A Method For Multiclass Lymphoma Classification Based on Morphological and Non-Morphological Descriptors

Tiago P. de Faria
Universidade Federal de Uberlândia
Uberlândia, Brasil
Email: tiagofaria@ufu.br

Marcelo Z. do Nascimento
Universidade Federal de Uberlândia
Uberlândia, Brasil
Email: marcelo.nascimento@ufu.br

Luiz G. A. Martins
Universidade Federal de Uberlândia
Uberlândia, Brasil
Email: lgamartinsr@ufu.br

***Abstract*—Lymphoma is one of the most common types of cancer and its treatment can be more effective if the disease variant is correctly diagnosed. Many works have been done using computer vision and machine learning to classify the lymphoma based on histological images. This work presents a method using simple descriptors and a decision tree-based ensemble classifier, aiming to maintaing the interpretability of the data and understand what information in most important to the classification task. We use morphological and non morphological descriptors extracted from the cells nuclei, a feature selection method based on principal component analysis (PCA), and a gradient boosting decision tree (GBDT) method for multiclass classification. Our approach achieves an average accuracy of 0.932. this result is close to those obtained in the state of the art, while it uses simpler descriptors and better interpretable classification models.**

***Index Terms*—Multiclass classification, feature selection, morphological and non-morphological descriptors, lymphoma.**

I. Introduction

Lymphoma is the group of cancer that develop in the lymphocytes, a type of blood cell that is responsible for the immunology of the organism [18]. It is one of the most common types of cancer [17] and can be divided between Hodkin lymphomas (HL) and non-Hodkin lymphomas (NHL). Twenty-eight variants of the disease are already known, but only three NHL represents 85% of all lymphoma cases: the mantle cell lymphoma (MCL), the follicular lymphoma (FL) and the chronic lymphocytic leukemia (CLL) [22]. The high occurrence frequency of this types make them important subjects of research.

An important step on the treatment of lymphomas is correctly diagnosing the type of the disease. Unfortunately, this crucial task is not trivial, even to specialists. The analysis of histological images can be used in this process, but inter and intra-pathologist variability in the diagnosis can occur, caused by the subjectivity of human analysis [13]. Computer aided diagnosis (CAD) systems are tools designed to facilitate physicians in the analysis of medical exams. Using computer vision and machine learning techniques, a CAD system can be created to classify histological images of lesions, which can speed up the diagnosis and mitigate the variation in the analysis.

The success of the various machine learning models depends on collecting high quality and reliable data. A gold standard annotated database of histological images of lesions of three types of NHL (MCL, FL and CLL) is presented in [23]. Several researches have been conducted using this database, addressing different approaches to extract useful information for classifying images between the cited types. The extraction of descriptors from medical images and their statistical analysis can lead to new features and aid in the discoveries regarding the subject diseases [5]. However, many works either do not segment the image to find information in the cell nuclei or they use complex descriptors, that can not be easily interpreted by a specialist [6], [13], [14]. In [17], the authors extract simple data from the cell nuclei related to their color and shape, but they only perform binary classification.

Many multiclass approaches with high accuracy for the lymphoma classification problem have been proposed in the literature. However, they often use more complex descriptors and machine learning methods with low levels of explainability [3], such as deep learning [4], [15]. In these models, the knowledge and logic used in classification are intrinsic, making it difficult to understand why and how a decision is made. On the other hand, when a specialist uses an automated system to help in his diagnosis, he needs more then just an accurate prediction. It is really important for a CAD to be transparent enough so the user can understand better what context leads to each diagnosis [3].

In this paper, we present an approach for the classification of lesions in NHL histological images. The proposed method uses images with nuclei detected by the method proposed by [17]. Morphological and non morphological descriptors are extracted from the segmented images, which are based on the form of these nuclei and on the brightness level of their pixels, respectively. Then, a principal component analysis (PCA) technique are used to select a subset of features. Finally, the gradient boosting decision tree (GBDT) method is employed to classify the lymphoma. The choice of algorithms aims to maintain the simplicity and transparency of the information used

The authors thanks the financial support of CNPq (145081/2019-2)

in the decision, being a step further to the creation of a fully interpretable model, able to support doctors in their diagnoses. The main contributions of this paper are: (i) The combination of simple and interpretable descriptors, a PCA-based feature selector and a DT-based ensemble method in a classification approach in order to demonstrate its discriminative capability in the different lesion classes; (ii) The investigation of morphological and non morphological descriptors in multiclass classification models, highlighting which descriptors are the most important for the task and discussing the implications of combining both types of descriptors.

II. Methodology

Fig. 1 illustrates the general flow of the proposed method. Initially, descriptors are extracted from the images using the method proposed in [17]. In preprocessing, a feature selection algorithm based on PCA is used to reduce the number of features, which are then used in the training of a GBDT.

The algorithms used were chosen aiming to create a simple prediction methodology, but that have achieved competitive performance in other classification problems [11], [27], [29].

A. Dataset and Image Descriptors

The dataset used in this work were proposed in [23] and is formed by 375 labeled images of NHL lesions, where their subtypes CLL, FL and MCL have 113, 140 and 122 images, respectively. They was created from a set of 30 histological samples (10 samples of each type of NHL), colorized with hematoxylin and eosin (H&E).

Aiming to obtain data from the images, we use the same segmentation and feature extraction method proposed in [17], which allows the detection of the cells nuclei in the images and the extraction of morphological and non-morphological descriptors. The process generates a set of 36 morphological descriptors based on statistical data about the geometric shape and size of cells nuclei. For each nucleus, nine morphological metrics (area, extent, perimeter, convex area, solidity, eccentricity, equivalent diameter, minor and major axis) are computed. Then, four statistical data (mean, median, mode and standard deviation) are calculated for each metric, considering all nuclei of each image. In addition, a set of 80 non-morphological descriptors are extracted based on the pixel brightness in the model RGB (Red, Green and Blue) and gray-scale channels of the image. For each nucleus, the mean, median, standard deviation, minimum and maximum values of pixel brightness in each of four channel are computed. From these data, we then calculated the mean value, median, standard deviation and mode for each image channel.

In order to compare them, experiments were conducted using each set of descriptors separately, as well as with a third group formed of both of them.

B. Preprocessing

A large number of features can negatively influence in the performance of classification algorithms [2]. Therefore, the descriptors extracted from the image are submitted to a feature selection algorithm in order to get a smaller subset that still contains enough information to separate the classes.

The traditional PCA algorithm reduces dimensionality by mapping the original features to new orthogonal variables, called principal components (PC), in order to maximize the variance [1]. This method can project the data in a less dimensional space, reducing the complexity of the classification problem. However, the number of descriptors that must be extracted from the image remains the same, since each main component is created from a linear combination of all features. In addition, the new attributes are more difficult to interpret.

In this work, a feature selection algorithm based on PCA [24] was used. The PCA-based selector chooses the K most relevant features of the database, that is, those having more residual variance in the principal components. This approach returns a subset formed by original descriptors (morphological and non-morphological). Unlike the traditional approach, where there is decomposition of features in a different data space, the selection method preserves the information of each feature of the selected subset. Therefore, it improves the model interpretability, favoring the analysis of the selected features by a human specialist and reduces the cost of image processing, since it requires a smaller number of descriptors. The value of K should be determined searching the best trade-off between the quantity of features used in the training stage and the quality of the classifier.

For evaluating this feature selection method, we also compared it's results with the decomposition approach.

C. Classification

In present work, we perform a multiclass classification for NHL using a decision tree-based ensemble, where the models are trained from the features selected in the preprocessing. An ensemble is formed by several simpler models, where the classification is made based on the individual decisions of each model (eg by simple voting). This technique has already proved to be more efficient than a single model in several applications [19].

One of the machine learning techniques commonly used in ensemble models is the decision tree (DT), due to its simplicity and fast training. Each tree node makes a decision based on a single attribute, guiding the algorithm to one of its child nodes and, when it is a leaf node, classifies the sample [21]. This simple structure allows the specialist to understand the criteria adopted in the decision and confirm the diagnosis. For example, Fig. 2 illustrates a piece of a DT-based classifier for NHL. As noted, when the non-morphological descriptor *StdDev_MedianPixelValueGr* (standard deviation of median value of pixel brightness of the cells nuclei in the green channel) is less than or equal to 13.82, the lesion is classified as CLL. Otherwise, it is necessary to check the morphological descriptor *Median_ConvexArea* (median of the convex area of all image nuclei) to decide between FL (≤ 98.5) and CLL (> 98.5).

GBDT is a method, where new decision trees are iteratively added to the ensemble in order to decrease its error. In GBDT,

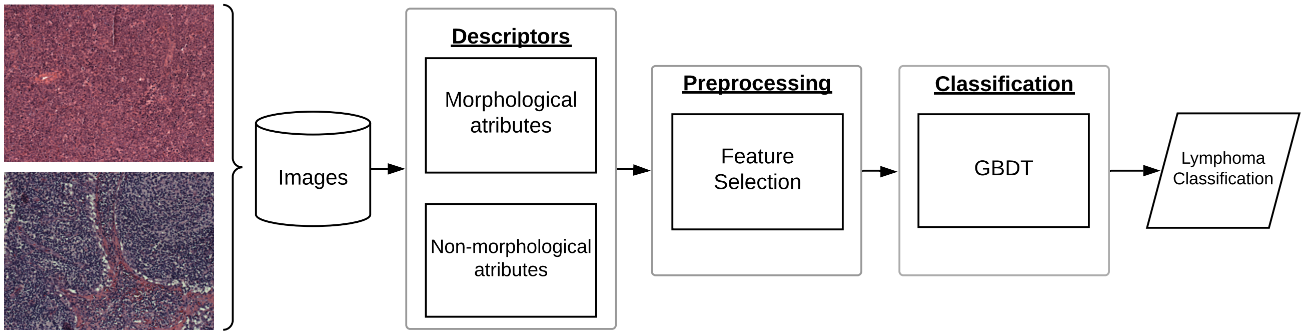
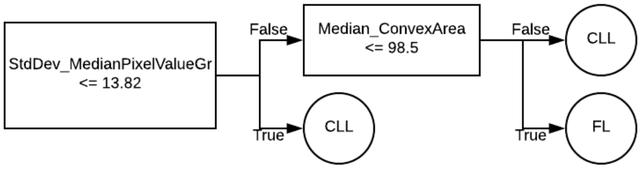


Fig. 1. Scheme of the proposed methodology for lymphoma classification.



the problem is modified between each iteration, focusing on the samples most likely to be classified incorrectly. After trained, the model uses a combination of the results of all trees to create its final decision.

The experiments were carried out using the GBDT implementation from the light gradient boosted machine (LGBM) library [10] and its default configuration for the algorithm's parameters.

III. Experiments

This section presents the experiments carried out to evaluate the performance of our methodology and perform an empirical analysis on the importance of the features on the classifiers decisions. In order to evaluate and compare the efficiency of morphological and non-morphological features in the classification of lymphomas, the models were constructed considering each set of descriptors separately and another one composed by both (features junction). Initially, we evaluated the influence of PCA-based feature selection method and their parameterization on the quality of the classifiers. Then, we analyzed the importance of each feature, and it's differences between the set of descriptors, highlighting patterns and behaviors that we judge as important knowledge and evidence for future works regarding this problem.

In the comparative analysis, 100 executions of each evaluated method were considered. For each one, 10% of the data was randomly selected for testing, while the other 90% was used in the training of the classification model. The separation of samples was performed in order to maintain the proportion of each class was the same between the training and test subsets. The algorithm performance is measured by the accuracy of the model [9]. We also used the Welch test [8], [28] with a significance level of 5% to confirm or refute the null hypothesis (H0) that the classification approaches have similar average performances. It is usually used to verify the hypothesis that two populations have the same average, without assuming equal variance between them. In other words, if the test returns a p-value less than 0.05 (H0 rejected), we determine that the difference between the models is statistically relevant with 95% confidence. Otherwise, we observed that the difference in the performance may have resulted from the random variation inherent to stochastic methods.

A. Analysis of the Feature Selection

In order to determine which is the smallest subset of descriptors able to training a model with good classification accuracy, a feature selection algorithm based on PCA were evaluated. The number of features selected for each set of descriptors was determined through multiple executions of the method, increasing the value of K from 3 to the total number of descriptors in the set. This incremental process stops when the improvement in model accuracy is considered insignificant (< 0.01). Thereby, we could find a good trade-off between number of features and classifier performance. Based on the experimental results, we determined the number of features to be selected for each set of descriptors. While 76 features were selected from the complete set (all descriptors), for the morphological and non-morphological sets, 33 and 16 descriptors were chosen respectively.

To compare the efficiency of classifiers generated from our feature selection approach, experiments using a well established dimensionality reduction technique (PCA-based feature decomposition algorithm) was also performed. Table I presents the performance achieved by the classifiers trained from the features set generated using these two methods for each sets of descriptors. The column ”Qty” shows the number of features used in both algorithms. The average accuracy and it's standard deviation for each approach is shown in columns ”Decomposition” and ”Selection”, respectively. The column ”All Features” shows the evaluations achieved by using all features, without any dimensionality reduction. The highlighted

TABLE I
MODEL'S ACCURACY WITH DIFFERENT PCA-BASED DIMENSIONALITY REDUCTION METHODS.

	Qty	Decomposition	Selection	All Features
All	76	**0.929 ± 0.041**	**0.932 ± 0.042**	0.913 ± 0.048
Morph.	18	0.718 ± 0.072	0.702 ± 0.072	0.720 ± 0.077
Non-Morph.	34	0.885 ± 0.050	**0.911 ± 0.042**	**0.917 ± 0.045**

values (bold) indicate the best performances with statistical significance (p-value ≤ 0.05). As can be noted, there was no significant difference in performance for the morphological set. Considering the non-morphological set, the selection-based method had a performance similar to the model using all features, and superior to the decomposition-based method. For the mixed set (all descriptors), both evaluated methods performed better than the approach using all resources, without significant difference between them. Therefore PCA-based feature selection is able to achieve a good performance, while allowing us to keep the interpretability of the original features unchanged.

B. Analysis of the Feature Importance

Aiming to provide insights about what kind of information can be more relevant on the classification of histological images of lymphoma, we present an analysis of the importance of each descriptor used in the training of the model. We defined the importance of a feature as the number of times it is used in a tree node in the 100 executions of the experiments. In each execution, the PCA-based selector algorithm was used to choose the features according to the quantity established in the previous section. If a feature is not selected, it's importance is 0 in that execution.

Fig. 3 shows the top-10 most important features for each descriptor set, ranked according to their degree of importance. Morphological and non-morphological descriptors are represented in different geometric shapes, so that it's easy to visualize which of them are being used in the set with both types together (mixed set). Arrows shows the change of position in ranking between features in their specific descriptor type set and in the mixed set.

Analyzing the ranks, it is possible to notice that the inclusion of new descriptor changes the importance of each feature. For example, the 4th most important feature on the set with all descriptors (standard deviation of the average pixel value on red channel) was placed at 37th when using only the non-morphological set. Therefore, the addition of morphological features increases the impact of this descriptor in the decision. It's also important to note that, despite the morphological set results in the worst performance in the classification, 2 of the top-10 most important characteristics in the mixed set are morphological. If we consider the top-15 most important features of the mixed set, 6 are morphological, indicating its contribution when used in combination with non-morphological features. These results are evidence that weaker descriptors (that initially has low discriminatory value) can be used in conjunction with stronger ones, improving the accuracy of the classifiers. This is specially useful on the lymphoma problem, where morphological descriptors are notably weaker, but advantageous to be used in explainable models, as they are simpler and easier to understand than non-morphological ones. The 6 morphological descriptors present on the top-15 features of the mixed set are all related to the area and the convex area of the cells nuclei, showing the importance of this type of information when combined with non-morphological features.

Fig. 4 shows the average importance level for different groups of descriptors. In Fig. 4(a) is presented the mean importance of non-morphological descriptors grouped according to color channel. The red and blue channel's descriptors are selected more often and consequently has more importance then those from the green and gray channels. However, although there are differences between the color channels, they all seem to contribute to the classification task. Figure 4(b) shows the mean importance of descriptors clustered with respect to the morphological metrics, composed eight groups of descriptors. Among them, soliditys' group seems to be way less relevant. Three descriptors of this group was selected at least once in the training phase, but they appear only 2092 times in average, considering a total of 637,239 tree's nodes in all 100 executions. This type of information can be taken in consideration when designing new descriptors for this problem.

IV. CORRELATED WORKS

Comparison with previous approaches for the lymphoma lesions classification can be somewhat difficult, since the available data and the solution's architecture varies a lot, and this has a big impact on performance [26]. In general, published studies report high accuracy ($> 90\%$) when using deep learning, but this technique often requires large datasets (2,560 to 850,000 images) or pretrained models [26].

Table II shows the performance of several correlated works that uses the same dataset employed in our research. The column "ACC" shows the accuracy of the method. The column "Ref" shows the reference to each work. The number of features and how it was extracted is in the columns "Qty" and "Features", respectively. For comparison, we show the results achieved from our method using a PCA-based selector and the mixed set of descriptors. This configuration was chosen since it achieved the best average accuracy.

All works presented in the table use more complex descriptors demanding more expensive image processing and hinder their model's interpretation by doctors, who are not familiar with this type of data. Furthermore, their methods are generally black-box, or at least, not trivial to be understood by non-specialists, such as SVM [6], [16], [25], neural networks [20] and polynomial classifiers [13], where the information is embedded in the model or requires an expensive training. In the other hand, our methodology uses simpler and easier to extract descriptors, and employs the GBDT algorithm, providing

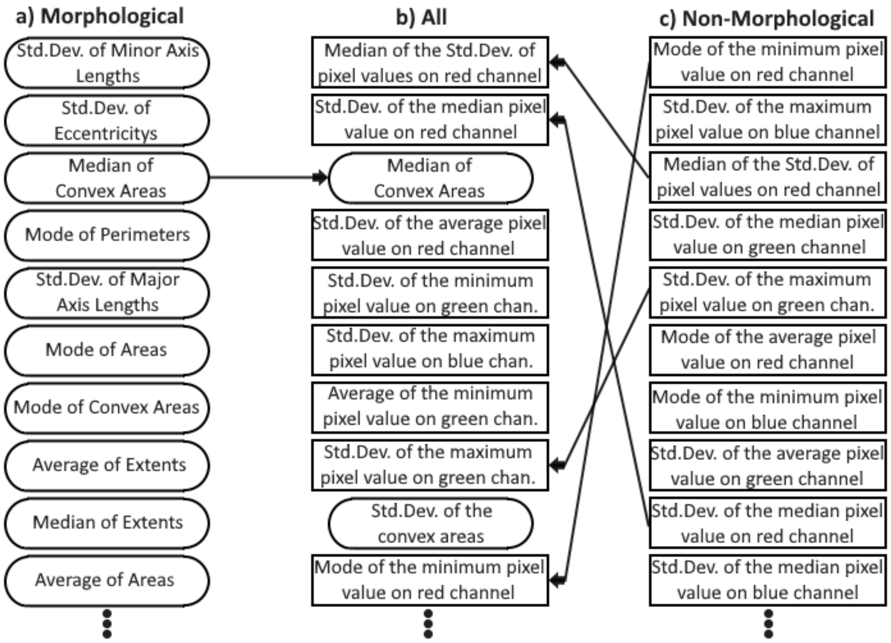

a) Morphological
Std.Dev. of Minor Axis Lengths
Std.Dev. of Eccentricitys
Median of Convex Areas
Mode of Perimeters
Std.Dev. of Major Axis Lengths
Mode of Areas
Mode of Convex Areas
Average of Extents
Median of Extents
Average of Areas
b) All
Median of the Std.Dev. of pixel values on red channel
Std.Dev. of the median pixel value on red channel
Median of Convex Areas
Std.Dev. of the average pixel value on red channel
Std.Dev. of the minimum pixel value on green chan.
Std.Dev. of the maximum pixel value on blue chan.
Average of the minimum pixel value on green chan.
Std.Dev. of the maximum pixel value on green chan.
Std.Dev. of the convex areas
Mode of the minimum pixel value on red channel
c) Non-Morphological
Mode of the minimum pixel value on red channel
Std.Dev. of the maximum pixel value on blue channel
Median of the Std.Dev. of pixel values on red channel
Std.Dev. of the median pixel value on green channel
Std.Dev. of the maximum pixel value on green chan.
Mode of the average pixel value on red channel
Mode of the minimum pixel value on blue channel
Std.Dev. of the average pixel value on green channel
Std.Dev. of the median pixel value on red channel
Std.Dev. of the median pixel value on blue channel

Fig. 3. Feature importance rank from the GBDT models trained using morphological, non-morphological or mixed set.

TABLE II
RELATED WORK BASED ON MULTICLASS CLASSIFIERS FOR THE LYMPHOMA DATABASE

Ref	Features	Qty	Classifier	ACC
[20]	Features based on the percolation theory	15	DECORATE	0.92
[14]	Color, Histogram, Texture, Wavelet and Binary Patterns information	50	C-RSPM	0.927
[6]	histogram; LBP; gist; curvelet; color correlogram and moments; wavelets;	200	SVM ensemble	0.955
[25]	IFV; LBP; HOG;GIST and CENTRIST	180	SVM	0.968
[13]	Fractal geometry features	18	HPG4	0.914
[4]	patch-level textural & statistic. feat.; color feat.; GoogLeNet pre-trained model	550	RF + CNN	0.991
[16]	Stationary Wavelet Transform	34,236	SVM	1
Ours	Morph. and non-Morph. features	76	RF	0.932

a fast training of classification DT-based ensembles. Despite the simplicity of the methodology, our approach achieved an accuracy close to several other recent methods in the literature.

V. CONCLUSIONS

In this work, we present a method for classifying NHL that employs interpretable and easy to extract descriptors and a classifier based on decision tree ensemble, where the knowledge used in decision making is explicit in the tree nodes. This features are important steps toward the creation of a method of classification that favors the interpretation of data by human specialists and, consequently, its use to aid the physician's diagnosis. Our approach proved to be efficient, returning good performance in the task of multiclass classification of histological images of patients with lymphomas. Experimental results indicate that it is possible to obtain an accuracy similar to related works from the selection of a subset of characteristics, both for the complete set and for the one formed only by non-morphological descriptors. In both scenarios, it was possible to build classifiers based on DT ensemble that reached good accuracy. The two best performances were achieved using 76 features from the mixed set (0.932 accuracy) and 34 descriptors of the non-morphological set (0.911 accuracy), respectively. Both models using the gradient boosting decision tree method as the classifier.

Although ensembles make it difficult to interpret the generated model, there are methods in the literature to create interpretable models from them, for example, a technique that transforms an ensemble of decision trees in a simpler and interpretable model based on decision rules is presented in [7]. Another approach is discussed in [12], where local feature importance is used to explain tree ensemble models. In future work, we intend to investigate such methods to further improve the interpretability of our model. Based on the descriptors analysis realized in present work, we see potential in reformulating or creating other types of interpretable

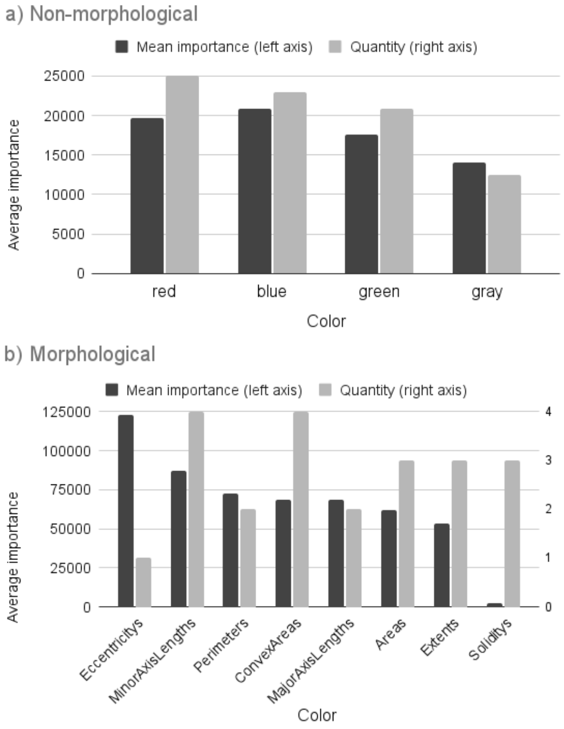


Fig. 4. Quantity and mean importance of descriptors

descriptors, increasing the model's accuracy while maintaining its features easy to understand. Future works may also evaluate the efficiency of this method on other histological classification problems.

References

[1] Abdi, H., Williams, L.J.: Principal component analysis. Wiley interdisciplinary reviews: computational statistics **2**(4), 433–459 (2010)

[2] Aggarwal, C.C.: Data mining: the textbook (2015)

[3] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., et al.: Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. Information Fusion **58**, 82–115 (2020)

[4] Bai, J., Jiang, H., Li, S., Ma, X.: Nhl pathological image classification based on hierarchical local information and googlenet-based representations. BioMed research Int. (2019)

[5] Beck, A.H., et al.: Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. Sci. Transl. Med. (2011)

[6] Codella, N., Moradi, M., Matasar, M., Sveda-Mahmood, T., Smith, J.R.: Lymphoma diagnosis in histopathology using a multi-stage visual learning approach. In: Medical Imaging 2016: Digital Pathology. vol. 9791, p. 97910H (2016)

[7] Deng, H.: Interpreting tree ensembles with intrees. Int. J. of Data Science and Analytics **7**(4), 277–287 (2019)

[8] Derrick, B., Toher, D., White, P.: Why welch's test is type i error robust. The Quantitative Methods in Psychology (2016)

[9] James, G., Witten, D., Hastie, T., Tibshirani, R.: An introduction to statistical learning, vol. 112. Springer (2013)

[10] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: Lightgbm: A highly efficient gradient boosting decision tree. In: Advances in neural information processing systems. pp. 3146–3154 (2017)

[11] Li, Z., et al.: Gbdt-svm credit risk assessment model and empirical analysis of peer-to-peer borrowers under consideration of audit information. Open Journal of Business and Management **6**(02), 362 (2018)

[12] Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.I.: From local explanations to global understanding with explainable ai for trees. Nature machine intelligence **2**(1), 56–67 (2020)

[13] Martins, A.S., Neves, L.A., de Faria, P.R., Tosta, T.A., Longo, L.C., Silva, A.B., Roberto, G.F., do Nascimento, M.Z.: A hermite polynomial algorithm for detection of lesions in lymphoma images. Pattern Analysis and Applications pp. 1–13 (2020)

[14] Meng, T., Lin, L., Shyu, M.L., Chen, S.C.: Histology image classification using supervised classification and multimodal fusion. In: 2010 IEEE Int. symposium on multimedia. pp. 145–152 (2010)

[15] Nannia, L., Ghidoni, S., Brahnam, S.: Ensemble of convolutional neural networks for bioimage classification. Applied Computing and Informatics (2020)

[16] Nascimento, M.Z., Neves, L., Duarte, S.C., Duarte, Y.A.S., Batista, V.R.: Classification of histological images based on the stationary wavelet transform. J. of Physics: Conference Series **574**, 012133 (jan 2015)

[17] do Nascimento, M.Z., Martins, A.S., Tosta, T.A.A., Neves, L.A.: Lymphoma images analysis using morphological and non-morphological descriptors for classification. Computer methods and programs in biomedicine **163**, 65–77 (2018)

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

[19] Oza, N.C., Tumer, K.: Classifier ensembles: Select real-world applications. Information fusion pp. 4–20 (2008)

[20] Roberto, G.F., Neves, L.A., Nascimento, M.Z., Tosta, T.A., Longo, L.C., Martins, A.S., Faria, P.R.: Features based on the percolation theory for quantification of non-hodgkin lymphomas. Computers in biology and medicine **91**, 135–147 (2017)

[21] Russell, S., Norvig, P.: Artificial intelligence: a modern approach

[22] Santos, F.P.S., Fernandes, G.S.: Linfomas não-Hodgkin. MedicinaNet (2008)

[23] Shamir, L., Orlov, N., Eckley, D.M., Macura, T.J., Goldberg, I.G.: Iicbu 2008: a proposed benchmark suite for biological image analysis. Medical & biological engineering & computing **46**(9), 943–947 (2008)

[24] Song, F., Guo, Z., Mei, D.: Feature selection using principal component analysis. In: Int. Conf. on System Science, Engineering Design and Manufacturing Informatization. vol. 1, pp. 27–30 (2010)

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

[26] Steinbuss, G., Kriegsmann, M., Zgorzelski, C., Brobeil, A., Goeppert, B., Dietrich, S., Mechtersheimer, G., Kriegsmann, K.: Deep learning for the classification of non-hodgkin lymphoma on histopathological images. Cancers **13**(10), 2419 (2021)

[27] Sun, R., Wang, G., Zhang, W., Hsu, L.T., Ochieng, W.Y.: A gradient boosting decision tree based gps signal reception classification algorithm. Applied Soft Computing **86** (2020)

[28] Welch, B.L.: The generalization of 'student's' problem when several different population variances are involved. Biometrika pp. 28–35 (1947)

[29] Yuan, Y., Li, S., Zhang, X., Sun, J.: A comparative analysis of svm, naive bayes and gbdt for data faults detection in wsns. pp. 394–399 (2018)