

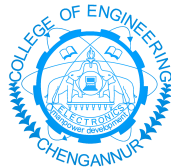
# Automatic identification and counting of blood cells using Deep learning

03CS6902 Mini Project

CHN20CSIP02 Karthika G Nath

karthikagnath619@gmail.com

M. Tech. Computer Science & Engineering (Image Processing)



Department of Computer Engineering

College of Engineering Chengannur

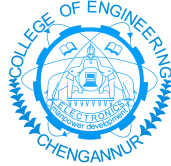
Alappuzha 689121

Phone: +91.479.2165706

<http://www.ceonline.edu>

[hod.cse@ceonline.edu](mailto:hod.cse@ceonline.edu)

College of Engineering Chengannur  
Department of Computer Engineering



C E R T I F I C A T E

This is to certify that, this report titled *Automatic identification and counting of blood cells using Deep learning* is a bonafide record of the work done by

**CHN20CSIP02 Karthika G Nath**

Second Semester M. Tech. Computer Science & Engineering (Image Processing)  
student, for the course work in **03CS6902 Mini Project**, under our guidance and supervision, in  
partial fulfillment of the requirements for the award of the degree, M. Tech. Computer Science &  
Engineering (Image Processing) of **APJ Abdul Kalam Technological University**.

Guide

Coordinator

Syama S  
Asst. Professor  
in Computer Engineering

Ahammed Siraj K K  
Associate Professor  
in Computer Engineering

Head of the Department

October 4, 2021

Dr. Smitha Dharan  
Professor  
in Computer Engineering

## Permission to Use

In presenting this mini project dissertation at College of Engineering Chengannur(CEC) in partial fulfillment of the requirements for a Postgraduate degree from APJ Abdul Kalam Technological University, I agree that the libraries of CEC may make it freely available for inspection through any form of media. I further agree that permission for copying of this dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the Head of the Department of Computer Engineering. It is understood that any copying or publication or use of this dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to CEC in any scholarly use which may be made of any material in this mini project dissertation.

---

Karthika G Nath.

## Statement of Authenticity

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at College of Engineering Chengannur(CEC) or any other educational institution, except where due acknowledgement is made in the report. Any contribution made to my work by others, with whom I have worked at CEC or elsewhere, is explicitly acknowledged in the report. I also declare that the intellectual content of this report is the product of my own work done as per the **Problem Statement** and **Proposed Solution** sections of the mini project dissertation report. I have explicitly stated the major references of my work. I have also listed all the documents referred, to the best of my knowledge.

Karthika G Nath.

## Acknowledgements

Primarily, I thank Lord Almighty for his eternal support through out my project work.

I express my sincere thanks to **Dr. Jacob Thomas V.**, Principal, College of Engineering Chengannur for extending all the facilities required for doing my project. My heartfelt words of gratitude to **Dr. Smitha Dharan**, Professor and Head of Department of Computer Engineering, for providing constant support.

Now I express my gratitude to my project co-ordinator **Mr. Ahammed Siraj K K**, Associate Professor in Computer Engineering and my project guide **Mrs. Syama S**, Assistant Professor in Computer Engineering who played a great role for valuable suggestions and expert guidance.

## **Abstract**

Blood testing is observed to be one of the most significant medical examination test to evaluate health condition. In pathology labs, different types of blood cells are counted to diagnose the diseases in patients including anemia, infection and leukemia. The blood constitutes mainly three types of cells such as red blood cells (RBCs), white blood cells (WBCs), and platelets. Traditionally blood cells are counted manually using haemocytometer along with other laboratory equipment's and chemical compounds, which is time intense, tedious, and entails lot of technical expertise. This work presents a deep learning approach for automatic identification and counting of three types of blood cells using 'you only look once' (YOLO) algorithm. YOLO is a state-of-the-art object detection classification algorithm. It requires only one forward propagation pass through the network to make a fast prediction for both image class and location. In this approach it automatically identify and count blood cells from a blood smear image using YOLO. To improve accuracy, the method employed KNN and IOU (Intersection of union) based method to remove multiple counting of the same object.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Proposed Project . . . . .	1
1.1.1	Problem Statement . . . . .	1
1.1.2	Proposed Solution . . . . .	1
<b>2</b>	<b>Report of Preparatory Work</b>	<b>3</b>
2.1	Literature Survey Report . . . . .	3
2.2	System Study Report . . . . .	4
<b>3</b>	<b>Project Design</b>	<b>6</b>
3.1	Project design . . . . .	6
3.2	Hardware & Software Requirements . . . . .	8
3.3	Anaconda . . . . .	8
<b>4</b>	<b>Implementation</b>	<b>9</b>
4.1	Preparation of blood smear image dataset . . . . .	9
4.2	Using pretrained yolo network by transfer learning . . . . .	10
4.3	Blood cell identification and counting method . . . . .	10
<b>5</b>	<b>Results &amp; Conclusions</b>	<b>11</b>
5.1	Conclusion . . . . .	12
	<b>References</b>	<b>13</b>

# Chapter 1

## Introduction

### 1.1 Proposed Project

Blood is the body fluid that delivers various substances such as nutrients and oxygen to cells and takes away metabolic waste from the cells. The blood accounts for 7–8% of total body weight. The human blood cells have three major components such as, white blood cells (WBCs), red blood cells (RBCs) and platelets. RBCs are the majority of the blood sample counts and are responsible for providing oxygen to the various vital organs of the body, determining blood type, and also carrying away the waste product. WBCs fight against infections and platelets help with blood clotting. In the medical field, the analysis of the blood sample of the patient is a critical task. A complete blood cell (CBC) count is an important test often requested by medical professionals to evaluate health condition. Blood cells are commonly counted using the Hemocytometer. To count blood cell, physician must view hemocytometer through a microscope and count blood cells using hand tally counter. This manual method of counting cells is tedious and time consuming and need a specialist to do this, but the automatic methods can overcome these problems.

Deep learning based object detection method can be used for automatic identification and counting of different blood cells. Deep learning, also known as deep structured learning is a class of machine learning algorithm based on artificial neural networks with representation learning. Deep learning model tries to transform its data into an abstract representation. It enables to extract the information from the layers present in its architecture[2].

#### 1.1.1 Problem Statement

This project aims at automatic identification and counting of blood cells through a deep learning approach using ‘you only look once’ (YOLO) object detection and classification algorithm.

#### 1.1.2 Proposed Solution

The proposed solution for the problem is to develop a deep learning approach for automatic identification and counting of three types of blood cells using ‘you only look once’ (YOLO) object detection and classification algorithm. YOLO detect all three types of blood cells, i.e, RBC, WBC, platelets simultaneously. It is a state-of-the-art object detection classification algorithm which requires only one forward propagation pass through the network to make a fast prediction. YOLO framework

is trained with a modified configuration dataset of annotated blood smear images to automatically identify and count red blood cells, white blood cells, and platelets. For training and testing of yolo model usually keras deep learning libraries are used. So, this can be modified using Tensorflow deep learning library which is more advanced and has increased functionality than keras and take less amount of time for execution.



## Chapter 2

# Report of Preparatory Work

With the development of machine learning techniques, image classification and object detection applications are becoming more robust and more accurate. Among the state-of-the-art object detection algorithms such as Region based convolutional neural network (R-CNN), you only look once(YOLO) object detection method is used here. YOLO framework is about three times faster than Faster R-CNN. Related works on different approaches in the automated counting process of blood cells are reviewed in this section.

### 2.1 Literature Survey Report

1. Acharya V., Kumar P, **Identification and red blood cell automated counting from blood smear images using computer-aided system**[3], Med. Biol. Eng. Comput., 2018

This work proposes an image processing technique for counting the number of RBC. It aims to examine and process the blood smear image, in order to support the counting of red blood cells and identify the number of normal and abnormal cells in the image automatically. Granulometric analysis is used to separate the red blood cells from the white blood cells. The red blood cells obtained are counted using the labeling algorithm and circular Hough transform. A computer-aided system is used to help to attain precise results. Here, the image is converted to gray scale further segmentation and morphological operations are performed to process the smear image. Counting of the cells are done using the labeling algorithm. The objects in the image are labeled using standard connectivity value. Then the algorithm constructs a matrix with same size as that of the binary image. The background pixels are labeled with a value of zero. Matrix will have non-zero values for the labeled foreground objects. A counter is set to count the number of pixels with label 1. The number of red blood cells was given by the counter value.

2. Sarrafzadeh O., Dehnavi A.M., Rabbani H., **Circlet based framework for red blood cells segmentation and counting**[4], IEEE Workshop on Signal Processing Systems, December 2015

This method introduces a novel method for RBCs segmentation and counting from microscopic images using Circlet Transform(CT) which operates directly on grayscale image.

Circular Hough Transform method is used for circular shape objects. CT decomposes an image into circles called circle with different radii and certain width. RCB detection using circle transform involves various methods such as Removing WBCs and background of image, Obtaining the mask of RBCs, Applying circle transform, Estimating minimum and maximum number of RBCs, Detection and Counting RBC using iterative soft thresholding and removing conflict RBCs.

3. Zhao J., Zhang M., Zhou Z, **Automatic detection and classification of leukocytes using convolutional neural networks**[5], Med. Biol. Eng. Comput., 2017

This method proposes an automatic detection and classification system for WBCs from peripheral blood images. It firstly detects WBCs from the microscope images based on the simple relationship of colors and the morphological operations. Then for each WBC, a granularity feature (pairwise rotation invariant co-occurrence local binary pattern) is extracted and SVM is applied on these granularity features to discern the eosinophil and basophil from other types of WBCs. CNN is used to extract remaining three types of WBCs i.e, neutrophil, monocyte and lymphocyte automatically.

4. Yan Li, Rui Zhu, Lei Mi, Yihui Cao, and Di Yao, **Segmentation of white blood cell from acute lymphoblastic leukemia images using dual-threshold method**[6], Computational and mathematical methods in medicine 2016.

This method proposes a dual-threshold method based on a strategic combination of RGB and HSV color space for white blood cell (WBC) segmentation. This method consists of three main parts: preprocessing, threshold segmentation, and postprocessing. In the preprocessing part, two images are obtained for further processing i.e, one contrast-stretched gray image and one H component image from transformed HSV color space. In the threshold segmentation part, a dual-threshold method is proposed for improving the conventional single-threshold approaches and a golden section search method is used for determining the optimal thresholds. For the postprocessing part, mathematical morphology and median filtering are utilized to denoise and remove incomplete WBCs.

5. Biswas, Soumen, and Dibyendu Ghoshal, **Blood cell detection using thresholding estimation based watershed transformation with Sobel filter in frequency domain**[7]. Procedia Computer Science 2016.

In this method a thresholding estimation algorithm has been proposed with watershed transforming Sobel filter in frequency domain for detection of different cells in microscopic image of blood. The proposed algorithm performs edge detection using Sobel filter in frequency domain.

## 2.2 System Study Report

Blood cell count plays an important role in the field of clinical medical diagnosis. It is used to evaluate overall health condition of persons. The accurate cell count helps in detecting potential diseases and related lesions, and determining the type and development of the

patient's diseases. There are three common types of blood cells are red blood cells (RBCs), white blood cells (WBCs), and platelets in the body. Blood cells are commonly counted using the Hemocytometer. It is a conventional method of placing the smear image under a microscope and the expertize will count the blood cells manually. It is a tedious and time consuming method.

To effectively automate the counting of blood cells, recently the deep-learning-based detection has been proposed and used successfully. The aim of this project is to develop a deep-learning approach for automatic identification and counting of blood cells using Yolo object detection algorithm. In this implementation three types blood cells such as RBC,WBC and platelets are detected and counted from annotated blood smear image using Yolo object detection method.

## Chapter 3

# Project Design

### 3.1 Project design

The main aim of the project is to use the object detection and classification algorithm YOLO to detect and count blood cells directly from smear image. In this method YOLO framework is trained with a modified configuration and annotated blood cells training images. It is used to detect all three types of blood cells simultaneously. The whole process is fully automated, fast, and accurate.

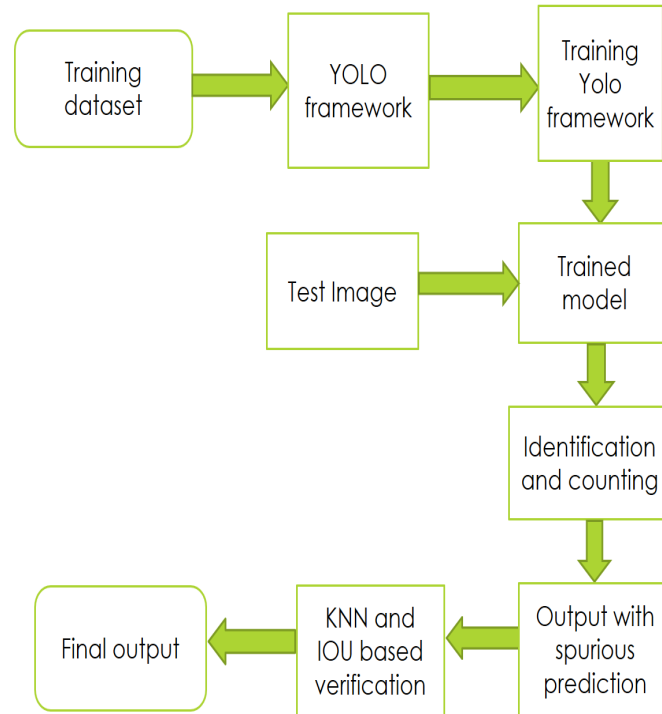


Figure 1: Block diagram of automatic blood cells identification and counting system

Inorder to develop a deep learning model for automatic counting and identification of blood cells, Yolo framework is trained using a dataset of annotated blood cell images. It includes Yolo training processes and operations. Overall the computer-aided system of detection and counting enables to count blood cells from smear images in faster rate.

**1. Yolo :-**

‘You Only Look Once’ in short YOLO is a state-of-the-art object detection classification algorithm. It treats object detection as a regression problem. It requires only one forward propagation pass through the network to make a fast prediction for both image class and location. Since, it only needs to look once at the image to detect all the objects, hence named ‘You only look once’. Yolo divides the input image into  $S \times S$  grid, where each grid cell predicts for bounding boxes and confidence score for the boxes. If the centre of an object falls into grid cell, then that grid cell is responsible for detecting that object. There are different versions of Yolo, here a tiny Yolo framework is used, as it is the fastest of all. Tiny YOLO uses 9 convolutional layers.

**2. Training :-**

Tiny YOLO configuration was trained for three classes consisting of WBC, RBC, and platelets for blood cell identification. YOLO predicts five values along with class probabilities for each anchor box. The values are the probability of having an object in a grid cell, x and y coordinates of the object, height, and width of the object. Here a training model is developed with a modified configuration where the final convolution layer for three outputs can be changed, identification of blood cells with an appropriate threshold, and count them from their labels.

**3. Identification and counting :-**

Four parameters are obtained from the YOLO model for each detected cell. They are the label of the cell, the confidence of being that cell, top left corner position, and bottom right corner position. The blood cells are counted using their label. The modified YOLO returns three kinds of labels ‘RBC’, ‘WBC’, and ‘Platelets’ depending on the detected cell. The total number of RBC in a smear image will be the total number of labels containing ‘RBC’, the total number of WBC will be the total number of labels containing ‘WBC’ and so on. In some cases, same platelet from two consecutive grid cells are detected as result platelets are counted twice. It can be resolved by using K-nearest neighbour (KNN) and intersection over union (IOU) in each platelet. Here, 10% of the overlap between platelet and its closest platelet is considered. If the overlap is larger than that, then ignore that cell as double count to get rid of spurious counting.

Algorithm for automatic blood cells identification and counting:

```

1: Import Test image
2: Import Trained weights
3: Predictions using object detection algorithm
4: Removing fake predictions using appropriate thresholds value
5: rbc = 0
6: wbc = 0
7: platelet = 0
8: # Loop over for all the cell predictions
9: for i in range (length of predictions) do
10:  (x1,y1) = top left coordinate of bounding box
11:  (x2,y2) = bottom right coordinate of bounding box
12:  label= label of cell
13:  # Checking for spurious overlapping platelet prediction
14:  if label == 'Platelets' then
15:    Find the nearest platelets using KNN
16:    Applying IOU to calculate the overlap between
      detected platelet and nearest platelet
17:    # allowing only 10%overlap
18:    if overlap > 10% then
19:      continue
20:    end if
21:  end if
22:  # Cell counting
23:  if label == 'RBC' then rbc ← rbc + 1
24:  else if label == 'WBC' then wbc ← wbc + 1
25:  else if label == 'Platelets' then platelet ← platelet + 1
26:  end if
27:  center = int ((x1 + y1)/2,(x2 + y2) / 2)
28:  radius = int ((x2-x1) / 2)
29:  Drawing circular bounding boxes label at the center
30:  i = i + 1
31: end for
32: Saving the image

```

## 3.2 Hardware & Software Requirements

Dataset	: Collected publicly available dataset.
Operating System	: 64-bit Operating System
Supporting software	: Anaconda Python 3.8
Processor	: Intel Core i5 9th Gen 2.40GHz
RAM	: 8GB
Monitor	: Any colour monitor

## 3.3 Anaconda

Anaconda is a free and open-source distribution of the programming languages Python and R. The distribution comes with the Python interpreter and various packages related to machine learning and data science.

## Chapter 4

# Implementation

The various stages of implementation involved in this project are:

- Prepare the blood smear image dataset.
- Using pretrained yolo network by transfer learning.
- Blood cell identification and counting method using test images.

### 4.1 Preparation of blood smear image dataset

To train the YOLO framework a publicly available dataset of annotated blood cell images called Blood Cell Count Dataset (BCCD) is used. For training 300 annotated blood smear images and for testing 60 images are used.

```
- <annotation>
  <folder>JPEGImages</folder>
  <filename>BloodImage_00000.jpg</filename>
  <path>/home/pi/detection_dataset/JPEGImages/BloodImage_00000.jpg</path>
  - <source>
    <database>Unknown</database>
  </source>
  - <size>
    <width>640</width>
    <height>480</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  - <object>
    <name>WBC</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    - <bndbox>
      <xmin>260</xmin>
      <ymin>177</ymin>
      <xmax>491</xmax>
      <ymax>376</ymax>
    </bndbox>
  </object>
  - <object>
    <name>RBC</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    - <bndbox>
      <xmin>78</xmin>
      <ymin>336</ymin>
      <xmax>184</xmax>
      <ymax>435</ymax>
    </bndbox>
  </object>
```

Figure 1: annotation of blood smear image

The above figure shows the annotation file which describes a file named BloodImage-00000, has a dimension of  $640 \times 480 \times 3$ . It has object tags which represent bounding boxes. The class is specified by name tag whereas the details of bounding boxes are represented by the 'bndbox'tag. The bounding box is described by the coordinates of its top-left(xmin,ymin) corner and its bottom-right(xmax,ymax) corner.

## 4.2 Using pretrained yolo network by transfer learning

For training the dataset for blood cell identification and counting, pretrained tiny yolo weights are used. To adapt the configuration for blood cell identification, three classes consisting of WBC, RBC, and platelets are used. YOLO predicts five values along with class probabilities for each anchor box. The values are the probability of having an object in a grid cell, x and y coordinates of the object, height, and width of the object. For using pretrained weights of yolo model, tensorflow deep learning libraries are used.

## 4.3 Blood cell identification and counting method

For blood cell identification and counting process, the trained weights and python packages such as openCV, Tensorflow are used. Yolo model for each detected cell gives four values. They are the label of the cell, the confidence of being that cell, top left corner position, and bottom right corner position. Since blood cells are close to circular in shape, thus circular bounding boxes are enclosed around each cells. It requires the conversion of the top left and bottom right coordinates to radius and centre of the circle.

Given the top left and bottom right coordinates are  $(x_1, y_1)$  and  $(x_2, y_2)$ , the centre point C and the radius r of the circle that enclosesthe cell can be calculated by,

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

$$r = (\frac{x_2 - x_1}{2})$$

The blood cells are counted using their label. The modified YOLO returns three kinds of labels 'RBC', 'WBC', and 'Platelets' depending on the detected cell. The total number of RBC in a smear image will be the total number of labels containing 'RBC', the total number of WBC will be the total number of labels containing 'WBC'and so on. In some cases, same platelet from two consecutive grid cells are detected as result platelets are counted twice . It can be resolved by using K-nearest neighbour (KNN) and intersection over union(IOUS) in each platelet. Here, 10% of the overlap between platelet and its closest platelet is considered. If the overlap is larger than that, then ignore that cell as double count to get rid of spurious counting.



## Chapter 5

# Results & Conclusions

In this method for automatic identification and counting of RBCs, WBCs, and platelets from blood smear image yolo algorithm is presented. For this, a publically available dataset(BCCD) of annotated blood smear images are used. The model is tested on a dataset of 60 images and observed that, it accurately identifies RBCs, WBCs, and Platelets. To avoid spurious counting of blood cells and to improve accurate detection, the method employed KNN and IOU(Intersection over union) based method. IOU can be computed as area of intersection divided over area of union of two bounding boxes.

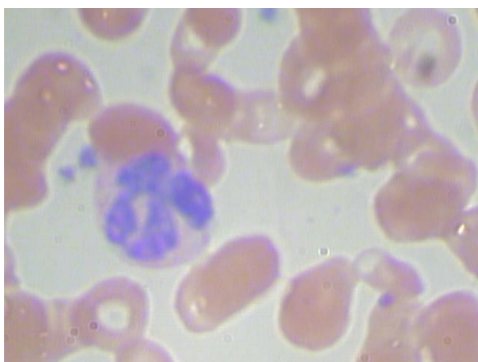


Figure 1: input blood smear image

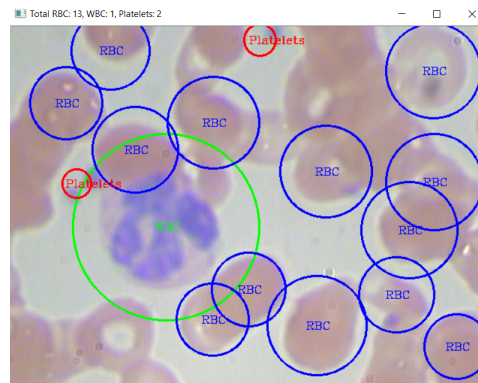


Figure 2: Blood cell detection output

## 5.1 Conclusion

Accurate cell counting is important in medical image analysis. In clinical applications, generally, various types of cells are manually counted, leading to a huge workload. With the Deep learning-based object detection method like YOLO, can automatically identify and count RBCs, WBCs, and platelets in a faster manner. It requires only one forward propagation pass through the network. Thus, the computer-aided system of detection and counting enables us to count blood cells from smear images in less than a second, which is useful for practical applications.

# References

- [1] Mohammad Mahmudul Alam and Mohammad Tariqul Islam : Machine learning approach of automatic identification and counting of blood cells. Healthcare Technology Letters, May 2019.
- [2] <https://en.wikipedia.org/wiki/Deep-learning>
- [3] Acharya V and Kumar P: Identification and red blood cell automated counting from blood smear images using computer-aided system. In Med. Biol. Eng Comput, August 2018.
- [4] Sarrafzadeh O., Dehnavi A.M., Rabbani H: Circlet based framework for red blood cells segmentation and counting, IEEE Workshop on Signal Processing Systems, December 2015
- [5] Zhao J., Zhang M., Zhou Z: Automatic detection and classification of leukocytes using convolutional neural networks, Med. Biol. Eng. Comput., 2017
- [6] Yan Li, Rui Zhu, Lei Mi, Yihui Cao, and Di Yao: Segmentation of white blood cell from acute lymphoblastic leukemia images using dual-threshold method, Computational and mathematical methods in medicine 2016
- [7] Biswas, Soumen, and Dibyendu Ghoshal: Blood cell detection using thresholding estimation based watershed transformation with Sobel filter in frequency domain. Procedia Computer Science 2016
- [8] <https://pjreddie.com/darknet/yolov2/>