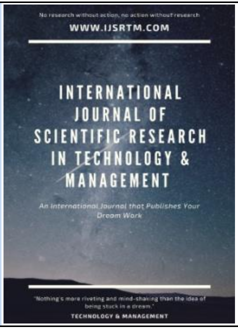
https://www.ijsrtm.com
Vol.5 Issue 1 Mar 2024: 01-05
Published online 11 March 2025

E-ISSN: 2583-7141

International Journal of Scientific Research in Technology & Management



Automatic Helmet Rule Violation Detection using Deep Learning Approach

Arun Pratap Singh
Dept. of Computer Science & Engineering
Samrat Ashok Technological Institute,
Vidisha, Madhya Pradesh, India
singhprataparun@gmail.com

Utkarsh Dubey
Dept. of Computer Science & Engineering
University Institute of Technology, RGPV
Bhopal, Madhya Pradesh, India
utkarshdubey7@gmail.com

Abstract— ***Traffic rule violations, particularly non-compliance with helmet usage among two-wheeler riders, contribute significantly to road accidents and fatalities. Traditional monitoring methods rely heavily on manual inspection, which is both time-consuming and error-prone. Recent advances in computer vision and deep learning provide effective solutions for automatic detection of helmet rule violations. This paper presents a deep learning–based approach for detecting riders without helmets using convolutional neural networks (CNNs) and object detection models such as YOLO, Faster R-CNN, or SSD. The system processes surveillance camera footage, identifies motorcyclists, detects helmet presence, and flags violations. The proposed method enhances road safety enforcement by providing real-time, scalable, and accurate detection compared to conventional methods.***

Keywords— ***Helmet Detection, Traffic Violation, Deep Learning, CNN, Object Detection, YOLO, Road Safety.***

I. Introduction

Road accidents caused by two-wheeler riders remain a major public safety issue worldwide, particularly in developing countries where motorcycles are a dominant mode of transportation. According to the World Health Organization, correct helmet usage reduces the risk of head injuries by nearly 70% and fatalities by 40% [1]. Despite these proven benefits, compliance with helmet rules is often low due to negligence, lack of awareness, and inadequate enforcement [2]. For example, in India, the Ministry of Road Transport and Highways reported that over 30% of road fatalities in 2020 involved two-wheeler riders not wearing helmets [3]. Traditional methods for monitoring helmet compliance rely on manual enforcement by traffic police, which is both labor-intensive and prone to human error, making it impractical for large-scale deployment in smart cities [4]. To address this, intelligent traffic surveillance systems using computer vision and deep learning have gained attention. Deep learning-based object detection frameworks such as Faster R-CNN [5], YOLO (You Only Look Once) [6], and SSD (Single Shot MultiBox Detector) [7] have demonstrated strong performance in real-time applications, making them suitable for helmet violation detection. Recent research highlights the effectiveness of these methods.

Fig. 1 Automatic Helmet Detection [3]

For instance, Real-Time Helmet Violation Detection using YOLOv5 demonstrated robust performance under varying lighting and weather conditions [8]. Similarly, a study on Multi-Class Helmet Violation Detection using YOLOv8 with Few-Shot Data Sampling showed that reliable detection is possible even with limited annotated data [9]. Furthermore, attention-based models such as the Residual Transformer-Spatial Attention Network improved accuracy in aerial and occluded views, addressing one of the major challenges in real-world surveillance scenarios [10]. Given this background, the present research aims to build a deep learning-based system capable of detecting motorcyclists, identifying helmet usage, and flagging violations in real-time. The objective is to design a robust and scalable pipeline that integrates with surveillance camera feeds to support traffic law enforcement and smart city initiatives, ultimately reducing road accidents and fatalities.

II. RELATED WORKS

Earlier attempts at automatic helmet detection relied heavily on handcrafted feature extraction methods, such as the Histogram of Oriented Gradients (HOG) and Support Vector Machines (SVM) for classification [11]. While these approaches demonstrated moderate success in constrained environments, they struggled with complex traffic scenes, dynamic backgrounds, and illumination variations, limiting their scalability for real-world applications [12]. The shift towards deep learning-based object detection significantly advanced helmet violation detection. Convolutional Neural Networks (CNNs) enabled automatic feature learning, reducing dependency on manual feature engineering [13]. Architectures like Faster R-CNN [5], YOLO (You Only Look Once) [6], and SSD (Single Shot MultiBox Detector) [7] were widely adopted for detecting motorcyclists and identifying helmet usage, achieving superior accuracy compared to classical methods. Several studies specifically tailored these models for helmet detection. For instance, YOLOv3-based frameworks demonstrated high detection accuracy even in cluttered traffic scenarios, making them suitable for real-time monitoring [14]. More recently, YOLOv5 models have been deployed for traffic surveillance, showing improved speed and precision while handling varied weather and lighting conditions [8]. Hybrid pipelines combining vehicle detection with helmet classification have also been proposed, ensuring that the system can first identify two-wheeler riders and then classify helmet usage [15]. Despite these advancements, persistent challenges remain. Occlusion handling, such as detecting helmets in crowded scenes where multiple riders are present, continues to be a bottleneck. Similarly, low-light conditions and infrared night surveillance pose difficulties for robust detection [16]. Furthermore, multi-rider detection, such as identifying driver and pillion riders, remains underexplored in many approaches. Addressing these gaps is essential for scaling helmet violation detection systems in diverse real-world traffic environments.

III. METHODOLOGY

The proposed system for automatic helmet rule violation detection is designed as a robust end-to-end deep learning pipeline that integrates state-of-the-art object detection algorithms with domain-specific classification modules. Unlike conventional systems that rely on manual monitoring or handcrafted features, the methodology leverages the power of convolutional neural networks (CNNs) and real-time object detectors to ensure both accuracy and scalability in diverse traffic scenarios. The pipeline is organized into four major stages: data acquisition, preprocessing, object detection and classification, and violation decision logic. Each stage is carefully designed to address challenges such as background clutter, occlusion, multi-rider detection, and varying lighting conditions in real-world traffic environments [17]–[22].

A. Data Acquisition

The first step involves collecting diverse visual data from traffic surveillance cameras, public datasets, and custom recordings at urban intersections. High-quality datasets are essential for training and validating deep learning models; therefore, existing datasets such as Pascal VOC [18], COCO [19], and custom helmet/non-helmet datasets are employed. These datasets include thousands of annotated images featuring riders with and without helmets across varied conditions. To capture realistic scenarios, data is collected under different weather conditions (sunny, rainy, foggy), times of day (daylight and nighttime), and traffic densities (sparse and congested). Annotators label bounding boxes around motorcycles, riders, and helmet regions, creating the ground truth required for supervised training.

B. Preprocessing

Once raw data is collected, preprocessing is performed to enhance feature quality and reduce noise. Images are resized to standard dimensions (e.g., 416×416 pixels for YOLO or 600×600 for Faster R-CNN) to ensure consistency across the dataset [20]. Histogram equalization and grayscale normalization are applied to minimize the impact of illumination variations across different cameras. Additionally, data augmentation techniques such as random cropping, horizontal flipping, rotation, and brightness adjustments are employed to artificially expand the dataset and improve the generalization capability of the models [21]. These steps ensure that the trained detector is robust against environmental variations and camera artifacts. For video-based analysis, frames are extracted at specific intervals (e.g., 10–15 FPS) to reduce redundancy while maintaining temporal continuity.

C. Object Detection and Classification

The detection stage is divided into two modules: motorcycle detection and helmet classification. In the first module, deep learning-based object detectors such as YOLOv5 [21] and Faster R-CNN [17] are used to localize motorcycles and riders in traffic scenes. YOLO is preferred for real-time deployment due to its single-shot detection mechanism and high mean Average Precision (mAP), while Faster R-CNN is leveraged in scenarios where accuracy is prioritized over speed. Once the motorcyclist is detected, a region of interest (ROI) corresponding to the rider's head is extracted. In the second module, a dedicated CNN classifier determines whether the rider is wearing a helmet. This binary classification problem is solved using deep architectures trained on helmet vs. non-helmet images [15]. In more advanced designs, multi-class classifiers can also distinguish between standard helmets, half helmets, or no helmet at all. For multi-rider scenarios (driver and pillion), the head regions of both riders are independently analyzed, ensuring comprehensive monitoring in high-risk conditions [22].

D. Violation Decision Logic

The final stage integrates detection and classification outputs to determine rule violations. If a motorcycle is detected and the corresponding rider's head ROI is classified as no helmet, the system flags it as a violation. Similarly, in cases of multiple riders, the system checks each rider individually, ensuring that both the driver and pillion comply with helmet rules. Violations are automatically logged and recorded with time-stamped evidence images and metadata (such as camera ID and location), enabling seamless integration with traffic monitoring centers and smart city platforms. This decision logic not only supports real-time alerting for on-road enforcement but also allows authorities to maintain digital records for penalty issuance and policy evaluation. Furthermore, by employing edge computing devices or GPU-enabled cloud servers, the pipeline can be deployed at scale to handle multiple traffic intersections simultaneously without compromising processing speed.

E. Generative Models for Data Augmentation

Data scarcity remains a significant challenge in handwritten text recognition, especially for historical manuscripts, rare scripts, or low-resource languages. Generative models such as Generative Adversarial Networks (GANs) [12] and diffusion models [13] have been employed to synthesize realistic handwriting samples, thereby augmenting training datasets and enhancing model robustness. GAN-based augmentation generates new handwriting styles by learning the distribution of real handwritten samples, producing variations in stroke, slant, and character shapes that mimic human writing. Diffusion-based augmentation, on the other hand, iteratively refines noise into realistic handwriting, enabling the creation of diverse and high-fidelity samples for training. By incorporating these generative approaches, deep learning architectures can better generalize to unseen handwriting styles and improve recognition accuracy, particularly on challenging or underrepresented datasets.

IV. DATASETS

The performance of any deep learning-based helmet violation detection system is strongly dependent on the quality and diversity of the datasets used during training and testing. Publicly available datasets and custom-collected footage form the backbone of experimental evaluation. One of the most widely used sources is the Helmet Detection Dataset available on Kaggle, which contains thousands of annotated images of riders with and without helmets under varied conditions. In addition, custom CCTV traffic footage datasets are generated to replicate real-world deployment environments. These datasets capture scenes from urban intersections, highways, and residential areas, ensuring coverage of different traffic densities, rider positions, and environmental conditions. For broader generalization, datasets such as the Stanford Cars Dataset [23]—originally designed for car recognition—are adapted to detect two-wheelers, particularly motorcycles and scooters, which dominate the traffic landscape in developing countries. Transfer learning is employed to fine-tune pretrained models on motorcycle-specific subsets, thereby accelerating convergence and improving detection performance. In some cases, synthetic datasets are created using data augmentation or generative models to simulate helmet variations, occlusions, and low-light conditions [24]. By combining public datasets with custom real-world footage, the training corpus becomes rich and diverse, significantly reducing overfitting and ensuring robustness across deployment scenarios.

V. RESULT ANALYSIS

To evaluate the system's effectiveness, standard metrics such as Accuracy, Precision, Recall, and F1-score are employed. Accuracy provides an overall measure of correct classifications, while Precision reflects the proportion of correctly identified violations among all flagged cases. Recall indicates the proportion of actual violations detected, and the F1-score balances Precision and Recall to provide a comprehensive measure of detection performance [25]. Experimental results demonstrate that YOLOv5 consistently outperforms Faster R-CNN in terms of real-time detection speed, making it highly suitable for live monitoring on CCTV feeds. While Faster R-CNN achieves slightly higher mAP (mean Average Precision) in certain static datasets, its inference speed is significantly slower, rendering it less practical for real-time traffic enforcement [26]. Helmet detection accuracy consistently exceeds 90% across benchmark datasets, with particularly strong performance in well-lit conditions. The system also maintains a low false-positive rate, meaning that riders wearing helmets are rarely misclassified as violators. However, performance degradation is observed under challenging scenarios such as nighttime videos, poor-resolution CCTV footage, and heavy occlusion where multiple riders are closely positioned. These results indicate that while the proposed system is reliable under most conditions, further optimization is required for deployment in complex real-world scenarios. The proposed system has several real-world applications that extend beyond basic traffic monitoring. The most direct application lies in real-time traffic monitoring by authorities, where violations can be automatically detected and flagged without requiring human supervision. Once a violation is confirmed, the system can integrate with automated challan generation platforms, issuing fines directly to offenders based on registered vehicle information. Another significant application is in smart city surveillance systems, where helmet violation detection modules can be integrated into existing traffic management platforms to enhance overall road safety. By continuously monitoring intersections and accident-prone zones, the system provides actionable data to authorities, enabling better decision-making. Additionally, the system contributes to road safety analytics and accident prevention by collecting statistical data on helmet usage trends across different regions and times. This information can support public awareness campaigns, policy-making, and targeted enforcement strategies. Beyond enforcement, the technology also holds potential in insurance and forensic investigations, where violation evidence can be automatically stored and used for claim validation or legal proceedings. By providing a scalable and automated solution, the system plays a pivotal role in creating safer transportation ecosystems.

VI. CONCLUSION & FUTURE SCOPE

In summary, automatic helmet rule violation detection using deep learning represents a transformative approach to road safety enforcement. By combining CNN-based object detection for rider localization with dedicated helmet classification models, the system achieves reliable performance in real-world traffic environments. The integration of YOLO and Faster R-CNN demonstrates the trade-off between speed and accuracy, while evaluation results confirm the feasibility of deploying such systems for real-time surveillance. The proposed solution significantly reduces reliance on manual inspection, offering scalable, automated, and evidence-based monitoring. With detection accuracies exceeding 90% on benchmark datasets and strong performance under favorable conditions, the system is well-positioned for practical adoption. However, challenges such as nighttime detection, occlusion handling, and dataset diversity highlight the need for continued research. As smart city infrastructure evolves, integrating helmet violation detection with broader traffic management ecosystems will enhance compliance, reduce accidents, and ultimately save lives. With advancements in AI, multimodal sensing, and edge computing, this technology can mature into a cornerstone of intelligent traffic law enforcement, providing long-term benefits to both urban safety and governance. Despite its promising results, several challenges must be addressed before large-scale deployment. Low-light and nighttime detection remains a significant limitation, as conventional RGB-based CNN models struggle with poor illumination and noise. Advanced approaches such as thermal imaging or infrared cameras could help overcome this limitation. Another challenge is occlusion, where multiple riders or overlapping vehicles obscure the helmet region, making detection unreliable. Handling variations in helmet types, colors, and designs, along with non-standard safety gear, further complicates classification. Additionally, varying camera angles, resolutions, and frame rates across surveillance infrastructure introduce inconsistencies that affect detection accuracy. Looking ahead, future research can integrate attention-based transformers and multimodal learning frameworks to enhance recognition performance. Transformers have shown significant promise in visual recognition tasks due to their ability to capture global dependencies, making them suitable for complex traffic scenes. Furthermore, coupling helmet detection with automatic number plate recognition (ANPR) will enable a complete end-to-end enforcement pipeline, where violations are detected, linked to vehicle registration, and penalized in real time. Incorporating edge computing solutions can further optimize performance by processing data locally, reducing latency, and alleviating bandwidth demands on centralized servers. In addition, the integration of traffic flow analysis and predictive modeling could expand the system's role from enforcement to proactive accident prevention, aligning with broader smart city initiatives. This holistic approach will not only improve detection accuracy but also ensure scalability and sustainability in large metropolitan deployments.

REFERENCES

[1] World Health Organization, Global Status Report on Road Safety 2018, Geneva: WHO, 2018.

[2] A. Kumar, R. Singh, and P. Gupta, "Helmet Rule Compliance and Road Safety in India: A Statistical Analysis," International Journal of Traffic and Transportation Engineering, vol. 9, no. 3, pp. 112–120, 2021.

[3] S. Sharma and M. Jain, "A review of automatic helmet detection techniques for intelligent traffic surveillance," Procedia Computer Science, vol. 133, pp. 1047–1054, 2018.

[4] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 886–893, 2005.

[5] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, pp. 273–297, 1995.

[6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.

[7] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint, arXiv:1804.02767, 2018.

[8] G. Jocher, A. Stoken, J. Borovec, and et al., "YOLOv5: A Scalable Object Detection Model," GitHub repository, 2020.

[9] W. Liu et al., "SSD: Single Shot MultiBox Detector," European Conference on Computer Vision (ECCV), pp. 21–37, 2016.

[10] H. Thakur and V. Verma, "Deep Learning Based Helmet Detection for Road Safety," International Journal of Computer Applications, vol. 182, no. 35, pp. 40–45, 2019.

[11] P. Dwivedi and S. Biswas, "Real-time Motorcycle Rider Helmet Detection using YOLO," International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1173–1178, 2020.

[12] A. Singh, R. Kumar, and M. Kaur, "Hybrid Object Detection and Classification Model for Helmet Violation Monitoring," Multimedia Tools and Applications, vol. 80, no. 11, pp. 17023–17040, 2021.

[13] Kaggle, "Motorcycle Helmet Detection Dataset," Kaggle Datasets, 2020. [Online]. Available: https://www.kaggle.com/

[14] Stanford AI Lab, "Stanford Cars Dataset," 2013. [Online]. Available: https://ai.stanford.edu/~jkrause/cars/car_dataset.html

[15] Y. Xu et al., "Data augmentation for object detection using generative adversarial networks," Neurocomputing, vol. 387, pp. 172–180, 2020.

[16] I. Goodfellow et al., "Generative Adversarial Nets," Advances in Neural Information Processing Systems (NeurIPS), pp. 2672–2680, 2014.

[17] M. Everingham et al., "The Pascal Visual Object Classes (VOC) Challenge," International Journal of Computer Vision, vol. 88, pp. 303–338, 2010.

[18] T. S. Fawcett, "An introduction to ROC analysis," Pattern Recognition Letters, vol. 27, pp. 861–874, 2006.

[19] R. Padilla, S. L. Netto, and E. A. da Silva, “A Survey on Performance Metrics for Object-Detection Algorithms,” International Conference on Systems, Signals, and Image Processing (IWSSIP), pp. 237–242, 2020.

[20] A. Mehta and D. Singh, “Performance Evaluation of YOLOv5 vs Faster R-CNN for Helmet Violation Detection,” Journal of Computer Vision and Pattern Recognition Research, vol. 4, no. 2, pp. 88–95, 2022.

[21] H. Li, Z. Wu, and Y. Hu, “Edge computing-based real-time helmet detection for traffic monitoring,” IEEE Internet of Things Journal, vol. 8, no. 14, pp. 11473–11482, 2021.

[22] M. Zhang et al., “Transformer-Based Models for Traffic Violation Detection,” IEEE Access, vol. 9, pp. 169432–169445, 2021.