Combination of CNN with GRU for Plate Recognition

Fucheng You¹, Yangze Zhao², Xuewei Wang³

¹Beijing Institute of Graphic Communication. No. 1, Xinghua Street (two section), Daxing District, Beijing, China
²Beijing Institute of Graphic Communication. No. 1, Xinghua Street (two section), Daxing District, Beijing, China
³Beijing Institute of Graphic Communication. No. 1, Xinghua Street (two section), Daxing District, Beijing, China

1263108136@qq.com, 1263108136@qq.com, 1263108136@qq.com

Abstract. License plate recognition has been a hot topic. Most of the existing license plate recognition solutions are mainly implemented through character segmentation and then recognition. However, these methods have been lacking in robustness and character segmentation has been a difficult problem to be solved perfectly. This thesis boils down the problem of character recognition to a problem of sequential learning. The convolution neural network is used for feature extraction to describe the high-level semantics of the image, and the GRU neural network is used as the sequence learning device to effectively model the internal relations of the sequence. Considering that the output sequence cannot be aligned with the input feature frame sequence, we use structured Loss. A background (Blank) category is also introduced to absorb the obfuscation of adjacent characters. The experimental training set of the paper is more than 10,000 plate data sets in the nearly real scene produced by human, and the test results of 99% accuracy can be achieved on hundreds of test sets.

1. Introduction

Automatic License Plate Recognition (ALPR) has been a frequent topic of research ((1), (2),(3)) due to many practical applications, such as automatic toll collection, traffic law enforcement, private spaces access control and road traffic monitoring. Traditional license plate recognition includes license plate character segmentation and character recognition. The main purpose of character segmentation is to segment the image of license plate character into individual characters to recognize it. (4) At present, the methods of character segmentation mainly include image vertical projection feature segmentation, matching algorithm based on template library and connected domain segmentation. The algorithm based on template library matching is too slow to meet the requirement because it needs to traverse every possible location. However, the connected domain algorithm has poor segmentation effect due to fuzzy characters, sticky character borders and close rivets. This paper combines CNN and GRU. We put the license plate recognition problem((5), (6),(7),(8),(9)) as a sequence modeling problem, through the use of sliding window slip license plate area, in order to each image of sliding window to extract the convolution features, so you can get the whole license plate image sequence of convolution characteristics, ((10),(11),(12),(13)) then using RNN (Recurrent Neural Network) Network and CTC (Connectionist Temporal Classification) method to get the final no segmentation recognition as a result, the process is simple and high in accuracy under the condition of stability. The following sections will introduce the network model, data set, results and summary of this article in turn.
2. Model Architecture

As is depicted above figure 1, Firstly, image is fed to CNN to extract image features. The next step is to apply Recurrent Neural Network to these features followed by the special decoding algorithm. This decoding algorithm takes LSTM outputs from each time step and produces the final labeling. The detailed architecture will be described below.

2.1. Feature Extraction
Model receives source image and extract image features by CNN. CNN produces tensor with shape 32*16*16. Due to equipment limitations and lack of data, we only designed smaller ones, including only two convolutional layer and two maxpooling layer. The feature map height equals 32, width equals to 8, and the depth is 16.

2.2. Reshape
Now we do reshape operation. It depends on different situations. Because we have to apply the feature to FC (fully connected layer) so we have to make feature match FC layer. we apply fully connected layer followed by softmax layer and get the vector of 32 elements. This vector contains probability distribution of observing alphabet symbols at each LSTM step. The choice will depend on the specific task.

2.3. Encoding
Encoding process transform feature vector into probability distribution. We apply GRU as encoding algorithm. The input feature vector is 32*32 and each vector of 32 elements is fed in GRU and outputs vector of 512 elements. So final shape of the early stage of GRU is (32, 512). And we feed the reshape vector into two identical GRU and add them to 1 dimension. After that we feed the addition result into another two same GRU network. This time we concatenate the two vectors of shape (32, 512). So we get output (32, 1024). The final probability distribution of observing alphabet symbols will be got by FC and activation layer.

2.4. Decoding
Decoding process is pretty simple. On the above diagram we have eight vectors of probabilities at each LSTM time step. Let’s take most probable symbol at each time step. As a result we obtain the string of eight characters—one most probable letter at each time step. Then we have to glue all consecutive repeating characters into one. In our example two “e” letters are glued to single one. Special blank character allows us to split symbols that are repeated in the original labeling. We added blank symbol to the alphabet to teach our neural network to predict blank between such case symbols. Then we remove all blank symbols. The detailed architecture is displayed as figure 2. And the result is as figure 3.
3. Dataset
On paper the experimental training sets supervise more than ten thousand copies of artificial approximate real scenarios license plate data sets, the style of the data set has been shown above. There are many Chinese characters in China, and it takes time and energy to arrange and obtain license plates. For convenience, the data set is made by taking foreign license plates as an example, namely the combination of English and Numbers. The experimental results of this paper have achieved 99% accuracy on hundreds of test sets. The images look like figure 4.

4. Result
The data set used in this experiment is artificial data set, so it is not compared with other algorithm results. The results obtained in the test set reached 99% of the test accuracy and achieved real-time monitoring effect, about 200ms per image.
5. Conclusion
This paper discusses the combination of CNN and GRU and considers the license plate recognition problem as a series modeling problem. Slip license plate area, through using the sliding window for each image of sliding window to extract the convolution features, so you can get the whole license plate image sequence of convolution characteristics, and then will get maps into LSTM encoder to encode, probability distribution vector collection to get letters, no segmentation to identify the final result by decoder implementation, process simple, and in the case of high accuracy remained stable. Finally, the test set shows superior results, indicating that the network generalization presented in this paper is good.

The future work mainly focuses on the use of Faster R-CNN for vehicle license plate detection in deep reference learning. Under the condition that the recognition accuracy remains unchanged, the location of the license plate in the image is detected. In order to make the results more persuasive, experimental data set will attempt to adopt the License Plate detection in the authoritative SSIG License Plate Character Segmentation Database data set. At the same time, in order to achieve better generalization effect and reduce computing power, we use the migration learning method to directly extract features under the picture data set of real scenes, and only train the GRU encoding and decoding process after feature extraction.

References
[1] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, “Automatic license plate recognition (ALPR): A state-of-the-art review,” IEEE Transactionson Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311–325, Feb 2013.
[2] C. Gou, K. Wang, Y. Yao, and Z. Li, “Vehicle license plate recognition based on registrations and restricted Boltzmann machines,” IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 4, pp.1096–1107, April 2016.
[3] O. Bulan, V. Kozitsky, P. Ramesh, and M. Shreve, “Segmentation and annotation-free license plate recognition with deep localization and failure identification,” IEEE Transactions on Intelligent Transportation Systems, vol. 18, no. 9, pp. 2351–2363, Sept 2017.
[4] Amin Safaei, Hongying L. Tang, Saeid Sanei. Real-time search-free multiple license plate recognition via likelihood estimation of saliency[J]. Computers and Electrical Engineering, 2016, 56.
[5] Tejendra Panchal, Hetal Patel, Ami Panchal. License Plate Detection Using Harris Corner and Character Segmentation by Integrated Approach from an Image[J]. Procedia Computer Science, 2016, 79.
[6] C. Gou, K.-F. Wang, Y.-J. Yao et al., "Vehicle license plate recognition based on extremal regions and restricted Boltzmann machines", IEEE Trans. Intell. Transp. Syst., vol. 17, no. 4, pp. 1096-1107, 2016.
[7] O. Bulan, V. Kozitsky, P. Ramesh, M. Shreve, "Segmentation-and Annotation-Free License Plate Recognition With Deep Localization and Failure Identification", IEEE Transactions on Intelligent Transportation Systems, no. 99, pp. 1-13.
[8] Samuel A. Babatunde. Review of strengthening techniques for masonry using fiber reinforced polymers[J]. Composite Structures, 2017, 161.
[9] Jaber Shabanian, Pierre Sauriol, Jamal Chaouki. A simple and robust approach for early detection of deflu dization[J]. Chemical Engineering Journal, 2017, 313.
[10] Khalid Aboura, Rami Al-Hmouz. An Overview of Image Analysis Algorithms for License Plate Recognition[J]. Organizacija, 2017, 50(3).
[11] Martijn H.J. Huismans, Mark van Heijl, R. Marijn Houwert, Tim K. Timmers, Ger van Olden, Egbert Jan M.M. Verleisdonk. Anteroinferior versus superior plating of clavicular fractures[J]. Journal of Shoulder and Elbow Surgery, 2016, 25(3).
[12] Yu Wang, Xiaojuan Ban, Jie Chen, Bo Hu, Xing Yang. Corrigendum to “License plate recognition
based on SIFT feature” [Optik 126 (2015) 2895–2901][J]. Optik - International Journal for Light and Electron Optics, 2016, 127(1).