Automatic identification and counting of blood cells using Deep learning

03CS6902 Mini Project
Design Report

CHN20CSIP02 Karthika G Nath
karthikagnath619@gmail.com
M. Tech. Computer Science (Image Processing)

Department of Computer Engineering
College of Engineering Chengannur
Alappuzha 689121
Phone: +91.479.2165706
http://www.ceconline.edu
hod.cse@ceconline.edu

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 diagonise 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

Chapter 1

Introduction

Blood is the body fluid that delivers various substances such as nutritions 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 [3].

1.1 Proposed Project

This project aims at developing a fully automated, fast, and accurate method for counting of blood cells through a deep learning approach using 'you only look once' (YOLO) object detection and classification algorithm.

1.1.1 Problem Statement

Traditionally the blood cells are counted manually by placing the blood smear under microscope which may sometimes lead to erroneous results.It is also tedious,time consuming and require a specialist to do this.Therefore,a fully automated method is proposed to overcome these problem using a deep learning approach.

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 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 blood smear images to automatically identify and count red blood cells, white blood cells, and platelets.

Chapter 2

Project Design

2.1 Project design

The main aim of the project is to 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.

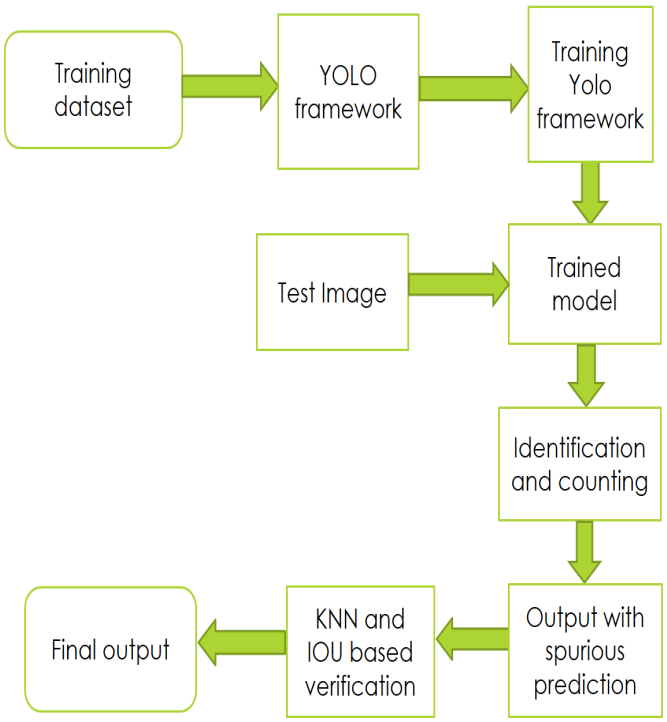

Figure 1: Block diagram of automatic blood cells identification and counting system

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,were 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 we choose to use Tiny YOLO 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 bethe 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:

```
Import Test image
Import Trained weights
Predictions using object detection algorithm
Removing fake predictions using appropriate thresholds value
rbc = 0
wbc = 0
platelet = 0
# 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) = bottom right coordinate of bounding box
    label= label of cell
    # Checking for spurious overlapping platelet prediction
    if label == 'Platelets' then
        Find the nearest platelets using KNN
        Applying IOU to calculate the overlap between
          detected platelet and nearest platelet
        # 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 == 'Patelets' then platelet ← platelet + 1
    end if
    center = int ((x1 + y1)/2,(x2 + y2) / 2)
    radius = int ((x2-x1) / 2)
    Drawing circular bounding boxes label at the center
    i = i + 1
end for
Saving the image
```

2.2 Hardware & Software Requirements

Operating System	: 64-bit Operating System
Supporting software	: Python
Processor	: Intel Core i5 9th Gen 2.40GHz
RAM	: 8GB
Monitor	: Any colour monitor

Chapter 3

Project Progress

Below are the work done so far:

1. Studied the reference paper.
2. Studied over the project area and problem domain
3. Made a literature survey over the topic
4. Studied about the main method used in the project.
5. Collected some data sets.
6. Made the design of the project.

3.1 Work Schedule

Below are the schedule of work(till August 10):
1.Identify suitable project domain and topic.
2.Study the reference paper in detail.
3.Prepare for IC and select a guide.
5.Conduct a literature survey on related works of the project topic.
6.Analys various methods used in this project.
7.Collect available datasets.
8.Make a design of the project.
9.Prepare for design presentation.

Work schedule for coming time period:
1.Start implementation.
2.Implement each module of the project.

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] 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.

[3] https://en.wikipedia.org/wiki/Deep-learning