

Segmentation of Neoplastic Cell Nuclei for Assisted Cell Labelling using Mask R-CNN

Ankit V Prabhu¹, Jenendar Bishnoi², R. I. Minu³

^{1,2} UG Students, Department of Computer Science, SRM Institute of Science & Technology, Kattankulathur, Chennai, India.

³Associate Professor, Department of Computer Science, SRM Institute of Science & Technology, Kattankulathur, Chennai, India.

Abstract

The paper proposes an instance segmentation technique to label and segment neoplastic cell nuclei from multiple instances of whole-slide images (WSI) using a contemporary neural network architecture called as the Mask Region based Convolutional Neural Network (Mask R-CNN). The main objective of this research is to automate the mentioned process by generating a pixel-wise binary mask capable of segmenting these instances and facilitating the advancement of intelligent systems in the field of medical imaging and computational pathology. Neoplastic cells are tumorous cells which pose detrimental damage to the cells around them and are the prologue for cancer development in organs. The problem of identifying these cells poses a bottle-neck in the research of cancer cure as it is an extremely tedious job to manually isolate these from the rest of the cells in the tissue. Hence, the automation of this process using deep learning (DL) based object-detection and segmentation techniques such as Mask R-CNN will allow researchers and pathologists to save valuable time otherwise consumed in manually identifying these nuclei. This time can instead be devoted to developing better cures by conducting more research. The paper also focuses on and highlights the best techniques and practices that can be employed while training a model for a task of such complexity. Using these techniques, a mean average precision (mAP) score of 0.756 and a binary panoptic quality (bPQ) score of 0.675 for neoplastic cells was achieved.

Keywords: Medical Imaging, Image Processing, Neoplastic cell, Deep Learning, Computer Vision, Segmentation, Computational Pathology, Mask R-CNN, Cancer Research.

I. INTRODUCTION

According to the 2018 Data Science Bowl [1] competition, cell nucleus identification is one of the first steps in the long procedure of synthesizing cures for diseases. This is because DNA of each cell is present in the nucleus and analysis of that determines how quickly and effectively a cure can be made. Researchers and pathologists spend hours trying to identify the required cell nuclei to gauge the level of progress made

each time a new experiment is conducted. If this process can be automated, it can allow researchers to track the cells and measure the level of treatment on them, significantly improving the number of experimentations that can be conducted in a particular time span and consequently, new and better cures for diseases such as cancer can be developed at a rapid pace.

With that being said, the process of cell identification and labelling is a complex task to automate, as stated by [2] and [3]. Firstly, the shape and size of each type of cell varies greatly. Secondly, sub-types of certain cells such as red-blood cells and white-blood cells are present which induces an added layer of complexity in the process of cell identification. Thirdly, the various staining techniques that are used to visualize cells in Computational Pathology (CPath) can generate varied results when coupled with different algorithms. Lastly, due to the complex structure and arrangement of cells, a lot of the substantial ones can be hidden or overlapped, resulting in outliers.

Deep Learning (DL), in recent years has achieved impressive results in the field of Computer Vision (CV) and image processing where certain models have outclassed even humans. The ILSVRC [4] has facilitated significant contributions in the development of Convolutional Neural Networks (CNN) which have completely changed the paradigm of CV from the ages of hand-engineered algorithms to complete end-to-end trainable DL models. Szegedy et al [5] proposed the CNN micro-architecture called as the “inception module” which allowed CNNs to be trained to greater depths yielding greater accuracies on complex problems. In this, they branched the inputs to the module into different sized kernels and later concatenated the feature maps depth-wise. Gaining inspiration from the micro-architecture models, [6] and later [7] were designed by Kaiming He et al. These worked on the concepts of residual blocks and identity mappings where the input to a series of convolution operations was added to the result of those operations. This solved the problem of vanishing gradients and allowed CNNs to be trained to depths of 100 or even 1000 layers. Today CNNs have massively increased in strength and

are able to tackle complex applications such as medical image diagnosis of novel diseases [8].

Mask R-CNN [9] is one such modern deep learning architecture which is able to perform object detection and image segmentation, allowing detected objects to not only be localized in images using a bounding-box but also be segmentable by generating a binary mask of the particular instance of that object. It combines novel region of interest (ROI) generation networks such as the Region Proposal Network (RPN) with a Convolutional Neural Network (CNN) feature extractor and region detector by using shared parameters across the convolutional operations. This makes the entire model end-to-end trainable and reduces the time for inference while still giving robust detection results. A model trained on this State-Of-The-Art (SOTA) paradigm would be:

- Capable of generalizing to different tissue types and detect neoplastic cells in WSI all on its own.
- Robust and accurate
- Able to localize the neoplastic cells with a binary mask which can further be used to extract cell data for post-processing
- Able to save a lot of time for histopathology researchers and speed up cure development

This paper gives a detailed description of the dataset used to train our model in section 3. This section will also describe the different modules of our implementations and highlight our model training techniques. Section 4 will describe the results obtained and section 5 will present our conclusion and remarks on this project.

II. RELATED WORK

Before DL [10] made its way through to medical imaging and cell segmentation, hand-engineered algorithms were used. P. Sankaran & V. K. Asari [11] used an adaptive thresholding technique to distinguish cell boundaries and applying a low-pass pre-processing filter while [12] proposed morphological techniques to enhance the existing water-shed methods to detect cells. Grab-cut [13] was also a popular choice for foreground segmentation.

As DL advanced, it was applied to more complex applications such as [21]. Ciresan et al [14] used a deep CNN model that was able to classify each pixel of an image based on a pixel-patch from its neighbours. [15] employed the use of fully convolutional networks to train semantic segmentation model which combined the abstract features from deeper layers with the local pixel-level features of the initial layers. This model gave better results than [13] and made it operational towards arbitrary dimensions. Drozdzal et al [20] further improved the accuracy and lowered overfitting of the U-Net model by

combining short skips similar to that of ResNet [7] with the long skips. Chen et al [19] proposed a hybrid model for faster and robust segmentation of small lung cancer. This model combined 3d convolutions which learned long-ranged 3d features along with the 2d convolutions which learn the short-range local features. A similar approach was used by [16] where they used a set of mixed convolutional blocks to better generalize on the dataset. It also made sure that their model adjusted quickly to the varying scenarios without compromising on accuracy.

The SOTA Mask R-CNN was built after a series of improvement to the CNN based object detection system. In the year 2014, Girshick et al published a paper [22] which was ground-breaking in the field of object detection. This model was called R-CNN and used selective search techniques to extract region proposals to detect objects. The proposed regions would then be converted into feature maps and classified using a SVM based approach. The drawback of this method was its speed so they proposed another faster system called as the Fast R-CNN [23]. This model was end-to-end trainable and used a Region of Interest (ROI) pooling layer to filter out the large number of ROIs generated by selective search. Later Girshick et al published another improvement to the Fast R-CNN model which was named Faster R-CNN [24]. In this they came up with two innovations which was baking the Region Proposal Network (RPN) into the architecture itself which eliminated the use of selective search and used anchors to generate proposal regions. The RPN used a set of two convolutional kernels to calculate and separate the foreground from the background and generate a bounding-box for the image. This work was further improved to generate the Mask R-CNN model.

In recent years Mask R-CNN has become a highly potent choice for segmentation models. [27] used this to segment cervical cancer cells. Their model used a shallow base network for faster inference speeds. Ma et al [26] used a Mask R-CNN to segment abnormal cells by adding attention-based mechanisms to their model to generate fixed sized ROIs. These ROIs were then combined with the original ROIs to generate more accuracy masks. A similar effort was made by [25] to utilize attention-based activation maps to train on partially labelled cell datasets without losing out on accuracy.

III. METHOD

This section describes the methodology used in the paper. It gives a detailed description of the dataset, model architecture, training techniques and hyperparameter values, in the following sub-sections.

III.I Dataset Explanation

Powerful CNN architectures almost always manage to overfit the data given to them due to surface statistical irregularities [28] and subsequently do not generalize well in real-life scenarios. Hence training a DL model for the task of cell nuclei segmentation should be facilitated by a generalized dataset of cells which enables the model to perform with the same efficiency regardless of the cell type and tissue area, while not localizing to a single region of the body.

Looking at and considering all the challenges associated with training a model for nuclei segmentation, a conclusion that the dataset must be just as comprehensive as the model, was reached. Hence, this paper works with the PanNuke dataset [29] to train the model. This dataset as briefly illustrated by Fig 1, consists of 200,000 labelled nuclei across five categories; Neoplastic, Non-Neoplastic Epithelial, Inflammatory, Connective and Dead. The speciality of this dataset is that it doesn't localize the data to one part or organ of the body but rather encompasses 19 different body tissues types. This pervasiveness allows the model to combat overfitting. Furthermore, since the dataset takes into account different body tissues it saves the tedious process of training multiple models for different tissue types. According to [30], nuclei detection done in this manner also facilitates tissue phenotyping in an effective manner. For the purpose and scope of our research we decided to work with only labelled neoplastic nuclei. After performing data pre-processing and extracting the neoplastic cells, we obtained a substantial number of labelled nuclei to train our model with. We decided to split 80% of the data for training and 20% for validation after randomly sampling 490 images as the test set. Detailed descriptions of our splits are given in Table 1.

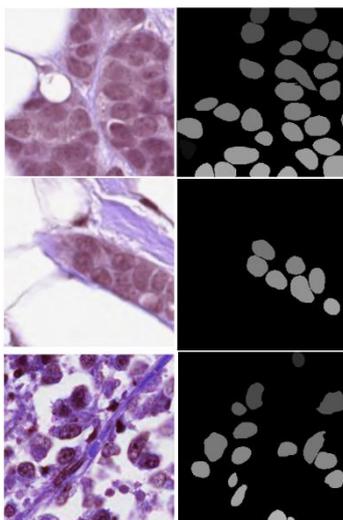


Fig 1. A collage showing the cell images (left) and instance wise segmented ground truth mask for neoplastic nuclei (right)

Table 1. Describes the number of images and their splits used for training, validation & testing

	Training	Validation	Testing	Total
Images	2950	750	490	4,190
Labelled Cell Nuclei	67,171	14,434	11,258	92,863

III.II Model Architecture

In this sub-section a brief description of the architecture of the Mask R-CNN model (Fig 2) and highlights of some of its features that enable it to perform the segmentation task is given.

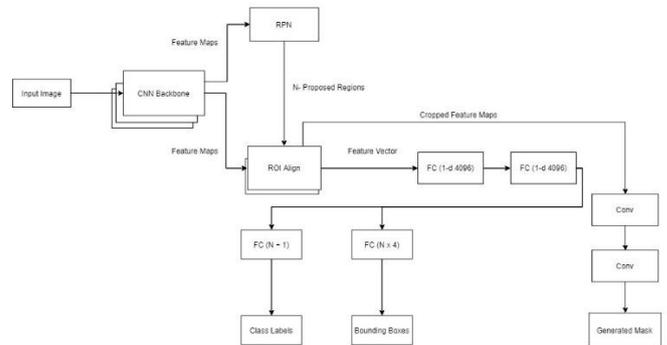


Fig 2. A flow diagram describing the various modules of the Mask R-CNN model used.

Backbone Network: The base network is responsible for feature extraction from images so it should be a fully convolutional neural network of any archetype such as ResNet or GoogLeNet. At the moment, looking at the complexity of the task, a ResNet 101 backbone is used. While it is not mandatory to use a fully convolutional network as at any time in the forward propagation process of a CNN one can just stop at a certain convolutional layer and retrieve the feature maps, it comes with the added advantage of accepting images of any spatial dimensions.

Region Proposal Network (RPN): The RPN is a small fully convolutional neural network that acts on the feature maps and generates proposals to feed into the deeper layers of the model. The RPN uses a set of anchors to generate these proposals and based on the scale and aspect ratio, a total of nine anchor points can be generated to crop proposal regions from the feature maps. This works similar to and thus alleviates the needs of having a complex feature pyramid to detect objects at varying scales. The RPN also generates an objectiveness score for each region and classifies it as either

positive, negative or neutral. Based on this score, top n proposals are selected and propagated to the next layer

Region of Interest Align Layer: This layer is an improvement over the ROI Pooling layer used in the Faster R-CNN structure. The job of this layer is to further reduce the ROIs from the RPN by cropping out fixed size feature vectors from the feature maps. This layer also facilitates the binary mask generation by aligning the proposals with the ground truth masks. After this the feature vectors propagate through three parallel layers

Fully Connected Branch (FC): This branch includes two out of three parallel propagation paths. These are dense fully-connected layers which provide two outputs at the end. One of them of size n+1 where n is the number of classes (one is added for background) provides the class labels for the dataset. The second FC output is of size 4xn and provides bounding-box coordinates namely, (x1, y1, x2, y2).

Convolutional Branch: This branch was proposed as an overhead improvement to the Faster R-CNN model by Girshick et al in [9]. This is a parallel path from the ROI Align layer which passes through a set of convolutional layers and outputs a binary mask at the end of it. The mask is then interpolated using nearest neighbour technique and overlaid on the input image.

III.III Implementation

To simplify the implementation and reduce complexity and time, Matterport's Mask R-CNN implementation [17] is used and the base classes are morphed as needed. This sub-section highlights all the training techniques and hyperparameter choices made to optimize the model. Note that the loss metrics used to train all aspects of the model (bounding-box regressors, class probabilities and mask generation) are used exactly as stated in [9] and are not mentioned explicitly.

Stochastic Gradient Descent [31] optimizer with a momentum value of 0.9 was used. To train the model, Transfer Learning [32] was used in the backbone network. To facilitate this, a pre-trained Mask R-CNN model which was trained on the MS COCO [33] dataset was used. While this dataset has no affiliation with cells, it still helps the CNN backbone network learn the features faster as the initial layers regardless of the dataset, learn similar kinds of rudimentary low-level features such as lines and shapes. While using Transfer Learning, first only the heads of the network encompassing the feature pyramid network, RPN and the Mask R-CNN bounding box and mask heads were trained for 20 epochs. This facilitates smoother overall training and lets the whole model in the later stages make better updates. Later the entire model along with the backbone was trained for 50+ epochs till the training plateaued.

Traditionally, while training model of such complexity, the learning rate (LR) is dropped by a factor of 10 each time the losses start to plateau. While this is very effective, a custom learning rate decay function was used in this paper which showed improved results in our experiments against using a sharp learning rate decay. A simple linear decay function which decayed the learning rate by a factor of approximately 16 between two sets of epochs was used. While training the heads, the LR was dropped linearly from 1e-3 to 9.9e-5 over 20 epochs. With the whole network training, the LR was dropped from 1e-4 to 3.33e-6 over 30 epochs. The formula of the learning rate decay is shown in Equation 1.

$$end_{lr} = start_{lr} * \left(1 - \left(\frac{current_{epoch}}{final_{epoch}} \right) \right)^{power} \quad (1)$$

In the Equation 1, end_lr is final decayed LR, start_lr is current LR at epoch, current_epoch is currently running epoch, final_epoch is max epoch to train model to, power is 1.0 (for decay to be linear in nature)

Since a ResNet101 backbone was used, the model started to overfit after the commencement of the second stage of training, irrespective of the size of the dataset. This is where some strong regularization techniques were used to train the model. A weight decay [34] value of 1e-4 and a gradient clipping normalization value of 5 was employed. Furthermore, Data Augmentation was used to improve validation accuracy at the expense of training accuracy. An image augmenter randomly flipped the training data along the horizontal and vertical axis and also rotated it by 10 degrees. This level of augmentation is possible to use in the Mask R-CNN model because of its translation invariance property [9]. Also, since the masks are aligned by the ROI Align layer at the end of the model, we decided to use a mini-mask function which resizes the mask to a smaller dimension to reduce computational power.

IV. RESULTS

Initial experiments: This sub-section highlights the initial experiments conducted with the Mask R-CNN model. At the beginning, the model was trained in three parts- heads, ResNet 4+ layer modules and the full model. Later it was found that training the model in two parts only gave better accuracy. Nevertheless, these results are included to show the effects of using augmentation even though the training data is sufficiently large. Fig 3 shows two stages of the training process. The training process without augmentation (Fig 3a) overfit by epoch 37, while the process using augmentation (Fig 3b) doesn't and gives considerably good results. This experiment proved the potency of data augmentation, which was used in the final experiment as well.

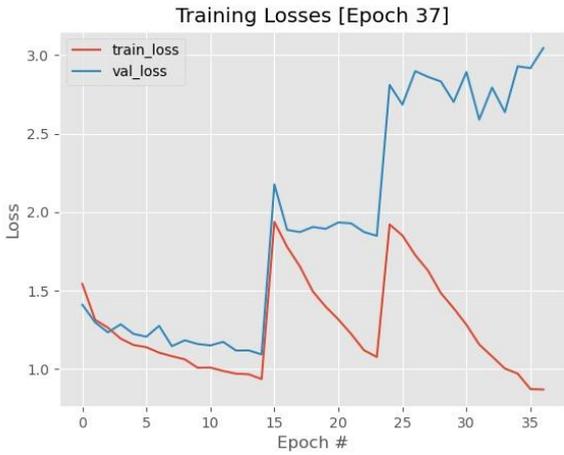


Fig 3a. Graph showing training results without the use of augmentation techniques.

compares it to the Mask R-CNN baselines given in the PanNuke paper [29].

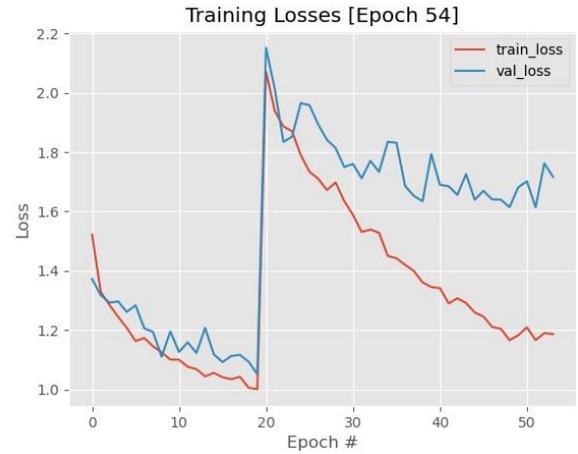


Fig 4. Final experiment results where all mentioned training techniques are used. Losses start to plateau after 47 epochs.

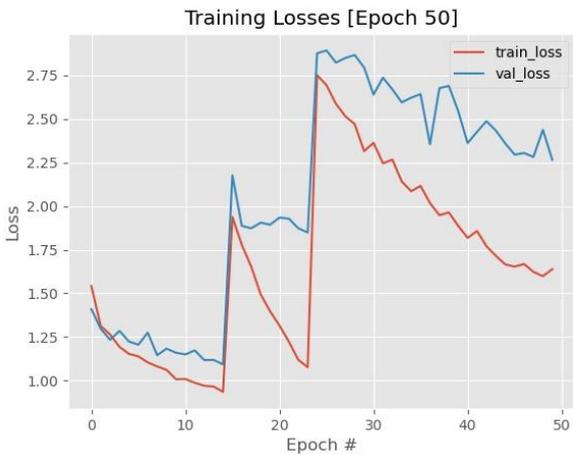


Fig 3b. Graph showing training results with the use of augmentation techniques.

Final results: Fig 4 shows the final experiment conducted which bore the best results. As shown training the model in a two-stage process generates better accuracy towards the end and the model remains relatively stable throughout the training process.

In order to evaluate the effectiveness of the model, two inference metrics, namely mean Average Precision (mAP) [35] and binary Panoptic Quality (bPQ) (a version of the PQ score that assumes all the cells belonging to one class and differentiating only between cells and background; which is sufficient for our case, having only two classes) [36] are used in this paper. The paper by S. Graham *et al* [18] explain the merits of using bPQ as being one of the modern metrics in evaluating cell nucleus segmentation tasks. For both these metric calculations only those predictions whose Intersection over Union (IoU) score was above 0.5 were considered. Table 2 describes the bPQ value across the various tissues and

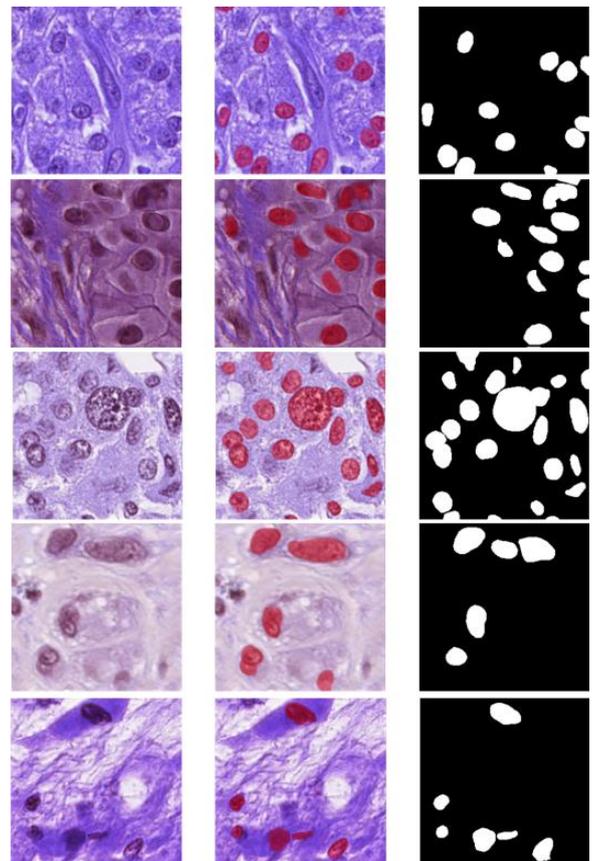


Fig 5. A collage showing output from model inference. (1st Column) Original cell nuclei image. (2nd Column) Detected and segmented neoplastic cells with an alpha mask outputted from the model. (3rd Column) Ground truth neoplastic nuclei binary mask thresholded to max pixel intensity for visibility

Table 2. Comparison table of bPQ scores of our Mask R-CNN model and the model mentioned in the PanNuke paper

Tissue Types	Our Model bPQ	PanNuke Mask R-CNN bPQ
Adrenal Gland	0.717	0.555
Bile-duct	0.652	0.557
Bladder	0.795	0.605
Breast	0.679	0.557
Cervix	0.673	0.548
Colon	0.586	0.460
Esophagus	0.740	0.569
Head & Neck	0.640	0.546
Kidney	0.657	0.509
Liver	0.742	0.609
Lung	0.612	0.513
Ovarian	0.744	0.578
Pancreatic	0.572	0.546
Prostate	0.735	0.579
Skin	0.667	0.502
Stomach	0.771	0.598
Testis	0.701	0.542
Thyroid	0.657	0.571
Uterus	0.585	0.559

We also compare the average PQ (aPQ) score of our model and the PanNuke Mask R-CNN model only for the task of neoplastic cell detection. Results are shown in Table 3

Table 3. A table showing comparison of the neoplastic cell detection average PQ score between the two models

Cell Type	Our Model aPQ	PanNuke Mask R-CNN aPQ
Neoplastic PQ	0.675	0.472

For the purpose of training this model and for inference, we used Nvidia’s RTX 2070 Super Max-Q GPU. The time taken to perform inference and segment the cell nuclei was 1.32 sec/per image.

Fig 5 shows a collage of some of the results obtained by our model.

IV. CONCLUDING REMARKS

As seen in Table 2 and 3, we think that our model shows improvement over the base Mask-RCNN model for the following stipulated reasons:

- Due to the differences in the inference technique and the data splits used
- Our model only tackles neoplastic cell nuclei while the base paper tackles four other classes
- The backbone used by us is more comprehensive than used in the base paper
- Best training practices used by us has improved the efficiency of the model

To make this research more thorough, we plan on using better and more efficient training strategies in the future. We would also test our model on more external datasets similar to PanNuke and fine-tune the model using these datasets.

Finally, we believe that the current model presented by us is very robust in its detection and should at least work as an assistive intelligent agent for researchers when they tackle the problem of neoplastic cell nuclei segmentation. The model being extremely fast in its inference time satisfies all of its proposed objectives and is a step forward in the use of intelligent systems for computational pathology.

REFERENCES

[1] <https://www.kaggle.com/c/data-science-bowl-2018>

[2] Mankar, B. S. Deshmukh and V. H., “Segmentation of Microscopic Images: A Survey,” International Conference on Electronic Systems, Signal Processing and Computing Technologies, pp. 362-364, 2014.

[3] Yongming Chen, K. Biddell, Aiyang Sun, P. A. Relue and J. D. Johnson, “An automatic cell counting method for optical images,” Proceedings of the First Joint BMES/EMBS Conference. 1999 IEEE Engineering in Medicine and Biology 21st Annual Conference and the 1999 Annual Fall Meeting of the Biomedical Engineering Society, vol. 2, p. 819, 1999.

- [4] Russakovsky, O., Deng, J., Su, H. et al, "ImageNet Large Scale Visual Recognition Challenge," *Int J Comput Vis*, pp. 211-252, 2015.
- [5] Szegedy, Christian and Liu, Wei and Jia, Yangqing and Sermanet, Pierre and Reed, Scott and Anguelov, Dragomir and Erhan, Dumitru and Vanhoucke, Vincent and Rabinovich, Andrew, "Going Deeper With Convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [6] He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [7] He K., Zhang X., Ren S., Sun J., "Identity Mappings in Deep Residual Networks," : Leibe B., Matas J., Sebe N., Welling M. (eds) *Computer Vision – ECCV 2016*. *ECCV 2016. Lecture Notes in Computer Science*, vol. 9908, 2016.
- [8] Soufi, Shervin Minaee and Rahele Kafieh and Milan Sonka and Shakib Yazdani and Ghazaleh Jamalipour, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Medical Image Analysis*, vol. 65, 2020.
- [9] He, Kaiming and Gkioxari, Georgia and Dollár, Piotr and Girshick, Ross, "Mask R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [10] Sánchez, Geert Litjens and Thijs Kooi and Babak Ehteshami Bejnordi and Arnaud Arindra Adiyoso Setio and Francesco Ciompi and Mohsen Ghafoorian and Jeroen A.W.M. {van der Laak} and Bram {van Ginneken} and Clara I., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017.
- [11] P. Sankaran and V. K. Asari, "Adaptive Thresholding Based Cell Segmentation for Cell-Destruction Activity Verification," *35th IEEE Applied Imagery and Pattern Recognition Workshop (AIPR'06)*, Washington, DC, USA, 2006, pp. 14-14, doi: 10.1109/AIPR.2006.9.
- [12] C. Di Rubeto, A. Dempster, S. Khan and B. Jarra, "Segmentation of blood images using morphological operators," *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, vol. 3, pp. 397-400, 2000.
- [13] Ronneberger, R. Bensch and O., "Cell segmentation and tracking in phase contrast images using graph cut with asymmetric boundary costs," *IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 1220-1223, 2015.
- [14] Ciresan, Dan C. and Giusti, Alessandro and Gambardella, Luca M. and Schmidhuber, Jrgen, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks," *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2013*, pp. 411-418, 2013.
- [15] Long, Jonathan and Shelhamer, Evan and Darrell, Trevor, "Fully Convolutional Networks for Semantic Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [16] C. Huang, H. Ding and C. Liu, "Segmentation of Cell Images Based on Improved Deep Learning Approach," *IEEE Access*, vol. 8, pp. 110189-110202, 2020.
- [17] Waleed Abdulla, "Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow", *GitHub repository*, 2017.
- [18] S. Graham, Q. D. Vu, S. E. A. Raza, A. Azam, Y. W. Tsang, J. T. Kwak, and N. Rajpoot, "Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images," *Medical Image Analysis*, vol. 58, p. 101563, 2019.
- [19] W. Chen, H. Wei, S. Peng, J. Sun, X. Qiao and B. Liu, "HSN: Hybrid Segmentation Network for Small Cell Lung Cancer Segmentation," *IEEE Access*, vol. 7, pp. 75591-75603, 2019.
- [20] Drozdal, Michal and Vorontsov, Eugene and Chartrand, Gabriel and Kadoury, Samuel and Pal, Chris, "The Importance of Skip Connections in Biomedical Image Segmentation," *Deep Learning and Data Labeling for Medical Applications*, pp. 179-187, 2016.
- [21] Prasanna Porwal et al, "IDRiD: Diabetic Retinopathy – Segmentation and Grading Challenge," *Medical Image Analysis*, vol. 59, 2020.
- [22] Girshick, Ross and Donahue, Jeff and Darrell, Trevor and Malik, Jitendra, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [23] Girshick, Ross, "Fast R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015.
- [24] Ren, Shaoqing and He, Kaiming and Girshick, Ross and Sun, Jian, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in Neural Information Processing Systems*, vol. 28, pp. 91-99, 2015.

- [25] L. Feng, J. H. Song, J. Kim, S. Jeong, J. S. Park and J. Kim, "Robust Nucleus Detection With Partially Labeled Exemplars," *IEEE Access*, vol. 7, pp. 162169-162178, 2019.
- [26] B. Ma, J. Zhang, F. Cao and Y. He, "MACD R-CNN: An Abnormal Cell Nucleus Detection Method," *IEEE Access*, vol. 8, pp. 166658-166669, 2020.
- [27] Kurnianingsih et al, "Segmentation and Classification of Cervical Cells Using Deep Learning," *IEEE Access*, vol. 7, pp. 116925-116941, 2019.
- [28] Bengio, Jason Jo and Yoshua, "Measuring the tendency of CNNs to Learn Surface Statistical Regularities," *CoRR*, vol. abs/1711.11561, 2017.
- [29] Gamper, Jevgenij and Koohbanani, Navid Alemi and Graham, Simon and Jahanifar, Mostafa and Khurram, Syed Ali and Azam, Ayesha and Hewitt, Katherine and Rajpoot, Nasir, "PanNuke Dataset Extension, Insights and Baselines," *arXiv preprint arXiv:2003.10778*, 2020.
- [30] Rajpoot, Sajid Javed and Arif Mahmood and Muhammad Moazam Fraz and Navid Alemi Koohbanani and Ksenija Benes and Yee-Wah Tsang and Katherine Hewitt and David Epstein and David Snead and Nasir, "Cellular community detection for tissue phenotyping in colorectal cancer histology images," *Medical Image Analysis*, vol. 63, 2020.
- [31] J. Kiefer. J. Wolfowitz. "Stochastic Estimation of the Maximum of a Regression Function." *Ann. Math. Statist.* 23 (3) 462 - 466, September, 1952. <https://doi.org/10.1214/aoms/1177729392>
- [32] Torrey, Lisa and Jude Shavlik. "Transfer Learning." *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, edited by Emilio Soria Olivas, et al., IGI Global, 2010, pp. 242-264. <http://doi:10.4018/978-1-60566-766-9.ch011>
- [33] Lin TY. et al. (2014) Microsoft COCO: Common Objects in Context. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) *Computer Vision – ECCV 2014*. *ECCV 2014. Lecture Notes in Computer Science*, vol 8693. Springer, Cham. https://doi.org/10.1007/978-3-319-10602-1_48
- [34] Anders Krogh, John A Hert. "A Simple Weight Decay Can Improve Generalization". In *Advances in Neural Information Processing Systems 4*, JE Moody SJ Hanson and RP Lippmann eds, Morgan Kauffmann Publishers, San Mateo CA, 1992
- [35] Beitzel S.M., Jensen E.C., Frieder O. (2009) MAP. In: LIU L., ÖZSU M.T. (eds) *Encyclopedia of Database Systems*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_492
- [36] A. Kirillov, K. He, R. Girshick, C. Rother, and P. Dollár, "Panoptic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9404–9413, 2019.