ROI-BASED DEEP LEARNING METHOD FOR VARIABLE-LENGTH LICENSE PLATE CHARACTER SEGMENTATION

Dan Wang[1], Bingshu Wang[2,], Yang Chen [1]*

[1] Department of Mathematics, Faculty of Science and Technology
University of Macau, Avenida da Universidade, Taipa, Macau, 999078 China.
Email: bohewan@126.com, yayangchen@umac.mo
[2] Department of Computer and Information Science, Faculty of Science and Technology
University of Macau, Avenida da Universidade, Taipa, Macau, 999078 China.
Email: yb77408@umac.mo

ABSTRACT

License plate character segmentation plays link function between license plate detection and recognition. It is based on the results of license plate detection and produces segmented characters for subsequent recognition module. Character is the region of interest (ROI) in the license plate. Previous works focus on fixed-length license plates and face challenging on the slant plates. To solve the problem, we propose a ROI-based deep learning method for variable-length license plate character segmentation. It treats all the characters as objects and converts character segmentation as object detection. This paper exploits Faster R-CNN to detect the possible character regions. Region Proposal Network is designed to provide sufficient proposals for character detection of license plate and full connected network can modify the candidate boxes and predict the character class simultaneously. Moreover, we create a dataset for character segmentation of license plate. Experimental results demonstrate that our method can achieve a high accuracy in three scenes compared with some state-of-the-art approaches.

Index Terms— ROI, Character segmentation, Faster R-CNN, License plate recognition

1. INTRODUCTION

Generally, intelligent license plate recognition [1, 2] can be divided into three parts: license plate detection [3, 4], character segmentation [5, 6, 7, 8], and character recognition [9, 10]. License plate character segmentation is part of license plate recognition system. It is an important between license plate detection and character recognition. Handling this problem is fundamental to perform important tasks in the field of intelligent transport systems [2, 11, 12] such as electronic toll collection and counterfeit plates detection.

—————————
Corresponding author: Bingshu Wang.

Fig. 1. Some examples of Macau license plates.

In the past decades, many researchers around the world have proposed different approaches to segment characters. Those methods can be classified as: projection analysis-based methods [13, 14, 15, 16], MSER-based segmentation methods [2, 17], connected components analysis-based methods[6], prior knowledge-based methods [18, 19] and template matching-based methods [5, 20]. Most of them focus on the local license plates that have a standard and a fixed-length number [6, 19, 5]. Our observation is that variable-length license plates are used in one region or country. For example, Fig. 1 presents some Macau license plates which have different number of characters and ratios.

In practice, license plates are always influenced by two factors: plate variations and environment variations [1]. Plate variations mainly includes rivet, text font, rotation, borders, etc. Environment variations mainly includes illumination changing and weather difference, for example, vehicle's headlights and hazed days. Thus, character segmentation faces big challenges in practical application.

Recently, deep learning-based techniques are exploited to segment characters[21, 22]. Abdussalam et al [21] proposed to use a Darknet-18 network for the segmentation of Chinese license plates with only 7-characters. Laroca et al [22] exploited a CR-NET [23] for character segmentation. Howev-
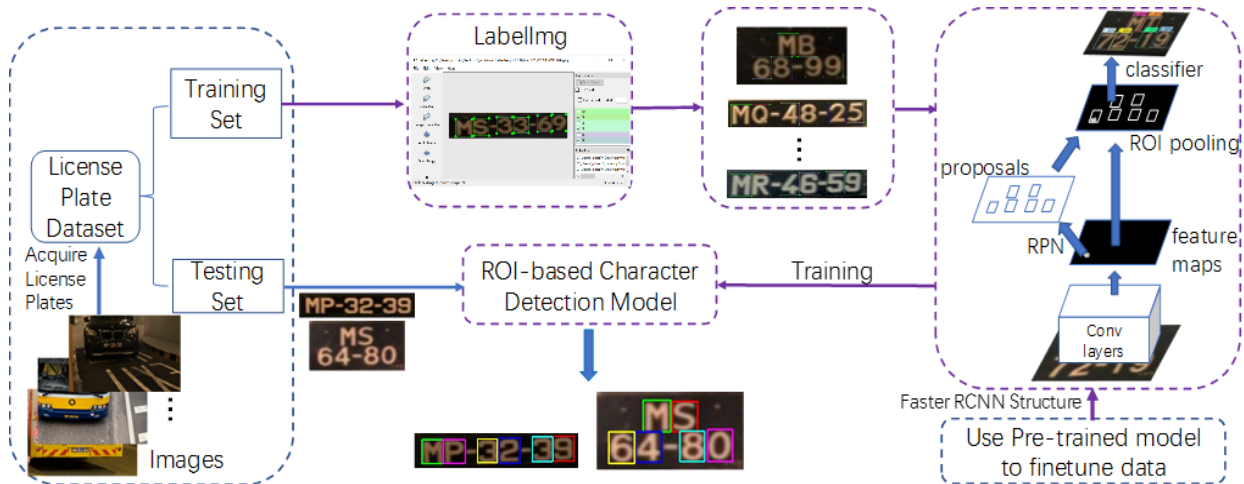
Fig. 2. The flowchart of ROI-based character segmentation.

er, they are only designed to segment fixed-length plates and lacks of flexibility, and they also require fixed-size images as inputs of Networks. In the real-world applications, the segmentation methods of variable-length license plates are required in most cases.

In this paper, we show that ROI-based deep learning model has significant performance on the character segmentation. We convert a traditional character segmentation problem into an object detection problem. Faster R-CNN [24] is chosen as a representative of ROI-based models due to its excellent detection accuracy on multiple datasets. ROI-based model can produce the detection boxes and object classes simultaneously. For segmentation, the boxes are our expected ROIs and can be analyzed . In addition, we create a dataset of character segmentation for training of ROI-based deep learning techniques [24, 25]. Experiments demonstrate the effectiveness of ROI-based character segmentation models.

2. THE PROPOSED METHOD

Figure 2 illustrates the flowchart of ROI-based character segmentation method. The first step is to build a dataset of license plate character segmentation with ground-truths manually. Then, these labeled images are fed into ROI-based model for training. In this paper, we choose Faster R-CNN as a typical representative. As a result, the trained model is used to test new samples to produce segmentation results.

2.1. Label Characters

To our knowledge, few character segmentation datasets are available so far. Thus, we build a dataset of license plate character segmentation. The license plates are collected from Macau. Figure 1 shows the characteristics of Macau license

plates. In Macau, all the license plates are printed by letters and digits. Thus, the total number of character classes is 36 (26 classes of letters and 10 classes of digits). We mark every character with a corresponding class by a tool of $LabelImg$.

2.2. ROI-based Character Segmentation

We try to solve the problem of license plate character segmentation through a ROI-based deep learning method. Unlike traditional segmentation methods using projection, template matching, edge detection or other techniques, the main idea of this paper is to detect all the characters of one plate through convolutional neural network. It is able to handle variable-length license plates without any prior knowledge.

In this paper, Faster R-CNN is selected as the representative of ROI-based models due to its high accuracy in the detection of Pascal VOC and COCO [26] datasets, which mainly includes four steps as following.

(1) Conv layers: Faster R-CNN uses a groups of $conv + relu + pooling$ layers to extract feature maps, which are shared by subsequent region proposal network and fully connected layers.

(2) Region Proposal Networks (RPN): this network is used to generate sufficient and potential region proposals for classification. It consists of two items: box-regression layer and box-classification layer. Box-regression layer is used to modify the anchors in order to get accurate locations of proposals. Box-classification layer is used to recognize a box to be foreground or background.

(3) ROI-pooling: this layer collect the features from Conv layers and the proposals from RPN. It converts different size of proposals into an unified form then feds the features into subsequent full connected layer for predicting classes.
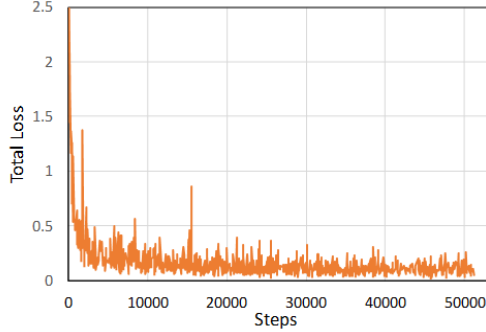
(4) Classification: It receives proposal feature map and

Fig. 3. The training loss of Faster RCNN for character segmentation and recognition model.

outputs the probability of classes through full connected layers and softmax layers. Meanwhile, it makes use of bounding box regression to adjust the locations of proposals to get more accurate boxes.

It should be noted that Region Proposal Network is very effective on the character detection. AdaBoost algorithm utilizes sliding windows and image pyramid to generate candidate regions and R-CNN uses Selective Search to generate proposals, both of which are time-consuming. Faster R-CNN exploits RPN directly to provide candidate regions, which improve the detection speed significantly.

2.3. Training Process

In this paper, we train our ROI-based character segmentation model by tuning a model "faster_rcnn_inception_v2" which is pre-trained on COCO dataset. The pre-trained model can be obtained from the website (https://github.com/tensorflow/models). During the training stage, a learning rate of 0.0002 and a momentum of 0.9 are adopted. The implementation uses Tensorflow. The total loss changing process is shown in Fig. 3. It can be seen from the figure that the training reach a convergence quickly. After 68630 steps training, we get a ROI-based model for character segmentation.

3. EXPERIMENTAL RESULTS

3.1. Dataset and Evaluation Metric

To verify the effectiveness of ROI-based character segmentation model, we conduct experiments on Macau license plates. The proposed method is compared with some state-of-the-art approaches: projection-based [15], MSER-based [17], CCA-based [6]. They are typical representatives in the filed of license plate character segmentation.

We collect license plate images from parking lots and stations in Macau to build a dataset. The images are divided into two parts: training set and testing set. Training dataset has
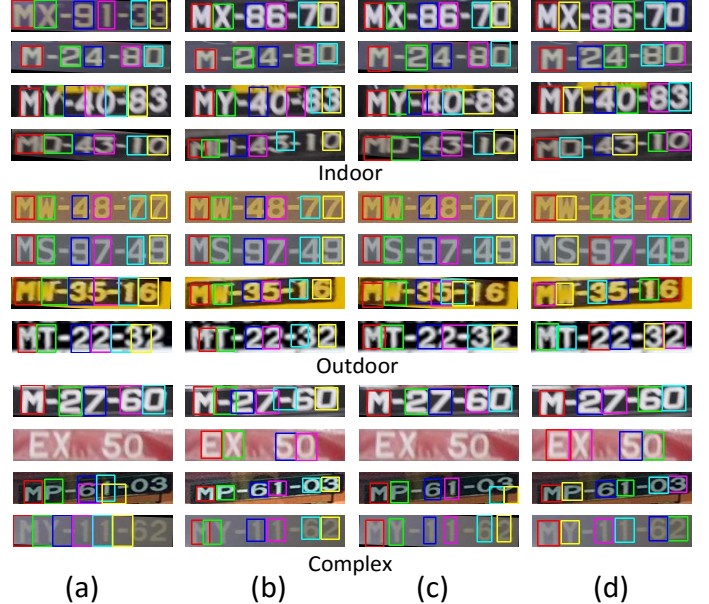

Fig. 4. Visual comparison of single line license plates. (a) Projection-based method [15], (b) MSER-based method [17], (c) CCA-based method [6], (d) Our method.

2200 license plate images. They are all labeled by a tool of $LabelImg$. Testing images involves three parts: $Indoor$ with 1000 images, $Outdoor$ with 830 images, and $Complex$ with 404 images. The images of $Indoor$ are captured from parking lots. The images of $Outdoor$ are collected from station and campus. The images of $Complex$ are captured with complex conditions such as strong illuminations and large slant, which are full of challenge.

Evaluation metric of character segmentation refers to Yang [17]:

$$Accuracy = \left(1 - \frac{errors}{total}\right) \times 100\% \qquad (1)$$

where $errors$ represents the number of license plates that are falsely segmented and $total$ represents the number of all license plates. Once any character in a license plate is segmented incorrectly, it will be regarded as a failure.

3.2. Visual Results

Visual results including $Indoor$, $Outdoor$, and $Complex$ are provided in Fig. 4 and Fig. 5. Figure 4 presents the comparisons of single line plates and Fig. 5 gives the comparisons of double line plates.

It can be seen from Fig. 4 that all the methods can generate good results when license plates are captured with good conditions, as shown in the first two rows of $Indoor$ part and $Outdoor$ part. When characters are connected with other objects (e.g., the yellow object in the third row), the compared
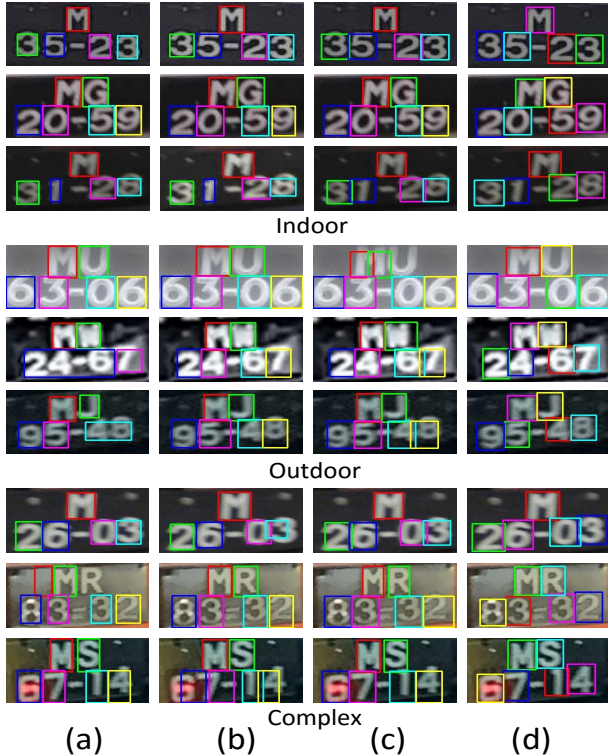
Indoor

Outdoor

Complex

(a) (b) (c) (d)

Fig. 5. Visual comparison of double line license plates. (a) Projection-based method [15], (b) MSER-based method [17], (c) CCA-based method [6], (d) Our method.

approaches may fail. It can affect the histogram projection of [15] and the connected region analysis of [6]. When the characters are touching, MSER-based approach [17] might keep the touching parts while ignoring independent characters. Our ROI-based model can overcome these problems well thanks to the proposals provided by RPN.

Since the $Outdoor$ images are affected by borders and hyphens, Projection-based approach [15] easily takes them into account and generates inaccurate boxes (see the third row in the $Outdoor$ and $Complex$ parts in Fig. 4). When the characters are independent and the background is homogenous, MSER-based approach [17] still provides good segmentation results, which can be seen in the first three rows in the $Outdoor$ part in Fig. 4.

The $Complex$ part includes touching characters, variable-length license plate, borders, obvious slant, and blurring characters. The compared methods cannot handle these situations. For the second row of $Outdoor$ part Fig. 4, the methods of [15] and [6] can not segment characters. On the contrary, our method can provide satisfactory results.

Figure 5 illustrates the visual comparisons of double line plates. Our method outperforms the compared approaches in multiple situations including slant, complex backgrounds, touching characters, and non-homogeneous illuminations.

Table 1. Quantitative comparison of the proposed method and some state-of-the-art approaches.

Scene	Methods			
	Projection-based [15]	*MSER-based [17]*	*CCA-based [6]*	*Our method*
Indoor	97.60%	89.50%	96.50%	**99.20%**
Outdoor	85.54%	92.41%	90.60%	**96.75%**
Complex	79.95%	91.58%	88.12%	**96.04%**

Fig. 6. Some segmentation failures.

3.3. Quantitative Comparisons

We present a quantitative evaluation for four methods in Table 1. For $Indoor$ set, Projection-based approach achieves a 97.6% accuracy among the compared methods. For $Outdoor$ and $Complex$ sets, the Projection-based method and CCA-based method decrease obviously while MSER-based method keeps steady in the accuracy. More variable-length license plates and touching characters are included in $Indoor$ scene and they are usually touching. Since MSER-based method has difficulty in dealing with them, thus the accuracy in $Indoor$ scene is lower than those of $Outdoor$ and $Complex$. Notably, our method achieves impressive results. For $Indoor$ scene, it reaches a high accuracy of 99.20%. Although the accuracy decreases for challenging $Outdoor$ and $Complex$ scenes, our method is still able to obtain over 96%.

We have to say that there are some situations that are hard to dealt with. Figure 6 gives several failure samples. For those extreme blurring and large slant images, our method faces challenges. A potential solution to solve this problem is to add more samples into training set. In addition, ROI-based deep learning method require much computation cost. Therefore, more optimization needs to be done in the future.

4. CONCLUSION

In this paper, we propose a ROI-based deep learning method for character segmentation of license plate. It exploits a series of convolutional operations to generate feature map, and it uses Region Proposal Network to provide candidate regions, which ensures sufficient boxes for characters. Our method can deal with multiple situations such as slant, complex backgrounds, touching characters, and non-homogeneous illuminations. Moreover, we create a dataset of license plate character segmentation. Experiments conducted on the dataset in-

dicate that our training model achieves higher accuracy than some state-of-the-art methods. For the complex image set, it can achieve more than 96% accuracy and is expected to be promising for application. Our method can be applied to other field of character segmentation, for example, providing a participant recognition for Marathon race.

Acknowledgment

5. REFERENCES

[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] G. Hsu, J. Chen, and Y. Chung, "Application-oriented license plate recognition," *IEEE transactions on vehicular technology*, vol. 62, no. 2, pp. 552–561, 2012.

[3] Y. Yuan, W. Zou, Y. Zhao, X. Wang, X. Hu, and N. Komodakis, "A robust and efficient approach to license plate detection," *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1102–1114, 2016.

[4] H. Xiang, Y. Yuan, Y. Zhao, and Z. Fu, "License plate detection based on fully convolutional networks," *Journal of Electronic Imaging*, vol. 26, no. 5, pp. 053027, 2017.

[5] L. Miao, "License plate character segmentation algorithm based on variable-length template matching," in *2012 IEEE 11th International Conference on Signal Processing*. IEEE, 2012, vol. 2, pp. 947–951.

[6] G. R. Gonccalves, S. P. G. da Silva, D. Menotti, and W. R. Schwartz, "Benchmark for license plate character segmentation," *Journal of Electronic Imaging*, vol. 25, no. 5, pp. 053034, 2016.

[7] P. Shivakumara, A. Konwer, A. Bhowmick, V. Khare, U. Pal, and T. Lu, "A new gvf arrow pattern for character segmentation from double line license plate images," in *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*. IEEE, 2017, pp. 782–787.

[8] P. Kim, K. Lim, and D. Kim, "Learning based character segmentation method for various license plates," in *2019 16th International Conference on Machine Vision Applications (MVA)*. IEEE, 2019, pp. 1–6.

[9] C. Chang and C. Lin, "Libsvm: A library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, pp. 27, 2011.

[10] C. L. Chen, Z. Liu, and S. Feng, "Universal approximation capability of broad learning system and its structural variations," *IEEE transactions on neural networks and learning systems*, vol. 30, no. 4, pp. 1191–1204, 2018.

[11] B. Wang, W. Zhu, S. Tang, Y. Zhao, and W. Zou, "Background subtraction using dual-class backgrounds," in *2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, 2016.

[12] B. Wang, Y. Zhao, and C. L. P. Chen, "Moving cast shadows segmentation using illumination invariant feature," *IEEE Transactions on Multimedia*, vol. Early Access, 2020.

[13] C. Wu, L. On, C. Weng, T. Kuan, and K. Ng, "A macao license plate recognition system," in *2005 International Conference on Machine Learning and Cybernetics*. IEEE, 2005, vol. 7, pp. 4506–4510.

[14] C. Wu and K. Ng, "A high accurate macau license plate recognition system [j]," *Journal of Macau University of Science and Technology*, vol. 2, no. 1, pp. 23–28, 2008.

[15] C. Pun and W. Ho, "An edge-based macao license plate recognition system," *International Journal of Computational Intelligence Systems*, vol. 4, no. 2, pp. 244–254, 2011.

[16] B. Wang and C. P. Chen, "License plate character segmentation using key character location and projection analysis," in *2018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*. IEEE, 2018, pp. 510–514.

[17] X. Yang, Y. Zhao, J. Fang, Y. Lu, Y. Zhang, and Y. Yuan, "A license plate segmentation algorithm based on mser and template matching," in *2014 12th International Conference on Signal Processing (ICSP)*. IEEE, 2014, pp. 1195–1199.

[18] F. Yang, Z. Ma, and M. Xie, "A novel approach for license plate character segmentation," in *2006 1st IEEE Conference on Industrial Electronics and Applications*. IEEE, 2006, pp. 1–6.

[19] W. Ho and C. Pun, "A macao license plate recognition system based on edge and projection analysis," in *2010 8th IEEE International Conference on Industrial Informatics*. IEEE, 2010, pp. 67–72.

[20] J. Tian, R. Wang, G. Wang, J. Liu, and Y. Xia, "A two-stage character segmentation method for chinese license

plate," *Computers & Electrical Engineering*, vol. 46, pp. 539–553, 2015.

[21] A. Abdussalam, S. Sun, M. Fu, H. Sun, and I. Khan, "License plate segmentation method using deep learning techniques," in *International Conference On Signal And Information Processing, Networking And Computers*. Springer, 2018, pp. 58–65.

[22] R. Laroca, E. Severo, L. Zanlorensi, L. 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.

[23] Sergio Montazzolli Silva and Claudio Rosito Jung, "Real-time brazilian license plate detection and recognition using deep convolutional neural networks," in *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. IEEE, 2017, pp. 55–62.

[24] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.

[25] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, and S. Guadarrama, "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7310–7311.

[26] T. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014, pp. 740–755.