

DETECTION OF WHITE BLOOD SAMPLE CELLS USING CNN

K.SONA¹, C.SRIRAGAVI², A.VIJAYA³ B.V.VARSHINI ⁴

¹K.SONA Student

²C.SRIRAGAVI Student

³ A.VIJAYA Student

⁴ B.V.VARSHINI Assistant professor Dept. of computer science and Engineering, RMD Engineering college, Tamil Nadu, India

Abstract - This paper presents a new classification system for white blood cells to recognize 4 types of white blood cells. For the segmentation of white blood cells from images, We can segment from an image a white blood cell. Convolution neural network has already demonstrated power in many fields of application and is accepted as a better approach by more and more people as a better approach than traditional models of machine learning. The implementation of Convolution Neural Networks (CNN), in particular, brings enormous benefits to the medical field, where the processing and analysis of a huge number of images is required. This paper implements a Convolution Neural Network for the classification of the four blood subtypes A CNN-based framework for the automatic classification of blood cells. Experiments are carried out on a dataset of 15k images of blood cells with their subtypes, and the proposed CNN approach generated improved results and reduced the rate of error compared to other models. A CNN model based on Deep Learning, where deep learning enhances the extraction capability and smooth scaling of features in case of increased parameters and 81 percent accuracy was achieved in the classification of WBCs.

Key Words: : White Blood Cells ,Deep Learning, Convolutional Neural Network.

1.INTRODUCTION

The microscopic inspection of blood provides diagnostic information concerning patients' health status. The differential blood count inspection results reveal a wide range of significant hematic pathologies. For example, the presence of infections, leukemia and certain specific types of cancers can be diagnosed based on the classification results and the white blood cell count. Experienced operators perform the traditional method for differential blood count. They use a microscope and count the percentage of each type of cell that is counted within a area of interest.

This manual process of counting is obviously very tedious and slow. Furthermore, the classification and accuracy of the cell may depend on the operators ' capabilities and experiences. Consequently, the need for an automated system of differential counting becomes inevitable. Recently, a number of different approaches have been proposed to implement a white blood cell recognition system based on image processing. White blood cell classification usually involves the following three stages: a white blood cell segmentation from an image, the extraction of effective features, and a classifier design. to some extent, the performance of an automatic white blood cell classification system depends on a good segmentation algorithm to segment white blood cells from their background. We extract three types of characteristics from the segmented cell region below. These characteristics are fed into three different neural networks for the classification of five white blood cell types. We extract three types of characteristics from the segmented cell region below. These characteristics are fed into three different neural networks for the classification of five white blood cell types.

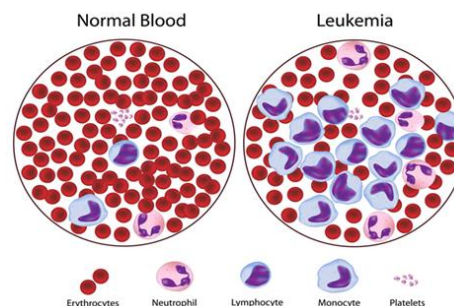


Fig.1 Leukemia blood

1.DATASET

Classify it as either polynuclear or mononuclear given stained image of a white blood cell. Note that while lymphocytes and monocytes are mononuclear, eosinophils and neutrophils are polynuclear. Leukocyte (white blood cells) evaluation is the primary step in diagnosing many diseases related to the blood. The evaluation of the five major leukocyte subtypes -Neutrophils, Lymphocytes, Eosinophils, Monocytes, and Basophils can help to identify various diseases. Manual counting involves white blood cell counting (WBC) done primarily by medical operators, whose accuracy is highly dependent on the skills of the operator. While the impedance based hematology analyzer has its advantages, it may mistakenly classify cell types as white blood cells as their primary classification parameters are limited to size and particle number. Therefore, precise, time saving diagnostic systems need to be introduced in order to accurately classify the count of WBC to determine different diseases.

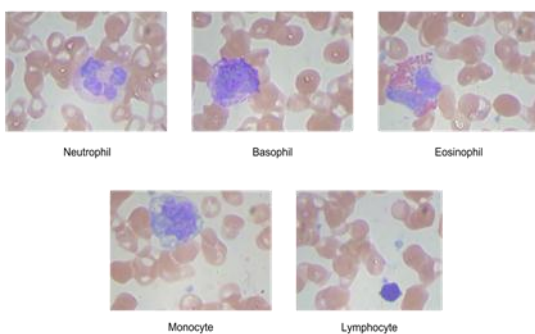


Fig.2 Dataset

2. RELATED WORK

A review based on segmentation techniques (Adollah et al., 2008) argues that conventional color - based methods and thresholding methods are simple to sacrifice accuracy, whereas methods such as region - growing can offer high accuracy with high computation costs. Some methods work directly on the RGB color space, while others work directly on HSI or CMYK color space, referring to the color - based segmentation methods. In general, methods based on the S - component outperform those based on the RGB.

By leveraging the CMYK color models, Putzu et al. (2013) attempts to build the feature vector. They find out that all the other components except white blood cells have some yellow color in them, while leukocytes show a good contrast in the CMYK color model's Y component. Young adopted four characteristics and a minimum distance classifier to classify 5 cell types [4]. Wavelet transform coefficients and artificial neural networks used by Sheik et al. to recognize white blood cells, red blood cells, and platelets [8]. Bikheth et al. selected 10 features and adopted a minimum distance classifier to build an automatic classification system that achieved a 91 percent correct classification rate for a 71 white blood cell database [6]. Piuri and Scotti proposed an automatic classification and detection system based on 23 morphological characteristics and a neural classification system [7]. A classification system was proposed in [5] based on own - cell and parametric characteristics. A system that achieved a 77 percent classification rate for the classification of white blood cells in the bone marrow was reported in [9].

Nilufar et al. proposed a system of classification based on joint histogram - based characteristics and a vector support machine [10]. Osowski et al. presented a genetic algorithm and a vector supporting machine for the recognition of blood cells in the bone marrow [11].

Rezatofighi et al. adopted morphological characteristics and textural characteristics extracted from local binary pattern (LBP) and then trained two types of neural networks for classification [1]. Tabrizi et al. adopted the main component analysis for selection of features and used a neural network of learning vector quantization to classify 5 types of white blood cells [2]. Ghosh et al. provided Naïve Bayes classifier with four statistically significant features to classify five types of white blood cells with an overall accuracy of 83.2 percent [3]. Each approach has its own considerations for adopting features and classifiers of what kinds.

Model Representation

CNN should be useful in classifying images and recognizing objects. It takes the raw pixels as input and produces an outcome indicating the probabilities that

the input belongs to different classes. Instead of implementing the fully connected structure in each layer, CNN imposes two additional layers, convolution and pooling, which can significantly reduce the parameter magnitude. The convolutionary operation entitles the Convolution layer to extract the features from the input images. Kernel modification generates the features that have the effects of the variant, such as object identification, edge detection, image sharpening, etc. The pooling procedure is also referred to as sub-sampling or down-sampling, which is intended to reduce the convolved characteristics produced by the convolution operator with the incentive to remain the significant information. There are various methods of pooling, such as maximum, average, summation, etc. In our method we use the maximize pooling, an example is shown in Fig. 2 A single node is connected to all nodes in the previous layer in a fully connected layer. Moreover, more than one hidden layer may apply, and the different classification operator may be used by the output layer.

i). Convolution

Model Working Convolutionary Neural Network is extensively used to classify images as it uses neighboring pixel information to effectively sample the image and then perform predictions resulting in high accuracy. They also use neural networks that can be scaled to large datasets. It includes a complex neural feed forward network that includes convolutions, pooling, and classification. The term convolution refers to calculating similarities between two functions when one function passes (or convolutes) over another function.

ii). Max pooling

When the image is too large, pooling is used to reduce the number of parameters, followed by training-based classification. A computer-based image is perceived as a collection of three-dimensional numbers or pixels. Width, depth and height. Thus, CNN's core operations are matrix multiplications. As a feature extraction and classification, the functioning of CNN can be divided into two parts. Convolution is primarily responsible for extraction of features. By

sliding a filter (feature vector) over input data, it creates a feature map. This is achieved by multiplying the matrix at each location to extract various parts of the image and summarize the result to a feature map. This operation is performed multiple times to obtain multiple feature maps using different filter values. This is the convolution layer's output. The output is non-negative and non-linear in the real world, so an activation function is applied to it. In this paper we add a layer of pooling between the layers of convolution to reduce the number of computations in the network by reducing the dimensionality. Various types of pooling such as, max pooling (taking the maximum of adjacent pixels after convoluting), average pooling (taking the average of adjacent pixels after convoluting) and sum pooling (taking into account all adjacent pixel values) can be used.

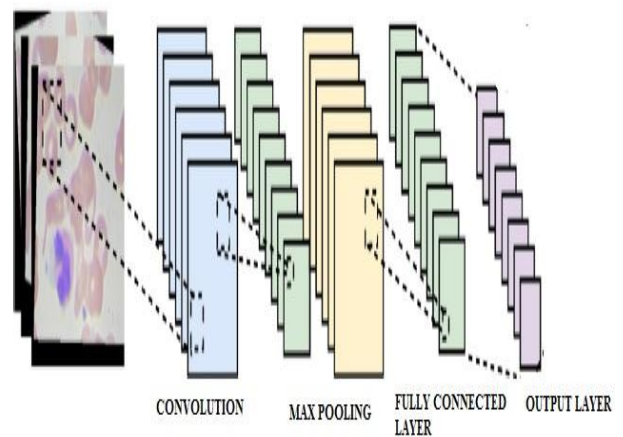


Fig 3. Convolutional Neural Network

Various types of pooling like, max pooling (taking the maximum of neighboring pixels after convoluting), average pooling (taking the average of neighboring pixels after convoluting) and sum pooling (considering all neighboring pixel values) can be used.

iii). Fully Connected Layers

One node is connected to all nodes in the previous layer in a fully connected layer. Moreover, more than one hidden layer may apply, and the different classification operator may be used by the output layer.

IV. Experimental results:

Setups

For this analysis , these setups have been utilized to achieve the result and can be seen below.

Processor :GPU

Language used to implement the project: Python

Tools Used: list of Python libraries are Pandas , Matplotlib, Scikit Learn , Tensor flow etc.

Development Environment Used :Google Collab.

As shown in Table 1, for precision, a value of 83% is obtained, which means that 83% of the times we have the correct (expected) result. Recall is a measure of a prediction model's ability to select instances from a data set of a certain class. A 78 percent recall average shows that 78 percent of the time the model has been able to correctly classify a specific class. F1 score transmits the balance between accuracy and memory. We get a 78 percent F1 score.

Table -1

METHOD CNN			
Class	Precision	Recall	Fscore
NEUTROPHIL	0.60	0.86	0.70
EOSINOPHIL	0.82	0.65	0.71
MONOCYTE	0.84	0.72	0.73
LYMPHOCYTE	1	1	1
AVERAGE	0.81	0.80	0.78

Table 1. Model results Cell Subtype preciseness Recall F1-score white corpuscle zero.57 0.88 0.69 white blood cell zero.96 0.53 0.68 white blood corpuscle zero.84 0.81 0.83 white corpuscle zero.97 0.92 0.94 Average / Total zero.83 0.78 0.78 As shown in table one, a worth of eighty three is obtained for preciseness, which suggests that eighty three of the days we have a

tendency to get the proper (expected) result. Recall may be a live of the flexibility of a prediction model to pick out instances of a definite category from a knowledge set. A recall average of seventy-eight shows that seventy-eight of the days the model was properly able to categorify a specific class. F1 score conveys the balance between preciseness and recall. we tend to acquire Associate in Nursing F1-score of seventy-eight.

3. CONCLUSIONS

In this paper, a classification model based on deep learning was implemented using Convolutional Neural Network to classify the image dataset into four WBCs — neutrophils, lymphocytes, eosinophils, and monocytes. The model achieved 81 percent accuracy on the dataset.

REFERENCES

- [1] Rezatofighi SH, Khaksari K, Soltanian-Zadeh H. Automatic recognition of five types of white blood cells in peripheral blood. Proceedings of the International Conference of Image Analysis and Recognition; 2010; pp. 161–172.
- [2] Tabrizi PR, Rezatofighi SH, Yazdanpanah MJ. Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10); September 2010; pp. 5593–5596.
- [3] Ghosh M, Das D, Mandal S, et al. Statistical pattern analysis of white blood cell nuclei morphometry. Proceedings of the IEEE Students' Technology Symposium (TechSym '10); April 2010; pp. 59–66.
- [4] Young IT. The classification of white blood cells. *IEEE Transactions on Biomedical Engineering.* 1972;19(4):291–298.
- [5] Yampri P, Pintavirooj C, Daochai S, Teartulakarn S. White blood cell classification based on the combination of eigen cell and parametric feature detection. Proceedings of the 1st IEEE Conference on Industrial Electronics and Applications (ICIEA '06); May 2006; pp. 1–4.

- [6] Bikhet SF, Darwish AM, Tolba HA, Shaheen SI. Segmentation and classification of white blood cells. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing; June 2000; pp. 2259–2261.
- [7] Piuri V, Scotti F. Morphological classification of blood leucocytes by microscope images. Proceedings of the IEEE International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA '04); July 2004; pp. 103–108.
- [8] Sheikh H, Zhu B, Tzanakou EM. Blood cell identification using neural networks. Proceedings of the IEEE 22nd Annual Northeast Bioengineering Conference; March 1996; pp. 119–120.
- [9] Umpon NT, Dhompongsa S. Morphological granulometric features of nucleus in automatic bone marrow white blood cell classification. *IEEE Transactions on Information Technology in Biomedicine*. 2007;11(3):353–359.
- [10] Nilufar S, Ray N, Zhang H. Automatic blood cell classification based on joint histogram based feature and Bhattacharya Kernel. Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers (ASILOMAR '08); October 2008; pp. 1915–1918.
- [11] Osowski S, Siroic R, Markiewicz T, Siwek K. Application of support vector machine and genetic algorithm for improved blood cell recognition. *IEEE Transactions on Instrumentation and Measurement*. 2009;58(7):2159–2168.