Estimation of Breast Cancer Density in Pathology Slides using a Convolution Deep Neural Network

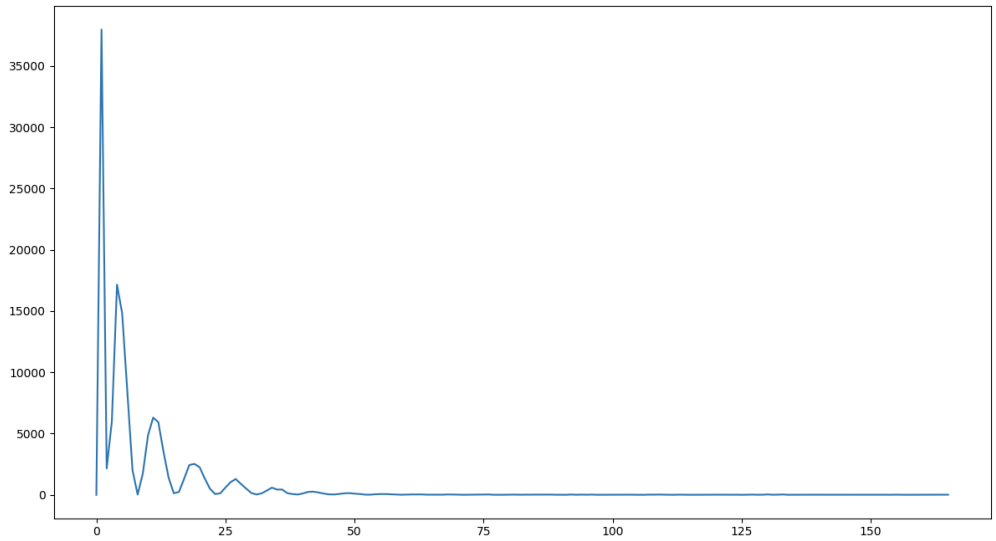
Introduction

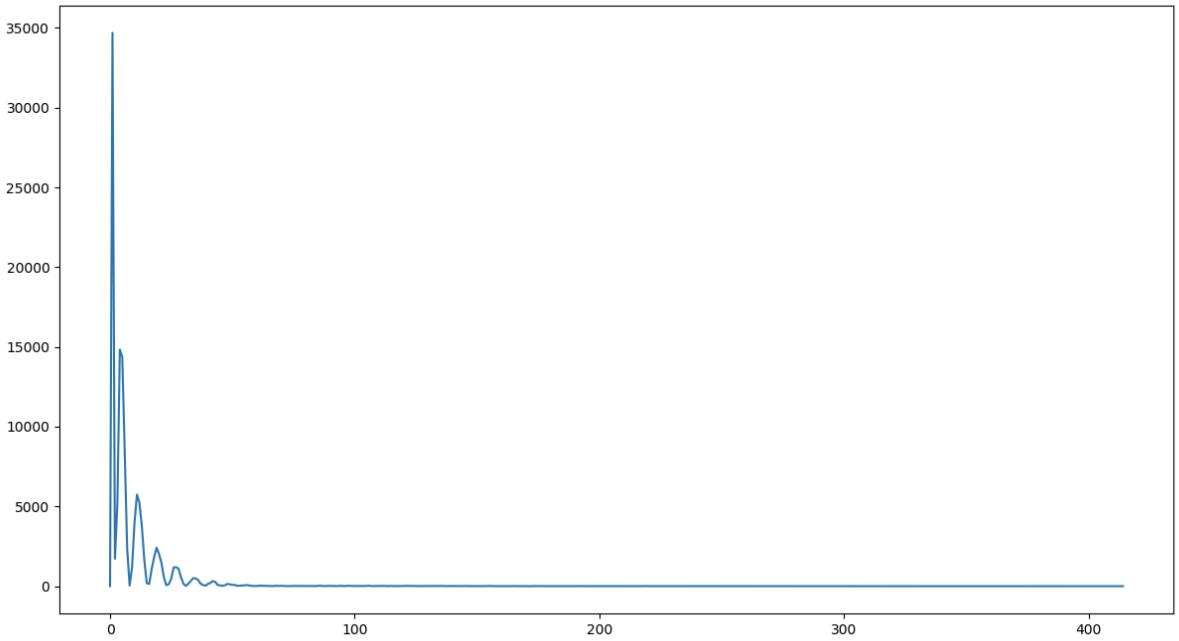
The goal of this project is predict the percentage of breast cancer cells in pathology slides of samples from breast cancer after surgical resection. When patients diagnosed with breast cancer, one of the first treatment option is surgical resection for this breast cancer. The success rate of surgical resection is high in the early stage of breast cancer. However, patients with advanced disease usually have poor prognosis. After a surgical resection, the tumor is usually send to the pathology lab and a analyzed via microscopy for staging and qualification of the type of cancer and how aggressive the cancer. This process are important to determine the type of treatment that patients needs after surgical resection, which usually include chemotherapy and radiations.

The aim of this project is develop conventional neural network algorithm to estimate the cancer cellularity for assessment of the tumor burden in these patients. This project would likely help to develop an automated method for analyzing histology patches extracted from whole

slide images and assign a score reflecting cancer cellularity in each. Developing an automated method for analyzing histopathology patches would likely improve the accuracy to quantify cancer cells as well as the efficiency of pathologists who have to identify and quantify the cancer burden. Furthermore, reproducibility of cancer cellularity scores is a concern in current practice. This report will summarize the data as well as the environment and experimental setup for achieving this goal.

In this project, I participated in pre-process the data and created the shallow convolutional network, dilation convolutional neural network, and VGG convolutional neural network. I also participated in the graphs as well as powerpoint slides. All codes and figures were uploaded in github repo. All conventional neural network architectures were performed very well in predicting and estimating cancer cellularity in the pathology slides. However, deeper networks performed better than shallow network. The following slides includes loss function for the shallow and deep neural networks.





I learned from this project that a classification model is more powerful than regression model. Deep neural network is performed very well for regression model. In this project, approximately 50% of the codes were copied from the internet.

In conclusion, this project was very helpful for me to learn how to apply convolution neural network in a regression problem.