Project Report for Project Machine Learning Algorithm for Cell Detection and Classification

Anirudh Bhupalarao [2018AAPS1242H]

Abstract :

To develop an image processing algorithm that calculates the number of white blood cells in an image. The white blood cells are generally circular in shape, small in size, and have grainy texture. If we find a way to differentiate between the different types of cells inside our urine it makes it much easier for doctors to monitor our health.This opens up unique opportunities for doctors to analyse the sample with much more accuracy and diagnose patients effectively. Even though the WBC cells in the images look very similar to the other cells present, the algorithm was able to detect the correct cells with good accuracy.

Approach Used :

As the project statement requires cell detection, the safest method that could be used is image processing. So, we use a method in OpenCV called HoughCircles Transform.

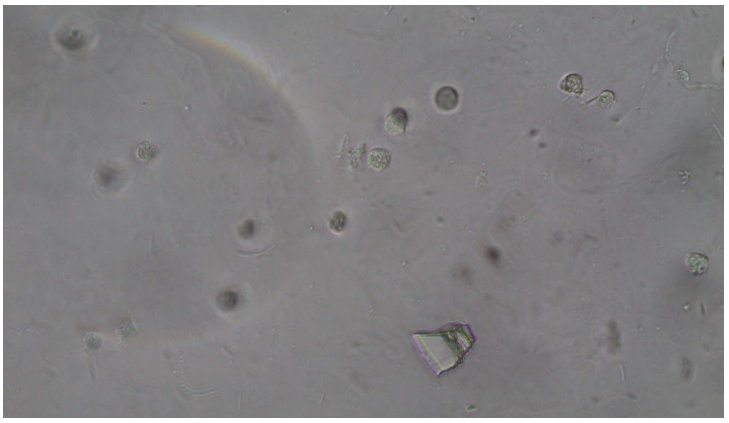
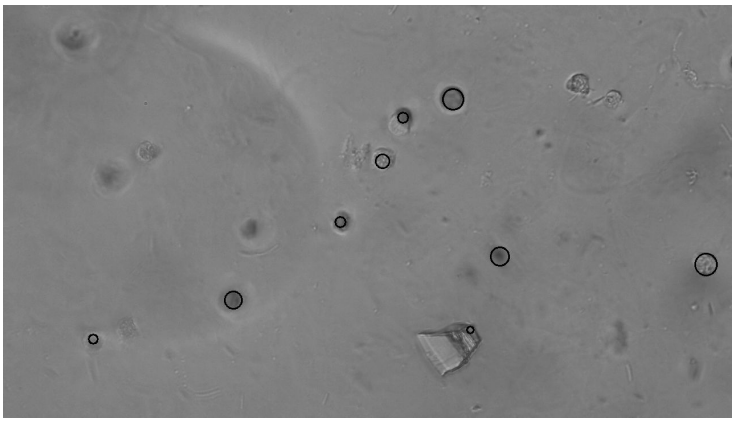
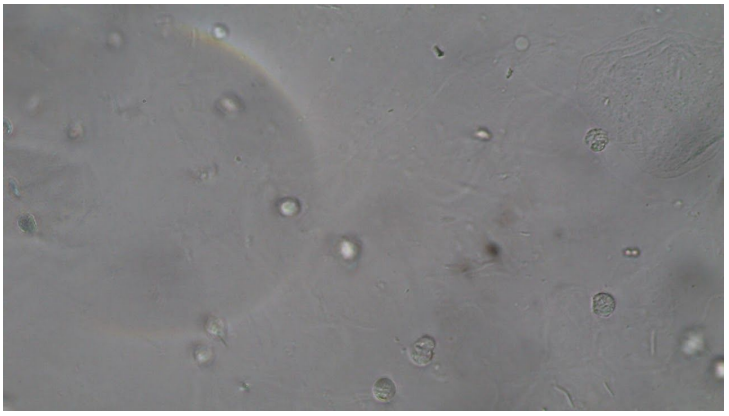
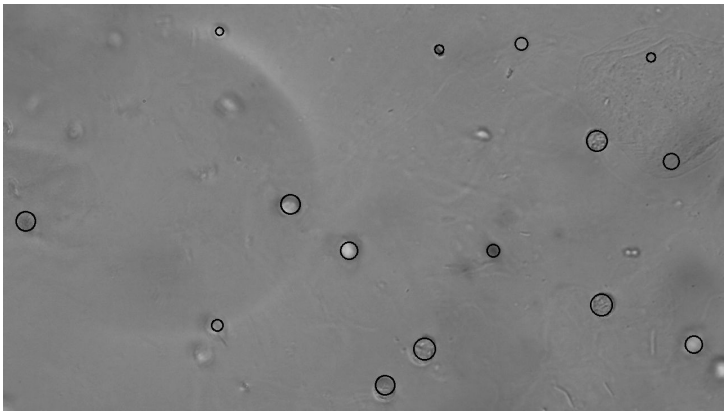
A circle is represented mathematically as $(x - x_{center})^2 + (y - y_{center})^2 = r^2$ where (x_{center}, y_{center}) is the center of the circle, and r is the radius of the circle. From the equation, we can see we have 3 parameters, so we need a 3D accumulator for hough transform, which would be highly ineffective. So OpenCV uses a trickier method, Hough Gradient Method which uses the gradient information of edges. The function we use here is cv2.HoughCircles().

There are multiple steps to this algorithm :-

(1) First initialize a matrix of dimensions rows*cols*maxRadius with zeros. (2) Preprocess the image into grayscale and apply blurring if necessary. (3) Loop through points and find values of r in the given ranges using a nested loop. (4) Pick the points in the accumulator matrix with the maximum value. These are strong points which indicate the existence of a circle with a, b and r parameters. This gives us the Hough space of circles. (5) Using the above circles as candidate circles, vote according to the image. The maximum voted circle in the accumulator matrix gives us the circle.

The HoughCircles() function has multiple parameters which we can vary according to our images to get appropriate results. These parameters are Detection Method (which is HoughGradient), dp, minDist, Param1, Param2, minRadius and maxRadius.

Results and Discussion :

As the above images show, the algorithm detects most of the white blood cells. However, it also detects extra objects that are close to circular. This is because of the parameters given in the HoughCircles method. The algorithm cannot detect the textures of the objects which can instead be done by Otsu's Method and findContours().

Conclusions and Recommendations :

If we fine tune the algorithm parameters according to the images we have, we can detect the WBCs perfectly with fewer errors. With more time we can implement thresholding to the images and use methods like Otsu's Method or other variance based methods to detect contours and textures so we can differentiate the cells properly.
We could also use Deep Learning algorithms like YOLO (You Only Look Once) for detection of the cells. YOLO is a DL model that utilizes Convolutional Neural Networks to train itself on a dataset of 1000s of images. Moreover, this approach will have less than 5% error as it is a very powerful algorithm and is used a lot in computer vision.

References :

[1] Kanade Takeo, Yin Zhaozheng, Bise Ryoma, "Cell Image Analysis: Algorithms, System and Applications", Carnegie Mellon University, Pittsburgh.
[2] Shitong Wang, Min Wang, "A New Algorithm NDA Based On Fuzzy Cellular Neural Networks for White Blood Cell Detection", IEEE Transactions on IT in Biomedicine, 2006.
[3] http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html#java