Drowsiness Detectors: Lightweight Models

By Cash, Chung, Luka, Matthew

ECE 228: Prof.

i think a table of model size and performance would be good (like they have in the source paper

Overview:

Detecting driver drowsiness is crucial for preventing road accidents. In this project, we propose a lightweight deep learning model that significantly reduces computational and memory demands without sacrificing detection accuracy, enabling practical real-time drowsiness monitoring in resource-constrained environments.

- Four common approaches to measuring drowsiness:
  monitoring steering wheel angle and lane deviation
  analyzing bio-signals such as EEG
  evaluating facial features and head position through image data
  combination of these methods
- This project focuses on image-based signals

Project Goal:

Though the above methods can achieve high accuracy in detecting drowsiness, they often require significant computational resources, limiting practical deployment in constrained environments. We propose a lightweight deep learning approach for drowsiness detection that reduces computational and memory costs while maintaining high accuracy.

Primary Paper

Florez et. al (2023), presents a CNN-based approach for driver drowsiness detection by analyzing regions of interest of facial features extracted from processed frames of video. For training data, the NITYMED database is used, which contains videos of drivers exhibiting drowsiness symptoms. Pre-processing is conducted to extract specific facial points and select a region of interest around the eyes, followed by transfer learning of three CNN architectures (InceptionV3, VGG16, ResNet50V2).

Accuracy on the testing set for all three CNN architectures reach 99%, file sizes varied between 100 MB - 500 MB, with inference time varying between 70-140ms. This motivates our project goal to implement a lightweight version of their framework in order to deploy in constrained, lower resource environments.

Data & Preprocessing

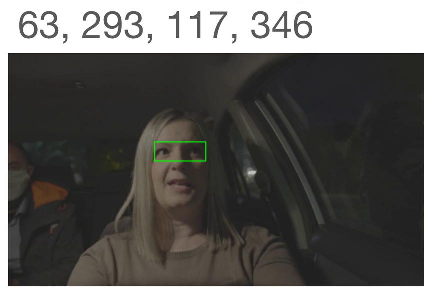
Following a similar data preprocessing setup as the source paper:
Input: 1920x1080p videos, 25 FPS.

Per-frame processing:
Uses lightweight, pre-trained MediaPipe FaceMesh model (15MB) to detect the facial image in the frame along with 468 facial landmarks.

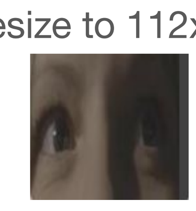
Applies region of interest (ROI) extraction algorithm from the paper for stable eye cropping.
Resizes output to 112×112 px, suitable for image CNN input.
Output: (video length x 25 frames) of (112 x 112 x 3) images.
Images are then manually sorted into "Drowsy" or "not-drowsy".
Data sourced from NITYMED (Night-Time Yawning-Microsleep Eyeblink Driver Distraction) Dataset: https://datasets.esdalab.ece.uop.gr/

original 1920x1080p image

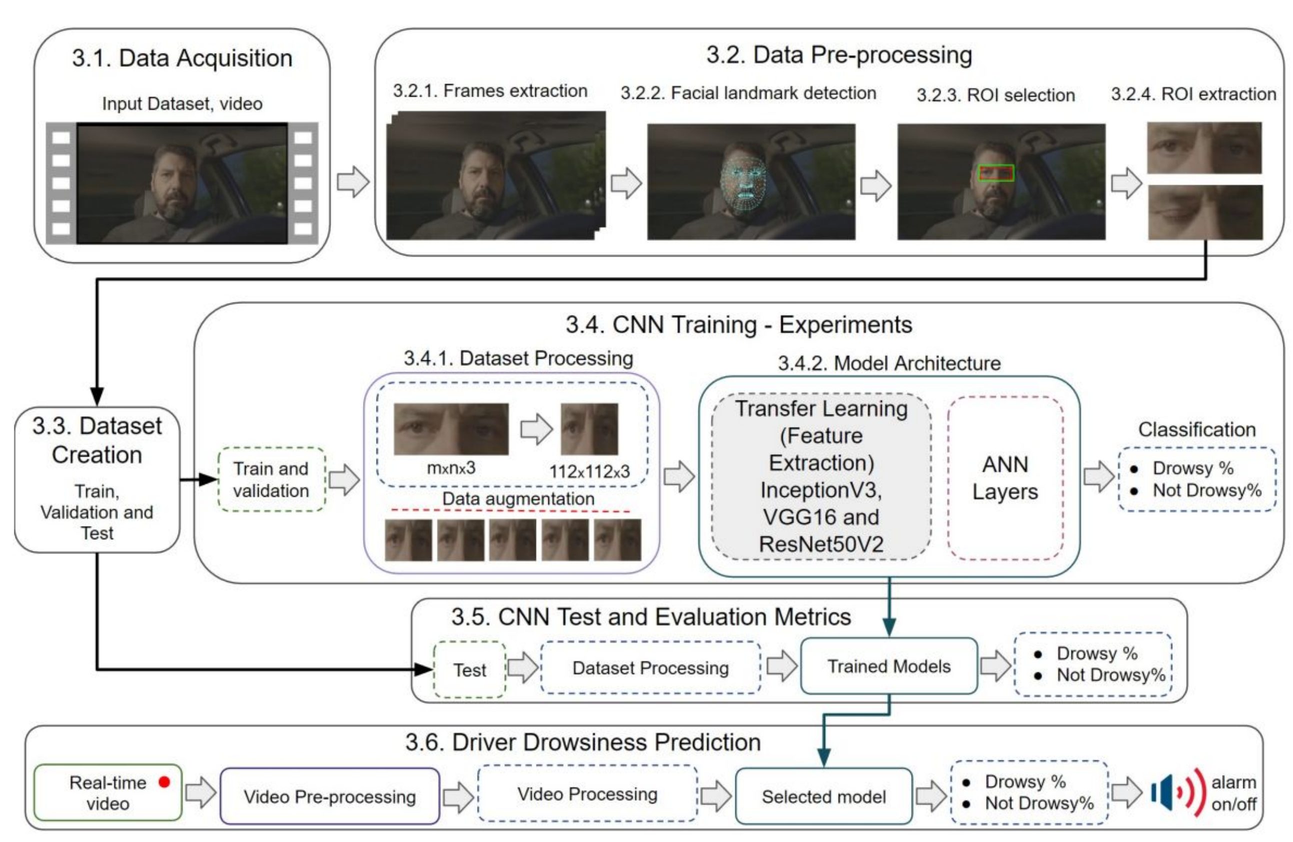
ROIS extracted: points 63, 293, 117, 346



Reshape image around ROI and resize to 112x112



Development



Florez, R., Palomino-Quispe, F., Coaquira-Castillo, R. J., Herrera-Levano, J. C., Paixão, T., & Alvarez, A. B. (2023). A cnn-based approach for driver drowsiness detection by real-time eye state identification. *Applied Sciences*, *13*(13), 7849.

Demonstration / Deployment

Live video from webcam is captured and processed in a three step pipeline:

1. Face Detection & Landmark Extraction: MediaPipe Facemesh
2. Pre-Processing:
3. Drowsiness Detection: MobileNetV3-S

The fir MediaPipe Facemesh ROI extraction, which is then forwarded through our retrained CNN to classify drowsiness. (luka can clarify how this works exactly as I'm not too sure). We can then provide a live overlay on the video of the driver's drowsiness probability.

Experimental Changes

Dataset Size Discrepancy:

- Original Paper: 6,860 files
- Our Implementation: 23,809 files

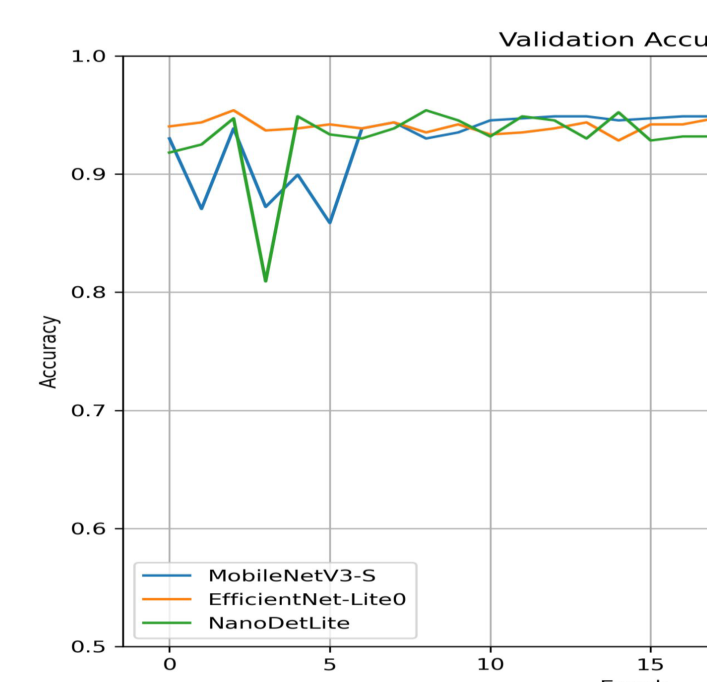
Conclusion

Reference

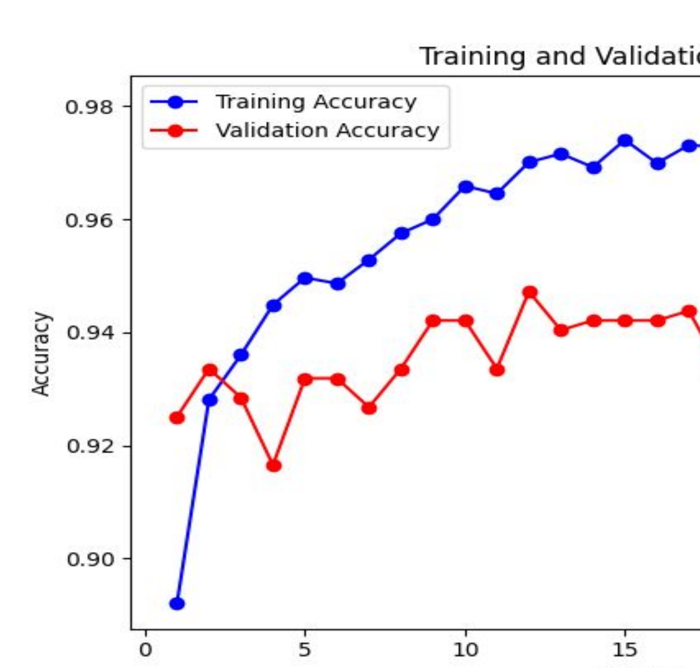
Data Collection/Graphical Analysis

EfficientNet-Lite0	MobileNetV3-S	NanoDetLite
4.7M	2.54M	0.95M
13.1MB	5.9MB	4.4MB

VGG16	InceptionV3	ResNet50V2
19.4M	29.9M	56.3M
		98MB



Inception V3



ResNet50V2