Malaria Cells Detection in Microscopy Images using Faster R-CNN based on Curriculum Learning

Salman Alsalman[1,2] S.ALSALMAN@UKY.EDU
[1] *University of Kentucky, Lexington, KY, United States*
[2] *Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia*

Abstract

Malaria is a disease that takes many lives all over the globe. Even though it is a deadly disease, it is still curable when it is diagnosed during early stages of the infection. The need to have reliable deep-learning algorithms for this matter is crucial. In this paper, we propose an approach to malaria cell detection in microscopic images using one of the benchmarks object detection algorithms, with employing some curriculum learning strategies. We perform our experiment on a dataset that consists of images that have been taken from multiple microscopes and multiple magnification levels. Our suggested method leverages curriculum learning to train detector models by gradually introducing more complex samples in order to improve detection performance. The implementation of the code can be found here: https://github.com/alsalman-s/CS685_Project

Keywords: malaria cell detection, Faster R-CNN, curriculum learning, microscopic images, deep learning, object detection, medical imaging

1. Introduction

Malaria remains a significant global health issue, with early detection and diagnosis being crucial for effective treatment (World Health Organization, 2020). Automated malaria cell detection in microscopic images can enhance diagnostic efficiency and accuracy. Recent advancements in deep learning, specifically in object detection algorithms, have shown promising results for such tasks. In this work, we aim to improve malaria cell detection by incorporating curriculum learning into the training process of a Faster R-CNN model (Ren et al., 2016). The primary goal is to design and implement an automatic malaria cell detection and classification system using the Faster R-CNN object detection algorithm. We propose to apply a curriculum learning approach to the training process, expecting to improve the model's performance compared to existing methods.

2. Related Work

Previous work on malaria cell detection has explored various state-of-the-art algorithms. The Faster R-CNN algorithm has demonstrated superior performance in terms of accuracy and mean average precision (mAP) (Sultani et al., 2022). The idea of curriculum learning is to expose the model to easy dataset samples and follow it with more complex ones (Soviany et al., 2022). Different studies on curriculum learning have been conducted to improve the performance of machine learning models. However, there is no specific way to apply it to one's problem. Thus, each study represents a different strategy based on its requirements. The common ground among them is to find a way to introduce different input samples during the training process of a model (Soviany et al., 2022).

3. Method

Our methodology is comprised of two primary components: the Faster R-CNN object detection model and the curriculum learning strategy.

The Detector: We employ the Faster R-CNN architecture utilizing the established deep learning library, Detectron2 (Wu et al., 2019). Figure 1 shows the architectural design and pipeline of the chosen detector for our experimentation. The Faster R-CNN detector encompasses a backbone convolution network layer, succeeded by a feature maps layer that functions as a region proposal network for the bounding boxes. We opted for (faster-rcnn-R50-FPN-3x) from Detectron2 as our pre-trained model with a base learning rate of 0.001 and 4000 maximum solver iterations. The model is 50 layers deep and pre-trained with the COCO dataset (Lin et al., 2014). The losses the model uses are the same as the original developed Faster R-CNN model in (Ren et al., 2016). In the Region Proposal Network (RPN), as shown in Figure 1, it applies a regression loss function computed with smooth L1 loss for bounding-box coordinates, and a classification loss function that uses binary cross-entropy loss to distinguish foreground objects from the background.

Regarding the curriculum learning strategy: We organize the training samples based on factors such as the microscope type, including high-cost microscope (HCM) and low-cost microscope (LCM) images, as well as the magnification levels, specifically 100x, 400x, and 1000x. To accomplish our objective, we utilize the M5-Malaria dataset (Sultani et al., 2022), which comprises 7,542 multi-cell microscopic images. The dataset is segregated into two microscope types and three magnification levels.

The model is initially trained on the least challenging samples and subsequently advances to more intricate samples as its performance improves. Consequently, we devised two curriculum learning strategies: First strategy, training the model on the LCM-1000x images then LCM-400x images. Second strategy, training the model on the HCM-1000x images then HCM-400x images. We did not use any of the 100x magnification levels because after our first experiments on the Faster R-CNN model with the regular learning method, as shown in Table 1, we noticed that the model did not perform as we hoped for. That is why we decided to focus our curriculum learning strategy toward 400x and 1000x samples only. In addition, we tested the models that were trained using the curriculum learning against LCM images only, since the main motivation of this research is to find a way to lower the cost of detecting malaria cells from blood smear images.

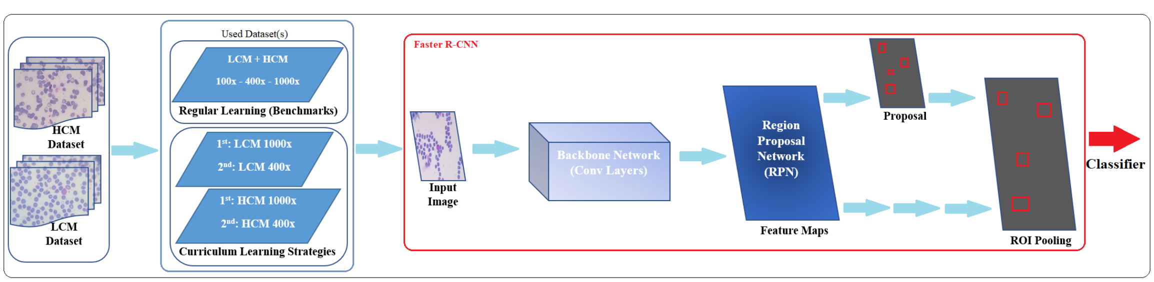


Figure 1: Detector Architecture Design and Pipeline

4. Experiments and Results

Our experimental procedure comprises three distinct stages: preprocessing the dataset, implementing and training the detector, and evaluating the detector.

Dataset preprocessing: As we mentioned in the method section, we used the M5-Malaria dataset (Sultani et al., 2022). The dataset consists of images that have been taken using two different microscopes, HCM and LCM, under three different magnification powers: 100x, 400x, and 1000x. To prepare and preprocess the dataset, we organize the M5-Malaria dataset into training, validation, and test sets, ensuring a balance between the various image types and magnification levels. We use the same split as used in the original paper (Sultani et al., 2022). They divided the dataset into 66.5% training, 3.5% validation, and 30% testing sets. The annotation files are initially in PASCAL Visual Object Classes (VOC) format (Everingham et al., 2010); hence, we must convert them to COCO format to facilitate their use in the Faster R-CNN detector from Detectron2. Then we implement and train the model.

Model Implementation and Training: As shown in Figure 1, the dataset was used in three different experiments. First, for the regular learning strategy. The Faster R-CNN was trained and tested without applying any modification on the dataset. Six models were implemented, trained, and tested. Table 1 shows the results of these models and their tests. Table 1 presents the outcomes of executing the model without employing the curriculum learning strategy. We conducted 36 distinct runs, training the model with different microscopes and magnification levels. Subsequently, we identified the highest results in terms of mean average precision (mAP%). In general, the results did not meet our expectations; nevertheless, the objective of this paper is to apply the curriculum learning strategy to enhance the results. Both HCM and LCM at 100x magnification yielded the lowest scores in terms of mAP, which is why we will not consider them during our experiment with curriculum learning. Evidently, the highest score for HCM occurs when training the model at 1000x and testing it with HCM at 1000x, which is 22.51. However, to achieve a more practical detector, we need to test the model on LCM during the second experiment for curriculum learning.

Training Dataset		Test Dataset					
		Microscope Type					
		LCM			**HCM**		
Microscope Type	**Magnification's Level**	**100x**	**400x**	**1000x**	**100x**	**400x**	**1000x**
LCM	**100x**	0.33	**5.72**	2.46	3.78	**7.46**	5.14
	400x	0.03	**12.17**	11.75	0.22	**5.85**	2.71
	1000x	0.01	6.10	**15.36**	0.31	**13.25**	6.64
HCM	**100x**	0.90	0.90	**3.47**	**7.42**	1.12	1.10
	400x	0.01	7.01	**10.65**	0.52	19.01	**20.49**
	1000x	0.00	0.30	**7.25**	0.00	6.71	**22.51**

Table 1: Results of using Faster R-CNN detector, trained and tested on multiple microscope types and magnification levels.

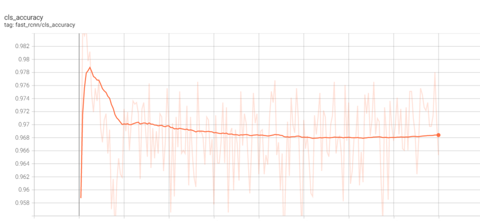
Table 2 displays the results of applying the curriculum learning approach. In comparison to the regular training approach, testing the model trained with LCM1000x-400x demonstrates a substantial improvement of approximately 26% when the model was tested on

LCM at 400x. Additionally, Table 2 indicates a slight improvement of approximately with the other model that was trained with HCM1000x-400x.

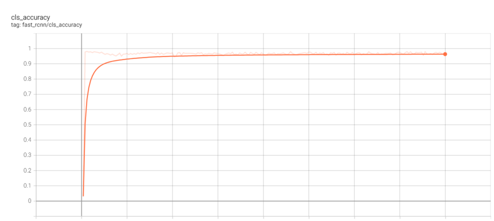
Training Dataset		Test Dataset: LCM, Magnification Level	
Microscope Type	**Curriculum Learning Strategy (Based on Magnification Level)**	**400x**	**1000x**
LCM	**{1000x → 400x}**	**15.64**	13.01
HCM	**{1000x → 400x}**	7.24	**11.11**

Table 2: Results of applying different curriculum learning strategy using Faster R-CNN detector trained and tested on multiple microscope types and magnification levels.

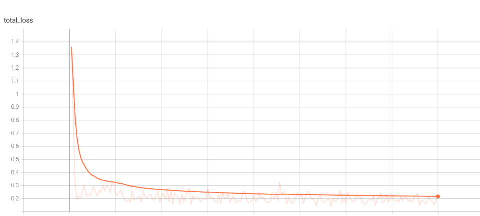
Results analysis: In general, the results of our experiments did meet our expectations; even though, we succeeded in improving the detector when it was tested with LCM-400x by around 26% as shown in Table 2. In addition, we want to bring attention to the fact that while training the model and validating it, the results were much higher than testing it with the testing datasets. Figure 2 shows the classification accuracy curve and the total loss curve for both the models that were trained using the regular learning approach and the curriculum learning strategy. Figure 3 presents an example prediction visualization of running the Faster R-CNN detectors with and without curriculum learning compared to ground truth images.



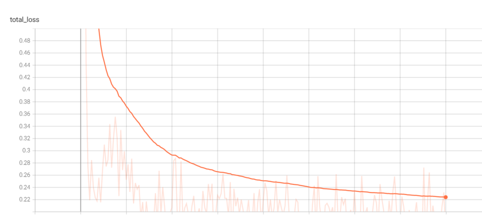
(a) CLS Accuracy (Regular Learning)



(b) CLS Accuracy (Curriculum Learning)

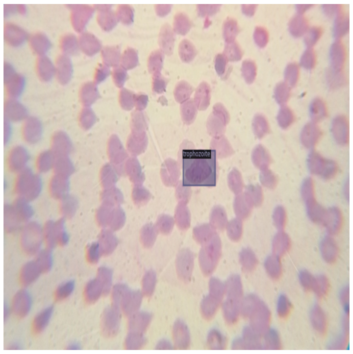


(c) Total Loss (Regular Learning)

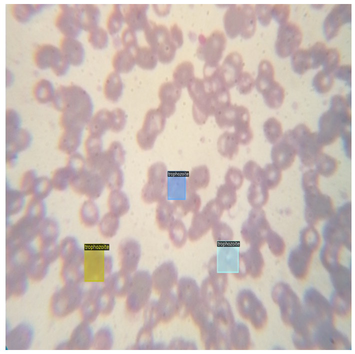


(d) Total Loss (Curriculum Learning)

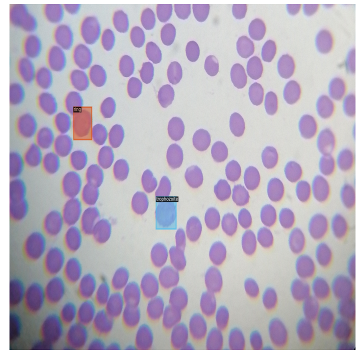
Figure 2: Faster R-CNN predictor's classification accuracy and total loss using regular learning and curriculum learning.



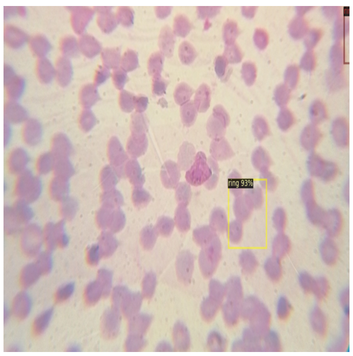
(a) Ground Truth Sample 1



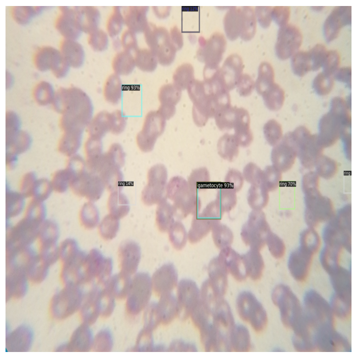
(b) Ground Truth Sample 2



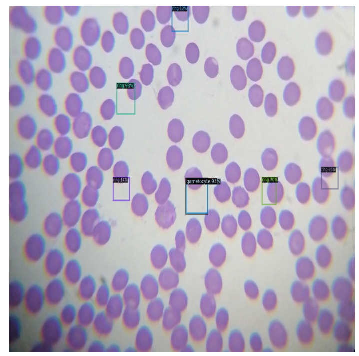
(c) Ground Truth Sample 3



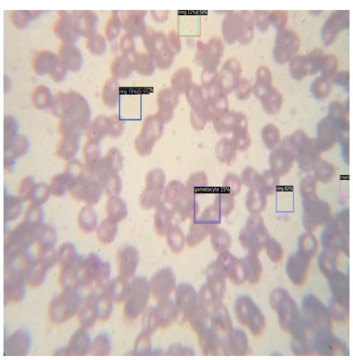
(d) Prediction on Sample 1



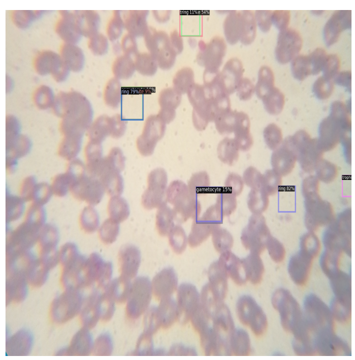
(e) Prediction on Sample 2



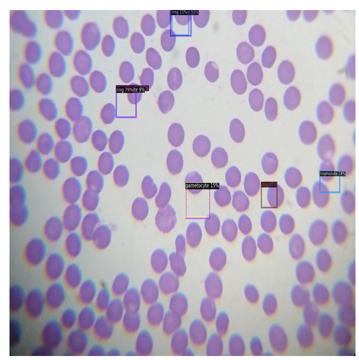
(f) Prediction on Sample 3



(g) Prediction using Curriculum Learning on Sample 1



(h) Prediction using Curriculum Learning on Sample 1



(i) Prediction using Curriculum Learning on Sample 1

Figure 3: Results of running Faster R-CNN predictor. First row, for ground truth images. Second row, for the prediction images from the model with regular learning. Third row, for the prediction images from the model with curriculum learning.

5. Conclusion

In this paper, we proposed a method for malaria cell detection in microscopic images using the Faster R-CNN object detection algorithm and curriculum learning. Our method aims to improve the detection performance by gradually introducing more complex samples during the training process. The results of our experiments have slightly improved the baseline results, contributing to more accurate and efficient malaria diagnosis. More experiments and fine-tuning need to be done in order to achieve better improvements in the results.

Acknowledgments

This paper describes research done at the University of Kentucky in the department of Computer Science. I am thankful to Dr. Abdullah-Al-Zubaer Imran for his guidance and cooperation.

References

Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John M. Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vis.*, 88 (2):303–338, 2010. URL http://dblp.uni-trier.de/db/journals/ijcv/ijcv88.html#EveringhamGWWZ10.

Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312, 2014. URL http://arxiv.org/abs/1405.0312.

Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks, 2016.

Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum learning: A survey, 2022.

Waqas Sultani, Wajahat Nawaz, Syed Javed, Muhammad Sohail Danish, Asma Saadia, and Mohsen Ali. Towards low-cost and efficient malaria detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 20687–20696, June 2022.

WHO World Health Organization. World malaria report 2020 : 20 years of global progress challenges. https://www.who.int/publications/i/item/9789240015791, 2020.

Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick. Detectron2. https://github.com/facebookresearch/detectron2, 2019.