IN CAB DRIVER DISTRACTION DETECTION

Nishanth Chennagiri Keerthi 24212022
Ashish Mohamed Usman 24201306

ACM40960 Projects in Maths Modelling, MSc Data and Computational Science, University College Dublin

Abstract

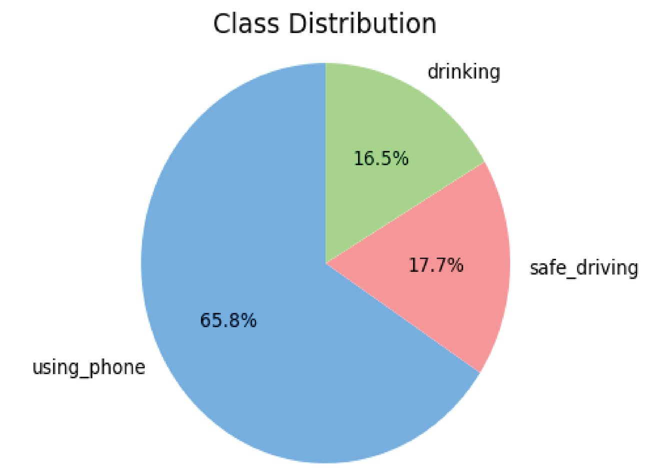
This project addresses the growing issue of **in-cabin driver distractions** by developing a deep learning-based image classification system to identify **unsafe driving behaviours**. Using a curated subset of the **State Farm Distracted Driver Detection dataset**, we focused on three key classes: **safe driving**, **using phone**, and **drinking**. A baseline Convolutional Neural Network was implemented and enhanced with a **custom CNN model** using dropout regularization and class balancing, achieving over **98% test accuracy** with balanced precision and recall. To extend its applicability, we built a **video-to-image inference pipeline** that flags and annotates distracted behaviour in uploaded driving videos, highlighting the potential of computer vision for improving road safety.

Motivation

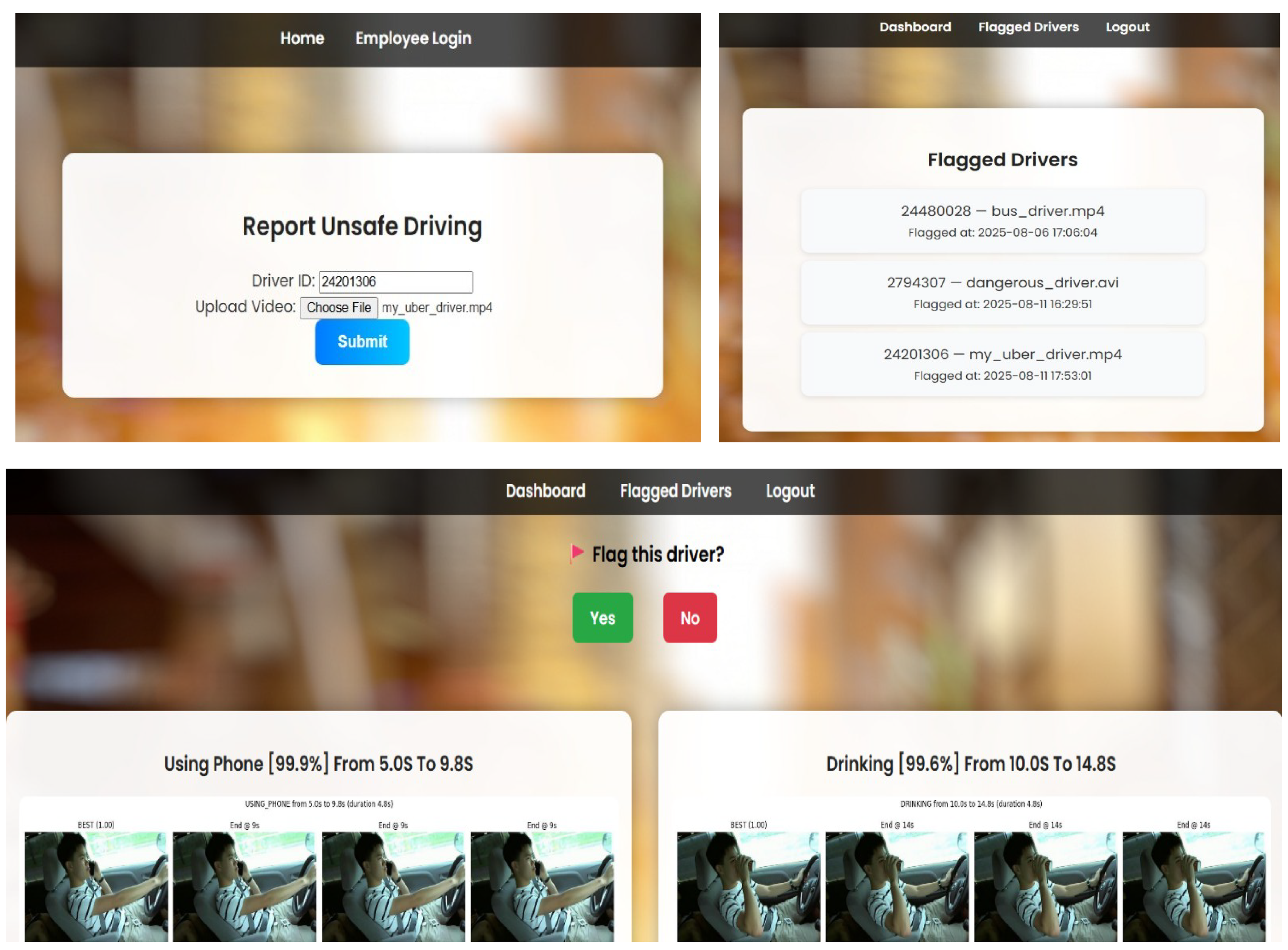
Distracted driving is a leading cause of **road accidents** worldwide, responsible for millions of crashes and preventable deaths each year. Actions like **using a phone**, **drinking**, or adjusting in-car controls reduce driver focus and raise accident risk. The **World Health Organization** reports over **1.19 million** road traffic deaths annually, with distraction as a major factor. In Ireland, the **Road Safety Authority** links **20–30%** of collisions to distraction, causing over **1,400** serious or fatal crashes yearly. Our project addresses this issue by using deep learning and computer vision to analyse in-cabin driving videos, automatically flagging distracted behaviour to help reviewers identify risks and take preventive action.

Dataset Used

Our project uses the **State Farm Distracted Driver Detection** dataset, focusing on three key safety behaviours: **safe driving**, **using a phone** (merged from multiple phone-related categories), and **drinking**. We worked with about **14,000 images**, splitting them into training,validation, and testing sets, each with metadata for quick access. The classes are imbalanced, with using a phone being the most common and drinking the least, which we kept in mind during model development.

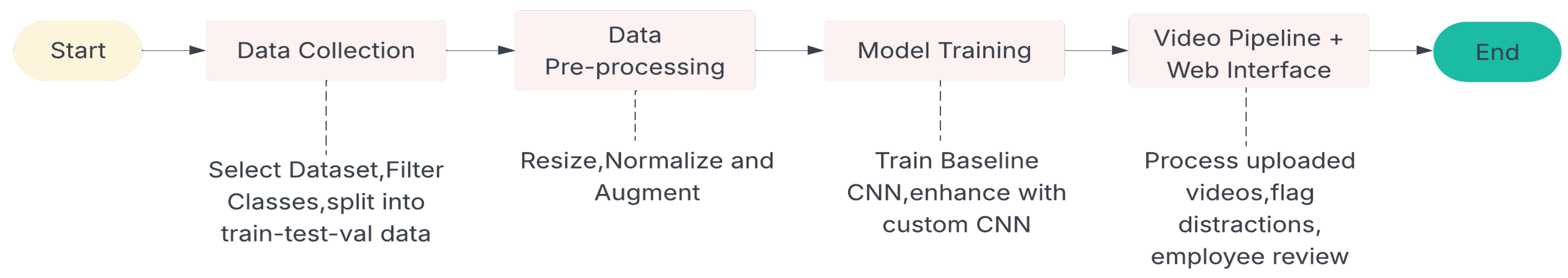


Video Pipeline & Webpage



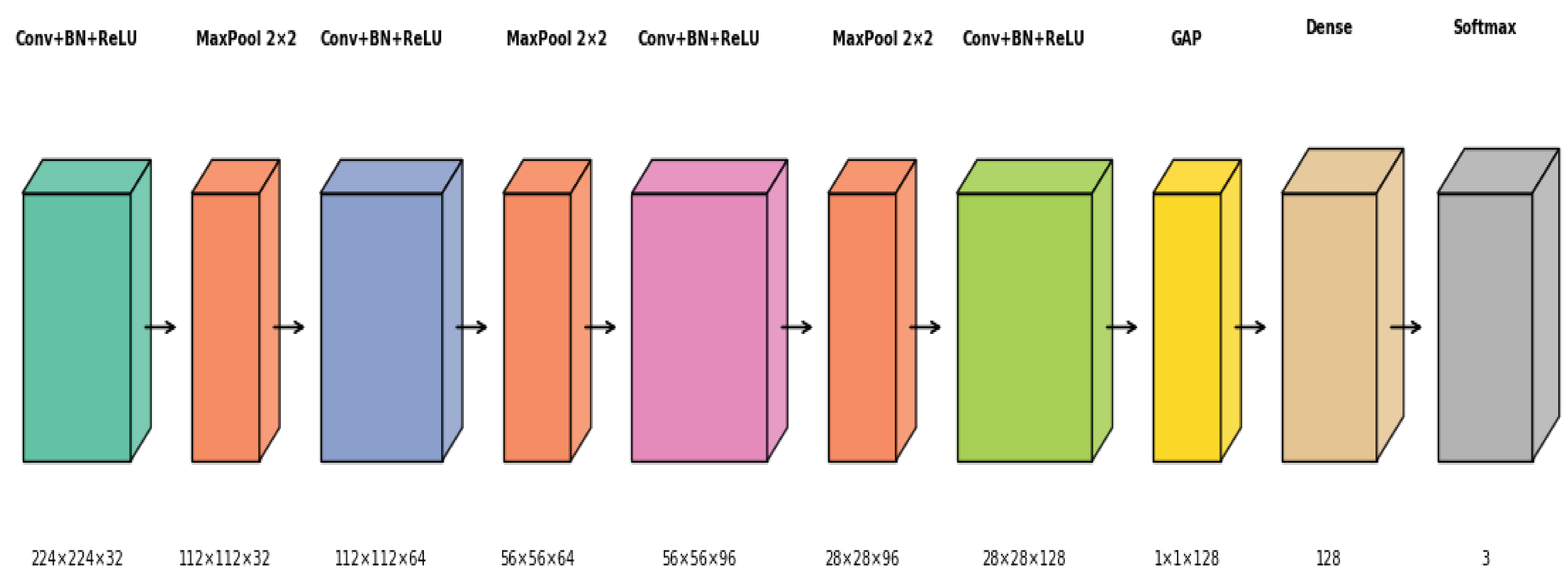
Methodology

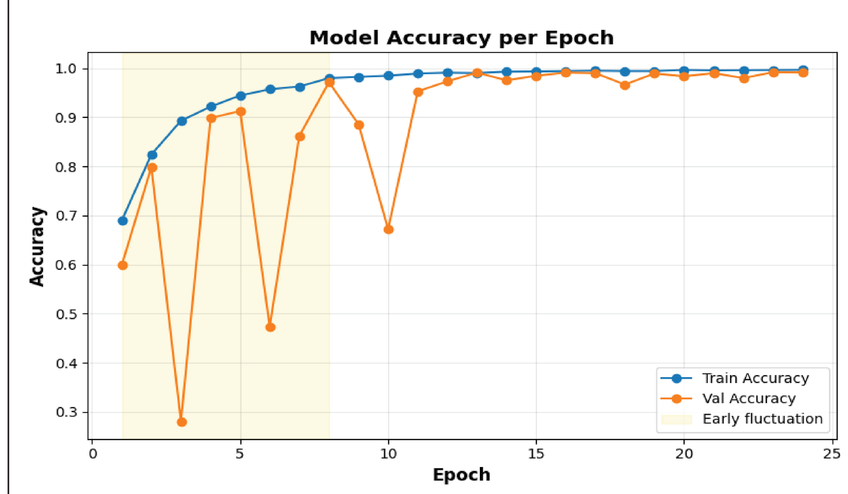
Data Collection: Used the State Farm Distracted Driver Detection dataset, focusing on three key classes: safe driving, using phone, and drinking, organised into train, validation, and test sets with metadata.
Preprocessing: Resized,normalised, and augmented images to standardise inputs and improve generalisation.
Model Training: Developed a baseline CNN and a custom CNN with dropout and class balancing, achieving over 98% test accuracy.
Video Pipeline: Extracted frames from uploaded videos, classified them, and flagged potential distractions for review.
Web Interface: Allowed public video uploads and provided reviewers with tools to assess flagged frames and record decisions.



Model Training

To handle **class imbalance**, we applied targeted augmentation to minority classes and used **class weights** during training, ensuring balanced performance across categories. The **custom CNN** has **four convolutional layers** with **batch normalization** and **ReLU activation**, **max pooling** for downsampling, **global average pooling**, and a dense layer with **regularization** and **dropout**,ending with a **softmax** output for the three driver behaviour classes.



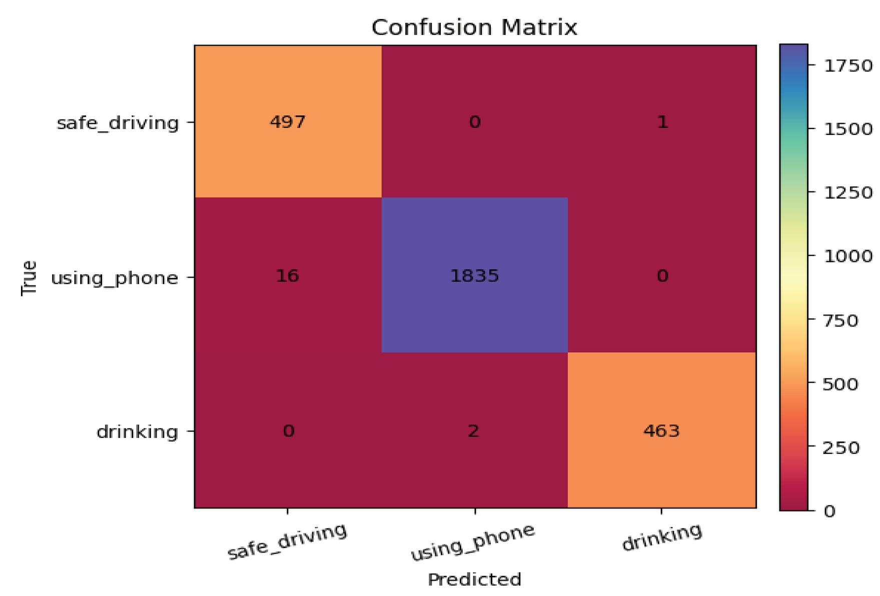


Class	Precision	Recall	F1-score	Support
safe_driving	0.969	0.998	0.983	498
using_phone	0.999	0.991	0.995	1851
drinking	0.998	0.996	0.997	465
accuracy				2814
macro avg	0.989	0.995	0.992	2814
weighted avg	0.993	0.993	0.993	2814

*The consistently high **macro F1 score** of **0.992** further confirms that the **model performs equally well across all classes**, demonstrating that class imbalance was effectively addressed.*

*The **training accuracy** increased **steadily**, while v**alidation accuracy** showed **early fluctuations** due to heavy **augmentation, regularisation, and class weightig effects**.After epoch 8, both **stabilised above 98%,** indicating strong generalisation and effective handling of class imbalance.*

*The **confusion matrix** on the right shows how well the model distinguishes between classes, with **strong diagonal values** indicating **high accuracy**.*



Correctly Classified Examples

Future Scope

Enable **real-time detection** within vehicles, allowing drivers to receive **quick alerts** that can help **prevent accidents**.
Combine video analysis with **sensor data** such as **eye-tracking** or **vehicle telemetry** to capture a more complete picture of driver behaviour.
Implement **privacy-focused** continuous learning, allowing the system to **improve over time** without storing or exposing **sensitive** driver footage.

SCAN FOR REPO

Nishanth Chennagiri Keerthi

Ashish Mohamed Usman

References:

1. Kaggle. (n.d.). State Farm Distracted Driver Detection. Retrieved from https://www.kaggle.com/competitions/state-farm-distracted-driver-detection
2. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778. https://doi.org/10.1109/CVPR.2016.90
3. World Health Organization. (2023). Global status report on road safety. Retrieved from https://www.who.int/publications/i/item/9789240066783