Speech Understanding Programming Assignment-3

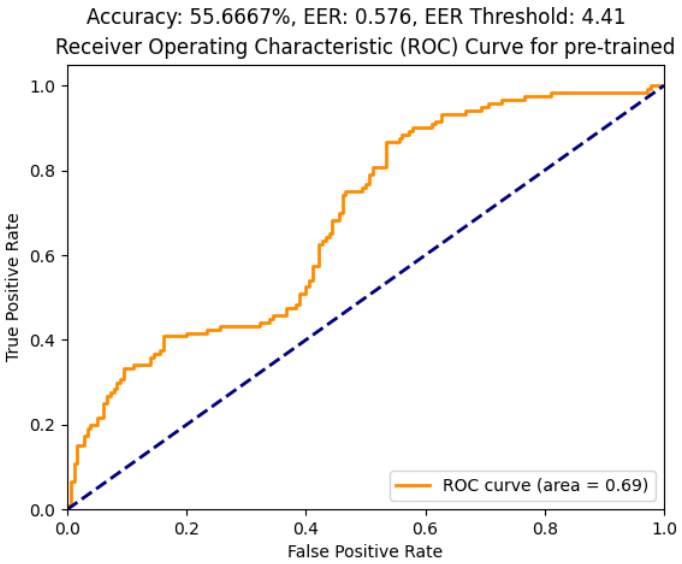
GitHub Repo	https://github.com/parasharharsh16/SU-PA3
Wandb Dashboard	https://wandb.ai/parasharharsh16/SU-Programming-Assignment-3?nw=nwuserparasharharsh16

Experiment Setup	
CPU	AMD Ryzen 7- 3801 MHz
GPU	RTX-4060
RAM	16 GB
Operating System	Linux-5.15.146.1
Framework	PyTorch

Testing Dataset Configuration	
Dataset Name	Custom Dataset
Testing Data Length	300
Sampling Rate	16000

Results on Pre-trained Model

Accuracy	55.66 %	**EER**	0.576	**EER Threshold**	4.41

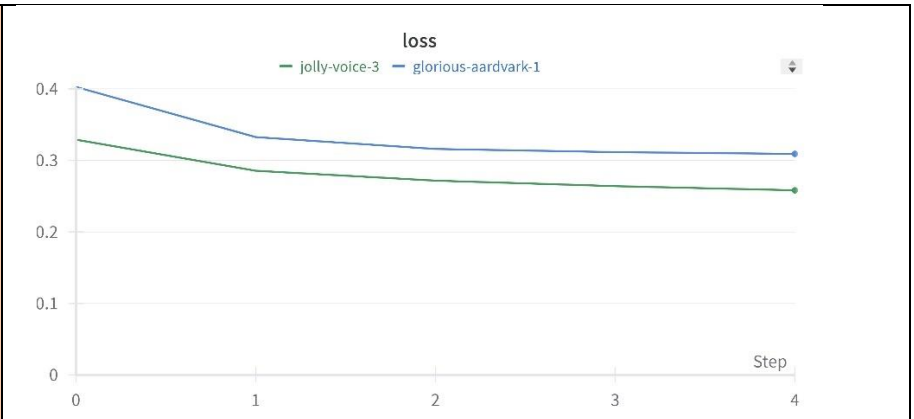
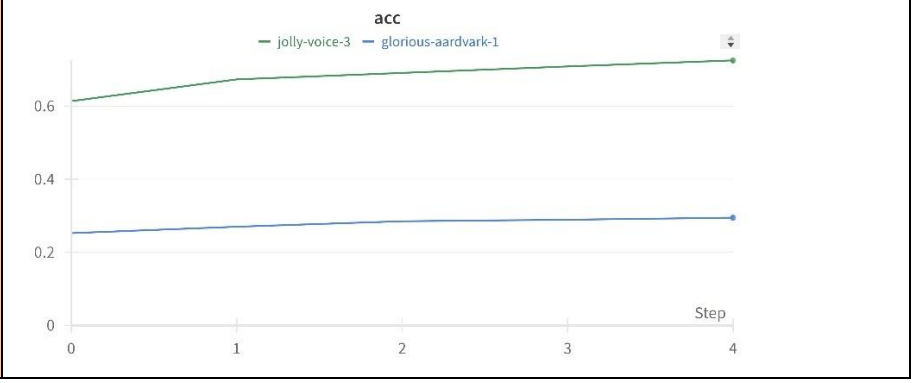


Fine-Tuning of Pre-Trained Model

Fine-Tuning Hyper Parameter	
Dataset Name	FOR Dataset
Training Data Length	13956
Sampling Rate	16000
Evaluation Data Length	2826
Epochs	5
Learning Rate	0.001
Optimizer	Adam
Batch Size	14

Speech Understanding Programming Assignment-3

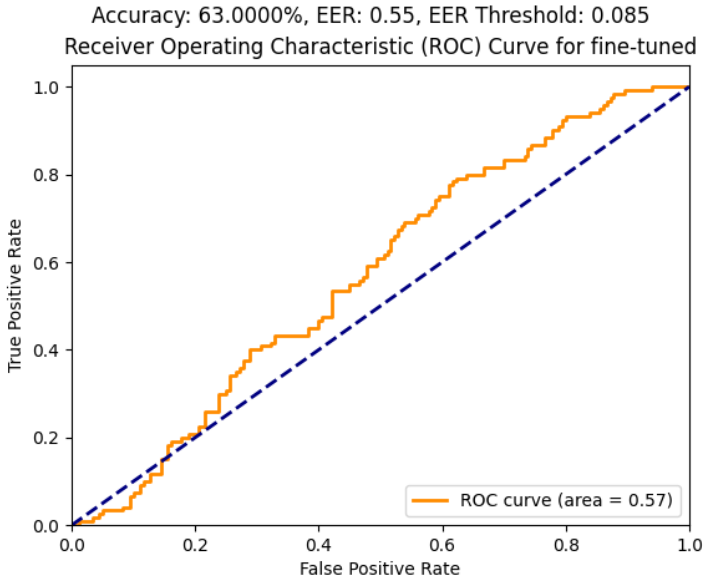
Fine-Tuning Training Matrices

Loss Plot	
Accuracy Plot	

For detailed experiment information, please navigate to [Wandb]

Results on Fine-Tuned Model

Accuracy	63 %	**EER**	0.55	**EER Threshold**	0.085



Speech Understanding Programming Assignment-3

Analysis

The fine-tuned model shows enhanced performance in terms of Accuracy (63% vs. 55.66%) and a reduced Equal Error Rate (0.55 vs. 0.576) when compared to the pre-trained model. But the area of the ROC curve sees a slight decrease from 0.69 to 0.57, which means a slight decline in the ability to differentiate between both classes (Real/Fake).

The decrease in the ROC Curve Area can be caused by a lower number of epochs (i.e. 5), and the fine-tuning of only the final two layers.