Journal of Contemporary Issues in Business and Government Vol. 26, No. 03, 2020
https://cibgp.com/

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

Classification of Blood Cell Types Using CNN

[1]D Navya Associate Professor,

dubbaka.navya@gmail.com,

[2]E Krishna Associate Professor,

krishna.cseit@gmail.com,

[3]Dr. J Rajaram Professor,

[4]S Raju Assistant Professor,

srajunayak@gmail.com

Department of CSE Engineering,
Pallavi Engineering College,
Kuntloor(V),Hayathnagar(M),Hyderabad,R.R.Dist.-501505.

Abstract

White blood cells, commonly referred to as leukocytes, play a critical part in human immunity development and maintenance. Classifying White Blood Cells plays a critical function in diagnosing sickness in a person. Using the classification, disorders including infections, allergies, anaemia, leukaemia, cancer, and the Acquired Immune Deficiency Syndrome (AIDS), which are caused by aberrations in the immune system, may be more accurately identified and treated as a result. To help haematologists identify the kind of White Blood Cells present in the human body and uncover the root cause of disorders, this categorization is necessary. There is now a lot of study being done in this area. A deep learning technology called Convolution Neural Networks (CNN) will be used to classify WBC pictures into four subtypes, namely neutrophil, eosinophil, lymphocyte, and monocyte, since classifying WBCs has enormous potential. In this work, we'll present the results of a number of experiments on the Blood Cell Classification and Detection (BCCD) dataset, which we used to train CNNs.

Keywords:Basophils, Eosinophil, Monocytes, Lymphocytes, and Neutrophils are all types of blood cells. A framework for deep learning called TensorFlow In the name of Keras function of softmax The Relufunction's. a kind of leukocyte Google's joint venture

Introduction

The human body's immune system relies heavily on white blood cells. Red Blood Cells (RBC) provide oxygen, White Blood Cells (WBC) are the immune system's face, and platelets are responsible for clotting in injured tissues [1]. [2] A healthy adult's white blood cells make about 1% of their total blood volume. In the human body, each kind of white blood cell has a specific purpose and protects against different infections or disorders, which is why they are found throughout the body. As a result, if they detect any of these in the blood, they assault them to offset any possible harm these elements may do in the body. WBCs are distinguished from other blood cell types by their large, lobed nucleus, which is the primary feature of the WBC structure. WBCs have cytoplasm and a cell wall, but no nucleus.

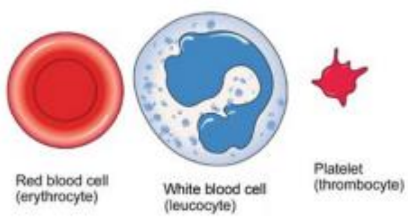

Fig. 1. Types of blood cells.

The human body has five primary types of WBC. There are only four categories: Basophils (0.4 percent roughly), Eosinopoiel (2.3 percent approximately), Monocytes (5.3 percent approximately), Lymphocytes (30 percent approximately), and Neutrophils (62 percent approxAQ3 imately) owing to data set limits (Fig. 1).

Eosinophil

Every day, depending on the season and the stage of life, one's body's Eosinophil count fluctuates somewhat. Two to four per cent of the total WBC count, Eosinophil may remain in circulation for eight to twelve days. The medulla, brain, gastrointestinal tract, and lymph nodes all contain these [2]. The bi-lobed nucleus and skin-red hue of the cell make it easy to identify as an Eosinophil. The nucleus's two lobes are linked by a thin thread of filament. 1.2 percent of the population is composed of monocytes. Monocytes make about 6–9% of the body's total WBC count in a healthy individual. Monocytes have

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

a lifespan ranging from a few hours to many days. T cells can more quickly eliminate these infections thanks to the work of monocytes, which also helps shorten the time it takes for antibodies in humans to produce a response. With specific traits, the nucleus of a kidney-shaped-roundish cell with skin red and some purple hue in it may be identified in the BCCD picture collection [2]. The number of lymphocytes in the blood is around 1.3 million. A healthy person's lymphocyte count is between 25 and 30 percent of their overall white blood cell count. The lymphatic system contains more of these cells than the bloodstream. B-cells and T-cells are the two kinds of cells that make up lymphocytes. These cells attack virus-infected cells in the human [3] body directly and also eliminate cancer cells via Author Proof Blood Cell Types Classification Using CNN 3. The nucleus of a lymphocyte is plainly shaped like a purple potato and eccentric in the BCCD dataset.

1.4 Neutrophils

Innate immunity includes neutrophils. Neutrophils account for around 60–70% of WBCs, making them the WBCs' most significant component. Antibiotic-resistant organisms such as bacteria and fungi are Neutrophils' primary prey. Skin-red cells include multi-lobed groundnut-shaped nuclei that are purple in hue. The nucleus usually has three to five lobes with a clear cytoplasm (Figs. 2 and 3)

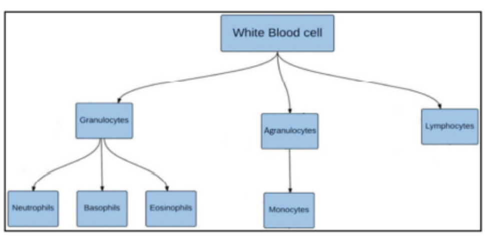

Fig. 2. Classification of WBCs.

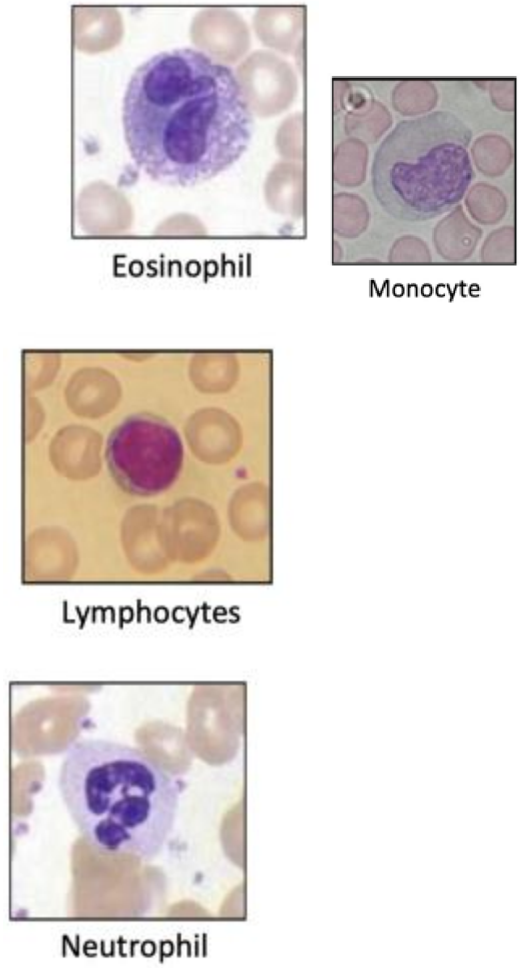

Fig. 3. Types of blood cells

Application of WBC Classification

Hematologist may utilise this categorization procedure to diagnose the patients in an efficient way. For instance, this categorization may be used to examine whether a per son's blood sample comprises of a certain kind of WBC. Generally in the lab oratory to categorise WBC several devices are utilised notably, flow cytometry. Author Proof 4 I. Singh et al. However, this instrument is not very precise hence to overcome this disadvantage an autonomous system incorporating image processing, feature extraction and certain deep learning methods has to be created for more accurate classification. The reduced level of Monocytes might be related to a lower number or absence of WBC in human body

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

which can be due to chemotherapy, bone marrow problem or bloodstream infection. The growing number of Monocytes categorization suggests that the cells are increasing in response to infections, sarcoidosis and Langerhans [6]. B cells and T cells present in the WBC are also poor which implies that chances inhibiting cancer cells is extremely low. If the likelihood of categorising blood cells into Lymphocytes is larger that suggests that the body is suffering from infection (bacterial, viral or other) or Cancer of the blood and lymphatic system. Defi ciency of Neutrophils signifies that a person is suffering from a disease known as Neutropenia. Without these cells human body cannot defend against infections.

Problem Statement

In medical research, one of the issues confronted is the identification and determining the amount of white blood cells. The fundamental cause is the amount of red blood cells. In a healthy adult, the WBCs make around 1 percent of the total blood volume. Due to this low fraction of the WBCs in blood, the identification of the WBCs become a difficulty which further toughens the work of identifying the subtypes of WBCs. The change in number of any subtype of categorization signifies that there is an issue in the body and the body is reacting to a kind of infection. The far ther prognosis of the illness may be assessed by the disease kind and can aid in prescription of therapy for that particular sickness. The standard technique of categorising the WBC kinds comprises evaluating the blood smeared slides under light and electron microscopy. The blood cells are dyed before to viewing them under a microscope [7]. For correct identification, the pathologists need to examine for the form of nucleus and evaluate the size compared to RBCs. Since this is a manual procedure it is prone to mistake and is time consuming. The major draw back with the manual categorization is the element of human error related with mechanical scanning of glass slide and the tradeoff between the picture quality and microscopic field of view (FOV) (FOV). To address these shortcomings, there is a need for an automated approach to categorise the white blood cells utilising an images of stained blood cells.

Blood Cell Count and Detection Dataset

For training the algorithm and executing the validations and tests, we utilised the Blood Cell Count and Classification (BCCD) data set. The data collection includes of enhanced photos of the blood cells in JPEG format. Apart from the blood cell photos, it is accompanied with cell type designations in a .csv file. There are a total of 12,500 labelled photos of the blood cells with around 3000 Author Proof Blood Cell Types Classification Using CNN 5 images for 4 types of WBCs namely Eosinophil, Lymphocyte, Monocyte and Neutrophil [3]. There are two primary folders in the data set named "dataset-master" and "dataset2-master". The "dataset-master" has 410 subtype tagged and augmented photos whereas the "dataset2-master" contains 2500 augmented images. In all, "dataset2-master" comprises 3000 photos for each sub classes whereas "data master" contains 88, 33, 21 and 207 images of the subtypes. In addition to this, the data set includes xml files corresponding to each picture which provides the attributes that may be used to generate bounding boxes on the photos to emphasise the RBC and WBC in the cell image. We are not utilising this element of the data set since our study is confined to identifying the cell picture on the basis of the WBC present in the image

Techniques Used

5.1 Convolution Neural Networks (CNN)

Depending on the picture resolution, the CNN algorithm creates a feature map from the image before classifying it. The CNN network is often made up of numerous hidden layers that are combined with the inputs from the preceding layer by multiplying or using a dot product. Using a feature map created from the picture and a filter/kernel, it is possible to execute operations on the image and get useful results. Edge detection, sharpening, blurring the whole/partial picture, and identification may all be done with these filters [14]. We lower the number of parameters by performing dimension reduction on the pooled output once the convolution method is complete. An ANN architecture is used to process a 3D resulting matrix after pooling. We use flattening on the resulting matrix to organise the 3D volume of numbers into a 1D vector in order to achieve convergence out of 3D input to a 1D array [9]. (See Figure 4)

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

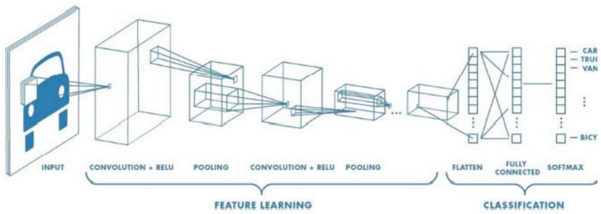

Fig. 4. A figure illustrates basic flow of CNN network

5.2 Linear Rectified Unit (ReLU) Mathematically, the activation function ReLU is defined as y = max (0, x). Extracting just the good aspects of a statement is what this technique is utilised for. When training a neural network, ReLU introduces nonlinearity into the process. Only positive values should be used to train our algorithm in this case. ReLU is simpler than other activation functions since it doesn't need any complicated calculations.

5.2 Rectified Linear Unit

(ReLU) The activation function y = max may be expressed theoretically as a form of ReLU activation function (0, x). Extracting just the good aspects of a statement is what this technique is utilised for. When training a neural network, ReLU introduces nonlinearity into the process. Only positive values should be used to train our algorithm in this case. ReLU is simpler than other activation functions since it doesn't need any complicated calculations.

$$P(y = j) \mid \theta^{j} = \frac{e^{\theta^{(j)}}}{\sum_{j=0}^{k} e^{\theta_k^{(j)}}}$$

KerasKeras

may operate on top of TensorFlow, an open-source neural network framework Because Keras contains implementations of convolutional neural networks and filters, it is employed in the blood cell training model implementation. Other CNN principles like pooling, joining to a highly connected network, dense, and so on may be done using the Keras framework's built-in functions.

TensorFlow

Machine learning models may be built using TensorFlow, which is a network of computation. Similarly to Keras, it is comprised of an ecosystem of tools that may be used to implement algorithms as libraries and create machine learning-related applications from the ground up. There are many abstract techniques that may be expanded and optimised for use in algorithms, much as Keras does. As part of TensorFlow, the architecture on which the application is executing, such as a CPU, TPU, or a GPU, may be selected.

Google Colab

Google Colab was used as a training and testing platform because of the constraints of personal computers. Colab is a Google project that offers a completely cloud-based Jupyter notebook environment for free. As a result of the tremendous computational effort needed to train the CNN network.

Implementation

The batch size is set to 128 in the implementation. The end of an era is signalled by the passage of all pictures across the network once. For the mean and standard picture, numpy combines all of the training photos and applies the mean and standard deviation functions to each pixel position. Mean image is removed from each batch of photographs after the calculation of the images Image results are then subtracted from standard image (Fig. 5)

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

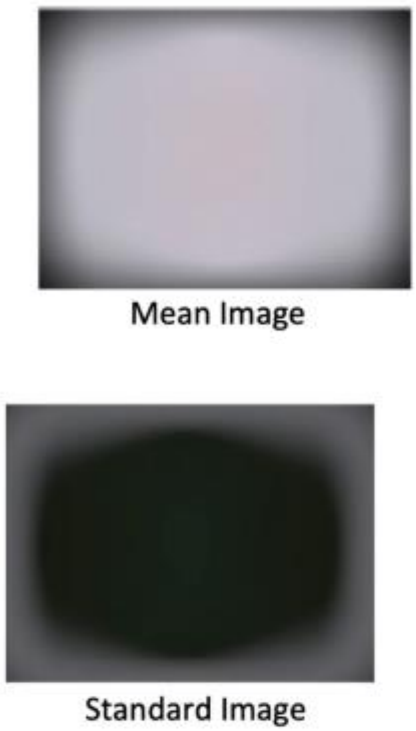

Fig. 5. Images of mean and standard deviation

After that, we separate our data into three sections: training, validation, and testing. The training data has a size of 8961 bytes, while the resulting output data has a matrix size of 8961 by four bytes. A dataset with a size of 996 is used for validation, and the resulting output dataset is 996 * 4 in size (dataset size times the number of classes). Three hidden layers and a single output layer are formed in the CNN model. Conv2D function of Keras developed a convolution kernel that is supplied to the layer input to construct a tensor of outputs for input layer. This layer uses convolution, RELU, Batch Normalization, Dropout, and pooling to extract features from the picture. In order to avoid overfitting, the Dropout is utilised, which randomly changes a proportion of input units to zero at each update during training. For feature extraction, four Convolution layers are constructed using three distinct filters with 16, 8 and 4 layers each. The Dense function of Keras is used to construct three dense hidden layers of 32, 16 and 8 dimensions once the flattening stage has been completed, and an output layer of 4 dimensions is created along with these layers [15]. The next step is to train the model after initialising it. Fit generator() of Keras is used to train the model. It is possible to construct a model summary after putting the training data into an optimizer such as RMSProp, which is an optimizer used to modify neural network properties like weights and learning rates in order to decrease losses (Fig. 6).

Challenges Faced

The major issue was that the system was taking a long time to train and test the model because of a lack of RAM. There were initially roughly 1000 iterations of the model, which would have taken around 8 to 10 hours to train the model and considerably longer for testing. For this, the number of runs was reduced to roughly 200 and the Jupyter platform for Python programming on the cloud, Google Colab, was utilised. 9

Future Work

Each picture in the dataset has a corresponding attribute file that may be accessed while studying the dataset. Bounding boxes' dimensions are specified in the xml file's attributes. On the basis of this information, any cells that appear in the picture may be identified as either a WBC/RBC/platelet. Because of the way the WBC, RBC, and platelet counts may be indicative of different disorders, the cell image count is useful in making medical diagnoses. The major rationale for this diagnosis is the decrease in platelet count that occurs after a dengue infection. 10

Conclusion

Hematologists may now use Convolutional Neural Network algorithms on microscopic pictures of cells to categorise white blood cells into subgroups. Through this categorization, the cells of a patient can be identified, and the kind of sickness they are suffering from may be determined. Images can be identified more reliably using this experiment's findings than they can using traditional lab procedures. For the test set, the accuracy was over 90. Consequently, an ideal model for medical analysis and applications involving the amount of white blood cells and subtypes of white blood cells may be developed using models with high computing capacities.

References

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2020.26.03.016

1. Maton, A.: Human Biology and Health. Prentice Hall, Englewood Cliffs (1993)

2. LaFleur-Brooks, M.: Exploring Medical Language: A Student-Directed Approach, 7th edn, p. 398. Mosby Elsevier, St. Louis Missouri (2008). ISBN 978-0-323-04950-4

3. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: Molecular Biology of the Cell, p. 1367. Garland Science, New York (2002) AQ4

4. Kampbell, N.A.: Biology. Benjamin Cummings, San Francisco (n.d.) AQ5

5. NCI Dictionary of Cancer Terms (n.d.). https://www.cancer.gov/publications/ dictionaries/cancer-terms/

6. Macawile, M.J., Quinones, V.V., Ballado, A., Cruz, J.D., Caya, M.V.: White blood cell classification and counting using convolutional neural network. In: 2018 3rd International Conference on Control and Robotics Engineering (ICCRE) (2018) Author Proof 12 I. Singh et al. 7. Al-Dulaimi, K., Chandran, V., Banks, J., Tomeo-Reyes, I., Nguyen, K.: Classification of white blood cells using bispectral invariant features of nuclei shape. In: 2018 Digital Image Computing: Techniques and Applications (DICTA) (2018)

8. Dertat, A.: Applied Deep Learning - Part 4: Convolutional Neural Networks, Medium, 13 November 2017. https://towardsdatascience.com/applieddeep-learning-part-4-convolutional-neural-networks-584bc134c1e2

9. Prabhu: Understanding of Convolutional Neural Network (CNN) - Deep Learning, Medium, 21 November 2019. https://medium.com/@RaghavPrabhu/ understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148

10. Dernoncourt, F.: What is batch size in neural network? Cross Validated, 01 June 1965. https://stats.stackexchange.com/questions/153531/what-is-batchsizein-neural-network

11. Daniel: What's is the difference between train, validation and test set, in neural networks? Stack Overflow, 01 July 1960. https://stackoverflow.com/questions/ 2976452/whats-is-the-difference-between-train-validation-and-test-set-in-neuralnetwo

12. LNCS Home Page. http://www.springer.com/lncs. Accessed 4 Oct 2017

13. A Comprehensive Guide to Convolutional Neural Networks - the ELI5 way. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neuralnetworks-the-eli5-way-3bd2b1164a53. Accessed 17 Dec 2018

14. ReLU. https://www.tinymind.com/learn/terms/relu. Accessed 30 Oct 2017

15. Di Ruberto, C., Putzu, L.: Accurate blood cells segmentation through intuitionistic fuzzy set threshold. In: 2014 Tenth International Conference on Signal-Image Technology and Internet-Based Systems, pp. 57–64, November 2014. https://doi. org/10.1109/SITIS.2014.43

16. Blood Cell Images. https://www.kaggle.com/paultimothymooney/blood-cells. Accessed 21 Apr 2018

17. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15(1), 1929–1958 (2014)