Environmental Noise Contamination Detection

Todd Schultz, Rahul Birmiwal, Sean Miller

Project Overview

Background

Regulatory agencies of aviation are looking to reduce the amount of noise generated by aircraft while the number of flights worldwide continues to increase. The noise certification process for aircraft is expensive and requires monitoring personnel. Due to the remote location of these tests, contamination of the audio data is possible from birds, insects, various wildlife, and road traffic.

Objective

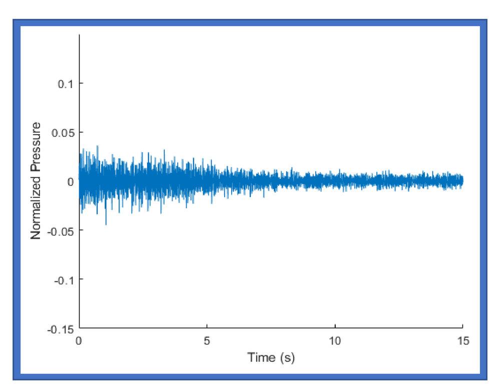
Research the feasibility of automating the detection of environmental noise contamination contained in acoustic measurements for aircraft community noise testing.

Data

This project leveraged 66 audio files that were provided by Boeing Test and Evaluation and 6 audio files from the US National Parks Service. Altogether we have 49 recordings of aircraft and 10 ambient recordings to make up our clean signal data. We have 13 different files to use as contamination. The data provided by Boeing have been anonymized through normalization.

Process

To generate the data set a clean signal file is first split into blocks of length *t*. If designated to be a file with contamination audio, an unclean signal is added at the desired signal-to-noise ratio to allow for finer-grained control. Features are then generated for each block and the resulting feature set can be handed to a model for training.



Methods

Our investigation into the viability of automated detection of environmental noise contamination is broken into two parts. A broad assessment of feature sets and models followed by a deeper investigation into three of the best performing pairs.

Broad Investigation

To get a general idea of how different feature sets and models perform on this classification task, our initial broad investigation's goal was to enumerate as many feature set, model pairings as possible while limiting the block size and holding the signal-to-noise ratio constant.

For this investigation, models were trained with their default hyperparameters and settings.

Deep Investigation

After selecting three of the best performing feature set, model pairs, our more thorough investigation includes each of the following steps to understand how they might behave in a more real-world scenario.

1. Optimize the models hyperparameters.
2. Randomize the signal-to-noise ratio.
3. Randomize the percent of a block that contains contaminated audio.
4. Randomize which blocks are contaminated.
5. Vary all of above parameters through Monte-Carlo simulation.

Broad Investigation Results

Feature Set	Block Length (s)	Logistic Regression	Fine Tree	Medium Tree	Coarse Tree	Boosted Trees Ensemble	Bagged Trees Ensemble	Linear SVM	Quadratic SVM	Cubic SVM	Fine Gaussian SVM	Medium Gaussian SVM	Coarse Gaussian SVM	Subspace KNN	Neural Net	CNN	LSTM NN
Modified MFCC	1	75.8%	83.1%	81.1%	75.6%	85.0%	86.8%	77.2%	87.3%	89.1%	85.7%	88.0%	80.4%				
Modified MFCC	2	78.3%	85.2%	82.5%	76.3%	86.3%	88.2%	79.0%	89.5%	***91.9%***	88.6%	89.3%	80.5%		86.7%		
Modified MFCC	1	78.0%	81.8%	78.9%	75.6%	85.0%	85.5%	79.0%	88.8%	89.7%	76.3%	88.1%	79.7%				
Modified MFCC	2	80.2%	85.2%	81.4%	74.3%	85.9%	88.3%	80.4%	***90.7%***	***92.8%***	80.9%	***90.3%***	79.7%		***92.5%***		61.8%
1/3 Octaves	1	70.7%	83.7%	80.2%	75.3%	85.1%	89.0%	70.3%	77.5%	81.2%	71.9%	75.4%	70.1%				
1/3 Octaves	2	72.8%	87.5%	82.8%	77.3%	87.5%	***91.9%***	71.7%	79.7%	83.8%	76.6%	76.5%	71.5%		71.5%		
Octaves	1	67.3%	82.0%	79.4%	72.7%	82.2%	86.3%	66.6%	76.4%	80.0%	78.0%	75.6%	66.8%				
Octaves	2	67.8%	84.7%	80.6%	73.7%	84.2%	89.0%	66.5%	77.0%	81.2%	81.1%	76.8%	66.9%		77.4%		
MFCC	2	76.3%	82.0%	78.2%	72.1%	83.5%	85.3%	76.8%	85.2%	87.4%	83.8%	87.0%	77.0%				
MFCC	1	74.1%	80.4%	78.9%	68.8%	81.6%	82.9%	75.3%	83.0%	84.6%	80.5%	84.8%	76.4%		77.5%		
FFT (25 Hz resolution)	0.041	61.3%	78.4%	75.8%	70.0%												
FFT (100 Hz resolution)	0.01	68.3%	76.8%	74.0%	69.6%												
CWT Scalogram	1															***92.9%***	
CWT Scalogram	3															85.1%	
CWT Scalogram	5															83.3%	
CWT Scalogram Bag of Features	1	81.0%	75.5%	74.2%	73.3%	82.1%	82.0%	83.1%	86.4%	87.1%	60.5%	86.1%	79.9%				
CWT Scalogram Bag of Features	3	80.7%	73.2%	73.7%	75.3%	84.8%	82.7%	86.7%	87.7%	88.6%	59.1%	87.9%	81.6%				
CWT Scalogram Bag of Features	5	55.0%	73.0%	74.1%	70.6%	82.3%	81.7%	88.0%	89.4%	***90.9%***	60.2%	87.7%	81.6%				
DWT (Coiflet2, 4 levels)	1	70.0%	***90.4%***	88.3%	78.1%	***91.2%***	***93.1%***	80.1%	86.2%	88.6%	84.8%	82.3%	76.0%		81.5%		76.4%
DWT (Coiflet2, 4 levels)	2	75.4%	***91.6%***	88.1%	80.1%	***91.0%***	***93.3%***	80.5%	87.3%	***90.5%***	85.8%	82.4%	77.6%		81.5%		78.2%
DWT (Coiflet2, 5 levels)	2	81.5%	***91.7%***	89.1%	79.3%	***92.3%***	***94.2%***	82.1%	88.0%	***90.6%***	87.2%	83.5%	78.8%		82.3%		
DWT (Coiflet2, 5 levels, Hampel Filter)	2	81.0%	***91.7%***	88.7%	79.3%	***92.1%***	***94.0%***	82.0%	88.9%	***91.5%***	87.6%	83.8%	76.8%	76.5%	84.9%		
DWT (Coiflet2, 4 levels, Hampel Filter)	2	70.1%	***92.1%***	88.6%	80.0%	***91.1%***	***93.5%***	80.7%	87.9%	***90.8%***	86.3%	82.7%	77.8%		83.1%		
DWT (Coiflet2, 3 levels, Hampel Filter)	2	67.1%	89.7%	87.5%	78.1%	89.6%	***92.9%***	79.0%	84.4%	52.0%	85.9%	81.1%	76.5%	81.3%	81.4%		
DWT (Debauchies4, 4 levels)	1	64.4%	90.1%	86.0%	81.1%	88.9%	***91.6%***	77.2%	80.6%	51.4%	81.5%	77.6%	66.1%	71.1%			
DWT (Haar, 4 levels)	1	53.8%	77.0%	73.1%	68.7%	76.4%	80.5%	71.5%	76.4%	58.7%	76.1%	73.9%	70.4%	67.1%			
DWT (Coiflet2, 4 levels, no entropy)	2	71.6%	***90.6%***	85.7%	81.6%	88.3%	***92.1%***	79.2%	57.1%	49.9%	76.3%	67.9%	61.0%	54.1%			
DWT (Coiflet2, 2 levels)	2	70.1%	89.0%	84.6%	78.9%	88.2%	***92.3%***	68.8%	71.8%	53.9%	77.7%	74.0%	61.5%	73.7%			

The three feature set, classifier pairs we selected from this investigation are as follows:

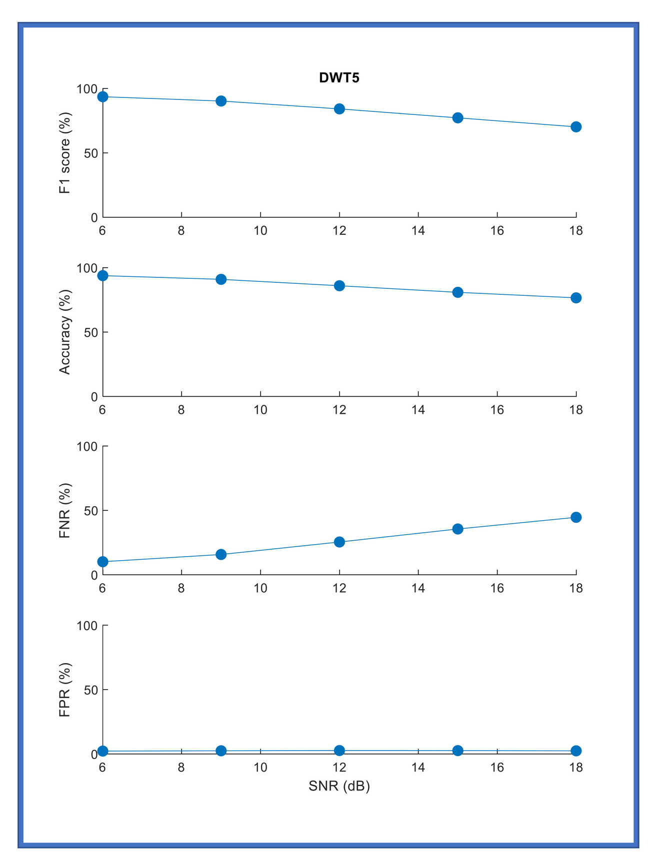
- **Model**: Bagged Trees
- Wavelets (Coiflet2, 5 level, T=2s)
  - Best overall with accuracy of 94.2%

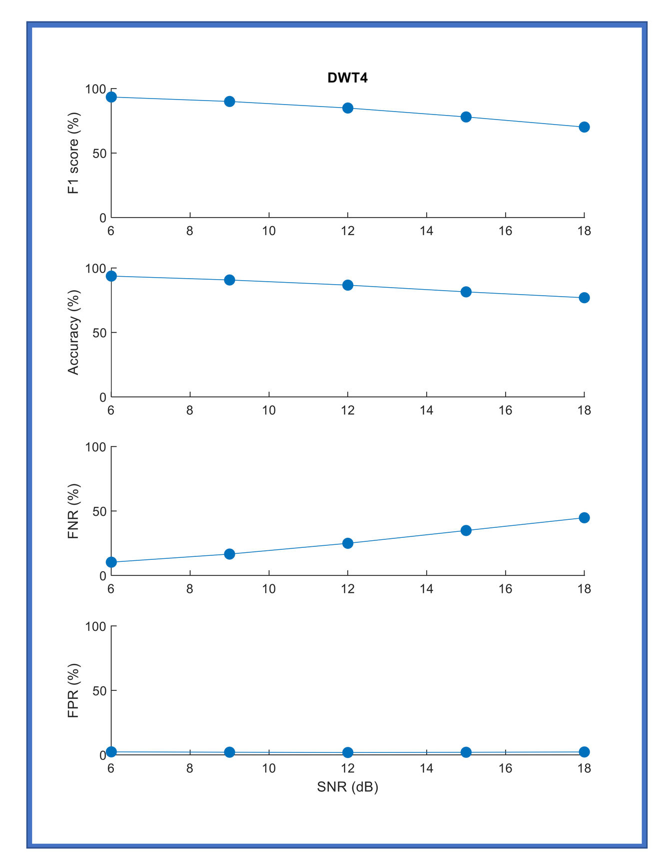
- **Model**: Bagged Trees
- Wavelets (Coiflet2, 4 level, T=2s)
  - Accuracy of 93.3%

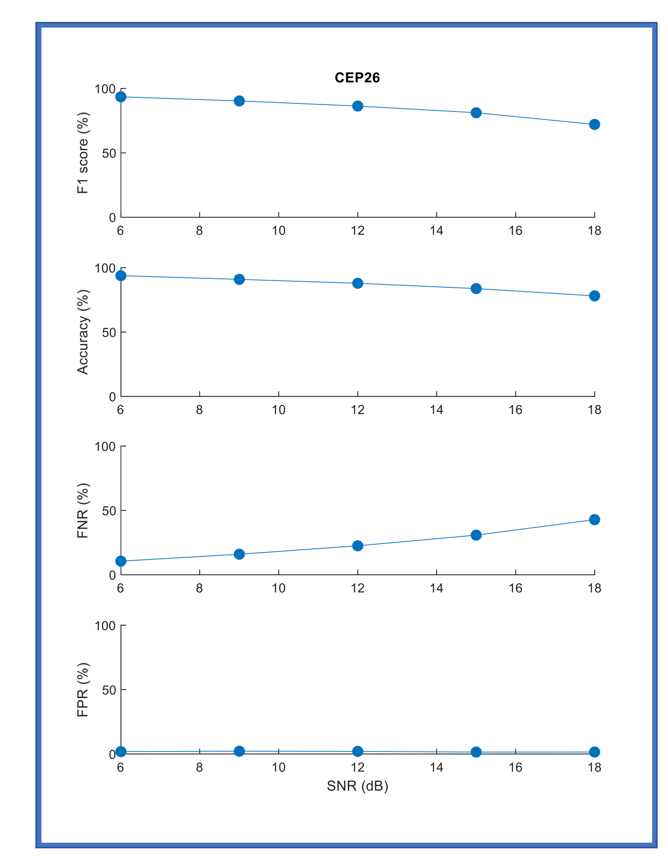
- **Model**: Cubic SVM
- Cepstral (26 feature T=2s)
  - Best non-NN, non-wavelet performer with accuracy of 92.8%

Deep Investigation Results

Signal-To-Noise Ratio – Parameter Sweeps







Performance summary of optimized feature set and classifier pairs for detailed study subjected to the randomized signal blocks contaminated

	F1 SCORE	ACCURACY	FNR	FPR
DWT5	93.4%	93.6%	10.4%	2.4%
DWT4	93.6%	93.8%	10.1%	2.2%
CEP26	93.6%	93.9%	10.3%	1.9%

Challenges

How To Turn Audio Into Features

We separated each audio file up into blocks of an equivalent length and then processed the features for each block. Blocks from which features were generated could contain overlapping audio.

How To Quantify Contamination Levels

The audio recordings that were used only had a single source of noise allowing us to combine aircraft or ambient signals with contamination to create a larger data set than the 66 files. This also allowed us to control the signal-to-noise ratio.

Limitations

Due to the small size of potential contamination audio, the models are not as accurate when classifying new sources of contamination.

One recommendation that we do have at this time is that we suggest that more data is provided to fully enumerate the possible contamination classes that would exist near the test site.

Recommendations

This section is currently a work in progress as we have not finalized our Deep Investigation.

Acknowledgements

We would like to thank our sponsor for the project, **Steven Underbrink** as well as our Data Science Capstone advisor: **Dr. Megan Hazen**, University of Washington.