Brazilian Automatic License Plate Recognition: On the Importance to Add Vehicle Detection Phase in a Deep Learning Approach

Lucas Mity Kamado[1], Murilo Varges da Silva[2]
[1, 2] *Federal Institute of São Paulo (IFSP)*, Birigui, Brazil
mity.k@aluno.ifsp.edu.br[1], murilo.varges@ifsp.edu.br[2]

***Abstract*—In recent years, surveillance cameras have found wide-ranging applications, particularly in automatic vehicle license plate recognition to enhance security. One of the major problems of performing vehicle license plate recognition is the wide variety of existing license plate standards around the world, only in Brazil there are two standards, in addition, automatic license plate recognition depends directly on the quality of the input image. The objective of this paper is to evaluate whether adding the vehicle detection step improves license plate recognition rates in a deep learning approach. The method was implemented using the YOLOv8s model in all stages, however, two different approaches were proposed, one with and the other without the vehicle detection stage, the database used was the UFPR-ALPR, which only presents plates of the old Brazilian standard of 3 letters and 4 numbers.**

***Index Terms*—ALPR, Deep Learning, YOLOv8s**

I. Introduction

The automatic recognition of license plates can be defined as a series of digital image processing and computer vision techniques that seek, from an image or video with vehicles, to extract the license plate information in the text format [1]. With the increased use of surveillance cameras in cities, the possible applications of this technology in society are countless, but we can highlight: assistance in monitoring highways and streets; detection and search of stolen vehicles; access monitoring to public and private places and among other applications [2].

One of the biggest problems in carrying out the task of recognizing a vehicle license plate is the wide variety of existing license plate standards, each country adopts a different standard, Brazil, for example, has two models of license plates currently (see Figure 1), therefore, developing a system that recognizes any type of license plate is a challenger, a large number of adaptations are necessary [3]. In addition, another adversity to be faced in this type of application is the quality of the license plate in the images, as it may not be ideally visible, due to natural causes, such as lack of light or small occlusions in the license plate, or even the physical conditions that the license plates are in.

That said, the present work seeks to explore the existing techniques for the recognition of license plates of the old model used in Brazil. For this, it focuses on deep machine learning techniques, such as YOLO ("You Only Look Once") [6], which have shown promising results compared to traditional computer vision techniques, in addition to being much more used in this type of task.

This work was supported by grant #2022/05407-0, São Paulo Research Foundation (FAPESP).

The implementation of the proposed method, according to the literature, is basically composed of three stages: i. the detection of the license plate in the image; ii. character segmentation and iii. character recognition [7]. However, we can find some methods that perform the license plate recognition task in a smaller or larger number of steps, nevertheless, when this number is greater, they are steps that seek to help or improve the performance of one of these three predefined basic steps, since when the number is smaller, some of these basic steps still remain included in the process but in an implicit form.

II. Materials and Methods

This section presents the database and the method used in the experiments at each stage of the proposed method to generate the results presented in Section III, from the choice of the database to the detection and recognition of the characters of the license plates.

A. UFPR-ALPR dataset

The UFPR-ALPR dataset, created in 2018, was obtained from recordings of 150 videos of 1 second each and with a frame rate equal to 30, thus producing a total of 4500 images. The videos were recorded from inside a car recording urban traffic, among the vehicles recorded by the cameras are cars, motorcycles, vans, trucks and buses. Some examples of images from the database can be seen in Figure 2. The 4500 images are divided into training, validation and test sets, with 1800 (40%), 900 (20%) and 1800 (40%) images, respectively.

As the images in this database were collected only from the Parana state vehicles, the distribution of letters is not balanced, as can be seen in Figure 3. The plates of each Brazilian state start with some particular letters, in this case, the plates of the Parana state vary from AAA-0001 to BEZ-9999.

B. Automatic license plate recognition

The automatic license plate recognition method was implemented using two approaches that differ in one step, vehicle detection. While the first approach uses a vehicle detection step, facilitating the license plate detection task, the second

Fig. 1. License plate models in Brazil, on the left, the old model, and, on the right, the Mercosul model. Upper plates for cars and lower plates for motorcycles. Source: [4], [5].

Fig. 2. Examples of images from the UFPR-ALPR database [8].

seeks to detect the license plate without prior vehicle detection. In Figure 4 you can see a detection example for each method. Considering that in YOLO, a fixed resolution of the input image in the neural network is defined, which, in most cases, is a lower resolution than the original image, the method with vehicle detection has an advantage, since vehicles are larger objects, therefore, they are less affected by the resizing of the image in the input of the neural network than the license plates. On the other hand, the license plate recognition speed is favorable for the second method, which has one step less.

One of the most used versions of YOLO is the YOLOv3 [9], a model launched in 2018, however, one of the last available versions is the YOLOv8 [10], launched in 2023. In view of this, it was decided to adopt the most recent YOLO model to implement all steps of automatic recognition of license plates. YOLOv8 basically has 5 models: YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l and YOLOv8x, and these models basically differ by the number of layers and size of the model. Trying to combine performance and speed, among them, YOLOv8s was chosen to compose the experiments of this work.

C. Vehicle detection

The vehicle detection step was trained with a neural network input dimension of 640x640, the default value of YOLOv8s. And the experiments were produced with a confidence rate

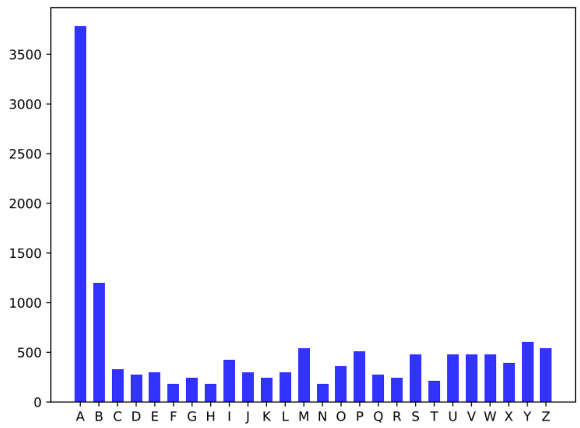

Fig. 3. Distribution of letters of license plates in the UFPR-ALPR database [8].

Fig. 4. Examples of license plate recognition using the two developed models. Bounding boxes in green correspond to the vehicle detection step, in red license plate detections and in blue segmentation and character recognition. Plate identification has been blurred in order to preserve and respect the privacy of vehicle owners.

threshold of 25%, also a default value. It was decided to keep the default values in this step, as it is the first step of license plate recognition. Generally, predictions with very low confidence rates tend to be false positives. In addition, vehicle detection has 2 prediction classes: car and motorcycle.

D. License plate detection

The license plate detection step was divided into two neural networks: plate detection without a previous vehicle detection and plate detection preceded by vehicle detection. The first, as it is the same case as the II-C section, a neural network that is having the first contact with the image, the training and testing parameters were the default values, 25% of confidence threshold and 640x640 of network input dimension.

The second license plate detection neural network had a different treatment from the first, the only similar parameter of the two was the neural network input dimension. This YOLOv8s preceded by vehicle detection has two advantages: the image that enters your network will already be cropped with the vehicle detection, and even this YOLOv8s was trained with the cropped image of vehicles, unlike the first model that detects the license plate from the whole image, and as it has the knowledge that the input image will be of a vehicle, the probability of having a license plate in this image is high. Thus, the confidence threshold of this YOLOv8s model may be low, as license plate detections with low confidence in an image of a car have a high chance of being a license plate. A confidence threshold of 0.1% was then defined for this neural network. If more than one license plate is detected on a vehicle, due to the low confidence threshold, only the license plate with the highest confidence rate is considered.

For both license plate detection models, the prediction class is only one, the old standard plate, since UFPR-ALPR does not have Mercosul standard plates.

E. Character segmentation and recognition

To carry out the segmentation and character recognition, only a detection neural network with a confidence threshold of 0.1% was used. Following the same reasoning as the license plate detection step preceded by the vehicle detection step, the chances of having 7 characters in a license plate prediction are high, if more than 7 characters are predicted on a license plate, only the 7 predictions with the highest confidence rates are considered. The character detection training was performed on clippings of plates from the original images, since, in the end, this neural network will always receive a plate to perform prediction. Different from all the other trained YOLOv8s models, this one was trained with a neural network input dimension of 384x384, because normally the license plates contained in the databases tend to be small objects. Character detection was trained with 35 classes: 10 numbers, from 0 to 9, and 25 letters, from A to Z excluding O.

After the character detections, a heuristic of equivalence between numbers and letters was used: in the first three characters, 5 is replaced by S, 7 by Z, 1 by I, 8 by B, 2 by Z, 4 by A, 6 by G and 0 by O, while in the last 4 characters, Q is replaced by 0, D by 0, Z by 7, S by 5, J by 1, I by 1, A by 4 and B by 8. This is possible since the old license plate used in Brazil is presented in the format of "LLLNNNN", where L are letters and N numbers.

In the character detection step to return the recognized license plate, it was necessary to define a way of ordering the predictions, as they are not obtained ordered. For this, two ways were defined to acquire the heuristic of what type of license plate was detected, since the disposition of the characters varies according to the type of vehicle, car or motorcycle, as can be seen in Figure 1, for the recognition of license plates with the vehicle detection step, the information of what type of vehicle was used, while in recognition without vehicle detection, the proportion of the detected license plate (width/height) was used, if it is greater than 2, treat it as a car, if it is smaller, treat it like a motorcycle. For cars, the sorting of characters is based on the X coordinate position of the centers of each character's bounding boxes, that is, sorting is in horizontal axis order. As for motorcycles, the median of all the Y coordinates of the detections is calculated, so that we can separate the characters into two groups, those above the Y median and those below, after being separated, just sort each group separately by the X coordinate, in the same way as the license plate, so that we can concatenate the two groups, with the upper group being the first characters and the lower group being the last.

F. Metrics

To evaluate the performance of YOLOv8s's neural networks at each stage of license plate recognition, the metrics presented below were used.

1) IOU (Intersection Over Union): is a method to compare detection bounding boxes. For this, the ratio between the intersection and the union of the predicted and actual bounding boxes is calculated, represented in equation (1). The IOU value ranges from 0 to 1, the closer to 1, the greater the similarity of the dimensions and positions of the bounding boxes, and the closer to 0, the less similar.

$$IOU = \frac{A \cap B}{A \cup B} \tag{1}$$

where A is the predicted bounding box area and B is the actual bounding box area.

2) Precision and Recall: Precision is calculated by dividing the true positives by anything that was predicted as a positive, Precision measures the number of instances that are relevant, out of the total instances the model retrieved, calculated by (2). Recall (or True Positive Rate) is calculated by dividing the true positives by anything that should have been predicted as positive, Recall measures the number of instances that the model correctly identified as relevant out of the total relevant instances. calculated by (3).

$$Precision = \frac{TP}{TP + FP} = \frac{Objects\ detected\ correctly}{All\ objects\ detected} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} = \frac{Objects\ detected\ correctly}{All\ objects\ labeled} \tag{3}$$

where TP (True Positive) are the correct predictions with an IOU greater than a threshold, usually 0.5, FP (False Positive) are the predictions without labels or that their IOU is below the threshold, meaning a bad prediction and FN (False Negative) are the dataset labels that had no prediction.

3) Accuracy: Accuracy is a simpler and more general measure, unlike the previous ones it does not use the IOU, from a sample it calculates the percentage of hits that the model obtained, this metric is calculated by (4).

$$Accuracy = \frac{Total\ number\ of\ hits}{Total\ number\ of\ items} \tag{4}$$

In the experiments, the performance of the detection and segmentation steps are measured by the precision and recall metrics, the IOU participates implicitly by helping with the calculations of the two measures mentioned, while the recognition steps, such as the final recognition of the license plate and character recognition, have their performance measured through accuracy metric.

III. Results

This section presents the results achieved by the methods used in the experiments and some comparisons with other works.

A. Evaluation and methods comparison

This comparative experiment analyzes the performance differences of a license plate recognition method with and without an additional vehicle detection step. The results obtained by the method with the additional step are shown in Table I, while the results of the other method without vehicle detection are shown in Table II. Although the segmentation and character recognition step was performed in only one step, the evaluation of this license plate recognition stage was divided into 2 evaluations: first, how well the characters were segmented regardless of the classification assigned to each one of them through recall (for a character to be considered a True Positive, it was enough to have an IOU greater than 0.5 in relation to some labeling in the dataset) and then, the ability to recognize or classify the segmented characters through accuracy was evaluated.

It is evident when comparing Table I with Table II that the difference in license plate recognition performance between the two methods was 2.22 percent points, this difference was relatively small, but it still denotes that the vehicle detection step in these cases of the test set acted as a hindrance instead of facilitating the steps that follow it. This is due to the error propagation phenomenon by forwarding the steps, since the method whose results are presented in Table I contains more steps, thus whether errors occur in the early steps, those errors affect the final decision.

TABLE I
RESULTS OBTAINED IN EACH STEP OF AUTOMATIC LICENSE PLATE RECOGNITION WITH VEHICLE DETECTION. RECALL FOR THE DETECTION AND SEGMENTATION STEPS AND ACCURACY FOR THE RECOGNITION STEPS.

Step	Recall	Precision	Accuracy
Vehicle Detection	95.09%	98.44%	-
License Plate Detection	92.82%	98.34%	-
Character Segmentation	90.94%	96.97%	-
Character Recognition	79.81%	-	-
License Plate Recognition	-	-	40.33%
$\geq$ 6 characters	-	-	71.88%

TABLE II
RESULTS OBTAINED AT EACH STAGE OF AUTOMATIC LICENSE PLATE RECOGNITION WITHOUT VEHICLE DETECTION. RECALL FOR THE DETECTION AND SEGMENTATION STEPS AND ACCURACY FOR THE RECOGNITION STEPS.

Step	Recall	Precision	Accuracy
Licence Plate Detection	94.32%	99.30%	-
Character Segmentation	92.18%	97.10%	-
Character Recognition	80.83%	-	-
License Plate Recognition	-	-	42.55%
$\geq$ 6 characters	-	-	71.66%

Some predictions were unique to each method. The license plate recognition approach with vehicle detection was able to make predictions, which were not necessarily correct, in situations where license plate detection is difficult, while the other method was not able to return any detection. In comparison, there were cases where the method with vehicle detection failed to detect the vehicle, it prevented the license plate detection. In contrast, the license plate recognition method without vehicle detection had the advantage of being able to detect the license plate regardless of whether or not there is a vehicle in the image.

As the character recognition step was where the greatest loss of efficiency was obtained, it makes sense to analyze it through a confusion mixture that is presented in Figure 5. The most alarming case that it is possible to notice is the letter N that was predicted in all its occurrences as the letter R, this can be justified by Figure 3, since the letter N is one of the letters that are less present in the databases, thus, training is impaired with an imbalance in the number of samples of each class in the database. Another clear example that we can relate Figure 3 with Figure 5 is the amount of true positive of the classes that are in greater quantity in the base, the letter A and B, visibly are one of the classes with the highest hits. That is, as expected, most of the characters with the lowest concentration in the database had a correct recognition rate below the desired level, thus highlighting the importance of having a balanced distribution of classes in a database.

B. Comparison with other works

This subsection aims to analyze the results obtained in this work with the results achieved by other works. From Table III, we can see that the vehicle detection of [8] obtained a recall of 100%, that is, all vehicles in the test set were successfully detected, therefore, this step practically did not propagate an error to the next step, unlike the model developed in this work. However, it is worth mentioning that the methods used by [8] were based on detailed tests and adjustments, according to their work, the vehicle detection step is adjusted by its confidence threshold, the lower it is, the greater the recall, however, the lower the precision. The problem with lowering the confidence threshold too much is that although there is an increase in the detection of true positives, as a result, there is also an increase in false positives, so a balance between precision and recall in these adjustments must be sought. Laroca et al. [8] also adjusts the confidence thresholds for the following steps of vehicle detection, however, it is more drastic, after the first step all the rest of the steps have confidence thresholds equal to zero, that is, it forces YOLO to return predictions, regardless of whether the confidence value is very close to zero. The problem with this is that any bad vehicle detection or even a false positive can lead to a false license plate detection.

Comparing directly, the work by [8] approached the character segmentation and recognition task as two separate steps, unlike the work reported in this paper, it first segments the characters and then recognizes them through two different neural networks, one to recognize letters and the other for numbers.

Finally, despite the excellent recall rates of [8] in the recognition steps, the three results had a decrease in license plate recognition accuracy, showing that having high character recognition rates does not mean good license plate recognition accuracy.

TABLE III
COMPARATIVE RESULTS. RECALL FOR THE DETECTION AND SEGMENTATION STEPS AND ACCURACY FOR THE RECOGNITION STEPS.

Step	Recall/Accuracy		
	Laroca *et al.* [8]	1*	2**
Vehicle Detection	100%	-	95.09%
License Plate Detection	98.33%	94.32%	92.82%
Character Segmentation	95.97%	92.18%	90.94%
Character Recognition	90.37%	80.83%	79.81%
License Plate Recognition	64.89%	42.55%	40.33%
$\geq$ 6 characters	87.33%	71.66%	71.88%

*: License plate recognition implemented without vehicle detection.
**: License plate recognition implemented with vehicle detection.

IV. CONCLUSION

Given what was exposed in this paper, it can be concluded that the vehicle detection step did not bring a gain in the steps that follow it in the performed experiments, since resulted in the lack of recognition of the license plates of some undetected vehicles. Despite this, there are still some reasons to choose to include the vehicle detection step in license plate recognition. In a real system, to prioritize safety, vehicle detection is essential, as we must remember that the main purpose of automatic license plate recognition is to find a way to identify vehicles by their license plates, thus, recognizing license plates without vehicles in a certain way can lead to greater problems. But also, as already mentioned, when we have applications in which the size of the license plates in the images are either small or vary greatly in size, the vehicle detection step can serve as a solution, since, in a way, when it works, it acts as a pre-processing step performing an image crop or an enlargement of a part of the image that is most relevant for license plate detection.

Concerning the segmentation and character recognition step, it was clear that performing this task with a single neural network is not the best alternative. Analyzing the results of this study, it is possible to note that using the heuristic of the position of the characters on the license plate to know whether it is a number or a letter to carry out equivalence exchanges between them is a limited technique. For example, if the wrong prediction of C and G by 6 is common, as it is only possible to choose one of the occurrences to perform the exchange of number by letter, normally the most frequent one is chosen, if it is chosen to exchange 6 by G, the character C continues to be predicted wrong. One solution would be to improve training to reduce the wrong predictions, however, the way that has shown better results is the division of the character recognition neural network in two, one for the recognition of letters and another for numbers, in this way, the wrong prediction of number by letter or letter by number is impossible.

ACKNOWLEDGMENT

The authors would like to thank São Paulo Research Foundation (FAPESP) grant #2022/05407-0.

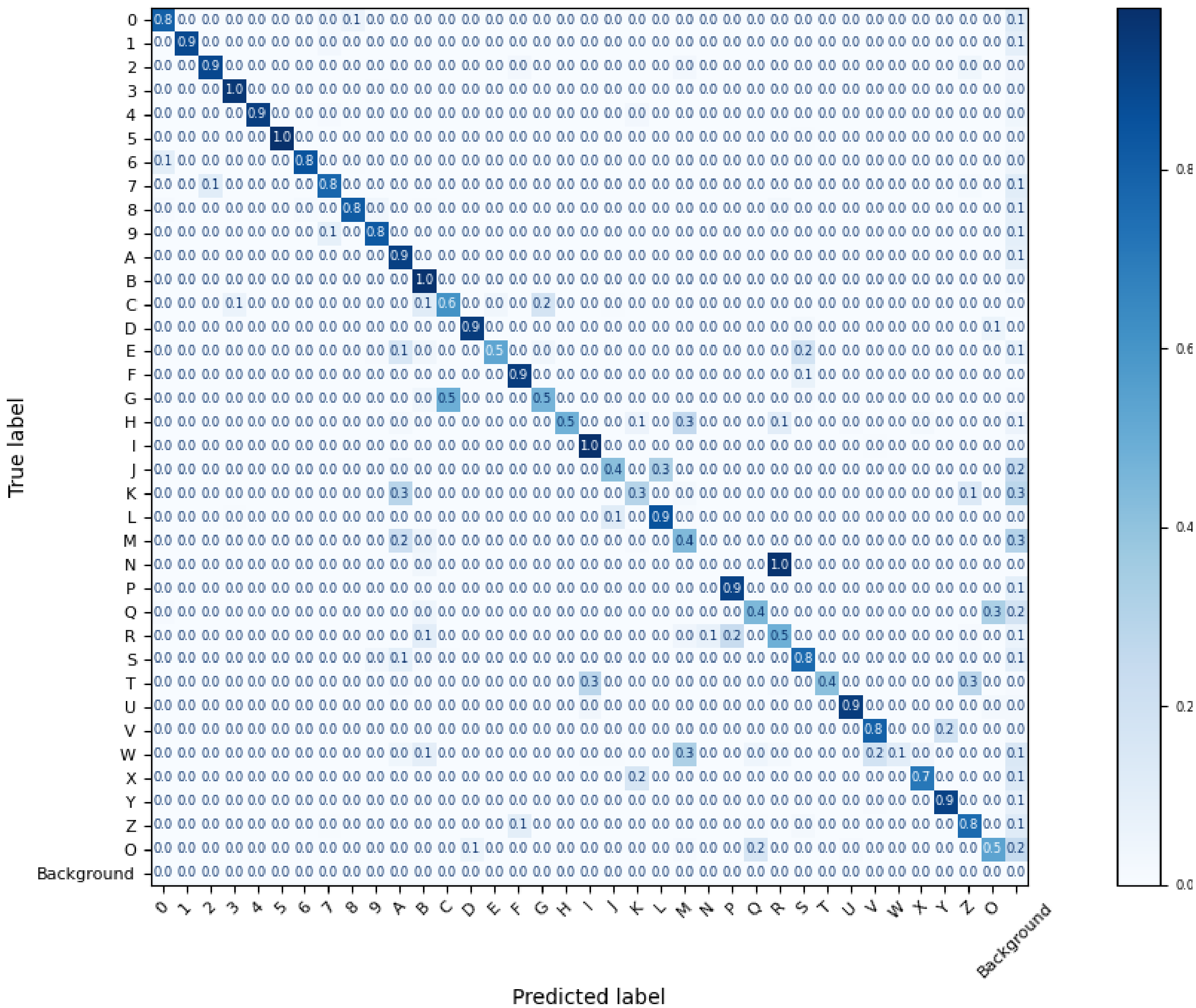

Fig. 5. Character recognition confusion matrix.

References

[1] P. Agarwal, K. Chopra, M. Kashif, and V. Kumari, "Implementing alpr for detection of traffic violations: a step towards sustainability," *Procedia Computer Science*, vol. 132, pp. 738–743, 2018.

[2] G. L. Corneto, F. A. da Silva, D. R. Pereira, L. L. de Almeida, A. O. Artero, J. P. Papa, V. H. C. de Albuquerque, and H. M. Sapia, "A new method for automatic vehicle license plate detection," *IEEE Latin America Transactions*, vol. 15, no. 1, pp. 75–80, 2017.

[3] C. Henry, S. Y. Ahn, and S.-W. Lee, "Multinational license plate recognition using generalized character sequence detection," *IEEE Access*, vol. 8, pp. 35 185–35 199, 2020.

[4] CONTRAN, "Resolução contran nº 231 de 15/03/2007," https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-contran/resolucoes/resolucao_231.pdf, 3 2007.

[5] ——, "Resolução contran nº 510 de 27/11/2014," https://www.gov.br/transportes/pt-br/assuntos/transito/conteudo-contran/resolucoes/resolucao5102014.pdf, 12 2014.

[6] 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*, 2016, pp. 779–788.

[7] J. Shashirangana, H. Padmasiri, D. Meedeniya, and C. Perera, "Automated license plate recognition: a survey on methods and techniques," *IEEE Access*, vol. 9, pp. 11 203–11 225, 2020.

[8] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz, and D. Menotti, "A robust real-time automatic license plate recognition based on the yolo detector," in *2018 international joint conference on neural networks (ijcnn)*. IEEE, 2018, pp. 1–10.

[9] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[10] G. Jocher, A. Chaurasia, and J. Qiu, "YOLO by Ultralytics," 1 2023. [Online]. Available: https://github.com/ultralytics/ultralytics