AI Powered Multimedia Platform

This project delivers a full-stack AI-powered multimedia processing platform with a Flask backend, FFmpeg integration, and both traditional and neural filters for audio and video.

Backend Architecture

A Python Flask server manages all core operations:

- Upload/Delete: Handles single video upload (MP4) and ensures only one file is stored at a time.

- Configure/Apply: Exposes REST endpoints to select and apply filters.

- Processing Pipeline: Uses FFmpeg via subprocess calls for audio extraction, filter chaining, and remuxing.

- Streaming: Serves processed or original video through a /stream endpoint with HTTP range support for smooth playback.

Audio Filter Implementation

Voice Enhancement

- Pre-emphasis: Implements $y[n]=x[n]-\alpha x[n-1]$ (α configurable) to boost high frequencies.

- Butterworth Band-Pass: Applies a 4th-order filter between 800–6000 Hz to isolate speech.

AI Noise Suppression

- Integrates the Intel® OpenVINO™ DenseUNet model for learned noise removal.

- Processes 16 kHz mono audio in fixed-length chunks, maintains state tensors across inferences, and merges cleaned audio back into the video.

Video Filter Implementation

Grayscale Conversion

- Uses FFmpeg's hue=s=0 filter to desaturate frames.

- Encodes with H.264/AAC and -movflags faststart for browser compatibility and seeking.

AI Integration Highlights

- Ambassador-exclusive OpenVINO model for real-time noise suppression.

- Model downloaded via Open Model Zoo tools and compiled on CPU (or GPU) with OpenVINO Runtime.

- Demonstrates stateful, chunked inference with minimal per-chunk latency.

In summary, this project unifies a lightweight web server, robust FFmpeg workflows, traditional DSP, and deep-learning–based noise suppression into a coherent multimedia processing service. The project can be found with all source codes and documentations here : https://github.com/MahadiAlif/OpenVINO_multimedia_processing