

# WHITE BLOOD CELL CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORKS WITH TRANSFER LEARNING

Paulsy Tharakan<sup>1</sup>,

<sup>1</sup>PG Student, Computer science and Engineering, Jyothi Engineering College, Thrissur, APJ Abdul Kalam Technical University,

E-mail: paulsytharakan1@gmail.com

Available online at: [www.ijcseonline.org](http://www.ijcseonline.org)

Received: , Revised: , Accepted: , Published:

**Abstract**— Classification of white blood cells play a vital role in predicting the diseases like Leukemia, Anemia etc. Traditional white blood cell classification systems use approaches such as manual inspection, Flow cytometry and Image processing which are found to be tedious, expensive and sometimes inaccurate. We propose a novel approach for classifying white blood cells using convolutional neural networks with transfer learning. The pre-trained convolutional network model Alexnet is used to develop the proposed system which is proven to be efficient with a much better accuracy than the traditional machine learning methods.

**Keywords**— White blood cell classification; Image processing; Convolutional neural networks; Transfer learning; Alexnet

## I. INTRODUCTION

In medical field, the analysis of white blood cells (WBC) in blood is of vital importance for diagnosing diseases. Blood cells begin their life as stem cells, and they mature into three main types of cells— Red blood cells, White blood cells and platelets. Particularly, changes in the distribution of the five types of white blood cells ( Basophil, Eusinoophil, Lymphocytes, Leukocytes and Monocytes) have a close connection with the condition of human immune system. To analyze the different components of WBCs deeply, segmentation and classification of the cells should be performed [1]. A computerized system for differential WBC counting and classification based on morphological features can make manual differential cell counts faster and less tedious for pathologists and laboratory professionals. Traditional approaches of WBC counting are manual microscopy, Flow cytometry and Image processing. Contemporary systems introduce machine learning approaches to the task of WBC classification. Previous WBC identification systems with machine learning involves successive dependent stages; pre-processing, segmentation, feature extraction, feature selection, and classification [2]. To increase the performance of WBC classification systems, proposed system make use of a transfer learning approach using convolutional neural networks. WBCs are classified with respect to the features of their shape, colour and texture. Since the chosen features affect the classifier performance, deciding which features are used in a specific data classification scenario is as important as the classifier itself. Researchers have previously proposed features to differentiate WBCs. For instance, Falcao et al. proposed an unsupervised

automatic WBC segmentation method based upon mathematical morphology and scale-space theories [3]. In the first step, the nucleus of the cell is extracted using the image foresting transform. After that, cytoplasm is segmented using basic operations such as thresholding and morphological opening based on the size distribution of the blood cells. S. Nazlibilek et al. converted an RGB blood smear image into a gray scale image and used Otsu's method to convert the gray scale image to a binary image [4]. The individual images were applied to a neural network-based classifier to classify the cells into the five types. These studies have proven the feasibility and effectiveness of image processing for WBC classification, and shows that they are more flexible, efficient, and reliable than the traditional manual method. However, most of the existing methods focus on two-dimensional images captured by traditional light microscopy and applying image processing based on spatial information to further segment and classify WBCs. Although these methods have decent performance, depending only on the shapes to classify WBCs is unreliable because there are often many variances in the shapes of the five kinds of WBCs. In addition, these methods require complex steps and laborious training of the classifiers to segment and classify WBC. The previous employed algorithms do not consider the characteristics of blood smear images such as their different components, appearance, light distribution and variation of staining intensities.

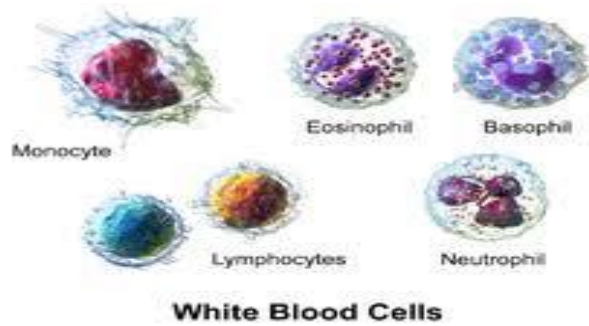


Fig 1 - Different WBCs types ( basophil, eosinophil, lymphocyte, monocyte and neutrophil)

The previous employed algorithms do not consider the characteristics of blood smear images such as their different components, appearance, light distribution and variation of staining intensities. The segmentation algorithms cannot be generalized for all WBCs as their different shapes and colours. The WBCs segmentation also requires the separation of each WBC components (nuclei, cytoplasm) and since each WBC has its own shape and colour nature as shown in Fig.1, it is difficult to have a general segmentation algorithm for all WBCs [5]. Feature selection reduces the feature vector dimension, and decreases the training and testing times. However, classification accuracy may be lowered due to FS process [6]. The multi-stage classification is also increasing the overall system complexity and consumes a lot of processing time. For all of the above reasons, the traditional methods does not achieve the real sense of expertise pathologist in WBCs identification, leaks robustness, and works on limited size datasets under strict conditions. These previous drawbacks create real motivations to employ the deep learning (DL) methodology because of its applicable advantages in our problem-solving. These advantages are as follows; there is no need to perform an enhancement stage for the input image since DL is insensitive to image quality [5], there is no need for neither segmentation stage nor hand-crafted features extraction stage since the features are already extracted through the convolutional concept, and finally the classification stage in DL systems is more simple and does not require a multi-stage of classification. DL is a recent machine learning theory which exponentially grows up in the last years. DL is an extension of artificial neural network (ANN). Convolutional neural network (CNN) was proposed by LeCun [8], which aims to generate learned filters by performing multi-convolution of the same input, through the multilayer network. CNNs have demonstrated as a powerful tool for image recognition, segmentation, detection and retrieval [9].

## II. RELATED WORK

### A. Previous classification approaches

Initially white blood cells were classified using manual microscopy in which pathologists observed blood samples through a microscope and classified them according to their analytical skills. Then technologies such as Coulter counters and Flow Cytometry came into picture. Former classified the WBCs by measuring current conductivity and the latter classifier them by measuring refraction rates on particles using laser lights. Automated approaches of WBC classification were introduced by Q. Wang, L. Chang et.al. [1]. It extracted morphological, statistical and spectral features through image processing and used SVM as the classifier. Then an artificial neural network approach was introduced by S. H Rezatofghi. This paper presented an image analysis system to recognize five groups of white blood cells in the peripheral blood they used the Gram-Schmidt orthogonalization method for the segmentation of the nucleus which can be categorized as the color-based method. Whilst applying three sets of Gram-Schmidt orthogonalizations was simple and had a short processing time, it reduced the susceptibility of the proposed method to color and intensity variations. Another novelty of the method was to propose an adaptive algorithm for finding an initial contour for the snake algorithm. Mazin Z. Othman et.al [11] proposed a Neural Network Classification of White Blood Cell using Microscopic Images. It involves several pre-processing steps such as segmentation, labelling and feature extraction. Based on features extracted such as shape, intensity and texture feature vectors are created. By using feed forward back propagation, data is trained and tested towards a test dataset. Analyzing the previous approaches we understand that they achieve high accuracy only under strict conditions and for small datasets. In [12], it was reported that the morphological features are susceptible to any errors in segmentation stage. All of the above drawbacks lead to a thought of bringing deep learning into the scenario of WBC classification.

### B. Convolutional neural networks

Convolutional Neural Networks (CNN) are akin to ordinary Neural Networks. They consist of neurons which have certain weights and biases. Each neuron accepts some inputs, executes a dot product and follows it with a non-linearity optionally. CNN architectures explicitly assume that the inputs are images and allow users to embed certain properties into the architecture. These then make the forward function more efficient to implement and vastly reduce the amount of parameters in the network. Basic CNN is a series of layers, and each layer of a CNN changes one set of

activations to another set of activations through a differentiable function. We use three main types of layers to build Convolutional neural network architectures: Convolutional layer, pooling layer and Fully-Connected Layer (exactly as seen in regular Neural Networks). We will stack these layers to form full Convolutional neural network architecture.

### C. Transfer learning

Transfer learning is a machine learning method where a pre-trained network model trained on one problem is used to resolve a second related problem. In transfer learning, we initially train a base CNN on a base dataset and task, and then we reuse the learned features, or transfer them, to a second target CNN to be trained on a target dataset and task. This method performs well if the features are common meaning that it is suitable to both base and target tasks, instead of specific to the base task. The concept is that this pre-trained model will act like a feature extractor. We remove the last layer of the network and replace it with our own classifier according to our problem space. We then freeze the weights of all the other layers and train the network normally. We use the pre-trained network called Alexnet [13] to transfer the learned weights. Number of filters increases as the network grows.

## III. METHODOLOGY

We propose a deep learning approach to the problem of classification of WBC using convolutional neural networks with transfer learning. There is no need of pre-processing steps as the features are extracted through convolutional concept itself. The inspiration behind the innovation of CNN is animal's visual cortex. It has small regions of cells which are sensitive to specific regions of the visual field. This idea was expanded upon by a fascinating experiment by Hubel and Wiesel in 1962 where they showed that some individual neuronal cells in the brain responded (or fired) only in the presence of edges of a certain orientation. The concept works as follows: Suppose we have an input image of size  $32 \times 32 \times 3$  where the 3 corresponds RGB colour space. Concept of filter is like moving a flash light over the input image. Both the input image and the filter are actually matrices of pixels where the filter size may be, say  $5 \times 5$ . Dot products of the filter and corresponding input image pixels are computed for each convolution. In this way, each convolutional layer will detect the edges and curves. Dataset used in our experiment was downloaded from Kaggle website which is a famous machine learning data set repository. We take the images in training data set, pass it through the pre-trained network Alexnet which consists of a series of convolutional, nonlinear, pooling (down sampling), and fully connected layers, and get an output. Layers of CNN detect low level features such as edges and curves. To detect the high level features, fully connected layers of the network are used. This layer primarily takes an input volume and outputs an N

dimensional vector where N is the number of classes that the program has to choose from. Architecture of the proposed system is shown in Figure 2.

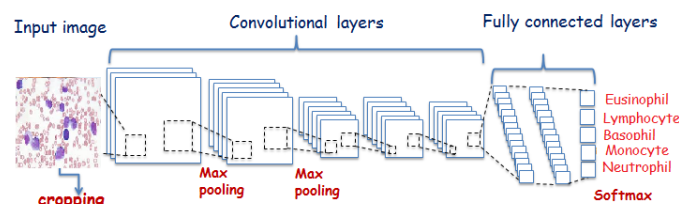


Figure 2 – Architecture of modified Alexnet for WBC classification

The way the system is able to adjust weights of the layers is through a training process called back propagation. It can be separated into 4 different steps, the forward pass, the loss function, the backward pass, and the weight update. A loss function can be defined in many different ways here we use negative log likelihood as the loss function. Finally, to test the accuracy of proposed method, we pass the images in our testing data set through the CNN. We compare the outputs to the ground truth and measure the accuracy of transfer learning approach.

## IV. RESULTS AND DISCUSSION

Comparing the results of transfer learning approach through pre-trained CNN with previous machine learning approaches in task of WBC classification; we understand that our approach performed much better than the traditional machine learning methods. Our approach has less number of pre-processing steps and is insensitive to errors in that steps as the features of images are extracted only through convolutional concepts. Results of experiment during the training and testing phases are shown in Figure 3 and Figure 4.

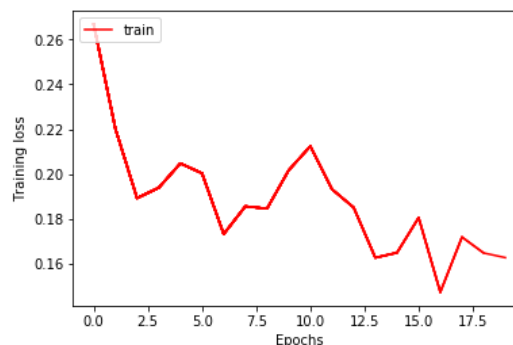


Figure 3 – Measurement of training loss during transfer learning

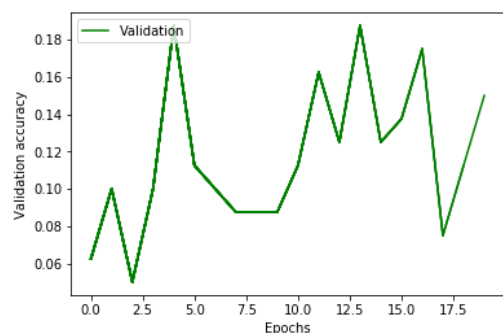


Figure 4 – Measurement of testing accuracy while testing the network

## V. CONCLUSION and FUTURE SCOPE

We used a basic convolutional network prototype to classify the WBC images in our dataset which attained a much better accuracy than traditional methods, based merely on the image data. The proposed model can be extended to more demanding problems with multiple categories, variety of lighting conditions, and cells of new kinds. In this way, machine learning can be pertained to solve problems in healthcare in a meaningful and valuable way.

## REFERENCES

- [1] Qian Wang, Li Chang, Mei Zhou, Qingli Li n, Hongying Liu, Fangmin Guo, "A spectral and morphologic method for white blood cell classification", Article history: Received, 26 January 2016.
- [2] A.X. Falcao, J. Stolfi, R. de Alencar and Lotufo, "The image foresting transform, theory, algorithms and application", IEEE trans, Pattern. Anal. Mach. Intell. 26, 2004(19-29).
- [3] Gholam Reza Abedini, Mohammad Firouzmand, Seyed Ali Razavi Ebrahimi, "Recognition and counting of WBCs using wavelet transform", www.ijettcs.org, Volume 2, Issue 5, September – October 2013.
- [4] S. H. Rezaatofghi and H. Soltanian-Zadeh, "Automatic recognition of five types of white blood cells in peripheral blood," Computerized Medical Imaging and Graphics, vol. 35, no. 4, pp. 333-343, 2011.
- [5] S. Dodge, and L. Karam. "Understanding how image quality affects deep neural networks." Quality of Multimedia Experience (QoMEX), 2016 Eighth International Conference on. IEEE, 2016.
- [6] A.I. Shahin, Yanhui Guo, K. M. Amin and Amr A. Sharawi, "White Blood Cells Identification System Based on Convolutional Deep Neural Learning Networks", *Computer Methods and Programs in Biomedicine*, 14 November 2017.
- [7] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2015.
- [8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of IEEE*, vol. 86, no. 18, p. 2278–2324, 1998.
- [9] Nisha Ramesh, Bryan Dangott<sup>1,2</sup>, Mohammed E. Salama<sup>1,2</sup>, Tolga Tasdizen, "Isolation and two-step classification of normal white blood

cells in peripheral blood smears", *Journal of Pathology informatics*, 16 March, 2012

- [10] Mazin Z. Othman, Thabit S. Mohammed, Alaa B. Ali, "Neural Network Classification of White Blood Cell using Microscopic Images", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 5, 2017
- [11] M. C. SU, C. Y. CHENG and P. C. WANG, "A neural-network-based approach to white blood cell classification," *The Scientific World Journal* 2014, pp. 1-9, 2014
- [12] L. Putzu, Caocci and C. Di Ruberto, "Leucocyte classification for leukaemia detection using image processing techniques," *Artificial intelligence in medicine*, vol. 3, no. 62, pp.179-191, 2014
- [13] Krizhevsky, A.; Sutskever, I. & Hinton, G. E. Pereira, F.; Burges, C. J. C.; Bottou, L. & Weinberger, K. Q. (Eds.) *ImageNet Classification with Deep Convolutional Neural Networks Advances in Neural Information Processing Systems 25, Curran Associates, Inc., 2012*, 1097-1105

## Authors Profile

Ms. Paulsy Tharakan pursued Bachelor of Technology from Jyothi Engineering college, University of Calicut, Kerala, India in 2016 and worked as a consultant at Amazon in 2017. She is currently pursuing Masters in Technology in Jyothi Engineering college under APJ Abdul Kalam Technical University.

