GLOBAL RESEARCH IMMERSION PROGRAM FOR YOUNG SCIENTISTS

Wageningen University & Research | Zhejiang University

Mike Geerts | supervision: Prof.dr.ir. Yu & Prof.dr.ir. Der Horng-Lee | mike.geerts@wur.nl

Detection of DRIVER DROWSINESS by Deep Learning

- BACKGROUND -

(WHO, 2024; Prajapat, 2022)

1.2 MLN TRAFFIC DEATHS a year, worldwide

Due to DRIVER DROWSINESS **20%**

There is the need for models that accurately detects drowsiness. Below, a cross-entropy loss function shows the discrepancy between actual labels (like 'drowsy' and 'non drowsy') and predicted probabilities. I aim to minimize the discrepancy, or so called 'loss'. The smaller the loss, the larger the model's accuracy. To detect drowsiness.

$$L(y,\hat{y}) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1-\ y_i)\log(1-\ \hat{y}_i)\,]$$

Also, I will show how Large Language Models (LLM) can handle data of different types (like eye tracking data, video recordings showing hand movements and lidar data showing road conditions).

- OBJECTIVES -

1) Improve accurate drowsiness detection
also in case of different types of data

2) Analyse performance of models
ResNet50 and MobileNetV2

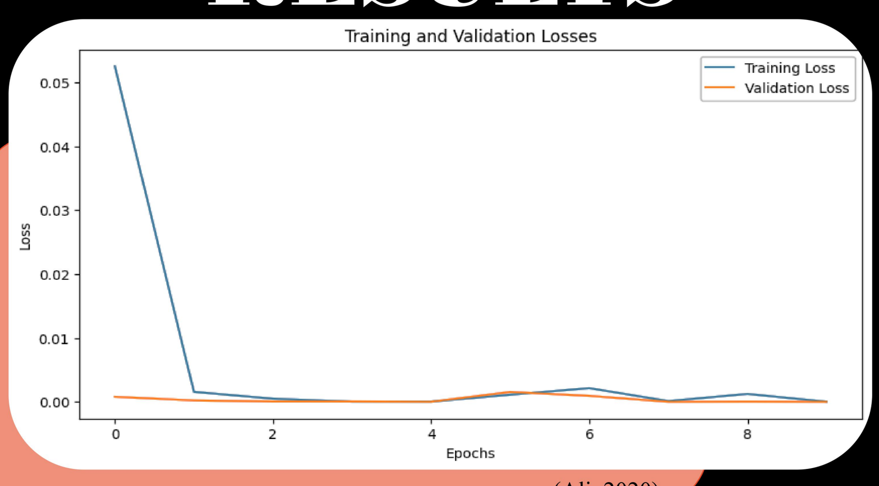
3) Correlate various data types
for comprehensive understanding of drowsiness

- METHODS -

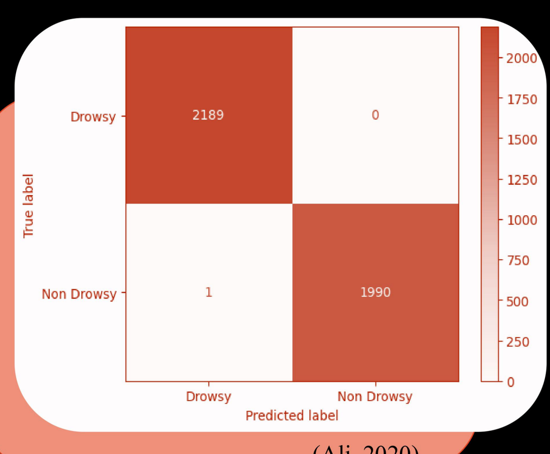
I will analyze the application of ResNet50 – in the research by Ali (2020) – and MobileNetV2 – Sayed (2024) – to compare their performance in drowsiness detection. I focus on these models due to their contrasting characteristics. ResNet50 and MobileNetV2 are both convolutional neural networks (CNNs) designed to analyze patterns in images. ResNet50 is known for achieving high accuracy. MobileNetV2 offers efficiency and fast processing. Comparing highlights which level of accuracy is achieved in each model.

System workflow	1 - Installation and importing Libraries	2 - Data Exploration	3 - Data preparation	4 - Model architecture	5 -Training performance
Ali (2020)	PyTorch and torchvision	Split: training, validation, and test sets Visualize samples labelled 'drowsy' or 'non drowsy'	Set up a way to load and transform images Apply data augmentation techniques (like resizing and normalization)	Pre-trained **ResNet50 model** Base layers unchanged; top layers are adapted to focus on drowsiness detection.	Train over 10 epochs **Validation Accuracy: 100.00% |Validation Loss: 0.0000** **Test Accuracy: 99.98% | Test Loss: 0.0009**
Sayed (2024)	TensorFlow and Keras	Split: training, validation, and test sets Use of ImageDataGenerator	ImageDataGenerator to load pictures and rescale pixel values Visualize sample images to check class distribution	Pre-trained **MobileNetV2 model** Base layers unchanged; top layers are adapted to focus on drowsiness detection.	Train over 10 epochs **Training Accuracy: 99.97%** **Training Loss: 0.0024** **Test Accuracy: 99.86% | Test Loss: 0.0056**

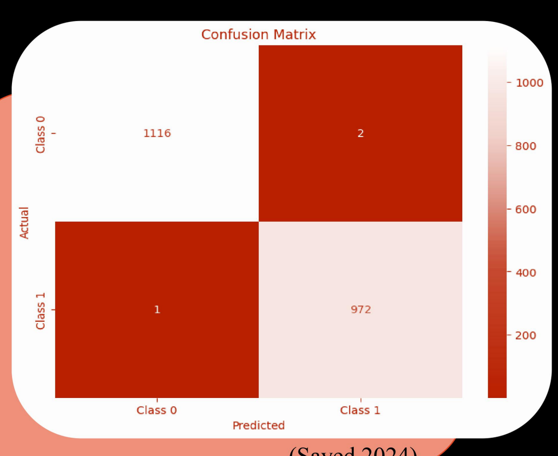
- RESULTS -



(Ali, 2020)



(Ali, 2020)



(Sayed 2024)

- CONCLUSION -

1) High accuracy / low 'loss' in both models
This is shown in the graph and confusion matrices on the left.

2) Lower loss in ResNet50 compared to MobileNetV2
This makes ResNet50 more reliable in drowsiness detection; MobileNetV2 requires fewer computational resources

- DISCUSSION -

I compare only two models. Not all deep learning models have been applied to driver drowsiness detection, so further research is recommended to understand the advantages and disadvantages of various models. Ali's results present validation accuracy, while Sayed's show training accuracy. Validation accuracy measures a model's performance on unseen data, indicating its potential to generalize. In contrast, training accuracy reflects performance on data the model was trained on. Therefore, MobileNetV2's ability to generalize may be overestimated. I also recommend further research:

1) use more extensive datasets and various types of data. Ali and Sayed focus on facial expressions in video data, but models should also analyze eye tracking, heart rate, video recordings of hand movements, meteorological data for weather prediction, and lidar data to measure road conditions. Gemini (Team Google, 2024) is useful for handling multiple data types and correlating them.
2) use LLMs like llama 3.1 for automatic data labeling and pattern recognition. The input could be a list with labels the model chooses from or an open question (prompt) to ask the model to give a label
3) use more labels like "semi-drowsy". This can lead to a more holistic view on driver states.

Ali, Y. (2020). *Driver drowsiness detection || PyTorch*. Driver Drowsiness Dataset (DDD). Retrieved on August 3, 2024 from https://www.kaggle.com/code/youssefyasserali/driver-drowsiness-detection-pytorch/notebook#6.-Evaluating-the-Model.

Gemini Team, Google. (2024). *Gemini: A family of highly capable multimodal models*. Retrieved on August 4 from https://assets.bwbx.io/documents/users/iqjWHBFdfxIU/r7G7RrtT6rnM/v0.

Prajapat, S. (2022). *Driver drowsiness using Keras*. Drowsiness_dataset, Driver drowsiness, prediction images. Retrieved on August 1, 2024 from https://www.kaggle.com/code/saurabhprajapat/driver-drowsiness-using-keras/notebook#Model.

Sayed, E. (2024). *Driver drowsiness detection (CNN ||MobileNetV2)*. Driver Drowsiness Dataset (DDD). Retrieved on August 3, 2024 from https://www.kaggle.com/code/esraameslamsayed/driver-drowsiness-detection-cnn-mobilenetv2/notebook#%F0%9F%94%97Evaluation-time.

World Health Organization. (2023). *Road traffic injuries*. Retrieved on August 1, 2024 from https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.