International Journal for Modern Trends in Science and Technology, 7(09): 07-12, 2021

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0709002
Available online at: http://www.ijmtst.com/vol7issue09.html

White Blood Cell Image Classification Using Deep Learning

Neerukattu Indrani[1], Chiraparapu Srinivasa Rao[2]

[1]PG Scholar, Department of Computer Science, SVKP & Dr K S Raju Arts & Science College, Penugonda, W.G.Dt., A.P, India.
[2]Associate Professor in Computer Science, SVKP & Dr K S Raju Arts & Science College, Penugonda, W.G.Dt., A.P, India.

To Cite this Article
Neerukattu Indrani and Chiraparapu Srinivasa Rao. White Blood Cell Image Classification Using Deep Learning. *International Journal for Modern Trends in Science and Technology* 2021, *7*, 0709006, pp. 07-12. https://doi.org/10.46501/IJMTST0709002

Article Info
Received: 09 August 2021; Accepted: 31 August 2021; Published: 01 September 2021

ABSTRACT

The microscopic inspection of blood smears provides diagnostic information concerning patients' health status. For example, the presence of infections, leukemia, and some particular kinds of cancers can be diagnosed based on the results of the classification and the count of white blood cells. The traditional method for the differential blood count is performed by experienced operators. They use a microscope and count the percentage of the occurrence of each type of cell counted within an area of interest in smears. Obviously, this manual counting process is very tedious and slow. In addition, the cell classification and counting accuracy may depend on the capabilities and experiences of the operators. Therefore, the necessity of an automated differential counting system becomes inevitable.

In this paper, CNN models are used. In order to achieve good performance from deep learning methods, the network needs to be trained with large amounts of data during the training phase. We take the images of the white blood cells for the training phase and train our model on them. With this method we achieved good accuracy than traditional methods. And we can generate the results within the seconds also.

KEYWORDS: Image Processing, Convolutional Neural Network, Deep learning, Histology

1.INTRODUCTION

White Blood cells are key players in the immune system of the human body. There are three broad classifications of blood cells-Red Blood Cells (RBC) that transport oxygen, White Blood Cells (WBC) the face of immune system and platelets that trigger blood clotting in damaged tissues [1]. White Blood Cells makes up 1% of the human blood in a healthy human adult. They are present throughout the body and each type of White Blood Cells have a certain function-ality in the human body and serves by protecting human body against various infections and diseases. If they detect any of these in the blood, they attack them to counter any potential damage these elements can cause in the body [2,6]. The structure of the WBC, predominantly, comprises of a large lobed nucleus that can be used to distinguish a WBC from other blood cell types. Apart from a nucleus, WBC consists of cytoplasm and cell wall.

There are major four categories of WBC in human body. However due to data set constraints we have classified data into four categories: Basophils (0.4% approximately), Eosinophil (2.3% approximately), Monocytes (5.3% approxi- mately), Lymphocytes (30%

approximately) [2,7] and Neutrophils (62% approximately)

1.1Eosinophil

The exact count of the Eosinophil in body keeps on changing during the day depending on the season and during different phases of the human body. On an average, Eosinophil amount to 2–4% of the total WBC count and can persist in the blood circulation for 8–12 days. These are found in the medulla, cortex region, lower gastrointestinal region and lymph nodes [2]. For recognizing the Eosinophil from the images, the cell is rounded in shape of skin-red color with a purple-colored bi-lobed nucleus. The lobes of the nucleus are connected by a thin strand.

1.2 Monocytes

The count of Monocytes present in a healthy human body varies between 6– 9% of total WBC count. The lifetime of Monocyte varies from hours to a day. Monocytes are also responsible for presenting these pathogens to T cells so that they can be easily killed and it helps reduce the response time of antibodies in humans. Monocytes can be recognized in the BCCD image dataset using certain features, the nucleus in it is a kidney shaped-roundish cell with skin red color and some purple color in it without any lobe [2].

1.3 Lymphocytes

Lymphocytes count present in a healthy human body is around 25–30% of total WBC count. These cells are present in lymphatic system than in blood. Lymphocytes consists of two types of cells, namely B-cells and T-cells. These cells responsible for directly killing virus infected cells in human [3] body and also eliminating cancer cells. Lymphocytes can be easily identified in the BCCD dataset by looking at the nucleus as it is clearly round purple colored potato and eccentric.

1.4 Neutrophils

Neutrophils are a part of the innate immune system. The WBCs consists of almost 60–70% of Neutrophils making it largest component of WBCs. The main targets of Neutrophils are bacteria and fungal pathogens. Neutrophils can be recognized as purple colored multi-lobed groundnut-shaped nucleus inside the skin-red colored cells. Generally, there are three to five lobes in the nucleus with a transparent looking cytoplasm (Figs. 2 and 3).

Application of WBC Classification

Hematologist can use this classification process to diagnose the patients in an effective manner. For instance, this classification can be used to check if a person's blood sample consists of a particular type of WBC. Generally in the lab- oratory to classify WBC some instruments are used namely, flow cytometry. However, this instrument is not very accurate so to overcome this disadvantage an automatic system using image processing, feature extraction and some deep learning techniques needs to be implemented for more accurate classification. The lower level of Monocytes can be due to a lower number or lack of WBC in human body which can be due to chemotherapy, bone marrow disorder or bloodstream infection. The increase number of Monocytes classification means that the cells are increasing in response to infections, sarcoidosis and Langerhans [6]. B cells and T cells present in the WBC are also deficient which means that chances inhibiting cancer cells is very low. If the chance of classifying blood cells into Lymphocytes is more that means that the body is suffering from infection (bacterial, viral or other) or Cancer of the blood and lymphatic system. Deficiency of Neutrophils means that a person is suffering from a problem known as Neutropenia. Without these cells human body cannot resist from infections.

2. LITERATURE REVIEW:

Classification of white blood cells from microscope images is a challenging task, especially in the choice of feature representation, considering intra-class variations arising from non-uniform illumination, stage of maturity, scale, rotation and shifting. In this paper, we propose a new feature extraction scheme relying on bispectral invariant features which are robust to these challenges. Bispectral invariant features are extracted from the shape of segmented white blood cell nuclei. Segmentation of white blood cell nuclei is achieved using a level set algorithm via geometric active contours. Binary support vector machines and a classification tree are used for classifying multiple classes of the cells. Performance of the proposed method is evaluated on a combined dataset of 10 classes with 460 white blood cell images collected from 3 datasets and using 5-fold cross validation. It achieves an average classification accuracy of 96.13% and outperforms other popular representations including local binary pattern,

histogram of oriented gradients, local directional pattern and speeded up robust features with the same classifier over the same data. The classification accuracy of the proposed method is also compared and benchmarked with the other existing techniques for classification white blood cells into 10 classes over the same datasets and the results show that the proposed method is superior over other approaches.[7]

Counting and analysis of blood cells allow the evaluation and diagnosis of a vast number of diseases. In particular, the identification of white blood cells and red blood cells is a topic of great interest to hematologists. Nowadays the observation of blood samples is still performed manually by skilled operators. This task is tedious, lengthy and repetitive, and the results accuracy heavily depends on the operator skills. Differently, the automated analysis by computer is performed quickly, requires only one image of the blood sample and provides precise results. One of the major steps on image analysis is segmentation, that subdivides images into meaningful parts. Thresholding is one of the most used technique and subdivides the image pixels on the basis of their intensity grey levels. There are many crisp techniques for calculating the threshold value of an image. Recently many intuitionistic fuzzy methods have been proposed to determine the optimal threshold value, showing better results for segmentation but not even better computational performance. In this paper we propose an intuitionistic fuzzy set approach for optimal threshold selection based on computations performed on the histogram. This method is then extended in order to perform multiple thresholds and in order to take into account possible local variations on the image. The proposed approach has been tested on peripheral blood images, in order to subdivide the various image components, showing excellent performance both for segmentation both in terms of speed[15]

Establishing an accurate count and classification of leukocytes commonly known as WBC (white blood cells) is crucial in the assessment and detection of illness of an individual, which involves complications on the immune system that leads to various types of diseases including infections, anemia, leukemia, cancer, AIDS (Acquired Immune Deficiency Syndrome) etc. The two widely used methods to count WBC is with the use of hematology analyzer and manual counting. Currently, in the age of modernization there has been numerous research in the field of image processing incorporated with various segmentation and classification techniques to be able to generate alternatives for WBC classification and counting. However, the accuracy of these existing methods could still be improved. Thus, in this paper we proposed a new method that could segment various types of WBCs: monocytes, lymphocytes, eosinophils, basophils, and neutrophils from a microscopic blood image using HSV (Hue, Saturation, Value) saturation component with blob analysis for segmentation and incorporate CNN (Convolutional Neural Network) for counting which in turn generates more accurate results[6]

3. PROPOSED SYSTEM:

Recently, deep learning-based approaches were shown to outperform conventional machine learning methods in many image analysis tasks, automating end-to-end processing. In the domain of medical imaging, convolutional neural networks (CNN) have been successfully used for White blood cell classification, cancer tumours, brain blood clots and so many more.In this work, we present an approach for microscopy image analysis for White blood cell classification. Our approach utilizes deep CNNs for feature extraction and gradient boosted trees for classification and, to our knowledge, outperforms other similar solutions.

- More accurate results than traditional methods
- We can generate results within seconds
- No experienced people need.
- We can deploy our model in emergency situations.
- Cost efficient also.

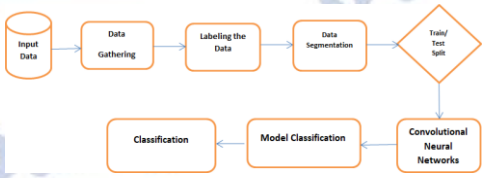


Fig1: Architecture of Proposed Work

4. METHODOLOGY:

Convolutional Neural Networks

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases)

to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area.

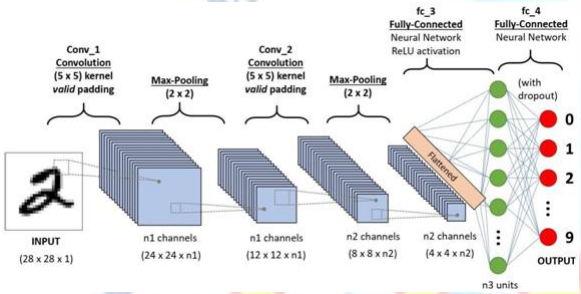


Fig. Convolutional Neural Networks Architecture

Convolution Layer

Convolution is the first layer to extract features from an input image. Convolution preserves the relationship between pixels by learning image features using small squares of input data. It is a mathematical operation that takes two inputs such as image matrix and a filter or kernel.

An image matrix (volume) of dimension (h x w x d)

A filter (fh x fw x d)

Outputs a volume dimension (h - fh + 1) x (w – fw + 1) x 1

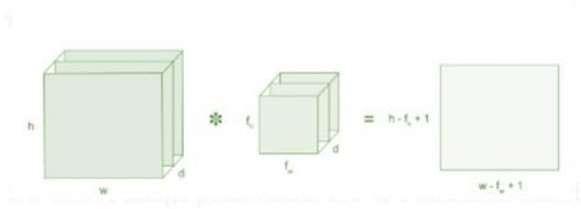


Fig: Image matrix multiplies kernel or filter matrix

Consider a 5 x 5 whose image pixel values are 0, 1 and filter matrix 3 x 3 as shown in below

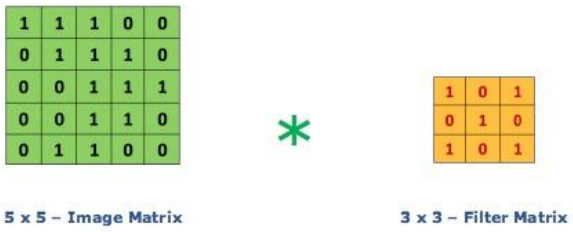


Fig. Image matrix multiplies kernel

Then the convolution of 5 x 5 image matrix multiplies with 3 x 3 filter matrix which is called "Feature Map" as output shown in below. Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters.

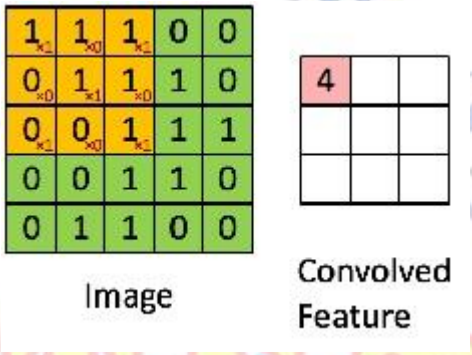


Fig. 3 * 3 Output Matrix

Strides

Stride is the number of pixels shifts over the input matrix. When the stride is 1 then we move the filters to 1 pixel at a time. When the stride is 2 then we move the filters to 2 pixels at a time and so on. The below figure shows convolution would work with a stride of 2.

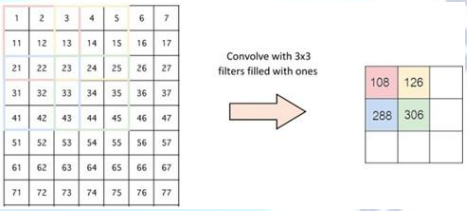


Fig: Stride of 2 Pixels

Padding

Sometimes filter does not fit perfectly fit the input image. We have two options:

Pad the picture with zeros (zero-padding) so that it fits

Drop the part of the image where the filter did not fit. This is called valid padding which keeps only valid part of the image.

Non Linearity (ReLU)

ReLU stands for Rectified Linear Unit for a non-linear operation. The output is $f(x) = \max(0,x)$.

Why ReLU is important : ReLU's purpose is to introduce non-linearity in our ConvNet. Since, the real

world data would want our ConvNet to learn would be non-negative linear values.

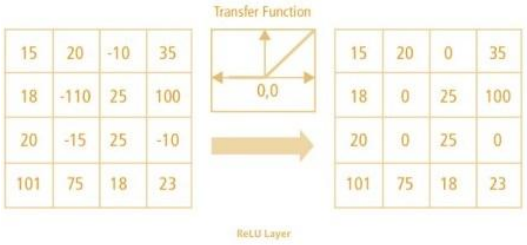


Fig. ReLU Operation

There are other non linear functions such as tanh or sigmoid can also be used instead of ReLU. Most of the data scientists uses ReLU since performance wise ReLU is better than other two.

Pooling Layer

Pooling layers section would reduce the number of parameters when the images are too large. Spatial pooling also called subsampling or downsampling which reduces the dimensionality of each map but retains the important information. Spatial pooling can be of different types:

Max Pooling

Average Pooling

Sum Pooling

Max pooling take the largest element from the rectified feature map. Taking the largest element could also take the average pooling. Sum of all elements in the feature map call as sum pooling.

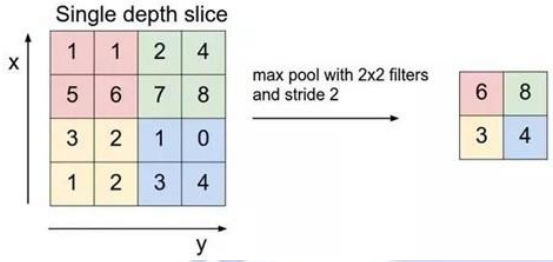


Fig. Max Pooling

Fully Connected Layer

The layer we call as FC layer, we flattened our matrix into vector and feed it into a fully connected layer like neural network.

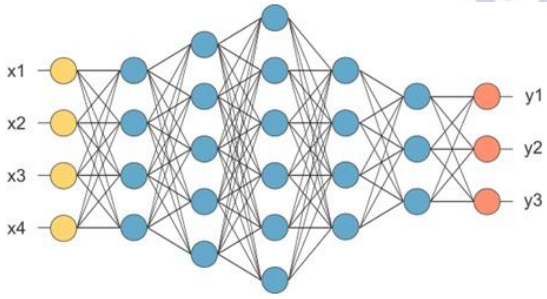
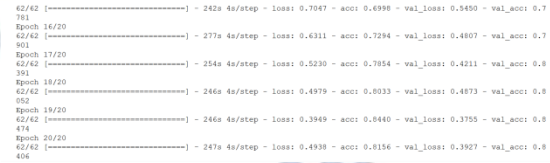


Fig: After pooling layer, flattened as FC layer

In the above diagram, feature map matrix will be converted as vector (x1, x2, x3, …). With the fully connected layers, we combined these features together to create a model. Finally, we have an activation function such as softmax or sigmoid to classify the outputs as cat, dog, car, truck etc.,

5. EXPERIMENTAL RESULTS:



We can see that the metrics improved when we increased the number of epochs to 10 from 1. The loss function for training dropped down to 0.865. The accuracy increased significantly to 80% for training. Similarly, a significant improvement can be seen in validation and test part as the loss function value dropped to 0.57 and 2.02, respectively.

	precision	recall	f1-score	support
EOSINOPHIL	0.29	0.18	0.22	499
LYMPHOCYTE	0.25	0.24	0.24	496
MONOCYTE	0.24	0.26	0.25	495
NEUTROPHIL	0.25	0.34	0.29	499
accuracy			0.25	1989
macro avg	0.26	0.25	0.25	1989
weighted avg	0.26	0.25	0.25	1989

```
: print('Test score:', score)
  print('Test accuracy:', acc)

Test score: 0.3941345512866974
Test accuracy: 0.8406234383583069
```

After we training 20 epochs of the cnn network we get the train accuracy of the 84%.And also for the test dataset also we get the results 84% and precision of the each class images also really good.

6. CONCLUSION:

This research helps the hematologist to classify White Blood Cells into their subtypes with the help of microscopic images of cell using Convolutional Neural Network techniques. This classification helps to distinguish the cells and check what type of disease a patient is suffering from. The results obtained from this experiment helps identify images in a robust way as compared to the orthodox lab methods. The good level of accuracy above 90 for the test set. Hence, when the model is trained with high computational abilities

present, a perfect model can be trained and can be applied in the medical analysis and applications dealing with the number of white blood cells and sub types of white blood cells.

On analyzing the dataset, there are attribute files attached with the dataset cor-responding to each image. In that xml file, the attributes specify the dimensions of bounding boxes. Based on that, every cell present in the image can be classified as a WBC, RBC or platelet. The cell image count helps in diagnosis of various diseases as an elevation or depreciation in the count of WBC, RBC and platelets can be a pointer to identify an ailment. The main reason for this diagnosis is that during a dengue infection, the number of platelets starts decreasing.

REFERENCES

[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.)

[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)

[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/applied-deep-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-batchsize-in-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-neural-netwok.

[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-neural-networks-the-eli5-way-bd2b1164a53. 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/bloodcells. 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)

ABOUT AUTHORS:

Neerukattu Indrani is currently pursuing MCA in SVKP & Dr K S Raju Arts & Science College, affiliated, to Adikavi Nannaya University, Rajamahendravaram. Her research interests include Data Structures Web Technologies, Operating Systems and Artificial Intelligent.

Ch.Srinivasa Raois a Research Scholar in the Department of Computer Science & Engineering at Acharya Nagarjuna University, Guntur, A.P, India. He is working as Associate Professor in SVKP & Dr K S Raju Arts&Science College, Penugonda, A.P. He received Master's degree in Computer Applications from Andhra University and Computer Science & Engineering from Jawaharlal Nehru Technological University, Kakinada, India. He Qualified in UGC NET and AP SET. His research interests include Data Mining and Data Science.