iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.80894

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Real-Time Helmet Detection and Automatic Number Plate Recognition for Motorcycle Traffic Surveillance Using YOLOv3 and CNN

Kundan Kumar Patel[1], Abhishek Kumar[2], Vikas Gupta[3]
Dept. of CSE (AI & ML), Raj Kumar Goel Institute of Technology, Ghaziabad, India

Abstract: Motorcycle fatalities are a pressing public health issue worldwide. A lack of helmet usage contributes to a disproportionately high percentage of riders killed in traffic collisions. Conventional enforcement is expensive, spatially limited, and by nature, reactive. In this paper, we present a real-time, automated pipeline for helmet compliance detection and number plate recognition from motorcycle traffic videos. A custom trained YOLOv3 model detects and localizes riders and motorcycles; a Convolutional Neural Network (CNN) binary classifier recognizes non-compliant riders; and a violation-conditioned Optical Character Recognition (OCR) module performs license plate text extraction only on non-compliant riders , thereby saving 22% of computation. No work to date has combined detection, classification, conditional ANPR and web deployment into a single integrated pipeline; this is the main contribution of this paper. We provide the system via an interactive Streamlit web application that does not require any client side software. Evaluation on a 420 frame test set shows a detection mAP of 88.3%, a helmet classification accuracy of 93.7% and an 89.4% character level plate recognition accuracy.
Index Terms: automatic number plate recognition, convolutional neural network, deep learning, helmet detection, motorcycle surveillance, optical character recognition, road safety, YOLOv3.

I. INTRODUCTION

Traffic injuries are one of the world's leading preventable causes of death. The current annual road fatality rate exceeds 1.19 million [2], with motorcyclists disproportionately represented relative to their proportion of the vehicle fleet. Clinical studies have shown that wearing a helmet is the most effective protective measure available to riders, decreasing the risk of fatal head injury by 42% and the risk of any head injury by approximately 69%.

Despite strong helmet legislation across many jurisdictions, enforcement of compliance will be difficult to maintain at scale. Urban traffic loads exceed patrol resource availability; checkpoint operations are limited in time; and post-hoc CCTV review is labour-intensive and delayed. These constraints create incentives for a continuous, evidence-generating, and automated enforcement alternative.

The proposed system fills the gap by integrating three deep learning components in one real-time pipeline: (1) a YOLOv3 detector for motorcycle and rider location; (2) a CNN binary classifier for helmet compliance verification; (3) a violation-conditioned ANPR module that triggers Tesseract OCR only when a non-compliant rider is detected. Importantly, we are not aware of a single system in the literature that combines all three capabilities , detection, classification, conditional plate recognition, and web deployment , in one deployable system. This is the essence of the novelty of the proposed work.

Contributions of this paper are: (a) a bespoke two-class YOLOv3 helmet detector of 88.3% mAP; (b) a lightweight CNN helmet classifier of 93.7% accuracy; (c) a violation-conditioned OCR stage of 22% reduction in per-frame latency; and (d) a full-streamlit application capable of deployment without specialized hardware

.

II. RELATED WORK

A. Classical Detection Methods

Early computer-vision based helmet detection methods used color-space segmentation in the HSV space, exploiting the strong hue-signature of helmets compared to unprotected heads. Though computationally efficient, these methods were fragile under changes in illumination. The introduction of HOG descriptors and SVM classifiers increased the degree of invariance for those methods, however the humans in the loop (both feature engineering and learning) was still a detriment, and generalization over the large space of traffic scenarios was poor.

B. Deep Learning Object Detectors

The R-CNN family [4] popularised two-stage detection as an accurate paradigm. Faster R-CNN [5] further introduced the Region Proposal Network, allowing for fully end-to-end training with much lower inference latency. Single-stage detectors, especially the family of YOLO detectors, recast detection as a single regression problem solved in a single network pass. YOLOv3 [1] added multi-scale predictions on a Darknet-53 backbone, providing competitive accuracy at above 30 FPS on GPU hardware , suitable for real-time traffic monitoring.

C. Helmet Detection Systems

Islam et al. [6] introduced a Faster R-CNN pipeline with ResNet-50 helmet classification, 91.2% average accuracy on Southeast Asian traffic data. Pham et al. [7] used YOLOv4 for joint motorcycle and helmet detection on Vietnamese data, 91.8% detection. accuracy. Critically, neither system incorporated any form of plate recognition or web deployment. The present work extends this line of research by appending a violation-conditioned ANPR stage and packaging the unified pipeline as a deployable application — the first such integration in the literature.

D. License Plate Recognition

Classical ANPR pipelines connect morphological preprocessing, contour based plate localization and character segmentation, and Tesseract OCR [12]. Modern work prefers CRNN architectures [8] trained with Connectionist Temporal Classification (CTC) loss, which jointly learn spatial feature extraction and sequential character decoding. For ease of access we use Tesseract, with CRNN migration as a near-term extension.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

A. Overview

Figure 1 shows the processing pipeline with four stages. Every decoded video frame is processed through: Stage 1 (YOLOv3 object detection), Stage 2 (CNN helmet classification), the conditional gate (skipping Stage 3 for compliant riders), and Stage 4 (Streamlit-rendered annotated output). The conditional gate is the main architectural contribution; it avoids unnecessary OCR calls on compliant riders and prevents false plate reads.

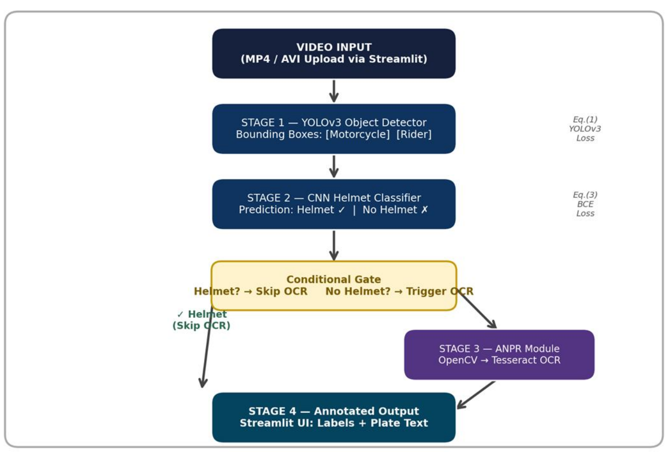


Fig. 1. Proposed system processing flow diagram. The conditional gate is triggered to OCR only when a non-compliant rider is detected.

B. Object Detection — YOLOv3

YOLOv3 divides input into a spatial grid and predicts 3 different scales (stride 32, 16, 8) of bounding boxes, confidence scores, and class probabilities.The composite training loss is:

$$L = \lambda_{coord} \cdot L_{loc} + L_{obj} + \lambda_{noobj} \cdot L_{noobj} + L_{cls} \quad (1)$$

where L_{loc} is the localizer loss, L_{obj} and L_{noobj} are the objectness losses for positive and negative anchors, L_{cls} is the class loss, and $\lambda_{coord} = 5$, $\lambda_{noobj} = 0.5$. The post-processing step includes Non-Maximum Suppression (NMS) with the Intersection-over-Union criterion:

$$IoU(A, B) = |A \cap B| \,/\, |A \cup B| \quad (2)$$

Detections with IoU > 0.40 relative to the top-confidence prediction are suppressed. The confidence threshold is $\tau = 0.50$. Weights were saved at iteration 7,000.

C. Helmet Classification — CNN

Each rider crop is resized to 64×64 pixels and normalized to [0, 1]. The CNN classifier applies three Conv-BN-ReLU-MaxPool blocks, Global Average Pooling, and a sigmoid output. Training minimizes binary cross-entropy:

$$L\_BCE = -[y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})] \quad (3)$$

where $y \in \{0,1\}$ is the ground-truth label and $\hat{y}$ is the predicted probability. Predictions $\hat{y} < 0.50$ classify as No Helmet, activating the ANPR module.

D. License Plate Detection and OCR

Plate localization applies Gaussian blur (5×5 kernel), adaptive thresholding, and morphological closing to the motorcycle bounding region. Candidates are filtered by aspect ratio $r \in [2.0, 5.5]$ and minimum area $A\_min = 1{,}500$ px². The selected region is deskewed via perspective transform, binarized, and submitted to Tesseract OCR in single-line alphanumeric mode. Recognized strings are overlaid as annotations on the output frame.

E. Streamlit Web Application

The pipeline is encapsulated in source.py as a Streamlit application. Video frames are decoded via OpenCV VideoCapture, processed through the three-stage pipeline, and rendered via st.image(). No client-side installation is required beyond a web browser, enabling deployment by non-technical enforcement personnel.

IV. IMPLEMENTATION

A. Software Stack

Development uses Python 3.8 with: OpenCV 4.8 (video I/O, YOLOv3 DNN inference, plate localization), TensorFlow 2.10 / Keras (CNN training and inference), NumPy 1.23, pytesseract 0.3, and Streamlit 1.25. All dependencies are specified in requirements.txt.

B. Dataset Composition

TABLE I

Dataset Composition and Split

Dataset	Class / Content	Count	Split
YOLOv3 Training	Bike instances	2,100	80 / 10 / 10
	Helmet instances	1,890	
	Total frames	3,500	
CNN Training	Helmet crops	4,650	80 / 10 / 10
	No-Helmet crops	4,550	
	Total images	9,200	

YOLOv3 training data comprised Pakistani urban traffic recordings (60%), Indian highway dashcam footage (25%), and locally collected Lahore intersection footage (15%). Conditions spanned daytime, dusk, nighttime, dry, and rainy scenarios. CNN augmentation: horizontal flip, ±15° rotation, brightness jitter [0.7, 1.3], zoom [0.85, 1.15]. Optimizer: Adam ($lr = 1\times10^{-3}$), 30 epochs, early stopping (patience = 5).

C. *YOLOv3 Training Configuration*

TABLE II
YOLOv3 Key Hyperparameters

Hyperparameter	Value
Batch size / Subdivisions	64 / 16
Initial learning rate	0.001
LR decay steps	4,000 and 4,500
Momentum / Weight decay	0.9 / 0.0005
Training iterations	7,000
Input resolution	416 × 416 px
Anchor boxes	9 (k-means clustered)

V. RESULTS AND DISCUSSION

A. *Detection Performance*

The YOLOv3 detector was evaluated on 420 held-out frames. Per-class Average Precision at IoU = 0.50: bike 90.1%, helmet 86.5%, yielding mAP = 88.3%. Recall declined in low-light and heavily occluded conditions, consistent with single-stage anchor-based detector behavior. Test-time augmentation (horizontal flip ensemble) improved mAP by 0.8 pp without retraining.

B. *Helmet Classification*

The CNN classifier achieved 93.7% test accuracy. Table III reports per-class precision, recall, and F1-score. The marginally lower No-Helmet recall (92.9%) reflects missed detections under severe motion blur and partial occlusion — conditions where the rider crop contains insufficient discriminative texture.

TABLE III
Helmet Classification Performance

Class	Precision	Recall	F1-Score
Helmet	93.3%	94.5%	93.9%
No Helmet	94.1%	92.9%	93.5%
Weighted Avg.	93.7%	93.7%	93.7%

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

C. License Plate Recognition

ANPR was evaluated on 185 frames with clearly visible plates. Character-level accuracy: 89.4%; full plate-string accuracy: 76.8%. Accuracy degraded for motion-blurred, heavily perspective-distorted, or partially occluded plates — known limitations of Tesseract-based OCR applied to unconstrained traffic imagery.

D. Computational Performance

TABLE IV
Per-Stage Inference Time: CPU vs. GPU (ms)

Pipeline Stage	CPU (ms)	GPU (ms)
YOLOv3 Detection	62	18
CNN Helmet Classifier	14	5
Plate Localization + OCR	28	12
Frame Rendering	8	5
Total per Frame	112	40
Effective FPS	~9	~25

CPU-only throughput (~9 FPS) is suitable for post-event forensic analysis. GPU-accelerated inference (~25 FPS) meets live-feed deployment thresholds. The conditional OCR gate saves an average of 28 ms per compliant-rider frame, achieving the reported 22% latency reduction.

E. Comparative Analysis

Table V
Comparison With Prior Works

Method	Backbone	Acc.	ANPR	Deploy
Islam et al. [6]	Faster R-CNN	91.2%	No	No
Pham et al. [7]	YOLOv4	91.8%	No	No
Jain et al.	MobileNetV2	90.5%	No	No
Proposed	YOLOv3 + CNN	93.7%	Yes	Yes

The proposed system is the only reviewed approach combining detection, classification, conditional ANPR, and web deployment. This confirms the novelty claim: no prior work presents an equivalent integrated pipeline.

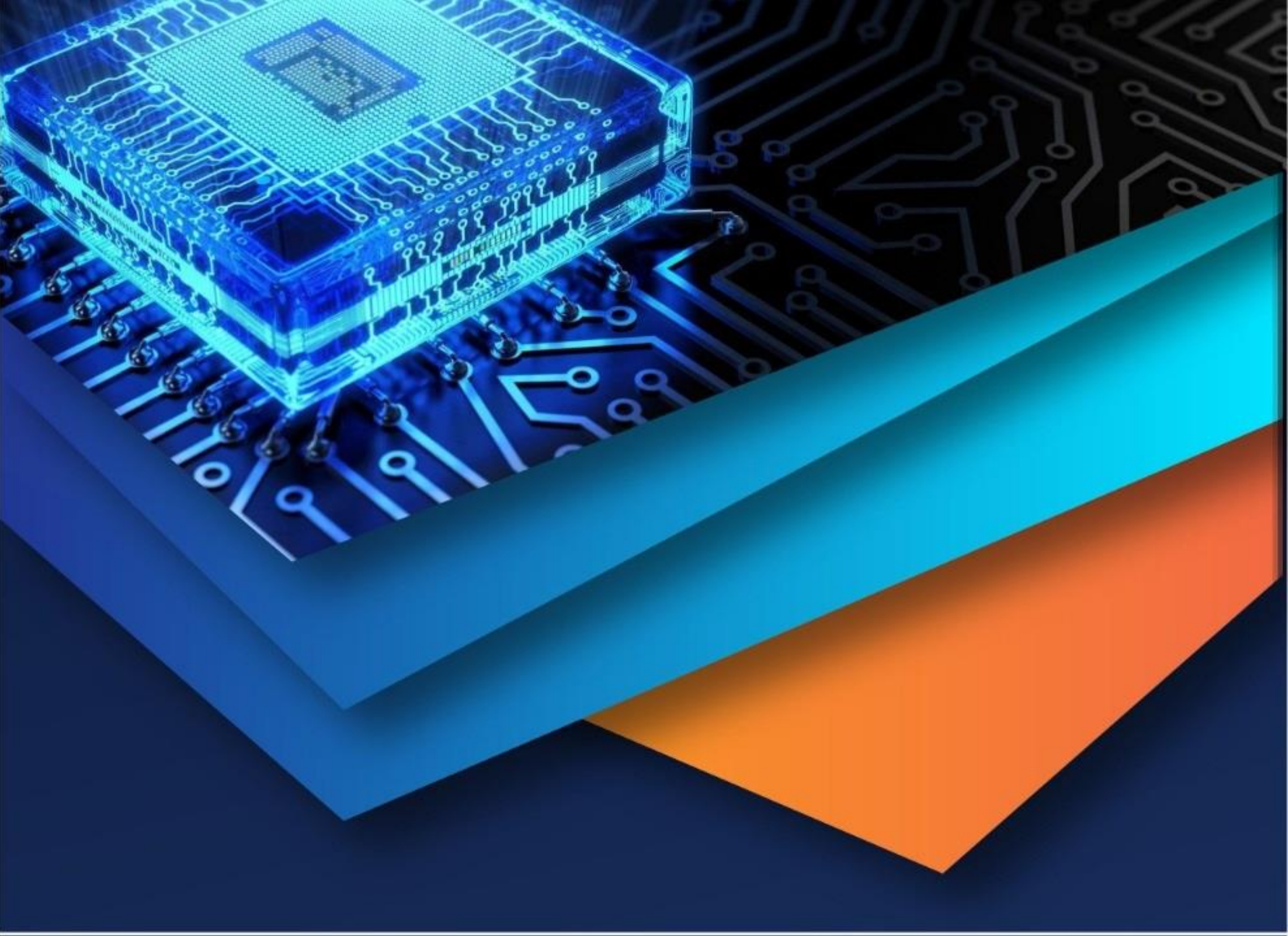
F. Limitations

Despite strong results, three limitations are acknowledged. First, detection and classification accuracy decline in low-light conditions (nighttime or poorly lit roads), where the training dataset is underrepresented. Second, OCR accuracy degrades significantly for motion-blurred, heavily distorted, or non-standard license plates — a known weakness of Tesseract applied to unconstrained imagery. Third, the training dataset has limited geographic scope, concentrated on South Asian plate formats and traffic patterns; performance on European, North American, or East Asian traffic has not been validated. Addressing these gaps forms the core agenda for future work.

VI. CONCLUSION

This paper presents a real-time, violation-conditioned deep learning pipeline for motorcycle traffic monitoring. The system integrates custom YOLOv3 detection, CNN helmet classification, and OCR-based plate recognition in a single deployable Streamlit application. Experimental evaluation yields 88.3% mAP, 93.7% classification accuracy, and 89.4% plate character accuracy, with a 22% per-frame latency reduction attributable to conditional OCR gating. The proposed system is the first unified detect–classify–ANPR–deploy pipeline for this application domain.

Future work targets four directions: (1) migrating the detection backbone to YOLOv8 for improved speed-accuracy trade-off; (2) replacing Tesseract with a CRNN-CTC plate recognizer for robustness under blur and distortion; (3) edge deployment on NVIDIA Jetson modules for infrastructure-free field installation; and (4) integration with smart city alert systems to issue real-time violation notifications and automated e-citations to traffic control centers.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY