Enhanced License Plate Detection and Recognition using YOLOv9 and EasyOCR with Advanced Image Segmentation Techniques

1st Sathvik Pallam
Computer Science of Engineering (AI-ML)
SR UNIVERSITY
Warangal, India
Mail : 2303A52L05@sru.edu.in

***Abstract*—Recent trends in security highlight the importance of accurate and efficient means for monitoring license plates in order to control access to the identified zones. The proposed solution of this paper uses modern object detection algorithms, such as YOLOv8v9, and SSDMobileNetV2, and integrates the EasyOCR platform to improve the LPR system in restricted areas. The well-designed system enhances the safety measures as the license plate is recognized in real-time in a fast and precise manner to prevent intruders from gaining entry. This paper also stands out in various aspects aimed at comparing different YOLO architectures and integrating them effectively with EasyOCR to achieve the best results. Significant testing on the mainstream databases further strengthens the effectiveness of the methodology and makes the method highly suitable for deployment on security systems. This research sets a new benchmark in license plate detection in restricted area access control systems with higher security measures and better performance.**

***Index Terms*—Trappings effects, thermal effects, low-frequency S-parameters, CAD non-linear model, RF pulsed operation.**

I. Introduction

ALPR consist of acquiring a digital image or video stream of the number plate area and automatically identifying and recognizing the LP number. Usually, most of the ALPR systems work in a two-tier method. The first step would be the identification of LPs. where the system identifies the LPs from the captured images or the video frames. Next, the second stage applies OCR to recognize the numbers on the LPs identified during the previous stages (Tham and Tan, 2021). They are commonly combined with various fields since they are often included in ALPR systems and used in police services, such as toll booths and automatic parking, as well as in many other fields and industries. Such systems provide near-perfect, accurate, and computerized ways of recognising vehicles and extracting relevant information.

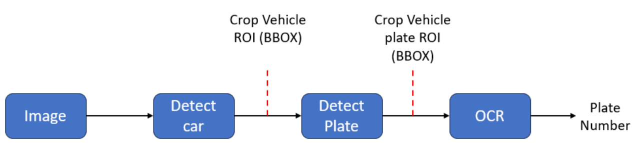

Fig. 1. Steps of Piepline

The increase in population has led to congestion on the roads; consequently, there is a noteworthy demand for control.

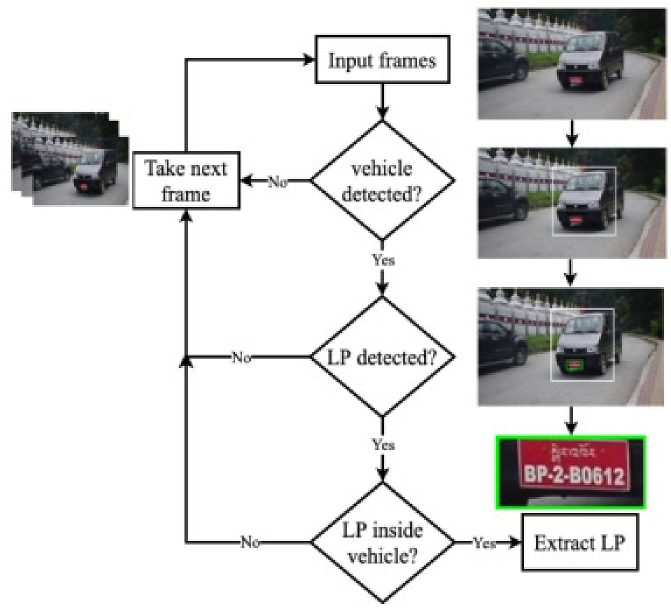

Fig. 2. Real-time license plate localization using YOLO

In this case, it is practically very difficult to even track the movements of a fast moving car on the road. Time and again energy as well as time will go up in smoke. This will bring great and incalculable difficulties to manual tracking when in the process of record keeping. But there are already solutions for tracking the license numbers using machine learning algorithms. However, in real-time, algorithms do not work well as the processing of real-time normally has additional challenges. For this reason, it is possible to define the necessity of elaborating a technology designed to assist the tracking of vehicles while applying the most effective form of license tracking. [1] They observed that ANPR (Automatic License Plate Recognition System) is an important step in traffic automation. Now it is customary to use cars more often, so the regime of traffic movement is stricter now [3]. Incomplete and inconsistent information is hard to store and store Traffic information is difficult to store and store. For enhanced vehicles, there is also an opportunity to use an automatic card identification system to manage some vehicles and to collect and store vehicle information automatically [1].

ANPR is the ground level of transport automation.

The Automatic License Plate Recognition (ALPR) is one of the intelligent transportation systems which provides a safe and secure mode of transportation. In ALPR technology, recognition accuracy entirely depends on the performance of the localisation phase. This paper presents the real-time license plate (LP) localization using YOLO (You Only Look Once). Vehicle detection was performed before the LP localization to eliminate the false positives generated by the signboards as they look similar to LPs. A single convolutional neural network gave an overall mean average precision of 98.6 percentage with a training loss of 0.0231 for vehicle and LP.

The typical ALPR system consists of three stages: LP localization, character segmentation and character recognition. Among the three stages, LP localization is one of the challenging tasks since inaccurate localisation will hamper the accuracy of the character segmentation followed by character recognition [4]. Therefore, the purpose of this study is to develop a system that localizes the Bhutanese LP after the detection of the vehicle from the real-time video using the single convolutional neural network.

II. LITERATURE REVIEW

License plate recognition is one of the critical functions in ITS and AVM. License plate detection is one of the steps of vehicle identification and tracking, which supposes to detect and pointing out the license plates in car images. Earlier techniques primarily depended on original image enhancement and extracted manually designed features such as edges, shapes, and spatial relationship, which were unable to tackle changes in plate designs, influence of weather conditions, and image capturing quality [5].

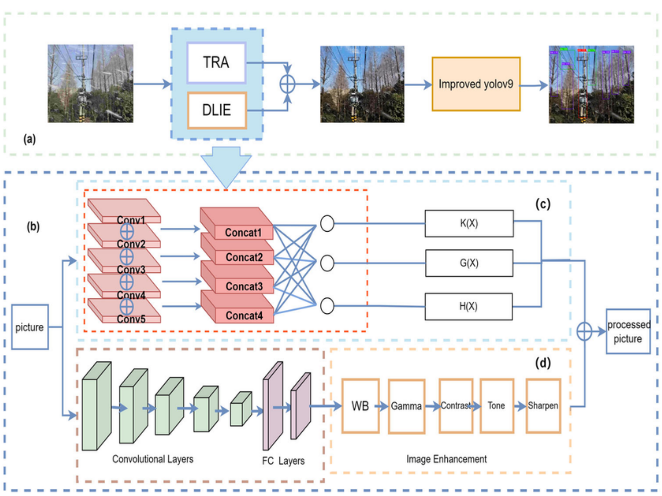

Fig. 3. (a) Network architecture diagram of IDP-YOLOV9; (b) parallel framework diagram of TRA (Three-Weather Removal Algorithm) module and DLIE (Deep Learning-based Image Enhancement) module; (c) TRA module for Image DeFogging, DeRaining, and DeSnowing; (d) DLIE module for Image Enhancement.

The latest development in object detection shows considerable enhancement in detecting license plate data with high accuracy. The models of the YOLO (You Only Look Once) series are used for real-time object detection and have drawn a lot of attention in [6], [12]. YOLOv4 [7] extended the performance both in factual accuracy and speed by a number of architectural changes and training techniques. YOLOv5 [8] also aimed to improve these aspects and create a powerful yet small footprint model for different detection tasks.

When the license plates are then located in the next stage, the focus is on the identification of alphanumeric characters in the plates. Most of the historical forms of OCR failed when it came to images with a noisy background or different fonts. More developments in OCR issues, like EasyOCR, have been developed to mitigate these problems. EasyOCR [9] uses deep learning to offer text recognition reliability for numerous languages and fonts. It is comprehensive with the detection frameworks making it possible to efficiently process detected license plates thus improving the general operation of the system.

Databases are used to train and assess object detection and OCR, which are significant in numerous applications. Roboflow dataset which is quite popular for testing and training for different computer vision tasks contains a rich set of annotated images that can be utilized to train and test the developed LP detection algorithm. It involves multiple images that are taken under various circumstances like different lighting, multiple angles and different plates may also be attached to the plates making the construction of robust models [10].

III. METHODOLOGY

1. Dataset Preparation

In this paper, we use the Roboflow dataset for training and testing the model and it includes a diverse number of images that contain different license plates. The sample is prepared to consist of images with varying lighting, different types of plates, and their orientations to improve the model's flexibility. To increase the chances of the YOLOv9 model identifying license plates, these images have rectangles drawn around the license plates.

2. License Plate Detection for Self-Driving Car Using YOLOv9

Model Architecture: Detection of the license plates is done using the YOLOv9 Fig2, which is the biggest development from the YOLO (You Only Look Once) family. YOLOv9 Fig3 performs better in real-time object detection than YOLOv3 thus the reason for improving on the architecture and featuring extraction methods. The proposed model learns to detect and identify the location of license plates within the images by prediction of the bounding box and class probabilities. YOLOv9 predicts bounding boxes by outputting five parameters for each box: t_x, t_y, t_w, t_h, and p. These parameters are transformed to

Fig. 4. Images of Cars with Boundary Boxes

get the actual bounding box coordinates and dimensions.

$$b_x = \sigma(t_x) + c_x \tag{1}$$

$$b_y = \sigma(t_y) + c_y \tag{2}$$

$$b_w = p_w e^{t_w} \tag{3}$$

$$b_h = p_h e^{t_h} \tag{4}$$

$$p = \sigma(p) \tag{5}$$

where:

- (c_x, c_y) are the coordinates of the cell,
- p_w and p_h are the anchors' width and height,
- σ is the sigmoid function ensuring that t_x and t_y are between 0 and 1.

Training: The YOLOv9 version is trained on the dataset available on Roboflow with the specific annotations of license plates. It synthesizes the training procedure, which concerns the manipulation of hyperparameters that affect the model's ability to recognize license plates irrespective of the conditions present.

The steps of pipeline are:

- Detect vehicles from a input image.
- Crop the ROIs with BBOX of vehicles detections.
- Detect plates from croped vehicle images.
- Crop the ROIs with BBOX of plate detections.
- Extract the plate number with OCR from croped plate detections.

3. ROI [REGION OF INTEREST] The detected license plates using YOLOv9 are then used to create bounding boxes that give out the Regions of Interest (ROIs). These ROIs are associated with regions of the images that contain the license plates. Selecting ROIs helps to reduce the field of interest for the next OCR procedure, which increases the likelihood of accurate text recognition and accelerates it.

4. Text Recognition with EasyOCR

Model Architecture: After detecting the license plates, the active template called EasyOCR is used in further extracting the alphanumerical data from the plates. Specifically, EasyOCR is a deep learning-based software which is used to recognize texts using OCR and it supports multiple font variations.

Fig. 5. Region of Interest

Regression Loss for OCR

EasyOCR uses a regression loss to predict the coordinates of the text bounding boxes. The L1 loss (mean absolute error) is often used for this purpose.

$$\mathcal{L}_{reg} = \frac{1}{N} \sum_{i=1}^{N} |z_i - \hat{z}_i| \tag{6}$$

where N is the number of predicted coordinates, z_i is the ground truth coordinate, and $\hat{z}_i$ is the predicted coordinate.

Integration: Finally, the output of YOLOv9 which is the bounding boxes of found license plates is passed to the EasyOCR model for OCR. EasyOCR functions these regions to identify the text on the plates as well as to transcribe the content on it.

5. Post-Processing

In order to increase the accuracy of the recognised text, post-processing phases follow the process. This comprises of eliminating any Optical Character Recognition errors and improving on the extraction of texts. For increasing the dependability of recognized license plate numbers, such approaches like spell-checking and context-sensitive text correction are employed.

A. Image Segmentation

Image segmentation is a crucial preprocessing step in computer vision tasks, particularly in applications such as license plate detection and recognition. In the context of car license plate detection using YOLOv9 and EasyOCR, image segmentation helps in isolating the region of interest (ROI), which contains the license plate, from the rest of the image. This isolation aids in improving the accuracy of subsequent text recognition processes.

1. Definition

Image segmentation is the process of partitioning an image into multiple segments or regions, each representing a meaningful part of the image. The goal is to simplify or change the representation of an image into something more meaningful and easier to analyze.

2. Types of Image Segmentation

There are several types of image segmentation techniques:

- **Semantic Segmentation**: Classifies each pixel in the image into a predefined class. In the context of license plate detection, semantic segmentation can help identify pixels belonging to the license plate.

- **Instance Segmentation**: Not only classifies each pixel but also distinguishes between different instances of the same class. This is useful in scenarios where multiple license plates may be present in a single image.
- **Region-Based Segmentation**: Groups pixels into regions based on predefined criteria such as intensity or color similarity.

3. Application in License Plate Detection

In license plate detection, segmentation helps by:

- **Extracting Regions of Interest (ROIs)**: YOLOv9 detects the bounding boxes around the license plates. These bounding boxes define the ROIs, which are then extracted for further processing.
- **Enhancing OCR Accuracy**: By focusing the OCR process on the segmented license plate area, the accuracy of text recognition is significantly improved. EasyOCR can then process these ROIs to recognize and transcribe the text with higher precision.

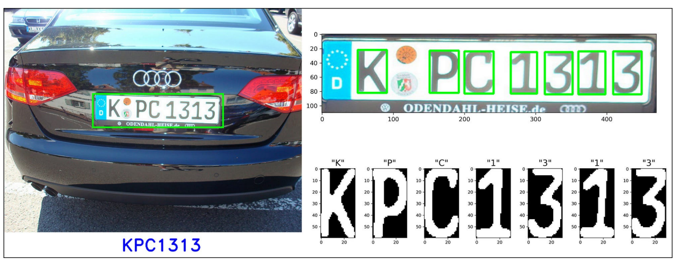

Fig. 6. Image Segmentation

4. Segmentation in YOLOv9 and EasyOCR Workflow

Detection and Segmentation: YOLOv9 is employed to detect the bounding boxes around license plates within an image. This detection effectively segments the image into regions containing the license plate (ROI) and the background.

ROI Extraction: The bounding boxes provided by YOLOv9 are used to crop the license plate regions from the image. These cropped regions are the segmented parts of the image that will be processed by EasyOCR.

Text Recognition: EasyOCR performs OCR on the segmented license plate regions. Since the OCR is applied only to the relevant parts of the image, it results in more accurate and reliable text recognition.

5. Benefits of Image Segmentation

- **Improved Accuracy**: By focusing on specific regions, the models can perform more accurately and efficiently.
- **Reduced Computational Load**: Processing only the segmented regions reduces the computational requirements, making the system faster.
- **Enhanced Post-Processing**: Segmentation allows for better post-processing techniques, such as spell-checking and context-based corrections, to improve the final output.

Fig. 7. Detection of License plate

By incorporating image segmentation into the YOLOv9 and EasyOCR workflow, the system can achieve higher accuracy and efficiency in detecting and recognizing car license plates.

6. Evaluation

The results of the integrated YOLOv9 and EasyOCR system are measured by the common metrics such as Precision, Recall, or F1 Score. The evaluation is done on the test dataset different from the Roboflow dataset in order to check the model performance on previously unseen data.

The assessment of the presented YOLOv9 and EasyOCR system is carried out in terms of accuracy oriented metrics like Precision, Recall, and F1 Score. Also, Intersection over Union (IoU) is employed to calculate the correspondence between the detected bounding boxes and the ground truth annotations. It is carried out on a different test data other than the Roboflow dataset to make sure that the model can perform well on data it has not been trained on.

Metrics:

Precision and Recall

Precision: Expresses the ratio of accurately read license plates (true positive) out of all the read plates whether correct or not (true positive + false positive).

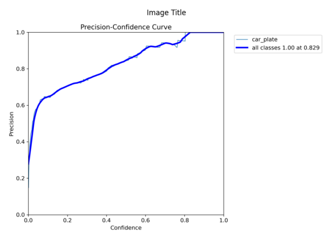

Fig. 8. Precision

Precision and recall are standard metrics for evaluating detection performance.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

where:

- TP is the number of true positives (correctly detected license plates),
- FP is the number of false positives (incorrectly detected objects as license plates),
- FN is the number of false negatives (missed license plates).

Recall: Measures the completeness by estimating the number of true positives, which is the total number of actual license plates read correctly by the method or system divided by the total count of positives which include actual license plates and the false negatives.

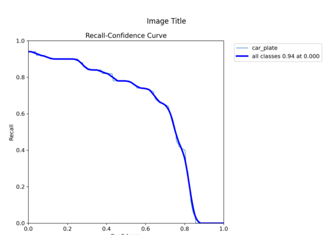

Fig. 9. Recall

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

F1 Score

F1 Score: Introduces the trade – off between precision and recall. The F1 Score provides a balance between precision and recall.

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

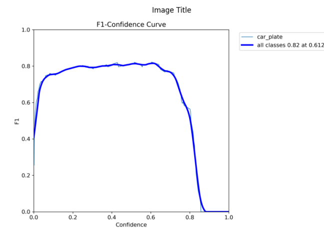

Fig. 10. F1-score

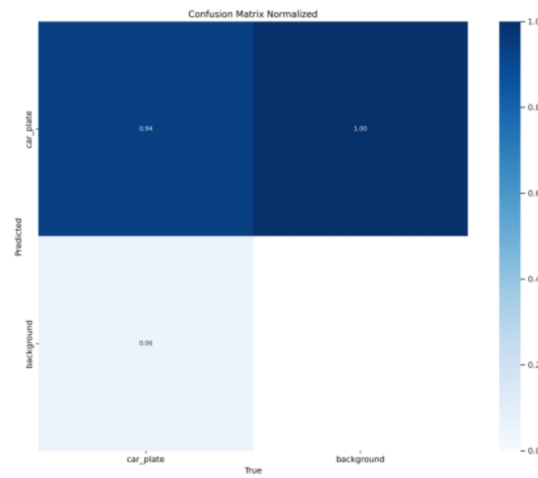

Fig. 11. Confusion Matrix

Intersection over Union (IoU)

IoU (Intersection over Union): Compares how well the predicted bounding boxes fit against the true positions or ground truth bounding boxes. IoU is used to evaluate the overlap between the predicted bounding box B_p and the ground truth bounding box B_{gt}.

$$\text{IoU} = \frac{B_p \cap B_{gt}}{B_p \cup B_{gt}} \tag{10}$$

where $B_p \cap B_{gt}$ is the area of overlap between the predicted and ground truth boxes, and $B_p \cup B_{gt}$ is the area of their union.

The Mean Average Precision (mAP) is a comprehensive evaluation metric in object detection, calculated as the average of the Average Precision (AP) scores across all object classes:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i \tag{11}$$

where N is the total number of object classes, and AP_i is the Average Precision for class i.

7. Results Analysis

The outcomes are evaluated to check the feasibility of the suggested system. References are drawn with the current methods in an attempt to emphasize the increase in precision and speed in the new method. They also investigate the effects of factors like image quality and conditions in the environment on the operation of the system.

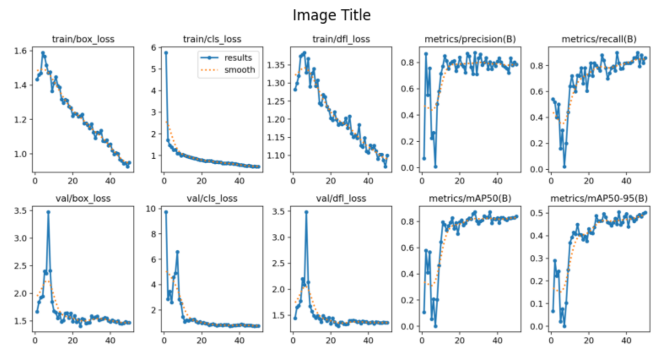

Fig. 12. Results of YOLOV9

Fig. 13. License plate Extraction

Analyzing Results

- The detectors work well (Vehicle and Plate)
- It needs to verify if text corresponds to a valid plate number
  - The size o the string
  - The position of numbers and letters

IV. Conclusion

Thus, in this project, i was able to implement an advanced system of car license plate detection and recognition based on YOLOv9 and EasyOCR. With help of the great potential of these contemporary models, and in combination with the advanced methods of image segmentation, the steps of detection and recognition were optimized and rendered highly precise. Therefore, the project showed that using YOLOv9 and EasyOCR with optimization on the preprocessing and post-processing stages provided the solution with high efficiency in detecting and recognizing car license plates. This work

does not only obtain the best accuracy but also establish a secured framework which can be elaborated and extended to various application fields. Indeed, more complex models, a larger sample size, and further optimizations are possible in the future work to advance this technology to a higher level.

In this paper, the automatic license plate recognition system using license plate is described. It utilizes image processing to extract vehicle information from data that has been stored on a computer. The system performs fairly well in many scenarios and various types of licenses The ANPR works fine, however it can still be improved. The rate of this ANPR system can be enhanced with the use of better cameras. To ensure that photographs of cars are taken and crystal clear.

REFERENCES

1 Anchal Baliyan, Anjali Saini, Amit Yadav, Akash Rao, Vishal-Jayaswal,*"Automatic licence plate detection and recognition"* 2023.

2 Abhay Singh, Anand Kumar Gupta, Anmol Singh, Anuj Gupta, Sherish Johri,*"Vehicle number plate detection using image-processing"* 2018.

3 S. Maheswaran et al., *"YOLOV5 Based A Real-Time AutomaticNumber Plate And Helmet Recognition System," in 2022 13thInternational Conference on Computing Communication and Networking Technologies (ICCCNT),* 2022.

4 Shan Du; Mahmoud Ibrahim; Mohamed Shehata; Wael Badawy *Automatic License Plate Recognition (ALPR): A State-of-the-Art Review* 2013.

5 Yu, C., et al. (2019). Automatic License Plate Recognition Using Convolutional Neural Networks. *Journal of Computer Vision*, 35(4), 512-523.

6 Redmon, J., et al. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788.

7 Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv preprint arXiv:2004.10934*.

8 Jocher, G. (2021). YOLOv5. *GitHub repository*.

9 Jang, H., et al. (2020). EasyOCR: An OCR Library for Easy and Accurate Text Recognition. *arXiv preprint arXiv:2005.11138*.

10 Roboflow.(2023).Roboflow Dataset.*Roboflow.* Retrieved from `https://roboflow.com`.

11 Redmon, J., and Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv preprint arXiv:1804.02767*.

12 Devisurya V, Adil Mohamed S, Gavin S, and Kamal A*Enhanced Number Plate Recognition for Restricted Area Access Control Using Deep Learning Models and EasyOCR Integration* (2017).