Journal of Critical Reviews

ISSN- 2394-5125 Vol 7 , Issue 9, 2020

CHARACTERISTICS DIAGNOSIS OF WHITE BLOOD CELLS USING CNN

A. vanathi[1], Dr.N. Naga malleswara rao[2] , Naga Subrahmanyeswari.N., PNS Lakshmi

Assoc Professor In Aditya Engineering College (A)& Research Scholar In Anu, Guntur
Orcid: 0000-0001-9127-0353
Professor, Rvr&Jc College Of Engineering(A),Guntur
Nnmr3654@Gmail.Com
Lecturer In Computer Science,A.S.D.Government Degree College For Women (A) Orcid Id: 0000-0001-9616-4242
Sr.Asst Professor, Dept Of Cse, Aditya College Of Engineering
Lakshmi_Cse@Acoe.Edu.In

Received: 23.03.2020 ***Revised: 24.04.2020*** ***Accepted: 25.05.2020***

Abstract
The traditional method for many disease diagnoses require blood sample analysis in which the white blood cell play a vital role. The White Blood Cell (WBC) defend the body against infections. The shape of the cell, the size of nucleus also helps in diseases diagnosing. The technique used for blood sample diagnosis may take a very long time. In some cases, the patient may die due to the delay of the reports and disease diagnosis so, we have proposed the idea of automated diagnosis of white blood cells and predicting related disorders. . In the present days, the research in medical field has begun to take machine learning and data mining techniques for the development of software that are related in extract of information. Researchers had made important changes on the development of systems which are efficient to scrutinize different types of medical images. Machine Learning enables to learn from data and improve its efficiency.

Keywords: While Blood Cells, Machine Learning, disorders

DOI: http://dx.doi.org/10.31838/jcr.07.09.204

INTRODUCTION

White blood cells are that which provide immunity that includes in protecting the body against disease. White blood cells are produced in the bone marrow which is called as hematopoietic stem cells. Leukocytes are foundall over body that includes in the blood and lymphatic system.It iscapable of portability and defends the body against infection and disease by ingesting foreign materials by destroying the harmful cells present in the body by producing the anti-bodies. Types of WBC's are neutrophils, eosinophils, basophils, lymphocytes,
and monocytes.

In the recent years with the improvement in technology, Machine Learning (ML) and Artificial Intelligence (AI) have developed fastly . Themechanizationof ML and AI have plays a vital role in medical field like image processing, computer-aided diagnosis, image interpretation, image segmentation, image retrieval and analysis of the image. Procedures of ML are extract information from the images and represents information effectively and efficiently. The ML and AI assists the doctors for the easy diagnose of diseases and prevent them rather than taking long time. These technique increases the ability of doctors and researchers for the better analysis of the disease and the usage of modern technology and the root cause of the disease.These techniques consists of conventional algorithms like Support Vector Machine, Neural Network, and deep learning algorithms such as Convolutional Neural Network , Recurrent neural Network , Extreme Learning Model, Generative Adversarial Networks etc. Machine Learning helps the computer to learn from previous data and improve its efficiency on the dataset. Machine Learning is a branch of Artificial Intelligence and it has many applications.The discovery of blood-based diseases frequently requires theidentifying and characterizing patient blood samples. we can use automated methods to detect and classify blood cells sub types have important medical implementation. The cell types are Eosinophil, Lymphocyte, Monocyte, and Neutrophil. To classify and detect blood cells sub types we use Convolution neural networks (CNN).

Related Works

Many methods have been proposed for the white blood cells detection using various algorithms. [1]One of the method is the WBC is detected on the basis of Artificial Bee colony algorithm. Since White blood cells can be identified in elliptical shape by using thedetection ofellipse method that we can identify the elements. The artificial bee colony (ABC) algorithm is used for the automatic detection of White blood cells in the colour blood cell images, by transforming it into an maximization problem. This method consists of boundary points as an ellipses in the blood smear. The main goal isto find the resemblance of a candidate ellipse with a real WBC on the cell image. Now the final process, which consists of the two neutral functions, they are candidate thresholds and the candidate ellipses which involves using the artificial bee colony algorithm (ABC).[2]Another method uses the Genetic Algorithm (GA) and Otsu threshold method. This method is rest on image equalization in HSV color space using Otsu threshold method.This process helps in identification of the WBCs in the microscope images of blood whichare based on the merging of Otsu method and GA. After this process the count of blood cells can be identified simply. After this type of blood cells images,it should be taken into consideration along with the WBC and RBCwhich can not be classified in the normal blood. The blood cells are treated in laboratories with different methods and services.[3]The other proposed system includes WBC segmentation algorithm which is based on the sparsityand geometry constraints. This segmentation is done by the nuclei by a sparsity constraint and an image representation method andthis features of nuclei more convenient, and therefore it can be detect better. In the next step, a geometrical constraint is used to detect cells. The main idea in this process is the usage of a model fitting procedure which is used to recover cells from weak and incomplete boundaries.

Proposed Method

The proposed system comprises of classification of WBC and detection of the disorders depending on the count. The WBC are of four subtypes namely neutrophils, monocytes, lymphocytes, Eosinophils. The WBC's are identified from the microscopic blood smear images and later they are classified into subtypes. Later depending on the count of the white blood cells the disorders are identified. The general range of the subtypes are as follows:

Table 1: Types of WBC

Type of WBC	Normal percentage of overall WBC count
neutrophil	55 to 73 percent
lymphocyte	20 to 40 percent
eosinophil	1 to 4 percent
monocyte	2 to 8 percent

METHOD

In this section the dataset used, data preprocessing, data augmentation and the CNN algorithm are described. The white blood cell images are collected and are classified based on their characteristics.

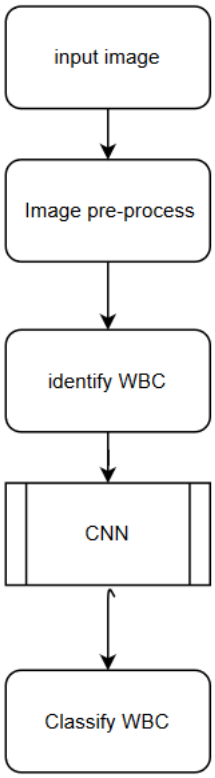

Figure 1: Flowchart of the proposed model

Data set:

The white blood cell images are obtained from the BCCD dataset which is available on the Kaggle.com website. The BCCD dataset consists of microscopic images of blood smear from which we have to distinct WBC's using data segmentation. The white blood cells are of four sub types namely neutrophils, monocytes, lymphocytes, Eosinophils.

Data Preprocessing:

The dataset are in the form of images, every image may contain noisy data than the required data.To remove the noisy data and to obtain more precise image image processing is needed. First image is converted to grey scale images to convert it into an array. Later the images are resized to remove noisy data.

Image Conversion and Augmentation: The microscopic images are converted to grey scale images using opencv module. The grey scale images are then resized to 60x60. Resizing the images removes the noisy data. The resized images are converted into an array using numpy module.

About Convolution Neural Network:

Convolution Neural Network (CNN) become commanding in different types computer works and it is showing interest towards a different types of domains, and radiology. It is a class

of deep learning methods. Convolution neural network which consists of several building blocks, they are convolution layers, pooling layers, and fully connected layers. It is mainly created to learn the features of spatial hierarchies across a back-propagation algorithm automatically.

In networks, CNN is the main grouping that is used to do image recognition and classification, detection of objects, face recognition. The CNN is widely used in the above areas.

The CNN image classifications takes an image(ex: dog, cat, tiger) as input, and it undergoes certain classification processes by using convolutional layers like pooling, full connected, Relu. Computers defines an input image in the form of an array with pixels. Based on the image resolution the computer defines an input image. By using image resolution it divides image pixels by considering certain factors height, width, dimension in the form of h*w*d.

Convolutional layer:
Convolutional layer is one of the important layers of CNN. It is the first layer that takes the characteristic of an input image. It uses the input small squares of the Convolution preservers the relationship between the pixelsby learning the image features. Convolution filters are applied by this layer simultaneously on the input image.After applying the filters on the input images, thefilters activates certain specific features. The output obtained from the previous layer will be act as an input to the next layer.

Strides:
Stride is so far the filter moves in every step along one direction. It is the number of pixels shifts for the input matrix. When the stride value is 1, then the filter will move to one pixel at a time. When the stride value is 2, then the filter will move two pixels at a time and so on.

Max Pooling:
Max pooling is an operation which is connected to CNN with convolutional layers. When this process is applied, max pooling reduces the proportions of the images. It is reduced by the number of pixels in the output which comes from the previous convolutional layer. Max pooling used for reducing the resolution of the output and helps to reduce ambiguity.

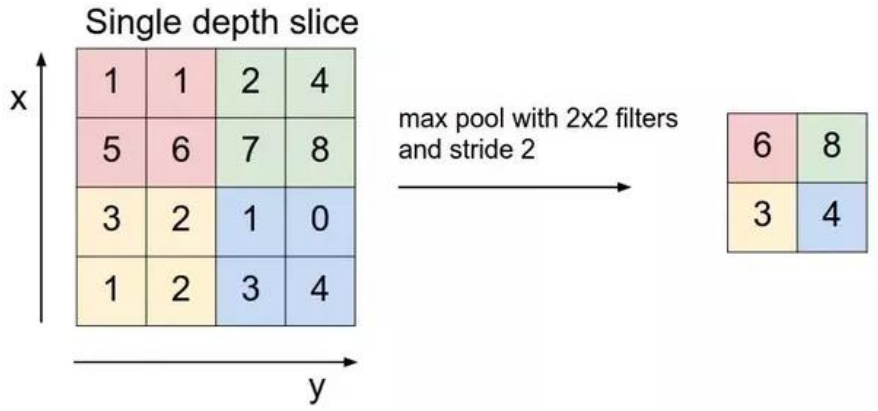

Figure 2: max pool with 2x2 filters

EXPERIMENT AND RESULT
The classification of wbc begins by taking the microscopic blood smear images as input . We have identified the WBc's from those images and further classified them depending on the shape. The CNN algorithm works on the grey scaled images which have been converted by using the opencv module. The opencv module helps in image preprocessing .The grey scaled images are cropped to determine the distinct WBC's. Later the images are converted into an array and feed to the CNN model.The dataset consist of approximately 3000 augmented images of all the four categories of white blood cells arranged in four folders. We have split the dataset into training and test data. The CNN model has got an accuracy of 97 %.

```
#count labels

sns.countplot(train_y,palette='Set3')
#we can see each categroy has equal data
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x2195d886b00>
```

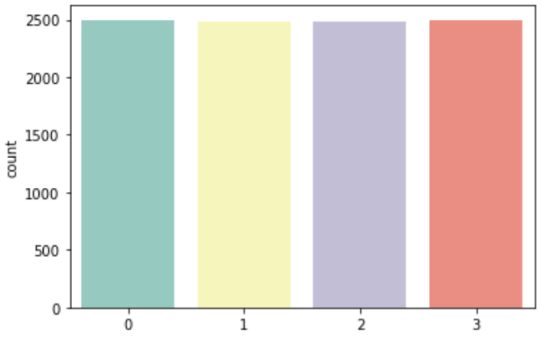

Figure 3: Bar graph representing the count of images of each class

Table 2: CNN model used for training

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 58, 58, 32)	896
max_pooling2d (MaxPooling2D)	(None, 29, 29, 32)	0
dropout (Dropout)	(None, 29, 29, 32)	0
conv2d_1 (Conv2D)	(None, 27, 27, 64)	18496
max_pooling2d_1 (MaxPooling2	(None, 13, 13, 64)	0
dropout_1 (Dropout)	(None, 13, 13, 64)	0
conv2d_2 (Conv2D)	(None, 11, 11, 128)	73856
max_pooling2d_2 (MaxPooling2	(None, 5, 5, 128)	0
dropout_2 (Dropout)	(None, 5, 5, 128)	0
flatten (Flatten)	(None, 3200)	0
dense (Dense)	(None, 64)	204864
dense_1 (Dense)	(None, 128)	8320
dense_2 (Dense)	(None, 64)	8256
dense_3 (Dense)	(None, 4)	260

Total params: 314,948
Trainable params: 314,948
Non-trainable params: 0

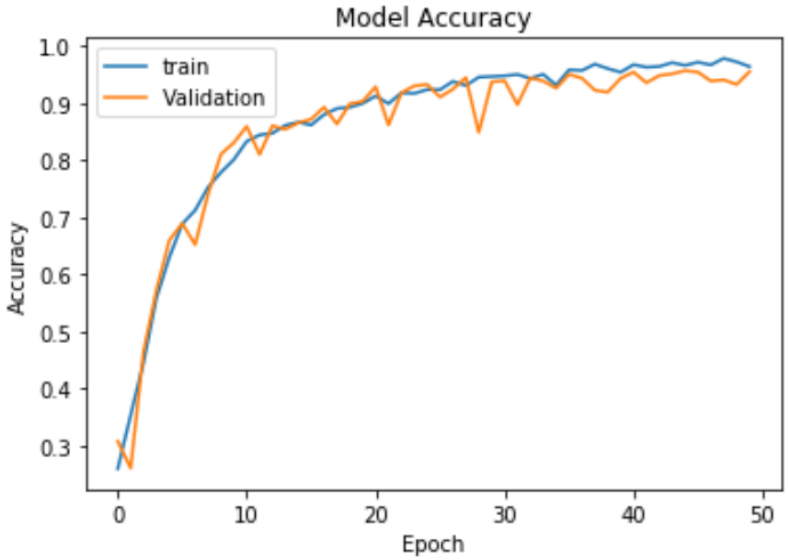

Figure 4: Accuracy graph of the model at different epoch values

CONCLUSION AND FUTURE SCOPE

In this paper , it shows the detail approach of detection of the blood cell subtype using the microscopic images . The images are taken and trained using the CNN algorithm .This model achieved the precision of 97.19% on the testing dataset. Deep learning proves significantly better than the traditional approach. The usage of the deep learning provides a scalable approach and enhanced extraction of the images. This method will reduce the work done by the pathologist. It reduces the manual errors and increases the accuracy. In the future , the enhancement of this model can be used for the automatic counting of the blood cells and the detection of the blood disorder present.

REFERENCES

1. Zheng Fu1, Ye Liu1, Haidong Hu2, Dongmei Wu1, Hao Gao1,"An Efficient Method of White Blood Cells Detection Based on Artificial Bee Colony Algorithm",IEEE,July-2017.
2. M. Khodashenas ,H. Ebrahimpour-komleh , A. M. Nickfarjam, "White Blood Cell Detection and Counting based on Genetic Algorithm", IEEE, March - April-2019.
3. Zhen Zhong ,Tao Wang , Kun Zeng , Xiaogen Zhou, "White Blood Cell Segmentation via Sparsity and Geometry Constraints", IEEE, November-2019.
4. OTSU N. "A threshold selection method from gray levelhistograms". IEEE Trans on SMC, 1979.