# NUCLEI DETECTION AND SEGMENTATION IN COLORECTAL TISSUE SAMPLES

## Background

Whole slide images (WSI) annotation is currently time-consuming, inconsistent, and prone to errors due to its manual nature. The large image size and tissue variability often result in incomplete or inaccurate labels making automated models less reliable. Automated methods also struggle to detect small objects like nuclei while over-segmenting larger regions

## Abstract

Our research presents a program to automate cell nuclei annotation in colon tissue whole slide images (WSI), a task currently done manually by medical professionals. The model addresses challenges like noise and overlapping structures, offering accurate detection and strong generalization. It enhances diagnostic efficiency and accuracy, and reduces manual workload in digital pathology.

## Methodology

### Method 1: K-Means Clustering

1. Load and convert image to grayscale and BGR
2. Remove darkest and lightest values to reduce impact of high contrast values
3. Extract darkest clusters
4. Apply Otsu's method to identify foreground and background
5. Apply K-Means Clustering to identify nuclei
6. FIlter unwanted large objects
7. Return image mask

### Method 2: Adaptive Thresholding

1. Load and convert image to grayscale and blur
2. Apply adaptive thresholding to identify small clusters of the same intensity
3. Tune block size to get predictions that are not too small or over-sized
4. Filter unwanted light clusters that are assumed to not be nuclei
5. Return image mask

## Analysis

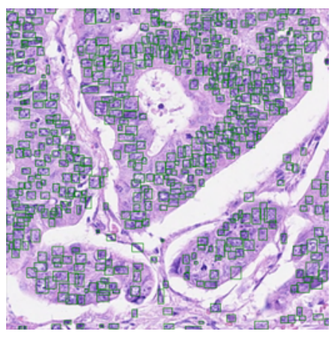
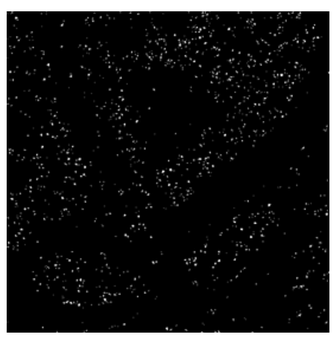
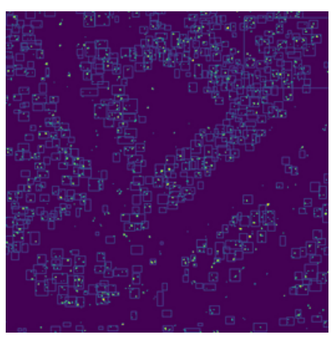


Image mask



Segmented image



Overlap Accuracy

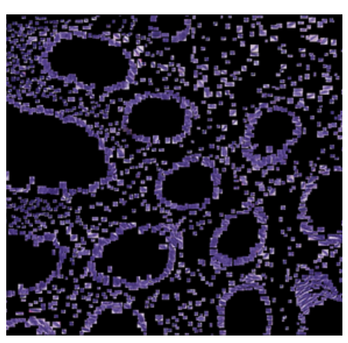
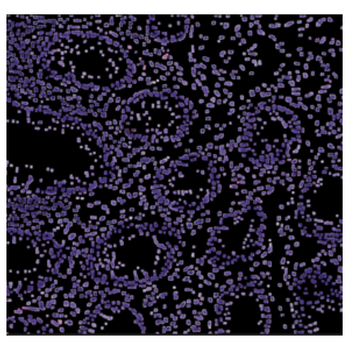
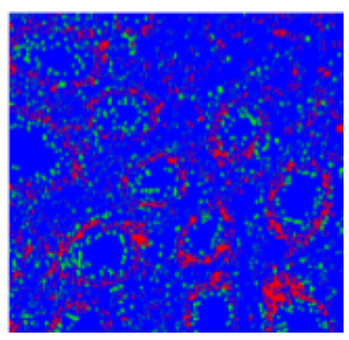


Image mask



Segmented image



Overlap Accuracy

## Results

Our programs achieved a 37% accuracy for the K-Means model and a 78% accuracy in the Adaptive Thresholding model within the processed images. This indicates that not all but some if not most of the objects were correctly identified and categorized. The Adaptive Thresholding program's performance can be considered for general use, but there is room for improvement, particularly in refining the detection of more challenging or less distinct objects. The accuracy for the K-Means model is likely low due to the difference in the nature of the image mask and the predictions .