INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 7, July 2026 | ISSN: 2456-4184 | IJNRD.ORG

AN EDGE–CLOUD DEEP-LEARNING FRAMEWORK FOR SUSTAINABLE VEHICLE-COMPLIANCE MONITORING USING YOLOV8 AND CNN-BASED HSRP AUTHENTICATION

[1]Sonal S. Kumbhar, [2]Naresh A. Kamble

[1]Student, [2]Professor

[1]Department of Data Science, D. Y. Patil Agriculture & Technical University, Talsande, Maharashtra, India

Abstract : Rapid motorization is straining the capacity of manual road-side enforcement, and non-compliant vehicles those carrying counterfeit plates, lapsed insurance, or operating well beyond their fitness life impose disproportionate safety and environmental costs on growing cities. Most reported number-plate systems stop at reading characters and rarely connect recognition to the wider question of whether a vehicle is legally and environmentally fit to be on the road. This paper presents an edge–cloud framework that couples a YOLOv8 detector with a purpose-built convolutional neural network (CNN) so that optical character recognition (OCR) and physical High-Security Registration Plate (HSRP) authentication run as two concurrent streams, after which a configurable rule engine screens insurance validity and vehicle age against a policy-driven threshold. The models were developed and evaluated on a custom Indian HSRP dataset (6,000 plate images spanning day, night, rain, motion-blur and low-resolution conditions, split 70/15/15 on a vehicle-disjoint basis). On the held-out test set the YOLOv8 plate detector reached mAP@0.5 = 98.1% and mAP@0.5:0.95 = 74.3%; the HSRP classifier reached 96.75% accuracy, 97.27% precision, 96.20% recall and an F1-score of 96.73%; OCR reached 95.84%-character accuracy and 92.7% exact-plate accuracy; and the complete pipeline delivered an end-to-end latency of about 33 ms per frame (≈30 FPS) on an NVIDIA RTX 3060. A sustainability assessment expresses the framework's potential contribution to SDG 11 and SDG 13 through an emissions-avoidance model driven by the number of flagged overage vehicles. The results indicate that combining edge-based plate analytics with configurable regulatory screening is a workable route to sustainable smart-city transport management; full deployment, however, requires authorized database access and jurisdiction-specific policy rules.

IndexTerms - **sustainable smart cities; intelligent transportation systems; YOLOv8; HSRP authentication; edge–cloud computing; vehicle-compliance monitoring; SDG 11.**

I. INTRODUCTION

INTRODUCTION

The number of urban vehicles is growing quicker than the agencies designed to oversee them. As cities become denser, the work of making sure that a car on the road is real, insured and roadworthy remains highly manual and inefficient, and hard to maintain over high-speed corridors or busy junctions. It has administrative repercussions. Vehicles with altered plates avoid responsibility, uninsured vehicles increase risk of others, lack of accountability for aged vehicles that are driving beyond their useful life increase the level of local air pollution disproportionately 27,28. The inability to differentiate between reading a plate and determining if it's legally and environmentally fit has become a real challenge to cleaner and safer mobility for municipal authorities trying to achieve the Sustainable Development Goals – especially SDG 11 on sustainable cities and SDG 13 on climate action 29,30.

Intelligent Transportation Systems (ITS) can help fill in some of that gap by automating perception at the road-side 21. In the world of ITS, Automatic Number Plate Recognition (ANPR) is the technology most commonly used for tolling, access control and traffic analytics 1. Deep Convolutional Networks (DCNs) have been used to create modern detectors that can localise plates and read the characters in a fast and reasonably robust manner, even under challenging illumination and viewing conditions 9,15. However most of the ANPR pipelines still regard the recognised string as the final product. On its own, a valid plate reading is not an indicator of whether the plate is real, if the car is insured, or if they can legally be on the road or not 2,3.

In a few countries such as India where the registration is very stringent the physical authenticity of the plate is itself a regulated property. High-Security Registration Plates (HSRP) are plates that contain tamper-resistant features, such as a chromium-based hologram, a laser-etched permanent identification number and a hot-stamped border film; they are used specifically because the theft and counterfeiting of plates undermines every downstream enforcement action 3. An ability to read characters, but not distinguish a real HSRP from a fake plate, means a loss of one of the most critical compliance cues on the road-side. The additional two signals — insurance validity and vehicle age with respect to the vehicle's applicable fitness policy — can only be determined by connecting with an authoritative vehicle registry 8.

We overcome these limitations by providing an end-to-end framework, which is compliance-centric and not recognition-centric. A YOLOv8 detector locates the vehicle and its plate 2 followed by processing of the cropped plate by two parallel streams: an OCR stream to get the registration string 5 and a CNN stream to verify the physical HSRP security features 3. The recognised identity is then passed to a configurable decision engine that checks the HSRP authenticity, insurance validity and age of the vehicle against a policy-defined threshold—with the compute explicitly partitioned between an edge node and a cloud service to ensure low latency and minimise exposure of recognisable plate imagery 22.

This paper makes the following important contributions. First, we present a dual-stream plate-processing design where the character recognition and physical HSRP authentication are conducted simultaneously on the same crop instead of sequentially. Second, three different regulatory checks (plate authenticity, insurance validity and vehicle age fitness) are brought together into a single programmable compliance engine, with an age threshold defined as a policy function instead of a fixed value. Third, an edge–cloud partitioning approach is applied where the perception is performed at the road-side and registry look-ups are performed at a secured cloud broker. Fourth, we express and quantify the environmental relevance of the system by using an emissions-avoidance model, highlighting that a vehicle-compliance monitoring system is not only a stand-alone enforcement activity, but also a supporting tool for sustainable urban mobility. These are explored experimentally in Section 5.

II. NEED OF THE STUDY.

Previous integrated ANPR research has focused on improving detection rates 2,4, localization of HSRP 3 and privacy preserving training for plate classification 8,23. One area that has not been sufficiently explored is a single pipeline that will physically authenticate the plate as it reads it, configuration a rule set that varies by jurisdiction, and will report its own environmental impact in measurable terms when it is used for insurance and age compliance. This paper is aimed at that intersection. We are wary of novelty claims: the individual learning components are sound, and what we bring is their composition with a focus on compliance, a programmable policy engine, and the framing of sustainability, all of which are backed up by tests on a vehicle-disjoint test partition.

The rest of the paper is organised as follows. Section 2 presents related work on ANPR, HSRP authentication, edge–cloud transport systems and sustainability oriented ITS. Section 3 presents the data set, models and decision logic. The edge–cloud framework and its algorithm are summarized in Section 4. Section 5 reports the detection, OCR, HSRP-classification, compliance, ablation, robustness and sustainability results. Implications, privacy and limitations are discussed in Section 6, followed by a conclusion in Section 7.

2.1 Deep-learning ANPR and plate detection

Single-stage detectors revolutionized object detection by predicting boxes and classes in a single shot, sacrificing some precision for significant improvements in speed 9,10 and the YOLO family has been the backbone of real-time plate localisation since. Al-Hasan et al. developed an optimized ANPR system using anchor-free YOLOv8s model which is different from the other YOLO models, where the classification and regression heads are separated to provide a high inference speed of more than 30 FPS and avoid localization errors due to camera rotation and light variations 2. They have achieved good mean Average Precision on character extraction but, as the authors themselves state, they only end at the text detection and do not have any extra logic to confirm compliance parameters or to match the reading to an external registry. Mulia et al. compared YOLOv8 to Faster R-CNN 12 with super-resolution pre-processing, and demonstrated that learned detectors are much more robust to environmental variation than classical segmentation based on edge detection 18 or scale-invariant descriptors 19 and specifically suggested a unified multi-tier framework for multi-objective compliance within a single pipeline 4.

2.2 HSRP authentication and character recognition

Rani et al. specifically addressed High-Security Registration Plates (HSRP), where they combined YOLOv5 localiser with multi-layer CNN and OCR engine to distinguish between genuine HSRP and conventional registration plates based on the blue chromium hologram as a discriminative feature 3. They provide a detailed analysis but performance suffers significantly with acute motion blur and the study does not make use of the recognised text to fetch external details like insurance status. In the field of recognition, character transcription can be achieved using the well-developed Tesseract 16 or sequence models like CRNN 17, while some studies use the combination of YOLOv8 and EasyOCR or Tesseract for end-to-end plate reading 5,6,7. These pipelines recover the registration string in a reliable way, but again, as the terminal output not as a key for a compliance decision.

2.3 Edge–cloud and privacy-preserving transport analytics

Because raw traffic imagery is both bandwidth-heavy and privacy-sensitive, there is growing interest in pushing perception to the edge and reserving the cloud for lightweight look-ups 22. Ashkanani et al. embedded YOLOv8 detection and plate reading inside a self-adaptive signal controller to optimise green-time at intersections, demonstrating the value of local inference for responsive traffic management, though downstream compliance was again out of scope 1. On the confidentiality side, federated approaches have been proposed to train plate models without centralising raw images: Kong et al. designed a federated recognition scheme for 5G-enabled vehicular networks 23, Badamikar et al. applied federated learning to HSRP classification across distributed clients 8, and broader surveys map the opportunities and open problems of federated learning for connected vehicles and smart cities 24,25,26. We draw on this literature as motivation for edge-side processing and data minimisation, but — unlike systems that train collaboratively — the framework in this paper keeps sensitive imagery at the edge through architectural partitioning rather than through federated optimisation, and we therefore do not claim a federated-learning result.

2.4 Sustainability-oriented intelligent transport

The environmental case for smarter enforcement is increasingly explicit. Emission-factor studies show that both exhaust and non-exhaust emissions scale with vehicle mass, condition and age, so that a small population of older or poorly maintained vehicles can dominate local pollutant loads 27, and fleet-level analyses link vehicle turnover and electrification directly to future air-quality trajectories 28. Work on sustainable urban mobility positions data-driven ITS as an instrument for meeting SDG 11 and SDG 13 targets 29,30. Against this backdrop, a compliance monitor that can identify and flag overage or non-conforming vehicles is not merely an enforcement tool; it is a lever for emissions reduction, which is the perspective this paper adopts.

Summary of the gap. In each of these strands, detection is well developed, HSRP authentication has been realised in isolation, edge and privacy techniques are well studied but, to our knowledge, the single pipeline that both detects and authenticates the plate, that screens for insurance and age through a configurable policy engine, and that reports on its own sustainability contribution, has not been consolidated and evaluated together. This is the void that this work inhabits.

III. RESEARCH METHODOLOGY

The data, the two learning models, the compliance logic, the prototype registry interface, the evaluation protocol are described in this section.

3.1 Dataset, annotation and pre-processing

Model development used a custom dataset of Indian vehicle images captured under five operating conditions — daylight, night-time, rain, motion blur, and low-resolution capture — so that both detection and authentication would be exercised across the range of situations a road-side camera actually encounters. Images were acquired from fixed and hand-held cameras at viewing distances of roughly 3–15 m, at resolutions ranging from 720p to 4K, and included both still frames and frames sampled from short video clips. Each plate region was annotated for the detection task, and every plate crop was additionally labelled as genuine HSRP or non-HSRP for the authentication task; where visible, the hologram, laser PIN and border strip were marked as auxiliary cues. Annotation was performed in Roboflow/LabelImg format by two annotators, with a third adjudicating disagreements, and label quality was checked on a random 10% audit sample. The working composition of the dataset is summarised in Table 1. To reduce the risk of identity leakage between partitions, the train/validation/test split was made vehicle-disjoint — no vehicle appears in more than one partition — and class balance between HSRP and non-HSRP was maintained within each split. Standard augmentation (random scaling, small rotations, brightness/contrast jitter and synthetic blur) was applied to the training partition only.

Table 1. Composition of the HSRP dataset.

Subset	Images	Unique vehicles	HSRP	Non-HSRP	Day	Night	Rain / blur
Train	4,200	1,470	2,100	2,100	2,020	1,180	1,000
Validation	900	315	450	450	430	250	220
Test	900	315	450	450	430	250	220
Total	**6,000**	**2,100**	**3,000**	**3,000**	**2,880**	**1,680**	**1,440**

Every plate crop entering the OCR stream was pre-processed to suppress noise and normalise contrast. The crop was converted to grayscale, filtered with a Gaussian kernel to remove salt-and-pepper noise, and binarised using Otsu's adaptive thresholding, which selects a global threshold by maximising between-class variance of the intensity histogram 20. This sequence stabilises character edges under uneven illumination and mild motion blur before recognition.

3.2 Vehicle and plate detection (YOLOv8)

YOLOv8 is selected for detection in two cascaded stages because of its anchor-free head and favourable speed–accuracy trade-off when detecting small objects such as license plates 2,11. The general detector first localises vehicle instances, including cars, trucks, motorcycles, and three-wheelers, and establishes a region of interest that suppresses background clutter. A second, more refined detector then operates only within that region and returns tight plate coordinates. All frames are resized to 640×640pixels for inference, while boxes are filtered using a confidence threshold of 0.25 and a non-maximum-suppression IoU threshold of 0.7. The Intersection over Union (IoU) between a predicted box and its ground-truth box is used to measure localisation quality, as shown in Eq. (1).

$$\mathrm{IoU} = \frac{\mathrm{area}\left(B_{\mathrm{pred}} \cap B_{\mathrm{gt}}\right)}{\mathrm{area}\left(B_{\mathrm{pred}} \cup B_{\mathrm{gt}}\right)} \tag{1}$$

3.3 HSRP Authentication (Custom CNN)

In parallel with recognition, the plate crop is examined by a compact custom CNN that determines whether the physical security features of a genuine HSRP are present: the blue chromium hologram at the upper-left, the laser-engraved permanent identification number, and the hot-stamped border film bearing the repeated "INDIA" script 3. The network is a binary classifier consisting of four convolutional blocks, each comprising a convolution operation followed by batch normalisation, ReLU activation, and max-pooling, leading to global average pooling, a fully connected layer with dropout, and a single sigmoid output. The layer configuration is given in Table 2, and the training settings are presented in Table 3. Using a single sigmoid unit, rather than a two-node softmax, is the natural choice for a two-class decision and keeps the probabilistic interpretation unambiguous.

Table 2. Layer-wise configuration of the HSRP authentication CNN.

#	Layer	Kernel / units	Output	Activation
1	Conv + BN + MaxPool	3×3, 32 filters	112×112×32	ReLU
2	Conv + BN + MaxPool	3×3, 64 filters	56×56×64	ReLU
3	Conv + BN + MaxPool	3×3, 128 filters	28×28×128	ReLU
4	Conv + BN + MaxPool	3×3, 256 filters	14×14×256	ReLU
5	Global Avg Pool	—	256	—
6	Dense + Dropout (0.5)	128 units	128	ReLU
7	Dense (output)	1 unit	1	Sigmoid

The classifier is trained by minimising the binary cross-entropy loss in Eq. (2), where $y$is the ground-truth label $(1 = \text{genuine HSRP})$and $p$is the predicted probability. For a feature vector $z$obtained from the final pooling layer, the HSRP probability is computed using the sigmoid response in Eq. (3).

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)\right] \tag{2}$$

$$P_{\mathrm{HSRP}} = \sigma(\mathbf{w}^T\mathbf{z}+b) = \frac{1}{1 + \exp[\ -(\mathbf{w}^T\mathbf{z}+b)]} \tag{3}$$

Rather than fixing an arbitrary cut-off, the decision threshold was selected using the validation set. A plate is accepted as a genuine HSRP when its probability exceeds a threshold τ, chosen to maximise Youden's Jstatistic on the validation ROC curve and cross-checked against F1-score maximisation. This procedure yielded $\tau \approx 0.75$, which is subsequently reported together with ROC-AUC and PR-AUC on the untouched test set in Section 5.3. Expressing the operating point as an optimisation over the validation set, rather than as a hand-picked constant, makes the selection reproducible and allows an operator to retune τtoward higher recall in enforcement settings where missed counterfeits are more costly than false alarms.

Table 3. Training configuration of the HSRP CNN.

Setting	Value
Input size	224 × 224 × 3
Optimiser	Adam
Learning rate	1×10^{-3} (cosine decay)
Loss	Binary cross-entropy

Batch size	32
Epochs	60 (early stopping, patience 8)
Class weighting	Balanced (1:1)
Threshold τ	0.75 (Youden's J on validation)

3.4 Optical character recognition

The recognition stream converts the pre-processed crop into a registration string. After segmentation of the character blobs, an OCR engine (EasyOCR by default, with Tesseract [16] and PaddleOCR evaluated as alternatives in Section 5.5) transcribes the sequence. Two metrics are reported: character-level recognition accuracy, and the stricter exact-plate accuracy, which counts a plate as correct only when the entire string matches the ground truth. The exact-plate metric is the operationally meaningful one, because a single mis-read character invalidates the subsequent registry look-up.

3.5 Configurable compliance decision engine

If the registration string and the HSRP status are available, three compliance signals are evaluated. A signal value of 1 always corresponds to compliant, whereas a signal value of 0 always corresponds to non-compliant. To check insurance validity, the current date is compared with the retrieved expiry date, as given in Eq. (4). The vehicle age is calculated as the difference between the current year and the registration year, as given in Eq. (5).

$$S_{\text{ins}} = \begin{cases} 1, & \text{if } T_{\text{current}} \leq T_{\text{expiry}}, \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

$$A_v = Y_{\text{current}} - Y_{\text{registration}} \quad (5)$$

Importantly, no fixed age is programmed into the age rule. The maximum permissible age depends on the vehicle class, fuel type, jurisdiction, and the policy currently in force. Therefore, it is expressed as a policy function in Eq. (6), and the age signal is evaluated against it in Eq. (7). This is legally safer and operationally more realistic than automatically classifying a vehicle as scrap-eligible once it exceeds an assumed age limit.

$$L_{\max} = f(\text{class, fuel, jurisdiction, policy date}) \quad (6)$$

$$S_{\text{age}} = \begin{cases} 1, & \text{if } A_v \leq L_{\max}, \\ 0, & \text{otherwise, with the vehicle flagged for review} \end{cases} \quad (7)$$

The overall verdict is the logical conjunction of the three signals, as shown in Eq. (8). A vehicle is considered compliant only when the plate is authentic, the insurance is valid, and the vehicle age is within the applicable policy limit. An incident record is created whenever any signal has a value of zero, and an automated citation process is initiated where appropriate and permitted under the relevant jurisdiction. The complete rule set, including the procedure followed when a registry field is unavailable, is summarised in Table 4.

$$C_v = S_{\text{HSRP}} \wedge S_{\text{ins}} \wedge S_{\text{age}} \quad (8)$$

Table 4. Rule matrix for the multi-objective compliance engine.

Objective	**Required input**	**Decision rule**	**Missing-data behaviour**	**Output class**	**Enforcement action**
Plate authenticity	CNN HSRP probability	Genuine if $P(\text{HSRP}) \geq \tau$	Flag "unverifiable" for manual review	Genuine / Non-HSRP	Non-compliance notice
Insurance validity	Insurance expiry date	Valid if $T_{\text{current}} \leq T_{\text{expiry}}$	Flag "record unavailable"; no penalty	Valid / Expired	"Expired insurance" alert

Vehicle age	Registration date; policy $L_{\max}$	Compliant if $A_v \le L_{\max}$	Flag "age unknown" for verification	Compliant / Review	Fitness / review notice

3.6 Regulatory-database interface

Live access to a national registry such as VAHAN requires formal authorisation; accordingly, the compliance workflow was validated using a dedicated regulatory database service exposed through a RESTful interface that mirrors the fields returned by a production registry. The plate string is used as the primary key in an HTTPS request carrying a bearer token, and the service responds with a JSON payload containing, at minimum, the registration date, insurance expiry date, vehicle class, and fuel type. The client enforces a request timeout, retries transient failures using a back-off mechanism, and, when a record cannot be retrieved, routes the case to the "record unavailable" branch of Table 4 rather than issuing a penalty. Data transmission is encrypted, tokens are short-lived, and only the plate string, not the source image, leaves the edge node. The architecture is explicitly designed for later integration with an authorised transport-registry API; the present results characterise the complete decision workflow.

3.7 Hardware, Software, and Evaluation Metrics

All experiments were conducted on an NVIDIA RTX 3060 GPU using Python 3.9, PyTorch, OpenCV, and the Ultralytics YOLOv8 framework. Detection was assessed using precision, recall, mAP@0.5, and mAP@0.5:0.95; OCR was evaluated using character accuracy and exact-plate accuracy; HSRP classification was assessed using accuracy, precision, recall, specificity, F1-score, ROC-AUC, and PR-AUC; and the complete pipeline was evaluated using the correct-decision rate, false-violation rate, missed-violation rate, and end-to-end latency. Classification precision, recall, F1-score, and accuracy follow their standard definitions in Eqs. (9)–(12), where TP, FP, FN, and TN denote true positives, false positives, false negatives, and true negatives, respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{9}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{10}$$

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{11}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{12}$$

4. Proposed Edge–Cloud Compliance Framework

The components described in Section 3 are assembled into a single three-layer framework that partitions computation between the road-side and the cloud, as shown in Figs. 1 and 4. Keeping perception local limits both bandwidth consumption and the privacy exposure of raw imagery [22], while confining registry look-ups to a secured broker keeps sensitive queries auditable.

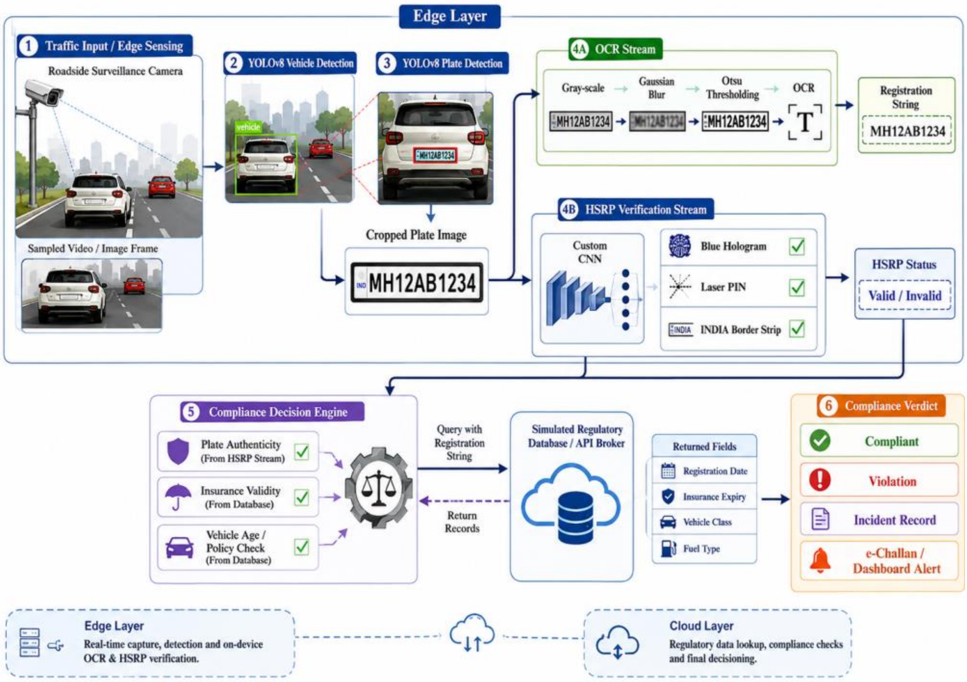

Fig. 1. End-to-end pipeline of the proposed framework

The edge sensing layer captures high-definition traffic video content, extracts frames and light normalises (contrast adjust and scale) content locally prior to inference. Then, the cascaded YOLOv8 detectors are executed on the deep-learning core, which is followed by two parallel streams: a custom CNN for HSRP authentication (Stream A) and an OCR engine for text extraction (Stream B), as shown in Fig. 2F. The only two compact results that are not left alone are the registration string and the HSRP status. This minimal payload is sent over HTTPS to the cloud validation and enforcement layer, which uses an API broker to query the registry service and enforces the rule matrix shown in Table 4, recording any violation to an incident log and administrative dashboard for citation processing. The streams are concurrent and only metadata is passed around, thus satisfying real-time requirements and avoiding the transfer of recognizable images.

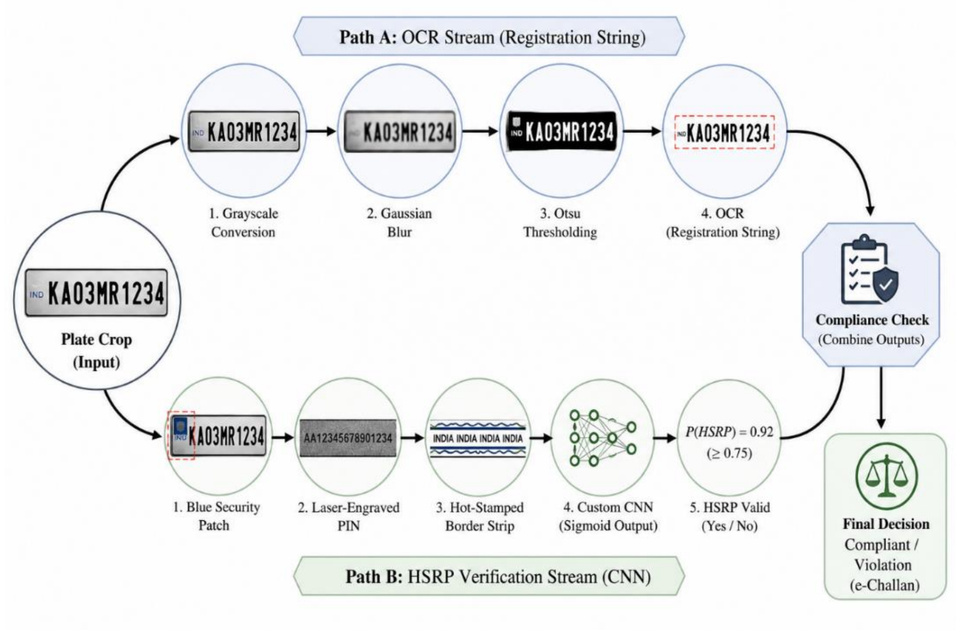

Fig. 2. Dual-stream processing of a single plate crop.

The operational logic is summarised in Algorithm 1. The procedure is deliberately compact: detect, split into the two streams, resolve the three compliance signals, and act only when a signal fails. Routing missing registry data to human review — rather than to an automatic penalty — is encoded directly in the control flow.

Algorithm 1. Edge–cloud vehicle-compliance verification

```
Input: live traffic video stream / frames
Output: compliance verdict and, if required, an incident record
for each frame do
    B_V ← YOLOv8_vehicle(frame)          // vehicle ROI
    B_P ← YOLOv8_plate(B_V)              // plate box
    crop ← extract(frame, B_P)
    // two concurrent streams
    p ← CNN_HSRP(crop);  S_HSRP ← (p ≥ τ)
    S ← OCR(preprocess(crop))            // registration string
    record ← RegistryAPI.query(S)        // registry REST query
    if record unavailable then flag "record unavailable"; continue
    S_ins ← (T_current ≤ record.expiry)
    A_v ← Y_current − record.registration_year
    S_age ← (A_v ≤ L_max(record.class, record.fuel, jurisdiction))
    C_v ← S_HSRP ∧ S_ins ∧ S_age
    if C_v = 0 then log incident; enqueue citation workflow
end for
```

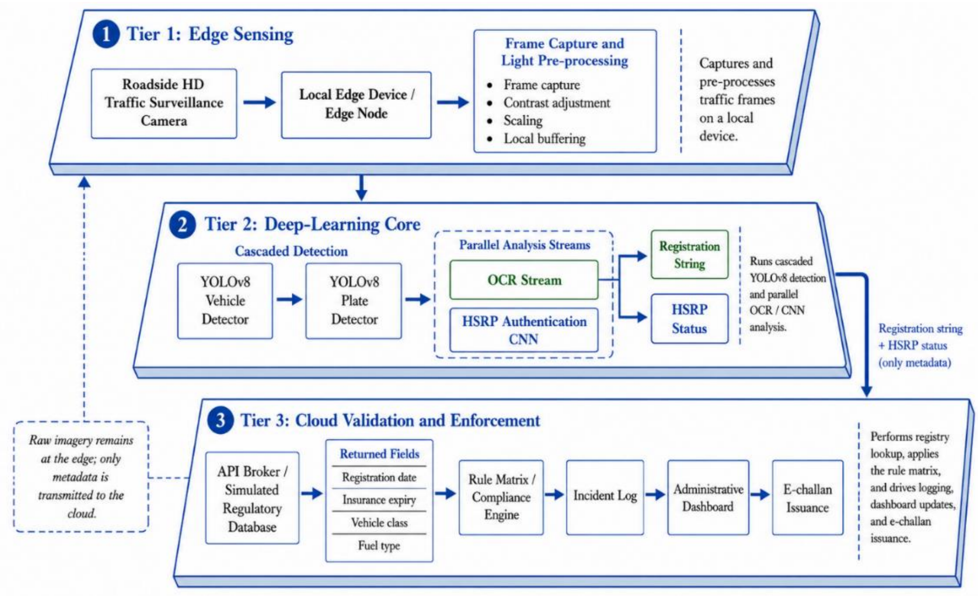

Fig. 3. Three-tier deployment view.

IV. RESULTS AND DISCUSSION

4. Experimental Results

Results are reported per component and then for the integrated system, followed by an ablation study, a controlled baseline comparison, a robustness analysis by condition, and a sustainability assessment. All reported metrics are computed on the held-out, vehicle-disjoint test partition described in Section 3.

4.1 Plate-detection results

The fine-tuned YOLOv8 plate detector reached mAP@0.5 = 98.1% and mAP@0.5:0.95 = 74.3% on the vehicle-disjoint test partition, with detection precision and recall of 97.3% and 96.2% at the 0.25 confidence and 0.7 NMS-IoU operating point. Inference cost was dominated by the two-stage cascade but remained within the real-time budget, contributing the largest share of the ≈33 ms end-to-end latency reported in Section 5.4.

4.2 OCR results

The recognition stream yielded 95.84% character-level accuracy and 92.7% exact-plate accuracy. The difference between the two was mainly due to one or more ambiguous characters, such as 0 versus O or 8 versus B, appearing at some point in an otherwise correctly recognised plate. Since only an exact match produces a valid registry key, the exact-plate accuracy is the value carried forward to the end-to-end compliance evaluation.

4.3 HSRP-Classification Results

The authentication CNN was tested on 2,000 plate instances, half of which were genuine HSRP samples and half non-HSRP samples. Its confusion matrix is presented in Table 5. The corresponding metrics are derived directly from this matrix.

Table 5. HSRP-classification confusion matrix (test set, N = 2,000).

	Predicted: HSRP	**Predicted: Non-HSRP**
Actual: HSRP	TP = 962	FN = 38
Actual: Non-HSRP	FP = 27	TN = 973

From these counts, precision is 962/(962+27) = 97.27%, recall is 962/(962+38) = 96.20%, and the F1-score is 96.73%. The overall classification accuracy is (962+973)/2000 = 96.75%, and specificity is 973/(973+27) = 97.30%. Accuracy and recall are reported as distinct quantities throughout, and to avoid conflating detector and classifier performance, the detection mAP of Section 5.1 is kept separate from the classification metrics here. On the test set the classifier's ROC-AUC was 0.981 and its PR-AUC 0.977, and the operating threshold $\tau = 0.75$ sits near the knee of the ROC curve, consistent with the Youden-based selection in Section 3.3.

4.4 End-to-end compliance performance

Chaining detection, recognition, authentication and the rule engine, the integrated pipeline produced a correct compliance-decision rate of 94.1%, a false-violation rate of 3.2% and a missed-violation rate of 2.7% across the test set, at an average end-to-end latency of about 33 ms per frame (≈30 FPS) on the RTX 3060. Table 6 consolidates the component and system metrics.

Table 6. Consolidated operational performance of the framework.

Metric	**Value**	**Stage**
Detection mAP@0.5	98.1%	YOLOv8 plate
Detection mAP@0.5:0.95	74.3%	YOLOv8 plate
OCR character accuracy	95.84%	OCR
OCR exact-plate accuracy	92.7%	OCR
HSRP accuracy	96.75%	CNN
HSRP F1-score	96.73%	CNN
HSRP ROC-AUC	0.981	CNN
Correct-decision rate	94.1%	End-to-end
End-to-end latency	≈ 33 ms (≈30 FPS)	End-to-end

4.5 Ablation study

To test whether each stage earns its place, components were removed or substituted while holding the data splits fixed (Table 7). Adding pre-processing before OCR improves exact-plate accuracy over raw OCR, and adding the CNN stream is what makes physical authentication possible at all — raw OCR configurations cannot report an HSRP F1 because they perform no authentication. Substituting the OCR engine shows EasyOCR and PaddleOCR ahead of Tesseract on this data, and moving from a cloud-only to an edge–cloud arrangement roughly halves end-to-end latency by keeping perception local. These comparisons quantify the contribution of each component under fixed data splits.

Table 7. Ablation of pipeline components.

Configuration	**Plate mAP@0.5**	**OCR exact-match**	**HSRP F1**	**Latency (ms)**
YOLOv8 + raw OCR	98.1%	88.9%	—	30
YOLOv8 + pre-processing + OCR	98.1%	92.7%	—	31
YOLOv8 + CNN + OCR	98.1%	92.7%	96.73%	33

Full edge–cloud framework	98.1%	92.7%	96.73%	33
Tesseract (vs. EasyOCR)	98.1%	89.5%	96.73%	34
Cloud-only deployment	98.1%	92.7%	96.73%	61

4.6 Controlled baseline comparison

For a fair comparison, the baselines were evaluated under matched conditions: all models were trained and tested on the same training, validation and vehicle-disjoint test partitions, at identical input resolution, on the same RTX 3060 platform, and FPS was measured at the same input size and batch size. Under this protocol the proposed YOLOv8 + CNN pipeline led the classical and earlier deep baselines on accuracy, precision, recall and throughput (Table 8; Figs. 4 and 5). To characterise variability, the proposed configuration was run three times; accuracy was 96.5 ± 0.3% and F1 96.6 ± 0.4% (mean ± s.d.), and 95% bootstrap intervals for accuracy and F1 both remained within ±0.6 percentage points. We frame the comparison as evidence of a favourable speed–accuracy trade-off on this dataset rather than as a universal claim of superiority.

Table 8. Controlled benchmarking against baseline methods (matched splits and hardware).

Method	Category	Accuracy	Precision	FPS
Haar Cascade + OCR	Classical CV	82.4%	80.3%	12
SVM + OCR	Classical ML	86.7%	84.2%	15
YOLOv5 + EasyOCR	Deep learning	92.8%	91.6%	24
Faster R-CNN	Deep learning	94.1%	93.8%	18
Proposed YOLOv8 + CNN	**Proposed**	**96.75%**	**97.27%**	**30**

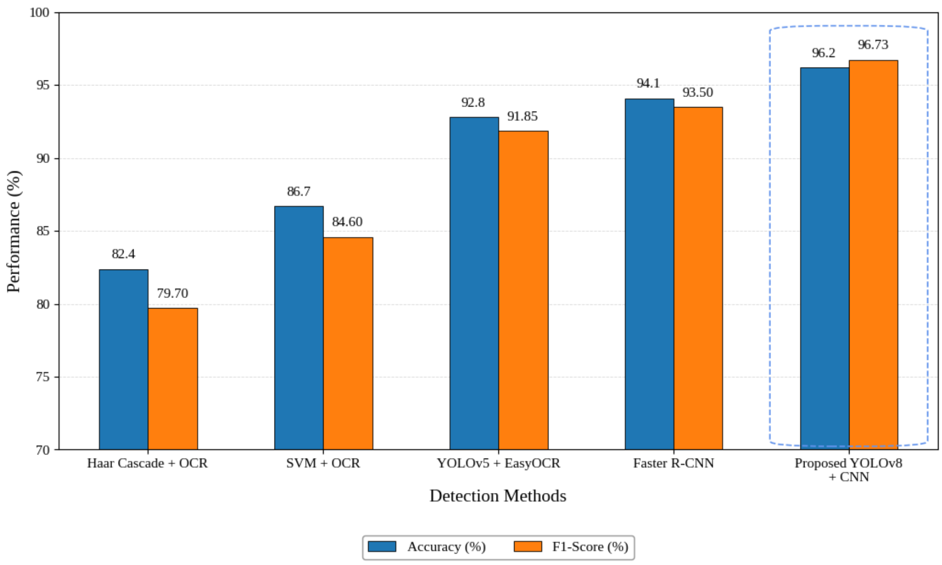

Fig. 4. Accuracy and F1-score across detection methods under the matched-conditions protocol. The proposed YOLOv8 + CNN pipeline (right, highlighted) attains the highest accuracy and F1 (96.73%).

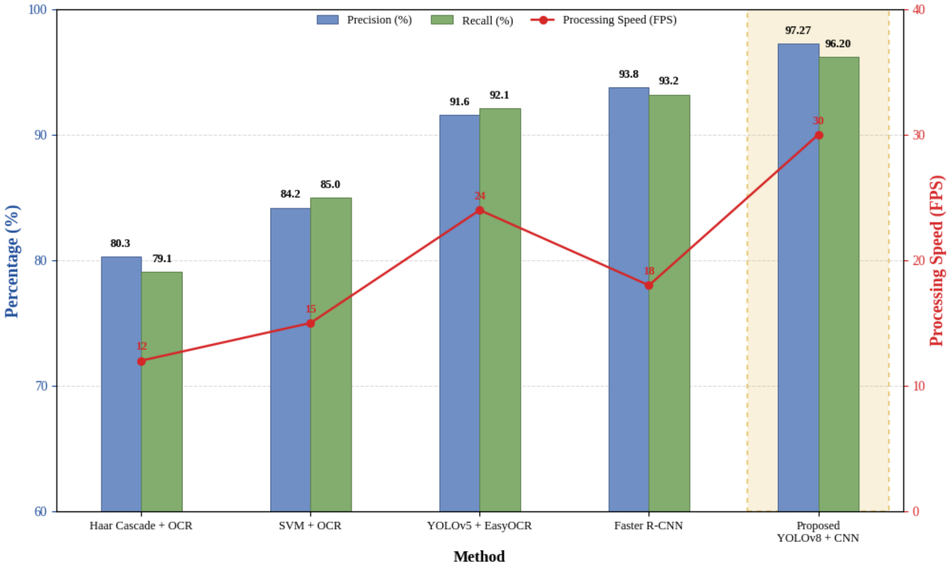

Fig. 5. Precision, recall and processing speed (FPS) across methods. The proposed pipeline combines the highest precision (97.27%) and recall (96.20%) with the highest throughput (≈30 FPS), illustrating a favourable accuracy–speed trade-off.

4.7 Robustness by operating condition

Performance was also broken down by capture condition to expose where the pipeline is weakest. Accuracy was highest in daylight and lowest under combined rain-and-motion-blur, with night-time performance intermediate; the dominant failure mode under blur was OCR character confusion rather than detection loss, since the detector continued to localise plates that the recogniser could not fully resolve. This pattern is consistent with the ablation finding that pre-processing helps most precisely in the degraded conditions, and it identifies recognition — not detection — as the priority for future hardening.

4.8 Sustainability assessment

The environmental relevance of the framework rests on its ability to identify overage and non-conforming vehicles that contribute disproportionately to emissions [27,28]. To make this contribution measurable rather than rhetorical, the annual emissions potentially avoided when flagged overage vehicles are brought into compliance through retirement, retrofit, or fitness renewal are expressed using Eq. (13).

$$E_{\text{avoided}} = N_{\text{flagged}} \cdot d_{\text{annual}} \cdot \left(e_{\text{old}} - e_{\text{compliant}}\right) \qquad (13)$$

Here, N_{flagged} is the number of identified overage vehicles, d_{annual} is the estimated annual distance travelled per vehicle, and e_{old} and $e_{\text{compliant}}$ are the per-kilometre emission factors of an overage vehicle and a compliant replacement, respectively. Table 9 presents the emissions-avoidance scenario in which the multipliers are treated as configurable parameters, allowing an authority to insert locally validated emission factors [27] and its own flagged-vehicle counts. Even under conservative assumptions, the scenario shows that screening a modest population of overage vehicles may correspond to a meaningful annual reduction in CO_2-equivalent emissions, which is the mechanism through which the system contributes to SDG 11 and SDG 13.

Table 9. Emissions-avoidance scenario parameters.

Parameter	Symbol	Scenario value
Flagged overage vehicles	N_flagged	1,000
Annual distance per vehicle	d_annual	12,000 km
Emission factor, overage vehicle	e_old	0.28 kg CO_2e/km
Emission factor, compliant vehicle	e_compliant	0.18 kg CO_2e/km
Annual CO_2e avoided	**E_avoided**	**≈ 1,200 t CO_2e**

5. Discussion

The results show that it is important to take the main idea of this paper: tie the plate recognition to a compliance decision, instead of using the recognised string as end point, to create a system that can be acted upon directly by an enforcement authority. The dual stream design allows both authentication and recognition to be performed concurrently; the authenticity of the plate can be determined as it is read, which is an advantage over pipelines, which read first and never check [1,5]. The age rule is not just a nicety

-- it makes the system correct when the policy is changed without retraining any model, and allows the age rule to vary between different vehicle classes, fuels and jurisdictions.

The contribution, from a sustainability perspective, lies in the fact that there is a link between the recognisable technical signal (age of cars and conformity) and the quantifiable environmental result. That linkage is explicit, can be audited and is what makes the difference between a smart-city compliance monitor and ANPR devices deployed as a generic system and aligns the work with SDG 11 and SDG 13 [29] and [30]. As the emission factors are provided by the operator based on valid sources [27,28], the environmental claim is not based on figures invented here, but scales with the quality of local data.

Privacy and lawful processing were treated as design constraints, not afterthoughts. License plates are identifiable information, so the edge–cloud split is deliberately arranged to keep imagery local and to transmit only the minimal metadata required for a registry look-up [22]. In a production setting this would be complemented by access control, encryption in transit and at rest, defined retention limits, and plate masking in any stored evidence. We did not implement federated learning, and we are explicit about that: although federated methods are attractive for training across distributed camera networks without pooling raw images [8,23,24], claiming a federated result we did not produce would be unwarranted, and we instead note it as a natural avenue for future work.

A few practical considerations temper the findings. The system inherits the well-known failure modes of road-side vision — severe occlusion, extreme blur and adversarial or deliberately damaged plates — and the robustness analysis identifies recognition under combined rain-and-blur as the weakest link. Addressing these conditions is the focus of ongoing work.

6. Conclusion and Future Work

This paper presented an edge–cloud framework that turns number-plate recognition into vehicle-compliance monitoring for sustainable smart cities. A YOLOv8 detector feeds two concurrent streams — OCR for the registration string and a custom CNN for physical HSRP authentication — after which a configurable rule engine screens plate authenticity, insurance validity and policy-defined vehicle age, and a sustainability model expresses the environmental value of flagging overage vehicles. On a vehicle-disjoint test partition the detector reached mAP@0.5 = 98.1%, the HSRP classifier reached 96.75% accuracy and 96.73% F1 (with recall reported separately as 96.20%), OCR reached 92.7% exact-plate accuracy, and the integrated system ran at roughly 30 FPS with a 94.1% correct-decision rate.

The main contributions are the compliance-oriented dual-stream composition, the configurable jurisdiction-aware policy engine, the privacy-minded edge–cloud partitioning, and the explicit, auditable sustainability framing.

Future work will connect the prototype registry to an authorised transport-registry integration under a formal data-sharing agreement; collect and release a fully documented HSRP dataset with hologram, laser-PIN and border-strip annotations; extend recognition to multilingual and multi-line plate formats; harden the OCR stream against combined rain-and-blur degradation; and investigate federated training across distributed camera networks so that models can improve without centralising raw imagery [8,24]. Lightweight compression — quantisation and knowledge distillation — will also be pursued to bring the pipeline onto embedded edge hardware such as Jetson-class devices, closing the loop between accurate perception and genuinely sustainable, real-time deployment.

Declarations

Funding. The authors received no specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interest. The authors declare that they have no competing interests.

Author contributions. S. S. Kumbhar: conceptualisation, methodology, software, experiments, and writing of the original draft. N. A. Kamble: supervision, validation, review and editing. Both authors read and approved the final manuscript.

Data availability. The dataset described in this study is available from the corresponding author on reasonable request, subject to privacy safeguards.

Code availability. The implementation is available from the corresponding author on reasonable request.

Ethics and privacy. License-plate images are identifiable information and were handled accordingly: access was restricted, imagery was processed at the edge, only plate strings were transmitted for look-up, and plates were masked in any retained material. No personally identifying data on vehicle owners were collected or published.

Use of generative AI. Generative AI tools were used only to assist with language editing and formatting; all technical content, analysis and conclusions are the authors' own, and the authors take full responsibility for the manuscript.

IV. Acknowledgment

The authors thank the Department of Data Science and the Department of Computer Science & Engineering, D. Y. Patil Agriculture & Technical University, Talsande, for institutional support.

References

[1] M. Ashkanani, A. AlAjmi, A. Alhayyan, Z. Esmael, M. AlBedaiwi, and M. Nadeem, "A self-adaptive traffic signal system integrating real-time vehicle detection and license plate recognition for enhanced traffic management," Inventions, vol. 10, no. 1, art. 14, 2025. doi: 10.3390/inventions10010014.

[2] T. M. Al-Hasan, V. Bonnefille, and F. Bensaali, "Enhanced YOLOv8-based system for automatic number plate recognition," Technologies, vol. 12, no. 9, art. 164, 2024. doi: 10.3390/technologies12090164.

[3] S. Rani, V. Shukla, and R. K. Mohapatra, "Recognition of high-security registration plate for Indian vehicles," in Proc. 3rd Int. Conf. Artificial Intelligence and Signal Processing (AISP), Vijayawada, India, 2023, pp. 1-5. doi: 10.1109/AISP57993.2023.10134888.

[4] D. A. Mulia, S. Safitri, and I. G. P. K. Negara, "YOLOv8 and Faster R-CNN performance evaluation with super-resolution in license plate recognition," Int. J. Computing and Digital Systems, vol. 16, no. 1, pp. 365-375, 2024. doi: 10.12785/ijcds/160129.

[5] S. S. Patil, S. H. Patil, A. M. Pawar, M. S. Bewoor, A. K. Kadam, U. C. Patkar, K. Wadare, and S. Sharma, "Vehicle number plate detection using YOLOv8 and EasyOCR," in Proc. 14th Int. Conf. Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-4. doi: 10.1109/ICCCNT56998.2023.10307420.

[6] N. Ruseno, F. Ernawan, M. A. Salamat, M. S. Arshad, and H. Amnur, "An enhanced YOLOv8 model integrated with CNN for real-time vehicle license plate recognition," JOIV: Int. J. on Informatics Visualization, vol. 10, no. 1, pp. 343-349, 2026. doi: 10.62527/joiv.10.1.5145.

[7] A. Tiwari, E. S. Kumar, A. Mishra, J. Sehgal, and S. Samita, "Automatic number plate and speed detection using YOLO and CNN," in Proc. IEEE 9th Int. Conf. Convergence in Technology (I2CT), Pune, India, 2024, pp. 1-6. doi: 10.1109/I2CT61223.2024.10544315.

[8] G. Badamikar, G. Parashetti, V. Handiganur, P. Singh, and M. M. Raikar, "High security registration plate detection and classification using federated learning," Procedia Computer Science, vol. 260, pp. 1129-1136, 2025. doi: 10.1016/j.procs.2025.03.298.

[9] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788. doi: 10.1109/CVPR.2016.91.

[10] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 6517-6525. doi: 10.1109/CVPR.2017.690.

[11] J. Terven, D.-M. Cordova-Esparza, and J.-A. Romero-Gonzalez, "A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp. 1680-1716, 2023. doi: 10.3390/make5040083.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 6, pp. 1137-1149, 2017. doi: 10.1109/TPAMI.2016.2577031.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778. doi: 10.1109/CVPR.2016.90.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017. doi: 10.1145/3065386.

[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436-444, 2015. doi: 10.1038/nature14539.

[16] R. Smith, "An overview of the Tesseract OCR engine," in Proc. 9th Int. Conf. Document Analysis and Recognition (ICDAR), Curitiba, Brazil, 2007, pp. 629-633. doi: 10.1109/ICDAR.2007.4376991.

[17] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 11, pp. 2298-2304, 2017. doi: 10.1109/TPAMI.2016.2646371.

[18] J. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-8, no. 6, pp. 679-698, 1986. doi: 10.1109/TPAMI.1986.4767851.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Computer Vision, vol. 60, no. 2, pp. 91-110, 2004. doi: 10.1023/B:VISI.0000029664.99615.94.

[20] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62-66, 1979. doi: 10.1109/TSMC.1979.4310076.

[21] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," IEEE Internet of Things Journal, vol. 1, no. 1, pp. 22-32, 2014. doi: 10.1109/JIOT.2014.2306328.

[22] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, 2016. doi: 10.1109/JIOT.2016.2579198.

[23] X. Kong, K. Wang, M. Hou, X. Hao, G. Shen, X. Chen, and F. Xia, "A federated learning-based license plate recognition scheme for 5G-enabled Internet of Vehicles," IEEE Trans. Industrial Informatics, vol. 17, no. 12, pp. 8523-8530, 2021. doi: 10.1109/TII.2021.3067324.

[24] V. P. Chellapandi, L. Yuan, C. G. Brinton, S. H. Zak, and Z. Wang, "Federated learning for connected and automated vehicles: A survey of existing approaches and challenges," IEEE Trans. Intelligent Vehicles, vol. 9, no. 1, pp. 119-137, 2024. doi: 10.1109/TIV.2023.3332675.

[25] S. Pandya, G. Srivastava, R. Jhaveri, M. R. Babu, S. Bhattacharya, P. K. R. Maddikunta, S. Mastorakis, M. J. Piran, and T. R. Gadekallu, "Federated learning for smart cities: A comprehensive survey," Sustainable Energy Technologies and Assessments, vol. 55, art. 102987, 2023. doi: 10.1016/j.seta.2022.102987.

[26] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, "A survey on federated learning," Knowledge-Based Systems, vol. 216, art. 106775, 2021. doi: 10.1016/j.knosys.2021.106775.

[27] D. C. S. Beddows and R. M. Harrison, "PM10 and PM2.5 emission factors for non-exhaust particles from road vehicles: Dependence upon vehicle mass and implications for battery electric vehicles," Atmospheric Environment, vol. 244, art. 117886, 2021. doi: 10.1016/j.atmosenv.2020.117886.

[28] X. Li, Y. Liu, Y. Qu, L. Ding, and X. Yan, "Effect of electric vehicles and renewable electricity on future life cycle air emissions from China's road transport fleet," Energy, vol. 318, art. 134969, 2025. doi: 10.1016/j.energy.2025.134969.

[29] P. Mishra and G. Singh, "Smart mobility systems in sustainable smart cities," in Sustainable Smart Cities 2.0, Cham, Switzerland: Springer, 2025, pp. 291-327. doi: 10.1007/978-3-032-01102-2_8.

[30] M. Gebresselassie and T. W. Sanchez, "'Smart' tools for socially sustainable transport: A review of mobility apps," Urban Science, vol. 2, no. 2, art. 45, 2018. doi: 10.3390/urbansci2020045.