2020 International Conference on Machine Vision and Image Processing (MVIP)
Faculty of Engineering, College of Farabi, University of Tehran, Iran, 19 & 20 Feb. 2020

Iranian License Plate Recognition using Deep Learning

Atefeh Ranjkesh Rashtehroudi
Department of Computer Engineering
University of Guilan
Rasht, Iran
atefehranjkesh@msc.guilan.ac.ir

Asadollah Shahbahrami
Department of Computer Engineering
University of Guilan
Rasht, Iran
shahbahrami@guilan.ac.ir

Alireza Akoushideh
Shahid-Chamran College
Technical and Vocational University
Tehran, Iran
akushide@afr.ac.ir

***Abstract*— Automated License Plate Recognition (ALPR) has many applications in intelligent transport system. The ALPR has three main steps, License Plate (LP) localization, segmentation and Optical Character Recognition (OCR). Each step needs different techniques in real condition and each technique has its specific characteristics. The LP localization techniques detect the LP after that segmentation algorithms should segment and isolate each character from each other. Finally, the OCR step is applied to recognize the separated characters. The final accuracy depends on the accuracy of each step. To improve the OCR step performance, we combine both segmentation and OCR steps as a single-stage using deep learning techniques such as the You Only Look Once (YOLO) framework. Our experimental results show that this proposed approach recognizes the Iranian LP characters with accuracy 99.2% compared to previous works.**

Keywords—optical character recognition, deep learning, YOLO, artificial neural network, support vector machine

I. INTRODUCTION

Due to the uniqueness of the License Plate (LP) in the vehicle, the Automatic LP Recognition system (ALPR) has become important research in the field of Intelligent Transformation System (ITS) [1]. The ALPR has many applications, such as traffic law enforcement, private spaces access control, road traffic monitoring, smart toll and speed control stations [2-5]. Noisy and dirty images, occlusion, different LP types and size, camera quality and different climatic and lighting conditions make it difficult to recognize text from the image [2, 3].

There are challenges in recognizing the LP of vehicles, as we know, the quality of the input image affects the result of the text recognition. Accordingly, pre-processing can improve efficiency. At various stages of preprocessing, we must address the challenges of image quality such as shadows and noise in different ways. To do this, they often use methods such as skew correction, image binarization or even removing unnecessary parts [3]. Of course, each of the above has its own methods that apply to the original image with respect to the amount of noise in them.

Feature extraction is one of the most basic steps in recognition text, so that applying the appropriate feature enables the classifier to be able to distinguish different classes with good accuracy. Of course, the size of feature vectors should also be considered. Researchers use a variety of features to recognize the LP, which may be based on statistical [7] size and shape [8], area [9], color [2], Scale Invariant Feature Transform (SIFT) [10], and other features [11].

In learning-based approaches for character recognition, Artificial Neural Network (ANN) is very prominent. After extracting the feature from the characters, ANN is used to detect them [12]. To train the ANN to recognize characters on the test dataset, a certain number of layers and neurons are typically used. Of course, different methods such as feedforward backpropagation, feedback, and feedback self-learning, may be used to learn the network [13]. Support Vector Machine (SVM) is a learning technique for classification, but nowadays it is mostly used in combination with other methods. For example, SVM is used to reduce LP candidate detection error and the execution time. In order to achieve the above goals, the combining SVM with Histogram of Oriented Gradients (HOG), K-nearest neighbor (KNN) and a combination of logistic regression with Random Forest and supervised K-means is used [14].

Deep learning (DL) techniques, provide conditions for the automatic selection of features from the image. Based on this, deep learning can be used to extract features and classifications [15]. Today, in most computer vision tasks, it is common to use deep learning techniques, especially methods that use some form of Convolutional Neural Network (CNN) to achieve the most advanced performance [16, 17]. So unlike the previous methods, we use DL methods to identify each of the LP character regions and then identify the characters using the Tesseract engine.

The remainder of this paper is organized as follows: In section 2 describes background and related works. Purpose approach is presented in section 3. We end it with the conclusions.

II. BACKGROUND AND RELATD WORKS

The LPs of vehicles may have different styles. The first challenge is the standardization of these plates. Iranian LP mainland has a strict standard, but it can be various in colors and the compound mode of characters. These are flexible in character count. There are characters from 0 to 9 with the Persian alphabet or without them or it may have specific symptoms but they have standard font and font size,

978-1-7281-6832-6/20/$31.00 ©2020 IEEE

but differences in weather conditions, light intensity, and contamination on the plaque surface cause differences in image quality. Fig. 1 shows some kinds of Iranian LPs. The target of this study is to implement an ALPR system focusing on recognition of all types of LPs.

As can be seen in Fig. 2, first LP should be localized after that, extracted area should be removed from detected LP. In the second step, extracting characters on the extracted LP should be segmented and finally segmented characters are recognized. In fact, the good performance of each step leads to the desired result. Most existing solutions make simple assumptions, namely that their primary data are extracted from fixed cameras with a certain angle and resolution, which may not work well in other datasets [2]. Therefore, we need to look for a way to achieve maximum efficiency in real data sets.The following steps are taken to achieve this goal:

(a) (b)

Figure. 1. Different type of Iranian LPs. (a) Difference background color and character. (b) Different quality of LPs.

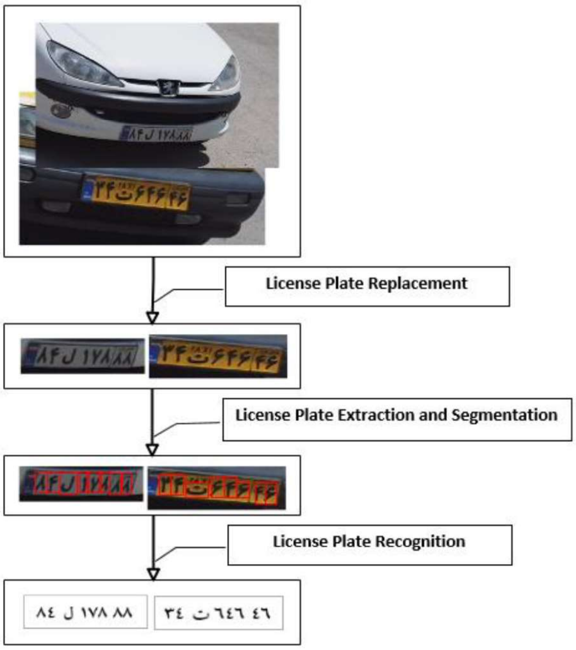


Figure. 2. Vehicle license plate recognition steps.

All deep learning structures are essentially multi-layer neural networks. What distinguishes these architectures from traditional neural networks is the way they threat input data. One of the features in most deep learning architecture models is the internal transformation of layers before reaching the desired output information. Some layers work on creating multiple intermediate input data based on input data. This creation can be based on several types of transformations, such as image filters or The denoising methods to enhance the quality of input data. YOLO was designed to work with object detection at first. It could detect objects with a good precision and good recall. YOLO can detect a relevant quantity of candidate areas in an image. In fact, YOLO is a framework that uses deep learning to achieve the desired results [19].

Some researchers have proposed DL methods based on segmentation free approaches [17]. They propose a CNN to perform feature extraction on the LP and a recurrent neural network to learn the sequential order of character features. Another study, proposed an approach for reading texts in the wild with a CNN, which explores deep learning for achieving high recognition results for natural images [17]. Some methods design deep CNNs to detect LPs with varying conditions, including position, lighting, occlusion and attempt to recognition various types of LPs such as size, background, font, and so on. Some approaches explore deep CNN at the pixel level. They suggested multi-digit number recognize from street view imagery [18].

In some studies, they propose a new robust real-time ALPR system based on the YOLO object detection convolutional neural networks with 19 convolutional layers and 5 max-pooling layers [19]. One of the biggest challenges in creating real-time ALPR systems is the need for expensive hardware. Therefore, we have to use a deep model of detection and recognition in a very short time, to achieve these goals; we choose to use a YOLO-based CNN network [16]. Others proposed a text detection model based on the YOLO Architecture that uses a fully convolutional deep neural network to identify text regions. The output of text regions identified by these approaches can then be used as input for further systems based on deep neural networks that perform text recognition [20]. In some studies, to obtain state-of-the-art recognition accuracy, it is recommended to use the YOLO detector for detecting plates and the CNN for recognition [19]. In [20], proposed a complete ALPR system that performs over a variety of scenarios and camera setup, it uses the YOLOv2 [21] to detect vehicle and introduce a novel CNN called WPOD-NET for LP detection and affine transformation regress then followed by an OCR module using a modified YOLO network to character recognition.

The OCR algorithm of any type of font, including handwritten characters, can be done easily today using free software such as Tesseract [21]. Some studies combine YOLO for detection and tesseract for recognition. For example, some researchers explored advanced DL techniques such as CNNs, Recurrent Neural Networks (RNNs), Long Short-term Memory (LSTM) on individual domains such as object recognition and then text reading modules. Their last end-to-end system was built and modified on deep convolutional network platform, by using YOLO network for object detection and Tesseract (LSTM) for character recognition and analysis. Both systems were tested and benchmarked on same set of test database for homogeneity in benchmarked results [24].

We used several operational phases to identify characters 1 to 9 (1000 samples in each class). In the pre-processing step, we used thresholding, noise reduction, image resizing and moving the center of mass (COM) to the center of the image. After the pre-processing step, we used Zernike moments to extract the feature, and finally, we used ANN and SVM in the classification step. Due to the rotation-invariant feature of the Zernike moments, we had difficulty recognition Persian numerals 7 and 8, and by adding the crossing count feature, we were able to increase the accuracy of the ANN by 99.1%.

In order to complete the dataset, we added to the previous dataset using images obtained from surveillance cameras, number 0 and letters on an Iranian LP. Fig. 3 shows the results of the evaluation of the numbers and letters on the LP of Iranian vehicles. Finally, by evaluating the new datasets, we get at 87% accuracy for the SVM and 90.2% for ANN.

Given that, there is no complete data set for Persian language texts, it is also difficult to obtain a balanced database and, on the other hand, achieving accurate modeling in learning methods depends on image quality, segmentation and feature extraction, so presentation a way to respond appropriately to these challenges is essential. In fact, we are looking for a way that can be used with better learning models for all plates without the need for segmentation and to select the appropriate learning features that ultimately achieve high accuracy in the identification domain.

III. Proposed Approach

As Fig. 4 shows, after receiving the extracted plates, we perform pre-processing, LP extraction and recognition operations. The proposed method does not use image segmentation, there are several reasons for this. As mentioned earlier, the extracted LPs may have problems such as size, un-normal, tilt, non-uniform brightness, and some other noises. To overcome all of these problems, using the pre-processing stage before segmentation algorithms [1]. Also, they need to be annotated after segmentation, which is a time-consuming task. So we try to use an algorithm that has some segmentation and annotation together. The image quality of all the plates is not the same, the camera's viewing angle or the light's angles cause the extracted plate to be tilted or shadowed. So the first step in the proposed method is to pre-process the plates. In this step, we first use Hough line transform methods to correct the tilting of the LPs [23], followed by the rotation filters as much as the angle required to fix the tilting image, by used the bilinear interpolation algorithm applied on the original image [24]. Finally, Bradley method [25] is used to remove the shadows. Therefore, after the pre-processing, the noise in the image of the LPs has been greatly reduced.

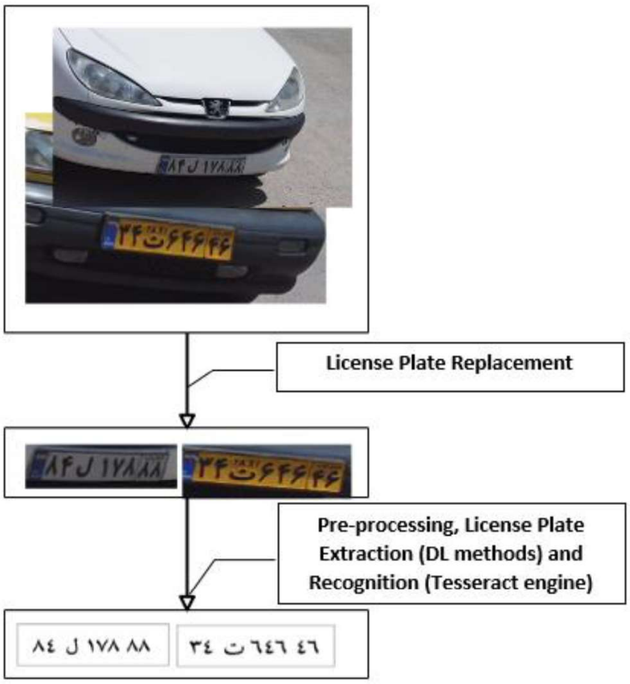


Figure. 4. Combination of segmentation and OCR steps as a single step using deep learning in the proposed approach.

The datasets used in this study are 1000 actual samples extracted from surveillance cameras. On the other hand, we know that one of the most important challenges in deep learning techniques is a large amount of data needed to train them. To address this challenge and to complement the dataset and multiply their numbers, we use augmentation methods on pre-processed images. This makes learning the model a better process. Give that many of the potential flaws in image pre-processing have been remedied.

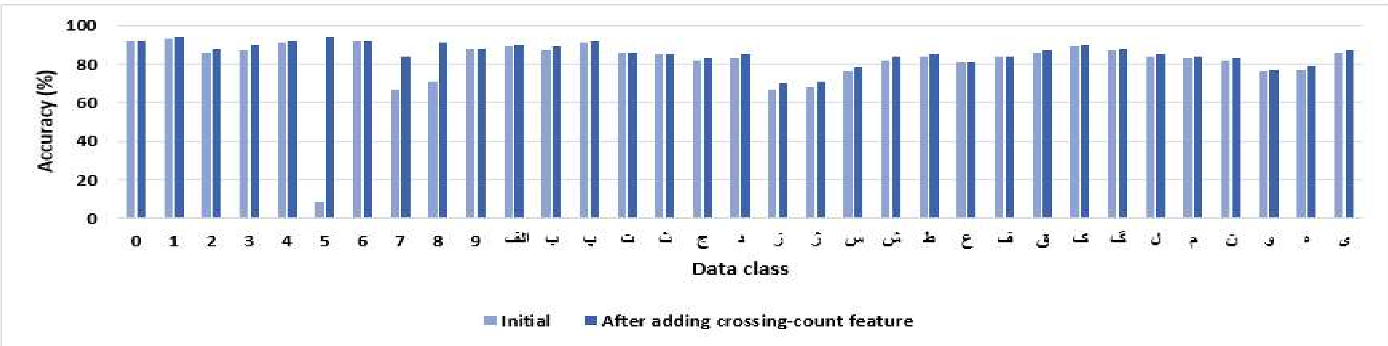


Figure. 3. Evaluating dataset using Zernike moments and crossing counts as extracted features

It is attempted to make changes in the angle, size, resolution, and other features of the augmentation methods. According to the learning algorithm used in the recognition steps, the data annotation is performed in Fig. 5. Accordingly, each of the four components in the Iranian LP is identified by a special label.

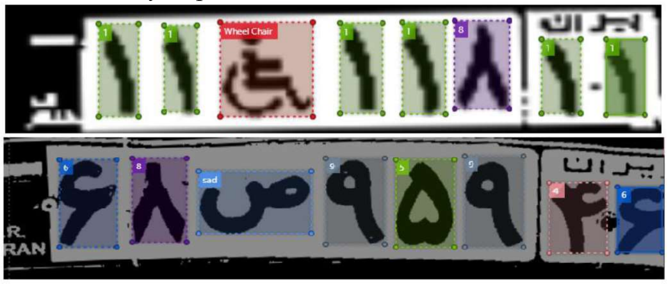


Figure. 5. Annotation of license plates samples.

Since segmentation and extraction of appropriate features lead to low-error model learning, so the hybrid model of YOLO and Tesseract will be efficient. Accordingly, the segmentation, extraction and recognition steps of LPs are combined, so that the deep learning model (YOLOv3) can identify the position of the texts. YOLOv3 use Darknet-53 [26] for feature extraction. Darknet-53 [26] is much more powerful than Darknet-19 [27] but still more efficient thanResNet-101 [28] or ResNet-152 [28]. YOLOv3 uses multi-scale prediction, which means it is detected on multiple scale feature maps. For this reason, the accuracy of target detection is improved. Its structure detail is shown in Fig. 6.

After detection the position of each component of the LPs, characters are recognized. As Fig. 7 shows, first the image of LP is passed into YOLO. Then, YOLO detects the required text regions and crops them out from the image. Later, we pass those regions one by one to Tesseract. Tesseract reads them, and make the result.

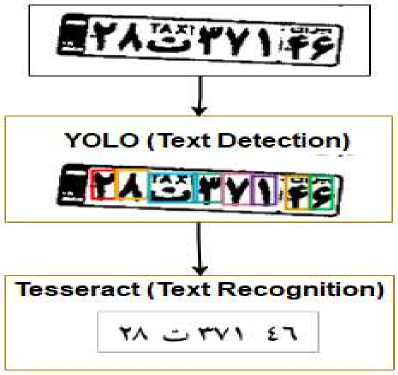
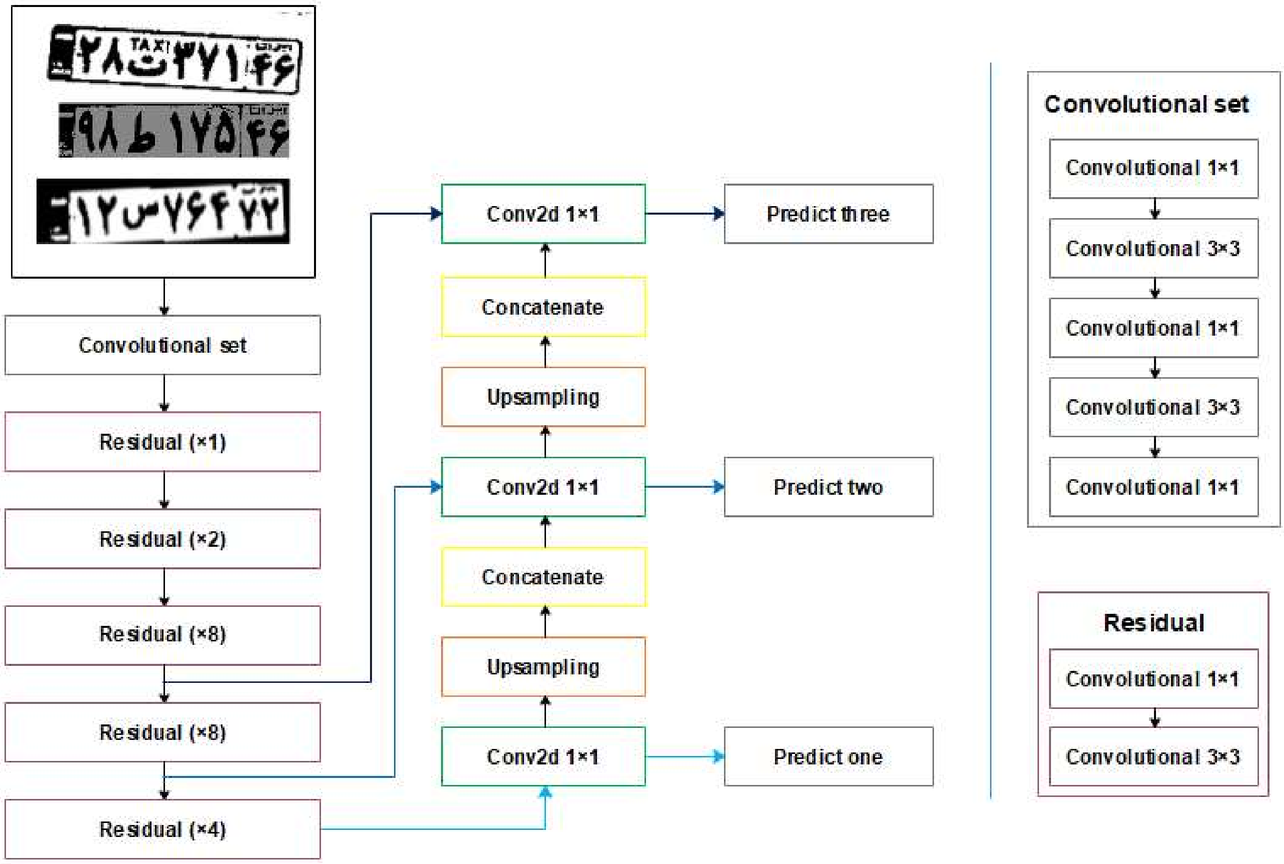


Figure. 7. Architecture of the conventional YOLOv3 framework.



c

Figure. 6. Architecture of the conventional YOLOv3 framework.

Using this method eliminates the challenges of image segmentation and also reduces the time needed for annotation of segmented images since the annotation used at this stage can be useful for text detection in YOLO. On the other hand, the use of Tesseract Engine for text recognition improves ALPR system performance by up to 99.2%. We implemented this project with the Python programming language.

IV. CONCLUSIONS

License Plate (LP) recognition is one of the important applications in the field of the intelligent transportation system. In this study, in order to improve the performance of LP recognition, we have combined two main steps, segmentation and OCR as a single step using deep learning. For the deep learning step, we have used the YOLO framework to detect the LP and Tesseract as OCR engine to LP recognition. This method works well than our previous method on segmented datasets of characters in which Zernike moments and crossing count combinations were used to extracting features.

REF ERENCE

[1] S. Du, M. Ibrahim, M. Shehata, and W. Badawy., "Automatic license plate recognition (ALPR): A state-of-the-art review," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311- 325, 2012.

[2] A. H. Ashtari, M. J. Nordin, and M, Fathy., "An Iranian license plate recognition system based on color features," IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 4, pp. 1690- 1705, 2014.

[3] L. W. Chen and Y. F. Ho, "Centimeter-grade metropolitan positioning for lanelevel intelligent transportation systems based on the internet of vehicles," IEEE Transactions on Industrial Informatics, vol. 15, no. 3, pp. 1474- 1485, 2019.

[4] K. N. Bui and J. J. Jung, "ACO-based dynamic decision making for connected vehicles in iot system," IEEE Transactions on Industrial Informatics, vol. 15, no. 10, pp. 5648- 5655, 2019.

[5] A. K. Keshari, R. Sharma, and M. J. Nigam., "Digitizing physical documents using optical character recognition," In 10th International Conference on Signal Processing Systems, vol. 11071, 2019.

[6] B.A. Fomani and A. Shahbahrami., "License plate detection using adaptive morphological closing and local adaptive thresholding", In 3rd International Conference on Pattern Recognition and Image Analysis, pp. 146-150, 2017.

[7] S. Naz, A. I. Umar, S. B. Ahmed, R. Ahmad, S. H. Shirazi, M. I. Razzak, and A. Zaman., "Statistical feature extraction for character recoginition using recurrent neural network," Pakistan Journal of Statistics, vol. 34, no. 1, 2018.

[8] N. N. Kyaw, G. Sinha, and K. L. Mon., "License plate recognition of myanmar vehicle number plates a critical review," IEEE 7th Global Conference on Consumer Electronics, pp. 771-774, 2018.

[9] F. Damayanti, S. Herawati, F. M. Ayu, A. Rachmad, and others, "Indonesian license plate recognition based on area feature extraction," Telkomnika, vol. 17, no. 2, pp. 620-627, 2019.

[10] M. Rizvi, H. Raza, S. Tahzeeb, and S. Jaffry., "Optical character recognition based intelligent database management system for examination process control," In 16th International Conference on Applied Sciences and Technology, pp. 500-507, 2019.

[11] S. Yu, B. Li, Q. Zhang, C. Liu, and M. Q.-H. Meng, "A novel license plate location method based on wavelet transform and emd analysis," Pattern Recognition, vol. 48, no. 1, pp. 114- 125, 2015.

[12] J. Jagtap, and S. Holambe., "Multi-style license plate recognition using artificial neural network for Indian vehicles," In International Conference on Information, Communication, Engineering and Technology, pp. 1-4, 2018.

[13] A. Saif, M. J. Hossain, M. Uzzaman, M. Rahman, M. Islam, and others., "Real time bangla vehicle plate recognition towards the need of efficient model a comprehensive study," International Journal of Image, Graphics and Signal Processing, vol. 10, no. 12, 2018.

[14] I. Astawa, I. G. N. B. Caturbawa, I. M. Sajayasa, and I. M. Atmaja., "Detection of license plate using sliding window, histogram of oriented gradient, and support vector machines method," Journal of Physics: Conference Series, vol. 953, 2018.

[15] M. Jaderberg, k. Simonyan, a. Vedaldi, and A. Zisserman., "Reading text in the wild with convolutional neural networks," International Journal of Computer Vision, vol. 116, no. 1, pp. 1-20, 2016.

[16] S. M. Silva, and C. R. Jung., "Real-time brazilian license plate detection and recognition using deep convolutional neural networks," In 30th Conference on Graphics, Patterns and Images, pp. 55-62, 2017.

[17] J. Wang, X. Liu, A. Liu, J. Xiao., "A deep learning-based method for vehicle license plate recognition in natural scene," APSIPA Transactions on Signal and Information Processing, vol. 8, 2019.

[18] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet., "Multi-digit number recognition from street view imagery using deep convolutional neural networks," arXiv preprint arXiv:1312.6082, 2013.

[19] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonccalves, W. R. Schwartz, and D. Menotti., "A robust real-time automatic license plate recognition based on the YOLO detector," In International Joint Conference on Neural Networks, pp. 1-10, 2018.

[20] S. M. Silva and C. R. Jung, "License plate detection and recognition in unconstrained scenarios," In Computer Vision, 2018.

[21] Smith, R. "An overview of the Tesseract OCR engine," In 9th International Conference on Document Analysis and Recognition, Vol. 2, pp. 629-633, 2007.

[22] J. Redmon and A. Farhadi, "YOLO9000: Better, faster, stronger," In IEEE Conference on Computer Vision and Pattern Recognition, pp. 6517-6525, 2017.

[23] J. Redmon, S. Divvala, R. Girshick, A. Farhadi., "You only look once: Unified, real-time object detection.," Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 779-788, 2016.

[24] C. Patel, A. Patel, D. Patel., "Optical character recognition by open source OCR tool tesseract: A case study," International Journal of Computer Applications, vol. 55, no. 10, pp. 50-56, 2012.

[25] R. Pan, X. Ma, and L. Wang., "An efficient method for skew correction of license plate," In 2nd International Workshop on Education Technology and Computer Science, pp. 90–93, 2010.

[26] L. Jun, J. Su, and X. Zeng, "A solution method for image distortion correction model based on bilinear interpolation," vol. 43, no. 1, 2019.

[27] D. Bradley and G. Roth, "Adaptive thresholding using the integral image," Journal of Graphics Tools, vol. 12, no. 2, pp. 13- 21, 2007.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ''You only look once: Unified, real-time object detection,'' IEEE Conference on Computer Vision and Pattern Recognition, pp. 779- 788, 2016.

[29] J. Redmon, and A. Farhadi, ''YOLOv3: An incremental improvement,'' 2018, [Online]. Available: https://arxiv.org/abs/1804.02767

[30] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' IEEE Conference on Computer Vision and Pattern Recognition, pp. 770- 778, 2016.