STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **16**, August 2026, pp 1086–1109.
Published online in International Academic Press (www.IAPress.org)

Multi-Variant Enhancement and OCR Selection for Robust License Plate Recognition in Hazy Conditions

Arafat[1,2*], Pulung Nurtantio Andono[1], Abdul Syukur[1], Affandy[1]

[1]*Faculty of Computer Science, Dian Nuswantoro University, Semarang, Indonesia*
[2]*Universitas Islam Kalimantan Muhammad Arsyad Al Banjari, Banjarmasin, Indonesia*

Abstract Automatic License Plate Recognition (ALPR) systems often experience performance degradation under adverse visual conditions such as haze, low contrast, and limited image resolution. This study proposes an adaptive ALPR framework that integrates Dark Channel Prior (DCP)-based dehazing, YOLOv11-based license plate detection, lightweight multi-variant image enhancement, and plate-aware OCR selection to achieve robust recognition under hazy conditions. After dehazing and plate detection, each cropped license plate image is processed with three enhancement variants: RAW, CLAHE-light, and SR2x + CLAHE-light. Each variant is recognized with PaddleOCR to generate multiple OCR candidates, which are then evaluated with a plate-aware scoring mechanism that combines OCR confidence, structural validity, and ambiguity penalties. Experimental evaluation on 945 hazy license plate images achieved a readability rate of 100.00%, an exact plate accuracy of 93.12%, and a character-level accuracy of 98.63%. Ablation analysis showed that the proposed framework improved exact plate accuracy by 3.70 percentage points over the strongest single-variant baseline, and statistical significance testing confirmed that the improvement was not due to random variation. Parameter sensitivity analysis further demonstrated that the proposed scoring mechanism remained stable across different weighting configurations. Although the complete pipeline requires an average processing time of 471.42 ms per image (2.12 FPS), the results indicate that adaptive multi-variant enhancement combined with structure-aware OCR selection provides a robust and accurate solution for ALPR under degraded hazy conditions without requiring retraining of either the detector or the OCR model.

Keywords Automatic License Plate Recognition, hazy images, adaptive enhancement, CLAHE-light, PaddleOCR, plate-aware OCR selection

DOI: 10.19139/soic-2310-5070-3818

1. Introduction

The development of Intelligent Transportation Systems (ITS) has driven the adoption of vehicle license plate recognition (Automatic License Plate Recognition / ALPR) as an essential component of traffic monitoring and data-driven law enforcement. ALPR systems enable automatic, real-time vehicle identification, playing a significant role in enhancing the efficiency of modern transportation management [1, 2]. However, ALPR performance is highly affected by environmental conditions. Atmospheric factors such as haze reduce contrast, degrade edge detail, and diminish image texture, thereby directly impacting the accuracy of license plate detection and recognition [3, 4]. In license plate OCR systems, a single character error can lead to a complete failure in end-to-end recognition, making resilience to visual degradation a crucial aspect [5].

Over the past two decades, Convolutional Neural Networks (CNN) have become the backbone of image recognition systems, including ALPR. The strength of CNNs lies in their ability to automatically extract spatial features and visual patterns without manual feature engineering [6]. CNN-based architectures such as YOLO,

*Correspondence to: Arafat (Email: p41202300070@mhs.dinus.ac.id). Faculty of Computer Science, Dian Nuswantoro University, Semarang, Indonesia.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Faster R-CNN, and SSD have been shown to achieve high accuracy and operate in real time for detecting small objects, such as vehicle license plates [7]. However, the performance of these networks remains highly dependent on the quality of input images. In foggy conditions, low contrast and changes in the intensity distribution degrade features in the early layers of CNNs, preventing the optimal extraction of edge information and character details [8].

To address atmospheric disturbances, various image preprocessing techniques have been developed. Dark Channel Prior (DCP) is one of the most widely used classic dehazing methods, which utilizes the statistical properties of outdoor images to estimate atmospheric light and transmission maps. The application of DCP has been shown to improve visibility and contrast, making features more representative for detection stages [3]. However, the effectiveness of DCP still depends on fog conditions and processing parameters, and does not always guarantee consistent improvement in OCR accuracy [10]. Based on this principle, DCP can estimate the transmission map and remove light scattering effects caused by fog particles. When DCP is applied before the CNN, the image's visual quality improves, enabling the network to extract more representative features [**?**]. Applying DCP before the detection or OCR stage can enhance feature readability and recognition stability in foggy images, but the extent of improvement depends on the dataset and fog level [10]. Additionally, variations in license plate quality due to high humidity and corrosion further degrade character recognition performance.

In addition to dehazing, local contrast enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) and super-resolution approaches are also commonly applied to clarify characters on low-resolution license plates [11]. However, fixed preprocessing approaches often result in a trade-off between visual quality and recognition accuracy. Excessive enhancement can increase noise or distort character shapes, thereby reducing OCR accuracy [12, 13]. In other words, there is no single optimal image enhancement strategy for all variations in visual conditions. This challenge is further amplified in tropical environments like Indonesia, which are characterized by high humidity, varying fog levels, limited CCTV camera quality, and diverse license plate conditions. Based on this literature review, most ALPR research still focuses on improving individual components, such as detection, dehazing, and OCR, in isolation. However, there is limited research that explicitly addresses ambiguities in OCR outputs arising from various image enhancement variants. In practice, some image enhancement techniques may produce OCR predictions that contradict one another, even at high confidence levels. The absence of a selection mechanism based on plate structure and character consistency remains a major research gap.

This study uses Dark Channel Prior (DCP)-based dehazing as the foundation to reduce the global fog influence at the dataset level, stabilizing the basic structure of characters before the detection process [14]. YOLOv11 has been reported to show competitive accuracy and real-time processing capability in small-scale object detection tasks [15]. However, even though license plates have been successfully cropped, the cropped images under foggy conditions often still exhibit low contrast, blurring, or limited resolution, thereby risking a reduction in OCR accuracy. To address these issues, this study proposes an Adaptive Multi-Variant Enhancement and OCR Selection approach. The proposed system integrates Dark Channel Prior (DCP)-based dehazing, license plate detection using YOLOv11, the creation of several lightweight image enhancement variants, and an adaptive OCR selection mechanism based on license plate structural characteristics. This approach performs parallel multi-variant evaluations to generate multiple OCR candidates and then selects the most valid and stable prediction.

The main contributions of this study are as follows: (1) Proposing an adaptive ALPR framework for hazy conditions that integrates DCP-based dehazing, YOLOv11-based license plate detection, lightweight multi-variant image enhancement, and a plate-aware OCR selection mechanism into a unified end-to-end pipeline. Unlike conventional ALPR systems that rely on a single dehazing–detection–OCR sequence, the proposed framework generates multiple OCR candidates from three complementary enhancement variants and dynamically selects the most reliable result on a per-image basis; (2) Analyzing the effects of three lightweight enhancement variants, namely RAW, CLAHE-light, and SR2x + CLAHE-light, on OCR robustness under degraded visual conditions, including failure-case analysis of CLAHE-light using image quality metrics and visual inspection; (3) Conducting comprehensive end-to-end evaluation on a public hazy license plate dataset, including ablation studies, statistical

significance testing, parameter sensitivity analysis, and runtime analysis, demonstrating that the proposed plate-aware selection mechanism consistently improves recognition performance without requiring retraining of either the detector or the OCR model.

The remainder of this paper is organized as follows. Section 2 reviews related studies on dehazing, image enhancement, license plate detection, OCR integration, and result selection strategies. Section 3 describes the proposed methodology, including dataset preparation, dehazing, detection, multi-variant enhancement, OCR processing, and plate-aware selection. Section 4 presents the experimental results, comparative analysis, ablation study, and error analysis. Finally, Section 5 concludes the paper and discusses its limitations and future work.

2. Related Works

Automatic License Plate Recognition (ALPR) is a multi-stage system typically consisting of object detection, image preprocessing, and character recognition. Under ideal conditions, deep learning-based architectures have achieved high accuracy and real-time performance. However, under degraded visual conditions such as fog, low contrast, and limited resolution, the system's performance declines significantly. To provide a systematic analysis, this literature review is organized into five main categories: (1) dehazing in traffic images; (2) contrast and resolution enhancement; (3) deep learning-based plate detection; (4) end-to-end OCR integration; (5) decision selection strategies at the OCR level. This thematic approach allows for a clear and logical identification of gaps in the research.

2.1. Dehazing in ALPR Systems

Fog causes light attenuation and scattering by atmospheric particles, reducing global contrast and eliminating edge details. The atmospheric image formation model is expressed as

$$I(x) = J(x)t(x) + A * (1 - t(x)) \tag{1}$$

where $I(x)$ is the observed image, $J(x)$ is the haze-free image, $t(x)$ is the transmission map, and A is the atmospheric light. Classical methods such as Dark Channel Prior (DCP) [3] are widely used to enhance the visibility of traffic images. This approach is effective in improving global contrast, but it is sensitive to extreme lighting conditions and can generate visual artifacts under certain conditions [16]. Deep learning-based approaches, such as GANs, have also been developed to produce more natural dehazing results [14], but they are computationally complex. Importantly, some studies have shown that improving visual quality through dehazing does not always directly translate into higher OCR accuracy [13]. This indicates that while dehazing is an important component, it is not a standalone solution for improving ALPR performance in foggy conditions.

2.2. Contrast Enhancement and Super-Resolution

Image enhancement techniques such as Contrast Limited Adaptive Histogram Equalization (CLAHE) are used to improve local contrast without significantly amplifying noise [17]. This approach has been reported to enhance character readability in low-light conditions [18, 19]. Additionally, super-resolution methods such as Real-ESRGAN [11] are used to improve image resolution. While these methods produce visually sharper images, generative-based approaches have the potential to alter the microstructure of characters, thereby reducing exact recognition accuracy. Alternatively, bicubic interpolation is often used because it is more stable and does not generate synthetic textures [20]. However, this process can still amplify noise and increase character ambiguity, such as distinguishing between O–0 and B–8. These findings suggest that no single enhancement method consistently outperforms others across all conditions. Effectiveness is highly dependent on the degree of image degradation and the distinct visual characteristics involved.

2.3. Deep Learning-Based Plate Detection

CNN-based object detection models, such as YOLO [7] and its more recent variants, YOLOv8 and YOLOv11 [15], have become a standard choice in modern ALPR systems. These models offer a balance between accuracy

and speed, particularly for detecting small objects such as license plates. However, detection performance remains highly dependent on the quality of the input image. In foggy conditions, changes in intensity distribution lead to a decline in feature quality during the early stages of feature extraction, which in turn affects detection accuracy [8]. Some studies combine image preprocessing before detection to improve system stability, but in general, they still rely on a single fixed preprocessing strategy.

2.4. OCR Integration in ALPR Systems

The integration of detection models such as YOLO and OCR systems such as PaddleOCR has been widely adopted in end-to-end ALPR systems [21]. Architectures such as CRNN with CTC decoding enable text recognition without explicit character segmentation [5]. Most approaches use a single preprocessed image as input to the OCR stage. If preprocessing is not optimal, the OCR may produce high-confidence predictions that are structurally incorrect. This highlights that OCR output uncertainty has not been explicitly modeled in many studies. Some approaches use simple ensemble methods such as confidence averaging or majority voting. However, these methods do not account for character-structure consistency or ambiguity among OCR candidates.

2.5. Result Selection and Research Gaps

The literature indicates that enhancing image quality does not always result in improved character recognition accuracy. Although some studies combine various enhancement techniques, the selection of OCR results is generally still performed simply and non-adaptively. Key limitations in previous research include: (1) the lack of an adaptive mechanism to compare multiple image variants before OCR; (2) OCR result selection has not been modeled as an inference process based on uncertainty; (3) character structure information has not been systematically utilized in decision-making. Most studies focus on improving individual ALPR components in isolation, without integrating uncertainty modeling and decision-level selection into a unified framework. Based on these gaps, this research proposes an adaptive multi-variant enhancement and OCR selection approach that generates multiple OCR candidates from various image variants and selects the most reliable result using an adaptive selection mechanism based on confidence, structural consistency, and character ambiguity. This approach shifts the focus from mere image-quality enhancement to candidate-based decision modeling, making the system more robust to variations in real-world visual conditions.

3. Methodology

This study proposes the Adaptive Multi-Variant Enhancement and OCR Selection framework to improve the performance of ALPR systems under fog-induced visual degradation. The proposed method consists of four main stages: dehazing, license plate detection, generation of multi-variant enhancements, and adaptive selection of OCR results. The system first enhances visibility using a dehazing method. Then, the license plate area is detected and cropped using an object detection model. The resulting plate regions are then processed through several image enhancement variants to generate diverse visual representations. Each variant is processed independently using OCR to generate text candidates. In the final stage, adaptive selection is performed to choose the most reliable OCR candidate. This approach is designed to overcome the limitations of single preprocessing by providing multiple OCR candidate results under various image representation conditions. Subsequently, the selection process is performed adaptively, taking into account OCR confidence, character consistency, and visual ambiguity. Thus, the proposed system improves robustness to fog-induced visual degradation without relying on a specific image enhancement strategy. The research methodology flow is shown in Figure 1.

3.1. Detection Training Dataset

In this study, the license plate detection model was trained on a dataset from Roboflow [22] comprising 10,100 license plate images. This dataset includes variations in license plate size, viewpoints, and image capture conditions, which support the model's learning process and make it more robust to real-world variations. All images

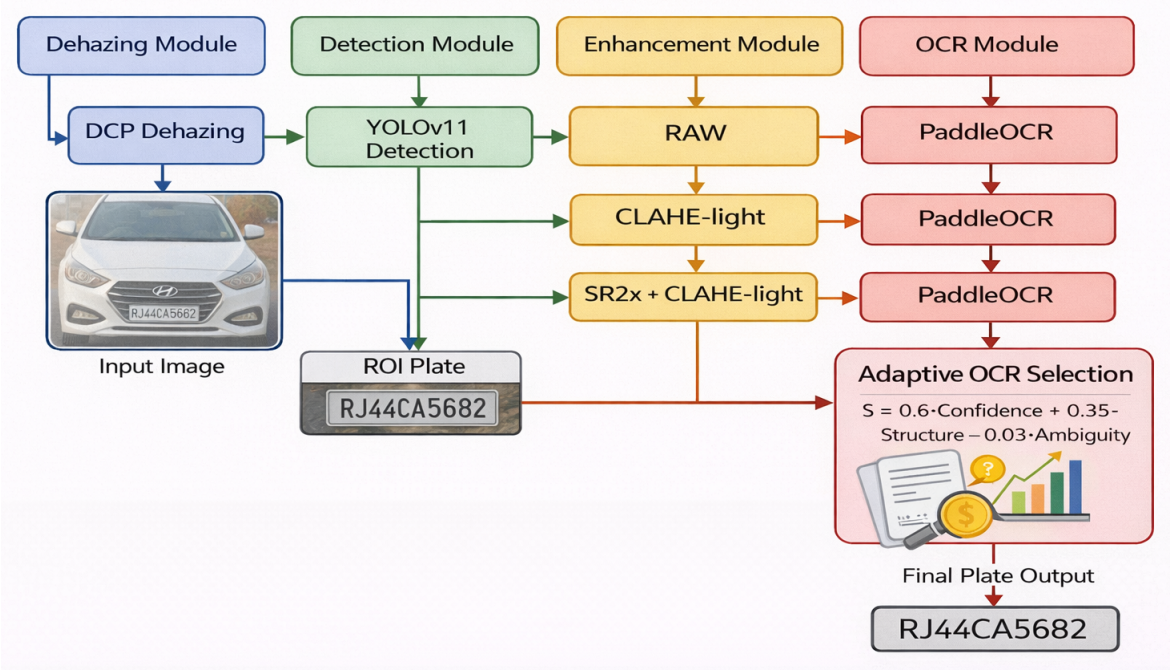

Figure 1. Overall Framework of the Proposed Method.

in the dataset were manually annotated to mark the license plate location using bounding boxes. Preprocessing steps were carried out to prepare the data before training, including image resizing, pixel-intensity normalization, and data augmentation to increase sample diversity and improve the model's generalization capability. The dataset was then converted into YOLOv11 format by extracting the bounding box coordinates and class labels into text-based annotation files, enabling efficient processing during training of the license plate detection model. A flowchart of the dataset collection, annotation, and preparation process through to the YOLOv11 model training stage is shown in Figure 2.

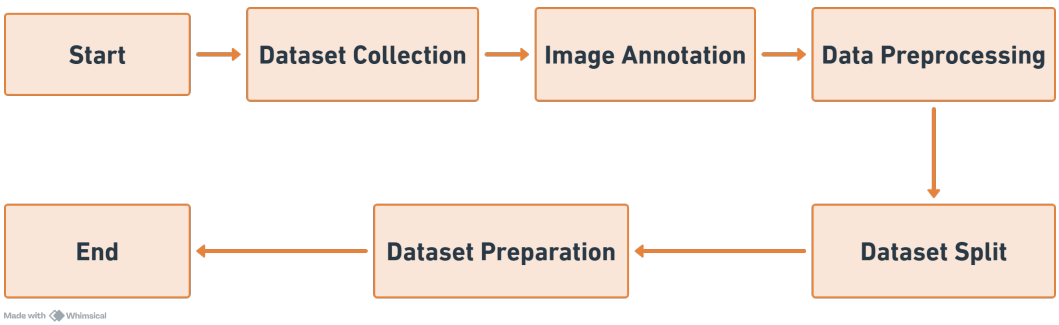

Figure 2. Obtaining a collection of datasets for YOLO.

3.2. System Processing Stages

3.2.1 Image Enhancement using Dark Channel Prior (DCP)

Dark Channel Prior (DCP) is an effective dehazing method that enhances image visibility by leveraging the dark channel characteristics of outdoor images [3]. This method assumes that in fog-free images, there are pixels with very low intensity in at least one color channel. In this study, DCP is used as a preprocessing step to improve the contrast and visibility of license plates, thereby enhancing YOLOv11 detection performance [2]. The process begins by separating the color channels (R, G, B) and extracting the dark channel through a minimum operation and erosion using a patch size of 15×15 pixels [3]. The atmospheric light, A, is estimated from the top 0.1% of the pixels with the highest intensity in the dark channel, and then iteratively updated using

$$A_i^{n+1} = \frac{1}{2}\left(A_i^n + R(x_{n+1})\right), \quad x \in \widetilde{R}, \quad i \in \{R, G, B\} \tag{2}$$

where A_i is the atmospheric light value for the i-th color channel, $R(x_{n+1})$ is the selected pixel intensity in the dark channel, and $\widetilde{R}$ is the set of candidate pixels with the highest intensity. The transmission map $t(x)$ is computed to model the scattering effect of fog with a haze retention parameter $\omega = 0.95$:

$$\widetilde{t}(x) = 1 - \omega \min_c \left(\min_{y \in \Omega(x)} \left(\frac{I^c(y)}{A^c} \right) \right) \tag{3}$$

where $I^c(y)$ is the pixel intensity in the c color channel, A^c is the atmospheric light in that channel, and $\Omega(x)$ is the local patch around pixel x. To prevent excessive noise in areas with thick fog, the transmission value is limited to a minimum threshold $t_0 = 0.1$. The transmission map is then refined using guided filtering with a radius parameter $r = 40$ and regularization $\epsilon = 10^{-3}$, by minimizing the following cost function [3]:

$$\min_t \left(t^T L t + \lambda \|t - \widetilde{t}\|^2 \right) \tag{4}$$

where t is the final transmission map, $\widetilde{t}$ is the initial estimate, L is the matting Laplacian matrix, and λ is the regularization parameter. The final step is reconstructing the haze-free image using the atmospheric light and the refined transmission map [14]. This process enhances contrast and clarifies object structures, thereby supporting detection and OCR performance in subsequent stages.

3.2.2 Non-local Dehazing

As a dehazing baseline, this study employs the non-local image dehazing approach to compare with the Dark Channel Prior (DCP) method. The non-local method utilizes long-range pixel correlations to model the fog distribution, while DCP estimates the transmission map using the dark channel prior to enhancing image visibility. The performance of both methods is evaluated using image quality metrics, including the Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), and execution time to measure computational efficiency. Generally, the dehazing model follows the atmospheric image formation model:

$$I(x) = J(x)t(x) + A\left(1 - t(x)\right) \tag{5}$$

where I(x) is the observed image, J(x) is the fog-free image, t(x) is the transmission map, and A is the global atmospheric light. The non-local approach estimates t(x) based on similarities between image regions, then reconstructs the clean image by inverting the atmospheric model. In this study, the non-local method is used as a baseline for comparison with the Dark Channel Prior (DCP) approach. All processing parameters are set fixed without specific tuning for each image to maintain consistency in evaluation. The experimental setup compares the two dehazing approaches implemented in this study: Dark Channel Prior (DCP) and non-local image dehazing. The performance of both methods is evaluated using image quality metrics such as the Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR), as well as execution time, to assess computational efficiency.

3.2.3 License Plate Detection Using YOLOv11

The next stage is license plate detection and cropping from the dehazed image using YOLOv11, which is selected for its balance between speed and accuracy in detecting small objects such as license plates [15]. YOLOv11 utilizes the Cross Stage Partial (CSP) architecture and the Path Aggregation Network (PANet) to improve spatial feature representation, and it is trained through transfer learning from pre-trained COCO weights that are subsequently fine-tuned on the ALPR dataset. The predicted bounding box location is calculated from the box position and size parameters, as shown in Equation 6, together with the confidence score for each detection. The output of this stage is a cropped license plate image, which is then passed to the adaptive enhancement and character recognition stages.

$$P = \sigma(t_x) + c_x, \qquad W = p_w e^{t_w} \tag{6}$$

where P represents the center position of the bounding box, W denotes the predicted width, and σ is the sigmoid activation function. Each prediction is accompanied by a confidence score indicating the probability that a license plate is present in the detected region. After the image is enhanced using DCP, YOLOv11 is applied to localize and extract the vehicle license plate, producing a cropped image that is ready for the next adaptive enhancement stage.

YOLOv11 is a convolutional neural network-based object detection model trained on a large-scale annotated dataset, enabling efficient and accurate detection of small objects such as license plates [23, 24]. The model training process is configured with an image size of 640×640 pixels, 100 epochs, and a batch size of 16. During inference, detection is performed using a confidence threshold of 0.25 to filter out low-confidence predictions, followed by Non-Maximum Suppression (NMS) with an IoU threshold of 0.7 to remove overlapping detections. These parameters ensure that only high-quality bounding boxes are retained.

3.2.4 Multi-Variant Enhancement

To address variations in visual conditions, the cropped license plate images are processed using several image enhancement variants in parallel. This approach aims to generate multiple visual representations that can provide different OCR results. The variants used in this study include:

1. RAW
   The RAW variant is the cropped license plate image without any enhancement processing. Mathematically, the RAW image can be expressed as:
   $$I_{\text{RAW}}(x, y) = I_{\text{ROI}}(x, y) \tag{7}$$
   where $I_{\text{ROI}}(x, y)$ is the pixel intensity of the license plate image after applying the Region of Interest (ROI) at coordinates (x, y). This variant is maintained as the baseline because, under certain conditions, the original image may have higher quality than the processed image. Using RAW helps avoid overprocessing, which can lead to character distortion or the loss of important information through image manipulation.
2. CLAHE-light (Contrast Limited Adaptive Histogram Equalization)
   To enhance local contrast in the image, CLAHE divides the image into small 8×8 blocks (tiles), that is, `tileGridSize = (8,8)`, and then performs adaptive histogram normalization on each block. A clip limit of 2.0 is used to limit contrast enhancement and prevent excessive noise amplification. This configuration is chosen to balance improved character visibility with preservation of image-structure stability. The clipped histogram is expressed as:
   $$H_c(i) = \min(H(i), T) \tag{8}$$
   where $H(i)$ is the original histogram at gray level i, and T is the clip limit. The cumulative distribution function (CDF) is then calculated as:
   $$\text{CDF}(i) = \sum_{j=0}^{i} H_c(j) \tag{9}$$

The pixel intensity transformation is performed using:

$$I_{\text{CLAHE}}(x,y) = \left\lfloor \frac{\text{CDF}(I(x,y)) - \text{CDF}_{\min}}{(M \times N) - \text{CDF}_{\min}} \times (L-1) \right\rfloor \tag{10}$$

where $M \times N$ is the number of pixels in a single tile, and L is the number of intensity levels, usually 256. This approach effectively enhances the intensity differences between characters, particularly in low-light or light-fog conditions, without causing overenhancement.

3. SR2x + CLAHE-light (Super-Resolution and Enhancement)
To handle low-resolution or blurred license plates, the image is upsampled using bicubic interpolation with a scale factor of $2\times$ to increase spatial resolution and enhance character details. Next, CLAHE-light is applied with the same parameters, namely `tileGridSize = (8,8)` and `clipLimit = 2.0`, to improve local contrast in the upsampled image. Bicubic interpolation calculates the new pixel value based on the 16 nearest neighboring pixels and can be expressed as:

$$I_{\text{SR}}(x',y') = \sum_{i=-1}^{2} \sum_{j=-1}^{2} w(i,j)\, I(x+i, y+j) \tag{11}$$

where $w(i,j)$ is the bicubic interpolation weight, and $I(x+i, y+j)$ is the pixel value of the neighboring pixels in the original image. Bicubic interpolation is chosen for its smooth and stable behavior and its ability to avoid synthetic texture reconstruction, which can introduce artifacts. After the upscaling process, CLAHE-light is applied to improve the local contrast in the higher-resolution image:

$$I_{\text{SR+CLAHE}} = \text{CLAHE}(I_{\text{SR}}) \tag{12}$$

The use of multiple enhancement variants aims to accommodate variations in visual conditions, with each method having different strengths and limitations in character recognition quality. For completeness, an additional benchmark experiment using Real-ESRGAN was also conducted. This experiment was not incorporated into the proposed multi-variant OCR selection framework and was evaluated separately to assess the effectiveness of a deep-learning-based super-resolution approach relative to the proposed enhancement strategy.

3.2.5 Character Recognition Using PaddleOCR

Each image resulting from multi-variant enhancement is processed using an Optical Character Recognition (OCR) system to extract text information from the license plate. In this study, PaddleOCR with the PP-OCRv4 model is used, as it is designed for text recognition in real-world scenarios while balancing accuracy and efficiency. The OCR configuration is as follows: (1) model: PP-OCRv4; (2) language: English (`language = en`); and (3) decoding method: Connectionist Temporal Classification (CTC) best-path decoding. For each image variant I_i, the OCR system generates a set of candidate text predictions along with their confidence scores. Formally, the OCR output can be expressed as:

$$R_i = \{(T_{i1}, C_{i1}), (T_{i2}, C_{i2}), \ldots, (T_{in}, C_{in})\} \tag{13}$$

where T_{ij} is the j-th text candidate from image variant I_i, and C_{ij} is the associated confidence score. For each image variant, the candidate with the highest confidence is selected as the representative result:

$$T_i = \arg\max_j (C_{ij}) \tag{14}$$

Thus, each image variant produces one text candidate T_i that is used in the next selection stage. The decoding method used is CTC best-path decoding, which selects the character sequence with the highest probability from

the model output. Mathematically, if y is the character label sequence and π is a possible path, then the output probability is defined as:

$$P(y \mid x) = \sum_{\pi \in B^{-1}(y)} P(\pi \mid x) \tag{15}$$

However, in the best-path decoding approach, the best path is selected directly as:

$$y^* = B\left(\arg\max_{\pi} P(\pi \mid x)\right) \tag{16}$$

where B is the mapping function that removes repeated characters and blank symbols. This approach is chosen because it has low computational complexity and is suitable for real-time systems. However, it also has limitations because it does not consider all possible paths, which can lead to errors for visually similar characters. Therefore, this study does not rely solely on a single OCR result, but instead utilizes multiple candidates from various image variants, which are further processed in the Adaptive Multi-Variant Enhancement and OCR Selection stage to improve overall accuracy.

3.2.6 OCR Text Normalization

Character recognition results produced by OCR often contain noise, invalid characters, and inconsistent formats due to variations in visual conditions and prediction errors. Such inconsistencies can reduce accuracy in both the evaluation and candidate-selection stages. Therefore, a text normalization process is applied to improve the consistency and reliability of OCR outputs before the adaptive selection stage. Given the raw OCR output T^{raw}, the normalization process produces a standardized text T^{norm} through the following transformation stages:

1. Removal of Non-Alphanumeric Characters
   Characters that are not part of the alphanumeric set are removed using a filtering function:

   $$T_1 = \{\, c \in T^{\text{raw}} \mid c \in [A\text{-}Z \cup 0\text{-}9] \,\} \tag{17}$$

   This step removes symbols, spaces, and noise characters that are irrelevant to the license plate structure.
2. Uppercase Letter Transformation
   All characters are converted to uppercase to ensure representation consistency:

   $$T_2 = \text{Uppercase}(T_1) \tag{18}$$

   This transformation is important because OCR may produce inconsistent combinations of lowercase and uppercase letters.
3. Removal of Noise Prefix
   OCR often generates additional characters at the beginning of the output that do not align with the expected license plate pattern. To address this issue, trimming is performed based on the appearance of a valid initial pattern. In general, this process can be modeled as:

   $$T_3 = \text{Trim}(T_2) \tag{19}$$

   where the Trim function removes leading characters until a valid starting pattern is found, for example, two consecutive uppercase letters as a region code. This approach is flexible and does not rely on a single fixed format.
4. Handling Character Ambiguity
   Some characters have high visual similarity and are often misinterpreted by OCR, such as:

   $$O \leftrightarrow 0, \quad I \leftrightarrow 1, \quad S \leftrightarrow 5, \quad B \leftrightarrow 8$$

To reduce these errors, character mapping is performed using a substitution function:

$$T^{\text{norm}} = f_{\text{amb}}(T_3) \tag{20}$$

where f_{amb} replaces ambiguous characters with a more stable form, for example:

$$O \to 0, \quad I \to 1$$

This approach improves result consistency without relying on strict contextual assumptions.

Overall, the normalization process can be formulated as:

$$T^{\text{norm}} = f_{\text{amb}}\Big(\text{Trim}\big(\text{Uppercase}(\text{Filter}(T^{\text{raw}}))\big)\Big) \tag{21}$$

This normalization process aims to: (i) reduce OCR output variation caused by visual noise; (ii) improve consistency in text representation across image variants; (iii) minimize errors caused by ambiguous characters; and (iv) facilitate the evaluation and selection of OCR candidates. Thus, the normalization stage becomes a crucial component in improving system stability before the adaptive OCR selection process. This normalization is designed to be domain-aware while remaining flexible to variations in license plate formats, so it does not limit the system's generalization capability.

3.2.7 Multi-Grammar Plate-Aware OCR Selection

The OCR stage produces multiple text candidates for each license plate because PaddleOCR is applied to three ROI image variants, namely RAW, CLAHE-light, and SR2x + CLAHE-light. Let $K = 3$ denote the number of image variants, resulting in OCR candidates $\{O_k\}$ for $k = 1, 2, 3$, where each candidate is represented as a fixed-length string of length $L = 10$:

$$O_k = \{C_{k,1}, C_{k,2}, \ldots, C_{k,10}\} \tag{22}$$

PaddleOCR also generates a confidence value per character, denoted by $p_{k,j}$, for each character position in candidate O_k, where $j = 1, \ldots, 10$ corresponds to the character positions in the string. To select the most reliable final recognition result, this study applies a plate-aware scoring and selection mechanism that leverages constraints on license plate structure. To improve generalization, this approach employs multi-grammar constraints. The plate structure is modeled as a set of position-based or regular-expression rules:

$$G = \{G_1, G_2, \ldots, G_n\} \tag{23}$$

where each G_i represents a valid pattern, for example, `[A-Z]{2}[0-9]{1,2}[A-Z]{0,3}[0-9]{1,4}`.

With this approach, structural validity is no longer restricted to a single template, but can accommodate various license plate formats from different regions. The structural validity score is then defined as:

$$V_k = \max_{G_i \in G} \text{ValidityScore}(O_k \mid G_i) \tag{24}$$

This formulation allows the system to select the OCR candidate that is most consistent with one of the valid grammars, thereby improving flexibility and generalization.

1. OCR Confidence Score
   The confidence score of candidate O_k is defined as the average character confidence:

$$C_k = \frac{1}{10} \sum_{j=1}^{10} p_{k,j} \tag{25}$$

The value $C_k \in [0, 1]$ represents the reliability of the OCR prediction at the character level.

2. Structural Validity Score (Fixed-Position Constraint)
The structural validity is calculated based on the alignment of characters at each position with the allowed character set T_j:

$$V_k = \frac{1}{10}\sum_{j=1}^{10}\mathbf{1}[C_{k,j} \in T_j] \tag{26}$$

where the indicator function $\mathbf{1}[\cdot]$ takes the value 1 if the condition is satisfied and 0 otherwise. One example of the valid set for each position is defined as:

$$T_j = \begin{cases} \{R\}, & j = 1 \\ \{J\}, & j = 2 \\ \{0, \ldots, 9\}, & j \in \{3, 4, 7, 8, 9, 10\} \end{cases} \tag{27}$$

For multi-format license plates, T_j can be adjusted according to different plate structures, such as those used in India or other formats considered in the experiment. With this formulation, V_k provides a partial score in the range $[0, 1]$ that reflects the proportion of characters matching the expected plate structure. The 10-position structure in Equation (26) represents a canonical validation template derived from the standard Indian license plate format. Since actual license plates may contain variable numbers of visible characters, each OCR candidate is first aligned to the template before structural validation is performed. This normalization is applied only for candidate scoring and ranking and does not modify the OCR output used for EPA and CLA evaluation. Importantly, the 10-position template is used only for structural scoring and candidate ranking and does not constrain the final OCR output length. Consequently, license plates with variable character lengths remain fully supported during both recognition and evaluation.

3. Character Ambiguity Penalty
To reduce errors caused by visually similar characters, the set of ambiguous character pairs is defined as:

$$A = \{(O, 0), (I, 1), (B, 8), (S, 5), (Z, 2)\} \tag{28}$$

The ambiguity penalty is calculated as the total number of ambiguous characters appearing in unsuitable positions or in candidates that still require normalization:

$$A_k = \sum_{j=1}^{10}\mathbf{1}[C_{k,j} \in A_{\text{amb}}] \tag{29}$$

where A_{amb} is the set of characters involved in ambiguous pairs.

4. Score Function and Final Output Selection
The final score of OCR candidate O_k is defined as a linear combination of the three components:

$$S(O_k) = \alpha C_k + \beta V_k - \gamma A_k \tag{30}$$

where $\alpha, \beta, \gamma > 0$ are balancing weights. In the initial implementation, the weights are set equally, that is, $\alpha = \beta = \gamma = 1$, to maintain simplicity and avoid tuning to the test data.
The final output is selected as the candidate with the maximum score:

$$\hat{O} = \arg\max_{k \in \{1,\ldots,K\}} S(O_k) \tag{31}$$

In the case of tied scores, a tie-breaking rule is applied by prioritizing V_k (structural validity) first, followed by C_k (confidence). This formulation ensures that OCR output selection depends not only on model confidence but also on license plate structural consistency and potential character ambiguities that frequently arise in foggy images. Therefore, the selection mechanism is measurable, reproducible, and readily integrated into an end-to-end evaluation framework.

4. Result and Discussion

4.1. Evaluation Dataset

The evaluation dataset used in this study comprises 945 vehicle license plate images captured under foggy conditions, obtained from Mendeley Data [22]. This dataset includes various license plate formats, including the Indian license plate format, which has variable character lengths and alphanumeric combinations. The study does not restrict the system to a single license plate template, thereby allowing a more realistic evaluation of its performance in handling structural variations under real-world conditions. All data are annotated with ground truth and processed without per-image parameter adjustments to ensure system consistency and generalization. To ensure experimental validity and prevent data leakage, the YOLOv11 detection model was trained on a separate dataset curated in Roboflow, comprising 10,100 license plate images. This training dataset was not used during the evaluation stage. All 945 foggy images from Mendeley Data were used exclusively as test data. This separation ensures that the reported performance reflects the system's ability to generalize to unseen conditions.

4.2. Testing Environment and Configuration

All experiments were conducted in a consistent computing environment to ensure a fair performance comparison across methods. The system was implemented in Python using object detection, image processing, and optical character recognition (OCR) libraries. License plate detection was performed using the Ultralytics framework (version 8.4.35). The implementation used in this study corresponds to the official Ultralytics release available through the Ultralytics Python package version 8.4.35. The detector was initialized from the pretrained YOLO11n model (`yolo11n.pt`) and subsequently fine-tuned on the training dataset used in this study. No modifications were made to the original detector implementation.

The OCR component was implemented using PaddleOCR version 2.7.0.3 (PP-OCRv4) with the `SVTR_LCNet` recognizer. The OCR configuration employed `det_db_thresh = 0.3`, `det_db_box_thresh = 0.6`, `det_db_unclip_ratio = 1.5`, `rec_batch_num = 6`, and `drop_score = 0.5`. Angle classification was enabled (`use_angle_cls = True`) with `cls_batch_num = 6`. All OCR experiments used English-language recognition and identical inference settings across all enhancement variants. The testing was performed on a system equipped with an Intel® Core™ i7-12650H processor, 16 GB RAM, and an NVIDIA RTX 3050 4 GB Laptop GPU. All detection, image enhancement, OCR, and plate-selection parameters were kept fixed without per-image adjustments to maintain consistency, reproducibility, and the system's generalization capability.

4.3. Evaluation Metrics

The performance of the ALPR system is evaluated using three key metrics that reflect recognition performance at both the plate and character levels: Exact Plate Accuracy (EPA), Character-level Accuracy (CLA), and Readability Rate (RR). This combination of metrics was chosen to provide a comprehensive view of recognition accuracy and system robustness under degraded image conditions. Exact Plate Accuracy (EPA) measures the percentage of license plates that are fully recognized correctly, that is, when all characters in the OCR result exactly match the ground truth. This metric represents end-to-end recognition performance and is defined as:

$$\mathrm{EPA} = \frac{N_{\mathrm{exact}}}{N_{\mathrm{total}}} \times 100\% \tag{32}$$

where N_{exact} is the number of correctly recognized plates, and N_{total} is the total number of test samples. Character-level Accuracy (CLA) evaluates recognition accuracy at the individual-character level, making it more sensitive to partial errors. This metric is calculated as:

$$CLA = \frac{\sum_{i=1}^{N} C_{\mathrm{correct},i}}{\sum_{i=1}^{N} C_{\mathrm{total},i}} \times 100\% \tag{33}$$

where C_{correct} represents the number of correctly recognized characters, and C_{total} is the total number of characters in the dataset. Readability Rate (RR) measures the proportion of license plate images that produce valid or readable

OCR outputs, thus reflecting the system's ability to avoid total recognition failure. Its formula is:

$$\mathrm{RR} = \frac{N_{\mathrm{readable}}}{N_{\mathrm{total}}} \times 100\% \tag{34}$$

All metrics are calculated automatically by comparing the recognition results against the ground truth without manual intervention, ensuring objective and consistent evaluation across all test samples.

4.4. Comparison of Dehazing Methods (DCP vs Non-local)

To evaluate image reconstruction quality prior to the detection and OCR stages, a comparison was conducted between two dehazing approaches: the Dark Channel Prior (DCP) method and a non-local image dehazing technique. The evaluation used Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) to measure restoration quality. As shown in Table 1, the DCP method achieves higher PSNR and SSIM compared to the non-local approach.

Table 1. PSNR and SSIM of two dehazing methods As shown in Table 1, the DCP method achieves higher PSNR and SSIM values than the non-local approach.

Table 1. PSNR and SSIM comparison of two dehazing methods

Method	**PSNR (dB)**	**SSIM**
Dark Channel Prior (DCP)	29.73	0.78
Non-local Image Dehazing	29.07	0.75

Table 1 shows that DCP yields higher PSNR and SSIM values compared to the non-local method. These results indicate better preservation of spatial structure and image contrast after the dehazing process, allowing crucial visual information for the detection and character recognition stages to be retained more effectively [25].

(a) Haze Image (b) Dark Channel Prior (c) Non-local Image Dehaze

Figure 3. Visual results of DCP and non-local dehazing.

These quantitative findings are visually confirmed in Figure 3. The DCP result displays clearer local contrast and more defined edge details, making the character structure on the license plate more readable. In contrast, the non-local approach can, in some cases, produce an over-smoothing effect, leading to a loss of fine texture and decreased color saturation.The distribution of PSNR and SSIM values in Figure 4 shows a consistent trend, where DCP maintains more stable reconstruction quality across all test samples. This indicates that dark channel-based transmission estimation is more effective at representing fog distribution than the non-local approach, which relies on aggregating similar patches.

The practical impact on the detection stage is shown in Figure 5. In the foggy images without preprocessing, YOLOv11 fails to localize the license plate. After applying DCP, the plate can be accurately detected and extracted. In contrast, the results from non-local dehazing remain unstable, indicating that the quality of image restoration affects the reliability of subsequent detection. Overall, these results emphasize that selecting the appropriate dehazing method is crucial in providing a reliable visual foundation for the detection and OCR stages in the ALPR system.

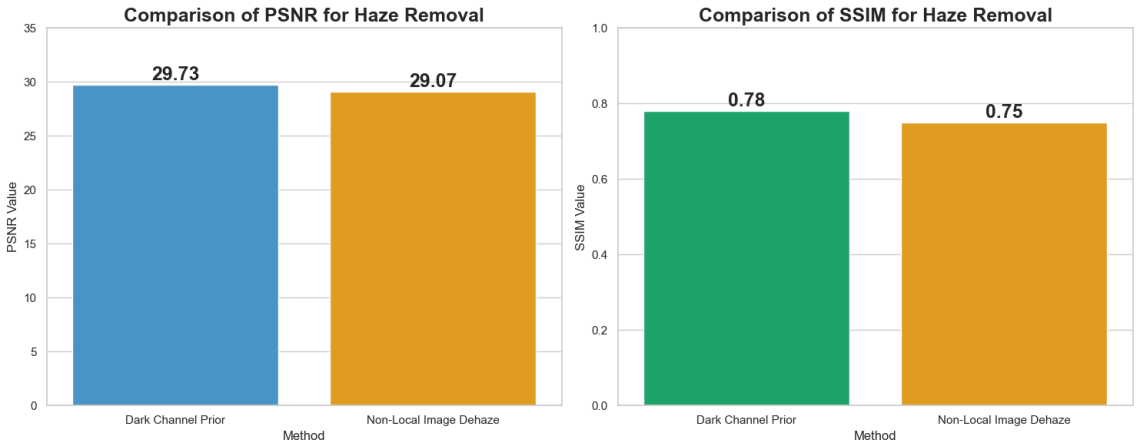

Figure 4. PSNR and SSIM Scores of DCP and Non-local Dehazing Methods.

(a) Haze Image (b) Dark Channel Prior (c) Non-Local Image Dehaze

Figure 5. Vehicle License Plate Detection Results.

4.5. License Plate Detection and Runtime Analysis

The license plate detection and cropping stage using YOLOv11 plays a crucial role in stabilizing the overall system performance. As illustrated in Figure 5, YOLOv11 accurately localizes license plates even when the input images are degraded by haze. The proposed method exhibits consistent detection performance under foggy conditions, as evidenced by the experimental results. The YOLOv11 model achieves a precision of 0.978, a recall of 0.952, an mAP@0.5 of 0.975, and an mAP@0.5:0.95 of 0.718, with an inference time of approximately 2.0 ms per image (approximately 500 FPS) for the detection stage alone, excluding the dehazing and OCR processes. These results satisfy real-time processing requirements for practical ALPR deployment. The normalized confusion matrix and the training–validation curves (Figures 6 and 7) further demonstrate a high true-positive rate, relatively low false negatives, and a stable, convergent training process without significant signs of overfitting. Overall, these findings confirm that YOLOv11 provides reliable and efficient license plate localization, forming a robust foundation for the subsequent enhancement and OCR stages.

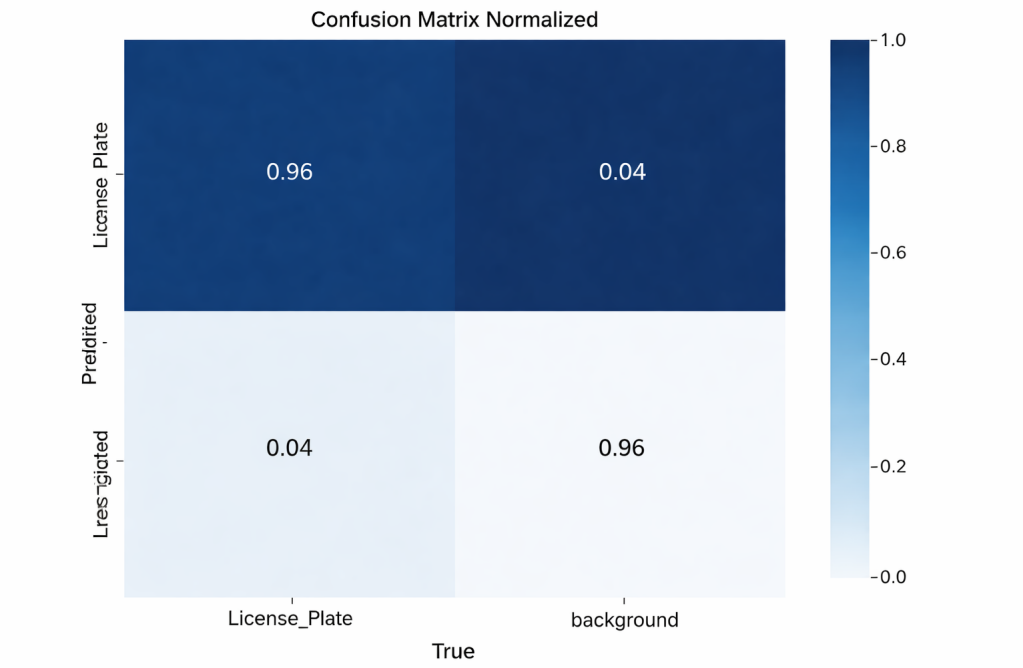

Figure 6. Normalized confusion matrix of YOLOv11.

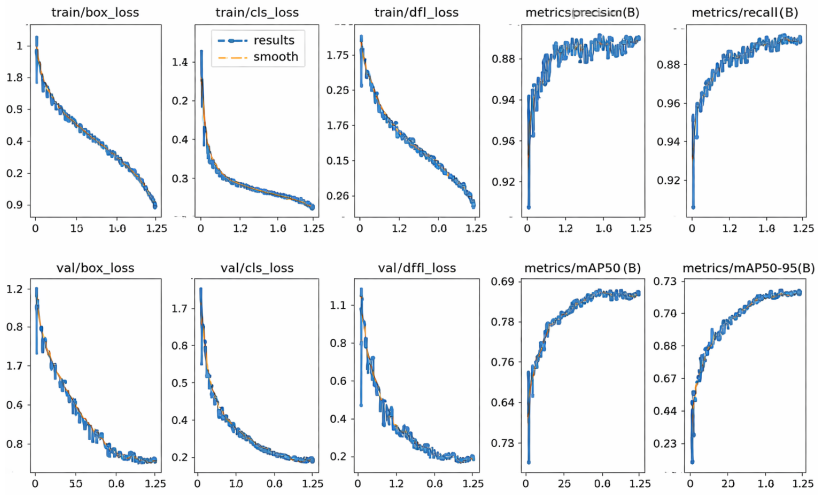

Figure 7. Training and validation curves of YOLOv11.

The detection performance comparison with YOLOv8 (Table 2) indicates that the proposed YOLOv11 achieves a higher mAP@0.5 (97.5% vs. 97.3%) and slightly better recall (95.2% vs. 94.8%). YOLOv11 is selected as the primary detection model because it provides a favorable balance between accuracy and computational efficiency, while still satisfying real-time requirements at 500 FPS, even though YOLOv8 has a slightly faster inference time (1.8 ms; 556 FPS). This performance improvement is mainly attributed to the Cross-Stage Partial (CSP)

architecture and the dynamic anchor allocation strategy, which improve detection sensitivity for small-scale objects such as vehicle license plates [15].

Table 2. Comparison of license plate detection performance

Model	Precision (%)	Recall (%)	mAP@0.5 (%)	Inference	FPS
YOLOv8	98.6	94.8	97.3	1.8 ms	556
YOLOv11 (proposed)	97.8	95.2	97.5	2.0 ms	500

4.5.1 End-to-End Runtime Analysis

To provide a realistic assessment of computational performance, an end-to-end runtime analysis was conducted on the complete ALPR pipeline using an Intel® Core™ i7-12650H CPU and an NVIDIA RTX 3050 Laptop GPU (4 GB VRAM). Runtime measurements were collected from all 945 test images. The average processing time (mean $\pm$ standard deviation) was calculated for each major stage of the system.

Table 3. Runtime Breakdown of Individual Pipeline Stages

Pipeline Stage	Mean $\pm$ Std (ms/image)	FPS
DCP Dehazing	26.32 ± 6.14	38.00
YOLOv11 Detection + Crop	26.07 ± 13.69	38.36
OCR without preprocessing	119.20 ± 76.83	8.39
CLAHE-light + OCR	138.21 ± 71.02	7.24
Real-ESRGAN + OCR	315.41 ± 291.61	1.74
Proposed Plate-Aware Selection	419.03 ± 156.96	2.39

Table 4. End-to-End Runtime Comparison

Method	Runtime (ms/image)	FPS
YOLO Crop + OCR (No DCP)	119.20	8.39
DCP + YOLO Crop (RAW)	171.59	5.83
DCP + YOLO Crop + CLAHE-light	190.60	5.25
DCP + YOLO Crop + Real-ESRGAN	575.95	1.74
Proposed Method	471.42	2.12

As shown in Table 4, the proposed framework requires an average end-to-end processing time of 471.42 ms per image, corresponding to approximately 2.12 FPS. The computational cost is primarily dominated by the multi-variant OCR stage, where three enhancement branches are processed independently and subsequently evaluated by the plate-aware selection mechanism. Compared with the strongest single-variant baseline (DCP + YOLO Crop (RAW)), which requires 171.59 ms per image, the proposed method introduces an additional computational overhead of approximately 299.83 ms per image. However, this overhead increases Exact Plate Accuracy from 89.42% to 93.12%, an absolute improvement of 3.70 percentage points. The results also show that Real-ESRGAN-based enhancement is computationally expensive, requiring 575.95 ms per image and achieving only 1.74 FPS. Despite producing visually sharper images, the additional processing cost does not translate into superior OCR accuracy compared with the proposed framework.

Conventional real-time video processing typically requires approximately 30 FPS (33.3 ms per frame), while practical interactive systems often target response times below 200 ms per image. The proposed framework exceeds both thresholds, operating at approximately 2.12 FPS with an average latency of 471.42 ms per image. Therefore,

although the YOLOv11 detector itself achieves approximately 500 FPS when evaluated independently, this value should not be interpreted as the performance of the complete ALPR system. The proposed framework is more appropriately characterized as an accuracy-oriented recognition system designed for robust operation in hazy conditions rather than as a strictly real-time ALPR solution.

4.6. Comparative Analysis of Image Enhancement Variants

After the cropping stage, the system generates three license plate image variants, namely RAW, CLAHE-light, and SR2x + CLAHE-light, as illustrated by the parallel branches in Figure 8. This A multi-variant strategy is designed to accommodate the heterogeneous levels of visual degradation commonly encountered in real-world data.

Figure 8. (a) RAW, (b) CLAHE-light, (c) SR2x + CLAHE-light.

By providing multiple lightweight enhancement options, the system avoids reliance on a single fixed preprocessing method that may not be optimal under all conditions. Instead, each variant offers complementary visual characteristics: RAW preserves the original structural information, CLAHE-light improves local contrast, and SR2x + CLAHE-light enhances spatial resolution and character clarity. This design enables the OCR module to operate on diverse image representations, thereby increasing the likelihood of accurate recognition under varying haze, blur, and low-resolution conditions. The RAW image preserves the original character stroke shapes with minimal distortion. In many cases, this variant alone is sufficient to achieve accurate character recognition, particularly when the structural information of the characters remains intact despite low contrast. The CLAHE-light variant enhances local contrast in a controlled manner, making the characters more distinguishable without significantly altering stroke geometry. Experimental results indicate that CLAHE-light often provides the best balance between visual enhancement and OCR stability, especially for license plates with reduced contrast due to haze.

In contrast, the SR2x + CLAHE-light combination produces visually sharper images. However, it may also introduce micro-distortions in character shapes due to interpolation and subsequent contrast amplification. These distortions can increase ambiguity between visually similar characters, such as O–0 or S–5, potentially reducing exact character recognition accuracy. These findings suggest that improved visual quality does not necessarily translate into better OCR performance. From a technical perspective, bicubic interpolation estimates pixel values using third-order polynomials derived from neighboring pixels, which can amplify high-frequency components and noise during upsampling. When combined with CLAHE, these artifacts can further exaggerate subtle contour distortions and reduce OCR prediction stability.

4.7. Adaptive Multi-Varian Enhancement and Plate-Aware Selection

After the license plate detection and cropping process using YOLOv11, the cropped image may still exhibit visual degradation due to residual fog, resolution differences, and the diverse physical conditions of the license plates. To address these issues, this study implements an adaptive multi-variant enhancement scheme that generates several lightweight image variants, namely RAW, CLAHE-light, and SR2x + CLAHE-light, without relying on a single image enhancement strategy. Each variant is processed independently by the OCR module to produce multiple recognition candidates. The system then applies a plate-aware scoring and selection mechanism to choose the most structurally reasonable OCR result.

This mechanism evaluates each candidate based on three main aspects: license plate format consistency, handling of common OCR character ambiguities (e.g., O–Q and 0–O), and recognition confidence level. With this approach,

the system does not rely solely on visual image quality, but also assesses the structural consistency of the OCR results against vehicle license plate rules.

Figure 9. OCR results for different enhancement variants.

Figure 9 presents a comparison of OCR results across various image enhancement variants. The RAW variant retains the natural character stroke patterns and is often sufficient under mild degradation conditions. CLAHE-light enhances local contrast in a controlled manner and generally provides the best balance between readability and OCR stability. In contrast, the SR2x + CLAHE-light combination produces a sharper visual appearance, but the upsampling process can also amplify high-frequency components and noise. Technically, bicubic interpolation estimates pixel values using polynomial combinations of neighboring pixels, which can strengthen edge artifacts. When followed by contrast enhancement, these minor distortions can increase character ambiguity and reduce exact recognition accuracy. These findings indicate that improved visual quality does not always correlate with improved OCR performance.

The quantitative performance of all evaluated methods is presented in Table 5. Among the compared approaches, the proposed DCP + Plate-Aware Selection method achieved the best overall results, with a Readability Rate (RR) of 100.00%, Exact Plate Accuracy (EPA) of 93.12%, and Character-Level Accuracy (CLA) of 98.63%. These results demonstrate that the proposed method provides the most robust and consistent recognition performance under hazy conditions. Table 5 reports both the enhancement variants used by the proposed framework and additional comparative baselines, including a DCP + Real-ESRGAN benchmark evaluated independently from the proposed selection pipeline.

Table 5. Summary of License Plate Recognition Performance Under Hazy Conditions

Method	**RR (%)**	**EPA (%)**	**CLA (%)**	p**-value**
YOLO Crop Without Preprocessing	90.37	80.42	97.04	< 0.001
DCP + YOLO Crop (RAW)	98.94	89.42	97.83	< 0.001
DCP + YOLO Crop + CLAHE-light	98.52	73.15	92.51	< 0.001
DCP + Real-ESRGAN	99.68	83.23	96.10	< 0.001
DCP + Plate-Aware Selection (Proposed Method)	100.00	93.12	98.63	–

Compared with the baseline YOLO Crop Without Preprocessing, the proposed method improved Exact Plate Accuracy from 80.42% to 93.12%, corresponding to an absolute gain of 12.70 percentage points. A clear improvement was also observed over DCP + YOLO Crop (RAW), with EPA increasing from 89.42% to 93.12%, and additional gains in RR from 98.94% to 100.00% and in CLA from 97.83% to 98.63%. This finding indicates that the proposed plate-aware selection strategy yields better final recognition performance than relying solely on a single OCR result from the raw enhanced plate image.

The comparison with enhancement-based alternatives further supports this conclusion. Although DCP + Real-ESRGAN achieved a high Readability Rate of 99.68%, its Exact Plate Accuracy remained at 83.23%, which is 9.89 percentage points lower than that of the proposed method. Likewise, DCP + YOLO Crop + CLAHE-light produced even lower performance, with 73.15% EPA and 92.51% CLA, despite applying additional contrast enhancement. This result suggests that stronger visual enhancement does not necessarily translate into better OCR accuracy. In some cases, enhancement may alter character morphology or introduce artifacts that negatively affect recognition.

For completeness, Real-ESRGAN was also evaluated as an additional comparative benchmark to assess the potential benefit of deep-learning-based super-resolution. However, Real-ESRGAN was not incorporated into the final multi-variant enhancement pipeline because it did not consistently improve OCR recognition accuracy. Therefore, the proposed framework uses the RAW and SR2x + CLAHE-light variants for candidate generation, while Real-ESRGAN is included only for comparative evaluation.

Overall, the results confirm that structural consistency and plate-aware candidate selection are more important than visual enhancement alone for improving ALPR performance in hazy scenes. While image enhancement can improve visual readability, final recognition accuracy depends more strongly on whether the selected OCR output preserves the license plate's expected alphanumeric structure. Therefore, the proposed combination of DCP-based restoration and plate-aware OCR selection offers a more stable and reliable framework for real-world ALPR deployment under degraded visual conditions.

(a) RAW (b) Real-ESRGAN (c) SR2X + CLAHE-light

Figure 10. Comparison of representative OCR failure cases.

Figure 10 shows examples of license plate character recognition failures across various image enhancement strategies. In the RAW image (Figure 10a), the license plate characters remain affected by residual fog and low contrast, leading to OCR errors. The application of Real-ESRGAN (Figure 10b) significantly improves visual sharpness but introduces subtle distortions in character shapes due to the super-resolution process, thereby decreasing exact recognition accuracy. Meanwhile, the combination of SR2x + CLAHE-light (Figure 10c) improves visual readability without increasing excessive detail complexity, but still produces character ambiguity under certain conditions. These findings reinforce the idea that visual quality enhancement alone does not always correlate with OCR robustness, highlighting the need for an adaptive, structure-based plate selection mechanism to achieve the most consistent recognition results.

In this study, the term "adaptive" refers to the framework's ability to dynamically select the most suitable OCR output among multiple enhancement variants on a per-image basis. Adaptiveness is achieved through candidate-level decision-making rather than through parameter learning or model retraining. Consequently, different enhancement variants may be selected for different license plate images depending on the quality and structural consistency of the OCR outputs.

4.7.1 Parameter Sensitivity Analysis

To evaluate the robustness of the proposed plate-aware scoring function, a sensitivity analysis was conducted by varying the weighting coefficients α, β, and γ while keeping the remaining parameters fixed. The analysis was performed on the complete evaluation dataset, comprising 945 hazy license plate images.

Table 6 presents the sensitivity analysis results of the scoring function weights.

Table 6 summarizes the Exact Plate Accuracy (EPA) obtained under different weight configurations. The highest EPA, 93.56%, was achieved with $\alpha = 0.5$, $\beta = 1.0$, and $\gamma = 1.0$. The default configuration ($\alpha = \beta = \gamma = 1.0$) achieved an EPA of 93.44%, representing a difference of only 0.12 percentage points. Across all tested configurations, EPA ranged from 93.33% to 93.56%, corresponding to a maximum variation of only 0.23 percentage points. These results indicate that the proposed scoring mechanism is highly robust to moderate variations in weighting coefficients and does not depend on precise hyperparameter tuning. Therefore, the effectiveness of the proposed framework primarily stems from the multi-variant OCR selection strategy and structural validation mechanism rather than carefully optimized weight values.

Table 6. Sensitivity Analysis of Scoring Function Weights

α	β	γ	**EPA (%)**
0.5	1.0	1.0	93.56
1.0	1.0	1.0	93.44
1.5	1.0	1.0	93.44
2.0	1.0	1.0	93.33
1.0	0.5	1.0	93.33
1.0	1.5	1.0	93.33
1.0	2.0	1.0	93.33
1.0	1.0	0.5	93.33
1.0	1.0	1.5	93.44
1.0	1.0	2.0	93.44

4.8. Analysis of CLAHE-Light Failure Cases

Although CLAHE-light was introduced to improve local contrast after the DCP restoration stage, the experimental results indicate that this enhancement strategy negatively affected OCR performance. To investigate the cause of this degradation, additional image quality analyses were conducted using Contrast-to-Noise Ratio (CNR) and Laplacian-based sharpness measurements.

Table 7 summarizes the image quality characteristics of representative failure cases generated by the DCP + CLAHE-light enhancement pipeline. As shown in Table 7, CLAHE-light increased the average sharpness from

Table 7. Image Quality Analysis of DCP + CLAHE-light Failure Cases

Metric	**RAW**	**CLAHE-light**	Δ
CNR	29.318	27.829	-5.08%
Sharpness	2239.04	5811.94	$+159.57\%$
Failure Cases	–	195	–

2239.04 to 5811.94, representing an improvement of approximately 159.57%. This result indicates that CLAHE-light successfully enhanced local edges and fine image structures. However, the average CNR decreased from 29.318 to 27.829, representing a reduction of approximately 5.08%. The lower CNR suggests that the enhancement process did not improve the separability of license plate characters from the surrounding background, despite the increase in visual sharpness.

A detailed inspection of the recognition results identified 195 failure cases in which the DCP + YOLO Crop (RAW) variant correctly recognized the license plate while the CLAHE-light variant produced incorrect predictions. Representative examples of these failure cases are presented in Figure 11. In the first example, the RAW variant correctly recognized the license plate, whereas CLAHE-light generated an incorrect prediction despite producing a higher OCR confidence score. In the second example, CLAHE-light failed to produce a valid OCR result, whereas the RAW variant still yielded a readable prediction. These examples demonstrate that higher visual contrast and OCR confidence do not necessarily correspond to higher recognition accuracy.

Visual examination further revealed that CLAHE-light frequently amplified residual artifacts, noise, and local intensity variations inherited from the DCP restoration stage. Although the resulting images appeared visually sharper, the enhanced artifacts often distorted character boundaries, altered stroke continuity, and reduced OCR robustness. Consequently, the increase in perceived image sharpness did not translate into improved recognition performance. Instead, excessive local enhancement reduced character–background discriminability and negatively affected OCR accuracy.

These findings explain why the DCP + YOLO Crop (RAW) variant achieved higher recognition accuracy than the CLAHE-light variant despite producing visually less sharp images. It should be noted that only the standard CLAHE configuration (`clipLimit = 2.0` and `tileGridSize = 8 × 8`) was evaluated in this study.

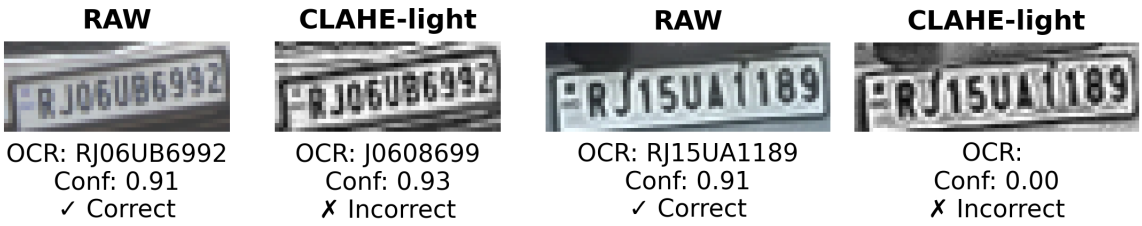

Figure 11. Representative failure cases of the DCP+CLAHE-light branch.

Therefore, the present results indicate that the observed degradation originates from the interaction between CLAHE enhancement and residual DCP artifacts, while the individual contributions of the clip limit and tile size remain subjects for future investigation. Nevertheless, CLAHE-light was retained as one candidate branch within the proposed multi-variant framework because the plate-aware selection mechanism prevents systematic performance degradation by selecting the OCR result with the highest confidence and structural validity score among all enhancement variants.

4.9. Statistical Significance Analysis

To determine whether the observed improvements are statistically significant, paired statistical tests were conducted over all 945 test images. For exact plate recognition accuracy (Readability), McNemar's test was applied using paired binary correctness outcomes between the proposed method and each baseline. For character-level accuracy (EPA), the Wilcoxon signed-rank test was employed on the per-image EPA scores.

Table 8 summarizes the results of McNemar's test.

Table 8. McNemar Test Results

Comparison	χ^2	**p-value**
Proposed vs No Preprocessing	101.15	< 0.001
Proposed vs RAW	26.88	< 0.001
Proposed vs CLAHE-light	182.34	< 0.001
Proposed vs Real-ESRGAN	82.05	< 0.001

As shown in Table 8, McNemar's test indicates statistically significant differences between the proposed method and all baseline approaches ($p < 0.001$). The largest difference was observed against CLAHE-light ($\chi^2 = 182.34$), while even the strongest baseline, RAW, remained significantly inferior ($\chi^2 = 26.88$).

Table 9 presents the results of the Wilcoxon signed-rank test. Similarly, the Wilcoxon signed-rank test results

Table 9. Wilcoxon Signed-Rank Test Results

Comparison	**Statistic**	**p-value**
Proposed vs No Preprocessing	349.5	1.10×10^{-25}
Proposed vs RAW	82.0	7.67×10^{-7}
Proposed vs CLAHE-light	344.5	7.37×10^{-40}
Proposed vs Real-ESRGAN	648.0	2.04×10^{-16}

presented in Table 9 demonstrate that the proposed method achieves significantly higher character-level recognition accuracy than all competing methods ($p < 0.001$). The comparison against the strongest baseline (RAW) yielded a test statistic of 82.0 with $p = 7.67 \times 10^{-7}$, confirming that the observed EPA improvement is unlikely to be due to random variation. These results provide statistical evidence that the proposed plate-aware enhancement selection strategy consistently improves both exact plate recognition and character-level OCR accuracy across the entire dataset.

4.10. Ablation Study

An ablation study was conducted to evaluate the contribution of each major component in the proposed framework, particularly the preprocessing strategy and the plate-aware OCR selection mechanism. The results are summarized in Table 4.

Table 10. Ablation Analysis of Method Components

Method	Readability (%)	Exact Accuracy (%)	Analysis
YOLO Crop Without Preprocessing	90.37	80.42	Without preprocessing, performance is still limited.
DCP + YOLO Crop (RAW)	98.94	89.42	Significant improvement in readability and exact recognition accuracy.
DCP + YOLO Crop + CLAHE-light	98.52	73.15	Contrast enhancement increased visual readability, but reduced exact recognition accuracy.
DCP + Real-ESRGAN	99.68	83.23	Readability increased, but exact accuracy remained lower than the RAW baseline.
DCP + Plate-Aware Selection (Proposed)	100	93.12	Best overall results with perfect readability and the highest exact accuracy.

The ablation results demonstrate that haze removal is the most influential preprocessing step for improving OCR performance. Compared with the YOLO Crop Without Preprocessing baseline, applying DCP dehazing increased readability from 90.37% to 98.94% and improved exact accuracy from 80.42% to 89.42%, confirming the importance of restoring image visibility prior to OCR. However, stronger image enhancement did not necessarily lead to better recognition accuracy. Although DCP + YOLO Crop + CLAHE-light and DCP + YOLO Crop + Real-ESRGAN + OCR achieved high readability scores of 98.52% and 99.68%, their exact accuracies decreased to 73.15% and 83.23%, respectively. These findings indicate that visual enhancement alone may improve perceptual image quality while simultaneously introducing artifacts or distortions that negatively affect character recognition. The proposed DCP + Plate-Aware Selection framework achieved the highest overall performance, reaching 100.00% readability and 93.12% exact accuracy. This result suggests that preserving structural consistency and selecting OCR outputs based on confidence and plate-validity constraints are more effective than relying solely on increasingly aggressive image-enhancement techniques.

To further evaluate whether the additional complexity of the proposed multi-variant framework is justified, a comparison was conducted against the strongest single-variant baseline, namely DCP + YOLO Crop (RAW). The baseline achieved an Exact Plate Accuracy (EPA) of 89.42% and a Character-Level Accuracy (CLA) of 97.83%, whereas the proposed framework achieved 93.12% EPA and 98.63% CLA. This corresponds to absolute improvements of 3.70 percentage points in EPA and 0.80 percentage points in CLA. To determine whether these improvements were statistically significant, McNemar's test and the Wilcoxon signed-rank test were conducted. As reported in Tables 8 and 9, the proposed framework significantly outperformed the DCP + YOLO Crop (RAW) baseline (McNemar: $\chi^2 = 26.88$, $p < 0.001$; Wilcoxon: $p = 7.67 \times 10^{-7}$). Therefore, the observed gains are unlikely to be caused by random variation.

Although the proposed framework incurs additional computational cost due to multiple enhancement branches and a plate-aware selection stage, the statistically significant improvement in recognition accuracy demonstrates that the added complexity yields measurable benefits over the strongest single-variant baseline. These findings support the effectiveness of the proposed multi-variant OCR selection strategy for robust license plate recognition under hazy conditions.

4.11. Error Analysis

Evaluation on 945 test images showed that the system produced readable OCR outputs for all samples, resulting in a 100% readability rate. The proposed method achieved an End-to-End Exact Accuracy of 93.12% (880/945). At the character level, the average accuracy reached 98.63%, corresponding to a character error rate of 1.37%. Error analysis showed that most recognition failures were caused by ambiguity between visually similar characters, particularly O vs 0, I vs 1, and B vs 8. These ambiguities occurred more frequently in low-quality or low-contrast plate images, where character boundaries were less distinct. A smaller portion of errors was associated with more severe visual degradation, such as dense haze, motion blur, or imperfect ROI cropping. In such cases, even minor character deviations were sufficient to cause exact plate mismatch.Overall, the remaining errors appear to be more strongly associated with limitations in input image quality than with the plate-aware selection mechanism itself. This suggests that, although the proposed selection strategy improves final recognition reliability, system performance is still constrained by the visual quality of the detected license plate region.

5. Conclusion

This study proposes an adaptive Automatic License Plate Recognition (ALPR) framework that combines DCP-based dehazing, lightweight multi-variant image enhancement, and a plate-aware OCR selection mechanism to improve recognition performance under hazy conditions. Experimental results on 945 hazy license plate images achieved 100.00% readability, 93.12% exact plate accuracy, and 98.63% character-level accuracy, outperforming both single-enhancement and super-resolution-based approaches. Ablation and statistical analyses further confirmed that the proposed multi-variant selection strategy provides significant accuracy improvements over the strongest baseline. Although the complete pipeline requires 471.42 ms per image (2.12 FPS), the results demonstrate that combining adaptive enhancement with structure-aware OCR selection offers a robust and effective solution for ALPR in degraded hazy environments.

5.1. Limitation

Although the proposed Multi-Variant Enhancement and Plate-Aware OCR Selection framework achieved substantial improvements in readability and recognition accuracy under hazy conditions, several limitations remain. First, the framework's performance is still affected by severe image degradation, including dense haze, motion blur, low contrast, and illumination variations. In addition, visually similar characters such as O/0, I/1, and B/8 may still lead to recognition errors, particularly in heavily degraded license plate images.

Second, the current framework relies on structural rules derived from the license plate format represented in the evaluation dataset. Consequently, its performance may decrease when applied to non-standard plates, temporary registrations, damaged plates, or license plate formats from other countries. Furthermore, the evaluation was conducted using a single hazy benchmark dataset, and additional cross-dataset validation is required to further assess the generalization capability.

Finally, the proposed framework introduces additional computational cost due to DCP dehazing, multi-variant enhancement generation, multiple OCR executions, and plate-aware candidate selection. Runtime analysis showed an average end-to-end latency of 471.42 ms per image (2.12 FPS) on an Intel® Core™ i7-12650H CPU and NVIDIA RTX 3050 Laptop GPU (4 GB VRAM). Although this overhead significantly improves recognition accuracy, it exceeds conventional real-time requirements and may limit deployment in high-throughput video analytics applications. Future work will focus on lightweight OCR architectures, more efficient candidate selection strategies, and broader cross-dataset evaluation while maintaining the recognition accuracy achieved by the proposed framework.

REFERENCES

1. C.-N. E. Anagnostopoulos, I. Anagnostopoulos, V. Loumos, and E. Kayafas, *A License Plate-Recognition Algorithm for Intelligent Transportation System Applications*, IEEE Transactions on Intelligent Transportation Systems, vol. 7, pp. 377–392, 2006.
2. R. Laroca, L. A. Zanlorensi, G. R. Gonçalves, E. Todt, W. R. Schwartz, and D. Menotti, *An Efficient and Layout-Independent Automatic License Plate Recognition System Based on the YOLO Detector*, CoRR, vol. abs/1909.0, 2019.
3. K. He, J. Sun, and X. Tang, *Single Image Haze Removal Using Dark Channel Prior*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2341–2353, 2011.
4. L. Zhang, P. Wang, H. Li, Z. Li, C. Shen, and Y. Zhang, *A Robust Attentional Framework for License Plate Recognition in the Wild*, IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 11, pp. 6967–6976, 2021.
5. 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 Transactions on Pattern Analysis and Machine Intelligence, vol. 39, pp. 2298–2304, 2015.
6. A. Younesi, M. Ansari, M. Fazli, A. Ejlali, M. Shafique, and J. Henkel, *A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends*, IEEE Access, vol. 12, pp. 41180–41218, 2024.
7. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You Only Look Once: Unified, Real-Time Object Detection*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788, 2016.
8. S. Yu, D. Seo, and J. Paik, *Haze Removal Using Deep Convolutional Neural Network for KOMPSAT-3A Multispectral Remote Sensing Imagery*, Engineering Applications of Artificial Intelligence, vol. 123, p. 106481, 2023.
9. W. Zhu, L. Yang, and X. Zang, *Application of Dark Channel Prior Principle to Licensed Plate Detection in Foggy Weather*, in Proceedings of the 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 2591–2595, 2018.
10. C. Khanna, B. Bisht, D. Kamshetty, H. Bhardwaj, and M. Bhutani, *Enhancing License Plate Recognition Using YOLO-NAS, YOLOv8, and SORT Algorithms*, Journal of Trends and Challenges in Artificial Intelligence, vol. 2, no. 1, pp. 19–26, 2025.
11. X. Wang, L. Xie, C. Dong, and Y. Shan, *Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data*, in Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 1905–1914, 2021.
12. H. Shi and D. Zhao, *License Plate Localization in Complex Environments Based on Improved GrabCut Algorithm*, IEEE Access, 2022.
13. R. Laroca, E. Cardoso, D. Lucio, V. Estevam, and D. Menotti, *On the Cross-dataset Generalization in License Plate Recognition*, pp. 166–178, 2022.
14. R. Li, J. Pan, Z. Li, and J. Tang, *Single Image Dehazing via Conditional Generative Adversarial Network*, in EEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018,pp. 8202–8211.
15. Sutikno, A. Sugiharto, and R. Kusumaningrum, *Enhanced Automatic License Plate Detection and Recognition Using CLAHE and YOLOv11*, Engineering, Technology & Applied Science Research, vol. 15, no. 1, pp. 20271–20278, 2025.
16. Arafat, P. N. Andono, A. Syukur, A. Affandy, and M. A. Soeleman, *Enhancing License Plate Recognition in Foggy Condition Based on YOLOv8 and Dark Channel Prior*, in 2024 International Seminar on Application for Technology of Information and Communication (iSemantic), pp. 540–546, 2024.
17. K. Zuiderveld, *Contrast Limited Adaptive Histogram Equalization*, in Graphics Gems IV, pp. 474–485, 1994.
18. A. Amin, R. Mumtaz, M. J. Bashir, and S. M. H. Zaidi, *Next-Generation License Plate Detection and Recognition System Using YOLOv8*, in HONET, pp. 179–184, 2023.
19. T. A. Shyaa and A. A. Hashim, *Superior Use of YOLOv8 to Enhance Car License Plates Detection Speed and Accuracy*, Revue d'Intelligence Artificielle, vol. 38, no. 1, pp. 139–145, 2024.
20. Y. Guo, X. Zhang, J. Liu, and H. Chen, *Energy-efficient Image Enhancement Using Bicubic Interpolation with Edge-aware Filtering*, Image and Vision Computing, vol. 140, p. 104837, 2024.
21. Z. Y. Lim, Z. J. Bong, Y. H. Pang, S. Y. Ooi, W. H. Khoh, and F. S. Hiew, *Integration of YOLOv8 and PaddleOCR for Car Plate Recognition System in Diverse Environments: Malaysia*, Journal of Engineering Science and Technology, vol. 20, no. 1, pp. 250–270, 2025.
22. D. Sharma, S. Sharma, and V. Bhatnagar, *Foggy-Hazy License Plates Images*, Mendeley Data, 2023.
23. N. Haque, S. Islam, R. A. Tithy, and M. S. Uddin, *Automatic Bangla License Plate Recognition System for Low-Resolution Images*, in STI, pp. 1–6, 2022.
24. D. A. Subhahan, S. R. Divya, U. K. Sree, T. Kiriti, and Y. Sarthik, *An Efficient and Robust ALPR Model Using YOLOv8 and LPRNet*, in ICRAIS, pp. 260–265, 2023.
25. D. Berman, *Non-Local Image Dehazing*, in CVPR, pp. 1674–1682, 2016.