International Journal of Scientific & Engineering Research Volume 12, Issue 10, October-2021
ISSN 2229-5518

Detection and Recognition System for Car Number License Plates using Convolution Neural Networks

Bao Thai Duong

Abstract— Automatic license plate recognition (ALPR) system is intelligent system to find the exact license plate by analyzing image and video data. The ALPR has three key steps: location of the license plate (LP), segmentation and recognition of optical character (OCR). In real conditions, each step requires different techniques, and each technique has its specific characteristics. The process of Vehicle License Plates (VLPR) is difficult because of changes in the view, shape, color, multiple formats and uniform lights when you acquire pictures. After segmentation algorithms, LP localization techniques will detect the LP and isolate each character from the other. Finally, the OCR step is used to recognize the individual characters. The overall reliability depends on the correctness of each step. Both segmentation and OCR steps are performed as single stage by methods of Deep Learning to improve the OCR step performance. The proposed method uses the YOLO Framework and OCR Tesseract tool to achieve plate locale and character segmentation. Test results proved the effectiveness of 99.2 percent compared to previous works has improved substantially.

Index Terms— Intelligent Transformation System, Optical character recognition, Deep learning, YOLO, Covolutional Neural Network.

1 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). There are many potential applications for the ALPR system such as security closed – circuit television (CCTV), speed control stations, vehicle parking, traffic management, toll enforcement, and many other applications.

The ALPR has three key steps: location of the license plate (LP), segmentation and recognition of optical character (OCR). Only the number plate is detected from the entire input image and processed further in the next step of character segmentation and recognition. Detecting the number plate is a non-trivial task, mainly due to the differences in the different types and formats of numerical plates... This failure thus affects the precision of the segmentation and recognition of the character.
The recognition of LP of vehicles, as we know, has challenges, because the input image quality affects the result of text recognition. Pre-processing can therefore enhance efficiency. We must address the challenges of image quality like shade and noise in different ways at various phases of preprocessing. In order to do this, methods like bug correction, binarization of images or even the deletion of unnecessary parts often are used [3]. Of course, each of the above methods has its own methods, which apply through relation to the noise in the original image.
Feature extractions are one of the most important steps in text recognition, allowing the classifier to distinguish different classes with good accuracy by applying the appropriate feature. Obviously, it should also consider the size of the feature vectors. Researchers use a variety of LP recognition features based on statistical [7] size and form [8], area [9], color [2], Scale Invariant Transform Feature (SIFT) [10] and other features. [11].

Artificial Neural Network (ANN) is a prominent feature in learning-based approaches for character recognition. ANN is used to detect the features after they have been extracted from the characters [12]. A certain number of layers and neurons are typically used to train the ANN to recognize characters on the test dataset. Of course, different methods for learning the network, such as feedforward backpropagation, feedback, and feedback self-learning, can be used [13].

Deep learning (DL) techniques enable the automatic selection of image features. Deep learning can be used to extract features and classifications based on this [15]. Deep learning techniques are now commonly used in most computer vision tasks, particularly methods that use some form of Convolutional Neural Network (CNN) to achieve the most advanced performance [16], [17]. Unlike the previous methods, we use DL methods to identify each of the LP character regions and then use the Tesseract engine to identify the characters.
The main purpose of this paper is to obtain a specific and reliable license plate localization and recognition system for Vietnamese's car license plate, mainly for parking security. This paper presents a CNN-based method for high accuracy, real-time car license plate.

This paper is organized as follows. Section 2 discusses an overview of the related works in section 2. A description of our approach for classification of vehicles is presented in section 3. Section 4 details the experimentation carried out on traffic videos along with the results. Finally, conclusions are given in section 5.

2 Related Works

2.1 LP detection techniques

Detecting the car before the plate is a common strategy in the LPD pipeline, aiming to reduce the search area and the number of false positives. Over the past decade, most research in the subject of license plate detection techniques has emphasized the method of image processing in the computer vision. Dhar et al. [15] projected a system deployment for LP recognition under the application of edge detection and CNN. The model was consumed with character segmentation in the form of pre-processing phase for LP analysis. In case of character segmentation, the newly developed technique finds edge prediction, morphological task and regional features. Thus, it is more effective for images with elegant backgrounds, but the images showed no impact by the comprised complexities. In 2016, Ullah et al. [17] concentrated on predicting LP based on the mathematical morphological attributes. The newly presented model could work every English LP that differs in shape and structure.

In recent times, a robust ability as well as discriminating energy of DL methods along with few technologies were developed with various DL approaches for LP recognition. Deep Learning approaches typically explore CNN models. Many of them use the successful YOLO architecture, presented by Redmon et al. [16], that performs detection and recognition in a single pass (originating the name You Only Look Once). A few methods using YOLO-based models for LP.

In Hsu et al. [22], the authors trained a YOLO network to directly recover the LPs instead of primarily detecting vehicles. In order to face the problem of detecting small objects, they opted to enlarge the network output granularity. However, their paper lacks information about the quality of the bounding boxes extracted, since they do not provide the IoU considered in the evaluation. Moreover, they did not evaluate the performance when vehicles are far from the camera, which generates small LPs.

Laroca et al. [21] employed two YOLO networks to detect vehicles and LPs, respectively, using an internal training set containing annotated car bounding boxes that was made publicly available after the publication. These approaches explore two different networks to perform LPD, requiring specialized datasets with extra annotation content (e.g., cars and LP rotation angle). This increases the training complexity and amount of memory required to store the networks – which, for instance, might be relevant in mobile applications.

In Dong et al. [16] was develop a network to detect license plates recovering all its 4-corner points. Such kind of detection has the advantage of permitting rectification of the detected license plates, which eases the work of the OCR. Due to the complexity of the task, the network was trained with more than 18k images from a private dataset.

2.2 License Plate Segmentation

A common approach for character segmentation is the use of binarization method to separate the foreground from the background, following the application of connected components analysis to extract the bounding boxes. This is highly dependent on the chosen binarization method and nonuniformity of the plate (e.g., due to textured background, low contrast between the background and the characters, or noneven illumination).

Recent LPS approaches also employ CNNs. In Selmi et al. [20] CNNs were used to filter and recognize character candidates after a segmentation process. Laroca et al. [21] used a YOLO-based network specially trained for LPS, not performing recognition. They argue that making the detection and recognition process together can worsen the results for the Brazilian scenario, where letters and digits occupy fixed positions inside the LP.

2.3 Character Recognition

Character recognition is a crucial task in ALPR systems since an error in a single character can invalidate the entire LP. 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]. In our approach, we changed the training parameters and by using image processing techniques we could achieve much better results.

3 PROPOSED MODEL

In this section, we described in detail how we detected plates from images and the usage of CRNN to recognize the plates. In previous sections, we divided the job into two main parts and then we prepared about images to train YOLOv4 [28] network to identify the position of a license plate. Basically, YOLO completes several algorithms at once including localization, orientation, sizing and segmentation with image processing.

After receiving the extracted plates, we perform preprocessing, LP extraction and recognition operations. As mentioned earlier, the extracted plates LPs have many problems such as size, un-normal, tilt, non-uniform brightness, and some other noises. To overcome all the problems, using the pre-processing stage before segmentation algorithms.

In addition, following segmentation, they must be annotated, which is a time-consuming job. So, we try to utilize an approach that combines segmentation and annotating.

Because the picture quality of all the plates differs, the extracted plate is tilted or shadowed due to the camera's viewing angle or the light's angles. The pre-processing of plates is the initial stage in the suggested procedure. In this phase, we correct the tilting of the LPs using Hough line transform methods [23], then apply rotation filters as needed to fix the tilting image using the bilinear interpolation approach [24]. To remove the shadows, the Bradley method [25] is applied. The noise in the image of the LPs has been considerably decreased as a result of the pre-processing.

The 1000 genuine samples taken from surveillance cameras were used in this study. On the other hand, we know that a great amount of data is required to train deep learning techniques, which is one of the most significant problems. We employ augmentation methods on pre-processed photos to overcome this difficulty and to complement and multiply the dataset's numbers. Learning the model becomes easier as a result of this. Given that many of the image pre-processing problems have been addressed.

Changes to the angle, size, resolution, and other aspects of the augmentation methods are attempted. The data annotation is performed in Fig. 5 according to the learning method employed in the recognition steps. As a result, each of the four components of the Iranian LP is labeled separately.

Because segmentation and extraction of relevant features result in low-error model learning, the YOLO and Tesseract hybrid model will be effective. As a result, the LP segmentation, extraction, and recognition phases are merged such that the deep learning model (YOLOv4) can determine where the texts are located. Darknet-53 [26] is used by YOLOv4 to extract features. Darknet-53 [26] is much more powerful than Darknet-19 [27], yet ResNet-101 [28] and ResNet-152 [28] are still more efficient. Multi-scale prediction is used by YOLOv4, which implies it is identified on several scale feature maps. As a result, the target detection accuracy has improved. Figure 6 depicts the structure in further detail.

Characters are recognized when the position of each component of the LPs has been determined. The image of LP is first transmitted into YOLO, as seen in Fig. 7. The relevant text portions are then detected by YOLO and cropped out of the image. Later, we cross through each of those places one by one on our way to Tesseract. Tesseract examines them and determines the outcome.

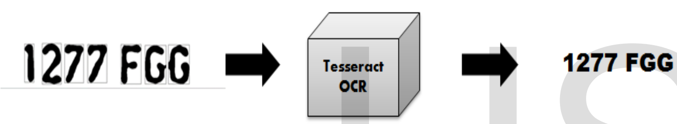

Fig. 1. Character recognition using Tesseract.

This method solves the problems of picture segmentation while also shortening the time required for annotation of segmented images because the annotation employed at this step can be useful for text identification in YOLO. The usage of Tesseract Engine for text recognition, on the other hand, boosts ALPR system performance by up to 99.2 percent. We used the Python programming language to complete this project.

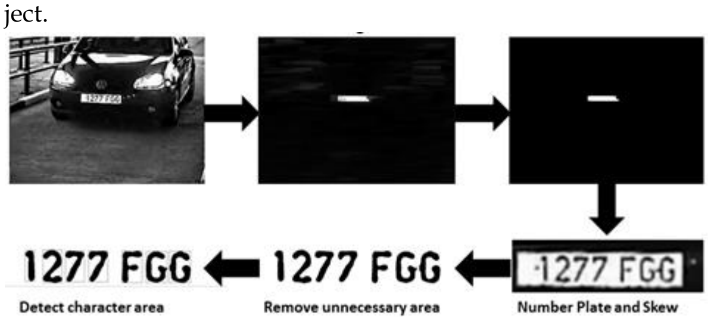

Fig. 2. Extract number of license plate.

4 Experimental Result and Discussion

4.1 Dataset

For experiments, we used half real, half synthetic data. First, we divided the dataset into two parts by randomly selecting training and test images as 85% to 15% respectively. We acquire YOLOv4 with 9500 different quality and background view images. However, for training the CRNN model we didn't require to have image of a car, but only license plate Using YOLO we cropped a license plate from the car on image and fed that plate to CRNN. The detection accuracy using YOLO was quite impressive with 99.8%.

4. 2 Training phase

All trainings were carried out on a Windows computer with a CPU Intel Core i5-7600, 3.50 GHz, and a GPU Nvidia GeForce GTX 1080Ti with 16 GB RAM. To train the YOLOv4 network, we utilized YOLOv4-416, which converts images to 416 x 416 pixels, with learning rate a =1e3, batch size = 64 on frozen layers and batch size = 8 on unfrozen layers, and Adam optimizer with a =13 for frozen layers and a =14 after unfreezing the layers [2]. Training in YOLO is more time consuming, and it takes roughly 25 hours for a network to reach its convergence.

Table 1. Recognition accuracy (%).

Method	Type 1	Type 2	Type 3	Type 4	Average
CNN	95.6	94.9	66.2	92.9	87.4
CNN and color inversion	99.7	99.4	79.6	99.5	94.6

As can be seen, inverting the colors improves recognition accuracy overall. Because the image quality is poor when we extract plates from Tesseract, and because there were many dark plates, inverting the color simply reverses the RGB values and focuses better sequence learning.

5 Conclusions

License Plate (LP) recognition is a significant application in the field of intelligent transportation systems. In this study, we used deep learning to integrate two major phases, segmentation and OCR, in order to increase the performance of LP recognition. For the deep learning step, we used the YOLO framework to identify the LP and Tesseract as an OCR engine to recognize the LP. This strategy outperforms our earlier method on segmented datasets of characters in which Zernike moments and crossing count combinations were utilized to extract features.

References

[1] Choo Kar Soon, Kueh Chiung Lin, Chung Ying Jeng and Shahrel, "*Malaysian Car Number Plate Detection and Recognition System*", Australian Journal of Basic and Applied Sciences, 6(3): 49-59, ISSN 1991-8178.

[2] Clausner, C., Antonacopoulos, A. & Pletschacher, "*Efficient and effective OCR engine training*", IJDAR 23, 73–88 (2020).

[3] Christian Clausner, Apostolos Antonacopoulos and Stefan Pletschacher. Efficient and effective OCR engine training, "*International Journal on Document Analysis and Recognition*", (IJDAR) volume 23, pages73–88(2020).

[4] Chung-cheng chiu, Min-yu ku và Chun-yi wang (2010), "*Automatic Traffic Surveillance System for Vision-Based Vehicle Recognition and Tracking*", Journal of information science and engineering, 26, 611-629.

[5] De Sousa Matos, F.M and R.M.C.R de Souza (2012), ”An image vehicle classification method based on edge and PCA applied to blocks”, Proceedings of the 2012 IEEE International Conference on Systems, Man and Cybernetics (SMC), October 14-17, 2012, IEEE, Seoul, South Korea, ISBN: 978-1-4673-1713-9, pp: 1688-1693.

[6] Hamad K , Kaya M . “*A Detailed Analysis of Optical Character Recognition Technology*”, International Journal of Applied Mathematics Electronics and Computers. 2016; (Special Issue-1): 249-244.

[7] P. N. Pavan Kumar, Mohit Jain, A. Pavan, M. Pramod Bhat and C (2013). Manjunath, “*Intelligent Collaborative Surveillance System”*”, 2013 International Conference on Advances in Technology and Engineering (ICATE). Mumbai, India: IEEE Publisher.

[8] Patel. Chirag, Patel. Atul and Patel. Dharmendra (2012), "*Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study*", International Journal of Computer Applications, vol. 55, issue 10, pp. 50-56.

[9] Robert Fisher, Simon Perkins, Ashley Walker, Erik Wolfart (2003), Hypermedia image processing reference (HIPR).

[10] Ray Smith, “An Overview of the Tesseract OCR Engine”, Document Analysis and Recognition, 2007.

[11] Sahar S. Tabrizi, Nadire Cavus (2016),”*A hybrid KNN-SVM model for Iranian license plate recognition*”, 2th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria.

[12] Tesseract OCR, https://github.com/tesseract-ocr. Accessed 20 April 2021.

[13] Tan, Chew Lim, et al 2002, “*Imaged document text retrieval without OCR*”, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 838-844.

[14] Vinh Du Mai, Duoquian Maio, Ruizhi Wang, Hongyn Zhang (2011),” *An Improved method for Vietnam License Plate Location*”, Segmentation and Recognition. 2011 International Conference on Computational and Information Sciences.

[15] P. Dhar, S. Guha, T. Biswas, and M. Z. Abedin, ‘‘*A system design for license plate recognition by using edge detection and convolution neural network*,’’ in Proc. Int. Conf. Comput., Commun., Chem., Mater. Electron. Eng. (IC4ME2), Feb. 2018, pp. 1–4.

[16] M. Dong, D. He, C. Luo, D. Liu, and W. Zeng, ‘‘A CNN-based approach for automatic license plate recognition in the wild,’’ in Proc. Brit. Mach. Vis. Conf., 2017, pp. 1–12.

[17] I. Ullah and H. J. Lee, ‘‘*An approach of locating Korean vehicle license plate based on mathematical morphology and geometrical features*’’ in Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI), Dec. 2016, pp. 836–840

[18] S. S. Omran and J. A. Jarallah, ‘‘*Iraqi car license plate recognition using OCR*,’’ in Proc. Annu. Conf. New Trends Inf. Commun. Technol. Appl. (NTICT), Mar. 2017, pp. 298–303

[19] Nguyen, T.; Nguyen, D, “*A New Convolutional Architecture for Vietnamese Car Plate Recognition*”, In Proceedings of the 2018 10th International Conference on Knowledge and Systems Engineering (KSE), Ho Chi Minh City, Vietnam, 1–3 November 2018; pp. 7–12

[20] Z. Selmi, M. Ben Halima, A.M. Alimi, “*Deep learning system for automatic license plate detection and recognition*”, in 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), IEEE, 2017, pp. 1132–1138, https://doi.org/10.1109/ICDAR.2017.187

[21] R. Laroca, E. Severo, L.A. Zanlorensi, L.S. Oliveira, G.R. Gonalves, W.R. Schwartz, D. Menotti, “*A robust real-time automatic license plate recognition based on the yolo detector* ”, International Joint Conference on Neural Networks (IJCNN) 2018 (2018) 1–10, https://doi.org/10.1109/IJCNN.2018.8489629

[22] J. Redmon and A. Farhadi, ”YOLO 9000: 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, “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, 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, 2017.

[28] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection”, Computer Vision and Pattern Recognition, 2020.