JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

ANALYSING MICROSCOPIC IMAGES OF PERIPHERAL BLOOD SMEAR USING DEEP LEARNING

[1]Shreyas Raju A, [2]Dr. Muzameel Ahmed,

[1]PG Student, [2]Associate Professor,
[1]Department Of ISE,
[1]DSCE,VTU,Bangalore, India,
[2]Department Of ISE,
[2]DSCE,VTU,Bangalore, India

Abstract : The analysis of blood smear is an important medical diagnostic test to evaluate the general state of health. Generally blood cells are counted manually by means of a hemocytometer alongside other laboratory equipment's and chemicals, this process requires a lot of time and effort. In this work a method is proposed using Machine learning technique for analysing the blood smear for classification of blood cell types like Erythrocytes known as Red Blood Cells (RBC), Leukocytes known as White Blood Cells (WBC), and Thrombocytes known as Platelets using "YOLO - You Only Look Once" state of the art object detection and classification algorithm. The YOLO network is trained with dataset collected from Blood Cell Count Dataset (BCCD) of blood smear images. The trained YOLO frame work will identify and classify the blood cells into respective blood cell category. The result obtained here has better accuracy and performance, compared with other CNN - Convolutional neural network architecture.

IndexTerms **-Platelets, White Blood Cells, Red Blood Cells, You Only Look Once, Blood Cell Count Dataset, Convolutional Neural Network.**

I.INTRODUCTION

Blood smear is sample of human blood that has been distributed on a glass slide and stained specifically. The common blood cell types found in human blood are erythrocytes, leukocytes and thrombocytes also known as RBC, WBC and Platelets respectively. The RBC are commonly produced by bone marrow, the cytoplasm of RBC is rich in hemoglobin which is responsible for red colour of the cells and blood, the RBC main function is to carry the oxygen from lungs to different tissues and organs of the human body, the normal healthy person RBC count will be between 4.7 to 6.1 million cells per micro litre (cells/mcl). The WBC are procured from multipotent cells in bone marrow known as hematopoietic stem cell, the WBC are found in immune system of the human body, the main functionality of WBC are, they fight against the contagions and disorders in human. The normal healthy person WBC count will be between 4,500 to 11,000 cells per micro litre. The platelets are particles of cytoplasm that are procured from megakaryocytes of bone marrow or lungs, the platelets react to bleeding from blood vessels injuries and will initiate the blood clotting. The normal healthy person platelet count will be between 150,000 to 450,000 cells per micro litre.

In this work, deep learning method for blood cell classification has been proposed. Convolutional neural network (R-CNN), YOLO regions are one of the most advanced object detection systems, among them YOLO frame work has been chosen which is quicker than R-CNN along with architecture of VGG -16. YOLO utilise a solitary neural network to estimate the bounding box and probability class straightly from the blood smear image in a single assessment, so YOLO frame work has been chosen for automatically identify and classify RBC's, WBC's and Platelets cells from datasets of blood smear. To boost the accuracy and performance, a method for verification has been implemented, so imitated counting of cells has been avoided by frame work.

Literatures Survey : In this research area, there are some works has been carried out by authors like

[1] Dheeraj Mundhra et.al. (2017) in their work they have proposed automated peripheral blood smear analysis system consists of an automated microscope and software component for analysis of the blood smear images. They have used methodology like U-net deep learning architecture for WBC and platelets, Image processing for RBC extraction. [2] Hany A. Elsalamony proposed an algorithm for detecting the different types of anaemia-contaminated red blood cells, using methodologies like Circular Hough transform (CHT), Watershed and Morphological tools. [3] Roopa B. Hegde et.al. (2018) they have summarised the method of image processing techniques for analysing blood smear images, they have categorise the methods into analysis of RBC, WBC and Platelets respectively. [4] Biji G et.al. (2017) they have proposed the methodology for detecting white blood cells in complicated blood smear images using Electromagnetism-like optimisation algorithm. [5] Vasundhara Acharya et.al. (2021) they have proposed model for blood cell segmentation in blood

smear images for identifying the stage of Acute Myeloid Leukemia using K-medoids algorithm, Modified Watershed Transform to Separate Touching Cells.

This literature survey conducted, has suggested that there is a room to improve the blood smear detection method, the paper focus on the proposed YOLO frame work that will detect 3 blood cell types simultaneously, without any greyscale conversion of blood smear images.

II. Data and Sources of data

Collected the related data sets of blood smear from publicly available data set known as complete blood cell (CBC) and Kaggle data base.

Types of data collected:

- Red Blood Cells - 11400 images.
- Platelets - 2348 images.
- White Blood Cells
- Neutrophil - 3330 images.
- Basophil - 1218 images.
- Eosinophil - 3117 images.
- Lymphocyte - 1214 images.
- Monocyte - 1420 images.

III Methodology

YOLO : In the area of object detection algorithms, in this work You only look once algorithm has been chosen, it is the state-of-the-art object detection and classification algorithm, it normally functions as a regression issue for object detection. Along the network, just one forwards propagation pass is necessary to quickly predict both image class and location. The YOLO frame work will divide the image into grids, by applying 7*7 grid lines and every grid will predicts bounding box and the confidence score. If an object's centre point falls within the grid cell, that grid cell is in charge of locating that object. The below figure shows deep learning-based system, proposed for identifying blood cell types.

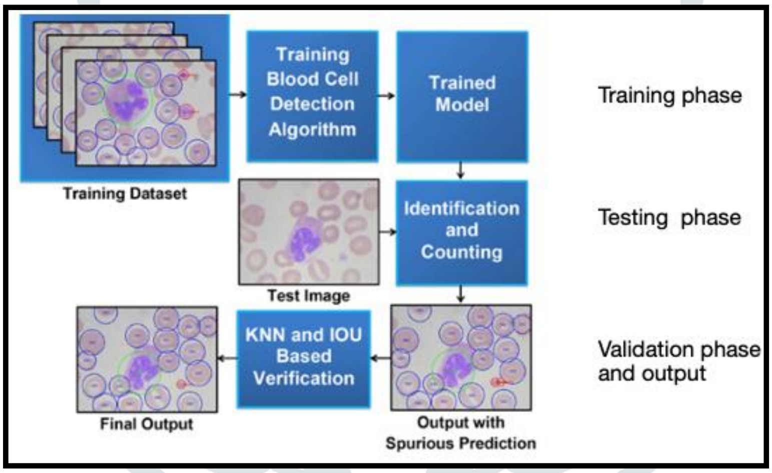

figure 1. block diagram of blood cell identification.

The proposed method i.e; YOLO frame work will detect 3 blood cell types simultaneously, without any greyscale conversion of blood smear images.

Training : In this YOLO frame work 3 classes has been defined they are RBC, WBC and Platelets for blood cell identification. Due to 3 classes, the filters in final convolution layer must be modified. So the number of filters can be calculated using equation

$$N_F = N_A \times (N_C + 5) \quad (1)$$

where N_A is anchor boxes i.e; 5 and N_C is number of classes i.e; 3 in the experiments, so N_F is found to be 40. The YOLO predicts 5 values, the values are the x and y co-ordinates, objects height and width, and probability value of having object in each grids.

During the training in every step, this method recorded loss and moving average loss, from the YOLO weights. Lower learning rates at later phases show higher convergence; two distinct learning rates — 10^-5 and 10^-7 were used. The proposed work has recorded the YOLO weights and model has been estimated, the learning curve for both learning rates has been calculated based on the loss function with help of vector values as shown in below figure 2.

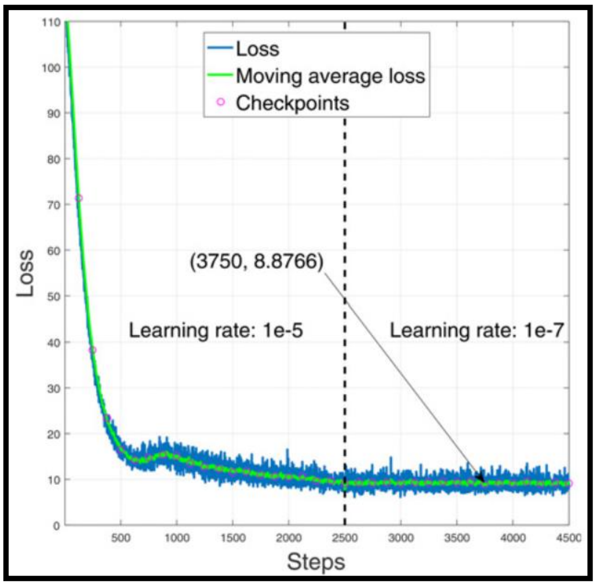

figure 2. learning curve

The proposed work doesn't misrepresent between the cells by classifying Platelet as either WBC or RBC or WBC as either RBC or Platelets. But in some rare case it was counting the same platelet cell twice, this issue has been resolved by applying Intersection over Union (IOU) and K-nearest neighbour (KNN) methods, this is shown in the figure 3.

```
# Loop over for all the cell predictions
for i in range(length of Predictions) do
    (x1, y1) = top left coordinate of bounding box
    (x2, y2) = bottomright coordinate of bounding box
    label = label of cell
    # Checking for spurious overlapping platelet prediction
    if lable == 'Platelets' then
        Find the nearest platelet using KNN
        Applying IOU to calculate the overlap between
detected platelet and nearest platelet
        # We are allowing only 10% overlap
        if overlap > 10% then
            continue
        end if
    end if
    # Cell counting
    if label == 'RBC' then rbc ← rbc + 1
    else if label == 'WBC' then wbc ← wbc + 1
    else if label == 'Platelets' then platelet ← platelet + 1
    end if
    center = int((x1 + y1)/2, (x2 + y2)/2)
    radius = int((x2 − x1)/2)
    Drawing circular bounding boxes with label at the center
    i ← i + 1
end for
```

figure 3. algorithm for KNN and IOU

From the implemented methodology 4 parameters has been detected in each cell, that are cell label, confidence value, and corner position of both top left and bottom right. Using the corner co-ordinates (x1,y1) and (x2,y2) respectively, using equation 2 and 3 C and r i.e; centre point and radius of the cell has been calculated.

$$C = \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2}\right) \qquad (2)$$

$$r = \left(\frac{x_2 - x_1}{2}\right) \qquad (3)$$

IV Experiments and Results

Our proposed work will automatically identify the RBC, WBC and Platelets, and the model has been tested for 180 images of blood smear. First, the model has been tested for validation data set of 60 blood smear images with different threshold, from that average absolute error has been calculated between the ground truth and estimated cells. With the different threshold values, error has been calculated using equation 4,

$$\varepsilon^{\text{cell}} = \frac{1}{N}\sum_{i=1}^{N} |x^{(i)}_{\text{groundtruths}} - x^{(i)}_{\text{estimated}}| \quad (4)$$

where cell is either RBC or WBC or platelets, N is validation dataset size i.e; of 60 images, x is the total count of cells. The obtained error values are shown in table 1 below.

Threshold,%	RBC	WBC	Platelets
20	5.650	0.083	0.217
25	4.417	0.050	**0.083**
30	3.450	0.033	0.083
35	2.750	**0.017**	0.083
40	2.500	0.050	0.083
45	2.183	0.100	0.100
50	2.133	0.150	0.100
55	**2.083**	0.200	0.117
60	2.100	0.33	0.150

table 1. average absolute error values at different threshold.

From the table 1 the appropriate threshold value for each cell types are noted
RBC - 55% (0.55).
WBC - 35% (0.35).
Platelets - 25% (0.25).
With the above threshold values the proposed work has achieved the accuracy of 96% for RBC, 86% for WBC and 96% for

	RBC	WBC	Platelets
Ground truths	857	61	55
Estimated	823	53	96
Accuracy	96%	86%	96%

Platelets,

table 2. accuracy for estimation of blood cell types.

The Table 3 shows the accuracy of detecting cells, the best accuracy for detecting RBC and platelets is obtained from YOLO architecture, by achieving 96% accuracy in detecting and counting RBCs and Platelets.

		RBC	WBC	Platelets	Execution Time, ms
	Ground truths	857	61	55	
YOLO	Estimated Accuracy	823 96%	53 86%	96 96%	60
VGG-16	Estimated Accuracy	1006 72.92%	61 100	60 90	106
ResNet50	Estimated Accuracy	952 79.8%	58 95.08%	48 87.27%	118
InceptionV3	Estimated Accuracy	889 87.75%	61 100	57 96.36%	130
MobileNet	Estimated Accuracy	588 74.24%	57 93.44	46 83.64	84

table 3. comparision of accuracy of detecting blood cell types using different cnn algorithm with yolo

The proposed work will identify the blood cell types on the basis of grid cells, that is shown in figure below

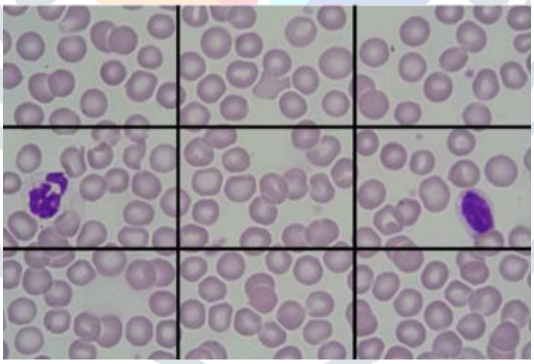

fig 4. image divide into 3*3 grids.

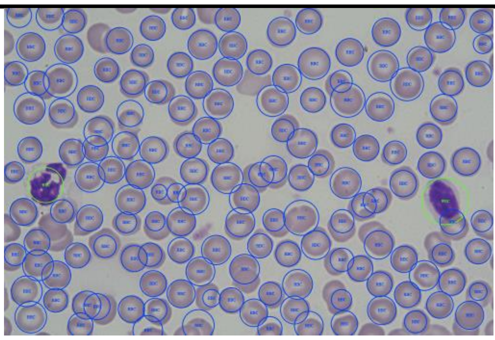

fig 5a. combined output of high resolution image

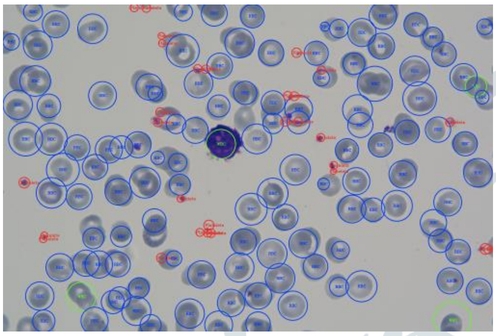

fig 5b. combined output of high resolution image

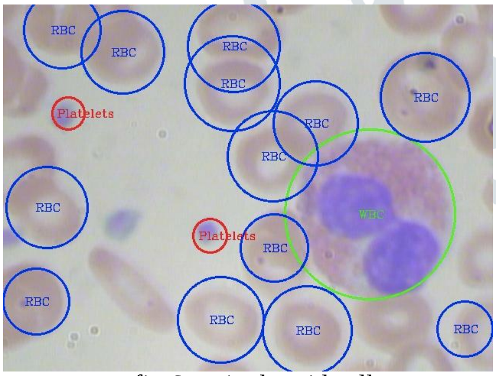

fig 6 a. single grid cell output

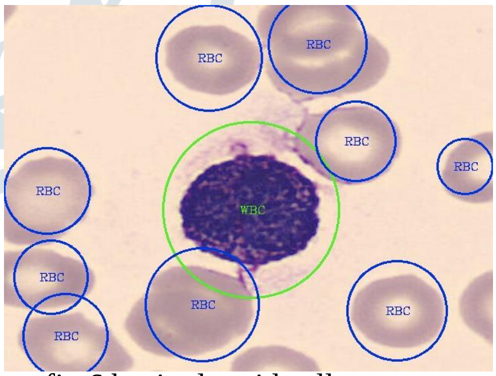

fig 6 b. single grid cell output

V Conclusion

In this proposed work of machine learning approach for blood smear analysis for identification of blood cell types is obtained from YOLO algorithm, and by implementing KNN and IOU based methods, the accuracy has been increased. This model is tested on publicly available dataset i.e; 360 blood smear images from Blood Cell Count Dataset, and 23,000 blood smear images from Kaggle, for this work train data of size 300 blood smear images of each cell type, and test data of size 180 blood smear images has been considered and the model is able to detect the blood cell types accurately at the threshold of 0.55 for RBC, 0.35 for WBC and 0.25 for platelets. And this model has been tested on both the data sets (BCCD and Kaggle) and it has competently preformed. With this performance of the proposed work, it can be said that, this method has potential to ease up manual process of blood smear analysis.

References:

[1] Dheeraj Mundhra, Bharath Cheluvaraju, Jaiprasad Rampure, and Tathagato Rai Dastidar, (2017) "Analyzing Microscopic Images of Peripheral Blood Smear using Deep Learning".

[2] Nicola Ritter, James Cooper : "Segmentation and Border Identification of Cells in Images of Peripheral Blood Smear Slides", this work was supported by a grant from the Division of Arts at Murdoch University.

[3] Hany A. Elsalamony (2016) "Healthy and unhealthy red blood cell detection in human blood smears using neural networks", Elsevier, Volume:32–41.

[4] Roopa B. Hegde , Keerthana Prasad , Harishchandra Hebbar , I. Sandhya (2018) "Peripheral blood smear analysis using image processing approach for diagnostic purposes ", Elsevier, BBE 257 1-14.

[5] Biji G, Assistant Professor, Dr.S.Hariharan, Professor (2017) "An Efficient Peripheral Blood Smear Image Analysis Technique For Leukemia Detection ", International conference on I-SMAC .

[6] Siraj Khan, Muhammad Sajjad, Tanveer Hussain, Amin Ullah (2020) "Traditional Machine Learning and Deep Learning Models for WBCs Classification in Blood Smear Images", IEEE Access.

[7] Aliyu Abubakar, Mohammed Ajuji, Ibrahim Usman Yahya (2021) "DeepFMD: Computational Analysis for Malaria Detection in Blood-Smear Images Using Deep-Learning Features", MDPI.

[8] Muhammad Imran Razzak, SaeedaNaz, "Microscopic Blood Smear Segmentation and Classification using Deep Contour Aware CNN and Extreme Machine Learning", CVPR workshop.

[9] Vasundhara Acharya, Vinayakumar Ravi, Tuan D Pham, Chinmay Chakraborty (2021) "Peripheral Blood Smear Analysis Using Automated Computer-Aided Diagnosis System to Identify Acute Myeloid Leukemia", Research Gate.

[10] Kokeb Dese, Hakkins Raj, Gelan Ayana, Tilahun Yemane, Janarthanan Krishnamoorthy, Wondimagegn Adissu (2021) "Accurate Machine-Learning-Based classification of Leukemia from Blood Smear Images ", Elsevier.

[11] Haneen T. Salah, Ibrahim N. Muhsen, Mohamed E. Salama, Tarek Owaidah, Shahrukh K. Hashmi (2019) "Machine learning applications in the diagnosis of leukemia: Current trends and future directions", ISLH, John Wiley and Sons.

[12] Siddhika Arunachalam (2020) "Applications of Machine Learning and Image Processing Techniques in the Detection of Leukemia ", International Journal of Scientific Research in Computer Science and Engineering, Vol.8, Issue.2, pp.77-82.[13] Mustafa Ghaderzadeh, Farkhondeh Asadi, Azamossadat Hosseini, Davood Bashash, Hassan Abolghasemi, Arash Roshanpour (2021) "Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images: A Systematic Review", Scientific Programming.

[14] Cruz D, Jennifer C, Valiente L.C. (2017) "Determination of blood components (WBCs, RBCs, and platelets) count in microscopic images using image processing and analysis", Environment and Management (HNICEM).

[15] Acharjee S, Chakrabartty S, Alam M.I. (2016) "A semiautomated approach using GUI for the detection of red blood cells", Conf. on Electrical, Electronics, and Optimization Techniques.

[16] Lou J, Zhou M, Li Q (2016) "An automatic red blood cell counting method based on spectral images ".

[17] Islam M.T., Aowal M.A., Minhaz A.T., "Abnormality detection and localization in chest X-rays using deep convolutional neural networks".