

RESEARCH

Open Access



Examination of blood samples using deep learning and mobile microscopy

Juliane Pfeil¹, Alina Nechyporenko^{1,2}, Marcus Frohme^{1*}, Frank T. Hufert³ and Katja Schulze⁴

*Correspondence:
marcus.frohme@th-wildau.de
¹ Molecular Biology
and Functional Genomics,
Technical University
of Applied Sciences,
Hochschulring 1,
15745 Wildau, Germany
Full list of author information
is available at the end of the
article

Abstract

Background: Microscopic examination of human blood samples is an excellent opportunity to assess general health status and diagnose diseases. Conventional blood tests are performed in medical laboratories by specialized professionals and are time and labor intensive. The development of a point-of-care system based on a mobile microscope and powerful algorithms would be beneficial for providing care directly at the patient's bedside. For this purpose human blood samples were visualized using a low-cost mobile microscope, an ocular camera and a smartphone. Training and optimisation of different deep learning methods for instance segmentation are used to detect and count the different blood cells. The accuracy of the results is assessed using quantitative and qualitative evaluation standards.

Results: Instance segmentation models such as Mask R-CNN, Mask Scoring R-CNN, D2Det and YOLACT were trained and optimised for the detection and classification of all blood cell types. These networks were not designed to detect very small objects in large numbers, so extensive modifications were necessary. Thus, segmentation of all blood cell types and their classification was feasible with great accuracy: qualitatively evaluated, mean average precision of 0.57 and mean average recall of 0.61 are achieved for all blood cell types. Quantitatively, 93% of ground truth blood cells can be detected.

Conclusions: Mobile blood testing as a point-of-care system can be performed with diagnostic accuracy using deep learning methods. In the future, this application could enable very fast, cheap, location- and knowledge-independent patient care.

Keywords: Mobile microscopy, Blood cell detection, Machine learning, Deep learning, Instance segmentation

Introduction

The potential of mobile microscopy for diagnostic purposes in human and veterinary medicine has already been investigated using different setups and for many applications. The utilisation varies between the diagnosis of infectious diseases such as tuberculosis, the identification of parasites (malaria, helminth infections) and the detection of viruses or bacterial spores [1]. The basis are primarily constructions based on a special hardware and a smartphone. A differentiation can be made between lens based [2–4], lensless [5] or digital holographic microscopy [6, 7] and ptychographic systems [8]. The lens based



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

systems consist of conventional microscopy components (lenses, objective, eyepiece) and their accuracy or resolution and the field of view are strongly dependent on the efficiency of the smartphone camera used. In lensless microscopy, the sample is usually placed very close to the camera sensor and illuminated vertically. To improve resolution, digital holographic microscopy not only illuminates the cells or objects to create a wave front, but also creates an interference pattern by generating a reference beam. With both techniques, the image must be reconstructed [5]. Overall for lensless and holographic systems commercial mobile systems are generally not available and the reconstruction of the images is algorithmically complex [6]. For ptychographic systems a reconstruction is still necessary but easier in comparison to the aforementioned construction.

The use of a mobile phone has the advantage that it can be used directly for image acquisition, storage and especially for evaluation. In addition, cloud resources can be accessed via Internet connection to perform particularly computationally intensive examinations in a timely manner and/or to contact a doctor or medical facility. The development and utilisation of reliable point-of-care (POC) diagnostic tools is of outstanding importance for the medical care of the population of infrastructural underdeveloped regions. The lack of specialized professionals and facilities makes the application of a low-cost smartphone-based mobile microscope particularly advantageous for timely and proper treatment [1].

In this context, blood testing is a simple but beneficial opportunity to assess the general state of health and detect various signs of disease. With the help of a microscopic blood examination a specialist can determine the number of red blood cells (RBC), white blood cells (WBC), platelets (PLT) and notice morphological changes. For example, a deficiency of RBC is an indicator of anemia, WBC provide information about a possible infection, the number of PLT is crucial for blood coagulation and morphological changes can provide information about sickle cell disease [9].

Many scientists have already worked on automated detection of different blood cells from microscopic images. Previous approaches use conventional image recognition for segmentation of different blood components [10]. These efforts are focused on methods such as thresholding [4, 11–15], watershed algorithm [11, 12], morphological filters [12, 14, 16], color-based conversion/segmentation [4, 12, 13, 15–17], histogram equalization [14], active contours [3, 11] and haar cascade [14]. The results might reach the desired accuracy, but these approaches are very error-prone towards minimal changes in image acquisition, such as contrast, exposure, resolution etc. [18]. In addition, they mostly focus on only one [3, 12, 14–17] or two [4, 10, 11] cell types and are rarely applied on a mobile device [13, 14]. Recent methods for image segmentation are increasingly based on machine learning (ML)/deep learning (DL) strategies due to their robustness, higher accuracy, flexible deployment and generalization ability [18–20]. A distinction can be made between simple object detection, semantic- and instance segmentation approaches [19, 20]. Instance segmentation allows an exact delineation of each cell and detection of its morphology. Therefore, it is a beneficial method to make particularly precise statements about the different blood cell types. Current research on blood cell segmentation of microscopic images focuses on networks such as YOLO [21] and region-based convolutional neural networks (R-CNN) [22] for object detection; U-Net [23], Seg-Net [24] and fully convolutional networks (FCN) [25] for semantic segmentation and Mask

R-CNN [26] for instance segmentation. There is only one published example of researchers Alam and Islam [27] using the YOLO network to detect all blood cell types (RBC, WBC, PLT). For training, they use the Blood Cell Count Dataset (BCCD) [28], which provides stained, microscopic blood images and annotations. They achieve 96.09% accuracy for detection of RBC, 86.89% for WBC and 96.36% for PLT. However, their object segmentation approach does not allow accurate delineation of the cell membrane and they require K-nearest neighbour (KNN) and intersection over union (IoU) based post-processing as the PLT are often detected twice. The Faster R-CNN [29] approach is used to segment either the WBC [30] or RBC [31] only. In publications on semantic segmentation of blood cells, U-Net [32–34], Seg-Net [35] and FCN [31] are appropriately adapted to likewise segment only one [33, 34] to two [32, 35] cell types. There are very few publications that establish instance segmentation using Mask R-CNN for RBC and WBC [36] or only WBC [37] detection. Dhieb et al. [36] achieve 92% accuracy for detection of RBC and 96% for WBC and Fan et al. [37] achieve 99% average accuracy for segmentation masks of WBC.

In order to train the DL algorithms, different publicly available databases such as BCCD [27, 29, 30], acute lymphoblastic leukemia image database (ALL-IDB) [34, 35, 38] or self-made microscopic images [31, 32] are used. However, all of them were acquired with a laboratory microscope and high quality setups. Only one paper uses image data acquired with a smartphone camera, but also generated with an automated laboratory microscope [32]. Consequently, there is no DL-instance segmentation approach for all three blood cell types, which uses microscopic images from a mobile setup whose images can be generated cheaply and quickly, but cannot match the quality of images from publicly available databases or laboratory microscopes.

Methods

Sample preparation and microscopy

Donor for all blood samples used in this study is the first author. A small drop of capillary blood was obtained by fingerprick and was taken up with the middle of a cover slip. Waldeck's Testsimplents® [39] were used to stain the different blood cell types. The cover slip with the drop of blood is placed on the prestained area of the slide. After 10 to 15 min the slide was microscopically evaluated. For that the Bresser Erudit DLX microscope [40] with an objective with $60\times$ magnification and a numerical aperture of 0.85 was used [41], resulting in a resolution of $0.32\ \mu\text{m}$. The eyepiece provides an additional $10\times$ enlargement resulting in a total magnification of $600\times$. For image acquisition the microscopic ocular was removed and a digital eyepiece camera with 5 megapixels [42] was inserted. This setup produces an equivalent magnification of $600\times$. A commercially available smartphone (Xiaomi Mi A2 [43]) was connected via USB and color microscopic images with a resolution of 1440×1080 pixels were taken with the mobile app OTG view [44]. Before the annotation was started all images were cut to a resolution of 1000×1000 pixels and 96 dots per inch (dpi) to provide a square, standard size for different DL networks. One pixel corresponds to $0.1\ \mu\text{m}$. With the goal of increasing the robustness of the DL algorithms, the infrared (IR)/ultraviolet (UV) blocking filter of the ocular camera was removed for some image acquisitions. This resulted in artificially noisy and color-shifted images, which could also be generated naturally by mobile

use at different locations. A rapid analysis and inexperience of the performing person could produce inaccurate images that are not perfectly focused and make it difficult to optimally set the lighting, which is also influenced by ambient parameters (artificial light indoors or sunlight outdoors).

Dataset and image labelling

The dataset contains 40 microscopic pictures and each recording includes approx. 150–200 RBC, 2–4 WBC and 10–15 PLT. For training and validation a total of 5101 RBC, 71 WBC, 432 PLT were used. Images were prepared and annotated with the software labelme [45] and CVAT [46]. 30 images were taken with the normal camera setup (high quality images, Fig. 1a) and for ten images the color filter of the camera was removed (low quality images, Fig. 1b).

All RBC, WBC and PLT were labelled with polygon masks. 24 images were chosen for the training set (18 high quality and six low quality images) and eight images for the validation set (six high quality and two low quality images). A threefold cross validation [47] was performed with different splits of these 32 images resulting in three training and three validation sets. The remaining eight images (six high quality and two low quality images) were used for the test set containing 997 RBC, 14 WBC and 75 PLT and served to confirm the results.

DL algorithms and training

For the targeted instance segmentation task of all blood cells, four different DL algorithms were implemented, optimised and the results evaluated. The well-known and extensively used Mask R-CNN [26] consists of 2 stages and is based on Faster R-CNN [29], which uses a region proposal network to predict bounding boxes for the different object classes. On top of this, Mask R-CNN predicts segmentation masks for the individual instances. Training and validation were done with the Mask R-CNN implementation of the toolbox MMDetection provided by Chen et al. [48]. Besides, Mask R-CNN

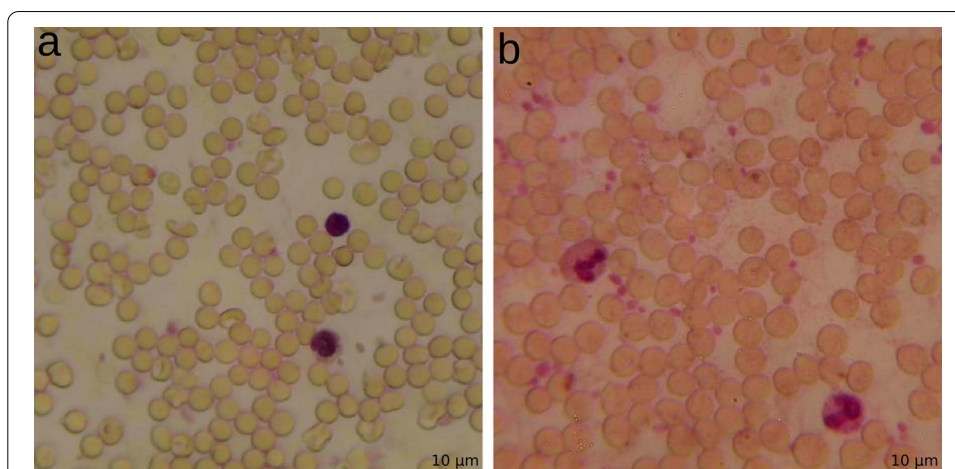


Fig. 1 Microscopic images of stained blood samples at 600 × magnification. The preparation and microscopy of the blood samples were performed as described in the section "Sample preparation and microscopy". **a** High quality image, **b** low quality image. RBC are stained light red, WBC and PLT are stained dark purple

is the predecessor of Mask Scoring R-CNN (MS R-CNN) [49]. This network significantly improves the accuracy of instance masks in the 2017 common objects in context (COCO) challenge [50]. The additional implementation of a network block termed as the MaskIoU head, which is trained with the quality of the predicted instance masks, improves the accuracy of the mask predictions. D2Det [51] is another two-stage detector based on the Faster R-CNN framework [29]. In addition, discriminative region of interest (RoI) pooling and dense local regression is applied for instance segmentation to improve accuracy and speed. YOLACT [52] is a one-stage framework for real time instance segmentation, which is characterized by an excellent inference speed, but therefore shows some drawbacks in segmentation accuracy. The network predicts a certain number of prototype masks that are generated by FCN [25] and calculates mask coefficients in parallel, which are multiplied together for each instance to create a linear combination of output masks. Training parameters for all frameworks have been customized for the existing graphics processing unit (GPU) infrastructure (GPU NVIDIA Tesla V100, DDR4-RAM 384 GB) and the task of blood cell instance segmentation. The following training parameters are common between adapted frameworks: training method stochastic gradient descent (SGD), momentum 0.9, backbone residual network with 101 layers (ResNet-101) [53] including feature pyramid network (FPN) [54] and augmentation methods such as resizing, random flips and change in hue etc. were applied. Furthermore, pre-trained weights based on ImageNet [55] and COCO datasets [50] were used for all frameworks. For YOLACT and D2Det the settings for Non-Maximum Suppression (NMS) were adjusted. In addition, anchor box size/scales were adapted for the detection of small objects and the number of possible detections per image respectively the number of trainable masks were increased. For YOLACT the number of predictions for NMS was changed from 200 to 400, the confidence threshold was decreased from 0.05 to 0.01 and the boxes threshold was modified from 0.5 to 0.1. For the anchor scales the configuration of YOLACT++ [56] was used. The number of masks was increased from 250 to 500 as well as the number of possible detections per image from 300 to 500. For D2Det the NMS threshold for the region proposal network (RPN) was decreased from 0.7 to 0.3 and the anchor scales were modified from 8 to 2. The number of detections per image were increased from 100 to 300. For MS R-CNN the number of possible detections per image were also increased in the MaskIoU head from 100 to 500. Other training parameters that vary between the different frameworks are shown in Table 1. The hyperparameters of the original networks and the modified values are listed. All networks were trained until no significant loss could be detected. The learning rate and

Table 1 Training parameters for Mask R-CNN, MS R-CNN, D2Det and YOLACT

Framework	# epoch		Learning rate		Batch size		Weight decay	Basis for pretrained weights
	Original	Modified	Original	Modified	Original	Modified		
Mask R-CNN	12	1000	0.02	0.02	2	2	0.0001	COCO [57]
MS R-CNN	3750*	1000*	0.02	0.00025	16	1	0.0001	ImageNet [58]
D2Det	24	3000	0.02	0.02	2	2	0.0001	COCO [59]
YOLACT	66.666*	1000–2000	0.0001	0.0001	8	2	0.0005	COCO [60]

* Calculated by: # epoch = number of iterations*batch size/number of training images

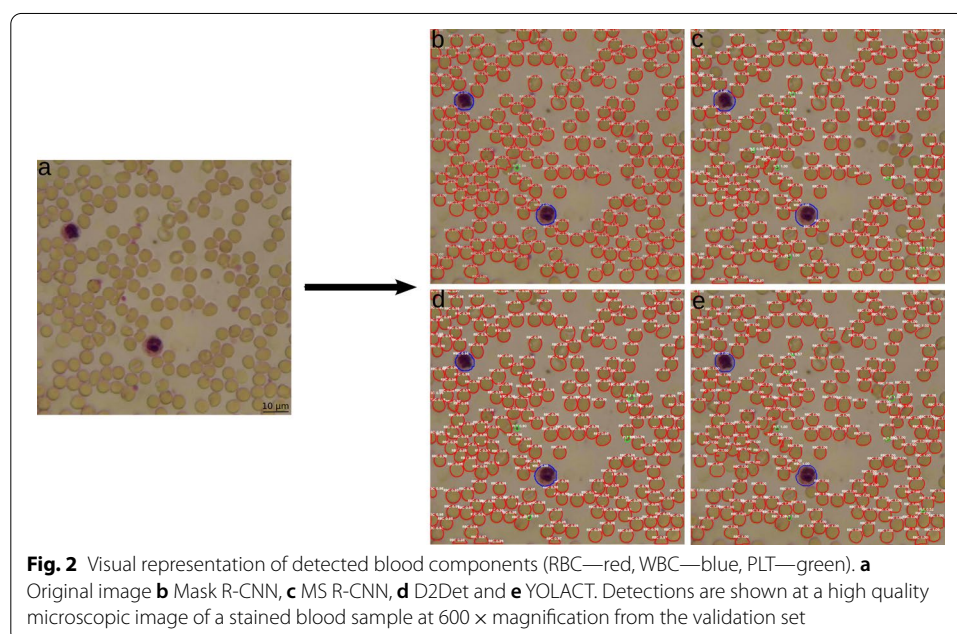
batch size were chosen according to the GPU memory. The weight decay was applied according to the defaults in the available code from Github.

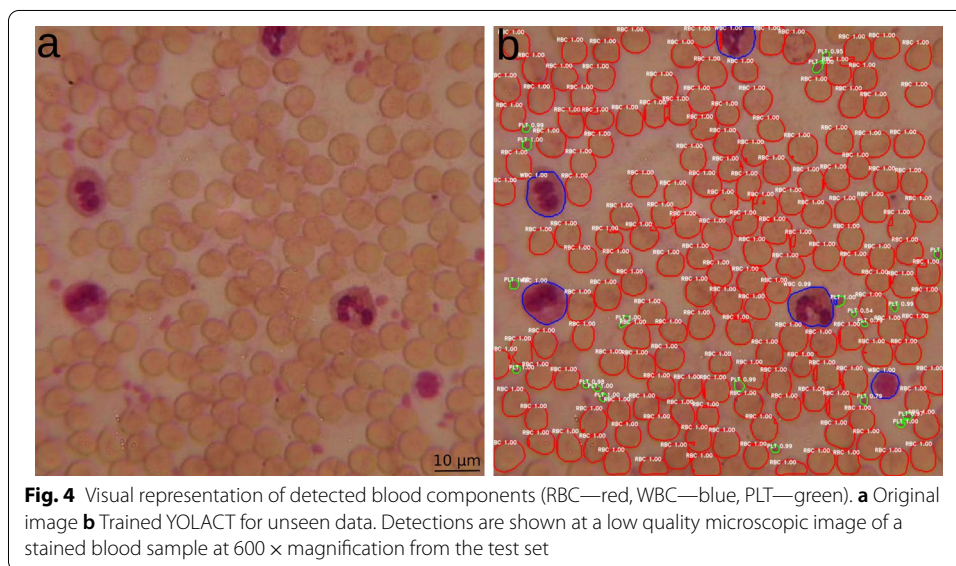
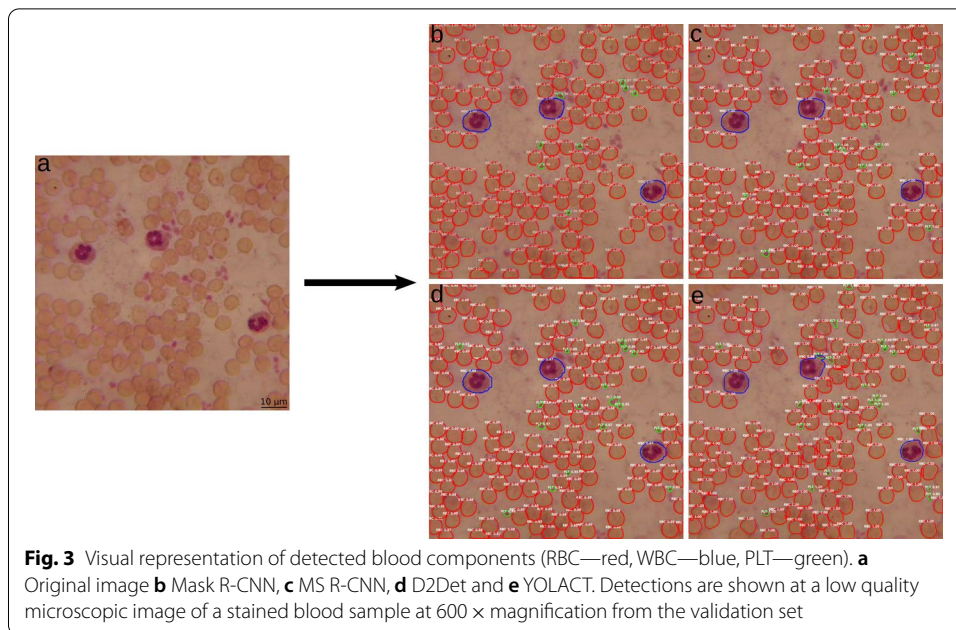
Results

In this research work, different instance segmentation frameworks were modified, trained, optimised and their performance evaluated. The goal was to achieve the most accurate segmentation and classification of the blood components RBC, WBC and PLT on microscopic images generated with a mobile setup and a smartphone. Supported by a threefold cross validation, three different models were trained for each DL framework. The performance of these models is evaluated in the following sections. The results for the validation sets and the test set are presented visually (representation of predicted detections, including segmentation masks) and in addition the findings are described qualitatively (mean average precision (mAP) and mean average recall (mAR) of segmentation masks) and quantitatively (number of detected blood components).

Visual results

Visually, the output has been unified for all trained frameworks. The predicted output masks were generated with an IoU threshold of 0.5. RBC are displayed in red, WBC in blue and PLT in green. Furthermore, each detection is labelled with the corresponding class and confidence score (quality of predicted mask and detected class). Figures 2 and 3 show the results of a cross validation output model for the frameworks Mask R-CNN, MS R-CNN, D2Det and YOLACT. In both Figures the detection results for all trained models basically do not show any difference between high and low quality images. All frameworks show a high confidence score for the different classes and accurate masks for WBC and PLT. The visual results for Mask R-CNN in Figs. 2b and 3b show several missing PLT detections. For trained MS R-CNN some RBC are not detected in Fig. 2c





and some PLT are missing in Fig. 3c. D2Det failed to detect some PLT in Figs. 2d and 3d. Trained YOLACT shows slightly noisy masks for RBC and a few misdetections at the image edges in both pictures.

In Fig. 4, a trained YOLACT model was used for a low quality image of the test set. The result also shows very good detection and segmentation output with no significant differences to results from the validation set.

Qualitative results

For the assessment of the segmentation results, the mAP and the mAR were determined with IoU's of 0.50 to 0.95 (step size 0.05) with a maximum of 100 detections aided by the

Table 2 mAP and mAR at IoU's of 0.50 to 0.95 for the validation (v) and test set (t)

Framework	Dataset	∅ mAP ± σ				∅ mAR ± σ			
		All	RBC	WBC	PLT	All	RBC	WBC	PLT
Mask R-CNN	v	0.43 ± 0.02	0.46 ± 0.02	0.68 ± 0.06	0.15 ± 0.01	0.46 ± 0.02	0.48 ± 0.02	0.72 ± 0.06	0.18 ± 0.02
	t	0.47 ± 0.01	0.60 ± 0.01	0.65 ± 0.02	0.16 ± 0.02	0.50 ± 0.01	0.63 ± 0.01	0.67 ± 0.02	0.18 ± 0.02
MS R-CNN	v	0.67 ± 0.01	0.53 ± 0.02	0.91 ± 0.01	0.56 ± 0.00	0.69 ± 0.01	0.54 ± 0.02	0.93 ± 0.01	0.59 ± 0.02
	t	0.48 ± 0.02	0.58 ± 0.01	0.57 ± 0.06	0.28 ± 0.01	0.53 ± 0.01	0.63 ± 0.01	0.61 ± 0.05	0.36 ± 0.02
D2Det	v	0.42 ± 0.04	0.56 ± 0.02	0.46 ± 0.05	0.24 ± 0.07	0.47 ± 0.04	0.60 ± 0.02	0.51 ± 0.05	0.30 ± 0.07
	t	0.44 ± 0.01	0.61 ± 0.02	0.45 ± 0.01	0.25 ± 0.03	0.49 ± 0.01	0.65 ± 0.02	0.50 ± 0.01	0.33 ± 0.04
YOLACT	v	0.45 ± 0.03	0.42 ± 0.02	0.71 ± 0.08	0.24 ± 0.05	0.49 ± 0.03	0.44 ± 0.02	0.74 ± 0.07	0.28 ± 0.09
	t	0.50 ± 0.01	0.56 ± 0.02	0.67 ± 0.02	0.28 ± 0.03	0.55 ± 0.00	0.60 ± 0.01	0.69 ± 0.02	0.37 ± 0.04

IoU's of 0.50 to 0.95 with maximum of 100 detections for trained Mask R-CNN, MS R-CNN, D2Det and YOLACT for the validation and test set

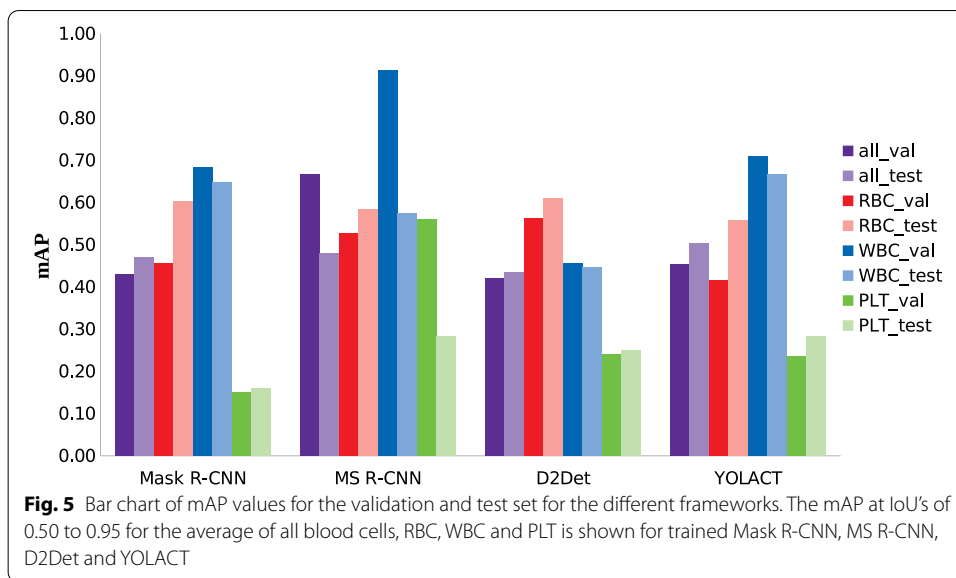
COCO evaluation code [61]. In Table 2, the average of these values and their standard deviations (obtained by cross validation) for the trained models on the validation and the test set are shown and the best scoring outputs are marked in bold. mAP and mAR are mainly stable for the different models of cross validation, as indicated by the low standard deviation of mostly less than 0.05. Moreover, the different frameworks show the same tendencies for both parameters, e. g. MS R-CNN performs best for WBC in the validation set or YOLACT for PLT in the test set.

For a more straightforward visualisation, the mAP for both data sets is shown in Fig. 5 (the mAR performs equivalently and is therefore not shown).

The difference between validation and test set of each trained framework is mostly only 0.01 to 0.05 of the respective mAP. Larger deviations are only found in the values for the RBC for Mask R-CNN and YOLACT and for the WBC and PLT for MS R-CNN. Despite this, the results indicate a good generalisation ability of the models. The visual results from the previous section (Figs. 2 and 3) are also well reflected. The mAP of all models ranges between 0.42 and 0.61 for the RBC and between 0.45 and 0.91 for the WBC, confirming the high accuracy of masks and confidence scores for these classes. YOLACT achieves the lowest value for the RBC with a mAP of 0.42, visually also detectable by misdetections and noisy masks. The best performance for all blood components is achieved with the MS R-CNN model for the validation set and with YOLACT for the test set. With Mask R-CNN, the PLT in particular are poorly segmented and detected, which is also evident in Figs. 2a and 3a. However, the other trained models also achieve a good average output. Larger differences and the worst segmentation performance for all models is evident in the PLT. The reason for this is that basically smaller objects are more difficult to recognize and all frameworks were pre-trained and optimised for the COCO dataset. The smallest objects in this dataset correspond to 4% of the image size [50]. The PLT in the blood images, with an average size of 25 × 25 pixels, correspond to only 2.5% of the image size.

Quantitative results

For the validation and the test set, the number of ground truth (GT) and predicted detections were also calculated for each class. In Table 3 the sum of the values for each cross validation for the trained models are shown and results closest to the GT are marked in



bold. Mask R-CNN and D2Det provide the best results on quantitative accuracy for the RBC. For WBC, Mask R-CNN and MS R-CNN show nearly 100% detection accuracy for both data sets. But again, the weak performance of Mask R-CNN for the detection of PLT is significant.

For evaluation purposes, the predicted detections in % of the GT for the validation and test set and the different frameworks are shown in Fig. 6. The average performance for all blood components is shown as a separate bar in violet.

Similar to the qualitative results, only minor deviations between validation and test set are evident for the individual networks and the standard deviation is also mostly below 5%. Larger differences are only found for the PLT, as their small size and the associated detection difficulty can lead to greater fluctuations. However, these results once again confirm the generalisability of the models. On average, YOLACT performs best for all types of blood cells, followed by MS R-CNN, D2Det and Mask R-CNN. The detection

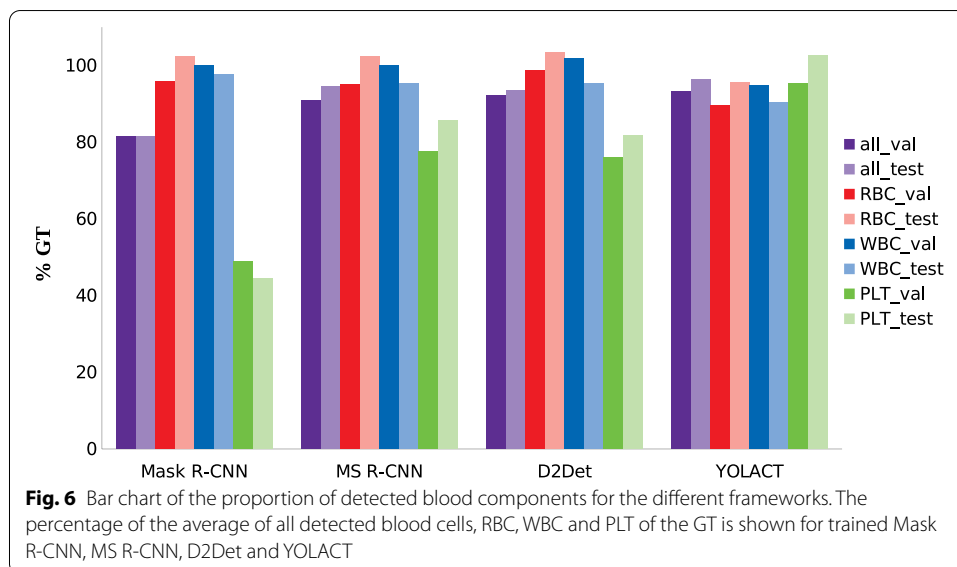
Table 3 Number of detected blood cells (RBC, WBC, PLT) for the validation (v) and test set (t)

Framework	Dataset	Σ # detections		
		RBC	WBC	PLT
Mask R-CNN	v	3662	54	158
	t	3058	41	100
MS R-CNN	v	3636	54	256
	t	3059	40	193
D2Det	v	3775	55	242
	t	3089	40	184
YOLACT	v	3427	51	311
	t	2861	38	231

Trained Mask R-CNN, MS R-CNN, D2Det and YOLACT

GT numbers: Σ validation sets RBC—3825, WBC—54, PLT—328

Σ test set RBC—2991, WBC—42, PLT—225



results for RBC and WBC show very good scores for all models and frameworks with values above 90%. Only the detection performance of YOLACT for the RBC is a bit weaker in comparison to the other models. However, this is also clearly visible in the visual results where some RBC were not detected at the edges of the image. However, this model is able to detect most PLT compared to the other frameworks and achieves over 95% detection accuracy for both data sets.

Discussion

In principle, all modified and trained models (Mask R-CNN, MS R-CNN, D2Det and YOLACT) achieve good to very good results in visual output, qualitatively and quantitatively, and could possibly be used for diagnostic purposes. The adjustment and optimisation of training parameters was absolutely necessary, as the original versions delivered very weak performance or were even unable to detect some blood cell types. To improve the relatively weak detection performance of Mask R-CNN (Ø 47% of GT for validation and test set) for PLT, further optimisation of the anchors to detect smaller objects should be performed. Furthermore, YOLACT shows very good results for the detection of WBC and PLT, but the performance for RBC is a bit weaker as these are often not detected at the image edges. Whether the detection of cells that are partly outside the field of view will be advantageous for diagnostic applications in the future has yet to be assessed. Otherwise, these cells could be ignored and would significantly improve the overall performance of the network. Pre-processing of the images with traditional image recognition algorithms such as thresholding, k-means clustering, contrast enhancement etc. are also conceivable to improve the overall performance.

The performance of the optimised instance segmentation models for the recognition of different blood components is also compared with existing similar research. Dhieb et al. [36] used the Mask R-CNN network as a basis for the instance segmentation of two blood components (no PLT) and achieved a quantitative accuracy of 92% for the detection of RBC and 96% for WBC. The detection accuracy of the models presented in this

publication is up to 7% higher for RBC (\emptyset 99% for Mask R-CNN and MS R-CNN of GT for validation and test set). For WBC, Mask R-CNN, MS R-CNN and D2Det achieve a better result for the detection performance (\emptyset 99% for Mask R-CNN and D2Det, \emptyset 98% for MS R-CNN of GT for validation and test set). The training and validation set in this work contain in total only 71 WBC while Dhieb et al. [36] used a dataset of 150 images with 24,000 cells [62]. However, due to applied cross validation including calculated standard deviations, the comparison between the results of the validation and test set allow a statistically significant statement. The researchers in the aforementioned work did not perform a qualitative evaluation of the segmentation results. Alam and Islam [27] use YOLO as a basis for object detection of all blood components and achieve a quantitative accuracy for the detection of RBC of 96.09%, 86.89% for WBC and 96.36% for PLT. Again, the detection accuracy of the models in this publication is up to \emptyset 3% higher for RBC and \emptyset 13% higher for WBC. For PLT, only YOLACT achieves a higher detection performance of \emptyset 99% of GT for validation and test set. The authors achieved a mAP of 0.6236 as a qualitative measure, thus being slightly higher than the MS R-CNN presented (\emptyset 0.57 for validation and test set). However, Alam and Islam [27] use only object detection and not, as in this work, significantly more advanced instance segmentation. Both publications also use datasets [28, 61] with images taken at $1000 \times$ magnification and standard microscopes, so that a significantly better resolution is available and the detection of the cells is consequently easier.

To further improve the presented models and increase their accuracy, the dataset should be enlarged to provide more images for training and validation. Furthermore, additional robustness can be generated by using other noise parameters to degrade the image quality, such as changing the illumination, using a camera with lower resolution etc.

Conclusions

In this research work, the performance of DL-based instance segmentation algorithms for the detection of all blood cell types on microscopic images taken with a mobile microscope and a smartphone is thoroughly investigated. Training and optimisation of parameters for Mask R-CNN, MS R-CNN, D2Det and YOLACT network architectures were conducted. After examining the visual, qualitative and quantitative results, MS R-CNN performs in total best and achieved a \emptyset mAP of 0.57 and \emptyset mAR of 0.61 for the validation and test set for the segmentation of all blood cell types and was able to detect \emptyset 93% of all cells from both sets. All frameworks surpass their source versions in terms of visual output, qualitative (mAP, mAR) and quantitative results and their feasibility and effectiveness were demonstrated. Although some smartphone-based microscopes are already commercially available, the presented solution is innovative and its deployment is advantageous because mobile use of the optical system is already conceivable (lightweight, rechargeable microscope, digital eyepiece camera and mobile phone). A suitable smartphone application has to be developed for a location-independent evaluation of microscopic blood images. Future work will investigate the applicability of these algorithms in such an application allowing mobile analysis to be performed directly at the POC. In addition, further reduction in the size of the hardware will also increase

mobility. Providing an autofocus and an automatic image capture are further adaptation options that enable easy utilisation regardless of location.

Abbreviations

ALL-IDB1: Acute lymphoblastic leukemia image database; COCO: Common objects in context; DL: Deep learning; dpi: Dots per inch; FCN: Fully convolutional neural network; FPN: Feature pyramid network; GPU: Graphics processing unit; GT: Ground truth; IoU: Intersection over union; IR: Infrared; mAP: Mean average precision; mAR: Mean average recall; ML: Machine learning; MS R-CNN: Mask Scoring R-CNN; NMS: Non-Maximum Suppression; PLT: Platelets; POC: Point-of-care; RBC: Red blood cells; R-CNN: Region-based convolutional neural network; RoI: Region of interest; RPN: Region proposal network; ResNet-101: Residual network with 101 layers; SGD: Stochastic gradient descent; UV: Ultraviolet; WBC: White blood cells.

Acknowledgements

Many thanks to all reviewers for their corrections, time and effort.

Authors' contributions

JP: Methodology, experimental design and data analysis, writing of the manuscript. AN: Support with discussion and significant revisions. MF: Conceptualization of the study, critical evaluation and discussion. KS: Support with evaluation metrics, discussion and conclusion. FH: Support with discussion and conclusion. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. Work for this manuscript was financed by the Federal Ministry of Education and Research (BMBF) of Germany under grant number 13FH209PX8. The funding body had no role in the design of the study, analysis and interpretation of data and in writing the manuscript.

Availability of data and materials

The dataset generated and analysed during the current study is not publicly available because the research project was in cooperation with the company Oculyze GmbH, but is available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the local Ethics Committee (protocol number EK 3/2021) of the Technical University of Applied Sciences, Wildau (Germany) in accordance with the Declaration of Helsinki. All participants signed an informed consent.

Consent for publication

Written informed consent for publication for images or other personal or clinical details was obtained from all participants.

Competing interests

There are no competing interests for the authors JP, AN, MF, FH and KS. The submitted work was performed in cooperation with KS from the company Oculyze GmbH, Wildau, Germany. They work in the field of mobile microscopy and computer vision and have therefore a vested interest in the success of this field.

Author details

¹Molecular Biology and Functional Genomics, Technical University of Applied Sciences, Hochschulring 1, 15745 Wildau, Germany. ²Kharkiv National University of Radio Electronics, Kharkiv, Ukraine. ³Institute for Microbiology and Virology, Brandenburg Medical School Theodor Fontane, Neuruppin, Germany. ⁴Oculyze GmbH, Mobile Microscopy and Computer Vision, Wildau, Germany.

Received: 20 October 2021 Accepted: 4 February 2022

Published online: 11 February 2022

References

1. Pfeil J, Dangelat LN, Frohme M, Schulze K. Smartphone based mobile microscopy for diagnostics. *J Cell Biotechnol.* 2018;4(1–2):57–65.
2. Skandarajah A, Reber CD, Switz NA, Fletcher DA. Quantitative imaging with a mobile phone microscope. *PLoS ONE.* 2014;9(5):e96906.
3. Zhu H, Mavandadi S, Coskun AF, Yaglidere O, Ozcan A. Optofluidic fluorescent imaging cytometry on a cell phone. *Anal Chem.* 2011;83(17):6641–7.
4. Zhu H, Sencan I, Wong J, Dimitrov S, Tseng D, Nagashima K, Ozcan A. Cost-effective and rapid blood analysis on a cell-phone. *Lab Chip.* 2013;13(7):1282–8.
5. McLeod E, Ozcan A. Microscopy without lenses. *Phys Today.* 2017.
6. Cacace T, Bianco V, Mandracchia B, Pagliarulo V, Oleandro E, Paturzo M, Ferraro P. Compact off-axis holographic slide microscope: design guidelines. *Biomed Opt Express.* 2020;11(5):2511–32.

7. Zhang Y, Koydemir HC, Shimogawa MM, Yalcin S, Guziak A, Liu T, Oguz I, Huang Y, Bai B, Luo Y, Luo Y. Motility-based label-free detection of parasites in bodily fluids using holographic speckle analysis and deep learning. *Light Sci Appl*. 2018;7(1):1–8.
8. Lee KC, Lee K, Jung J, Lee SH, Kim D, Lee SA. A smartphone-based fourier ptychographic microscope using the display screen for illumination. *ACS Photonics*. 2021;8(5):1307–15.
9. Ballard MC. Atlas of blood cells in health and disease. Atlanta: US Department of Health and Human Services, Public Health Service, Centers for Disease Control; 1987.
10. Adollah R, Mashor MY, Nasir NM, Rosline H, Mahsin H, Adilah H. Blood cell image segmentation: a review. In: 4th Kuala Lumpur international conference on biomedical engineering, 2008. Springer, Berlin, pp. 141–144
11. Habibzadeh M, Krzyzak A, Fevens T, Sadr A. Counting of RBCs and WBCs in noisy normal blood smear microscopic images. In: Medical imaging 2011: computer-aided diagnosis 2011; vol. 7963, p. 79633l. International Society for Optics and Photonics.
12. Sharif JM, Miswan MF, Ngadi MA, Salam MS, bin Abdul Jamil MM. Red blood cell segmentation using masking and watershed algorithm: a preliminary study. In: 2012 international conference on biomedical engineering (ICoBE), 2012; pp. 258–262. IEEE.
13. Kit CY, Tomari R, Zakaria W, Nurshazwani W, Othman N, Safuan SN, Ang Jie Yi J, Tan Chun Sheng N. Mobile based automated complete blood count (auto-CBC) analysis system from blood smeared image. *Int J Electr Computer Eng*. 2017;7(6):10.
14. Moravapalle UP, Deshpande A, Kapoor A, Ramjee R, Ravi P. Blood count on a smartphone microscope: Challenges. In: Proceedings of the 18th international workshop on mobile computing systems and applications, 2017; pp. 37–42.
15. Hegde RB, Prasad K, Hebbar H, Singh BM. Development of a robust algorithm for detection of nuclei and classification of white blood cells in peripheral blood smear images. *J Med Syst*. 2018;42(6):1–8.
16. Safuan SN, Tomari MR, Zakaria WN. White blood cell (WBC) counting analysis in blood smear images using various color segmentation methods. *Measurement*. 2018;116:543–55.
17. Wu J, Zeng P, Zhou Y, Olivier C. A novel color image segmentation method and its application to white blood cell image analysis. In: 2006 8th international conference on signal processing, 2006, vol 2. IEEE.
18. O'Mahony N, Campbell S, Carvalho A, Harapanahalli S, Hernandez GV, Krpalkova L, Riordan D, Walsh J. Deep learning vs. traditional computer vision. In: Science and information conference 2019; pp. 128–144. Springer, Cham.
19. Minaee S, Boykov YY, Porikli F, Plaza AJ, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: a survey. *IEEE Trans Pattern Anal Mach Intell* 2021.
20. Zhao ZQ, Zheng P, Xu ST, Wu X. Object detection with deep learning: a review. *IEEE Trans Neural Netw Learn Syst*. 2019;30(11):3212–32.
21. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 779–788.
22. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 580–587.
23. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, 2015, pp. 234–241. Springer, Cham.
24. Badrinarayanan V, Kendall A, Cipolla R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(12):2481–95.
25. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431–3440.
26. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2961–2969
27. Alam MM, Islam MT. Machine learning approach of automatic identification and counting of blood cells. *Healthcare Technol Lett*. 2019;6(4):103–8.
28. BCCD Dataset. Roboflow. 2021. <https://public.roboflow.com/object-detection/bccd>. Accessed 30 March 2021.
29. Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*. 2015
30. Xia T, Jiang R, Fu YQ, Jin N. Automated blood cell detection and counting via deep learning for microfluidic point-of-care medical devices. In: IOP conference series: materials science and engineering, 2019, vol. 646, no. 1, p. 012048. IOP Publishing.
31. Bailo O, Ham D, Min Shin Y. Red blood cell image generation for data augmentation using conditional generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2019.
32. Mundhra D, Cheluvraju B, Rampure J, Dastidar TR. Analyzing microscopic images of peripheral blood smear using deep learning. In: Deep learning in medical image analysis and multimodal learning for clinical decision support, 2017, pp. 178–185. Springer, Cham.
33. Zhang M, Li X, Xu M, Li Q. Automated semantic segmentation of red blood cells for sickle cell disease. *IEEE J Biomed Health Inform*. 2020;24(11):3095–102.
34. Li H, Zhao X, Su A, Zhang H, Liu J, Gu G. Color space transformation and multi-class weighted loss for adhesive white blood cell segmentation. *IEEE Access*. 2020;8:24808–18.
35. Tran T, Kwon OH, Kwon KR, Lee SH, Kang KW. Blood cell images segmentation using deep learning semantic segmentation. In: 2018 IEEE international conference on electronics and communication engineering (ICECE), 2018, pp. 13–16. IEEE.
36. Dhieb N, Ghazzai H, Besbes H, Massoud Y. An automated blood cells counting and classification framework using mask R-CNN deep learning model. In: 2019 31st international conference on microelectronics (ICM), 2019; pp. 300–303. IEEE.

37. Fan H, Zhang F, Xi L, Li Z, Liu G, Xu Y. LeukocyteMask: An automated localization and segmentation method for leukocyte in blood smear images using deep neural networks. *J Biophotonics*. 2019;12(7):e201800488.
38. Labati RD, Piuri V, Scotti F. All-IDB: the acute lymphoblastic leukemia image database for image processing. In: 2011 18th IEEE international conference on image processing, 2011, pp. 2045–2048. IEEE. <https://doi.org/10.1109/ICIP.2011.6115881>
39. Waldeck. <https://www.waldeck-ms.de/en/waldeck-brands/>. Accessed 06 April 2021.
40. Bresser. <https://www.bresser.de/en/Microscopes-Magnifiers/BRESSER-Erudit-DLX-40-1000x-Microscope.html>. Accessed 06 April 2021.
41. Bresser <https://www.bresser.de/en/Microscopes-Magnifiers/Accessories/Objectives/BRESSER-DIN-Objective-60x.html> Accessed 13 December 2021.
42. AliExpress. <https://www.aliexpress.com/item/32943465268.html>. Accessed 06 April 2021.
43. Mi. Xiaomi Mi A2. <https://www.mi.com/global/mi-a2/>. Accessed 06 April 2021.
44. Teamforce Tools (2020). OTG View (Version 3.7) [Mobile App]. Google Play Store. <https://play.google.com/store/apps/details?id=com.dingtai.snakecamera&hl=de&gl=US>
45. Wada K. Labelme: Image polygonal annotation with python. 2016. GitHub repository: <https://github.com/wkentaro/labelme>. Accessed 2020, Version: 4.4.0.
46. Computer Vision Annotation Tool (CVAT). 2018. Github repository: <https://github.com/openvinotoolkit/cvat>. Accessed 2021, Version: 1.7.0
47. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Stat Surv*. 2010;4:40–79.
48. Chen K, Wang J, Pang J, Cao Y, Xiong Y, Li X, Sun S, Feng W, Liu Z, Xu J, Zhang Z. MMDetection: Open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155. 2019. GitHub repository: <https://github.com/open-mmlab/mmdetection>. Accessed 2021, Version: 2.10.0.
49. Huang Z, Huang L, Gong Y, Huang C, Wang X. Mask scoring r-cnn. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019 (pp. 6409–6418). GitHub repository: https://github.com/zhuang22/maskscoring_rcnn. Accessed 2020.
50. Lin TY, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P, Zitnick CL. Microsoft coco: common objects in context. In: European conference on computer vision 2014 Sep 6 (pp. 740–755). Springer, Cham. <https://cocodataset.org/>. https://doi.org/10.1007/978-3-319-10602-1_48
51. Cao J, Cholakkal H, Anwer RM, Khan FS, Pang Y, Shao L. D2det: Towards high quality object detection and instance segmentation. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 11485–11494. GitHub repository: <https://github.com/JialeCao001/D2Det>. Accessed 2020, Version: 1.1.0.
52. Bolya D, Zhou C, Xiao F, Lee YJ. Yolact: Real-time instance segmentation. In: Proceedings of the IEEE/CVF international conference on computer vision 2019 (pp. 9157–9166). GitHub repository: <https://github.com/dbolya/yolact>. Accessed 2020, Version: 1.2.
53. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition 2016, pp. 770–778.
54. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition 2017, pp. 2117–2125.
55. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition, 2009; p. 248–55. <https://doi.org/10.1109/CVPR.2009.5206848>
56. Bolya D, Zhou C, Xiao F, Lee YJ. Yolact++: Better real-time instance segmentation. arXiv preprint arXiv:1912.06218. 2019 Dec 3. GitHub repository: <https://github.com/dbolya/yolact> Accessed 2021
57. Mask R-CNN pretrained weights. https://download.openmmlab.com/mmdetection/v2.0/mask_rcnn/mask_rcnn_r101_fpn_1x_coco/mask_rcnn_r101_fpn_1x_coco_20200204-1efe0ed5.pth. Accessed: 2021
58. Mask Scoring R-CNN pretrained weights. <https://dl.fbaipublicfiles.com/detectron/ImageNetPretrained/MSRA/R-101.pkl>. Accessed: 2021
59. D2Det pretrained weights <https://drive.google.com/open?id=1rsYWWJ7zJ7-sSWz5q6aiuGFJ55bduSDo>. Accessed: 2021
60. YOLACT pretrained weights. <https://drive.google.com/file/d/1tvqFPd4bJtakOlmn-ula492g2qurRChj/view?usp=sharing>. Accessed: 2021
61. COCO API. 2014. GitHub repository: <https://github.com/cocodataset/cocoapi>. Accessed 2020.
62. Medical Image & Signal Processing Research Center (MISP) and Department of Pathology at Isfahan University of Medical Sciences Dataset. <http://misp.mui.ac.ir/fa/red-blood-cells-rbcs>. Accessed 30 June 2021.

Publisher's Note

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