Spiro Sound

Jake Garrison, Farshid Salemi Parizi

University of Washington · University of Washington Electrical Engineering · ubicomplab

Motivation

The standard way to measure a patient's lung function is by visiting a pulmonologist and performing a spirometry test that involves exhaling into a $5000 machine which plots the airflow verse volume of air. The pulmonologist extracts a few key points from this curve and can diagnose obstructed lung conditions. The main value that determines obstruction is the FEV1, or total volume of air exhaled within one second.

Dataset and Approach

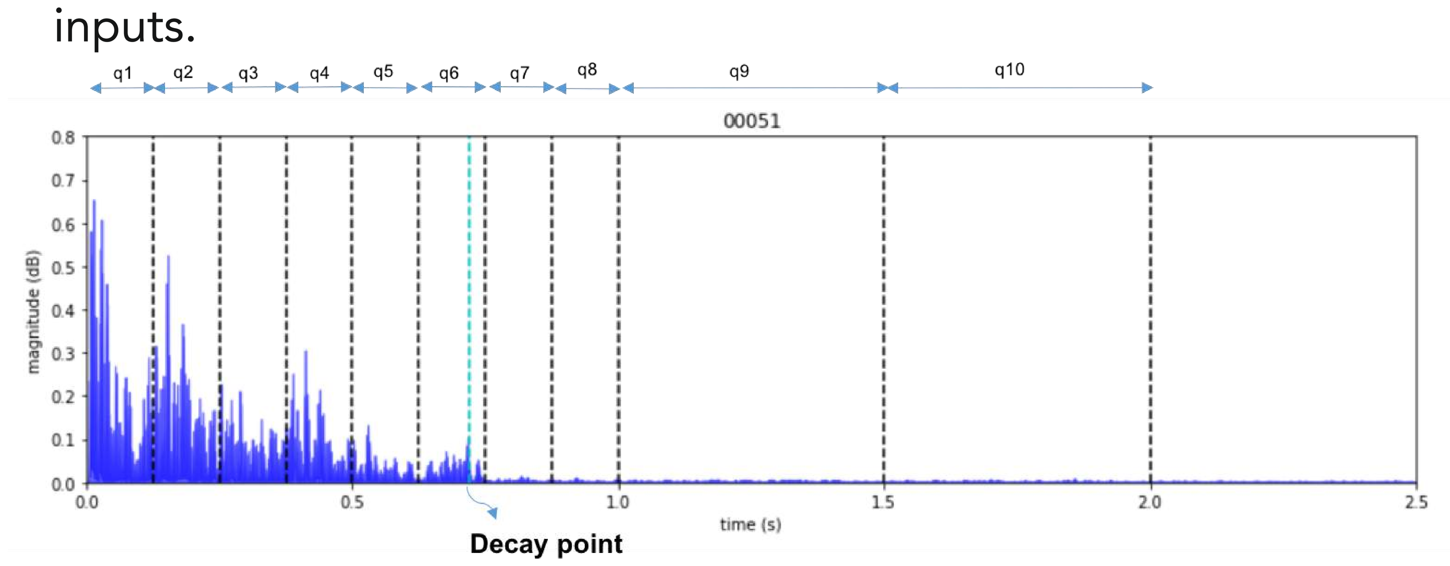
We have a dataset of several audio recordings from a variety of healthy and unhealthy patients with multiple audio recordings per patient. In addition, each audio recording has ground truth data to use as regression points. We have used two approaches to determine lungs obstruction.

1. Classical regression techniques to predict the FEV1 using methods such as random forest, gradient boost, SV, Ridge and KNN.
2. Deep learning using spectrogram input to classify on a scale of 1-16 of a patient's lung function. The output layer probabilities are then used to regress to a FEV1 predication

Input Features

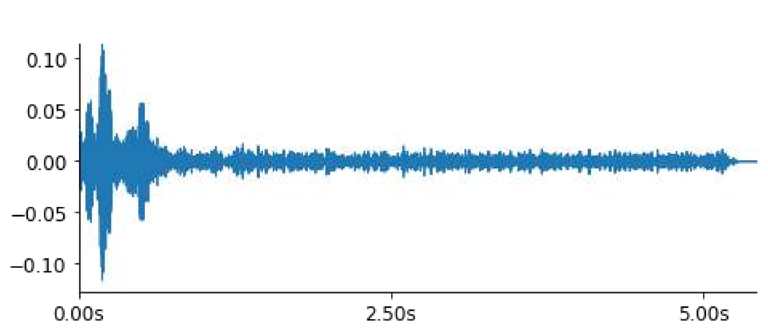
Machine Learning

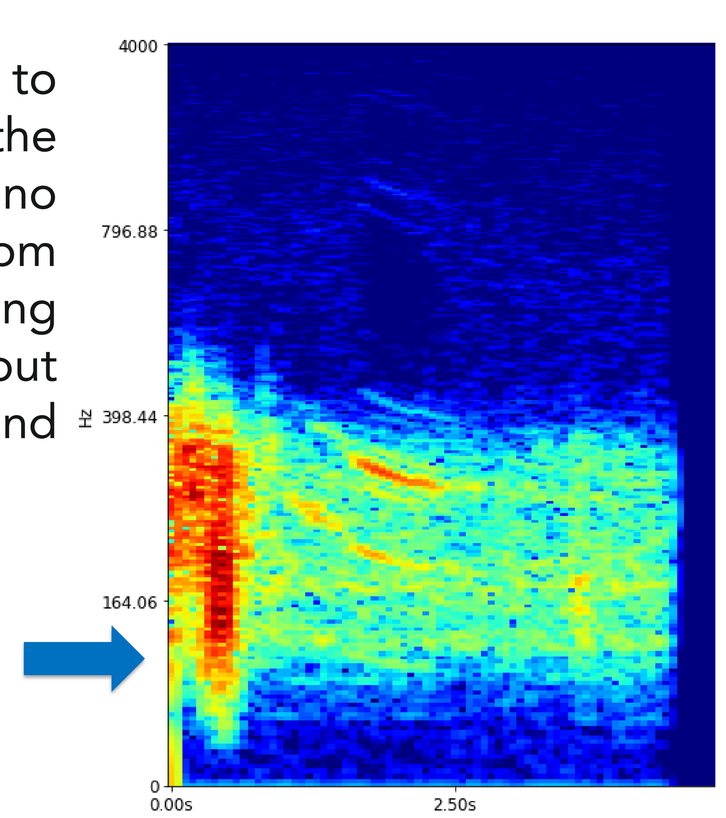
For the machine learning approach we used the audio signal's amplitude envelope to extract **temporal features** like the total area under the sound, the peak and etc. Also we included **patient features** like gender, smoker and etc. to our feature inputs.



Deep Learning

For deep learning we wanted to see how well we could predict the FEV1 with **only the audio data**, no prior information like gender from the patient. We ended up using the melpower spectrograms, but would like to try other inputs and architectures

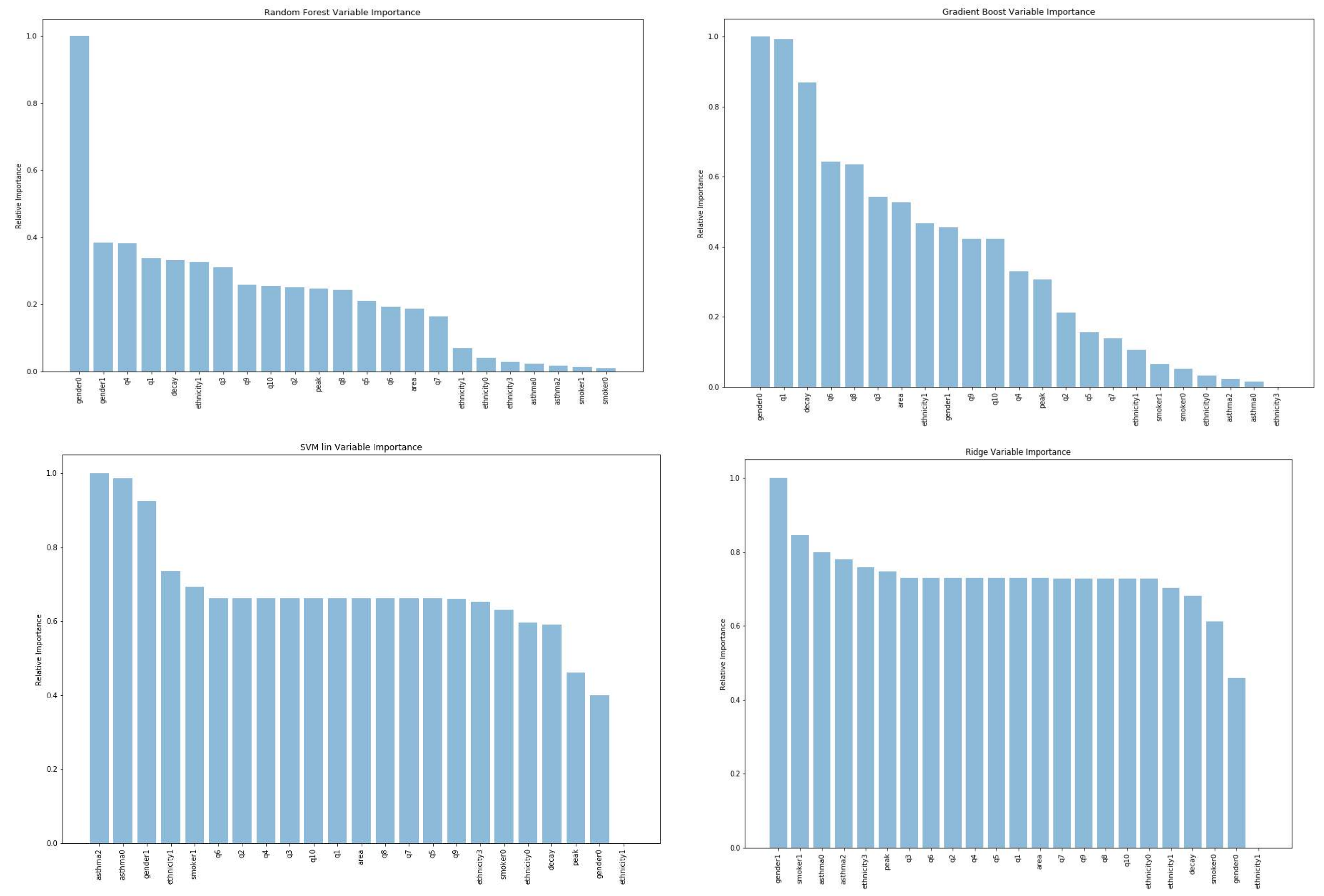




Machine Learning Approach

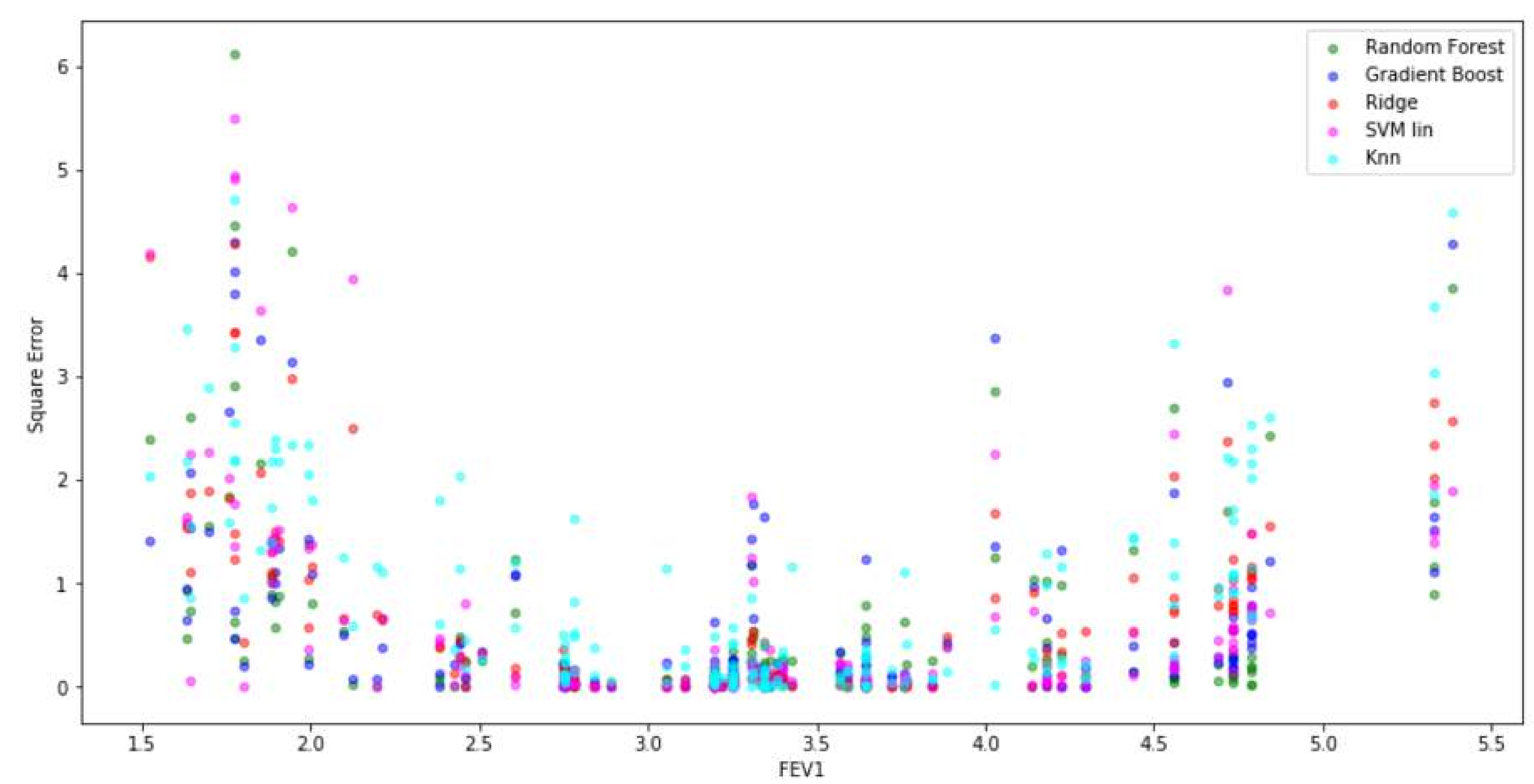
Technique	Random Forest	Gradient Boost	SVM Linear	Ridge	KNN
MSE	0.677	0.717	0.733	0.733	0.875

Feature Importance



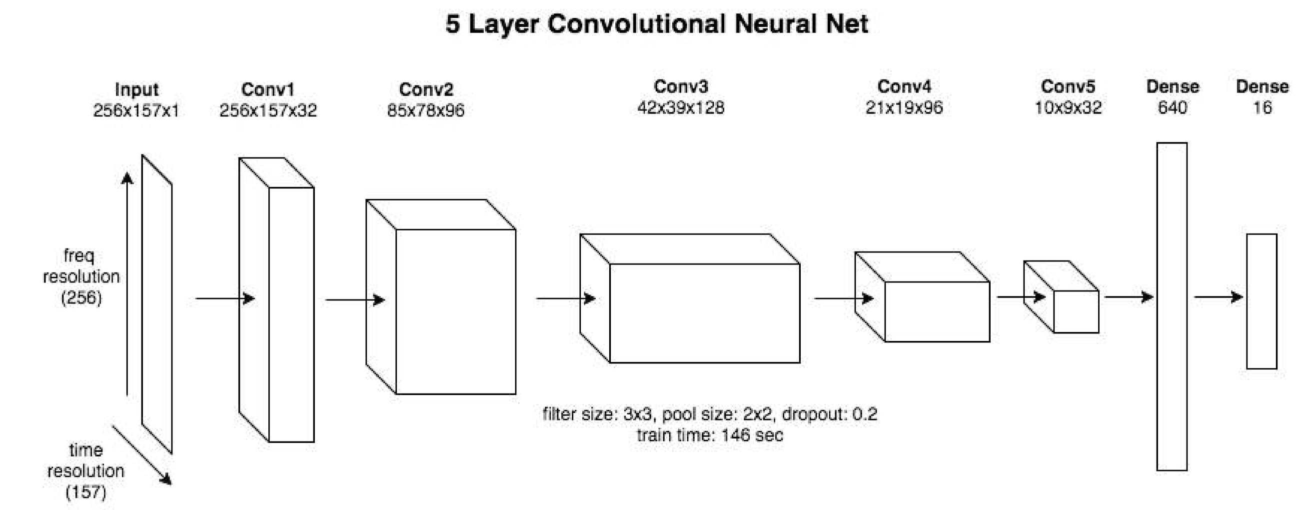
Squared Error

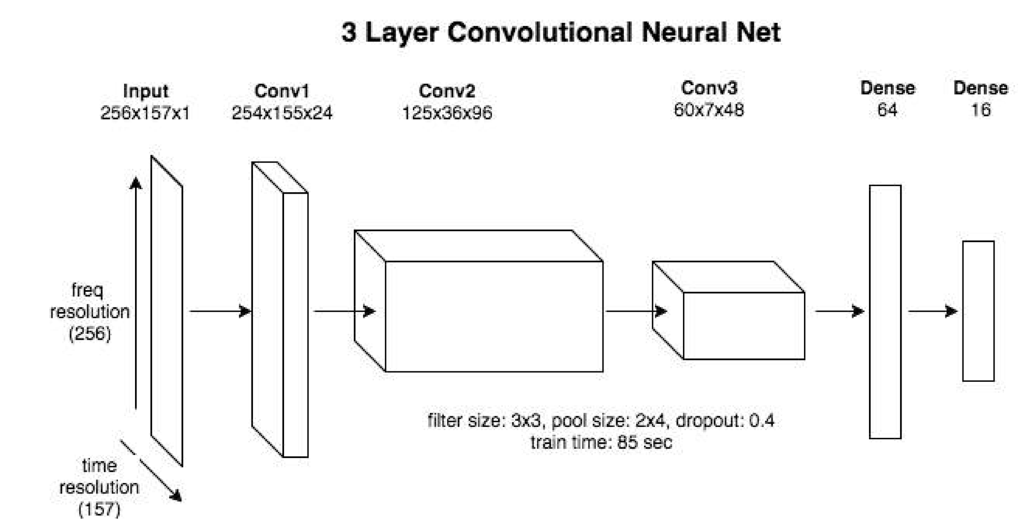
Squared error for each model at each test FEV1 value



Deep Learning Approach

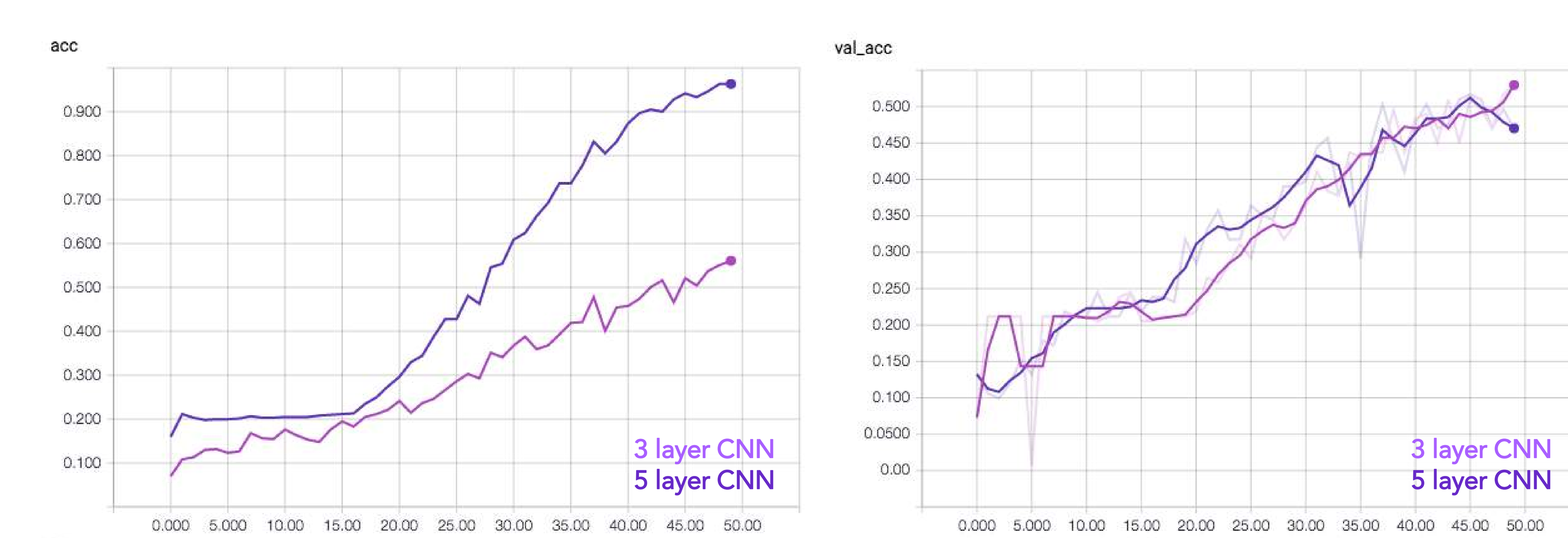
Technique	Classification Accuracy	Regression MSE	Train Time (seconds)
3 Layer	0.47	1.08	85 sec
5 Layer	0.52	1.14	146 sec





Training and Validation accuracy

We used multilayered CNN as a classifier to predict the FEV1 from 16 classes (1.25 to 5.5 in steps of 0.25), optimizing for accuracy. Using the softmax layer of probabilities, we regressed to a single FEV1 value and then computed MSE from this.



Bland-Altman Error Plots