Driver Drowsiness Detection Using CNN

Sakshamdeep Singh
50475857
ssingh86@buffalo.edu

Kamala Krishnan
50478764
kamalakr@buffalo.edu

Shashank Vivek Baratwaj
50479686
svivekba@buffalo.edu

Abstract

Drowsiness is a critical contributor to driving accidents and fatalities worldwide, highlighting the urgent need for effective drowsiness detection systems. While several methods have been developed to detect drowsiness, the image processing technique has emerged as the most accurate and efficient approach. Deep learning has revolutionized image processing, making it a promising tool for real-time drowsiness detection in driving situations. In this study, we aim to leverage deep learning algorithms to develop a drowsiness detection system for drivers. Our primary objective is to create a system that can accurately detect drowsiness in real-time and alert drivers to prevent potential accidents. To achieve this goal, we plan to use state-of-the-art models and techniques in deep learning and image processing. The proposed system's implementation and evaluation will involve assessing the accuracy and efficiency of the system by utilizing various evaluation metrics, including precision and recall, latency, false positive rate, and usability. Overall, our project's anticipated outcomes will be a reliable and effective drowsiness detection system for drivers that can be widely implemented to enhance road safety.

1 Motivation

The proposed project is inspired by the urgent need to address the growing concern of driver drowsiness, which has contributed to an alarming rate of road accidents. According to a report by the National Highway Traffic Safety Administration in the USA, drowsy driving resulted in a staggering 72,000 crashes, 44,000 injuries, and 800 fatalities in 2013 alone. Despite significant efforts to mitigate this issue, the number of fatalities caused by drowsy driving continues to escalate each year. To tackle this pressing problem, our project aims to develop a drowsiness detection system, leveraging a custom Convolutional Neural Network architecture. Our innovative solution will promptly alert drivers who are at risk of falling asleep, potentially preventing accidents and saving lives. By implementing our proposed system, we hope to make roads safer for all drivers and passengers. Our team is committed to delivering a practical solution to a growing concern, and we believe that our proposed project will significantly contribute to reducing the number of fatalities caused by drowsy driving.

2 Introduction

2.1 Background

Drowsiness detection has been a subject of increasing interest in recent years due to its direct impact on road safety. Various techniques have been proposed to detect drowsiness, including physiological measurements, vehicle-based measures, and behavioral monitoring. However, these methods often suffer from limitations such as invasiveness, reliance on additional hardware, and lack of real-time applicability.

With the advent of deep learning and advancements in computer vision, image processing techniques using Convolutional Neural Networks (CNNs) have shown significant potential in addressing these limitations. CNNs have been successfully applied to various tasks in image recognition, including object detection, facial recognition, and activity recognition. As such, they present a promising approach for real-time drowsiness detection based on facial features and eye activity.

2.2 Objectives

Despite the existence of several drowsiness detection methods, a large number of accidents and fatalities continue to occur due to drowsy driving. There is an urgent need for a real-time, non-invasive, and accurate drowsiness detection system that can be widely implemented in vehicles to improve road safety.

The primary goal of this project is to design and develop a drowsiness detection system using a CNN architecture that can accurately and efficiently recognize drowsy drivers based on facial features and eye activity. The proposed system should be capable of alerting drivers in real-time to prevent potential accidents, thus contributing to a safer driving environment.

To address this problem, the following research questions will guide the project:

- How can a CNN architecture be designed and optimized for real-time drowsiness detection based on facial features and eye activity?
- How does the performance of the proposed drowsiness detection system compare to existing methods in terms of accuracy, efficiency, and real-time applicability?
- What are the practical challenges and limitations associated with implementing the proposed drowsiness detection system in real-world driving scenarios?

2.3 Scope and Limitations

2.3.1 Scope

The scope of this project is focused on the following aspects:

- Development of a drowsiness detection system using a custom CNN architecture and YOLO (You Only Look Once) for real-time detection based on facial features and eye activity.
- Integration of YOLO to enhance the speed and efficiency of the drowsiness detection system, allowing for rapid identification of drowsy drivers.
- Utilization of publicly available dataset for training and validating the CNN. The datasets contain images of drivers with varying degrees of drowsiness. We ended up creating a custom data for training the YOLO model.
- Evaluation of the system's performance using standard metrics such as accuracy, precision, recall, F1-score, and latency. These metrics will be used to compare the performance of the proposed system to existing drowsiness detection methods.
- Exploration of strategies to improve the model's robustness and generalization, such as data augmentation, transfer learning, and model fine-tuning.
- Analysis of the system's potential for real-world implementation, considering factors such as computational resources, compatibility with existing hardware, and user acceptance.

2.3.2 Limitation

- Data quality and diversity: The performance of the proposed system largely depends on the quality and diversity of the training data. If the dataset does not cover various driving conditions, demographics, and drowsiness levels, the system may suffer from biases and limited generalization capabilities.
- Environmental factors: The system's performance may be affected by environmental factors such as lighting conditions, occlusions, and camera angles. The project may not fully address these challenges, resulting in degraded performance under certain conditions.

3 Relevant Work

In the field of drowsiness detection, there have been significant developments in recent years. While some commercial products exist for detecting drowsiness, particularly those implemented by automotive manufacturers, there is a need for greater accuracy in predicting and monitoring this condition. This is because existing vehicle-based methods may not provide a comprehensive analysis of the driver's condition, resulting in false alarms or missed detections. To address this issue, researchers have developed state-of-the-art models that incorporate a variety of different methods for detecting drowsiness. These methods include physiological, behavioral, and vehicle-based approaches, which are combined in what is known as hybrid methods. These hybrid methods have shown promising results in accurately detecting drowsiness in drivers, making them an attractive option for developing a robust drowsiness detection system.

In our proposed project, we aim to build on this existing work by implementing a deep learning-based approach using YOLO for real-time drowsiness detection. To achieve this, we integrated a custom dataset, which we created using our own faces in awake and drowsy states, with a publicly available dataset of open and closed eyes. This integration provides a more diverse and representative dataset for training and validating the YOLO model, ensuring robust performance across various driving conditions and driver demographics.

Our approach stands out by leveraging the real-time capabilities of YOLO, which enables rapid detection of drowsy drivers, making it highly suitable for practical applications. By combining the strengths of YOLO with the custom dataset, we aim to develop a highly accurate and efficient drowsiness detection system that can overcome the limitations of existing methods, such as false alarms and missed detections. Ultimately, our project seeks to contribute to the advancement of drowsiness detection systems, making roads safer for all drivers and passengers.

4 Methodology

4.1 Initial Approach and Limitations

4.1.1 Publicly Available Dataset

In our first attempt to build a driver drowsiness detection system, we used a publicly available dataset from Kaggle.

`https://www.kaggle.com/datasets/dheerajperumandla/drowsiness-dataset`

This dataset comprised four classes of images: open eyes, closed eyes, Yawning and Non-Yawning drivers. We employed a Convolutional Neural Network (CNN) model to perform classification on this dataset.

While the initial approach seemed viable, we quickly realized that the dataset's structure and the resulting classification did not make practical sense for real-world applications. The limitations of the publicly available dataset stem from its fragmented representation of drowsiness indicators, which do not capture the full complexity of driver drowsiness. This disjointed approach to drowsiness detection can lead to inaccurate classifications, particularly when applied to real-world driving situations. A more comprehensive and context-aware approach, such as using a custom dataset that includes a range of drowsiness indicators and various real-life driving conditions, is needed to develop a more effective and reliable driver drowsiness detection system.

4.1.2 Practical Challenges and Limitations

The limitations of the publicly available dataset are significant, particularly when attempting to create a practical and effective driver drowsiness detection system. The two classes of open eyes and closed eyes consist of annotated images of eyes alone, while the other two classes, drowsy and non-drowsy drivers, focus on yawning and non-yawning behaviors. This separation of features leads to several challenges and limitations:

1. **Incomplete representation of drowsiness:** Relying solely on eye states (open or closed) for drowsiness detection ignores other crucial indicators of drowsiness, such as yawning

or microsleep episodes. Conversely, using yawning as the sole criterion for drowsiness classification can also be misleading, as a person might yawn due to other factors, like boredom or environmental conditions, even if they are not drowsy. Remove?

2. **Inaccurate classification in real-time scenarios:** Since the classification of drowsiness is based on isolated features (eyes and yawning), the model may produce incorrect results in real-world driving situations. For example, a person might be opening their eyes while yawning, leading the model to misclassify them as awake, despite the presence of drowsiness.
3. **Lack of context:** The publicly available dataset does not account for various factors that can influence the appearance of the eyes or yawning behavior, such as lighting conditions, head orientation, and facial expressions. This lack of context may result in suboptimal model performance when applied to real-life scenarios, where drivers exhibit diverse behaviors and appearances.

4.2 YOLO Model with Custom Dataset

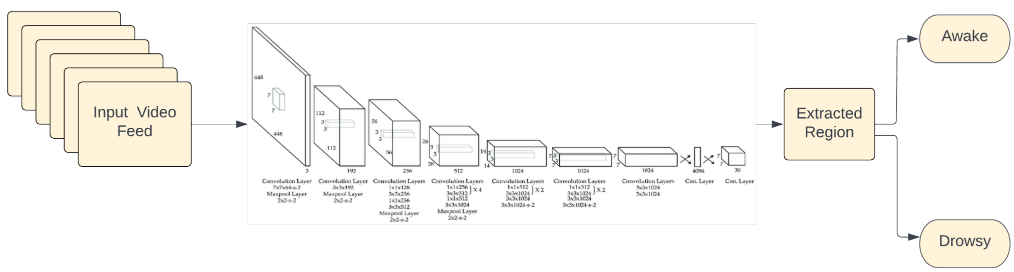


Figure 1: YOLO Architecture

4.2.1 Dataset Collection and Annotation

In order to address the limitations of the publicly available dataset and better represent the target problem, we created a unique dataset featuring images of our team members in both awake and drowsy states. The process of developing this custom dataset involved several key steps to ensure a diverse and representative collection of images:

1. **Image collection:** We gathered a total of 84 images (64 for training and 20 for testing) of each team member under varying lighting conditions and head orientations. This diverse set of images aimed to accurately represent the different scenarios that the drowsiness detection system would encounter in real-world conditions.
2. **Annotation:** Using the LabelImg tool, we manually annotated each image with labels of either 'awake' or 'drowsy', taking into account the appearance of the eyes, facial expressions, and other relevant features that could indicate the state of alertness or drowsiness in the subject.
3. **Data preprocessing:** To ensure compatibility with the input requirements of the YOLO model, we resized and normalized the images in the dataset. This preprocessing step aimed to standardize the image data, allowing the model to learn more effectively from the dataset and improving its overall performance.
4. **Data splitting:** To evaluate the performance of our custom YOLO model, we divided the dataset into training and validation sets. This partitioning enabled us to assess the model's accuracy and generalizability, providing valuable insights into its effectiveness in detecting drowsiness in real-world scenarios.

By developing a custom dataset with diverse and representative images, we aimed to create a more robust and accurate drowsiness detection system. This dataset served as a valuable resource for training and validating our YOLO model, ultimately contributing to the development of a more effective and reliable solution for driver drowsiness detection.

4.2.2 Model Training and Validation

We loaded the pre-trained YOLOv8 model from Ultralytics as a starting point and then trained it our custom drowsiness detection model. The results indicate a high performance of our custom YOLO model in detecting drowsiness, with mAP50 values of 0.995 and 0.995 on the training and validation sets, respectively.

4.2.3 Real-time Implementation

To evaluate the effectiveness of our drowsiness detection system in real-time, we used the OpenCV library, specifically the 'cv2.VideoCapture' function. This function allows us to capture video frames from a camera in real-time, which can then be fed into our YOLOv8 model for drowsiness detection. The advantage of using 'cv2.VideoCapture' is that it provides a simple and efficient way to obtain video input for our model, making it possible to test the model's performance in realistic scenarios. By analyzing the video frames in real-time, our drowsiness detection system can identify signs of drowsiness and generate alerts promptly, potentially preventing accidents caused by drowsy driving.

In summary, our final approach to driver drowsiness detection utilized a custom dataset and a YOLOv8 model, which addressed the practical challenges and limitations of our initial attempt. The custom dataset, containing images of team members in both awake and drowsy states, provided a more representative and meaningful input for the model. The YOLOv8 model, combined with real-time video processing using the OpenCV library, enabled us to effectively detect driver drowsiness and issue timely alerts, demonstrating the potential of our system in real-world driving scenarios.

5 System Architecture and Workflow

5.1 Software and Hardware Requirements

The project implementation relied on the following hardware and software specifications:

1. **Processor:** 12th Gen Intel(R) Core(TM) i7-12700H, with a base clock speed of 2.30 GHz.
2. **Memory:** 16.0 GB of installed RAM, with 15.7 GB usable.
3. **Web Camera:** Full HD Integrated Web Camera (720p)
4. **Operating System:** 64-bit version with x64-based processor.
5. **Graphics:** NVIDIA GeForce RTX 3050 GPU, with 4GB dedicated video memory and 2048 CUDA Cores.

The powerful hardware, particularly the NVIDIA GPU, facilitated efficient parallel processing and accelerated the CNN training on the extensive dataset. For software requirements, Python was the primary programming language, supported by deep learning libraries such as Pytorch, Ultralytics, and OpenCV for image processing tasks.

5.2 System Design and Workflow

The drowsiness detection system combines custom and publicly available datasets, a custom-trained YOLOv8 model, and real-time video processing. The system architecture and workflow involve the following components:

1. **Data Collection and Preprocessing:** The system utilizes a custom dataset created by capturing and annotating images. Data preprocessing encompasses resizing images, data augmentation, and normalization to ensure YOLOv8 model compatibility.
2. **Model Training:** The YOLOv8 model is trained on a custom dataset. The training process refines the model's parameters to accurately identify drowsiness-related features such as closed eyes and yawning. Model performance is assessed using metrics like precision, recall, and F1-score.
3. **Real-time Drowsiness Detection:** The trained YOLOv8 model is integrated into a real-time video processing pipeline using OpenCV's 'cv2.VideoCapture' function. This enables the system to capture video frames from a camera, pass them through the YOLOv8 model for

inference, and display drowsiness detection results on the original frame in real-time. The system is designed to alert drivers promptly, potentially preventing accidents and saving lives.

4. **System Evaluation:** The drowsiness detection system's performance is evaluated using metrics such as accuracy, latency, and false positive rate. The system's usability and effectiveness in real-world scenarios are examined to ensure widespread adoption and enhanced road safety.

The system architecture and workflow emphasize state-of-the-art deep learning techniques and real-time video processing to develop a reliable and effective drowsiness detection system for drivers. This system aims to reduce accidents and fatalities caused by drowsy driving, ultimately contributing to safer roads for all drivers and passengers.

6 Results and Discussion

The performance of the custom-trained YOLOv8 model was evaluated on both the training and validation sets. The results are presented below:

6.1 Training Set Evaluation

Class	Images	Instances	Box(Precision)	Recall	mAP50	mAP50-95
all	64	64	0.996	1	0.995	0.901
drowsy	64	41	0.994	1	0.995	0.928
awake	64	23	0.998	1	0.995	0.874

6.2 Validation Set Evaluation

Class	Images	Instances	Box(Precision)	Recall	mAP50	mAP50-95
all	20	20	0.993	1	0.995	0.8
drowsy	20	13	1	0.999	0.995	0.704
awake	20	7	0.986	1	0.995	0.895

The results show that our custom YOLOv8 model demonstrates high performance in detecting drowsiness on both training and validation sets. The model achieves high precision, recall, and mAP50 scores, indicating its ability to accurately identify drowsy and awake drivers.

6.3 Precision-Recall Confidence Curves

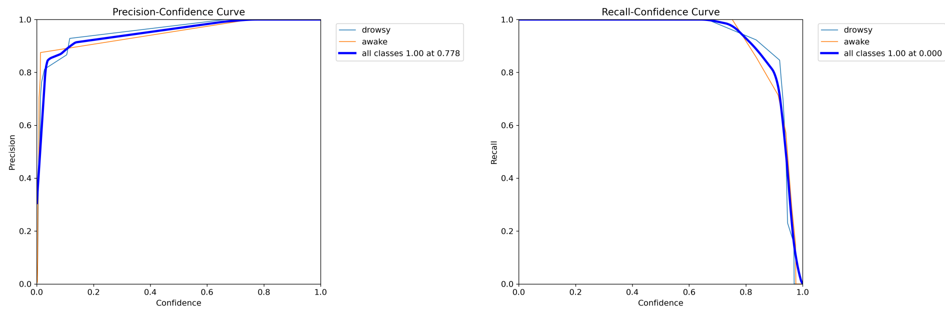
The Precision-Recall Confidence curves help visualize the trade-off between precision and recall at different confidence levels. In our project, the curves indicate that the model maintains high precision across varying recall levels, confirming its capability to detect drowsiness effectively.

6.4 Real-Time Drowsiness Detection

The YOLOv8 model was integrated into a real-time video processing pipeline, enabling real-time drowsiness detection using captured video frames. Sample images from the real-time implementation demonstrate that the model can accurately detect and classify drivers as awake or drowsy, with confidence scores boxed around the face.

6.5 Training and Validation Losses

The grouped graph showing train/val loss, train/dfl loss, and val/dfl loss metrics provides insights into the model's learning process. The graph reveals that the model's training and validation losses decrease as the training progresses, suggesting that the model is learning to recognize drowsiness-related features effectively.



(a) Precision Confidence Curve (b) Recall Condifence Curve

Figure 2: Precision and Recall Graphs

Figure 3: Real Time Detection done using LabelImg

6.6 Summary of Findings

Overall, the results demonstrate the effectiveness of our custom-trained YOLOv8 model in detecting drowsiness in real-time. The high precision, recall, and mAP50 scores on the training and validation sets, coupled with the promising performance in real-time video processing, indicate that our drowsiness detection system can potentially improve road safety and save lives by alerting drowsy drivers.

7 Conclusion

The custom-trained YOLOv8 model developed for driver drowsiness detection has demonstrated promising results in accurately classifying drivers as awake or drowsy. The high precision, recall, and mAP50 scores on the training and validation sets, as well as the effective real-time video processing, suggest that our drowsiness detection system has the potential to significantly improve road safety and save lives by alerting drowsy drivers before accidents occur. The system's effectiveness in detecting drowsiness and its real-time performance make it a valuable tool for enhancing safety on the roads.

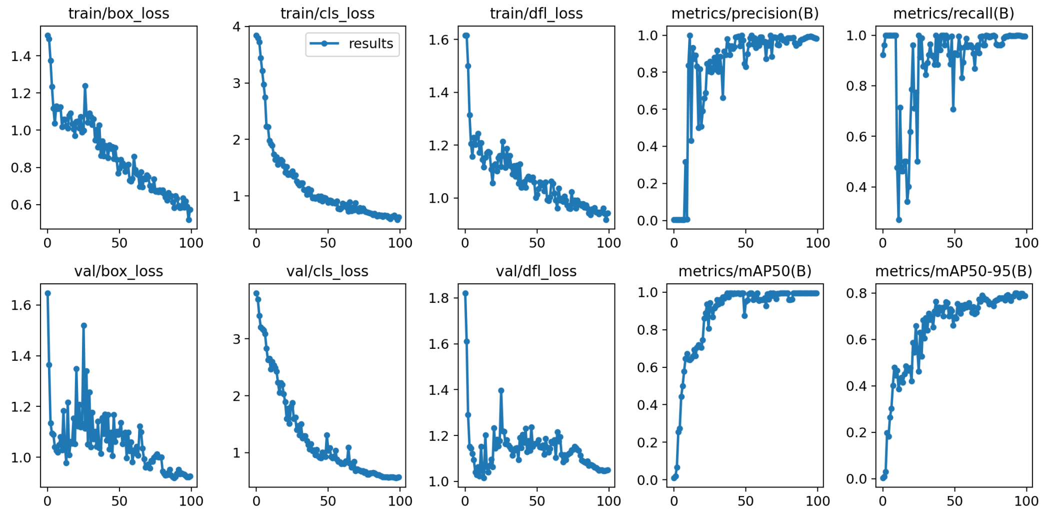


Figure 4: Overall Results

8 Future Work

Despite the encouraging results, there is still room for improvement and potential future research directions to further enhance the performance and applicability of the drowsiness detection system:

1. **Expanding the dataset:** The performance of the model could be further improved by incorporating a more diverse dataset, including images of drivers from various ethnicities, age groups, and with different facial features, to ensure the model's robustness and generalizability.
2. **Multi-modal data fusion:** Integrating additional physiological and contextual data, such as heart rate, head pose, or vehicle speed, could enhance the system's overall accuracy and reliability in detecting drowsiness and help reduce false alarms.
3. **Developing a user-friendly interface:** Designing a more intuitive and user-friendly interface for the real-time drowsiness detection system would improve its usability, making it more accessible to a wider range of users, such as fleet operators, taxi drivers, or individual car owners.
4. **Evaluating the system in real-world driving scenarios:** Conducting extensive field tests and evaluations in real-world driving conditions could provide valuable insights into the system's effectiveness, usability, and impact on driver behavior, ultimately contributing to the further refinement and optimization of the system.

By pursuing these future research directions, the driver drowsiness detection system can be further refined and optimized, ensuring its widespread adoption and contributing to safer roads for all drivers and passengers.

References

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016* url=https://arxiv.org/pdf/1506.02640.pdf

[2]M. S. Basit, U. Ahmad, J. Ahmad, K. Ijaz and S. F. Ali, "Driver Drowsiness Detection with Region-of-Interest Selection Based Spatio-Temporal Deep Convolutional-LSTM," *2022 16th International Conference on Open Source Systems and Technologies (ICOSST), Lahore, Pakistan, 2022, pp. 1-6, doi: 10.1109/ICOSST57195.2022.10016825.*

[3] M. Elham Walizad, M. Hurroo and D. Sethia. "Driver Drowsiness Detection System using Convolutional Neural Network," *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*

[4] Y. Suresh, R. Khandelwal, M. Nikitha, M. Fayaz and V. Soudhri, "Driver Drowsiness Detection using Deep Learning," *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*

[5] Rateb Jabbar, Khalifa Al-Khalifa, Mohamed Kharbeche, Wael Alhajyaseen, Mohsen Jafari, Shan Jiang, "Real-time Driver Drowsiness Detection for Android Application Using Deep Neural Networks Techniques", *Procedia Computer Science, Volume 130, 2018, Pages 400-407, ISSN 1877-0509*