IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

Prediction of Disease from Blood Microscopic Analysis Classification using Deep Learning Algorithm

Gokulakrishnan.V[1], Manikandan. S[2], Santhosh. R[3], Saravanan. S[4], Vinoth. R[5]
Assistant Professor, Department of Computer Science and Engineering[1]
Students Department of Computer Science and Engineering[2,3,4,5]
Dhanalakshmi Srinivasan Engineering College (Autonomous), Perambalur, India

Abstract: *Traditional blood disease detection based on visual inspection of blood smears using a microscope is time consuming, error-prone, and limited by the physical acuity of the hematologist. To facilitate clinical decision-making, an automated optical image processing system is required. Leukemia is a kind of cancer that is distinguished by the aberrant development of immature, abnormally shaped white blood cells (WBC) known as "blasts." Leukemia is a cancer of the white blood cells (WBCs) that affects the bone marrow and/or blood. A timely, safe, and accurate diagnosis of leukemia at an early stage is critical to treating and preserving patients' lives. The white blood cells in a blood smear are often examined under a microscope to determine the diagnosis.Many machine learning have been created to diagnose various illnesses, such as leukemia, and to deliver a high mis-classification error rate. As a result, we may use a deep learning system to identify microscope pictures for White Blood Cell Count study. The WBC differential counting system was divided into two modules: detection and classification. The detection module first processed the raw bone smear pictures, detecting all WBCs including red blood cells, platelet counts, staining impurities, and so on. The discovered cells were then sent into the categorization module.The categorization module was divided into two phases. In the first step, we separated a large number of cells, such as crushed cells, degenerated cells, and so on, that are not employed in the diagnosis for leukemia. The countable WBCs were then presented for multi-class separation using the Convolutional neural network technique in the second step.*

Keywords: Blood disorders, White blood cells, Classification, Machine learning, Deep learning, Leukemia

I. INTRODUCTION

Cancer is the uncontrolled growth and development of aberrant cells. Cells develop, mature, perform their designated role, and finally die under normal conditions. In order to replace damaged cells and maintain proper cellular function, the body continually regenerates new cells. Cells can grow and proliferate in an unorganized and uncontrollable manner. Cells can fail to grow properly, causing them to operate abnormally. Cells may die abnormally. When cells become malignant, one or more of these events may occur. Leukemia is a bone marrow malignancy that affects blood-forming cells. These abnormal, immature cells increase blood and organs. They are unable to perform regular blood cell activities. Hematologists seek for minute structural and histological alterations inside the cells of the a blood film to describe the leukemia and its kind. There are two forms of leukemia: acute and chronic. Acute leukemia is a quickly progressing disease that causes an increase in the number of faulty lymphoid cells in the blood of the patient. A normal marrow test will indicate a large quantity of leukemic cells and a small number of healthy white blood cells. People with acute leukemia may experience fatigue, bruise susceptibility, and infections. Chronic leukemia, on either hand, often advances slowly. In the early stages, leukemic cells remain functioning and perform their original duties, and they become severely weakened in the later phases. Patients report nausea and fatigue, and the initial diagnostic was founded on irregular blood test findings. If left untreated, leukemia cells will outnumber healthy blood cells, limiting systemic function. Acute and chronic leukemia are further subdivided based on the afflicted cell types. This second categorization identifies whether the leukemia is mononuclear or lymphoid in nature. Myeloid leukemia cells clump together to form "sarcomas," extramedullary myeloid tumors, granulocytic sarcomas, or chloromas, while lymphoid leukemia cells clump together to cause lymph node enlargement.. Figure 1 depicts the four type of leukemia.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

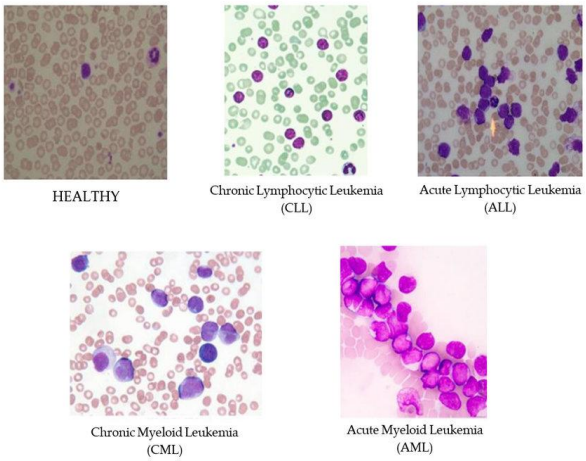

Fig 1: Leukemia Types

II. RELATED WORKS

A. M. Abdeldaim et al,…[1] focused on lymphocyte cells that are impacted by lymphoblastic malignancy. The primary goal of this study is to detect lymphocytes by segmenting microscopic pictures, and then to diagnosis (classify) every segmented cell as normal or abnormal. The published ALL-IDB2 set of data was used in this chapter to evaluate the efficiency of the suggested detection method. The dataset comprises 260 cell pictures, of which 50% are normal and 50% are afflicted by ALL. The suggested method begins with segmentation, which includes RGB to CMYK colour model conversion, contrast stretching through histogram equalisation, thresholding, noise reduction, and background removal. Second, some characteristics of each cell were retrieved. Color, texture, and form characteristics are the three types.Third, to enhance classification performance, three data normalisation approaches (z-score, min-max, and grey-scaling) were used to each extracted feature. Finally, several classifiers (K-Nearest Neighbour, Naive Bayes, Support Vector Machines, and Decision Trees) were employed to evaluate the proposed system's efficiency.

M. Nassar et al,…[2] presented a A workflow combining image flow cytometry and deep learning to analyse, categorise, and create WBC differential counts for use by clinicians in illness diagnosis and monitoring. We tested six machine-learning classifier and found that the best one, GB coupled with random under sampling, can categories WBCs into the three major kinds with a 97 percent average F1-score. Furthermore, we demonstrated that Gradient Boosting paired with random under sampling can identify lymphocytes with just an average cross validation F1-score of 78 percent, revealing for the first time that lymphocytes may be discriminated morphologically. The methodology provided here improves on the current state-of-the-art flow cytometry approach by categorising cells without the use of fluorescent markers.This may help to eliminate physical disruptions to the cells and speed up and strengthen the sample preparation operation.

R. B. Hegde et al.,…[3] demonstrated WBC categorization utilising classic image processing and deep learning methods. Both approaches performed equally well, with 99 percent overall accuracy and sensitivity. The accuracy of classification in classical image processing approaches is dependent on the precision of feature selection and feature extraction. This is eliminated by using the deep learning approach. It learns the feature on its own, regardless of image alterations, but it requires a huge amount of labelled data and robust infrastructure. Because of the data availability and the reduced size of cell pictures required by the network, CNN may be employed for WBC categorization. The current study's future scope would include developing a strong segmentation approach to accommodate fluctuations in peripheral blood smear pictures.We produced comparable results using a standard image processing technique that is simpler and requires fewer CPU resources.

Pau Rodríguez, et.al,…[4] proposed a modular feedforward focus technique that is quick, does not affect the fundamental CNN model and is trained via SGD. As a result, the suggested model may be used to supplement any pretrained architecture, such as residual neural nets (ResNets) or VGG, and it is built to operate in parallel with the primary architecture, with no computational time overhead. Attention modules are positioned at various depths of the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

CNN, providing testing technique based on the local data, such as finely grained features, extracted from of the main CNN feature activations at various levels of abstraction. The attention gates then adjust the appropriate network output classification performance using these local predictions. For example, after paying attention to the blue sirens, the suggested model corrects the forecast "sedan" to "police-car."This paradigm, however, adds extra processing processes, resulting in higher overhead. Alternatively, our solution may be run in parallel with the upgraded design, reducing the computing cost.

Jianpeng Zhang, et.al,…[5] propose the ARL-CNN model for skin cancer detection in dermoscopy pictures, which combines residual learning with an unique attention learning technique to increase DCNN discriminative representation capacity. The innovative attention learning process is intended to build attention maps for lower layers by using feature maps learnt by higher levels. On the ISIC-skin 2017 dataset, we tested our model. Our findings demonstrate that the proposed ARL-CNN algorithm can flexibly concentrate on the discriminative regions of skin lesions, achieving state-of-the-art skin lesion classification performance. Each ARL block improves its discriminative representation capacity by combining residual learning and innovative attention learning processes. Instead of employing additional learnable layers, the suggested attention learning technique seeks to capitalise on DCNNs' inherent self-attention capacity, i.e. leveraging feature mappings learnt by a high layer to build a closely monitoring for a low layer.

Lenin Shen, et.al,…[6] proposed a DCR module includes cross residue shortcuts to improve information transmission within deep CNN layers, and it was used to classify HEp-2 cell staining patterns. To train neural learning with minimal data, a class-balanced deep learning strategy was effectively deployed. The display of feature maps reveals that a DCR module with more CC may collect finer picture information and generate more discriminative features. Without utilising any external training data, our technique produced state-of-the-art results on two data sets, namely 80.8 percent ACA in the ICPR 2012 data and 85.1 percent MCA in the I3A task 1 dataset. Computer-aided systems attempt to evaluate HEp-2 pictures automatically and assist clinicians in diagnosing as much as feasible. The majority of methods in the literature are divided into three stages. Image pre-processing, intensity normalisation, and noise reduction are common in the initial step. The second stage is primarily concerned with extracting features such as Scale Invariant Transform (SIFT), Local Binary Pattern (LBP), and Histogram of Gradient (HOG), among others. Although local descriptors with their extensions are the most often used ways, some methods employ Bag of Words (Bow) or the Fisher Vector architecture to express these descriptors. Using the features from stage two and their related ground-truth labels, classifier such as Nearest Neighbour (k-NN), SVM Classifier (SVM), and naive Bayes are trained in the third step. Four similar contests were organised in connection with Global Forum on Pattern Recognition (ICPR) as well as the World Summit on Image Processing (ICIP) to assess state-of-the-art systems in order to enhance the performance of cell categorization systems.

Jaroonrut Prinyakupt, et.al,…[7] proposed WBC segmentation method applied to two datasets, with outcomes compared to the gold standard manually segmented by a haematologist. Both yield greater than 90% accuracy. This procedure is quick, dependable, and effective. As a result, white blood cell morphologies may be retrieved and compared in linear and naive Bayes classifiers. The linear classifier outperforms the naive Bayes classifier somewhat. Furthermore, the five categories of white blood cells may be categorised with excellent accuracy. It really should be noted that now the photos in dataset 2, which was downloaded through the CellAvison programme, are not from a conventional camera. They have clearly been trimmed to show just the white blood cell. The resolution, on the other hand, has been approximated based on the actual RBC size. Nonetheless, testing on two picture datasets of varying resolutions showed that the proposed segmentation procedure may be calibrated to handle varied images or formats as much as the resolution is known.

Cesar Mauricio, et.al,…[8] The relevance of computer vision in performing routine processes that are carried out physically through observation was proved by delivering an objective assessment with high precision in short periods. The blood samples were obtained using the methodology specified in the hematologic analysis manuals, in which staining plays an important role since it emphasises the white blood cells of other objects, enabling for the valid documentation of the white blood cells through photographs. Although there were issues with just some samples during picture capture, they must take into consideration the space of the region you are watching with the microscope's objective, given that identification of white blood cells is done in fields where the bacteria are not grouped. The use of Gaussian radial base function networks in tandem with morphometric descriptors and the distance among objects

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

produced better results in the edge detection of nuclei than other techniques used, owing to highlights and extracts that removed completely the nuclei whilst also retaining its shape on pictures directly in colour space.

Shahd T. Mohamed, et.al,…[9] compares PSO and GSA were used to teach the Optimized Multilayer Feedforward Network (OFNN) to categorise Agranulocytes (lymphocytes and monocytes) and Granulocytes (neutrophils, eosinophils, and basophils). To offer unique information on white blood cells, we employed the statistical characteristics mean, variance, skewness, kurtosis, area, and perimeter. The OFNNs tested varied iteration numbers as well as particle and agent counts when using PSO and GSA. Experiment findings suggest that increasing the number of Atoms and Agents in swarm algorithms improves classification accuracy. OFNN achieved classification accuracy of up to 100% for Agranulocytes and 96.67% for Granulocytes.To achieve the lowest training error, it is critical to solve the slow convergence of traditional neural networks and search for weight training in Conventional Neural Network (FNN). This research provides an Optimized Multilayer Feedforward Network (OFNN) for the categorization of tiny white blood cell (WBC) pictures. Particle swarm optimization (PSO) and the Gravitation Search Algorithm (GSA) have been used to training the feed forward neural network and find the weights of the FFN in order to attain a low error rate and a high classification rate. The OFNN is used to properly categorise white blood cells into Agranulocytes, which include lymphocytes and monocytes, and Granulocytes, which include neutrophils, eosinophils, and basophils. The segmented cells' form properties are used to train the OFNNs.

Hamed Talebi, et.al,…[10] developed a viable approach for recognising WBCs. Over 260 distinct colour photos, our method's identification power and accuracy were examined, indicating its superiority over other approaches. White blood cells (WBCs) are a key component in blood cancer detection, but their basic features in microscopic pictures, such as non-uniform colors/illuminances, varied forms, sizes, and textures, can make recognition and classification jobs challenging. Furthermore, overlapping WBCs in bone marrow imaging and proximity to red blood cells are found as causes of categorization mistakes. In this research, we attempted to segment distinct regions of medical pictures using the Nave Bayes clustering approach, followed by the TSLDA classification, which is provided by features obtained from the covariance descriptor, yielding an accuracy of 98.02 percent.

III. FEATURES EXTRACTION

The feature extraction procedure entails extracting image parameters to characterise the microscopic characteristics of leukaemia and making a judgment based on these parameters. Medical professionals rely on the characteristics of leukaemia. For feature selection, the method of diagnostic used is critical. Asymmetry and pigmented network, for example, are characteristics in contour rule as well as pattern analysis, respectively. The visual evaluation of leukaemia diagnosis is tough since the contents of information in pictures is complex and requires competent physicians for interpretation. Prior to categorization, a crucial procedure called feature selection is carried out. Its goal is to lower the computational burden of classification by reducing the number of extracted feature descriptors.However, this reduction is not insignificant because it eliminates duplication, which may have a detrimental impact on discriminatory power.

3.1 Color Based Feature Extraction

Color is the most obvious and crucial aspect that humans detect when seeing an image. Because colour information is more sensitive to human eye than grey levels, colour is the first choice employed for feature extraction. One frequent approach for representing colour contents is the colour histogram. The methods proceed in a similar manner: selecting a colour space, representing colour attributes, and matching algorithms. The colour histogram is indeed the most often used approach for obtaining an image's colour characteristic. It depicts the image from the a different angle. It displays the frequency analysis of picture colour bins. It counts and saves comparable pixels.Color histograms are classified into two types: global colour histograms and local color histograms. The color histogram is offered as a global colour descriptor that examines each statistical color frequencies in an image. It is used to address problems involving translation, rotation, and angle of view changes. Local color histograms concentrate on specific areas of a picture. Local color histograms take into account the geographical extent of pixels, while global color histograms do not. Color histograms are useful in picture indexing and retrieval because they are simple to generate and insensitive to slight alterations in the image. Apart from these benefits, it has two big disadvantages. First, no consideration is given to total spatial information.Second, the distribution is not robust as well as unique since two separate photos with identical

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

colour distributions produce similar histograms, whereas the same view photographs with varying light exposure produce different histograms. Color histograms are described as a set of bins, with each bin denoting the likelihood of pixels in a picture being of a specific colour. A colour histogram is described as a vector for a given image:

$$H = \{H[0], H[1], H[2],\ldots, H[i],\ldots, H[N]\}$$

If I is a colour in the colour histogram which corresponds to something like a subcube in the RGB colour space, H[i] defines the number of dots in colori in that picture, and N denotes the number of bins in the histogram, i.e., the number of colours in the colour model used.

Typically, each pixel in the image is allocated to a bin of the image's colour histogram, thus the value of each bin in an image's colour histogram is the number of pixels with the same corresponding colour. Color statistics should be adjusted before comparing photos of various sizes. H' denotes the normalised colour histogram.

$$H' = \{H'[0], H'[1], H'[2],\ldots, H'[i],\ldots, H'[N]\}$$

where H I = H[i]/P and P is the maximum number of pixel in the image (the remaining variables are defined as before). A perfect colour space quantization assumes that unique colours should not be distributed to the same sub-cube and that related colours must be distributed to the same sub-cube. Using fewer colours reduces the likelihood of similar colours being assigned to separate bins while increasing the likelihood of diverse colours being allocated to the very same bins, and the incredibly powerful of the pictures decreases to a greater extent. Colorhistograms with a high number of bins, on the other hand, will include more content of photos, reducing the probability of distinct colours being allocated to the same bins.. However, they improve the probability of identical colours being allocated to different bins, the amount of metadata stored, and the time required to calculate the distance between colour histograms. As a result, deciding how many bins to employ in colour histograms involves a trade-off.

3.2 Shape Based Feature Extraction

The shape of a picture is an important basic property that is utilised to define its content. However, due of noise, occlusion, and arbitrary distortion, the form is frequently damaged, and the object detection problem has become more complicated. Shape characteristics, that are either based on shape boundary data or boundary plus internal content, are the foundation of shape representation. For object identification, several sorts of form characteristics are constructed and assessed based on how correctly those shape attributes allow one to obtain comparable objects from a database. Shape descriptors should be able to successfully locate comparable forms from the database, whether they are affinity altered shapes such as twisted, translated, flipped, scaled, and so on, for acceptable retrieval accuracy.The form descriptors should also be able to discover faulty shapes, noise-affected shapes, and shapes acceptable by humans for shape comparison and retrieval. This is referred to as the resilience requirement. A shape descriptor must be able to do image retrieval for all sorts of forms rather than just a few, and that it should be program agnostic. Low computing complexity is a key feature of the shape descriptor. By integrating fewer picture attributes in the computing process, computation problem is reduced and the shape descriptor becomes more robust. In this case, low computing complexity implies clarity and stability.For form retrieval applications, several shape representation and description approaches have been developed. Shape description and description approaches are split into two groups based upon whether form characteristics are derived from the contour alone or from the entire shape region.

1. Contour based methods.
2. Region based methods.

Each strategy is further subdivided into two approaches: structural and global. The structural and global techniques differ depending on whether the form is depicted as a whole rather than in parts.

3.3 Texture Based Features

Texture is a significant picture characteristic and a potent geographical descriptor that aids in retrieval. Texture cannot discover comparable photos on its alone, but it may be used to distinguish texture images against non-textured ones but then paired with another visual property, such as colour, to improve retrieval.

Textural features are : Statistical measures

- Entropy
- Homogeneity

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

- Contrast

Used to extract the texture features such as

Contrast : Size contrasting or Local intensity fluctuations on the diagonal will contribute to P(i,j), i=j. The following formula may be used to compute the contrast between a neighbourhoodpixels:

Contrast = $\sum_{n=0}^{G-1} n^2\{\sum_{i=1}^{G}\sum_{j=1}^{G} P(i,j)\}, |i-j| = n$

Correlation: Connection (correlation) is indeed the size of the grey level inter-pixel linearly depending on each pixel's relative position. Correlation may be determined using the formula:

Correlation = $\sum_{i=0}^{G-1}\sum_{j=0}^{G-1}\frac{\{i*j\}XP(i,j)-(\mu_x-\mu_y)}{\sigma_x X\sigma_y}$

where $\mu_x, \mu_y, \sigma_x, \sigma_y$ is mean and P(i,j) is standard deviation values

Homogeneity: Also known as HOM in the short term. It sends the value calculated by the severity of dispersion of the GLCM elements towards the GLCM diagonal. The diagonal GLCM has a value of 1 and a range of [0,1]. Homogeneity weight values are the inverse of contrast weighting factor, with weight decreasing exponentially away from the diagonal.

Homogeneity = $\sum_{i,j=0}^{N-1} P(i,j)/R$

Energy: Because energy is utilised to do labour, orderliness follows. It makes advantage of the texture of a picture to determine ordering. In GLCM, it returns the total of square components. It is not the same as entropy. Whenever the window is properly organised, the efficiency value is high. As Energy, the square root of the ASM (Angular Second Moment) texturing component is utilised. It has a range of [0 1]. Its value is set since it is a constant picture. The energy equation is determined as follows:

Energy = $\sum_{i,j=0}^{N-1} P(i,j)$^2

Normalized each providing an excellent by dividing each providing an excellent but by average of same constituent of the same patient's healthy blood cell, and it performed better than the other normalisation procedure. Texture characteristics are retrieved using generalised co-occurrence matrices, which are cell image extensions of the co-occurrence matrix.CNN hierarchical structure is effectively demonstrated by efficient analysis of observable representations. The fundamental issue in such visual stimuli is the emergence of inter-class as well as the objects of form variations modelling. The excitable data are represented as two-dimensional graphs with hundreds of wavelength bands. Differentiating relative classes with the human eye is challenging since each class curve has its unique visual form that differs from other class (e.g., gravel and self-blocking bricks). CNNs can solve competitive challenges and outperform humans by employing spectral signatures; this skill drives researchers to investigate the idea of using CNNs for HSI categorization. The efficient examination of observable representations successfully demonstrates CNN hierarchical structure. The formation the inter-class as well as the subjects of shape variation modelling is the key challenge in such visual stimuli. The excitable data is shown as this double graphs with hundreds of wavebands. Differentiating related classes with the human eye is difficult since each class curve has a distinct visual shape that differs from the others (e.g., gravel and self-blocking bricks). By exploiting spectral characteristics, CNNs could solve competitive tasks and outperform humans; this ability motivates academics to study the notion of using Convolutions for HSI classification.

IV. CONVOLUTIONAL NEURAL NETWORK ALGORITHM

CNN hierarchical structure is effectively demonstrated by efficient analysis of observable representations. The fundamental issue in such visual tasks is the appearance of inter-class and also the objects of form variations modelling. The hyper-spectral data are represented as two-dimensional lines with hundreds of spectral bands. Differentiating relative classes with the human eye is challenging since each class curve has its unique visual form that differs from these other classes (e.g., gravel and self-blocking bricks). CNNs can solve competitive challenges and outperform humans by employing spectral signatures; this skill drives researchers to investigate the idea of using CNNs for HSI categorization. The CNN differs depending on how the network is trained and how the convolution layer and maxpooling are implemented.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

Construction of the CNN Model

function CNN_MODEL (θ, [n1–5])
layerType = [convolution, max-pooling, fully-connected, fully-connected];
layerActivation = [tanh(2), max(),softmax()]
model = new Model();
for$i$=1 to 4 do
layer = new Layer();
layer.type = layerType[i];
layer.inputSize = ni
layer.neurons = new Neuron [ni+1];
layer.params = θi;
model.addLayer(layer);
end for
return model;
end function

Training of the CNN Model

First, the variables ITERmax, ERRmin, BATCHEStraining, and SIZEbatch are set, which represent the learning rate, maximum number of iterations, minimal error, training batches, and batch size, respectively. According to $n1$ and $n5$; compute $n2$, $n3$, $n4$, $k1$, $k2$.
Randomly generate the weights θ of the CNN;
cnnModel = InitCNNModel(θ, [n1–5]);
iter = 0; err = +inf;
while err >ERRmin and iter<ITERmax do
err = 0;
forbatch = 1 to $BATCHES_{training}$ do
[$\nabla\theta J(\theta)$, $J(\theta)$] = cnnModel_train (TrainingDataset,
TrainingLabelsets), and θ should be updated;
err = err + mean($J(\theta)$);
end for err = err/$BATCHES_{training}$;
iter++;
end while
Save parameters θ of the CNN
These networks differ depending on the spectral data channel size as well as the output classes produced based on the feedback obtained from cell data. The suggested technique overcomes irregular boundary separation in blood picture classification by extracting form and texture data. The proposed work involves a phase of training and testing. Train the numerous blood cell pictures using simple image processing procedures on the training side. In the testing stage, input the photo and execute preprocessing procedures to remove noise from the image, features extracted, and sort the data using a CNN method with a higher accuracy rate.

V. EXPERIMETNAL RESULTS

The system's performance may be evaluated utilizing KAGGLE Blood cell datasets. For analysing system performance, many performance indicators such as accuracy, sensitivity, specificity, error rate, and precision may be generated.
True positive (TP): the number of true positives resulting from a perfect positive prediction.
False positive (FP): the amount of false positives resulting from poor positive prediction.
True negative (TN): the number of true negatives resulting in a complete negative forecast.
False negative (FN): the number of real negatives resulting from poor negative prediction.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

Accuracy

Accuracy (ACC) is calculated by dividing the total amount of perfect predictions by the total quantity of test data. It can alternatively be written as 1 - ERR. The best possible accuracy is 1.0, while the poorest possible accuracy is 0.0.

$$ACC = \frac{TP+TN}{TP+TN+FN+FP} \text{ x } 100$$

ALGORITHM	ACCURACY
Naives Bayes algorithm	50%
Support Vector Machine	65%
Convolutional neural network algorithm	80%

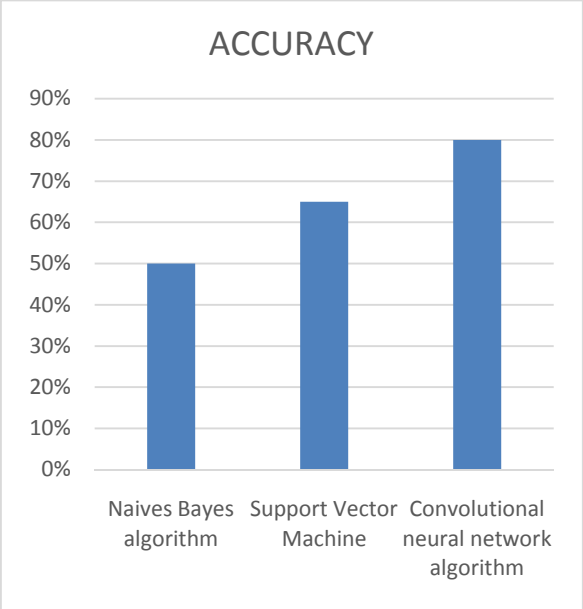

Figure 3: Performance chart

According to the graph above, the BPNN algorithm outperforms the existing computational modelling algorithms in terms of accuracy.

VI. CONCLUSION

Using microscopic pictures of the cells, this study employs Convolutional Neural Network techniques to assist hematologists in classifying White Blood Cells into subgroups. This categorization assists in cell identification and determining the sort of illness afflicting a patient. When contrasted to machine learning techniques, the findings of this investigation assist in more reliably identifying photographs. The test set has a high level of accuracy (more than 80%). As a consequence, so when model is trained with powerful computational skills, a perfect model can be developed and used in diagnostic testing and apps dealing with the number of white blood cells and subtypes of white blood cells.

REFERENCES

[1]. Abdeldaim, Ahmed M., et al. "Computer-aided acute lymphoblastic leukemia diagnosis system based on image analysis." Advances in Soft Computing and Machine Learning in Image Processing. Springer, Cham, 2018. 131-147.

[2]. Nassar, Mariam, et al. "Label-free identification of white blood cells using machine learning." Cytometry Part A 95.8 (2019): 836-842.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 1, May 2023

[3]. Hegde, Roopa B., et al. "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images." Biocybernetics and Biomedical Engineering 39.2 (2019): 382-392.

[4]. Rodriguez, Pau, et al. "Pay attention to the activations: a modular attention mechanism for fine-grained image recognition." IEEE Transactions on Multimedia 22.2 (2019): 502-514.

[5]. Zhang, Jianpeng, et al. "Attention residual learning for skin lesion classification." IEEE transactions on medical imaging 38.9 (2019): 2092-2103.

[6]. Shen, Linlin, Xi Jia, and Yuexiang Li. "Deep cross residual network for HEp-2 cell staining pattern classification." Pattern Recognition 82 (2018): 68-78.

[7]. Prinyakupt, Jaroonrut, and CharnchaiPluempitiwiriyawej. "Segmentation of white blood cells and comparison of cell morphology by linear and naïve Bayes classifiers." Biomedical engineering online 14.1 (2015): 1-19.

[8]. Rodríguez Barrero, Cesar Mauricio, Lyle Alberto Romero Gabalan, and Edgar Eduardo Roa Guerrero. "A novel approach for objective assessment of white blood cells using computational vision algorithms." Advances in Hematologic 2018 (2018).

[9]. Mohamed, Shahd T., et al. "Optimized feed forward neural network for microscopic white blood cell images classification." The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019) 4. Springer International Publishing, 2020.

[10]. Talebi, Hamed, et al. "High accuracy classification of white blood cells using TSLDA classifier and covariance features." arXiv preprint arXiv:1906.05131 (2019).

[11]. Gupta, Deepak, et al. "Bat-inspired algorithm for feature selection and white blood cell classification." Nature-Inspired Computation and Swarm Intelligence. Academic Press, 2020. 179-197.

[12]. Wong, Annika, et al. "Analysis of Vision-based Abnormal Red Blood Cell Classification." arXiv preprint arXiv: 2106.00389 (2021).

[13]. Liang, Gaobo, et al. "Combining convolutional neural network with recursive neural network for blood cell image classification." IEEE access 6 (2018): 36188-36197.

[14]. Ma, Li, et al. "Combining DC-GAN with ResNet for blood cell image classification." Medical & Biological Engineering & Computing 58.6 (2020): 1251-1264.

[15]. Sahlol, Ahmed T., Philip Kollmannsberger, and Ahmed A. Ewees. "Efficient classification of white blood cell leukemia with improved swarm optimization of deep features." Scientific Reports 10.1 (2020): 1-11.

[16]. Deshpande, Nilkanth Mukund, Shilpa Shailesh Gite, and RajanikanthAluvalu. "Microscopic Analysis of Blood Cells for Disease Detection: A Review." Tracking and Preventing Diseases with Artificial Intelligence (2022): 125-151.

[17]. Khalil, Ahmed J., and Samy S. Abu-Naser. "Diagnosis of Blood Cells Using Deep Learning." International Journal of Academic Engineering Research (IJAER) 6.2 (2022).

[18]. Sharma, Sarang, et al. "Deep learning model for the automatic classification of white blood cells." Computational Intelligence and Neuroscience 2022 (2022).

[19]. Abas, Shakir Mahmood, Adnan Mohsin Abdulazeez, and Diyar QaderZeebaree. "A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia." Indones. J. Electr. Eng. Comput. Sci 25.1 (2022): 200-213.

[20]. Wang, Ching-Wei, et al. "Deep learning for bone marrow cell detection and classification on whole-slide images." Medical Image Analysis 75 (2022): 102270.