Nuclei Segmentation Using Deep Learning and Hyperspectral Imaging for High Accuracy

Gusain, Hridaanshu
Shaker High School
Latham, New York
hridaanshu@gmail.com

Mestha, Lalit K. *Life Fellow IEEE*
Adjunct Faculty, Rensselaer Polytechnic Institute,
110 8th St, Troy, NY 12180
lkmestha@gmail.com

***Abstract*— Computerized microscopy image analysis plays a crucial role in computer aided nuclei segmentation/diagnosis. Machine learning methods have emerged as a powerful tool in medical investigation and clinical practice. Recently studies have been conducted on training neural networks utilizing large datasets of Red, Green, Blue (RGB) images to produce segmentations of nuclei on human tissue. This has attracted considerable attention in the biomedical informatics industry as these large datasets are difficult to acquire and validate. This study provides an analysis of the differences in nuclei segmentation of human tissue using RGB images and Hyperspectral images (HSI) using the UNET Convolutional Neural network architecture. We introduce the popular neural network specifically designed for biomedical segmentation problems and we summarize current deep learning achievements in tasks such as detection and segmentation. The results of this research qualify the potential usage of hyperspectral imaging to improve the quality of neural network segmentations given a smaller dataset of HSI volumes. We explain the network architecture, the principles of the convolutional neural network and describe how we were able to improve its efficiency and accuracy. In addition, we discuss the challenges and the potential trends of future research in biomedical image analysis using machine learning methods.**

I. Introduction

The cells presented in Figure 1 are very similar in visual appearance and structure. However, there are microscopic differences that occur in benign cells and malignant cells, which can be exploited to differentiate and diagnose the cell types and their properties to be cancerous. Considerable academic study is required for pathologists to become proficient in identification of cancerous cells. However, due to human error, false diagnoses can be made [1]. With new machine learning methods of segmenting nuclei in images of human tissue, diagnosis time can be expedited with increased accuracy and efficiency [2][3].

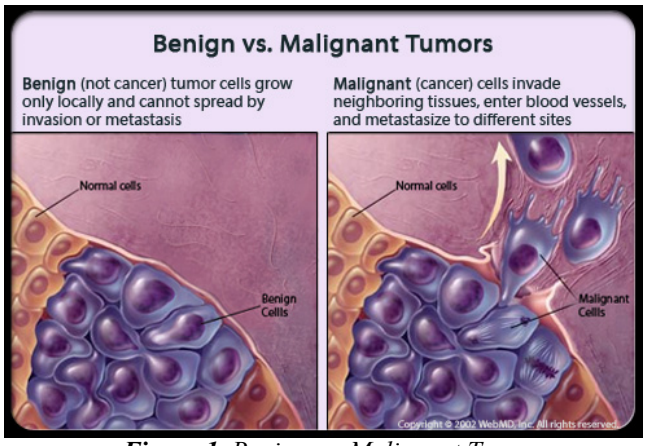


Figure 1. *Benign vs. Malignant Tumors*
Source: Adapted from [4]

The findings presented in this study will improve nuclei segmentation of pathological images for cancer detection with precision and accuracy under a few seconds utilizing machine learning methods. Current methods of segmentation utilize RGB images which consist of 3 color channels: Red, Green and Blue (broad wavelengths centered at 630 nanometers, 535 nanometers, and 435 nanometers) [5][6]. This research proposes an alternative method of training neural networks using Hyperspectral imaging (HSI). Hyperspectral image volumes from a microscope contain multiple color wavelengths with narrow bands in the electromagnetic spectrum from 400-1000 nanometers. Such images are utilized in aerial photography, astronomy, and agriculture because they allow specific data about the chemical makeup of a subject to be identified [7][8]. This study proposes the application of this imaging system to the semantic segmentation process of nuclei for cancer detection. HSI images allow specific chemicals and cell structure to be identified in the cell, creating a larger range of data for the network to train on [9]. The images are captured from a hyperspectral-capable microscope and cannot be viewed directly on a computer screen (unless they are modified to Red, Green and Blue bands) as common displays are not capable of displaying images with more than three channels *Figure 2.*

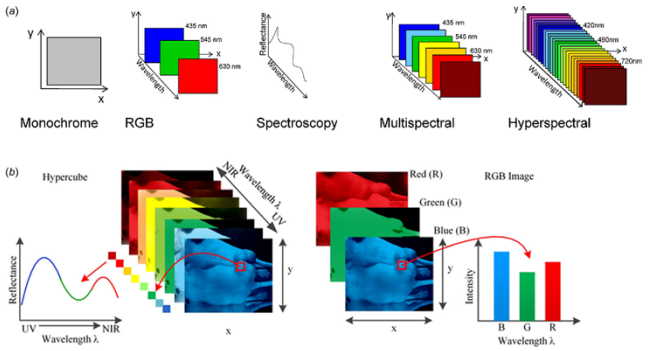
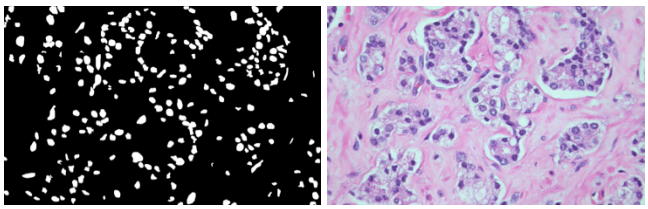


Figure 2. *Hyperspectral image example*
Source: Adapted from [10]

II. Neural Network Architecture

The Universal Network Architecture was utilized in this study because it was developed specially for biomedical image segmentation problems [11]. The first half comprises of a feature contraction path to apprehend details and the second half, an expansive path to obtain precise location details. The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two 3x3 2D-convolutions (unpadded convolutions), each followed by a dropout to prevent overfitting. A rectified linear unit (ReLU) follows with a 2x2 max pooling operation with stride 2 for down-sampling. At each down-sampling step we double the number of feature channels. The expansion process is executed with transposed convolution kernels shrinking feature map numbers and enlarging the dimensions of the image. The maps of features from the contraction process are concatenated with the expansion path layers to reduce pattern detail loss. A 1x1 convolution is then applied to produce a segmented image. The output from the neural network is a probability mask of the locations of nuclei within the image.

III. Image Preparation



a: Manually annotated ground truth mask(grayscale)

b: Corresponding cell field

Figure 3. *Annotated ground(left) truth with corresponding cell field(right)*

Controlled testing was required to compare maximum accuracy for the RGB and HSI dataset. To obtain ground truth masks, RGB fields (Figure 3b) were manually annotated and validated using the Label-Studio software (Figure 3a). Each RGB tissue field had a replica in the Hyperspectral data format, so the same masks were used for both imaging types. Each tissue field was 966x606 pixels from the prostate but were cropped to smaller patches to improve training efficiency.

The spectral python package enabled specific bands to be extracted from the entire data cube. Bands that contained pixels majorly representative of connective tissue were not selected to train on because chemical structure of nuclei were invisible. Without nuclear structure data, network quality will decrease.

Various dimensions of the UNET architecture were tested with the following patch sizes (length in pixels × width in pixels): 128×128, 256×256, and 512×512. 64×64 patch sizes rendered the field of view too small for the network to detect individual instances of nuclei and 1024×1024 yielded the issue of not enough training data as the fields were only 966×606.

The third dimension of an image represents the number of bands (color channels) in the image. The HSI model was trained using the HSI images with 1 to 12 color channels to test improvement compared to the RGB model with 3 color channel images.

Each image was input as multi-dimensional Tensor of 64-bit floating point numbers. Figure 5 contains examples of hyperspectral images, (Figure 4a) RGB images (Figure 4b), and their corresponding ground truth mask (Figure 4c).

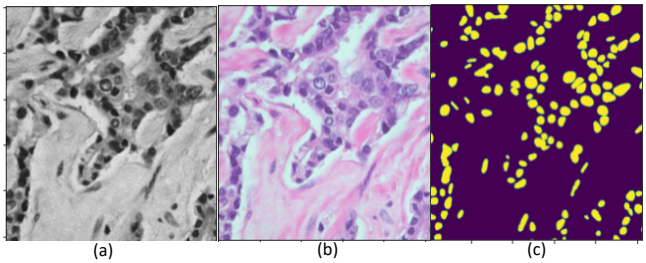


Figure 4. *Single band HSI image, RGB image, and ground truth examples [HSI image example(left). Corresponding RGB image(middle). Corresponding ground truth (right)]*

IV. Network Training

The network was initially trained with 128×128, 256×256 and 512×512 patches of the tissue fields to test overfitting on different datasets. After training on RGB and 3-band HSI images for 25, 30, and 50 epochs, it was determined that 512×512 patch sizes were the most efficient, had less overfitting, and provided the highest average IoU (**Eq. (1)**). 128×128 and 256×256 patch sizes cropped out individual nuclei within the patch and caused the network to have a compromised training dataset.

$$IoU = \frac{TP}{TP + FP + FN} \quad (1)$$

Where TP=True Positive; FP=False Positive; FN=False Negative;

Different dataset sizes were tested from 5 fields to 85 fields to test the effect of data size on model accuracy. In practice, large datasets are difficult to acquire and require vast amounts

of time to annotate and validate. Therefore, if HSI provides comparable accuracy to the RGB model trained on a smaller dataset, it would prove to be a more efficient use of training resources.

To prevent testing the network on known images, the batches were split by 80% training and 20% testing. The model's optimizer was Adam, as it is an extended form of stochastic gradient descent and is commonly utilized in computer vision models [12]. The loss function this research utilized was binary cross entropy [13].

The network was allowed to train for a maximum of 200 epochs but had a check-pointer to monitor validation accuracy and saved the model with the highest validation accuracy to prevent overfitting. After training, the mean IoU and F-score (**Eq. (2)**) was calculated for all testing images in the dataset.

$$F - score = \frac{2 * TP}{2 * TP + FP + FN} \quad (2)$$

V. Results

With a dataset of just 5 fields, the HSI model achieved 75% IoU and 85% F-score with 10 color channels. With a dataset of 5 fields, the RGB dataset achieved 66% IoU and 80% F-score. Both training methods caused saturation in training loss and validation loss as epochs increase (Figure 5). As the dataset size increased, both models achieved an accuracy of 81% when trained on 45 fields.

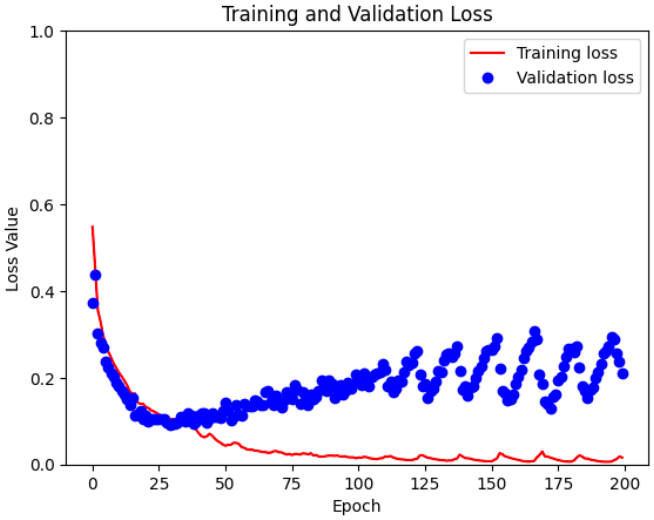


Figure 5. *Training and Validation loss curves while training for 200 epochs*

The maximum IoU achieved from the HSI dataset was 81% and an F-score of 91% when trained on 36 fields with 10 color channels. The RGB segmentation model produced a maximum IoU of 81% and a maximum F-score of 90% when given a dataset of 45 fields. IoU and F-score saturated as the number of training fields increased.

Predictions from both RGB and HSI models with the highest average IoU are displayed in Figures 6a and 6b. They are overlayed on RGB images for ease of viewing. The predictions were processed through a function to extract the contours of the nuclei to represent the morphology of the nuclei distribution using white lines. The RGB model predicted nuclei as interconnected, and the centroids are not individually depicted (Figure 6a). In the HSI segmentation, these faults are less prevalent and there are more distinct nuclei predicted in the patch of tissue (Figure 6b).

Figure 7a is the RGB version of the input image and Figure 7b is the ground truth mask. Figure 7c and 7d are corresponding predictions from the HSI model, and RGB model when trained on 5 fields. The predictions from the Hyperspectral model (Figure 7c) are made with a higher confidence interval, are less connected, and have less noise compared to the predictions made by the RGB model (Figure 7d).

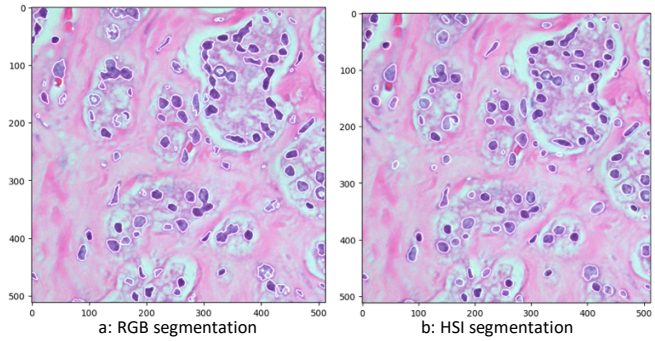


Figure 6*. RGB vs. HSI prediction contour overlayed on RGB test image.*

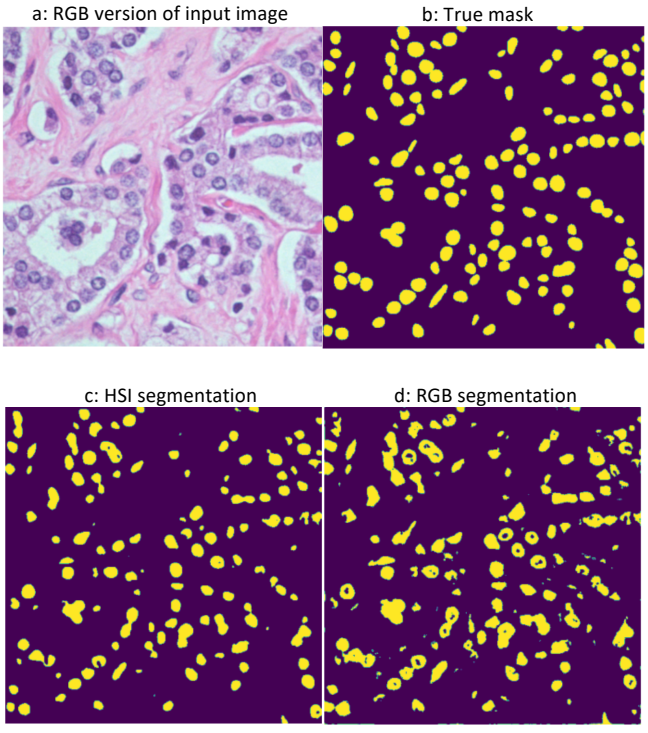


Figure 7. *Examples of RGB and HSI predictions compared to corresponding ground truth*

Different HSI datasets were trained on with 2 to 12 color channels. Each patch was 512×512 to maximize IoU as described in Section IV. When training, the same automatic maximum callback method was used to prevent overfitting. The training epochs ranged from 15-105 epochs according to the number of color channels per image (more color channels required more epochs). The IoU threshold peaked at 80.9% as depicted in Figure 10 and F-score peaked at 91% with 10 color bands with 105 epochs of training. As number of color channels increased after 10, IoU decreased, indicating training saturation.

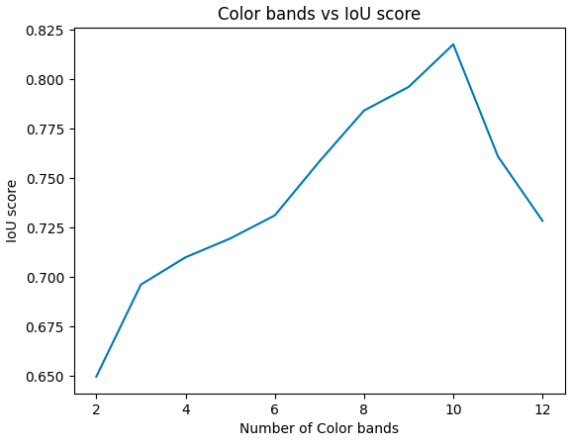


Figure 8. *Graph of HSI color bands vs. IoU score*

VI. Conclusion and Future Work

All the testing in this study proves that Hyperspectral images improve the accuracy of a neural network model for nuclei segmentation when the given dataset limited in size. As the number of color bands increases, there is an increase in segmentation accuracy until a point of overfitting. The model trained on hyperspectral images provides a segmentation of higher quality compared to the model trained on RGB images. It distinguishes individual nuclei separately and with a higher confidence interval, whereas the RGB model depicts separate nuclei as connected and with a low confidence interval. The Hyperspectral model also provides a segmentation that is less noisy and is visually cleaner for a pathologist to use in practice.

Methods such as unsupervised learning, or active learning can be employed in the future to train on larger data size without ground truth masks. Unsupervised learning will enable the neural network to learn patterns in an image without the need for ground truth masks, which will allow it to recognize specific cell structure based on different chemical signals from the hyperspectral images. Active learning, when implemented in the future with hyperspectral imaging, can allow a pathologist to work alongside a neural network in a collaborative manner to train the network on specific classification problems. The network will prompt a pathologist to verify its segmentation, and when prompted with a validation accuracy, it will continue its back-propagation methods. A Multi-Dimensional UNET can also be used, as it allows for more efficient training on a large volume image (60 or higher color channels).

This research presents an optimistic path for the methods used to segment and analyze images in the histopathological field of biology with the use of hyperspectral imaging to increase the number of color channels in a training dataset. The ability to discern different biological structures in a cell based on multiple color channels in the electromagnetic spectrum proved to increase the quality of neural network output.

References

[1] M. J. Fromer, "Study," *Oncology Times*, vol. 27, no. 22, pp. 25–26, Nov. 2005, doi: https://doi.org/10.1097/01.cot.0000291164.08133.9a.

[2] R. Pandey, R. Lalchhanhima, and Ksh. Robert Singh, "Nuclei Cell Semantic Segmentation Using Deep Learning UNet," *Nuclei Cell Semantic Segmentation Using Deep Learning Unet*, Dec. 2020, doi: https://doi.org/10.1109/acts49415.2020.9350516.

[3] K. Bera, K. A. Schalper, D. L. Rimm, V. Velcheti, and A. Madabhushi, "Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology," *Nature Reviews Clinical Oncology*, vol. 16, no. 11, pp. 703–715, Aug. 2019, doi: https://doi.org/10.1038/s41571-019-0252-y.

[4] "Microbiology: Tumor Cell Lines," *Blogspot.com*, 2023. https://3.bp.blogspot.com/-cYljfmf-T-M/WI6-_mXc2JI/AAAAAAAAAhs/CtilAWSLpnE2amuctlCo-0iVf4UEerL5wCEw/s640/Picture1.png (accessed Nov. 24, 2023).

[5] P. Naylor, M. Laé, F. Reyal, and T. Walter, "Nuclei segmentation in histopathology images using deep neural networks," *IEEE Xplore*, Apr. 01, 2017. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7950669 (accessed Jan. 27, 2022).

[6] F. Xing and L. Yang, "Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review," *IEEE Reviews in Biomedical Engineering*, vol. 9, pp. 234–263, 2016, doi: https://doi.org/10.1109/rbme.2016.2515127.

[7] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 29–43, 2002, doi: https://doi.org/10.1109/79.974724.

[8] C. Guilloteau, T. Oberlin, O. Berné, É. Habart, and N. Dobigeon, "Simulated JWST Data Sets for Multispectral and Hyperspectral Image Fusion," *The Astronomical Journal*, vol. 160, p. 28, Jul. 2020, doi: https://doi.org/10.3847/1538-3881/ab9301.

[9] P. N. Hedde, R. Cinco, L. Malacrida, A. Kamaid, and E. Gratton, "Phasor-based hyperspectral snapshot microscopy allows fast imaging of live, three-dimensional tissues for biomedical applications," *Communications Biology*, vol. 4, no. 1, Jun. 2021, doi: https://doi.org/10.1038/s42003-021-02266-z.

[10] N. Mehta, S. Shaik, R. Devireddy, and M. R. Gartia, "Single-Cell Analysis Using Hyperspectral Imaging Modalities," *Journal of Biomechanical Engineering*, vol. 140, no. 2, Jan. 2018, doi: https://doi.org/10.1115/1.4038638.

[11] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *Lecture Notes in Computer Science*, vol. 9351, pp. 234–241, 2015.

[12] D. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *Computer Science*, 2014, doi: https://doi.org/10.48550/arXiv.1412.6980.

[13] U. R. Dr. A, "Binary cross entropy with deep learning technique for Image classification," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 5393–5397, Aug. 2020, doi: https://doi.org/10.30534/ijatcse/2020/175942020.