Gesture Recognition Project Overview

Phase 1: Technical Plan Presentation

Authors: Bora ILCI & Kaan Emre KARA

Group 10 | Instructor: Valeriya Khan

Date: May 12, 2025

Background & Objectives

- Design, implement, and evaluate a neural-network model for hand gesture recognition.
  - Compare model performance against established methods.
  - Document comprehensive methodology, datasets, and results.
  - Deliver final report and presentation with comparative analysis.

Algorithm Selection & Rationale

1. CNN + Temporal Pooling

- Simplifies temporal modeling by aggregating features without complex recurrence.
- Leverages frame-wise feature learning for clear spatial representation.

2. 3D-CNN (C3D)

- Learns joint spatiotemporal features directly from frame volumes.
- Provides a performance benchmark for more advanced models.

3. Transfer Learning + LSTM (Optional)

- Utilizes pretrained visual representations to reduce training time.
- LSTM refines temporal dependencies for improved sequence modeling.

Dataset Selection & Rationale

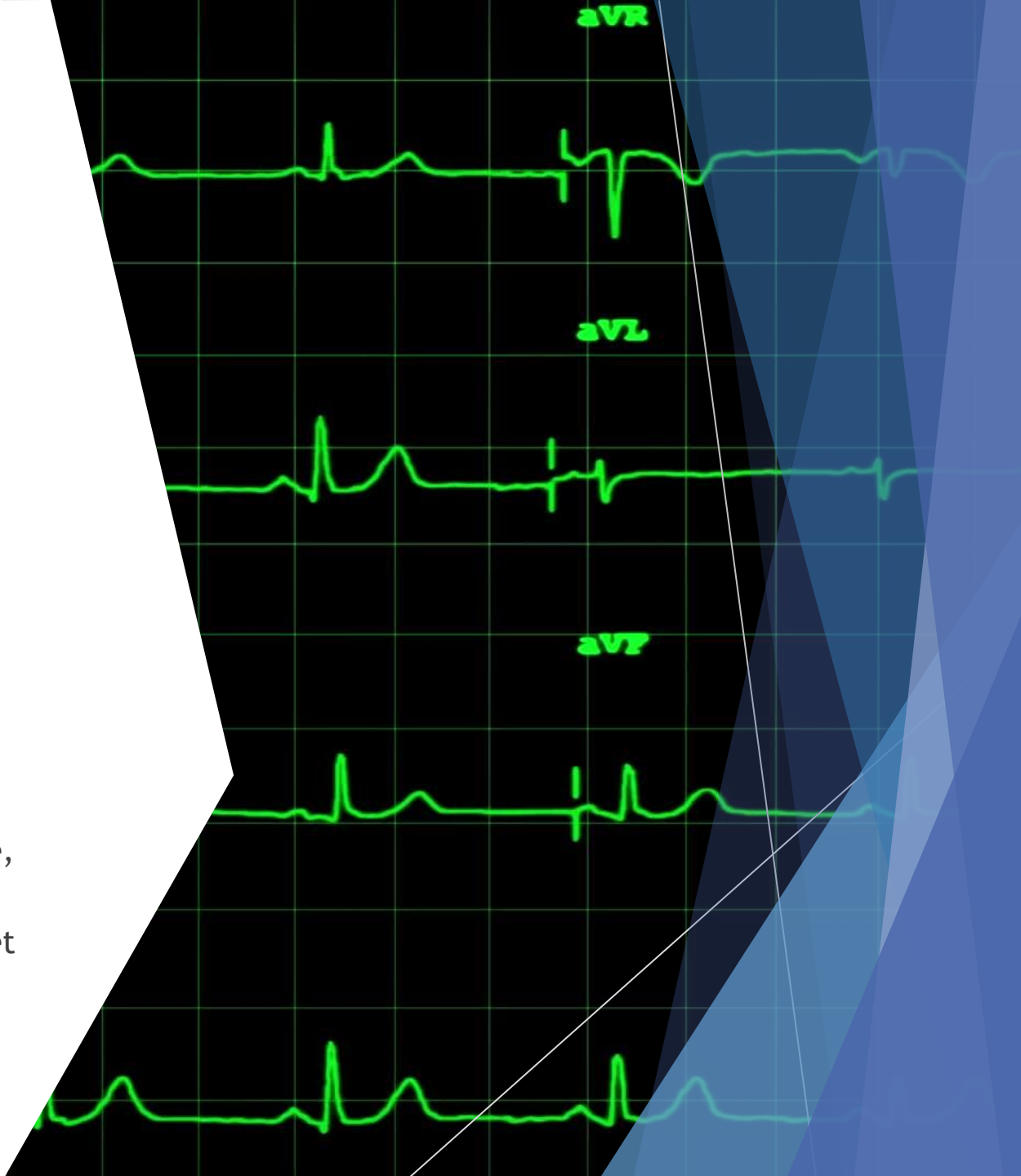
- Primary: LeapGestRecog
  - Provides balanced classes across 10 gestures ensuring fair evaluation.
  - Frame-based format eliminates video decoding complexity.
  - Moderate dataset size (~2,000 sequences) enables rapid iteration on standard hardware.
- Secondary (Future Phases):
  - SHREC dataset adds depth modality for richer feature exploration.
  - Custom webcam data tests model robustness to real-world variance.

Libraries & Technical Details

- Python 3.8+, TensorFlow 2.x (Keras API)
  - OpenCV, NumPy & pandas for data handling
  - scikit-learn for splitting & metrics
  - Matplotlib & Seaborn for visualization
  - TensorBoard for monitoring
  - flake8 for code style compliance

Experimental Plan

- 1. Preprocess & split dataset (70/15/15 stratified)
- 2. Train CNN + Temporal Pooling baseline
- 3. Train C3D baseline for spatiotemporal comparison
- 4. Conduct hyperparameter sweep (LR, batch size, pooling method)
- 5. Evaluate on test set and compare metrics



Visualization & Metrics

- Learning curves: loss & accuracy over epochs
- Normalized confusion matrix heatmap
- Bar charts of per-class precision, recall, F1
- Sample frame sequences with predicted labels
- Efficiency table: parameters, FLOPs, inference time

Implementation & Reproducibility