

DETECTION OF BLOOD CANCER CELLS USING SVM ALGORITHM

Gaddamidi Ramyasri¹, Gottipati Bhavana², V Rakesh³

¹ Student, Information Technology, B V Raju Institute of Technology, Telangana, India

² student, Information Technology, B V Raju Institute of Technology, Telangana, India

³ Associate Professor, Information Technology, B V Raju Institute of Technology, Telangana, India

ABSTRACT

Cancer is a fatal illness. A type of cancer known as blood cancer targets the lymphatic, bone marrow, or blood systems. Hematologists do a blood test to diagnose it, during which certain blood cell types are examined. Currently, there are numerous varieties of blood cancer cells like Leukaemia, Lymphoma, Myeloma, etc. Leukemia is a type of blood cancer that develops when White Blood Cells (WBCs) are abnormal or immature. We only looked at acute myelogenous leukaemia, which is one of the blood cancer types that falls under the category of acute leukaemia and most frequently affects adults. Leukemia must be automatically recognized since when doctors see abnormal blood cells under a microscope using CBC (Complete Blood count), it requires a lot of manual work and time so it may harmful to the patient. In recent years, Support Vector Machines (SVM) have gained popularity in the field of medical image analysis due to their ability to accurately classify and segment data. In this study, we propose an approach to detect blood cancer cells using SVM. The proposed method involves preprocessing of the blood smear images to extract features, such as shape, texture, and color. The extracted features are then used as input to train the SVM classifier, which classifies the cells as normal or abnormal. The performance of the SVM classifier is evaluated using metrics such as sensitivity, specificity, accuracy, and F1-score. The proposed approach has the potential to aid in the early detection and diagnosis of blood cancer, which could improve the overall survival rate of patients.

Keyword: SVM, CBC, Leukaemia, White blood cells.

1. INTRODUCTION

Blood cancer is the condition that leads to the greatest losses in humans. Blood cancer is a form of cancer that affects the blood, bone marrow, or lymphatic systems. The bone marrow, where blood is made, is where the majority of blood cancers, also known as hematologic tumours, begin. Blood cancers develop when abnormal blood cells begin to expand uncontrollably and interfere with the regular blood cells' ability to fend off infection and generate new blood cells. It is diagnosed by hematologists by a blood test in which certain blood cell types are inspected. There are many possible ways for a human to effect with blood cancer like increasing age, substandard immune system, consuming food items with more chemicals which effects our blood cells, radiation treatments, genetic disorders etc. Leukemia accounts for more than 3.5% of all new cancer cases in the United States, where more about 60,000 new cases were recorded in 2018.

Therefore, the majority of the diagnosing process has utilized machine learning techniques over the past two decades to increase the life expectancy of patients and provide the results of the diagnosis in a timely manner. Machine learning is a branch of Artificial Intelligence. The discipline of image analysis greatly benefits from machine learning. Machine learning approaches are used to evaluate visual data for issues such as picture segmentation, object recognition, and scene understanding.

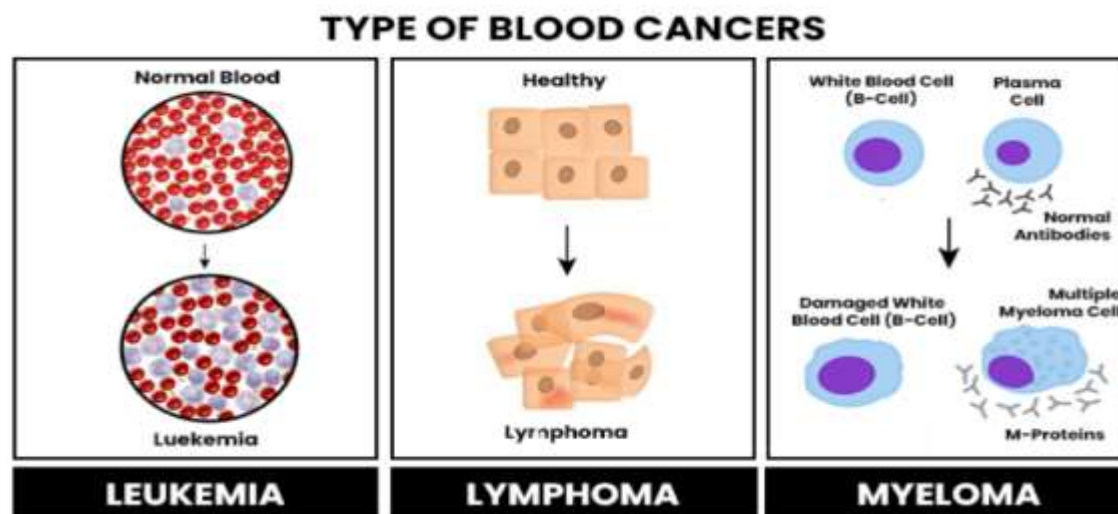


Fig-1: Types of blood cancers

Fig-1 displays the various blood cancers as well as the differences between healthy blood tissues and cells and cancerous tissues.

1.1 Support Vector Machine

SVM (Support vector Machine) is a supervised machine learning method that solves problems including classification and regression. It is mostly applied for classification. Cancer cells are detected via classification after undergoing many processes such as image pre-processing, image segmentation, and feature extraction.

There are two SVM subtypes. There are linear SVM and non-linear SVM.

Linear SVM: When a dataset can be divided into two groups using only a single straight line, it is said to be linearly separable, and the linear SVM classifier is used to separate the data into these two classes.

Non-linear SVM: When a dataset cannot be identified using a straight line, it is said to be non-linear, and the classification algorithm utilized is known as a non-linear SVM classifier. SVM classification utilizes vector machine algorithms and hyperplane. By extending the decision boundary's margin, SVM carries out capacity control. SVM technique trains the dataset using support vectors and hyperplane by utilizing the blood samples dataset. **Hyperplane:** In n-dimensional space, there may be several lines or decision boundaries used to separate the classes, but we must choose the optimum decision boundary that best aids in classifying the data points. The hyperplane of SVM is a term for this optimal boundary.

Support Vector: Support vectors are the data points or vectors that seem to be closest to the hyperplane and have the most influence on where the hyperplane is located. These vectors are termed support vectors because they support the hyperplane. Using the processed image, the model is trained using the SVM technique to determine if the patient has blood cancer.

2. METHODOLOGY

The SVM algorithm is used in the proposed system for the identification of blood cancer cells, which is based on microscopic pictures of blood samples. For the purpose of putting this proposed approach into practice, we employed Google colab as a software tool. 50 photos of healthy blood cells and 50 images of infected blood cells were captured using a microscope to create 100 tiny photographs of blood cells. This classification is followed by few steps:

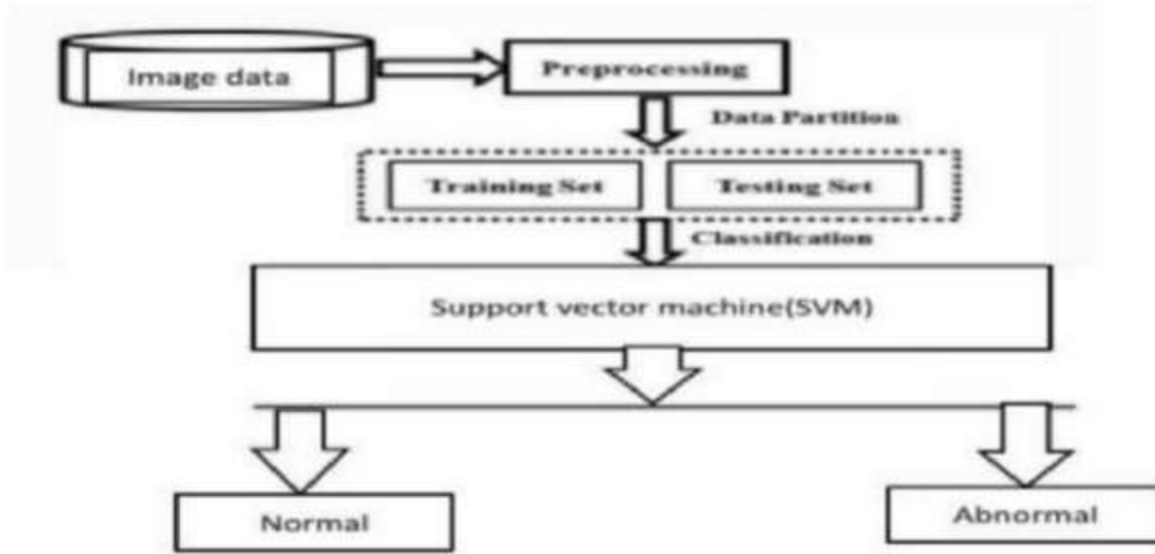


Fig-2: Architecture of Proposed System

Fig. 2 shows the proposed System's architecture. This figure shows the system's suggested process step by step.

2.1 Acquisition of Images

The initial phase involves gathering microscopic photographs of blood cells made up of various pixels, which will subsequently be represented as a 2D matrix.

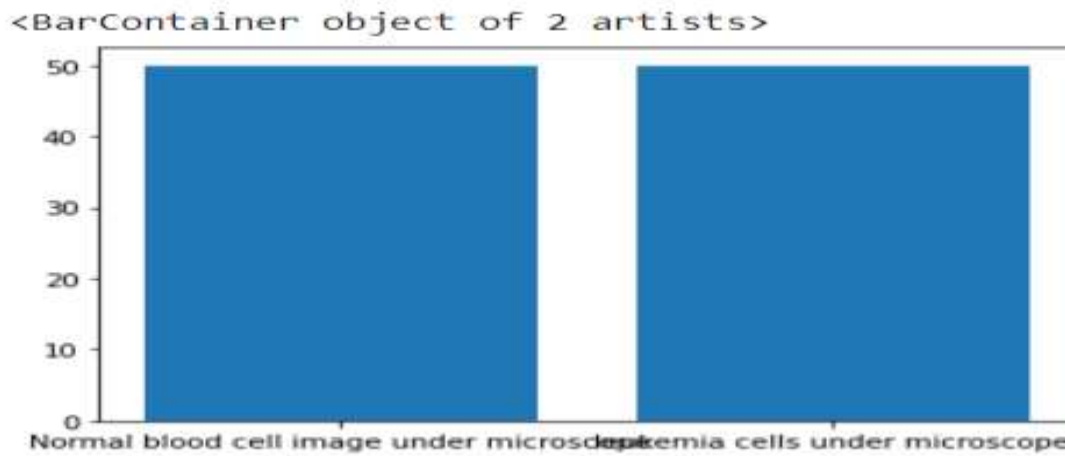


Chart -1: Number of images in two datasets

2.2 Pre-Processing the image

Noise from the input image is primarily removed during the pre-processing step (data). The pre-processing stage in this system primarily includes two activities.

- i)Resize: Although the input photos are of varying sizes, during the pre-processing stage, all input images are converted into fixed-size pixels. All of the input photographs are reduced in size in the suggested system to (150,150,3) pixels.

ii) Image Flattening: Flattening is the process of converting a 2D matrix into a 1D array. The flattened data are used in further classification processes.

2.3 Split the dataset into training and testing data

We can split our pre-processed dataset which undergone feature extraction and segmentation stages into a training dataset and a test dataset when we do a train-test split. The test dataset can be used to test the model after it has been trained using the training set. As a result, you may build your models upon this training dataset and check their accuracy on the test set that hasn't been seen. This stage involves dividing the full data set into training and testing data. Testing data makes for 30% of the proposed system.

2.4 Training and Fitting the model

Building the model will take place in this phase. In SVM, building the model uses some kernel parameters like C and gamma, where C is a "Regularisation parameter" that is used to adjust the model's tolerance for allowing data points to be incorrectly classified in order to reduce Generalisation error. In this proposed system, a linear kernel is considered for classification. Linear kernel is nothing but separating the data linearly by a single line which is used for classifying the hyperplane. Training the data also uses the extracted features of the data.

2.5 Classification

The outcome is classified at this stage using a trained model. When a new blood sample is submitted for analysis, it may be predicted if the sample is made up of normal blood cells or whether it could have blood cancer

3. RESULTS

The proposed system's results are shown in the figures below.

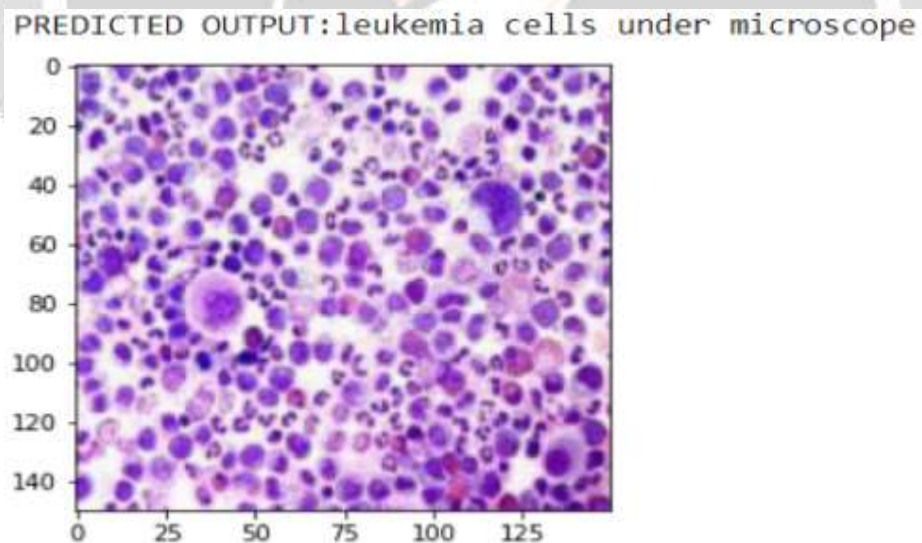


Fig-3: Predicted outcome of Lueukemia cells

PREDICTED OUTPUT: Normal blood cell image under microscope

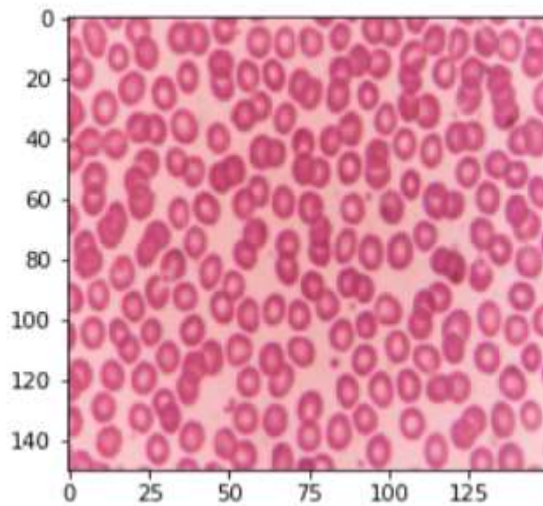


Fig-4: Predicted outcome of normal blood cells

4. CONCLUSIONS

A crucial step in the diagnosis of blood-related disorders is the detection of blood cancer cells. Early diagnosis of the illness has the potential to improve treatment and life expectancy. This paper suggested an approach for classifying blood cancer cells from a sample of human blood utilizing pre-processing, feature extraction techniques, and model-building with the SVM algorithm. The accuracy achieved by the suggested approach was up to 85.33%. This suggested system has the ability to identify between normal and diseased blood cells.

5. REFERENCES

- [1]. Sachin Paswan, Yogesh Kumar Rathore, "Detection and Classification of Blood Cancer from Microscopic Cell Images Using SVM KNN and NN Classifier" ,International Journal of Advance Research, Ideas and Innovations in Technology.
- [2]. Sudharshan Duth, Amrita Vishwa Vidhyapeetham, "Detection of Blood Cancer-Leukemia using K-Means Algorithm".
- [3]. S. Mohapatra, D. Patra, and S. Satpathy, "Unsupervised blood microscopic image segmentation and unsupervised blood microscopic image segmentation and leukemia detection using color-based clustering," International Journal of Computer Information Systems and Industrial Management Applications, vol. 4, pp. 477–485, 2012.
- [4]. Jakkrich Laosai, Kosin Chamnongthai, "Acute leukemia classification by using SVM and K-Means clustering", 2014 International Electrical Engineering Congress (iEECON).
- [5]. Mustafa Ghaderzadeh, Farkhondeh Asadi, Azamossadat Hosseini, Davood Bashash, Hassan Abolghasemi, and Arash Roshanpour, "Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images: A Systematic Review".
- [6]. Nimesh patel, Ashutosh Mishra, "Automated Leukaemia Detection Using Microscopic Images".
- [7]. Cristianini, N., and J. Shawe-Taylor. "An Introduction to support vector machines and other kernel-based learning methods" New York: Cambridge University Press, 2000.

- [8]. R. B. Hegde, K. Prasad, H. Hebbar, B. M. K. Singh, and I. Sandhya, "Automated decision support system for detection of leukemia from peripheral blood smear images," *Journal of Digital Imaging*, vol. 33, pp. 361–374, 2019.
- [9]. E. M. Nagiub Abdelsalam, K. F. Hussain, N. M. Omar, and Q. T. Ali, "Acute myeloid leukemia diagnosis using deep learning," *Clinical Lymphoma Myeloma and Leukemia*, vol. 19, p. S206, 2019.

