Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
18/2 (2025), 141-153. DOI: http://dx.doi.org/10.21609/jiki.v18i2.1320

Indonesian License Plate Detection and Recognition System using Gaussian YOLOv7

Juan Thomas Wijaya* and Aniati Murni Arymurthy†

Faculty of Computer Science, University of Indonesia, Depok, 16424, Indonesia

*E-mail: *juan.thomas@ui.ac.id, †aniati@cs.ui.ac.id*

Abstract

In recent years, Automatic License Plate Recognition (ALPR) systems have garnered attention in computer vision research. However, practical applications face challenges such as inconsistent lighting, diverse license plate designs, and environmental variations, which increase the complexity of the task and lead to more false detections. To address these issues, we proposed Gaussian YOLOv7 for license plate detection and character recognition within ALPR systems, along with the Spatial Transformer Network (STN) for rectifying license plate orientation, aiming to enhance performance and adaptability to real-world scenarios. Additionally, we introduced a novel dataset for Indonesian ALPR systems to ensure robust detection and a balanced class distribution. Evaluation results indicate that Gaussian YOLOv7 improves precision and reduces false positives by 37.5% in the detection stage, albeit with poorer performance in other metrics. Conversely, the implementation of STN results in decreased character recognition accuracy, underscoring its limited effectiveness. Despite these challenges, Gaussian YOLOv7 excels in license plate rectification, achieving a recall of 83.8% and reducing false positives by 50.13% compared to YOLOv7. Moreover, post-processing techniques introduced by our approach further enhance precision by 5.3% and recall by 1%. Overall, our approach offers promising advancements in Indonesian ALPR systems, addressing fundamental challenges and enhancing performance.

Keywords: *detection, license plate, recognition, rectification, YOLO*

1. Introduction

In recent years, Automatic License Plate Recognition (ALPR) systems have gathered significant attention in the field of computer vision [1]. These systems offer the ability to detect and identify vehicle license plates automatically, finding extensive applications in areas such as traffic surveillance, parking management, and electronic traffic law enforcement (ETLE) [2]. In Indonesia, law enforcement officers have initiated a transition toward electronic traffic law enforcement, deploying 270 static cameras and 806 mobile cameras nationwide to monitor traffic violations since September 2022 [3]. Despite this advancement, vehicle identification processes often remain manual, primarily handled by Regional Traffic Management Centre (RTMC) back office personnel. This manual approach proves inefficient and lacks real-time capabilities. Moreover, traditional parking systems in smaller Indonesian cities heavily rely on parking attendants' manual entry of license plate numbers, resulting in time inefficiencies and increased error risks [4]. Therefore, adopting ALPR systems emerges as a promising solution to enhance vehicle identification efficiency across parking and ETLE systems.

ALPR systems generally undergo two stages: license plate detection and character recognition. Implementing these stages poses unique challenges. In the detection stage, license plates may vary in size, shape, and color. At the same time, various environmental factors, such as image orientation, lighting, weather conditions, and locations, add complexity to the task. On the other hand, character recognition involves identifying the diverse character compositions found on license plates, which vary across countries, encompassing different alphabets like Latin, Arabic, and Chinese, each with unique fonts and formatting arrangements. Therefore, developing ALPR systems in Indonesia requires a focus on achieving high-speed and accurate detection and recognition of license plates across various real-world scenarios, adjusted to fit the specific characteristics of Indonesian license plates.

One prominent solution to achieve enhanced license plate detection performance is through deep learning approaches, which have gained traction in object detection due to superior accuracy and speed compared to conventional methods. Recent studies use real-time Convolutional Neural Networks (CNNs) like You Only Look Once (YOLO) and Single-Shot Detector (SSD) for ALPR systems [2, 4, 5, 6, 7]. These models offer significant advantages, particularly in real-time processing, due to their ability to simultaneously detect multiple objects in an image. As for now, YOLOv7, which was introduced in 2022, stands out as one of the best-performing algorithms, offering a balance of speed and detection accuracy that surpasses other real-time object detection algorithms [8]. However, applying deep learning algorithms to ALPR applications faces challenges, particularly in addressing lighting and environmental variations in input images, which can affect license plate detection accuracy. These variations become more pronounced in complex scenarios, leading to increased uncertainty in detection bounding boxes. Consequently, this increases the likelihood of false positives, which remains a critical challenge in most deep learning-based object detection algorithms. Hence, the Gaussian modeling approach introduced in [9] emerges as a viable solution to mitigate false positives in object detection tasks.

Meanwhile, in the character recognition stage, several approaches are employed, utilizing YOLO [2, 6], as well as TesseractOCR and EasyOCR [7], which are libraries for pre-trained character recognition. Although these libraries offer advantages by eliminating the need for training, they are not specifically designed to recognize Indonesian license plates. On the other hand, other deep learning-based approaches in Indonesia also encounter challenges due to imbalanced datasets [4, 6]. The dataset's imbalance results from the collection process being limited to a single region in Indonesia, which may result in a bias towards characters that appear more frequently in the dataset during the recognition stage. Consequently, a vehicle license plate recognition system must be specifically tailored to recognize Indonesian license plates according to their distinctive characteristics and layout to ensure optimal performance in detecting Indonesian vehicle license plates. Moreover, the character recognition stage faces challenges in complex scenarios such as skewed perspectives, uneven lighting, and adverse weather conditions, making it crucial to rectify detected license plates before the recognition stage to improve character recognition accuracy [10].

Given these challenges, we propose an approach for Indonesian license plate detection and character recognition based on YOLOv7, with the following contributions:

1) Introduce a novel dataset of Indonesian vehicles and license plates, capturing diverse and more balanced scenarios across various regions in Java, Indonesia.
2) Develop a Gaussian modeling approach for YOLOv7 to reduce the number of false positives in license plate detection and recognition.
3) Utilize Spatial Transformer Network (STN) to rectify license plate orientation, thereby enhancing ALPR system performance.
4) Propose post-processing methods for YOLOv7 character recognition outputs involving removing duplicates and applying character switching techniques following Indonesian vehicle license plate formats.

The proposed dataset includes more realistic images taken from mobile cameras, which incorporate a wide range of real-world conditions such as varying angles, lighting, and motion, thereby enhancing the robustness of the ALPR system. This approach not only improves the system's performance in dynamic environments but also ensures broader applicability, such as for Electronic Traffic Law Enforcement (ETLE). Additionally, our system is designed to detect license plates on multiple vehicles simultaneously, leveraging the capability of YOLOv7 for multi-object detection. Furthermore, we will evaluate several image quality assessment (IQA) methods to improve the ALPR system's accuracy in filtering out blurry license plate images before the character recognition stage. Through this combination, we aim to achieve better license plate detection and character recognition performance, particularly in complex environmental conditions in Indonesia.

2. Related Works

In this paper section, we outline the framework of automatic license plate recognition (ALPR), which encompasses the automated processes of detecting, rectifying, and recognizing license plates within images. The typical ALPR system comprises two main subtasks: license plate detection, which aims to identify and localize license plates within images, and license plate recognition, which aims to recognize the characters within the license plate. Additionally, license plate rectification serves as an additional subtask aimed at enhancing character recognition performance. In this section, we will present an overview of the key tasks within ALPR systems, including license plate detection, rectification, and recognition methods.

2.1. License Plate Detection

Several classical image processing approaches, such as edge processing using histograms and short object elimination, as well as edge detection using Sobel edge detector, have been employed [11, 12]. These classical approaches have the advantage of localizing plates without needing prior training. However, their limitations lie in their ability to handle complex environmental conditions, images with multiple vehicle license plates, and their inability to operate in real-time.

Recent advancements in ALPR systems have primarily revolved around the utilization of CNN-based real-time object detection algorithms, notably You Only Look Once (YOLO) [13] and Single-Shot Detector (SSD) [14]. Several studies have explored the application of YOLO-based algorithms across diverse case studies. In Brazil, the FRV/LPD-NET [2], based on YOLO, was integrated into an ALPR system to detect vehicles and license plates sequentially. Similarly, in Bangladesh, YOLOv3 [15] was employed for license plate detection on Bangladeshi license plates, albeit encountering reduced detection speed in the presence of multiple vehicles in a single frame [5]. In Indonesia, investigations employing YOLOv2 [16] and YOLOv4 [17] for license plate detection have been conducted, revealing promising detection capabilities but limited by dataset complexity, primarily collected from parking areas, thus restricting their applicability to specific scenarios such as parking systems [4, 6].

Moreover, studies leveraging SSD with MobileNet V1 as the base network for detection achieved satisfactory license plate detection accuracy, albeit facing challenges related to hyperparameter adaptability to dataset changes, necessitating iterative optimization [7]. Additionally, another method, such as ALPRNet, was developed to detect mixed-style license plates from various regions, employing two Fully Convolutional One-stage Object detectors (FCOS detectors) constructed on ResNet50 with Feature Pyramid Network (FPN) [18]. While ALPRNet offers simultaneous detection and character recognition capabilities, its evaluation was confined to datasets with less complex environmental conditions, underscoring the need for further performance assessment.

2.2. License Plate Rectification

Classical image processing techniques, such as homography transformation [19, 20], have been utilized in vehicle license plate rectification systems. These techniques utilize a homography matrix to adjust camera perspective distortion by transforming the distorted image onto a flat plane, thereby rectifying the license plate image. Despite their significance in achieving accurate character extraction and recognition, these classical methods are hampered by their inability to support real-time operations, compounded by their susceptibility to variations in lighting conditions and image contrast.

Several deep learning approaches, including WPOD-NET [21] and the perspective rectification network [22], have been devised to rectify vehicle license plates. WPOD-NET integrates YOLO architecture for localization and Spatial Transformer Network (STN) [23] for rectification, allowing for the acquisition and application of spatial transformations such as translation, rotation, and scaling to the license plates. Despite their effectiveness in improving character recognition accuracy under complex environmental conditions and real-time processing capabilities, deep learning methods encounter challenges with extreme plate distortion, limited dataset variation, and resource-intensive training.

2.3. License Plate Recognition

License plate recognition using classical image processing, such as template matching [11], has been employed to automatically recognize vehicle license plate characters. These approaches offer the benefit of localizing and recognizing characters requiring extensive training datasets. However, they suffer from slow processing time and lack real-time functionality. Additionally, their effectiveness is constrained to particular image orientations, reducing their practical application efficiency.

Several studies have employed YOLO-based architectures as a deep learning approach for license plate recognition. In a Brazilian case study, character segmentation and recognition are performed using YOLO-VOC [16] architecture and additional character sorting techniques. This character sorting technique is explicitly adjusted for Brazilian license plates to enhance recognition accuracy. In another study, YOLOv3 [15] and YOLOv5 [24] were employed for character recognition on Bangladeshi and Indonesian license plates, respectively. These two case studies encountered challenges due to dataset collection being limited to a single city, resulting in an imbalanced dataset issue that adversely affected the character recognition performance for license plates from another region in the same country.

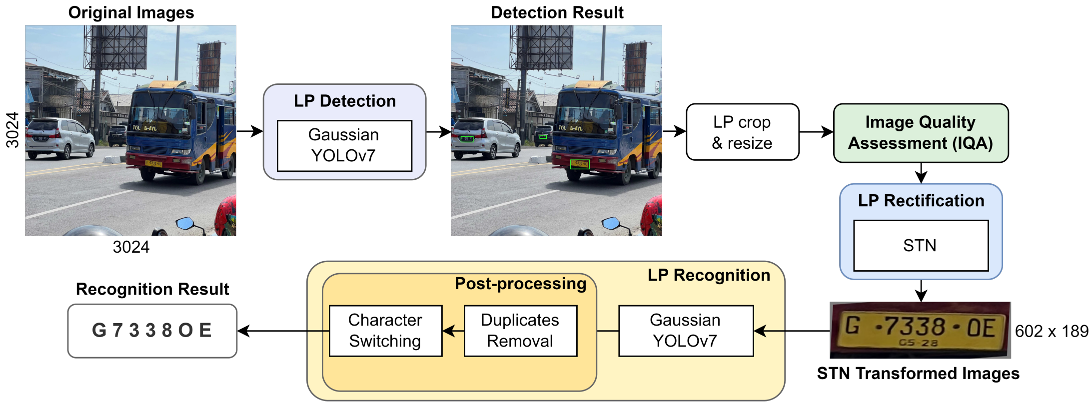

Figure 1. Proposed system architecture.

Additionally, several researchers have utilized pre-trained libraries such as Tesseract OCR and Easy OCR for character recognition [7]. These libraries present certain advantages by eliminating the need for training and exhibiting the capability to recognize various types of characters. However, they are not specialized for recognizing license plate characters and their specific fonts, which may result in misclassifications of characters sharing similar shapes.

3. Proposed Method

The proposed framework for Indonesian license plate detection and recognition integrates Gaussian YOLOv7 for the detection of multiple license plates and subsequent character recognition, along with Spatial Transformer Network (STN) for license plate rectification prior to the character recognition stage, as depicted in Fig. 1. Gaussian YOLOv7, which is based on the state-of-the-art YOLOv7 architecture [8], incorporates Gaussian modeling [9] within the detection layer to estimate variance and filter out uncertainty during object detection. Additionally, STN is utilized to learn and apply spatial transformations on license plates for the license plate rectification stage. In this system, Image Quality Assessment (IQA) methods are further employed to filter out blurry license plate images before the rectification stage. This critical step ensures that excessively blurry images, to the extent that their characters cannot be reliably identified even by human observers, are excluded from further processing. Post-processing steps, including duplicate removal and character switching, are then applied to the output of the license license plate character recognition system. Notably, our system is designed to detect license plates on multiple vehicles simultaneously, leveraging the capability of YOLOv7 for multi-object detection. This section will provide a detailed explanation of the proposed method.

3.1. Gaussian YOLOv7

The proposed Gaussian YOLOv7 is constructed using the YOLOv7 architecture [8] as its foundation, as illustrated in Fig. 2. The YOLOv7 structure comprises an input layer, backbone, feature pyramid network, and detection head. The detection process with YOLOv7 involves several stages. Initially, RGB images of size 640×640 are fed into the YOLOv7 backbone for feature extraction, producing three feature maps sized 80×80×512, 40×40×1024, and 20×20×1024. These feature maps are then merged within the feature pyramid network to combine feature information across different scales. Subsequently, object detection occurs at different grid sizes (80×80, 40×40, and 20×20) after applying of the RepConv (reparameterized-convolution) module. This RepConv module is divided into training and inference phases, employing branch structures to enhance feature extraction during training and streamline detection during inference. Finally, the detection process leverages implicit knowledge introduced in YOLOR, integrating implicit and explicit knowledge to enhance overall performance across various tasks.

In YOLOv7 architecture, the CBS module comprises convolution, batch normalization (BN), and sigmoid linear unit (SiLU) operations with varying kernel ($\boldsymbol{k}$) and stride ($\boldsymbol{s}$) values. Convolution with $\boldsymbol{k} = \mathbf{1}$ and $\boldsymbol{s} = \mathbf{1}$ is utilized for channel adjustment, $\boldsymbol{k} = \mathbf{3}$ and $\boldsymbol{s} = \mathbf{1}$ for feature extraction, and $\boldsymbol{k} = \mathbf{3}$ and $\boldsymbol{s} = \mathbf{2}$ for downsampling. Besides the CBS module, YOLOv7 incorporates the efficient layer aggregation network (ELAN) module, which controls the shortest and longest gradient paths to continuously enhance network learning capabilities and robustness. Additionally, the YOLOv7 detection head employs the spatial pyramid pooling and cross-stage partial connection (SPPCSPC) module to obtain different sensitivity

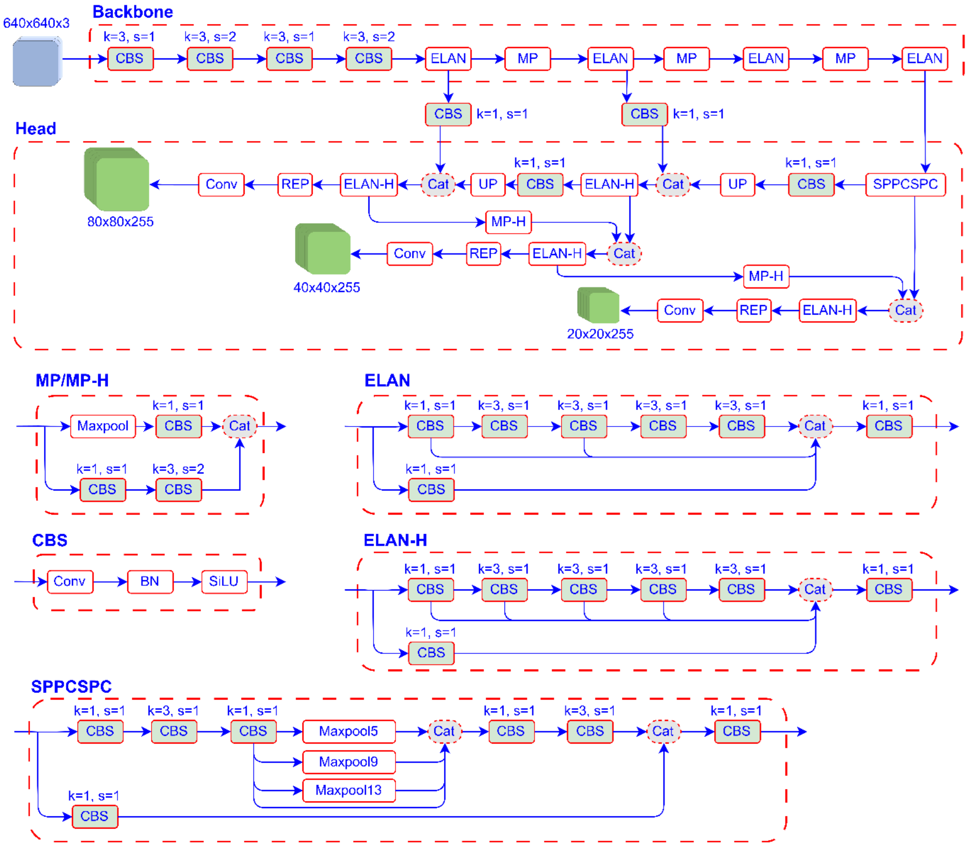

Figure 2. YOLOv7 architecture.

fields through max pooling, improving the differentiation of large and small objects.

The Gaussian approach [9] is based on the principles of the Gaussian distribution. This distribution is used to model uncertainty, which is prevalent in object detection tasks due to factors such as image noise and ambiguous object boundaries. Integrating Gaussian distributions into object detection frameworks enables effective modeling and management of uncertainty in object localization. This approach is achieved by incoporating Gaussian parameters, such as mean and variance. The mean represents the central location of the object within the bounding box, while the variance indicates the degree of uncertainty associated with the predicted object location. Consequently, implementing this approach within the YOLOv7 framework allows for the filtering of uncertainty and the prevention of false positive predictions during vehicle license plate detection and character recognition.

In YOLO-based object detection, predictions within the detection module typically include bounding box coordinates $(\boldsymbol{t_x, t_y, t_w, t_h})$, objectness score $(\boldsymbol{P_{obj}})$, and classes score $(\boldsymbol{P_{class}})$. However, unlike objectness and classes score, which range from 0 to 1 to determine object detection, bounding box coordinates provide only coordinate values without indicating detection confidence. Additionally, the objectness score lacks information about bounding box uncertainty. To address this limitation, the Gaussian approach is performed by modifying the detection module by modeling each bounding box coordinate as mean $(\boldsymbol{\hat{\mu}})$ and variance $(\boldsymbol{\hat{\Sigma}})$, thereby enriching predictions by encompassing the variance of the bounding box.

In the inference stage of YOLOv7, the mean and variance values of the bounding box are further transformed using equations (1) and (2):

$$\begin{aligned} \mu_{t_x} &= 2\sigma(\hat{\mu}_{t_x}) - 0.5 + c_x, \\ \mu_{t_y} &= 2\sigma\left(\hat{\mu}_{t_y}\right) - 0.5 + c_y, \\ \mu_{t_w} &= p_w\left(2\sigma(\hat{\mu}_{t_w})\right)^2, \\ \mu_{t_h} &= p_w\left(2\sigma(\hat{\mu}_{t_h})\right)^2 \end{aligned} \tag{1}$$

$$\begin{aligned} \Sigma_{t_x} &= \sigma(\hat{\Sigma}_{t_x}), \quad & \Sigma_{t_y} &= \sigma\left(\hat{\Sigma}_{t_y}\right), \\ \Sigma_{t_w} &= \sigma(\hat{\Sigma}_{t_w}), \quad & \Sigma_{t_h} &= \sigma(\hat{\Sigma}_{t_h}) \end{aligned} \tag{2}$$

where $\boldsymbol{c_x}$ and $\boldsymbol{c_y}$ represent the x and y grid coordinates, $\boldsymbol{p_w}$ and $\boldsymbol{p_h}$ denote the width and height of the anchor, and $\boldsymbol{\sigma(x)}$ is the sigmoid function producing output between 0 and 1. Through these equations, the YOLOv7 algorithm transforms the bounding box center points $(\boldsymbol{\hat{\mu}_{t_x}, \hat{\mu}_{t_y}})$ from grid cell coordinates into bounding box coordinates in the image, while also transforming the width and height of the bounding box through integration with anchor dimensions. Additionally, the objectness and class scores are also transformed to produce values within the range of 0-1.

Based on the variance values obtained, the uncertainty in localization can be used to determine detection criteria by adjusting the objectness score based on the average variance as shown in Equation (3):

$$P_{obj} = P_{obj} \cdot (1 - \bar{\Sigma}) \tag{3}$$

where $\bar{\Sigma}$ represents the average variance for each predicted bounding box coordinate. This equation reduces the objectness score as the average uncertainty of the predicted bounding box increases. Consequently, the detected bounding boxes can be filtered out if the resulting detection criterion from the multiplication of objectness and each class score fails to meet a predefined threshold. This approach prioritizes more reliable predictions, enhances object detection accuracy, and mitigates the impact of noisy data during training, thereby improving the performance of the YOLOv7 algorithm.

Furthermore, to incorporate the mean and variance values, the Gaussian approach modifies the loss function of the bounding boxes, transitioning from the CloU loss to the Negative Log Likelihood (NLL) loss. The NLL loss measures the disparity between predicted probabilities and ground truth labels. This metric computes the negative logarithm of the predicted probabilities for the correct class and penalizes deviations from the correct classification. The NLL loss for each bounding box is formulated using Equation (4):

$$L_{NLL_x} = -\sum_{i=i}^{W}\sum_{j=1}^{H}\sum_{k=1}^{K} \gamma_{ijk} \log\left(PDF(x_{ijk}) + \varepsilon\right) \tag{4}$$

where L_{NLL_x} represents the NLL loss of the t_x coordinates, W and H denote the number of grids for each width and height parameter, and K represents the number of anchors. $PDF(x_{ijk})$ is the probability density function of the Gaussian distribution. Meanwhile, γ_{ijk} is calculated based on the ratio of width and height of the target bounding box, and ε is a value added to achieve numerical stability for the logarithmic function, typically set to 10^{-9}. In computing the NLL loss, the probability density function of the Gaussian distribution is formulated with Equation (5):

$$PDF(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} \tag{5}$$

where μ represents the mean or expectation of the distribution, while σ denotes the standard deviation, with the variance expressed as σ^2. The utilization of NLL loss can enhance the model's robustness to noisy data by penalizing the loss attributed to data uncertainty during training, thus contributing to improved performance.

3.2. Spatial Transformer Network

For license plate rectification, we employed Spatial Transformer Network (STN) [23], which is a type of neural network architecture capable of learning to perform spatial transformations on input images, such as translation, rotation, and scaling, to enhance the overall system performance. STN comprises three primary components: the localization network, grid generator, and sampler. The localization network predicts the desired spatial transformation parameters through convolutional layers, max-pooling operations, and ReLU activation functions. A regressor then processes these parameters to compute the affine transformation matrix (θ), which is responsible for the spatial manipulation of the input image. Subsequently, the grid generator generates a grid of spatial coordinates, mapping original image pixels to their transformed positions based on the θ matrix. Finally, the grid sampler applies 2D transformations to the input image, resulting in an output image with rectified orientation at a resolution of 94×24.

To preserve image resolution quality in the rectification process, modifications were implemented in the STN algorithm through several stages. Initially, all vehicle license plate images were resized to dimensions of 602×189. Subsequently, the STN's localization network received input image feature maps adjusted to the STN input resolution of 94×24, yielding the transformation matrix θ as an output. Lastly, the grid was generated based on the θ matrix but at a resolution of 602×189. These adaptations enable the license plate rectification system to uphold optimal resolution in the license plate images, ensuring their suitability for subsequent stages, particularly for the license plate character recognition stage.

3.3. Image Quality Assessment

Image Quality Assessment (IQA) is a quantitative process aimed at measuring the quality of images to predict how well the measured images can be perceived by human observers [25]. It comprises two types: Full-Reference (FR) IQA and No-Reference (NR) IQA, each employing distinct methods for evaluating image quality. Unlike FR IQA, which requires undistorted reference images for comparison, NR IQA does not rely on reference images. This characteristic makes NR IQA suitable for scenarios where obtaining reference images is difficult or impractical, such as in License Plate Image Quality Assessment (LP IQA) tasks. In this study, we employed two NR IQA methods – Laplacian Variance and BRISQUE – that are suitable for integration into our Python-based system.

Laplacian Variance is an NR IQA method employed to measure the sharpness or contrast changes within an image. This model computes the variance of the image after applying a Laplacian filter [26]. In images with poor clarity, where edges are poorly defined, the Laplacian filter tends to produce nearly uniform values, resulting in lower variance. Consequently, higher variance values indicate greater detail or sharpness within the image.

In the Laplacian Variance method, the Laplacian filter serves as a second derivative operator formulated by Equation (6). Meanwhile, the variance of the Laplacian is expressed in Equation (7):

$$\boldsymbol{L} = \frac{\boldsymbol{1}}{\boldsymbol{6}}\begin{bmatrix} \boldsymbol{0} & \boldsymbol{-1} & \boldsymbol{0} \\ \boldsymbol{-1} & \boldsymbol{4} & \boldsymbol{-1} \\ \boldsymbol{0} & \boldsymbol{-1} & \boldsymbol{0} \end{bmatrix} \tag{6}$$

$$\boldsymbol{LV} = \sum_{m}^{M}\sum_{n}^{N}(|\boldsymbol{L}(\boldsymbol{m},\boldsymbol{n})| - \bar{\boldsymbol{L}})^{2} \tag{7}$$

where $\boldsymbol{M}$ and $\boldsymbol{N}$ represent the dimensions of the input image, $\boldsymbol{L}(\boldsymbol{m},\boldsymbol{n})$ denotes the convolution of the input image with the Laplacian filter $\boldsymbol{L}$, and $\bar{\boldsymbol{L}}$ signifies the mean of absolute values of $\boldsymbol{L}(\boldsymbol{m},\boldsymbol{n})$, formulated by Equation (8).

$$\bar{\boldsymbol{L}} = \frac{\boldsymbol{1}}{\boldsymbol{N}\cdot\boldsymbol{M}}\sum_{m}^{M}\sum_{n}^{N}|\boldsymbol{L}(\boldsymbol{m},\boldsymbol{n})| \tag{8}$$

Meanwhile, the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [27] is an NR IQA method known for its ability to assess image quality across various types of distortions in the spatial domain. BRISQUE operates by extracting statistical features from images, including mean, variance, skewness, and kurtosis of locally normalized luminance coefficients. These features are then used to establish relationships with perceived image quality, eliminating the need for reference images. Initially, BRISQUE extracts a set of natural scene statistics (NSS) features from the input image, capturing statistical regularities found in natural images from locally normalized image patches. These features are then input into a support vector machine (SVM) regressor to predict the quality score of an image based on its NSS features, ultimately providing an output indicative of the perceived quality of the input image.

In our methodology, we manually adjust the thresholds for IQA methods to establish the most suitable threshold for our task. This adjustment is necessary to filter out blurred LP images with unrecognizable characters. By aligning these thresholds with human perceptions of image quality, we ensure that only images with sufficiently clear and distinguishable license plate characters proceed for further processing. This step is crucial because it enhances the effectiveness of our system, particularly in scenarios with variable image clarity due to factors like motion blur or camera distance. Our emphasis on efficiency and real-time performance necessitates maintaining fixed thresholds, streamlining the image quality assessment process, and minimizing complexities to ensure optimal system performance.

Furthermore, we have modified our approach to prioritize the Region of Interest (ROI) within the IQA process, specifically targeting the central 60% section of license plate images. This adjustment is aimed at enhancing the precision of the IQA procedure, given that license plate images commonly encompass a surrounding frame. This frame is particularly noticeable in plates with contrasting colors to the vehicle colors. The presence of these frames within the images has the potential to influence IQA evaluations. Thus, by restricting the assessment to the ROI, we ensure a more precise evaluation focused exclusively on the license plate itself, thereby minimizing the likelihood of interference from surrounding frames.

3.4. Post-processing

As previously discussed, certain characters exhibit visual similarities, leading to potential misclassification issues such as distinguishing between the numeral "0" and the letter "O," which can significantly reduce character recognition accuracy. Hence, implementing post-processing steps becomes necessary to mitigate character recognition errors. Additionally, license plates in each country possess unique characteristics that distinguish them from those of other countries. Consequently, post-processing steps must be customized to leverage the specific characteristics of license plates used in a particular country. These

Figure 3. Examples of images captured in the proposed dataset.

customized post-processing techniques are essential to refine the character recognition process, ensuring accurate identification of license plate characters and enhancing the system's overall performance.

Indonesian license plates exhibit distinct characteristics, structured into a format comprising area/registration code, registration number, and supplementary letter series [28]. The area code denotes the vehicle's registration region, while the registration code reflects the vehicle's designated usage, both represented by 1-2 letter characters. Additionally, the registration number encompasses numerical characters ranging from one to four digits. Furthermore, the letter series on Indonesian license plates varies from zero to three letter characters Indonesian license plates also include a second row indicating the expiration date, which must be excluded from the license plate recognition system. Moreover, Indonesian license plates are available in five primary color variations, predominantly assigned to private vehicles, featuring either a white background with black text (post-June 2022) or a black background with white text (pre-June 2022).

To refine the detected bounding boxes of Indonesian license plates, several post-processing steps are performed. Initially, bounding boxes with center coordinates located within the lower 40% of the image, corresponding to the expiration date area, are removed. Subsequently, a sorting process is applied to arrange the detection results from left to right. This initial step is essential to ensure the character recognition results align with the desired Region of Interest (ROI). Furthermore, duplicates may arise in the detection output of YOLOv7. Thus, duplicate removal is performed considering the layout of the license plate. This step started by determining the first index ($\boldsymbol{idx_{first}}$), separating the area/registration code and registration number, and the second index ($\boldsymbol{idx_{last}}$), separating the registration number and letter series. Given the variable character lengths of Indonesian license plates, index determination involves calculating the distance between the x-coordinates of each bounding box. Bounding boxes exceeding a predefined threshold distance from the preceding box are assigned as the index.

Table 1. Swap rules for area/registration code and letter series.

Original Character	Changed to
0	O
1	I
2	Z
3	E
4	A
5	S
6	G
7	Z
8	B
9	B

Table 2. Swap rules for registration number.

Original Character	Changed to
C, D, O, Q, U	0
I, J, L	1
Z	2
E	3
A	4
S	5
G	6
T	7
B, H, P, R	8
-	9

Upon identifying the index that separates the area/registration code, registration number, and letter series, the removal of duplicates can be prioritized based on character classes. Specifically, the area/registration code and letter series must comprise letters, while the registration number must consist of digits. When duplicate bounding boxes are detected, with one containing a digit and the other a letter, the removal steps perform filtering to ensure that the characters for the area/registration code, which are positioned before $\boldsymbol{idx_{first}}$, and the characters for the letter series positioned after $\boldsymbol{idx_{last}}$, are letters (A-Z). Meanwhile, the characters for the registration number situated between $\boldsymbol{idx_{first}}$ and $\boldsymbol{idx_{last}}$ are digits (0-9). Moreover, if both characters belong to the same type (letters or digits), the bounding box with the highest confidence is selected. Additionally, the character switching technique is employed if there are still characters with mismatched classes, such as digits detected in the area/registration code or letter series or letters found in the registration number. This technique involves swapping digits with letters or vice versa,

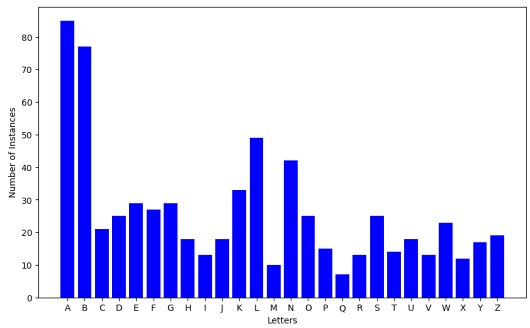

Figure 4. Letters distribution in the testing dataset.

according to predefined conversion tables, as shown in Table 1 and Table 2, which specify the characters to be interchanged.

4. Experiments

4.1. Dataset

The dataset presented in this study is specifically designed to reflect real-world scenarios encountered in ALPR systems accurately. To achieve this, we collected images under diverse conditions while ensuring a more balanced representation of the vehicle license plate registration area. The inclusion of images under varied conditions aims to enhance the robustness and effectiveness of the ALPR system in accurately identifying license plates. Concurrently, using a balanced dataset aims to refine character recognition accuracy on license plates. The proposed dataset comprises 1000 images depicting vehicles and their respective license plates, as shown in Figure 3. These images were taken from vehicles navigating through regular urban traffic and intercity highways. We captured these images using the rear camera of a mobile phone, which has a resolution of 9 megapixels and follows a 1:1 aspect ratio.

To ensure diversity in our dataset, we gathered images depicting scenarios involving varying numbers of vehicles, backgrounds, lighting, and weather conditions. The dataset includes images captured under bright, dark, and rainy conditions, depending on the lighting and weather at the capture time. Additionally, to create a balanced dataset, we collected images from various locations across Java, including Jakarta, Tegal, Surabaya, Malang, and along the Trans-Java toll road. By incorporating these variations, we aim to enhance the dataset's representation of real-world situations and improve the ALPR system's effectiveness under different conditions.

Our dataset was then partitioned as follows: 80% for training, 10% for validation, and 10% for testing. We then proceeded to annotate the dataset following established practices documented in related literature, ensuring consistency and compatibility with recognized approaches. For the license plate detection stage, the dataset was annotated with a single class, namely "LP_Indonesia." Through this partitioning, we obtained a ground truth consisting of 1857 vehicle license plates in the training dataset, 229 in the validation dataset, and 244 in the testing dataset. The cropped license plates, which were based on the ground truth labels from the detection dataset, were utilized for both the image quality assessment and character recognition stages. During the recognition stage, any blurry images were filtered out from the dataset. As a result, the dataset consists of 212 images, each annotated with 36 classes, encompassing 10 digit classes and 26 letter classes. The distribution of letter classes within the testing dataset is shown in Figure 4.

As depicted in Figure 4, although there is still some class imbalance within the dataset, the distribution of classes is not dominated by a single class. This distribution indicates a more balanced representation across various classes, although not all classes are represented equally. This distribution of letter classes corresponds to the area/registration codes prevalent in the cities where the dataset was collected. For instance, the area/registration code "B" predominates in Jakarta, "G" in Tegal, "L" in Surabaya, and "N" in Malang. This geographic diversity contributes to the balanced distribution of classes within the dataset, enhancing its representativeness of real-world scenarios.

4.2. Experimental Settings

This research implemented a modified version of the YOLOv7 algorithm, namely Gaussian YOLOv7, to address false detections in license plate detection and recognition tasks. Gaussian YOLOv7 was developed using the PyTorch library, with modifications to the detection layer and loss function compared to the original YOLOv7. To evaluate the performance of the Gaussian approach, our proposed method was compared with the original YOLOv7 as the baseline. During the detection stage, both models were trained on the training dataset with a batch size of 16 and for 100 epochs. Their performance was evaluated using metrics such as precision, recall, and F1-score, alongside numerical evaluations based on true positives (TP), true negatives (TN), and false negatives (FN).

For the recognition stage, the performance of Gaussian YOLOv7 was further compared with EasyOCR, a pre-trained library for character recognition. Since annotation of the training dataset had not been performed, pre-training of both YOLOv7 and Gaussian YOLOv7 was conducted for 50 epochs on a character dataset comprising 200 training images and 20 validation images for each character, utilizing a batch size of 16. Subsequently, training was carried out on a dataset of vehicle license plates from Indramayu, consisting of 235 training images and 75 validation images, for 100 epochs with a batch size of 16. This dataset comprised images of vehicle license plates with black text on a white background, and the training was extended to 100 epochs. To assess the performance of the character recognition methods, a testing dataset containing 148 images of vehicle license plates with black text was utilized. The evaluation was conducted using precision, recall, and F1-score metrics, along with numerical evaluations of TP, TN, and FN. Moreover, the influence of post-processing techniques and STN on character recognition performance was assessed. For STN, pre-trained weights collected from joint training of STN with LPRNet on the CCPD dataset were utilized.

Additionally, we evaluated the effectiveness of integrating IQA methods to filter out blurred detected license plates before the recognition stage. Therefore, we defined blur thresholds of Laplacian Variance and BRISQUE methods. The blur threshold for Laplacian Variance was set at 15.0, with values lower than the threshold indicating higher levels of blurriness. On the other hand, the blur threshold for the BRISQUE method was set at 70.0. Images with BRISQUE scores lower than the threshold were deemed acceptable for character recognition purposes.

5. Results & Analysis

In this section, we will evaluate the proposed system through several steps, including license plate detection, image quality assessment, license plate rectification and license plate recognition.

5.1. License Plate Detection

In the detection stage, the performance of the vehicle license plate detection system using Gaussian YOLOv7 was compared with YOLOv7 as the baseline. The evaluation was conducted on a testing dataset consisting of 100 images, with a total of 244 Ground Truth vehicle license plates. Evaluation metrics comparing the performance of both methods are presented in Table 3, while numerical evaluations can be found in Table 4.

Based on Table 3, it is evident that Gaussian YOLOv7 exhibits better performance in terms of precision, indicating more accurate detection with a lower rate of false positives. This observation is further supported by Table 4, where detection using Gaussian YOLOv7 reduces the number of false positives by 37.5% compared to YOLOv7. However, further evaluation reveals that despite the improvement in precision and reduction in false positives achieved by Gaussian YOLOv7, several other evaluation metrics still favor YOLOv7. Numerical evaluation indicates that Gaussian YOLOv7 has yet to succeed in increasing true positives, experiencing a decrease of 11.67% compared to YOLOv7. Additionally, YOLOv7 demonstrates significantly better recall, with only 4 false negatives compared to 32 for Gaussian YOLOv7. This comparison suggests that although Gaussian YOLOv7 contributes to reducing false positives, YOLOv7 still provides superior overall performance.

Table 3. Evaluation of detection results on testing dataset.

Model	P	R	F1
YOLOv7	0.909	**0.984**	**0.945**
Gaussian YOLOv7 (ours)	**0.934**	0.869	0.900

Table 4. Numerical evaluation of detection result.

Numerical evaluation	YOLOv7	Gaussian YOLOv7	Variation Rate (%)
TP	240	212	-11.67
FP	24	15	-37.5
FN	4	32	+700

These findings indicate that for single-class detection, the reduction in false positives with Gaussian YOLOv7 is accompanied by an increase in false negatives. This is attributed to Gaussian YOLOv7 imposing a greater penalty on uncertainty, leading to a decrease in the number of detections. This underscores the nuanced trade-off between precision and recall in the context of Gaussian YOLOv7's performance.

5.2. Image Quality Assessment

Before the recognition stage, the quality of vehicle license plate images was assessed using two Image Quality Assessment (IQA) methods: Laplacian variance and BRISQUE. This stage aimed to provide a comprehensive understanding of the effectiveness of IQA methods in filtering out blurry images within license plate recognition applications. These methods were employed to evaluate 244 license plate images from the detection's testing dataset, which comprised 200 sharp images and 44 blurry images. Quantitative evaluation using accuracy metrics was then conducted to assess the performance of the IQA methods. Additionally, the average processing time for each method was compared. The evaluation results of the IQA methods are presented in Table 5.

Table 5. IQA methods evaluation.

Methods	Accuracy	Avg. time
Laplacian Variance	0.885	**2.5 ms**
BRISQUE	**0.901**	24.7 ms

The results indicate that BRISQUE outperforms Laplacian Variance in distinguishing between sharp and blurry images. Numerical evaluation reveals that BRISQUE yields a higher number of accurate detections for both sharp and blurry images compared to Laplacian Variance. However, BRISQUE incurs significantly higher computational overhead, with image processing times nearly 10 times longer than those of Laplacian Variance. Consequently, it can be inferred that for real-time applications, such as ALPR systems, Laplacian Variance is more suitable due to its relatively good accuracy combined with faster processing speed.

5.3. License Plate Recognition

In the license plate recognition stage, we assessed the performance of Gaussian YOLOv7 compared to YOLOv7 and EasyOCR as the baseline. EasyOCR is a deep learning-based algorithm trained to support over 80 languages capable of recognizing characters from images. Therefore, we utilized pre-trained libraries like EasyOCR as a baseline for character recognition to evaluate the performance of general character recognition algorithms against our method, which was specifically trained to recognize characters from Indonesian vehicle license plates. Meanwhile, Gaussian YOLOv7 was evaluated against YOLOv7 to examine the performance of the Gaussian approach in a multi-class object detection scenario. Additionally, the performance of STN in rectifying vehicle license plates was evaluated based on character recognition results from STN-rectified images. This method was compared with character recognition from original resized images. The evaluation results are presented in Table 6.

Based on the findings presented in Table 6, the quantitative evaluation indicates that the implementation of STN in the license plate rectification stage resulted in a decrease in character recognition performance, with up to a 6.3% drop in recall values. This decline can be attributed to the suboptimal performance of the STN algorithm in rectifying license plate images, primarily due to its training on a different dataset. This observation is further supported by Figure 5, which showcases several examples of rectified license plate images.

From Figure 5, it is evident that while STN is capable of rectifying license plate images, the improvements achieved are not substantial. Additionally, STN often introduces distortions to the orientation of license plate images that were initially straight, thereby significantly compromising character recognition performance. Given that most license plate images have only minor orientation distortions, incorporating STN into the ALPR system tends to increase the overall training cost and inference time without significantly improving performance.

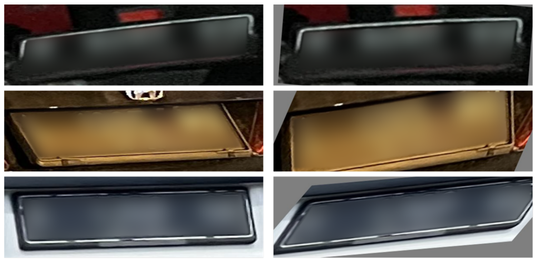

Figure 5. Qualitative results of the rectification stage: (a) Original images, (b) Images rectified using STN.

Regarding the recognition models, Gaussian YOLOv7 demonstrated superior performance compared to YOLOv7 and EasyOCR, achieving a recall of 83.8% for non-rectified images. This improved performance of Gaussian YOLOv7 over EasyOCR could be attributed to EasyOCR's lack of a specific design for detecting Indonesian license plate characters, leading to false detections outside the license plate character set. Conversely, both YOLOv7 and Gaussian YOLOv7, which were explicitly trained for license plate character detection, did not exhibit such false detections.

These findings also highlight that employing the Gaussian approach for multi-class object detection tasks, such as license plate recognition, results in superior performance compared to the YOLOv7 algorithm. This superiority stems from Gaussian YOLOv7's ability to filter out uncertainty that could lead to false detections. Consequently, Gaussian YOLOv7 minimizes the occurrence of false positives during the character recognition stage by up to 50.13% for non-rectified images compared to YOLOv7. Notably, in multi-class object detection scenarios, false positives in one class could lead to false negatives in other classes. As such, Gaussian YOLOv7 also demonstrates a decrease in the number of false negatives, resulting in higher precision and recall compared to YOLOv7.

Furthermore, the post-processing steps we proposed enhance the performance of both YOLOv7 and Gaussian YOLOv7 across all evaluation metrics. This improvement arises from the fact that both YOLOv7 and Gaussian YOLOv7 exhibit multiple detections on a single character. By employing post-processing techniques, we eliminate duplicate detections, leading to a

Table 6. Evaluation of recognition results on testing dataset.

Dataset	Model	TP	FP	FN	P	R	F1
Original	EasyOCR	856	176	228	0.83	0.79	0.809
	YOLOv7	811	359	309	0.693	0.724	0.708
	YOLOv7 + post-processing (ours)	830	153	254	0.844	0.766	0.803
	Gaussian YOLOv7 (ours)	**929**	140	179	0.869	0.838	0.853
	Gaussian YOLOv7 + post-processing (ours)	919	**78**	**165**	**0.922**	**0.848**	**0.883**
STN	EasyOCR	844	181	240	0.823	0.779	0.8
	YOLOv7	744	437	365	0.63	0.671	0.65
	YOLOv7 + post-processing (ours)	762	214	322	0.781	0.703	0.740
	Gaussian YOLOv7 (ours)	**890**	167	225	0.842	0.798	0.82
	Gaussian YOLOv7 + post-processing (ours)	872	**85**	**212**	**0.911**	**0.804**	**0.854**

significant reduction in the number of false positives. However, numerical evaluations indicate a decrease in the number of true positives. This is due to the YOLOv7 and Gaussian YOLOv7 algorithms sometimes detecting one character with two bounding boxes of the same class, giving the impression of detecting more correct characters. Furthermore, the character switching technique in post-processing reduces the number of false detections by swapping characters into numbers or vice versa, depending on the character's position. Thus, this post-processing technique improves performance, with a precision increase of 5.3% and a recall improvement of 1% for non-rectified images for the Gaussian YOLOv7 model we proposed.

6. Conclusion

In this paper, we proposed Gaussian YOLOv7 as a pivotal component for refining the processes of license plate detection and character recognition within ALPR systems. By deploying Gaussian YOLOv7, we aimed to mitigate the occurrences of false detections encountered in conventional YOLOv7 models. Additionally, we integrated several IQA techniques to filter out blurred images before the character recognition stage. Moreover, we conducted an in-depth evaluation of STN to rectify license plate distortions, aiming to ascertain its impact on character recognition performance. This comprehensive approach sought to advance the capabilities of ALPR systems, contributing to more accurate and reliable license plate recognition in real-world applications.

Moreover, to facilitate research and development in the field of ALPR systems, we introduced an Indonesian vehicle and license plate dataset specifically designed to replicate diverse real-world scenarios. This dataset was collected under varying lighting, weather, and environmental conditions across multiple locations in Java to ensure a more balanced representation of vehicle license plate registration areas. The dataset's images were captured from vehicles navigating through regular urban traffic and intercity highways, providing researchers and developers with a comprehensive resource for enhancing ALPR technology for practical applications.

Evaluation results show that Gaussian YOLOv7 demonstrated improved precision and a notable reduction in false positives by 37.5% compared to YOLOv7 in the detection stage. However, it struggled to increase true positives and exhibited inferior recall performance. This balance between precision and recall highlighted the nuanced nature of Gaussian YOLOv7's performance in single-class detection, where reducing false positives led to more false negatives. Moreover, when assessing IQA methods for the detected license plates, Laplacian Variance emerged as more suitable due to its better balance between accuracy and processing time, compared to BRISQUE, which, although more accurate, took significantly longer to process, making it less practical for real-time ALPR applications.

Additionally, implementing the STN in the license plate rectification stage led to a noticeable decrease of up to 6.3% in character recognition accuracy. This decline was observed due to STN's tendency to distort initially straight license plates, indicating its limited effectiveness in improving ALPR systems. Regarding model performance, Gaussian YOLOv7 outperformed YOLOv7 and EasyOCR in license plate recognition, achieving a recall of 83.8% for non-rectified images. This superiority was attributed to Gaussian YOLOv7's tailored design for Indonesian license plate characters, effectively reducing false detections. Furthermore, Gaussian YOLOv7 significantly reduced false positives by up to 50.13% compared to YOLOv7, demonstrating its ability to filter out uncertainty, especially in multi-class detection scenarios. Additionally, post-processing techniques notably refined the performance of Gaussian YOLOv7, particularly in reducing false positives and enhancing precision, showing a precision increase of 5.3% and a recall improvement of 1% for non-rectified images.

References

[1] Y. Y. Lee, Z. Abdul Halim and M. N. Ab Wahab, "License Plate Detection Using Convolutional Neural Network–Back to the Basic With Design of Experiments," in *IEEE Access*, vol. 10, pp. 22577-22585, 2022, doi: 10.1109/ACCESS.2022.3153340.

[2] S. M. Silva and C. R. Jung, "Real-time license plate detection and recognition using deep convolutional neural networks," in *Journal of Visual Communication and Image Representation*, vol. 71, p. 102773, 2020. doi: 10.1016/j.jvcir.2020.102773.

[3] Indonesia. Pusiknas Bareskrim Polri, "ETLE Ditegakkan untuk Disiplinkan Perilaku Berkendara Masyarakat," 2022.

[4] Joshua, J. Hendryli and D. E. Herwindiati, "Automatic License Plate Recognition for Parking System using Convolutional Neural Networks," *2020 International Conference on Information Management and Technology (ICIMTech)*, Bandung, Indonesia, 2020, pp. 71-74, doi: 10.1109/ICIMTech50083.2020.9211173.

[5] N. Saif et al., "Automatic License Plate Recognition System for Bangla License Plates using Convolutional Neural Network," *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*, Kochi, India, 2019, pp. 925-930, doi: 10.1109/TENCON.2019.8929280.

[6] S. L. Muslimah, N. Ibnazhifi, and K. Blessinda, "Automatic Indonesian License Plate Recognition with YOLO as Object Detector," in *Iconic Research and Engineering Journals*, vol. 5, no. 6, pp. 184-190, 2021.

[7] N. Awalgaonkar, P. Bartakke and R. Chaugule, "Automatic License Plate Recognition System Using SSD," *2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA)*, Goa, India, 2021, pp. 394-399, doi: 10.1109/IRIA53009.2021.9588707.

[8] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Vancouver, Canada, 2023, pp. 7464-7475. doi: 10.1109/CVPR52729.2023.00721.

[9] J. Choi, D. Chun, H. Kim and H. -J. Lee, "Gaussian YOLOv3: An Accurate and Fast Object Detector Using Localization Uncertainty for Autonomous Driving," *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 2019, pp. 502-511, doi: 10.1109/ICCV.2019.00059.

[10] Y. Wang, P. Bian, Y. Zhou, and P. Chau, "Rethinking and Designing a High-Performing Automatic License Plate Recognition Approach," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8868-8880, 2022. doi: 10.1109/TITS.2021.3087158.

[11] P. Hidayatullah et al., "License plate detection and recognition for Indonesian cars," in *International Journal on Electrical Engineering and Informatics*, vol. 8, no. 2, pp. 331–346, 2016. doi: 10.15676/ijeei.2016.8.2.7.

[12] K. P. P. Aung, K. H. Nwe and A. Yoshitaka, "Automatic License Plate Detection System for Myanmar Vehicle License Plates," *2019 International Conference on Advanced Information Technologies (ICAIT)*, Yangon, Myanmar, 2019, pp. 132-136, doi: 10.1109/AITC.2019.8921286.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779-788. doi: 10.1109/CVPR.2016.91.

[14] W. Liu et al., "SSD: Single Shot MultiBox Detector," *Computer Vision – ECCV 2016*, ser. Lecture Notes in Computer Science, vol. 9905. Springer, Cham, 2016, pp. 21-37. doi:10.1007/978-3-319-46448-0_2.

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

[16] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," 2016, arXiv:1612.08242.

[17] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," 2020, arXiv:2004.10934.

[18] Q. Huang, Z. Cai and T. Lan, "A Single Neural Network for Mixed Style License Plate Detection and Recognition," in *IEEE Access*, vol. 9, pp. 21777-21785, 2021, doi: 10.1109/ACCESS.2021.3055243.

[19] D. P. Sihombing, H. A. Nugroho and S. Wibirama, "Perspective rectification in vehicle number plate recognition using 2D-2D transformation of Planar Homography," *2015 International Conference on Science in Information Technology (ICSITech)*, Yogyakarta, Indonesia, 2015, pp. 237-240, doi: 10.1109/ICSITech.2015.7407810.

[20] M. . Alhussein, K. . Aurangzeb, and S. I. . Haider, “Vehicle License Plate Detection and Perspective Rectification”, *ELEKTRON ELEKTROTECH*, vol. 25, no. 5, pp. 47-56, Oct. 2019. doi: 10.5755/j01.eie.25.5.24356.

[21] S. M. Silva and C. R. Jung, "License Plate Detection and Recognition in Unconstrained Scenarios," *Computer Vision – ECCV 2018*, vol. 11216. Springer, Cham, 2018, pp. 36. doi: 10.1007/978-3-030-01258-8_36.

[22] H. Xu, Z. -H. Guo, D. -H. Wang, X. -D. Zhou and Y. Shi, "2D License Plate Recognition based on Automatic Perspective Rectification," *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2021, pp. 202-208, doi: 10.1109/ICPR48806.2021.9413152.

[23] M. Jaderberg et al., "Spatial Transformer Networks," in *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[24] G. Jocher, "YOLOv5 by Ultralytics," 2020. doi: 10.5281/zenodo.3908559.

[25] Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861.

[26] J.L. Pech-Pacheco, G. Cristobal, J. Chamorro-Martinez, and J. Fernandez-Valdivia, "Diatom autofocusing in brightfield microscopy: A comparative study," *Proceedings of the 15th International Conference on Pattern Recognition (ICPR-2000)*, vol. 3, pp. 314-317, 200, doi: 10.1109/ICPR.2000.903548.

[27] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," in *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.

[28] Indonesia. Kepala Kepolisian Negara Republik Indonesia. "Peraturan Kepolisian Negara Republik Indonesia Nomor 7 Tahun 2021 Tentang Registrasi dan Identifikasi Kendaraan Bermotor." 2021.