

A new QR Code Recognition Method using Deblurring and Modified Local Adaptive Thresholding Techniques

Junnian Li, Biao Hu and Zhengcai Cao*, *Member, IEEE*

Abstract—Quick Response (QR) Code would be easily affected by motion blur. In this study, a new recognition method for blurry QR Code is presented. First, a motion deblurring algorithm based on generative adversarial networks (GAN) is applied to perform the blind motion deblurring on single QR Code image for improving the quality of QR Code pictures. Then a modified local adaptive thresholding algorithm applying integral image technology is utilized in order to acquire binary QR Code pictures. Finally, this algorithm would be compared with some previous methods. Furthermore, for testing this algorithm, a database which contains motion blurry QR Code images from both public database and real scenes is collected. Experimental results show that this proposed algorithm perform well on motion blurry QR Code recognition missions.

I. INTRODUCTION

Quick Response (QR) Code is a kind of popular two-dimensional code. In 1994, it was first created in Japan and used in the automotive industry. As presented in Fig. 1, QR Code usually contains a common structure consisting of data module, version module, position module, format information, etc. Nowadays, QR Code has experienced an extremely huge increase because they are widely used in numerous applications such as high-speed machine reading, data storage, website redirection, etc. Due to their simple structure and easily obtaining ways, QR Code is also applied in some complicated environments such as warehouses monitoring. In these warehouses, usual recognition machines may not recognise QR Code accurately and quickly because of the complexity of their environments [1].

Much effort has been made on QR Code recognition. Zhang et al. [2] designed an identification algorithm that recognizes the QR Code to extend the recognition limit in practical situations. Wang et al. [3] proposed an optimal threshold method for reading QR Code encoding information based on maximum likelihood criterion. Although these previous studies perform well in the recognition of 1D or 2D barcodes, they may be severely influenced in several certain scenes. For example, in some warehouses, the recognition of QR Code might be influenced by some motion or focal-blurred, caused by camera shakes, image motion or out-of-focus. Some important regions of QR Code may be affected

This work is supported by the National Key R&D Program of China 2019YFB1312102, National Natural Science Foundation of China under (Grant No.91748102, No.91848103) and the Beijing Leading Talents Program (Grant No.Z191100006119031)

Junnian Li and Biao Hu are with College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, China.

*Zhengcai Cao is with College of Information Science and Technology, Beijing University of Chemical Technology, Beijing, 100029, China, corresponding author (e-mail: giftczc@163.com).

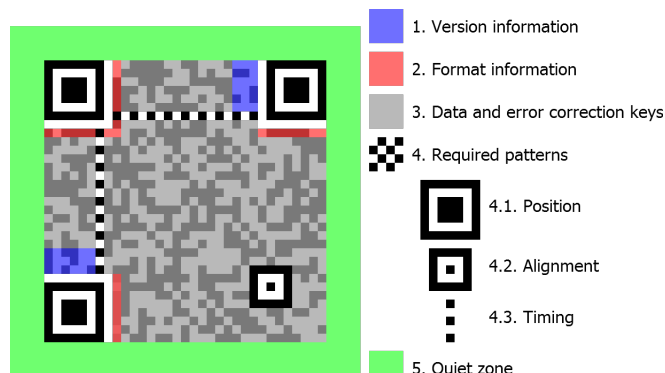


Fig. 1. Structure of QR Code (version 7)

by motion blur and significant information such as position module would be lost. The loss of vital information would seriously harm the visual quality of QR Code pictures. As presented in Fig. 2, under this situation, the structure of QR Code images would be severely wrecked. It can be seen that under this circumstance, QR Code recognition becomes highly challenging to deal with.

In the area of computer vision and machine learning, motion deblurring problems have been widely studied. Traditional algorithms [4]–[7] concentrate on estimating a single blurry kernel in order to recover original images. For example, Levin [4] segmented the image into layers with different blurs and modeled the expected derivatives distributions. U. Schmidt et al. [5] derived a discriminative model cascade and proposed a discriminative approach. They also trained their model by loss minimization and generated their own training data. Zhang et al. [6] presented a algorithm of scale space perspective on single image blind deblurring. They also presented a cascaded scale space expression. Pan et al. [7] introduced a solution which could simultaneously estimate object segmentation and camera motion. In [7], both segmentation and deblurring are evaluated and optimized in one framework. However, these methods are not specially designed for QR Code recognition. Because the deblurring process of QR Code should be real-time, traditional algorithms would not have superior performance.

With the rapid development of deep learning, over the past several years, plenty of algorithms [8]–[11] by applying CNN have been proposed by generating large-scale database and learning the blurry pattern precisely. Gong et al. [8] presented a framework which directly estimated the motion flow from the input blurry photograph and recovered the

clear photograph from the estimated motion flow. Tao et al. [9] investigated the coarse-to-fine scheme and proposed a framework called SRN-DeblurNet (Scale-recurrent Network) for deblurring mission. They also evaluated their algorithm on large-scale deblurring databases with different kinds of motion. Zhang et al. [10] introduced a deep hierarchical multi-patch framework on blind deblurring through a fine-to-coarse hierarchical expression. Kupyn et al. [11] proposed an end-to-end structure called DeblurGAN. This structure is basically based on both the conditional GAN and the content loss.

QR Code images captured by camera may be influenced by highlight spots, camera calibration errors and low quality. These phenomena may result in decoding difficulties of QR Code. Because of this, image preprocessing is a significant process of recognition of QR Code. During this process, graying and binarization are inevitable parts of image segmentation for QR Code recognition. After image-graying, binarization can be applied to distinguish foreground and background. Each pixel of QR Code images could be replaced with a black or a white pixel due to its intensity.

Otsu [12] proposed a method which can automatically select the optimal threshold based on the discriminant criterion. However, because Otsu's method is noise-sensitive, it cannot be well applied in many different environments. Sauvola et al. [13] presented a new method for adaptive image binarization. In Sauvola's paper, two methods were presented to compute a local threshold for each pixel of the image. Soft decision method which includes noise filtering and signal tracking abilities was introduced for background and pictures. And specialized binarization method was raised in terrible environments to separate vital information from background. Lazzara et al. [14] improved Sauvola's method by implementing an efficient multi-scale fulfillment for both small and large things within a document. In [14], an efficient way of how to perform Sauvola's method was also described in details.

In our study, a new method for QR Code recognition under complicated blurry circumstances is given. First, a motion deblurring algorithm based on generative adversarial networks (GAN) is applied to restore clean QR Code pictures from motion-blur ones. Then a modified local adaptive thresholding algorithm is utilized in order to acquire binary QR Code pictures. We would validate this algorithm on a database which is collected from both public database and real scenes. To evaluate the proposed algorithm, it will be compared with several previous algorithms on the same database. An ablation experiment shows that the proposed algorithm perform well than other methods both qualitatively and quantitatively. The robustness of QR Code recognition could be effectively improved by applying the proposed algorithm especially under sophisticated motion-blur environments.

The remainder of this paper is organized as follows. First, the framework of proposed algorithm is presented in Section II. In this section, both motion deblurring algorithm and the modified local adaptive thresholding algorithm are described

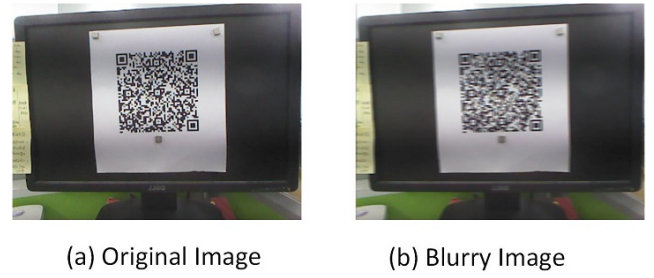


Fig. 2. Original and Blurry images: (a) Original image. (b) Blurry images.

in details. Then implementation details and ablation experimental results are presented in Section III. Conclusions are discussed at last.

II. DEBLURRING AND MODIFIED LOCAL ADAPTIVE THRESHOLDING

In this section, the proposed QR Code recognition method is described in details. The complete framework of this method is presented in Fig. 3. Here, the architecture of deblurring method based on GAN is introduced at the beginning. The model of deblurring method is trained on public database and tested on our own database in order to acquire clean QR Code images from motion-blur ones. Then the modified local adaptive thresholding algorithm applying integral image technology is presented.

A. Deblurring Methods

QR Code images would inevitably face the low quality and visibility problems affected by motion blur. Motion blur is often generated by camera shakes and causes significant information loss of QR Code. Therefore, QR Code would be challenging to be decoded.

The common mathematical non-uniform blur model that would be usually applied in previous single-image blind motion deblurring projects is presented via

$$I_B = k(M) * I_S + N \quad (1)$$

where I_B represents the blurred image, I_S is the original clear image, $*$ depicts the convolution, N is described as additive noise, $k(M)$ are unknown blur kernels (which is also called point spread function, PSF) determined by M , and M is depicted as motion field.

Goodfellow et al. [15] introduced the structure of GAN and it has already been applied in numerous computer vision areas such as image dehazing, style transfer, object recognition, et al. In this paper, GAN is used to generate deblurring QR Code images. The framework of GAN defines two significant competitive modules: generation module and discrimination module. Assuming that the generation module is represented by G and the discrimination module is represented by D . In the space of arbitrary functions G and D , there is a special solution, where G is responsible for restoring the training data distribution, D is used to evaluate the results produced by each iteration of G , and D 's goal is to evaluate as much as possible to obtain the

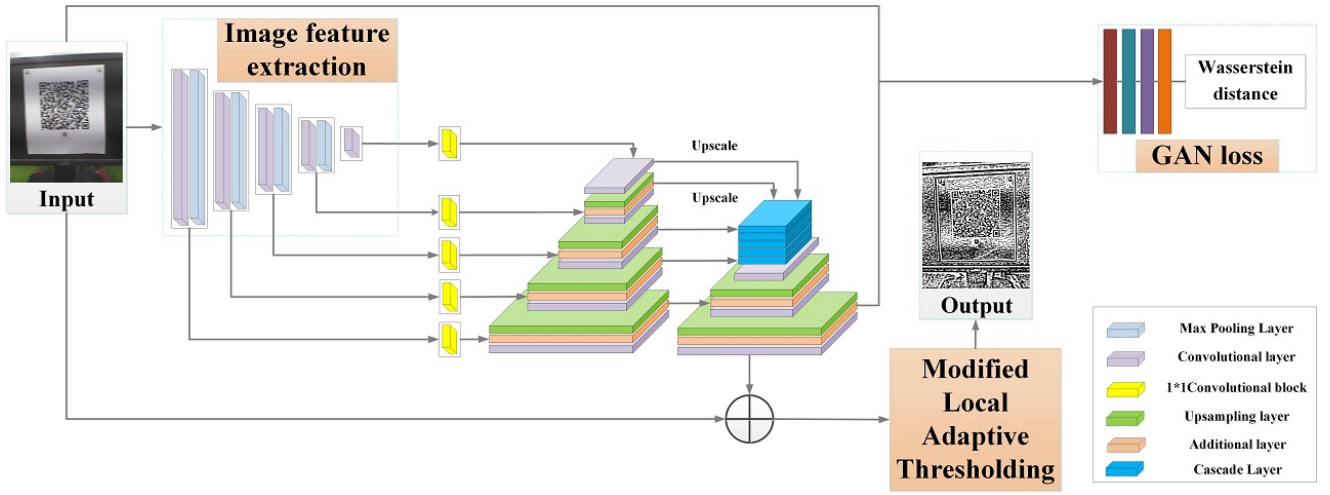


Fig. 3. The proposed Algorithm Structure. It includes Deblurring Method and Modified Local Adaptive Thresholding

true distribution. The distribution generated by D is more realistic than the “high imitation distribution” generated by G . The mathematical model of training process of D and G is written via

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))] \quad (2)$$

where the data distribution and the model distribution is respectively depicted by $p_{data}(x)$ and $p_z(z)$. Within the formula $\mathbb{E}_{x \sim p_{data}(x)} [\log D(x)]$, $D(x)$ describes the probability that x judges it to be real data through D . Within the formula $\mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$, $G(z)$ describes the image data generated within G by noise z . The probability of the authenticity of image data generated by G is described by $D(G(z))$. Here, the goal of G optimization is mainly making the result of $D(G(z))$ close to 1. The goal of D optimization is mainly making the result of $D(x)$ close to 1 and the result of $D(G(z))$ close to 0.

Deblurring method first uses Feature Pyramid Networks (FPN) for QR Code image features extraction. The framework of FPN contains both a bottom-up pathway and a top-down pathway. The bottom-up pathway is applied for QR Code image features extraction and is designed using general convolutional network. Furthermore, the bottom-up pathway is also applied for downsampling the spatial resolution. Simultaneously, semantic information within this is extracted and compressed. Meanwhile, the top-down pathway is used for reconstructing higher spatial resolution by upsampling spatially coarser from semantically rich layers.

In our study, a framework including FPN architecture is presented. Within this framework, five final feature maps from different layers are extracted as the output. Then these features are up-sampled to the $\frac{1}{4}$ original input size. After that, these features are concatenated into one tensor and semantic information from different levels will be recorded on that tensor. Two upsampling and convolutional layers are additionally added to acquire clear QR Code images with original input size. Except that, artifacts would be reduced

and the performance of the proposed algorithm would be improved obviously. In order to focus on the residue between the input motion blurry images and the output clear ones, a direct skip connection is added.

Compared with framework of DeblurGAN-v2 [16], batch normalization layer is added at the end of each convolutional layer in order to increase the generalization of the proposed algorithm. As we know, the training based on large-scale database may result in overfitting of data. In order to reduce this phenomenon, batch normalization is applied to normalize the data of all layers before the activation function are used to the same data distribution. Then these data would be transferred to the next layer. This can effectively reduce the “jitter” of data, and the output in the middle of each convolutional layer will be more stable, effectively reducing the risk of data overfitting. The use of batch normalization can also help us use a large initial learning rate and make the learning rate decay faster. This will make the algorithm converge faster than the original and will obviously reduce the training time. The use of batch normalization will also speed up the training of deep convolutional neural networks, and the generalization ability of the algorithm will be much better than before.

Inspired by recently proposed improved Wasserstein GAN [17], we design the loss function for our deblurring method. In this study, the loss function of discrimination module (D_loss function) is described via

$$L(D) = -\mathbb{E}_{x \sim P_r} [D(x)] + \mathbb{E}_{x \sim P_g} [D(x)] + \lambda \mathbb{E}_{x \sim P_{\tilde{x}}} [||\nabla_x D(x)|| - 1]^2 \quad (3)$$

where $-\mathbb{E}_{x \sim P_r} [D(x)] + \mathbb{E}_{x \sim P_g} [D(x)]$ is original critic loss and $\lambda \mathbb{E}_{x \sim P_{\tilde{x}}} [||\nabla_x D(x)|| - 1]^2$ is gradient penalty. The loss function of generation module (G_loss function) is described via

$$L(G) = -\mathbb{E}_{x \sim P_g} [D(x)] \quad (4)$$

The overall loss is constructed due to both recently proposed perceptual distance “content” loss L_X and mean-square-error (MSE) loss L_P . The G_loss function in this study

is defined via

$$L(G) = 0.01 * L_{adv} + 0.006 * L_X + 0.5 * L_P \quad (5)$$

within the above loss function, both global and local discriminator losses are in L_{adv} . It can be found that L_P would help us correct color and texture distortions. The training process of D_loss and G_loss function is shown in the experimental section.

B. Modified Local Adaptive Thresholding

Transforming QR Code images into binary ones can save significant information and reduce the amount of redundant data at the same time. So binarization is a vital method for the recognition of QR code images. In this paper, modified local adaptive thresholding algorithm using integral image technique by referencing to [18] is applied to fulfill this mission. First, QR Code images are transformed into grayscale ones. Then the adaptive threshold would be computed. After that, each pixel of grayscale images is sorted into two classes: black or white.

1) Image Graying

First, all of the QR Code images are transformed into grayscale ones by using the standard mathematical model which is described via

$$Gray = 0.3 * R + 0.59 * G + 0.11 * B \quad (6)$$

where R, G, B represent pixel of the red, green and blue channels.

2) Threshold Computing

The threshold of the image would be computed in order to separating each pixel of the grayscale images. If the pixel of the QR Code image is larger than the threshold, it would be set to white. Otherwise it would be set to black. The process is described via

$$I_b(x,y) = \begin{cases} 255, I_s(x,y) \geq T \\ 0, I_s(x,y) < T \end{cases} \quad (7)$$

where T is the threshold that would be computed. I_s is the input colored image which would be changed into the filtered grayscale image I_b . The threshold should be computed accurately because it would influence the result of the binarization. The way of how to calculate threshold is described as follows.

For a moving $w * w$ window, R_s is defined as the adaptive threshold over the rectangle R. The mathematical model of R_s is described via

$$R_s = I(p_2, q_2) - I(p_1, q_2) - I(p_2, q_1) + I(p_1, q_1) \quad (8)$$

where $p_1, p_2 = p + \frac{w}{2}$ and $q_1, q_2 = q + \frac{w}{2}$. In this equation, I depicts integral value at any point (p, q) . Within R, the number of pixels is computed via

$$C = (p_2 - p_1) * (q_2 - q_1) \quad (9)$$

The representation of adaptive thresholding is described via

$$R_s * (1 - T) \geq I(p, q) * C \quad (10)$$

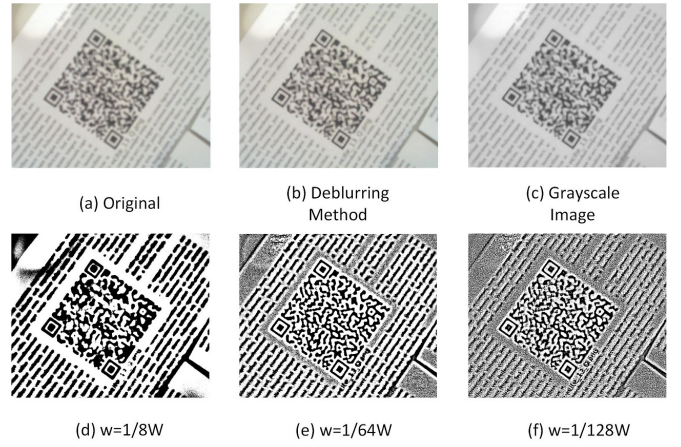


Fig. 4. The comparison results of w:(a) Original Image. (b) After Deblurring Method. (c) Grayscale Image. (d) $w = \frac{1}{8}W$. (e) $w = \frac{1}{64}W$. (f) $w = \frac{1}{128}W$. (e) could be directly decoded, (d) and (f) could not be directly decoded.

where T can be defined by ourselves. In this paper, T is set as 0.15. Here, if the intensity of each pixel is lower than R_s , it is set to black, otherwise it is set to white.

The parameters which are used for our algorithm are $w = \frac{1}{64}W$, where W is the size of input images and w depicts maximum width of the window. According to previous method, w is set due to practical scenes. In this study, according to experimental results, when windows size $w = \frac{1}{8}W$ or $w = \frac{1}{128}W$, QR Code images cannot be easily decoded. However, when $w = \frac{1}{64}W$, experimental results show that better results can be obtained. An ablation experiment is presented in Fig. 4.

In the end, the threshold is calculated accurately by using the proposed algorithm and binary QR Code images are generated successfully. In general, binary QR Code images could help us decode QR Code images more efficiently and quickly.

3) Integral Image Computing

In our study, the definition of the integral images would be applied. For an input QR Code photograph, its integral image $I(p, q)$ is defined as the image whose intensity $i(p, q)$ at the position (p, q) equals to the summation of intensities of the whole pixels upper left of that position within the original input photograph $h(p, q)$. The mathematical model of intensity at the position (p, q) can be written via

$$I(p, q) = \sum_{i=0}^p \sum_{j=0}^q h(p, q) \quad (11)$$

III. EXPERIMENTS

A. Dataset

A public database [19] of challenging pictures of QR Code is applied in our study, which is specially created for QR Code images recognition. We select 50 motion blurry images provided by this public database for testifying the proposed algorithm. Then we collected a database which contains motion blurry QR Code pictures from real scenes.

We also randomly select 50 motion blurry images provided by the database collected by ourselves. So, the test set includes totally 100 pictures provided by both public and our own database. The test results between our proposed algorithm and other previous methods are shown in the ablation comparison section.

In the first row of Fig. 5, several original images from the database collected by ourselves could be seen. And their motion blurry images which are created with MATLAB are in the second row. In our own database, QR Code images contain some chemical information. The degradation of QR Code images affected by motion blur is dependent on their motion angle θ and length L . So there are 450 motion blurry QR Code images in our database wherever L takes in the range $[1, 15]$ and θ takes in the range $[20, 180]$ in increments of 20. We primarily focus on the accuracy of the decoding of QR Code, so we select all motion blurry QR Code images that could not be directly decoded by using our decoding program to testify the performance of all methods.

B. Implementation Details

Our processor possesses Intel(R) Core(TM) i7-9750H CPU@2.60GHz, NVIDIA GeForce GTX 1660 Ti and 16 GB memory. The deblurring method is performed by using Python and PyTorch. During the training process, Adam optimizer is applied and learning rate is set to 10^{-4} for 200 epochs. The public dataset GoPro [20] is first utilized to train the proposed deblurring method with its 2103 image pairs including both motion blurry images and original ones. And the database we collected is applied as test set. The training model is obtained to deblur motion blurry QR Code images and the deblurring results are presented in the third row of Fig. 5. The training process of D_loss function, G_loss function, “content” loss function L_X and L_{adv} function is presented in Fig. 6.

Then the modified local adaptive thresholding method is applied for transforming deblurring QR Code images into binary ones. In the fourth row of Fig. 5, the binarization results could be seen. Within this binary method, the parameters are set as $w = \frac{1}{64}W$ and $T = 0.15$. As presented in Fig. 4, by setting the parameter $w = \frac{1}{64}W$, better binarization results can be acquired. In our study, the binarization method is performed by using MATLAB.

In the end, QR Code images are decoded with our own decoding program performed with MATLAB. Information within these images is presented in the last row of Fig. 5.

C. Ablation Comparison

Different deblurring algorithms may affect the robustness of QR Code recognition. For example, vital information of QR Code images such as position module would be lost during the process of deblurring. In this section, several different deblurring methods involving traditional and recently proposed CNN methods are compared by using the same database: (a) Sparse Blind Deblurring & adaptive thresholding; (b) Adaptive Thresholding; (c) DMPHN & adaptive thresholding; (d) Our proposed algorithm. Our test

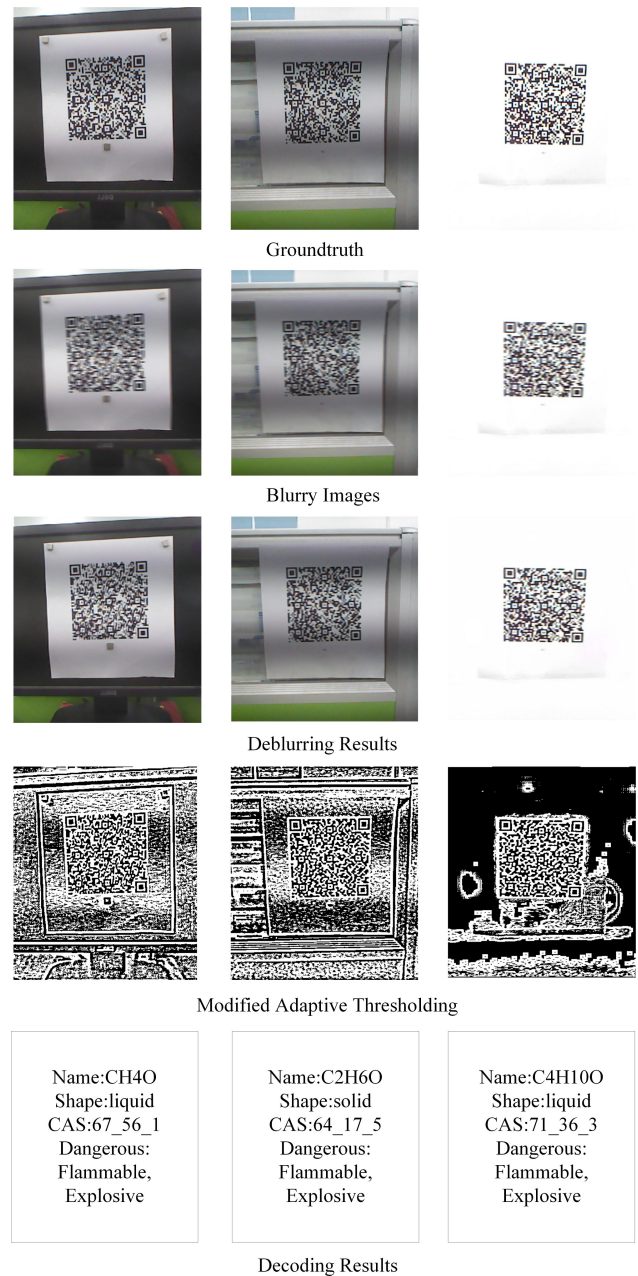


Fig. 5. Experiment Process. Row 1: Groundtruth QR Code images. Row 2: Blurry Images. Row 3: Deblurring Method Results. Row 4: Modified Local Adaptive Thresholding Results. Row 5: Information of these QR Code.

set includes 100 pictures provided by both public and our own database.

As shown in Table 1, under same conditions, the decoding accuracy of QR Code of traditional deblurring methods is obviously lower than recently proposed CNN-based methods, and its processing time of single image is also longer than these CNN-based methods. We can find that during the processing of traditional deblurring methods, some vital information of QR Code pictures lost. However, CNN-based deblurring methods restore more clear pictures and be more robust in motion blurry conditions. The accuracy of proposed algorithm is higher than any other algorithms, which means

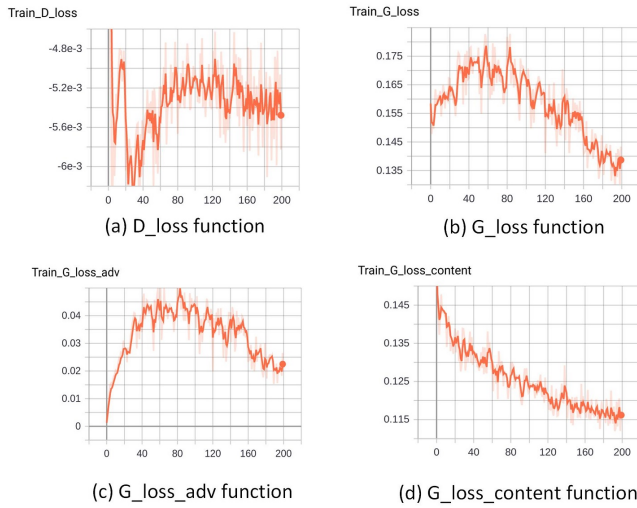


Fig. 6. The training process of D_loss and G_loss function: (a) D_loss function. (b) G_loss function. (c)G_loss_adv function. (d)G_loss_content function.

the proposed algorithm behaves more robust in sophisticated motion blurry scenes.

TABLE I
ACCURACY COMPARISON RESULTS

Algorithm	Total Pictures	Accuracy (%)	Time (s)
Sparse Blind Deblurring[6] & adaptive thresholding	100	25	433.75
Adaptive Thresholding	100	26	30.25
DMPHN[10] & adaptive thresholding	100	35	30.37
Our proposed algorithm	100	46	30.95

IV. CONCLUSIONS

This study introduces a new QR Code recognition method using deblurring and modified local adaptive thresholding techniques under motion blurry conditions. First, motion blurry QR Code images provided by both public and our own database are selected as test set for evaluating our proposed algorithm. Deblurring QR Code images are acquired by using high-quality deblurring algorithm based on GAN. Then modified local adaptive thresholding method is applied in order to transform deblurring QR Code images into binary ones. In the end, binary QR Code images are decoded and recognition accuracy of different motion deblurring methods both traditional and CNN-based are compared. An ablation experiment shows that the proposed algorithm could make the detection of QR Code more accurate in same conditions.

It behaves more robust than any other methods in motion blurry conditions. The robustness and processing time of proposed algorithm would be continually improved in the future.

REFERENCES

- [1] Y. Cheng, Z. Fu, B. Yu, "Improved visual secret sharing scheme for QR code applications," *IEEE Transactions on Information Forensics and Security*, vol.13, no.9, pp. 2393–2403, 2018.
- [2] J. Zhang, Z. Yin, L. Zhao, "Extending the Recognition Limit of Two-Dimensional Bar Code," *IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS)*, pp. 170–176, 2019.
- [3] W. Xuan, C. Peng, C. Fang-Fang, et al, "Research on the Optimal Threshold of QR Code Recognition Based on Maximum Likelihood Criterion," *IEEE 4th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, pp. 107–111, 2018.
- [4] A. Levin, "Blind motion deblurring using image statistics," *IEEE International Conferences on Neural Information Processing Systems*, pp. 841–848, 2007.
- [5] U. Schmidt, C. Rother, S. Nowozin, et al, "Discriminative non-blind deblurring," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 604–611, 2013.
- [6] H. Zhang, J. Yang, "Scale adaptive blind deblurring," *IEEE International Conference on Neural Information Processing Systems*, pp. 3005–3013, 2014.
- [7] J. Pan, Z. Hu, Z. Su, et al, "Soft-segmentation guided object motion deblurring," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 459–468, 2016.
- [8] D. Gong, J. Yang, L. Liu, et al, "From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 2319–2328, 2017.
- [9] X. Tao, H. Gao, X. Shen, et al, "Scale-recurrent network for deep image deblurring," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 8174–8192, 2018.
- [10] H. Zhang, Y. Dai, H. Li, et al, "Deep Stacked Hierarchical Multi-patch Network for Image Deblurring," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 5978–5986, 2019.
- [11] O. Kupyn, V. Budzan, M. Mykhailych, et al, "Deblurgan: Blind motion deblurring using conditional adversarial networks," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 8183–8193, 2018.
- [12] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol.9, no.1, pp. 62–66, 1979.
- [13] J. Sauvola, M. Pietikäinen, "Adaptive document image binarization," *Pattern recognition*, vol.33, no.2, pp.225–236, 2000.
- [14] G. Lazzara, T. Gaud, "Efficient multiscale Sauvolas binarization," *International Journal on Document Analysis and Recognition (IJ DAR)*, vol.17, no.2, pp. 105–123, 2014.
- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al, "Generative adversarial nets," *IEEE International Conference on Neural Information Processing Systems*, pp. 2672–2680, 2014.
- [16] O. Kupyn, T. Martyniuk, J. Wu, et al, "Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 8878–8887, 2019.
- [17] I. Gulrajani, F. Ahmed, M. Arjovsky, et al, "Improved training of wasserstein gans," *IEEE International Conference on Neural Information Processing Systems*, pp. 5767–5777, 2017.
- [18] K. Peuwun, K. Woraratpanya, K. Pasupa, et al, "Modified Adaptive Thresholding Using Integral Image," *International Joint Conference on Computer Science & Software Engineering*, pp. 1–5, 2005.
- [19] M. Dubskä, A. Herout, J. Havel, "Real-time precise detection of regular grids and matrix codes," *Journal of Real-Time Image Processing*, vol.11, no.1, pp. 193–200, 2016.
- [20] S. Nah, T. H. Kim, K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 3883–3891, 2017.