Audio Enhancement project

Improving audio quality through denoising: Audio Enhancer

The project enhances input audio using deep learning techniques for denoising. It involves converting audio to a spectrogram and applying mask elimination.

Methodology/Technology

The project employs supervised learning with the WAMR dataset and SpeechBrain's fine-tuned unidirectional LSTM with ReLU activation. It uses MSE loss function on the spectrogram of actual and predicted data.

Results

Final evaluation metric:
Short-Time Objective Intelligibility (STOI)

Score: 0.8

Technological achievement

The project's success in denoising input audio highlights to potential of deep learning in audio signal processing task with implications for various applications

Impact

Improved versions of this can enhance call quality in telecom and optimize performance in voice assistants, offering cost-effective solutions.

Future Goals

Future developments can be done through GANs and LLMs that can take this project all together to the next level.