# **CSCE 5222: Feature Engineering Project Proposal**

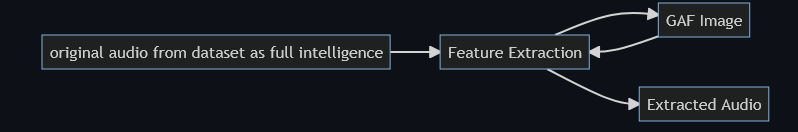
## **Project Title and Team Members**

* Team Members: Andrew Fausak, Andy Fausak, Andre Sharp, Mica Haney
* Project Title: Audio Analysis Using Gramian Angular Field (GAF) and FAST

## **Idea description**

* Summarizing text and speech is complex. Moreover, performing search methods to find similarity between works is difficult. Having a convenient way to produce feature extraction from a given set of text or audio files to represent overall structure and content of the work is desired. Using NLP tools to feed GAF generator for creating a special feature tag representing the overall structure and content of a work is the goal of this effort. This will in itself provide NLP Analysis of audio file and convert to GAF with FAST compatibility to ultimately generate a visual object representing the content of text or audio file.

## **Goals and Objectives:**

* Can GAF be used in a QR-code-like rendering for the identification characterization of audio files (mp3/wav), and creating a unique “fingerprint”?
* The plan is to convert audio
* (as full intelligence)-> GAF image -> Feature Extraction -> audio (as reference feature)  
  
* Identify usability with FAST keypoints for audio files, and ability to detect plagiarism between audio files.
* Identify possibility of converting a GAF back to an audio file.

## 

## **Motivation**

* There is not any research on this yet, as far as we can tell. It seems like a fun project. Converting cata from one domain of information representation to another can provide useful information that would not otherwise be noticable, such as in spectrograms. People have hidden "text" and images in sound files that can be identified only by using specific graphical tools. Using graphical processing techniques to find hidden information in audio is an interesting topic with the potential to improve the current methods of working with audio.

## **Significance**

* This is significant as it is a new way of considering audio files. By doing so more information may be extracted from the audio, which will allow for better handling of audio tasks. If the proposed method does extract information from sound that is not usually or easily picked up by current audio processing techniques, then adding such features and procedures could notably improve results involving audio, particularly in machine learning models. Additionally, if the proposed technique works then this is a potentially novel way of performing speech recognition (pending further research).

## **Literature Survey**

* There was not any other research on this topic we were able to find. There is a lot of research on using a time series with GAF, however, nothing on audio files.

## **Features**

* Initial Audio Files Dataset
* Initial Generated Text

## **GAF Image Features**

* FAST keypoints identification with GAF
* Plagiarism Detection with GAF
* Final Generated Audio File
* Final Generated Text

## **Expected Outcome**

* Identify the usefulness of a GAF for audio files, what is the loss in audio? Does this make a difference in understanding the audio? Does it improve the audio? At the end of the project we expect to be able to answer these questions. We also expect to find that we will be able to extract and/or generate text from audio files where this is relevent to the file.

## **References**

* <https://www.mdpi.com/1263342>
* Johann Faouzi and Hicham Janati. pyts: A python package for time series classification. Journal of Machine Learning Research, 21(46):1−6, 2020.

## **Github Link:**

* <https://github.com/dallasai/csce5222-project>
* <https://github.com/dallasai/csce5222-project/blob/main/Project-Proposal.md>