Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 7, No.1, July : 2025

DRIVING SAFETY MONITORING SYSTEM BASED ON DROWSINESS DETECTION: A DEEP LEARNING APPROACH USING YOLO

Sudharani Bangarimath, PG Student, Department of Computer Science and Engineering (Data Science) D. Y. Patil Agriculture and Technical University,Talsande Kolhapur,Maharashtra,India.
Dr. Sangram T. Patil, Professor, Department of Computer Science and Engineering (Data Science) D. Y. Patil Agriculture and Technical University,Talsande Kolhapur,Maharashtra,India.
Dr. Jaydeep Patil, Associate Professor, Department of Computer Science and Engineering (Data Science) D. Y. Patil Agriculture and Technical University,Talsande Kolhapur,Maharashtra,India.

ABSTRACT
Traffic accidents kill and injure most people. Traffic accident injuries kill a million people annually, according to the WHO. Lack of sleep, rest, or fatigue can cause drivers to fall asleep, endangering themselves and others. This research focuses on driver sleepiness detection and response. The artificial intelligence represents advanced technology, offering solutions to create a majority of vision-based applications. In this research, a driver drowsiness detection system has been developed utilizing the deep learning Yolo v8 algorithm. An image dataset comprising two classes, namely drowsiness and non-drowsiness, is utilized for training with Yolov8. Image classification is conducted based on the condition of the driver's eyes, specifically whether they are open or closed. The validation of the trained model indicates an mean Average Precision (mAP) of 80% and an F1-Score of 97.5%.

Keywords:
Deep Learning, Driver Drowsiness,Yolo.

I. Introduction

Driver drowsiness leading to accidents presents a considerable challenge on a global scale. Fatigue or drowsiness in drivers significantly diminishes their reaction times and decision-making capabilities, potentially leading to dangerous situations such as falling asleep while driving, which can result in accidents. In low- and middle-income countries, a staggering 92% of global road fatalities take place, despite these regions accounting for approximately 60% of the world's vehicles [1]. It estimates that driver fatigue contributes to approximately 20% of road accidents worldwide. Data from multiple nations underscores the gravity of the issue. Three main techniques are commonly employed to detect drowsiness.

A report from the Indian National Safety Council reveals that drowsy driving was associated with almost 100,000 accidents, 71,000 injuries, and 1,500 fatalities in 2017. Data reveals that 21% of these incidents were linked to fatigued drivers, resulting in errors. The impact of fatigue as a contributing factor is frequently undervalued among the various elements that can result in accidents, indicating that the true figure may still be rather conservative. Insufficient infrastructure and fatigue play significant roles in the occurrence of disasters in developing countries. Unlike substances like alcohol and narcotics that have distinct indicators and simple testing methods, fatigue poses considerable difficulties for detection and monitoring.

Furthermore, there are issues associated with the utilization of mobile and touchscreen devices during driving and prolonged use, impacting not just younger individuals but also adults. An effective strategy to address this issue may include pinpointing incidents linked to driver fatigue and requiring drivers to recognize it when necessary. The former entails considerably elevated expenses and introduces more substantial difficulties in implementation. The latter is contingent up on the former, given that the advantages of prolonged driving are considerable. With the rising demand for labor, wages tend to rise as well, resulting in an influx of individuals joining the workforce. The main element driving this issue is clearly financial incentives, which are leading a growing number of drivers to make detrimental

decisions, frequently compromising their own safety and that of others. For example, driving a vehicle during the night while feeling fatigued. This is mainly because drivers frequently do not recognize the serious risks associated with driving while tired. Some countries have regulations limiting the maximum driving hours for operators; however, this approach alone does not sufficiently tackle the problem [2].
In the current era of technological advancement, Machine Learning and Its various algorithms are making a decision making systems automated [3]. The growth of private transportation is accelerating swiftly traveling a long distance by car can often become quite tedious and monotonous. Extended periods of driving without sufficient rest and sleep significantly impact the driver's level of attentiveness. Operating a vehicle while experiencing fatigue is a possibility for exhausted individuals. Drowsiness can lead to severe and potentially fatal incidents that may result in death at any time. Ongoing assessment of the driver's alertness is essential to avert occurrences of this kind, and the driver must be alerted if drowsiness is identified. The CNN-based Yolov8 Model has been utilized for our project. This approach has the potential to greatly reduce the incidence of accidents and preserve lives[4].

II. Literature Review

[5] Explores the effectiveness of physiological indicators in predicting driver drowsiness; these indications include heart rate variability (HRV), blink rate, blink percentage, and electro dermal activity (EDA) signals. Half of the 30 participants drove in a non-monotonous setting and the other half in a monotonous one as part of the controlled simulated driving situations used in the study. For HRV, EDA was trained using a 1D-convolutional neural network (1D-CNN), and eye tracking was accomplished with the help of a convolutional recurrent neural network (CRNN). These three deep learning methods were utilized. Both the HRV-Based Model and the EDA Based Model performed admirably when it came to classifying tiredness.
[6] Create an RMSCM that includes a CNN-Bi-LSTM submodule, a residual structure branch, and a multi-scale convolution sub-module. Highly discriminative lengthy short-term temporal feature information can be effectively extracted using RMSCM. Our second proposal is a method for assigning weights to the relevance of the edges of a spatial graph and a dynamic graph convolution module that use a brain region partitioning strategy to learn the intrinsic spatial features between electrodes. Thirdly, we create a feature fusion module (FFM) that combines spatial and short-term temporal information via channel attention. As it learns about the fused features, FFM ranks the importance and relevance of each channel. To acquire the projected driver fatigue state, the fused spatio-temporal features are then sent into the classification module.
[7] Used a common image dataset that tracked drivers' eye movements to carry out the study. Integrating the best features from Light Gradient-Boosting Machine (LGBM) and the Visual Geometry Group (VGG-16), we presented a new feature creation approach based on transfer learning. In order to provide valuable transfer features, the suggested VGLG (VGG16-LGBM) method uses spatial features extracted from input eye image data.
[8] The DLID3-ADAS method, which is based on deep learning, is an intelligent driver drowsiness detection system. The goal of the DLID3-ADAS method is to improve driver alertness and, by extension, road safety. By utilizing the Shuffle Net method, the DLID3-ADAS technology is able to extract complicated characteristics from images. We also use the Northern Goshawk Optimization (NGO) technique to pick the best hyper parameters for the Shuffle Net model. Last but not least, a model from an extreme learning machine (ELM) is employed to accurately identify and categorize drivers' states of sleepiness.
[9] Using a temporal-frequential attentional convolutional neural network (TFAC-Net) to identify tiredness in drivers using only one channel of electroencephalogram (EEG) data as input. In particular, a concurrent spectral-temporal representation is initially generated using the continuous wavelet

technique in order to extract the possibly useful information from single-channel EEG. The next step is to use the temporal frequential attention mechanism to bring to light crucial time frequency regions related to the driver's mental state. In the end, it is thought that an adaptive feature fusion module might be used to re-calibrate and incorporate the most important feature channels for the prediction.

[10], using a shallow CNN architecture to detect driver sleepiness. Identifying faces and extracting eye regions is accomplished using the 68-point facial landmark identification approach. Utilizes a convolutional neural network (CNN) design that is shallow, meaning it uses fewer layers and parameters, to identify sleepy drivers when training data is scarce. Eyelid closure is one of the most important visual clues for drowsiness detection, and feature extraction aims to extract these signals. It is also possible to detect driver sleepiness using transfer learning models like VGG19, ResNet50, MobileNetV2, and InceptionV3. There were two datasets used: Dataset-1 and Dataset-2.

[11]A convolutional neural network with the best possible structure for detecting driver drowsiness was found using a search technique for network construction. The following are some of the contributions to this study: To start, in order to fill gaps in current datasets and improve our ability to detect driver tiredness, we have taken frames from multiple films that exhibit different stages of sleepiness and labeled them as drowsy or natural. Secondly, a genetic algorithm, namely FER-2013, was employed to get an ideal structure for a convolutional neural network, which encompasses several aspects such as the number of layers, kind of objective function, and more. In the third step, transfer learning is used. Here, the genetic algorithm's optimized network is trained on the sleepy dataset, and it is regarded as a feature extraction component.

[12], identifying lethargy by utilizing EEG and ECG data in conjunction with a fuzzy neural network and 2D convolutional neural networks. After convolutional neural networks were employed to extract features from the EEG and ECG signals, a fuzzy neural network was employed for feature classification and drowsiness detection.

[13], A method for detecting driver drowsiness based on video analysis of facial features, including local facial regions and landmarks (VBFLLFA), is suggested in this research. In order to make the most of the important sleep-related facial features while avoiding redundant data, we get the head movement data from facial landmark analysis and the eye and mouth movement data from local facial areas. On top of that, spatial filtering utilizing the common spatial pattern (CSP) technique is implemented to enhance sample class discrimination. The temporal and spatial information can be effectively extracted using a two-branch multi-head attention (TB-MHA) module.

[14], an MSGPT deep learning model, which is capable of autonomously detecting driver exhaustion from start to finish. In order to facilitate the interaction between global and local features, we first build an intra-inter-scale cascade framework using Transformer and MC-Patch Embed. A global prompt token is then included into the framework. Second, a mixed token is designed by aggregating the output from the intra-scale, which comprises rich low-level feature information for multiscale, in order to efficiently integrate intra-scale and inter-scale feature information. In addition, multi-head self-attention (MSA) incorporates a new learnable query to bring the computational complexity down to a linear level. When it comes to the classification evaluation metrics of EEG based fatigue driving, MsGPT significantly outperforms other methods. This is demonstrated in experiments run on the SEED-VIG and SADT datasets with both intra- and inter subject settings.

III. Proposed Work

This study implements deep learning based AI model to detect driver drowsiness using kaggle dataset. The proposed work uses Yolov8 pre-trained model and it is being trained by custom data. The weights are transferred from trained model to predict unknown data. The proposed system architecture is shown below in the figure 1.

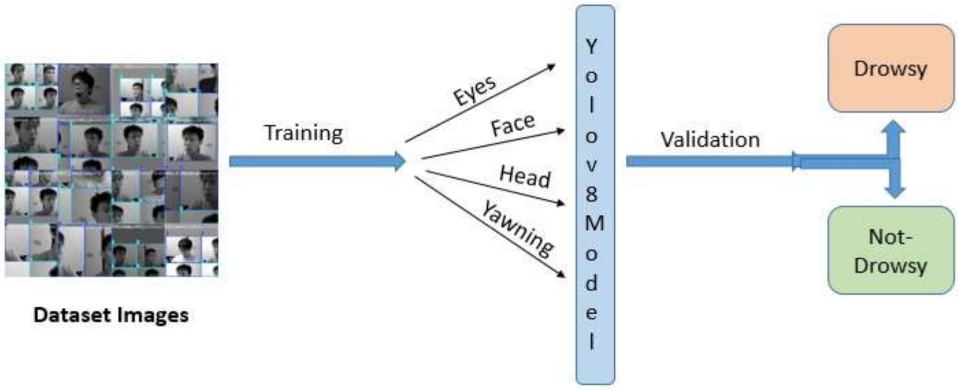

Figure: 1. Proposed System Architecture.

In this methodology pretrained yolo model updated with new weights which has been calculated from custom data through transfer learning. The transfer learning is a powerful machine learning (ML) technique that allows quickly to retrain a model on new data without having to restart the network. By freezing the layers with the weights of the initial layers and only altering the parameters of subsequent layers, transfer learning significantly cuts down on the time and computational resources required for training.

Benefits of transfer learning

•Training time reduction
•Resource Efficiency
•Generalization of model
•Improve in accuracy

A. Data set

Eleventeen participants over five conditions (BareFace, Glasses, Night-BareFace, Night-Glasses, and Sunglasses) make up the training dataset. There is a one-minute recording of each subject's sequence, which includes yawning and slow blink rate with nodding. There are two main scenarios that are recorded for approximately 1.5 minutes each. One scenario involves a mixture of symptoms associated to sleepiness (yawning, nodding, slow blink rate), while the other involves a combination of actions unrelated to sleepiness (talking, laughing, glancing to both sides). A total of ninety driving movies filmed by the remaining eighteen individuals are included in the testing and assessment datasets; these videos feature sleepy and non-drowsy drivers in a variety of situations [15].

B. Data Preparation

Here dataset utilized consisting of image data, from which individual frames were extracted and categorized into two classes: 'drowsy' and 'non-drowsy'. A data preparation strategy was utilized for DL-based object identification models like YOLO. Bounding boxes surrounding important areas of interest, such the eyes and other facial characteristics, were painstakingly designated on the frames. These comments were formatted according to the model in question; for example, YOLO-based models used the YOLO format. By using this labeling procedure, we made sure that the models could learn to recognize the right patterns and traits of sleepiness.

C. Yolov8 Model

In computer vision, YOLO (You Only Look Once) refers to a common collection of object detection models utilized for object classification and detection in real-time. YOLO's primary characteristic is its single-stage detection method, which is intended to identify objects accurately and in real time. YOLO processes the entire image in a single pass, which makes it faster and more effective than two-stage detection models like R-CNN, which initially suggest regions of interest before classifying these regions [15].

•YOLOv8 has better accuracy than previous YOLO models.
•Training of YOLOv8 will be probably faster than the other two-stage object detection models.
•Overall, YOLOv8's high accuracy and performance make it a strong contender for your next computer vision project.

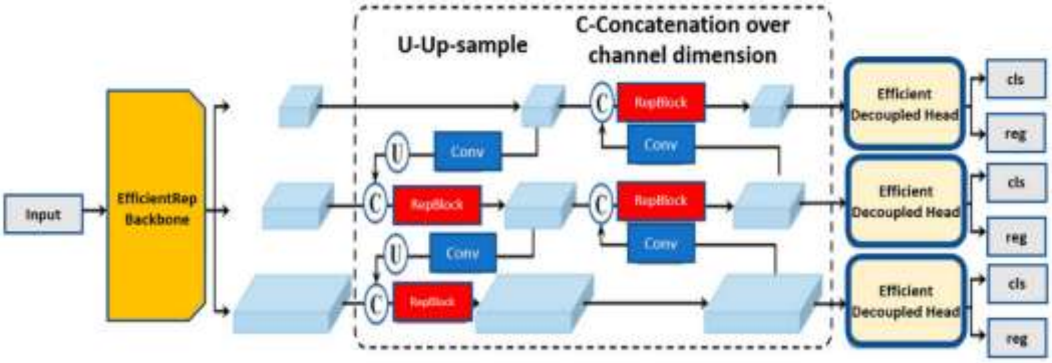

Figure: 2. Yolov8 Architecture.

D. Model Algorithm

Here, Model algorithmic steps are defined in algorithm 1.

Algorithm 1 Proposed Model Steps

1: Procedure
Step 1: Prepare the Dataset
Step 2: Load the pre-trained Yolov8 model
Step 3: Transfer the weights from pre-trained model
Step 4: Train the Yolov8 with custom dataset
Step 5: Validate the trained model
2: End Procedure

```
# Load a model
from u l t r a l y t i c s import YOLO
# b u i l d a new model f rom YAML
model = YOLO( " yolo8n . yaml " )
# load a p r e t r a i n e d model
model = YOLO( " yolo8n . p t " )
# b u i l d f rom YAML and t r a n s f e r we i g h t s
model = YOLO( " yolo8n . yaml " ) .
l o a d ( " yolo8n . p t " )
# Tr a i n t h e model
r e s u l t s = model . t r a i n ( d a t a =" cus tom . yaml " , epochs =100 , imgsz =640)
```

The dataset and model setup for training detection models are defined by the Ultralytics framework using a YAML file format. An illustration of the YAML format needed to define a detection dataset is provided here:

```
p a t h : / { y o u r l o c a l p a t h } / d a t a
t r a i n : images / t r a i n
v a l : images / v a l
t e s t : images / t e s t
# Cl a s s e s
nc : 2
```

names: [’ drowsy ’ , ’Not Drowsy ’]

E. Implementation

The proposed model is implemented using Python language and run on GPU in google colab.

•Environment selected is cloud based
•Hardware Requirement is NVIDIA T4
•Dataset is divided into valid, train and test folders
•Each folder contains images and labels folders
•The image folder contains all images
•The label folder contains the data label

F. Performance Metrics

1) Mean Average Precision (mAP): It measures the average precision (AP) scores across all object classes and summarizes the model’s overall ability to correctly identify objects of different classes. The mAP ideal value of good model should be above 50%.

2) F1 Score Curve (F1 curve.png): This curve represents the F1 score across various thresholds. Interpreting this curve can offer insights into the model’s balance between false positives and false negatives over different thresholds.

$F1 = 2 *[(Precision * Recall)/(Precision + Recall)]$

3) Confusion Matrix (confusion matrix.png): The confusion matrix provides a detailed view of the outcomes, showcasing the counts of true positives, true negatives, false positives, and false negatives for each class.

	Predicted Positive	Predicted Negative
Actual Positive	TP	FP
Actual Negative	FN	TN

Table 1 confusion matrix

•TP= True Positive
•FP = False Positive
•FN = False Negative
•TN = True Negative

IV. Experimental Analysis

The trained model is validated on nearly 2500 images and the performance is measured with respect accuracy, f1 score. The confusion matrix is derived to visualize the count of actually predicted and not predicted. The various performance diagram are shown as below.

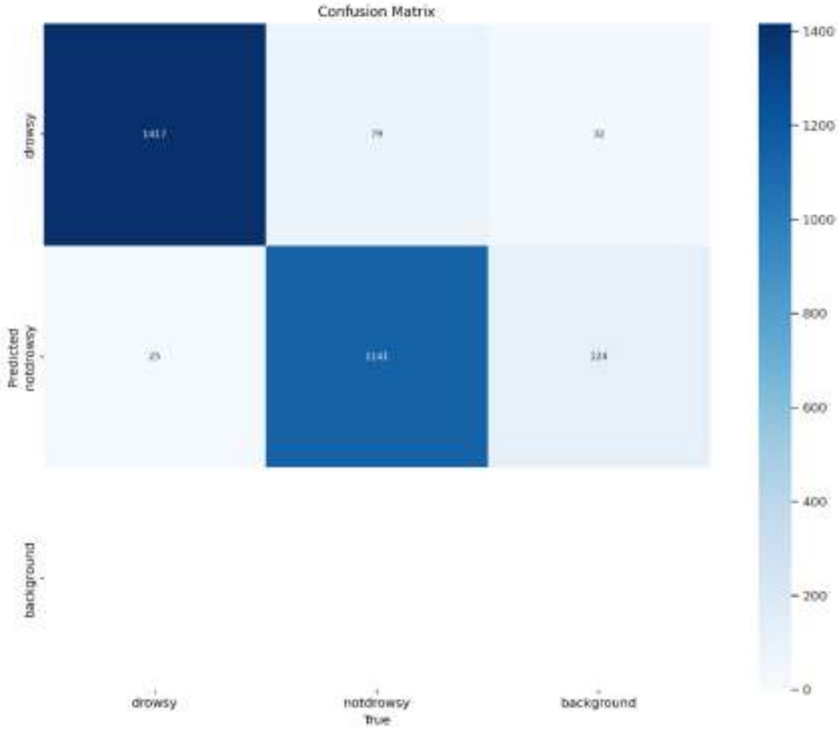

Figure: 3. Confusion Matrix

The confusion matrix in the Figure No. 3 shows that totally 1417 images are predicted drowsy accurately and 1141 images not drowsy accurately, so nearly we got 96.85% of accuracy.

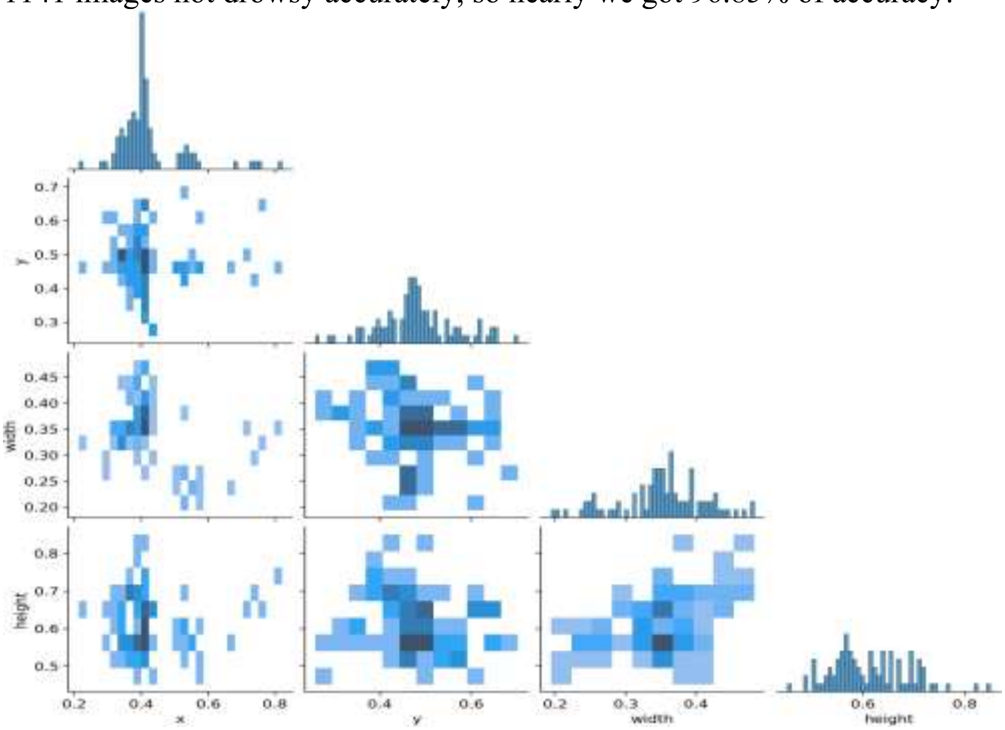

Figure: 4. Labels Correlogram

Figure No. 4 gives insights about labels correlation. the labels are in xywh space. In the trained model it shows positive correlation between labels.

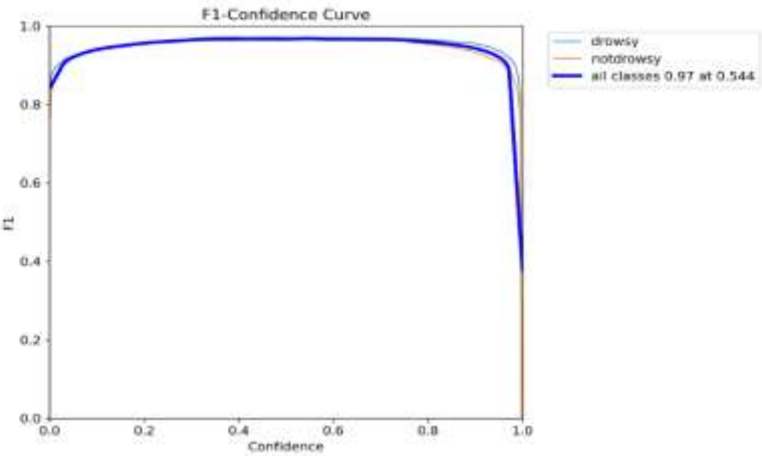

Figure: 5. F1-Score

The F1 Score in the Figure No. 5 shows nearly 97% score is attained for both drowsy and not drowsy images.

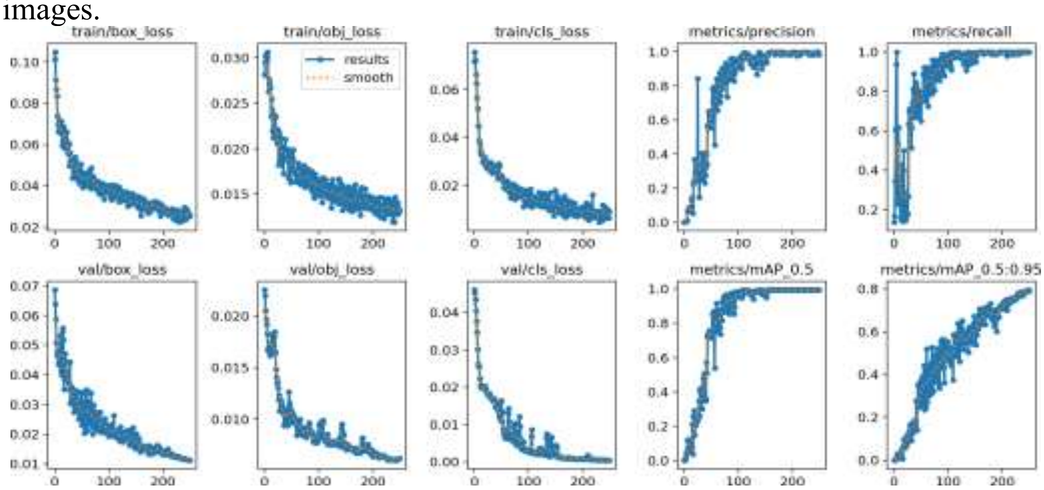

Figure: 6. Precision,Recall and mAP

The performance graphs of precision and recall are shown in the Figure No. 6. The performance for training and validation data is measured. The graphs shows mean average precision(mAP),precision and recall during training vs validation.

Figure No. 7 Validation Output

The Figure No. 7 shows the images validated from the trained model.

V. Conclusion

A CNN-yolov8-based driver drowsiness detection system is a promising technology that could increase road safety by warning drivers when they are becoming distracted or drowsy. The technology analyzes the driver's face and eyes to look for sleepiness indicators like yawning and drooping eyelids. In order to reduce the number of fatalities brought on by tiredness and prevent traffic accidents, drowsiness detecting systems can be installed in all automobiles. The proposed study introduced a unique method for detecting driver tiredness by utilizing image processing through a Convolutional Neural Network (CNN) and the YOLOv8 algorithm. The outcomes show how successful YOLOv8 is in real-time situations and showcase its improved detection skills. This study shows how AI has advanced significantly in the field of public safety and emphasizes how important cutting-edge deep learning algorithms are to reducing the dangers of fatigued and dis- tracted driving. Finally proposed trained model achieved 80% of mAP(mean Average Precision) and 97.5% of F1-score.

VI. Future scope

The Driver Drowsiness Detection System has a great deal of room to grow and advance as technology advances. As AI and machine learning grow more popular, future systems can leverage larger datasets and more reliable algorithms to achieve better accuracy and efficiency. Integration with advanced driver-assistance systems and autonomous vehicles offers opportunities to increase safety by enabling cars to recognize tiredness and take appropriate action, such as slowing down or stopping. Enhancements in edge computing could benefit the system by enabling low-latency real-time processing on embedded devices, even at remote places. IoT network integration may provide centralized analytics and monitoring to improve traffic safety and urban mobility as smart cities proliferate. The system's usability and effectiveness may also be enhanced by customization choices like adaptive thresholding that are dependent on driving patterns and user behavior.

References

[1]WHO, "Road traffic injuries." https://www.who.int/news-room/factsheets/ detail/road-traffic-injuries, 2023.
[2]A. K. Jain, V. Sharma, S. Goel, R. G. Tiwari, A. Vajpayee, and R. Bhandari, "Driver drowsiness detection using deep learning," in 2023 3rd International Conference on Intelligent Technologies (CONIT), pp. 1–6, IEEE, 2023.
[3]J. Pujari and A. Hake, "Comparative study of various machine learning techniques for diabetes prediction," in 2025 International Conference on Artificial Intelligence and Data Engineering (AIDE), pp. 793–798, IEEE, 2025.
[4]D. Ojha, A. Pawar, G. Kasliwal, R. Raut, and A. Devkar, "Driver drowsiness detection using deep learning," in 2023 4th International Conference for Emerging Technology (INCET), pp. 1–4, IEEE, 2023.
[5]J. Alguindigue, A. Singh, A. Narayan, and S. Samuel, "Biosignals monitoring for driver drowsiness detection using deep neural networks," IEEE Access, vol. 12, pp. 93075–93086, 2024.
[6]F. Hu, L. Zhang, X. Yang, and W.-A. Zhang, "Eeg-based driver fatigue detection using spatio-temporal fusion network with brain region partitioning strategy," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 8, pp. 9618–9630, 2024.
[7]H. A. Madni, A. Raza, R. Sehar, N. Thalji, and L. Abualigah, "Novel transfer learning approach for driver drowsiness detection using eye movement behavior," IEEE Access, vol. 12, pp. 64765–64778, 2024.

[8]E. Yang and O. Yi, "Enhancing road safety: Deep learning-based intelligent driver drowsiness detection for advanced driver-assistance systems," Electronics, vol. 13, no. 4, p. 708, 2024.
[9]P. Gong, P. Wang, Y. Zhou, X. Wen, and D. Zhang, "Tfac-net: A temporal-frequential attentional convolutional network for driver drowsiness recognition with single-channel eeg," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 7, pp. 7004–7016, 2024.
[10] M. Venkateswarlu and V. Rami Reddy Ch, "Drowsydetectnet: Driver drowsiness detection using lightweight cnn with limited training data," IEEE Access, vol. 12, pp. 110476–110491, 2024.
[11] Y. Jebraeily, Y. Sharafi, and M. Teshnehlab, "Driver drowsiness detection based on convolutional neural network architecture optimization using genetic algorithm," IEEE Access, vol. 12, pp. 45709–45726, 2024.
[12] G. Yogarajan, R. N. Singh, S. A. Nandhu, and R. M. Rudhran, "Drowsi- ness detection system using deep learning based data fusion approach," Multimedia Tools and Applications, vol. 83, no. 12, pp. 36081–36095, 2024.
[13] L. Yang, H. Yang, H. Wei, Z. Hu, and C. Lv, "Video-based driver drowsiness detection with optimised utilization of key facial features," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 7,pp. 6938–6950, 2024.
[14] P. Zhao, C. Lian, B. Xu, and Z. Zeng, "Multiscale global prompt transformer for eeg-based driver fatigue recognition," IEEE Transactions on Automation Science and Engineering, vol. 22, pp. 270 2711, 2025.
[15] NTHU, "Driver drowsiness dataset." https://www.kaggle.com/datasets/banudeep/nthuddd2, 2016.