SUAS Detection and Classification

Classical and Deep Learning Based Approaches to Signal Analysis

[CONTROLLED UNCLASSIFIED]

US Air Force Research Laboratory
Information Directorate

June 11, 2024

Agenda

1. Introduction
2. Methodology
3. Future Work

Introduction

Motivation

- Detection and classification of commercially available SUAS devices presents great difficulties
- Audio signals recorded in real-world scenarios contain noise from various sources which increases the difficulty of identifying the presence of a UAS
- An automated, robust audio separation algorithm increases the likelihood of successful classification via supervised machine learning approaches

Digitization + Documentation

- Streamlined workflow, centralized code, and compiled data via GitHub and Google Drive
- AudioSeparation.ipynb - contains basic NMF functionality, ability to plot and save results, and combine specific components

Methodology

Data

- Original dataset consists of audio signals recorded by the DADS (Drone Acoustic Detection System) and DARA (Drone Acoustic Recording Array) devices during the ESCAPE II Data Collection
- Audio files collected during SUAS flights contain various sources of noise i.e., insect noises, wind, or chatter
- Each signal has been broken down into 5-second chunks and labeled with the SUAS device in the air during the recording
- Presence of SUAS in each chunk approximated based on whether the energy (np.sum(signal^2)) is greater than or equal to the 25th percentile of energy from the full-length signal
- Performed this process for each audio file, yielding a labeled dataset for future use with supervised machine-learning approaches

Aircraft Information

				
Aircraft	**Inspired Flight (USA)**	**DJI**	**DJI**	**DJI**
Model	IF 1200	Matrice 600	Phantom 4 Pro V2	Mavic Air 2
Type	VTOL	VTOL	VTOL	VTOL
Use	Payload Carrier, Red Target	Red Target	Red Target	Red Target

Table: Escape II Data Collection SUAS devices deployed

Non-negative matrix factorization (NMF)

Problem Statement

Given a non-negative matrix $V \in \mathbb{R}^{m\times n}$, find non-negative matrices $W \in \mathbb{R}^{m\times r}$ and $H \in \mathbb{R}^{r\times n}$ such that:

$$V \approx WH$$

where r is the rank of the factorization

Interpretation

- W represents a basis matrix:
  - Each column of W represents a basis vector in the space of the original data V
- H represents a coefficient matrix:
  - Each column of H represents the temporal activation coefficients to express the corresponding basis vector

NMF For Audio Source Separation

- Used to decompose the original signal into N spectral components, each capturing a relevant feature
- The dot product of W and H yields an approximation of V:

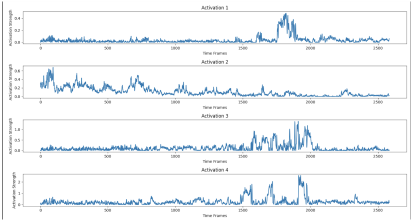
$$W \odot H \approx V$$

- NMF allows you to select specific spectral components and add/remove them from a recreated signal:

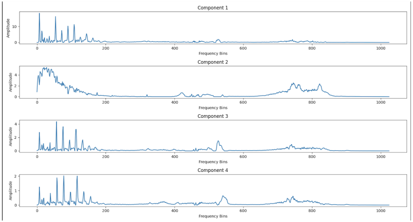
$$V \approx W[:,1] \cdot H[1,:] + W[:,2] \cdot H[2,:] + \dots$$

  where $W[:,j]$ represents the j-th basis vector and $H[j,:]$ represents the coefficients for the j-th basis vector

Visualization



(a) First four columns of H (temporal activation coefficients)



(b) First four basis vectors of W

Non-negative Matrix Factorization (NMF) Optimization

The goal of NMF is to minimize the difference between the original matrix V and the product of the basis matrix W and the coefficient matrix H. This is typically achieved by minimizing a cost function, such as the Frobenius norm of the difference:

$$\min_{W,H} \frac{1}{2}||V - WH||_F^2 \quad \text{subject to} \quad W, H \geq 0$$

The optimization problem is solved using iterative update rules for W and H:

$$H_{ij} \leftarrow H_{ij} \frac{(W^T V)_{ij}}{(W^T WH)_{ij}}$$

$$W_{ij} \leftarrow W_{ij} \frac{(VH^T)_{ij}}{(WHH^T)_{ij}}$$

These update rules are applied alternately until convergence or a maximum number of iterations is reached. The non-negativity constraints on W and H ensure that the factorization yields interpretable and meaningful components.

Flowchart goes here

NMF for Audio Source Separation

In the context of audio source separation, NMF is applied to the magnitude spectrogram of the audio signal. The spectrogram is obtained by performing a Short-Time Fourier Transform (STFT) on the time-domain signal:

$$V = |\mathrm{STFT}(x)|$$

where x is the time-domain audio signal and V is the resulting magnitude spectrogram. NMF decomposes the spectrogram into a set of spectral basis vectors W and their corresponding temporal activation coefficients H:

$$V \approx WH$$

Each basis vector in W captures a specific spectral pattern, while the corresponding row in H represents the temporal activation of that pattern.
To separate specific sources or remove noise, relevant basis vectors can be selected and used to reconstruct a modified spectrogram:

$$\tilde{V} = \sum_{j \in J} W_{:,j} H_{j,:}$$

where J is the set of indices corresponding to the selected basis vectors. The modified spectrogram $\tilde{V}$ can then be used to synthesize the separated or denoised audio signal using the inverse STFT.

Flowchart goes here

Matrix Embedding and Optimization [IN PROGRESS]

The generalized noise type matrices are first horizontally concatenated to form a noise basis matrix W_{noise}:

$$W_{\text{noise}} = [W_{\text{type1}}; W_{\text{type2}}; W_{\text{type3}}]$$

where W_{type_i} represents the basis matrix for the i-th noise type. The noise basis matrix W_{noise} is then embedded into a larger basis matrix W along with randomly initialized signal basis vectors W_{signal}:

$$W = \left[\begin{array}{ccc|c} \boxed{W_{\text{type1}}} & \boxed{W_{\text{type2}}} & \boxed{W_{\text{type3}}} & \boxed{W_{\text{signal}}} \end{array} \right]$$

The corresponding activation matrix H is also initialized with *entirely* random values:

$$H = \left[\begin{array}{c} \boxed{H_{\text{noise}}} \\ \boxed{H_{\text{signal}}} \end{array} \right]$$

During the optimization process, the signal basis vectors W_{signal} and the entire activation matrix H are updated iteratively using the multiplicative update rules, while the noise basis vectors W_{noise} remain fixed:

$$H \leftarrow H \odot \frac{W^T V}{W^T (WH)} \qquad W_{\text{signal}} \leftarrow W_{\text{signal}} \odot \frac{V H_{\text{signal}}^T}{(WH) H_{\text{signal}}^T}$$

where $\odot$ represents element-wise multiplication and division. By fixing the noise basis vectors and optimizing only the signal basis vectors and activation matrix, the algorithm aims to separate the noise and drone signals effectively.

Future Work

Future Work

- Perform unsupervised clustering on the data
- Refine NMF-based approach to enhance SUAS signature and reduce noise
- Modify NMF loss function to handle the frequency and time dimensions differently, prevent bleeding in the time domain
- Change the transform used prior to NMF from STFT -> Orthogonal Mode Decomposition

Overarching Goals

- Develop a robust and automated noise removal algorithm to enhance SUAS detection and classification via machine learning
- Deploy a CNN for classifying the specific SUAS device present in a given audio signal. Eventually, use a rank-3 tensor convolution for spatial data fusion.
- Explore techniques to confound machine learning models trained to detect and classify SUAS devices
- Investigate GANs and the potential to generate synthetic SUAS or noise signals when paired with the NMF approach