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.
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 \textit{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 feeds the features into subsequent full connected layer for predicting classes.

4. Classification: It receives proposal feature map and
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 field 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 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]:

\[
\text{Accuracy} = \left(1 - \frac{\text{errors}_{\text{total}}}{\text{total}}\right) \times 100\% 
\]  

where \(\text{errors}\) represents the number of license plates that are falsely segmented and \(\text{total}\) 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
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 homogeneous, 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>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Indoor</td>
<td>97.60%</td>
<td>89.50%</td>
<td>96.50%</td>
<td>99.20%</td>
</tr>
<tr>
<td>Outdoor</td>
<td>85.54%</td>
<td>92.41%</td>
<td>90.60%</td>
<td>96.75%</td>
</tr>
<tr>
<td>Complex</td>
<td>79.95%</td>
<td>91.58%</td>
<td>88.12%</td>
<td>96.04%</td>
</tr>
</tbody>
</table>

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.

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 deal 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

We really thank Evan EdjeElectronics’ work for his code sharing (https://github.com/EdjeElectronics). This work was supported by the Macau Science and Technology Development Fund under grant number FDCT 023/2017/A1 and by the University of Macau under grant number MYRG 2018-00125-FST.

5. REFERENCES


