Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.98549526

UPGRADED DROWSINESS DETECTION WITH DEEP LEARNING FOR ADAS

Mayank Agrawal
Masters in Computer Applications, Amity University Noida, AIIT Department
Uttar Pradesh India, agarwal.mayank515@gmail.com

Dr. Monika Sharma
Associate Professor, Information Technology, Amity University Noida,
msharma5@amity.edu

Abstract - Detecting motorist drowsiness is crucial for ensuring road safety. Traditional methods for detecting tiredness rely on manual feature extraction and rule-based systems, which frequently lack precision and generalizability. This paper presents an improved fatigue detection system for automobiles that leverages the strength of deep learning methods. Deep convolutional neural network (CNN) architecture is utilised to automatically learn and extract features from driver-related data, such as facial expressions, eye movements, and physiological signals. The network is trained on a large dataset of labelled examples comprising multiple fatigue levels. The CNN optimises its internal parameters through backpropagation to accurately detect and classify fatigue levels in real-time. The proposed system accomplishes greater precision and reliability than conventional methods. Experimental results demonstrate that the approach based on deep learning outperforms existing systems in terms of detection precision, adaptability to a variety of parameters, and real-time performance. Implementing deep learning in drowsiness detection systems has the potential to significantly improve road safety by providing fatigued drivers with opportune warnings and alerts, thereby mitigating the risks associated with lethargic driving. Future research can concentrate on incorporating additional contextual information and multi-modal data sources to further enhance the performance and reliability of the system.

Keywords – Deep Learning, Drowsiness Detection, Artificial Intelligence, Object Detection, Yolo

I. INTRODUCTION

Modern vehicles are increasingly equipped with advanced driver assistance systems (ADAS), which provide drivers with tools to enhance safety and prevent accidents. A crucial component of ADAS is drowsiness detection system, which can warn drivers when they are becoming fatigued and at risk of falling unconscious behind the wheel. Drowsiness is a significant issue for drivers, as it can impair the

reaction times, judgement, and decision-making skills, resulting in severe accidents. Traditional techniques for detecting drowsiness, such as monitoring pulse rate or tracing eye movements, have limitations in terms of accuracy and efficacy. To enhance drowsiness detection, researchers have turned to artificial intelligence (AI) techniques. AI-based drowsiness detection systems can accurately detect when a driver is becoming drowsy by analyzing data from multiple sensors, such as cameras and accelerometers.AI-based systems can also utilize machine learning algorithms to learn from the driver's behavior and increase detection accuracy over time. In addition, deep learning algorithms can analyze facial expressions to detect signals of fatigue that are difficult to observe with conventional methods. The combination of AI and ADAS has the potential to substantially reduce accidents caused by drowsiness. AI-based drowsiness detection systems can enhance road safety and save lives by alerting drivers when they are becoming sleepy and at risk of falling unconscious. In this paper, we are focusing on the implementation of AI by using object detection algorithms in detecting drowsiness in order enhance safety on the road.

Fig. 1. Accidents caused by drowsy drivers [& Layne, PLLC]

II. LITERATURE REVIEW

Human error causes 94%–96% of automobile accidents, according to the NHTSA. in 2016. At least 90% of automotive crashes are triggered by human error, according to many studies. 20% of drivers have passed out while driving in the last year, and 40% have done so at least once (Matine and & Coleman, LLC).

Advanced Driver Assistance Systems (ADAS) remove human error in vehicle operation. Driver performance is improved by ADAS systems' advanced technology. ADAS uses a variety of technologies for sensors to detect the vehicle's surroundings and inform the driver about them or take action. The driver's eyesight, sensibility, and decision-making are enhanced by ADAS sensors. Are you night-sighted? RADAR can. Before reversing, can you echolocate a youngster behind your car? However, SONAR sensors can. Do you have 360-degree vision? cameras and LiDAR technology can. Always aware of your latitude and longitude? Your car may get this and other data from numerous sets of global positioning satellites ("What Is ADAS (Advanced Driver Assistance Systems)"). AI, image processing, and IoT are frequently utilised to improve road safety since they autonomously acquire and assess data without human mistake("Understanding the Potential of Emerging Digital Technologies for Improving Road Safety") .Computer vision is used to recognise things in photos and videos. Most object detection techniques use machine learning or deep learning. Humans can instantly spot things of interest in photos and movies. Computer-based object detection mimics this cognition. ADAS use object detection to recognise driving lanes and pedestrians for road safety. Surveillance footage and image retrieval benefit from object detection. Several methods can detect objects. Convolutional neural networks (CNNs) like R-CNN and YOLO v2 automatically recognise objects in photos ("What Is Object Detection?"). CNN-based object detectors benefit recommendation systems. High-performance object detection using YOLO models. Each grid in YOLO detects items inside itself. Data streams enable real-time object detection. They use little calculation. Generalised Object Detector will have pre-training and bounding box classification and prediction brains. Backbones support GPU and CPU platforms. For the Dense and Sparse prediction object detectors, the Head may be one-stage

(YOLO, SSD, RetinaNet) or two-stage (Faster R-CNN). Recent object detectors capture feature data in the Neck between the backbone and the head. YOLOv4 uses CSPDarknet53 as a backbone and SPP block to extend the receptive field and separate critical features without affecting network performance. PAN combines backbone parameters. YOLOv4 uses the anchor-based YOLOv3 cranium (Gutta).

III. METHODOLOGY

The process for creating an AI-based drowsiness detection model is as follows:

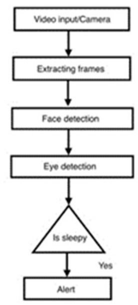

Fig. 2. Drowsiness detection flowchart

A. Data Collection : Collecting data is the initial phase in creating an artificial intelligence-based drowsiness detection system. In our research, we collected 200 high resolution images taken in different lighting conditions.

B. Data Pre-processing : The second stage is data pre-processing, which includes labelling images, changing their orientation, cropping the image, applying grayscale, etc. We have utilised roboflow, a tool for image labelling. It is a free instrument that accelerates the image labelling process.

Fig. 3. Dataset used in the study

C. Selection of object detection algorithm : Object detection is a computer vision technique that enables the identification and localization of objects within an image or video. With this type of identification and localization, object detection can be used to accurately label and count objects in a scene, as well as determine and monitor their precise locations. Consider an image containing two animals and a person, for instance. Object detection enables us to simultaneously classify and locate instances of the detected objects within an image. Multiple object detection algorithms exist, including region-based convolutional neural networks (R-CNN), faster R-CNN, single shot detectors (SSD), and you only look once (YOLO). We utilised the YOLOv5 object detection model in our research due to its rapid computation and processing speed, particularly in real time.

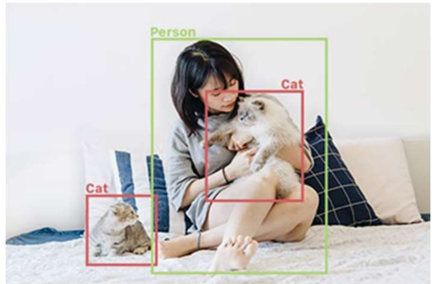

Fig. 4. Object detection

[“OBJECT DETECTION GUIDE | FRITZ AI.” OBJECT DETECTION GUIDE | FRITZ AI, WWW.FRITZ.AI/OBJECT-DETECTION.]

D. Model Training: Training the model is the most important aspect of developing AI for drowsiness detection. To train the model, several parameters must be determined. The batch size and image size come first, followed by the epoch values and model weight. We utilized a 320x320 image with 16 batches and 300 epochs. The Yolo model is available in multiple sizes; we chose a small one for our research.

E. Model Evaluation: The model generates multiple metrics, such as the f1 curve, confusion matrix, recall curve, etc., after successful training. To evaluate the performance of the model, the f1 curve, mean average precision (mAP), and recall values were utilized. In addition, we evaluated the model using photographs, videos, and a live camera feed.

IV. IMPLEMENTATION

A. Developing drowsiness detection using YOLOv5 involves the following implementation steps:

1. Installing and importing dependencies : The first step is to import and install dependencies such as PyTorch, Numpy, Matplotlib, and CV2, and then use the git command to clone the model. After integrating dependencies, the requirements.txt file from the Yolov5 repository was installed.

2. Loading the model : The Yolov5 model from the Ultralytics repository was imported into a variable using the torch.hub.load command.

3. Making initial detections : To verify that the model has been accurately loaded, we performed some preliminary OpenCV detections on photos and through the web camera. The model could distinguish objects such as automobiles, buses, people, etc.

4. Labelling of images : For supervised machine learning, labelled data are essential. We utilised roboflow, an open-source software for image labelling that made the process of image labelling quicker and more accurate.

5. Training on custom dataset : Our custom dataset of 100 awake and 100 drowsy images was utilised to train the model. The sizes of the image and batch were 320 and 16, respectively. Yolov5s was trained for approximately 24 hours with over 300 epochs

6. Loading the final model : After training the model, we loaded it and conducted detections on images of drowsy and awake individuals. We also utilised CV2 to create webcam-based detections.

B. Challenges faced during implementation phase

1. Deciding the appropriate dataset size : Choosing the optimal dataset size is a difficult task. Training on an insufficiently-sized dataset will not produce the intended results. Nevertheless, training on a large dataset, such as 4000 images, will require significant time and resources. Furthermore, training on a large dataset necessitates a robust CPU and GPU.

2. Configuring the train command : Train command requires multiple parameters, including image size, batch size, epoch value, etc. Choosing the ideal values requires effort. In addition, the folder structure of YOLOv5 is not straightforward, and it is sometimes necessary to search for the correct path to the dataset.yaml file, which is essential for effectively executing the train command.

3. Exiting from web camera window : Although using open cv to access a web camera is straightforward. However, exiting the launched window does not always work. Additionally, the live camera feed will become obstructed if your CPU and GPU are underpowered.

V. RESULTS

The dataset used in training consists of 200 images captured under different lighting conditions and labelled "awake" or "drowsy." This dataset is intended for use in drowsiness detection-related machine learning applications. The dataset could be utilised for training and evaluating machine learning models that can distinguish between conscious and drowsy states based on visual signals from images. The dataset could be analysed further to ascertain how illumination affects the accuracy of drowsiness detection systems. Mean average precision (mAP), precision, recall, and the model's f1 score are utilised to evaluate the model's performance. These metrics provide an approximation of the model's performance under various conditions.

Metrics	Results
mAP	99.5%
precision	99.4%
recall	95.0%

Fig. 5. Metrics of our model

mAP calculates the average precision (AP) for every class and then takes the mean across all classes. The AP is a measure of how well the model retrieves pertinent search results. A mAP score of 99.5% indicates that our model performs exceptionally well on the task we evaluated it for. Precision is defined as the ratio of true positives to the sum of true positives and false positives. In other terms, it quantifies the proportion of accurate positive predictions made by the model. A precision of 99.4% indicates that our model has an exceptionally low rate of false positives, i.e., it makes very few erroneous positive predictions. Recall is the ratio of true positives to the sum of true positives and false negatives, or Recall = TP / (TP + FN). In other terms, it assesses the proportion of positive instances correctly identified by the model. A recall of 95% indicates that our model can accurately identify 95% of the positive instances in the dataset.

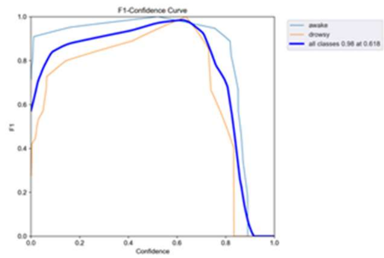

Fig. 6. F1 Curve of yolov5 model

An F1 curve is a plot of the F1 score for a binary classification model as a function of the classification threshold. The threshold represents the probability score above which the model classifies a sample as positive, and below which the sample is classified as negative. In a multi-class classification problem, the F1 curve can be computed for each class, and it plots the F1 score for that class as a function of the classification threshold. The phrase "all classes 0.98 at 0.618" suggests that the F1 score for all classes in the multi-class classification problem is 0.98 at a classification threshold of 0.618. This means that the model has high precision and recall for all classes at this particular threshold value.

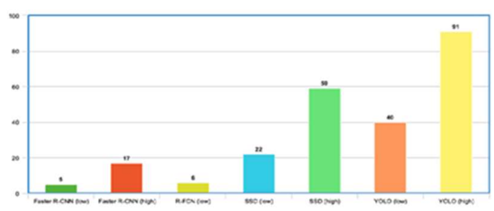

Fig. 7. Comparison of object detection algorithms frames per second https://www.datacamp.com/blog/yolo-object-detection-explained

Frames per second (FPS) is a crucial metric for drowsiness detection in automobiles because it impacts the accuracy and responsiveness of the object detection system used to monitor the driver's behaviour. Object detection systems for drowsiness detection typically analyse video footage from a camera mounted on the dashboard or steering wheel of the vehicle using computer vision algorithms. These algorithms use real-time video processing to detect and monitor the driver's facial features, such as their eyes and mouth, as well as to detect signs of drowsiness such as drooping eyelids or yawning. The greater the FPS, the greater the frequency with which the object detection system can capture and analyse images of the driver's visage.

This improves the accuracy and reliability of drowsiness detection, as the system is now able to detect even subtle changes in the driver's facial expressions that indicate lethargy. If the FPS is too low, however, the object detection system may overlook crucial details or movements of the driver's visage, resulting in inaccurate or delayed detection of drowsiness. This can be hazardous because it may cause accidents or other road safety hazards. YOLO is incredibly quick due to the absence of complex pipelines. It can compute 45 frames per second (FPS) of images. Moreover, YOLO achieves more than twice the mean Average Precision (mAP) of other real-time systems, making it an excellent candidate for real-time processing. With 91 frames per second, YOLO is far superior to other object detectors, as shown in the Fig. 8.

Fig. 8. Training time comparison

One of the most significant enhancements to the YOLOv5 architecture is the addition of the Focus layer, represented by a single layer, which replaces the first three layers of YOLOv3. This integration decreased the number of layers and parameters and increased both forward and reverse speed without significantly affecting the mAP.

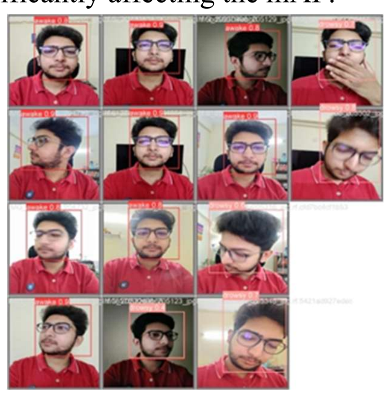

Fig. 9. Model results on validation images

In object detection tasks, a bounding box with a score of 0.9 for the awake class and a red colour suggests the model is confident that it has detected a "awake" class object in the image. The score indicates the degree of trust in the model's prediction, with a score of 0.9 indicating a high level of confidence. The colour of the bounding box can be used to indicate the class of the detected object; in this case, red indicates the "awake" class. Our machine learning model's drowsiness detection accuracy is extremely high, based on the metrics . The following is a concise summary of what these metrics indicate: Mean average precision (mAP) of 99.5% is a common evaluation metric for object detection models that evaluates the average precision across various levels of recall. A mAP of 99.5% indicates that our model can detect drowsiness with high precision and recall. Precision of 99.4%: Precision is the proportion of true positives (i.e., correctly identified instances of lethargy) among all positive instances. A precision of 99.4% indicates that our model rarely misidentifies non-drowsy instances as drowsy, or that it has a very low rate of false positives. Recall of 95.0%: Recall assesses the proportion of true positives among all actual positive instances (whether identified correctly or erroneously). A recall of 95.0% indicates that our model correctly identifies the majority of drowsy instances and has a low rate of erroneous negatives, i.e., it overlooks actual drowsy instances only rarely. These metrics indicate that our model detects drowsiness with a high level of precision and recall. These high performance values may be beneficial in real-world settings where detecting fatigue is crucial for safety, such as in transportation and heavy machinery operation. However, it is essential to bear in mind that the performance of the model can vary depending on the specific use case and dataset, and it is always recommended to evaluate the model's performance in a variety of scenarios before deploying it in real-world applications. Depending on the specific implementation and configuration, faster R-CNN has attained a mean average precision (mAP) of up to 42.1%. This is less than the mAP of 99.5% which we have achieved with our model, indicating that your model may outperform Faster R-CNN in terms of overall detection accuracy.

Algorithms	**Speed**	**Accuracy**	**Ease of Implementation**
Faster RCNN	Bad	Good	Bad
SSD	Good	Good	Bad
Yolo	Good	Good	Good

Fig. 10. Comparison table of object detection algorithms

Single Shot Detector (SSD) and You Only Look Once (YOLO) are significantly faster than the Faster R CNN algorithm. Faster R CNN is based on two shot detectors, which requires two image analysis stages. In addition, Faster R CNN can only produce 7 frames per second on a single GPU, which is less than the other algorithms. In terms of precision, the algorithms perform comparably. However, in terms of implementation simplicity, YOLO is superior to the other two. YOLO can be implemented with significantly less code and is simpler to implement than the other two algorithms. YOLOv8 is the most recent state-of-the-art YOLO model for object detection, image classification, and instance segmentation. Ultralytics devised YOLOv8, as well as the influential and industry-defining YOLOv5 model. Compared to YOLOv5, YOLOv8 includes numerous architectural and developer experience modifications and enhancements. As of the writing of this post, YOLOv8 is undergoing active development, with Ultralytics working on new features and responding to community feedback. When Ultralytics releases a model, it receives long-term support: the organisation collaborates with the community to improve the model.

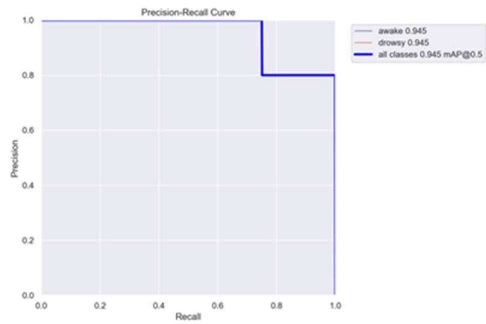

Fig.11. Precision Recall Curve of Yolov8s

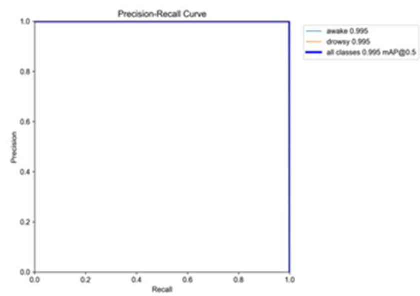

Fig. 12. Precision Recall Curve of Yolov5s

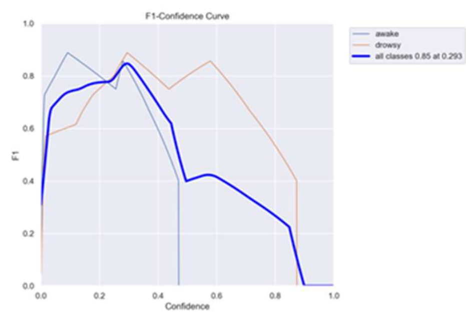

Fig. 13. F1 curve of Yolov8s model

The most recent Yolo model created by Ultralytics is Yolov8. Yolov8 has numerous advantages over previous Yolo models, including a shorter training period, a more straightforward folder structure, and greater accuracy. Performance improves with a greater precision-recall (PR) curve. Thus, YOLOv5 with a PR curve value of 0.995 is better than YOLOv8 with 0.945.The precision-recall curve is a graphical illustration of the trade-off between precision (the model's ability to properly forecast positive examples) and recall (the model's ability to identify all positive instances) for different categorization thresholds. The curve summarises the model's performance across thresholds and can be used to evaluate its efficacy.The model's precision increases with recall. YOLOv5 has higher precision and recall for positive cases. It suggests YOLOv5 detects objects or classes of interest more accurately and reliably.

Model	F1 Score	PR Curve Score
Yolov8s	0.85	0.945

Yolov5s	0.98	0.995

Fig. 14. Comparison between Yolov5s and Yolov8s

However, after training Yolov5 and Yolov8 on the same dataset, we discovered that Yolov5 had a higher f1 score than the most recent Yolov8. Yolov5 received a score of 0.98 while Yolov8 earned 0.85.It simply means that Yolov5 has shown higher accuracy than the most recent Yolov8 model.

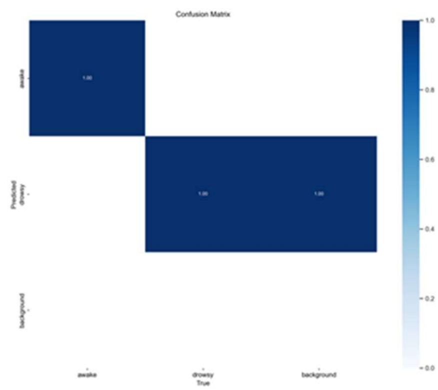

Fig. 15. Confusion matrix of yolov5s

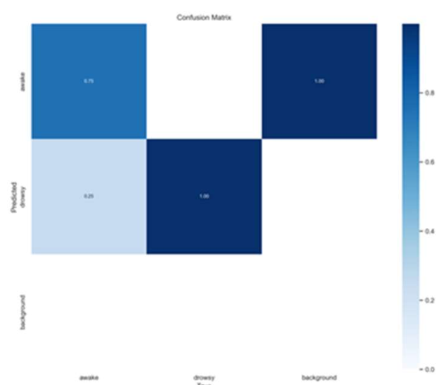

Fig.16. Confusion matrix of yolov8s

Actual versus predicted values are represented visually in the Confusion Matrix. This table-like structure assesses the efficacy of our Machine Learning classification model. Confusion matrix of both the models are relatively same. However for yolov8s 25% of awake images are predicted as drowsy.

VI. CONCLUSION

The detection of drowsiness is an integral component of advanced driver assistance systems. There are several methods for developing a fatigue detection system, including sensor-based monitoring, monitoring of physiological signals, and computer vision-based methods. However, the computer vision approach is frequently regarded as effective for detecting fatigue due to its ability to analyse facial features and eye movements, which are significant indicators of drowsiness. Utilising visual data, it detects patterns and symptoms of fatigue in real time.

Here are a few reasons why computer vision is well-suited for detecting drowsiness: Non-intrusive: Methods based on computer vision can capture and analyse visual signals without physical contact with the driver. This feature is advantageous because it does not interfere with the driving experience or necessitate the attachment of additional sensors or devices to the driver. Computer vision techniques can analyse video streams in real time, allowing for continuous monitoring of the driver's facial expressions and eye movements. Real-time monitoring enables timely detection of drowsiness events and immediate driver alerts. Analysis of the face and eyes reveals that drowsiness is frequently accompanied by facial and eye-related changes, such as drooping eyelids, protracted eye closure, and altered facial expressions. These features can be detected and analysed by computer vision algorithms to accurately identify indicators of fatigue. Adaptability to diverse environments: Computer vision techniques can be trained on large datasets containing illumination conditions, driver appearances, and camera angles that vary. This adaptability enables the system to perform well in a variety of real-world driving scenarios and to accommodate various drivers. YOLO (You Only Look Once) is a well-known object detection algorithm that has earned recognition for its real-time performance and precision. The strengths of YOLO make it suitable for fatigue detection in computer vision-based methods: Rapidity and productivity: Yolo's rapid inference pace makes it suitable for real-time applications such as slumber detection. It processes images holistically and predicts bounding boxes and class probabilities directly in a single pass, resulting in efficient processing with minimal computational overhead. Accuracy and object detection capabilities: Yolo can detect and localise multiple objects in an image with high precision. This is advantageous for drowsiness detection systems because it enables the detection of pertinent facial landmarks and ocular regions for analysing drowsiness-related characteristics. Adaptability and generalisation: Yolo's capacity to acquire complex representations and patterns from large datasets enables it to generalise well to various object categories and environments. This adaptability is advantageous in slumber detection systems because it allows the algorithm to account for variations in facial appearance, illumination conditions, and other factors. According to our research, Yolo is a highly effective solution for detecting drowsiness. Moreover, a side-by-side comparison of Yolov 5 and Yolov 8 revealed that Yolov 5 surpassed Yolov 8 in terms of accuracy, indicating that Yolov 5 is one of the most viable options for fatigue detection.

VII. REFERENCES

1. Matine, David, and Caroselli & Coleman, LLC Beachler. “What Percentage of Car Accidents Are Caused by Human Error? | Pittsburgh Law Blog.” Caroselli, Beachler & Coleman, L.L.C., 22 Sept. 2021, www.cbmclaw.com/what-percentage-of-car-accidents-are-caused-by-human-error.

2. “What Is ADAS (Advanced Driver Assistance Systems).” Data Acquisition | Test and MeasurementSolutions, dewesoft.com/blog/what-is-adas.

3. “Understanding the Potential of Emerging Digital Technologies for Improving Road Safety.” Understanding the Potential of Emerging Digital Technologies for Improving Road Safety - ScienceDirect, 28 Dec. 2021, https://doi.org/10.1016/j.aap.2021.106543.

4. “What Is Object Detection?” What Is Object Detection? - MATLAB & Simulink, www.mathworks.com/discovery/objectdetection.html.

5. Gutta, Surya. “Object Detection Algorithm—YOLO V5 Architecture.” Medium, 4 Aug. 2021, medium.com/analytics-vidhya/object-detection-algorithm-yolo-v5-architecture-89e0a35472ef.

6. “What Is Transfer Learning? Exploring the Popular Deep Learning Approach.” Built In, 25 Aug. 2022, builtin.com/data-science/transfer-learning.

7. blog.roboflow.com/yolov5-improvements-andevaluation.

8. “Real-time Drowsiness Detection System for an Intelligent Vehicle Publication | IEEE Xplore, ieeexplore.ieee.org/abstract/document/4621125.

9. https://www.analyticsvidhya.com/blog/2022/05/drowsiness-detection-system/

10. www.yourwvinjuryattorneys.com/car-accidents/drowsy-driving-responsible-for-21-percentof-fatal-accidents.

11. Spiceworks.” Spiceworks, www.spiceworks.com/tech/iot/articles/what-is-adas.