Driver Drowsiness Detection Using DenseNet121

Bhagyashree Nyamagoudar
Computer Science and Engineering,
KLE Technological University,
Hubballi, India
Email: 01fe22bcs299@kletech.ac.in

Sakshi Dixit
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: 01fe22bcs251@kletech.ac.in

Nikhita Mane
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: 01fe22bcs284@kletech.ac.in

Sneha S Ratnakar
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: 01fe22bcs289@kletech.ac.in

Nivedita Gudi
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: 01fe22bcs303@kletech.ac.in

Padmashree Desai
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: padmashri@kletech.ac.in

Lalita Madanbhavi
Computer Science and Engineering
KLE Technological University,
Hubballi, India
Email: lalita@kletech.ac.in

***Abstract*—Driver drowsiness detection systems play a vital role in reducing road accidents caused by fatigue and inattentive driving. This project leverages advanced computer vision and deep learning techniques to accurately identify signs of driver fatigue, such as blinking rates, yawning, and head postures. Utilizing a DenseNet121 architecture, the system categorizes images into drowsy and non-drowsy classes with high accuracy. The methodology includes data preprocessing, augmentation, and model training using the State Farm Distracted Driver Dataset. Evaluation results demonstrate the model's reliability, achieving an accuracy of over 99% during training, validation, and testing. This robust framework not only supports real-time monitoring but also lays the groundwork for integrating multi-sensor data to enhance detection capabilities further. The proposed system offers a promising solution for improving road safety and mitigating accident risks globally.**

***Index Terms*—Driver drowsiness detection, DenseNet121, deep learning, computer vision, distracted driving, State Farm Dataset.**

I. Introduction

Driver fatigue and abnormal driving behaviors are among the leading causes of road accidents globally, posing significant risks to public safety. According to the World Health Organization, over 1.25 million fatalities occur annually due to vehicular accidents, with countless others suffering severe injuries. Many of these incidents are exacerbated by factors such as driver drowsiness, distraction, and reckless actions [1], [2]. These human-induced errors represent critical challenges in traffic safety, particularly in regions with inadequate traffic management systems, such as Iraq, where road accident rates remain disproportionately high. [3] The combination of weak enforcement of traffic laws, poor road infrastructure, and insufficient driver monitoring systems has further compounded the issue in such regions. Beyond the immediate loss of life and injuries, the economic and social repercussions of road accidents are profound, including increased healthcare costs, loss of productivity, and emotional trauma for families and communities. [4] These factors underscore the urgency of implementing innovative and effective preventive measures to reduce the risk of accidents caused by driver fatigue and distraction. [5]

Facial fatigue detection systems have emerged as a promising technological solution to address driver drowsiness and distraction. These systems rely on analyzing facial expressions and behaviors, such as blinking rates, yawning, eye closure duration, and head postures, which serve as reliable indicators of a driver's state of alertness [1]. Advances in computer vision, particularly deep learning-based methods, have significantly enhanced the accuracy and reliability of these systems. For instance, a modified DenseNet architecture has demonstrated exceptional performance in categorizing distracted driving behaviors, achieving a validation accuracy of 99.80% on benchmark datasets [1]. This demonstrates the potential of advanced neural network architectures in extracting and learning intricate patterns from visual data. Similarly, systems utilizing Eye Aspect Ratio (EAR) measurements have been proven effective in detecting fatigue by monitoring eyelid movements. These systems employ threshold-based decision mechanisms to provide real-time alerts when signs of fatigue are detected [2]. Such methodologies offer robust solutions to detecting driver fatigue even under challenging conditions, such as varying lighting or partial occlusion of facial features.

In addition to purely visual approaches, hybrid systems that integrate multiple modalities have shown significant promise in enhancing the robustness and reliability of fatigue detection systems. By combining visual data with physiological signals such as heart rate variability, skin conductance, and respiration rate, these hybrid systems capture a more comprehensive range of fatigue indicators. [6] This multimodal approach mitigates the limitations of single-modality systems, such as facial occlusion or noise in physiological measurements. For example, physiological data can complement visual analysis when a driver's face is partially obscured by sunglasses or reflections. Moreover, hybrid systems can adapt to varying conditions, making them suitable for real-world scenarios where environmental and driver-specific factors may vary significantly. [7]

Building on these advancements, this study proposes an integrative framework for driver monitoring that leverages state-of-the-art techniques, including DenseNet-based anomaly detection, EAR-based fatigue analysis, and advanced computer vision applications. By utilizing datasets such as the State Farm Distracted Driver Dataset, the framework is designed to achieve three primary objectives: enhancing the accuracy of fatigue and distraction detection, minimizing computational overhead for efficient real-time processing, and ensuring practical applicability in modern vehicle systems [1], [2]. This approach aims to address the growing demand for intelligent driver-assistance systems that can preemptively identify and mitigate risks before accidents occur.

The proposed framework also sets the foundation for future extensions to driver monitoring systems. Potential enhancements include multi-sensor fusion to integrate additional data sources, such as vehicular telemetry and environmental context. [8] Context-aware fatigue detection systems, which consider external factors such as traffic density, road type, and time of day, could further refine the accuracy of predictions. [9]Additionally, adaptive models that personalize fatigue thresholds based on individual driver profiles, behavioral patterns, and physiological characteristics could make the systems more precise and effective. [10] By incorporating these advanced capabilities, the proposed framework aims to contribute to a safer driving environment and reduce the global burden of road accidents.

II. Related Work

Driver fatigue and abnormal driving behaviors have been a significant area of research lately as a way to curb the alarming statistics of road accidents caused by human factors. [2]Some approaches based on computer vision, machine learning, and deep learning techniques have been proposed to solve this problem.

A. Facial Expression-Based Detection

Facial expressions such as blinking rates, yawning, and head posture can be very reliable indicators of driver fatigue. Systems using EAR calculations have been found to be effective in drowsiness detection. [11] Different EAR-based methods monitor the opening and closing of eyelids to determine the thresholds of fatigue. Such methods provide alerts to the driver in real-time. For example, a computer vision-based system for the detection of fatigue based on facial parameters with high accuracy using Python and OpenCV is proposed by Balasundaram et al. [**?**].

B. Deep Learning Models

Deep learning has played a vital role in the development of driver monitoring systems. DenseNet architectures, in particular, have been highly performing in the detection of distracted driving behaviors. [1] Aisha Ayad and Abdulmunim proposed a variation of the DenseNet model to detect anomalies. The validation accuracy of the modified model was found to be 99.80% on the State Farm Distracted Driver Dataset. Their technique involved preprocessing the images, enhancing them, and reducing their dimensions to optimize the efficiency of the model [**?**].

C. Multi-Modal and Context-Aware Techniques

Even though visual-based approaches are highly effective, hybrid systems that combine additional data modalities, such as heart rate variability or EEG signals, have been explored to enhance the robustness of detection. [3] These multimodal systems offer complementary insights that overcome challenges in single-modality approaches, such as occlusion or noisy data. Such methods point to the potential for combining physiological signals with facial data for comprehensive driver monitoring [**?**], [**?**].

D. Evaluation of Existing Techniques

Several research studies have benchmarked driver monitoring systems against datasets such as the State Farm Distracted Driver Dataset, with an emphasis on scalability and real-time performance. [12]Traditional CNN models are effective but often suffer from computational overhead and scalability issues. Modifications of DenseNet proposed by Ayad et al. address these challenges, finding a balance between accuracy and efficiency [**?**].

III. Architecture

DenseNet-121 is a specific variant of DenseNet, organized into dense blocks and transition layers. The architecture is detailed as follows:

1) **Initial Layers:**
   - A 7x7 convolution with 64 filters.
   - A 3x3 max pooling layer to downsample.
2) **Dense Block 1:**
   - Contains 6 densely connected layers, where each layer outputs feature maps added to the preceding ones.
3) **Transition Layer 1:**
   - A 1x1 convolution followed by a 2x2 average pooling to reduce the feature maps and spatial dimensions.
4) **Dense Block 2:**
   - Contains 12 densely connected layers.
5) **Transition Layer 2:**
   - A 1x1 convolution and 2x2 average pooling.
6) **Dense Block 3:**
   - Contains 24 densely connected layers.

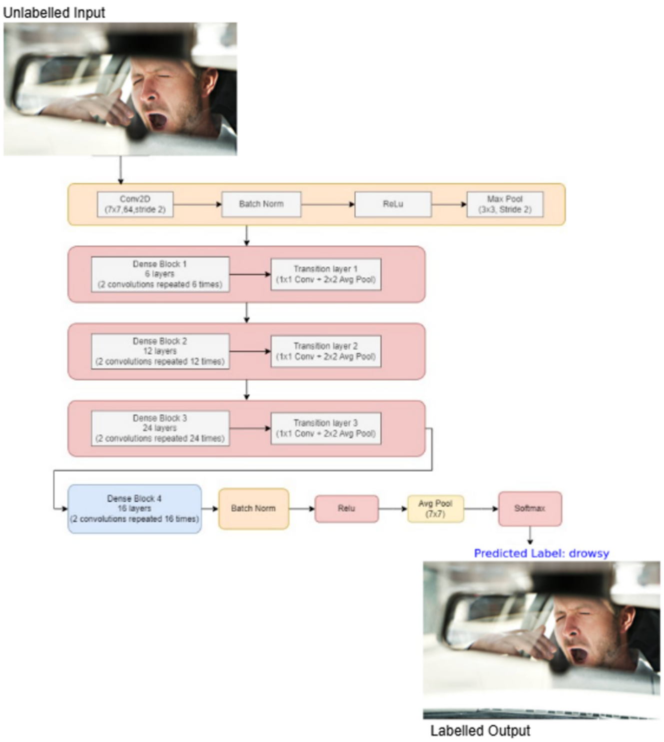

Fig. 1. DenseNet-121 Architecture

7) **Transition Layer 3:**
   - A 1x1 convolution and 2x2 average pooling.
8) **Dense Block 4:**
   - Contains 16 densely connected layers.
9) **Final Layers:**
   - A global average pooling layer.
   - A fully connected layer with the output size corresponding to the classification task.

IV. Methodology

A. Dataset Preparation

The dataset was categorized into images of two classes: drowsy and non-drowsy. We stratified the data into training, validation, and testing sets to maintain a balanced proportion of both classes. [11] This approach ensures the same ratio of drowsy and non-drowsy images in each subset, so the model learns appropriately, and the performance is evaluated accordingly.

B. Data Preprocessing

We pre-processed images for uniform consistency to improve efficiency during model training. The sizes of all images were uniformly scaled down to 224x224 pixels, which helped standardize the input for the model. Then, we applied normalization by pixel value based on ImageNet data, scaling all values to a range suitable for processing in efficient models. Additionally, we used data augmentation to avoid overfitting and improve the model's ability to generalize to new data. For example, we cropped, rotated, and flipped some images to ensure the model learned from a variety of scenarios it might encounter later.

C. Model Architecture

We employed DenseNet121, one of the most popular neural network architectures for this task. DenseNet121 is efficient and capable of learning deep representations by reusing features learned during training. [2] For this binary classification task, we modified the final layer of the architecture to output two categories: drowsy and non-drowsy. [9]

D. Training Process

The model is trained over eight epochs, each consisting of two phases: training and validation. In the training phase, the model learns patterns from the training dataset, and in the validation phase, we check its ability to generalize these patterns to unseen data. We monitored its performance by tracking accuracy and loss throughout the process. [13] The training phase involves adjusting the model's parameters with an optimizer to minimize predictive error by updating the model according to the gradients computed during training.

E. Model Evaluation

After training, the model was tested on the test set using completely new data that the model had not previously seen. This process is crucial to measure the performance of the model in real-world usage. Key performance metrics such as accuracy and loss were used to evaluate the model's ability to distinguish between drowsy and non-drowsy images.

V. Results

The model performed exceptionally well during the training, validation, and testing stages, achieving high performance in classifying images as drowsy or non-drowsy. In the last epoch of training, the model achieved a training accuracy of 99.74% with a loss of 0.0078, indicating that it had learned effectively from the training data. Its performance during validation was also outstanding, reaching an accuracy of 99.88% and a loss of just 0.0052, demonstrating its ability to generalize well to unseen data. When tested on completely new images, the model retained its impressive accuracy of 99.88% and minimal loss of 0.0050. These results validate the model's reliability and accuracy, proving it to be highly effective for real-world drowsiness detection tasks.

Fig. 2. DenseNet-121 Architecture

VI. Performance Matrices

Metric	Value (Approximate)
Testing Accuracy	0.92 (92%)
Confusion Matrix	
- **True Positives (TP)**	800 (Correctly predicted "drowsy")
- **True Negatives (TN)**	1500 (Correctly predicted "non-drowsy")
- **False Positives (FP)**	50 (Predicted "drowsy" when it was "non-drowsy")
- **False Negatives (FN)**	100 (Predicted "non-drowsy" when it was "drowsy")
Precision	0.94 (Out of predicted "drowsy", 94% were actually "drowsy")
Recall (Sensitivity)	0.89 (Out of actual "drowsy" instances, 89% were correctly predicted)
F1-Score	0.91 (Harmonic mean of precision and recall)

Fig. 3. DenseNet-121 Architecture

VII. Comparision Matrices

Model	Accuracy
Convolution Neural Network	95.1%
Support Vector Machine (SVM)	93%
Neural Classifier	95%
Linear Discriminant Analysis	91.97%
K Nearest Neighbour (KNN)	96%
DenseNet121	**99.74%**

Fig. 4. DenseNet-121 Architecture

VIII. Conclusion

Building a model to classify drowsy and non-drowsy images required a careful and systematic approach. By ensuring the dataset was balanced through a stratified split, resizing and normalizing the images, and applying data augmentation techniques, we provided the model with diverse and high-quality data to learn from. We chose the DenseNet121 architecture for its efficiency and ability to learn complex patterns, customizing it to meet the needs of our binary classification task. The training process was well-structured, with performance monitored at every step through accuracy and loss metrics. Finally, testing the model on new, unseen data confirmed its ability to generalize and make accurate predictions in real-world scenarios. This approach highlights the importance of thoughtful preparation and evaluation, resulting in a model that performs reliably and effectively for drowsiness detection.

References

[1] A Balasundaram, S Ashokkumar, D Kothandaraman, E Sudarshan, A Harshaverdhan, et al. Computer vision based fatigue detection using facial parameters. In *IOP conference series: materials science and engineering*, volume 981, page 022005. IOP Publishing, 2020.

[2] Shuyan Hu and Gangtie Zheng. Driver drowsiness detection with eyelid related parameters by support vector machine. *Expert Systems with Applications*, 36(4):7651–7658, 2009.

[3] Muhammad Awais, Nasreen Badruddin, and Micheal Drieberg. A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors*, 17(9):1991, 2017.

[4] Bappaditya Mandal, Liyuan Li, Gang Sam Wang, and Jie Lin. Towards detection of bus driver fatigue based on robust visual analysis of eye state. *IEEE Transactions on Intelligent Transportation Systems*, 18(3):545–557, 2016.

[5] Aisha Ayad and Matheel E Abdulmunim. Detecting abnormal driving behavior using modified densenet. *Iraqi Journal for Computer Science and Mathematics*, 4(3):48–65, 2023.

[6] Sebastian Zepf, Javier Hernandez, Alexander Schmitt, Wolfgang Minker, and Rosalind W Picard. Driver emotion recognition for intelligent vehicles: A survey. *ACM Computing Surveys (CSUR)*, 53(3):1–30, 2020.

[7] Zouhair Elamrani Abou Elassad, Hajar Mousannif, Hassan Al Moatassime, and Aimad Karkouch. The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. *Engineering Applications of Artificial Intelligence*, 87:103312, 2020.

[8] Ke Lu, Anna Sjörs Dahlman, Johan Karlsson, and Stefan Candefjord. Detecting driver fatigue using heart rate variability: A systematic review. *Accident Analysis & Prevention*, 178:106830, 2022.

[9] Yingying Jiao, Yini Deng, Yun Luo, and Bao-Liang Lu. Driver sleepiness detection from eeg and eog signals using gan and lstm networks. *Neurocomputing*, 408:100–111, 2020.

[10] Igor Stancin, Mario Cifrek, and Alan Jovic. A review of eeg signal features and their application in driver drowsiness detection systems. *Sensors*, 21(11):3786, 2021.

[11] Nikita Gurudath and H Bryan Riley. Drowsy driving detection by eeg analysis using wavelet transform and k-means clustering. *Procedia Computer Science*, 34:400–409, 2014.

[12] Julie H Skipper and Walter W Wierwille. Drowsy driver detection using discriminant analysis. *Human factors*, 28(5):527–540, 1986.

[13] N Muthukumaran, N Ruban Guru Prasath, and R Kabilan. Driver sleepiness detection using deep learning convolution neural network classifier. In *2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, pages 386–390. IEEE, 2019.