Tissue and Cell 76 (2022) 101761

Contents lists available at ScienceDirect

Tissue and Cell

journal homepage: www.elsevier.com/locate/tice

A comparative assessment of deep object detection models for blood smear analysis

Kabyanil Talukdar [a], Kangkana Bora [a,*], Lipi B. Mahanta [b], Anup K. Das [c]

[a] Department of Computer Science & IT, Cotton University, Guwahati, 781001, Assam, India
[b] MCSD Division, Institute of Advanced Study in Science and Technology, Guwahati, 781038, Assam, India
[c] Arya Wellness Centre, Guwahati, 781032, Assam, India

ARTICLE INFO

Keywords:
WBC
Platelet
Deep learning
Object detection
Classification
Localization
Precision
Recall
Inference

ABSTRACT

A blood smear is a common type of blood test where blood sample is taken from a patient, smear is made from the sample followed by observation of red blood cells, white blood cells and platelets. A pathologist carefully observes the sample and manually counts the number of RBC, WBC and platelets. This entire process from creating a smear to manually counting each element is tedious and susceptible to human errors. That is why, with the advancement of deep learning, various object detection techniques have become useful for automating the process and mitigating human errors in blood smear analysis. This work presents a comparative assessment of three different object detection models namely Faster R-CNN, EfficientDet D3 and CenterNet Hourglass, and presents their respective inference results. The three models have been compared using the COCO evaluation metrics to identify the best model performance for the given task. It is observed that out of the three models, the Faster R-CNN model performs the best in detecting WBCs and platelets in microscopic blood smear images with an average precision of 99.4%. Critical tasks like medical image processing require accurate predictions to prevent unintended ramifications. Therefore, while slower in terms of inference time, Faster R-CNN is the go-to model where accuracy is the priority. The work is also compared with the existing work in this domain to prove its efficiency.

1. Introduction

Human blood consists of four main components - plasma, RBC, WBC, and platelets. Among them, plasma is the largest and makes up around 55 % of its total content. Plasma carries hormones, proteins, nutrients, and waste materials to and from different parts of the body. RBC is made in the bone marrow and contains Hemoglobin, which carries oxygen from the lungs to the entire body. On the other hand, WBCs are part of our body's immune system (Almezhghwi and Serte, 2020). These cells help our body to fight against infections and diseases. WBCs can engulf pathogens (virus, bacteria, fungi etc.) and prevent our body from getting infected. There are five types of WBCs in human body - basophils, neutrophils, eosinophils, monocytes and lymphocytes. WBCs normally range between 4300 and 10,800 cells per cubic millimeter of blood. On the other hand, Platelets are very tiny blood cells that help form clots to stop bleeding. In case a blood vessel in the human body gets damaged, it fires signals and in response to that signal, the body sends groups of platelets in that particular vessel to form clots to stop further bleeding. Counting these components with accuracy is crucial to detecting potential diseases. Platelet count is essential in identifying diseases such as malaria, dengue, yellow fever, etc. Therefore, an immediate diagnosis of platelet count is needed to determine the state of a patient's health. In a healthy human body, there exist around 150,000–400,000 platelets per microliter of blood (Ghoshal and Bhattacharyya, 2014).

Many techniques have been introduced to achieve automation in blood smear image processing. Researchers have been applying machine learning methods to achieve the addressed task. However, with the advancement of deep learning, it has significantly outperformed the former. That is why, we have conducted the survey on the application of deep learning techniques in this domain. On the basis of our survey, it was observed that Mohammad Mahmudul Alam et al., (Alam and Islam, 2019) proposed a deep learning based object detection model based on YOLO framework with VGG-16 architecture to detect different blood cells; achieving an accuracy of 96.06 % for RBCs, 86.89 % for WBCs and

* Corresponding author.
E-mail addresses: kabyaniltalukdarindia@gmail.com (K. Talukdar), kangkana.bora@cottonuniversity.ac.in (K. Bora), lbmahanta@iasst.gov.in (L.B. Mahanta), dranupdas@gmail.com (A.K. Das).

https://doi.org/10.1016/j.tice.2022.101761
Received 12 October 2021; Received in revised form 21 January 2022; Accepted 18 February 2022
Available online 22 February 2022
0040-8166/

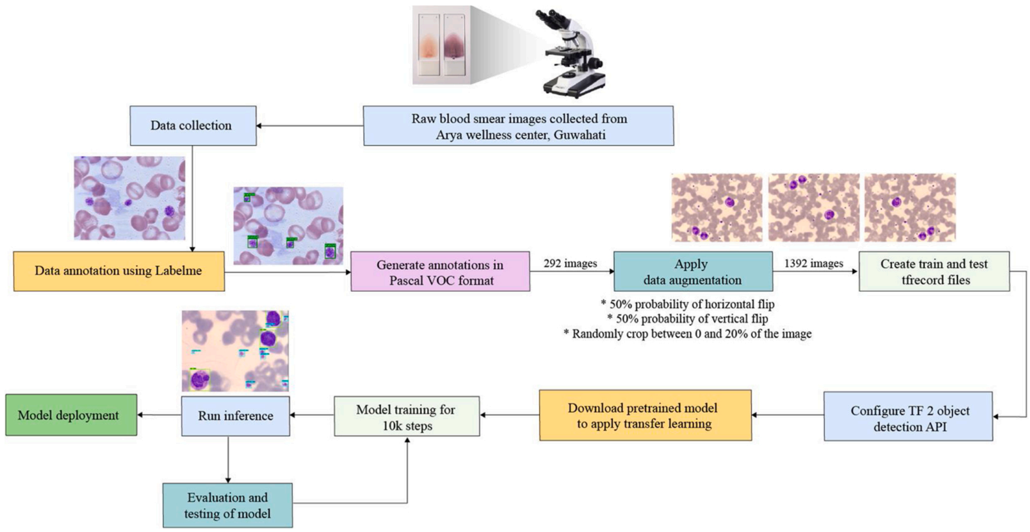

Fig. 1. Workflow diagram of the proposed work.

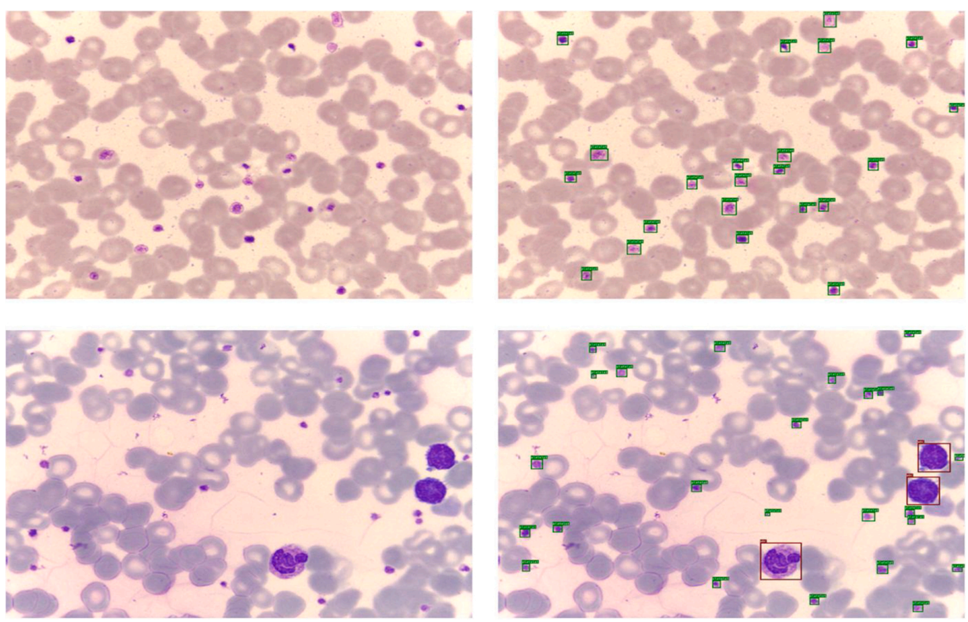

Fig. 2. Sample raw images (left) and their annotated versions (right). The green bounding boxes in the right side images are used for annotating platelets and red bounding boxes are used for annotating WBC.

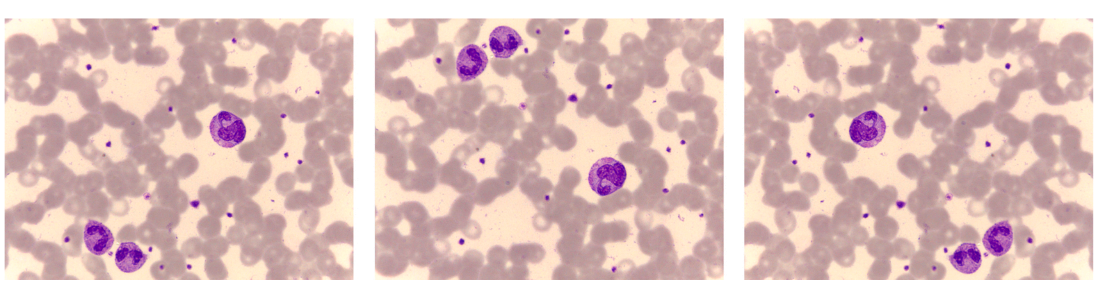

Fig. 3. Raw image (left), its horizontally flipped augmentation (middle), and its vertically flipped augmentation (right).

96.36 % for platelets. Ensaf Hussein Mohamed et al., (Mohamed et al., 2020) proposed a method to automatically identify and classify WBCs in a microscopic image into four different categories, using MobileNet-224 as feature extractor with logistic regression; achieving accuracy of 97.30 %. Zhengfen Jiang et al., (Jiang et al., 2021) proposed a new deep learning method called Attention-YOLO for drawing better bounding boxes in detecting different blood cells with improved accuracy. Hüseyin Kutlu et al., (Kutlu et al., 2020) proposed a work on detecting

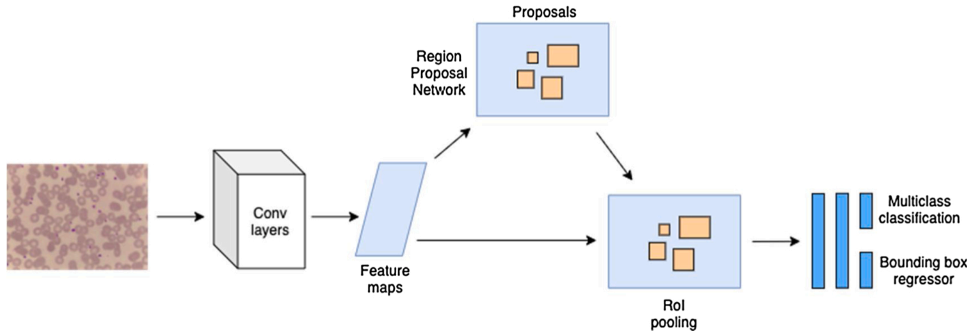

Fig. 4. Faster R-CNN model architecture.

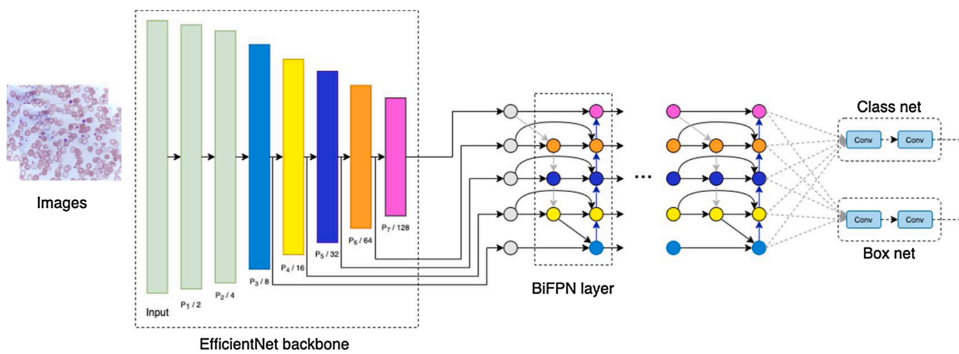

Fig. 5. EfficientDet D3 model architecture.

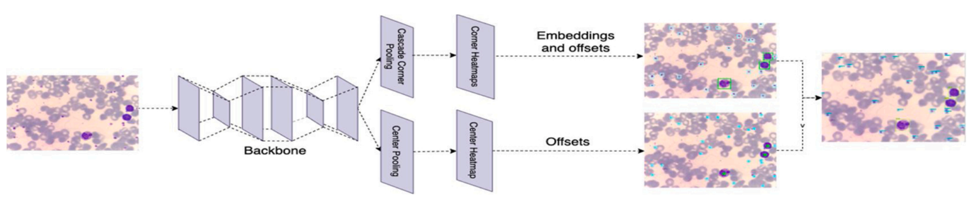

Fig. 6. CenterNet model architecture.

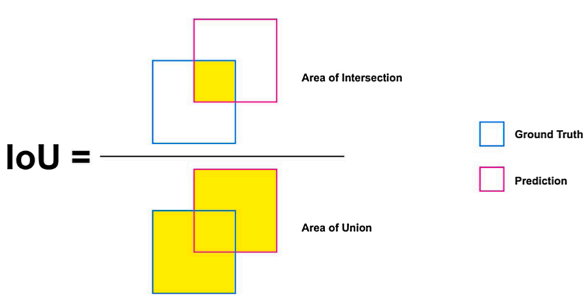

Fig. 7. Intersection over Union (IoU) in object detection.

Table 1
Average precision metrics and recall measures of FR-CNN model. Bold value is indicating the best AP and AR percentage values and its corresponding IoU.

IoU	AP Percentage	Max Detections	AR Percentage
0.50:0.95	71.6 %	1	38.8%
0.50	99.4 %	10	65.3%
0.75	80.7 %	100	76.4%

white blood cells using region based convolutional neural networks (R-CNN) with ResNet50 as feature extraction layer; which achieved 100 % success in determining WBC cells. Another study of counting while blood cells in bone marrow images was carried out by Da Wang et al. (Da et al., 2021) which used Faster R-CNN and a Feature Pyramid Network to detect and count WBCs which achieved 98.8 % accuracy when trained and tested on the dataset of The Second Affiliated Hospital of Zhejiang University. Similarly, Tiancheng (Xia et al. (2019)) carried out a study to investigate in-vitro detection of white blood cells using Faster R-CNN with transfer learning. This particular study showed a new pathway to apply deep learning based object detection into microfluidic point-of-care medical devices. Faster R-CNN has shown promising results in detecting blood cells in Malaria infected blood images as well, as suggested by the study by Jane (Hung et al. (2019)) to recognize infected cells and their stages. Another study proposed by Nalla (Praveen et al. (2021)) used YOLO v3 model to localize and classify the white blood cells with 99.2 % accuracy, trained and tested on the public BCCD dataset. Object detection models have also been used in detecting Leukemia, which is one of the deadliest diseases in humans. One of such studies was done by Samir (Abou El-Seoud et al. (2020)) which was focused on using CNN to detect and classify normal WBCs. Luis Claudio Soto-Ayala et al. (Soto-Ayala and Cantoral-Ceballos, 2021) presented their development of a deep learning model to identify anomalies in blood cells, which helped address the absence of tools to detect cancer at an early stage.

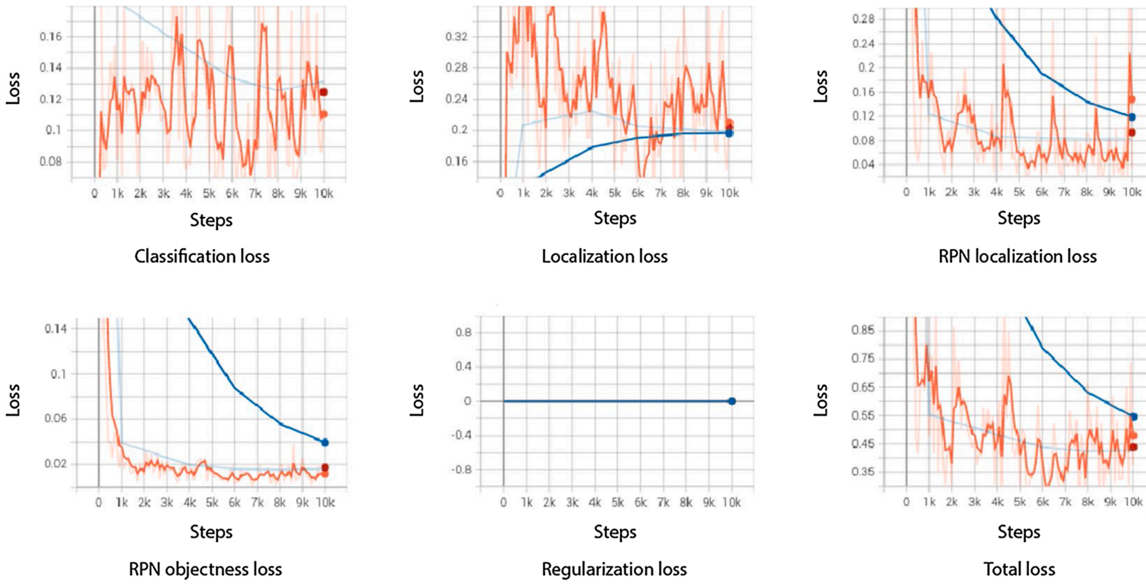

Fig. 8. Training(orange), validation (blue) and testing(red) loss of FR-CNN model after applying smoothing through TensorBoard to produce more consistent curves. The light cyan and light orange curves represent the actual loss curves respectively. Decreasing values of the loss is an indication of the models learning abilities. Decaying loss indicates better learning.

Table 2
Average precision and recall metrics of EfficientDet D3 (without tuning). Here, AP and AR values are not satisfactory (below the range of 89 % and 67 % respectively).

IoU	AP Percentage	Max Detections	AR Percentage
0.50:0.95	60.5 %	1	35.3%
0.50	88.2 %	10	58.6%
0.75	70.8 %	100	66.6%

Table 3
Tuned hyperparameters for EfficientDet D3.

Hyperparameters	Default value	Tuned value
fixed_shape_resizer	keep_aspect_ratio_resizer	Width: 896, height: 896
use_dropout	N/A	True
dropout_keep_probability	N/A	0.8
aspect_ratios	1.0, 2.0, 0.5	1.0
max_detections_per_class	100	30
max_total_detections	100	60
classification_weight	1.0	1.2
localization_weight	1.0	0.8
learning_rate_base	0.08	0.001
warmup_learning_rate	0.001	0.0001
max_number_of_boxes	100	60

Manually observing blood smears under a microscope yields results that depend on the observer's skills and make it susceptible to human error. This work aims to automatically develop an object detection system to detect WBCs and platelets from sample blood smear images. This work presents the Faster R-CNN (FR-CNN), EfficientDet D3 and CenterNet Hourglass object detection models, and compares their results to find the best fit model. Real-world implementation of the proposed work will result in accurate test results. It would also reduce the workload, the dependency of the operator-specific skills, and the time taken for the analysis.

1.1. Contribution of the study

The proposed work aims to provide a comparative assessment of three state-of-the-art object detection models in performing blood smear analysis. This work is the first in its domain to provide a comparative assessment among FR-CNN, EfficientDet D3 and CenterNet Hourglass. Some common blood tests carried out regularly in patients as well as during diagnosis including Complete Blood Count (CBC) test require pathologists to manually count each cell. Based on that counting, the disease of the patient is diagnosed. This work automates the process of detecting and counting of WBCs and platelets. It also finds the best model to achieve the task. Therefore, future research in the domain would be able to leverage the data presented in this work to decide which models to use in order to get the best results. The novelty of the work lies in the application framework where deep learning is deployed in object detection of blood components name WBC and Platelets.

Rest of the work has been organized as follows - Section 2 focuses on the methodology used for performing the experiments, results and discussion is explained in Section 3. Finally, conclusions are drawn in last section.

Table 4
Average precision and recall metrics of EfficientDet D3 fine tuned. Highest precision value achieved is 96.1 % and recall value is 65.8 % which is very low.

IoU	AP Percentage	Max Detections	Percentage
0.50:0.95	60.3%	1	35.3%
0.50	96.1 %	10	57.6%
0.75	61.2%	100	65.8 %

2. Methodology

The overview of the proposed methodology is displayed in Fig. 1. The proposed methods detect WBCs and platelets in microscopic blood smear images and counts each elements' number. We used 292 microscopic blood smear images from Arya Wellness Center, Guwahati, Assam, India, following all ethical protocols and with patient consent. The microscope used to capture the images was Leica (Leica ICC 50 HD microscope, 24-bit colour depth). The resolution of each image was 2048 × 1536 pixels. The raw images were annotated manually using a labelling software named LabelMe with the help of doctors. Blood elements were labelled into WBCs and platelets and images were exported to Pascal VOC format. Fig. 2 displays the raw and annotated version of

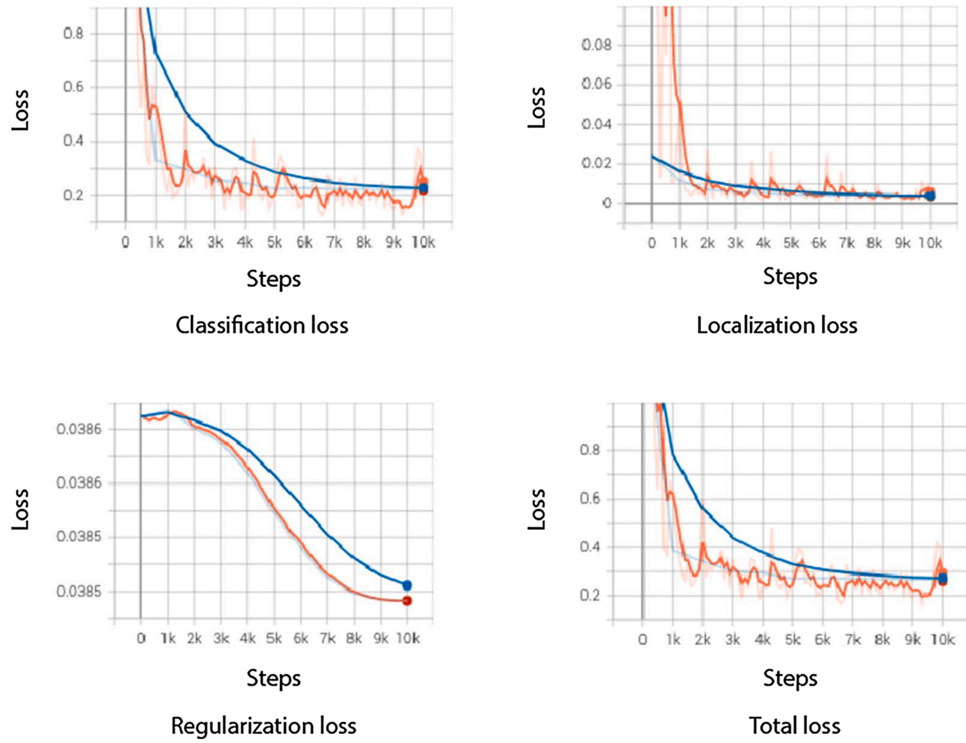

Fig. 9. Training(orange), validation(blue) and testing(red) loss of EfficientDet D3 fine tuned after applying smoothing through TensorBoard to produce more consistent curves. The light cyan and light orange curves represent the actual loss curves respectively.

Table 5
Average precision metrics of CenterNet Hourglass. The precision value is very satisfactory which is around 99.2 % with 0.50 as IoU. AR value achieved is 80.2 % corresponding to the IoU 0.50:0.95.

IoU	AP Percentage	Max Detections	AR Percentage
0.50:0.95	75.5%	1	40.3%
0.50	99.2 %	10	68.1%
0.75	85.2%	100	80.2 %

two blood smear images.

The exported image dataset was then preprocessed in the following steps -

A Image pre-proprecessing:
upperRomanB Auto orient
upperRomanC Augmentations:
- Brightness (-35 % to +35 %), Salt & Pepper noise (up to 5% of pixels), Horizontal flip, Vertical flip, Crop (0–20 % zoom)

D Train / Test split:
I Training set 87 % (1215 images)
II Validation set 9% (119 images)
III Testing set 4% (58 images)

Once the image pre-processing was complete, a total of 1392 images have been generated. Augmenting the images with Salt & Pepper noise in up to 5% of pixels in the training images improves model generalization. This type of noise randomly adds black and white pixels throughout the image and significantly deteriorates the image clarity. Due to which, the model is able to generalize better in unseen low quality images. Fig. 3 displays an example of how the images in the dataset have been augmented. The Tensorflow Object Detection API 2.0 is used to train the object detection models, which read images as TFRecord files. Therefore, the train, validation and test images is converted into train.tfrecord, validation.tfrecord and test.tfrecord files respectively.

In object detection, the main task of a model is to draw bounding boxes around detected objects of interest and to classify each object independently. As there could be variable number of detected objects, standard convolutional neural networks with a fully connected layer fail to facilitate an adequate architecture. That is why, algorithms like region based convolutional neural network have emerged (Yamashita et al., 2018). In this work, three models have been trained and compared their results – FR-CNN, EfficientDet D3 and CenterNet Hourglass.

FR-CNN was proposed by Shaoqing Ren et al., (Ren et al., 2016) that leverages a fully connected convolutional network called Region Proposal Network (RPN), which drastically improved existing region proposal computation bottleneck. In this model, the input images pass through a collection of convolutional layers and produce feature maps. Then the feature maps are passed through Regional Proposal Network (RPN), which is a network for proposing regions where an object could be located. Each pixel in the feature maps is called an anchor and there are multiple aspect ratios by which an anchor can be generated, such as 1:1, 1:2, 2:1 etc. The task of RPN is to label the anchor boxes as either foreground of background. The RPN has a classifier and a regressor to check if the anchor has objects and adjust the bonding boxes so that it fits the relevant object. The output of the RPN is a bunch of object proposals with bounding boxes which are passed to RoI pooling layer to make them fixed sized. Finally, the output is fed to another classifier to classify the object and regressor to better fit the bounding box to the object. Faster R-CNN replaces the selective search algorithm with RPN, reducing the region proposal time from 2 s to 10 milliseconds per image. Which makes it a candidate model for our work. Fig. 4 displays the architecture of Faster R-CNN model. With the use of RPN, Faster R-CNN performs much faster than its predecessors R-CNN and Fast R-CNN, which makes it usable for real time object detection.

AI has recently seen applications in devices with limited resources, such as embedded devices and self-driving cars. With that, model efficiency has become increasingly important. The family of EfficientDet

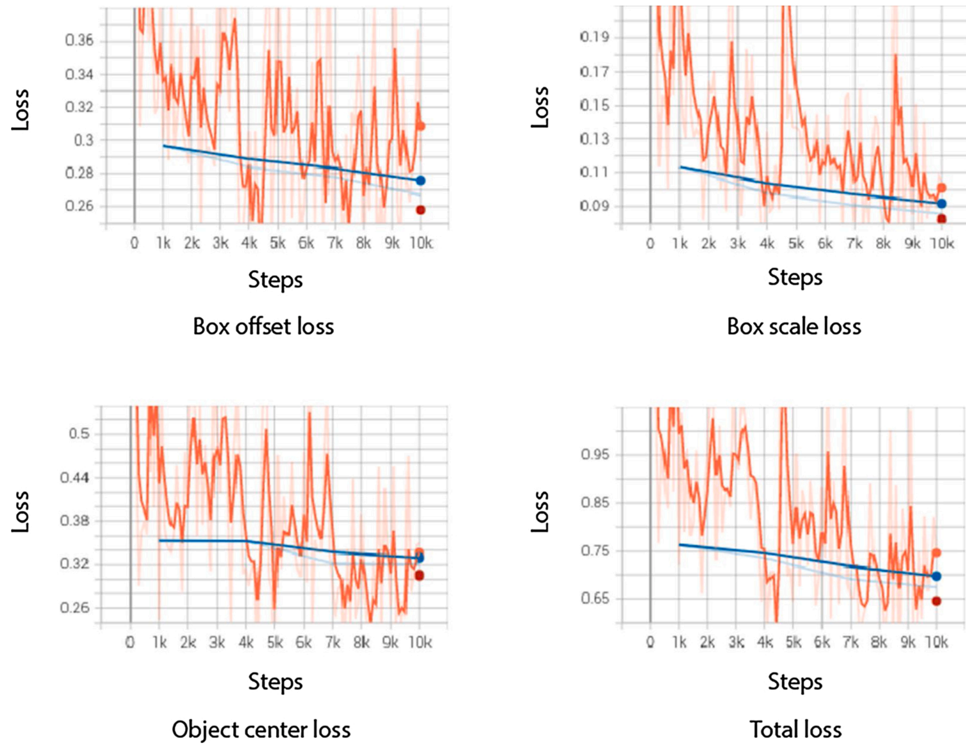

Fig. 10. Training(orange), validation(blue) and testing(red) loss of CenterNet Hourglass after applying smoothing through TensorBoard to produce more consistent curves. The light cyan and light orange curves represent the actual loss curves respectively.

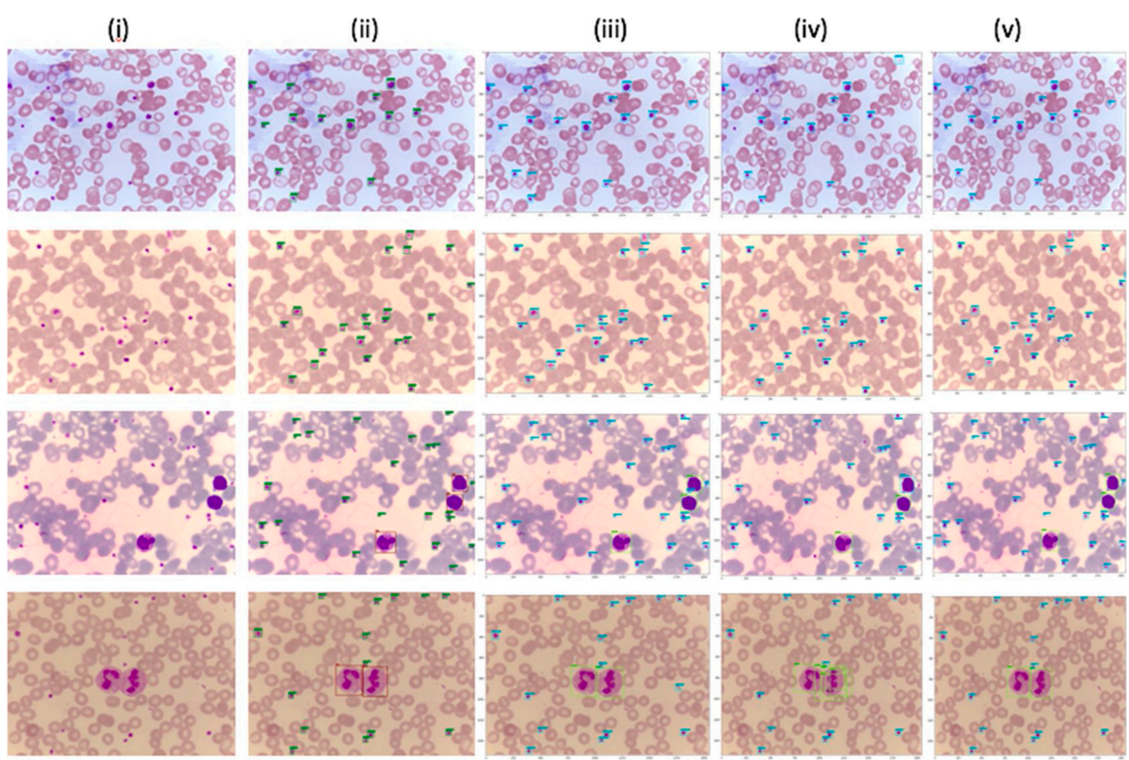

Fig. 11. Raw microscopic smear image (Column i), labelled ground truth image (Column ii), model predicted image of FR-CNN (Column iii), fine tuned EfficientDet D3 (Column iv) and of CenterNet Hourglass (Column v). Bounding box for platelets are in blue colour and bounding box for WBC is green colour.

models was developed keeping efficiency, accuracy and scalability in mind. It was developed by Mingxing Tan et al., (Tan et al., 2020) at Google. It was developed to increase model efficiency while being significantly smaller than the existing state-of-the-art models, so that it could be run on limited computational devices. EfficientDet was built on top of EfficientNet (Tan and Le, 2020) and it leveraged Weighted Bi-directional Feature Pyramid Network (BiFPN) as feature extractor. EfficientDet is a family of model ranging from D0 to D7. The architecture

Table 6
Side-by-side comparison of model performance. It is observed that FR-CNN achieved the highest precision of 99.4 %, but CenterNet Hourglass achieved a recall value of 81.2 % which is highest among all the other compared models. But the recall of fine tuned FR-CNN is 0.6 % low compared to CenterNet Hourglass. Also, precision indicates that best model for the object detection is FR-CNN. Testing the models on ALL-IDB public dataset yielded satisfactory results.

	Generated dataset		**ALL-IDB Public Dataset**	
Model	AP @ IoU	AR @ MaxDets	AP @ IoU	AR @ MaxDets
FR-CNN	99.4 % at 0.50	76.4% at 100	96.5 % at 0.50	70.4% at 100
EfficientDet D3 fine tuned	96.1 % at 0.50	65.8 % at 100	91.3% at 0.50	51.2% at 100
CenterNet Hourglass	99.2 % at 0.50	80.2 % at 100	95% at 0.50	71.7% at 100

of EfficientDet consists of a backbone network which extracts features from the input image. One of the limitations of the existing feature networks is the one way flow of information, which is why information of previous layers is lost in the later layers of the model. The proposed BiFPN bi-directional feature network is able to preserve the low level features flowing from the different levels of the backbone network. The final piece of the network is the box and class prediction network which predicts the object class and regresses the bounding box respectively. Fig. 5 displays the architecture of EfficientDet.

CenterNet was developed by Kaiwen (Duan et al. (2019)) in April 2019 which proposed a new paradigm of object detection without drawing any anchor boxes, contrary to what the existing state-of-the-art object detection models did. CenterNet was developed on top of CornerNet (Law and Deng, 2019) to address the issue of CornerNet's corner keypoint pairs' weak ability of catching the global information of an object, due to which CornerNet could not determine which pairs of keypoints should be grouped into an object. CenterNet became the state-of-the-art keypoint-based object detection model until YOLO V5 and Cascade R-CNN were proposed. Centernet treats the center of a box as both an object and a key point, and then uses this predicted center to find the bounding box's coordinates/offsets. The convolutional network processes an image and generates heatmaps for various keypoints. The heatmap peaks are associated with a particular class it belongs to. The network also predicts the width and height of the bounding box for each center. Every center has its own bounding box width and height. This method eliminates the Non-Maximal Suppression step in post processing. CenterNet uses the centers, bounding box dimensions and class probabilities to detect objects. Fig. 6 shows the architecture of CenterNet model.

The three trained models were evaluated using the COCO evaluation metrics. The COCO evaluation metrics show us Average Precision, Average Recall over multiple IoU thresholds. IoU tells us the overlap between the ground truth box and the prediction box (Fig. 7). If the IoU value is equal or greater than the threshold, that particular prediction is considered to be a True Positive (TP). If the IoU is less than the threshold, that particular prediction is considered to be a False Positive (FP). In object detection tasks, an IoU of 0.5 is considered good enough.

Once the TPs and FPs have been calculated, we can obtain Precision and Recall scores.

$$Precision = \frac{TP}{TP + FP} Recall = \frac{TP}{TP + FN}$$

From the precision and recall scores, we can obtain Average Precision (AP) values over different thresholds. The COCO evaluation metrics calculates AP for IoU from 0.5 to 0.95 with step size of 0.05, AP for IoU = .50 and AP for IoU = 0.75. We can also obtain Average Recall (AR) values from COCO evaluation metrics. The Tensorflow Object Detection API calculates all of them automatically.

Tensorflow Object Detection API works with pipeline configuration files, which contain model parameters. Fine tuning can be achieved by changing the configuration values such as input image dimension, learning rate, augmentation, dropout, IoU threshold, batch size etc. During training, a script can be run in the terminal in parallel to validate the model performance as it is being trained. Likewise, the test evaluation of the model can be achieved by running a script in the terminal once the training has finished. The results are stored in a specified directory as Tensorflow events. These event files can be fed into Tensorboard in order to visualize the metrics. We have trained, validated and tested the models in the above mentioned procedure.

3. Results and discussions

3.1. Faster R-CNN

Faster R-CNN has been trained with Inception ResNet V2 (Szegedy et al., 2016) as feature extractor for 10,000 steps with batch size 2. We can observe from Table 1 that it achieves 71.6 %, 99.4 % and 80.7 % Average Precision (AP) considering 0.50:0.95, 0.50 and 0.75 IoU respectively. It shows high precision of the model. Table 1 also lists the Average Recall (AR) metrics displaying good recall of the model. Fig. 8 shows different training, validation and testing losses which are decreasing over time, indicating that the model is learning. We have implemented learning rate decay, which initially increases and then decreases over time during training with the use of cosine learning rate decay with warm up. After training is complete, we run inference on test images and get promising results, as displayed in Fig. 11.

3.2. EfficientDet D3

We trained the EfficientDet D3 model for 10,000 steps with batch size 2 and initially achieved 88.2 % accuracy. With hyper-parameter tuning, the accuracy increased to 96.1 %.

Table 7
Comparison with the existing works in the same domain.

Author	Methodology	Aim (Region of Interest)	Dataset used	Average Precision (AP)	Remarks/ Observation
Muhammad et al.(Alam and Islam, 2019)	YOLO	RBC WBC Platelet	BCCD	62.36 %	AP can be improved
Ensaf Hussein (Mohamed et al. (2020))	DenseNet	WBC	BCCD	86 %	AP can be improved
Zhengfen (Jiang et al. (2021))	Attention-YOLO	RBC WBC Platelet	BCCD	94.3 %	AP can be improved
Hüseyin (Kutlu et al. (2020))	FR-CNN	WBC	BCCD + LISC	74 %	Only concentrating on WBC
Sun (Cheng et al. (2019))	FR-CNN	WBC	Own dataset	98.16 %	Only concentrating on WBC
Prayag (Tiwari et al. (2018))	DCLNN	Subtypes of WBC	Own dataset	88 %	Only concentrating on WBC
Proposed	FR-CNN	WBC Platelet	Own dataset	99.4 %	Best performance for detecting WBC and platelets

3.2.1. EfficientDet D3 before hyper-parameter tuning

When trained on default parameters, EfficientDet D3 model achieves average precision of 60.5 %, 88.2 % and 70.8 % with IoU thresholds of 0.50:0.95, 0.50 and 0.75 respectively. Similar to the Faster R-CNN model, we applied cosine learning rate decay in this model as well. Table 2 lists the AP metrics of the model. We noticed a drop in terms of AP compared to the previous model. The model's AR metrics are also displayed in Table 2. We observed again that the AR scores performed poorly than the Faster R-CNN model.

We observed a significant performance drop on this model without hyper-parameter tuning. That is why we decided to apply hyper-parameter tuning to improve the model performance.

3.2.2. EfficientDet D3 with hyperparameter tuning

We retrained the same model with different hyperparameters for 10,000 steps with batch size 2. The tweaked parameters are listed in Table 3.

The tuned model performed significantly better than the non-tuned model, but still it is not good compared to FR-CNN. As we can see in Table 4, the AP scores increased from the previous model to 96.1 %. However, we observed a decrease in AR for this model. We applied cosine learning rate decay in this model as well. Table 4 also displays the AR metrics. Fig. 9 shows the training, validation and test losses of the model. We observed from the results that, even though the EfficientDet model achieves higher benchmark scores than the Faster R-CNN model, it underperforms in this particular task. However, a possible scope of improvement would be to train the model for more steps. Fig. 11 shows the visual output of the object detection model.

3.3. CenterNet hourglass

We then trained CenterNet Hourglass model for 10,000 steps with batch size 2. Here we observe a significant improvement over the EfficientDet model. We can see from Table 5 that the AP scores are similar to the Faster R-CNN model; with 99.2 % on 0.50 IoU. The AR metrics in Table 5 also resemble the Faster R-CNN model's scores. Although very shaky, the losses also gradually slope downwards indicating that the model is learning. The different losses are displayed in Fig. 10. During training, this model also uses cosine learning rate decay to bring down the learning rate gradually. We observe the inference run on the model in Fig. 15 which shows promising accuracy.

4. Comparative assessment of three models viz. Faster RCNN, fine tuned EfficientDet D3 and CenterNet hourglass

In this work, a comparative assessment has been forwarded of three object detection models trained to detect objects in microscopic blood smear images. The evaluated results indicate that, among the three, FR-CNN performs the best in terms of accuracy with 99.4 %. While EfficientDet is the current state-of-the-art model for object detection with 51.2 COCO mAP, we observed that training the model for 10,000 steps with batch size 2 yields an accuracy of 96.1 % only. Even though EfficientDet is the fastest object detector among the three, it requires the most computing resources for training. On the other hand, the CenterNet Hourglass model yields much better accuracy of 99.2 %, leveraging it's novel keypoint-based architecture for better object detection.

We have also tried to evaluate the model performances on a benchmark dataset named as ALL-IDB (ZZZZZ, 2022). The model performances can be observed from Table 6. An AP of 96.5 % is achieved in the benchmark ALL-IDB dataset which is very satisfactory. This indicates the models consistency while working on two different datasets. So it can be concluded that the approach adopted for object detection is acceptable with respect to performance as well source variations.

5. Comparative assessment with existing works

Table 7 gives a comparative overview on all existing works in this particular domain. It is really challenging to compare with the existing models because of many reasons including the unavailability of the generated dataset the researchers are using and also unavailability of model parameters to regenerate their work. But we are performing a comparative assessment based their methodology and their outputs on their respective datasets. We are only concentrating on the papers which have used deep learning methodology.

6. Conclusion

Many common blood tests in patients include steps like manually detecting and counting blood elements which is time consuming, tedious and susceptible to human errors. That is why we have leveraged state-of-the-art deep learning models to automate the process and mitigating human errors in blood smear analysis. The proposed work provides a comparative assessment of three state-of-the-art object detection models, namely Faster R-CNN, EfficientDet D3 and CenterNet Hourglass in blood smear analysis. From our observation, we have found that the best model for this task is Faster RCNN. However, a lot of factors like dataset size, training time, fine tuning etc. come into play to determine the best result. Therefore, future research in the domain would be able to leverage the data presented in this work to decide which models to use in order to get the best results and improve on top of it. The novelty of the work lies on both the technical grounds where models have been fine tuned as well as the application framework where deep learning is deployed in object detection.

Medical image processing is a domain where accuracy is much important than inference time. From this we come to a conclusion that, when inference time is not a concern, Faster R-CNN is still the best model among the three for blood smear analysis.

Data availability

Data will be made available on request.

Author statement

Kabyanil Talukdar and Kangkana Bora has performed the experiments and designed the models. Anup K das is the domain expert (pathologist) who has guided in understanding the problem from biological aspects, generated the dataset and annotated the ground truth. Lipi B Mahanta conceptualized the main idea and lead the team and also helped in experimental evaluation and manuscript writing.

References

Almezhghwi, Khaled, Serte, Sertan, 2020. Improved classification of white blood cells with the generative adversarial network and deep convolutional neural network. Comput. Intell. Neurosci. 2020, 1–12. https://doi.org/10.1155/2020/6490479.

Ghoshal, Kakali, Bhattacharyya, Maitree, 2014. Overview of platelet physiology: its hemostatic and nonhemostatic role in disease pathogenesis. The Scientific World Journal 1–16. https://doi.org/10.1155/2014/781857.

Alam, Mohammad Mahmudul, Islam, Mohammad Tariqul, 2019. Machine learning approach of automatic identification and counting of blood cells. Healthc. Technol. Lett. 103–108. https://doi.org/10.1049/htl.2018.5098.

Mohamed, Ensaf H., El Behaidy, Wessam H., Khoriba, Ghada, et al., 2020. Improved white blood cells classification based on pre-trained deep learning models. J. Commun. Softw. Syst. 16, 37–45. https://doi.org/10.24138/jcomss.v16i1.818.

Jiang, Zhengfen, Liu, Xin, Yan, Zhuangzhi, et al., 2021. Improved detection performance in blood cell count by an attention-guided deep learning method. OSA Continuum 4, 323–333. https://doi.org/10.1364/OSAC.413787.

Kutlu, H.üseyin, Avci, Engin, Özyurt, Fatih, 2020. White blood cells detection and classification based on regional convolutional neural networks. Med. Hypotheses 135, 1–18. https://doi.org/10.1016/j.mehy.2019.109472.

Ren, Shaoqing, He, Kaiming, Girshick, Ross, et al., 2016. Faster R-CNN: Towards Real-time Object Detection With Region Proposal Networks arXiv:1506.01497v3.

Tan, Mingxing, Pang, Ruoming, Le, Quoc V., 2020. EfficientDet: Scalable and Efficient Object Detection arXiv:1506.01497v3.

Duan, Kaiwen, Bai, Song, Xie, Lingxi, et al., 2019. CenterNet: Keypoint Triplets for Object Detection arXiv:1904.08189v3.

Yamashita, R., Nishio, M., Do, R.K.G., et al., 2018. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629. https://doi.org/10.1007/s13244-018-0639-9.

Cheng, Sun, Suhua, Yang, Shaofeng, Jiang, 2019. Improved faster RCNN for white blood cells detection in blood smear image. 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI) 2019. https://doi.org/10.1109/ICEMI46757.2019.9101445, 1677 - 1682.

Tiwari, Prayag, Qian, Jia, Li, Qiuchi, et al., 2018. Detection of subtype blood cells using deep learning. Cogn. Syst. Res. 52, 1036–1044. https://doi.org/10.1016/j.cogsys.2018.08.022.

Szegedy, Christian, Ioffe, Sergey, Vanhoucke, Vincent, et al., 2016. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning arXiv:1602.07261v2.

Tan, Mingxing, Le, Quoc V., 2020. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks arXiv:1905.11946v5.

Law, Hei, Deng, Jia, 2019. CornerNet: Detecting Objects As Paired Keypoints arXiv:1808.01244v2.

ALL-IDB Acute Lymphoblastic Leukemia Image Database for Image Processing, Department of Computer Science - Università degli Studi di Milano https://homes.di.unimi.it/scotti/all/.

Da, Wang, Hwang, Maxwell, Jiang, Wei-Cheng, et al., 2021. A deep learning method for counting white blood cells in bone marrow images. BMC Bioinformatics 22, 94. https://doi.org/10.1186/s12859-021-04003-z.

Xia, Tiancheng, Jiang, Richard, Fu, Yong Qing, et al., 2019. Automated blood cell detection and counting via deep learning for microfluidic Point-of-care medical devices. In: IOP Conference Series: Materials Science and Engineering, 646. https://doi.org/10.1088/1757-899X/646/1/012048.

Hung, Jane, Ravel, Deepali, Lopes, Stefanie C.P., et al., 2019. Applying Faster R-CNN for Object Detection on Malaria Images arXiv:1804.09548v2.

Praveen, Nalla, Punn, Narinder Singh, Sonbhadra, Sanjay Kumar, et al., 2021. White Blood Cell Subtype Detection and Classification arXiv:2108.04614v3.

Abou El-Seoud, S., Siala, M.H., McKee, G., 2020. Detection and classification of white blood cells through deep learning techniques. International Journal of Online and Biomedical Engineering (iJOE) 16 (15), 94–105. https://doi.org/10.3991/ijoe.v16i15.15481.

Soto-Ayala, L.C., Cantoral-Ceballos, J.A., 2021. Automatic blood-cell classification via convolutional neural networks and transfer learning. Ieee Lat. Am. Trans. 19 (12), 2028–2036. Retrieved from. https://latamt.ieeer9.org/index.php/transactions/article/view/5051.