



Article

Character Segmentation and Recognition of Variable-Length License Plates Using ROI Detection and Broad Learning System

Bingshu Wang ¹, Hongli Xiao ¹, Jiangbin Zheng ¹, Dengxiu Yu ² and C. L. Philip Chen ^{3,*}

¹ School of Software, Northwestern Polytechnical University, Xi'an 710129, China; wangbingshu@nwpu.edu.cn (B.W.); xiaolili@mail.nwpu.edu.cn (H.X.); zhengjb@nwpu.edu.cn (J.Z.)

² Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an 710129, China; yudengxiu@126.com

³ School of Computer Science and Engineering, South China University of Technology, Guangzhou 510641, China

* Correspondence: philip.chen@ieee.org

Abstract: Variable-length license plate segmentation and recognition has always been a challenging barrier in the application of intelligent transportation systems. Previous approaches mainly concern fixed-length license plates, lacking adaptability for variable-length license plates. Although objection detection methods can be used to address the issue, they face a series of difficulties: cross class problem, missing detections, and recognition errors between letters and digits. To solve these problems, we propose a machine learning method that regards each character as a region of interest. It covers three parts. Firstly, we explore a transfer learning algorithm based on Faster-RCNN with InceptionV2 structure to generate candidate character regions. Secondly, a strategy of cross-class removal of character is proposed to reject the overlapped results. A mechanism of template matching and position predicting is designed to eliminate missing detections. Moreover, a twofold broad learning system is designed to identify letters and digits separately. Experiments performed on Macau license plates demonstrate that our method achieves an average 99.68% of segmentation accuracy and an average 99.19% of recognition rate, outperforming some conventional and deep learning approaches. The adaptability is expected to transfer the developed algorithm to other countries or regions.

Keywords: character segmentation; ROI detection; deep transfer learning; cross class removal; broad learning system



Citation: Wang, B.; Xiao, H.; Zheng, J.; Yu, D.; Chen, C.L.P. Character Segmentation and Recognition of Variable-Length License Plates Using ROI Detection and Broad Learning System. *Remote Sens.* **2022**, *14*, 1560. <https://doi.org/10.3390/rs14071560>

Academic Editors: Xiangtao Zheng, Fulin Luo and Qi Wang

Received: 11 February 2022

Accepted: 21 March 2022

Published: 24 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Automatic license plate recognition (ALPR) plays an important role in intelligent transportation systems [1–4]. The core of ALPR is to recognize license plate characters that constitute an identification of a vehicle [5–7]. For each country or region, it has its own standards for license plates. If only one standard is required, license plates can be produced with a fixed length. In practice, due to historic or political reasons, more than one type of license plates are used in one country or region. Each type of license plate has its own fixed length and character distributions. Thus, it results in various representations of license plates [8–11].

Generally, two factors have influences on ALPR. The first factor is that there are many variable-length license plates. Some samples are shown in Figure 1. Four, five, or six characters are printed on one plate. Significant differences appear between single-row and double-row license plates. The second is that there are many variations between characters such as color, font, space mark, and screws. Notably, some letters are similar with some digits, for example, 'A' and '4', 'Q' and '0', and 'B' and '8'.

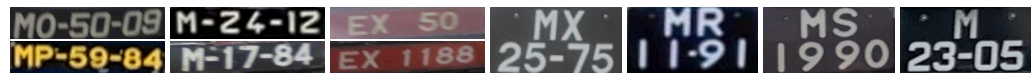


Figure 1. Some examples of variable-length license plates. Different numbers, colors, and fonts are covered.

Previous works mainly focus on fixed-length license plates with one style. They include projection-based approaches [12–16], MSER-based approaches [17,18], connected components analysis-based approaches [19–21], and template matching-based approaches [22–24]. Although these approaches have achieved good results in given scenes, they lack of flexibility and adaptability to handle variable-length license plates. Thus, investing variable-length license plate segmentation and recognition is necessary and valuable.

The objective of this paper is to propose a solution to solve the problems of variable-length license plate segmentation and recognition. Recently, deep learning-based methods have achieved surprising results in the field of ALPR [4]. However, some problems still exist, as shown in Figure 2. The first is the cross-class problem. For one character on the license plate, more than one region might be located. The overlapped regions have the same class in most cases. One potential reason is that detection models of deep neural network may be restricted by its region proposal network. The overlapped regions also have different classes. One possibility is that some subregions of complex character (e.g., ‘M’ and ‘W’) are probably detected by deep learning models.

The second problem is false positives or missing detections. Figure 2b sheds light on some background regions that look similar to characters, and these regions are easily detected as characters. It should be noted that some real characters are not detected by deep learning methods. One of the main reasons seems to be that deep learning methods are data-driven, and a limited number of samples cannot ensure that trained models cover all application scenes.

Thirdly, recognition errors may occur due to the similarity between letters and digits, for example, digit ‘1’ and letter ‘l’. The reason may be derived from the absurd structure design of license plates. Designers initially do not consider the minor differences between some letters and digits carefully. It is impracticable to transform all types of license plates into a unified type due to reasons with respect to time and cost.

In this paper, we propose a solution for the segmentation and recognition of variable-length license plates. Our work is based on the results of license plate detection [11]. Given one license plate, regions of interests (ROIs) are detected firstly. Then, cross class is removed and template matching is carried out to overcome missing detections. Lastly, two-fold broad learning systems are exploited for recognition. The main contributions include the following:

- We propose a method that regards character segmentation problem as an object detection problem. A deep transfer learning model based on Faster-RCNN framework is trained to generate sufficient ROIs;
- A strategy of cross class removal of character is designed to reject overlapped ROIs, and a mechanism of template matching is designed to predict missing characters. They can improve segmentation accuracy significantly;
- This paper exploits two-fold broad learning systems for character recognition. It can largely reduce recognition errors caused by the confusion between letters and digits.

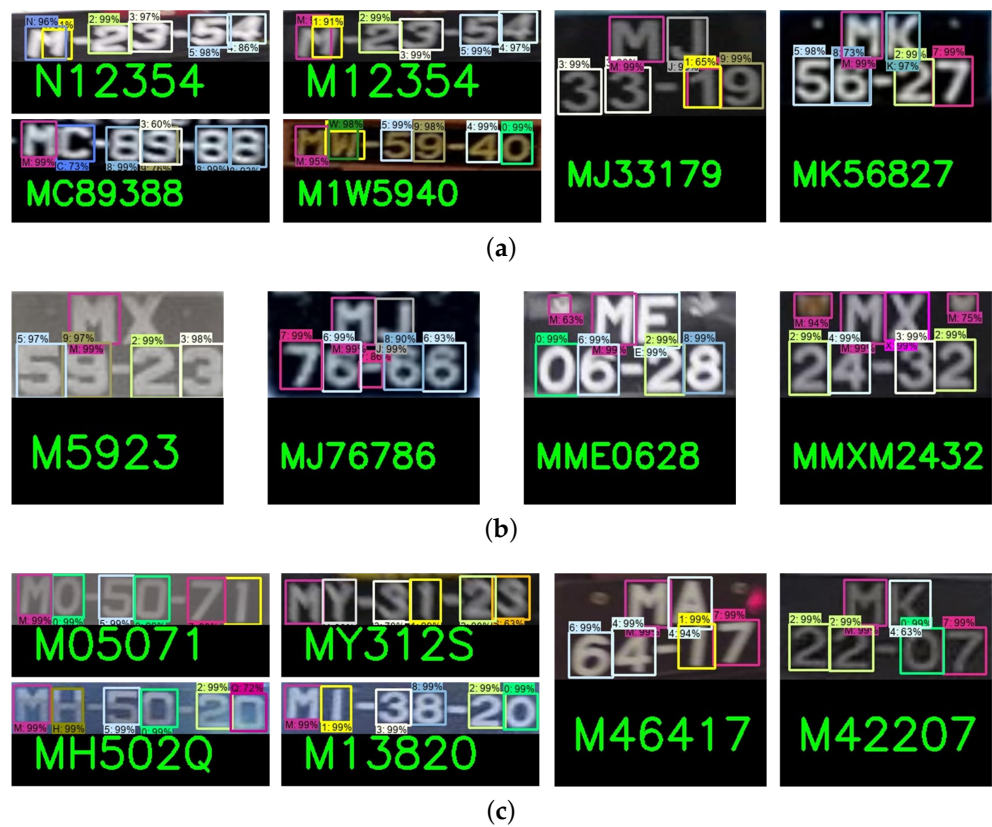


Figure 2. Some problems of character segmentation and recognition generated by the ROI detection approach. (a) Cross class problem. (b) False positives or missing detections. (c) Recognition confusion between letters and digits.

2. Related Work

Over the past decade, many algorithms were proposed for license plate character segmentation (LPCS). They can be divided into conventional methods and deep learning methods.

2.1. Conventional Methods

Generally, conventional LPCS methods are classified as follows: projection-based methods [12–16], MSER-based methods [17,18], CCA-based methods [19–21], and template matching-based methods [22–24].

Pun et al. [13] explored an edge-based method for ALPR. Vertical projection is used to extract character separators by using the pixel intensities followed by morphological erosion and trimming operations. However, it mainly tackled some simple scenes for Macau license plates with a 95% recognition rate. Ariel et al. [16] employed vertical and horizontal coordinates to delimit characters for Argentine license plates, which reached an accuracy of 96.49%. Our previous work [14] exploited a key character and used projection analysis to segment single-row or double-row license plates in Macau. However, only fixed-length license plates were processed.

Yang et al. [17] employed a Maximally Stable Extremal Region (MSER) to detect candidate regions and located the positions of characters using the template matching approach. It was able to process two kinds of Chinese license plates. However, it faces challenges in touching characters. Kim et al. [20] exploited a super-pixel based degeneracy factor to realize the adaptive binarization of character regions. Then, CCA is used to obtain segmented characters. It is claimed to be effective for local illumination change scenes. Goncalves et al. [19] presented a detailed review about character segmentation. They hold the view that the (near) optimal effectiveness of LPCS is very important for

ALPR. CCA was selected to segment license plate characters of Brazilian plates. Miao [22] explored the characteristics of horizontal and vertical projections for LPCS by applying a variable-length template matching strategy. Although the templates have multiple sizes, the number of license plate characters is fixed. Kim et al. [24] firstly detected one common character from an anchor image and then used a classifier to determine the license plate type. Structural information was utilized to estimate the approximate locations of other characters according to the type. However, if the common character was not found, the method failed.

In addition, there are some other methods developed for character segmentation task [25–29]. Most conventional methods primarily focus on fixed-length license plates for which their characters are frontal and structured information is clear.

2.2. Deep Learning Methods

Most deep learning methods for ALPR depend on Convolutional Neural Network (CNN).

Musaddid et al. [30] used some techniques including CNN, sliding window, and a bounding box refinement for the character segmentation of Indonesian license plates. Sliding window provides candidate regions that will be classified by CNN. However, this process requires increased computation costs. Silva et al. [29] proposed a flexible approach for ALPR and agreed to use detection-based methods for character segmentation, especially for multi-row characters. The culmination some prior knowledge can help improve the recognition rate.

Duan et al. [31] developed an end-to-end CNN classification model to detect consecutive characters. However, it is limited to imbalanced data. Selmi et al. [32] employed Mask-RCNN to extract object instance regions and classified 38 classes of characters. It might avoid the cross-class problem, but it cannot handle the confusion problem between letters and digits. An ROI-based deep learning method [33] was proposed for the character segmentation of variable-length license plates, which achieved higher accuracy than some conventional methods for LPCS. However, the cross-class problem and missing detection problem remain to be solved. In addition, there are many other methods proposed for ALPR [34–38]. As far as object recognition is concerned, many algorithms including neural networks have been proposed [39–44]. Notably, it is reported that the broad learning system (BLS) [43,44] has advantages over deep learning methods [40,42] in training efficiency.

In summary, deep learning methods have been verified to have advantages over conventional methods in the field of LPCS. ROI detection-based methods are expected to segment characters well for variable-length license plates [33]. In this paper, Faster-RCNN framework [45] and InceptionV2 structure [46] are employed for variable-length LPCS, and the broad learning system is selected for character recognition.

3. The Proposed Method

The outline to our method is illustrated in Figure 3. It mainly has three stages. Stage A aims to provide candidate character regions, i.e., ROIs. Stage B is designed to remove cross class and reduce missing detections. With the potential candidate regions, two-fold broad learning systems are utilized for character recognition in Stage C.

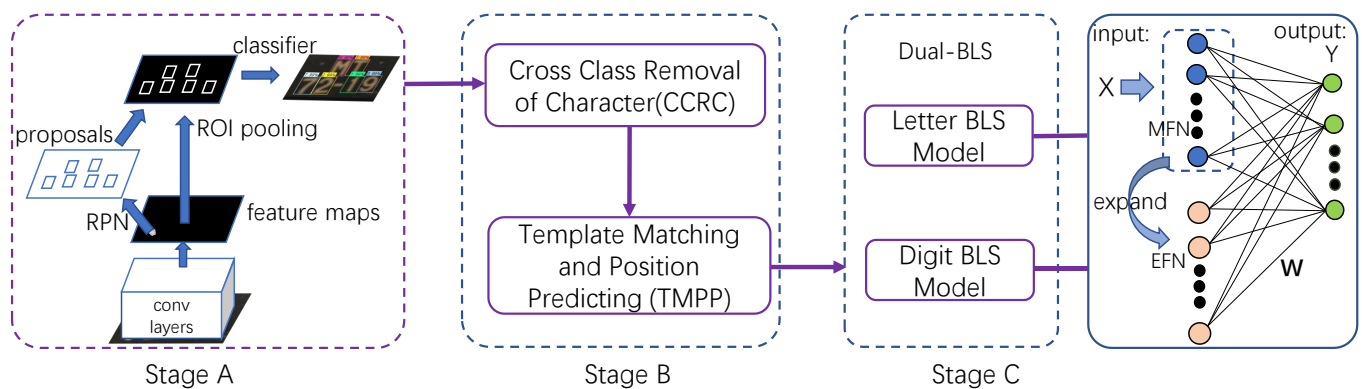


Figure 3. The flowchart of the proposed method. It includes three stages: stage A is employed to generate ROIs, stage B is designed to remove cross class and predict positions of missing characters, and stage C is designed to handle the problem of recognition confusions.

3.1. Deep Learning for ROI Detection

In this section, the character segmentation problem is converted into an object detection problem because each character can be regarded as an object. We train a deep transfer learning model to realize the function. It is built up on Faster-RCNN with InceptionV2 structure, which encompasses four steps.

1. Extract feature maps through Conv layers. This step is designed by a series of *conv + relu + pooling* layers. The last feature map is shared by the subsequent operations such as region proposal network (RPN) and ROI pooling.
2. Generate proposals by RPN. RPN is utilized to provide a group of potential character proposals for classification. Each proposal is an anchor. To generate sufficient proposals, three aspect ratios (0.5, 1.0, and 2.0) and four scales (0.25, 0.5, 1.0, and 2.0) are selected. RPN includes a box-regression layer and a box-classification layer. It determines whether a box belongs to object or background, and it also adjusts the positions of a box.
3. Obtain fixed-dimension of features by ROI pooling. It receives a feature map from step 1. and converts different dimension of proposals into a fixed-dimension feature map by using the max-pooling operation. It generates a fixed dimension of the feature map for multi-dimension feature maps.
4. Proposal classification. It receives feature maps of many proposals and outputs classification results. Meanwhile, the positions of proposals are predicted by bounding box regression. Finally, the objects and their locations are generated.

It is pretty clear that RPN is of importance to produce effective proposals, which ensures detection accuracy. Considering the small sizes for most license plates, a good trade-off can be achieved by the straightforward pass of four steps. Moreover, a traditional convolution network used in Faster-RCNN requires many computations and faces representational bottleneck. To address these issues, the InceptionV2 structure is employed [46].

In this paper, our ROI detection model is transferred from a model that is pre-trained on a COCO dataset (https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf1_detection_zoo.md, accessed on 1 January 2022). All training images for character detection are labeled manually by a tool called “LabelImg” (<https://github.com/tzutalin/labelImg/>, accessed on 1 January 2022). Finally, a group of candidate characters with locations and scores was generated.

3.2. Proposed Techniques for ROIs Processing

3.2.1. Proposed Cross Class Removal of Character

Although the ROI detection technique provides many regions, overlapped regions may occur, as illustrated in Figure 2a. This easily results in redundant detections and

causes erroneous recognitions. This is called the cross-class problem. To solve the problem, the redundant regions need to be removed. Thus, a strategy of Cross Class Removal of Character (CCRC) is proposed as follows.

For one license plate, firstly, its candidate regions are sorted by x coordinate ascending. Then, all the values of Intersection over Union (IoU) between adjacent regions are computed. IoU is used to compare with a threshold τ . If one IoU value is larger than τ , the two regions are merged into one new region. It is effective for single-row license plates. When it comes to double-row license plates, they are divided into two rows. After that, each row is handled individually. In practice, $\tau = 0.25$ may provide satisfactory results. The process of obtaining accurate segmented characters is described in Algorithm 1.

Algorithm 1 Cross Class Removal of Character.

Input: Candidate regions $C = \{C_1, C_2, \dots, C_k\}$ with their locations.

Output: Updated regions $UC = \{UC_1, UC_2, \dots, UC_m\}$.

- 1: Do ascending sort according to x position for all candidate regions $C' = \{C'_1, C'_2, \dots, C'_k\}$.
- 2: Compute Intersection over Union (IoU) for adjacent regions in C' : $IU = \{IU_1, IU_2, \dots, IU_{k-1}\}$.
- 3: **for** $j = 1$ to $k - 1$ **do**
- 4: **if** ($IU_j > \tau$) **then**
- 5: Merge the two regions C'_j, C'_{j+1} into one region UC_t by MER-Minimum Enclosing Rectangle.
- 6: Add UC_t into UC .
- 7: **else**
- 8: Add C'_j into UC directly.
- 9: **end if**
- 10: **end for**
- 11: **return** UC

Be aware that all the detected classes for candidate regions will not be used. All candidate regions will be recognized by subsequent broad learning systems.

3.2.2. Proposed Template Matching and Position Predicting

Some missing detections or false positives may be generated by the ROI detection technique. To address this problem, a Template Matching and Position Predicting (TMPP) algorithm is proposed, which is based some prior knowledge such as the basic structural information of license plates.

The flowchart of TMPP is presented in Figure 4. For one single-row license plate, its third character is missing by the ROI detection method (Figure 4a). Then, all the positions of detected characters are mapped into one labeled figure in Figure 4b. Next, the figure is compared with the license plate template library. The most matching template will be selected for predicting positions (Figure 4c). The position of the missing character is predicted by its neighboring region in Figure 4d. Finally, the segmentation results are updated, as shown in Figure 4e. More results processed by TMPP are given in Figure 5.

For the efficiency of template matching method, experiments are conducted on 2234 images (including single-row and double-row license plates). It takes about 6.26 ms for the TMPP strategy to reduce the false positives and missing detections.

As far as the adaptability of template matching is concerned, it depends on the practical problems and the standards of the country or region. The design of template matching library is supposed to refer to the standards. In our manuscript, Macau license plates have many forms due to historical or political reasons. The establishment of license plate template library can provide valuable structural information for reference. When TMPP is applied to other countries or regions, a similar library is expected to be built for the segmentation task.

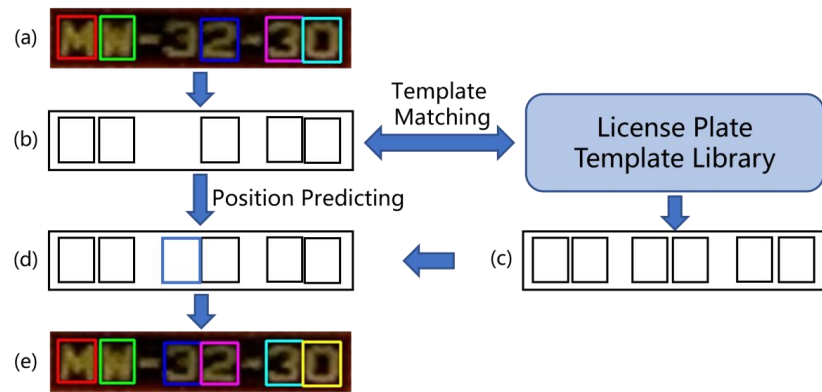


Figure 4. The flowchart of handling missing detection by the proposed TMPP. (a) is the detection result generated by ROI-based method, (b) shows the rectangles of detected regions, (c) is the selected template from license plate template library that matches with (b), (d) represents predicted result by template matching, (e) shows final detection result.

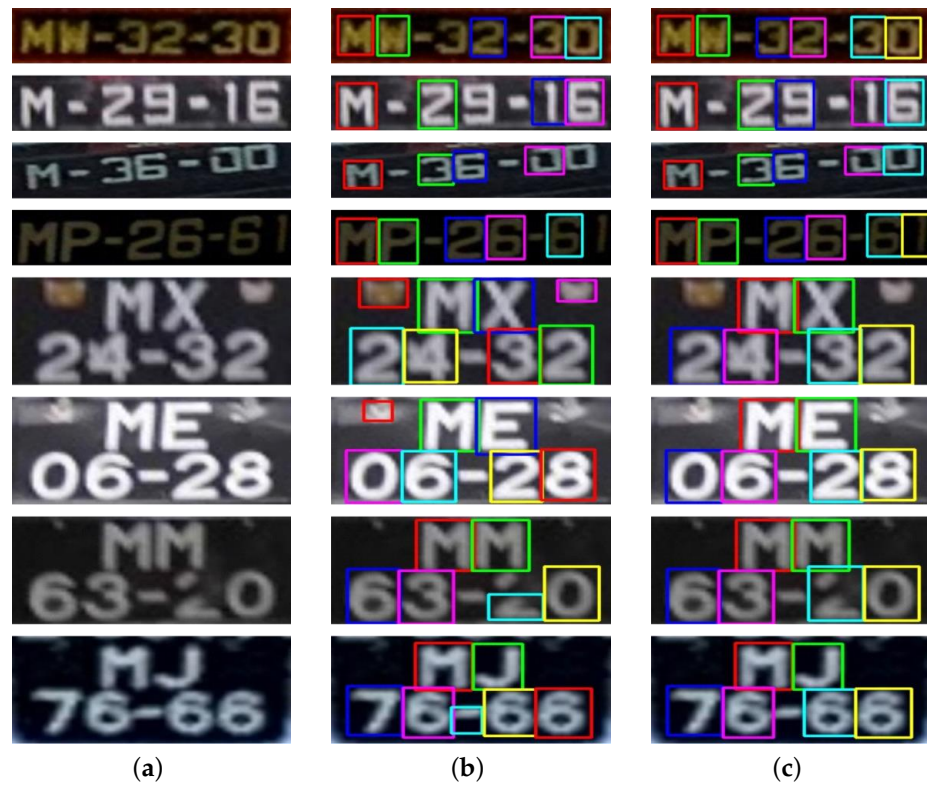


Figure 5. Some examples to show the function of TMPP. (a) Input images. (b) False positives or missing detections. (c) The results processed by TMPP.

3.3. Designed Dual Broad Learning Systems for Recognition

In this section, to realize the classification goal, a broad learning system is exploited. It is built up on a flat neural network [43], which is the main characteristic. Figure 6 elaborates on the detailed structure.

For the classification task, at first input images are transformed into random features called “mapped feature nodes”; then, all the mapped feature nodes are expanded to enhanced features called “enhanced feature nodes”. The design is a dependable method for extracting essential features from the wider perspective.

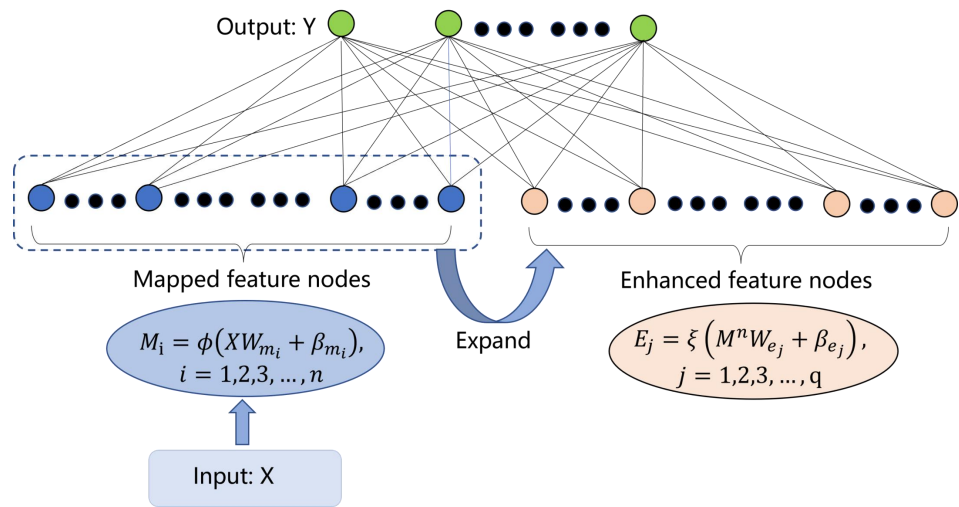


Figure 6. The structure of BLS. It encompasses mapped feature nodes and enhanced feature nodes.

Mapped feature nodes are defined by the following:

$$M_i = \phi(XW_{m_i} + \beta_{m_i}), i = 1, 2, \dots, n \tag{1}$$

where W_{m_i} and β_{m_i} are random weights generated from specific distributions, and ϕ is a mapping function. Then, a sparse auto-encoder [41] is employed to finetune the mapped features and explore more essential features. After the mapping operations, n groups of feature nodes are produced, which can be denoted by $M^n \equiv [M_1, \dots, M_n]$. After that, M^n is used to expand to enhanced features:

$$E_j = \zeta(M^n W_{e_j} + \beta_{e_j}), j = 1, 2, \dots, q \tag{2}$$

where ζ is a nonlinear activation function, e.g., *tansig*. W_{e_j} and β_{e_j} are random weights generated from specified distribution. Herein, the first q groups of enhanced feature nodes can be denoted by $E^q \equiv [E_1, \dots, E_q]$.

With mapped feature nodes and enhanced feature nodes obtained, they are both connected to the output layer:

$$\begin{aligned} Y &= [M_1, M_2, \dots, M_n, E_1, E_2, \dots, E_q]W \\ &= [M^n | E^q]W \end{aligned} \tag{3}$$

where W represents the entire weight, and Y is the output layer. The term W can be computed by $W \triangleq [M^n | E^q]^+ Y$, where $[M^n | E^q]^+$ can be derived from the pseudoinverse of the ridge regression approximation.

In practice, when the current BLS model cannot learn a task well, an effective solution is to add feature nodes: mapped or enhanced. This is the function of incremental learning for BLS, updating the model without retraining it from the start.

When one mapped feature node is added, $M_{n+1} = \phi(XW_{m_{n+1}} + \beta_{m_{n+1}})$, the new mapped features can be expressed by $M^{n+1} \equiv [M_1, \dots, M_{n+1}]$. As a result, enhanced feature nodes are updated as $E^{ex_j} \triangleq [\zeta(M^{n+1}W_{ex_1} + \beta_{ex_1}), \dots, \zeta(M^{n+1}W_{ex_j} + \beta_{ex_j})]$, where W_{ex_j} and β_{ex_j} , $j = 1, 2, \dots, q$ are the weights generated randomly. If one enhanced feature node is added, it can be expressed by $E_{q+1} = \zeta(M^{n+1}W_{ex_{q+1}} + \beta_{ex_{q+1}})$.

Herein, we denote $A_n^q \triangleq [M^n | E^q]$ and $A_{n+1}^{q+1} \triangleq [M^{n+1} | E^{ex_q} | E_{q+1}]$. The updated weights can be calculated by the following:

$$(A_{n+1}^{q+1})^+ = \begin{bmatrix} (A_n^q)^+ & -DB^T \\ & B^T \end{bmatrix} \tag{4}$$

$$W_{n+1}^{q+1} = \begin{bmatrix} W_n^q - DB^T Y \\ B^T Y \end{bmatrix} \quad (5)$$

where $D = (A_n^q)^+ [M_{n+1} | E^{exq} | E_{q+1}]$:

$$B^T = \begin{cases} (C)^+ & \text{if } C \neq 0 \\ (1 + D^T D)^{-1} D^T (A_n^q)^+ & \text{if } C = 0 \end{cases} \quad (6)$$

and $C = [M_{n+1} | E^{exq} | E_{q+1}] - A_n^q D$.

To recognize characters and avoid confusion between letters and digits, we trained two broad learning systems to classify letters and digits, respectively. These are called dual broad learning systems. Letters and digits are the most widely used characters around the world. Thus, the trained models can be applied to other countries or regions. Moreover, it is also expected to train more models for recognition task when more categories of characters are added.

4. Experimental Results

4.1. Dataset, Evaluation Metric, and Parameter Setting

To evince the effectiveness of the proposed method, we provide experimental results on Macau license plates. The license plate images are collected from parking lots and bus stations for training, validation, and testing. The training set includes 1913 images, and the validation set includes 287 images. They are labeled manually by one tool “labellmg”. The testing set includes 2234 images that are divided into three categories, as shown in Table 1: *Indoor*, *Outdoor*, *Complex*. Among the three test sets, *Complex* is the most challenging set. The dataset has been open to the public: https://github.com/BookPlus2020/VL_LPR (accessed on 1 January 2022). Related segmentation and recognition models are also available for download. Details have been released on the Github website.

Table 1. The dataset of Macau license plate for test.

	Indoor	Outdoor	Complex	Sum
Number	1000	830	404	2234

In this paper, the projection-based approach [13], MSER-based approach [17], CCA-based approach [19], ROI-based approach [33] are selected for quantitative and visual comparisons. They are typical conventional approaches or deep learning techniques, and they run on a PC with Windows 10 OS and Intel Xeon CPU E5-1650 V2. Their parameters refer to [33]. For our method, the parameter setting is presented as follows. For the ROI detection stage, the maximum of proposals for RPN is 300, the momentum is 0.9, the learning rate is 0.0002, and the number of training steps is 200k. For CCRC, $\tau = 0.25$ is selected. For the BLS model, the number of mapped feature nodes is 600, and the number of enhanced feature nodes is 5600. Experiments are conducted on a PC with windows 10 OS, Intel Core i7-10700F CPU, Tensorflow 1.5, and NVIDIA Geforce GTX 1660 Super with 6 GB memory.

With respect to the evaluation metric of character segmentation, if there are false positives or missing detections for one license plate, it will be regarded as a failure; thus, the metric refers to the following [17]:

$$\eta = \left(1 - \frac{\text{errors}}{\text{total}}\right) \times 100\% \quad (7)$$

where *errors* is defined as the number of failures, and *total* is the number of all license plates. Similarly with segmentation accuracy η , if any character is recognized erroneously, license plate recognition will be regarded as a failure.

4.2. Quantitative Analysis

We present a quantitative segmentation comparison in Table 2. The complexity of the scenes *Indoor*, *Outdoor*, *Complex* is ranges from easy to difficult. For the *Indoor* scene, the Projection-based approach [13] achieved 97.60% accuracy, outperforming the MSER-based approach [17] and CCA-based approach [19]. The ROI-based approach [33] is 2.7% larger than that of [13]. This indicates the effectiveness of the ROI-based approach for character segmentation. With the addition of CCRC, there is a slight 0.1% improvement. It is also notable that, when TMPP is utilized, accuracy reaches an encouraging result of 100%.

Compared with the *Indoor* set, the segmentation accuracy of all the approaches on the *Outdoor* set decreases except for the MSER-based approach. Since the MSER-based approach has difficulty in dealing with touching characters that appear many times in the *Indoor* set, its segmentation rate is very low. Notably, the *Complex* set is the most challenging set because the highest accuracy among all the conventional approaches is no more than 92%, while the ROI detection approach achieves an obvious value of 96.75%. In a collaboration with the designed CCRC and TMPP, the ROI detection approach has a striking improvement. The proposed method, a combination of ROI, CCRC, and TMPP, reaches a 99.25% for *Complex* set and 99.68% of average value. We experimentally show that the proposed method surpasses the compared approaches on three test sets.

Table 2. Quantitative segmentation comparison (%) between our method and the compared approaches. The strategy used in our method is ROI+CCRC+TMPP.

Scene	Indoor	Outdoor	Complex	Sum
Projection-based [13]	97.60	85.54	79.95	87.69
MSER-based [17]	89.50	92.41	91.58	91.16
CCA-based [19]	96.50	90.60	88.12	91.74
ROI-based [33]	99.20	96.75	96.04	97.32
ROI+CCRC	99.30	97.71	97.27	98.34
ROI+CCRC+TMPP	100	99.52	99.25	99.68

We also provide experiments for recognition comparison, as shown in Table 3. The ROI-based detection approach [33] is implemented by Faster-RCNN with InceptionV2 structure, with an accuracy of average 88.14%. It recognizes letters and digits together and is easily confused by the cross-class problem. In addition, it also faces the false positives problem. We compare several methods on ablation research, in a task involving character recognition, and found that the combination of ROI, CCRC, TMPP, and a two-fold BLS is able to achieve better performance than others in Table 3. CCRC can promote the performance of [33] slightly. With the introduction of dual broad learning systems, the recognition rate is promoted from 89.03% to 98.21% significantly. In a collaboration with the designed TMPP, the recognition accuracy reaches an average value of 99.19%, with every set higher than 98.5%. It is clearly noted that the improvement is obvious. Thus, the proposed method outperforms [33] in the recognition task of Macau license plates.

Table 3. Quantitative comparison between character recognition methods. The last row ROI+CCRC+TMPP+BLS is the proposed method for ALPR task.

Scene	Indoor	Outdoor	Complex	Sum
ROI-based [33]	87.6	88.43	88.86	88.14
ROI+CCRC	88.2	89.46	89.85	89.03
ROI+CCRC+BLS	99.2	97.71	96.78	98.21
Ours	99.7	98.92	98.51	99.19

4.3. Efficiency Analysis

We also present a running time comparison between the proposed approach and the compared methods, as shown in Table 4. Two sizes, 320×64 and 256×128 , are selected

to test the efficiency of methods. Obviously, the larger the image's size, the more time it takes. It can be seen from Table 4 that conventional approaches require less time to segment characters. For example, to process a 256×128 image, projection-based, MSER-based and CCA require 3.65 ms, 6.73 ms, and 4.18 ms, respectively. The ROI-based method needs more time (125.49 ms), because it requires much time in the forward propagation of a deep neural network.

It should be noted that, for character segmentation, it takes about 0.015 ms for the CCRC module and 6.20 ms for the TMPP module to process one license plate. Compared with ROI detection, CCRC and TMPP take less time. One of the main reasons is that only a few logistic computations are carried out. Although it needs more time than that of conventional methods, the addition of CCRC and TMPP modules helps improve segmentation accuracy significantly. For recognition tasks, the trained digit BLS or letter BLS needs about 5.09 ms to process one image with a size of 32×32 pixels. The running time of BLS is based on the selection of parameters. In this paper, it is enough to use 600 mapped feature nodes and 5600 enhanced feature nodes. Compared with conventional methods, the segmentation accuracy and recognition rate of the deep learning method increase at the cost of computations.

Table 4. Running time (ms) comparison between the proposed approach and the compared methods.

Image Size	320×64	256×128
Projection-based [13]	2.31	3.65
MSER-based [17]	4.03	6.73
CCA-based [19]	2.53	4.18
ROI-based [33]	117.49	125.49

4.4. Visual Results and Discussion

Visual comparison includes two parts: segmentation results and recognition results. Figure 7 illustrates the variable-length segmentation results of single-row license plates. Conventional methods such as the Projection-based approach [13] and CCA-based approach [19] have difficulties in handling variable-length license plates. The approaches [13,19] even fail to locate the positions of license plates with red backgrounds. The MSER-based approach [17] can produce good results for proper images. However, it may fail to deal with touching characters (e.g., the result at the second row and third column). The ROI-based approach [33] offers a powerful new method to segment characters. However, it easily brings about the cross-class problem (from the fourth row to the eighth row in Figure 7), missing detections (from the 9th row to the 12th row in Figure 7), and false positives (from the sixth row to the ninth row in Figure 8). Benefiting from CCRC and TMPP, the proposed method outperforms the compared approaches in terms of handling these complex problems.

Figure 9 exhibits some visual recognition results. The odd rows are generated by [33], while the even rows are generated by our method. There are classification confusions between letters and digits for the approach [33]. For example, in the *Indoor* part, 'S' and '3' are at the first row and the first column, 'I' and '1' are at the first row and the second column, and 'X' and '4' are at the third row and the first column. Figure 9 shows that our method produces the correct results by classifying letters and digits separately.

The cross-class problem is also covered in Figure 9; for example, in the *Outdoor* part, the license plate with real number 'MW5940' at the first row and the fifth column is recognized as 'M1W5940'. Thanks to the design of CCRC, the redundant '1' is excluded and does not appear in our recognition result. One example of false positive is presented in the third row and the fifth column in the *Complex* part. The real number is 'ME0628', but the license plate is recognized as 'MME0628' by [33] due to a false positive at the top left. In contrast, the false positive is removed by using TMPP.

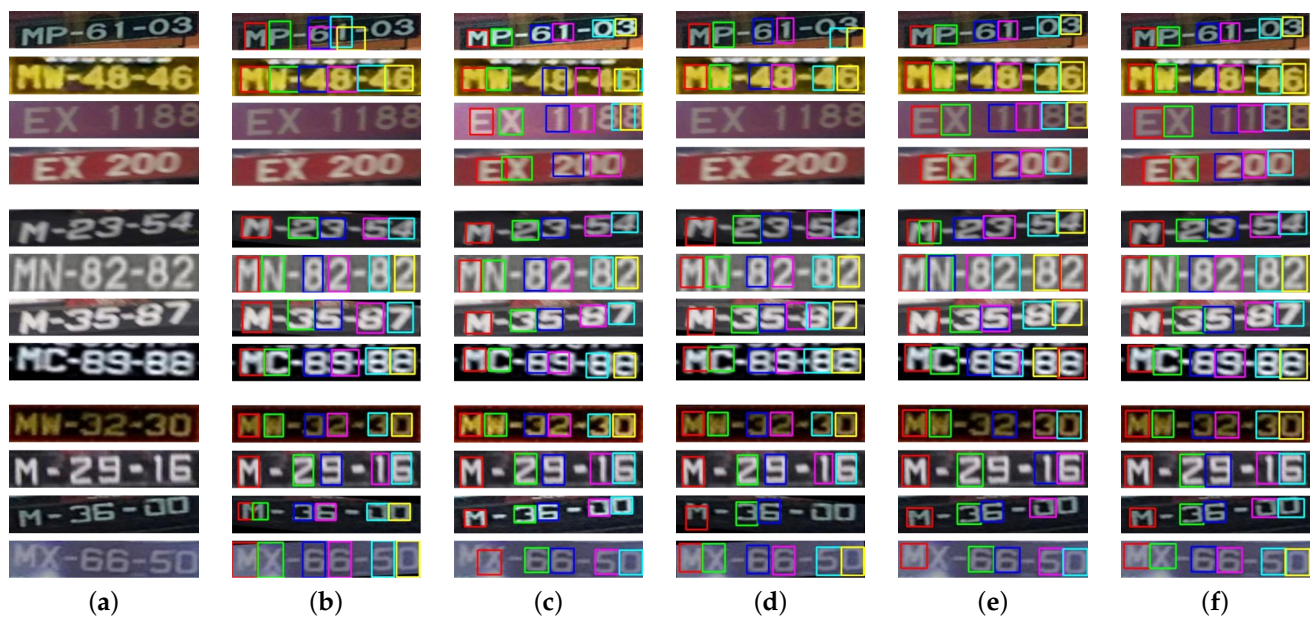


Figure 7. Some segmentation results of single-row license plates. (a) Input images. (b) Projection-based [13]. (c) MSER-based [17]. (d) CCA-based [19]. (e) ROI-based [33]. (f) Ours.

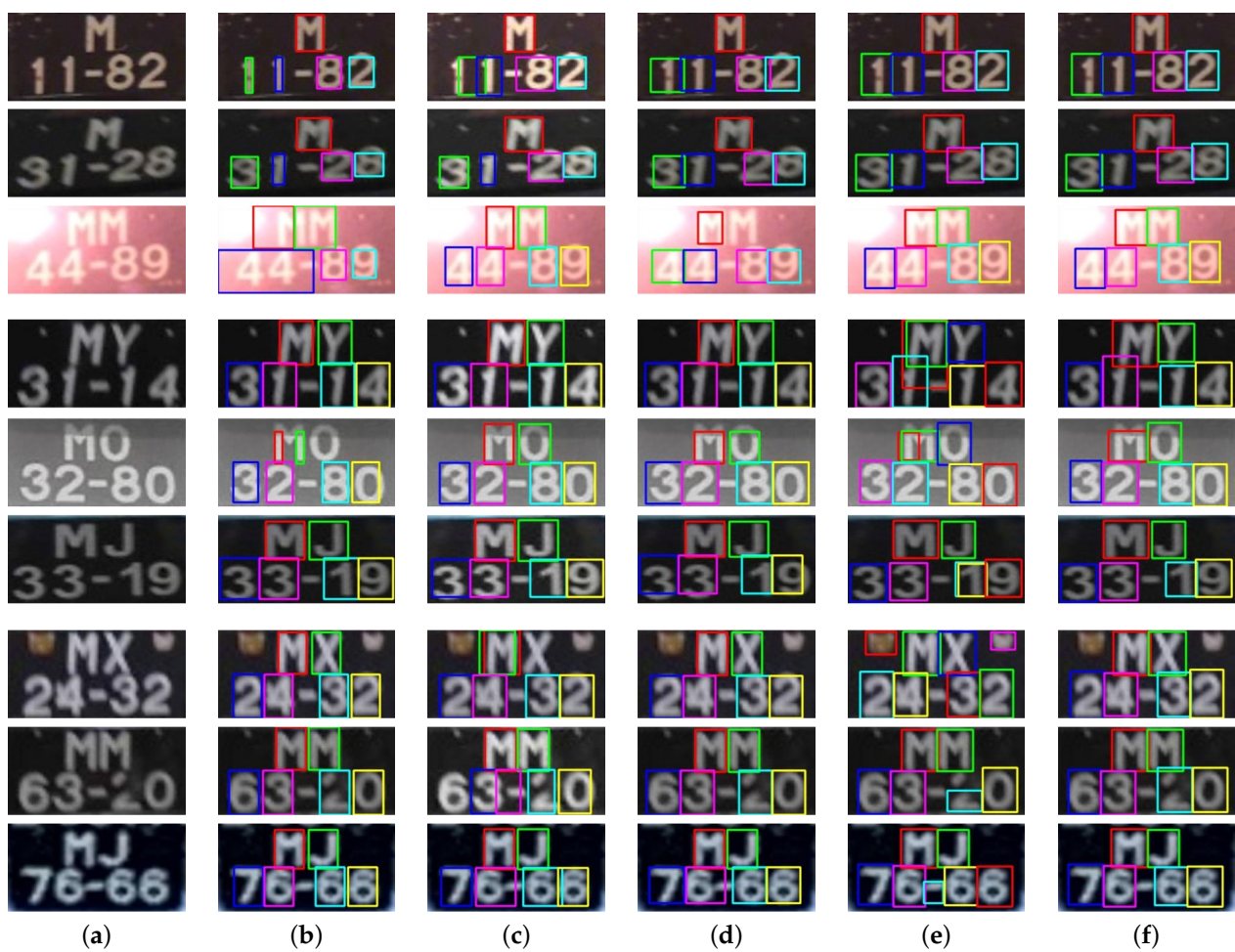
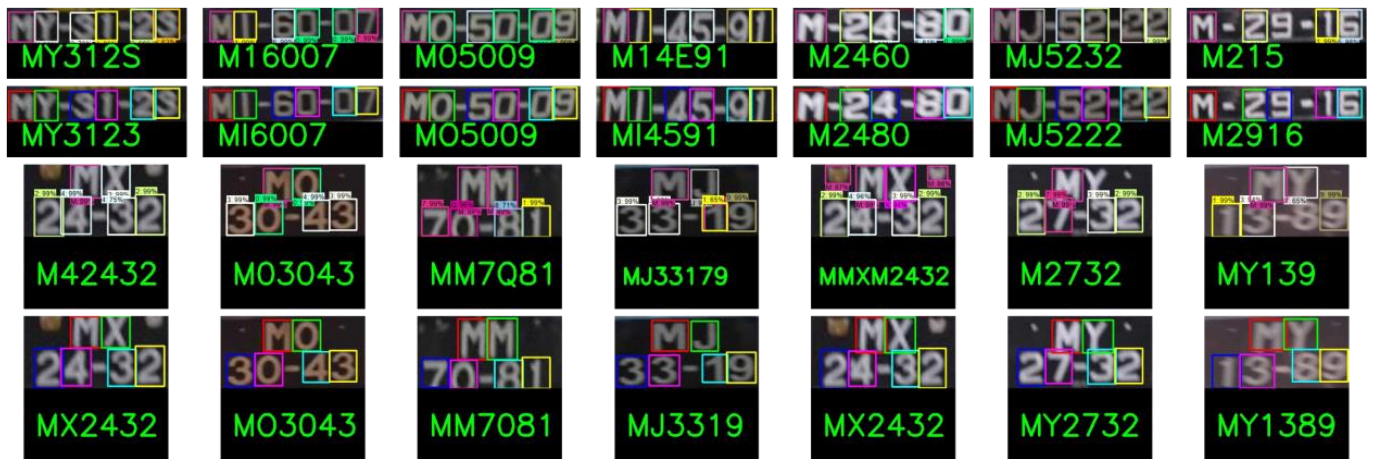


Figure 8. Some segmentation results of double-row license plates. (a) Input images. (b) Projection-based [13]. (c) MSER-based [17]. (d) CCA-based [19]. (e) ROI-based [33]. (f) Ours.



(a)



(b)



(c)

Figure 9. Some comparison of recognition results between ROI-based approach [33] and ours. Three sets are covered. The odd rows are generated by [33], and the even rows are generated by the proposed method. (a) Indoor; (b) outdoor; (c) complex.

Quantitative and visual results indicate the effectiveness of proposed method. Actually, there are some failures, as shown in Figure 10. They are complex or extreme cases for slanting, touching characters, and low-quality image. These cases are inevitable for realistic applications. An potential solution to deal these cases is to use image sequences. When

one car drives close to or leaves the camera, a series of frames can be captured. Thus, to boost the performance of an ALPR, image sequences can be employed with statistical information for recognition tasks. Two examples are illustrated in Figure 11. Nine of ten frames are recognized successfully, and the license plate number is expected to be obtained accurately.



Figure 10. Some segmentation failures including missing detections and incorrect segmentation.

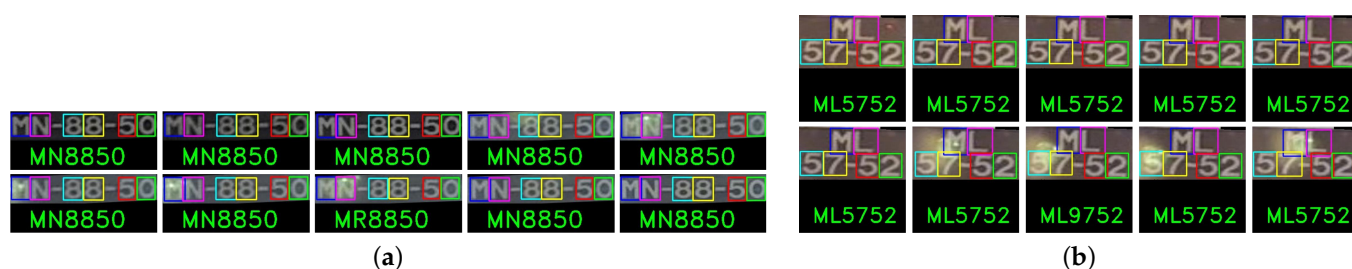


Figure 11. The recognition of license plates using image sequences. For each example, (a) MN8850 and (b) ML5752, among the 10 frames, 9 frames are successful and only 1 failure occurred (at 2nd row, 3rd column) because of a strong reflection of illumination.

5. Conclusions

In this paper, we propose a method using ROI detection and BLS for variable-length LPR. Our method is based on the Faster-RCNN framework, and it aims to solve several problems of ROI detections and generate correct recognition results. The designed strategy of Cross Class Removal of Character aims to reduce overlapped detections. The developed mechanism of template matching and position predicting is used to reduce false positives and estimate the possible positions of missing characters. Specially, we employ a dual broad learning system to tackle the confusions of character recognitions. Experiments conducted on Macau license plates demonstrate that our method achieves an average 99.68% accuracy for the character segmentation task, and an average 99.19% accuracy for the character recognition task. With the promising adaptability, the proposed strategies are expected to become translated into other regions for which their license plates have similar characteristics with Macau's. In the future, it would be interesting to process image sequences by integrating these modules, and the computation costs should be reduced with more investigations.

Author Contributions: Conceptualization, B.W. and C.L.P.C.; methodology, B.W.; validation, B.W. and H.X.; formal analysis, D.Y. and C.L.P.C.; investigation, B.W. and J.Z.; resources, J.Z. and D.Y.; data curation, B.W. and C.L.P.C.; writing—original draft preparation, B.W.; writing—review and editing, H.X.; visualization, B.W. and H.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China, Youth Fund, under number 62102318, and in part by the Fundamental Research Funds for the Central Universities, under number G2020KY05113. The work was also funded by the National Key Research and Development Program of China under number 2019YFA0706200 and 2019YFB1703600.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Du, S.; Ibrahim, M.; Shehata, M.; Badawy, W. Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *23*, 311–325. [\[CrossRef\]](#)
2. Lin, M.; Liu, L.; Wang, F.; Li, J.; Pan, J. License Plate Image Reconstruction Based on Generative Adversarial Networks. *Remote Sens.* **2021**, *13*, 3018. [\[CrossRef\]](#)
3. Wu, S.; Zhai, W.; Cao, Y. PixTextGAN: Structure aware text image synthesis for license plate recognition. *IET Image Process.* **2019**, *13*, 2744–2752. [\[CrossRef\]](#)
4. Zhang, L.; Wang, P.; Li, H.; Li, Z.; Shen, C.; Zhang, Y. A robust attentional framework for license plate recognition in the wild. *IEEE Trans. Intell. Transp. Syst.* **2020**, *22*, 6967–6976. [\[CrossRef\]](#)
5. Anagnostopoulos, C.N.E. License plate recognition: A brief tutorial. *IEEE Intell. Transp. Syst. Mag.* **2014**, *6*, 59–67. [\[CrossRef\]](#)
6. Shashirangana, J.; Padmasiri, H.; Meedeniya, D.; Perera, C. Automated license plate recognition: A survey on methods and techniques. *IEEE Access* **2020**, *9*, 11203–11225. [\[CrossRef\]](#)
7. Qin, S.; Liu, S. Efficient and unified license plate recognition via lightweight deep neural network. *IET Image Process.* **2020**, *14*, 4102–4109. [\[CrossRef\]](#)
8. Dun, J.; Zhang, S.; Ye, X.; Zhang, Y. Chinese license plate localization in multi-lane with complex background based on concomitant colors. *IEEE Intell. Transp. Syst. Mag.* **2015**, *7*, 51–61. [\[CrossRef\]](#)
9. Huang, Q.; Cai, Z.; Lan, T. A new approach for character recognition of multi-style vehicle license plates. *IEEE Trans. Multimed.* **2020**, *23*, 3768–3777. [\[CrossRef\]](#)
10. Min, W.; Li, X.; Wang, Q.; Zeng, Q.; Liao, Y. New approach to vehicle license plate location based on new model YOLO-L and plate pre-identification. *IET Image Process.* **2019**, *13*, 1041–1049. [\[CrossRef\]](#)
11. Chen, C.P.; Wang, B. Random-positioned license plate recognition using hybrid broad learning system and convolutional networks. *IEEE Trans. Intell. Transp. Syst.* **2020**, *23*, 444–456. [\[CrossRef\]](#)
12. Ho, W.Y.; Pun, C.M. A macao license plate recognition system based on edge and projection analysis. In Proceedings of the 2010 8th IEEE International Conference on Industrial Informatics, Osaka, Japan, 13–16 July 2010; pp. 67–72.
13. Pun, C.M.; Ho, W.Y. An edge-based Macao license plate recognition system. *Int. J. Comput. Intell. Syst.* **2011**, *4*, 244–254.
14. Wang, B.; Chen, C.P. License plate character segmentation using key character location and projection analysis. In Proceedings of the 018 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC), Jinan, China, 14–17 December 2018; pp. 510–514.
15. Xie, H. License plate character segmentation algorithm in intelligent IoT visual label. In Proceedings of the 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 14–17 July 2019; pp. 110–117.
16. Ariel, O.D.G.; Martín, D.F.; Ariel, A. ALPR character segmentation algorithm. In Proceedings of the 2018 IEEE 9th Latin American Symposium on Circuits & Systems (LASCAS), Puerto Vallarta, Mexico, 25–28 February 2018; pp. 1–4.
17. Yang, X.; Zhao, Y.; Fang, J.; Lu, Y.; Zhang, Y.; Yuan, Y. A license plate segmentation algorithm based on MSER and template matching. In Proceedings of the 2014 12th International Conference on Signal Processing (ICSP), Hangzhou, China, 19–23 October 2014; pp. 1195–1199.
18. Hsu, G.S.; Chen, J.C.; Chung, Y.Z. Application-oriented license plate recognition. *IEEE Trans. Veh. Technol.* **2012**, *62*, 552–561. [\[CrossRef\]](#)
19. Gonçalves, G.R.; da Silva, S.P.G.; Menotti, D.; Schwartz, W.R. Benchmark for license plate character segmentation. *J. Electron. Imaging* **2016**, *25*, 053034. [\[CrossRef\]](#)
20. Kim, D.; Song, T.; Lee, Y.; Ko, H. Effective character segmentation for license plate recognition under illumination changing environment. In Proceedings of the 2016 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 7–11 January 2016; pp. 532–533.
21. Abderaouf, Z.; Nadjia, B.; Saliha, O.K. License plate character segmentation based on horizontal projection and connected component analysis. In Proceedings of the 2014 World Symposium on Computer Applications & Research (WSCAR), Sousse, Tunisia, 18–20 January 2014; pp. 1–5.
22. Miao, L. License plate character segmentation algorithm based on variable-length template matching. In Proceedings of the 2012 IEEE 11th International Conference on Signal Processing, Beijing, China, 21–25 October 2012; Volume 2, pp. 947–951.
23. Lu, Y.; Zhao, Y.; Fang, J.; Yang, X.; Zhang, Y. A vehicle license plate segmentation based on likeliest character region. In Proceedings of the 2014 Seventh International Symposium on Computational Intelligence and Design, Hangzhou, China, 13–14 December 2014; Volume 2, pp. 258–262.
24. Kim, P.; Lim, K.T.; Kim, D. Learning based character segmentation method for various license plates. In Proceedings of the 2019 16th International Conference on Machine Vision Applications (MVA), Tokyo, Japan, 27–31 May 2019; pp. 1–6.
25. Wu, C.; On, L.C.; Weng, C.H.; Kuan, T.S.; Ng, K. A Macao license plate recognition system. In Proceedings of the 2005 International Conference on Machine Learning and Cybernetics, Guangzhou, China, 18–21 August 2005; Volume 7, pp. 4506–4510.
26. Ariff, F.; Nasir, A.S.A.; Jaafar, H.; Zulkifli, A. Character segmentation for automatic vehicle license plate recognition based on fast k-means clustering. In Proceedings of the 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 9 November 2020; pp. 228–233.

27. Shivakumara, P.; Konwer, A.; Bhowmick, A.; Khare, V.; Pal, U.; Lu, T. A new GVF arrow pattern for character segmentation from double line license plate images. In Proceedings of the 2017 4th IAPR Asian Conference on Pattern Recognition (ACPR), Nanjing, China, 26–29 November 2017; pp. 782–787.
28. Rahman, R.; Pias, T.S.; Helaly, T. Ggcs: A greedy graph-based character segmentation system for bangladeshi license plate. In Proceedings of the 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Istanbul, Turkey, 22–24 October 2020; pp. 1–7.
29. Silva, S.M.; Jung, C.R. A flexible approach for automatic license plate recognition in unconstrained scenarios. *IEEE Trans. Intell. Transp. Syst.* **2021**. [[CrossRef](#)]
30. Musaddid, A.T.; Bejo, A.; Hidayat, R. Improvement of character segmentation for indonesian license plate recognition algorithm using CNN. In Proceedings of the 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 5–6 December 2019; pp. 279–283.
31. Duan, N.; Cui, J.; Liu, L.; Zheng, L. An end to end recognition for license plates using convolutional neural networks. *IEEE Intell. Transp. Syst. Mag.* **2019**, *13*, 177–188. [[CrossRef](#)]
32. Selmi, Z.; Halima, M.B.; Pal, U.; Alimi, M.A. DELP-DAR system for license plate detection and recognition. *Pattern Recognit. Lett.* **2020**, *129*, 213–223. [[CrossRef](#)]
33. Wang, D.; Wang, B.; Chen, Y. Roi-based deep learning method for variable-length license plate character segmentation. In Proceedings of the 2020 3rd International Conference on Computer Science and Software Engineering, Beijing, China, 22–24 May 2020; pp. 102–106.
34. Danilenko, A. License plate detection and recognition using convolution networks. In Proceedings of the 2020 International Conference on Information Technology and Nanotechnology (ITNT), Samara, Russia, 26–29 May 2020; pp. 1–6.
35. Zou, Y.; Zhang, Y.; Yan, J.; Jiang, X.; Huang, T.; Fan, H.; Cui, Z. A robust license plate recognition model based on bi-LSTM. *IEEE Access* **2020**, *8*, 211630–211641. [[CrossRef](#)]
36. Raza, M.A.; Qi, C.; Asif, M.R.; Khan, M.A. An adaptive approach for multi-national vehicle license plate recognition using multi-level deep features and foreground polarity detection model. *Appl. Sci.* **2020**, *10*, 2165. [[CrossRef](#)]
37. Chowdhury, P.N.; Shivakumara, P.; Jalab, H.A.; Ibrahim, R.W.; Pal, U.; Lu, T. A new fractal series expansion based enhancement model for license plate recognition. *Signal Process. Image Commun.* **2020**, *89*, 115958. [[CrossRef](#)]
38. Silva, S.M.; Jung, C.R. Real-time license plate detection and recognition using deep convolutional neural networks. *J. Vis. Commun. Image Represent.* **2020**, *71*, 102773. [[CrossRef](#)]
39. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 1–27. [[CrossRef](#)]
40. Feng, S.; Chen, C.P.; Zhang, C.Y. A fuzzy deep model based on fuzzy restricted Boltzmann machines for high-dimensional data classification. *IEEE Trans. Fuzzy Syst.* **2019**, *28*, 1344–1355. [[CrossRef](#)]
41. Tang, J.; Deng, C.; Huang, G.B. Extreme learning machine for multilayer perceptron. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 809–821. [[CrossRef](#)] [[PubMed](#)]
42. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
43. Chen, C.P.; Liu, Z. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture. *IEEE Trans. Neural Netw. Learn. Syst.* **2017**, *29*, 10–24. [[CrossRef](#)]
44. Chen, C.P.; Liu, Z.; Feng, S. Universal approximation capability of broad learning system and its structural variations. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *30*, 1191–1204. [[CrossRef](#)]
45. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28 (NIPS 2015), Montreal, QC, USA, 7–12 December 2015; Volume 28.
46. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.