Automated Tumor Proportion Score Analysis

Ajay Singh Panwar
Department of Material science and Engineering

Dr. Tushar Sandhan
Department of Electrical Engineering

Indian Institute of Technology Kanpur

INDRODUCTION

This project aims to develop a machine learning method that calculates the tumor proportion score (TPS) to predict the benefit of immunotherapy for patients. The Tumor Proportion Score (TPS) is a measure used to evaluate the expression of the PD-L1 protein on the surface of tumor cells. TPS is determined as the percentage of PD-L1 positive stained tumor cells (TCs) present in total number of TCs.

$$TPS = \left(\frac{Total\ number\ of\ PD-L1\ positive\ TCs}{Total\ number\ of\ viable\ TCs} \right) \times 100$$

RESEARCH OBJECTIVE

1. Implement a machine learning algorithm for the segmentation of tumor regions utilizing thresholding-based segmentation, specifically employing Otsu's thresholding method.
2. The primary objective of this project is to segment positively stained tumor regions and predict the mask. This is achieved by applying thresholding, followed by segmentation using the watershed algorithm and contour detection to facilitate mask prediction.
3. we implement a machine learning technique for counting positive and total tumor cells. For this purpose, we use contour detection and draw circles around each detected object for annotation.
4. Finally, the objective is to calculate the Tumor Proportional Score (TPS) based on the segmented and annotated information.

METHODOLOGY AND RESULTS

- **Data-Set :-** We obtained microscopic images of tumors stained with PDL1 and H&E directly from doctors. Additionally, we conducted image preprocessing on these images. This preprocessing involved careful cropping of the images to suit our specific purposes and manual marking of positive strained regions within the tumor region.

- **Tumor Segmentation:-** For the segmentation of tumor regions in a grayscale image, we employ a sequence of image processing techniques. Initially, the image undergoes enhancement using Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance contrast and improve visualization of subtle details. Following enhancement, a Gaussian blur is applied to achieve smoother transitions and reduce noise within the image. Subsequently, Otsu's thresholding method is utilized to segment the tumor region by automatically determining an optimal threshold based on the image's histogram. This step effectively separates tumor areas from the background, facilitating clearer delineation. To refine the segmentation, morphological operations are then employed to eliminate small objects and noise from the binary image. These combined techniques ensure robust and accurate identification of tumor regions, laying a solid foundation for further quantitative analysis and clinical interpretation of medical images.

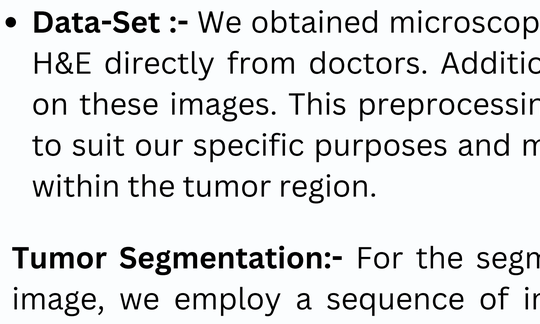

figure 1: tumor image

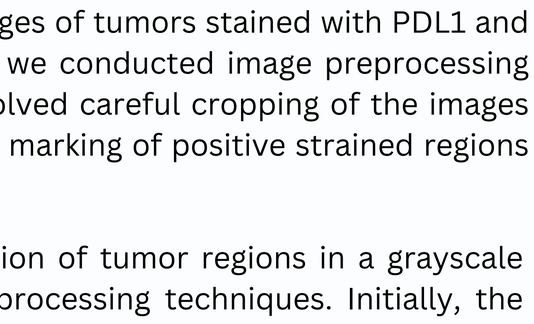

figure 2: segmented image

- **Mask Predection :-** we implement a multi-step approach to segment positive strained tumor regions from a given image. Initially, the image is loaded and converted to grayscale. Using Otsu's thresholding method, we create a binary image where tumor regions are highlighted against a dark background. Next, the distance transform is computed to determine the distance of each pixel from the nearest zero-valued pixel in the binary image. This distance map is used to identify local maxima, which serve as markers for the watershed - segmentation algorithm. Watershed segmentation is applied to delineate distinct tumor regions based on these markers, refining the segmentation further. Contours are then detected within each segmented region, and bounding circles are drawn around the detected contours to highlight and annotate individual tumor areas. Finally, Contrast Limited Adaptive Histogram Equalization (CLAHE) and Gaussian blur to enhance contrast and smooth transitions, followed by additional thresholding and noise reduction to produce a refined segmentation result.. This comprehensive approach combines thresholding, watershed segmentation, contour detection, and image enhancement techniques to accurately identify and characterize tumor regions in medical images.

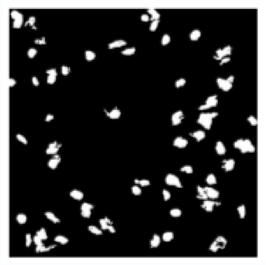

figure 3: Predicted Mask

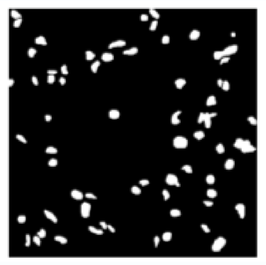

figure 4: Actual Mask

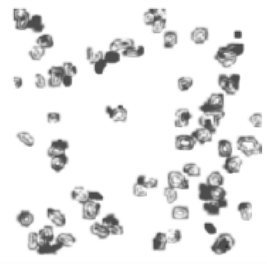

figure 5: Difference

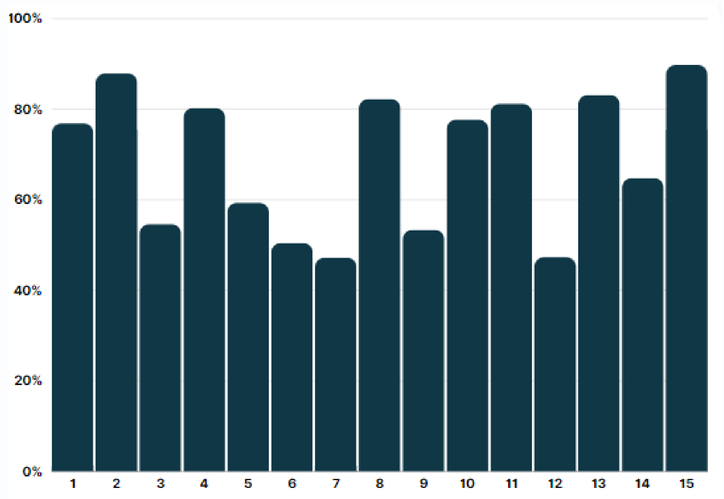

figure 6: Accuracy plots of Mask predection code on different images.

- **Cell Counting and tps calculation :-** we perform bounding box detection on a microscopic image of tumor cells. Contours are detected within each segmented region using the OpenCV library, and bounding rectangles are drawn around these contours to enclose and annotate each identified tumor area. This approach combines thresholding, and contour detection techniques to achieve accurate localization and characterization of tumors in medical images.

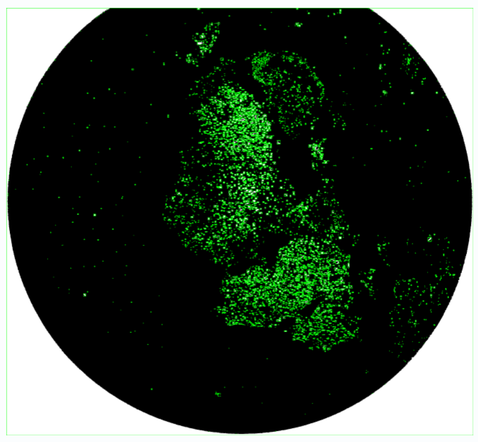

figure 7: boundary box cell counting

CONCLUSION

Our project demonstrates the effectiveness of a machine learning based approach to calculating the Tumor Proportion Score (TPS) for predict□ing the benefit of immunotherapy in cancer patients. Through a comprehensive methodology that includes various segmentation techniques, image enhancement, and mask prediction, we have achieved significant advancements in the automation and accuracy of TPS calculation . Finally, we will perform tumor proporion score(TPS) calculation.