A3: Real-time Activity Sensing

Machine Learning and Sensing | Spring 2025

Assignment Overview

Goal:

- Deploy a real-time inference system
- Compare ML vs DL in real-time

Submit:

- Code (.ipynb or .py)
- requirements.txt
- Two demo videos (ML + DL model real-time inference)

A3.1– Real-Time Acoustic Inference (20 pts)

Objective: Deploy real-time system using best model from A1, and DL model from A2 (or state-of-the-art DL model).

Steps:

- Mic audio → sliding window → inference → live display
- Record 2-3 minutes for each video including the 5 main activities (randomly).
- Display:
  - Waveform (real-time)
  - Predictions + confidence (ML + DL)
  - Inference time (ms)

Reference:

- Mic input: PyAudio, python-sounddevice
- GUI: Tkinter, PyQt
- Ubiquitous:
  https://github.com/FIGLAB/ubicoustics/blob/master/example_liveprediction_simple.py
  https://github.com/FIGLAB/ubicoustics/blob/master/example_liveprediction_detail.py

A3.1– Real-Time Acoustic Inference (20 pts)

Other Models:

1. Wav2Vec 2.0
2. AST (Audio Spectrogram Transformer)
3. AudioMAE

Pre-trained models

Model	Finetuning split	Dataset	Model
Wav2Vec 2.0 Base	No finetuning	Librispeech	download
Wav2Vec 2.0 Base	10 minutes	Librispeech	download
Wav2Vec 2.0 Base	100 hours	Librispeech	download
Wav2Vec 2.0 Base	960 hours	Librispeech	download
Wav2Vec 2.0 Large	No finetuning	Librispeech	download
Wav2Vec 2.0 Large	10 minutes	Librispeech	download
Wav2Vec 2.0 Large	100 hours	Librispeech	download
Wav2Vec 2.0 Large	960 hours	Librispeech	download
Wav2Vec 2.0 Large (LV-60)*	No finetuning	Libri-Light	download

Write-up Summary

A3.1 Report:

- Describe the pipeline: audio capture → processing → inference → display.
- How did you handle buffering and overlapping windows?
- What were the typical inference times for each model?
- Was one model noticeably faster or more stable?
- What kind of interface did you build? How did you visualize predictions and confidence?
- How well did the system respond to different environments (quiet, noisy, echo, etc.)?

Submission Checklist

Code: Notebook or script + requirements.txt

video_ml.mp4: Real-time demo with ML model (A1/A2)

video_dl.mp4: Real-time demo with DL model (A3.1)