International Journal of Arts Humanities and Social Sciences Studies
Volume 10 Issue 05 ǁ May 2025
ISSN: 2582-1601
www.ijahss.com

Using deep learning approaches to segment, detect, and extract the cellular nuclei from fluorescence microscopy images

Zhixuan WANG, Yunkun SONG
Aquinas International Academy
Sen delta International Academy

Abstract—*The precise segmentation of cell nuclei in fluorescence micros copy images plays a multitude of key roles including characterizing cells, quantifying the nuclei, identifying cancerous cells and measuring the expression level of a gene. Although there is significant research conducted on the use of convolutional neural networks for nuclei segmentation, there is a need for a comprehensive study on how these models work on a common data set for bench marking purposes. Also, in order to evaluate them in the real world, it is imperative to test their performance on unseen data using a common metric. In this paper, I compare the performance of a wide variety of neural network architectures proven to perform well on cell nuclei segmentation tasks.*

Index Terms—*Fluorescence microscopy, Mask R-CNN, nuclei segmentation, U-Net.*

I. Introduction

Fluorescence micros copy is an optical microscopy technique that uses fluorescence to image 3D sub cellular structures [1,2]. It is an indispensable tool lying in a biologist’s arsenal, enabling the system aticspatiotemporald is section of life’s molecular machines [3]. Accurate three-dimensional segmentation of biological specimens is important to characterize and quantify cells, nuclei, or other microscopic structures. The accurate segmentation also contributes to cancer research, since cancerous cell shave differently shaped nuclei owing to them being in a constant state of mitosis (always dividing). Segmentation also plays a key role in high-speed and high-performance applications such as the analysis of the role of cells and the expression level of agene, which can be calculated by measuring the transcribed M RNA and the expressed protein [4]. One such task is the problem of cell nuclei segmentation. The accurate segmentation of nuclei can contribute to the development of a system that is successful in automating the analysis of microscopic images for pathology detection. The major characteristics of tumors are linked with the nuclei of the cells. It is hence vital to segment the image, to isolate the nuclei from the rest of the cytoplasm and elements coming from preparation of the image, similar to the staining ink [5].

The segmentation of cells or the nuclei of cells is a vital step with regards to the analysis of biomedical microscopy images [6]. Multiple medical analysis can be performed based on accurate cell nuclei segmentation, including studying the observable characteristics of cells[7],the segmentation of blood cells [8] and cell type classification [9]; which aids in providing essential diagnostic information to researchers and doctors a like. However, considering how time-intensive and labor con summing this task is, the conventional image processing techniques still being employed for it are often unable to achieve the optimal performance due to a wide array of reasons, such as the inability to deal with diverse images [6].

In recent years, with an explosion in the development of deep learning-based techniques, convolutional neural networks have had astounding success- regularly surpassing human performances–for hard tasks such as speech recognition [10] and vision [11]. Practically every discipline in science and engineering, ranging from astronomy to high energy physics [12] has been greatly impacted. CNNs are particularly relevant and vital in fields that interpret or produce images. An example would be medical imaging, which applied deep CNNs to image processing of magnetic resonance images and computed tomography [13,14]. Fluorescence microscopy, too, has greatly benefitted from these advances with deep CNNs being applied to deconvolution, super-resolution, as well as image segmentation, classification, and phenotyping [15].

Convolutional neural network shave been used by scientific researchers for many years with the first successful application being the LeNet [22] for hand-written digits recognition. CNNs have been widely used in segmentation problems. For instance, in [24] a segmentation technique was developed that used a CNN structure with a max-pooling operation (which preserves features while down sampling the feature maps) to segment neuron membranes for electron microscopy images. The use of a convolutional neural network with a support vector machine to detect Tyrosine Hydroxylase-containing cells in zebrafish brain images from wide-field microscopy is investigated in [25]. AU-Net using two- dimensional convolutional neural networks with an encoder-decoder architecture to segment neuronal structures in electron microscopy and cells in light microscopy is also proposed [26]. A nuclei segmentation method on his to pathology images using deep CNN with selection-based sparse shape model is described in [27].

However, there is a need for a comprehensive study on how these models work on a common dataset for bench marking purposes. Also, in order to evaluate them in the real world, it is imperative to test their performance on unseen data using a common metric.

In this paper, I survey the different available types of neural networks for cell nuclei segmentation. The experiments were run on the data set provided by the Kaggle 2018 Data Science Bowl, which challenged participants to segment the nuclei in 2D microscopy images without any manual alteration. The participants were provided with a training set of images containing the nuclei long with the corresponding solutions (segmentation masks for the nuclei) and test sets of images for which they had to generate the segmentations. The nuclei segmentation methods we real so challenged to generalize by operating on unseen data without any additional annotation/ training.

II. Related Work

Image segmentation has a multitude of applications in the field of medical analysis, including image processing, facial recognition, and wayfarer detection [16].

Image segmentation has two main objectives.

1. The first objective is to decompose the image into parts for further analysis. In straightforward cases, the environment might be controlled well enough so that the segmentation process reliably extracts only the parts that need to be analyzed further. For example, in the domain of color, an algorithm could be used for segmenting ahu man face from a color video image. This type of segmentation would be dependable, given that the person’ s clothing or room in the background does not have the same color components as a human face.

2. The second objective is to perform a change of representation. The image pixels need to first be arranged and ordered into units of a higher level that are either more efficient or more meaning ful for further analysis. A critical issue that lies in this field is whether this image segmentation can be performed for multiple different domains using general bottom-up methods that do not use any special domain knowledge.

With the rapid development of deep learning-based neural network architectures, the potential to employ these neural networks in cell nuclei segmentation is being investigated by multiple researchers who are incorporating such techniques into their work. For example, the implementation of an U-Net encoder decoder architecture for faster image segmentation was proposed by Ronne berger et al. Developed along with tensor flow[17], being used for Radio Frequency Inter Ferenc mitigation using deep convolutional neural networks in the original implementation. The use of VGG-style neural networks combined with feature pyramid networks [18] to predict the masks of cells was proposed by C. Hernándezet.al. T. Tran et. al. and achieved an accuracy of 89.45%whileusingaSeg-Net [8] to segment the different types of blood cells in smeared peripheral images of blood. The use of convolutional neural networks to predict region so fab normality in the cells in the Pap smear test was proposed by F. Araújoet.al [19]. As reported in [6], in the domain of cell nuclei segmentation or the segmentation of multiple cells, the U-Net [20] encoder decoder architecture-based algorithms are the most commonly used as they are usually the ones to achieve state-of-the-art segmentation results with relative ease. For instance, a method that employed both a U-Net and a Mask R-CNN was proposed by

R. Hollan diet. al. [21] to predict the segmentation of the nuclei of cells. The accuracy of this algorithm out performed that of the other 739 other submissions to the 2018 Kaggle Data Science Bowl. The algorithm that achieved first place also used the U-Net as a fundamental baseline model.

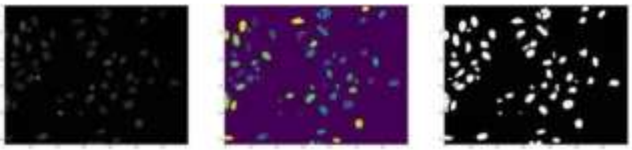

Fig.1.Samplecellnucleisegmentation

III. Methodology

A. *Data set Description*

The 2018 Data Science Bowl Kaggle dataset contains a large number of segmented cell nuclei images. It comprises 670 training images and 65 test images. These images were acquired under a variety of conditions and vary in the cell type, magnification, and imaging modality (bright fieldvs. fluorescence).

The data set is represented by a set of folders. Each of the folders contains 2 sub-folders: “images” and “masks”. Sub-folder “images” contains a single file which is an image of a tissue scan. Sub-folder “masks” contains several. Png files —one file per an notated nucleus. A mask file is a black-white picture with the size of the original tissue scan, where white pixels depict the location of the nucleus. The original image is in the left top corner, the rest are mask images. The models were trained using the K eras deep learning API in tensor flow on google colab using a **Nvidia** T4 GPU.

B. *Models*

In this study, I choose from a wide variety of neural network architectures proven to perform well on cell nuclei segmentation tasks. In specific, we describe four models which have demonstrated high performance when evaluated on the curated dataset and apply that on unseen data.

1. Simple CNNs

Simple convolutional neural networks are powerful visual models that yield hier archies of features. Simple convolutional networks take input of arbitrary size and produce correspondingly sized output with efficient inference and learning.

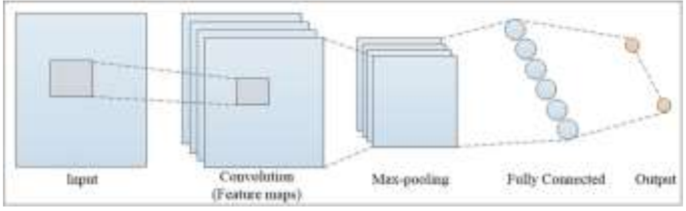

Fig.2. Simple CNN architecture

According to the basic structure of Convolutional Neural Networks, the outputs from the prior layers are used as input for the subsequent layers. This architecture ensures there use the feature sin side the model, therefore it provides better performance on different computer vision tasks which in empirically investigated on different data sets in [32]. In my nuclei segmentation model, I use one Batch Normalization layer to standardize the inputs to the next layer and reduce the number of training epochs required [33], followed by6 fully connected denselayersanda2dimensionalconvolutional output layers that gives us the predicted segmented mask for each input image. I use a kernel size of (3,3) and a sigmoid activation function for the output layer.

2. U-Net

The U-Net encoder- decoder implementation was proposed by Ronne berger et al. developed along with tensor flow [34]. The code has been developed and used for Radio Frequency Inter Ference mitigation using deep con volutional neural networks [39]. The network can be trained to perform image segmentation on arbitrary imaging data. The model starts with blocks of con volution layers and max pooling layers followed by deconvolution layers and recreating the image step by step. The output is similar to input image size with segmented or labeled masks.

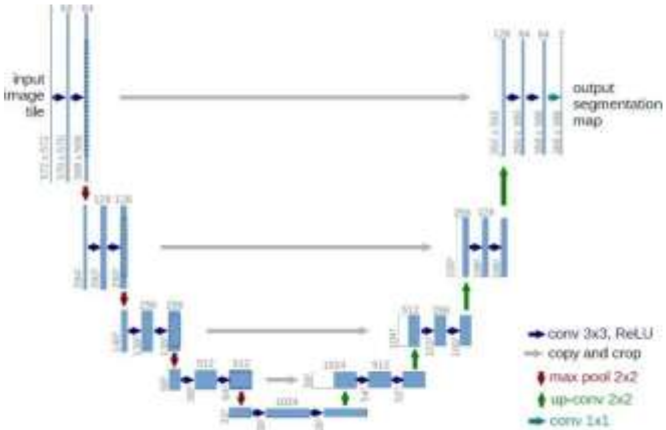

Fig.3. U- Net encoder decoder architecture

The U-Net follow sacon volve-resample process. The Convolution layers convolve through the image and geta detailed observation of the image. The model can be made more detailed by increasing the Convolution layers. I also used a few max pooling layers to consider maximum values and to reduce architecture size. The Deconvolution layers are generated by using the convolution transpose function in tensor flow which per forms the reverse of convolution. The output is generated by gradually recreating the image using the con volution layers used previously. I use 19 layers of convolution and 4 deconvolution layers in my model. I use a sigmoid activation function for the output layer and ana dam optimizer since using Adagra das the optimizer resulted in early stopping as the loss was stagnant and the optimizer was not suitable for the particular problem in hand.

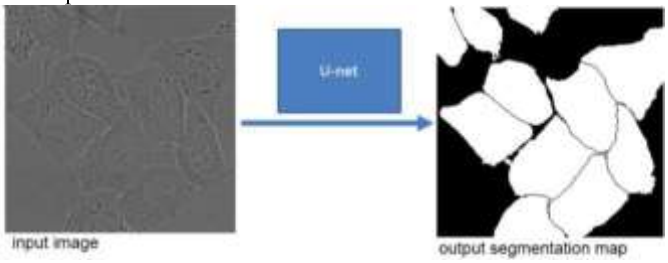

Fig.4.SamplesegmentationbyU-Net

3. Res Net

Deeper neural networks are more difficult to train. A residual learning framework used to ease the training of networks that are substantially deeper than those used previously. Res Nets learner sidual functions with reference to the layer inputs, instead of learning unreferenced functions. Instead of hoping each few stacked layers directly fit a desired underlying mapping, they stack residual blocks on to pof each other to form network: e.g. aResNet-152hasonehundred and fifty- two layers using these blocks.

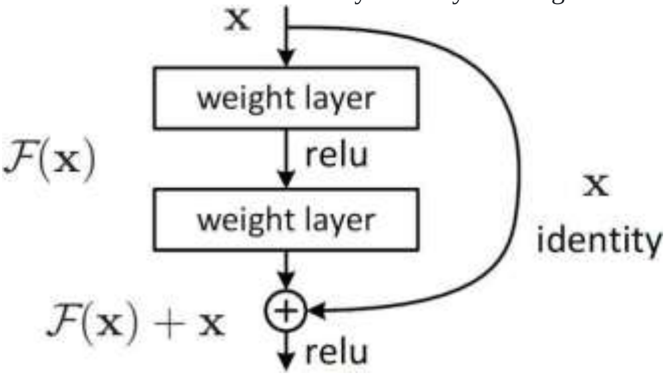

Fig.5.ResNet architecture

Comprehensive empirical evidence has been provided showing that these residual networks are easier to optimize and can gain accuracy from considerably increased depth [35]. The depth of representations is extremely important for visual recognition tasks, helping Res Nets obtain a 28% relative improvement on the COCO objected tection dataset. There are 20000 steps in training and check points are created at then do f1000 th step. The model uses pretrained model weights done for COCO dataset.

4. MaskR-CNN

Mask R-CNN is essentially annex tension of Faster R- CNN. Faster R-CNN is widely used for object detection tasks. For a given image, it returns the class label and bounding box coordinates for each object in the image. The Mask R-CNN framework is built on top of Faster R-CNN. So, for a given image, Mask R-CNN, in addition to the class label and bounding box coordinates for each object, will also return the object mask.

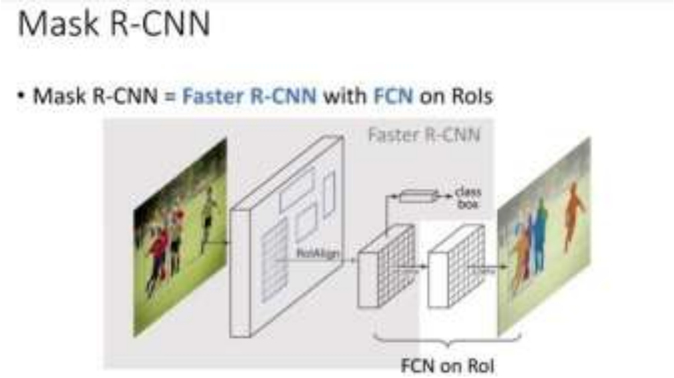

Fig.6.MaskR-CNN architecture

Mask R-CNN extends Faster R-CNN [36] by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. When run without modifications on the original Faster R-CNN architecture, the Mask R- CNN authors [37] realized that the regions of the feature map selected by RoI Pool were slightly misaligned from the regions of the original image, leadings to inaccuracies. The model uses pretrained model weights from COCO dataset.

C. *Metrics*

$$I(X) = \sum_{v \in V} X_v \times Y_v$$

$$U(X) = \sum_{v \in V} X_v + Y_v - X_v \times Y_v$$

$$IoU = \frac{I(X)}{U(X)}$$

$$loss = 1.0 - IoU$$

Fig.7.IoUmetricequation

Inter section over union is a common metric for assessing performance in semantic segmentation tasks. In a sense, IoU is to segmentation what an F1 score is to classification. Both are non-differentiable, and not normally optimized directly. Optimizing cross entropy loss is a common proxy for these scores, that usually leads to decent performance, provided everything else has been setup correctly,
e.g regularization, stopping training at an appropriate time.

IV. Experiment Sand Results

The training set I used comprises 670 images and my testsetconsistsof65images.Thedetailed evaluations of all the models can be seen in Table1 below.

The metric I used to evaluate the performance of different neural network architecture son the segmentation data set is the Inter section over union (IoU) measure [38]. IoU is known to be a good metric for measuring the overlap between two bounding boxes or masks, hence chosen for this task of cell nuclei segmentation. If the prediction is completely correct, IoU= 1. The lower the IoU, the worse the prediction result.

Table I Comparison of Results on Dataset

Method	Training acc.	Test acc.	mI OU score
Simple CNN	93.10 %	93.70%	62%
Res Net	96.30%	96.10%	77%
U-Net	99.75%	98.00%	83%
MaskR-CNN	99.68%	99.20%	88%

Using the simple CNN architecture, the nuclei were spottedandatrainingaccuracyof93.10%,test accuracyof93.70%andmeanIoUscoreof62%was achieved. Although this was the lowest performing model, I showed that convolutional networks by themselves, trained end-to-end, pixels-to-pixels, exceed the state of-the-art in semantic segmentation.

While CNNs work well with images, they lack efficiency with larger datasets, so I tried to apply a segmentation model on the dataset.

The U-Net encoder-decoder architecture was the first segmentation model I applied. It proved to be the second-best performing model for nuclei segmentation. The Mean Intersection over Union scorewas83%witha99.75%trainingaccuracyand a 98% test accuracy as in Table 1. The contraction and expansion layers are convolution and up sampling layers and hence the image is recreated with segmented masks similar to the image input size.

The Mask R-CNN model performed better than the U-Net model, giving a higher mean IoU score of 88% and an accuracy of 99.20% on the test dataset. The masks generated were clearer than those generated from the U-Net architecture.

Finally, the Res Net model generated nuclei masks for select images and it was less efficient than the U- Net and the Mask R-CNN architectures, with a mean IoU score of 77% and an accuracy of 96.10% on the test dataset. This model was very resource exhaustive and took a day to complete training the network.

V. Conclusion And Future Work

The main goal of this problem was to understand various segmentation models and compare their efficiency to find the best suite done for the given problem. The architectures I tried are CNN, FCN, U-Net, Mask RCNN and Resnet-101.The MaskR- CNN architecture outperformed the other models in the task of cell nuclei segmentation as seen in Table 1.Theimageswiththemasksweregeneratedclearly and the segments were spotted correctly. The way I could improve on this problem is to spend time on gathering more training data given the split of training and test images.

The underlying problem with these techniques however is that while they produce good results in 2 Dimages, these techniques cannot utilize the z- directional information in a volume. Furthermore, one of the most daunting underlying problems in using convolutional neural networks is the extremely large amount of training images required. One of the ways to address this issue is through the use of data augmentation methods where linear and nonlinear transforms are done on the training data to create “new” training images. Some of the typical transformations include spatial flipping, warping and other deformations or contrast transformation [26,28]. While these methods may help, there is a still limitation if there are only a few 3D ground truth image volumes.

Moreover, distinguishing the segmented nuclei from the different sub cellular structures present in the cell is aided by the use of convolutional neural networks when they are trained with many training images where the nuclei are manually segmented and annotated thus providing 2-dimensional ground truth image volumes. Al though the a forementioned deep learning algorithms demonstrate efficiency in the task of cell nuclei segmentation and are considered state-of-the-art segmentation algorithms, they rely heavily on complex operations such as difficult bagging techniques of processing of images and model ensemble, which make the applications of such algorithms in real clinical situations nearly impossible and thus impede their use fulness. This is because clinics are often equipped with very little computing power. For instance, the proposed method in [29] employs two U- Net encoder-decoder architectures to make the final mask predictions, which helps increase the accuracy significantly while paying the price of a higher computing load during inference. Further, the extensive use of simple bright- or dark-field microscopy can generate large number of images which need to be processed and analyzed rapidly, which requires the segmentation algorithms to be light-weighted while maintain in ga high accuracy [30], with the ability of running fast with constrained computing resources, since transferring a large number of images to the cloud server will generally cause along uploading time [31]. Therefore, it is still necessary to develop new deep learning-based algorithms to increase the accuracy of segmentation, especially with limited computing power.

References

[1] *Vonesch, Cédric, et al. "The colored revolution of bioimaging." IEEE signal processing magazine 23.3 (2006): 20-31.*

[2] *Dunn, Kenneth W., et al. "Functional studies of the kidney of living animals using multicolor two-photon microscopy." American Journal of Physiology-Cell Physiology 283.3 (2002): C905-C916.*

[3] *Lichtman, Jeff W., and José-Angel Conc hello. "Fluorescence micros copy." Nature methods 2.12 (2005): 910-919.*

[4] *Gharipour, Amin, and Alan Wee-Chung Liew. "Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting." Pattern Recognition 58 (2016): 1-11.*

[5] *Hiremath, P. S., and Y. Humnabad Iranna. "Automated cell nuclei segmentation and classification of squamous cell carcinoma frommicroscopicimagesofesophagustissue."2006InternationalConferenceon Advanced Computing and Communications. IEEE, 2006.*

[6] *Caicedo, JuanC.,etal."Evaluationofdeeplearningstrategiesfornucleussegmentation in fluorescence images." Cytometry Part A 95.9 (2019):952-965.*

[7] *Bougen‐Zhukov, Nicola, et al. "Large‐scale image‐based screening and profiling of cellular phenotypes. "CytometryPartA91.2(2017):115-125.*

[8] *Tran, Thanh, et al. "Blood cell images segmentation using deep learning semantic segmentation." 2018 IEEE International Conference on Electronics and Communication Engineering (ICECE). IEEE, 2018.*

[9] *Liu, Ying, and Feixiao Long. "Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning." ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Springer, Singapore,2019. 113-121.*

[10] *Choi, Keunwoo, et al. "Convolutional recurrent neural networks for music classification." 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2017.*

[11] *Szegedy, C., Iofe, S., Vanhoucke, V. & Alemi, A. A. Inception-v4, inception-res net and the impact of residual connections on learning. AAAI 4, (12 (2017).*

[12] *Radovic, Alexander, et al. "Machine learning at the energy and intensity frontiers of particle physics." Nature 560.7716 (2018): 41-48.*

[13] *Xu, Yan, et al. "Deep learning of feature representation with multiple instance learning for medical image analysis."2014 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2014.*

[14] *Jin, Kyong Hwan, et al. "Deep convolutional neural network for inverse problems in imaging." IEEE Transactions on Image Processing 26.9(2017): 4509-4522.*

[15] *Moen,Erick,etal."Deeplearningforcellularimageanalysis."Naturemethods (2019): 1-14.*

[16] *Yuheng, Song, and Yan Hao. "Image segmentation algorithms overview." ar Xiv preprint arXiv:1707.02051 (2017).*

[17] *Falk, Thorsten, et al. "U-Net: deep learning for cell counting, detection, and morphometry." Nature methods 16.1 (2019): 67-70.*

[18] *Hernández, Carlos X., Mohammad M. Sultan, and Vijay S. Pande. "Using deep learning for segmentation and counting within microscopy data." ar Xiv preprint arXiv:1802.10548 (2018).*

[19] *Araújo, Flávio HD, et al. "Deep learning for cell image segmentation and ranking." Computerized Medical Imaging and Graphics 72(2019): 13-*

21.

[20] *Ronne berger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for medical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.*

[21] *Hollandi, Reka, et al. "A deep learning framework for nucleus segmentation using image style transfer." Bio Rxiv (2019): 580605.*

[22] *LeCun, Yann, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86.11 (1998): 2278-2324.*

[23] *Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Image net classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.*

[24] *Ciresan,Dan,etal."Deep neural networks segment neuronal membranes in electron microscopy images." Advances in neural information processing systems. 2012.*

[25] *Dong, Bo, et al. "Deep learning for automatic cell detection in wide-field microscopy zebra fish images." 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI). IEEE, 2015.*

[26] *Ronne berger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.*

[27] *Xing, Fuyong, Yuanpu Xie, and Lin Yang. "An automatic learning-based framework for robust nucleus segmentation." IEEE transactions on medical imaging 35.2 (2015): 550-566.*

[28] *Fu, Chichen, et al. "Nuclei segmentation of fluorescence microscopy images using convolutional neural networks." 2017 IEEE 14thInternational Symposium on Biomedical Imaging (ISBI 2017). IEEE,2017.*

[29] *Our Solution, 4th Place on the Private LB. https://www.kaggle.com/c/data-science - bowl- 2018/discussion/55118\#latest-527734. Accessed May 2019.*

[30] *Long, Feixiao. "Microscopy cell nuclei segmentation with enhanced U-Net." BMC bioinformatics 21.1 (2020): 1-12.*

[31] *Shi, Weisong, et al. "Edge computing: Vision and challenges. "IEEE internet of things journal 3.5 (2016): 637-646.*

[32] *Huang,Gao,et al. "Densely connected convolutional networks. "Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.*

[33] *Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift."ar XivpreprintarXiv:1502.03167 (2015).*

[34] *Ronne berger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.*

[35] *He,Kaiming,etal."Deepresiduallearningforimagerecognition."ProceedingsoftheIEEEconferenceoncomputervisionandpattern recognition. 2016.*

[36] *Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.*

[37] *He, Kaiming, et al."Maskr-cnn."Proceedingsof the IEEE international conference on computer vision. 2017.*

[38] *Reza to fighi, Hamid, et al. "Generalized inter section overunion: Ametric and a loss for bounding box regression." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.*

[39] *Akeret, Joel, et al. "Radio frequency interference mitigation using deep convolutional neural networks." Astronomy and computing 18 (2017):35-39.*