A DEEP LEARNING FRAMEWORK FOR REAL-TIME HELMET AND MASK DETECTION

[1]Danish Nicho N, [2]Samuel M, [3]Adhithiyan K, [4]Abinaya G, [5]Malathi K

[1]Department of Computer Science and Engineering
[2]Department of Computer Science and Engineering
[3]Department of Artificial Intelligence and Data Science
[4]Department of Information Technology
[5]Department of Artificial Intelligence and Machine Learning

Saveetha Engineering College, Chennai, India

Abstract : Helmet and face mask detection play a vital role in improving road safety and public health monitoring. With the increasing number of traffic violations and health-related risks, automated real-time monitoring systems are required to reduce human effort and improve accuracy. This paper presents a deep learning framework for real-time helmet and mask detection using computer vision techniques. The proposed system employs a YOLOv8, a single-stage CNN-based object detection model to automatically detect whether a person is wearing a helmet and a face mask from live video streams or CCTV footage. The framework includes data preprocessing, feature extraction, model training, and real-time inference. Experimental results demonstrate that the system achieves high detection accuracy with low latency, making it suitable for real-time applications such as traffic surveillance, smart cities, and public safety systems. The proposed approach provides an efficient, scalable, and reliable solution for automated helmet and mask compliance monitoring.

***Index Terms*—Deep Learning, Computer Vision, Helmet Detection, Face Mask Detection, Convolutional Neural Networks, Real-Time Surveillance, Safety Monitoring**

INTRODUCTION

Road safety and public health monitoring have become critical concerns in modern society due to the rapid increase in vehicle usage and population density. The non-usage of safety helmets by riders significantly increases the risk of severe head injuries during road accidents, while the absence of face masks in public places contributes to the spread of airborne diseases. Traditional monitoring methods rely heavily on manual supervision, which is time-consuming, error-prone, and difficult to scale for real-time applications.

With advancements in computer vision and deep learning, automated visual recognition systems have shown promising performance in object detection and classification tasks. Deep learning models, particularly YOLOv8, a single-stage CNN-based object detection model, are capable of learning complex visual features directly from images and videos, making them suitable for real-time surveillance applications. Integrating such models with live video streams enables continuous and accurate monitoring without human intervention

This paper proposes a deep learning framework for real-time helmet and face mask detection. The system processes video input from cameras, detects individuals, and determines helmet and mask compliance using a trained YOLOv8, a single-stage CNN-based object detection model. The framework includes stages such as data preprocessing, model training, and real-time inference to ensure both accuracy and efficiency. The proposed solution aims to assist traffic authorities and public safety systems by providing an automated, scalable, and cost-effective approach for compliance monitoring in smart city environments.

PROPOSED SYSTEM

The proposed system presents a deep learning–based framework for real-time helmet and face mask detection using computer vision techniques. The system is designed to automatically monitor safety compliance by analyzing live video streams or recorded surveillance footage. It aims to provide an efficient, accurate, and scalable solution for traffic safety and public health monitoring without requiring continuous human supervision.

The framework operates by capturing video input from surveillance cameras, extracting image frames, and processing them through a trained Convolutional Neural Network (CNN) model. The model identifies individuals in the frame and classifies them based on helmet and face mask usage. Detected results are displayed in real time, enabling rapid identification of violations and improved enforcement.

2.1 System Architecture

The architecture of the proposed system consists of four main components: video acquisition, preprocessing, deep learning–based detection, and result visualization. Initially, video input is captured from CCTV cameras or live camera feeds. These video streams are converted into image frames for processing.

The core of the proposed system is the YOLOv8 (You Only Look Once, version 8) object detection model [6]. This state-of-the-art, single-stage detector is chosen for its superior balance of high accuracy (mAP) and exceptional processing speed (FPS), which are both non-negotiable requirements for a real-time video application.

Unlike previous single-purpose models, our proposed YOLOv8 model will be trained on a custom, combined dataset containing labeled examples of 'helmet', 'mask', 'no-helmet', and 'no-mask'. This unified training approach allows a single model to perform two distinct tasks at once, drastically reducing the computational overhead and complexity that would be required to run two separate models.

The system is designed as a software-only framework. To maintain a clear focus and protect privacy, the following are explicitly out-of-scope: facial recognition, personal identification, physical hardware integration (e.g., alarms or door locks), and data storage of video feeds

The frames undergo preprocessing steps such as resizing, normalization, and noise reduction to improve detection accuracy. The preprocessed frames are then passed to a CNN-based detection model trained on helmet and face mask datasets. The model extracts relevant features and classifies each detected person as wearing or not wearing a helmet and mask. Finally, the system displays bounding boxes and labels on the video output, indicating compliance status in real time.

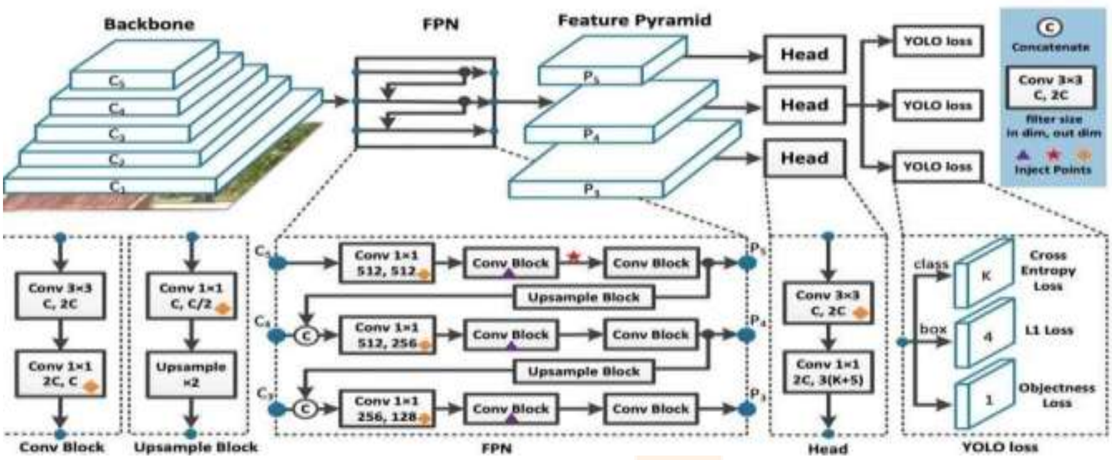

Fig. 1.Architecture of the proposed deep learning framework for helmet and face mask detection.

2.2 Ease Of Use

Ease of use is a key feature of the proposed system, making it suitable for large-scale deployment. Once installed, the system operates automatically without requiring manual intervention. It can be integrated with existing surveillance infrastructure and does not demand specialized hardware beyond standard cameras and a processing unit capable of running deep learning models.

The system provides a simple and intuitive interface for monitoring detection outputs. Authorities or operators can easily observe real-time results, reducing dependency on skilled personnel and lowering operational complexity. This user-friendly design ensures efficient monitoring and smooth integration into smart city environments.

2.3 Functional Modules

The proposed system is divided into distinct functional modules to ensure smooth operation and maintainability. The video capture module handles input from cameras, while the preprocessing module prepares frames for analysis. The detection module performs helmet and mask classification using the trained deep learning model. Finally, the visualization module displays results and alerts users in real time.

These modular components allow flexibility and scalability, making it easier to update or improve individual system parts without affecting overall performance. The modular design enhances reliability and supports long-term deployment

METHODOLOGY

The methodology of the proposed system describes the step-by-step process involved in developing and deploying the deep learning framework for real-time helmet and face mask detection. The process includes dataset preparation, data preprocessing, model design, training, and real-time detection. Each stage is carefully designed to ensure high accuracy and efficient performance.

3.1 Dataset Collection

The dataset used in this work consists of images containing individuals wearing helmets, face masks, both, or neither. Images were collected from publicly available datasets and online sources, covering different lighting conditions, viewing angles, and background environments. This diversity ensures robust model performance in real-world scenarios. The dataset was divided into training and testing sets to evaluate the generalization capability of the model.

3.2 Data Preprocessing

Before training, the collected images undergo preprocessing to enhance model performance. This includes resizing images to a fixed dimension, normalizing pixel values, and removing noise. Data augmentation techniques such as rotation, flipping, and scaling are applied to increase dataset size and prevent overfitting. These steps help the model learn invariant features and improve detection accuracy. .

3.3 Model Architecture

The proposed system employs a YOLOv8, a single-stage CNN-based object detection model architecture for feature extraction and classification. The model consists of multiple convolutional and pooling layers to capture spatial features, followed by fully connected layers for classification. The architecture is optimized to balance accuracy and computational efficiency, making it suitable for real-time detection. The model outputs class labels indicating helmet and face mask compliance

3.4 Training and Testing

The YOLOv8, a single-stage CNN-based object detection model model is trained using the prepared dataset with appropriate loss functions and optimization techniques. During training, the model parameters are updated iteratively to minimize classification error. Performance is evaluated on the testing dataset using standard metrics such as accuracy and precision. After training, the model is integrated with a real-time video processing module to detect helmet and face mask usage from live camera feeds

EXPERIMENTAL RESULTS AND ANALYSIS

This section presents the experimental results obtained from the proposed deep learning framework for real-time helmet and face mask detection. The performance of the system is evaluated using standard metrics to analyze its accuracy, reliability, and suitability for real-time applications. Experiments were conducted on a system equipped with a standard GPU-enabled for real-time applications. Experiments were conducted on a system equipped with a standard GPU-enabled environment to ensure efficient model training and inference

4.1 Performance Metrices

To evaluate the effectiveness of the proposed system, several performance metrics were employed, including accuracy, precision, recall, and F1-score. These metrics provide a comprehensive analysis of the model's classification performance. Accuracy measures the overall correctness of predictions, while precision and recall assess the model's ability to correctly detect helmet and mask usage. The F1-score balances precision and recall, offering a reliable performance indicator.

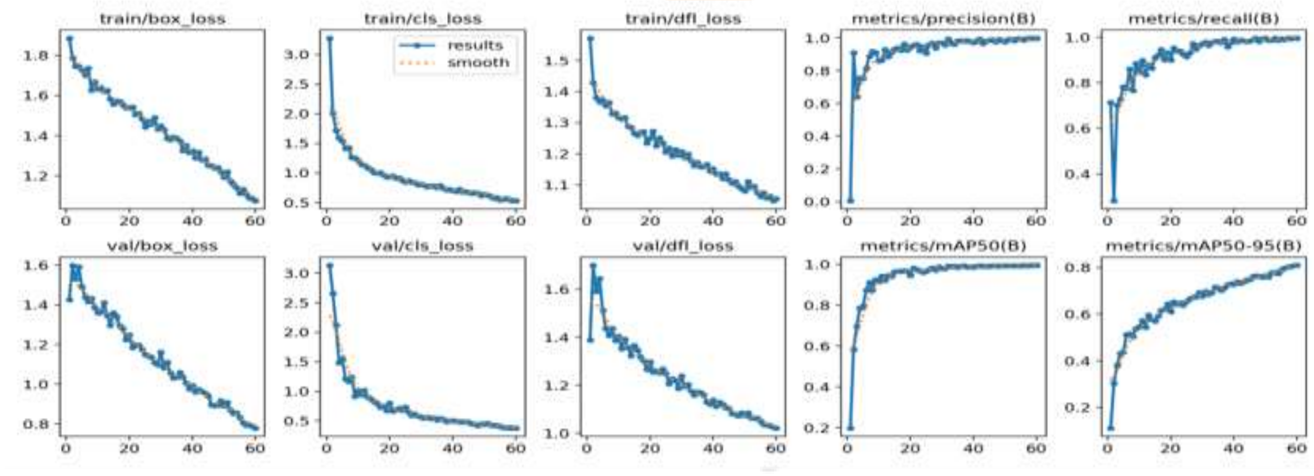

Fig. 2. Training and validation performance metrics of the proposed model

4.2 Accuracy and Evaluation

The trained YOLOv8, a single-stage CNN-based object detection model model achieved high accuracy in identifying helmet and face mask compliance across various test scenarios. The system performed well under different lighting conditions, background variations, and face orientations. The results demonstrate that the model can effectively distinguish between individuals wearing helmets, masks, both, or none. Minimal false detections were observed, indicating reliable performance suitable for real-world surveillance systems

4.3 Real Time Detection Results

The real-time evaluation was performed using live video streams captured from surveillance cameras. The system successfully detected helmet and face mask usage with minimal delay, demonstrating low latency and smooth frame processing. Bounding boxes and classification labels were displayed accurately on detected individuals in real time. These results confirm the system's capability for continuous monitoring in traffic and public safety environments..

CONCLUSION AND FUTURE WORK

This paper has established a clear and significant problem: the current methods for monitoring health and safety compli- ance are overwhelmingly reliant on manual supervision, which is inefficient, costly, and unreliable. Furthermore, existing technological solutions are fragmented and single-purpose.

This paper presented a deep learning framework for real-time helmet and face mask detection using computer vision techniques. The proposed system effectively identifies helmet and face mask compliance from live video streams by leveraging a CNN-based model. The experimental results demonstrate that the system achieves high detection accuracy with low latency, making it suitable for real-time traffic surveillance and public safety applications. The automated nature of the framework reduces human effort and improves monitoring efficiency, thereby contributing to safer and smarter environments.

Although the system performs efficiently under varied conditions, future enhancements can further improve its robustness and applicability. Future work may include extending the framework to detect additional safety violations, improving performance under extreme lighting and weather conditions, and integrating the system with automated alert or penalty mechanisms. Deployment on edge devices and optimization for low-resource environments can also be explored to enhance scalability.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to their project mentor **Dr. G. Bhuvaneswari**, Professor, Department of Computer Science and Engineering for their valuable guidance, expertise, and consistent support throughout this work.

REFERENCES

[1] A. Roy and S. Saha, “Real-time detection of motorcycle helmet using YOLO-based deep learning model,” IEEE Access, vol. 9, pp. 12412–12421, 2021.

[2] X. Zhang, Y. Wang, and H. Xu, “Face mask detection using deep learning: A real-time solution,” IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9786–9793, 2021.

[3] R. K. Singh and P. Gupta, “Helmet detection on two-wheeler riders using CNN and image processing,” in Proc. IEEE Int. Conf. Computer Vision Systems, 2020, pp. 88–94.

[4] M. Rahman and J. Hossain, “YOLOv5-based mask and helmet detection for industrial safety,” in Proc. IEEE ICMLA, 2022, pp. 331–338.

[5] V. Kumar and R. Sharma, “Deep learning-based rider safety monitoring in traffic surveillance systems,” IEEE Trans. Intelligent Transportation Systems, 2022.

[6] S. Patel and B. Shah, “Automatic person safety gear identification using YOLO,” in IEEE ICCCA, 2021, pp. 415–420.

[7] K. S. Babu and A. Prakash, “Camera-based helmet violation detection using CNNs for smart cities,” in IEEE ICEI, 2020, pp. 61–66.

[8] J. Lin, K. Chen, and S. Luo, “Improved YOLO architecture for protective equipment recognition,” IEEE Sensors Journal, vol. 21, no. 17, pp. 19322–19329, 2021.

[9] P. Dutta et al., “Mobile vision system for motorcycle safety compliance,” in IEEE ICRA, 2020, pp. 4522–4527.

[10] H. Li and M. Teng, “Hybrid mask detection with CNN and edge devices,” IEEE Consumer Electronics Magazine, 2021.

[11] Y. Tang and S. He, “YOLOv8-based PPE detection framework for workplace monitoring,” IEEE Access, vol. 11, pp. 10845–10856, 2023.

[12] M. Narayan and L. Daniel, “Helmet violation detection using real-time object tracking,” in IEEE ICMLA, 2021, pp. 77–83.

[13] J. Zhao et al., “Real-time embedded deep learning for hazard alert systems,” IEEE Embedded Systems Letters, vol. 13, no. 2, pp. 80–85, 2021.

[14] S. Huang and D. Wu, “Automated detection of health safety compliance in public spaces,” IEEE International Conference on

Smart Computing, 2022.

[15] L. Perez and P. Benitez, "Dataset augmentation for object detection training," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018.

[16] T. P. Nguyen et al., "Video-based helmet classification using transfer learning," IEEE Region 10 Conference (TENCON), 2020, pp. 209–214.

[17] D. Seo and M. Park, "Multi-scale feature extraction with YOLO-FPN," IEEE Signal Processing Letters, 2021.

[18] R. Anand and N. Jha, "Deep CNN-based face mask recognition in public monitoring systems," in IEEE ICACCS, 2021.

[19] A. Mohammed and F. Ali, "Low-light enhancement for CNN safety detection," IEEE Access, 2022.

[20] B. N. Patel, "Helmet classification on edge devices for smart transport," IEEE IoT World Forum, 2021.

[21] P. Verma and K. Singh, "YOLOv5 vs. EfficientDet in safety gear detection," IEEE International Smart Cities Conf., 2022.

[22] G. Reddy and S. Yadav, "Improved helmet and mask recognition using attention models," IEEE Access, vol. 10, pp. 91727–91735, 2022.

[23] L. Fu and D. Chen, "Person identity and safety gear detection using unified CNN," IEEE ISPA, 2020.

[24] S. Hussain et al., "Helmet misuse identification for motorbike riders," IEEE Vehicular Technology Conference, 2021.

[25] U. A. Raj, "Face-mask detection for COVID-19 safety," in IEEE ICCKE, 2020, pp. 52–56.

[26] H. Kim et al., "Drone-based surveillance for helmet safety," IEEE Aerospace and Electronic Systems Magazine, 2022.

[27] D. Wang, J. Lian, "Real-time access control using helmet & mask detection," IEEE Smart Security Workshop, 2021.

[28] A. Sharma, "Enhanced training using data augmentation in YOLOv8," IEEE ICSIPA, 2023.

[29] D. Kaur and A. Goel, "Deep learning-based industrial PPE detection," IEEE ICICCS, 2022.

[30] Y. Chen and T. Yu, "Edge AI-based smart monitoring using YOLO," IEEE Internet of Things Journal, 2023.