International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

Comparison of YOLOv3, YOLOv4 and YOLOv5 Performance for Detection of Blood Cells

Ramya. A[1], Venkateswara Gupta Pola[2], Dr. Amrutham Bhavya Vaishnavi[3], Sai Suraj Karra[4]

[1]Ramya. A
Dept. of Computer Science and Engineering
Marri Laxman Reddy Institute of Technology and Management

[2]Venkateswara Gupta Pola
Dept. of Computer Science and Engineering
SASTRA University

[3]Dr. Amrutham Bhavya Vaishnavi
Bachelor of Dental Surgery
Malla Reddy Institute of Dental Sciences

[4]Sai Suraj Karra
Dept. of Mechanical Engineering (Mechatronics)
Mahatma Gandhi Institute of Technology

Abstract - *Blood cell count plays a vital role within the field of clinical diagnosing. In the recent times," the deep - learning-based detection method YOLO" has been proved be a novel method to count the blood cells and platelets effectively. Albeit its efficiency, the YOLO detection method has a few limitations like insufficient positioning of the bounding boxes and in distinguishing overlapping objects, in order to overcome these limitations, we propose a brand new deep-learning-based method, termed Attention-YOLO. Attention - YOLO is achieved by adding the channel attention mechanism and the spatial attention mechanism to the feature extraction network. By using the filtered and weighted feature vector to switch the initial feature vector for residual fusion, Attention-YOLO can help the network to boost the detection accuracy. The experimental results suggest that the Attention-YOLO has a higher detection performance in somatic cell count without introducing too many additional parameters compared to YOLO network. The popularity accuracy of cells (RBCs, WBCs, and platelets) in Attention-YOLO has an improvement of 6.70%, 2.13%, and 10.44%, respectively, and in addition to that the mean Average Precision (mAP) demonstrated an improvement of 7.14%. The purpose of this paper is to compare the performance of YOLO v3, v4 and v5 and conclude which is the best suitable method.*

Key Words*: *Machine Learning, YOLO, Blood cells, Deep learning

1. INTRODUCTION

Blood cell count is one of the crucial parts of the biopsy. The proper cell count helps in diagnosing potential diseases and associated lesions, also early detection of the underlying pathology and in-turn help in proper treatment. The three common types of blood cells are red blood cells (RBCs), white blood cells (WBCs), and platelets. The role of RBCs is to supply oxygen, while WBCs play an immune role, and platelets are involved in hemostasis. Generally, Normal RBC range is 4.5 to 6.5 million cells per micro of blood, Normal WBC range is 4500 to 11000 WBC per micro liters of blood and Normal platelet count is 1.5 L to 4.5 L Platelets per micro liter of blood, which is considerably a large amount. The quality manual detecting method is incredibly cumbersome and is liable to manual errors.

In the automatic reckoning method of blood cells, 2 totally different strategies, i.e., the image process strategies and deep-learning (DL)-based strategies are used for unit measurement. The first stage which involves the image process strategies unit wide uses the automated cell recognition technology-supported Hough rework, and watershed technique supported distance transformation, etc. But these image process strategies have some issues in cell detection, such as the accuracy of sleuthing cells cannot meet the necessities in areas with high cell overlap. By victimization convolutional neural network (CNN) as a classifier, Habibzadeh et al. proposed a corpuscle classification technique, which could mechanically classify WBCs into one in all 5 sorts from biology pictures. However, the speed of this system is slow. By employing a class-conscious erythrocyte patch extraction technique followed by a shape-invariant erythrocyte patch normalization technique, Xu et al. reduced the machine value throughout each of the coaching and learning procedures and classified RBC of ill patients. Moreover in 2019, with the help of YOLO (you only look once), Alam and Islam mechanically

established and counted RBCs, WBCs, and platelets, with the detection speed of a second. However, the YOLO technique has difficulties in characteristic overlapping objects and positioning the bounding box. Together, in cell experiments, blood cells might overlap each other, and seem as clusters within the image, which concludes that YOLO noticeon technique is not efficient to accurately detect every cell.
To overcome the restrictions, we've proposed a replacement DL-based technique, termed Attention-YOLO, which is achieved by adding the channel attention mechanism and abstraction attention mechanism to the feature extraction network. Furthermore, by victimization of the filtered and weighted feature vector to exchange the initial feature vector for residual fusion, Attention-YOLO is helpful in increasing the detection accuracy of the network. To evaluate the performance of the projected technique, a series of experimental knowledge square measures are used. The investigational results illustrated that Attention-YOLO has better performance in corpuscle counting without introducing extra parameters than standard YOLO detection technique. The Attention -YOLO has an improved recognition accuracy of cells (RBCs, WBCs, and platelets) by about 6.7%, 2.13%, and 10.44% respectively and therefore mean Average preciseness (mAP) has an improvement by 7.10%. This paper demonstrates the methodology in the following manner: In one section a pair of, the strategies and dataset used will be discussed? the experimental results for various versions of yolo is measured. Followed by the conclusions in section five.

2. METHODOLOGY

2.1YOLO network structure and detection principle

YOLO treats the detection task as a regression downside and has been wide utilized in image process fields. Recently, several versions of YOLO (e.g., YOLOv1, YOLOv2, and YOLOv3) have additionally been planned. Compared to YOLOv1 and YOLOv2, YOLOv3 has the following advantages:

1) The improved classification performance on complicated datasets.
2) The increased quantity of data within the feature map.
3) The deeper network layers

We've used YOLOv3 as it is advanced. Briefly, YOLOv3 uses Darknet-53 because of the feature extraction network, and adopts the strategy just like the feature pyramid network. It will directly use the initial input pictures and annotations for coaching. As a result, it saves computing resources. In the detection task, firstly, an image is sent as an input into the feature extraction network, and extracted feature vectors are obtained which are then sent to a structure just like the feature pyramid network, and therefore the grid cell is obtained at three scales. Furthermore, every grid cell predicts three bounding boxes, and every bounding box predicts a vector P, as follows:

(1) $P=(tx+ty+tw+th)+P0+(P1+P2+\cdots+Pn)$

with

(2) $P0=\Pr(\text{Object})\times \text{IOU}_{\text{predtruth}}$

where tx, ty, tw, th is that the coordinates associated with the bounding box. $Pr(Object)$ represents the likelihood that the thing is within the prediction box. IOU reflects the accuracy of the object's position. Finally, the non-maximum suppression is performed on the generated prediction to get the ultimate prediction result.

2.2Attention-YOLO network structure and detection principle

It is to be noted that in YOLOv3 detection processes, every region in entire feature map is treated equally. That is, every space is taken into account with identical contribution to the ultimate detection. However, in experiments, blood cells could overlap one another, or seem as clusters within the image.

Based on the above concerns, to enhance the detection accuracy, the eye mechanism module is introduced into the network. channel attention mechanism will filter and weight options in channel dimension, that is useful for up the detection performance. The spatial attention mechanism will model options on the spatial relationship, which might supplement the point relationship info that the channel attention mechanism cannot get. Considering that: 1) Convolutional Block Attention Module (CBAM) combines each house and channel attention mechanisms, that permits to quickly notice necessary feature from various options, suppresses the moot or unimportant options, and improve the potency and accuracy of necessary options processing; 2) CBAM will be handily embedded in network structures, and therefore the end-to-end coaching will be dispensed while not dynamical the initial network structure, during this work, CBAM is chosen and introduced into YOLOv3 network. In detail, as shown in Fig. 1, the input feature maps are subjected to world most pooling and world average pooling operations, therefore they are sent to the multilayer perceptron (MLP) for channel info filtering. After that, the MLP output options supported the coefficient element-wise for generating the channel attention feature map. consecutively, the most pooling and average pooling are performed on the channel dimension on top of that, the output of the 2 operations is combined and a feature descriptor is obtained. Finally, the convolution operation is employed to encrypt, and therefore the spatial attention map is obtained.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

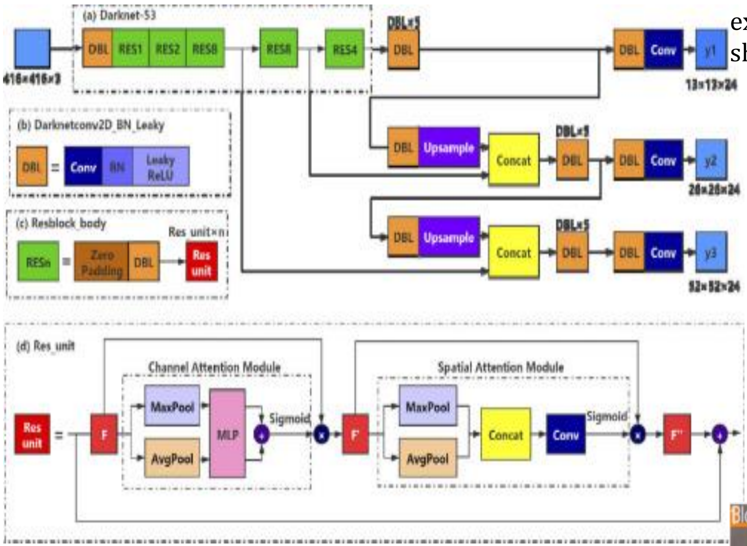


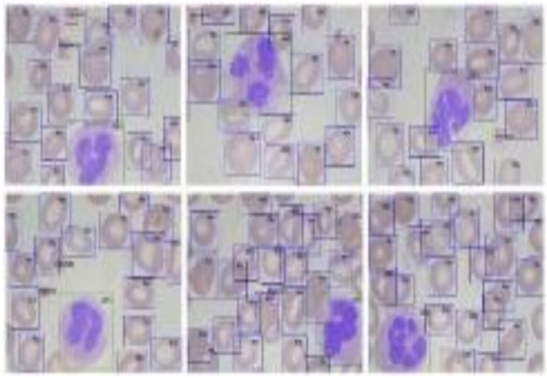
Figure 1: - Attention-YOLO network structure

attention module is introduced into the residual module.
It should be noted that the initial implementation of the YOLO configuration is trained for eighty totally different categories. once mistreatment YOLO for blood count, the quantity of categories is modified from eighty to three (RBCs, WBCs, and platelets). Corresponding, the quantity of filters within the final convolutional layer is modified to twenty-four. relating, the quantity is calculated by,

□□=□□×(□□+5) $NF=NA\times(NC+5)$ with Na being the quantity of anchor boxes and Old North State being the quantity of categories. Here, Na and Old North State are set to three, severally.

3. Blood cells dataset

The dataset used for this work could be a public blood corpuscle dataset (Blood Cell Count Dataset, BCCD), that is offered on https://github.com/Shenggan/BCCD_Dataset. BCCD contains 874 cell pictures and four, 870 annotations (RBCs, WBCs, and platelets). The image resolution is 416×416 pixels. The selection of those datasets is target-hunting by the very fact that they're open supply and absolutely accessible to the analysis community and also the general public.
It ought to be recognized that in a very BCCD dataset, some pictures contain RBCs, however the provided annotation file doesn't contain the corresponding RBCs. The mismatches might have an effect on the accuracy of the analysis. to deal with the matter, during this work, supported the antecedental provided annotations, we have a tendency to manually annotate the dataset once more, that is achieved by LabelImg image annotation software package (https://github.com/tzutalin/labelImg). Here, the annotation operation follows 2 principles: 1) For the native cells at the sting of the image, if the realm is a smaller amount than half-hour of the complete cell, it'll not be labelled; For the extremely adherent cells if the cell overlap exceeds eightieth, only 1 cell is marked. Figure a pair of shows some labelled pictures manually.



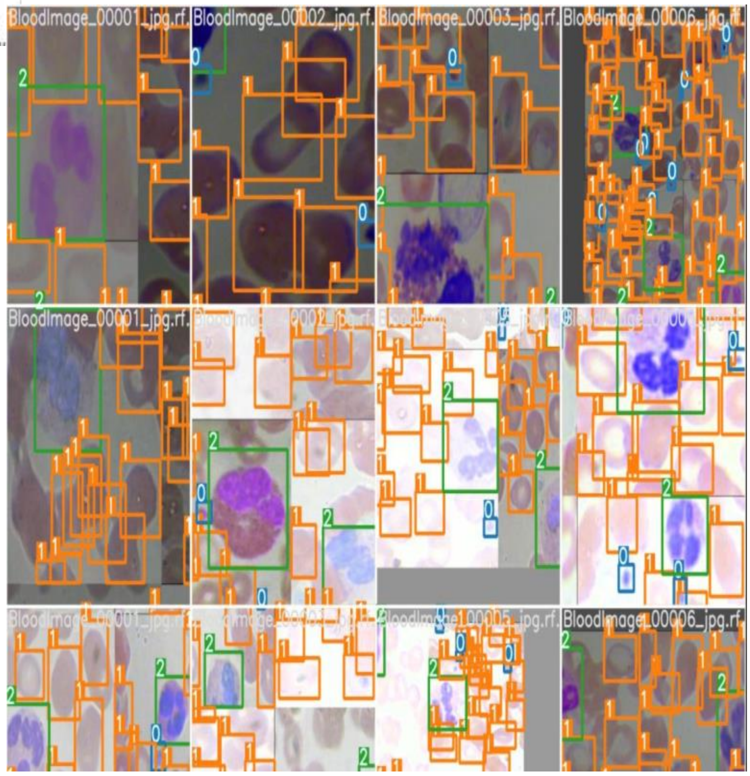


Figure 2: - The above is sample annotated dataset

4. RESULTS DISCUSSIONS

This is the part which is most awaited as the results of YOLOv3, YOLOv4 and YOLOv5 will be compared. The Comparison is done using the tensor board module in which we can enable the metrics and compare the results. The training is done with 100 epochs on YOLOv3, YOLOv4 and YOLOv5 respectively.

COMPARISON TABLE

	Precision	Recall	Acuracy
YOLOv3	0.71	0.87	0.86
YOLOv4	0.82	0.88	0.89
YOLOv5	0.84	0.89	0.91

Figure 3: - Comparison of parameters of YOLO v3, v4 and v5

The below figures are cumulative graphs of performance characteristics for YOLO v3, v4 and v5 respectively.

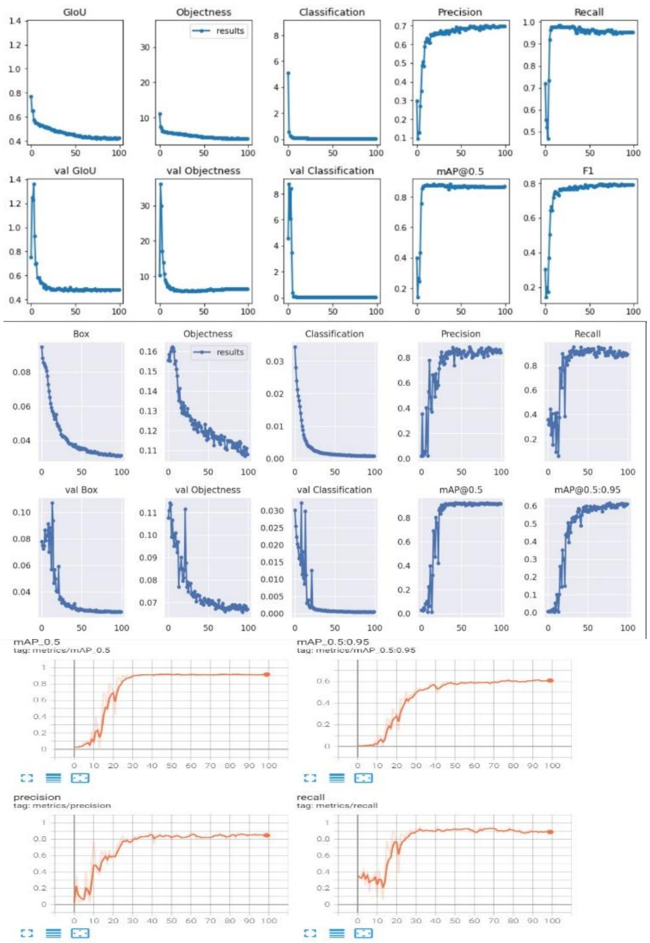


Figure 4: - The above are cumulative graphs of performance characteristics for YOLO v3, v4 and v5 respectively.

5. CONCLUSION

The best suitable version now is YOLO v5, but the time taken for training is much lower in YOLOv3 compared to v4 and v5. And when taken in consideration v4 is having some good performance characteristic and also training time is relatively lower than that of v5. So, we can choose v4 and make the hyper-parameters better and make the model more lucrative in nature.

Accurate cell reckoning is vital in medical image analysis. In clinical applications, generally, varied forms of cells are manually counted, resulting in an laborious work. The DL-based detection methodology, such as YOLO, will mechanically establish and count RBCs, WBCs, and platelets. However, this YOLO methodology has difficulties in identifying overlapping objects and positioning the bounding box. the aim of this paper is to boost cell detection accuracy, that is achieved by adding channel attention and spatial attention mechanisms to the feature extraction network.

It can be observed from the experimental results that compared with the quality YOLO methodology, the projected Attention-YOLO methodology achieves the higher performance in police work RBCs, WBCs, and platelets, in terms of RA and mAP, wherever the RA of RBCs, WBCs, associate degreed platelets have associate degree improvement and mAP has an improvement. Additionally, the results from another dataset demonstrates the efficacy of Attention-YOLO model, where RBC, WBC, and platelets are effectively detected.

It should note that for the DL-based methodology, a large quantity of coaching knowledge will be required to improve the generalization ability and efficiency of the model. In this work, the coaching set has 765 pictures. the Attention-YOLO model will observe alternative datasets, where RA has been declined, usually caused by the lean coaching datasets. On the other hand, corpuscle datasets are solely custom trained. In clinical applications, the categories of cells are varied and complicated, therefore differing kinds of cell knowledge need to be collected and used. Moreover, during this work, the arrogance threshold is set by calculating the typical absolute error. The ways for obtaining the arrogance threshold could be by improving the detection performance of the projected methodology. The dataset utilized in this work is unbalanced, the detection performance of the Attention-YOLO methodology can be improved by completely processing these unbalanced knowledges. Due to the addition of the eye mechanism, the detection time of Attention-YOLO has enhanced by concerning twenty milliseconds compared to YOLO. These connected works are going to be investigated in our future study.

REFERENCES

[1] Y. M. Alomari, S. N. H. S. Abdullah, R. Z. Azma, and K. Omar, "Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection method," Comput. Math. Method. Med. 2014, 1–17 (2014). [**CrossRef**]

[2] S. Acharjee, S. Chakrabartty, M. I. Alam, N. Dey, V. Santhi, and A. S. Ashour, "A semiautomated approach using GUI for the detection of red blood cells," 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT) (2016), pp. 525–529.

[3] Z. Ejaz, C. A. Ahmad, and H. Justin, "Automatic red blood cell detection and counting system using Hough transform," Am. J. Pharm. Sci. **5**(8), 267–271 (2018). [**CrossRef**]

[4] V. Acharya and P. Kumar, "Identification and red blood cell automated counting from blood smear images using computer-aided system," Med. Biol. Eng. Comput. **56**(3), 483–489 (2018). [**CrossRef**]

[5] S. M. Mazalan, N. H. Mahmood, and M. A. A. Razak, "Automated red blood cells counting in peripheral blood smear image using circular Hough transform," 2013 International Conference on Artificial Intelligence, Modelling and Simulation (AIMS) (2013), pp. 320–424.

[6] J. Hari, A. S. Prasad, and S. K. Rao, "Separation and counting of blood cells using geometrical features and distance transformed watershed," 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR) (2014), pp. 77–82.

[7] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," Med. Image. Anal. **42**, 60–88 (2017). [**CrossRef**]

[8] S. Q. Liu, S. D. Liu, W. D. Cai, H. Y. Che, S. Pujol, R. Kikinis, D. G. Feng, and M. J. Fulham, "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," IEEE. Trans. Biomed. Eng. **62**(4), 1132–1140 (2015). [**CrossRef**]

[9] "Improved detection performance in blood cell count by an attention-guided deep learning method" Zhengfen Jiang, Xin Liu, Zhuangzhi Yan, Wenting Gu, and Jiehui Jiang.