

# White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models

Nurasyeera Rohaziat<sup>1</sup>, Mohd Razali Md Tomari<sup>2</sup>, Wan Nurshazwani Wan Zakaria<sup>3</sup>, Nurmiza Othman<sup>4</sup>

Faculty of Electrical and Electronic Engineering  
Universiti Tun Hussein Onn Malaysia  
Malaysia

**Abstract**—There are several types of blood cancer. One of them is Leukaemia. This is due to leukocyte or white blood cell (WBCs) production problem in the bone marrow. Detection at earlier stage is important so that the patient is able to get a proper treatment. The conventional detection and blood count method is less efficient and it is done manually by pathologist. Thus, there will be a long line to wait for the results and also delay the treatment. A faster detection procedure and technique will have high impact on the real time diagnostic. Fortunately, these problems are able to overcome by making the blood test procedures automatic. One of the effort is the development of deep learning for WBCs detection and classification. In computer aided WBCs detection, the You Only Look Once (YOLO) based platform present a promising outcome. However, the investigation of optimal YOLO structure remains vague. This paper investigate the effect of the deep learning based WBCs detection using You Only Look Once version 3 (YOLOv3) with different pretrained Convolutional Neural Network (CNN) model. The models that been tested are the Alexnet, Visual Geometry Group 16 (VGG16), Darknet19 and the existing YOLOv3 feature extraction model, the Darknet53. The architecture consist of the bounding box for class prediction, feature extraction, and additional convolutional layers. It was trained with 242 WBCs images from Local Initiatives Support Corporation (LISC) dataset. The final outcome shows that the YOLOv3 architecture with Alexnet as its feature extractor produced the highest mean average precision of 98% and have better performance than the other models.

**Keywords**—Alexnet; darknet19; darknet53; detection; VGG16; white blood cells; YOLO

## I. INTRODUCTION

The human body immunization system depends on the white blood cells condition. Normal WBCs consist of Basophil, Eosinophil, Lymphocyte, Monocyte and Neutrophil. Immune system will be affected if there are any abnormality detected in the WBCs. Blood cancer such as leukemia is most common WBCs abnormality. Patient with this type of cancerous cell will have problem to fight virus and bacteria in the body system and weaken the immune system [1] [2]. It also affect the production of red blood cells and platelets in the bone marrow. There are several type of leukemia. An Acute Lymphoblastic Leukemia (ALL) is the one of it, which have abnormal lymphocyte which called leukocytes. This condition is crucial for early detection in order to establish a proper treatment for the patient. A typical diagnose method is the WBCs blood count which provide the data for the immune system and any blood related disease [3]. The conventional

method for leukemia diagnose are by bone marrow biopsy, lymph node biopsy, flow cytometry, lumbar puncture, lab test and also image tests, which is very challenging. However, the current automated system is depending on the application image processing, segmentation, feature extraction and finally the classification steps. Unfortunately, this method have an optimization issue [4]. A simple automated blood count method have been developed by using Convolutional Neural Network (CNN) architecture. The microscopic images are fed directly into the architecture for the classification and produce an output result [5]. On top of that, the neural network (NN) method have been evolve over the years and become the basic of a faster detection method.

The YOLOv3 detection method which been implemented in this project utilized the fundamental of neural network. This detection method practice a deep learning method for localization step which by using bounding box prediction, instead of the common sliding window search method [6]. The training process uses the sum of squared error loss and the logistic regression analysis for the objects' score prediction. The score will become 1 if the bounding box is overlap with the ground truth prior than another bounding box. Next, for the prediction of the classes, the independent logistic classifier and the cross-entropy loss are implemented. Then, as for the step prediction, three scale sizes are used. Meanwhile, for the feature extraction, Darknet-53 is used and followed by a few layers of convolutional layers. Lastly it produced a labeled output image.

This project investigated the implementation of different pretrained models (Alexnet, VGG16, Darknet-19) as the YOLOv3 feature extractor. LISC dataset images were used during the training and testing of the system. The finding of this research is the effect of different feature extractor on the detection rate and the detection average precision. The Alexnet as the feature extractor showed the highest detection mean average precision. Thus, this model can improve the existing blood smear detection method.

## II. YOU ONLY LOOK ONCE (YOLO)

Initially, the You Only Look Once (YOLO) was first introduced by J. Redmond et al. [7]. The YOLO detection system is illustrated in Fig. 1. In essence, the system will first resize the image, then undergo the convolutional network, and lastly the non-max suppression layer. The outcome is the detected labeled image. In the paper, the object detection is shown as a regression problem as the spatially separated the

bounding boxes and the associated class probabilities. From a full image, a single neural network is able to predict both the bounding boxes and the class probabilities with one evaluation. The speed of the model is at 45 frames per second. On the other hand, the Fast YOLO (smaller version) has higher process speed which is at 155 frames per second and had attained higher mean average precision in comparison with the other real time detectors.

The convolutional layers network of the YOLO detection is shown in the Fig. 2. It consist of 24 convolutional layers, two fully connected layers and interchanging 1x1 convolutional layers for feature space reduction from the preceding layers.

The convolutional layers of the YOLO detection network are illustrated in Fig. 2. It consist of 24 convolutional layers, two fully connected layers with 1x1 convolutional layers

alternately. This is to reduce the feature space from preceding layers. The overall YOLO model is trained together with the loss function that directly resembled the performance of the detection.

The authors presented another paper in the same year with YOLOv2 which is the upgraded from the previous YOLO model [8]. It is also named YOLO9000 due to its ability to detect 9000 or more objects' categories in real-time. The Darknet-19 was implemented in the architecture as the classifier. It consist of 19 convolutional layers and five maximum pooling layers. The model is able to process different image sizes while performing a balance of speed and accuracy. After that, the author updated the YOLOv2 to YOLOv3 [9]. In this new version, the Darknet19 had been upgrade to Darknet53 as the feature extractor. Its improvement include the extractor shortcut connection, feature map upsampling and concatenation.

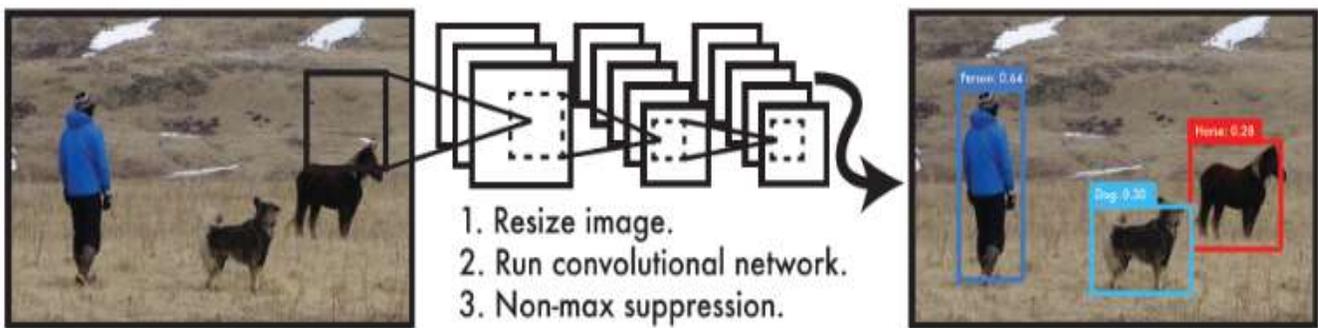


Fig. 1. YOLO Detection System [7].

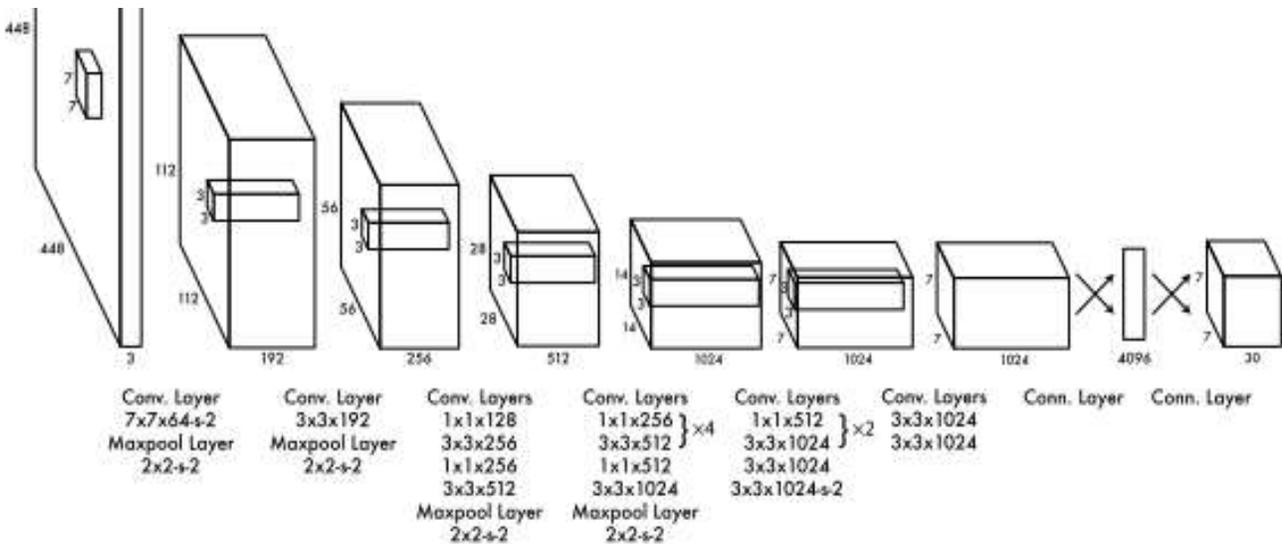


Fig. 2. YOLO Architecture [7].

#### IV. PRE-TRAINED MODELS

There are many types of pretrained model that been implemented in any machine learning architecture. Models like Alexnet, VGG16, and Darknet that had been trained with Imagenet and the weight is kept and ready to use for future implementation. Most of these models are CNN based.

Alexnet consist of five convolutional layers (only layer 1, 2 and 5 are followed by the max pooling) and three full connected layers [10]. All the inner layers applied Rectified Linear Units (ReLu) function as it activation function and softmax activation function in the final output layer. The application of the ReLu activation function speed up the training time up to six times on the Canadian Institute for Advanced Research 10 (CIFAR-10) dataset, instead of using the tanh activation function. Other than that, Alexnet allowed a multi graphics processing unit (GPU) training. Thus, it able to train bigger model and also shorten the training time. The conventional CNN pooling process will pool output from its neighbor groups of neurons without overlapping. Nevertheless, overlapping pooling was introduced in Alexnet which shows 0.5% loss reduction and the model less likely to overfit. The summary of the Alexnet architecture is tabulated as in Table I and the architecture diagram is as in Fig. 3.

The VGG16 model has more layers compare to the Alexnet structure. VGG16 architecture is illustrated in the Fig. 4. It was introduced in 2014 and to be an improvement of the Alexnet [11]. The large kernel-sized filters in Alexnet is

replace with a multiple 3x3 kernel-sized filters. The model had achieved 92.7% test accuracy in ImageNet. The structure has 16 layers depth and the VGG16 model is summarized in Table II. The input of the first convolutional layers is 224x224 RGB image and passed through the layers of convolutional layers then it applied maximum maximum pooling layer together with the 3x3 sized kernel filter. This produced a smaller image with dimension of 112x112x64. Then, it followed by two more convolutional layer, 3x3 sized 128 feature maps. Next, the maximum pooling with same size. Convolutional layers with 3x3 sized filter with 256 feature maps followed by maximum pooling are in the fifth and sixth layers. In the seventh to twelfth layer, there are two groups of three 3x3 sized 512 filters convolutional layers and maximum pooling layer. The last reduced size output is 7x7x512. Then, the output from the convolutional layer is flatten through the fully connected layers. Lastly, the layer of softmax function. All the hidden layers consist of the ReLu activation function.

Another CNN base model is the Darknet-19. It is initially used in YOLOv2 [8]. The model used filters and after each pooling step, there will be a couple of channel. The global average pooling is used for making the prediction and the filters for feature representation compression between convolutions. In order to stabilize the training, increase the convergence timing, and to make the model batch regulated, the paper used Batch Normalization technique. The Darknet19 consist of 19 convolutional layers and five maximum pooling layers. The summary of the model is tabulated in Table III.

TABLE I. SUMMARY OF ALEXNET ARCHITECTURE

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	227x227x3	-	-	-
1	Convolution	96	55x55x96	11x11	4	Relu
	Max Pooling	96	27x27x96	3x3	2	Relu
2	Convolution	256	27x27x96	5x5	1	Relu
	Max Pooling	256	13x13x256	3x3	2	Relu
3	Convolution	384	13x13x384	3x3	1	Relu
4	Convolution	384	13x13x384	3x3	1	Relu
5	Convolution	256	13x13x256	3x3	1	Relu
	Max Pooling	256	6x6x256	3x3	2	Relu
6	FC	-	9216	-	-	Relu
7	FC	-	4096	-	-	Relu
8	FC	-	4096	-	-	Relu
Output	FC	-	1000	-	-	Softmax

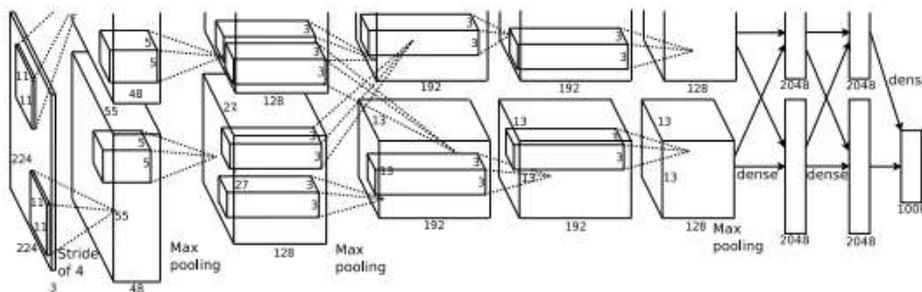


Fig. 3. Alexnet Architecture.

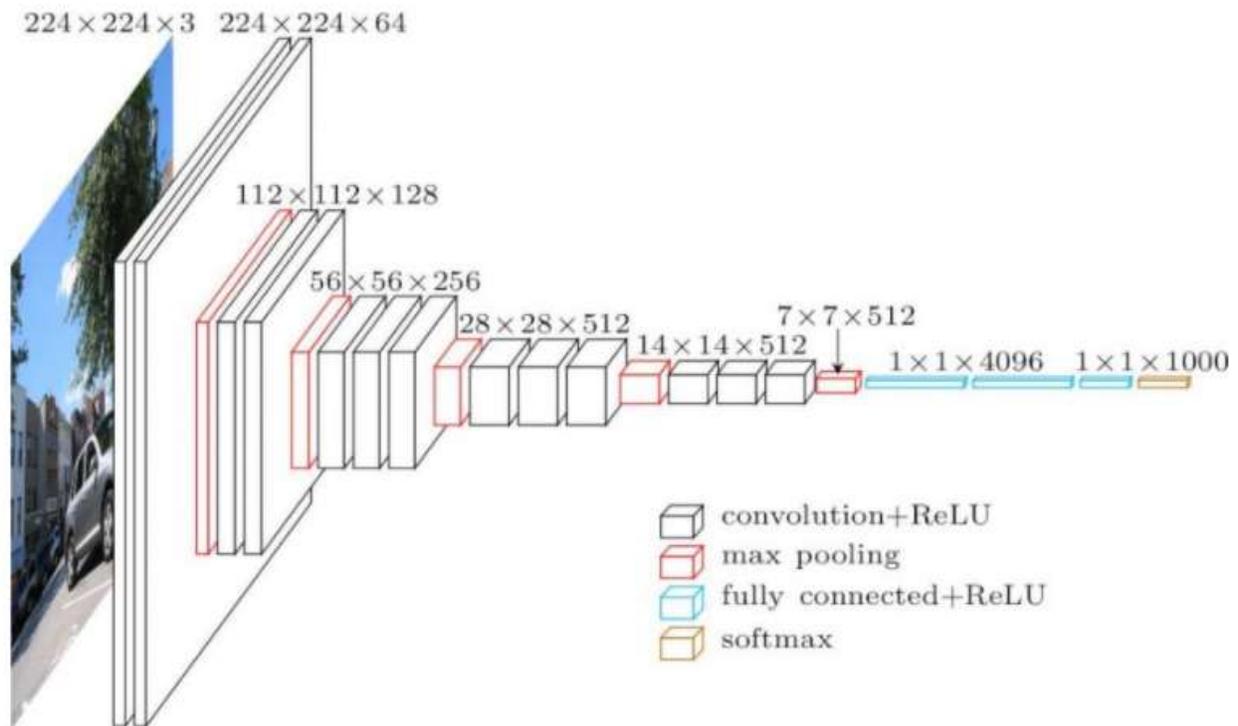


Fig. 4. VGG16 Architecture.

TABLE II. SUMMARY OF VGG16 ARCHITECTURE

Layer	Feature Map	Size	Kernel Size	Stride	Activation	
Input	Image	1	224x224x3	-	-	
1	2 x Convolution	64	224x224x64	3x3	1	Relu
	Max Pooling	64	112x112x64	3x3	2	Relu
3	2 x Convolution	128	112x112x128	3x3	1	Relu
	Max Pooling	128	56x56x128	3x3	2	Relu
5	2 x Convolution	256	56x56x256	3x3	1	Relu
	Max Pooling	256	28x28x256	3x3	2	Relu
7	3 x Convolution	512	28x28x512	3x3	1	Relu
	Max Pooling	512	14x14x512	3x3	2	Relu
10	3 x Convolution	512	14x14x512	3x3	1	Relu
	Max Pooling	512	7x7x512	3x3	2	Relu
13	FC	-	25088	-	-	Relu
14	FC	-	4096	-	-	Relu
15	FC	-	4096	-	-	Relu
Output	FC	-	1000	-	-	Softmax

TABLE III. SUMMARY OF DARKNET19 ARCHITECTURE

Type	Filter	Size / Stride	Output
Convolutional	32	3x3	224x224
Max Pooling		2x2/2	112x112
Convolutional	64	3x3	112x112
Max Pooling		2x2/2	56x56
Convolutional	128	3x3	56x56
Convolutional	64	1x1	56x56
Convolutional	128	3x3	56x56
Max Pooling		2x2/2	28x28
Convolutional	256	3x3	28x28
Convolutional	128	1x1	28x28
Convolutional	256	3x3	28x28
Max Pooling		2x2/2	14x14
Convolutional	512	3x3	14x14
Convolutional	256	1x1	14x14
Convolutional	512	3x3	14x14
Convolutional	256	1x1	14x14
Convolutional	512	3x3	14x14
Max Pooling		2x2/2	7x7
Convolutional	1024	3x3	7x7
Convolutional	512	1x1	7x7
Convolutional	1024	3x3	7x7
Convolutional	512	1x1	7x7
Convolutional	1024	3x3	7x7
Convolutional	1000	1x1	7x7
Average pool	-	Global	1000
Softmax	-	-	-

## V. RELATED WORKS

The detection and classification method of the WBCs has been studied widely in medical and also engineering field. One of the study presented a Regional Convolutional Neural Network (RCNN) which was trained by transfer learning using Alexnet, VGG16, Googlenet and Resnet50 [12]. The Resnet50 transfer learning shows the highest performance.

The author also have tested another CNN based architecture that tested on different pretrained models as its feature extractor with the Extreme Learning Machine (ELM) classifier [13]. The output from these feature extractors were then combined and the minimum redundancy maximum relevance method was used to select the efficient features. Finally, the ELM was enabled. The results shows accuracy rate of 96.03%. There are three studies that had been done. The Alexnet-ELM method which used ELM as classifier for the features at the fully connected layers of each CNN models. It obtained an accuracy rate of 95.29%. Then, the performance of classifier were tested and the Resnet model achieved 95.2% accuracy rate. Lastly, the paper studied the CNN - Minimum Redundancy Maximum Relevance (MRMR) - ELM method

on the WBCs data. MRMR feature selection algorithm was used for features combination at the last layers of the tested models. This method achieved the accuracy rate of 96.03%.

A project also had demonstrated that Alexnet has the best performance as feature extraction for WBCs type classification in comparison with Lenet, and VGG16 architectures [14]. The Discrete Transform (DT), quadratic discriminant analysis (QDA), linear discriminant analysis (LDA), Support vector machine (SVM), k- nearest neighbors (kNN) with Alexnet also been compared to a softmax classifiers and the highest accuracy is the combination of QDA-Alexnet which is 97.78%.

Additional application other than WBCs detection, Alexnet also been implemented in the detection and classification of the red blood cell (RBCs) [15]. The designed framework was able to classify 15 types of RBCs. The results obtained were: 95.92% accuracy, 77% sensitivity, 98.82% specificity, and 90% precision.

A different comparison had been made between the VGG16 and Resnet50 for WBCs classification [16] and the Resnet50 achieved 88.29% of accuracy. One of the project that utilized VGG16 is by M. Shahzad [17]. The framework starts with feeding the original images and ground truth images to the preprocessing stage. This include labeling of the pixel-level and conversion of RGB-Grayscale. The VGG16 later fed into the system as a feature extractor. Then the training process begin. The system accuracies are 97.45% for RBCs, 93.34% for WBCs, and 85.11% for platelets. Meanwhile, another paper also had compared the utilization of CNN models and Alexnet had perform better than GoogleNet and Resnet-101 [18]. A paper had implement image processing algorithm based for preprocessing and together with VGG16 as its classifier [19]. The experiment achieved 95.89% accuracy. In addition, a paper had use the concept of capsule for the classification model of WBCs [20]. The developed model proved that it had higher precision value than using the Resnet and VGG model.

A paper by K. Almezghwi also apply the VGG16, Resnet and Densenet in the project. The paper studied the generative adversarial networks (GAN) and image transformation operation for data augmentation together with the deep neural networks for feature extraction [21]. The outcome of the experiments is that the highest accuracy was achieved by using the Densenet-169 as the feature extractor which is 98.8%.

## VI. METHODOLOGY

The project starts with preparing the hardware, software and datasets. The hardware used for this experiment is the Intel® Core™ i5-5200U CPU @ 2.20GHz processor. Whereas the Spyder by Anaconda software is used for all the programming activity include training, detecting and result analysis.

### A. Datasets

This project worked with the dataset from LISC which is a public dataset. There are five image categories which are Basophil, Eosinophil, Lymphocytes, Monocytes and the

Neutrophil. Images are parted into training and testing as in Table IV. The number of training images is as 53, 39, 52, 48, and 50, respectively. Whilst the number of testing images is five for each category and eight images containing multiple WBCs type. The preparation of the datasets starts with annotating the images. The labelling process is done in the Microsoft’s Visual Object Tagging Tool (VoTT) software. After completing the preparation of datasets, the experiment continue with the training and detecting process.

**B. YOLOv3**

The main structure of this project is the YOLOv3 image detection. It starts with the forming of bounding box. Next, followed by the class prediction, prediction across scales, feature extraction, and the convolutional layer. Based on the YOLOv3 architecture in Fig. 5, it used a multi-scale method for the multiple target detection. The finer grid cell enable for smaller target detection. In every grid, three bounding boxes are predicted. Then, it will predict the three categories and five basic parameters which are x, y, w, h, and c.

Then, the multilabel classification will predict the classes in each bounding box. The classes will be predicted by using the binary cross-entropy loss during the training. Three sizes scales will then go through feature extraction. In addition, the Darknet-53 model will supported by additional convolutional layers. The feature map from the previous two layers were upsampled by double and concatenation the feature map from the prior layer in the network which produced a semantic and finer grained data.

Fig. 6 shows the flow process of the YOLOv3. The dataset was split into training set and testing set. The training set was

then undergo a bounding box step, class prediction and the prediction across scales. Later, the features were extracted by the Darknet53 model before it pass through convolutional layer and produced output. The steps were repeated by replacing the extractor with different CNN models. This project tested on the Alexnet, VGG16 and Darknet19 model which shows in Fig. 7.

**C. Evaluation of the Classifier Test Data**

The performance of the architectures was determined based on its accuracy, precision and sensitivity ratios as in equation (1), (2), and (3), respectively.

$$Accuracy (Acc.) = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

$$Precision (Pre.) = \frac{TP}{TP+FP} \tag{2}$$

$$Sensitivity \text{ or } Recall (Rec.) = \frac{TP}{TP+FN} \tag{3}$$

TABLE IV. NUMBER OF IMAGE DATASET IN EACH WBCS CATEGORY

Images	Training	Testing
Basophil	53	5
Eosinophil	39	5
Lymphocytes	52	5
Monocytes	48	5
Neutrophil	50	5
Mixed	-	8
Total	242	33

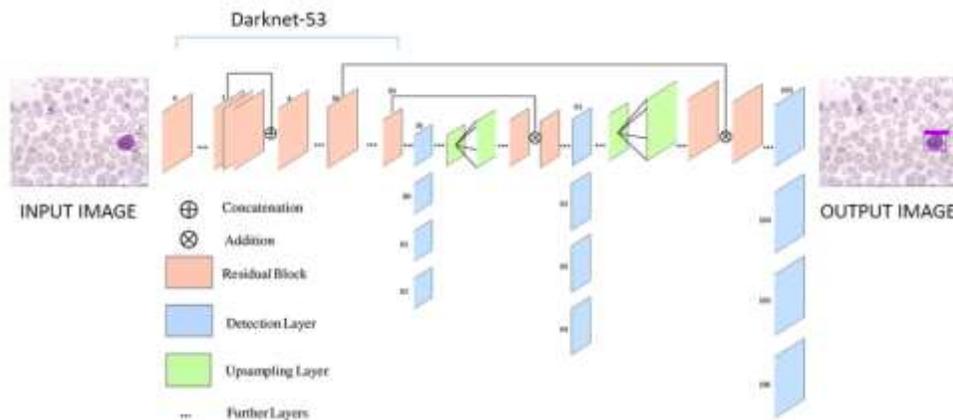


Fig. 5. YOLOv3 Architecture for WBCs Detection.

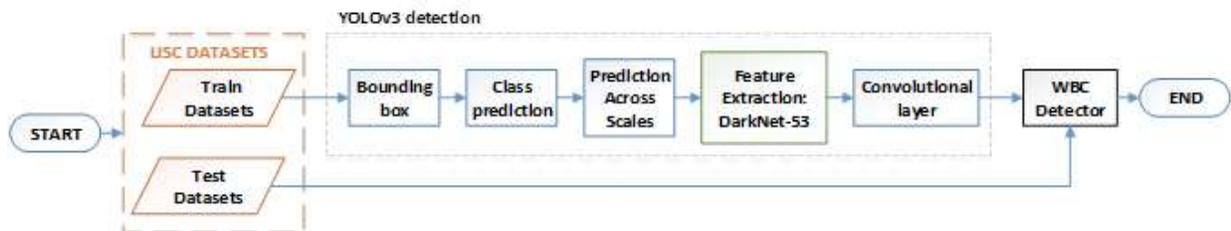


Fig. 6. YOLOv3 Process Flow.

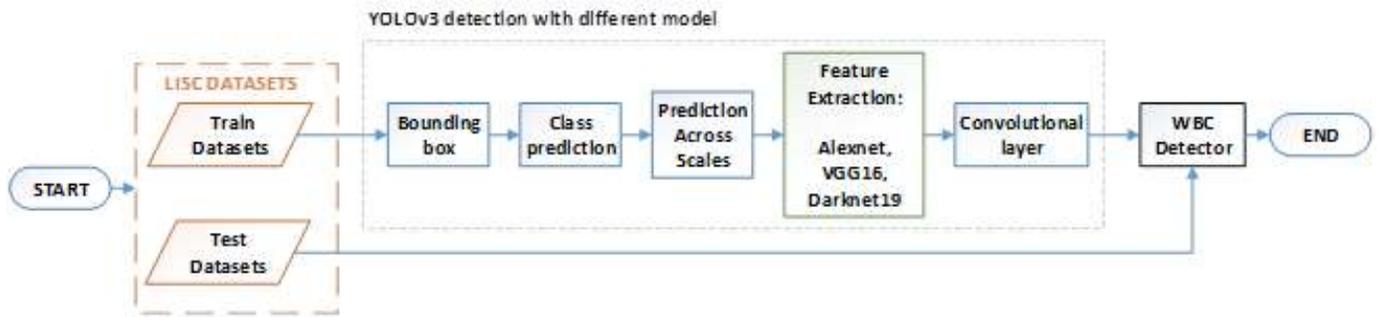


Fig. 7. YOLOv3 with different Model Process Flow.

The abbreviations are define as: TP = true positive, TN = true negative, FP = false positive, FN = false negative.

Then, the mean average precision (mAP) value is obtained by using (4).

$$AP = \int_0^1 Pre.(Rec.) dRec. \quad (4)$$

### VII. RESULTS AND DISCUSSION

The Table V presented the results of the training loss, validation loss, testing time, detection rate, and the detection mean average precision for each model used. The training loss and validation loss for all for models were between 13.8492 to 14.9078 and 13.4578 to 14.3617 respectively. Alexnet reached the lowest number of training loss, validation loss and also the

testing time. Based on the tabulated data, Alexnet has the highest detection mean average precision, 98%. However, the VGG-16 achieved the lowest, 96.61%.

The summary of mean average precision of each WBCs type is presented in Table VI. All four model had detected Basophil and Neutrophil with highest mAP. Meanwhile, only Darknet-19 achieved 100% of mAP. The lowest value of mAP, was the detection Monocyte cell by VGG16, which only 75%. The visual characteristic of the basophil and neutrophil are clearly able to differentiate compared to the visual characteristic of the monocyte. The output images were presented in the Table VII. The images were labelled with predicted class.

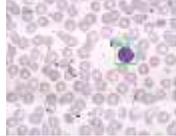
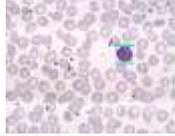
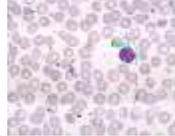
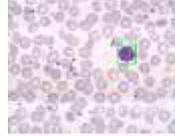
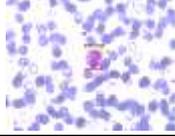
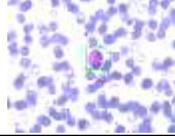
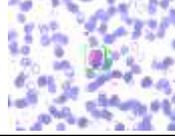
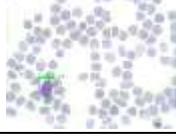
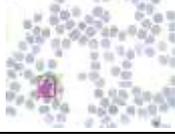
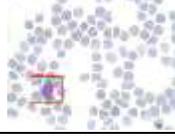
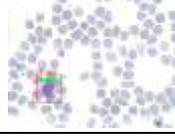
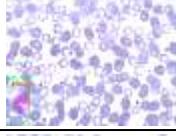
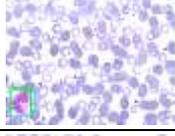
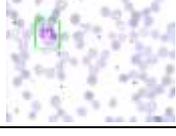
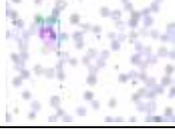
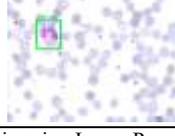
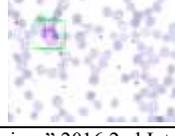
TABLE V. TRAINING AND DETECTION RESULTS

Model	Training Loss	Validation Loss	Testing time	Detection rate	Detection Mean Average Precision (mAP / %)
AlexNet	13.8492	13.4578	74.4sec	0.4FPS	98%
VGG-16	14.9078	13.9120	103.1sec	0.3FPS	93.06%
Darknet-19	14.5619	14.2912	105.9sec	0.3FPS	96.61%
Darknet-53	14.0242	14.3617	78.5sec	0.4FPS	96.19%

TABLE VI. THE MEAN AVERAGE PRECISION VALUE OF EACH CELL TYPE IN EVERY TEST

Cell Type	Model			
	AlexNet	VGG-16	Darknet-19	Darknet-53
Basophil	1	1	1	1
Eosinophil	1	0.9	1	0.94
Lymphocyte	1	1	0.83	1
Monocyte	0.9	0.75	1	0.87
Neutrophil	1	1	1	1
<b>Mean Average Precision (mAP / %)</b>	98%	93.06%	96.61%	96.19%

TABLE VII. IMAGES OUTCOME

Images	Model			
	AlexNet	VGG-16	Darknet-19	Darknet-53
Basophil12				
Eosinophil22				
Lymphocyte12				
Monocyte38				
Neutrophil35				

### VIII. CONCLUSION

The YOLOv3 architecture originally implement the Darknet53 as its feature extractor. This paper investigated the effect of using different feature extractor (Alexnet, VGG16, and Darknet19). The outcome of the experiment shows that the YOLOv3 with Alexnet as feature extractor obtained the highest mean average precision which was 98%. It also have the lowest losses and shortest testing time. As a result, it demonstrated to have higher performance compared to other models. On the other hand, more improvement is able to apply as a continuous project, such as to train on greater number of dataset, testing on several different source of datasets and also to use GPU to increase the training speed.

### ACKNOWLEDGMENT

This research is supported by a Fundamental Research Grant Scheme (FRGS) FRGS/1/2019/TK04/UTHM/02/6 (Vot. No. K191) sponsored by the Ministry of Higher Education (MOHE) Malaysia.

### REFERENCES

- [1] B. J. Bain, "The Nature of Leukaemia, Cytology, Cytochemistry and the FAB Classification of Acute Leukaemia," in *Leukaemia Diagnosis*, Fifth., B. J. Bain, Ed. John Wiley & Sons Ltd, 2011, pp. 1–63.
- [2] A. Vora, *Childhood Acute Lymphoblastic Leukemia*. Springer, 2017.
- [3] B. Martin, "Leukemia," in *Harmful Interaction between the Living and the Dead in Greek Tragedy*, U. Popat and J. Abraham, Eds. Demos Medical Publishing, 2011, pp. 191–194.
- [4] R. G. Bagasjvara, I. Candradewi, S. Hartati, and A. Harjoko, "Automated Detection and Classification Techniques of Acute

Leukemia using Image Processing : A Review," 2016 2nd Int. Conf. Sci. Technol., pp. 35–43, 2016.

- [5] F. Novoselnic, R. Grbic, I. Galic, and F. Doric, "Automatic White Blood Cell Detection and Identification Using Convolutional Neural Network," *Proc. Int. Conf. Smart Syst. Technol.* 2018, SST 2018, pp. 163–167, 2018.
- [6] A. Gautam, "Automatic Classification of Leukocytes using Morphological Features and Naïve Bayes Classifier," 2016 IEEE Reg. 10 Conf., pp. 1023–1027, 2016.
- [7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 779–788, 2016.
- [8] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 6517–6525, 2017.
- [9] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," 2018.
- [10] A. Krizhevsky, I. Sutskever, and Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Handb. Approx. Algorithms Metaheuristics*, pp. 1–1432, 2007.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.
- [12] H. Kutlu, E. Avci, and F. Özyurt, "White blood cells detection and classification based on regional convolutional neural networks," *Med. Hypotheses*, vol. 135, no. October 2019, p. 109472, 2020.
- [13] F. Özyurt, "A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine," *Soft Comput.*, vol. 24, no. 11, pp. 8163–8172, 2020.
- [14] M. Togacar, B. Ergen, and M. E. Sertkaya, "Subclass separation of white blood cell images using convolutional neural network models," *Elektron. ir Elektrotehnika*, vol. 25, no. 5, pp. 63–68, 2019.

- [15] H. A. Aliyu, M. A. A. Razak, R. Sudirman, and N. Ramli, "A deep learning alexnet model for classification of red blood cells in sickle cell anemia," *IAES Int. J. Artif. Intell.*, vol. 9, no. 2, pp. 221–228, 2020.
- [16] S. Vathanaavaro, S. Tungjitnob, and K. Pasupa, "White Blood Cell Classification: A Comparison between VGG-16 and ResNet-50 Models," p. 2.
- [17] M. Shahzad, A. I. Umar, M. A. Khan, S. H. Shirazi, Z. Khan, and W. Yousaf, "Robust Method for Semantic Segmentation of Whole-Slide Blood Cell Microscopic Images," *Comput. Math. Methods Med.*, vol. 2020, 2020.
- [18] M. Sharma, A. Bhave, and R. R. Janghel, "White Blood Cell Classification and Counting Using Convolutional Neural Network," *Adv. Intell. Syst. Comput.*, vol. 900, pp. 135–143, 2019.
- [19] C. B. Wijesinghe, D. N. Wickramarachchi, I. N. Kalupahana, L. R. De Seram, I. D. Silva, and N. D. Nanayakkara, "Fully Automated Detection and Classification of White Blood Cells," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2020-July, pp. 1816–1819, 2020.
- [20] Y. Liu, Y. Fu, and P. Chen, "WBCaps: A Capsule Architecture-based Classification Model Designed for White Blood Cells Identification," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 7027–7030, 2019.
- [21] K. Almezghwi and S. Serte, "Improved Classification of White Blood Cells with the Generative Adversarial Network and Deep Convolutional Neural Network," *Comput. Intell. Neurosci.*, vol. 2020, 2020.