosen because it presented the best performance in the separation of the classes which the data have the behaviour non-linear.

2.6. Evaluation of performance

For the statistical analysis, we employed the Students t-test to measure the significance of the features for discriminating between the information from the CLL, FL and MCL images. We used the McNemar's statistical test to evaluate the classification algorithms

Table 6

The *p*-values for comparisons between PL and the other classifiers for different feature selection algorithms.

Hypothesis	Selection algorithm	Features	CLL-FL ρ	CLL-MCL ρ	FL-MCL ρ
	ANOVA	morphological	3.36E-005	9.27E-006	7.16E-008
		non-morphological	0.0049125225	9.62E-001	0.9619767493
		combination	0.0115147722	2.5136E-006	0.1156694431
	WRS	morphological	3.36E-005	1.44E-005	5.57E-007
PL vs SVM		non-morphological	0.0250495925	9.62E-001	0.9619767493
		combination	0.0319296016	1.47E-006	0.1156694431
	AB	morphological	2.21E-005	9.09E-007	2.03E-007
		non-morphological	9.27E-006	0.9619767493	0.9619767493
		combination	0.0150435225	0.0150435225	0.0506305012
	ANOVA	morphological	1.47E-006	5.57E-007	2.03E-007
		non-morphological	0.0403698405	0.9619767493	0.9619767493
		combination	0.1668118624	0.0005011547	0.0506305012
	WRS	morphological	5.93E-006	1.21E-007	3.36E-005
PL vs DT		non-morphological	0.0403698405	0.9619767493	0.9619767493
		combination	0.2333286016	1.44E-005	0.0952047047
	AB	morphological	3.76E-006	7.16E-008	2.36E-006
		non-morphological	0.0115147722	0.9619767493	0.9619767493
		combination	0.0250495925	3.36E-005	0.1668118624
	ANOVA	morphological	0.0003495796	3.50E-004	1.44E-005
		non-morphological	0.1668118624	0.9619767493	0.4176854972
		combination	0.4176854972	7.59E-005	0.0952047047
	WRS	morphological	7.59E-005	3.50E-004	7.59E-005
PL vs RF		non-morphological	0.1668118624	9.62E-001	0.4176854972
		combination	0.1668118624	2.21E-005	0.2333286016
	AB	morphological	0.0026697499	3.76E-006	3.76E-006
		non-morphological	0.0250495925	0.5354219419	0.9619767493
		combination	0.1980334747	0.0010038484	0.1156694431

over the dataset. When the ρ-value is 0.05 or less are, it is considered as statistically significant [48].

In order to evaluate the classification of the lesion, we applied quantitative performance measures considering the following metrics: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). Positive (P) and negative (N) sample elements were grouped according to some special features. These groups simplified computing statistical information to evaluate separability properties described in [49]. These properties are:

- Accuracy (AC): the proportion of correct predictions related to the total number of evaluated cases

$$AC = \frac{TP + TN}{P + N};\qquad(17)$$

- Sensitivity (SE): the proportion of positive cases correctly identified

$$SE = \frac{TP}{TP + FN};\qquad(18)$$

- Specificity (SP): the proportion of negative cases correctly identified

$$SP = \frac{TN}{TN + FP}.\qquad(19)$$

Another important performance measure is the area under the ROC curve (AUC), described in [50,51]. This measure relates the success rate to the failure rate of groups. Both rates range from zero to one. The ROC curve is plotted on the $x - y$ plane, in which the sensitivity varies along y, whereas the specificity varies along x. One takes the AUC as an evaluation of the classifier. A classification test is called perfect when its AUC equals 1.0, and is called perfectly wrong when its AUC equals 0.0.

In our experiments, all tests were evaluated with the accuracy and AUC, which is considered to be the most significant measure for evaluating the classification algorithms [52].

3. Results and discussion

The dataset of lymphoma images was evaluated by applying the techniques described previously, and in this section the results are explained and discussed. The next Section 3.1 explains our proposed procedure concerning the average values for the number of features for the three feature selection methods. At this point we are going to consider the steps for the classification of the lesions by using different algorithms (Section 2.5).
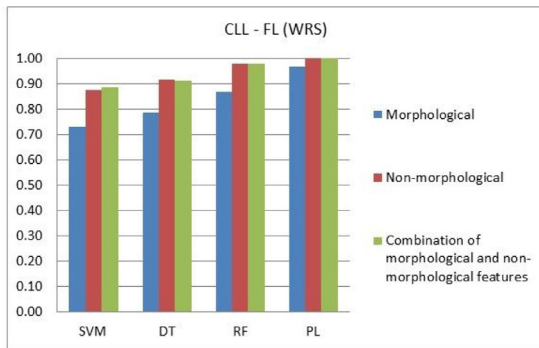
3.1. Feature selection

In Fig. 6(a)–(c), the average values and standard deviation are shown respectively for three classes of features, morphological, non-morphological and the subspace that combines this information, after the application of the feature selection algorithms.

In Fig. 6(a), one can see that the morphological features for the CLL-MCL group obtained the same behaviour in relation to the reduction of the number of features after applying the WRS and AB algorithms. The non-morphological features were also analysed separately by the feature selection methods, and Fig. 6(b) shows that the methods were able to reduce the number of initial features (80 descriptors). The AB algorithm presented significant results when compared to the other methods, which provided a smaller number of features for the different groups of lymphoma subtypes (CLL-FL and FL-MCL). For the CLL-MCL group, the ANOVA and WRS techniques yield a lower number of attributes, approximately 9 features. The use of information obtained from morphological and non-morphological features provided a reduction in the initial number of descriptors.
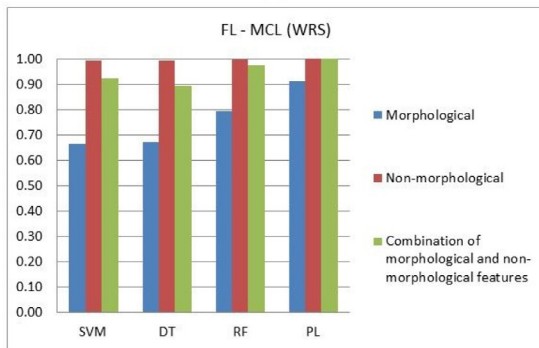
3.2. Impact from features sets

The Students t-test method was used for each feature of the groups and the p-values obtained from each can be used as a measure of how effective it is at separating classes. Fig. 7 presents the
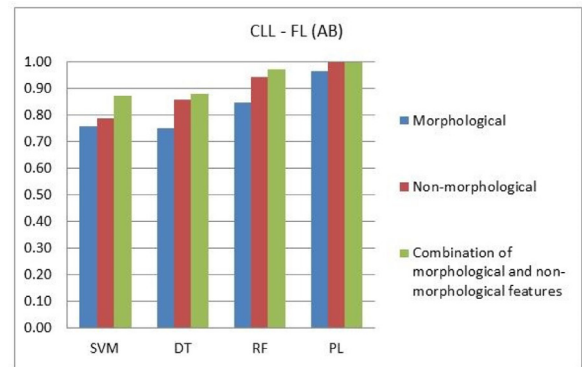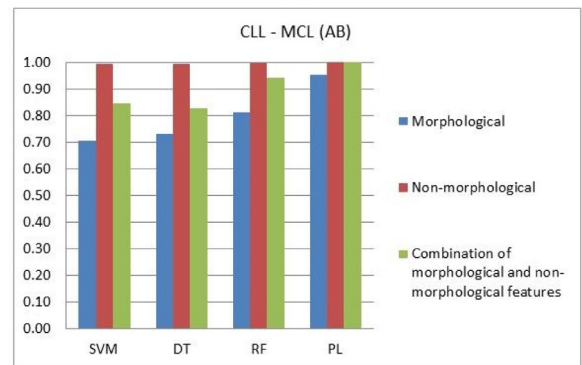
(a)
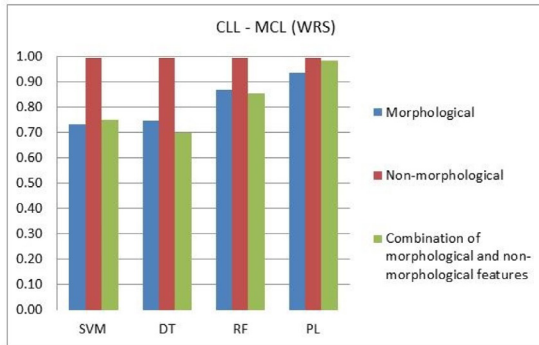
(b)

(c)

Fig. 9. Results of the AUC metric with the WRS feature selection technique and the classifiers: (a) CLL-FL; (b) CLL-MCL and (c) FL-MCL.

(a)

(b)

(c)

Fig. 10. The AUC metric obtained with the AB method and the classifiers with the investigated lesions groups: (a) CLL-FL; (b) CLL-MCL and (c) FL-MCL.

empirical cumulative distribution function (CDF) for the dataset p-values.
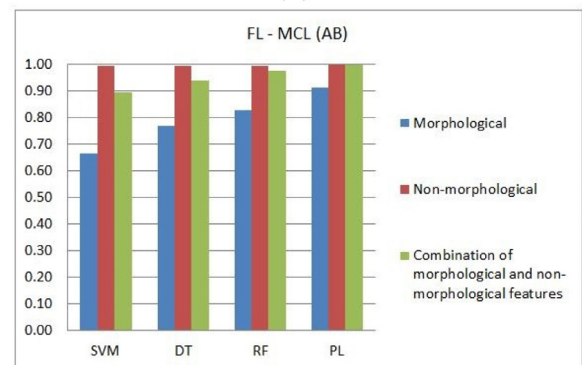
In Fig. 7(a), one notes that more than 60% of the features have p-values close to zero and over 70% of the features have p-values smaller than 0.05 for the morphological features. For the non-morphological features, the CLL-FL and FL-MCL data groups have more than 80% of the features containg p-values smaller than 0.05, which can be used for discrimination of the information concerning each class of lesion. Only the CLL-MCL data group resulted in 10% of relevant information, which can be applied to the discrimination of the features. These data show that there are morphological descriptors that have an adequate separation between the groups investigated. For non-morphological features, the CLL-MCL group presents a number of descriptors with lower statistical relevance in relation to the other groups. This behaviour is maintained with the subspace that combines this information (see Fig. 7(c)).

Tables 3–5 present the results for the set of features. Discriminative performances were evaluated by the SE, SP and AC metrics.

On Table 3, one observes that the results are improved using the PL algorithm and the feature selection algorithms. It is evident that the PL algorithm resulted in the best values of AC for the CLL-FL group (AC = 96%). The PL algorithm demonstrated a superior performance when compared to other classifiers from the FL-MCL, CLL-FL and CLL-MCL groups. For instance, the value of AC with the SVM, DT and RF algorithms and the ANOVA and WRS techniques were on average 20% lower when compared to the PL classifier and the selection algorithms. The use of the DT and ANOVA algorithms applied to the classification of the morphological features presented the worst results for the FL-MCL group.

The non-morphological features show more relevant results when compared to the morphological features (see Table 4). These

Table 7

Comparison of performance from different approaches for lymphoma tissues.

Reference	Feature extraction	Feature selection	Classifier	AC	AUC
[1]	Fourier, Chebyshev and Wavelet transforms	Fisher linear discriminant, minimum redundancy maximum relevance and Fisher/ Correlation	Weighted neighbor distance Nave Bayes network	99.0	–
[15]	Color Dominant, Color Histogram, Co-occurence, Wavelet, Tamura, Gabor and LBP.	Chi-Square	Collateral Representative Decision Tree and Support Vector Machine	92.7	–
[16]	Stationary Wavelet Transform	Analysis of Variance	Support Vector Machine	100.0	–
[54]	Color histogram, edge histogram, Wavelet transforms LBP, color LBP, color multiscale LBP,gist, curvelets, color correlogram, color moments, image statistics and wavelet		Support Vector Machine and Caffe deep learning	95.5	–
[17]	A high-dimensional multi-modal descriptor based in Fisher vector with LBP, HOG, census transform and GIST features	LDA, generalized discriminant analysis, full matrix learning and dictionary learning	Support Vector Machine	96.8	–
[18]	A FV descriptors with local features: dense SIFT, deep belief networks and a CNN model	Separation-guided dimension reduction	Support Vector Machine	97.9	0.99
[55]	Percolation theory		Logistic	–	0.99
Our method	Color channel, morphological features	Analysis of Variance, Wilcoxon rank sum and AnsariBradley	Decision Tree, Random Forest, Support Vector Machine and Polynomial	100.0	1.00

features usually present in species descriptions of internal information of each nuclei present on the histological image. On Table 4, all classifiers showed observable relevant results concerning the CLL-MCL group for a histological analysis system. The PL algorithm presented results with greater differences in the CLL-FL group.

The strategy of a subspace that combines the morphological and non-morphological features is present on Table 5. These experiments showed that part of the results obtained with the PL algorithm for different lymphoma achieved accuracy rates higher than 96%. However, some experiments have shown that there are classification algorithms in which the information combination provided accuracy values lower than the results obtained only with the non-morphological descriptors. For the CLL-MCL group, the result obtained with the WRS and SVM algorithms presented a 73% accuracy rate. The value is less than the result obtained only from the measurement with the non-morphological descriptors (AC = 99%).

Fig. 8–10 show a comparison between AUC obtained using the feature selection methods and classifiers. The AUC values presented in these Figures show that non-morphological descriptors are more relevant than morphological for different combinations of classifier and feature selection methods. The strategy of a subspace that combines morphological and non-morphological features also contributed toward obtaining relevant AUC values.

The results showed that the PL algorithm provided AUC superior than 0.90 for different comparisons of the lesions and the feature selection methods. Then, we applied the McNemar's test for comparisons of the PL algorithm and the other classifiers. The *p*-values are shown on Table 6 for the evaluation with the cross-validation technique. In these experiments, if the *p*-value for a hypothesis is larger than 0.05, then this hypothesis is rejected [53]. The *p*-values that express statistically significant differences between given pairs of models are marked in italic. Considering the values of pair (PL *versus* SVM) with the selection algorithms (Table 6), only the FL-MCL group was rejected for all the subspace that combines this information. This behaviour is also noted with non-morphological features for FL-MCL group. When we analysed the other pairs that were making comparisons between the classifiers, they present similar results in relation to the FL-MCL group. This analysis showed a statistically irrelevant behaviour among the

classifiers for the FL-MCL group when dealing with the feature selection algorithms for the evaluation of the non-morphological descriptors and the subspace that combines morphological and non-morphological features.

3.3. Comparison with previous works

In literature there are methods have been developed to investigate the classification of lymphoma tissues. In Table 7, we present a quantitative comparison in terms of maximum AUC and AC values, that can be seen as complementary. We detail the extraction method, feature selection and classifier used in the original works about the dataset of lymphoma. We also include the results obtained with our method. Part of the methods in Table 7 lead to an almost ideal system. We believe that the different methodologies are rather complementary than ratable.

4. Conclusions and future scope

This study presented a computer aided biopsy analysis support system of lymphoma for the classification of CLL, MCL and FL images based on morphological and non-morphological features. This work investigated the different descriptors and their performance with the classification algorithms. The main goal is to provide a second reading for specialists in order to aid in lesion diagnostics. In order to develop our technique, we presented several experiments that helped to choose adequate combinations from among the different descriptors. The obtained results showed that the algorithms presented relevant information for the detection of lymphoma subtypes. The PL classifier provided a good discriminative performance: CLL-MCL (AUC = 0.906); CLL-FL (AUC = 0.891); and, FL-MCL (AUC = 0.859). Based on these values for AUC, we can state that the proposed technique using non-morphological descriptors and feature selection is a relevant tool for classification of lymphoma tissues concerning digital histology.

In this study, we also analysed the morphological information obtained from the images, with results for AC between 94% and 96%. The use of the ANOVA method was relevant to the investigation of lymphoma subtypes. The strategy of using a subspace that combines the information for features showed that the values of

AC or AUC could be improved when the PL algorithm is used for classification. However, the use this strategy for other classification algorithms presented lower results than the PL algorithm.

The PL algorithm is a supervised classifier and presents pertinent results with non-linear data. This algorithm expands the input data in a superior space dimension, in a manner that allows for the adequate separation between the analyzed classes. In this study different combinations were concatenated from the features of dimensions in order to obtain the result. There are new parameters that can be applied with the objective of evaluating the behavior of the algorithm.

The statistical evaluation showed that the classifiers did not perform similarly when the morphological features were analysed for the feature selection methods. Only in the comparison between SVM versus PL were all the feature selection algorithms relevant for evaluation among the descriptors investigated in this study. Then, this study shows that the information subspace along with the use of non-morphological features presented similar results for the different classifiers.

The main drawbacks are present in the steps addressed to distinguish the nuclei cell, as well as some limitations of the PL classifier. For instance, the proposed method was not able to separate the overlapping lymphocytes and perform the distinction between nuclei and cytoplasm regions. In stage classification, the limitations of the PL classifier are related to the processing time of the algorithm as it is of an exponential order. Therefore, the higher the number of features involved, higher will be the processing time. In this manner, it is not feasible at the present moment to use this in real time for routine diagnostics. Future studies can be addressed to investigate the feature sets that allow thorough discrimination for the improvement of the effectiveness and efficiency of classification processes. Despite these drawbacks, the proposed method was capable of providing relevant results and can be applied as a second reading for specialists in order to aid the diagnosis of MCL, FL and CLL groups.

Acknowledgment

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2018.05.035.

References

[1] N.V. Orlov, W.W. Chen, D.M. Eckley, T.J. Macura, L. Shamir, E.S. Jaffe, I.G. Goldberg, Automatic classification of lymphoma images with transform-based global features, IEEE Trans. Inf. Technol. Biomed. 14 (4) (2010) 1003–1013.

[2] P. Desbordes, C. Petitjean, S. Ruan, Segmentation of lymphoma tumor in pet images using cellular automata: A preliminary study, IRBM 37 (1) (2016) 3–10.

[3] M.d.A. Martins, F.J. Carrilho, V.A.F. Alves, E.A.d. Castilho, G.G. Cerri, C.L. Wen, Clínica médica, Clínica médica, Manole, 2009.

[4] M. da Saúde Instituto Nascional do Câncer Brazil, Estimativa ano 2018: Incidência de câncer, 2018, (http://www.inca.gov.br/estimativa/2018/). Last accessed 22 May 2018.

[5] A. Society, Cancer statistics center, 2018, (https://www.cancerstatisticscenter.cancer.org). Last accessed 22 May 2018.

[6] K.R. Shankland, J.O. Armitage, B.W. Hancock, Non-hodgkin lymphoma, The Lancet 380 (9844) (2012) 848–857.

[7] F.P.S. Santos, G.S. Fernandes, Linfoma de hodgkin, 2017, (http://www.medicinanet.com.br/conteudos/revisoes/2146/linfoma_de_hodgkin.htm). Last accessed 22 May 2018.

[8] M. Oger, P. Belhomme, M.N. Gurcan, A general framework for the segmentation of follicular lymphoma virtual slides, Comput. Med. Imag. Graph. 36 (6) (2012) 442–451.

[9] M. Veta, J.P. Pluim, P.J. Van Diest, M.A. Viergever, Breast cancer histopathology image analysis: a review, IEEE Trans. Biomed. Eng. 61 (5) (2014) 1400–1411.

[10] J.M. Haggerty, X.N. Wang, A. Dickinson, C.J. O'Malley, E.B. Martin, Segmentation of epidermal tissue with histopathological damage in images of haematoxylin and eosin stained human skin, BMC Medical Imaging 14 (1) (2014) 7.

[11] L. Shamir, N. Orlov, D.M. Eckley, T.J. Macura, I.G. Goldberg, Iicbu 2008: a proposed benchmark suite for biological image analysis, Med. Biol. Eng. Comput. 46 (9) (2008) 943–947.

[12] M.F.A. Fauzi, M. Pennell, B. Sahiner, W. Chen, A. Shana'ah, J. Hemminger, A. Gru, H. Kurt, M. Losos, A. Joehlin-Price, et al., Classification of follicular lymphoma: the effect of computer aid on pathologists grading, BMC Med. Inf. Decis. Making 15 (1) (2015) 115.

[13] O. Sertel, J. Kong, G. Lozanski, A. Shana'ah, U. Catalyurek, J. Saltz, M. Gurcan, Texture classification using nonlinear color quantization: application to histopathological image analysis, in: Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, IEEE, 2008, pp. 597–600.

[14] O. Sertel, J. Kong, U.V. Catalyurek, G. Lozanski, J.H. Saltz, M.N. Gurcan, Histopathological image analysis using model-based intermediate representations and color texture: Follicular lymphoma grading, J. Signal Process. Syst. 55 (1-3) (2009) 169.

[15] T. Meng, L. Lin, M.-L. Shyu, S.-C. Chen, Histology image classification using supervised classification and multimodal fusion, in: Multimedia (ISM), 2010 IEEE International Symposium on, IEEE, 2010, pp. 145–152.

[16] M.Z.d. Nascimento, L. Neves, S.C. Duarte, Y.A.S. Duarte, V.R. Batista, Classification of histological images based on the stationary wavelet transform, in: Journal of Physics: Conference Series, vol. 574, IOP Publishing, 2015, p. 012133.

[17] Y. Song, W. Cai, H. Huang, D. Feng, Y. Wang, M. Chen, Bioimage classification with subcategory discriminant transform of high dimensional visual descriptors, BMC Bioinf. 17 (1) (2016) 465.

[18] Y. Song, Q. Li, H. Huang, D. Feng, M. Chen, W. Cai, Low dimensional representation of fisher vectors for microscopy image classification, IEEE Trans. Med. Imag. 36 (8) (2017) 1636–1649, doi:10.1109/TMI.2017.2687466.

[19] M.T. McCann, J.A. Ozolek, C.A. Castro, B. Parvin, J. Kovacevic, Automated histology analysis: Opportunities for signal processing, IEEE Signal Process. Mag. 32 (1) (2015) 78–87.

[20] E.D. Pisano, S. Zong, B.M. Hemminger, M. DeLuca, R.E. Johnston, K. Muller, M.P. Braeuning, S.M. Pizer, Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms, J. Digital imag. 11 (4) (1998) 193–200.

[21] R. Eustice, O. Pizarro, H. Singh, J. Howland, Uwit: Underwater image toolbox for optical image processing and mosaicking in matlab, in: Underwater Technology, 2002. Proceedings of the 2002 International Symposium on, IEEE, 2002, pp. 141–145.

[22] MATLAB, version 7.10.0 (R2010a), The MathWorks Inc., Natick, Massachusetts, 2010.

[23] R.C. Gonzalez, R.E. Woods, Digital Image Processing (3rd Edition), Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[24] P. Soille, Morphological Image Analysis: Principles and Applications, Springer-Verlag New York, Inc., 2003.

[25] G. de Melo, V. Gomes, C. Baccili, L. de Almeida, A.d.C. Lima, A robust segmentation method for counting bovine milk somatic cells in microscope slide images, Comput. Electron. Agric. 115 (2015) 142–149.

[26] F. Meyer, Topographic distance and watershed lines, Signal Process. 38 (1) (1994) 113–125.

[27] C. Solomon, T. Breckon, Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab, Wiley, 2011.

[28] R. Haralick, L. Shapiro, Computer and Robot Vision, vol. 1, Addison-Wesley, 1992.

[29] A. Mencattini, M. Salmeri, G. Rabottino, S. Salicone, Metrological characterization of a CADx system for the classification of breast masses in mammograms, IEEE Trans. Instrum. Measur. 59 (11) (2010) 2792–2799.

[30] B. Surendiran, A. Vadivel, Feature selection using stepwise ANOVA, discriminant analysis for mammogram mass classification, Int. J. Recent Trends Eng. Technol. 3 (2) (2010) 55–57.

[31] Z. Ma, J.M.R. Tavares, Effective features to classify skin lesions in dermoscopic images, Expert Syst. Appl. 84 (2017) 92–101.

[32] B.H. Cho, H. Yu, K.-W. Kim, T.H. Kim, I.Y. Kim, S.I. Kim, Application of irregular and unbalanced data to predict diabetic nephropathy using visualization and feature selection methods, Artif. Intell. Med. 42 (1) (2008) 37–53.

[33] J. Demšar, Statistical comparisons of classifiers over multiple data sets, J. Mach. Learn. Res. 7 (2006) 1–30.

[34] M.U. Akram, A. Tariq, S.A. Khan, M.Y. Javed, Automated detection of exudates and macula for grading of diabetic macular edema, Comput. Methods Programs Biomed. 114 (2) (2014) 141–152.

[35] M.M.R. Krishnan, P. Shah, C. Chakraborty, A.K. Ray, Statistical analysis of textural features for improved classification of oral histopathological images, J. Med. Syst. 36 (2) (2012) 865–881.

[36] R.B. Oliveira, J.P. Papa, A.S. Pereira, J.M.R. Tavares, Computational methods for pigmented skin lesion classification in images: review and future trends, Neural Comput. Appl. 29 (3) (2018) 613–636.

[37] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[38] R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers, San Mateo, CA, 1993.

[39] S. Keerthi, S. Shevade, C. Bhattacharyya, K. Murthy, Improvements to platt's smo algorithm for svm classifier design, Neural Comput. 13 (3) (2001) 637–649.

[40] T.M. Mitchell, Machine Learning, first ed., McGraw-Hill, Inc., New York, NY, USA, 1997.

[41] J.R. Quinlan, Induction of decision trees, Mach. Learn. 1 (1) (1986) 81–106, doi:10.1023/A:1022643204877.

[42] J.R. Quinlan, C4. 5: Programs for Machine Learning, Elsevier, 2014.

[43] L. Breiman, J. Friedman, R. Olshen, C. Stone, Classification and Regression Trees, Wadsworth and Brooks, Monterey, CA, 1984.

[44] V.N. Vapnik, An overview of statistical learning theory, Neural Netw. IEEE Trans. 10 (5) (1999) 988–999.

[45] T. Shanableh, K. Assaleh, Feature modeling using polynomial classifiers and stepwise regression, Neurocomputing 73 (10) (2010) 1752–1759.

[46] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, John Wiley & Sons, 2012.

[47] M.Z. Do Nascimento, A.S. Martins, L.A. Neves, R.P. Ramos, E.L. Flores, G.A. Carrijo, Classification of masses in mammographic image using wavelet domain features and polynomial classifier, Expert Syst. Appl. 40 (15) (2013) 6213–6221.

[48] K.S.S. Bhatia, A.C.L. Lam, S.W.A. Pang, D. Wang, A.T. Ahuja, Feasibility study of texture analysis using ultrasound shear wave elastography to predict malignancy in thyroid nodules, Ultrasound Med. Biol. 42 (7) (2016) 1671–1680.

[49] S. Dua, H. Singh, H.W. Thompson, Associative classification of mammograms using weighted rules, Expert Syst. Appl. (2010) 9250–9259.

[50] J. Davis, M. Goadrich, The relationship between precision-recall and roc curves, in: Proceedings of the 23rd International Conference on Machine Learning, ACM, 2006, pp. 233–240.

[51] T. Saito, M. Rehmsmeier, The precision-recall plot is more informative than the roc plot when evaluating binary classifiers on imbalanced datasets, PloS one 10 (3) (2015) e0118432.

[52] J. Huang, C.X. Ling, Using auc and accuracy in evaluating learning algorithms, IEEE Trans. Knowl. Data Eng. 17 (3) (2005) 299–310, doi:10.1109/TKDE.2005.50.

[53] B. Trawinski, M. Smetek, T. Lasota, G. Trawinski, Evaluation of Fuzzy System Ensemble Approach to Predict from a Data Stream, Springer International Publishing, Cham, pp. 137–146. doi:10.1007/978-3-319-05458-2_15.

[54] N. Codella, M. Moradi, M. Matasar, T. Syeda-Mahmood, J.R. Smith, Lymphoma diagnosis in histopathology using a multi-stage visual learning approach, SPIE Medical Imaging, International Society for Optics and Photonics, 2016. 97910H–97910H

[55] G.F. Roberto, L.A. Neves, M.Z. Nascimento, T.A. Tosta, L.C. Longo, A.S. Martins, P.R. Faria, Features based on the percolation theory for quantification of non-hodgkin lymphomas, Comput. Biol. Med. 91 (2017) 135–147.