CV701 Assignment 4 – Task 1 Report

Mashrafi Monon

November 29, 2025

1 Introduction

This report summarizes the implementation and evaluation of the facial keypoint detection model for Task 1. The dataset consists of celebrity faces with 68 annotated landmarks. Our objective is to regress the landmark coordinates and infer the facial emotion (negative/neutral/positive) from the detected keypoints.

2 Methodology

2.1 Data Pipeline

Describe the dataset splits, custom transforms (resize, random flip, normalization), and dataloader configuration.

2.2 Model

Summarize the ResNet18-based regression head, training hyperparameters, loss (Smooth L1), optimizer (AdamW), and scheduler (Cosine Annealing).

2.3 Emotion Heuristic

Explain the rule-based classifier operating on predicted landmarks (mouth width/height, smile curvature normalized by inter-ocular distance).

3 Experiments

3.1 Training Configuration

List key hyperparameters (batch size, epochs, learning rate, device).

3.2 Validation Metrics

Present MAE/RMSE/NME curves, mention best epoch, and reference figures/tables.

3.3 Test Results

Summarize the metrics stored in `artifacts/task1_hpc/metrics.json` and emotion distribution from the test predictions CSV.

4 Qualitative Analysis

Include visualizations or descriptions of correctly predicted landmarks and common failure cases (e.g., occlusions, profile faces).

5 Conclusion and Future Work

Wrap up Task 1 highlights and outline next steps for Task 2 (deployment + optimization).

References