2D License Plate Recognition based on Automatic Perspective Rectification

Hui Xu[1,2], Zhao-Hong Guo[1,2], Da-Han Wang[3], Xiang-Dong Zhou[1], Yu Shi[1]

[1]Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences

[2]Chongqing School, University of Chinese Academy of Sciences

[3]Xiamen University of Technology, China

Introduction

◆ License Plate Recognition(LPR)

- The plates in the images are likely to be distorted due to shooting angles, which directly affect the recognition of license number.
- License plates(LPs) include many kinds, varying in background color, length of numbers and character arrangement. The diversity of Chinese LPs is a challenge to LPR. The numbers consists of Latin letters, digit and Chinese characters.

Motivations

- **Existing LPR methods**

(a) Character-based

(b) RNN-based

Existing LPR methods can be roughly categorized into two groups: character-based and RNN-based.

- **Character-based** methods relies on character segmentation, which is unreliable to illumination, pose and noise in the image.

- **RNN-based** methods can only recognize plates with no or minimal deformations.

Motivations

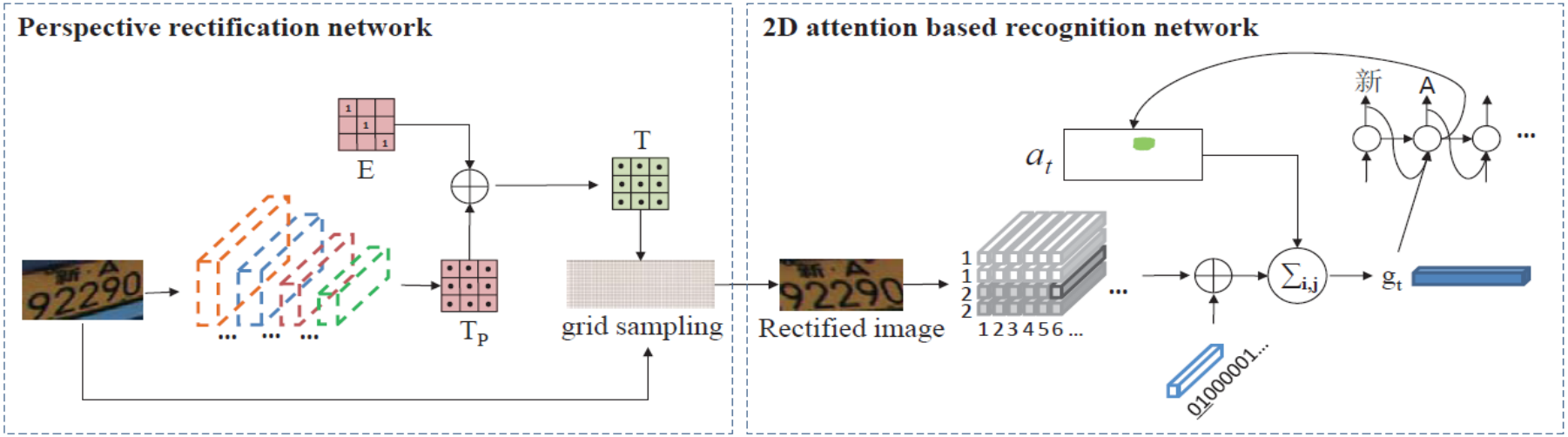
◆ **Existing Irregular text recognition methods**

- A text rectification module is always designed for irregular text recognition. However, the universal text rectification methods are designed for various irregular text, which can not achieve good effectiveness in LPR.

- LPs can **only produce perspective deformation** in the image due to its rigid body property.

✓ we propose a novel method consisting of a novel perspective rectification network (PRN) and a location-aware 2D attention based recognition network for the LPR task.

Proposed Method



- **Perspective rectification network (PRN)** estimates the perspective transformation of an input LP image automatically and generates the rectified LP image.
- **2D attention based recognition network** identify both single-line and double line LPs with 2D attention mechanism.

Proposed Method

◆ Perspective rectification network (PRN)

PRN includes two steps: perspective transformation prediction and grid sampling.

TABLE I
ARCHITECTURE OF PERSPECTIVE RECTIFICATION NETWORK

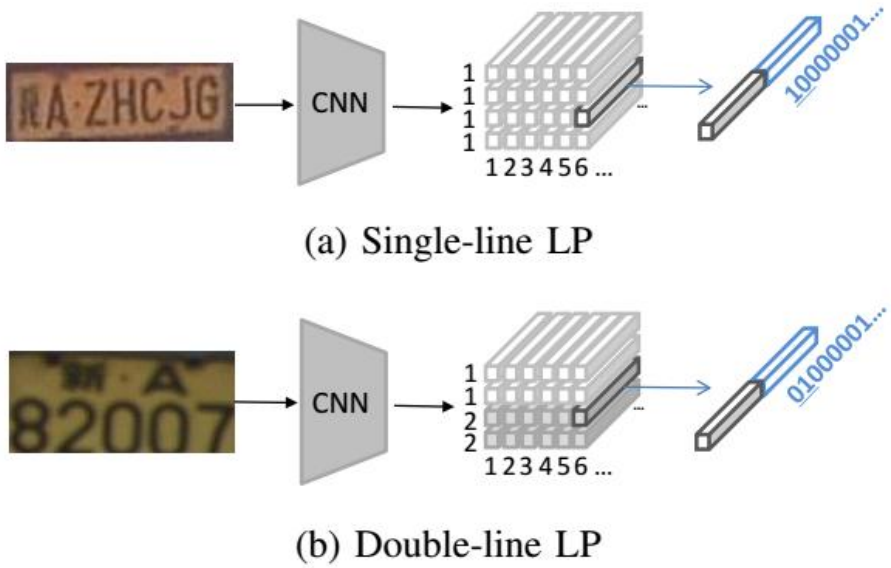
Layer Name	Configurations	Size
Input	-	1 × 32 × 100
Convolution	c:64, k:3 × 3, s:1 × 1, pad:1	64 × 32 × 100
MaxPooling	k:2 × 2,s:2 × 2	64 × 16 × 50
Convolution	c:128,k:3 × 3, s:1 × 1, pad:1	128 × 16 × 50
MaxPooling	k:2 × 2,s:2 × 2	128 × 8 × 25
Convolution	c:256,k:3 × 3, s:1 × 1, pad:1	256 × 8 × 25
MaxPooling	k:2 × 2,s:2 × 2	256 × 4 × 12
Convolution	c:512,k:3 × 3, s:1 × 1, pad:1	512 × 4 × 12
AvgPooling	k:2 × 2,s:2 × 2	512 × 1 × 1
fc1	256	256
fc2	9	9

c,k,s,pad represent channel,kernel, stride and padding sizes respectively.

- **Perspective transformation prediction**: The rectification network directly predicts an offset matrix $\mathbf{T_p}$, and the perspective transformation **T** is calculated.
- **Grid sampling**: The sampling grid is generated by **T** and properties of perspective transformation.

Proposed Method

◆ 2D attention based recognition network



(a) Single-line LP

(b) Double-line LP

- 2D attention mechanism is adopted to predict license numbers sequence directly from the 2D feature maps.
- the one-hot encoding of the simplified spatial coordinates is used to make α_t sensitive to single-line or double-line.

$$r_j = \begin{cases} 1, & single-line \\ \left[\frac{j}{2} + 0.5\right], & double-line \end{cases}$$

$$H = W_h h_{i,j} + W_s s_{t-1} + W_i f_i + W_{r_j} f_{r_j} + b$$

$$e_{t,i,j} = V_a{}^T tanh(H)$$

$$\alpha_{t,i,j} = \frac{exp(e_{t,i,j})}{\sum_{i,j}(exp(e_{t,i,j}))}$$

Experiments

◆ Datasets

- **Synthetic data** are generated randomly with various brightness, chroma, clarity and angles of view.
- **CLPD** dataset contains about 260,000 Chinese single-line LPs collected from different security and surveillance cameras
- **CCPD** is collected in the parking lot of Hefei province of China.
- **DLTD** is a private double-line plates dataset collected from surveillance cameras of traffic crossroads.
- **SYSU-ITS** dataset is a public LP image set, which is provided by OpenITS.

(a) Synthetic data (b) CLPD (c) CCPD (d) DLTD (e) SYSU-ITS

Experiments

◆ Rectification module

- The ablation study is carried out to validate the performance gain of the rectification component.
- The rectified LP images of different rectification methods are shown to validate the visualization performance of PRN.

TABLE II
PERFORMANCE OF RECTIFICATION MODULE

Rectification	Recognition	CLPD	T-CCPD
NULL	2D-Attn	98.12%	72.72%
TPS[10]		98.33%	80.78%
MORN[11]		98.22%	74.27%
PRN (ours)		**98.5%**	**82.2%**

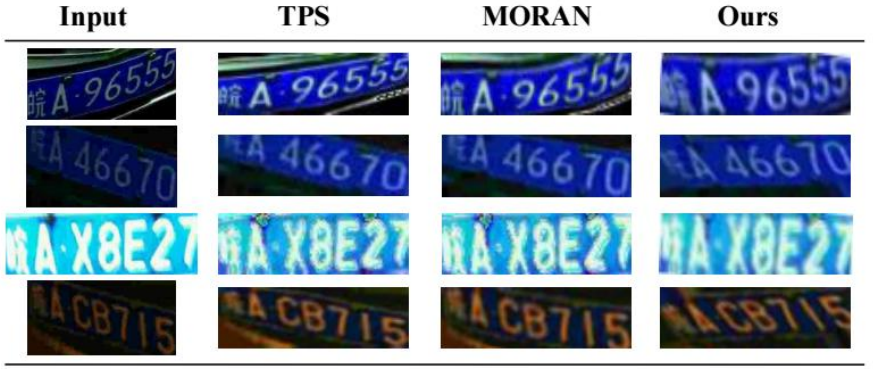


Fig. 4. Rectified images generated by different methods.

Experiments

◆ Recognition module

Input	Method	Rectified Image	Attention Output	Prediction
 皖A0J201	Baseline			A0J2D1
	2d-Attn	-		苏A0J201
	Ours			皖A0J201
 皖AZ2Z46	Baseline			豫AZ224G
	2d-Attn	-		1AZ2Z4G
	Ours			皖AZ2Z46
 新A92290	Baseline			92290
	2d-Attn	-		新A922290
	Ours			新A92290
 HEHE HAHA	Baseline			HAHA
	2d-Attn	-		HEHEHAHA
	Ours			HEHEHAHA

- The recognition module performs well on both single-line and double-line plates.
- The overall model is improved by the perspective rectification module and location-aware attention.

TABLE III
PERFORMANCE OF RECOGNITION MODULE, WHERE RECT. REPRESENTS RECTIFICATION MODULE.

Rect.	Recognition	CLPD	T-CCPD	D-SYSU
-	1D-Attn [10]	97.4%	59.7%	87.72%
	2D-Attn [34]	98.12%	72.72%	93.8%
LPRNet[8]		97.3%	67.41%	-
Yu [18]		94.5%	68.4%	89.1%
Baseline [10]		97.71%	79.26%	90.4%
Ours		**98.7%**	**83.1%**	**94.2%**

Conclusions

- We have proposed a novel 2D license plate recognition method based on automatically perspective rectification for LPR.

- We have improved the recognition performance of the license plates with heavy deformations in the images.

- In the future work, the LP detection and the model efficiency are the pivotal research.

THANK YOU !