



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJIRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Special Issue 2, December 2021

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.569

A Review on Automatic Classification of Non-Hodgkin Lymphoma

N. Sivaranjini¹, Dr. M.Gomathi²

Asst. Professor, Department of Computer Science, Auxilium College (Autonomous), Vellore, Tamil Nadu, India¹

Asst. Professor, Department of Computer Science, Government Arts and Science College, Kumarapalayam, Erode, Tamil Nadu, India²

ABSTRACT- Lymphoma is a cancer disease that initiates in lymphatic system, during infection fighting cells known as lymphocytes. The proof of identity of genetic and environmental factors is very vital in developing different strategies of cancer prevention. Many different methods and techniques have been considered for the processing and exploration of cancer expression profiling datasets. The accuracy of results are often reliant on to the measurement of approaches in practical. Several data processing methods have been proposed for the prediction of the Non Hodgkin Lymphoma Disease (NHL) cancer. This review presents the various techniques used for classification of NHL. The objective of this analysis is to evaluate the various techniques utilized in classifying the problem and to increase the accuracy of predicting the cancer in early stage.

KEYWORDS: Lymphoma, Data Mining, Neural Network, Prediction, Gene Data, Microarray, Segmentation.

I. INTRODUCTION

Cancer may be a doubtless serious disease affected principally by environmental factors that change genes encoding vital cell controlling proteins. The unusual cell behavior ends up in plenty of abnormal cells that destroy traditional tissue can spread into vital organs leading to disseminated disease. The histo-pathologic analysis is the most significant diagnostic criteria which analyses the morphologic features of tumor under microscope.

The tumor is divided into sections and slides and each slide is stained with hematoxylin and eosin. Lymphomas grow within the glands or nodes of the lymphatic system, a network of nodes, and organs, specifically in the spleen, tonsils, and thymus, that filter bodily fluids and turn out infection-fighting white blood cells, or lymphocytes. Lymphomas also are referred to as solid cancer, may additionally occur in specific organs like the abdomen, breast or brain. Lymphomas are also called as extra nodal lymphomas. Lymphoma is a fatal disease affecting the lymphatic system and it is classified in different sub-types are Chronic Lymphocytic Leukemia (CLL), Follicular Lymphoma (FL), and Mantle Cell Lymphoma (MCL). The cancers are unit sub-classified into two categories: Hodgkin and Non-Hodgkin lymphoma. The occurrence of Reed-Sternberg cells in Hodgkin cancer distinguishes Hodgkin cancer from Non-Hodgkin cancer.

II. RELATED WORK

A. Stages

There are four stages, the initial stage cancer is found in one of the lymph systems. The second stage is found in two or a lot of teams of body fluid nodes in two or a lot of teams of body fluid nodes on identical aspect of the diaphragm. The third stage cancer is in node areas on either side of the diaphragm. Fourth stage has developed wide into a minimum of one organ exterior the lymph system, like the bone marrow, liver, or lung. The symptoms of NHL are swollen glands, in neck, armpit or groin, cough, shortness of breath, fever, night sweats, fatigue, weight loss, itching.

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

B. Diagnosis

The diagnosis of lymphoma is done by excisional lymph node biopsy. NHLD within the glands or nodes concerned liquid body substance nodes are submandibular, higher cervix chain and posterior triangle areas. Cells grow out of control and become aggressive. Lymphomatous nodes inclined to be expanded with a minimum diameter of ten millimeter or higher in node size. Additional criteria for differentiating lymphomatous nodes are shape, border, echogenicity, and echogenic hilum. The following examinations are done to collect data sets through Lymph node biopsy, bone marrow aspiration, MRI, PET scan, molecular test, blood test.

Nuclei size	Mitotic Figures
≤ 1.5 then small	counts/40 \times field of ≤ 5 considered low
1.5-2.0 then medium	counts/40 \times field of 6–19 considered medium
≥ 2.0 then large	counts/40 \times field of ≥ 20 high

C. Machine Learning Techniques for NHLD Classification

Data Mining is the process of extracting the patterns from the huge data sets by combining the several data processing techniques. The data mining algorithms were broadly used to classify the cancer disease within the initial stage. There are many recent techniques were used to classify the cancer illness like supervised and unsupervised classification ways.

Data processing procedures supply great assurance to find patterns hidden within the data, that will facilitate the clinicians in decision making and analysis of varied applied data processing classification techniques is acceptable and might facilitate the medical professionals in assessment for early diagnosis. Advances in molecular classification of tumor might play a vital role in cancer treatment.

III. LITERATURE SURVEY

Yu Guo., et al [1] proposed the method that encodes PET images used as spatial images and sensory system images by executing supervised tumor separation and image rebuilding. This method was related with the traditional feature extraction method. Training set of 140 patient data images and testing set of 46 patient data images were used with 4-fold cross valuation was performed to evaluate.

Matheus G. Ribeiro., et al [2] analyzed the influence of color normalization in the classification of lymphoma images. This technique combines multidimensional fractal methods, curvelet changes and Haralick structures. The decision tree, random forest, support vector machine, naive bayes and kstar is used for evaluation. The three common lymphoma classes CLL, FL and MCL were analyzed. The tests were done using 10-fold cross-validation and the result was a rate 96% accuracy using random forest classifier.

Huiyan Jiang., et al [3] anticipated hierarchical prediction model for three NHLD extreme images, including CLL, FL and MCL. In this model all images are converted to grayscale and then classified into 130 non-overlapped patches with 100 \times 100 pixels. An unsupervised feature extraction method called the sparse autoencoder, is applied. For the prediction of NHLD as CLL, FL and MCL, 680-dimension feature set is applied. The label of each NHLD image is predicted by 130 patches.

Ahmed E-S. Negm., et al [4] introduced an automatic system with an user friendly interface to distinguish between the types of diffuse NHL cells based on structural features like size, perimeter and circularity. Digital microscopic images to measure structural parameters number of cells for image processing is used, such as overlapping, cell distortion that mark the sensitivity.

Borges H., et al [6] presents a evaluation of attribute selection methods to reduce the number of genetic factor to be

evaluated. The analysis of attributes subset attempts to define the consistency level on the attributes set tries to see the consistency level on the attribute subset. Bayesian net, Naive Bayes, k-NN. The corresponding classifier algorithmic rule utilized in the choice is used to classify attributes set. The result shows that subsets generated through the choice attribute the quantity of property classified instances is higher.

Belkacem-Boussaid., et.al [7] projected an automated method to acknowledge centroblast cells from non-centroblast cells for computer aided analysis of FL tissue samples. Training and testing of a Quadratic Discriminant Analysis classifier is used for prediction. A set of features direct few geometric and color features as training data for Quadratic Discriminant Analysis classifier. This method identifies the centroblast (CB) item with previous data. Therefore, the summary of Principal Component Analysis within the spectral region to identify color texture features. Geometric and texture features were used to bring the classification. Experimental results on actual follicular lymphoma disease pictures reveal that the combined feature enhanced the performance of the system considerably.

Siddharth., et.al [8] described an automatic classification to detect follicles in immunohistochemical (IHC)-stained tissue. The system utilizes texture and color feature to imitate the method of human expert to identify follicle regions. In CD10 and CD20-stained tissues, the regions of interest contain brown hues parted by blue or background regions. Large kernel size with median filter for leveling the gray scale image is found by the reduction in texture differences inside the follicle. The texture energy using co-occurrence matrix energy metrics were used as a feature vector to identify follicle region. The cluster corresponding to follicle region is detected by K-means clustering algorithm. The median filtered pixel value of all pixels assigned to the cluster. The result associates texture information and color information from the image for identifying follicle regions accurately 87.11%.

Nikita V., et.al [9] proposed an empirical study classifies the three types of NHLD CLL, FL, and MCL. The computer vision method is used for quantitative classification of image content. Two-stage approach was applied, the spectral planes were extracted by transforming the outer level raw pixels at first stage. The inner level, multipurpose global features were calculated from every spectral plane by the identical feature set, at the second stage. Hematoxylin and eosin stains were applied to the specimens. Common image features were analyzed from these color spaces using three different classifiers, through Weighted Neighbor Distance (WND) CHARM's, Radial Basis and a Naive Bayes Network.

Tim Peters., et.al [10] provides a supervised classification of tissue samples from microarray data by combining feature selection method and cross-entropy method. This hybrid algorithm feature selection based on cross-entropy simulates a competitive environment. The implementation of this algorithm allows sets of genes to separate classes in the training set to be derived by emergent properties of iterations. For gene selection linear discriminant analysis was used.

IV. DISCUSSION

The above survey provides the detailed description of classification of Lymphoma using various data mining techniques as depicted in Table 1.

S. No	Author Name	Methods Used	Dataset Used	Merits	Results	Year
1	Yu Guo, Pierre Decazes, Stéphanie Becker, Hua Li, SuRuan	SVM classifier	186 PET images	A deep neural network is used to encode PET images as spatial representations and modality representations by supervised tumor region segmentation	Accuracy 75%	2020
2	Matheus Ribeiro, Leandro A. Neves,	Color Normalization, Random forest	173 Tissue samples of H & E stained images	features were extracted from the curvelet sub-images. Random forest method classifies differentiate 3	Accuracy 85.03%	2018

	Guilherme F. Roberto, Tha'ina A. A. Tosta, Alessandro S. Martins and Marcelo Z. do Nascimento	classifier		classes (MCL, FL, CLL)		
3	Huiyan Jiang, Zhongkuan Li, Siqu Li, Fucan Zhou	SAE feature extraction method, ELM & SVM classifier	374 pathological NHL H & E stained images	Classification is done by an extreme learning machine (ELM) to classify FL from NHL. Then, CLL and MCL are divided into two classes using an SVM by determined via the predictive labels statistics of its 130 patches	Accuracy 97.6%	2018
4	Ahmed S.Negma, Ahmed H. Kandila, and Osama A. Hassana	Association rules, Rule induction and Deep learning	Blood Count samples	Classified as tumor blood and other blood diseases	Accuracy 80.83% using Deep Learning	2017
5	K. Belkacem-Boussaid, M. Pennell, G. Lozanski, A. Shana'ah, M. N. Gurcan	Quadratic Discriminate Analysis	Centroblast (CB) cells	Distinguishes CB from non-CB cells in follicular lymphoma using morphometric and color texture features in the spectral domain	Graphical User Interface, Sensitivity Accuracy 81.8%, and Specificity of Accuracy 86.4%	2010
6	Nikita V. Orlov, Wayne W. Chen, David Mark Eckley, Tomasz J. Macura, Lior Shamir, Elaine S. Jaffe, and Ilya G. Goldberg	Feature Selection Techniques	Microarray data	Indicative of lymphoma malignancies and classifying these malignancies by type	Classification Accuracy 99%	2010
7	Siddharth Samsi, Gerard Lozanski, Arwa Shana'ah, Ashok K. Krishnamurthy, and Metin N. Gurcan	K-means Clustering, Watershed Algorithm	Centroblast Image	Combines color and texture information from the image for identifying follicle regions	Accuracy 87.11%	2010 Accuracy 87.11 %

15th & 16th September 2021

Organized by

Department of Computer Science, Parvathy's Arts and Science College, Dindigul, Tamilnadu, India

8	Tim Peters, David W. Bulger, To-Ha Loi, Jean Yee Hwa Yang, and David Ma,	Linear discriminant analyses, Two-Step Cross-Entropy feature selection	Microarray data	Simple statistical method for assessing	Supervised classification of tissue samples from microarray data using Cross-Entropy feature selection	2011
---	--	--	-----------------	---	--	------

V. CONCLUSION

In this survey many data processing methods are mentioned in classification of lymphoma prediction. The data processing methods like Neural Network(NN), Bayesian Net, Naive Bayes, K-NN, T-test, Euclidean distances, Decision Trees, Feature Gene selection methods, Pattern recognition, Segmentation, Cross-Entropy, Clustering algorithms, fuzzy ARTMAP, FusionNET, UNet, ResNet, etc. The information gain is used to decrease a bias to multi valued attributes by taking the number of attributes. Finally, the 3D image datasets are used to classify the lymphoma disease chosen to improve the cancer classification. Each algorithmic program has importance and implemented on the behavior of the data, on the basis of classification of lymphoma. Various classification techniques area unit extremely acceptable and may facilitate the medical professionals for early diagnosis. This analysis work will be stretched to implement the innovative classification of image selection with higher accuracy and prediction.

REFERENCES

- [1] Guo, Y., Decazes, P., Becker, S., Li, H., & Ruan, S. (2020, April). Deep Disentangled Representation Learning of Pet Images for Lymphoma Outcome Prediction. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp. 1-4). IEEE.
- [2] Jiang, H., Li, Z., Li, S., & Zhou, F. (2018, October). An effective multi-classification method for nhl pathological images. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 763-768). IEEE.
- [3] Ribeiro, M. G., Neves, L. A., Roberto, G. F., Tosta, T. A., Martins, A. S., & Do Nascimento, M. Z. (2018, October). Analysis of the influence of color normalization in the classification of non-Hodgkin Lymphoma images. In *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)* (pp. 369-376). IEEE.
- [4] ES Negm, A., H Kandil, A., & A EF Hassan, O. (2017). Decision Support System for Lymphoma Classification. *Current Medical Imaging*, 13(1), 89-98.
- [5] Belkacem-Boussaid, Kamel, et al. "Extraction of color features in the spectral domain to recognize centroblasts in histopathology." 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2009.
- [6] Lisboa, Paulo JG, et al. "Data mining in cancer research [application notes]." *IEEE computational intelligence magazine* 5.1 (2010): 14-18.
- [7] Belkacem-Boussaid, Kamel, et al. "Effect of pathologist agreement on evaluating a computer-aided assisted system: recognizing centroblast cells in follicular lymphoma cases." 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE, 2010.
- [8] Samsi, Siddharth, et al. "Detection of follicles from IHC-stained slides of follicular lymphoma using iterative watershed." *IEEE Transactions on Biomedical Engineering* 57.10 (2010): 2609-2612.
- [9] Orlov, Nikita V., et al. "Automatic classification of lymphoma images with transform-based global features." *IEEE Transactions on Information Technology in Biomedicine* 14.4 (2010): 1003-1013.
- [10] Peters, Tim, et al. "Two-step cross-entropy feature selection for microarrays—power through complementarity." *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 8.4 (2011): 1148-1151.
- [11] Zillner, Sonja. "Using image metadata for automatic patient staging in Lymphoma." 2011 IEEE International Conference on Information Reuse & Integration. IEEE, 2011.



- [12] Kharya, Shweta. "Using data mining techniques for diagnosis and prognosis of cancer disease." arXiv preprint arXiv:1205.1923 (2012).
- [13] Yoon, Hee-Jin, and Joon S. Lim. "Lymphoma cancer classification using NEWFM based Filtering method." 2013 International Conference on Information Science and Applications (ICISA). IEEE, 2013.
- [14] Patil, Bhagyashri G., and Sanjeev N. Jain. "Cancer cells detection using digital image processing methods." International Journal of Latest Trends in Engineering and Technology 3.4 (2014): 45-49.
- [15] Arutchelvan, K., and Dr R. Periyasamy. "Cancer Prediction System Using Data mining Techniques." International Research Journal of Engineering and Technology 2.8 (2015).
- [16] Singh, Neelam, and Santosh Kumar Singh Bhadauria. "Early Detection of Cancer Using Data Mining." International Journal of Applied Mathematical Sciences 9 (2016): 47-52.
- [17] Amaya, H. A., et al. "Biological target volume delineation for a pediatric non-Hodgkin lymphoma: A semi-automated segmentation." 2018 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE). IEEE, 2018.
- [18] Bolluyt, Elijah, Alexandra Comaniciu, and Bogdan Georgescu. "Tumor Burden Assessment In Lymphoma Patients: Hierarchical Analysis Of Whole Body CT." *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019.
- [19] Brancati, Nadia, et al. "A Deep Learning Approach for Breast Invasive Ductal Carcinoma Detection and Lymphoma Multi-Classification in Histological Images." *IEEE Access* 7 (2019): 44709-44720



**Impact Factor:
7.569**



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



www.ijirset.com

Scan to save the contact details