



Extending Lightweight Driver FER: A Video-Based Approach

Aditya Chaudhary, Saadi Ibtissam, Douglas W. Cunningham

Department of Graphical Systems, Univ. BTU Cottbus-Senftenberg, Cottbus, Germany

1. INTRODUCTION

Uddin (2025) [1] presents the DALDL model, optimized for efficient facial emotion recognition (FER) in real-time ADAS. This research extends DALDL by integrating temporal aggregation for video-based FER, enhancing dynamic emotion detection.

- Leverages SqueezeNext with DAC block.
- Compact: 0.75M params, 3.9ms inference.
- Handles 3-5 frame video sequences.
- Boosts ADAS reliability in real-world driving.



Static FER Limited Context



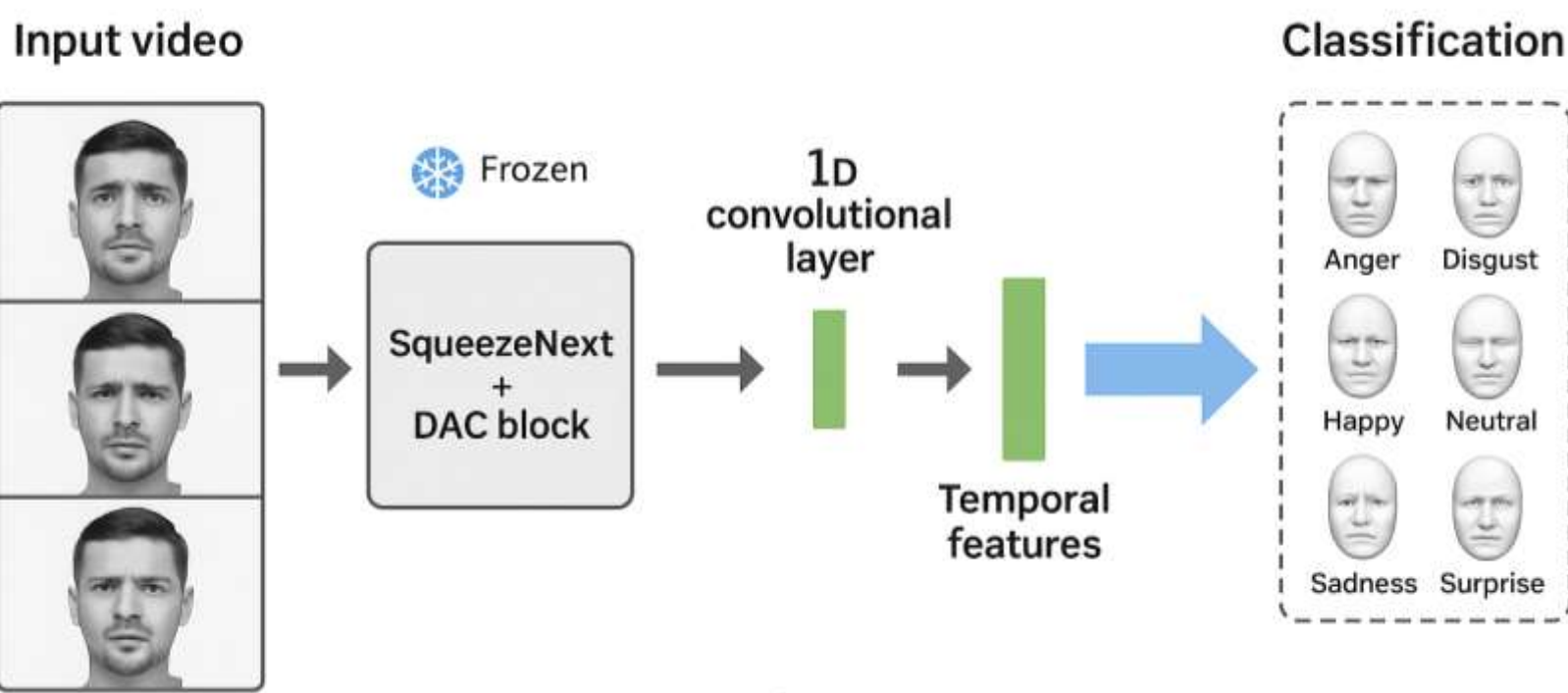
Video FER – Temporal Dynamics Captured



2. RESEARCH QUESTION

DALDL processes static images, missing temporal dynamics of driver emotions. Can video-based FER with temporal aggregation enhance ADAS reliability by capturing expression transitions in real-world driving scenarios?

3. METHODOLOGY



- **Objective:** Process 3-5 frame videos with DALDL.
- **Approach:** DALDL for spatial features, 1D Conv for temporal aggregation, classify 7 emotions
- **Why DALDL?** Lightweight, accurate, ADAS-fit.
- **Dataset:** KMU-FED (drivers' FEs, NIR camera).
- **Data Prep:** Group 3-5 frames, preprocess
- **Training:** 80-20 split, 55 epochs, Adam.