

Audio-Visual Emotion Detection using YOLO & Frequency Analysis -

3rd place winner in Ingenius 2024 Hackathon

Overview

This project combines **audio** and **visual** analysis to detect emotions from video files. By using **YOLOv8** for object detection and **Librosa** for audio frequency analysis, the system classifies emotions based on both visual cues (e.g., detecting dogs and movement) and audio patterns (e.g., howling or other frequency bands).

Key Features

- **Object Detection:** Detects dogs and other relevant objects in video frames using YOLOv8.
- **Audio Frequency Analysis:** Analyzes the audio for dominant frequencies (such as howling sounds) to classify emotions.
- **Emotion Classification:** Determines emotions like sadness, happiness, relaxation, and anger based on the combined visual and audio cues.

Technologies Used

- **YOLOv8:** For real-time object detection to identify moving objects and animals.
- **Librosa:** For extracting and analyzing frequencies from the audio track in the video.
- **MoviePy:** For extracting audio from video files.
- **OpenCV:** For real-time video frame processing.
- **NumPy:** For handling numerical data during audio analysis.

Prerequisites

Before running this project, make sure to install the required libraries.

Python Libraries:

- `ultralytics`
- `librosa`
- `moviepy`
- `opencv-python`
- `numpy`
- `matplotlib`

You can install them using `pip`:

```
pip install ultralytics librosa moviepy opencv-python numpy matplotlib
```

Contributors:

- Ingenius hackathon team:
 - Sanath
 - Neranjana
 - Souriesh