Lung Cancer Detection

Group:- 31

BACKGROUND

Lung cancer, also known as lung carcinoma, is a malignant lung tumor characterized by uncontrolled cell growth in tissues of the lung.

Worldwide in 2012, lung cancer occurred in 1.8 million people and resulted in 1.6 million deaths.

This makes it the most common cause of cancer-related death in men and second most common in women after breast cancer.

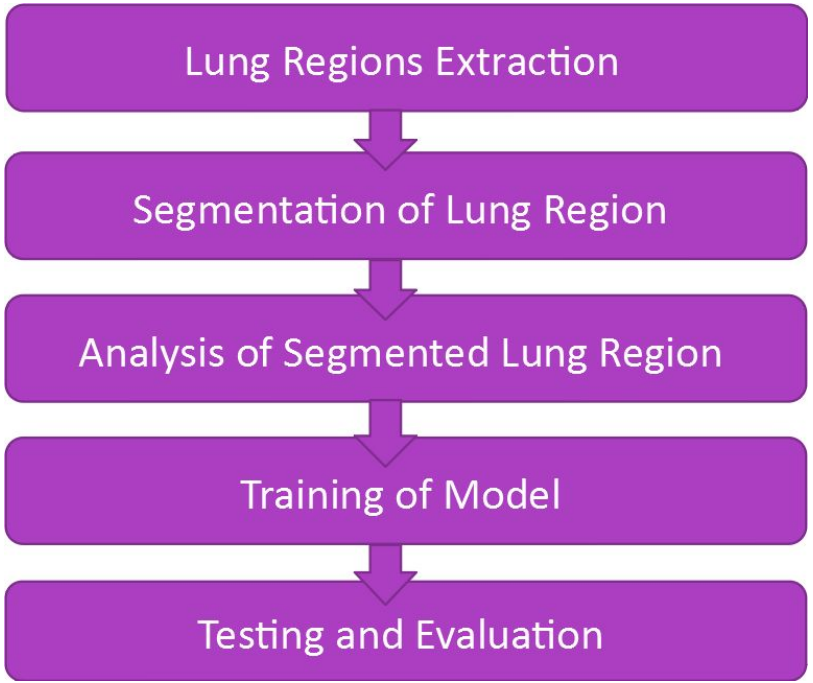
However, only 15% of diagnosed lung cancers are at this early stage.

The survival rate for the cancer patient can be increased by detecting the occurrence of cancer in earlier stages.

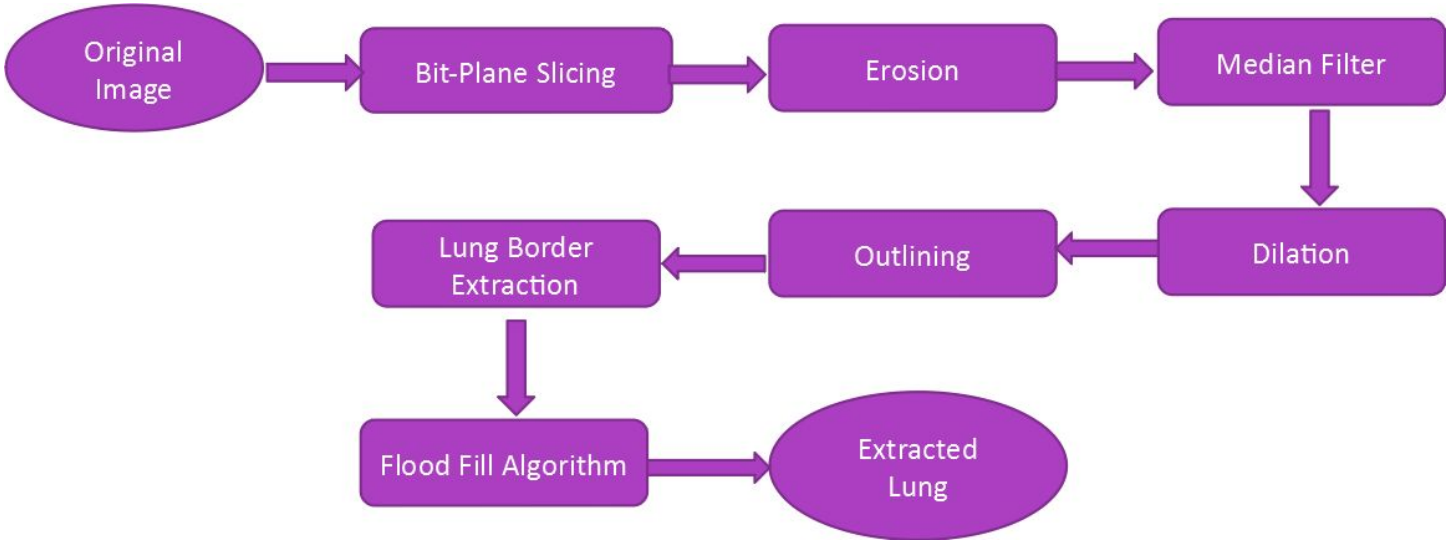
Problem Statement

Predict lung cancer from the analysis of CT images of chest

Methodology



Lung Region Extraction



Lung Regions Segmentation

Identifying the region of interest(ROIs) which helps in determining the cancer region.

Using K-Means(K=8) to do the clustering.

Getting 8 candidates of lung nodule

Features Extraction

Extracting three kinds of features from a CT image:

Area of the candidate region

The Maximum Drawable Circle (MDC) inside the candidate region

Mean intensity value of the candidate region

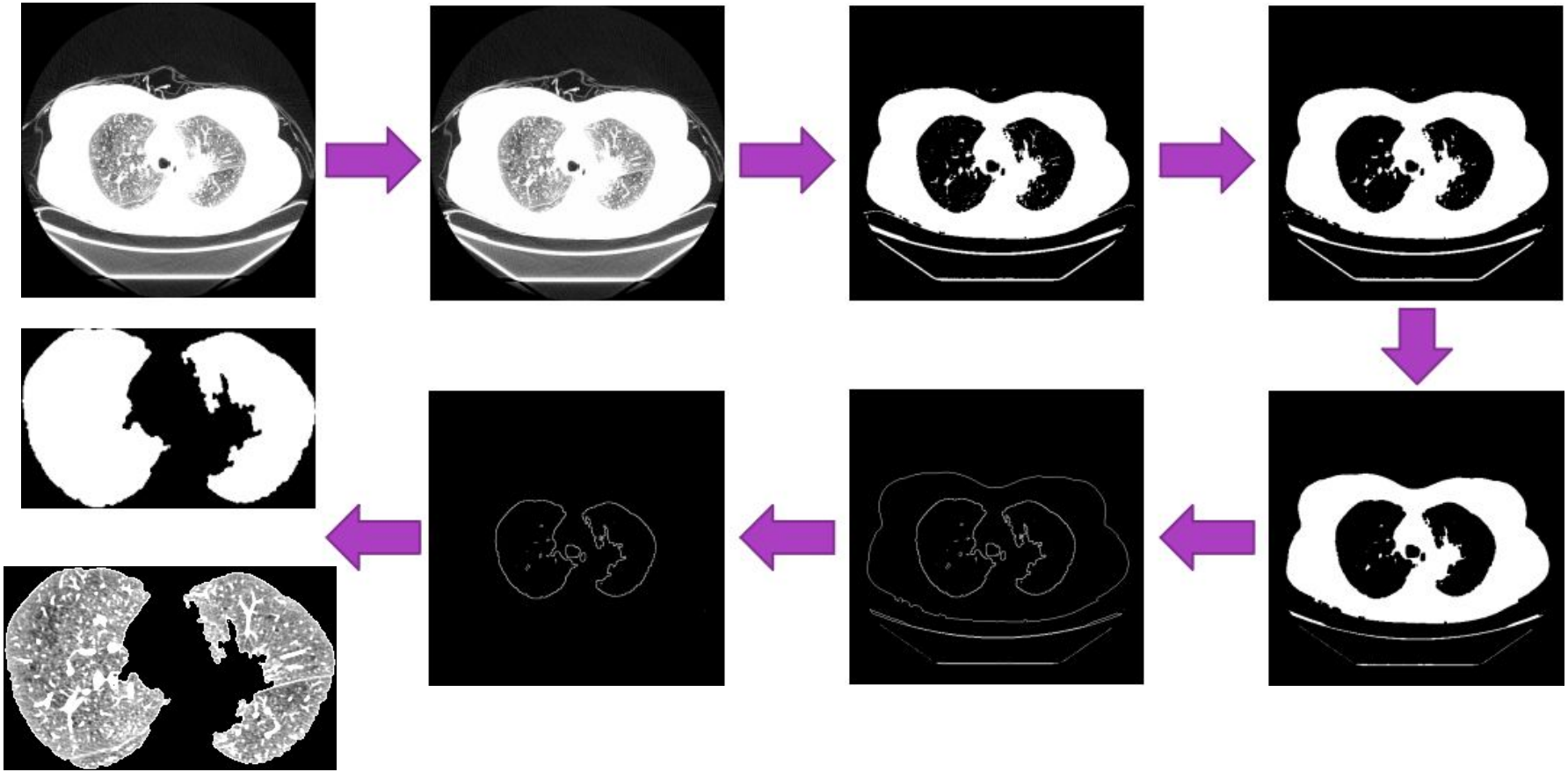
Getting a 3x8 matrix(3 features * 8 nodule candidates).

Taking an average of 8 candidates and getting a 3x1 matrix to represent this CT image

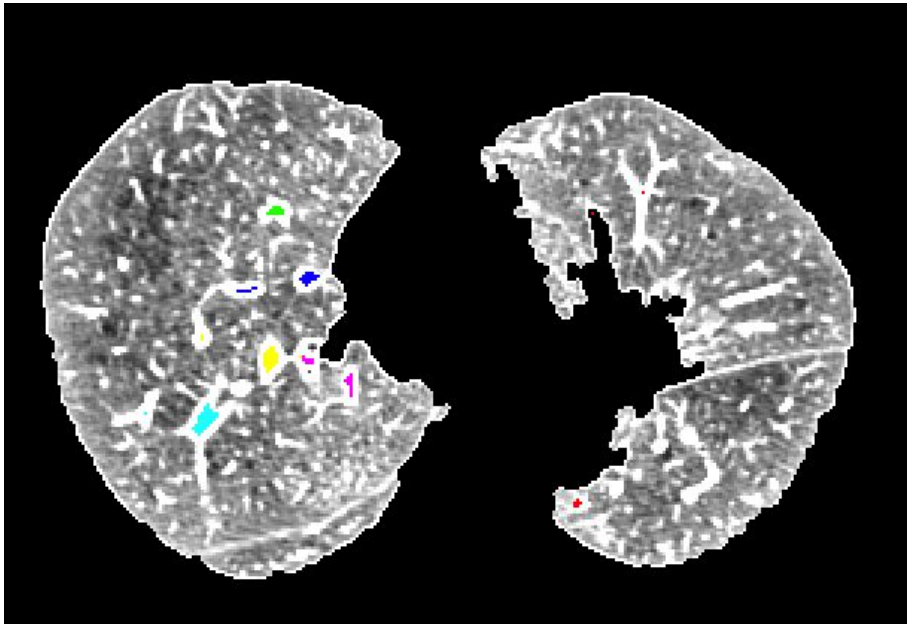
Model Training and Testing

We trained a Neural Network with 10 neurons to gain accuracy of 70.5% and have found a trained Neural Network Which Detects the false positive nodules accurately by 83.81%.

Results:lung extraction



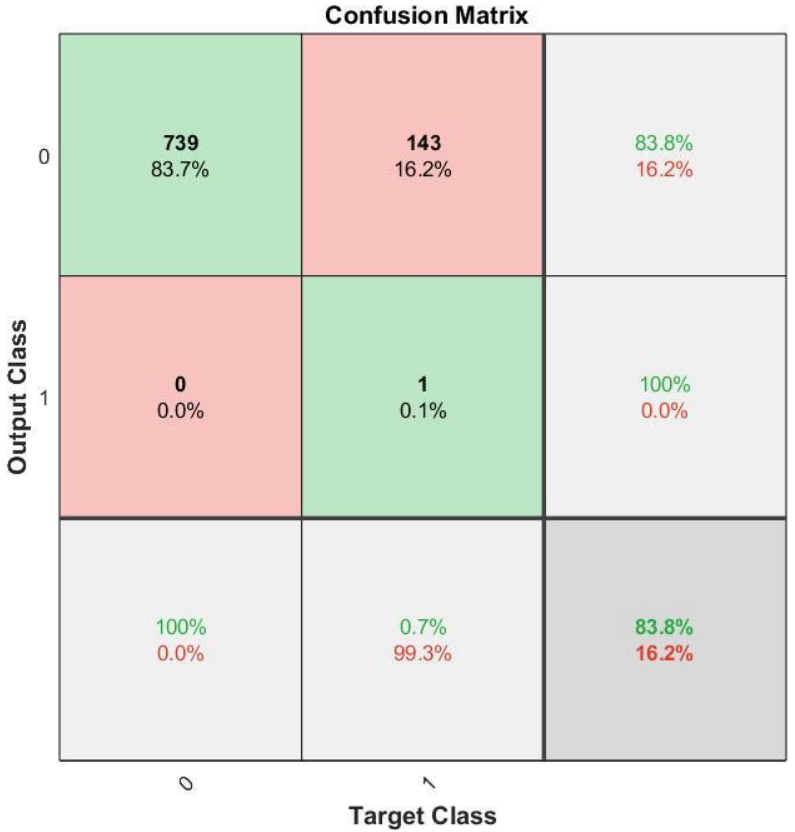
Results:Segmentation



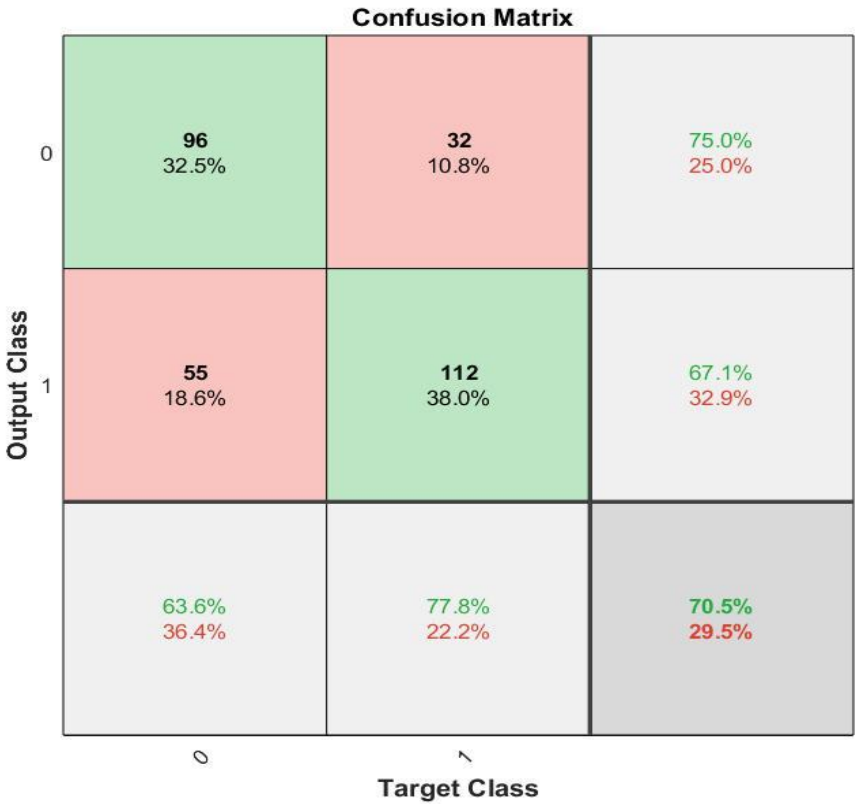
Analysis

Pre-trained network which predict false

Positive nodules



Self Trained Network



Difficulties and solutions

- CT images were not clean.
  - We used Erosion, Medium filtering, Dilation to obtain a clean image
- Extraction of both the lung region
  - We first eliminated the longest boundaries (outermost boundaries) than the next to largest boundaries are the two lung regions.
  - After that we flood filled the lung regions followed by cropping the two lung regions from the clean CT image.
- Feature Extraction:-We observed about 800 patients. For each patient we observed 10 CT images of different levels of lung and took 3 features from 8 nodules in each image.
  - We took mean of all the corresponding features so as to get 3 features for each patient.
- We tried many different training functions for Neural Network.
  - Trainlm(Levenberg-Marquardt) gave the best result out of all.
- We tried different sizes of hidden layer and found hidden layer with 10 neurons worked the best giving accuracy of 70.5%

References

- https://github.com/AustinNeverPee/LungImageAnalysis
- https://en.wikipedia.org/wiki/Erosion_(morphology)
- https://en.wikipedia.org/wiki/Dilation_(morphology)
- https://edoras.sdsu.edu/doc/matlab/toolbox/images/morph13.html

Team

- Kushal Jain
- T.Ooha
- Pratyush
- Sashikanth