北京邮电大学
Beijing University of Posts and Telecommunications

JOINT LICENSE PLATE SUPER-RESOLUTION AND RECOGNITION IN ONE MULTI-TASK GAN FRAMEWORK

Minghui Zhang, Wu Liu, Huadong Ma
Beijing Key Lab of Intelligent Telecommunication Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876, China
{mhzh, liuwu, mhd}@bupt.edu.cn

Introduction:

- We propose **an end-to-end deep learning-based license plate super-resolution and recognition system** for unconstrained urban surveillance scenes, which is different from the traditional process of LPR.
- We propose **a multi-task discriminative network** in MTGAN to combine the super-resolution and recognition in an adversarial manner to enhance each other.
- We propose **a generative network** in MT-GAN to combine prior knowledge from data distribution and domain knowledge of license plate to generate the spatial corresponding and high-resolution plate images.

Approach:

1. Faster RCNN based Detection

- We employ Faster RCNN to detect the license plates in images captured in urban surveillance.

2. GN for Super-Resolution

- We adopt GN to generate standard license plates for each detected license plate.
- We adopt a hierarchy of K residual blocks as the backbone of GN.
- After the K residual blocks, a deconvolutional layer with stride of two is adopted to enlarge the resolution of the feature map.
- At the tail of the network, we use a deconvolutional layer with three filters of 1×1 to generate the three-channel high-resolution images.

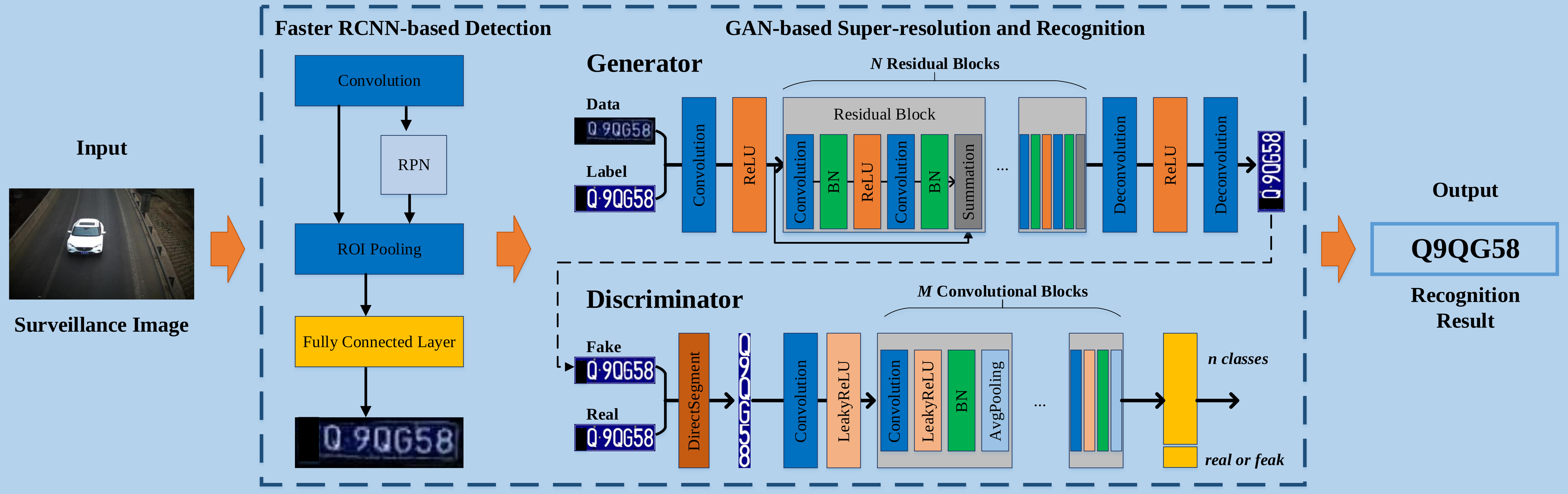

Fig. 1. The overview of the proposed joint license plate super-resolution and recognition system based on MTGAN.

3. DN for Segmentation and Recognition

- The Direct Segmentation Layer (DSL) segments the each license plate into 6 character blocks as follows:

$$W_i = W * a_i\,, \qquad H_i = H\,.$$

- The backbone of the discriminator employs M convolution layers and a fully connected (FC) layer (M = 8 in our implementation).
- At last, a sigmoid function is connected to the tail of the network as a classifier.

4. Adversarial Loss Function

- For the DN, the classification loss function adopts the Binary Cross Entropy (BCE) between the label and the output of DN for both real images and fake images. The concrete formulation is:

$$L_D = L_{BCE} = L_R + L_F = -1/B * N \sum_i (t_i * \log(o_i) + (1 - t_i) * \log(1 - o_i))\,.$$

- For the GN, we utilize the Mean Squared Error (MSE) and the return loss of the fake images from DN as follows:

$$L_{MSE} = 1/WH \sum_{x=1}^{W} \sum_{y=1}^{H} (I_{g(x,y)} - G(I_l)_{(x,y)})^2\,,$$
$$L_G = L_{MSE} - L_F\,.$$

Evaluation and Conclusions:

1. Dataset

- We construct a license plate dataset, containing 12,170 license plates, synthesized standard license plates, and labels.

2. Evaluation

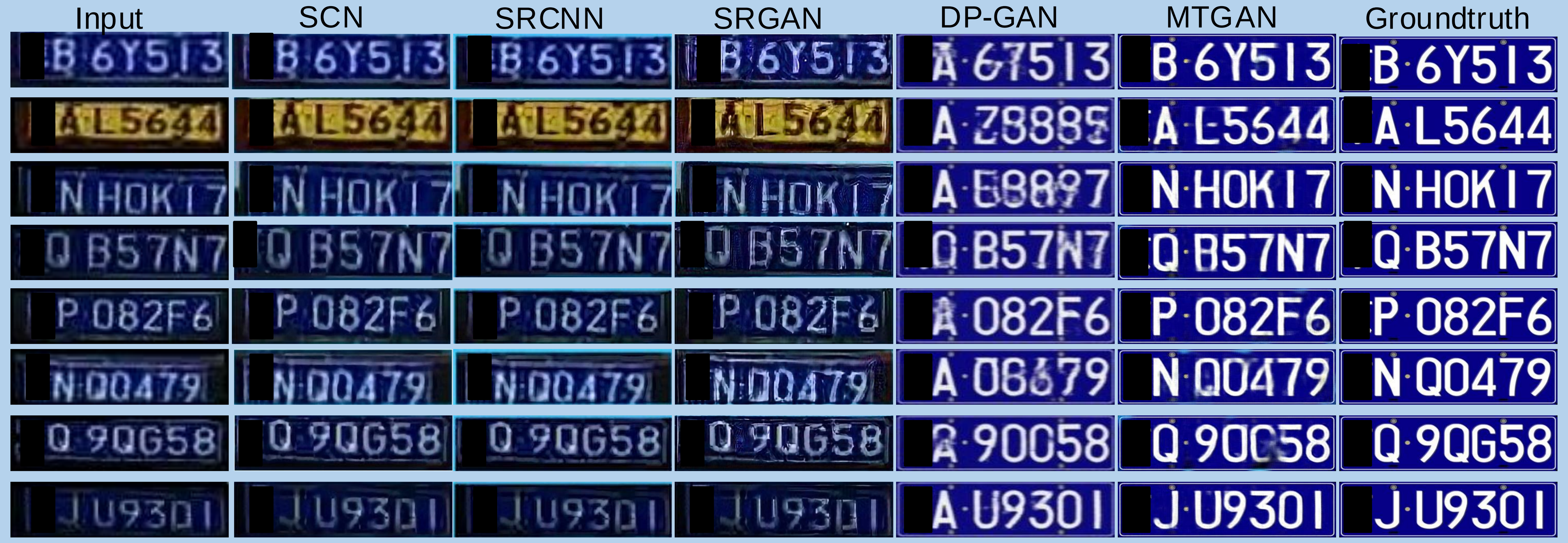

Fig. 2. The examples of different super-resolution approaches.

Method	Origin +EasyPR	SRCNN +EasyPR	GN +EasyPR	**MTGAN (Ours)**
Accuracy	36%	43%	54%	**75%**

Table 1. The recognition accuracy of different methods.

3. Conclusion

- We propose a deep learning-based and end-to-end LPR system for unconstrained urban surveillance scenes, which jointly sharpens the license plates and recognizes the characters in one multi-task based GAN framework.