Extending Lightweight Driver FER

A Novel Video-Based Approach

Aditya Chaudhary

BTU Cottbus

May 5, 2025

Introduction: Uddin (2025)

- Paper: “A Novel Lightweight Deep Learning Approach for Drivers’ Facial Expression Detection”
- Model: Dual Attention Lightweight Deep Learning (DALDL)
- Key Features:
  - SqueezeNext + Dual Attention Convolution (DAC)
  - Accuracy: 91.5% (CK+), 88.9% (KMU-FED)
  - Lightweight: 0.75M parameters, 3.9 ms inference
- Goal: Real-time emotion detection for ADAS

Motivation for Extension

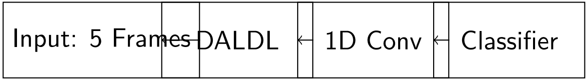
- **Limitation**: DALDL processes static images
  - Misses temporal dynamics of driver emotions
  - Less effective for continuous monitoring
- **Motivation**:
  - Video-based FER captures expression transitions
  - Enhances ADAS reliability in real-world driving
- **Proposed Extension**: Add temporal aggregation for video sequences

Proposed Extension: Video-Based FER

- **Objective**: Enable DALDL to process short video sequences (3-5 frames)
- **Approach**:
  - Retain DALDL (SqueezeNext + DAC) for spatial features
  - Add lightweight 1D convolutional layer for temporal aggregation
  - Classify 7 emotions: happy, surprise, anger, sad, fear, disgust, neutral
- **Dataset**: KMU-FED (simulate sequences) or AFEW

Methodology

- **Data Preparation**:
  - Group 3-5 frames into sequences
  - Preprocess: resize, normalize, histogram equalization
- **Model Modification**:
  - DALDL per frame: [512, 1, 1] features
  - 1D Conv: aggregate temporal features
  - Classifier: 7-class output
- **Training**: 80-20 split, 55 epochs, Adam



Expected Outcomes

- **Improved Accuracy**: 1-3% increase over static DALDL
  - CK+: $>$91.5%, KMU-FED: $>$88.9%
- **Maintained Efficiency**:
  - Inference time: $<$5 ms per sequence
  - Model size: 3.3 MB
- **Enhanced Robustness**: Better detection of expression transitions
- **ADAS Impact**: Reliable real-time driver monitoring

Future Work

- Validate with video datasets (e.g., AFEW)
- Test in real driving environments
- Optimize for embedded platforms (e.g., NVIDIA Jetson)
- Address privacy via data anonymization
- Explore multimodal inputs (e.g., gaze, physiological signals)

Conclusion

- Uddin (2025) provides a lightweight, efficient FER model
- Extension enables video-based FER, enhancing ADAS
- Small, impactful addition with minimal overhead
- Positions research for real-world driver safety applications

Reference: Uddin, J. (2025). *Designs*, 9(2), 45.