REAL-TIME DRIVER DROWSINESS DETECTION SYSTEM

[1]CHENNA LAVANYA, [2]Dr.GADI. HARITHA RANI
[1]Student, [2]Associate Professor
[12]DEPARTMENT OF COMPUTER APPLICATIONS
[12]RAJAHMAHENDRI INSTITUTE OF ENGINEERING AND TECHNOLOGY, Rajahmahendry

Abstract—**Driver fatigue and drowsiness pose significant risks to road safety, necessitating advanced monitoring systems to mitigate accidents. This research addresses this concern through the implementation of a Real-Time Driver Drowsiness Monitoring system using Vision Transformer (ViT). The study involved fine-tuning a Vision Transformer (ViT) model with additional layers on a dataset comprising 84900 images of open and closed eyes, enabling accurate detection of driver drowsiness. The motivation for drowsiness detection stems from the critical role of attentiveness in driving safety. Existing alarm systems often rely on conventional methods, and the research seeks to enhance this by incorporating state-of-the-art ViT technology. Achieving an impressive accuracy of 98.8%, the real-time monitoring system uses the ViT model to detect driver drowsiness promptly. When drowsiness is detected, an alarm is triggered to alert the driver, promoting immediate attention, and reducing the risk of accidents. The model's performance is also thoroughly evaluated, demonstrating high precision, recall, and F1-Score for both closed and open-eye states. This study benefits the field of driver safety by providing a reliable and effective real-time drowsiness monitoring method, thus enhancing road safety, and reducing the incidence of fatigue-related accidents**

Index Terms—**REAL-TIME ,DRIVER ,DROWSINESS, DETECTION ,SYSTEM**

I. Introduction

Ensuring everyone's safety on the road depends critically on the alertness of drivers. According to estimates from the World Health Organization, 1.3 million people die in accidents each year. Twenty to fifty million more people suffer non-fatal injuries, many of which result in lifelong disability. The main reasons for traffic accidents are tiredness, drowsiness, and inattentive driving. As a result, a system that can alert drivers to fatigue and help them become more vigilant by sending out timely warnings may be able to reduce the number of accidents that occur, save money, and reduce personal suffering.

The research introduces an innovative approach to advance driver safety through the fine-tuning of the Vision Transformer (ViT) model on a dataset for the accurate analysis of driver drowsiness. By delving into advanced computer vision, the study seeks to develop a robust system for proactive detection of driver drowsiness, ultimately contributing to accident prevention and overall road safety improvement. It is worthemphasizing the preference for Vision Transformer (ViT) over traditional algorithms like CNN or KNN due to ViT's exceptional ability to capture intricate long-range dependencies in images. To ensure reliable training and testing, the model is trained using a large, labeled dataset approach, with the data divided into 80% for training, 10% for validation, and 10% for testing.

SYSTEM ANALYSIS

EXISTING SYSTEM:

In existing driver drowsiness detection systems, various approaches have been explored to identify signs of fatigue among drivers. Early methods relied heavily on **Convolutional Neural Networks (CNNs)**, such as modified versions of **VGG-16** and **VGG-19**, achieving up to **95% accuracy** using datasets like ZJU. Traditional computer vision techniques, like **Haar Cascade Classifiers**, were also utilized to detect **eye blink rates** and **yawning frequency**, achieving **90% accuracy** for blink detection but only around **60%** for yawning, showing significant limitations. Several models based on **Support Vector Machines (SVMs)** were introduced, focusing on eye image characteristics like **circularity** and **black ratio**, reaching an accuracy of about **91.3%**. Other studies applied **Fisher's Linear Discriminant Analysis (FLDA)** and **Bayesian classifiers** alongside SVMs for improved detection based on visual behavior.

Some researchers explored the use of **gradient statistics** and machine learning to detect eye closure, achieving high accuracy rates, especially in cases without eyeglasses. Facial landmark detection methods, particularly using **Dlib** and calculations of **Eye Aspect Ratio (EAR)** and **Mouth Aspect Ratio (MAR)**, were proposed and recorded accuracy rates as high as **96.71%**. Beyond vision-based systems, physiological signal-based approaches were tested, such as using a **single EEG channel** analyzed by **Artificial Neural Networks (ANN)** and SVMs, where ANN showed superior performance. Lightweight deep learning models, particularly **Lightweight CNNs**, were also introduced to reduce model size while maintaining high accuracy of about **98.7%**. Additionally, simpler techniques like the **PERCLOS method** (percentage of eye closure over time) using **Viola-Jones** face detection achieved moderate success. Some systems combined handcrafted image descriptors with traditional supervised classifiers but only achieved limited accuracy around **75.90%**. Despite their contributions, these existing systems often faced challenges such as limited robustness under varied lighting conditions, lower real-time performance, high model complexity, and lack of scalability in real-world scenarios.

DISADVANTAGES OF EXISTING SYSTEM:

Low Accuracy with Simple Methods:

Haar Cascade + basic measurements like yawning detection achieved

only 60% accuracy.

High Dependency on Lighting Conditions:

Some computer vision methods (like Viola-Jones, Haar Cascades) are sensitive to lighting changes.

Limited Feature Extraction:

Classical methods (e.g., EAR, PERCLOS) extract only simple features, missing subtle drowsiness signs.

Heavy Models for Real-Time:

Traditional CNNs (VGG16/VGG19)areheavy forreal-time implementation without optimization.

Poor Generalization:

Models trained on small or limited datasets may not perform well across different faces, skin tones, glasses, or lighting variations.

Limited Real-Time Feedback:

Some systems lacked instant alarming mechanisms or had delays in detecting drowsiness.

EEG-based Methods:

While accurate, EEG-based systems are intrusive (require wearing electrodes), impractical for drivers.

Algorithm:

- **CNNs** (VGG-16, VGG-19)
- Bayesian Classifier
- Lightweight Convolutional Neural Network (Lightweight CNN)

1.1. PROPOSED SYSTEM:

The proposed system in this research focuses on developing a **Real-Time Driver Drowsiness Detection System** using a **Vision Transformer (ViT)** model. To begin with, a large and comprehensive dataset containing approximately **84,900 images** of open and closed eyes was curated. The dataset was carefully organized to handle class imbalance through **random oversampling** of the minority class, ensuring equal representation of both open and closed-eye images. In order to enhance the model's robustness to various real-world conditions, **data augmentation** techniques such as random rotations, random resized cropping, and sharpness adjustments were applied to the images.

For model development, the researchers selected the **Vision Transformer (ViT)** architecture, specifically the pre-trained **google/vit-base-patch16-224-in21k** model, due to its superior ability to capture long-range dependencies and subtle patterns in images. The ViT model was further customized by adding **fully connected layers** with hidden units of 128, 256, and 384 neurons, each followed by **ReLU activation functions** and a **dropout rate of 0.5** to avoid overfitting. Training was performed using a small **learning rate (0.00001)** and appropriate **batch sizes** (64 for training and 32 for validation/testing), with a **weight decay** of 0.02 and a **learning rate warm-up strategy** to ensure stable convergence.

The real-time implementation of the system involved capturing live video streams from a **webcam** using **OpenCV**, followed by **face detection** using a **Haar Cascade Classifier**. Once a face was detected, the system extracted the eye region and passed it through the fine-tuned ViT model to predict the eye state (open or closed). If the model predicted a closed eye with a probability greater than **0.5** consistently, an **alarm sound** was triggered to alert the driver immediately. The proposed system demonstrated exceptional performance with an accuracy of **98.8%**, providing a highly reliable and efficient method for preventing accidents caused by driver drowsiness.

ADVANTAGES OF PROPOSED SYSTEM:

Higher Accuracy:

Achieved **98.8% accuracy**, which is higher than many existing methods.

Better Feature Extraction:

ViT captures **long-range dependencies** and **subtle facial features** better than CNN or classical methods.

Improved Generalization:

Fine-tuning on a **large and balanced dataset** made the model more robust to different faces, lighting conditions, and eye states.

Real-Time Implementation:

The system processes live video feeds with very little delay, making it **practical for real-world driving scenarios**.

Alarm Integration:

Immediate **audio alerts** are triggered when drowsiness is detected, increasing driver

safety.

Efficient Use of Data Augmentation:

Through rotations, cropping, resizing, and sharpness adjustments, the model becomes more **adaptive to real-world variations** like different head poses or lighting.

Ease of Interpretation:

Using label mapping (label2id, id2label), the system easily interprets model outputs into human-readable classes ("open eye", "closed eye").

Scalability:

Vision Transformer architecture is **scalable** to different input sizes and more complex real-world datasets

Algorithm:

- Vision Transformer (ViT)
- OpenCV

1.2. FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are,

- ♦ **ECONOMICAL FEASIBILITY**
- ♦ **TECHNICAL FEASIBILITY**
- ♦ **SOCIAL FEASIBILITY**

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available. Only the customized products had to be purchased.

II. System Analysis

2.1 Existing System The current safety systems in railway and powerline infrastructure primarily rely on manual inspections and conventional sensor-based monitoring. These systems are reactive in nature and often fail to predict faults in advance. Railway track inspection involves human visual checks or vibration sensors, which are time-consuming and prone to human error. Similarly, in powerlines, visual and thermal inspections are conducted manually or with basic automated alerts. Maintenance schedules are typically time-based rather than condition-based. There's limited integration of real time data analytics. The systems lack adaptability to sudden environmental changes. Traditional approaches are not scalable for large and remote infrastructure. Alerting mechanisms are basic and often delayed. As a result, accidents and failures are not always preventable. 2.1.1 Disadvantages of the Existing System 1) Manual Dependency: Prone to human errors and fatigue, reducing reliability. 2) Delayed Fault Detection: Reactive approach fails to prevent many accidents. 3) Scalability Issues: Inefficient for large-scale infrastructure monitoring. 4) Lack of Real-Time Analytics: Data is not utilized effectively for predictive maintenance. 2.2 Proposed System The proposed system introduces machine learning (ML)-based models to ensure proactive monitoring and fault detection in railway and powerline systems. It integrates sensors, drones, and computer vision with AI algorithms to detect defects in real time. ML models are trained to recognize anomalies like rail cracks, powerline sagging, or vegetation interference. These systems are cloud-integrated for centralized data analysis and continuous learning. Real-time alerts and predictive 2 analytics help in preventive maintenance. IoT-enabled sensors feed live data to the ML model. The architecture supports adaptive learning from historical and live data. Automation minimizes human involvement and improves accuracy. This enhances response times and reduces operational risks. Overall, it improves infrastructure safety and sustainability significantly.

III. SYSTEM DESIGN

3.1 SYSTEM ARCHITECTURE:

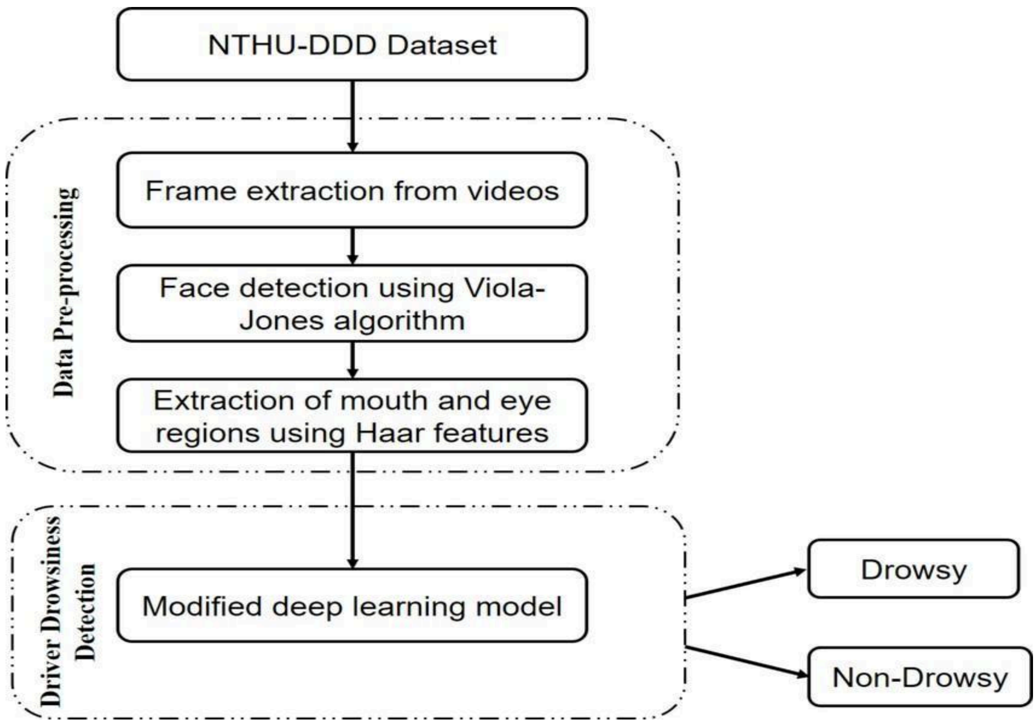

Fig-3.1: System Architecture

A. INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.
2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.
3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

B. OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system’s relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2. Select methods for presenting information.
3. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- Convey information about past activities, current status or projections of the
- Future.
- Signal important events, opportunities, problems, or warnings.
- Trigger an action.
- Confirm an action.

3.2 DATA FLOW DIAGRAM:

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.
4. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

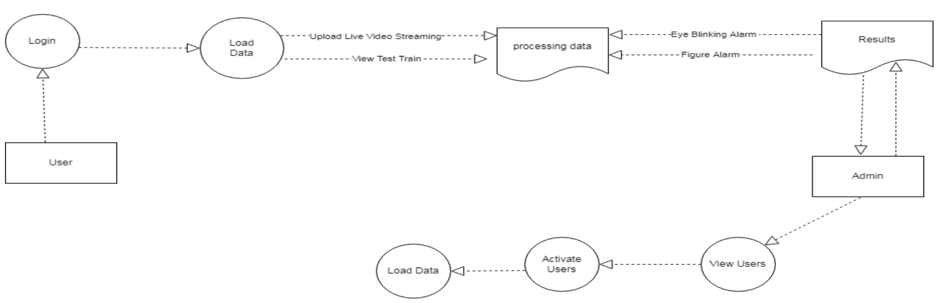

Fig-3.2: Data Flow Diagram

3.3 UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of

IV. CONCLUSIONS

The research has successfully developed a real-time Driver Drowsiness Monitoring system for drivers, employing a combination of image processing techniques and a Vision Transformer (ViT) with additional layers to discern the drowsiness of drivers. With an impressive accuracy rate of 98.8%, the model, trained on an 80-10-10 data split, showcases its robustness in identifying drowsiness. Furthermore, the integration of an alarming sound feature enhances the system's practicality by promptly capturing the driver's attention. With its proactive monitoring of driver attentiveness, this innovative solution represents a major advancement in road safety.

References

[1] M. Cao, “Competitive Advantages and Challenges of E-Commerce,” Adv. Econ. Manag. Polit. Sci., vol. 47, pp. 114-119, December 2023.
[2] S. C. S. C. Yong, R. T. Huan, W. S. Poh, M. Osman, and D. C. W. Ng, “Assessing the Factors Influencing Consumer Behaviour in E-Commerce Platforms,” Int. J. Multidiscip. Appl. Bus. Educ. Res., vol. 4, pp. 3725 3735, October 2023.
[3] Y. Zhang, “Utilizing machine learning algorithms for consumer behaviour analysis,” Appl. Comput. Eng., vol. 49, pp. 213-219, March 2024.
[4] M. Shen, C. S. Tang, D. Wu, R. Yuan, and W. Zhou, “JD.com: Transaction-Level Data for the 2020 MSOM Data Driven Research Challenge,” Manuf. Serv. Oper. Manag., vol. 26, pp. 2-10, February 2024.
[5] A. Biswas, “Feature Engineering: Categorize your data using One-Hot Encoding,” Medium, June 2024.
[6] H. Blockeel, L. Devos, B. Frénay, G. Nanfack, and S. Nijssen, “Decision trees: from efficient prediction to responsible AI,” Front. Artif. Intell., vol. 6, July 2023.
[7] N. Donges, “Random Forest: A Complete Guide for Machine Learning,” BuiltIn, March 2024.
[8] J. Brownlee, “Logistic Regression for Machine Learning,” Machine Learning Mastery, December 2023.
[9] M. H. Hassan, “Tuning Model Hyperparameters With Random Search,” Medium, November 2023.