A new approach for microscopic diagnosis of malaria parasites in thick blood smears using pre-trained deep learning models

Rose Nakasi¹ · Ernest Mwebaze² · Aminah Zawedde¹ · Jeremy Tusubira¹ · Benjamin Akera¹ · Gilbert Maiga¹

Received: 6 December 2019 / Accepted: 1 June 2020 / Published online: 20 June 2020
© Springer Nature Switzerland AG 2020

Abstract

One of the deadly endemic diseases in sub-Saharan Africa is malaria. Its prevalence is promoted by lack of sufficient expertise to carry out accurate and timely diagnosis using the standard microscopy method. Where lab technicians are available, the results are usually subjective due to variations in expert judgement. To address this challenge, prompt interventions to improve disease control are needed. The emerging technologies of machine learning that can learn complex image patterns have accelerated research in medical image analysis. In this study, on a dataset of thick blood smear images, we evaluate and compare performance of three pre-trained deep learning architectures namely; faster regional convolutional neural network (faster R-CNN), single-shot multi-box detector (SSD) and RetinaNet through a Tensorflow object detection API. Data augmentation method was applied to optimise performance of the meta architectures. The possibility for mobile phone detector deployment was also investigated. The results revealed that faster R-CNN was the best trained model with a mean average precision of over 0.94 and SSD, was the best model for mobile deployment. We therefore deduce that faster R-CNN is best suited for obtaining high rates of accuracy in malaria detection while SSD is best suited for mobile deployment.

Keywords Deep learning · Malaria detection · Thick blood smear · Mobile detector

1 Introduction

Malaria is one of the top causes of death in sub-Saharan Africa [1]. Of the 438,000 malaria cases registered, an estimated 92% resulted in deaths; two-thirds of which occurred among children under 5 years of age [1]. In highly endemic areas like Uganda, Malaria is the leading cause of death accounting for over 27% of loss of lives [2]. Therefore, detecting malaria parasites is key to malaria diagnosis as it contributes immensely to the prevention and treatment of the deadly disease [3].

The conventional approach for the diagnosis of malaria is microscopy [3, 4] that requires interpretation of the results by a skilled technician. Using this approach, a

sample of blood is drawn from a patient and smeared on a glass slide stained with Giemsa. This blood sample can be either thin or thick. A thick smear is more efficient for parasite detection with an 11 times higher sensitivity rate [5]. Although higher accuracy has been recorded with thick blood smears, most of the research in automation of malaria diagnosis has been done with thin blood smears thus calling for more research on thick smears [6].

Furthermore, there are few skilled lab technicians in the highly endemic areas to effectively interpret the microscopy diagnosis results. When they do, the interpretation of the results is affected by the level of human variability in observation of the parasites which leads to bias in

Supported by SIDA.

✉ Rose Nakasi, g.nakasirose@gmail.com | ¹College of Computing and Information Sciences, Makerere University, Kampala, Uganda. ²Makerere AI Research, Kampala, Uganda.

SN Applied Sciences (2020) 2:1255 | <https://doi.org/10.1007/s42452-020-3000-0>

the results [7]. This can however be improved through automation.

Automation using image analysis in the field of malaria microscopy with thick blood smears is mainly by use of shallow/traditional machine learning methods that require expertise and skills to conduct hand engineering of malaria parasite features [8]. This has however been improved by recent developments in research using deep learning models that have provided more accurate results that do not require hand engineering of the features [9]. This study extends work done in [8] which investigated the use of a customised deep learning convolutional neural network (CNN) model for object detection of malaria parasites in thick blood smears, Tuberculosis bacilli and intestinal parasites. However, customised CNN requires back and forth fine-tuning of the model layers and huge volumes of annotated data for improved accuracy which may not be available for medical image analysis [10, 11].

A recently introduced technique that can mitigate some of these challenges is transfer learning. In transfer learning, representations learned on a large dataset can be transferred to a smaller model from a totally different data domain. The representations learned from the bigger dataset can be viewed as extracted features or good initialisations for the smaller model. In this study, we adopt and implement the use of pre-trained transfer learning meta architectures and feature extractors [12] and apply them to the task of malaria pathogen detection from thick blood smears.

This study investigates the feasibility of using a new approach based on selected pre-trained deep learning models and evaluates their performance towards malaria parasite detection in thick blood smear with a small image dataset. A comparative analysis of the performance and suitability for mobile deployment for the different models is presented.

The rest of the article is organised as follows: a short review of machine learning models for malaria pathogen detection in microscopy diagnosis is presented in Sect. 2. In Sect. 3, we discuss materials, experiments and methods. Results and a brief discussion are presented in Sects. 4 and 5, respectively. We conclude by studying the performance, strengths and failures of the models in Sect. 6.

2 Related work

Previous work on computational diagnosis of Malaria has been reviewed [6] in which the authors recommend further improvements in detection accuracy. However, most of these attempts reviewed are based on hand engineering of feature extraction techniques which requires skills and expertise [8]. With a good feature extraction approach,

detection accuracy is improved compared to conventional microscopy [13].

Pre-trained deep learning models have the ability to deal with heterogeneity of different datasets and cater for scenarios where data are limited especially in medical image analysis [14]. Existing huge computer vision datasets like ImageNet provide the base datasets from which transfer learning pre-trained models can be trained. Pre-trained models in this case serve as a feature extractor to a new model trained on a dataset with fewer images, and can be applied to different tasks of computer vision to improve performance [15].

A study by [16] applied pre-trained deep learning models on the classification task of differentiating parasitised and uninfected red blood cells in thin blood smear images. The study compared different models including AlexNet, VGG-16, Xception, ResNet-50 and DenseNet-121 for the classification task. A level-set-based algorithm to detect and segment the RBCs was applied and attained an optimum accuracy score of 0.9. A similar task of classifying parasitised and uninfected red blood cells in thin blood smears using pre-training deep models like LeNet, AlexNet and GoogleNet is also given in [17]. The authors report a classification accuracy of above 95% for all deep learning models, a superior performance to 92% achieved by support vector machine (SVM).

Although the reported performance in the literature is outstanding, previous research on transfer learning based on deep learning models for malaria diagnosis has mostly been evaluated on classification tasks [10] of thin blood smear images. Thin blood smear images, unlike the thick blood smear, are less sensitive and usually associated with missing malaria parasite due to low parasitemia [6]. Scaling up the approach of transfer learning for object detection [18] in thick blood smears would provide more insights and considerable results for parasite detection [7].

3 Experiments

3.1 Data collection

The data collection procedure follows a similar protocol from previous work on microscopic malaria diagnosis [8]. In this study, 643 images of thick malaria blood smears with dimensions 750×750 pixels were collected from Mulago Referral Hospital in Uganda. These were taken using a smart phone camera that was attached on the microscope eyepiece using a 3D printable adapter as described in [8]. In order to implement our approach, it was necessary to obtain the ground truth of the images, and this was done with the help of expert lab technicians in Mulago Hospital who manually drew bounding boxes

around the malaria parasites using an open source tool Labellmg [19]. The annotations were saved in the Pascal VOC format [20] with an image and its corresponding xml file defining the annotations of objects in the image. The corresponding xml file in this case contained coordinates of bounding boxes of parasites labelled in the images. Every image considered contained at least one bounding box signifying presence of a parasite in the image. The annotated images were randomly split into a train and test set in the ratio of 9:1, respectively, and the combined dataset was encoded in the record format that is optimised for processing with Tensorflow [21].

3.2 Data pre-processing

In the data pre-processing, extensive data augmentation was applied. The primary purpose of the augmentation was to induce variation in the images and forcing the models during training not to get stuck in local minimums thus increasing generalisation as the models inherently are forced to learn a broader spectrum of spatial relationships. Subsequently, we adopted the default random horizontal flip augmentation strategy which flips an image horizontally with a probability, p and comes embedded as a helper function in the pre-processing stage of the Tensorflow Object Detection API for the models used [22].

3.3 Model selection and training

We used different oriented object detection algorithms in our experiments. There are various state-of-the-art algorithms that have been successfully used for computer vision in the literature [12]. Our choice of pre-trained deep learning models was to leverage the advantages of transfer learning such as shorter training time and better generalisation even for small datasets. For this study, we specifically based our experiments on three model architectures: (a) faster regional convolutional neural networks (faster R-CNN) [23], (b) single-shot detector (SSD) and (c) RetinaNet.

These models are object detection models, they detect and recognise the class and location of the parasites in the image. They are accurate in localising the box containing the object and show the level of confidence for detection of each target parasite. Moreover, transfer learning is a technique to translate learned weights from one domain to another and in this study, we used the weights of three models pre-trained on the Microsoft Common Objects in Context (COCO) dataset and apply them to our medical dataset. The models were developed using the Tensorflow Object Detection API [24] which provides multiple implementations of the current state-of-the-art deep learning models along with their learned weights.

3.3.1 Faster R-CNN

Faster R-CNN is one of the stellar object detection algorithms. It involves the generation of region proposals by selective search and then a CNN-based network is adapted for both classification of object class and detection of the location of the bounding box for the target object. This makes the object detection task faster. The algorithm has a region of interest (RoI) layer that is used for extraction of features and to classify objects with bounding-box regression to obtain the estimated targets. The Tensorflow Object Detection API [24] provides implementations of the faster R-CNN model built on both ResNet50 and ResNet101 architectures.

3.3.2 Single-shot multibox detector (SSD)

With SSD, only one shot is taken to detect multiple target objects within an image [25]. The choice of SSD in this study was due to its faster inference and its capability for mobile deployment. It discretises the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. At prediction time, the network generates scores for the presence of each object category in each default box and produces adjustments to the box to better match the object shape.

Additionally, the network combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes. SSD is simple relative to methods that require object proposals because it completely eliminates proposal generation and subsequent pixel or feature re-sampling stages and encapsulates all computation in a single network. This makes SSD easy to train and straightforward to integrate into systems that require a detection component. The Tensor Flow Object Detection API provides multiple implementations of the SSD model built on the MobileNetV2 and Inception architectures.

3.3.3 RetinaNet

RetinaNet is a single, unified network composed of a backbone network and two task-specific sub-networks [26]. The choice of RetinaNet in this study was due to its published attributes; being able to match the inference speed of SSD MobileNet models and give predictions with high accuracy similar to faster R-CNN. The backbone is responsible for computing a convolutional feature map over an entire input image. The first subnet performs convolutional object classification on the backbone's output and the second subnet performs convolutional bounding-box regression. Tensorflow Object Detection API [24] provides

implementations of the RetinaNet model built on the SSD ResNetFPN architecture.

3.4 Implementation details

All the above models were trained and tested on an Ubuntu system with a 5th Gen I7 Intel Core processor and 16 GB RAM, a GPU-enabled Nvidia GTX 1060 Graphical Processing Unit (GPU) with 6 GB RAM, Python 2.7, with Tensorflow back-end.

Training the models was done using the open-source library Tensorflow [21]. This enables convenient transfer learning from pre-trained models. Tensorflow also provides an API to predict bounding boxes for detection of objects. The object detection models were acquired from the Tensorflow Model Zoo [22]. The model zoo is a holding repository for all current models in both object detection and image classification and pre-trained on a COCO dataset [22]. We trained our models for up to 50,000 time steps to enable us monitor all the variations that may occur during the training over time.

The Tensorflow Object detection API also enables specification of different training options and parameters by setting of options in a configuration file. The training parameters used for each of the models are specified in Table 1.

For training of machine learning models, having a large batch size is usually desirable, although for this experiment, we used a batch size of 1 due to the limited memory resources available to execute the training job. While conducting experiments, it was observed that the SSD models detected bounding boxes and peaked their learning quite early in the training at about 10,000 time steps compared to the other models. To counter this and to reduce the rate of optimisation, the learning rate of the SSD models was set much lower than that of faster-RCNN and RetinaNet. The rest of the configuration parameters were left default as set in the Tensorflow object detection API [22].

Evaluation metrics used are the standard COCO metrics, the mean average precision (mAP@0.5) [27] for measuring how well the object detection works. Precision and recall were used to show how well the algorithm predicts the presence of parasites in an image.

4 Results

4.1 Inference time and mAP performance evaluation

To ascertain the level of performance of the state-of-the-art pre-trained models used in this study for malaria parasite detection in thick blood smear, the models were trained and evaluated using mAP and the inference time in milli seconds (ms).

Table 2 shows the results of mAP performance across all models in the study. The results indicate higher performance for the two-stage detectors (faster R-CNNs) over the one-stage detectors. Two-stage detectors attain more than 90% mAP accuracy compared to one-stage detectors (SSD and RetinaNet). This is majorly because two-stage detectors are able to localise objects of interest resulting into a higher Intersection over Union (IoU) thus a higher mAP. For one-stage detectors like SSD MobileNetV2 used in this study, we realise that it performs with even a lower mAP compared to SSD InceptionV2. This is because SSD MobileNet uses depth-wise convolutions which significantly reduce the number of parameters, and consequently this affects its performance.

Additionally as seen in Table 2, the inference time of the faster R-CNN models is higher than that of other models. This indicates that two-stage detectors like the faster R-CNN are computationally more expensive in

Table 2 mAP@0.5 and inference time performance for pre-trained models used in the study

Algorithm	Steps	mAP@0.5	Inference time (ms)
Faster R-CNN (ResNet101)	8960	0.9448	42.80
Faster R-CNN (ResNet50)	15,463	0.9306	18.93
Faster R-CNN (InceptionV2)	23,937	0.9254	6.79
SSD (MobileNetV2)	1091	0.3119	5.31
SSD (InceptionV2)	2063	0.8397	2.24
RetinaNet (SSD ResNetFPN)	49,500	0.7746	13.79

The bold values show the highest scores for each column

Table 1 Training parameters used for each of the models

Parameter	Faster RCCN	SSD	RetinaNet
Learning rate	0.002	0.0002	0.002
Batch size	1	1	1
Input image size	750 × 750	750 × 750	750 × 750
Optimiser	Momentum optimiser	Momentum optimiser	Momentum optimiser
Metrics	COCO metrics	COCO metrics	COCO metrics
IOU threshold	0.5	0.5	0.5

comparison with the others, which has implications for deployment on devices such as mobile phones.

To fine-tune the network to detect the spatial features of the malaria parasites, all models were trained for 50,000 time steps though each model had its own level of attained optimal performance (lowest loss and highest mAP) at respective steps indicated in Table 2. These are steps at which the model converges with early stopping.

4.2 Precision and recall performance

To evaluate the performance of the each model architecture used in this study, the performance was further evaluated in terms of precision and recall on a validation set of 64 images with 373 parasites. Table 3 gives the results of these experiments. Generally, we observe that the two-stage detectors exhibit a low precision as compared to one-stage detectors. Two-stage detectors were also associated with a higher number of false positives. One reason we conjecture for this is probably due to poor or missing annotations of some images by the expert lab technicians. This implies that some of the false positives were actual true positives thus biasing faster R-CNN model performance.

Table 3 Malaria parasite precision and recall detection results for different architectures

Algorithm	Steps	Precision	Recall
Faster R-CNN (ResNet101)	8960	0.7791	0.8981
Faster R-CNN (ResNet50)	15,463	0.7461	0.9062
Faster R-CNN (InceptionV2)	23,937	0.7229	0.9303
SSD (MobileNetV2)	1091	0.7978	0.3914
SSD (InceptionV2)	2063	0.9150	0.3753
RetinaNet (SSD ResNetFPN)	49,500	0.8697	0.6086

The bold values show the highest scores for each column

As seen in Table 3, the performance of the faster RCCN and SSD-based deep learning models varied to a less extent in terms of precision and to a great extent in terms of recall. The best performing SSD ResnetFPN (RetinaNet) had a precision of 0.9385 and a recall of 0.3271, whereas the better performing faster R-CNN had a precision of 0.7791 and a recall of 0.8981. The biggest difference between the models was in terms of false negatives, which consequently lead to a large difference in recall as well. Additionally, the models suffered from many false positives, which lead to the differences in the precision and again, this could have majorly been caused by poor or missing annotations.

The combined high values of recall and remarkable precision by the faster RCCN models clearly suggest that faster RCCN models could be used successfully in the malaria parasite detection task as compared to other models for transfer learning.

4.3 Parasite detection results

The performance of the algorithms was evaluated and tested based on the bounding box regression and class score of each malaria pathogen in thick blood smear images. Figure 1 shows examples of how each model (faster RCCN, SSD and RetinaNet) was able to detect the malaria pathogens and their respective locations in the test image.

The predicted results were compared with the ground truth (annotated image) using mAP@0.5. In this study, we found the best detection results are generated with faster R-CNN. The transfer learning approach used in this study promises a fast and efficient way for malaria pathogen detection since it provides comparable performance with a limited number of training examples.

Fig. 1 Results for detection with faster R-CNN (b), SSD (c) and RetinaNet (d) for a sample image (a). The black circles indicate False Negatives and the black rectangles in (b–d) indicate false positives and likely mis-annotations

4.4 Deployment on a mobile smartphone

This work is part of a larger project to enable smartphone-based diagnosis of malaria by mounting the smartphone on the eye piece of a microscope. In several health centres in Uganda where the data were collected and broadly in Africa, the number of microscopes out-numbers the available lab technicians. Automating this step of testing means quicker diagnoses and better health outcomes particularly for under-resourced rural health facilities.

The results from the experiments and the current smartphone infrastructure support SSD algorithms because of faster inference times and a small footprint in terms of size of model as well as amount of RAM required to run such a model. The relatively higher precision of SSD models means they are possibly suitable candidate for deployment. Given that a slide is comprised of multiple images that are processed through the model, a high precision ensures that we get a reliable signal to make the next decision in the diagnosis of malaria which is usually the presence or absence of malaria and the severity of disease. Future work will focus on improving the recall of these models.

5 Discussion of results

In this research, we have proposed the use of the current state-of-the-art deep learning model architectures for the task of malaria pathogen detection in thick blood smear images. This introduces an applicable solution for point of care microscopic disease diagnosis which shows both the class and location of the pathogens and the degree of detection confidence. The experiments carried out also highlight the possible trade-offs when extending this work to deployment on smartphones.

Our goal was to investigate if the state-of-art pre-trained deep learning models for transfer learning can be applied to malaria pathogen detection and if so which is the most suitable architecture for this task. Thus, the experimental results and comparisons between various deep learning architectures have demonstrated how transfer learning-based detection can be translated to successfully detect malaria pathogens in thick blood smear images.

Based on the experiments, it is also evident that there is indeed a speed vs accuracy trade-off. We observe that faster R-CNN models offer the best accuracy when given enough resources resulting in better performance, though it takes a lot longer while training and gives a high inference time in comparison with SSD and RetinaNet. It is worth noting that much as SSD models are a bit unstable with lower accuracy, they are much faster in detection and

smaller than all the other models which accounts for their ease of deployment on a low-resourced mobile phone.

6 Conclusion and future work

In this paper, we have presented an approach for malaria parasite detection in thick blood smear images using state-of-the-art pre-trained deep learning models for transfer learning. We have realised successful implementation of the different models on a small dataset of thick blood smear images with reasonable and promising results of mAP, inference time and precision accuracy. Moreover, with some models like SSD MobileNet promising a prototype for scalability and accessible deployment for use in a real-time environment on devices with limited computational power (mobile smart phones) while remaining moderately accurate and reliable for the end user. Our results thus conclude that the state-of-the-art pre-trained deep learning models have capability for providing accurate and speedy detection with a reasonable level of confidence with possible extensions to deployment on smartphones.

On the other hand, we noticed that although the proposed approach of using the current state-of-the-art deep learning meta architectures through transfer learning shows an outstanding performance on the evaluated cases, it also presents detection challenges in some cases, especially where the detection confidence is below 60% as compared to our earlier hand-engineered custom deep learning model with a precision accuracy of above 90% [8]. Further more, there is some pattern variations between ground truth annotations with what was predicted resulting in some false positives or lower average precision. While this could be due to poor annotations, in future work, we want to investigate these issues a lot more.

Additionally in future studies, for better results, a customised and modified pre-trained algorithm on a similar dataset like the one used in this study may need to be developed to fine tune the layers to reduce on the complexity of the default models.

Acknowledgements The authors thank Ministry of Health, Uganda for providing authorisation to use the data. They also acknowledge the laboratory Technicians in Mulago Referral Hospital, Uganda especially Dr. Alfred Andama and Vincent Wadda for granting us the medical and technical support in data acquisition, annotation and interpretability of results.

Funding The first author of this study is funded in part by the Swedish International Development Cooperation Agency (SIDA) and Makerere University under SIDA Contribution No. 51180060 as student STD0404. The Grant is part of the European and Developing Countries Clinical Trials Partnership (EDCTP2) programme supported by

the European Union. The authors would like to appreciate for the funding.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical standard We comply with the ethical standards through possession of a working IRB (Institutional Review Board) letter to access data from health facilities.

References

- WHO: World malaria report. World Health Organisation, Geneva, Switzerland (2016)
- Ministry of Health, U.: Malaria bulletin. Malaria Control Bull 1(1) (2016)
- Beadle C, Long G, McElroy P, Hoffman S, Weiss W, Oloo A (1994) Diagnosis of malaria by detection of *Plasmodium falciparum* hrp-2 antigen with a rapid dipstick antigen-capture assay. Lancet 343(8897):564–568
- WHO: The role of laboratory diagnosis to support malaria disease management: focus on the use of rapid diagnostic tests in the areas of high transmission. WHO report (2004)
- Dowling M, Shute G (1966) A comparative study of thick and thin blood films in the diagnosis of scanty malaria parasitaemia. Bull World Health Organ 34:249
- Rosado L, Correia da Costa L, Elias J, Cardoso D (2016) A review of automatic malaria parasites detection and segmentation in microscopic images. Anti-Infect Agents 14(1):11–22
- Poostchi M, Silamut K, Maude R, Jaeger S, Thomas G (2018) Image analysis and machine learning for detecting malaria. Transl Res. 194:36–55
- Quinn J, Nakasi R, Mugagga P, Byanyima P, Lubega W, Andama A (2016) Deep convolutional neural networks for microscopy-based point of care diagnosis. In: Proceedings of international conference on machine learning for health care, vol 50
- Bengio Y, Vincent P, Courville A (2013) Representation learning: a review and new perspectives. IEEE Trans Pattern Anal Mach Intell 35(8):1798–1828. <https://doi.org/10.1109/tpami.2013.1206.5538>
- Najafabadi M, Villanustre F, Khoshgoftaar T, Seliya N, Wald R, Muharemagic E (2015) Deep learning applications and challenges in big data analytics. J Big Data. <https://doi.org/10.1186/s40537-014-0007-7>
- Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge
- Weiss K, Khoshgoftaar T, Wang D (2016) A survey of transfer learning. J Big Data 3:9. <https://doi.org/10.1186/s40537-016-0043-6>
- EliteDataScience: Modern machine learning: strengths and weaknesses. elitedatascience.com (2017)
- Ravishankar H, Sudhakar P, Venkataramani R, Thiruvankadam S, Annangi PBN, Vaidya V (2017) Understanding the mechanisms of deep transfer learning for medical images. GE Global Research, Bangalore, India
- Razavian A, Azizpour H, Sullivan J, Carlsson S (2014) CNN features off-the-shelf: an astounding baseline for recognition. IEEE computer society conference on computer vision and pattern recognition workshops, Piscataway
- Rajaraman S,K,S, Antani SM, Silamut K, Hossain MA, Richard J, Maude R, Jaeger S, Thomas GR (2018) Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ 6:e4568. <https://doi.org/10.7717/peerj.4568>
- Dong Y, Jiang Z, Shen H, David PW, Williams LA, Reddy V, Benjamin W, Bryan A (2017) Evaluations of deep convolutional neural networks for automatic identification of malaria infected cells. In: IEEE EMBS international conference on biomedical and health informatics, BHI 2017. IEEE, Piscataway, pp 101–104. <https://doi.org/10.1109/BHI.2017.7897215>
- Szeliski R (2011) Computer vision algorithms and applications. Springer, London, pp 10–17
- Tarlin D (2019) Labelimg. GitHub repository. <https://github.com/tzutalin/labelimg>. Accessed 25 June 2019
- P.N. of Excellence: Pascal voc. <http://host.robots.ox.ac.uk/pascal/VOC/>. Accessed on 12 July 2019 (2019)
- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M, Levenberg J, Monga R, Moore SMDG, Steiner B, Tucker P, Vasudevan V, Warden P, Wicke M, Yu Y, Zheng X, Brain Y (2016) Tensorflow: a system for large-scale machine learning. In: OSDI'16 Proceedings of the 12th USENIX conference on operating systems design and implementation, pp 265–283
- GoogleAI (2019) Tensorflow object detection preprocessor. https://github.com/tensorflow/models/blob/master/research/object_detection/core/preprocessor.py#L443. Accessed 17 June 2019
- Ren S, He K, Girshick R, Sun J (2016) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39:1137–1149
- GoogleInc (2019) Tensorflow. https://github.com/tensorflow/models/tree/master/research/object_detection/models. Accessed 18 June 2019
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C, Berg A (2016) Ssd: Single shot multibox detector. In Proceedings of the European conference on computer vision—ECCV, Amsterdam, The Netherlands, pp 21–37
- Lin T, Goyal P, Girshick R, He K, Dollar P (2018) Focal loss for dense object detection. arXiv:1708.02002v2 [cs.CV] 7
- Cartucho J (2019) map. <https://github.com/Cartucho/mAP>. Accessed on 12 July 2019

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.