Multilingual and Skew License Plate Detection Based on Extremal Regions

Rongqiang Qian†, Bailing Zhang*, Frans Coenen† and Yong Yue*
†Department of Computer Science
University of Liverpool
Liverpool, L69 3BX, United Kingdom
*Department of Computer Science and Software Engineering
Xi'an Jiaotong-Liverpool University
Suzhou, P.R.China, 215123

***Abstract*—License Plate Detection (LPD) is an important component of many applications involving security and traffic surveillance. Despite current progress, lots of hurdles remain in the way of a robust LPD system. This is particularly true for the detection of license plates with different layouts and skew angles. In this paper, a novel LPD system is proposed for detecting English and Chinese license plates with large skew angles. The proposed system consists of three main stages: 1) a character proposal module to find candidate characters based on Extremal Regions (ERs); 2) feature extraction and classification relies on Convolutional Neural Network (CNN); 3) license plate detection by region linking. The new method much improves on the robustness of existing approaches by leaving out character segmentation. The performance of the proposed license plate localization algorithm is verified using different datasets of vehicle images, including a large field-captured dataset, a skew and tilt dataset, a 12 countries dataset and two benchmark datasets. For Chinese civilian vehicles, the accuracy of plate localization is over 98.3%.**

***Keywords:*—License plate Detection, multilingual license plate, skew license plate, extremal regions, convolutional neural network.**

I. Introduction

License Plate Detection (LPR) technology enables automatic detection and recognition of the registration numbers of vehicles via digital imaging. It is a key component for many security and intelligent transportation systems. Examples include car park access control, electronic highway toll payment systems, suspect vehicle analysis and tracking, and automatic identification of expired registrations. LPR has been studied for many years [1], [2], [3], with a variety of methods proposed for different kinds of license plates. Several review papers provide an important source of information [4], [5], [6], [7].

There is a great deal of variability between license plates in different countries with regards to the design, colors, characters used, and layout. Previously published LPR work has been mainly centered upon a small number of the countries, including: Australia, Taiwan, Japan, Korea, certain European countries, and mainland China. On the other hand, most of the reported LPR systems were based on a common structure, *i.e.*, the consecutive composition of license plate detection, character segmentation, and recognition. Effective segmentation is a critical step for LPR as the last step character classification depends largely on the quality of the segmentation output. However, segmentation is difficult in practice, especially when dealing with inherently noisy, low spatial resolution images such as those produced in various poor weather conditions. The most common practice of character segmentation is based on histogram analysis and thresholding [4], [5], [6], [7].

In summary, there are still many challenges to be met in order to produce a robust LPR system that able adapt to the variability of the environments and different demands. There exists many real-world difficulties. For example, a captured vehicle image could be at a poor level of resolution due to the large distance between camera and the vehicle, or the poor lighting and low contrast which may in turn come from overexposure, reflections or shadows. Part of the plate can be obscured due to discoloration or dirt. Variations in the angle between a camera and the vehicle can produce perspective projection distortion of a license plate. Compared with most western countries, Chinese license plates are more challenging because of the complex combination of different characters on a license plate, including Chinese characters, English letters and digits, and many sub-categories of license plates with distinct colors and layouts.

The main motivation of this paper is to present a multilingual and skew license plate detection system with main contributions include:

- A novel algorithm which simultaneously completes LPD and character segmentation.
- A multilingual LPD system that able to detect both English and Chinese license plate.
- A robust LPD system that able to detect extreme skewed license plate.

The rest of this paper is organized as follows: Section 2 presents the detailed introduction of the proposed LPD system; Experimental results will be provided in Section 3, followed with conclusion in Section 4.

II. Approach

The overall LPD detection system is demonstrated in Fig. 1. In the first stage, a large number of candidate regions that may contain license plate characters are generated by region generators. In the second stage, all the generated regions are passed to a CNN for feature extraction and classification. In the third stage, license plates are detected based on a region linking strategy.

Fig. 1. System overview

A. Region Proposal

ERs [8] are connected components of an image binarised at a certain threshold. In our system, an input image (RGB) is converted into a set of binary images by applying pre-set thresholds, followed with connected component analysis for generating ERs. In order to reduce computational cost, some regions with small areas ($< 10 \sim 20$ pixels) are ignored. The main motivation of generating different binary images and subsequent regions is to compensate for the negative effects of varying brightness and contrast, therefore, high recall rete can be achieved in this stage.

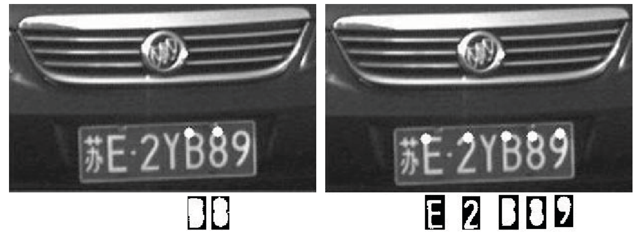
B. Classification

The network applied in this stage is similar to the famous VGG networks [9], which is illustrated in Fig. 2. The network consists of four convolution layers followed by fully connection layers and softmax layer. Each convolution layer followed with non-linear activation layer and max pooling layer. ReLU [10] is employed as the activation function for convolutional layers and full connection layers. Dropout [10] is adopted for preventing over-fitting. The final softmax layer has 2 outputs, corresponding to character and background. All the input regions are firstly resized to $48 \times 48 \times 3$. After this classifier, the label of regions are achieved, and then all the positive and negative regions are passed to next stage.

C. Region Linking

The layout of license plates were specially designed in most countries. For example, all the Chinese license plates are composed of Chinese characters, digits and English letters. For the majority of civilian vehicles, the first character is Chinese, representing provincial level divisions. To the right of the Chinese character is Latin alphabet character representing the municipality or county. The remaining part is a combination of five numbers or letters. Therefore, in order to detect a Chinese license plate, it is sufficient to only consider for non-Chinese characters to position the letters and digits on the plate, and then infer the complete plate location based on the prior knowledge about the layout of Chinese license plates.

For each binary threshold used in region proposal stage, a set of bounding boxes are obtained and denoted as $\text{Rect}_n^r = \{x_n^r, y_n^r, w_n^r, h_n^r\}$. The positive regions from previous stage can also generate a set of bounding boxes $\text{Rect}_m^c = \{x_m^c, y_m^c, w_m^c, h_m^c\}$. For example, there are 2 filtered regions in Fig. 3(a) and 5 filtered regions in Fig. 3(b) respectively. The region linking procedure can be described in the following steps.



(a) (b)

Fig. 3. Examples of filtered regions.

(1) Assume the position and size of current region (hypothesis character) defined by a rectangle $(x_m^c, y_m^c, w_m^c, h_m^c)$, where (x, y) is the coordinate of top left corner of the rectangle, and w and h are corresponding width and height. The next region, similarly denoted as $(x_n^r, y_n^r, w_n^r, h_n^r)$, is localized as in the following.
   (a) Search towards the right for a matched region with similar height, with $|h_n^r - h_m^c| < 0.5h_m^c$, and the horizontal and vertical differences between the two regions, i.e., $|x_m^c - x_n^r|$, $|y_m^c - y_n^r|$, being less than the height of current hypothesis character. If such a region is found, the searching process will continue. Otherwise, move to next step.
   (b) Search towards the left for a matched region with similar height, with $|h_n^r - h_m^c| < 0.5h_m^c$, and the horizontal and vertical differences between the two regions, i.e., $|x_m^c - x_n^r|$, $|y_m^c - y_n^r|$, being less than the height of current hypothesis character. If such a region is found, the searching process will continue. Otherwise, move to next step.

   Some examples are provided in following Fig. 4 to further explain the above two steps. In Fig. 4, four examples are used to demonstrate the detection processes. In Fig 4(a), the algorithm first searches towards the right with two similar regions found, and then searches towards the left finding three similar regions. In Fig. 4(b), the algorithm first searches towards the right with one similar region found, and then searches towards the left with four similar regions located. In Fig. 4(c), the algorithm first searches towards right finding five similar regions, and then search towards the left without detecting any similar regions. In Fig. 4(d), the algorithm first searches towards right without finding any similar region, and then searches towards the left with five similar regions detected.
(2) Count the number of regions. If the number is less than six, then the current region is not a character on license plate. If the number is six, and the spacing between the leftmost character and its right neighbor conforms to the standard, it can be concluded that the last six characters are from a license plate. If the number is larger than six, then attempt to find a region such that the spacing

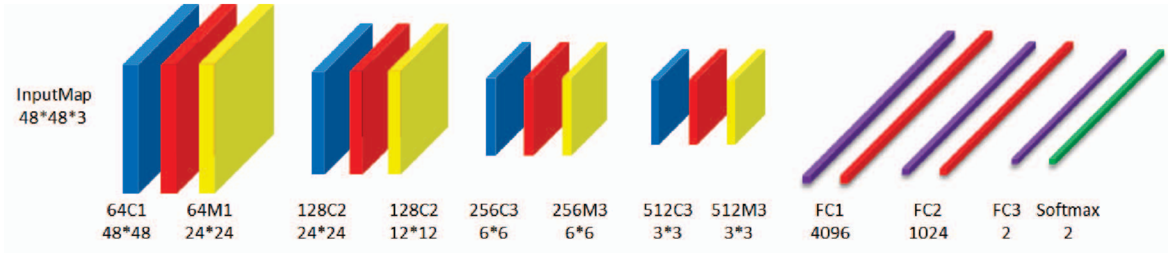
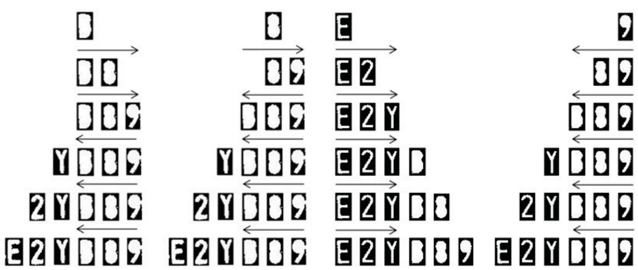


Fig. 2. Network architecture

between it and the right neighboring character conforms to the standard. If such a region exists, the region together with its right-side five consecutive characters will be deemed as the characters on a license plate. Otherwise the current region is not a character on license plate.

(3) The position of the last six consecutive characters will be used to infer the position of Chinese character on the plate.

(4) Since a same license plate may be detected several times by using different regions, it is necessary to remove those duplicate license plates. In our algorithm, this is simply implemented by removing the redundant license plates with the same position.



(a) (b) (c) (d)

Fig. 4. Explanation of the inference process for a hypothesis license plate. (a) and (b): license plate detection starts from B and 8. (c) and (d): license plate detection starts from E and 9.

III. Experiment

A. Implementation Details

To thoroughly evaluate the performance of proposed LPD system, a computer with CPU Xeon 3.3GHz and 16GB memory was employed. The program runs on a 64-bit Windows system with CUDA 7.0, Matlab 2015a and MatConvNet [11] installed.

B. Field-captured Dataset

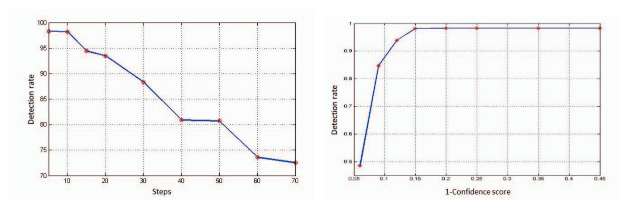
The first set was provided by police, which were field-captured from surveillance cameras at different traffic intersections in Suzhou, China. This image set consists of 1039 different type of images with size 1360×1024 pixels. The distribution between the daytime and nighttime capture time of the images was 85% and 15%, as shown in Table I.

TABLE I
Data distribution of field-captured dataset.

	Car	Bus	Truck	Van
Daytime	244	233	172	230
Nighttime	18	80	42	20

As the LPD algorithm depends on two parameters, namely, the interval of thresholds for binarization and the confidence scores of CNN. Therefore, we evaluate the performance based on these two parameters.

As explained, the chief motivation for generating multiple binary images from a single original image with different thresholds is to compensate for the large illumination variability, particularly due to the change from day to night. The practice will also minimize the negative effect from different background colors of the license plates (white, yellow and blue). From our experience, a range from 10 to 240 is sufficient for threshold adaptation. In other words, the minimum and maximum values of the threshold are 10 and 240 respectively, within which the threshold will be recursively updated for binarization. We also experimented with different steps for the looping from minimum value to maximum value. More specifically, a step of 5 means that 46 thresholds, i.e., $10, 15, 20, \cdots, 230, 235, 240$ will be applied to generate 46 binary images for the subsequence template matching, while a step of 70 means only four thresholds, namely, $10, 80, 170, 220,$ will be exploited. With the 1039 sample images, we recorded the detection rates with the different steps; the results are illustrated in Fig. 5(a).



(a) (b)

Fig. 5. Detection rates with different parameters. (a) Illustration of the detection rates with the changes of steps in the range [10,240]. (b) Illustration of the detection rates with changes of confidence scores of CNN.

It is obvious that relatively small steps of $5 \sim 10$ bring high detection rates ($> 98\%$). In this range, the performance is almost flat, steps smaller than 5 will not bring any improvement while calculation cost is raised rapidly. When the step size increases (> 10), the performance drops significantly. Consequently, an empirical value of 10 is a feasible step. Confidence score is another important factor that critically influences the detection performance. The detection rates under different confidence scores are demonstrated in Fig. 5(b). The results show that a score 0.8 is a reasonable value. When score larger than 0.8 is adopted, many characters are filtered out, thus giving low detection rate. On the contrary, score smaller than 0.8 also decrease the detection speed because regions that do not correspond to any characters would be treated as positive.

With the above empirical step of 10 for image binarization and confidence score of 0.8 for CNN, the detection results corresponding on field-captured dataset are given in Table II. The accuracies of LPD are about 98.3% and 98.1% for daytime images and nighttime images respectively. This demonstrates the advantage of the proposed algorithm, *i.e.*, changes in the illumination do not apparently affect the performance.

The overall localization rate of over 98.3% is indeed encouraging. The result also indicates that the performance does not change much for different types of vehicles. The main reasons for the 1.8% of images which are failed in the experiment are: (i) the two neighboring characters become connected after binarization; (ii) part of a character is missed; or (iii) the boundary of the characters is obscured. These problems are generally from vehicles with rusty plate or dirty plate.

TABLE II
RESULTS OF FIELD-CAPTURED DATASET.

	Daytime	Night	Overall Rates (%)
Car	98.0	94.4	97.7
Bus	99.1	98.9	99.0
Truck	95.9	100	96.7
Van	99.6	95.0	99.2
Overall Rates (%)	98.3	98.1	98.3

C. LP and Caltech Cars 1999 Dataset

To further verify the advantages of our proposed LPD algorithm, the comparison experiments were conducted with four published LPD algorithms on two benchmark datasets, the LP dataset [12] and Caltech Cars 1999 [13]. The first dataset contains 410 Chinese vehicle images and bears varied imaging conditions such as resolution, illumination and viewing angles. The second dataset has 126 images, each containing a U.S. license plate with a cluttered background.

The first compared algorithm is the Principal Visual Word (PVW) method [12], which locate license plates by principal visual word, discovery and local feature matching. Other three algorithms were cited and compared in [12], including hybrid license plate extraction based on line detection and the construction of weighted edge map, denoted as HLPE, license plate detection in coarse-to-fine based on vertical edge detection and mathematical morphology, denoted as LPE , and license plate detection based on edge statistics and morphology, denoted as ESM.

In [12], the metric of evaluation for the two benchmark datasets was defined by (i) *high level-true*, *i.e.*, license plate is totally encompassed by the bounding box and $A \cap B / A \cup B \geq 0.5$, where A is the detected region and B is the ground truth region; (ii) *low level-false*, *i.e.*, the license plate is totally missed by the bounding box; and (iii), *middle level-partial*, namely the remaining results excluded by the above two types.

For proposed algorithm, the above metric needs to be revised because our LPD algorithm aims at simultaneously detecting individual characters and the whole plate. The metric is redefined as (i) *high level-true*, which means license plate is completely localized, and all of the characters are segmented; (ii) *low level-false*: the license plate is totally missed by the bounding box; and (iii) *middle level-partial*, which refers to the situation that only part of the license plate has been localized and the segmented characters are incomplete.

The accuracy definition [12] of *True*, *Partial*, *False* and *False Positive Rate (FPR)* are formulated as

$$
\begin{aligned}
True &= \frac{TP}{TP + Partial_{TP} + FN + FP} \\
Partial &= \frac{Partial_{TP}}{TP + Partial_{TP} + FN + FP} \\
False &= \frac{FP + FN}{TP + Partial_{TP} + FN + FP} \\
FPR &= \frac{FP}{TP + Partial_{TP} + FP}
\end{aligned}
$$

where the TP is the true positive number, $Partial_{TP}$ stands for partial true positive number, FP denotes false positive number, and FN is false negative number (miss detection).

With above metric of evaluation, the comparison results are summarized in Tables III and IV, from the LP dataset and Caltech Cars 1999 datasets respectively. For the LP dataset, it is obvious that proposed algorithm performs the best in terms of the accuracy. Both the partial and false detection rates are lower than all the other four algorithms. However, the FP 3.4% from our algorithm is higher than the result 1.0% from the PVW method. This is mainly because many vehicle images in the LP dataset have advertisement or telephone numbers painted on the vehicles, which are prone to be detected as license plates. For the Caltech Cars 1999 dataset, the true detection rate of our approach is 88.4%, which is higher than all the other methods compared. The partial detection rate 6.2% is a little high, because of the low resolution of images in Caltech Cars 1999 dataset.

TABLE III
COMPARISON RESULTS OF LP DATASET

Approach	Accuracy			FPR
	True	*Partial*	*False*	
HLPE	80.8%	6.6%	12.6%	12.6%
LPE	84.6%	1.5%	13.9%	7.6%
ESM	74.6%	7.3%	18.1%	17.9%
PVW	93.2%	0.3%	6.5%	1.0%
Proposed	94.3%	0%	5.7%	3.4%

TABLE IV
COMPARISON RESULTS OF CALTECH CARS 1999 DATASET

Approach	Accuracy *True*	*Partial*	*False*	FPR
HLPE	61.6%	9.8%	28.6%	28.6%
LPE	58.0%	1.8%	40.2%	29.5%
ESM	68.7%	2.7%	28.6%	25.9%
PVW	84.8%	1.8%	13.4%	4.5%
Proposed	88.4%	6.2%	5.4%	2.4%

Some images in the LP dataset were captured with skew and tilt angles. As shown in Fig. 6 (top), our algorithm can successfully detect the license plates. Shadow was another negative factor that may affect LPD performance. For weak shadow, there will be no influence on detection result as demonstrated by the examples in Fig. 6 (middle). However, when the shadow increases, it may causes detection failure. As illustrated in Fig. 6 (bottom), a plate with strong shadow result in incomplete characters.

Fig. 6. Examples of LP dataset. Top: skew license plates, middle: weak shadow license plates, bottom: strong shadow license plates.

D. Skew and Tilt Dataset

In many situations, there exists considerable variation in the capturing angles between a camera and the target license plate, which will result in distorted license plate images. The license plate regions obtained from such images are far from perfect rectangles. Most of the previously proposed LPD methods failed to detect such distorted plates. A possible solution is perspective rectification as discussed in many computer vision problems, which is however computational intensive in finding a transformation matrix that may rectify the perspective projection distortion.

To evaluate the performance of the proposed algorithm for skewed and tilted license plates, we collected a set of vehicle images with controlled capturing angles as illustrated in Fig 7.

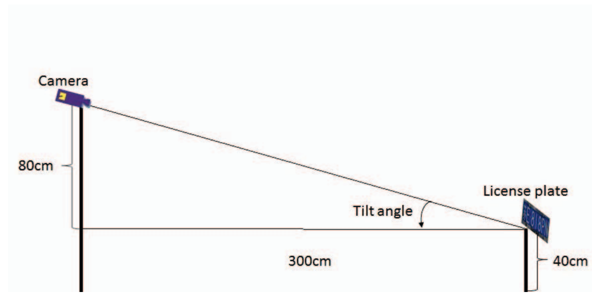


Fig. 7. Explanation of capturing system for collecting images with skewed and tilted license plates.

Fig. 8. Examples of skew and tilt dataset. Top: skew degree $\pm 70^o$, bottom: skew degree $\pm 75^o$.

To change the tilt degrees, the camera was fixed on a tripod and the distance to ground is around $120cm$. As the perpendicular distance between the license plate and ground is $40cm$ and the horizontal distance between camera and plate is $300cm$, the tilting angle between the camera and plate is around 15 degree. To obtain vehicle images with different skew angles, we define the angles towards the left as negative and the viewing angles towards the right as positive. The sampling angles are -75^o,-70^o,-65^o,-60^o,-50^o,-40^o,-30^o,-20^o,-10^o, 0^o, 10^o, 20^o, 30^o, 40^o, 50^o, 60^o, 65^o, 70^o, 75^o. For each angle, three images were shot with tilting degrees 14^o, 15^o and 16^o.

In the 57 skew and tilt images, 54 of license plates were correctly detected, which means 94.7% detection rate. Our system failed to detect license plate when the skew degree raised to 75^o (1 failed) and -75^o (2 failed), respectively. This indicates that the proposed algorithm is tolerant to skew degree upto $\pm 75^o$ when the tilting degree is around 15^o. Some of the examples are displayed in Fig. 8.

E. 12 Countries Dataset

The majority of the published work on LPD algorithms were proposed for a particular type of license plate in a specific country or region. This lack of extendibility is well-known problem in LPD. The method proposed in this paper is able to overcome this problem due to the flexible design of the connected component linking stage.

In this section, we applied the proposed LPD algorithm to a dataset found from the Internet, it contains of 171 licensed vehicle images, which belong to 12 different countries as illustrated in Fig. 9. The detection results are given in Table V, the total detection rate is 96% on this 12 countries dataset.

Fig. 9. Examples of images from 12 different countries, including Australia, Austria, Canada, Croatia, France, Germany, Israel, Italy, Spain, Portugal, UK, and USA.

TABLE V
RESULTS OF 12 COUNTRIES DATASET.

	Australia	Austria	Canada	Croatia
#Images	10	11	10	13
#Detected	9	10	10	13
	France	Germany	Israel	Italy
#Images	11	20	7	11
#Detected	10	20	7	10
	Spain	Portugal	UK	USA
#Images	10	10	11	47
#Detected	10	10	10	45

IV. CONCLUSION

In this paper, a robust LPD system is proposed for English and Chinese license plates with large skew angles. The main contributions include: (i) a novel algorithm which simultaneously completes LPD and character segmentation; (ii) a multilingual LPD system that able to detect both English and Chinese license plate; (iii) a robust LPD system that able to detect extreme skewed license plate. Extensive experiments have been presented with different kind of datasets, yielded convincing results.

REFERENCES

[1] P. Comelli, P. Ferragina, M. N. Granieri, and F. Stabile, "Optical recognition of motor vehicle license plates," *IEEE Transactions on Vehicular Technology*, vol. 44, no. 4, pp. 790–799, Nov 1995.

[2] R. Zunino and S. Rovetta, "Vector quantization for license-plate location and image coding," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 1, pp. 159–167, Feb 2000.

[3] T. Naito, T. Tsukada, K. Yamada, K. Kozuka, and S. Yamamoto, "Robust license-plate recognition method for passing vehicles under outside environment," *IEEE Transactions on Vehicular Technology*, vol. 49, no. 6, pp. 2309–2319, Nov 2000.

[4] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos, and E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 3, pp. 377–392, Sept 2006.

[5] C. N. E. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still images and video sequences: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 377–391, Sept 2008.

[6] C. N. E. Anagnostopoulos, "License plate recognition: A brief tutorial," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 1, pp. 59–67, Spring 2014.

[7] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (alpr): A state-of-the-art review," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 2, pp. 311–325, Feb 2013.

[8] J. Matas and K. Zimmermann, "A new class of learnable detectors for categorisation," in *Image Analysis: 14th Scandinavian Conference, SCIA 2005, Joensuu, Finland, June 19-22, 2005. Proceedings*, H. Kalviainen, J. Parkkinen, and A. Kaarna, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 541–550.

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *2015 International Conference on Learning Representations (ICLR)*, 2015.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.

[11] A. Vedaldi and K. Lenc, "Matconvnet: Convolutional neural networks for matlab," in *Proceedings of the 23rd ACM International Conference on Multimedia*, ser. MM '15. New York, NY, USA: ACM, 2015, pp. 689–692. [Online]. Available: http://doi.acm.org/10.1145/2733373.2807412

[12] W. Zhou, H. Li, Y. Lu, and Q. Tian, "Principal visual word discovery for automatic license plate detection," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4269–4279, Sept 2012.

[13] Online, "Caltech plate dataset," 2003. [Online]. Available: http://www.vision.caltech.edu