Report

Github repo - https://github.com/saikrishnareddysanda/SU_PA_3

Setup

- All the models and datasets mentioned in the question are placed in google drive and loaded in collab.
- Installed Miniconda in collab as the pretrained models (SSL_W2V on LA and DF) supported a particular version of torch which can be installed in only python 3.7.

Task 1&2 (Report AUC, EER and analyze)

- The code loads the custom dataset by reading audio files and their corresponding labels from the dataset directory structure.
- It then creates a torch dataset using a custom dataset class called SoundDataset, which handles tasks such as loading audio files, preprocessing, and padding audio data to a fixed length.
- For inference, we load a pre-trained model for SSL Anti-spoofing, Which predicts whether the given audio is Fake or Real.

Results

Receiver Operating Characteristic (ROC) Curve
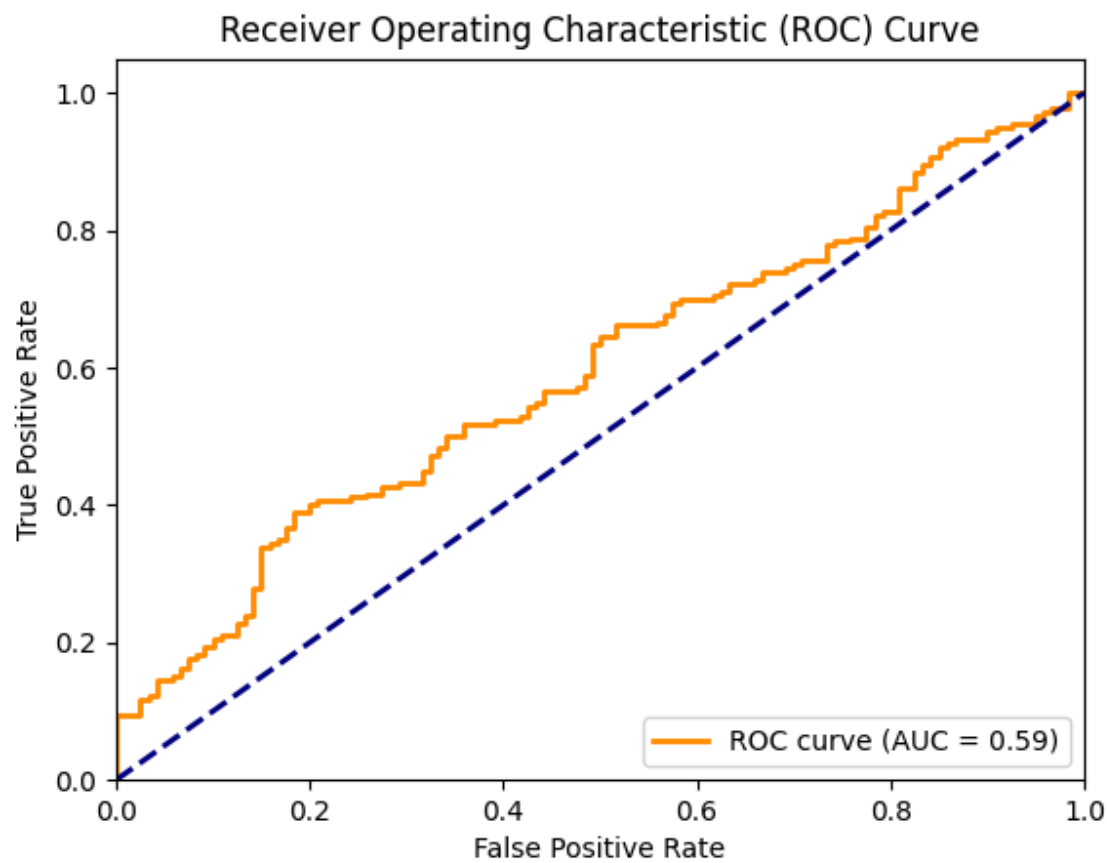
Analysis

AUC	0.59
EER	0.45

- The AUC (Area Under the ROC Curve) value of 0.59 suggests that the model's ability to distinguish between fake and real audio is slightly better than random chance, but still relatively low.
- The EER (Equal Error Rate) of 0.45 indicates that at the chosen operating point, the model has an equal false acceptance rate and false rejection rate. However, this value suggests suboptimal performance, as nearly half of the samples are misclassified at this threshold.
- The relatively low AUC and high EER values imply that the model may not effectively capture the distinguishing features between fake and real audio.

- Further fine tuning the model is required.