Computational Design of Optimal Machine Learning Algorithm for Cancer Detection in Histopathologic & PET Scans

Kokan, Aisha (School: University High School of Indiana)

The use of machine learning (ML) and image recognition holds tremendous potential in the treatment of cancers like Lymphoma. Histopathologic images derived from biopsies are used to diagnose cancer. PET scans are used as a key input to decide on the staging, prognosis, and post-treatment course of action in the case of relapse. The goal of this experiment was to design and train an ML algorithm that utilizes a data set to assess if artificial intelligence and ML can recognize and distinguish a histopathologic image or a PET scan image as cancer positive or negative. The training data sets were increased in increments of twenty-five thousand images. Additionally, different types of classifiers including convolutional neural network (CNN), support vector machine, random forest, and decision tree were used. From there, over 200 iterations were set up with varying parameters. The results of this experiment showed that the optimized model was a CNN containing an increased data set and optimized parameters at about 70% maximum capacity. This resulted in increased ML and optimal recognition of the presence of cancer, achieving an area under the receiver operating curve (ROC AUC) of .85 with a sensitivity of 80% and specificity of 20%. The use of ML in histopathologic and PET scan image analysis to detect cancer allows a physician to spend more time interacting with a patient and better prescribe treatments. Future enhancements to the algorithm could be used to aid oncologists in tracking risks of secondary cancers, tracking chemotherapy-related side effects, helping recommend long-term imaging, and blood tests follow-ups with a patient decades after treatment.