CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

Albert Qin; Samyak Surti

Project Number

S0826

Project Title

A Machine Learning Based Approach to Decrease the Lung Cancer Malignancy Detection Threshold

Abstract

Objectives/Goals
When creating our Machine Learning model, our objective was to be able to correctly identify the malignancy of the Lung presented based on the CT scan images. This was achieved through three main criteria. These included maximizing the classification accuracy of the model, minimizing the model's training time, and minimizing the model's overall run-time. We also wanted to measure the accuracy of our model in comparison to Google's Inception model. Our hope is that our model can be utilized in doctor's offices to forego the need to do biopsies or any other stressful surgeries.

Methods/Materials
Equipment: PC equipped with Nvidia GeForce GTX 980 GPU, Laptop with Nvidia GeForce GTX 1050. Software Components: Tensorflow - Google's Machine Learning and Artificial Intelligence Library for Python, CT scan image database acquired from a cancer imaging database, Google's Inception Deep Learning model

Results
After testing our model, we achieved around a 96% accuracy when classifying the cancer the patient had as being Benign, Malignant, or Metastatic, based on the testing data of CT scan images. The data that we used was sectioned off into training and testing sets to insure that the images in the testing sets haven't been seen by the model during training. This would result in an unbiased accuracy output. When compared to Google's Inception Deep Learning model, we consistently achieved around a 20 to 30% higher accuracy. These results were especially surprising, as we were able to achieve very high accuracy from a simplistic model in contrast to Inception which is extremely complex.

Conclusions/Discussion
Based on the accuracy we achieved with our model, we want to refine it so that doctors, especially radiologists and pulmonologists, can utilize this software to make an accurate diagnosis of the patient's condition without having to perform any biopsies. This will get rid of any stress or anxiety that is often attached to such procedures. By getting an accurate diagnosis of malignancy of the cancer early on, respective action can be taken without any delay, giving the patient peace of mind.

Summary Statement

By creating a simple Convolutional Neural Network, we were able to achieve a surprisingly high accuracy when diagnosing the malignancy of the patient's Lung Cancer.

Help Received

I received help from my dad in understanding some of the fundamentals of Machine Learning.