

# Detection and Classification for Blood Cancer – A Survey

Kuntal Barua<sup>1</sup>, Prasun Chakrabarti<sup>2</sup>

<sup>1</sup>Research Scholar, Dept. of Computer Science & Engineering, Sir Padampat Singhania Univesity, Udaipur-331601, Rajasthan, India

<sup>2</sup>Professor & Head, Dept. of Computer Science & Engineering, Sir Padampat Singhania Univesity, Udaipur-331601, Rajasthan, India

**Abstract**— The paper entails an idea to develop an automated method of analysis of AML blast cell images and to include in image-processing software, which enables the haematologist to diagnose AML more effectively and efficiently. Haematologists often face difficulties identifying the subtypes of AML, due to the similarities of their morphological features. Following AML detection, blast cells need to be classified into M3 or one of the other subtypes. The reason for targeting M3 is that its treatment differs from the treatment of the rest, requiring All-Trans-Retinoic-Acid (ATRA) to be added to the initial chemotherapy.

**Keywords**— Blood cancer, ALL, AML, CML, Haematologist.

## I. INTRODUCTION

Blood cancer is cancer of the blood cells. Blood cancer named as leukemia, is a type of cancer of the blood or bone marrow categorized by an irregular amount of undeveloped white blood cells called "blasts." It is a thick term covering a compilation of diseases. According to American Cancer Society it is approximated that 48,610 persons (27,880 men and 20,730 women) will be detected with and 23,720 men and women will terminate of blood cancer. In turn, it is part of the even broader set of diseases disturbing the blood, bone marrow, and lymphoid system, which are all known as hematological neoplasm. Over time, blood cancer cells can crowd out the normal blood cells. This can lead to serious problems such as anemia, bleeding, and infections. Blood cancer cells can also spread to the lymph nodes or to other organs and cause pain or swelling. There are several different types of blood cancer. It starts in bone marrow, the tissue (soft) inside most bones. Bone marrow is cells where blood cells are made. White blood cells help your body fight infection. Red blood cells carry O<sub>2</sub> to all parts of our body. Platelets help our blood clot. When you have blood cancer, the bone marrow starts to make a lot of white blood cells (abnormal), called blood cancer cells. They don't do the work of normal white blood cells. They grow faster than normal cells, and they don't stop growing when they should. Over time, blood cancer cells can crowd out the normal blood

cells. This can lead to serious problems such as anemia, bleeding, and infections. Blood cancer cells can also spread to the lymph nodes or other organs and cause swelling or pain. There are several different types of blood cancer. In general, blood cancer is grouped by how fast it gets worse and what kind of white blood cell it affects. It may be acute or chronic. Acute blood cancer gets worse very fast and may make you feel sick right away. Chronic blood cancer gets worse slowly and may not cause symptoms for years. It may be lymphocytic or myelogenous. Lymphocytic (or lymphoblastic) blood cancer affects white blood cells called lymphocytes. Myelogenous blood cancer affects the other type of cells that normally become granulocytes, red blood cells, or platelets. The four main types of blood cancer are: Acute lymphoblastic blood cancer, or ALL, Acute myelogenous blood cancer, or AML, Chronic lymphocytic blood cancer, or CLL and Chronic myelogenous blood cancer, or CML.

## II. CAUSES AND RISK FACTORS OF BLOOD CANCER

The satisfactory causes of blood cancer are unidentified and in most cases its unsettled why blood cancer has developed. Research into possible causes is going on all the time. Like other cancers, blood cancer isn't transferable and can't be approved on to other people. There are several of factors that may amplify a person's risk of budding blood cancer. Having a scrupulous hazard factor doesn't denote you will definitely get this category of disease and personnel lacking any recognized risk factors can still develop it. The recognized risk factor of generate this type of cancer i.e. blood cancer are clarify here.

(a) **Exposure to radiation:** People who exposed to high level of release, such as nuclear developed accidents, have a main risk of developing blood cancer than people who have not been exposed. On the other hand, a small numeral of people in the UK will be uncovered to emission levels high adequate to augment their risk.

(b) **Smoking:** Smoking increase the risk of initial blood cancer. This may be due to the intense levels of benzene in cigarette smoke.

(c) **Exposure to benzene:** In very unusual cases, blood cancer may begin due to the long term contact to benzene (and possibly other solvents) used in industry.

(d) **Cancer treatments:** Now and then, a few anti-cancer treatments such as chemotherapy or radiotherapy can be a basis for blood cancer to build up after some years of this behavior. The risk increase when persuaded types of chemotherapy drugs are mutual with radiotherapy. While blood cancer develops since of earlier anti-cancer treatment, this is called lower blood cancer or treatment related blood cancer.

(e) **Blood disorders:** People with certain blood disorders, such as myelodysplasia or myeloproliferative disorders have a distended risk of initial AML.

(f) **Genetic disorders:** People with certain hereditary disorder, excluding Down's syndrome and Franconia's anemia, have an inflated risk of embryonic blood cancer.

**Blood** is fundamental to human life. An average human body is approximately 70 litres of which five litres are blood. Biologically, blood is essential for maintaining homeostasis, that is keeping the body's status stable. This refers to temperature regulation, hydration and ion concentration. The main blood functions include.

- (a) Delivery of nutrients from the digestive system to all parts of the body.
- (b) Transport of oxygen from the lungs to all parts of a body.
- (c) Transport of CO<sub>2</sub> from all parts of the body to the lungs.
- (d) Waste products transportation from cells to the external environment, especially via the kidneys.
- (e) Maintaining an ongoing "discussion" of it is components with tissue fluids and keeping electrolyte balance.
- (f) Protecting the body against attack from foreign organisms through the white blood cells and antibodies.
- (g) Defending the body against injury or illness using the inflammatory response.
- (h) Preventing serious haemorrhage by the clotting process.
- (i) Maintaining the body's temperature by circulating heat.

**White Blood Cell** is larger than a red blood cell. White blood cell composition and concentration in the blood gives important information and plays a crucial role in the diagnosis of different diseases. White blood cells fall into 5 categories: Neutrophil, Eosinophil, Basophil, Monocyte and Lymphocyte, shown in the Table. These cells provide the greatest

defense against infections, and their individual concentrations can help specialists to distinguish between the presence or not of severe pathologies. The size of white blood cells will be used for the filtering process.

### III. PROBLEM STATEMENT

The microscopic study of human blood has led to the conclusion that a set of methods, including microscope colour imaging, segmentation, classification, and clustering can allow the identification of patients suffering from leukaemia . Machine learning is one of the methods used in image processing for detection of blast cells. This research focuses on AML and the classification of the eight subtypes mentioned previously, M0 to M7. This is an important issue since they require different treatment. All subtypes share extensive similarities which makes identification under a microscope difficult, time-consuming and physically tiring for the haematologists.

### IV. LITERATURE REVIEW

In a paper named "A Survey on Image Segmentation Techniques Used In Blood cancer Detection [1]" by Mashiat Fatma, Jaya Sharma, it has been observed that Image segmentation commonly known as partitioning of an image is one of any image processing technique's intrinsic parts. In image processing step, the digital image of choice is segregated into sets of pixels on the basis of some pre-defined and pre-selected standards or measures. There have been presented many algorithms for segmenting an image (digital). This research paper presents a general review of algorithms that have been presented for the purpose of image segmentation.

Then in the next paper named "A fuzzy-Neural Approach for Blood cancer Cancer Classification [2]" by Dr. B.B.M. Krishna Kanth, it had been stated that using the Fuzzy Hyper sphere neural network (FHSNN) classifier for the dis-crimination of acute lymphoblastic blood cancer (ALL) and acute myeloid blood cancer (AML) subtypes present in the blood cancer dataset. Prior to classification as the number of genes are larger in number compared to the samples which is available in the microarray datasets hence, to find the best features(genes) for classification dimensionality reduction methods such as SNR (Signal-to-Noise Ratio), Class-Separability, Wilcoxon rank sum statistic and Fisher Ratio are used. The experimental results show that his FHSNN is able to achieve 100% accuracy with much fewer genes than the previously published methods did.

Here in other paper named "Classification of Blood cancer Blood Samples Using Neural Networks [3]" by Malek Adjouadi, Melvin Ayala, Mercedes Cabrerizo, Nuannuan Zong, Gabriel Lizarraga And Mark Rossman, it had been seeing

that A novel artificial neural network (ANN) algorithm is proposed for optimizing the classification of multidimensional data, focusing on acute blood cancer samples. The programming tool established around the ANN architecture focuses on the classification of abnormal vs. normal blood samples, namely acute lymphocytic blood cancer (ALL) and acute myeloid blood cancer (AML). Here were 220 blood samples considered with 60 abnormal samples and 160 normal samples. The algorithm produced too high (sensitivity) results that improved up to 96.67% in ALL classification with increased data set size. With this type of accuracy, the tool (programming) provides information to medical doctors in the form of diagnostic references for disease (specific) states that are considered for this study. The results obtained prove that a NN classifier (neural network) can perform remarkably well for this type of flow-cytometry data. Even more significant is the fact that the evaluations (experimental) in the testing phase reveal that as the ALL data considered is gradually increased from small to big data sets, the more accurate are the classification results.

Detection Of Blood cancer In Human Blood Sample Based On Microscopic Images: A Study [4] by Fauziah Kasmin, Anton Satria Prabuwo, Azizi Abdullah said that This paper describes a preliminary study of developing a detection of blood cancer types using microscopic blood sample images. Analyzing images is very important as from images, diseases can be detected and diagnosed at earlier stage. From there, further actions (controlling, monitoring and prevention of diseases) can be done. Images are used as they do not require expensive testing, lab equipments and are cheap also. The system will focus on white blood cells disease, blood cancer. The system will use features in microscopic images and examine changes on texture, geometry, statistical-analysis and color. Changes in these features will be used as a classifier input. A literature review has been done and a learning method called Reinforcement Learning, is proposed to classify types of blood cancer.

Acute Blood cancer Classification using Bayesian Networks [5] by Abdel Nasser H. Zaied, 2 Mona G. Hebshy, 3 Mohamed A. Saleh, here it had been observed that two models for constructing acute blood cancer classifiers using the Signal-to-Noise Ratio (SNR) gene selection method in conjunction with the BNs (Bayesian Networks) have been proposed. In the first model, genes of the acute blood cancer training dataset are ranked using the SNR method and then top ranked genes are selected and used to construct the acute blood cancer BN classifier. In the second model, genes of the acute blood cancer training dataset are clustered using the a clustering technique named k-means clustering and then genes of each cluster are ranked using the Signal-to-Noise Ratio method after that top ranked

genes from clusters (gene) are selected and used to construct the acute blood cancer BN classifier. From the experimental evaluation, then the productive results showed that the classification accuracies achieved by the acute blood cancer classifiers constructed according to either of these two models are compared as good with the classification accuracies achieved in other studies. Here the results also indicated that a second model is much better than the first model.

Pooja Deshmukh, Prof.C.R.Jadhav wrote in A Survey on Detection of Blood cancer Using White Blood Cell Segmentation [6] that white blood cells analysis is generally done for diagnosis of various diseases. One of those diseases is Acute lymphoblastic blood cancer (ALL). ALL is detected by observing morphological changes in white blood cells. Morphological analysis along with classification and segmentation techniques helps to detect blood cancer at early stage and accurate detection. There are number of classification techniques which can be used to classify WBC's into different classes as per their respective features. Segmentation techniques segments nucleus and cytoplasm from each WBC and feature extraction process accurate features from nucleus and cytoplasm for accurate result.

Comparative Study Using Weka for Red Blood Cells Classification [7] by Jameela Ali Alkrimi, Hamid A. Jalab, Loay E. George, Abdul Rahim Ahmad, Azizah Suliman, Karim Al-Jashamy presented a comparative study for various techniques for classifying the RBCs as normal anemic or abnormal using WEKA which is an open source consists of different machine learning algorithms for data mining applications. The tested algorithms are Radial Basis Function NN (neural network), Support vector machine, and K-Nearest Neighbors algorithm. Here 2 sets of combined features were utilized to classify blood cells images. The first set which is consists of geometrical features was used to identify whether the tested blood cell has a spherical shape or non-spherical cells. While the second set, mainly consist of textural features was used to recognize the types of the spherical cells.

Gene Expression Based Acute Blood cancer classification - a neuro fuzzy approach [8] by B. B. M. Krishna Kanth, U. V. Kulkarni & B. G. V. Giridhar proposed the Modified Fuzzy Hypersphere Neural Network (MFHSNN) for the discrimination of acute lymphoblastic blood cancer (ALL) and acute myeloid blood cancer (AML) in blood cancer dataset. Dimensionality reduction methods which are like Spearman Correlation Coefficient and Wilcoxon Rank Sum Test are used for gene selection. The performance of the Modified Fuzzy Hypersphere Neural Network system is encouraging when benchmarked against those of Support vector machine (SVM) and the K-nearest neighbour (KNN) classifiers. A classification accuracy of 100% has

been achieved using the MFHSNN classifier using only two genes.

Morteza Moradi Amin, Saeed Kermani, Ardeshir Talebil, Mostafa Ghelich Oghli discussed on a paper named “Recognition of Acute Lymphoblastic Blood cancer Cells in Microscopic Images Using K-Means Clustering and Support Vector Machine Classifier [9]” that Acute lymphoblastic blood cancer is the most common form of pediatric cancer which is classified into three L1, L2, and L3 and could be detected through screening of blood and bone marrow smears by pathologists. Due to time-consuming and sameness of the procedure, a computer based system is acquired for convenient detection of Acute lymphoblastic blood cancer. Microscopic images are acquired from blood and bone marrow smears of patients with Acute lymphoblastic blood cancer and normal cases. After applying image pre-processing, cells nuclei are categorized by k-means algorithm. Then geometric and statistical features are extracted from nuclei and finally these cells are divided to cancerous and noncancerous cells by means of support vector machine classifier with 10 fold cross validation. These cells are also typed into their subtypes by multi-Support vector machine classifier. Classifier is being evaluated by these parameters: specificity, sensitivity and accuracy which values for cancerous and noncancerous cells 95%, 98% and 97%, respectively. These parameters are also used for evaluation of cell sub-types which values in mean 97.3%, 84.3% and 95.6%, respectively. The results show that proposed algorithm could achieve an performance which is acceptable for the diagnosis of Acute lymphoblastic blood cancer and its sub-types and can be used as an assistant diagnostic tool for pathologists.

A review paper named “Review - Detection Of Types Of Acute Blood cancer [10]” by Shrutika Mahaja, Snehal S. Golait, Ashwini Meshram, Nilima Jichkan proposed a system uses features of microscopic images by auditing changes like texture, geometry, colour and statistical analysis of images. These changes will be used as a classifier input. The presented process shows how effective an automatic morphological method to identify the Acute Lymphocytic Blood cancer (ALL) by microscope images of blood samples. At first the system individuates the leucocytes present in others blood cells, after that it admits the lymphocyte cells (cells that causes acute blood cancer), evaluation regarding morphological indexes from cells is done and lastly classification for the presence of the blood cancer is done.

Sulaja Sanal studied in Automated detection of acute lymphocytic blood cancer - a survey [11] Some of the techniques of image processing for detection automatically of ALL (Acute Lymphocytic Blood cancer). Automated detection of ALL minimizes the overhead of manual blood smear

processing. The efficiency also increases by using the automated techniques. Many new techniques are introduced in this field for enhancement and among that image processing techniques are most popular nowadays. ALL is the most common type of blood cancer in children. It is fatal if left untreated. The early detection of ALL is an factor for the right treatment. The manual checking of blood smear is time consuming and depends on the operator’s ability. Hence automatic techniques are introduced to enhance the performance. In this paper different automated method used for the detection of ALL is described.

Gene Expression Based Blood cancer Sub-Classification Using Committee Neural Networks [12] by Mihir S. Sewak, Narender P. Reddy and Zhong-Hui Duan described that Analysis of gene expression data gives an objective and fluent technique for sub-classification of blood cancer. The purpose of the present study was to design a committee NN based classification systems to subcategorize blood cancer gene expression data. In this study, a binary classification system is being described which was considered to differentiate acute lymphoblastic blood cancer from acute myeloid blood cancer. A ternary classification system which classifies blood cancer expression data into three subparts including B-cell acute lymphoblastic blood cancer, T-cell acute lymphoblastic blood cancer and acute myeloid blood cancer was also developed. In each classification system gene expression profiles of blood cancer patients were first subjected to sequenced simple pre-processing steps. The committee neural network system was later evaluated using data not used in training. The binary classification system divided microarray gene expression profiles into two categories with 100 percent accuracy and the ternary system correctly predicted the three subclasses of blood cancer in over 97 percent of the cases.

In Acute Lymphocytic Blood cancer Detection from Blood Microscopic Images [13], Sulaja Sanal, Lashma. K, Viji Balakrishnan said that, the classification of blood cells is important for the evaluation and diagnosis of most bug in medical diagnosis systems. Acute Lymphocytic Blood cancer (ALL) is a type of childhood blood cancer which is mostly seen in children below 7-8 years. It may be harmful if left untreated and causes death. Detection of Acute Lymphocytic Blood cancer may be done through the analysis of white blood cells which is also called as leukocytes. Usually the analysis of blood cells is functioned manually by skilled operators. This manual techniques have great drawbacks, such as slow study and a non-standard accuracy. It all depends on the operator’s skill. Hence many automatic systems are using in order to evaluate and classify the blood cells, but most of them produce only partial results. The main steps of this work are image pre-processing, WBC extraction,



separation of adjacent white blood cells, the process of feature extraction and classification. Image pre-processing is done by converting RGB images into lab colour space images. It is done to enhance the visual presence of the image and to reduce the memory requirements. Then only white blood cells are identified by using fuzzy C means grouping algorithm. Adjacent WBCs are a major threat while performing the process of feature extraction in the later stages. For avoiding that, separation of adjacent leukocytes is executed by using Marker based watershed distribution. For feature extraction, the features of WBC such as area, energy, entropy etc. are studied. To detect whether the patient has blood cancer or not, where a neuro-fuzzy classifier is used.

Based on gene expression, in Cancer classification based gene expression by neural networks [14], H.P. Hu, Z.J. Niu, Y.P. Bai and X.H. Tan have classified 53 colon cancer patients into two groups: relapse and no relapse. Some samples were taken from every patient, and gene information was derived. Of the 53 samples examined, 500 genes were studied through analyses by S-Kohonen, BP, and SVM neural networks. Classification accuracy obtained by S-Kohonen NN reaches 91%, which was more accurate than classification by BP and SVM NN. In this paper, the results show that S-Kohonen neural network is more plausible for classification and has a certain feasibility and validity.

## V. RESEARCH GAP

The motivation behind the research in this work, that is, the difficulties faced by haematologists in identifying subtypes of AML by visual inspection of microscopic images. There are eight subtypes of AML, with different morphological features. However, given that they also share extensive similarities diagnosis can be time consuming. The aim of this research is the development of an automated method for the detection of blast cells in images of blood smear and bone marrow microscope slides and their subsequent classification into AML subtypes. The most important issue is the classification of an image into M3 or the group of the remaining subtypes, as AML M3 requires different treatment. The approach proposed in this work are based on Artificial Intelligence and include Cellular Automata, Heuristic Search and Neural Networks for auto-detection and classification of blast cells. A software application will be developed to assist haematologists to diagnose AML more effectively and efficiently.

## VI. CONCLUSIONS

The essence is to develop an automated method of analysis of AML blast cell images and to include in image-processing software, which enables the haematologist to diagnose AML more effectively and efficiently. Haematologists often face difficulties identifying the subtypes of AML, due to

the similarities of their morphological features. Following AML detection, blast cells need to be classified into M3 or one of the other subtypes. The reason for targeting M3 is that its treatment differs from the treatment of the rest, requiring All-Trans-Retinoic-Acid (ATRA) to be added to the initial chemotherapy.

## ACKNOWLEDGMENTS

The authors wish to acknowledge all researchers and their valuable work on blood cancer from where they get an idea to develop an automated method of analysis of AML blast cell images and to include in image-processing software, which enables the haematologist to diagnose AML more effectively and efficiently. Specially we would like to thank Mr. Waidah Ismail from Brunel University for support.

## REFERENCES

- [1] Mashiat Fatma, Jaya Sharma, A Survey on Image Segmentation Techniques Used In Blood cancer Detection, *Int. Journal of Engineering Research and Applications* www.ijera.com ISSN : 2248-9622, Vol. 4, Issue 5( Version 6), May 2014, pp.66-71.
- [2] Dr. B.B.M. Krishna Kanth, A fuzzy-Neural Approach for Blood cancer Classification, *International Journal of Scientific Research* Volume : 2 | Issue : 11 | November 2013 • ISSN No 2277 – 8179.
- [3] Malek Adjouadi, Melvin Ayala, Mercedes Cabrerizo, Nuannuan Zong, Gabriel Lizarraga And Mark Rossman, Classification of Blood cancer Blood Samples Using Neural Networks, *Annals of Biomedical Engineering*, Vol. 38, No. 4, April 2010. pp. 1473–1482.
- [4] Fauziah Kasmin, Anton Satria Prabuwo, Azizi Abdullah, Detection Of Blood cancer In Human Blood Sample Based On Microscopic Images: A Study, *Journal of Theoretical and Applied Information Technology* ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195.
- [5] Abdel Nasser H. Zaied, 2 Mona G. Hebishy, 3 Mohamed A. Saleh, Acute Blood cancer Classification using Bayesian Networks, *Journal of Emerging Trends in Computing and Information Sciences* VOL. 3, NO.10 Oct, 2012 ISSN 2079-8407.
- [6] Pooja Deshmukh, Prof.C.R.Jadhav, A Survey on Detection of Blood cancer Using White Blood Cell Segmentation, *International Journal of Modern Trends in Engineering and Research (IJMTER)* Volume 02, Issue 12, [December – 2015] ISSN (Online):2349–9745 ; ISSN (Print):2393-8161.
- [7] Jameela Ali Alkrimi, Hamid A. Jalab, Loay E. George, Abdul Rahim Ahmad, Azizah Suliman, Karim Al-Jashamy, Comparative Study Using Weka for Red Blood Cells Classification, *World Academy of Science, Engineering and Technology - International Journal of Medical, Health, Biomedical, Bioengineering and Pharmaceutical Engineering* Vol:9, No:1, 2015.
- [8] B. B. M. Krishna Kanth, U. V. Kulkarni & B. G. V. Giridhar, Gene Expression Based Acute Blood cancer classification - a neuro fuzzy approach, *International Journal of Biometrics and Bioinformatics, (IJBB)*, Volume (4): Issue (4).
- [9] Morteza Moradi Amin, Saeed Kermani, Ardeshir Talebi1, Mostafa Ghelich Oghli, Recognition of Acute Lymphoblastic Blood cancer Cells in Microscopic Images Using K-Means Clustering and Support Vector Machine Classifier, *Journal of Medical Signals & Sensors* Vol 5 | Issue 1 | Jan-Mar 2015.
- [10] Shrutika Mahaja, Snehal S. Golait, Ashwini Meshram, Nilima Jichkan, Review - Detection Of Types Of Acute Blood cancer, *International Journal of Computer Science and Mobile Computing*, Vol.3 Issue.3, March- 2014, pg. 104-111.

- [11] Sulaja Sanal, Automated detection of acute lymphocytic blood cancer - a survey, International Journal of Engineering Research and General Science Volume 3, Issue 3, Part-2, May-June, 2015 ISSN 2091-2730.
- [12] Mihir S. Sewak, Narender P. Reddy and Zhong-Hui Duan, Gene Expression Based Blood cancer Sub-Classification Using Committee Neural Networks, Libertas Academia Freedom to Research - Bioinformatics and Biology Insights 2009.
- [13] Sulaja Sanal, Lashma. K, Viji Balakrishnan, Acute Lymphocytic Blood cancer Detection from Blood Microscopic Images, International Journal of Engineering Research and Technology (IJERT) ISSN: 2278-0181, Vol 4 Issue 09, September 2015.
- [14] H.P. Hu, Z.J. Niu, Y.P. Bai and X.H. Tan, Cancer classification based on gene expression using neural networks, Genet. Mol. Res. 14 (4): 17605-17611 (2015).