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].

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...
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 efficient quality-aware license plate recognition (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 \( P \) 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.

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-

---

**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 \( P = \{P^{(j)}_i : i \in [1, I], j \in [1, \delta(i)]\} \)

1. Obtain the \( i \)-th plate stream \( P_i = \{P^{(j)}_i : j \in [1, \delta(i)]\} \)
2. Set \( \bar{a}_i \) as the highest quality awareness score for \( P_i \)
3. Set \( \bar{r}^{(f)}_i \) as the final recognition result for each \( P^{(j)}_i \)

foreach \( i = 1 \) to \( I \)

for \( j = 1 \) to \( \delta(i) \)

if \( j = 1 \) then

7: Produce the quality awareness score \( a^{(1)}_i \)
8: Produce the plate recognition results \( r^{(1)}_i \)
9: Initialize \( \bar{a}_i \leftarrow a^{(1)}_i \), \( \bar{r}^{(1)}_i \leftarrow r^{(1)}_i \)

else

11: Produce the quality awareness score \( a^{(j)}_i \)
12: if \( a^{(j)}_i > \bar{a}_i \) then
13: Produce the plate recognition results \( r^{(j)}_i \)
14: Update \( \bar{a}_i \leftarrow a^{(j)}_i \), \( \bar{r}^{(j)}_i \leftarrow r^{(j)}_i \)
15: else
16: Assign \( r^{(j)}_i \leftarrow r^{(j-1)}_i \)
17: end if
18: end if
19: end for
20: end for
21: Get \( \bar{a} = (\bar{a}_1, \bar{a}_2, \cdots, \bar{a}_I)^T \), \( \mathcal{R} = \{\bar{r}^{(j)}_i\} \)

**Output:** The final plate recognition results \( \bar{r}^{(j)}_i \) for each \( P^{(j)}_i \)
putting 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 \( P_i \), with the same identity \( i \), the deep feature maps are generated by
\[
\tilde{P}_i^{(j)} = G(P_i^{(j)}), \quad F_q^{(j)} = \mathcal{N}_q(\tilde{P}_i^{(j)}), \quad F_r^{(j)} = \mathcal{N}_r(P_i^{(j)}),
\]
where \( \tilde{P}_i^{(j)} \) represents the grayscale image of \( P_i^{(j)} \) 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, \( F_q^{(j)} \) will be processed by the fully connected network \( \mathcal{N}_q^r \) while \( F_r^{(j)} \) is also further encoded as the feature sequences by RNNs. In practice, the LSTMs are exploited as the encoder to alleviate gradient vanishing or exploding. After that, the visual attention mechanism is introduced during the decoding phase which generates recognition results. This process can be formulated as
\[
a_i^{(j)} = \mathcal{N}_q^r(F_q^{(j)}), \quad r_i^{(j)} = \mathcal{N}_r(F_r^{(j)}),
\]
where \( a_i^{(j)} \) and \( r_i^{(j)} \) 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.,
\[
a_i^{(j)} = \min_{k \in [1, K]} \{ \text{softmax}(y_{i,k}^{(j)}) | y_{i,k}^{(j)} \in \{ y_{i,1}^{(j)}, \ldots, y_{i,K}^{(j)} \} \},
\]
where \( K \) means the plate sequence length and \( y_{i,k}^{(j)} \) is the soft-max input for the \( k \)-th character prediction. With the recognition labels \( \theta_i^{(j)} \), the loss functions can be formulated as
\[
\mathcal{L}_r = - \sum_{k=1}^{K} \ln P(\theta_i^{(j)} | \theta_i^{(j)}_{1:k-1}), \quad \mathcal{L}_q = |a_i^{(j)} - a_i^{(j)}|.
\]
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 \( a_i^{(j)} \) is generated based on the recognition results \( r_i^{(j)} \).

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 \( \tilde{N}_q^r \) can efficiently perform online plate quality awareness in a short period of time. Furthermore, observing that the inference of \( \mathcal{N}_q^r \) is more frequent than the recognition module, it is reasonable to compress the quality awareness module. Thus knowledge distillation [18] is exploited to compress \( \mathcal{N}_q^r \) into the lightweight \( \mathcal{N}_q^r \) and \( F_q^r(j) = \mathcal{N}_q^r(\tilde{P}_i^{(j)}) \). Specifically, based on EQ-LPR-S, we modify the loss function \( \mathcal{L}_q \) to \( \mathcal{L}_q^* \) by introducing additional constraints, given by
\[
\mathcal{L}_q^* = |a_i^{(j)} - a_i^{(j)}| + \|F_q^{(j)} - F_r^{(j)}\|_2,
\]
where \( \ell_2 \) norm is employed to minimize the element-wise distance between \( F_q^{(j)} \) and \( F_r^{(j)} \). In the formulation, \( F_q^{(j)} \) aims to fit the feature representation in \( F_r^{(j)} \) 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.
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) while 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.

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


