

We selected breast cancer histology images to be used for identification of cancer cellularity. Breast cancer is the most common cancer in women and it is the second most common cause of cancer death in the United States. Having a automated quantification of tumor burden in the histopathological samples of breast cancer patients would help identify patients who have responded to neoadjuvant therapy. The database used to train this model is the SPIE-AAPM-NCI BreastPathQ dataset which is available through the the international society for optics and photonics(SPIE), along with the American Association of Physicists in Medicine (AAPM), and the National Cancer Institute (NCI). The dataset contains 2600 color images of slide sections which are labeled with the cellularity percentage. We will train a convoluted neural network, most likely customized to be used for regression analysis and will implement it in PyTorch. The performance of the network will be assessed by mean square root error and Rsquared tests. The data has been obtained and the network design will start week of April 8th. The tentative finish time is April 20th.