Vehicle Plate Number Identification Available Models

DL-Project
Supervisor, Hantao Zhang, The University of Iowa, Iowa City, IA 52242

Nov 4, 2018

1. Convolutional Neural Networks

Convolution Neural Networks is one of the most deep learning models, it can be trained via back propagation algorithms. In recent years, it outstanding ability in image detection has attract lots of attention. Thus as a detection task, CNN is one of our selected models for vehicle detection.

The intended our CNN structure is composed of three convolution, three pooling and two fully-connected layers, slightly modified based on the structure proposed by Stark et al. (2015). The last layer outputs the probability distributions for all digits for which we can compute the prediction and uncertainty of the prediction.

For the reason that the gathered car template in our dataset is coming from European car template, which composed by 7 valid information position with upp-letters or digital numbers(0-9). Thus for each position, we have $26 + 10 = 36$ possible choices. We use 0-1 vector to represent the target for each template position. In this way, the output layer of our CNN models has $36 \times 7 = 252$ nodes.

Furthermore, as our vehicle plate number identification can be identified as a classification task, the last layer outputs the probability for all upp-letters and digits. To predict a digit, we consider the corresponding 36 neurons and normalize their sum to 1. The loss function is defined in terms of the cross-entropy between the label and the network output. For numerical stability the activation functions of the final layer are rolled into the cross-entropy calculation using softmax_cross_entropy_with_logits and sigmoid_cross_entropy_with_logits.

2. Recurrent Neural Networks and LSTM

Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition. Since the car template can also be viewed as a sequence with 7 characteristics thus RNN can be applied in our task.

All RNNs have feedback loops in the recurrent layer. This lets them maintain information in memory over time. But, it can be difficult to train standard RNNs to solve problems that require learning long-term temporal dependencies. This is because the gradient of the loss function decays exponentially with time (called the vanishing gradient problem). To solve this problem, LSTM units which includes a 'memory cell' that can maintain information in memory for long periods of time has been proposed. A set of gates is used to control when information enters the memory, when it's output, and when it's forgotten. This architecture lets them learn longer-term dependencies.

Li and Shen (2016) proposed a structure that based on BRNN sequence labeling. Due to the computational ability and time restrict, we simplified their method with multi-layer RNN and LSTM.

References

Hui Li and Chunhua Shen. Reading car license plates using deep convolutional neural networks and lstms. *arXiv preprint arXiv:1601.05610*, 2016.

Fabian Stark, Caner Hazırbas, Rudolph Triebel, and Daniel Cremers. Captcha recognition with active deep learning. In *GCPR Workshop on New Challenges in Neural Computation*, volume 10, 2015.