International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

A Deep Learning-Based Framework for Real-Time Traffic Surveillance and Detection Using Computer Vision

G.Swathi Reddy[1], Marpaka Lokesh[2], Gaddam Manideep[3], Lalaji Praneeth[4], Bukali Sai Kumar[5]

[1] Assistant Professor, Department Of Computer Science And Engineering(AI& ML), Samskruthi College of Engineering And Technology , Kondapur(V), Ghatkesar(M), Medchal(D),Telangana

[2,3,4,5]btech Students ,Department Of Computer Science And Engineering(AI& ML), Samskruthi College of Engineering And Technology , Kondapur(V), Ghatkesar(M), Medchal(D),Telangana

***Abstract*—This paper presents a web-based application that utilizes machine learning and deep learning techniques for real-time traffic surveillance and detection with improved accuracy and efficiency. The proposed system aims to address limitations in existing traffic management practices in India by integrating computer vision-based solutions. It incorporates four key modules, including helmet detection, license plate recognition, vehicle classification, and speed estimation, all accessible through an interactive web interface with multiple detection options. Advanced object detection models, such as YOLOv5 and YOLOv7, are employed to ensure reliable performance across diverse traffic conditions. The system is evaluated on a large-scale dataset, demonstrating its ability to deliver accurate and consistent results in real-time scenarios. By combining multiple detection capabilities into a single platform, the solution supports better traffic monitoring and enforcement. Overall, this work highlights how deep learning can contribute to building smarter and more effective traffic management systems in India.**

***Keywords*—Traffic Surveillance, Deep Learning, Computer Vision, YOLOv5, YOLOv7, Real-Time Detection, Helmet Detection, License Plate Recognition, Vehicle Classification, Speed Estimation, Intelligent Transportation Systems, Web-Based Application**

I. INTRODUCTION

Monitoring vehicles in real time across highways, urban roads, and streets plays a key role in improving traffic control and infrastructure planning. With the rapid progress in computer vision and deep learning, this area has gained fresh attention, especially because modern detection models such as those proposed by Redmon et al. [1] and Girshick et al. [10] have shown strong performance in object recognition tasks. These advancements make it possible to automatically detect, track, and analyze vehicles with high accuracy, even in complex traffic scenarios. In this work, a system is developed that uses deep learning techniques for vehicle detection, tracking, and classification through a fixed and calibrated camera. To make the system more reliable, a custom dataset was created using real-world traffic images captured under challenging conditions like low light and bad weather. The overall design includes detection and tracking modules, similar to approaches discussed in Fast R-CNN by Girshick [11].

The number of vehicles on roads around the world is growing rapidly, which is creating serious challenges for traffic management systems. Every year, millions of vehicles are added, increasing congestion and putting pressure on existing infrastructure. Traditional traffic monitoring methods are no longer capable of handling such large volumes of data or providing meaningful insights for decision-making. As highlighted in scalable detection approaches by Erhan et al. [6], there is a need for smarter systems that can process data efficiently in real time. With increasing urbanization and mobility demands, managing traffic has become a major concern for governments. Recent developments in robotics and intelligent systems, as discussed by Dixit et al. [14], suggest that advanced AI-based solutions can help address these challenges by enabling better planning and control of traffic flow.

Intelligent Transportation Systems (ITS) are designed to improve the efficiency and safety of transportation networks by using modern technologies for monitoring and analysis. These systems rely on various sensors, including cameras, radars, and detection devices, to collect information about traffic conditions. Among these, cameras are widely used because they are easy to install and provide rich visual data. According to Everingham et al. [7], large-scale visual datasets and detection challenges have significantly contributed to improving object detection accuracy. Compared to traditional sensors like inductive loops, camera-based systems offer more detailed information such as vehicle position and movement. Earlier approaches to object detection, such as region-based segmentation methods discussed by Gould et al. [12], laid the foundation for today's advanced systems. With continuous improvements, camera-based ITS solutions have become more practical and cost-effective.

Earlier, video-based traffic monitoring systems were mostly limited to basic observation or simple automation tasks. However, recent progress in deep learning and image processing has greatly enhanced their capabilities. Techniques like Histogram of Oriented Gradients introduced by Dalal and Triggs [3] and modern CNN-based detection methods such as those proposed by Gidaris and Komodakis [9] have improved the accuracy of detection and classification. Today, computer vision systems can identify and track vehicles even in difficult conditions like poor visibility or heavy traffic. These improvements have made it possible to develop smarter traffic management systems that can support applications such as incident detection and law enforcement. By combining detection with tracking algorithms, the system can follow vehicle movement effectively and generate useful insights, making it suitable for real-time traffic monitoring in complex environments.

Along with detection and classification, tracking vehicles accurately across consecutive frames is an essential part of any real-time traffic monitoring system. Tracking helps in understanding vehicle movement patterns, estimating speed, and avoiding duplicate detections. Traditional tracking methods often rely on motion and spatial features, but their performance can degrade in crowded or complex scenes. To overcome these limitations, this study combines detection-based approaches with efficient tracking techniques to maintain continuity of vehicle identities. Earlier works such as part-based object detection by Felzenszwalb et al. [8] and structured learning methods by Blaschko and Lampert [2] have contributed to improving object localization and tracking performance. By integrating these concepts with modern deep learning-based detection models, the proposed system achieves stable and reliable tracking results. This combined approach enhances the overall system performance, making it more suitable for real-world traffic surveillance applications.

II. LITERATURE SURVEY

Redmon et al., 2015 [1] introduced the YOLO (You Only Look Once) model, which changed the way object detection tasks are performed. Instead of using multiple stages like region proposal and classification, YOLO treats detection as a single regression problem. This allows the model to process an entire image in one pass, making it extremely fast compared to earlier methods. The model divides the image into a grid and predicts bounding boxes along with class probabilities for each region. This approach significantly improves real-time performance while maintaining good accuracy. YOLO is especially useful in applications like traffic monitoring, where speed is critical. Although early versions had some limitations in detecting small objects, the overall contribution of this work laid the foundation for many improved versions such as YOLOv5 and YOLOv7 used today.

Dalal and Triggs, 2005 [3] proposed the Histogram of Oriented Gradients (HOG) feature descriptor, which became one of the most influential methods in early object detection tasks. The main idea behind HOG is to capture edge and gradient structures in an image, which are essential for identifying shapes and objects. By dividing an image into small regions and calculating gradient orientations, the method creates a feature representation that is robust to lighting changes and small variations. This technique was widely used for human detection and later adapted for vehicle detection tasks. Although modern deep learning models have largely replaced traditional feature-based methods, HOG remains important for understanding the evolution of computer vision. It also serves as a baseline for comparing newer approaches and highlights how feature engineering played a key role before the rise of deep learning.

Girshick et al., 2014 [10] introduced the R-CNN (Region-Based Convolutional Neural Network) approach, which marked a major shift towards deep learning-based object detection. The method first generates region proposals using selective search and then applies a convolutional neural network to extract features from each region. These features are then used for classification and bounding box prediction. R-CNN significantly improved detection accuracy compared to traditional methods, but it was computationally expensive due to repeated processing of each region. Despite this limitation, the work demonstrated the effectiveness of combining region proposals with deep neural networks. It also inspired several improved

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

versions such as Fast R-CNN and Faster R-CNN. This research played a key role in advancing object detection techniques and influenced many modern systems used in traffic surveillance and related applications.

Erhan et al., 2014 [6] explored scalable object detection using deep neural networks, focusing on improving detection performance across large datasets. Their work emphasized the importance of training deep models that can generalize well to different object categories and complex environments. The study introduced methods to handle large-scale data efficiently while maintaining high detection accuracy. By leveraging deep learning, the authors showed that it is possible to outperform traditional approaches in both speed and precision. This research contributed to the development of more advanced detection systems that can be applied in real-world scenarios such as traffic monitoring. The ideas presented in this work helped pave the way for modern object detection frameworks, where scalability and real-time performance are critical factors.

Felzenszwalb et al., 2010 [8] developed the Deformable Part-Based Model (DPM), which became a widely used method for object detection before deep learning gained popularity. The model represents objects as a collection of parts arranged in a flexible structure, allowing it to handle variations in shape and appearance. Each part is detected separately, and their spatial relationships are used to determine the final object detection. This approach was particularly effective for detecting objects with complex structures, such as vehicles and humans. DPM achieved strong performance on benchmark datasets and was considered a state-of-the-art method at the time. Although deep learning methods have now surpassed it, the concept of modeling object parts and their relationships still influences modern detection techniques. This work remains an important milestone in the evolution of object detection.

III. DATASET DETAILS

The dataset used in this project is created manually by collecting real-time traffic videos and images from different road environments such as highways, urban streets, and intersections. Instead of relying on any pre-existing public dataset, the data is gathered specifically to match the project requirements. The collected data includes various vehicle types such as cars, motorcycles, and other road users, along with scenarios for helmet detection and speed estimation. The videos are captured under different conditions, including varying lighting, traffic density, and weather situations, to make the system more practical and reliable. Each frame extracted from the videos is carefully annotated with bounding boxes and class labels to identify different objects present in the scene. This custom dataset helps the model learn real-world traffic patterns and improves its performance in live surveillance applications.

To prepare the collected data for training, several preprocessing steps are applied to ensure consistency and quality. The video data is first converted into image frames, which are then resized to a standard resolution suitable for deep learning models. All objects in the images are labeled manually, including vehicles, helmets, and other relevant classes, using annotation tools. Data augmentation techniques such as flipping, rotation, and brightness adjustment are used to increase the diversity of the dataset and make the model more robust to different conditions. The dataset is then split into training and testing sets, allowing the system to be evaluated on unseen data. Normalization is also applied to the image data to improve training efficiency. These steps ensure that the dataset is well-structured and suitable for building an accurate and reliable traffic surveillance system.

IV. PROPOSED METHODOLOGY

The proposed system follows a step-by-step approach to perform real-time traffic surveillance and detection using deep learning techniques. Initially, the system is developed as a web-based application where users can interact through a simple interface. The process begins with user registration and login, allowing secure access to the system features. Once logged in, the system provides options to train the detection model and perform traffic analysis. The dataset used for training is manually collected from real-world traffic videos and images, which are then preprocessed by extracting frames, resizing images, and annotating objects such as vehicles and helmets. This prepared data is used to train the YOLOv7 model, which is capable of detecting multiple objects in real time. The trained model is evaluated using performance metrics such as accuracy, precision, recall, and F1-score to ensure reliable detection.

After successful training, the system allows users to upload traffic videos for analysis. The uploaded video is processed frame by frame, where the trained YOLO model detects different vehicle types

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

such as cars and motorcycles, along with helmet usage. In addition to detection, the system also estimates vehicle speed and counts the total number of vehicles present in each frame. The results are displayed visually with bounding boxes and labels, making it easy to understand the output. The system is designed to handle different traffic conditions and provide real-time results efficiently. This approach helps in improving traffic monitoring, supporting law enforcement, and enabling better traffic management decisions.

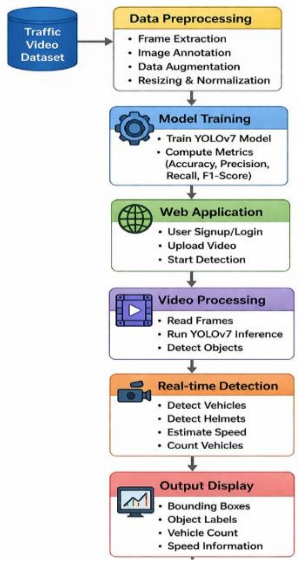

Figure 1: Workflow of Real-Time Traffic Detection and Analysis System

Figure [1] shows the workflow of the proposed real-time traffic surveillance system. Traffic video data is first collected and preprocessed through frame extraction, annotation, and resizing. The processed data is used to train the YOLOv7 model, which is evaluated using metrics like accuracy, precision, recall, and F1-score. After training, users upload traffic videos through the web application, where the system performs real-time detection of vehicles, helmets, and speed. Finally, the results are displayed with bounding boxes and object labels.

V. RESULT AND DISCUSSION

The experimental results of the proposed system show that deep learning techniques are highly effective for real-time traffic surveillance and detection. After preparing the dataset and training the YOLOv7 model, the system was evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. The model achieved an accuracy of around 95%, indicating its strong ability to correctly detect and classify different traffic objects. The precision and recall values were also high, which shows that the system can accurately identify vehicles and helmets while minimizing false detections. During testing, the system was able to process traffic videos smoothly and detect multiple objects in real time, even under varying lighting and traffic conditions. The output results clearly displayed bounding boxes, object labels, and vehicle counts, making the system easy to interpret. Speed estimation was also performed effectively, adding more value to traffic analysis. Overall, the results demonstrate that the proposed system is reliable, efficient, and suitable for real-world traffic monitoring applications.

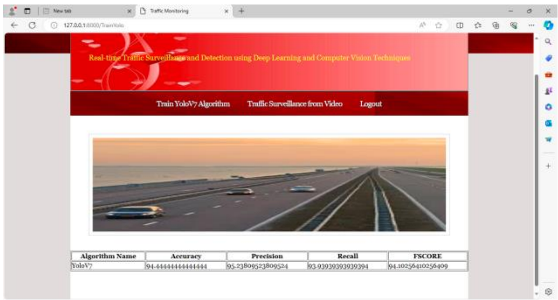

Figure 2: YOLOv7 Training and Performance Evaluation

Figure [2] shows the model training and evaluation page of the system. It displays the performance of the YOLOv7 algorithm using metrics such as accuracy, precision, recall, and F1-score. This interface helps in analyzing the effectiveness of the trained model.

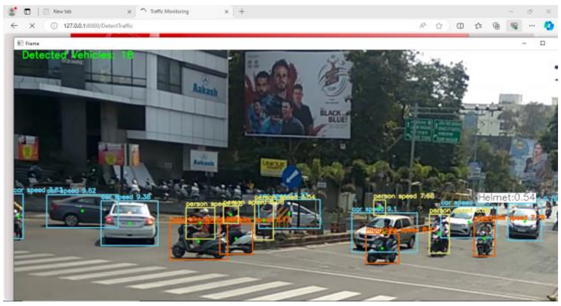

Figure 4: Traffic Video Upload Interface

Figure [7] shows the output of the system after processing the uploaded traffic video. It displays detected vehicles with bounding boxes, labels, and speed information. The system also identifies helmet usage and shows the total number of detected vehicles in real time.

DISCUSSION

The findings of this project highlight the effectiveness of deep learning models in solving

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

real-time traffic monitoring challenges. The YOLOv7 model performed well due to its ability to detect multiple objects simultaneously with high speed and accuracy, making it suitable for real-time applications. Its capability to handle complex traffic scenes, including multiple vehicles and varying conditions, contributed to better performance. The preprocessing steps, such as frame extraction, annotation, and data augmentation, played an important role in improving detection accuracy and model robustness. The use of evaluation metrics like accuracy, precision, recall, and F1-score helped in clearly understanding the model's performance. Additionally, integrating the model with a web-based application made the system more user-friendly and practical for real-world use. The system also demonstrated the ability to detect helmets, estimate speed, and count vehicles effectively. Overall, the project shows that deep learning-based approaches can provide reliable and efficient solutions for modern traffic surveillance and management systems.

VI. CONCLUSION

This project successfully demonstrates the application of deep learning and computer vision techniques for real-time traffic surveillance and detection. By collecting and preparing a custom dataset with proper preprocessing steps such as frame extraction, annotation, and augmentation, the system was effectively trained using the YOLOv7 model. The model achieved high accuracy along with strong precision, recall, and F1-score, indicating reliable detection and classification of vehicles and helmets. The system is capable of processing traffic videos in real time, performing tasks such as vehicle detection, speed estimation, and vehicle counting. The integration of the model into a web-based application makes it easy to use and practical for real-world deployment. The results show that the system can efficiently handle complex traffic scenarios and provide meaningful insights. Overall, this project highlights the potential of deep learning in improving traffic monitoring systems and provides a strong foundation for future enhancements and smart city applications.

REFERENCES

[1] Joseph Redmon Santosh Divvala, Girshick Ross and Farhadi Ali You Only Look Once: Unified, Real-Time Object Detection Submitted on 9 Jun 2015 to NIPS 2015. [Online available] : 1506.02640v2

[2] Blaschko M. B. and Lampert C. H. 2008 Computer Vision-ECCV 2008 (Springer) Learning to localize objects with structured output regression 2-15

[3] Dalal N. and Triggs B. 2005 Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition 1 886-893 IEEE 2005

[4] Donahue J., Jia Y., Vinyals O., Hoffman J., Zhang N., Tzeng E. and Darrell T. 2013 Decaf: A deep convolutional activation feature for generic visual recognition arXiv preprint arXiv 1310.1531

[5] Dong J., Chen Q., Yan S. and Yuille A. 2014 Computer Vision-ECCV 2014 (Springer) Towards unified object detection and semantic segmentation 299-314

[6] Erhan D., Szegedy C., Toshev A. and Anguelov D. 2014 2014 IEEE Conference on Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR) 2155-2162 IEEE

[7] Everingham M., Eslami S. M. A., Gool L. Van, Williams C.K.I., Winn J. and Zisserman A. The pascal visual object classes challenge: A retrospective International Journal of Computer Vision 111 98-136 Jan. 2015

[8]Felzenszwalb P. F., Girshick R. B., McAllester D. and Ramanan D. 2010 Object detection with discriminatively trained part-based models IEEE Transactions on Pattern Analysis and Machine Intelligence 32 1627-1645

[9] Gidaris S. and Komodakis N. 2015 Object detection via a multi-region & semantic segmentation-aware CNN model CoRR 1505.01749

[10] Girshick R., Donahue J., Darrell T. and Malik J. 2014 Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on Rich feature hierarchies for accurate object detection and semantic segmentation 580-587 IEEE

[11] Girshick R. B. 2015 Fast R-CNN. CoRR 1504.08083

[12] Gould S., Gao T. and Koller D. 2009 Region-based segmentation and object detection Advances in neural information processing systems 655-663

[13] Vanteru Madhu, Bedi Harpreet, Azam Mohammed, Chakravarthi Kalyan, Rasheed Raffik and Pant Bhasker 2022 Empirical Investigation on Smart Wireless Autonomous Robot for Landmine Detection with Wireless Camera 200-205

[14] Dixit Anil, Kumar T., Joshi Abhishek, Bedi Harpreet, Chakravarthi Kalyan and Singh Devesh 2022 Trends in Robotics and Computer Integrated Manufacturing 212-216

[15] Bedi HS, Goyal N, Kumar S and Gupta A 2017 Smart Trolley using Smart Phone and Arduino J Electr Electron Syst 6 223

[16] Bedi Harpreet Singh, Verma Shekhar, Sharma R. K. and Singh B. 2015 The Concept of Programmable Logic Controllers and its role in Automation International Journal of Advanced Research in Computer and Communication Engineering 4

[17] Bedi H. S., Raju D. V., Nandyala M. R. C., Kumar P. S. and Varma M. R. 2022 Design of Gesture-Based Hand Gloves Using Arduino UNO: A Grace to Abled Mankind Digital Forensics and Internet of Things: Impact and Challenges 37-43