A SINGLE IMAGE SUPER-RESOLUTION AIDED AUTOMATED LICENSE PLATE RECOGNITION SYSTEM

Xing Yao

Biomedical Engineering Department, Duke University, Durham, NC 27705 USA

ABSTRACT

Automated license plate recognition (ALPR) technology is more and more widely adopted in application scenarios such as traffic law enforcement, parking lot management, and crime tracking. However, ALPR from low resolution license plate images still remains challenging. In this study, we proposed an end-to-end single image super-resolution aided ALPR recognition system, which contains 5 stages: 1) super-resolution reconstruction based on efficient sub-pixel convolutional neural network (ESPCN), 2) license plate extraction, 3) tilt correction, 4) character split and 5) character recognition. The robustness and feasibility of the proposed system was evaluated on the custom designed datasets and achieved accuracy of 100%, 96% and 87% for the LP segmentation, character segmentation and single character verification tasks, respectively.

***Index Terms*—** Automated license plate recognition (ALPR), Single image super resolution (SISR), Convolutional neural network (CNN)

1. INTRODUCTION

ALPR is the core component of intelligent transportation systems which is now widely used in traffic management, vehicle tracking, parking lot management, and electronic toll collection. A typical ALPR system implements license plate detection, character segmentation and character verification through various image processing and computer vision technologies [1]. The system performance heavily rely on the quality of input image and faces challenges such as image blur, low contrast and low resolution caused by weather factors, illumination, imaging equipment and imaging distance in actual application scenarios.

Image super-resolution reconstruction (SR) technology can enhance image resolution and restore lost high-frequency information. Basically, SR can be roughly divided into two types based on the number of LR input: multi-frame SR (MFSR) and single image SR (SISR).

MFSR reconstructs a super-resolution single-frame image from multiple continuous images(ref) via modeling the acquisition process of LR observation images and using a regularization method to construct a priori constraints for HR images, and is widely used in video based LPR. However, the demand for multiple LR inputs restricts the application of MFSR in multiple scenes based on a single snapshot.

Compared with MISR, SISR aims to restore a HR output from one of its LR versions and is much more popular because of its high efficiency [2]. Traditional SISR mainly based on smoothing and interpolation techniques [3], such as bicubic [4] and cubic spline interpolation [5]. Recently, with the development of machine learning, many learning based SR algorithms are proposed. For example, SRCNN [6] inputs LR image pre-processed by bicubic interpolation into a three-layer convolutional network to adjust the nonlinear mapping to achieve HR reconstruction. Based on this, FSRCNN [7] achieves an end-to-end reconstruction by coupling a deconvolutional layer enlarge the feature map in the end of the network. VDSR [8] introduces the global residual error into the SR method to significantly improve training speed. ESPCN [9] uses sub-pixel convolutional layer to directly extract features from low-resolution image sizes, constructing an end-to-end fast super-resolution reconstruction method.

In this paper, we proposed a 5-stage SISR aided ALPR system (Fig. 1) which aim to achieve license plate recognition on the Chinese vehicle images with blue background license plates in low resolution condition. For given LR vehicle images, a color-feature based segmentation algorithm was adopted to extract license plate (LP) patterns and then the tilt of LP images was corrected via radon transform. Later, LR LP patterns were feed into a trained efficient sub-pixel neural network (ESPCN) to reconstruct HR LP images and were evaluated by mean squared error (MSE), peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). Next, multiple image processing methods were applied to remove irrelevant information from images and extract pictures of LP numbers. Then we used an improved projection-based segmentation algorithm to achieve characters segmentation. Finally, template matching technology was used to verify the LP number by calculating the normalized cross-correlation (NCC) degree between the LP characters and the template. The main contributions of this work are as following: 1) Designed an end-to-end pipline for LR condition ALPR which couple ESPCN as SR reconstruction methods; 2) Introduced mean filtering into the tilt correction algorithm based on radon transform to improve the algorithm performance 3) Improved

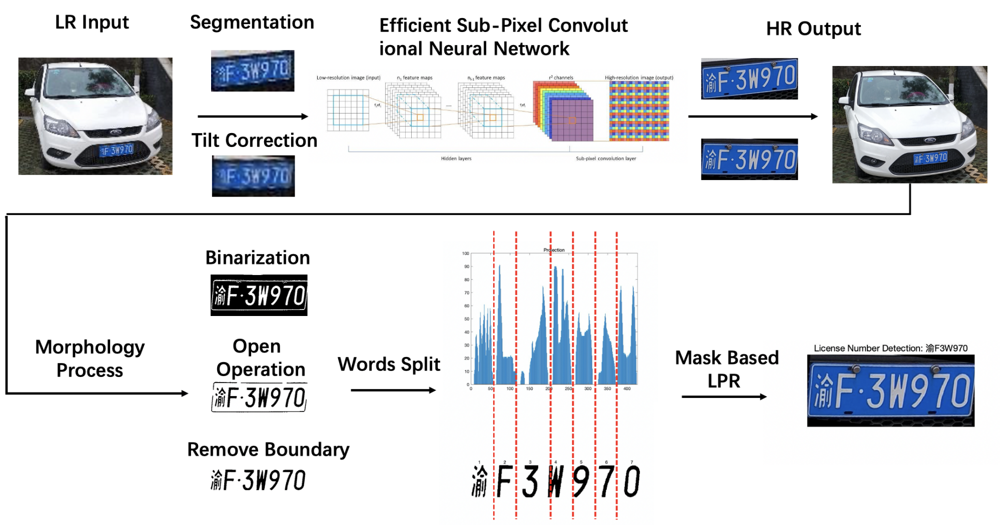

Fig. 1. The pipeline of ESPCN-LPR system

the robustness of character segmentation by a new projection algorithm based on adaptive threshold. The system achieved accuracy of 100%, 96% and 87% for the LP segmentation, character segmentation and single character verification tasks respectively on the custom-designed datasets.

2. PROPOSED METHOD

This section introduces our proposed end-to-end ALPR system, which includes the following 5 stages: 1. LP detection and tilt correction; 2. SR reconstruction; 3. morphology processing; 4. character split and 5. LP recognition.

2.1. LP Detection and tilt correction

Due to the limitation of calculation speed and memory size, many LP segmentation methods are based on gray-scale or binary image processing and ignore the color features [10]. In this work, we mainly performed detection for Chinese vehicle license plates with blue background and white numbers. Based on the strong prior knowledge, we adopted a color feature-based segmentation algorithm which can directly filter the pixels with a certain target color by the HSV properties and determine boundaries of the blue area corresponding to the license plates. Since without the need for more complex color distance calculations, the computational efficiency is improved. It is particularly effective and robust for blue and yellow license plates and achieved 100% accuracy on the custom data set. But for the pure colors, since their hue and saturation are meaningless and two filtering indicators are missing, the filtering performance is not very satisfactory.

Tilt correction (Fig.2(c)) was achieved via using radon transform to determine the straight line L (Fig.2(a)) in the LP image and determine the corresponding rotation angle θ by finding the maximum value on the sinogram (Fig.2(b)). In order to eliminate the interference of bright spots located at other tilt angles in S, we introduced the mean filtering:

$$S' = S * H \tag{1}$$

where H is the mean filtering kernel, $*$ is the convolution operation and S' is the processed sinogram.

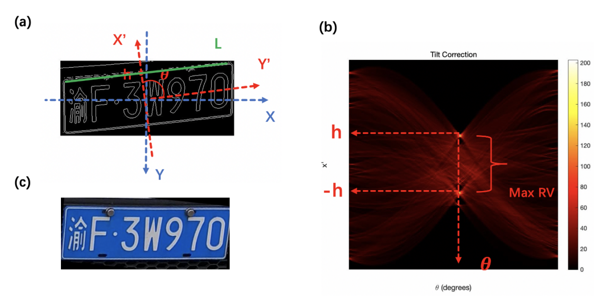

Fig. 2. Tilt correction via radon transform

2.2. Single image super-resolution

During the SISR process, an ESPCN was deployed to reconstruct the detected LR LP image into HR output. Our ESPCN contains 4 convolutional layers in which $(k1, n1) = (5, 64), (k2, n2) = (3, 64), (k3, n3) = (3, 32)$ and $(k4, n4) = (3, C \times r \times r)$. Where r = 3 is the upscaling factor, n_l and k_l are the feature maps and the filter size at layer l, respectively. $n_0 = C$. In the 4th layer l_4, periodical shuffle is used to convert $n_4 = C \times r \times r$ feature maps with size $H \times W \times C$ into a $Hr \times Wr \times C$ HR image. Compared to [9], we added one more layer and use the relu instead of tanh as the activation function, which achieved better performance even though training the model for fewer epochs.

2.3. Morphology Processing

The rivets, license plate boundaries, scratches and stains on LP images always affect the subsequent character extraction process. Therefore, before extracting characters, multiple image processing methods are needed to remove irrelevant information from the picture.

After binarizing the image (Fig.3(a)), Erosion was applied to eliminate boundary points and shrinking the boundary inward:

$$A \ominus B = \{X, Y | (B)_{x,y} \subseteq A\} \tag{2}$$

Where A is a binary image, B is the erosion kernel and (x, y) is the position of the origin of B in A.

Next, expansion was used to expand the boundary outward and to fill holes in objects:

$$A \oplus B = \{X, Y | (B)_{x,y} \cup A \neq \phi\} \tag{3}$$

The process of first erosion and then expansion is called open operation. As shown in Fig.3(b), open operation eliminated the small objects and shadow in binary images, and basically maintained the original shape and area of the target while smoothing the boundaries. After that, a horizontal histogram was applied to find the relevant area containing the characters and to further remove the upper and lower boundaries (Fig.3(c)).

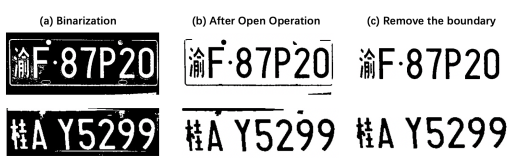

Fig. 3. The Morphology Processing: (a) binarization, (b) open operation, (c) removing boundaries via horizontal projection

2.4. Characters Segmentation

Many characters segmentation algorithms have been proposed in the past few years, such as [11], [12], [13]. In this study, we proposed an improved algorithm based on the projection-based segmentation, which typically uses vertical histogram to find the valleys separating characters and segment at these valleys [14]. We added adaptive vertical and horizontal threshold. The vertical threshold can dynamically adjust the depth of the valley, thereby improving the algorithm's resistance to interference information in the horizontal direction. Adaptive horizontal threshold can exclude small peaks caused by irrelevant patterns.

2.5. Characters Verification

Finally, we used template matching technology to recognize the segmented characters and convert them into documents. Normalized cross correlation (NCC) was adopted to represent the matching degree between characters and templates. Given a $M_x \times M_y$ image f and a $N_x \times N_y$ template t, the NCC γ is defined as:

$$\gamma = \frac{\sum_{x,y}(f(x,y) - \hat{f})(t(x,y) - \hat{t})}{\sqrt{\sum_{x,y}(f(x,y) - \hat{f})^2 \sum_{x,y}(t(x,y) - \hat{t})^2}} \tag{4}$$

Where $f(x, y)$ and $t(x, y)$ are pixel values of f and t in position (x, y), $\hat{f}$ and $\hat{t}$ are the global mean values of f and t, respectively.

3. EXPERIMENTAL EVALUATION

This evaluation section includes the following parts: 1) introduction of datasets, 2) - 4) the quantitative evaluation of LP detection, SISR reconstruction, characters segmentation and verification, respectively.

3.1. Datasets

In this work, 3 custom designed datasets are used for training the network, testing the LP detection performance and evaluating the character segmentation and verification algorithms.

BSDS500: Berkeley Segmentation Data Set and Benchmarks 500 as ground truth for ESPCN training.

CLPD-2: 24 HR Chinese vehicle images and 24 corresponding LR images form a test dataset for LP detection.

CLPD-3: 50 HR Chinese VLP images and 50 corresponding LR images form a test dataset for character segmentation and verification algorithms.

3.2. LP Detection

The performance of LP detection was evaluated on ***CLPD-2*** and achieve 100% accuracy. Fig.4 shows some LP detection results in LR condition.

Fig. 4. LP Detection results in LR condition

Table 1. Quantitative Evaluation for Fig.5

Mean deviation	PSNR	MSE	SSIM
$(ESPCN - Bicubic)_{aver}$	9.05	- 676.12	0.59

3.3. SISR Reconstruction

The ESPCN was trained on ***BSDS500*** by 100 epochs and the SISR reconstructed performance were evaluated by MSE, PSNR and SSIM. Fig.5 demonstrates part of the reconstruction results and Table.1 shows the mearn deviation of corresponding three indexes of quantitative assessment between ESPCN and bicubic interpolation. It can be seen that reconstruction results of ESPCN is significantly better than the results of bicubic interpolation as baseline.

Fig. 5. Part of reconstruction results. From the first to third column : HR input, Bicubic interpolation (Degrade) and SISR results (ESPCN)

3.4. Characters segmentation and verification

The performance of the whole system is evaluated on ***CLPD-3*** from three perspectives: the character segmentation accuracy (CS.ACC), the single character recognition accuracy (Words.ACC), and the overall recognition accuracy (Whole.ACC) (Table.2). Compared with LR results, the SISR aided system has been greatly improved. High CS.ACC (0.95) indicates the proposed system had good performance on image processing and character segmentation. Also the template matching algorithm has a high recognition rate for Arabic numerals and English letters (Words.ACC = 0.87), but a higher error rate for Chinese character recognition which mainly caused by the misrecognition of similar characters (Fig.6). The low-resolution system loses a lot of target information in the image processing stage due to the blurring of the input itself, noise and the adhesion between the characters, which affects the subsequent process.

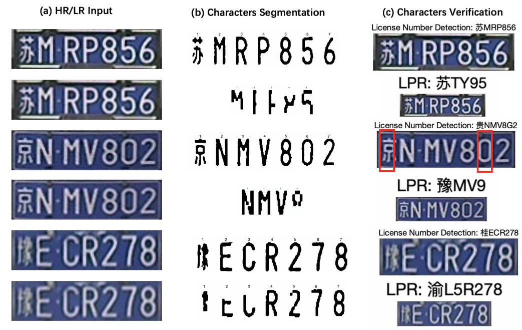

Fig. 6. Part of characters segmentation and verification results. From (a) - (c) : HR/LR inputs, Segmentation and Verification results

Table 2. Quantitative evaluation for the system

Resolution	CS.ACC	Words.ACC	Whole.ACC
LR	0.54	0.15	0.00
HR (proposed)	0.95	0.87	0.55

4. CONCLUSION

This paper proposes an end-to-end license plate detection system including license plate detection, high-resolution reconstruction, character segmentation, and character confirmation to recognize Chinese motor vehicle license plates under low-resolution conditions. By combining SISR and improved image processing algorithms, the system stable and well achieved the expected goal via realizing 100%, 96% and 85% accuracy in license plate recognition, character segmentation and single character verification on the custom data set.On this basis, the system has room for further improvement. Due to the special principle of color feature-based segmentation algorithm, it is determined that the recognition is mainly for the white plate on the blue background of the family small car. In the follow-up work, the applicability of the algorithm can be improved by combining the boundary and shape features. Also, different network structures, such as GAN [15] can can be used to further improve high-resolution reconstruction, but this requires a large number of high-quality data sets as support. For LP verification, the single-template matching algorithm is limited by the number of templates, which leads to low robustness, especially for characters with more complex patterns. The multi-template matching algorithm has low computational efficiency and heavily depends on the templates' quality. In contrast, the CNN based method have the characteristics of high efficiency and stability, which is also the direction of the next work.

5. REFERENCES

[1] Yule Yuan, Wenbin Zou, Yong Zhao, Xinan Wang, Xuefeng Hu, and Nikos Komodakis, "A robust and efficient approach to license plate detection," *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1102–1114, 2016.

[2] W. Yang, X. Zhang, Y. Tian, W. Wang, J. Xue, and Q. Liao, "Deep learning for single image super-resolution: A brief review," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 3106–3121, 2019.

[3] K. Su, Q. Tian, Q. Xue, N. Sebe, and J. Ma, "Neighborhood issue in single-frame image super-resolution," in *2005 IEEE International Conference on Multimedia and Expo*, 2005, pp. 4 pp.–.

[4] Hsieh Hou and H Andrews, "Cubic splines for image interpolation and digital filtering," *IEEE Transactions on acoustics, speech, and signal processing*, vol. 26, no. 6, pp. 508–517, 1978.

[5] Robert Keys, "Cubic convolution interpolation for digital image processing," *IEEE transactions on acoustics, speech, and signal processing*, vol. 29, no. 6, pp. 1153–1160, 1981.

[6] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang, "Learning a deep convolutional network for image super-resolution," in *European conference on computer vision*. Springer, 2014, pp. 184–199.

[7] Chao Dong, Chen Change Loy, and Xiaoou Tang, "Accelerating the super-resolution convolutional neural network," in *European conference on computer vision*. Springer, 2016, pp. 391–407.

[8] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1646–1654.

[9] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1874–1883.

[10] A. H. Ashtari, M. J. Nordin, and M. Fathy, "An iranian license plate recognition system based on color features," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 4, pp. 1690–1705, 2014.

[11] Vojtěch Franc and Václav Hlaváč, "License plate character segmentation using hidden markov chains," in *Joint Pattern Recognition Symposium*. Springer, 2005, pp. 385–392.

[12] Xin Fan, Guoliang Fan, and Dequn Liang, "Joint segmentation and recognition of license plate characters," in *2007 IEEE International Conference on Image Processing*. IEEE, 2007, vol. 4, pp. IV–353.

[13] Xin Fan and Guoliang Fan, "Graphical models for joint segmentation and recognition of license plate characters," *IEEE Signal Processing Letters*, vol. 16, no. 1, pp. 10–13, 2008.

[14] J. Jagannathan, A. Sherajdheen, R. M. Vijay Deepak, and N. Krishnan, "License plate character segmentation using horizontal and vertical projection with dynamic thresholding," in *2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN)*, 2013, pp. 700–705.

[15] Jin Zhu, Guang Yang, and Pietro Lio, "How can we make gan perform better in single medical image super-resolution? a lesion focused multi-scale approach," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 1669–1673.