WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Comparison of Algorithms for Blood Cancer Detection

Ajeya Bendigeri[1], Akshay Kurabet [2], Ayush Kumar Singh [3], Iranna Inamati [4], Vilas Jarali [5]

[1,2,3,4] Student, Department of Computer Science and Engineering, Angadi Institute of Technology and Management, Belagavi-590009, India

[5] Assistant Professor, Department of Computer Science and Engineering, Angadi Institute of Technology and Management, Belagavi-590009, India

A B S T R A C T

In this paper, Leukemia is also called as blood cancer Detection. Leukemia is a fatal cancer, affecting all ages and it commence in bone marrow. Manual diagnosis from microscopic images is slow and imprecise. There are four main types blood cancer Acute lymphoblastic leukemia (ALL), Acute myeloid leukemia (AML), Chronic lymphocytic leukemia (CLL) and Chronic myeloid leukemia (CML). The classification can be done by using a machine learning classifier called SVM (Support Vector Machine) classifier. This paper analysis the type of blood cancer using microscopic images with help of image processing techniques.

Keywords: Leukemia, Machine Learning, Detection, Support Vector Machine, Image processing

1. Introduction

Leukemia, a very deadly form of blood cancer, is caused by the malignant proliferation of abnormal white blood cells (WBC). Within the complex network of red blood cells, white blood cells and platelets, these components play an important role in oxygen transport, immune defense and general circulation. Understanding this disease depends on accurately identifying the affected white blood cells, especially in cases such as acute lymphoblastic leukemia (ALL), which is characterized by lymphocyte dysregulation. Cancer is a fatal disease that is often caused by the accumulation of genetic disorders and various pathological changes. Cancer cells are abnormal areas that often grow in any part of the human body and are life-threatening. Although the category is various aspects such as complex history, misdiagnosis and treatment, which are the main causes of death.

Divided into acute and chronic forms, leukemia presents unique challenges. Acute leukemia, like ALL, can be fully cured within three months if properly treated. On the other hand, chronic leukemia develops gradually. A worryingly significant proportion of childhood cancers, about 25%, are caused by ALL, a subtype that begins in the lymph nodes and first appears in the bone marrow. The goal is to analyze and review cancer detection using machine learning techniques in cancer leukemia. research emphasizes how machine learning in a guided, unsupervised and deep way can help cancer diagnosis and the healing process. Learning techniques The sample on the right consists of abnormally increased white blood cells. These various features can be used for detection leukemia with a machine learning module.

The irregular development and reduced efficiency of white blood cells in leukemia underscores the critical need for accurate detection and timely intervention. This introductory overview lays the groundwork for an in-depth exploration of the disease and the evolving landscape of diagnostic and therapeutic approaches. Microscopic blood tests are considered the main method for diagnosing leukemia. Analysis of blood samples is the most common way to detect leukemia, but not the only method. In addition, other methods such as molecular cytogenetics, long-range inverse polymerase chain reaction and array-based comparative genomic hybridization also require a lot of work and time to detect leukemia.

Research has proposed many methods to detect leukemia cells in red blood cells in the past. Therefore, this article combines techniques such as Core thresholding, edge detection, and watershed transformation in acute myeloid leukemia (AML). A leukemia cell segmentation system is proposed to evaluate the accuracy of the technique of the analysis of the result.

2. Methodology

The methodology for blood cancer detection will depend on the specific algorithms being compared and the characteristics of the data. Here are a few general steps that might be involved in this process:

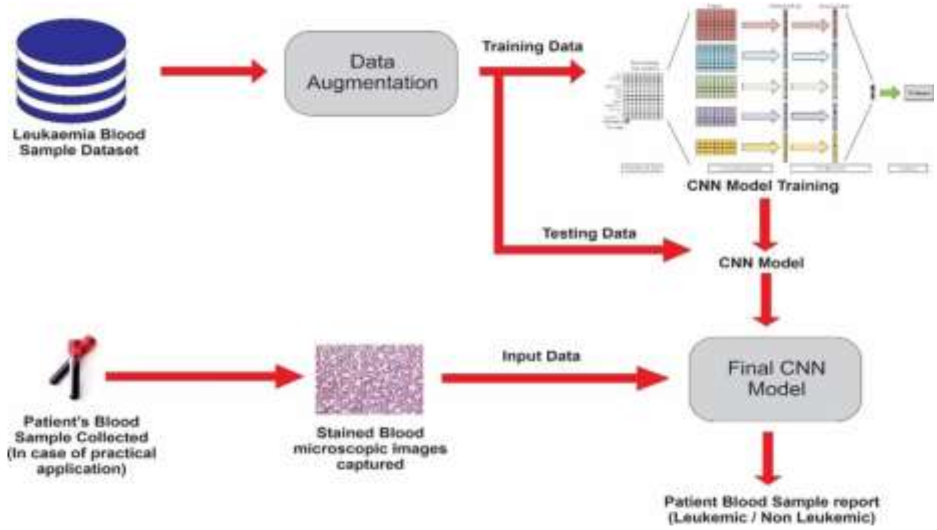

Fig. 1 – Block diagram of Blood Cancer Detection using ML

2.1 DATA COLLECTION

Collecting data for the detection of blood cancers using microscopic imaging involves the collection of a varied and representative set of images that record different types and types of blood cancer cells. This usually involves obtaining microscopic images of blood samples from patients with various forms of blood cancer, such as leukemia. Images must cover multiple conditions to ensure model robustness and accuracy.

The data collection process involves obtaining patient consent and working with medical facilities to obtain appropriate pathology slides or digital images. These images may come from a blood sample, bone marrow biopsies, or other relevant sources. Ensuring a well-annotated dataset with accurate labels for normal and cancer cells is crucial for training and validating machine learning models.

2.2 DATA AUGMENTATION

Data augmentation for blood cancer detection with microscopic images involves artificially expanding the dataset by applying various transformations to the existing images. This technique aims to improve the robustness of the model and improve its generalizability. Common plugins include rotate, translate, zoom, and change brightness and contrast. By introducing different variations of the model, it can better recognize different patterns and features in microscopic images, which ultimately improves its accuracy and reliability in detecting blood cancers such as leukemia.

2.3 CNN MODEL TRAINING

As a deep learning framework, CNN model training is used to detect blood cancer from microscopic images using Convolutional Neural Networks (CNN). The process involves feeding the CNN with a labeled dataset of microscopic images, where each image is associated with a corresponding blood cancer situation. During training, CNN learns to automatically extract hierarchical features from images through convolutional layers. The model is refined by adjusting weights and biases based on differences between predicted and actual labels. This iterative loss-based process continues until the CNN accurately identifies patterns suggestive of blood cancer in new, unseen microscopic images. The goal of training is to optimize the CNN's ability to generalize and make accurate predictions beyond the training dataset.

2.4 PATIENT'S BLOOD SAMPLE

Diagnosing blood cancer with microscopic images involves collecting samples of the patient's blood and examining them under a microscope. The process involves preparing thin blood samples or bone marrow slides, which are then visualized under a microscope. The images taken reveal the red blood cells and platelets. These microscopic images are important inputs for diagnostic algorithms such as image processing and deep learning models. By analyzing the morphology of cells and detecting possible abnormalities, such as irregularities in cell shape, size or distribution, these methods can help identify and classify cases of blood cancer, such as leukemia. The use of patient blood samples combined with microscopic imaging facilitates early.

2.5 STAINED BLOOD MICROSCOPIC IMAGES

Stained microscopic images of blood are essential for detecting blood cancers because they undergo a specific staining process that improves the visibility of cellular structures under the microscope. Common stains, such as Wright-Giemsa or May-Grünwald-Giemsa, highlight different components of blood cells, helping to identify abnormalities. These stained images play a crucial role in the diagnostic process, allowing pathologists and automated algorithms to distinguish between different types of blood cells and detect abnormalities in their morphology.

The enhanced contrast and color variations of the stained images provide valuable information for accurate and detailed analyses, contributing to the effective detection of blood cancers, including leukemia. In conclusion, microscopic images of stained blood improve the visibility of cellular details, facilitating more accurate analysis and diagnosis in the detection of blood cancers by microscopic imaging.

2.6 CLASSIFICATION

The main objective is to categorize the cotton leaves into the categories and healthy and "sick" based on income figures. This includes building a labeled dataset, extracting relevant features from images and training machine learning algorithms such as CNNs, SVMs, Random Forests and KNNs. The trained model is validated, tested and, if possible, put into practice. Continuous monitoring and possible retraining will improve the model and its accuracy over time, promoting early detection and proactive management of crop health.

3. Comparison Table

Algorithm	Study	Dataset	Training	Accuracy	Precision	Recall	Time Complexity	Other Details
K-means Clustering		60 pretested samples	Nearest Neighbor (kNN) and Naive Bayes	92.8%	Not Provided	Not Provided	Not Provided	The algorithm involves several steps, including image preprocessing, morphological operations, image segmentation
Decision Trees, Random Forest, XGBoost	Symptom Analysis using a Machine Learning approach for Early Stage Lung Cancer	Lung Cancer Data from data.world's data catalogue	15-fold cross-validation	94.9%	Not provided	Not provided	Not provided	Identified key factors for lung cancer in different age groups (Youth, Working Class, Elderly)
Convolutional Neural Networks (CNN)	ALL and MM	SN-AM dataset	Not provided	97.2%	Not provided	Not provided	Not provided	Outperforms traditional machine learning methods (SVMs, Decision Trees, Random Forests, Naive Bayes)
Gradient Boosting Decision Tree (LightGBM) Support Vector Machine (SVM)	Acute Lymphoblastic Leukemia (ALL)	10661 images	Not provided	70%	Not provided	Not provided	Not provided	Color-based, Geometrical, Statistical, Haralick Texture, Image Moments, Local Binary Pattern, Presence of Adjacent Cells.
LAB color-based thresholding algorithm	Detection of leukemia cells, specifically Acute Myeloid Leukemia (AML), in blood cell images.	Microscopic images of blood cells obtained from a light microscope with a 40x objective and stained by Wright-Giemsa stain.	Not provided	40%	Not provided	Not provided	Not provided	Thresholding, morphological operations, edge detection, and watershed transform.

Fig. 2 – Comparison Table

Among the more significant indicators for assessing a machine learning algorithm's effectiveness is accuracy. The definition of it is the amount of data that was correctly predicted. The SVM method displayed a precision rate of 70% in this investigation, while the CNN algorithm obtained the greatest accuracy rate of 97.2%. The ratio of actual positive results to the total number of expected positive results is known as precision. The CNN algorithm exhibited the greatest accuracy rate of 94.9%, while the decision tree and algorithm using random forests showed the highest precision rate of 97.2%. Furthermore, the thresholding method based on LAB color had the lowest accuracy, at 40%.A set of K-means

The LAB color-based thresholding methodology had the lowest recall rate (40%) and the CNN algorithm also had the greatest accuracy rate (97.2%).

The LAB color-based thresholding approach had the lowest recall rate (40%) and the CNN algorithm again had the greatest accuracy rate (97.2%). In all three categories, the CNN algorithm fared better overall than the other approaches. It is crucial to remember that the success of these algorithms could be influenced by a number of variables, including the quantity and caliber of the dataset, the caliber of the microscopic images used for examination, and the algorithms' level of complexity.

K-Means In machine learning and data science, clustering is the method of unsupervised learning used to solve clustering problems. We will study the definition of the K-means clustering procedure, how it works and its Python implementation in this topic.

The process in a decision tree starts at the root node in order to project the class for the given dataset. This algorithm follows the branch and advances to the next node through the comparison of the values of the root property with the value of the record (actual dataset) attribute.

SVM is an efficient supervised method that operates best on complex but smaller datasets. Though Support Vector Machines, often known as SVMs, are useful for classification as well as regression uses they function best in the former.

4. Conclusion

In conclusion, this study explores the use of machine learning algorithms, particularly focusing on Support Vector Machine (SVM) and Convolutional Neural Networks (CNN), for the detection of blood cancer, specifically leukemia. The research emphasizes the importance of accurate and timely diagnosis in leukemia cases, considering the challenges of manual analysis of microscopic images.

The comparison table in the study evaluates the performance of SVM, Decision Tree, Artificial Neural Network, and CNN, with CNN emerging as the top performer in accuracy, precision, and recall. While highlighting the potential of machine learning in blood cancer detection, the study acknowledges factors like dataset size and image quality influencing algorithm performance.

In conclusion, the findings suggest that machine learning, especially CNN, shows promise in early and accurate blood cancer detection, contributing to improved patient outcomes. Ongoing research is crucial to enhance algorithm effectiveness and support more effective disease detection and management strategies.

5. Reference

[1] Likhith N.G, Uzair Alam, Dr. Jagruthi H , LEUKEMIA CANCER DETECTION USING DEEP LEARNING, , Information Science And Engineering, BNM Institute Of Technology, Bengaluru, Karnataka, India, | volume:5, Issue:05, May-2023

[2] Akash Holker1 , Naveena Hegde2 , Rakesh M P3 , Yallalinga4 , Sushmitha M5, "Blood Cancer Detection Using Machine Learning", Department of Information Science, Acharya Institute Of Technology, Bengaluru, India Vol. 9, Issue 6, June 2022.

[3] Dharani T, Hariprasath S, “Diagnosis of Leukemia and its types Using Digital Image Processing Techniques” Department of ECE Saranathan College of Engineering Trichy, India.

[4] N. H. M. Daud1 , R. A. A. Raof1,2, M. K. Osman3 , N. H. Harun4, “Segmentation Technique for Nucleus Detection in Blood Images for Chronic Leukemia”, School of Computer & Communication Engineering, University Malaysia Perlis, Pauh Putra Campus, Perlis, Malaysia.

[5] Supriya Mandal, Vani Daivajna, Rajagopalan V, “Machine Learning based System for Automatic Detection of Leukemia Cancer Cell”, Mirafra Software Technologies Pvt Ltd Bangalore, India.

[6] Naresh Khuriwal, Nidhi Mishra, “Breast Cancer Detection From Histopathological Images Using Deep Learning”, Department of Computer Engineering Poornima University, Jaipur, India |25 November 2018.

[7] DEEPIKA KUMAR1 , NIKITA JAIN1 , AAYUSH KHURANA1 , SWETA MITTAL1 , SURESH CHANDRA SATAPATHY2 , ROMAN SENKERIK3, AND D.JUDE HEMANTH.4, “ Automatic Detection of White Blood Cancer from Bone Marrow Microscopic Images Using Convolutional Neural Networks”, Department of Computer Science and Engineering, Bharati Vidyapeeth’s College of Engineering, New Delhi,110063, India 2 School of Computer Engineering, Kalinga Institute of Industrial Technology, Deemed to be University, Odisha, India 3 Faculty of Applied Informatics, Tomas Bata University in Zlin, nám. T. G. Masaryka 5555, Zlin, 76001, Czech Republic 4Department of ECE, Karunya Institute of Technology and Sciences, Coimbatore, 641114, India.

[8] 1Sachin Kumar, 1Sumita Mishra, 1Pallavi Asthana, 2Pragya, “Automated Detection of Acute Leukemia using K-mean Clustering Algorithm”, 1Amity University, Lucknow Campus, 2 Department of Chemistry, MVPG College, Lucknow University.

[9] Saif Ali, Aneeqa Tanveer, Azhar Hussain, Saif Ur Rehman, “Identification of Cancer Disease Using Image Processing Approahes”, University Institute of Information Technology, PMAS Arid Agriculture University, Rawalpindi, Pakistan.

[10] Atharva Bankar, Kewal Padamwar, Aditi Jahagirdar, “Symptom Analysis using a Machine Learning approach for Early Stage Lung Cancer”, Dr. Vishwanath Karad MIT WPU Pune, India.