AI Audio Enhancement Research

Project Status & Future Roadmap

A comprehensive overview of our progress in digital audio processing and AI model development

What is this project about

Can we use AI to upscale lossy audio files to near-lossless audio quality?

Benefits:

- Reduce streaming bandwidth if native to streaming devices

- Optimize disk space for donwloaded audio files

- Allow lossless streaming on Bluetooth devices

- Highly distributed IoT networks for audio processing

Tools

AI

- Model - 1D U-Net archtecture
  - PyTorch for ML
- Hosted through Kaggle

Audio Processing

- Librosa
  - Strong audio processing library for python

What Went Well

Deep Understanding of Digital Audio Processing

- Extensive research into **digital audio compression** techniques

- Comprehensive study of **digital audio conversion**

- Strong foundation in **lossy vs. lossless** audio format differences

AI Model Development Progress

- Successfully initiated the **beginning stages** of our AI model

- Established baseline architecture for audio quality enhancement

Challenges Encountered

Dataset Limitations

- **Lack of diverse datasets** for comprehensive training

- Limited availability of high-quality audio samples

AI Interpolation Training

- Current output of AI model needs refinement on interpolation methodology

- Use original lossless metadata for interpolation parameters

Defintion upon definition

- The deeper you go the more definitions for more things you find

Key Features:

- Audio quality visualization

- Comparative analysis between audio files

- AI upscaling output

Future Plans & Roadmap

AI Model Enhancement

- Create/find data sets

- Focus on improving model's ability to properly interpolate

- Optimize for native implementation on less powerful devices

Streaming Simulation

- Create **mock streaming service** to simulate natural streaming conditions

- Create client that can playback all version of the audio, display graphs and also the metadata of files.

Key Learnings & Technical Insights

Digital Audio Processing Mastery

- Comprehensive understanding of **audio compression algorithms**
- In-depth knowledge of **perceptual audio coding**
- Expertise in **quality assessment methodologies**

Research Methodology

- Systematic approach to AI model development
- Evidence-based evaluation of audio enhancement techniques
- Iterative improvement process for optimal results

Sprint 1 overview

- Week 1: Defined **terminology** - began **audio evaluation**

- Week 2: **More terminology** - Continued evaluation - **audio I/O** - Began **AI model development**

- Week 3: *Even more* **terminology** - Meta-data handling. Differentiating data sets

- Week 4: **Surprise! more terminology** - Graph Construction/Modeling - **AI training**

Metrics

- Total individual Lines of Code (LoC): ~1200

- Number of individual features completed: **2_**

- Number of individual requirements completed: **5_**

- Individual burndown rate (%): **33%__**

Next Steps & Immediate Actions

Sprint 2

1. **Acquire additional datasets** from diverse sources

2. **Refine AI model architecture** to support advanced interpolation

3. Create **Mock streaming** frontend and backend

Technical Implementation Overview

Backend Architecture

- Audio file processing - chunking for streaming

- Python-based processing pipeline

Analysis Components

- Flutter/dart front end implementation

- Audio preprocessing and normalization both on front-end and back-end

- User visualization/listening platform