A Lung Cancer Prediction and Detection System Using Nodule Based Methods and Machine Learning Algorithms

Nallamalli, Sathvik (School: Olympia High School)

Lung cancer is the leading cause of cancer-related deaths in the world. The need exists to develop an innovative diagnostic and detection measure to promptly detect the presence and type of cancer This sparked the interest to develop an innovative lung cancer detection algorithm, LCDetect. It uses complex machine learning methods and nodule analysis to detect the presence of cancer, malignity, type and stage of cancer. It uses the lung CT scan to provide a thorough diagnosis and reduce the need for PET and biopsy to prevent late detection. The algorithm works in three modules; image preprocessing, image detection, and a convolutional neural network model (CNN). Preprocessing includes noise removal, normalization, image filters, segmentation and augmentation. Extracted regions of interest are passed to image detection for feature extraction of the nodules after pixel calculations. Based on the extracted features, the layers of CNN categorize the lung cancer using location-based analysis. After several optimization techniques such as backpropagation and regression, a thorough statistical analysis was performed. Then, the algorithm was deployed on Azure Web Services. The system was trained and tested across 50,000 datasets and patient CT scans from local radiologists. The final algorithm passed with an accuracy of 98%. LCDetect is an innovative and fully functional solution to accurately detect for adenocarcinoma, squamous cell, non-small cell, and small cell lung cancer and predict the stage. Using low-level computational techniques and the input CT scan, LCDetect satisfies the goals of accuracy and performance to reduce detection time.

Awards Won:
Fourth Award of $500