SegMed : Analysis on Automated Nuclei Segmentation Methods

Akhil Mokkapati Naga Anjaneyulu Kopalle Sai Sugeeth Kamineni Vijay Sai Kondamadugu
Indiana University Bloomington
{akmokka, nakopa, skaminen, vikond}@iu.edu
https://github.com/naga-anjaneyulu/SegMed

Abstract

In the medical domain, nuclei segmentation is one of the key steps in various disease diagnosis and drug discovery processes because the nucleus is a fundamental particle of the basic living unit i.e., cell. Medical researchers are highly dependent on the quality of nuclei segmentation to observe any biological reaction on a living organism. Automated nuclei detection techniques are rapidly developing but these methods are confined to organ-specific or image specific such as specific stain based images. In this report, we provide the analysis of a few nuclei segmentation techniques and their feasibility on multi-organ data sets along with performance, time, and infrastructure requirements. One of the main motivations of this review is to contribute to the Human Bio-molecular Atlas Program (HuBMAP). HuBMAP is to develop an open and global platform to map healthy cells in the human body.

1. Introduction

In humans, the proper functioning of organs and tissues is dependent on interaction and spatial organization of all the cells. Scientists estimate that there are 30+ trillion cells in a human body, so the HuBMAP project which is aided by the National Institutes of Health has planned to come up with relationships among these cells in an adult human body. HuBMAP will build the framework necessary to construct the tools, resources, and cell atlases needed to determine how the relationships between cells can affect the health of an individual. This analysis of nuclei segmentation methods to detect nuclei of any organ from a histopathological image is an initial contribution for over a 7-year long project.

Digitally detecting nuclei using histopathology images has attracted much attention in both research and clinical practice because the manual assessment is labor-intensive and prone to inter-observer errors. Although already there are different methods of automated histopathological imaging techniques existing, we aim at reviewing the methods that work for multi organs tissues. It is very difficult to achieve robust and accurate nucleus segmentation for various reasons. First, these microscopic images are cluttered with noise and blurry regions. Next, when it comes to different organs, there are variants in size, shape, and inter-cellular region features. Sometimes, there are even overlaps of nuclei in these images.

In this paper, we explore different techniques that cover both conventional image processing techniques as well as deep learning approaches on nuclei segmentation of digital microscopic histopathological images from three different data sources. We will discuss each of these methods elaborately along with their results. Although the same data sets are used for all the listed different methods, each method further employs preprocessing techniques based on its requirements. We will broadly focus on the recent Convolutional Neural Network (CNN) based techniques used for segmentation.

2. Background and related work

Most state-of-the-art nucleus segmentation techniques use watershed segmentation, morphological processing, color-based thresholding, active contours,

and their variants along with a multitude of pre and post-processing techniques. However, such methods fail to generalize across a wide spectrum of tissue morphologies due to inter and intra nuclear color variations in crowded and chromatin sparse nuclei.

Other techniques are based on machine learning as they can be trained to recognize nucleus shape and color variations. One type of learning-based methods use hand-engineered features such as color-texture, blue ratio, color histograms, Laplacian of Gaussian response, geometric features from the gradient or Hessian profiles and other image characteristics in standard learning-based models to segment nuclei and non-nuclei regions [1]-[4].

Another set of learning-based methods use CNNs, which have outperformed hand-crafted feature-based techniques. These techniques estimate a probability map of the nuclear and non-nuclear (two-class) regions based on the learned nuclear appearances [5]-[6] but have not demonstrated to work on multi-organ tissue images. In CNN based techniques there has been a further improvement to segment nuclei on multi-organ data by using advanced architectures that explicitly find inter-nuclear boundaries using the third class of pixels to separate crowded nuclei [7]-[8].

3. Methods

3.1. Data description

We have done our analysis of multi-organ nuclei segmentation on three different data sets.

The first one is MoNuSeg [12] data which has annotated tissue images of several patients with tumors of different organs. This data set is organized into training and testing data sets as it is part of MoNuSeg Grand Challenge. Training data of MoNuSeg contains 30 images and around 22000 nuclear boundary annotations and the test set has 14 images with 70000 nuclear boundary annotations. The annotations of this data set are not in an image format but are provided as XML files for each image. We have parsed through XML of each image to create masks for our detection methods.

The second data set is PSB [10] which is a crowd sourced image annotation for nucleus detection and segmentation that includes annotations from experts, automated methods, and the crowd. This data set has 810 histopathological images annotated in the form of binary masks. Although it has expert level annotations they are confined to only 10% of the data set. Since we wanted to use a complete set of 810 images, we have used annotations done by mid-level contributors.We have chosen contributor level 2 data set since it has few ground truth discrepancies.

The third data set, "TNBC," used in our work was first introduced in Naylor, et al. [11]. This data set consists of H&E stained, triple-negative breast cancer (TNBC) tissue slide images of 11 different patients. Here the images of size 512 X 512 were cropped from whole-slide images each to generate a total of 50 images.

3.2. Data pre-processing

3.2.1 MoNuSeg

As mentioned earlier training data of MoNuSeg contains 30 images and around 22,000 nuclear boundary annotations and the test set has 14 images with 70,000 nuclear boundary annotations. These manual annotations for nucleus in XML files have been used to generate segmentation masks by drawing polygons using the annotated coordinates.

Once we have the segmented masks, we have performed canny edge detection to detect nucleus edges and draw boundaries around each nucleus and generate the segmented masks with nucleus boundaries.Then , we label the background, boundaries, and nucleus pixels as 0, 1, 2 respectively, and generate single-channel images. These images are used as ground truth for training our unsupervised-segment model, nuclei boundary model, and mask-RCNN model.

3.2.2 PSB

This data set already has 810 histopathological images annotated in the form of binary masks, so we performed canny edge detection and pixel-level classification to generate ground truth for training UNet2, Unsupervised-SegMed model, NB model.

3.2.3 TNBC

This data set consists of H&E stained, with a total of 50 images and their corresponding binary masks, so we performed canny edge detection and pixel-level classification to generate ground truth for training Unsupervised-SegMed model, Nuclei Boundary model.

Further, each model employs various other data processing techniques for augmentation, generating patches which are discussed later.

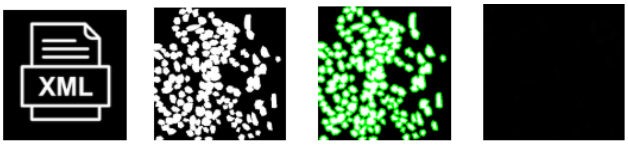


Figure 1: Data preprocessing steps: (a) XML annotations for nuclei, (b) Generated binary mask, (c) Canny Edge detection, (d) Pixel labelling (this image seems to be black because it has pixel values 1-3 for 3 classes

3.3. Segmentation Models

From previous work we can observe that nuclei segmentation methods can be broadly categorized as unsupervised/image processing based, Machine learning based which include CNNs in addition to two or three class problems. In this paper, we briefly discuss each of these methods which are trained on all the data sets.

3.3.1 Unsupervised method

This unsupervised technique consists of four steps: normalization, unsupervised color deconvolution, intensity thresholding, and post-processing to segment nuclei from the background. Figure 1 shows sample outputs after each of this step. This unsupervised method is adapted using the library HistomicsTK, specially developed for histology images and available at `https://github.com/DigitalSlideArchive/HistomicsTK`.

a. Color normalization: We employ Reinhard color normalization [14] to convert the color characteristics of all images into standard form by computing the mean and standard deviations of a target image in LAB space using a fixed reference image.

b. Color deconvolution: PCA-based 'Macenko'[15] method is used to perform unsupervised color deconvolution to separate the normalized image into two stains, i.e., hematoxylin and eosin. We project pixels onto a best-fit plane, wherein it selects the stain vectors as percentiles in the angle distribution of the corresponding plane. These two stain images are separated from the normalized image with the correct stain matrix for color deconvolution.

c. Intensity thresholding: To segment nuclei properly, we apply intensity thresholding in the eosin stain image wherein the intensity distribution is homogeneously distinct from the background. By converting the eosin stain image into a binary image with constant global thresholding, the cells are roughly segmented.

d. Post-Processing: In this step, objects with fewer pixels than minimum area threshold will be removed from the binary image. Then to remove small/thin protrusions, minimum apriori information is used to detect contours independent of their shape as done in [16]. To further improve the results, a morphological opening operation is applied to separate few touched nuclei.

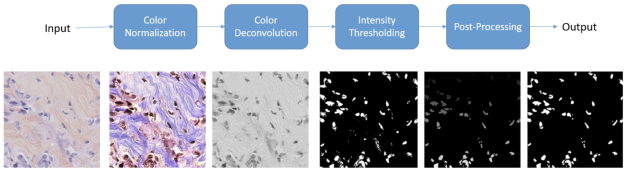


Figure 2: This figure shows the steps used in unsupervised model. From left to right, a) Input tissue image, b) Reinhard color normalized image, c) eosin stain image after deconvolution, d) after intensity thresholding, e) removing objects with minimum nucleus area threshold, f) final output of this unsupervised method

3.3.2 U-Net 2 Class

U-Net [17] is a highly successful architecture for instance based segmentation problems. This model uses the U-Net architecture to classify the foreground and backgrounds of medical images. The encoding layers are used to extract different levels of contextual feature maps. The decoding layers are designed to combine these feature maps produced by the encoding layers to generate the desired segmentation maps. Loss Function is Binary Cross Entropy. Evaluation Metric is Mean-IOU. Early stopping is also used on validation metrics so as to stop over-fitting. The architecture can be seen in Figure 3. The first part functions just like a traditional CNN but the second part start to upscale the image using the feature maps found, with finally ending up with an output map of same as size as that of the input. This models tries to predict only two classes, foreground (nuclei) and background (cytoplasm).

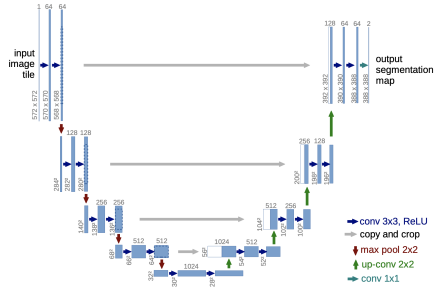


Figure 3: U-Net Architecture. Source: [17]

3.3.3 Mask R-CNN

Mask R-CNN [8] is a highly efficient and robust architecture built based on Faster RCNN. It is being used for nuclei detection in histopathological images due to its success in semantic segmentation. It is a two stage network that uses backbone networks like ResNet along with FPN and RPN to extract feature maps, region proposals that are fed to a FCN network for classification, bounding box and Mask detection. Model adopts Multi task RPN class, RPN bbox, MRCNN class, MRCNN bbox and MRCNN mask loss components.It uses L1-smooth for bounding-box regression and Cross entropy loss for classification. We have adopted the Mask RCNN Network developed by Inom Mirzaev for 2018 Data Science Bowl competition with modifications. `https://github.com/mirzaevinom/data_science_bowl_2018`

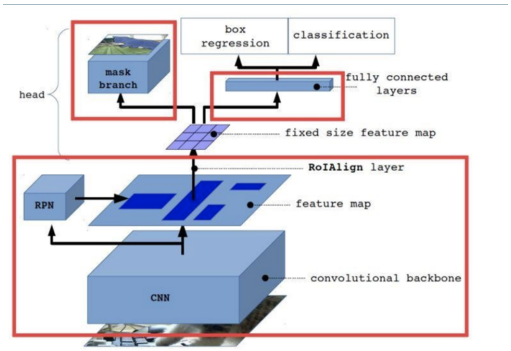


Figure 4: Mask R-CNN Architecture. Source: [8]

3.3.4 Nuclei-Boundary Model

NB model predicts the category of all the pixels of an image with only one pass using Fully convoluted network. This model merges two stages of nuclei detection and then extracting the nucleus area by extracting the nuclei and their edges at the same time. This model is based on UNet architecture [6], uses generated image patches for feeding the network and later uses overlapping patching technique for generating predicted images.

Some of the augmentation techniques used for generating requisite sample data are random elastic transformation, rescale, affine transformation, shift, flip and rotate. These augmentation methods are applied to the corresponding ground truth patches for the sake of consistency But, this model is different from the UNet2 model because this is a 3 class model - background, boundary, inside classes. The size of the input layer is 128X128. The weight of each convolutional layer is initialized by glorot uniform and bias is initialized to 0 as mentioned in [6]. The activation function for all the convolutions is scaled exponential linear units (SELUs) (equation 1).

$$selu(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leq 0 \end{cases} \quad (1)$$

As mentioned in the UNet2 model, the extracted features from the encoding layers are used for generating the segmented maps. The three channels in the output image have dimensions similar to that of the input image .Their values represent the probabilities of each pixel being in background, boundary or inside respectively.

This model tries to minimize the categorical softmax cross entropy loss between predictions and ground truth images. The loss function used is in equation (2).

$$L = \sum_i \sum_j W_{i,j} \log\left(p_{t(i,j)} i, j\right) \tag{2}$$

4. Results

4.1. Unsupervised

This simple four steps method runs well on H&E stain images with few constants fixed for similar kinds of data set, which takes less than 30 seconds per image. Based on observation on several images, it gives mixed results. Few images work as good as the one in figure 4, which can detect nuclei centers and boundaries along with segmentation. When a simple mean IOU is calculated it gives a value between 0.38 to 0.6 on different data sets.

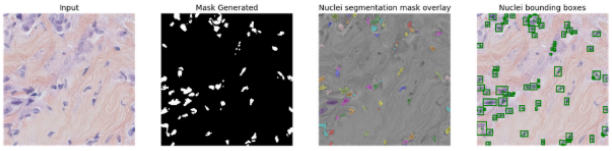
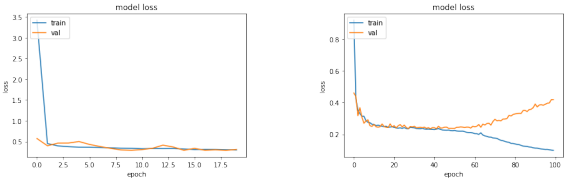


Figure 5: Unsupervised method on one of the sample data: (a) Input Tissue Slide image, (b) Generated binary mask, (c) Mask overlayed on Input image, (d) Nuclei centers detected with boundary box

4.2. U-Net 2 Class

Here we can see that the loss is decreasing for both training and validation for up to 40 epochs. But after which validation loss starts to increase due to over-fitting.



(a) 20 Epochs (b) 100 Epochs

Figure 6: Train and Validation Loss over 20 and 100 epochs for U-Net 2 Class method

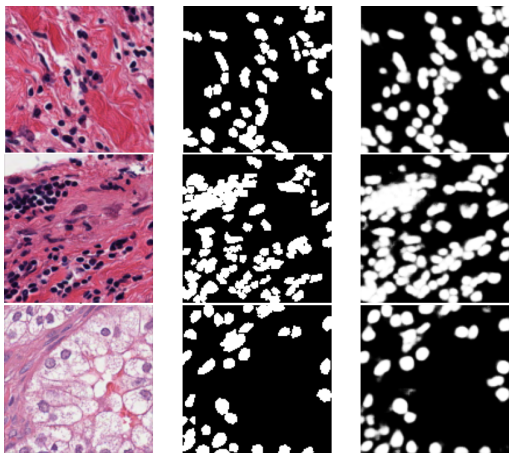


Figure 7: Results on three different inputs for U-Net 2 class method. First column has inputs, second has ground truths and the last column has predicted nuclei segmentation

In figure 6, we can see the input, ground truth and prediction images side by side for 3 examples. The average Mean IOU we got for the test data was 0.633. Although the prediction mask doesn't seem to be precise, the average Mean IOU seems to be decent.

4.3. Mask R-CNN

Mask RCNN with 'Resnet101' as the backbone and top-down pyramid size of 256 is trained on MoNuSeg data with 'Imagenet' weights. With 60 million parameters, Adam optimizer and a batch size of 2 images,it has taken around 54hrs for 100 epochs on IU high-performance computing DL Nodes. Even though the multi-task loss has reduced from 7.52 to 1.1 but the network has failed to perform well on the test images. Seems like the model is suffering from over-fitting on

training data and as the network is computationally intensive, tuning parameters was extremely difficult.

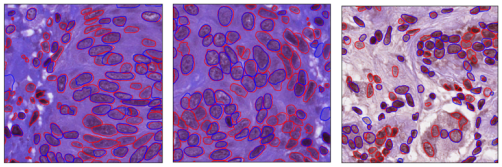


Figure 8: Mask RCNN Results

4.4. Nuclei-Boundary Model

We have trained the nuclei boundary model on MoNuSeg data for 100 epochs, by generating over 6,000 patches from 1000 x 1000 images,feeding the network in batches of size 40. The patches size over 128 x 128. The loss has reduced significantly from 4.23 to 0.293. Whereas, on TNBC data, we ran for 50 epochs and generated over 5000 image patches of 128 x 128 from 512 x 512 , feeding the network with similar batch sizes. For PSB data, we have trained our model on 6,460 patches of similar batch and patch sizes, which were extracted from images of size 400 x 400. Here, the loss has reduced from around 0.44 to 0.315 for over 100 epochs.It was mentioned in [5] that larger patch sizes yielded better results, we did not try for larger sizes due to GPU memory constraints.

With limited parameter tuning, the requisite data preprocessing Nucleus boundary model produces decent results across various data sets. Even though the predicted images have some noise, the model was able to detect nuclei better than other models. For example, when we observe results on the PSB data where the ground truth is not really great, but the model was able to detect most of the nucleus. This is some sort of blessing in disguise, on one hand, we were able to detect the nuclei areas, on the other hand this prevented us from defining a suitable evaluation strategy.

5. Discussion

Because of the change from Tensorflow 1.x to Tensorflow 2.x, we had to change a lot of the codes to be able to run them. This meant replacing many deprecated functions with new ones.

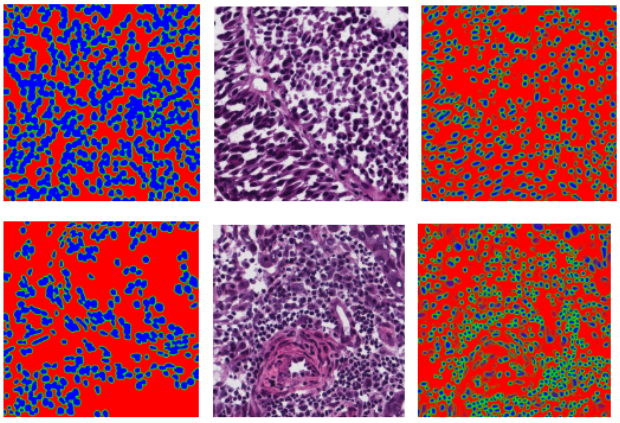


Figure 9: Results of Nuclei-Boundary model on MoNuSeg data. First column is ground truth, next is original image and the last column has predicted nuclei segmentations

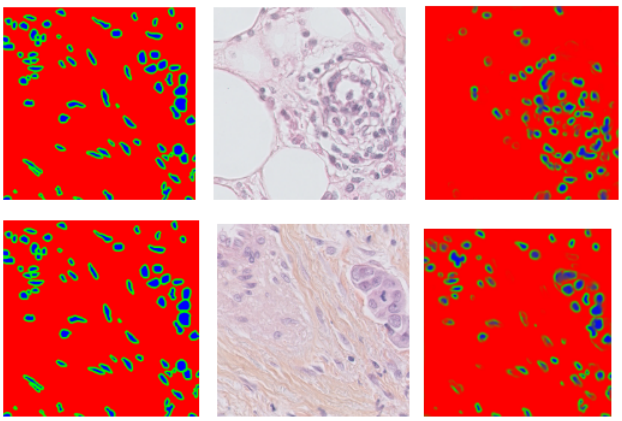


Figure 10: Results of Nuclei-Boundary model on TNBC data. First column is ground truth, next is original image and the last column has predicted nuclei segmentations

5.1. Unsupervised

This method is a quick and good starter for nuclei segmentation without any training. Few limitations of this method are: it is specific to H&E stained images, few constants such as minimum nuclei area, foreground intensity threshold, etc. have to be provided to modulate it for multi-organ histology images.

In this technique we had a lot of discussions with

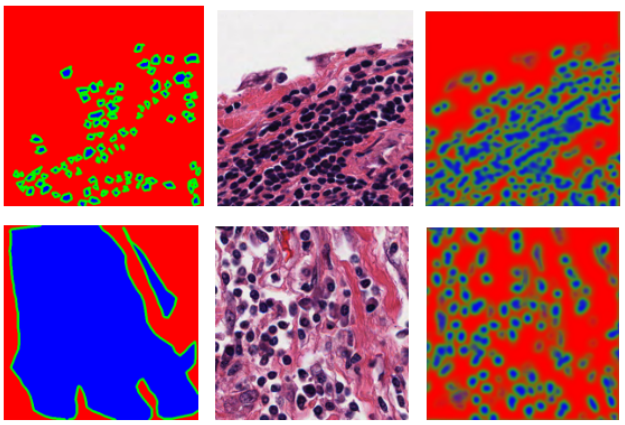


Figure 11: Results of Nuclei-Boundary model on PSB data. First column is ground truth, next is original image and the last column has predicted nuclei segmentations

the HistomicsTK development team on GitHub to fix their library which was very crucial for this unsupervised method. One new addition to this method is testing this technique on three of our data sets.

5.2. UNet2

The UNet2 model's code [https://github.com/ChristopherBui3/Metis_Projects/tree/master/Project_5] had many bugs that we had to fix for getting accurate results, like order of the data being matched between inputs and outputs, model evaluation etc. The model evaluation metrics and early stopping criteria were not working accurately. We had made changes such that these issues were fixed. The results look good for a 2 class problem. We can see that the boundaries of the nuclei are blurry. This is the drawback of using a 2 class method. This problem can be rectified by using a 3 class problem where even the boundary is predicted accurately as a separate class.

5.3. Mask R-CNN

As mentioned earlier,this model is suffering from over fitting problem and has huge computation costs due to 60 million parameters. This could be solved by training using some of the augmentation techniques employed in NB model.

5.4. Nuclei Boundary Model

As mentioned earlier, for the Nuclei Boundary Model we were unable to define a generalized evaluation strategy, since we did not have individual nucleus masks. We have used the mean IOU to evaluate the results for this model but, it resulted in 0.23. We were rather surprised because we felt the nuclear boundary model performed better than the other on a wide range of data sets. So, one of the major limitations of this model is defining meaningful evaluation metrics. As mentioned earlier, discrepancies in the crowd-sourced data (PSB) is probably one of the reasons, MoNuSeg results were much better for 100 epochs than the PSB data for 100 epochs and PSB's data is of dimensions 400 x 400 which might have resulted in similar patches and might have over-fitted the model whereas, in MoNuSeg data, with images size of 1000 x 1000 might have resulted in a diverse set of patches.

6. Conclusion

One of the major contributions in the project is studying various nuclei segmentation models, using image processing techniques to enable models to work on various data sets of histopathological images, various kinds of ground truth annotations. The NB model was originally tested on the Breast Cancer data set, but we were successful in replicating the results for tissue images of several patients with tumors of multiple organs, crowd sourced histopathological images, triple-negative breast cancer tissue slide images, etc.Based on our findings, we can conclude that we can use crowd sourced data for medical image analysis and given histopathological H&E stain images,and their corresponding segmented masks, we can perform medical image analysis such as nuclie segmentation.

Additionally, we have also explored unsupervised methods with the primary motivation being to develop a model for nuclei segmentation on any kind of H&E stained histopathological images.

It would be worth mentioning the multiple challenges we had to overcome during the course of our project which demanded considerable effort, most of

which was to make the models work. Some of them include painful GPU configurations, code fixes for enabling support for tensorflow-gpu, incorporation of data sets for augmentation techniques, working with HistometicsTK open source community to help resolve issues.

i.)[https://github.com/DigitalSlideArchive/HistomicsTK/issues/847#issuecomment-61752762]

ii.)[https://github.com/DigitalSlideArchive/HistomicsTK/issues/616#issuecomment-61923824]

We do have some unfinished business, particularly on defining a meaningful evaluation strategy. Considerable effort went into devising evaluation metrics, ways to evaluate models when individual nucleus masks are unavailable, but we weren't successful so far, this is something which we would like to focus our further research on.

7. References

1. Abdolhoseini, Mahmoud, et al. "Segmentation of Heavily Clustered Nuclei from Histopathological Images." Nature News, Nature Publishing Group, 14 Mar. 2019, www.nature.com/articles/s41598-019-38813-2.

2. Coelho, Luís Pedro, et al. "Nuclear Segmnentaion in Microscopic Cell Images: A Hand-Segmented Dataset and Comparision of Algorithms." Proceedings. IEEE International Symposium on Biomedical Imaging, U.S. National Library of Medicine, 2009, www.ncbi.nlm.nih.gov/pmc/articles/PMC2901896/.

3. Deep Adversarial Training for Multi-Organ Nuclei Segmentation in Histopathology Images,.https://www.researchgate.net/publication/327971066_Deep_Adversarial_Training_for_Multi-Organ_Nuclei_Segmentation_in_Histopathology_Images

4. Tsou, Chi-Hsuan, et al. "A Heuristic Framework for Image Filtering and Segmentation: Application to Blood Vessel Immunohistochemistry." Analytical Cellular Pathology (Amsterdam), Hindawi Publishing Corporation, 2015, www.ncbi.nlm.nih.gov/pmc/articles/PMC4707018/.

5. Md Zahangir Alom Chris Yakopcic "Nuclei Segmentation with Recurrent Residual Convolutional Neural Networks based U-Net (R2U-Net)". https://ieeexplore.ieee.org/document/8556686

6. Cui, Yuxin, et al. "A Deep Learning Algorithm for One-Step Contour Aware Nuclei Segmentation of Histopathology Images." Medical & Biological Engineering & Computing, vol. 57, no. 9, 2019, pp. 2027–2043., doi:10.1007/s11517-019-02008-8. https://arxiv.org/pdf/1803.02786.pdf

7. X. Xu et al., "Quantization of Fully Convolutional Networks for Accurate Biomedical Image Segmentation," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018, pp. 8300-8308. https://arxiv.org/pdf/1803.04907.pdf

8. K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 2980-2988. https://arxiv.org/abs/1703.06870

9. "About." HuBMAP Consortium, hubmapconsortium.org/about/

10. Irshad, H, et al. "Crowdsourcing Image Annotation for Nucleus Detection and Segmentation in Computational Pathology: Evaluating Experts, Automated Methods, and the Crowd." Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, U.S. National Library of Medicine, 2015, www.ncbi.nlm.nih.gov/pubmed/25592590.

11. Nuclei Segmentation in Histopathology Images Using Deep Neural Networks - IEEE Conference Publication, ieeexplore.ieee.org/document/7950669.

12. N. Kumar et al., "A Multi-Organ Nucleus Segmentation Challenge," in IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1380-1391, May 2020, doi: 10.1109/TMI.2019.2947628

13. B. Hu, Y. Tang, E. I. Chang, Y. Fan, M. Lai and Y. Xu, "Unsupervised Learning for Cell-Level Visual Representation in Histopathology Images With Generative Adversarial Networks," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 3, pp. 1316-1328, May 2019, doi: 10.1109/JBHI.2018.2852639

14. E. Reinhard, M. Adhikhmin, B. Gooch and P. Shirley, "Color transfer between images," in IEEE Computer Graphics and Applications, vol. 21, no. 5, pp. 34-41, July-Aug. 2001, doi: 10.1109/38.946629.

15. Macenko, Marc, et al. “A Method for Normalizing Histology Slides for Quantitative Analysis.” A Method for Normalizing Histology Slides for Quantitative Analysis — Proceedings of the Sixth IEEE International Conference on Symposium on Biomedical Imaging: From Nano to Macro, 1 June 2009, dl.acm.org/doi/10.5555/1699872.1700155.

16. Wienert, Stephan, et al. “Detection and Segmentation of Cell Nuclei in Virtual Microscopy Images: A Minimum-Model Approach.” Nature News, Nature Publishing Group, 11 July 2012, www.nature.com/articles/srep00503.

17. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convo- lutional networks for biomedical image segmentation,” in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, 2015, pp. 234–241.