

Multiple Mixed Modes: Single-Channel Blind Image Separation

Tiantian Yin^{1,2}, Yina Guo^{1,3,*}, and Ningning Zhang²

Abstract

As one of the pivotal techniques of image restoration, single-channel blind source separation (SCBSS) is capable of converting a visual-only image into multi-source images. However, image degradation often results from multiple mixing methods. Therefore, this paper introduces an innovative SCBSS algorithm to effectively separate source images from a composite image in various mixed modes. The cornerstone of this approach is a novel triple generative adversarial network (TriGAN), designed based on dual learning principles. The TriGAN redefines the discriminator's function to optimize the separation process. Extensive experiments have demonstrated the algorithm's capability to distinctly separate source images from a composite image in diverse mixed modes and to facilitate effective image restoration. The effectiveness of the proposed method is quantitatively supported by achieving an average peak signal-to-noise ratio exceeding 30 dB, and the average structural similarity index surpassing 0.95 across multiple datasets.

Keywords

Discriminator, Image Restoration, SCBSS, TriGAN

1. Introduction

Single-channel blind source separation (SCBSS) is a technique primarily used to isolate source signals from a single-channel mixed signal [1], and the mixing methods of mixed signals mainly include linear and convolution mixing, which play a vital role in denoising and restoration processes across various fields. These fields include medical research [2], image processing [3], speech processing [4], video processing [5], traffic signals [6] and other fields.

Historically, traditional SCBSS algorithms have primarily addressed the linear hybrid mode, such as nonnegative matrix factorization (NMF) [7] and independent component correlation algorithm (ICA) [8]. These methods grounded in the principles of linear equations to separate the source signal, are proficient in recovering the source signal from linearly mixed signals. However, they fall short in handling recovering the source signal from convolutionally mixed signals, which present a more complex challenge compared to linear mixing. To address this gap, researchers have further proposed a regression-based method [9]. This method learns the complex mapping relationship between mixed signal and source signal through the robust learning capabilities of deep neural networks. Despite their effectiveness, a

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received June 16, 2022; first revision January 5, 2023; accepted January 21, 2023.

* Corresponding Author: Yina Guo (zulibest@163.com)

¹ Electronics and Communication Engineering, Taiyuan University of Science and Technology, Taiyuan, China (zulibest@163.com, 1319606385@qq.com)

² Upgrading Office of Modern College of Humanities and Sciences of Shanxi Normal University, Linfen, China (yinttd@163.com, Z1287101973@163.com)

³ Guangdong Shantou University, Guangdong, China (zulibest@163.com)

significant limitation of these methods is their dependency on a predefined convolution mixing matrix. Consequently, any alterations in the mixing matrix render a trained model incapable of separating the new test set., highlighting a critical area for further research and development in SCBSS.

The SCBSS algorithms mainly include: the blind deconvolution algorithm proposed by Fan et al. [10] and Stoller et al. [11] was used to recover the source signal. Building upon this, Lin and Gao [12] introduced a blind source separation (BBS) algorithm combined with a high-order spectrum. While this algorithm exhibits some capability in separating mixed convolutional signals, it is marked by high computational demands and inefficient. Meanwhile, the application of convolutional neural networks [13] and fully connected neural networks [14] has been explored for separating mixed source signals and facilitating blind deconvolution. In terms of audio, recurrent neural networks [15] have shown proficiency in separating speech from mixed noise. In addition, automatic encoders [16] have been employed for supervised separation of source signals. In 2014, after the generative adversarial network (GAN) was proposed [17], Subakan and Smaragdis [18] presented a GAN-based SCBSS algorithm in the audio field. However, this method requires prior knowledge of the mixing matrix type and assumes that the mixing matrix and the source signal share the same distribution for training. Addressing the challenge of an unknown mixing matrix, Kong et al. [19] proposed a synthesis decomposition (S-D) algorithm utilizing deep convolutional generic adversarial networks (DCGAN). This approach, which does not require prior knowledge of the convolutional mixing matrix, has achieved notable success. However, this method is only for the source image separation of convolution mixing images. Aiming at the separation of multi-channel source signals, Liu et al. [20] estimated the source signal and mixing matrix through a reconstruction approach based on the minimum error of the observation signal and Bayesian maximum a posteriori estimation method.

In the actual process of composite image restoration, source images are often mixed through various mixing methods. To solve this problem, a closed-loop triple generative adversarial network (TriGAN) structure is constructed in this paper grounded in the dual learning concept, which learns the mapping relationship between the composite image and the source image, and breaks through the internal mathematical model disparities of source image separation caused by different mixing methods. The discriminator continuously updates the generator with feedback information until the generator reaches the optimal solution. Unlike previous models, TriGAN's discriminator calculates the loss at the granular level of 1×1 pixel block, utilizing the least square method to separate the difference between the image and the pixel block of the source image. In this way, the SCBSS of different mixed images is realized, thereby enhancing source image restoration. The rest of the paper is organized as follows. Section 2 presents mathematical models of the two main hybrid approaches in SCBSS. Section 3 elucidates the functioning of the TriGAN discriminator and outlines its training procedure. To demonstrate the effectiveness of the position of the discriminator proposed in this paper, Section 4 shows the efficiency of the discriminator under various conditions. Additionally, the experimental results in the second part of Section 4 reveal the effectiveness of TriGAN. Section 5 gives the conclusion.

2. Mathematical Model of SCBSS

Through extensive research efforts over the years, SCBSS models can be classified into two types: the linear mixing model and the convolution mixing model. The linear mixing model is widely recognized

as the more common and mature method within the SCBSS domain. Notably, other complex mixing methods can be transformed into this model through mathematical transformation. The focus of the algorithm presented in this paper is on scenarios involving two source images, denoted as $S = \{S_1, S_2\}$. These source images S_1 and S_2 undergo a linear addition process to form a linearly mixed image X . The mathematical model representing this linear combination is expressed as follows:

$$X(t) = A_1 S_1(t) + A_2 S_2(t), \quad (1)$$

where the mixture matrix is represented by A .

Contrasting with linear mixing, convolution mixing presents a more complex scenario in SCBSS. The key distinction lies in its mixing approach: rather than a straightforward linear relationship, convolution mixing involves a matrix convolution operation. In situations where the mixed image is the sole observational image X , the convolution mixed SCBSS mathematical model can be expressed as:

$$X(t) = (\alpha_1 * S_1)(t) + (\alpha_2 * S_2)(t). \quad (2)$$

The symbol "*" indicates convolution operation:

$$X(t) = \int_{R^d} [\alpha_1(t-v)S_1(v) + \alpha_2(t-v)S_2(v)]dv, \quad (3)$$

where R^d is the Euclidean space and α denotes the convolution mixing matrix. When only the observed image X is available, the SCBSS algorithm normally solves for the remaining unknowns by assuming one of the unknowns in mathematical solutions for the mixing matrix A , the convolutional mixing matrix α and the source image S . This process often presupposes knowledge of the mixing matrix type, utilizing it to resolve for the source image S .

The solution process varies with the type of the mixing matrix. Consequently, this variability poses a challenge for SCBSS algorithms in addressing the separation of source images under multiple hybrid methods. This paper endeavors to transcend these mathematical model constraints of SCBSS. By leveraging the inherent feature information of the blended image, the study aims to surmount the variations inherent in mixing methods. This approach seeks to accomplish comprehensive, thereby enhancing the versatility of the SCBSS algorithm in the realm of image restoration.

3. Proposed TriGAN

Building upon the foundation of the dual learning generative adversarial network (DualGAN), TriGAN consists of three GANs, which involves three image domains instead of dealing with the direct transformation problem of two image domains, diverging from the traditional approach of handling the direct transformation between two image domains. Instead, TriGAN employs a cyclic network structure to facilitate the learning of mapping relationships across these domains: from the visible image domain X to source image domain S_1 , source image domain S_1 to source image domain S_2 , and source image domain S_2 to visible image domain domain X .

The generators of TriGAN retain the structural essence of the original GAN, but introduce a redefined operational mode for the discriminators. The generators, denoted as $G_{X \rightarrow S_1}$, $G_{S_1 \rightarrow S_2}$, and $G_{S_2 \rightarrow X}$ are

mirrored by their corresponding discriminators $D_{X \rightarrow S_1}$, $D_{S_1 \rightarrow S_2}$, and $D_{S_2 \rightarrow X}$. Among them, the network architecture of the three generators is consistent, with each generator having the identical number of downsampling and upsampling layers. Additionally, mirror downsampling and upsampling layers are integrated between the generators. This inclusion, along with the implementation of jump connections, form a U-shaped structure that facilitates the sharing of low-level information between the input and generated image, thus enabling rapid convergence of the generators. The structure of the three discriminators follows the same pattern, except that, instead of using the DualGAN discriminators, a new discriminator is distributed over each pixel block of the whole image, creating a loss function between the blocks. This captures high frequency features more effectively on a pixel by pixel basis, making full use of the texture, color and style information inherent in the visible image.

In order to better and effectively learn the mapping relationship between two image domains for the deep learning network, the discriminator $D_{X \rightarrow S_1}$, $D_{S_1 \rightarrow S_2}$, and $D_{S_2 \rightarrow X}$ adopts the core principle of the least square method to calculate the error between each pixel of the generated image F and the real image R . The total error for each pixel block being calculated during the training process as follows:

$$\sum (R - F)^2. \quad (4)$$

Real image $R \in \{X, S_1, S_2\}$, F is generated by three generators corresponding to it.

The initial Gaussian distribution generates images randomly during the generation, so the errors are incurred randomly and these fluctuate up and down around the true value. At the point where the total error is small, it gets closer to the true value. The minimal total error is obtained when the derivative of R equals 0.

$$\frac{\partial \sum (R-F)^2}{\partial R} = 2 \sum (R - F)^2 = 0. \quad (5)$$

Then TriGAN is using this concept as a loss function, replacing the loss function of DualGAN, and the objective function of TriGAN can be expressed as follows:

$$\min_D V_{\text{TriGAN}}(D) = \frac{1}{2} E_{S \sim P_{\text{data}}(R)} [(D(R) - \epsilon)^2] + \frac{1}{2} E_{F \sim P_F(F)} [(D(G(F)) - \epsilon)^2]. \quad (6)$$

In the discriminator's objective function, the real data and the generated data are encoded, respectively. The discriminator $D_{X \rightarrow S_1}$, $D_{S_1 \rightarrow S_2}$, $D_{S_2 \rightarrow X}$ calculates the loss in pixels, and after its objective function reaches an optimal value, the generator is fine-tuned to create images that are increasingly akin to the domains X , S_1 , S_2 . The loss function for the new discriminator captures the distance of the image from the decision boundary, whereas allowing the more distant data to receive a penalty term in proportion to the distance, Therefore, for the discriminator's gradient to converge to zero, the generator image must closely approximate the real image's position. By replacing DualGAN's loss function with this new method, TriGAN mitigates instability issues. The training process in TriGAN begins with fixing the generator and then focuses on training the discriminator:

$$\begin{aligned} \min_D V_{\text{TriGAN}}(D) &= \frac{1}{2} E_{S \sim P_{\text{data}}(R)} [(D(R) - \epsilon)^2] + \frac{1}{2} E_{F \sim P_F(F)} [(D(G(F)) - \epsilon)^2] \\ &= \frac{1}{2} \int_S P_{\text{data}}(R) [D(R) - \epsilon]^2 dS + \frac{1}{2} \int_F P_F(F) [D(G(F)) - \epsilon]^2 dF. \end{aligned} \quad (7)$$

The calculated optimal solution of the discriminator is:

$$D^*(R) = \frac{\epsilon P_{data(R)} + \epsilon P_F(F)}{P_{data(R)} + P_F(F)}. \quad (8)$$

Once the discriminator attains its optimal state, the discriminator is fixed, and the generator is trained until its objective function also reaches an optimal solution.

The training procedure for TriGAN is summarized in Algorithm 1.

Algorithm 1. TriGAN training procedure

Require: The mixed image domain X , source image domain S_1 and source image domain S_2 are established.

Generator: $G_{X \rightarrow S_1}$, $G_{S_1 \rightarrow S_2}$, $G_{S_2 \rightarrow X}$. Discriminator: $D_{X \rightarrow S_1}$, $D_{S_1 \rightarrow S_2}$, $D_{S_2 \rightarrow X}$. Number of iterations: n_{critic} , Batch size = 120.

1. Randomly initialize the generator and discriminator.
 2. m groups of image samples $x^{(1)}, \dots, x^{(m)}, s_1^{(1)}, \dots, s_1^{(m)}, s_2^{(1)}, \dots, s_2^{(m)}$ are extracted in pixels from the real data space of image domain X , image domain S_1 and image domain S_2 .
 3. **Repeat**
 4. **for** $n = 1, \dots, n_{critic}$ **do**
 5. image sample: $\{x^{(m)}\}_{m=1}^{120} \subseteq X$, $\{s_1^{(m)}\}_{m=1}^{120} \subseteq S_1$, $\{s_2^{(m)}\}_{m=1}^{120} \subseteq S_2$
 6. fix the generator and update the discriminator
 7. the objective function of the discriminator reaches its optimal solution
 8. image sample: $\{x^{(m)}\}_{m=1}^{120} \subseteq X$, $\{s_1^{(m)}\}_{m=1}^{120} \subseteq S_1$, $\{s_2^{(m)}\}_{m=1}^{120} \subseteq S_2$
 9. fix the discriminator and update the generator
 10. the objective function of the generator reaches its optimal solution
 11. **end for**
 12. Continue the training process until convergence is achieved.
-

4. Experimental Results

In this paper, the MNIST dataset, ancient Chinese character image dataset [21] and RESIDE dataset were employed to ascertain the efficacy of TriGAN model. In addition, each experiment was repeated 50 times, with the average of the training results being considered for analysis.

4.1 Proposed Discriminator Works

To verify the efficiency of which the discriminator of TriGAN calculates the error between the generated image F and the real image R at a granular level of 1×1 pixels experiments were conducted based on the MNIST dataset. The normalized image size was 28×28 pixels. In this paper, the image samples were divided into six cases of 1×1 , 2×2 , 4×4 , 7×7 , 16×16 , and 28×28 pixels, respectively in order to compare the correlation between their separated images and source images and to evaluate the work efficiency of TriGAN under the condition of different image sample sizes. The correlation between S and S' is:

$$R_{SS'}(\theta) = \sum_{t=-\infty}^{\infty} S(t)S'(t + \theta) \quad (9)$$

Table 1 shows the correlation between S and S' of 1×1 , 2×2 , 4×4 , 7×7 , 16×16 , and 28×28 pixels. A 1×1 approach yields a mean correlation of 0.9072. The 2×2 , 4×4 , 7×7 , 16×16 , and 28×28 pixels approach achieves lower and lower correlation means of 0.8970, 0.7769, 0.5659, 0.1895 and 0.1180, respectively.

Table 1. Correlation between source signal S_1, S_2 and separated signal S'_1, S'_2 under different image sample division units

Unit	Correlation between S_1 and S'_1	Correlation between S_2 and S'_2	Correlation mean
1×1	0.9138	0.9007	0.9072
2×2	0.8987	0.8954	0.8970
4×4	0.7714	0.7824	0.7769
7×7	0.5506	0.5812	0.5659
16×16	0.1967	0.1824	0.1895
28×28	0.1160	0.1200	0.1180

Bold indicates the best results.

Stronger correlation indicates higher degree of image similarity. Table 1 illustrates that the separation image S' has the highest correlation with the source image S when the method in this paper divides the image samples to calculate the loss in 1×1 pixels. The correlation between the separation image S' and the source image S is similar when the segmentation units of the image samples are of the same size. As the segmentation unit size increases, the correlation decreases. This suggests that the pixel-level loss calculation employed by the TriGAN discriminator is more effective than the traditional GAN's global loss calculation and DualGAN's texture-based loss calculation.

In addition to the change of the working unit of the TriGAN discriminator, the loss function of the TriGAN discriminator has also changed. In order to verify that the new loss function is effective, an experiment was conducted replacing the original loss function of DualGAN with the new loss function. The incorporation of this new loss function into DualGAN resulted in a significant improvement in the restoration of synthetic images. In this section, the proposed loss function is applied in DualGAN to solve image-to-image translation. We carry out experiments on RESIDE dataset as starting research for this challenging problem and show the effectiveness of the proposed new discriminator. The dataset, known as real single image defogging (RESIDE), is a large-scale resource designed to enable fair evaluation and comparison of single image defogging algorithms. The experimental results are shown in Fig. 1.

Fig. 1 shows RESIDE dataset for image-to-image separate experiment (campsite): I_{com} (1) (2) (3) (4) the composite image, I_{sou} (1) (2) (3) (4) the source image, I_{DGs} (1) (2) (3) (4) DualGAN separate image [22], I_{NDs} (1) (2) (3) (4) this paper proposed new discriminator works globally separate image, I_{NDs} 1×1 (1) (2) (3) (4) this paper proposed new discriminator works 1×1 pixel unit separate image.

The first and second column of Fig. 1 demonstrate the composite and source image from the RESIDE dataset. The third column shows the separate image using DualGAN which has been notably successful in image-to-image for two image domains. The fourth column indicates that the new discriminator proposed in this paper works alone on a global scale. Fig. 1 indicate that this new discriminator generates images more closely resembling the source images compared to the original DualGAN.

The new discriminator addresses the issues of poor image generation quality and unstable training. The

traditional loss function in DualGAN does not allow the generator to continue to generate images that the discriminator discriminates as real images, especially when these images are still notably different from the real image. Utilizing the least squares method, the new approach calculates the distance of the image from the decision boundary, assigning a penalty term proportional to the distance to more distant data. This method ensures that the discriminator’s gradient approaches zero, compelling the generator to produce images that more closely align with the real image’s location, as shown in Fig. 2.

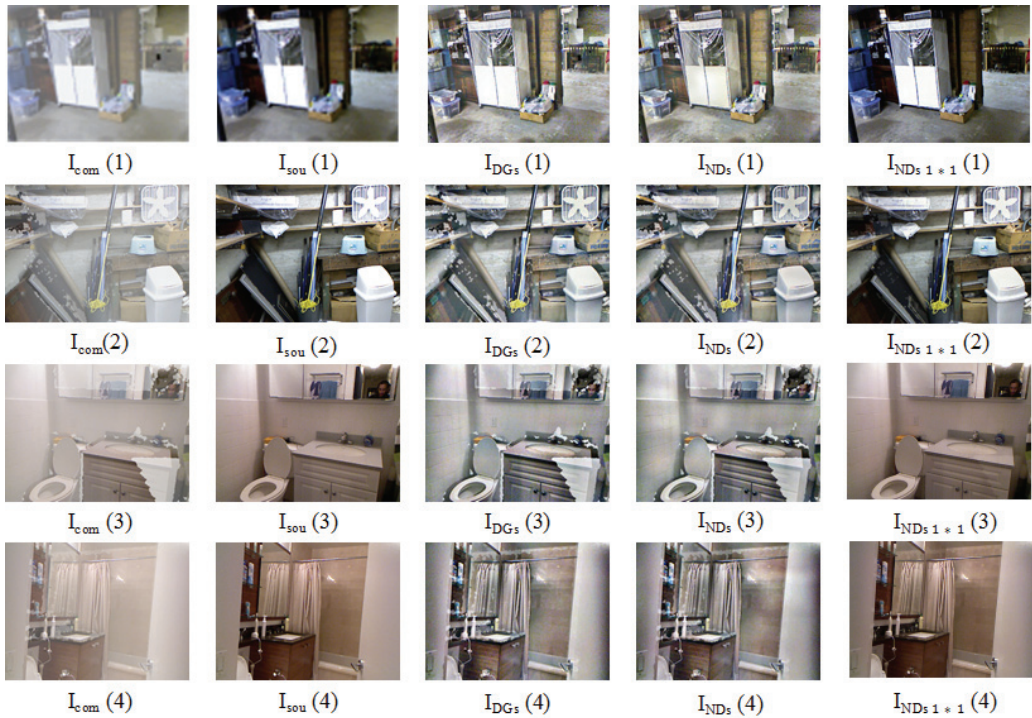


Fig. 1. Restoration result graph of composite image.

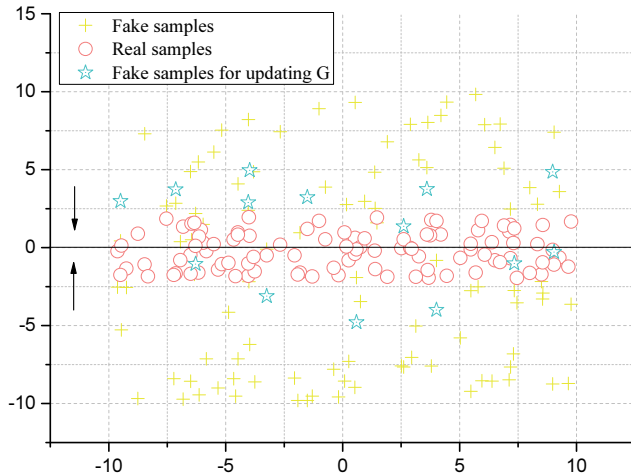


Fig. 2. Proximity of fake samples to real samples and the loss function decision boundary.

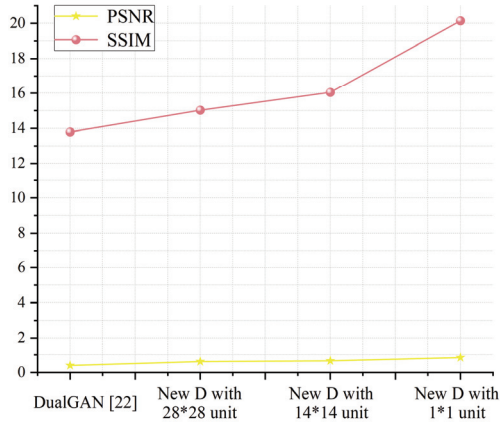


Fig. 3. PSNR and SSIM result graph.

Fig. 3 shows 500 synthetic images based on the RESIDE dataset, in which the image-to-image separation experiment was completed. A total of four methods were compared.

1) DualGAN [22]: This method uses the standard DualGAN approach for image separation.

2) New D with 28×28 unit: Here, the original DualGAN discriminator is replaced with a new discriminator, operating at a 28×28 pixel unit scale (global operation).

3) New D with 14×14 unit: This approach also involves replacing DualGAN's original discriminator, but with the discriminator's operational unit modified to 14×14 pixels.

4) New D with 1×1 unit: The final method features the new discriminator with an operational unit of 1×1 pixel, as proposed in this paper.

4.2 TriGAN Separates Different Mixed Images

Trigan can separate both the convolutionally mixed image and the linearly mixed image. The separation results are shown in Figs. 4 and 5. Experiments in this paper were conducted on 600 pairs of randomly selected images in the MNIST dataset. The MNIST dataset was mainly composed of handwritten numeral images and their corresponding labels. There are 10 types of images, ranging from 0 to 9, with a total of 10 Arabic numerals. These images were blended into 600 convolutionally blended images and 600 linearly blended images. After applying TriGAN to separate the mixed images, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used to evaluate the degree of deviation of each separated image from its source image, and the average value of 600 groups of experimental PSNR and SSIM were calculated, respectively. I edited for clarity here.

Please review to make sure your intent has been maintained.

Table 2. Based on 600 pairs of convolutionally/linearly mixed images, the average PSNR and average SSIM in different methods

		ICA [16]	S-D [19]	Proposed method
Convolution mixing	PSNR (dB)	8.7	23.3	31.0447
	SSIM	0.38	0.87	0.95
Linear mixing	PSNR (dB)	15.3	26.4	34.8251
	SSIM	0.46	0.92	0.97

Table 3. Based on 600 pairs of loss of ancient Chinese characters images, the average PSNR and average SSIM in different methods

	DCSS [21]	Proposed method
Average PSNR (dB)	29.2	32.6
Average SSIM	90.3	96.7

Fig. 4 shows MNIST dataset convolutionally mixing image experiment (campsite): I_{cm} (1) (2) (3) the convolutionally mixing image, I_{sou1} (1) (2) (3) the source image \mathcal{S}_1 , I_{sep1} (1) (2) (3) the separate image \mathcal{S}'_1 , I_{sou2} (1) (2) (3) the source image \mathcal{S}_2 , I_{sep2} (1) (2) (3) the separate image \mathcal{S}'_2 . Fig. 5 shows MNIST dataset linearly mixing image experiment (campsite): I_{lm} (1) (2) (3) the linearly mixing image, I_{sou1} (1) (2) (3) the source image \mathcal{S}_1 , I_{sep1} (1) (2) (3) the separate image \mathcal{S}'_1 , I_{sou2} (1) (2) (3) the source image \mathcal{S}_2 , I_{sep2} (1) (2) (3) the separate image \mathcal{S}'_2 .

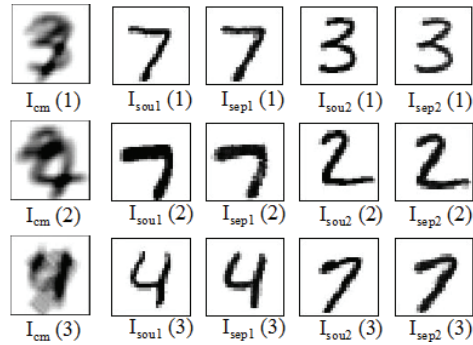
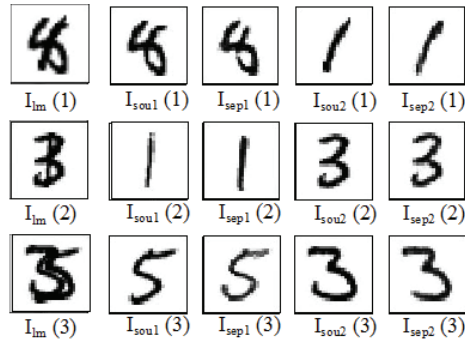
**Fig. 4.** Convolutionally mixed image restoration for MNIST datasets.**Fig. 5.** Linearly mixing image restoration for MNIST datasets.

Image-restoration represents a significant application of SCBSS. Yin et al. [21] presented the restoration of ancient Chinese characters using SCBSS and creating a specialized dataset for ancient Chinese characters. Fig. 6 shows sample datasets, comprising five sets of ancient Chinese character image sets randomly selected from the database, each training set has 4,096 images, 2,048 images of ancient Chinese characters and occlusion respectively, and the test set had 512 images, with an equal split of 256 images each for ancient Chinese characters and occlusion. This paper leverages the ancient Chinese character dataset to restore ancient Chinese character images, differentiating the source from these images and contrasting the results with the single-channel blind deconvolution algorithm based on deep convolution generating adversarial network (DCSS) method proposed by Yin et al. [21].

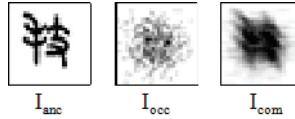


Fig. 6. A sample of the ancient Chinese character datasets: I_{anc} ancient Chinese characters image, I_{occ} occlusion image, I_{com} composite image.

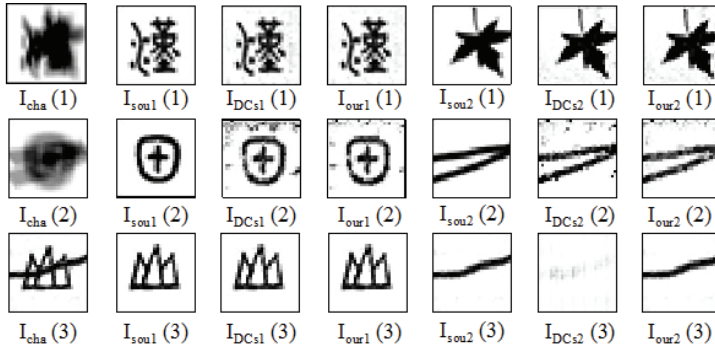


Fig. 7. Restoration results of ancient Chinese character datasets.

Fig. 7 shows the experiment with the ancient Chinese character datasets, illustrating the following: I_{cha} (1) (2) (3) is the ancient Chinese character image; I_{sou1} (1) (2) (3) refers to the ancient Chinese characters source image; I_{DCsl} (1) (2) (3) represents the ancient Chinese characters separate image by DCSS; I_{our1} (1) (2) (3) is the ancient Chinese characters separate image by this paper method; I_{sou2} (1) (2) (3) is the occlusion source image; I_{DCsl2} (1) (2) (3) denotes the occlusion separate image by DCSS and I_{our2} (1) (2) (3) is the occlusion separate image by this paper method.

5. Conclusion

In this paper, a novel TriGAN closed-loop structure is constructed. It is an attempt to surpass the constraints of the SCBSS mathematical model, making full use of the inherent feature information of the blended image and overcoming the inherent variation among ways of mixes and fulfil the separation process of the source image. The experimental results demonstrate the generality of the SCBSS algorithm in image restoration. A new discriminator is used to calculate the pixel loss of the generated image. This methodology enables the separation of source images from a single blended image without prior knowledge of the mixing matrix, a notable breakthrough in the field. The experimental results show that this algorithm is applicable to convolutionally mixed image and linearly mixed image, and outperforms other blind source separation algorithms. In addition, it has yielded exceptional results in the restoration of ancient Chinese characters, significantly improving their restoration effect.

Acknowledgement

This paper is funded by National Natural Science Foundation of China (Grant No. 61301250), China Scholarship Council (Grant No. [2020]1417), Key Research and Development Project of Shanxi Province (Grant No. 201803D421035), Natural Science Foundation for Young Scientists of Shanxi Province

(Grant No. 201901D211313), Shanxi Scholarship Council of China (Grant No. HGKY2019080 and 2020-127), Scientific and Technological Innovation Programs of Higher Education Institutions in Shanxi (STIP) (Grant No. 2022L641), and Shanxi Province Postgraduate Excellent Innovation Project Plan (Grant No. 2021Y679).

References

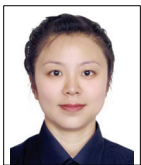
- [1] W. Fu, X. Zhou, and B. Nong, "The research of SCBSS technology: survey and prospect," *Journal of Beijing University of Posts and Telecommunications*, vol. 40, no. 5, pp. 1-11, 2017.
- [2] R. Hephzibah, H. C. Anandharaj, G. Kowsalya, R. Jayanthi, and D. A. Chandy, "Review on deep learning methodologies in medical image restoration and segmentation," *Current Medical Imaging*, vol. 19, no. 8, pp. 844-854, 2023. <https://doi.org/10.2174/1573405618666220407112825>
- [3] G. Liu, W. Tian, Y. Luo, J. Zou, and S. Tang, "A windowed-total-variation regularization constraint model for blind image restoration," *Journal of Information Processing Systems*, vol. 18, no. 1, pp. 48-58, 2022. <https://doi.org/10.3745/JIPS.04.0233>
- [4] J. Han, Y. Shi, Y. Long, and J. Liang, "Exploring single channel speech separation for short-time text-dependent speaker verification," *International Journal of Speech Technology*, vol. 25, pp. 261-268, 2022. <https://doi.org/10.1007/s10772-022-09959-8>
- [5] W. H. Lai and S. L. Wang, "RPCA-DRNN technique for monaural singing voice separation," *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2022, article no. 4, 2022. <https://doi.org/10.1186/s13636-022-00236-9>
- [6] B. Wu, T. Cheng, T. L. Yip, and Y. Wang, "Fuzzy logic based dynamic decision-making system for intelligent navigation strategy within inland traffic separation schemes," *Ocean Engineering*, vol. 197, article no. 106909, 2020. <https://doi.org/10.1016/j.oceaneng.2019.106909>
- [7] D. Kitamura, H. Saruwatari, K. Shikano, K. Kondo, and Y. Takahashi, "Music signal separation by supervised nonnegative matrix factorization with basis deformation," in *Proceedings of 2013 18th International Conference on Digital Signal Processing (DSP)*, Fira, Greece, 2013, pp. 1-6. <https://doi.org/10.1109/ICDSP.2013.6622812>
- [8] P. Dey, U. Satija, and B. Ramkumar, "Single channel blind source separation based on variational mode decomposition and PCA," in *Proceedings of 2015 Annual IEEE India Conference (INDICON)*, New Delhi, India, 2015, pp. 1-5. <https://doi.org/10.1109/INDICON.2015.7443723>
- [9] Z. Zhang, "Image restoration based on principal component analysis," *Journal of Changchun Normal University (Natural Science Edition)*, vol. 38, no. 6, pp. 21-23, 29, 2019. <https://doi.org/10.3969/j.issn.1008-178X.2019.06.005>
- [10] Z. C. Fan, Y. L. Lai, and J. S. R. Jang, "SVSGAN: singing voice separation via generative adversarial network," in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, 2018, pp. 726-730. <https://doi.org/10.1109/ICASSP.2018.8462091>
- [11] D. Stoller, S. Ewert, and S. Dixon, "Adversarial semi-supervised audio source separation applied to singing voice extraction," in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, 2018, pp. 2391-2395. <https://doi.org/10.1109/ICASSP.2018.8461722>
- [12] X. Lin and Y. Gao, "An improved resampling particle filter blind separation algorithm," *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*, vol. 31, no. 4, pp. 502-508, 2019. <https://doi.org/10.3979/j.issn.1673-825X.2019.04.010>
- [13] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142-3155, 2017. <https://doi.org/10.1109/TIP.2017.2662206>

- [14] E. M. Grais, G. Roma, A. J. Simpson, and M. D. Plumbley, "Two-stage single-channel audio source separation using deep neural networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 9, pp. 1773-1783, 2017. <https://doi.org/10.1109/TASLP.2017.2716443>
- [15] C. Zheng, X. Zhang, T. Cao, J. Yang, M. Sun, and Y. Xing, "Blind enhancement algorithm for throat microphone speech based on LSTM recurrent neural networks," *Journal of Data Acquisition and Processing*, vol. 34, no. 4, pp. 615-624, 2019. <https://doi.org/10.16337/j.1004-9037.2019.04.005>
- [16] T. Wang, C. Qiu, J. Yu, and J. Xu, "Natural image classification based on improved SAE network," *Information Technology*, vol. 2016, no. 8, pp. 1-4,8, 2016. <https://doi.org/10.13274/j.cnki.hdzt.2016.08.001>
- [17] M. Mirza and S. Osindero, "Conditional generative adversarial nets," 2014 [Online]. Available: <https://arxiv.org/abs/1411.1784>.
- [18] Y. C. Subakan and P. Smaragdis, "Generative adversarial source separation," in *Proceedings of 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Calgary, Canada, 2018, pp. 26-30. <https://doi.org/10.1109/ICASSP.2018.8461671>
- [19] Q. Kong, Y. Xu, W. Wang, P. J. Jackson, and M. D. Plumbley, "Single-channel signal separation and deconvolution with generative adversarial networks," 2019 [Online]. Available: <https://arxiv.org/abs/1906.07552>.
- [20] T. Liu, T. Yin, Z. Gong, and Y. Guo, "Research on single-channel blind deconvolution algorithm for multi-source signals," *Journal of Electronics & Information Technology*, vol. 44, no. 1, pp. 230-236, 2022. <https://doi.org/10.11999/JEIT200933>
- [21] T. Yin, T. Liu, and Y. Guo, "Research on single-channel blind deconvolution algorithm for ancient Chinese character restoration," *Small Microcomputer System*, vol. 42, no. 2, pp. 414-417, 2021. <https://doi.org/10.3969/j.issn.1000-1220.2021.02.033>
- [22] Z. Yi, H. Zhang, P. Tan, and M. Gong, "DualGAN: unsupervised dual learning for image-to-image translation," in *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017, pp. 2849-2857. <https://doi.org/10.1109/ICCV.2017.310>



Tiantian Yin <https://orcid.org/0000-0003-4578-4243>

She received an M.E. degree in Taiyuan University of Science and Technology, Taiyuan, China, in 2021. She is currently a teacher at Upgrading Office of Modern College of Humanities and Sciences of Shanxi Normal University. Her research interests include signal processing, image processing, and computer vision.



Yina Guo <https://orcid.org/0000-0002-0998-2448>

She received the B.Sc. degree in 2002 from China University of Mining and Technology, the M.E. degree in 2007, and the Ph.D. degree in 2014, all from Taiyuan University of Science and Technology, China. So far, she has written/co-authored more than 40 papers and two books, and six patents and four soft books have been recognized by relevant Chinese institutions. Her research interests include blind source separation, biological signal processing and phase recovery.



Ningning Zhang <https://orcid.org/0000-0002-6404-0218>

She received an M.E. degree in Northwest Normal University, Lanzhou, China, in 2021. She is a teacher at Upgrading Office of Modern College of Humanities and Sciences of Shanxi Normal University. Her research interest covers signal processing and image processing artificial intelligence.