

Fusion network for blur discrimination

Yumeng Tian,^a Mingzhang Luo,^{a,b} and Luoyu Zhou^{© a,b,*}

^aYangtze University, Electronic and Information School, Jingzhou, China

^bYangtze University, National Demonstration Center for Experimental Electrical and Electronic Education, Jingzhou, China

Abstract. Blurry image discrimination is a challenging and critical problem in computer vision. It is useful for image restoration, object recognition, and other image applications. In previous studies, researchers proposed a discrimination method based on hand-extracted features or deep learning. However, these methods are either pure data driven by deep learning or over-simplified assumptions on prior knowledge. As a result, a discrimination method is proposed for distinguishing sharp images and blurry images based on a fusion network. The proposed method can automatically discriminate and detect blur without performing image restoration or blur kernel function estimation. Actually, the blur and the noise are extracted by the improved VGG16 network and texture noise extraction algorithm, respectively. Then the fusion network integrates the advantages of deep learning and hand-extracted features, and achieves ultimate high-accuracy discrimination results. Rigorous experiments performed on own dataset and other popular datasets with a number of blurry images and sharp images, including RealBlur dataset, BSD-B dataset, and GoPro dataset. The results show that the proposed method outperforms with an accuracy of 98% on our own dataset and 94.8% on the other dataset, which satisfies the requirements of the image applications. Similarly, we have compared our method with state-of-the-art methods to show its robustness and generalization ability. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JEI.30.3.033030](https://doi.org/10.1117/1.JEI.30.3.033030)]

Keywords: blur discrimination; VGG16 network; texture noise parameters; fusion network.

Paper 210106 received Feb. 27, 2021; accepted for publication Jun. 7, 2021; published online Jun. 18, 2021.

1 Introduction

Digital images have become an indispensable core information carrier in the fields of computer vision and artificial intelligence. However, during the processes of image acquisition and image transmission, the image is inevitably contaminated by blur. Actually, blur is almost an omnipresent effect on natural images. Blur discrimination is great benefit to the subsequent image processing, including depth estimation, image quality assessment, information retrieval, image restoration, and others.¹⁻⁴ As a result, blur discrimination has already become an important problem in image and video processing systems.

Although blur discrimination has attracted much attention in recent years, most previous work focuses on solving the deblurring problem. On the contrary, more general blur discrimination is seldom explored and still far from practical applications. Various prior knowledge can be extracted from the statistics of natural images.⁵⁻⁹ Among this prior knowledge, dark channel prior deserves a special mention due to its restoration performance.⁸ Pan et al.⁸ found that most image patches in the sharp image contained some dark pixels and these pixels were not dark when averaged with neighboring high-intensity pixels during the blur process. This feature, called a dark channel, is a great benefit to deblurring. However, blur can decrease the dark channel, but the dark channels of some blurry images are more than those of some sharp images. As shown in Fig. 1, it is obvious that the dark channels of the blurry image are more than those of the sharp image. Therefore, it is impossible to discriminate the blur only by dark channel or other single prior knowledge.

*Address all correspondence to Luoyu Zhou, luoyuzh@yangtzeu.edu.cn



Fig. 1 The dark channel of sharp image and blurry image: (a) sharp image, (b) dark channel of (a), (c) blurry image, and (d) dark channel of (c).

In the last 20 years or so, several practicable methods are proposed for blur discrimination.^{10–16} Most of them employ a two-step strategy. First, some low-level blur-related features are hand-crafted based on various empirical image statistics in gradient, frequency, and other domains. Then, a binary classifier is used for blur discrimination. Therefore, a crucial issue of blur discrimination is to achieve useful blur features. Although the hand-crafted features are simple and low-dimensional, their discriminative and expressive capabilities are still improved. The latest advances have shown that the deep learning can extract superior deep features for blur discrimination,^{17,18} despite poor generalization capability.

Inspired by the above research, this paper designs a fusion network structure, which integrates the classification network and the texture noise extraction algorithm. The classification network is proposed by improving the VGG16 network. The noise extraction algorithm is improved based on wavelet estimation method. Experimental results demonstrate that the proposed method is successful for blur discrimination. In this paper, several contributions are summarized as follows.

- We propose a fusion network that not only fuses the advantages of the existing discrimination methods based on hand-extracted features and deep learning, but also prompts the robust convergence of the discrimination network. In the process of jointly training the entire network, the proposed method only demands a small number of training samples relative to other deep convolutional neural networks methods and achieves the superior discrimination performance.
- The classification network is proposed by improving the VGG16 network with the additional dropout layers, which can suppress the overfitting problem of the original network.
- The texture noise extraction algorithm is introduced by improved wavelet estimation method, which can effectively solve noise disturbance and improve discrimination accuracy.

The rest of this paper is organized as follows. The related works on blur discrimination are presented in Sec. 2. Three important parts of our proposed method are detailed in Sec. 3. Experimental results and analyses are presented in Sec. 4, and conclusion is given in Sec. 5.

2 Related Work

Blur discrimination is a challenging and long-studied topic in image processing and analysis. As far, the blur discrimination methods can be categorized into two groups: methods based on hand-extracted features and methods based on deep learning, which are discussed in this section.

2.1 Methods Based on Hand-Extracted Features

The methods based on hand-extracted features are proposed based on image statistical features. Shi et al.¹⁰ proposed a discrimination method using local filtering space, Fourier transform, and image gradient. These features are adaptive to blur scales in different images. Xu et al.¹¹ proposed several blur features using different image statistical information, including color, image gradient, and spectral information. Khan et al.¹² proposed a blur discrimination method by frequency-based multi-level fusion transformation, which could detect and classify the blur and non-blur by single image processing. Rugna and Konik¹³ observed that the blur was insensitive

to low-pass filtering. They utilized this feature to judge whether a given image was blurry or not. Liu et al.¹⁴ focused their attention on low-level features and proposed a blur discrimination method through some image features, including local auto-correlation congruency, gradient histogram span, spectrum slope, and maximum saturation. Teo and Zhan¹⁵ proposed a detection method for the blurry image by integrating image-derived features and position and orientation system-derived features. Wang et al.¹⁶ proposed a blur detection method for iris image based on local features, which are generated by radial symmetry transform and support vector machine. Gueraichi and Serir¹⁷ proposed a simple model for blur discrimination, which is based on discrete cosine transform associated to support vector machine. The experimental results are somewhat convincing.

These methods based on hand-extracted features are flexible in the extraction of prior knowledge, but suffer from over-simplified assumptions on prior knowledge.

2.2 Methods Based on Deep Learning

With the development of deep learning, several discrimination methods based on deep learning have been proposed in recent years. Huang et al.¹⁸ studied to learn discriminative blur features via deep convolutional neural networks. They designed an effective network with several feature extraction layers and one binary classification layer, which could accurately achieve patch-level blur likelihood. Zhao et al.¹⁹ studied a multi-stream, bottom-top-bottom, fully convolutional network for blur detection. However, their proposed network only detected defocus blur. Wang et al.²⁰ proposed a fast blur detection method for both motion and defocus blur using an end-to-end deep neural network. It can also detect joint motion and defocus blur and costs little time to implement the network. Zeng et al.²¹ proposed multiple convolutional neural networks (ConvNets) for automatically learning the most locally relevant features of defocus blur. The features related on motion blur and the other blur were not discussed in their paper. Szandała²² proposed a deep convolutional neural network as well as Laplacian method for determining whether an image is blurry or not and showed that deep convolutional neural network has considerable potential for blur discrimination.

In a word, deep learning methods benefiting from end-to-end training enjoy fast speed and powerful learning ability in handling blur features. However, deep learning models may lack the guidance of prior knowledge and be limited by poor generalization ability. The advantages and disadvantages of these discrimination methods have been presented in Table 1.

3 Proposed Method

The proposed method consists of three parts. First, the improved VGG16 network model is used for blur discrimination. Second, the noise parameter is obtained by the introduced noise extraction algorithm. Finally, a fusion network is designed and trained to generate a discriminative model, which integrates the advantages of data-driven deep learning and guidance of prior knowledge, and achieves a high-accuracy discrimination result. The overall flowchart of our proposed method is shown in Fig. 2.

3.1 Improved VGG16 Network

To the best of our knowledge, convolutional neural networks have been widely used in computer vision, including image classification, image segmentation, and the other applications.

Table 1 The advantages and disadvantages of these discrimination methods.

| Method | Advantages | Disadvantages |
|--|--|---|
| Methods based on hand-extracted features | Flexibility in the extraction of prior knowledge | Over-simplified assumptions on prior knowledge |
| Methods based on deep learning | End-to-end method, powerful learning ability | Pure data driven and lack of prior knowledge guide, poor generalization ability |

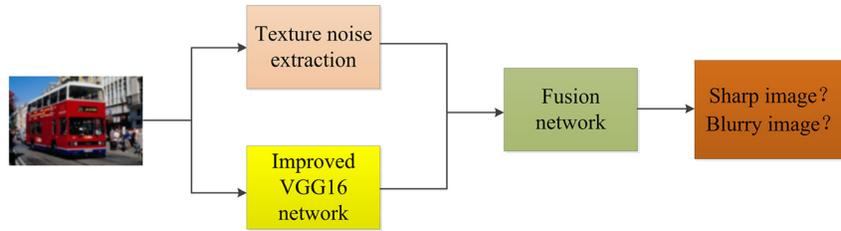


Fig. 2 Flowchart of our proposed discrimination network.

Therefore, we propose a convolutional neural network method to solve the blur discrimination problem. In the discrimination process, blur boundaries do not need to be manually specified. Moreover, it is almost impossible to mark out specific boundaries, which is one of the main limitations of traditional methods. These boundaries hide in the intrinsic prior knowledge of training samples, which can be learned by the dense connected convolutional neural network.

The Visual Geometry Group network (VGGNet) is a classical convolutional neural network,²³ which uses a smaller convolution kernel but deeper network level to extract more small features. It consists of 16 weight layers (13 convolution layers and 3 fully connected layers), which accepts 3-channel RGB image as an input. Moreover, a convolution sequence is formed by stacking a series of 2 or 3 convolution layers (using 3×3 convolution kernel). Each convolution sequence is followed by a maximum pooling layer with 2×2 window size and stride 2. The number of channels in the last three fully connected layers is 4096, 4096, and 1000, respectively. Finally, a softmax classifier with 1000 labels is used for classification. In this section, we chose the 16-layer VGG (VGG16) model as a pre-trained model. Moreover, we modified and developed it to fulfill our requirements for blur discrimination.

It is found that VGG16 network has relatively small loss function when it is directly applied in classification. However, its generalization ability still needs to be improved. In this paper, we adopt a dropout layer to further optimize the generalization ability.

The dropout layer is first proposed in Ref. 24. The principle of dropout layer is to randomly make the weights of some nodes in a hidden layer at a certain ratio stop working during model training. Those nodes that do not work can be temporarily regarded as not being part of the network structure. However, their weights must be retained, and thus the parameters will not be too large. The essence is that when the network extracts the features from training set, it will abandon some features to improve the generalization ability of the network. According to the linear algebra theory, the smaller the parameters are, the simpler the model is, and the less likely it is overfitting. Therefore, we added two dropout layers to suppress the overfitting. The flowchart of improved VGG16 network is shown in Fig. 3. The red box denotes the additional dropout layers and FC-2 layer. The parameters of dropout layer can be adjusted more optimally according to the network evaluation results. In addition, we only need to classify two categories (blurry image and sharp image), and we adjust the structure of VGG16 network to adapt to the number of categories. That is, the number of neurons in the output layer is adjusted to 2 (denoted as FC-2 in Fig. 3).

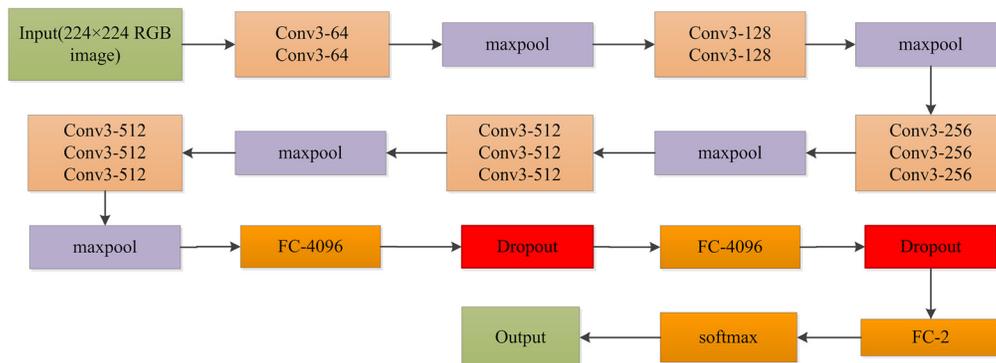


Fig. 3 Flowchart of improved VGG16 network.

3.2 Texture Noise Extraction Algorithm

In Sec. 2.1, it is found that the discrimination results of the improved VGG16 model are sometimes wrong when the image is contaminated by noise. Therefore, the texture noise parameters are introduced as a training element to improve the accuracy. There are many ways to estimate the noise parameters using some statistical characteristics of the image.²⁵⁻²⁷ In our paper, we chose to use wavelet transform to estimate the noise parameters, which was originally presented by Donoho and Johnstone.²⁸ The wavelet estimation algorithm transforms the image into wavelet domain, including low-frequency sub-band coefficients and high-frequency sub-band coefficients. Low-frequency sub-band coefficients reflect the basic information and high-frequency sub-band coefficients reflect noise, edges, and the other texture features. Based on this theory, the simple and effective noise estimation method²⁹ is shown as

$$\sigma = \frac{\text{Median}(|Y(i, j)|)}{0.6745}, \quad (1)$$

where $|Y(i, j)|$ represents the amplitude of high-frequency sub-band coefficients. Median represents median of signal. σ represents the estimated noise parameter and reflects image noise level. However, it is found that the estimated values are all over estimation, especially for the images with low-level noise. The main reason is that all high-frequency sub-band coefficients are taken as noise in origin wavelet estimation. Nevertheless, the high-frequency coefficients include noise, edges, and the other texture features. To decrease this over estimation, we proposed the texture noise extraction algorithm by patch-based wavelet estimation. The detailed flow of our proposed texture noise extraction algorithm is given in Algorithm 1.

By selecting different image patches and achieving their estimated noise parameters. The minimum of these estimated noise parameters represents that the image patch possesses fairly minimal texture details, and thus it is considered as the ultimate noise parameter. In this case, the influence of image texture can be eliminated as much as possible.

To demonstrate superior of the noise extraction algorithm, we test the algorithm on four different images, which are shown in Fig. 4. These images are contaminated with different noises and the noise standard deviations are 5, 8, 11, 15, 20, and 25. The estimated results of our proposed texture noise algorithm by patch-based wavelet estimation (PWE) algorithm and origin wavelet estimation (OWE) algorithm are both shown in Table 2. It is found that the average error of PWE is much smaller than those of OWE, which has shown superior of the proposed texture noise extraction algorithm. The proposed algorithm takes full advantage of prior knowledge and introduces it into the fusion network. It can increase the adaptability of the network to noise and improve discrimination results on the images contaminated with noise.

Algorithm 1 Steps for the texture noise extraction algorithm.

START:

Input: **A** (the given image) and **[m, n]** (the size of image)

For k = 1:5;

m0 = random(1)*(m – 64); Generating random numbers

n0 = random(1)*(n – 64); Generating random numbers

Pa = A(m0:m0 + 64, n0:n0 + 64, :); Randomly select a patch from the given image

Y = wavelet(Pa); Wavelet transform for the selected patches;

Obtain the estimated noise parameters $\sigma(k)$ using Eq. (1);

End For

$\sigma_{\text{ult}} = \min(\sigma)$; Achieve the minimum of five estimated noise parameters

FINISH



Fig. 4 The samples used for verifying the texture noise extraction algorithm.

Table 2 Comparison of the estimation values.

| Noise estimation methods | Real noise | 5 | 8 | 11 | 15 | 20 | 25 | Average error | |
|--------------------------|------------|-----------|-------|-------|-------|-------|-------|---------------|-------|
| Fig. 4(a) | OWE | Estimated | 7.34 | 10.25 | 13.18 | 17.07 | 21.65 | 26.81 | |
| | | Error | 46.8% | 28.1% | 19.8% | 13.8% | 8.3% | 7.2% | 20.7% |
| | PWE | Estimated | 5.33 | 7.69 | 11.28 | 15.44 | 19.91 | 25.88 | |
| | | Error | 6.6% | 3.9% | 2.5% | 2.9% | 0.5% | 3.5% | 3.3% |
| Fig. 4(b) | OWE | Estimated | 8.66 | 11.24 | 14.06 | 17.81 | 22.64 | 27.40 | |
| | | Error | 73.2% | 40.5% | 27.8% | 18.7% | 13.2% | 9.6% | 30.5% |
| | PWE | Estimated | 6.68 | 9.58 | 12.37 | 17.08 | 22.14 | 25.10 | |
| | | Error | 33.6% | 19.7% | 12.5% | 13.9% | 10.7% | 0.4% | 15.1% |
| Fig. 4(c) | OWE | Estimated | 11.40 | 13.70 | 16.30 | 19.63 | 24.06 | 28.55 | |
| | | Error | 128% | 71.3% | 48.2% | 30.9% | 20.3% | 14.2% | 52.1% |
| | PWE | Estimated | 7.80 | 9.89 | 12.79 | 16.14 | 21.10 | 25.76 | |
| | | Error | 56.0% | 23.6% | 16.3% | 7.6% | 5.5% | 3.0% | 18.7% |
| Fig. 4(d) | OWE | Estimated | 5.89 | 8.85 | 11.96 | 16.08 | 20.91 | 25.97 | |
| | | Error | 17.8% | 10.6% | 8.7% | 7.2% | 4.6% | 3.9% | 8.8% |
| | PWE | Estimated | 5.24 | 7.89 | 10.71 | 14.55 | 20.55 | 25.13 | |
| | | Error | 4.8% | 1.4% | 2.6% | 3.0% | 2.8% | 0.5% | 2.5% |

3.3 Fusion Network

To integrate the advantages of the improved VGG16 network and texture noise extraction algorithm, a fusion network is designed in this section. The fusion network is a back propagation (BP) neural network. We chose BP neural network to build a fusion network because of its strong non-linear mapping ability, high self-learning, adaptive ability, and fault tolerance ability.

BP neural network, a multi-layer feedforward network, is trained by the error BP algorithm.³⁰ Its main characteristic is that the signal propagates forward and the error propagates backward. The learning rule of BP neural network is the steepest descent method,³¹ which continuously adjusts the weights and biases of the network to minimize the sum of squared errors of the network. Moreover, the loss function used in this paper is the cross-entropy function, which is shown as follows:

$$\text{loss} = - \sum_{i=1}^n y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i), \quad (2)$$

where loss represents the loss function, n represents the number of categories, equaling to 2 in this network (blurry images and sharp images), \hat{y}_i represents the predicted probability, and y_i represents the true sample label. This loss function is an effective and popular loss function and can be minimized to approximate the real results.³²

The fusion network integrates the deep learning results and hand-features (texture noise), and comprehensively discriminates whether the image is blurry or not. It will increase the discrimination accuracy, especially for the images contaminated with noise.

3.4 Implemented Specifics

The improved VGG16 network demands that the size of input image is 224×224 . Therefore, the testing image must be cropped to size of 224×224 . However, it will lead to some random errors because of local texture features of images. Therefore, we crop out five sub-images from different regions, including the upper left corner, lower left corner, middle, upper right corner, and lower right corner. For sub-image 1 or 4 in Fig. 5, it is easy to make incorrect discrimination because of slightly blurry background regions. Therefore, through achieving average of discrimination values in different regions, it will reduce the random error and improve discrimination results. This will overcome the influence of the local texture features. The fusion network integrates the average blur probability and the texture noise parameter, and then achieves the final discriminative results. The implementation specifics of the overall discrimination method are given in Algorithm 2.

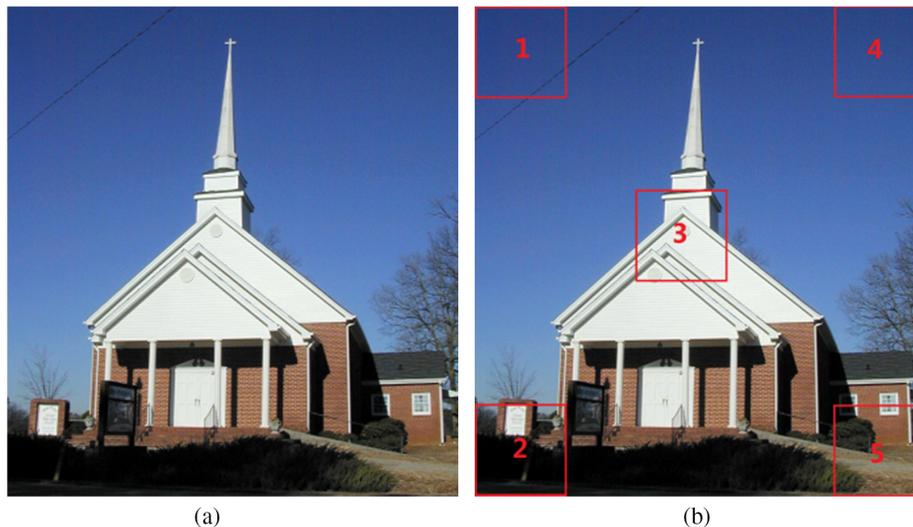


Fig. 5 Samples of cropped sub-images: (a) original image and (b) cropped sub-image.

Algorithm 2 Implementation specifics of the overall discrimination method.

START:

Input: **A** (the given image)

For $i = 1:5$;

Achieving cropped image **A(i)**;

Initial blur probability **P(i)** is obtained by inputting **A(i)** into the network (Fig. 3);

Average blur probability **Pm** = mean(**P(i)**);

End For

Noise parameter σ is obtained by Algorithm 1;

Final discrimination result is obtained by inputting **Pm** and σ into the fusion network.

FINISH

4 Experimental Results

4.1 Experimental Dataset and Training Results

The experimental environment of this paper is Windows 10 version 64-bit operating system, including Intel Core i5 2.5 GHz, Memory 16 GB, NVIDIA GTX1650Ti, CUDA version 10.1 and CUDNN version 7.6.

The training dataset and the testing dataset are, respectively, divided into two categories: blurry images and sharp images. To ensure the diversity of samples and the robustness of the model, we build our own blurry image dataset with multiple parameters and multiple types. The sharp images are, respectively, blurred with different Gaussian parameters, motion blur parameters, and different noises. Then the blurry image dataset is generated with a total of more than 290 different blur types. These parameters of blur are listed in Tables 3 and 4. Moreover, some blurry images are download from internet or taken with mobile phone to enrich the dataset. The samples of sharp images and blurry images are shown in Figs. 6 and 7, respectively.

In the training process, learning rate is set as 0.0001; max iteration is set as 13000; and batch size is set as 20. The accuracy curve and loss curve have been given in Figs. 8 and 9, which shows our method can achieve superior training results.

Table 3 Gaussian blur parameters.

| Types | Blur radius (r) | | | | | | | |
|---------------|---------------------|---|---|---|---|---|---|----|
| Gaussian blur | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 10 |

Table 4 Motion blur parameters.

| Blur length | Blur angle | | | | | |
|-------------|------------|-----|-----|-----|-----|-----|
| 5 | 30 | 45 | 60 | 75 | 90 | 105 |
| 7 | 15 | 45 | 75 | 105 | 135 | 165 |
| 9 | 15 | 40 | 65 | 90 | 115 | 140 |
| 11 | 10 | 35 | 60 | 85 | 110 | 135 |
| 13 | 0 | 45 | 70 | 100 | 130 | 150 |
| 15 | 90 | 120 | 135 | 150 | 165 | 180 |
| 18 | 0 | 30 | 45 | 60 | 90 | 135 |
| 25 | 0 | 45 | 60 | 90 | 135 | 160 |



Fig. 6 The samples of sharp images from our datasets.

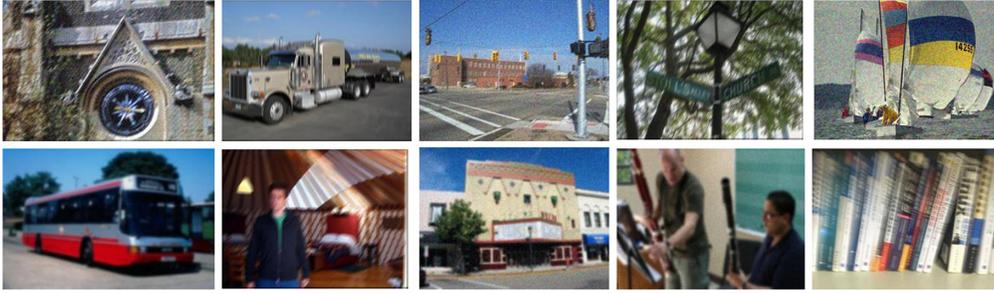


Fig. 7 The samples of blurry images from our datasets.

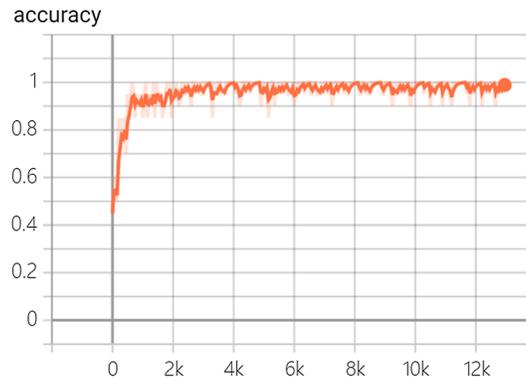


Fig. 8 Training results: the accuracy curve.

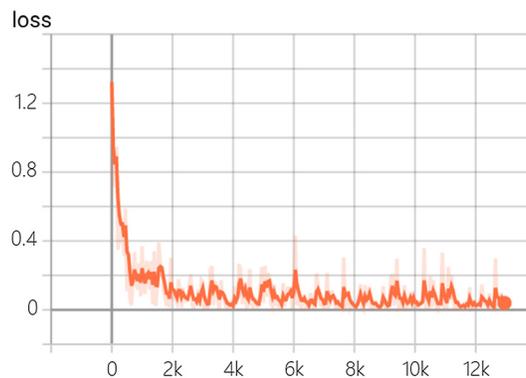


Fig. 9 Training results: the loss curve.

4.2 Performance Evaluation

In this paper, we evaluate the proposed method using four performance indices,³³ including precision, recall, *F1*-score, and accuracy. First, precision and recall are as follows:

$$\text{Precision} = \frac{T_p}{T_p + F_p}, \quad (3)$$

$$\text{Recall} = \frac{T_p}{T_p + F_N}, \quad (4)$$

where *T* and *F* are true and false, respectively, which means that the result is correct or not; *P* and *N* are positive and negative, respectively, which means that the result is considered to be “positive class” or “negative class.” T_p , referred to as “true positive,” represents the number of

instances that actually positive class is predicted into positive class. F_p , referred to as “false positive,” represents the number of instances that actually negative class is predicted into positive class. F_N , referred to as “false negatives,” represents the number of instances that actually positive class is categorized into negative class.

When evaluating the results, we hope that both precision and recall are high, but if precision increases, recall often decreases in most cases. In fact, they are contradictory, and we simultaneously use a new indicator, which takes into account both precision and recall to achieve the high balance, which is shown as follows:

$$F1 = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (5)$$

In addition, accuracy refers to how closely a measurement or observation comes to “true value.” Therefore, we use accuracy to visually observe the correct ratio of the proposed network and compare with other approaches. It is defined as follows:

$$\text{Accuracy} = \frac{N_{\text{correct}}}{N_{\text{total}}}, \quad (6)$$

where N_{correct} means the correct results and N_{total} denotes the total samples.

4.3 Experimental Results on Our Testing Dataset

We collected totally 300 blurry images with different blur and 300 sharp images to form our testing dataset. The sharp images and the blurry images are, respectively, regarded as the positive class to calculate the corresponding precision, recall, $F1$ -score, and accuracy. As listed in Table 5, we can clearly conclude that our method is better than the original VGG16 neural network in all evaluation indicators. Compared with the original VGG16 network, the improved VGG16 network with single sub-image can increase the accuracy by 15% (0.723 to 0.873). Moreover, the improved VGG16 network with average of five sub-images can increase the accuracy to 0.937, which has shown that the average of discriminative values can indeed reduce the random error and improve discrimination results. Finally, combining the texture noise extraction algorithm, we can achieve the satisfactory discriminative results (the accuracy is 0.980). In addition, we compared the accuracy with other discrimination approaches, as shown in Table 6.

For the methods based on hand-extracted features, the maximum accuracy of Liu’s method about blur/sharp discrimination is 75.2% when $\eta_a = 0.4$. Xu et al.’s method has an accuracy rate of 86.6% for blur/sharp discrimination. Teo and Zhan’s method proposed to use b_{motion} and post-derived features to discriminate whether the image is blurry or not, and its accuracy is 74.5%. The accuracy of Rugna and Konik’s pixel-based method for blur identification is about 90.3%. For the methods based on deep learning, Huang et al. used a CNN for blur and sharp discrimination with an accuracy rate of 75.7%.

Table 5 Comparison of test results.

| Evaluation parameters | Improved VGG16 network | | | | | | | |
|-----------------------|------------------------|-------|-----------------------------------|-------|-------|-------|--|-------|
| | Original VGG16 network | | Single sub-image (random cropped) | | | | Fusion of texture noise and improved VGG16 | |
| | Blur | Sharp | Blur | Sharp | Blur | Sharp | Blur | Sharp |
| Precision | 1.00 | 0.64 | 0.92 | 0.84 | 0.97 | 0.91 | 0.99 | 0.97 |
| Recall | 0.45 | 0.99 | 0.82 | 0.93 | 0.90 | 0.97 | 0.97 | 0.99 |
| $F1$ -score | 0.62 | 0.78 | 0.87 | 0.88 | 0.93 | 0.94 | 0.98 | 0.98 |
| Accuracy | 0.723 | | 0.873 | | 0.937 | | 0.980 | |

Table 6 Comparison with other approaches.

| Method | Teo and Zhan ¹⁵ | | Liu et al. ¹⁴ | | Xu et al. ¹¹ | | Rugna and Konik ¹³ | | Huang et al. ¹⁸ | | Ours | |
|-----------|----------------------------|-------|--------------------------|-------|-------------------------|-------|-------------------------------|-------|----------------------------|-------|-------|-------|
| | Blur | Sharp | Blur | Sharp | Blur | Sharp | Blur | Sharp | Blur | Sharp | Blur | Sharp |
| Precision | 0.73 | 0.76 | 0.74 | 0.76 | 0.85 | 0.88 | 0.92 | 0.85 | 0.79 | 0.73 | 0.99 | 0.97 |
| Recall | 0.78 | 0.71 | 0.77 | 0.73 | 0.89 | 0.84 | 0.88 | 0.89 | 0.70 | 0.81 | 0.97 | 0.99 |
| F1-score | 0.75 | 0.73 | 0.75 | 0.74 | 0.87 | 0.86 | 0.90 | 0.87 | 0.74 | 0.77 | 0.98 | 0.98 |
| Accuracy | 0.745 | | 0.752 | | 0.866 | | 0.903 | | 0.757 | | 0.980 | |

As a result, the accuracy of our approach is better than the existing approaches regardless of any blur type. The main reason is that the traditional hand-extracted methods are severely limited. The extracted prior knowledge cannot express the blur property of image very well. For example, Liu's method²² used local auto-correlation congruence, gradient histogram span, spectrum slope, and maximum saturation to achieve the discrimination results. They proposed that a blurry image usually had a large spectrum slope while a sharp image, contrarily, corresponded to a small spectrum slope. However, the spectrum slope was proposed based on the same image scene. For different scene, a blurry image may have a small spectrum slope, so these hand-extracted features do not fit all blurry images. The same example for the dark channel is discussed in Sec. 1 and Fig. 1. On the other hand, the deep learning methods only depend on pure data driven without considering guide of prior knowledge (noise effect). However, noise is random and easy to change the pixel distribution of the image, which will decrease discrimination accuracy.

By contrast, we make full use of the strong classification ability of deep learning network and then introduce texture noise. Moreover, by computing the average discrimination result of several cropped sub-images, the influence of the local texture features can be overcome. Therefore, we can achieve the satisfactory discrimination result.

4.4 Experimental Results on Other Testing Datasets

To further demonstrate the discrimination performance and generalization ability, we test our method on different and popular datasets for image discrimination, image deblurring, and image quality evaluation, including RealBlur, BSD-B, and GoPro. These datasets contain a large number of blurry images and sharp images. RealBlur dataset is a large-scale dataset of real-world blurry images, which is generated by Rim et al. In Ref. 34, BSD-B dataset is a synthetic dataset, which is generated from the BSD500 segmentation dataset.^{35,36} GoPro dataset is also a synthetic dataset generated in Ref. 36. They all contain a large number of blurry images and sharp images, which are used for training in image deblurring and image quality assessment. We randomly select 300 pairs of blurry and sharp images from each dataset. The test results on three datasets are shown in Table 7.

Table 7 Comparison of test results on other testing datasets.

| Evaluation parameters | Dataset | | | | | |
|-----------------------|----------|-------|-------|-------|-------|-------|
| | RealBlur | | BSD-B | | GoPro | |
| | Blur | Sharp | Blur | Sharp | Blur | Sharp |
| Precision | 0.951 | 0.973 | 0.942 | 0.979 | 0.906 | 1.000 |
| Recall | 0.973 | 0.950 | 0.980 | 0.940 | 1.000 | 0.897 |
| F1-score | 0.962 | 0.961 | 0.961 | 0.959 | 0.951 | 0.946 |
| Accuracy | 0.962 | | 0.960 | | 0.948 | |



Fig. 10 Sharp images in GoPro dataset look blurry in terms of subjective vision: (a) derives from the directory of `GoPro\train\GOPR0385_11_00\sharp\000128.png`; (b) derives from the directory of `train\GOPR0380_11_00\sharp\000157.png`; (c) derives from the directory of `\test\GOPR0384_11_05\sharp\004008.png`; and (d) derives from the directory of `\train\GOPR0868_11_01\sharp\000262.png`.

It is found that accuracy of these datasets is all greater than 94%, which again demonstrates our method has superior generalization ability and satisfactory robustness. Considering blurry images and sharp images separately, all precision and recall are greater than 94% except GoPro dataset. Precision of blurry images and recall of sharp images in GoPro dataset are both about 0.90. It illustrates that a certain number of sharp images are predicted to be blurry. This is mainly because that some testing images look blurry in terms of subjective vision, but these images are labeled as sharp images in GoPro dataset. The actual image quality is contradictory to the label in GoPro dataset. Some samples and the original directory in GoPro dataset are shown in Fig. 10.

5 Conclusions and Future Work

In this paper, a method is proposed for blur discrimination based on a fusion network. First, the VGG16 network is improved to achieve blur probability. Then the texture noise parameters can be extracted by the proposed noise extraction algorithm. Finally, the fusion network integrates the blur probability and noise parameters to achieve superior discrimination results. Actually, the proposed method combines data driven with guide of prior knowledge and make deep learning effective. Extensive experiments performed on own dataset and other popular blurring datasets with a number of blurry images and sharp images, including RealBlur dataset, BSD-B dataset, and GoPro dataset. We use four evaluation indices to evaluate the proposed method and achieve satisfactory discrimination results. The experiment demonstrates that the proposed method can obtain superior performance and be applied to many applications.

The limitation of this work is that the parameters of additional dropout layers are achieved by a lot of trials. Actually, these parameters can be determined by image texture features. Moreover, the method can only discriminate whether the image is blurry or not, but whether it is Gaussian blur, motion blur, or other blur types cannot be discriminated. They will be discussed and studied in future work.

Acknowledgments

We would like to thank the National Natural Science Foundation of China under Grant No. 61901059 and Hubei Provincial Excellent Young and Middle-Aged Scientific and Technological Innovation Team Project in Colleges and Universities under Grant No. T2020007.

References

1. Y. Chen et al., "Research of improving semantic image segmentation based on a feature fusion model," *J. Ambient Intell. Hum. Comput.* (2020).
2. L. Zhou et al., "Fraction-order total variation image blind restoration based on self-similarity features," *IEEE Access* **8**, 30346–30444 (2020).
3. Y. Chen et al., "The improved image inpainting algorithm via encoder and similarity constraint," *Vis. Comput.* (2020).
4. S. Qiu et al., "The infrared moving target extraction and fast video reconstruction algorithm," *Infrared Phys. Technol.* **97**, 85–92 (2019).

5. L. Xu, S. Zheng, and J. Jia, "Unnatural L0 sparse representation for natural image deblurring," in *IEEE Conf. Comput. Vision and Pattern Recognit.* (2013).
6. Z. Zha et al., "Image restoration via simultaneous nonlocal self-similarity priors," *IEEE Trans. Image Process.* **29**(8), 8561–8576 (2020).
7. T. Michaeli and M. Irani, "Blind deblurring using internal patch recurrence," *Lect. Notes Comput. Sci.* **8691**, 783–798 (2014).
8. J. Pan et al., "Blind image deblurring using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**(10), 2315–2328 (2018).
9. Y. Bai et al., "Single-image blind deblurring using multi-scale latent structure prior," *IEEE Trans. Circuits Syst. Video Technol.* **30**(7), 2033–2045 (2020).
10. J. Shi, L. Xu, and J. Jia, "Discriminative blur detection features," in *IEEE Conf. Comput. Vision and Pattern Recognit.* (2014).
11. W. Xu, J. Mulligan, and D. Xu, "Detecting and classifying blurred image regions," in *IEEE Int. Conf. Multimedia and Expo* (2013).
12. M. A. Khan, S. A. Irtaza, and A. Khan, "Detection of blur and non-blur regions using frequency-based multi-level fusion transformation and classification via KNN matting," in *Int. Conf. Math., Actuarial Sci., Comput. Sci. and Stat.* (2019).
13. J. D. Rugna and H. Konik, "Blur identification in image processing," in *IEEE Int. Joint Conf. Neural Network Proc.* (2006).
14. R. Liu, Z. Li, and J. Jia, "Image partial blur detection and classification," in *IEEE Conf. Comput. Vision and Pattern Recognit.* (2008).
15. T. Teo and K. Zhan, "Integration of image-derived and pos-derived features for image blur detection," *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **XLI-B1**, 1051–1055 (2016).
16. J. Wang et al., "Blurring detection based on selective features for iris recognition," in *Int. Conf. Rob. and Rehab. Intell.*, pp. 208–223 (2020).
17. R. Gueraichi and A. Serir, "Blurred image detection in drone embedded system," in *5th Int. Conf. Adv. Technol. Signal and Image Process.* (2020).
18. R. Huang et al., "Multiscale blur detection by learning discriminative deep features," *Neurocomputing* **285**, 154–166 (2018).
19. W. Zhao et al., "Defocus blur detection via multi-stream bottom-top-bottom fully convolutional network," in *CVPR* (2018).
20. X. Wang et al., "Accurate and fast blur detection using a pyramid M-shaped deep neural network," *IEEE Access* **7**, 86611–86624 (2019).
21. K. Zeng et al., "A local metric for defocus blur detection based on CNN feature learning," *IEEE Trans. Image Process.* **28**(5), 2107–2115 (2019).
22. T. Szandała, "Convolutional neural network for blur images detection as an alternative for Laplacian method," in *IEEE Symp. Ser. Comput. Intell.* (2020).
23. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition" (2014).
24. G. E. Hinton et al., "Improving neural networks by preventing co-adaptation of feature detectors," *Comput. Sci.* **3**(4), 212–223 (2012).
25. G. C. Temes and J. Silva, "Simple and efficient noise estimation algorithm," *Electron. Lett.* **40**(11), 640–642 (2004).
26. B. Kumar, "Mean-median based noise estimation method using spectral subtraction for speech enhancement technique," *Indian J. Sci. Technol.* **9**, 1–6 (2016).
27. V. M. Kamble, M. R. Parate, and K. M. Bhurchandi, "No reference noise estimation in digital images using random conditional selection and sampling theory," *Vis. Comput.* **35**(1), 5–21 (2019).
28. D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika* **81**(3), 425–455 (1994).
29. M. Jansen, *Noise Reduction by Wavelet Thresholding*, Springer Press (2001).
30. C. Jung et al., "Novel input and output mapping-sensitive error back propagation learning algorithm for detecting small input feature variations," *Neural Comput. Appl.* **21**, 705–713 (2012).

31. Y. Pu and J. Wang, "Fractional-order global optimal back-propagation machine trained by an improved fractional-order steepest descent method," *Front. Inf. Technol. Electron. Eng.* **21**, 809–833 (2020).
32. L. Ma and G. Sofronov, "Change-point detection in autoregressive processes via the cross-entropy method," *Algorithm* **13**(5), 128 (2020).
33. R. Tripathi, S. Jagannathan, and B. Dhamodharaswamy, "Estimating precisions for multiple binary classifiers under limited samples" (2021).
34. J. Rim et al., "Real-world blur dataset for learning and benchmarking deblurring algorithms," *Lect. Notes Comput. Sci.* **12370**, 184–201 (2020).
35. D. Martin et al., "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. Eighth IEEE Int. Conf. Comput. Vision* (2001).
36. S. Nah, T. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Int. Conf. Comput. Vision and Pattern Recognit.* (2017).

Yumeng Tian received her BS degree in electronic information science and technology from Xi'an Technological University, Xi'an, China, in 2018. She is currently pursuing her ME degree in signal and information processing at Yangtze University, Jingzhou, China. Her current research interests include image/video restoration, image classification, and image quality assessment.

Mingzhang Luo received his BS degree in electronic instrument and measurement from Jiangnan Petroleum Institute, Jingzhou, China, in 2001, and his PhD in earth exploration and information technology from Yangtze University, Jingzhou, China, in 2012. He is currently a professor with the School of Electronics and Information, Yangtze University, China. His current research interests include data acquisition and big data analysis.

Luoyu Zhou received his BS degree in optical information science and technology from the University of Science and Technology of China, Hefei, China, in 2008, and his PhD in optical engineering from Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun, China, in 2013. He is currently an associate professor with the School of Electronics and Information, Yangtze University, China. His current research interests include image processing, computer vision, and artificial intelligence.