EQ-LPR: EFFICIENT QUALITY-AWARE LICENSE PLATE RECOGNITION

Cong Zhang, Qi Wang, Xuelong Li*

School of Computer Science and Center for OPTical IMagery Analysis and Learning (OPTIMAL),
Northwestern Polytechnical University, Xi'an 710072, Shaanxi, P.R. China.

ABSTRACT

License plate recognition (LPR) has attracted considerable attention due to its widespread applications in real life. Although numerous approaches based on image processing have been presented in the past few years, it is still an urgent issue to perform the LPR task efficiently in complex and unconstrained scenarios. To remedy this problem, an efficient quality-aware license plate recognition algorithm is proposed by introducing the siamese networks for plate stream recognition and quality awareness in the traffic videos. Moreover, we explore three progressive architectures for efficient and accurate recognition. Knowledge distillation is adopted to compress the quality awareness network and make it lightweight. Extensive experiments have demonstrated the impressive performance and efficiency of the proposed method.

Index Terms— License Plate Recognition, Siamese Networks, Quality Awareness, Knowledge Distillation

1. INTRODUCTION

Recently, license plate recognition (LPR) has received considerable research attention in the fields of image processing and computer vision. LPR system plays a vital role in the maintenance of intelligent transportation systems (ITS), which has been widely utilized in numerous real-world applications such as traffic law enforcement, access control, and automatic highway toll collections [1, 2, 3]. In general, license plates are specially designed for the vehicle identification in traffic management systems, each of which involves several diverse numbers and characters. An available LPR system can accurately recognize each character of the localized license plates with various fonts, colors, and backgrounds in different administrative regions. Therefore, considering these properties, it is reasonable to implement LPR systems based on image processing algorithms.

In the past few decades, considerable academic researches on LPR have been published in the community, which can be roughly divided into traditional methods [4, 5] and deep learning-based methods [6, 7]. Thanks to the development of deep learning in object detection, the performance of license plate detection has been dramatically improved yet robust and efficient plate recognition is still an urgent task to be solved [8, 9, 10]. Regarding plate recognition, there are two different approaches, *i.e.*, segmentation-based methods [7] and segmentation-free methods [6], where the latter generally consider plates as text sequences without character segmentation. Then they exploit convolutional neural networks (CNNs) followed by recurrent neural networks (RNNs) to build the encoder-decoder model for sequence recognition. In practice, license plate segmentation is really a challenging task that is extremely sensitive to character distortion, occlusion, and illumination variation [6]. Consequently, segmentation-free plate recognition algorithms have achieved better performance, attracting tremendous research interests in recent years [6, 11, 12].

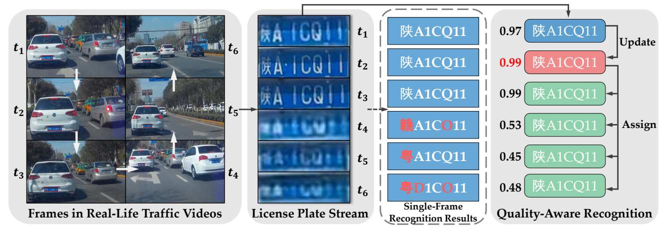


Fig. 1. There are six frames involved with the same license plate but at different times in a traffic video. The red characters of the single-frame recognition results indicate incorrect recognition that can be avoided by the proposed EQ-LPR.

These previous works contribute to the development of LPR systems, some of which have been applied in several explicit conditions. However, there are the following significant application limitations in these methods. (1) Most existing LPR algorithms perform on individual images captured by the sophisticated camera equipment in strictly specific environments like parking lot entrances. Nevertheless, motion blur, defocus, and perspective distortion usually occur in real-life unconstrained traffic scenarios, *e.g.*, intelligent driving, where many algorithms may not work well. (2) Despite considerable academic study on LPR systems, there are only a few video-based approaches to model temporal information explicitly [13]. Thanks to the development of video capture devices such as dashboard cameras, it is convenient and low-cost to acquire real-life traffic videos that involve more valuable information for LPR than single images. (3) Previous

*Corresponding author.

This work was supported by the National Natural Science Foundation of China under Grant U1864204, 61773316, U1801262, and 61871470.

existing video-based LPR techniques mainly focus on license plate super-resolution [14] and majority voting-based recognition [10], which introduce additional computational complexity and reduce efficiency. However, it is reasonable to evaluate each frame quality in license plate videos. As the prior knowledge, plate quality implies the degree of distortion that can be employed to guide recognition. As illustrated in Fig. 1, for a joint frame stream composed of the same plate but at different times, clearer plate frames enable higher recognition accuracy. Thus introducing plate quality estimation and temporal contextual information from videos may improve the robustness and effectiveness of the LPR system especially in complex unconstrained scenarios.

Against the above issues, an online **e**fficient **q**uality-aware **l**icense **p**late **r**ecognition (EQ-LPR) algorithm is proposed in this paper. With a plate stream based on the tracklet, it first automatically evaluates the image quality of each plate frame and then recommends the recognition result of the current highest quality frame as the final decision. It is noteworthy that EQ-LPR works online without video future information that can be simply embedded as an efficient and robust recognition module in the real-life video-based systems.

The main contributions of this paper are as follows: (1) The siamese architecture is developed for quality-aware plate recognition, one subnetwork towards online plate stream quality awareness while the other for sequence recognition based on the selected frames by the former. (2) The quality awareness network is further compressed for algorithm efficiency via knowledge distillation. (3) Extensive experiments have presented the competitive performance of the proposed method and also demonstrate its effectiveness and efficiency.

2. METHODOLOGY

In this section, we first give the overview of the proposed algorithm and then introduce its implementation details.

2.1. Overview

Generally speaking, an LPR system available in real-life complex traffic scenarios can be divided into three steps or components: license plate detection, tracking, and recognition [13]. As mentioned above, there are several existing detection and tracking approaches for video-based LPR systems, which have shown their impressive performance in unconstrained and harsh environments [7, 10, 15, 16, 17]. However, previous video-based methods usually ignore the impact of plate quality on recognition, some of which introduce excessive temporal redundancy in neighboring frames, leading to efficiency loss. To remedy these problems, we focus on efficient plate stream recognition where the potential plate regions have been favorably localized and linked with their assigned identities by the well-trained detector and tracker, respectively. The lightweight quality awareness network is developed for efficient stream quality estimation and plate frame recommendation. Moreover, motivated by [6, 12], we perform segmentation-free plate recognition controlled by the quality awareness network, which avoids segmentation errors caused by distortion and blur in low-quality frames. For a clear illustration, denote the cropped plate regions as $\boldsymbol{\mathcal{P}}$ while I and $\delta(i)$ represent the maximum identity and the total frame number of the plate stream with the current identity i, respectively. Then the process of EQ-LPR is summarized in Algorithm 1. To implement EQ-LPR, we explore three different but progressive frameworks depicted in Fig. 2, whose design patterns and training strategies will be described in detail.

Algorithm 1 Efficient Quality-aware LPR (EQ-LPR)

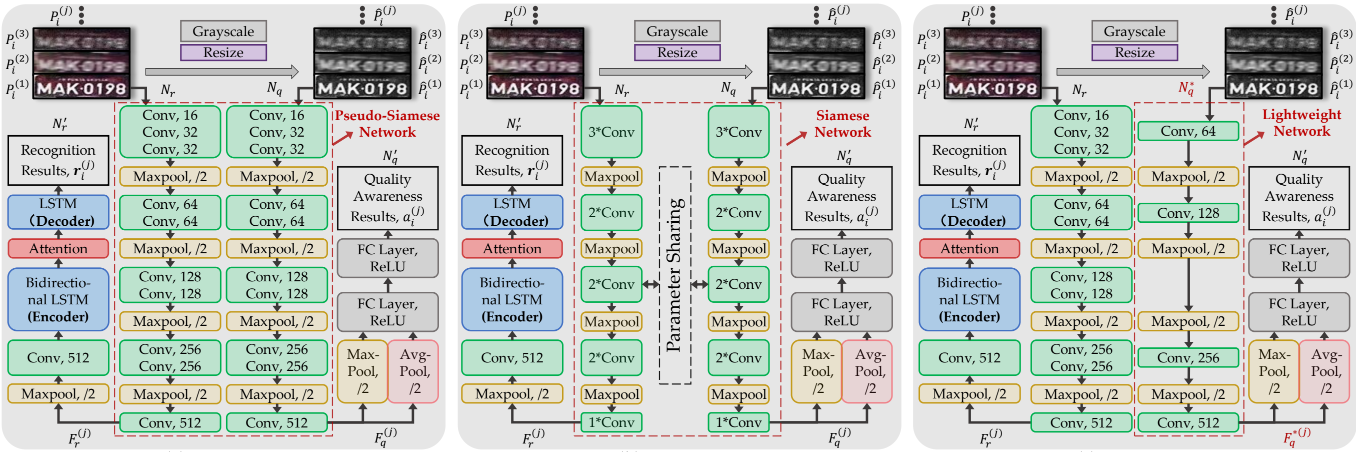
Input: The detected, linked and cropped license plate regions with their respective identities from the videos, formulated by $\boldsymbol{\mathcal{P}} = \{\mathcal{P}_i^{(j)} : i \in [1, I], j \in [1, \delta(i)]\}$

Obtain the i-th plate stream $\boldsymbol{\mathcal{P}}_i = \{\mathcal{P}_i^{(j)} : j \in [1, \delta(i)]\}$
Set $\tilde{a}_i$ as the highest quality awareness score for $\boldsymbol{\mathcal{P}}_i$
Set $\tilde{\boldsymbol{r}}_i^{(j)}$ as the final recognition result for each $\mathcal{P}_i^{(j)}$
for $i = 1$ **to** I **do**
    for $j = 1$ **to** $\delta(i)$ **do**
        if $j = 1$ **then**
            Produce the quality awareness score $a_i^{(1)}$
            Produce the plate recognition results $\boldsymbol{r}_i^{(1)}$
            Initialize $\tilde{a}_i \leftarrow a_i^{(1)}$, $\tilde{\boldsymbol{r}}_i^{(j)} \leftarrow \boldsymbol{r}_i^{(1)}$
        else
            Produce the quality awareness score $a_i^{(j)}$
            if $a_i^{(j)} > \tilde{a}_i$ **then**
                Produce the plate recognition results $\boldsymbol{r}_i^{(j)}$
                Update $\tilde{a}_i \leftarrow a_i^{(j)}$, $\tilde{\boldsymbol{r}}_i^{(j)} \leftarrow \boldsymbol{r}_i^{(j)}$
            else
                Assign $\tilde{\boldsymbol{r}}_i^{(j)} \leftarrow \tilde{\boldsymbol{r}}_i^{(j-1)}$
            end if
        end if
    end for
end for
Get $\tilde{\boldsymbol{a}} = (\tilde{a}_1, \tilde{a}_2, \cdots, \tilde{a}_I)^T$, $\boldsymbol{\mathcal{R}} = \{\tilde{\boldsymbol{r}}_i^{(j)}\}$

Output: The final plate recognition results $\tilde{\boldsymbol{r}}_i^{(j)}$ for each $\mathcal{P}_i^{(j)}$

2.2. Siamese Networks for Quality-Aware LPR

As shown in Fig. 2 (a), the pseudo-siamese neural network is first developed for quality-aware plate stream recognition, named EQ-LPR-P. To be specific, there are two separate networks without parameter sharing, one perceiving the grayscale image quality of each frame in the stream and the other recognizing the RGB plate image with the best quality. These two networks have the same structure CNN models to extract deep discriminative features which will be fed into their target-specific subnetworks for quality awareness and sequence recognition, respectively. However, the pseudo-siamese architecture EQ-LPR-P requires more com-



(a) EQ-LPR-P (b) EQ-LPR-S (c) EQ-LPR-E

Fig. 2. Overview of three different but progressive architectures for EQ-LPR, which involve the pseudo-siamese network, siamese network and compressed lightweight network respectively. All convolutional layers are configured with 3×3 kernels.

puting resources without the shared CNN feature extractor. Therefore, we also introduce the common siamese network to perform the proposed EQ-LPR, as depicted in Fig. 2 (b). Compared to EQ-LPR-P, EQ-LPR-S shares all parameters for the two tasks during the feature extraction phase. More importantly, both of EQ-LPR-P and EQ-LPR-S are developed in this work to explore two fundamental issues: (1) The plate recognition performance improvement by sharing discriminative features with plate quality estimation. (2) The model compression effectiveness of EQ-LPR-S for higher running speed and lower computational complexity. Formally, given the plate stream $\mathcal{P}_i^{(j)} \in \boldsymbol{\mathcal{P}}_i$ with the same identity i, the deep feature maps are generated by

$$\widehat{\mathcal{P}}_i^{(j)} = \mathcal{G}(\mathcal{P}_i^{(j)}),\ \mathcal{F}_q^{(j)} = \mathcal{N}_q(\widehat{\mathcal{P}}_i^{(j)}),\ \mathcal{F}_r^{(j)} = \mathcal{N}_r(\mathcal{P}_i^{(j)}), \quad (1)$$

where $\widehat{\mathcal{P}}_i^{(j)}$ represents the grayscale image of $\mathcal{P}_i^{(j)}$ that has been resized to the fixed size to avoid color interference.

As illustrated in Fig. 2, with the discriminative features encoded by the siamese network, $\mathcal{F}_q^{(j)}$ will be processed by the fully connected network $\mathcal{N}_q'$ while $\mathcal{F}_r^{(j)}$ is also further encoded as the feature sequences by RNNs. In practice, the LSTMs are exploited as the encoder to alleviate gradient vanishing or exploring. After that, the visual attention mechanism is introduced during the decoding phase which generates recognition results. This process can be formulated as

$$a_i^{(j)} = \mathcal{N}_q'(\mathcal{F}_q^{(j)}),\ \boldsymbol{r}_i^{(j)} = \mathcal{N}_r'(\mathcal{F}_r^{(j)}), \quad (2)$$

where $a_i^{(j)}$ and $\boldsymbol{r}_i^{(j)}$ denote the produced quality awareness score and recognition results, respectively. Both EQ-LPR-P and EQ-LPR-S can be automatically trained via the alternate and adaptive learning strategy. Concretely, the minimum softmax prediction is regarded as the ground truth of $a_i^{(j)}$, *i.e.*,

$$\alpha_i^{(j)} = \min_{k\in[1,\mathcal{K}]} \{softmax(y_{i,k}^{(j)}) | y_{i,k}^{(j)} \in \{y_{i,1}^{(j)}, \cdots, y_{i,\mathcal{K}}^{(j)}\}\}, \quad (3)$$

where $\mathcal{K}$ means the plate sequence length and $y_{i,k}^{(j)}$ is the softmax input for the k-th character prediction. With the recognition labels $\boldsymbol{\theta}_i^{(j)}$, the loss functions can be formulated as

$$\mathcal{L}_r = -\sum_{k=1}^{\mathcal{K}} \ln P(\boldsymbol{\theta}_{i,k}^{(j)} | \boldsymbol{\theta}_{i,1:k-1}^{(j)}),\ \mathcal{L}_q = |\alpha_i^{(j)} - a_i^{(j)}|. \quad (4)$$

It should be noted that the optimizations for $\mathcal{L}_r$ and $\mathcal{L}_q$ are performed separately and alternately rather than joint training since $\alpha_i^{(j)}$ is generated based on the recognition results $\boldsymbol{r}_i^{(j)}$.

2.3. Efficient Lightweight Quality Awareness Network

Both EQ-LPR-P and EQ-LPR-S can achieve quality-aware plate recognition based on the siamese architectures. However, considering efficiency and real-life applications of LPR, we innovatively explore the model compression for the quality awareness network, as shown in Fig.2 (c). Intuitively, the lightweight $\widetilde{\mathcal{N}_q}$ can efficiently perform online plate quality awareness in a short period of time. Furthermore, observing that the inference of $\widetilde{\mathcal{N}_q}$ is more frequent than the recognition network, it is reasonable to compress the quality awareness module. Thus knowledge distillation [18] is exploited to compress $\mathcal{N}_q$ into the lightweight $\mathcal{N}_q^*$ and $\mathcal{F}_q^{*(j)} = \mathcal{N}_q^*(\widehat{\mathcal{P}}_i^{(j)})$. Specifically, based on EQ-LPR-S, we modify the loss function $\mathcal{L}_q$ to $\mathcal{L}_q^*$ by introducing additional constraints, given by

$$\mathcal{L}_q^* = |\alpha_i^{(j)} - a_i^{(j)}| + \|\mathcal{F}_q^{*(j)} - \mathcal{F}_r^{(j)}\|_2, \quad (5)$$

where ℓ_2 norm is employed to minimize the element-wise distance between $\mathcal{F}_q^{*(j)}$ and $\mathcal{F}_r^{(j)}$. In the formulation, $\mathcal{F}_q^{*(j)}$ aims to fit the feature representation in $\mathcal{F}_r^{(j)}$ via supervised learning and the training strategy follows the description above.

3. EXPERIMENTS

Extensive experiments are conducted in this section to evaluate the performance of the proposed algorithm. Moreover, we present and analyze its effectiveness and efficiency by comparison with other methods based on the quantitative results.

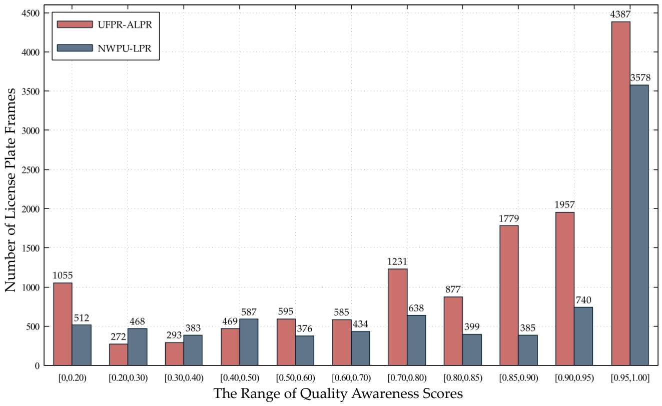


Fig. 3. The distribution of quality awareness scores in two different datasets, which are normalized to the range of $[0, 1]$.

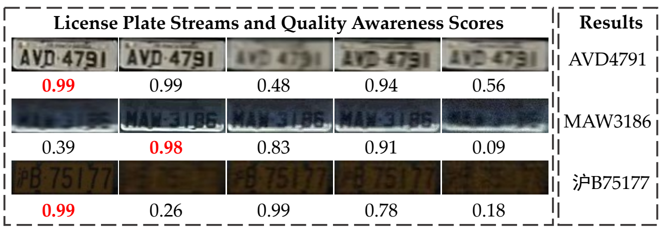


Fig. 4. Exemplar recognition results of the proposed algorithm EQ-LPR-E on different datasets and the red values indicate the final awareness scores of the plate streams.

3.1. Datasets and Implementation Details

To comprehensively evaluate the proposed LPR algorithm, two different datasets are exploited in the experiments. The first one is UFPR-ALPR dataset released in [10], which consists of $4,500$ license plate images captured inside the driving vehicles. All images in this dataset are obtained from 150 real-life traffic videos, corresponding to 150 license plates and each plate involving 30 frames at different times. Moreover, they are acquired with three different cameras, which ensures the diversity of plate images. In practice, for the fair competition, we follow the protocol division proposed in [10], *i.e.*, 60% for training and 40% for testing. Note that data augmentation is adopted in both the training and testing phases to avoid overfitting and expand the dataset. The second dataset, namely NWPU-LPR, is composed of $8,500$ images after data augmentation, corresponding to $1,700$ different plates and each plate involving 5 frames. With various distortions like noise, these images are acquired in harsh environments in different provinces of China. In the evaluation, $6,000$ images are utilized for training and others for testing.

3.2. Performance Evaluation and Effectiveness Analysis

Extensive experiments are conducted to evaluate the performance of the proposed three different architectures. Note that only accurately predicting the entire plate sequence means the correct recognition. As shown in Table 1, the proposed architectures outperform these existing methods on the UFPR-ALPR dataset. Furthermore, the third architecture EQ-LPR-E takes only 43 ms per frame ($23.26\,fps$) yet its recognition accuracy is 11.81% higher than the method in [10] with only additional 15 ms. It makes sense even for some real-time applications. However, since most images in the UFPR-ALPR dataset are captured in the driving scenarios with slow speeds, we specifically build the NWPU-LPR dataset to better demonstrate the effectiveness of the proposed efficient quality awareness network. In this dataset, image distortions such as motion blur and illumination variance usually occur, making it more challenging than the first one. As illustrated in Fig. 3, poor-quality images occupy a higher proportion in the dataset, leading to difficult recognition tasks. For the NWPU-LPR dataset, we take two public APIs toward Chinese LPR as comparison methods [21, 22] and our algorithms also achieve better recognition accuracy in Table 2. Concretely, all architectures of EQ-LPR outperform the comparison methods in recognition accuracy, which demonstrates the effectiveness of the quality awareness network for video-based LPR systems. Moreover, the lightweight EQ-LPR-E architecture shows its efficiency and speed benefits with the acceptable reduction in recognition performance, which is of constructive value to the real-life applications. The exemplar recognition results on the two different datasets are exhibited in Figure. 4, the first two rows and the last one corresponding to the UFPR-ALPR dataset and NWPU-LPR dataset, respectively.

Table 1. Recognition performance and speed comparison of different methods on UFPR-ALPR dataset.

Methods	Recognition Accuracy (%)	Time (ms)
Sighthound [19]	47.39	–
OpenALPR [20]	50.94	–
Laroca *et. al* [10]	78.33	**28**
EQ-LPR-P (Ours)	88.58	96
EQ-LPR-S (Ours)	**90.31**	66
EQ-LPR-E (Ours)	**90.14**	**43**

Table 2. Recognition performance and speed comparison of different methods on NWPU-LPR dataset.

Methods	Recognition Accuracy (%)		Overall Time Per Frame (ms)
	Excluding Chinese Characters	Including Chinese Characters	
EasyPR [21]	77.14	70.60	–
BaiduLPR [22]	92.33	87.25	–
EQ-LPR-P (Ours)	93.09	91.21	106
EQ-LPR-S (Ours)	**96.07**	**93.86**	**81**
EQ-LPR-E (Ours)	**95.27**	**92.91**	**49**

4. CONCLUSIONS

In this work, an efficient quality-aware license plate recognition algorithm is proposed for online and fast LPR systems especially in complex scenarios, which integrates quality awareness and plate sequence recognition into one network. Moreover, it involves three different but progressive architectures based on the siamese networks or knowledge distillation. Competitive experimental results on two datasets demonstrate the efficiency and effectiveness of EQ-LPR.

5. REFERENCES

[1] Francisco Delmar Kurpiel, Rodrigo Minetto, and Bogdan Tomoyuki Nassu, "Convolutional neural networks for license plate detection in images," in *IEEE International Conference on Image Processing*, 2017, pp. 3395–3399.

[2] Shan Du, Mahmoud Ibrahim, Mohamed Shehata, and Wael 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, 2013.

[3] Qi Wang, Junyu Gao, and Yuan Yuan, "Embedding structured contour and location prior in siamesed fully convolutional networks for road detection," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 1, pp. 230–241, 2018.

[4] Jinn-Li Tan, Syed AR Abu-Bakar, and Musa M Mokji, "License plate localization based on edge-geometrical features using morphological approach," in *IEEE International Conference on Image Processing*, 2013, pp. 4549–4553.

[5] Yule Yuan, Wenbin Zou, Yong Zhao, Xin'an 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, 2017.

[6] Hui Li, Peng Wang, and Chunhua Shen, "Toward end-to-end car license plate detection and recognition with deep neural networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 3, pp. 1126–1136, 2019.

[7] Sérgio Montazzolli Silva and Cláudio Rosito Jung, "License plate detection and recognition in unconstrained scenarios," in *European Conference on Computer Vision*, 2018, pp. 593–609.

[8] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, , no. 6, pp. 1137–1149, 2017.

[9] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi, "You only look once: Unified, real-time object detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.

[10] Rayson Laroca, Evair Severo, Luiz A Zanlorensi, Luiz S Oliveira, Gabriel Resende Gonçalves, William Robson Schwartz, and David Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector," in *International Joint Conference on Neural Networks*, 2018, pp. 1–10.

[11] Jianlin Wang, He Huang, Xusheng Qian, Jinde Cao, and Yakang Dai, "Sequence recognition of chinese license plates," *Neurocomputing*, vol. 317, pp. 149–158, 2018.

[12] Hui Li, Peng Wang, Mingyu You, and Chunhua Shen, "Reading car license plates using deep neural networks," *Image and Vision Computing*, vol. 72, pp. 14–23, 2018.

[13] Nicolas Thome, Antoine Vacavant, Lionel Robinault, and Serge Miguet, "A cognitive and video-based approach for multinational license plate recognition," *Machine Vision and Applications*, vol. 22, no. 2, pp. 389–407, 2011.

[14] Wu Liu, Xinchen Liu, Huadomg Ma, and Peng Cheng, "Beyond human-level license plate super-resolution with progressive vehicle search and domain priori GAN," in *ACM International Conference on Multimedia*, 2017, pp. 1618–1626.

[15] Hilário Seibel, Siome Goldenstein, and Anderson Rocha, "Eyes on the target: Super-resolution and license-plate recognition in low-quality surveillance videos," *IEEE Access*, vol. 5, pp. 20020–20035, 2017.

[16] Thidarat Pinthong, Worawut Yimyam, Narumol Chumuang, and Mahasak Ketcham, "License plate tracking based on template matching technique," in *IEEE International Symposium on Communications and Information Technologies*, 2018, pp. 299–303.

[17] Gaurav Kumar, Amal Barman, and Madhumita Pal, "License plate tracking using gradient based segmentation," in *IEEE Region 10 Conference*, 2019, pp. 1737–1740.

[18] Adriana Romero, Nicolas Ballas, Samira Ebrahimi Kahou, Antoine Chassang, Carlo Gatta, and Yoshua Bengio, "Fitnets: Hints for thin deep nets," in *International Conference on Learning Representations*, 2015.

[19] Syed Zain Masood, Guang Shu, Afshin Dehghan, and Enrique G Ortiz, "License plate detection and recognition using deeply learned convolutional neural networks," *arXiv preprint arXiv:1703.07330*, 2017.

[20] "OpenALPR API," `http://www.openalpr.com/cloud-api.html`.

[21] "EasyPR," `https://github.com/liuruoze/EasyPR`.

[22] "BaiduLPR API," `https://cloud.baidu.com/doc/OCR/OCR-API.html`.