Grenze International Journal of Engineering and Technology, January Issue

Smart Monitoring System for Driver Alertness

Smita Bhagwat[1], Siddhesh Chaudhari[2], Atharv Dhole[3], Arya Deshmukh[4], Om Chavan[5], and Shishir Chandanshiv[6]
[1-7]Vishwakarma Institute of Technology, Pune, India
Email: smita.bhagwat@vit.edu, siddhesh.chaudhari241@vit.edu, atharv.dhole241@vit.edu, arya.deshmukh241@vit.edu, shishir.chandanshiv24@vit.edu, om.chavan24@vit.edu

***Abstract*— Driver fatigue remains a major contributor to road accidents worldwide, resulting in significant losses in both human lives and economic terms. Traditional detection methods based on physiological signals or behavioral cues often fall short of reliability under diverse real-world conditions.**
In this work, we present a non-intrusive driver drowsiness detection system that integrates deep learning with advanced computer vision techniques. A ResNet50V2-based model, fine-tuned using the Driver Drowsiness Detection (DDD) dataset, is employed to monitor key facial features such as head orientation, eyelid closure, and yawning frequency for accurate driver state classification. The proposed system achieves an overall detection accuracy of 97% by leveraging data augmentation, transfer learning, and regularization strategies. With its high efficiency and robustness, the system is well-suited for real-time deployment in next-generation in-vehicle safety applications.

***Index Terms*— driver drowsiness detection, non-intrusive, behavioral analysis, deep learning, ResNet50V2, real-time monitoring.**

I. Introduction

Driver drowsiness and fatigue are significant contributors to road accidents, with drowsiness-induced inattention being a major cause, according to the NHTSA [1]. The objective of this project is to create a non-intrusive driver drowsiness detection system based on deep learning. Utilizing Convolutional Neural Networks (CNNs), i.e., the ResNet50V2 model, we categorize drivers as "alert" and "drowsy" classes. The approach is for everyday driving conditions and has potential in personal vehicle safety, commercial fleet management, and autonomous driving systems.
In contrast to conventional physiological and vehicle-based approaches, which are invasive or unreliable, computer vision-based behavioural approaches offer a feasible solution.
Deep learning-based methods, especially CNNs, have made great progress in computer vision applications.[7] ResNet50V2, with better pre-activation and batch normalization, provides better performance in the task of face recognition. By fine-tuning this model on our DDD dataset, we are making it more generalizable to real-world scenarios.
This project helps in making roads safer through an efficient driver monitoring system minimizing drowsy driving accidents.[13] The rest of this paper outlines current literature, methodology, experimental results, conclusions, and areas for future research.

Grenze ID: 01.GIJET.12.1.46_1

II. Literature Review

Driver drowsiness detection has been a significant research area for over a decade, with initial approaches relying on manually engineered features such as eye aspect ratio, blink frequency, and yawning detection [2]. While these early methods established foundational benchmarks, they often lacked robustness under varying lighting conditions and occlusion[9]. Sensor-based techniques, such as EEG and ECG monitoring, were explored but proved to be intrusive and impractical for real-world deployment[3]. A Real-Time Driver Drowsiness Detection System Using Visual Information was proposed to overcome the limitations of traditional sensor-based systems [12].

The advancement of deep learning has shifted the focus toward automatic feature extraction from raw image data enabling the development of non-intrusive driver monitoring systems. Convolutional Neural Networks (CNNs) like VGGNet, Inception, and ResNet have been applied to driver drowsiness detection with promising results[4]. Among these, ResNet architectures are particularly effective due to their residual learning mechanism, which enables deep networks to learn efficiently by addressing the vanishing gradient problem. The ResNet50V2 model, an enhanced variant, introduces pre-activation and improved batch normalization, which enhances convergence and classification performance[5].

Despite these advancements, several challenges remain in developing a system that performs reliably across diverse real-world driving scenarios[8].

Most existing models rely on small datasets and lack generalizability. The enhanced ResNet50V2 model remains underutilized in this domain. Many studies also do not incorporate sufficient data augmentation and regularization.[10]. Our research bridges these gaps by fine-tuning ResNet50V2 on the DDD dataset with robust augmentation techniques for improved real-world performance.

III. Methodology: Proposed System Overview

This section provides a full description of the proposed methodology, which comprises some important components, including dataset preparation, model structure, training process, and evaluation. All these components are explained in considerable detail in the subsequent subsections.

A. Preprocessing

Both the YAWDD dataset and the Driver Drowsiness Detection (DDD) dataset comprise thousands of images of faces collected under various conditions such as fluctuating lighting, angles, and even occlusions such as sunglasses. A clean preprocessing step needs to be adopted to bring the model to operation [15]. The implemented preprocessing pipeline for this research paper is as follows:

Rescaling and Normalization: All the images are rescaled before training so that the pixel intensity values will range in the interval [0,1]. Normalization is then done to enhance learning stabilization and faster and more reliable convergence in model optimization.

Image Resizing: All images are resized to 224 by 224-pixel dimensions to fulfill the input dimensions requirement of the ResNet50V2 network architecture.

Data Augmentation: Heavy data augmentation is utilized to increase model robustness to real-world data variances, and to mitigate overfitting. This includes transformations such as rotation, translation, scaling, shearing, zooming, and horizontal flipping. All augmentations introduce variances in the training dataset, thereby increasing the generalized out-of-sample performance of the model on non-processed datasets.

B. Model Architecture

This research paper employs a design based on ResNet50V2's architecture that has been pre-trained on ImageNet. We have altered the architecture for binary classification (alert vs drowsy) in the following ways:

Base Network: The pre-trained ResNet50V2 model was used as the feature extractor.[7] We did not keep the original classification head, but we kept the convolutional base.

$$X \in R^{H \times W \times C} \quad \text{----(1)}$$

In (1)dimensions of a multi-channel image or tensor

Global Average Pooling (GAP): A GAP layer is added to the base model output. For a feature map a GAP layer compute:

$$y_k = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{i,j,k} \quad \text{----(2)}$$

In (2) the GAP layer simplifies the spatial dimensions, helping combine the feature maps into a single vector per channel.
Dropout and Batch Normalization: A 50% dropout layer is applied to combat overfitting while training by randomly deactivating neurons during training. Then a batch normalization layer is applied to normalize the output layer of the earlier layer to stabilize the learning.
Dense Layers with L2 Regularization: The model adds one dense layer of 256 neurons that uses
ReLU activation and L2 regularization.[11] L2 regularization adds a term to the loss which is:

$$Loss_{total} = Loss_{data} + \lambda||W||_2^2 \quad \text{----(3)}$$

In (3) L2 regularization
To wrap up, we add a dense layer with one neuron and a sigmoid activation function, which gives the probability of the drowsy state.

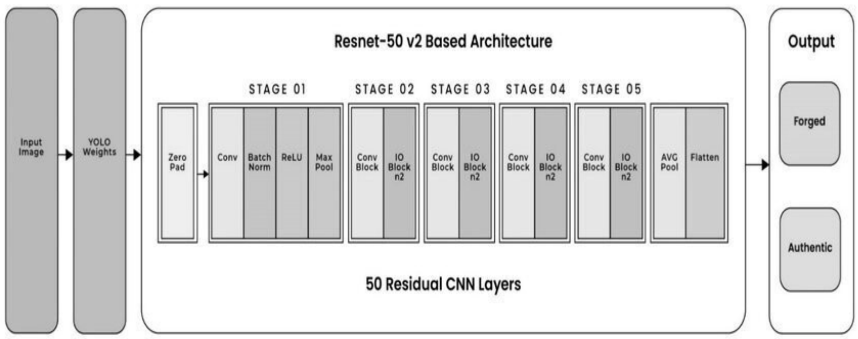

Fig 1: ResnetV2 Architecture

Layer Freezing for Transfer Learning: To take full advantage of the learned representations from ImageNet, we freeze the majority of the base model's layers and allow only the last 40 layers to be trainable[14]. This reduces the chances of overfitting and speeds up convergence.

C. Training Procedure

The model is trained with the Adam optimizer, a learning rate of 1×10−41 \times 10^{-4}1×10−4, and binary cross-entropy as the loss function. The model is trained for a number epochs, with early stopping and learning rate decay to limit the risk of overfitting. The data is split into three portions; 70–80% of the data for training, 10–15% of the data for validation, and 10–15% of the data for testing. Our pseudocode (in a high-level format) of the overall training process is shown in the Figure 2.

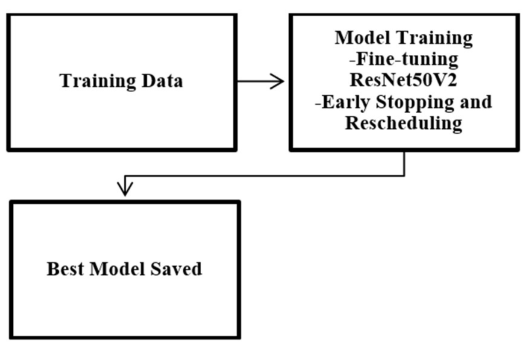

Fig 2: Model Training Procedure

D. Mathematical Foundations:

The success of ResNet50V2 is attributed to its residual learning framework. Each residual block computes an output as:

$$y = F(x, \{W_i\}) + x \quad \text{----(4)}$$

$F(x, \{W_i\})$

In (4) Residual Learning Function
where denotes the residual mapping. In ResNet50V2, a pre-activation variant is used—where batch normalization and ReLU activation precede the convolution operation—resulting in improved gradient flow during backpropagation.
The Adam optimizer used during training is defined by the following equations:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad \text{----(6)}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad \text{----(7)}$$

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad \text{----(8)}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad \text{----(9)}$$

$$\theta_{t+1} = \theta_t - \alpha \frac{\widehat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad \text{-----(10)}$$

In (10) Adam (Adaptive Moment Estimation) Optimization Algorithms

IV. Results and Discussion

The performance of the developed model was evaluated using a set of standard metrics.
Accuracy: measures the proportion of total predictions that were correct.
Precision and Recall: focuses specifically on how well the model identifies drowsy drivers while minimizing incorrect detections.
F1-Score: provides a single score that balances precision and recall, giving a clearer picture of overall performance.
Specificity: Specificity evaluates the model's ability to correctly recognize drivers who are alert,
ensuring that normal driving states are not mistakenly flagged.
All experiments were conducted on a workstation powered by an NVIDIA Tesla V100 GPU, using TensorFlow 2.0 and CUDA 11.2 for the software environment. The DDD and YAWDD datasets were split into three groups: 80% for training, and 10–15% each for validation and testing. The model was trained over thirty epochs, with early stopping and learning rate adjustment strategies used to prevent overfitting and help the model converge more effectively.
Table I shows the results obtained on the test set. As seen from the metrics, the model performed consistently well across different evaluation measures.

TABLE I

Metric	Value
Training Accuracy	99.73%
Validation Accuracy	98.04%
Precision	93.70%
Recall	92.80%
F1-Score	91.21%

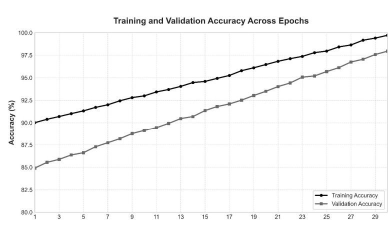

Fig 3: Training And Validation Accuracy Graph

V. Future Scope

Real-Time Integration: Enhancing the implementation framework of embedded automotive systems.
Multi-Modal Fusion: Utilization of other information sources (steering behaviour, vehicle telemetry) to support system overall reliability improvement.

Dataset Augmentation: Gathering additional diverse real-world data to increase the generalizability and decrease bias of the model.
Edge Computing Optimization: A thorough review of light-weight network structures aimed at enabling deployment in energy-efficient environments.

VI. Conclusion

In this paper, we have demonstrated an end-to-end driver drowsiness detection system utilizing the ResNet50V2 architecture. Utilizing transfer learning, heavy data augmentation, and state-of-the-art regularization methods, methodology of this research paper is to achieves elevated accuracy as well as robustness in driver state classification with the help of the DDD dataset. The in-depth experimental analysis vindicates the viability of this system for practical utilization in driver monitoring and safety.

Acknowledgement

We would also like to express gratitude to the participants, whose effort was crucial for the successful outcome of this research on the Driver Drowsiness Detection System. Further, we would like to express gratitude to our project supervisor, Prof. Smita Bhagwat, for her continuous support and guidance throughout the course of this project. We also thank the Department of Engineering, Sciences, and Humanities (DESH), Vishwakarma Institute of Technology, Pune, for having provided the resources and infrastructure needed for completion of this project. Lastly, we would also like to mention the researchers and inventors whose earlier work has been both our source of inspiration and useful reference for this present study.

References

[1] Y. Albadawi, A. Al-Ayoub, and M. A. S. Kamal, "Real-Time Machine Learning-Based Driver Drowsiness Detection Using Visual Features," Journal of Imaging, vol. 9, no. 5, p. 89, May 2023.
[2] S. E. Klauer, T. A. Dingus, V. L. Neale, J. D. Sudweeks, and D. J. Ramsey, "The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data," National Highway Traffic Safety Administration, Report No. DOT HS 810 594, 2006.
[3] M. Sahayadhas, R. Sundaraj, and S. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," Sensors, vol. 12, no. 12, pp. 16937–16953, 2012.
[4] F. Safarov, F. Sultanov, S. Hoey, M. H. Khan, and T. Whangbo, "Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety," Sensors, vol. 23, no. 14, p. 6343, Jul. 2023.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Identity Mappings in Deep Residual Networks," in Proc. Eur. Conf. Comput. Vis. (ECCV), Amsterdam, Netherlands, Oct. 2016, pp. 630–645.
[6] H. Alabdulkarem, S. Cherkaoui, E. Pérez-López, and C. Gonzalez, "A Deep-Learning Approach to Driver Drowsiness Detection," Safety, vol. 9, no. 3, p. 63, Sep. 2023.
[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.
[8] Y. Albadawi and M. A. Awad, "A Review of Recent Developments in Driver Drowsiness Detection Systems," Sensors, vol. 22, no. 5, p. 1781, 2022.
[9] S. A. Khan, S. J. Gilani, S. Ahmad, and H. M. K. Qureshi, "Driver Drowsiness Detection System Using Visual Features: A Survey," IEEE Access, vol. 8, pp. 90735–90745, 2020.
[10] A. Kaur and S. Kaur, "Driver Drowsiness Detection Using Deep Learning Techniques: A Survey," in Proc. Int. Conf. Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2019, pp. 1–6.
[11] R. Flórez, F. Palomino-Quispe, R. J. Coaquira-Castillo, and J. C. Herrera-Levano, "A CNN-Based Approach for Driver Drowsiness Detection by Real-Time Eye State Identification," Appl. Sci., vol. 13, no. 13, 2023.
[12] R. Sahayadhas and S. Murugappan, "A Real-Time Driver Drowsiness Detection System Using Visual Information," in Proc. IEEE Int. Conf. Intelligent Transportation Systems (ITSC), 2010.
[13] J. Léger and J. C. A. Pépin, "Economic and Societal Impact of Sleepiness: A Major Public Health Problem," Sleep Med. Clin., vol. 14, no. 2, pp. 221–232, 2019.
[14] F. T. Negassi, I. Yassin, A. Bashir, M. Awawdeh, and T. Faisal, "Systematic Development of Real-Time Driver Drowsiness Detection System Using Deep Learning," IAES Int. J. Artif. Intell., vol. 11, no. 1, pp. 148–155, Mar. 2022.
[15] DDD Dataset, "Driver Drowsiness Detection Dataset," Kaggle, 2018. [Online]. Available: https://www.kaggle.com/andrewmvd/driver-drowsiness-detection