Real-time facial expression recognition in videos

Lewis Bentley

2655220B

Proposal

Motivation

Facial expression recognition has applications in diverse fields, including human-computer interaction, mental health monitoring, and security systems. Recognizing expressions in real time from videos enhances the ability to interpret human emotions in dynamic environments. This project aims to improve the latency and efficiency of real-time emotion recognition systems, which are critical for applications where quick and reliable feedback is required.

Aims

This project will develop a system that can accurately detect and classify facial expressions in real-time from video streams. The system will leverage machine learning algorithms and pre-trained models to process video frames and identify emotions with minimal latency. Success will be measured by the system's accuracy, speed, and reliability under varied conditions.

Progress

- *Preprocessing pipeline for converting video frames to grayscale images (48x48) implemented.*
- *Initial neural networks (NNs) for emotion recognition completed.*
- *RAVDESS dataset prepared and loaded for training and testing.*
- *Basic CNN model implemented and trained on the dataset; initial accuracy metrics recorded.*
- *Basic emotion recognition functionality tested with static images.*
- Tested several pre-trained models on the static images and recorded accuracy metrics.

Problems and risks

Problems

- Model overfitting to data after very little iterations.

- Limited computational power for training more complex models, leading to longer training times.

Risks

- Overfitting on the limited dataset due to small size, lack of diversity and lab like scenarios.
- High latency when processing high-resolution video frames.
- Real-time inference accuracy may drop in low-light, high-noise scenarios or face occlusion scenarios.

Plan

Week 1: Test pre-trained models on datasets.

- **Tasks**:
    - Evaluate model performance on static images from datasets
- **Deliverable**: Performance metrics and insights on pre-trained model generalizability.

Week 2-3: Implement video-to-frame processing pipeline.

- **Tasks**:
    - Integrate pre-trained models for frame-level emotion recognition.
    - Begin testing the system with static videos to ensure proper functionality.
- **Deliverable**: A pipeline capable of processing video inputs frame by frame and classifying emotions.

Week 4: Build a context window for video-level emotion recognition.

- **Tasks**:
    - Implement a temporal context window to analyse and aggregate emotions from the last *X* frames.
    - Test and refine strategies for aggregating frame-level predictions (e.g., majority voting, weighted averages).
- **Deliverable**: A functional context window for video-level emotion recognition.

Week 5: Optimize real-time video processing.

- **Tasks**:
    - Adapt the pipeline for real-time video input with minimal latency.

    - Perform preliminary real-time testing to balance speed and accuracy.
- **Deliverable**: Optimized system capable of real-time video processing and emotion recognition.

Week 6: Conduct initial system testing and fine-tuning.

- **Tasks**:
    - Evaluate performance under controlled conditions, such as varying lighting and noise.
    - Refine the system based on test results and address any observed bottlenecks.
- **Deliverable**: Improved system with reliable performance under test conditions.

Week 7: Conduct comprehensive testing under diverse conditions.

- **Tasks**:
    - Test the system on videos with diverse scenarios, including low-light and occlusions.
    - Measure the impact of context window size on accuracy and latency.
- **Deliverable**: Detailed performance metrics and refined context window implementation.

Week 8: User testing and feedback collection.

- **Tasks**:
    - Conduct user testing to evaluate the system's usability and accuracy in real-world settings.
- **Deliverable**: Summary of user feedback and prioritized improvements list.

Week 9: Final implementation and polishing.

- **Tasks**:
    - Address feedback and improve system robustness and user experience.
- **Deliverable**: Fully polished, real-time emotion recognition system.

Week 10: Write up and prepare the final report.

- **Tasks**:
    - Submit a complete draft for supervisor review.
- **Deliverable**: Final draft of the report ready for feedback and revisions.