Nuclei Detection With Histopathological Images of Skin Tissue Through Machine Learning

Afya Shaikh - EDIT ML 2024

Introduction

The fight against cancer relies heavily on the ability to detect cancerous nuclei swiftly and accurately which is important for successful treatment and saving lives. Cancerous skin tissue, in particular, are known for their aggressive and rapid nature, making timely and precise diagnosis vital. Cancer appears abnormal because their cell nuclei is often larger and darker due to an excess of DNA. Detecting these changes early is difficult, as cancer cells can develop deep within tissues, making them harder to find with traditional methods.

Traditional methods of examining tissue samples under a microscope, which have been the standard for detecting cancer, often have limitations. These techniques can be slow and prone to human error, which can negatively affect patient care.

This research looks at using machine learning technology to make cancer detection better. By studying images of tissue samples, the goal is to speed up and improve how well we can identify cancerous cells. The project aims to create a machine learning model that can automatically spot cancerous cells in these images, offering a helpful tool in the fight against one of the toughest diseases in medicine.

References/Acknowledgments

Thank you to Dr. Joshua Levy, Onyinyechi Owo, and Dr. Minh-Khang Le for all the help and support throughout my project!

Results

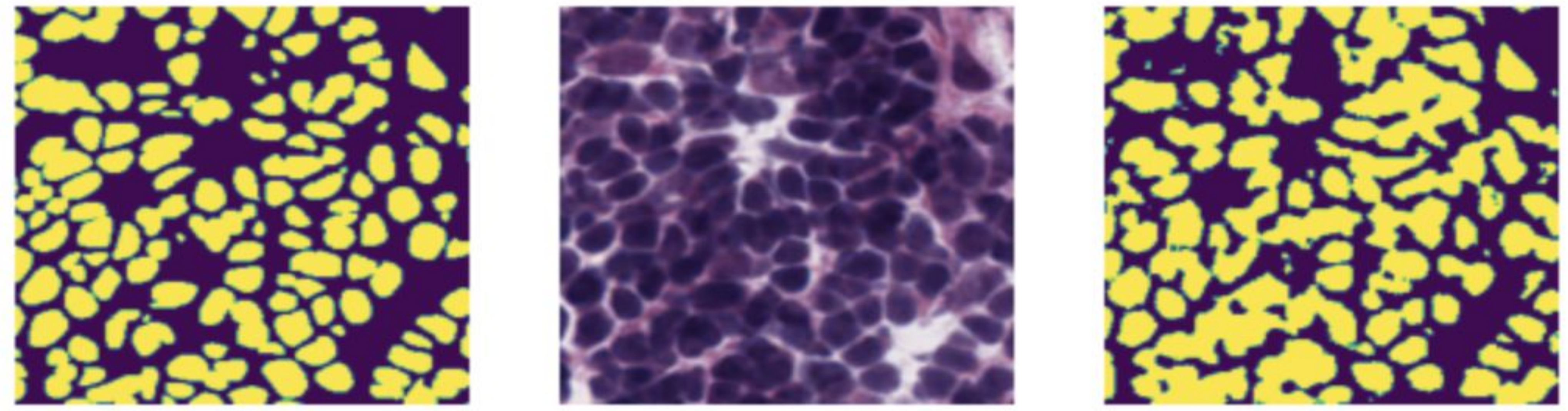

Figure 1. Human Mediastinum mask binary image, tissue image, and machine learning image

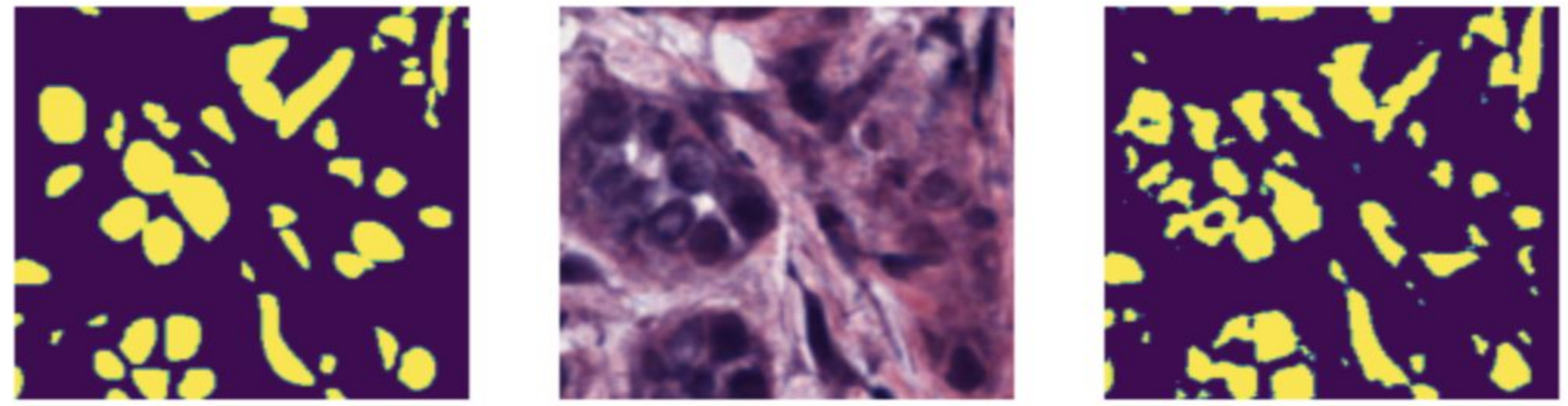

Figure 2. Human Thyroid Gland mask binary image, tissue image, and machine learning image

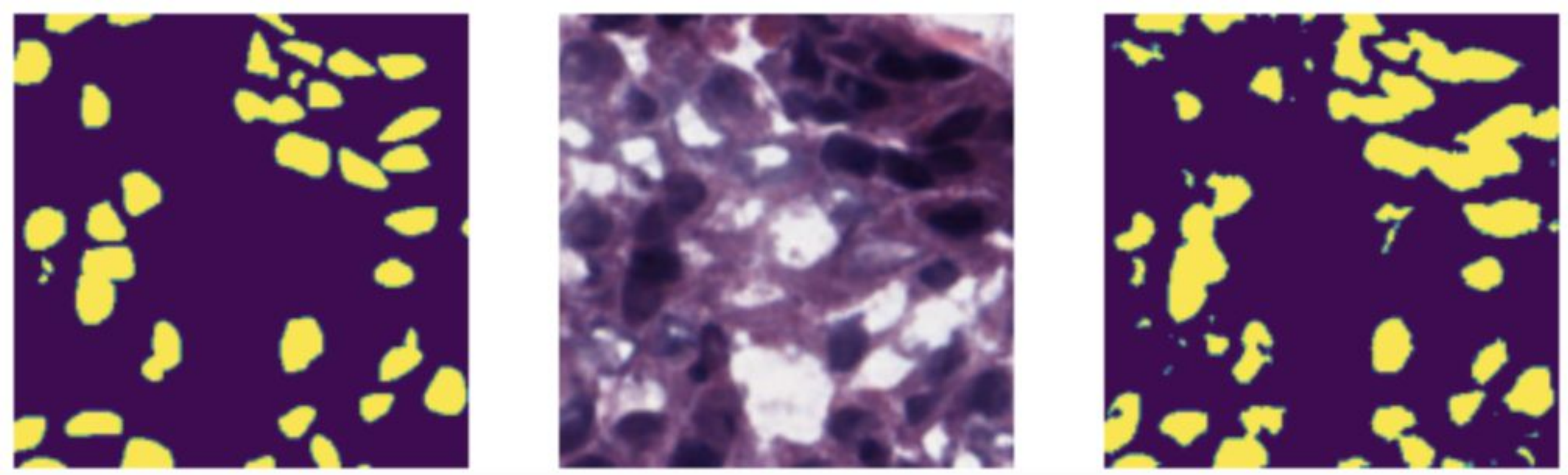

Figure 3. Human Larynx mask binary image, tissue image, and machine learning image

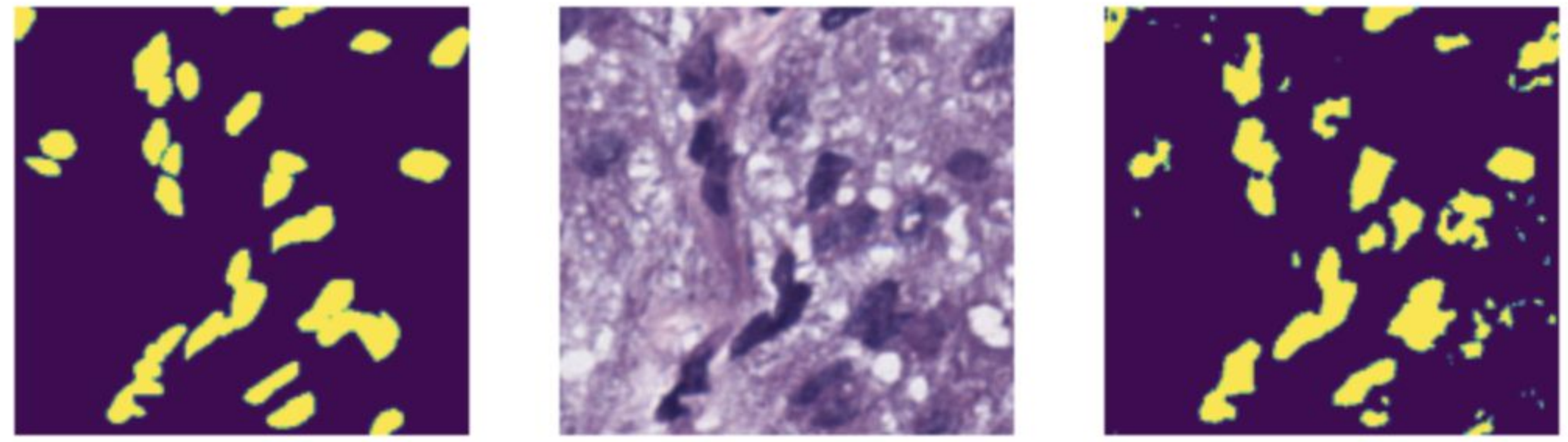

Figure 4. Human Adrenal Gland mask binary image, tissue image, and machine learning image

	precision	recall	f1-score	support
0	0.95	0.88	0.92	1145799
1	0.74	0.88	0.81	427065
accuracy			0.88	1572864
macro avg	0.85	0.88	0.86	1572864
weighted avg	0.90	0.88	0.89	1572864

Figure 5. The accuracy of the machine learning models cancer detection.

Methods

Materials

I used a dataset of 30 H&E-stained frozen sections from The Cancer Genome Atlas (TCGA), covering ten human organs. The dataset includes over 8,000 annotated nuclei and various segmentation masks.

Preparation

Images were standardized to a uniform size and pixel scale. Each image was divided into quarters to facilitate more effective training of the machine learning model. This approach ensured diverse tissue representation for unbiased learning.

Protocol

Training was conducted using a machine learning model implemented in Jupyter Notebook. The dataset was split into training and validation sets, with cross-validation used to assess model performance. Different organ tissues were included to avoid bias.

Measurements and Calculations

Model performance was evaluated using precision, recall, F1 score, and accuracy. These metrics assessed the model's ability to correctly identify cancerous nuclei.

Conclusion

The machine learning model significantly improved cancerous nuclei detection over traditional methods. Using advanced segmentation techniques, it enhanced both accuracy and speed in identifying cancerous cells in histopathological images. This boost in efficiency and precision can greatly improve skin cancer detection and treatment planning, ultimately benefiting patient care and outcomes.

Future work will focus on integrating feedback mechanisms for ongoing model improvement, expanding dataset diversity to detect hidden cancers and identify their original sites, and refining algorithms. These steps aim to further enhance detection, treatment planning, and overall patient care and outcomes.