It's a Matter of Life and Breath: A Novel Computer Diagnosis Tool for Early Detection of Lung Cancer SPNs towards Big Data Analysis

Jain, Aditya

Lung cancer is the leading cause of cancer related deaths in the world, yet many deaths can be avoided if malignant solitary pulmonary nodules (SPNs) are detected at an early stage of development. Current diagnosis is dependent on human intervention, which is inefficient and inconsistent. Radiologists must manually detect, classify, and contour SPNs. Additionally, patient CT scans remain on large databases and are never checked to see if a doctor missed nodules or utilized for nodule growth tracking. My project aims to solve these problems, by developing a novel automated diagnosis tool to classify malignant lung cancer SPNs at an early stage of development and be applied on patient databases to constantly improve itself, while flagging potentially cancerous nodules. My hypothesis was a MATLAB tool combining image pre-processing algorithms with further neural networking could have an accuracy of at least 80%. I first developed preprocessing methodology to detect SPNs. I experimented with many neural networks to determine which would result in best accuracy for classifying malignant lung cancer images and settled on conjugate feedforward backpropagation for training my tool. After testing my program with 1200 biopsy-confirmed CT Scans in the DICOM medical imaging format, I calculated the overall accuracy of the machine as 95.2%, exceeding my hypothesis. I further calculated sensitivity and specificity to be 96% and 95% and verified my work with Statistical and Uncertainty Analysis. The developed methodology can be used for both diagnostic and screening purposes, paving the path for big data analysis. My future work will focus on adding 3D volumetric analysis and testing my program on hospital databases.

Awards Won:
Fourth Award of $500