



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

**Volume 9, Issue 6, June 2021**

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 7.542**



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

# Hematologic Disease Detection Using Deep Learning

PROF. Pallavi Tekade<sup>1</sup>, Tejal Landge<sup>2</sup>, Bhakti Barhate<sup>3</sup>, Shruti Deokate<sup>4</sup>, Harish Kalshetty<sup>5</sup>

Dept. of I.T., JSPM's Rajarshi Shahu College of Engineering, Savitribai Phule Pune University, Pune, India<sup>1</sup>

BE Student, Dept. of I.T., JSPM's Rajarshi Shahu College of Engineering, Savitribai Phule Pune University, Pune, India<sup>2,3,4,5</sup>

**ABSTRACT:** For several diseases, blood analysis is an important indicator; it includes many parameters that are an indication of particular blood diseases. Blood has several functions that are central to survival, including supplying oxygen to cells and tissues. Providing essential nutrients to cells. Blood has many secrets that affect human life. This variation could be detected by the values of parameters inside blood analysis tests. Parameter that contributes to precisely identifying the disease should be identified to predict the disease according to the blood analysis test. Therefore, developing a deep learning model for predicting the Hematologic (Blood) disease and applying modern technological tools for helping physicians to improve the accuracy of disease diagnoses. Prediction of disease like Malaria, leukemia, sickle cell anaemia, and myeloma depend on many key factors related to blood parameters (red blood cell count, WBC count, platelets, and blood smear) In the past few years, lots of work and research papers are published which predicts the result of blood-related diseases by using the algorithms like support vector machine, random forest, etc. By analysing all those research papers, we have selected key factors that affect the results of blood disease and came up with a CNN algorithm that best fits this data and gives maximum accuracy. CNN convolves learned features with input data, and uses 2D convolutional layers, making this architecture well suited to processing 2D data, such as images. CNN's extract the features from the image automatically using numerous hidden layers.

**KEYWORDS:** blood disease, pre-processing, classifier algorithm, feature extraction, Convolutional neural network (CNN), etc

## I. INTRODUCTION

Blood has several secrets that affect human life. This variation could be detected by the values of parameters inside blood analysis tests. As we know, Blood tests are widespread so most doctors may recommend blood tests to predict the health level of the patient's body. Applying modern technological tools for helping physicians to improve the accuracy of disease diagnosing, become one of the hot topics of research, especially deep learning and artificial intelligence algorithms. The main aim of this research is to use deep learning methods for detecting blood diseases according to the blood test values, some techniques are performed for finding the most suitable algorithm that maximizes the prediction accuracy.

By using deep learning and image processing technique our motive is to detect four categories of hematologic disease i.e. Malaria, leukemia, sickle cell anaemia, and myeloma.

### 1.1 OBJECTIVE

The main objective of this research is to use deep learning methods to detect blood diseases (Malaria, leukemia, sickle cell anaemia, and myeloma) according to the values of the blood tests; many methods are carried out to find the most effective algorithm that maximizes the accuracy of prediction. The main objective of this research is to use machine learning methods for detecting blood diseases according to the blood tests values, numerous techniques are performed for finding the best suitable algorithm that maximizes the detection accuracy

## II. LITERATURE SURVEY

(Hanif & Iqbal, 2019) The number of patients with the haematological disease is expanded significantly and it should be diagnosed on time. To diagnose haematological diseases which are mostly low cost and give inaccurate diagnostic results, pathologists use conventional diagnostic tests. Complete blood count CBC is used to finding the existence of diseases. Machine learning is used to predicting haematological diseases.

This classification is used to diagnose diseases that categorize into four classes. Anaemia, Leukaemia, Thalassemia, and healthy patients. Five machine learning algorithms were used with all features and reduced features in this

investigation. The most efficient algorithm is the random forest with the highest accuracy at 98.59 % with the lowest error rate of 0.06%. Findings show that the first indicator for blood disease is haemoglobin.

(Nigatu et al., 2020)cardiovascular diseases (CVD) are one of the major causes of death in the world, which may harm the endothelium cells which may lead to atherosclerosis and cardiac arrhythmias. Blood pressure is a significant parameter and indicator in cardiovascular disease, patients with CVD who have many risk factors such as hypertension, stress, and obesity have been increasing. That’s why, it is important in the field of cardiovascular disease prevention to predict those at risk of cardiovascular diseases in the general population.

(Leonardo et al., 2019) Various diseases cause distinctive changes in the molecular composition of bio fluids such as human blood. Therefore, a sufficiently sensitive and specific blood analysis could be used for disease detection. Mainly, physiological phenotypes (health as well as disease states) are reflected by minor changes in the concentration of several, possibly thousands, of different molecules in blood which cover a wide concentration dynamic range. A method that could thus quantitatively detect diverse molecular groups of blood simultaneously (such as e.g., proteins, metabolites, carbohydrates) would be generally very beneficial.

(Muthu Lakshmi & Kavitha, 2019)cardiovascular disease (CVD) is a prolonged dysfunction caused by a weakening in cardiac physiology. It results in about 31% of death worldwide. Between CVDs, myocardial ischemia leads to constraint in blood supply to heart tissues. There is a need to develop an effective computer-aided detection system to reduce fatality. In this research, an attempt is made to complete mass screening of myocardial ischemic subjects and left ventricle volume estimation from cardiac magnetic resonance images using a deep convolutional neural network with Levenberg-Marquardt learning.

### III. METHODOLOGY / APPROACH

#### 3.1 EXISTING SYSTEM AND DISADVANTAGES

In existing system there were no computerized system to identify the blood disease used in real time scenario. Earlier they were using Support Vector Machine, Decision Tree and Bayes Classifier as machine learning algorithms. They got 60%, 67% and 72% respectively. All these studies were in traditional techniques which used image processing and machine learning (ML) algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Decision tree.

In SVM linear after training and testing the accuracy, classification the data had a problem of hardware capacity to process it. SVMs are most useful for small data sets with less outliers.

#### 3.2 PROPOSED SYSTEM AND ADVANTAGES

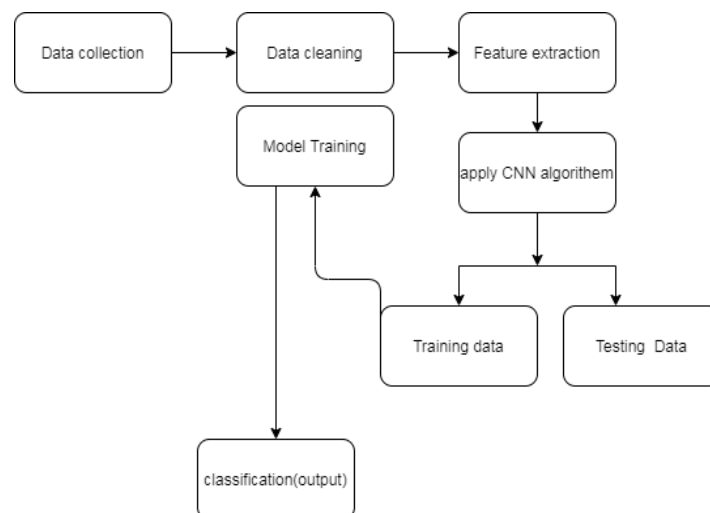


Figure 1: Advance System Architecture

In advanced methods which used Deep Learning, mainly Convolutional Neural Networks (CNNs) which is the most extensively used in the field of blood diseases detection since it is highly accurate, fast, and has the least cost. In addition, it analyses a number of recent works that have been introduced in the field including the size of the dataset, the used methodologies, the obtained results, etc. In conclusion, based on the conducted study, it can be concluded that the proposed system CNN was achieving huge successes in the field whether regarding features extraction or classification task, time, accuracy, and had a lower cost in the detection of diseases.

In hematologic disease detection system, we have to deal large number of data sets for 4 diseases and by comparing efficiency of all the algorithms and by comparing their accuracy we decide to use CNN algorithm. CNN algorithm gave us accuracy about to 95% which is higher than other algorithms

The Blood Diseases Detection, the identification of blood-based diseases often includes identifying and characterizing patient blood samples. We have automated methods for the detection and classification of blood cell subtypes that have important medical applications.

#### Advantages:

- 1) Confirmation of infected individuals is essential to manage and contain the virus successfully. It would be difficult to find the actual rates of cases without reliable testing. Therefore, it is important to find what these available tests can and can't do to use them correctly.
- 2) Secure and efficient system.
- 3) The leading objective of our paper is to enhance the value of blood disease detection

### 3.3 ALGORITHM

#### CNN:

A Convolutional Neural Network (CNN) is a class of neural networks that concentrates in processing data that has a grid-like topology, such as an image. The pre-processing required in a CNN is considerably lower as compared to other classification algorithms. While in primeval methods filters are hand-engineered, with enough training, CNN can learn these characteristics. The architecture of a CNN is equivalent to that of the connectivity pattern of Neurons in the Human Brain and was encouraged by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a constrained region of the visual field known as the Receptive Field. A assembly of such fields overlay to cover the entire visual area.

#### Types of layers:

**Input Layer:** In Input Layer, this layer holds the raw input of the image with width, height, and depth

**Convolution Layer:** This layer calculates the output volume by calculating dot product among all filters and image patches.

**Activation Function Layer:** This layer will apply an element-wise activation function to the output of the convolution layer. Some common activation functions are used.

**Pooling Layer:** This layer is periodically inserted in the comments and its main function is to decrease the size of volume which makes the computation fast decreases memory and also prevents overfitting.

**Fully-Connected Layer:** This layer is a neural network layer that takes input from the preceding layer and computes the class scores and outputs the 1-D array of size equal to the number of classes.

#### Technique:

##### Image pre-processing

Image pre-processing is the name for operations on images at the lowest level of abstraction whose aim is an improvement of the image data that suppress undesired distortions or enhances some image features important for further processing. It does not increase image information content. Local illumination can be enhanced using gradient filters, local histogram equalization, and rank filters. Blur and focus enhancements. Many well-known filtering methods for sharpening and blurring may be employed at the pre-processing stage

### IV. CONCLUSION AND FUTURE WORK

In this paper, we studied essential techniques for modelling the human process in many disciplines, especially in the medical field, because of the high availability of data. One of the essential disease detectors is blood analysis; as it



contains many parameters with different values that indicate definite proof for the existence of the disease. The deep learning algorithm accurateness depends mostly on the quality of the dataset, that's why, a high-quality dataset is collected and verified by experts. This dataset is used for training the classifiers for obtaining high accuracy. Adopting an image processing approach in blood cell image diagnosis could lead to a fundamental change in differential diagnosis and reduce the late detection of disease which may cause death. We can expand the scope of this system to another type of disease. (Covid-19 ,cancer , etc.)

## REFERENCES

- [1] Lewontin, Richard C. It ain't necessarily so: The dream of the human genome and other illusions. New York Review of Books, 2001.
- [2] Feldman, Eric A., Eric Feldman, and Ronald Bayer, eds. Blood feuds: AIDS, blood, and the politics of medical disaster. Oxford University Press, USA, 1999.
- [3] Fekkes, Minne, et al. "Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms." *Pediatrics* 117.5 ;2006: 1568-1574.
- [4] ESHRE, The Rotterdam, and ASRM-Sponsored PCOS Consensus Workshop Group. "Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome." *Fertility and sterility* 81.1 ;2004: 19-25.
- [5] Schalm, Oscar William, Nemi Chand Jain, and Edward James Carroll. *Veterinary hematology*. No. 3rd edition. Lea &Febiger., 1975.
- [6] Allison, James E., et al. "A comparison of fecal occult-blood tests for colorectal-cancer screening." *New England Journal of Medicine* 334.3 ;1996: 155-160.
- [7] Park, Sang Hyuk, et al. "Establishment of age-and gender-specific reference ranges for 36 routine and 57 cell population data items in a new automated blood cell analyzer, Sysmex XN-2000." *Annals of laboratory medicine* 36.3 ;2016: 244-249.
- [8] Cabitza, Federico, Raffaele Rasoini, and Gian Franco Gensini. "Unintended consequences of machine learning in medicine." *Jama* 318.6 ;2017: 517-518.
- [9] Darcy, Alison M., Alan K. Louie, and Laura Weiss Roberts. "Machine learning and the profession of medicine." *Jama* 315.6 ;2016: 551-552.
- [10] Jiang, Min, et al. "A study of machine-learning-based approaches to extract clinical entities and their assertions from discharge summaries." *Journal of the American Medical Informatics Association* 18.5 ;2011: 601-606.
- [11] Lison, Pierre. "An introduction to machine learning." ;2015.
- [12] Michalski, Ryszard S., and Yves Kodratoff. "Research in machine learning: Recent progress, classification of methods, and future directions." *Machine learning*. Morgan Kaufmann, 1990. 3-30.



**INNO**  **SPACE**  
SJIF Scientific Journal Impact Factor  
**Impact Factor: 7.542**



**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
**INDIA**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

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



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details