Automatic License Plate Detection Using Deep Learning Techniques

Bharath P Hegde, Mahalakshmi, S.
Information Science Engineering, BMSIT Bangalore, India
Corresponding author: *E-mail: bharathasuke@gmail.com

**Abstract:** Automatic License Plate Recognition (ALPR) systems capture a vehicle’s license plate and recognize the license number and other required information from the captured image. ALPR systems have numbers of significant applications: law enforcement, public safety agencies, toll gate systems, etc. The goal of these systems is to recognize the characters and state on the license plate with high accuracy. ALPR has been implemented using various techniques. Traditional recognition methods use handcrafted features for obtaining features from the image. Unlike conventional methods, deep learning techniques automatically select features and are one of the game changing technologies in the field of computer vision, automatic recognition tasks and natural language processing. Some of the most successful deep learning methods involve Convolutional Neural Networks. This technique applies deep learning techniques to the ALPR problem of recognizing the state and license number from the USA license plate. Existing ALPR systems include three stages of processing.

**Key words:** ALPR, Deep Learning, Convolution Neural Networks (CNN)

**Received on:** 07-04-2017
**Accepted on:** 27-05-2017

**Introduction**
Recognition of an on-road vehicle using its license plate is an important task performed by several intelligent transportation systems around the world. This task is known as Automatic License Plate Recognition (ALPR) and plays an important role in many real application scenarios such automatic toll collection, access control in private parking lots, stolen vehicles identification and traffic surveillance.

Recently new approaches have been proposed to perform ALPR in an efficient way. However, there are still many problems that can be explored using modern techniques, example simultaneous recognition of multiple vehicles and vehicle recognition in low-light environments and in high-speed highways with low quality samples.

ALPR approaches are commonly subdivided into multiple smaller and simpler tasks that are executed sequentially

(i) Image acquisition
(ii) Vehicle location
(iii) License plate detection
(iv) Character segmentation
(v) Optical character recognition (OCR)
Although some approaches perform vehicle tracking, they do not use all captured information to recognize the characters. Instead, they select only a single frame to perform the recognition, based on some defined rule making the method more sensitive to noise and prone to recognition errors.

The system proposed uses an real-time approach to perform ALPR. One of the main concerns is to avoid the need to embedding high-cost computers on the highways. This could make the system unfeasible to be employed in the real-world applications. Although there are some works in the literature providing outstanding results in computer vision tasks using techniques based on deep learning these are too computationally expensive and need computers with high processing power, usually provided by GPU cards. In addition, they need huge set of examples for training. Therefore, the proposed system does not intend to utilize Deep Learning approaches in our ALPR system.

This work proposes a temporal redundancy approach to perform ALPR based on multiple frames instead of selecting only a single frame that can be executed in real-time. Whereas Deep Learning is a well-known technique in the machine learning community, to the best of knowledge, the main contributions of this work can be pointed as follows:

- A new real-time framework to perform ALPR using spatiotemporal information
- Two post-processing techniques to improve the final accuracy of the ALPR system
- A public dataset of vehicles classified/labelled according to their appearance.

2. Deep learning vs conventional learning
Deep learning techniques are powerful machine learning models that accomplish fabulous performance on the learning problems like object recognition and speech recognition.

Conventional learning extracts features from the available input and then classifies features from the previous step with a classifier. Feature extraction is a procedure of examining and getting valuable data from the given input data. Some feature extractors in the field of computer vision are scale-invariant feature transform (SIFT) and histogram of oriented gradients (HOG). A typical learning methodology is a handcrafted feature extractor, where the features are selected manually and the selection of features is a difficult task which depends on the application. Designing good feature extractors is a painful job.

![Figure 2.1: Block diagram of conventional learning mode](image)

By contrast, deep learning neural networks can be trained using a labeled training set which have the potential to compute the network parameters. Deep learning models surpass the capabilities of conventional learning models. Deep learning techniques extract the relevant features from input data and learn from the training set itself. Fig. 2.2 illustrates a typical deep learning model.

3. Experimental study
3.1. Preprocessing of license plate
Preprocessing helps to enhance the visual appearance of the license plate and
removes the noise or other unwanted distortion. Preprocessing is generally an important step in image processing systems, and it helps computationally in further stages of the system. We use anisotropic diffusion and histogram equalization to achieve higher quality segmentation.

3.1.1. Anisotropic Diffusion
Anisotropic diffusion (AD) aims at lowering the image noise without removing significant information from the image such as edges, lines or other details. Anisotropic diffusion is one of the pioneering works in reducing the noise using partial derivatives. AD smooths the texture successfully without degrading the boundaries and small structures within the image. Unlike conventional smoothing filters, anisotropic diffusion is like an adaptive technique which smooths the image inside a region, but leaves the boundaries or untouched edges.

3.1.2. Histogram Equalization
Histogram equalization improves the contrast of the input image. This is important to ALPR because it helps in enhancing the image when most of the pixels in the image are confined to region. Basic histogram equalization uses global contrast which is not a good one for all the images.

3.2. Binarization
Binarization is the process of converting a gray scale image to a binary image. The image consists of pixel values 0 or 1, where 0 indicates black and 1 or 255 indicates white. Binarization is an important step in improving the quality of extraction of license number and state portion from the number plate. Proper binarization method is necessary for separating the foreground and background pixels in the image. Primary step in binarization is selection of optimal threshold value. A pixel with value less than threshold is classified as background pixel and with value greater than as foreground pixel.

4. Convolution neural networks
The Convolutional Neural Network (CNN), first proposed by LeCun in 1988, is a neural network model incorporating the following ideas: receptive fields, weight sharing and subsampling. It is a special type of multilayer perceptron trained in supervised mode using backpropagation. It is one of the most successful machine learning architectures in computer vision and has achieve state-of-the-art results in tasks as character recognition, object recognition, face detection and pose estimation, speech recognition, license plate recognition, image preprocessing and segmentation tasks. A CNN learns complex, high dimensional features from a large number of examples which makes it an obvious candidate for pattern recognition tasks. CNN architectures ensure some degree of shift, scale and distortion invariance using some of the features such as local receptive fields, subsampling, shared weights etc. Unlike conventional pattern recognition tasks one of the benefits of CNN’s lie in extraction features itself and which uses images as input for training, testing the network. The CNN builds complex features from large collection of examples and the complexity of learning features with more layers. This is done by successively convolving the input image with filters to build up a hierarchy of feature maps. The hierarchy results in complex features and learning, as well as translational and distortion invariant. The whole CNN can be expressed as a score.
function where raw image pixels are given as input on one end and determine the category or class score at the other end.

4.1. Convolutional layer
As previously mentioned, a convolutional layer takes an input image and convolves it with kernels to produce several two-dimensional planes of neurons called feature maps. Each element of a feature map is obtained by convolving the respective kernel with units in the neighborhood in the previous layer. These outputs obtained after each convolutional layer are then summed up together with a trainable bias term which is then passed to an activation function to obtain each unit of a feature maps. The input image could consist of hundreds or thousands of pixels but convolutional layers act as feature extractor to extract features such as corners, edges, endpoints or nonvisual features by convolving the input with kernels consisting of weights. As the weights are shared, the number of parameters to train the neural network are reduced. This also reduces the memory necessary to store these parameters during execution. The convolution operation in each convolutional layer makes a CNN translational and distortion invariant i.e. when the input image is shifted the output feature map will be shifted in the same amount as input. The number of kernels in each convolution layer depends upon the number of feature maps and varies from architecture to architecture.

4.2. Performance optimization of the Convolution Neural Network
There are several techniques that can improve the performance of a CNN. The optimization techniques help in following ways
• Avoiding the local minima
• Prevent overfitting
• Improving accuracy or reducing error rate

There are several factors that affect the Convolution Neural Networks performance as discussed below:
1. Datasets
2. Back propagation
3. Cross Validation
4. Weight Initialization
5. Dropout Method
6. Number of convolution Layers

5. Results and discussions
As discussed, segmentation is one of the important steps of our system. The overall accuracy of the system is heavily influenced by this stage. Difficulties faced in this stage are
1. Unwanted objects such as handicapped symbols, state symbols
2. Illumination problems
3. Blurry images

Problems arose while dealing with the degraded images. When binarized the characters are overlap and extraction of characters fails most of the time.

The segmentation is evaluated by testing the segmentation algorithm on a database of 395 images. The database of images also includes degraded images. To compute the accuracy, 394 are considered. Out of 394 images 372 images are segmented satisfactorily and 22 images failed. Some of the failed images segmented partially. This gives an accuracy of 94.4% for license number segmentation. The state portion has been extracted correctly for most of these images.
### Table 1: Accuracies of different Implementations

<table>
<thead>
<tr>
<th>SL No</th>
<th>Type of system</th>
<th>Datasets</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Open ALPR</td>
<td>Open Source</td>
<td>93.44%</td>
</tr>
<tr>
<td>2</td>
<td>Open ALPR</td>
<td>Closed API</td>
<td>94.55%</td>
</tr>
<tr>
<td>3</td>
<td>2-Convolution Layer CNN</td>
<td>Open Source</td>
<td>92.14%</td>
</tr>
<tr>
<td>4</td>
<td>3-Convolution Layer CNN</td>
<td>Open Source</td>
<td>92.32%</td>
</tr>
<tr>
<td>5</td>
<td>4-Convolution Layer CNN</td>
<td>Open Source</td>
<td>97.92%</td>
</tr>
</tbody>
</table>

### 6. Conclusion

#### 6.1 Conclusion

This thesis has investigated the use of deep learning techniques in the field of automatic license plate detection system. We proposed a system which consists of two stages: character segmentation and character recognition. An important aspect of this system is to incorporate a hybrid binarization technique which helps to improve quality of segmentation. This technique helps in the segmentation of degraded images.

Another important aspect incorporated in this research is using convolutional neural networks for feature extraction and license plate recognition. The recognition task includes recognition of license number and state information from the license plate. Our study also includes how the data format impacts the recognition system and different techniques to optimize convolution neural networks. Our experiments show an overall accuracy of 81.1% using gray scale images. In general training of cnn takes 1-6 days, which depends on the neural network architecture. Our results show the usage of deep learning techniques in the field of ALPR system. This system can be used for practical use by the following improvements.

1. Usage of real time data to train the convolution neural networks. Real time data includes images collected from the toll gates, parking lots and other agencies etc.
2. Inclusion of localization of license plate, de-skewing stage.
3. Localizing and segmenting the state information automatically.

#### 6.2. Limitations

The system shows satisfactory results in segmentation stage and recognition stage. There are certain limitations on the system as it is developed with certain assumption. The limitations are as follows:

1. The state information is assumed to be at upper portion of license plate. The license plate with state information at lower portion of license plate fails, but recognizes license number successfully.
2. For binary images ‘0’, ‘o’ and ‘q’ are considered as same class.
3. Recognition of unwanted symbols is ignored in this research and is included in future study.
4. ‘0’ and ‘o’ are considered as same class for gray scale images because of similar geometric structure.
5. Need for more data.

#### 6.3. Future work

The automatic license plate recognition system proposed in this research has several limitations. Most major being that the state information position is assumed to be at top part of license plate. Though most of the plates consists of state information at the upper part of license plate, the proposed system will not be able to recognize the state information if the position of state information is changed. Inclusion of recognizing
unwanted symbols improves the accuracy of the system. Future work includes localization and detection of state information. It also might locate and detect the license number from the license plate. Future work also includes implementation of license plate segmentation using deep learning techniques, expected to have good accuracy. Implementation of an alpr system completely using deep learning techniques might consume lot of execution time and is challenging due to limited availability of datasets. Deep learning techniques are providing efficient solutions in the field of artificial intelligence, pattern recognition etc with the incorporation of graphic processing units.

Acknowledgment
I would like to profoundly thank Management of BMS Institute of Technology & Management for providing such a healthy environment for the successful completion of paper work. It gives me immense pleasure to thank Dr. Manjunath T. N. Head of the Department for his constant support and encouragement. Also, I would like to express my deepest sense of gratitude to my guide Mrs. S. Mahalakshmi Assistant professor, Department of Information Science & Engineering for her constant support and guidance throughout the paper work.

References