One to multiple mapping dual learning: Learning multiple signals from one mixture

Ting Liu [a,1], Wen-wu Wang [b,1], Xiao-fei Zhang [a], Yi-na Guo [a,*]

[a] Department of Electronic Information Engineering, Taiyuan University of Science and Technology, Taiyuan, 030024, China
[b] The Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, Surrey GU2 7XH, UK

ARTICLE INFO

ABSTRACT

Single channel blind source separation (SCBSS) refers to separating multiple sources from a mixture collected by a single sensor. Existing methods for SCBSS have limited performance in separating multiple sources and generalization. To address these problems, an algorithm is proposed in this paper to separate multiple sources from a mixture by designing a parallel dual generative adversarial network (PDualGAN) that can build the relationship between a mixture and the corresponding multiple sources to achieve one-to-multiple cross-domain mapping. This algorithm can be applied to a variety of mixtures including both instantaneous and convolutive mixtures. In addition, new datasets for single channel source separation are created which include the mixtures and corresponding sources for this study. Experiments were performed on four different datasets including both one-dimensional and two-dimensional signals. Experimental results show that the proposed algorithm outperforms state-of-the-art algorithms, measured with peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), source-to-distortion ratio (SDR), source-to-interferences ratio (SIR), relative root mean squared error (RRMSE) and correlation.

1. Introduction

Single channel blind source separation (SCBSS) aims to separate multiple sources from a mixture collected by a sensor which has a variety of applications in speech [1], image and EEG signal processing [2][3], biomedical engineering [4][5][6], multi-source phase retrieval [7][8][9], and eddy current pulsed thermography [10]. For example, due to the anisotropies in the cosmic microwave background (CMB) radiation, it is often required to recover the CMB component as accurately as possible from a noisy mixture [11]. The SCBSS problem is an extreme case of an underdetermined source separation problem, which is an inherently ill-posed problem in the sense that only a single-channel mixture is available, and there is no prior knowledge about the sources and mixing matrix which leads to SCBSS being a very pathological and challenging problem.

Several traditional methods have been proposed for SCBSS such as the methods based on optimal filters (e.g. Wiener filter) [12][13], empirical mode decomposition and independent component analysis (EMD-ICA) [14][15], ensemble empirical mode decomposition and independent component analysis (EEMD-ICA) [16][17], and non-negative matrix factorization (NMF) [18][19]. These methods work well for instantaneous mixtures. However, they are also limited to various conditions. For example, filter based methods may require prior knowledge of source signals or transmission channels [13]. The intrinsic mode functions (IMFs) obtained by the EMD and EEMD are prone to modal aliasing [41]. ICA assumes that the sources are statistically independent of each other [42], and NMF assumes that all the elements in the matrix are non-negative constraints [18].

Recently, deep learning methods have been used as a solution for this problem. For example, auto-encoders (AEs) [37] have been proposed for supervised source separation, however, traditional AEs usually use a fully connected layer, which will cause the loss of image spatial information [43]. Singing Voice Separation generative adversarial network (SVSGAN) [38] has been proposed for separating voice sources with time-frequency masking. The Wasserstein-GAN [20] and a two-stage approach [39] is proposed for speech separation, where in the first stage, a dereverberation mask (DM) is applied to dereverberation from the mixture, in the second stage, the ideal ratio mask (IRM) is used to separate sources from the dereverberated mixture. This method involves the use of two DNNs which incur a high computational cost.

---

* Corresponding author.
 E-mail addresses: B20190029@stu.tyust.edu.cn (T. Liu), w.wang@surrey.ac.uk (W.-w. Wang), S20190505@stu.tyust.edu.cn (X.-f. Zhang), zulibest@tyust.edu.cn (Y.-n. Guo).
 [1] Ting Liu and Wen-wu Wang contributed equally to this work.
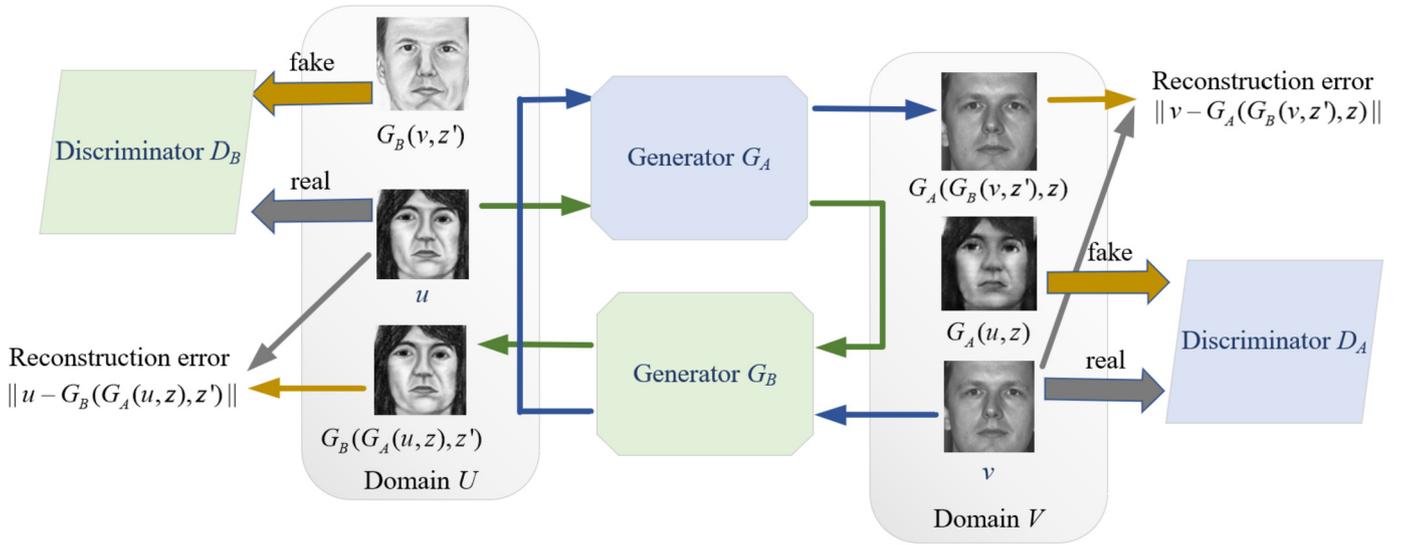
Fig. 1. The working process of DualGAN for image-to-image translation. Image $u \in U$ is translated to domain V using G_A. Similarly, $v \in V$ is translated to U using G_B.

Evolving multi-resolution pooling convolutional neural network (E-MRP-CNN) [40] is proposed for monaural singing voice separation (MSVS). Synthesis-decomposition (S-D) [21] is proposed for separating black and white images such as MINIST images from a mixture, however, they are not suitable for processing color images. In addition, these methods have degraded performance in separating multiple sources (more than three sources) and generalization to unseen scenarios, and they mainly focus on processing signals in a certain field.

In real life, the signal captured by a sensor could be degraded by a number of conditions. For example, the electroencephalogram (EEG) signals are often mixed with interference signals of electrocardiogram (ECG), electromyography (EMG), and eye movement artifacts (EOG). The remote sensing images are often obscured by clouds and fog. Due to the erosion of time, the loss of ancient Chinese characters often occurs in unearthed cultural relics and ancient books. Fingerprint images obtained from criminal investigation scenes often have multiple fingerprints overlapping with each other. In a noisy acoustic environment, multiple speakers are talking simultaneously, which raises the challenge of separating multiple sources from a mixture. Although many methods have been proposed, separating multiple sources from a single channel mixture remains an open challenge.

The aim of this paper is to address the problem of separating multiple sources from a single channel mixture, which is either a one-dimensional or a two-dimensional signal, with either an instantaneous or convolutive mixing model, and improve its generalization performance. Inspired by dual learning and GAN [23] which can be used to build one-to-one mapping to achieve image-to-image translation, we design a parallel dual generative adversarial network (PDualGAN) to achieve one-to-multiple mapping, and formulate SCBSS as a data conversion problem in different domains where multiple sources are separated from a mixture in terms of the mapping between a mixture and the corresponding sources. Our novel contributions are as follows:

Model. We have formulated a unified model for the instantaneous mixing model and convolutive mixing model.
PDualGAN algorithm. A new algorithm is proposed by introducing PDualGAN, where the model training is performed using the mixtures and corresponding multiple sources. The Wasserstein GAN gradient penalty (WGAN-GP) is applied in the loss function. Our algorithm can be applied to both one-dimensional and two-dimensional signals. In addition, different mixtures are used as

experimental data to evaluate the effectiveness and generalization performance.
Datasets. New datasets are created for this study which are composed of two parts: the mixtures and corresponding original sources. Each mixture is generated by using randomly generated mixing matrices and multiple sources with different weights. The datasets created could be valuable for researchers working in the image, speech, and EEG area of learning based source separation.

The remainder of this paper is organized as follows. Section 2 describes the background. Section 3 builds a one-to-multiple mapping model for the SCBSS problem. Section 4 presents our PDualGAN algorithm for the problem of SCBSS. Section 5 shows the experimental results. Section 6 concludes the paper and draws future research directions.

2. Background

A generative adversarial network (GAN) is a deep learning model and one of the most promising methods for unsupervised learning in recent years. It is composed of a generator and a discriminator, where the generator takes random noise as input, and then generates corresponding samples, while the discriminator distinguishes the distribution of the candidate signal generated by the generator from the true data distribution. The model produces good output through adversarial learning of the generator and the discriminator [26]. Inspired by GAN and dual learning [27] from natural language translation, a dual generative adversarial network (DualGAN) [23] is developed for one-to-one unlabeled data from two domains. The original GAN has the limitation in that it can only learn to translate data from domain U to those in domain V, but the DualGAN can learn to invert the task.

Given two sets of unpaired and unlabeled data selected from domains U and V, the task of the DualGAN [23] is to firstly learn a generator $G_A: U \rightarrow V$, which is a mapping from $u \in U$ to $v \in V$, secondly, the dual task is to train an inverse generator $G_B: V \rightarrow U$, i.e. which is a mapping from $v \in V$ to $u \in U$. This is realized with two GANs which have the same structure. The original GAN learns the generator G_A and discriminator D_A that discriminates between the fake and real data of domain V. Similarly, the dual GAN learns the generator G_B and a discriminator D_B. The overall working process is shown in Fig. 1 [23].

The data from domain $u \in U$ is translated to that in domain V with G_A. The fitting degree of $G_A(u, z)$ (z is random noise) is eval-

uated by D_A. Then, $G_A(u, z)$ is translated back to domain U with G_B, and output $G_B(G_A(u, z), z')$ (where z' is also random noise) as a reconstruction of u. Similarly, $v \in V$ is translated to domain U as $G_B(v, z')$ with G_B, and then reconstructed as $G_A(G_B(v, z'), z)$ with G_A. The discriminator D_A is trained with v as true data and $G_A(u, z)$ as fake data, however, D_B takes u as true data and $G_B(v, z')$ as fake data. The generators G_A and G_B are trained and optimized to output fake samples to cheat the corresponding discriminators D_A and D_B, and to minimize the reconstruction error $\| v - G_A(G_B(v, z'), z) \|$ and $\| u - G_B(G_A(u, z), z') \|$.

The cross-entropy loss function of the original GAN [26] is substituted by the loss function of Wasserstein GAN (WGAN) [28], which performs better in generator convergence and data quality, and in improving the stability of the network. The loss function applied in D_A and D_B can be written as

$$l_A^d(u, v) = D_A(G_A(u, z)) - D_A(v), \qquad (1)$$

$$l_B^d(u, v) = D_B(G_B(v, z')) - D_B(u), \qquad (2)$$

where $u \in U$ and $v \in V$.

The same loss function is applied in generators G_A and G_B as they have the same objective which adopts the L_1 distance to measure the reconstruction losses, as follows

$$\begin{aligned} l^g(u, v) = &\lambda_U \| u - G_B(G_A(u, z), z') \| \\ &+ \lambda_V \| v - G_A(G_B(v, z'), z) \| \\ &- D_B(G_B(v, z')) - D_A(G_A(u, z)) \end{aligned} \qquad (3)$$

where λ_U and λ_V are both constant parameters, which are typically set to a value within [100, 1000] [23].

Clearly, by training DualGAN, unlabeled and unpaired data can be converted from U to the corresponding data in V because the data of the two domains have some similar characteristics. Inspired by DualGAN, single channel blind source separation can be addressed by converting the mixtures to the corresponding multiple sources which have similar characteristics (for example, for images, each mixture has similar texture, style, and shape to the corresponding sources). In this paper, we develop a PDualGAN to train multiple DualGANs simultaneously and convert the mixtures to corresponding multiple sources using the PDualGAN algorithm.

3. Mathematical model

SCBSS is an extremely underdetermined problem of Blind source separation (BSS) that only one observed signal can be collected. BSS was first proposed by Herault and Jutten [22]. It refers to the analysis of original signals from multiple observed mixed signals. Assuming that the sources $s(t) = [s_1(t), s_2(t), ..., s_M(t)]^T$ are M independent signals, and $x(t) = [x_1(t), x_2(t), ..., x_N(t)]^T$ are N independent observation signals, that is

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ ... \\ x_N(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1M} \\ a_{21} & a_{22} & ... & a_{2M} \\ ... & ... & ... & ... \\ a_{N1} & a_{N2} & ... & a_{NM} \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ ... \\ s_M(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ ... \\ n_N(t) \end{bmatrix}, \qquad (4)$$

where A is a $N \times M$ matrix, $n(t) = [n_1(t), n_2(t), ..., n_N(t)]^T$ is noise. In practical application, it is assumed that there is no noise, so the (4) can be simplified as

$$x(t) = As(t). \qquad (5)$$

On this basis, Leon Cohen [44] first pointed out that SCBSS requires the separability of the source signal in time and frequency.

Then Hopgood and Rayner [45] studied and proposed that the separation can be realized when the sources are separable after transformation in the transform domain, besides, they defined the mathematical model and provide a theoretical basis for SCBSS. According to the mixing mode, the instantaneous mixing model and the convolutive mixing model are defined as

Instantaneous mixing model. For instantaneous mixing model, $x(t)$ can be defined as

$$x(t) = a_1(t) s_1(t) + a_2(t) s_2(t) + ... + a_N(t) s_N(t), \qquad (6)$$

where $x(t)$ is an observed mixture at discrete time t, $s_i(t)$ is the ith source, and $a_i(t)$ is the ith mixing coefficient, $i = 1, 2, ..., N$. (6) can be simplified as

$$x(t) = \sum_{i=1}^{N} a_i(t) s_i(t). \qquad (7)$$

Convolutive mixing model.

Similarly, for convolutive mixing model, the observed mixture $x(t)$ can be defined as

$$x(t) = \sum_{i=1}^{N} a_i(t) * s_i(t), \qquad (8)$$

where $s_i(t)$ is the ith source, $a_i(t)$ is the ith mixing filter, $i = 1, 2, ..., N$, and $*$ denotes convolution operation which can be defined as

$$a_i(t) * s_i(t) = \sum_{v=-\infty}^{+\infty} a_i(t - v) s_i(v), \qquad (9)$$

where t is the amount of displacement of $a(-v)$. Therefore, (8) can be described as

$$x(t) = \sum_{i=1}^{N} \sum_{v=-\infty}^{+\infty} a_i(t - v) s_i(v). \qquad (10)$$

Single channel blind source separation needs to separate multiple sources from a mixture, it is a very challenging problem that has the following characteristics:

(1) Only a mixture is available and the sources and mixing matrixes are unknown which made it an ill-posed and underdetermined problem in mathematics.

(2) There is no prior knowledge and representation of the sources and mixing matrix.

(3) In a practical situation, due to the reflection and time delay of the sources in the propagation process, the observed mixture is not only generated by instantaneous mixing but also more similar to that generated by convolutive mixing. Therefore, it is very important to solve the problem of single channel blind source separation with two mixing modes.

We can see from (7) and (10) that both the instantaneous mixing model and the convolutive mixing model can be regarded as a form of matrix multiplication, which maps multiple sources to a mixture. The existing methods mainly focus on a certain field with a single mixed mode. The aim of the proposed PDualGAN algorithm is to address the problem of separating multiple sources from a single channel mixture, which can be applied to both one-dimensional and two-dimensional signals, with either an instantaneous or a convolutive mixing model.
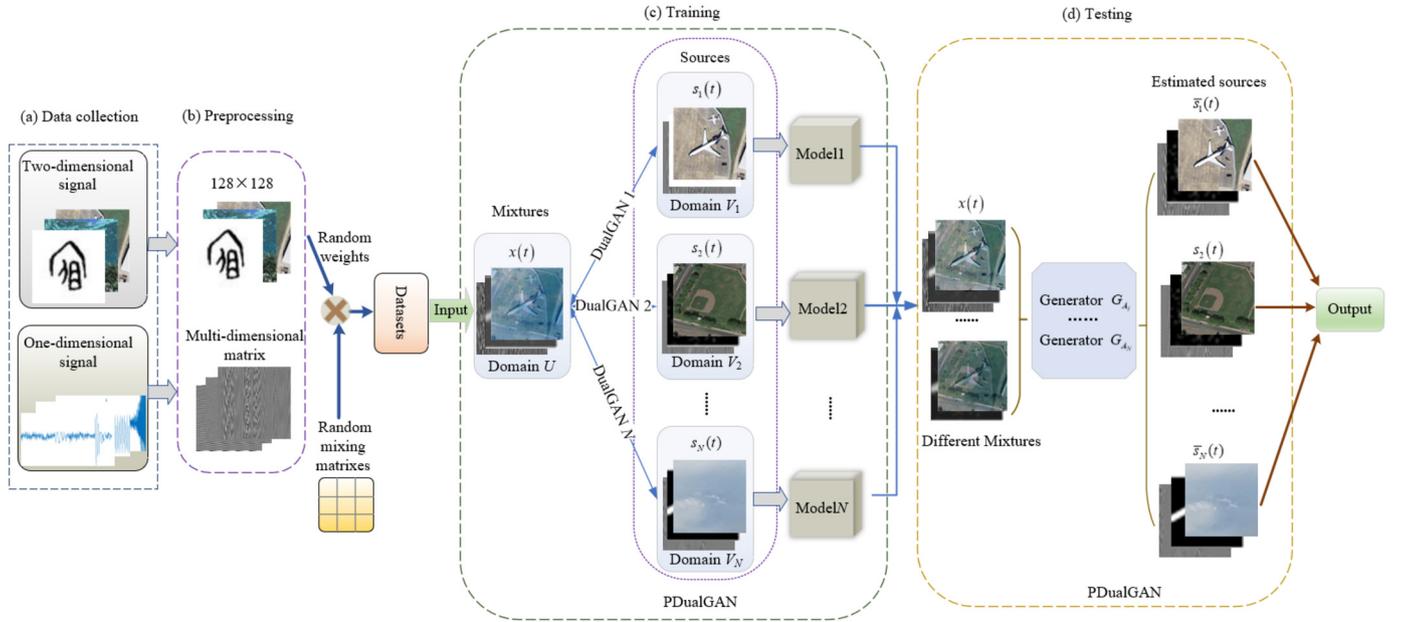
Fig. 2. The overall architecture of the PDualGAN based on the instantaneous and convolutive mixing model, including four parts: the data collection (a), both two-dimensional signals and one-dimensional signals are collected. The data preprocessing (b) and the building of one-to-multiple datasets. The training of PDualGAN (c), where N DualGANs are trained simultaneously, learn the mapping between the mixture $x(t) \in U$ and the corresponding sources $s_1(t) \in V_1$, $s_2(t) \in V_2$, ..., $s_N(t) \in V_N$ sequentially. The testing of PDualGAN (d), where different mixtures are generated by random mixing matrices and multiple sources of different weights (different from the mixing weights in training). $x(t)$ are used as the input for estimating the N sources $\bar{s}_1(t)$, $\bar{s}_2(t)$, ..., $\bar{s}_N(t)$ as the output.

4. PDualGAN algorithm

In this section, we present a PDualGAN algorithm to address the problem of separating multiple sources from a mixture and apply it to one-dimensional and two-dimensional signals. As shown in Fig. 2. Firstly, we collect the one-dimensional and two-dimensional signals. Secondly, the data is preprocessed and the size of two-dimensional signals is adjusted to 128×128, 16384 points are clipped and normalized of one-dimensional signal, and converted into the multi-dimensional matrix, then different mixtures are generated with random mixing matrices and multiple sources with different weights, and the datasets are constructed as the input of PDualGAN. Thirdly, we transform the problem of SCBSS into a data conversion problem in different domains by using PDualGAN to train N DualGANs simultaneously with the mixtures (either instantaneous or convolutive mixtures) and corresponding sources. The mixtures are generated by using randomly generated mixing matrices and multiple sources with different weights (i.e. different sources have different proportions), and the mixing matrices satisfy the standard normal distribution. The final step is to test the effect of the trained model with different mixtures generated by mixing the randomly selected sources with different weights.

4.1. Theoretical basis of PDualGAN algorithm

GAN can generate arbitrary data by random noise to produce clearer and real samples by training the generator and discriminator in an adversarial way. Based on this, A DualGAN is developed by introducing dual learning, which can generate unknown signals with random input and learn to invert the task. DualGANs can be applied to cross-domain image-to-image translation (such as photo-to-sketch conversion) because it implicitly assumes that the structure between input and output images is alignment [23].

Similarly, for the problems described in section 3, a mixture and the corresponding sources also have some similar characteristics, for a two-dimensional signal, each mixed image has some common features (texture, style, shape, cluster, etc.) with the corresponding multiple individual images; For one-dimensional signals,

each mixed signal has some common features (frequency, waveform, amplitude, etc.) with the corresponding multiple individual signals. Therefore, we consider designing a PDualGAN which includes N DualGAN to build one-to-multiple mapping which can be applied to both one-dimensional and two-dimensional signals.

Combined with the characteristics of DualGAN, in our algorithm, no specific domain knowledge or pre-trained domain representation is needed, but the features between the mixtures and multiple corresponding sources are searched to establish the mapping relationship, so, it is also applied to instantaneous mixing and convolutive mixing model.

The reconstruction error measures the disparity between the original sources and the reconstructed signals. PDualGAN contains N DualGANs, which are of the same structure as shown in Fig. 3. The mixtures are sampled from U and the corresponding sources are sampled from V_1, V_2, ..., V_N respectively. The primary task of our PDualGAN is to build the one-to-multiple mapping from $x(t) \in U$ to $s_1(t)$, $s_2(t)$, ..., $s_N(t) \in V_1$, V_2,..., V_N, respectively.

4.2. Network configuration

Each DualGAN has identical network architecture for G_{A_i} and G_{B_i} (G_{A_i} and G_{B_i} are the generators of the ith DualGAN, $i = 1, 2, ..., N$). The generator has the equal number of upsampling and downsampling layers, with skip connections between them, forming a U-shaped net [29][30], and such a structure enables low-level information to be shared between the input and output. In addition, z and z' are provided only in the form of dropout and applied to multiple layers of generators at both training and testing phases, but they are not explicitly provided. For discriminators, the Markovian PatchGAN architecture [31] is applied, which is effective in capturing local high-frequency features, and its effectiveness has been verified on various conversion tasks [27]. Furthermore, it requires fewer parameters, and therefore runs faster than conventional GAN. This scheme will be used for our signal separation tasks.

In the generators, 8 convolution layers ($g_1 \sim g_8$) are included in down-sampling layers and 8 deconvolution layers ($f_1 \sim f_8$) are in-
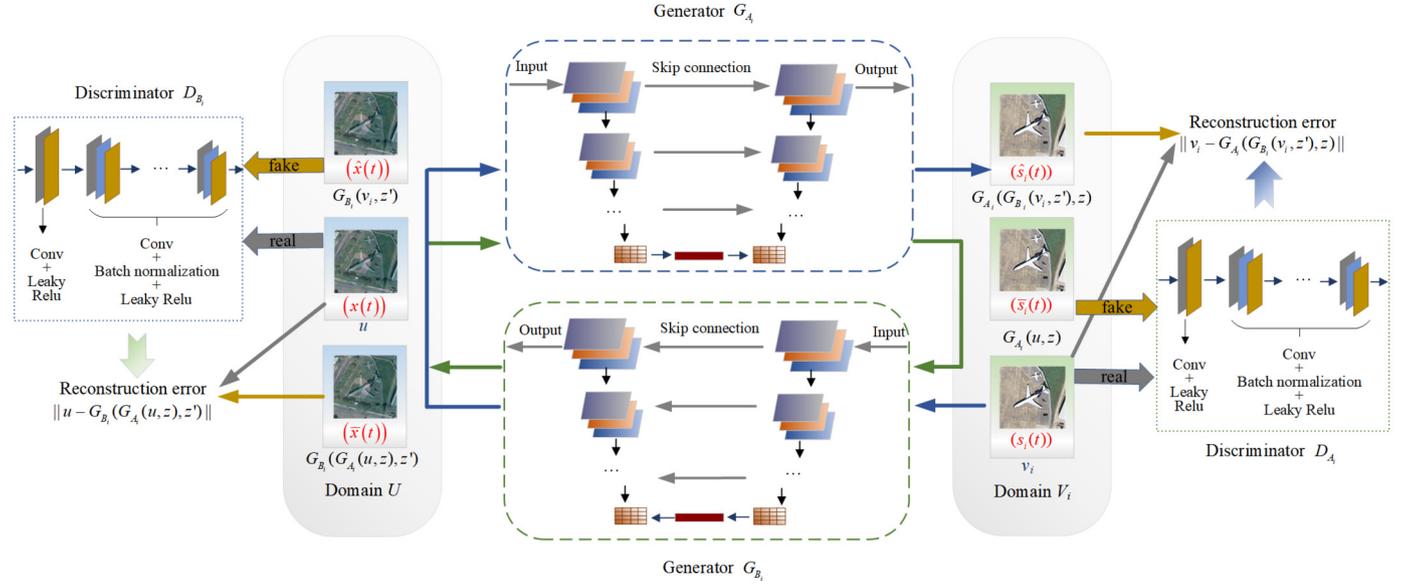
Fig. 3. The working process of each DualGAN. $\bar{s}_i(t) \in V_i$ is the fake data generated by the generator G_{A_i} from the mixture $x(t) \in U$, while $\bar{x}(t) \in U$ is reverse generated data by the generator G_{B_i}. Similarly, $\hat{x}(t) \in U$ is the fake data generated by the generator G_{B_i} from the sources $s_i(t) \in V_i$, and $\hat{s}_i(t) \in V_i$ is reverse generated data by the generator G_{A_i}. $x(t)$ (i.e. u), $s_i(t)$ (i.e. v_i), $\bar{x}(t)$ (i.e. $G_{B_i}(G_{A_i}(u,z),z')$), $\bar{s}_i(t)$ (i.e. $G_{A_i}(u,z)$), $\hat{x}(t)$ (i.e. $G_{B_i}(v_i,z')$), $\hat{s}_i(t)$ (i.e. $G_{A_i}(G_{B_i}(v_i,z'),z)$).

Table 1
The setting of the parameter of generators of the PDualGAN network. L_{nam}, O_{siz}, O_{dim}, C_{ker}, C_{str}, P_{str} and T_{str} represent the name of Layers, output size, output dimension, convolution kernel, convolution stride, pooling stride, and transpose convolution stride.

Generators											
Down-sampling layers						Up-sampling layers					
L_{nam}	O_{siz}	O_{dim}	C_{ker}	C_{str}	P_{str}	L_{nam}	O_{siz}	O_{dim}	C_{ker}	C_{str}	T_{str}
g_1	64×64	64×2	(3,3)	2	2	f_1	1×1	64×16	(3,3)	2	2
g_2	32×32	64×4	(3,3)	2	2	f_2	2×2	64×16	(3,3)	2	2
g_3	16×16	64×8	(3,3)	2	2	f_3	4×4	64×16	(3,3)	2	2
g_4	8×8	64×8	(3,3)	2	2	f_4	8×8	64×16	(3,3)	2	2
g_5	4×4	64×8	(3,3)	2	2	f_5	16×16	64×8	(3,3)	2	2
g_6	2×2	64×8	(3,3)	2	2	f_6	32×32	64×4	(3,3)	2	2
g_7	1×1	64×8	(3,3)	2	2	f_7	64×64	64×2	(3,3)	2	2
g_8	1×1	64×8	(3,3)	2	2	f_8	128×128	3	(3,3)	2	2

Table 2
The setting of the parameter of discriminators of the PDualGAN network. L_{nam}, O_{siz}, O_{dim}, C_{ker}, C_{str} represent the name of Layers, output size, output dimension, convolution kernel, convolution stride.

Discriminators				
L_{nam}	O_{siz}	O_{dim}	C_{ker}	C_{str}
d_1	64×64	64×2	(3,3)	2
d_2	32×32	64×4	(3,3)	2
d_3	16×16	64×8	(3,3)	2
d_4	4×4	64×16	(3,3)	2
d_5	2×2	1	(3,3)	2

cluded in up-sampling layers. In the discriminators, 5 convolution layers ($d_1 \sim d_5$) are included. The setting of parameter of generators and discriminators in PDualGAN network is shown in Table 1 and Table 2 respectively.

4.3. Training

As the momentum-based methods (such as Adam) would occasionally lead to instability, we use mini-batch stochastic gradient descent (SGD) and apply the RMSProp solver which is known to perform well on highly nonstationary signals [28]. The sigmoid cross-entropy loss of the traditional GAN is locally saturated and may cause the gradient to disappear. However, the Wasserstein

GAN gradient penalty (WGAN-GP) [28] loss is differentiable almost everywhere, resulting in a better discriminator which can provide more reliable gradient information.

As shown in Fig. 2 (c) and Fig. 3, for a mixture, a generator G_{A_i}: $U \rightarrow V_i$ in the ith DualGAN is learned by mapping the mixture u to a corresponding source $G_{A_i}(u,z)$ generated, while the dual task is to train an inverse generator G_{B_i}: $V_i \rightarrow U$ that maps a generated source $G_{A_i}(u,z)$ to a generated mixture $G_{B_i}(G_{A_i}(u,z),z')$, where z and z' are random noise signals. The N corresponding sources are simultaneously generated by N DualGANs (i.e. $i = 1, 2, ..., N$) from the same mixture with the identical structure.

For an original source, a generator G_{B_i}: $V_i \rightarrow U$ in the ith DualGAN is learned by mapping the source v_i to a generated mixture $G_{B_i}(v_i,z')$, while the dual task is to train an inverse generator G_{A_i}: $U \rightarrow V_i$ that maps a generated mixture $G_{B_i}(v_i,z')$ to a generated source $G_{A_i}(G_{B_i}(v_i,z'),z)$, where z and z' are random noises. The N different sources generate the same mixture by N DualGANs (i.e. $i = 1, 2, ..., N$) with an identical structure.

The discriminator D_{A_i} discriminates the real source v_i of domain V_i and the fake outputs $G_{A_i}(u,z)$. The discriminator D_{B_i} discriminates the real mixture u of domain U and the fake outputs $G_{B_i}(v_i,z')$. Note that D_{A_i} and D_{B_i} are the discriminators of the ith DualGAN, $i = 1, 2, ..., N$.

The same loss function is applied in each DualGAN for generators G_{A_i} and G_{B_i} which is defined as

---

Algorithm 1 Training of the PDualGAN Algorithm.

Input: The mixture $x(t) \in U$, original sources $s_i(t) \in V_i$, $i = 1, 2, ..., N$, λ_U, λ_{V_i}, initial learning rate, batch size, clipping parameter.

Output: Estimated sources $s_i(t) \in V_i$, $i = 1, 2, ..., N$.

1: Each mixture $x(t)$ (i.e. u) is mapped to N corresponding sources $s_1(t), s_2(t), ..., s_N(t)$ (i.e. $v_1, v_2, ..., v_N$) by N generators $G_{A_1}, G_{A_2}, ..., G_{A_N}$, and N original sources are mapped to the same mixture by N generators $G_{B_1}, G_{B_2}, ..., G_{B_N}$ simultaneously. The results are saved.

2: Optimize the loss function of generators G_{A_i} and G_{B_i} in each DualGAN [23].

$$l^{G_i}(u, v_i) = \lambda_U ||u - G_{B_i}(\{G_{A_i}(u, z), z'|all\ i\})||$$
$$+ \lambda_{V_i} ||v_i - G_{A_i}(G_{B_i}(v_i, z'), z)||$$
$$- D_{B_i}(G_{B_i}(v_i, z')) - D_{A_i}(G_{A_i}(u, z))$$

3: Optimize the loss function of discriminators D_{A_i} and D_{B_i} in each DualGAN [23].

$$l^D_{A_i}(u, v_i) = D_{A_i}(G_{A_i}(u, z)) - D_{A_i}(v_i)$$

$$l^D_{B_i}(u, v_i) = D_{B_i}(G_{B_i}(v_i, z')) - D_{B_i}(u)$$

4: End.

$$l^{G_i}(u, v_i) = \lambda_U ||u - G_{B_i}(\{G_{A_i}(u, z), z'|all\ i\})||$$
$$+ \lambda_{V_i} ||v_i - G_{A_i}(G_{B_i}(v_i, z'), z)|| \qquad (11)$$
$$- D_{B_i}(G_{B_i}(v_i, z')) - D_{A_i}(G_{A_i}(u, z)),$$

where $u \in U$, $v_i \in V_i$, and λ_U, λ_{V_i} are two constant parameters.

The corresponding loss functions applied in D_{A_i} and D_{B_i} are defined as:

$$l^D_{A_i}(u, v_i) = D_{A_i}(G_{A_i}(u, z)) - D_{A_i}(v_i), \qquad (12)$$

$$l^D_{B_i}(u, v_i) = D_{B_i}(G_{B_i}(v_i, z')) - D_{B_i}(u), \qquad (13)$$

where $u \in U$, $v_i \in V_i$.

The one-to-multiple mapping between a mixture $u \in U$ and N sources $v_1, v_2, ..., v_N \in V_1, V_2, ..., V_N$ is built by training the proposed PDualGAN.

4.4. Testing

As shown in Fig. 2 (d), after training the proposed PDualGAN, we save the trained model and parameters which are then used in the test stage. The mixtures with different weights are converted to corresponding multiple sources to achieve separation of multiple sources from the mixture.

5. Numerical experiments

In this section, experiments are conducted to demonstrate the performance of the proposed PDualGAN algorithm. Both one-dimensional and two-dimensional signals are used in the experiment. In addition, taking the separation of four sources from a mixture as an example, we compare the PDualGAN algorithm with state-of-the-art baseline algorithms.

Experimental settings. We train the discriminators n_{critic} steps, then one step on generators. The number of critic iterations per generator iteration n_{critic} can be set to 2-4, λ_U and λ_{V_i} are all set to 1000, an initial learning rate is set to 0.00005, and the batch size is assigned as 1. The clipping parameter is set in [0.01, 0.1], and λ_U, λ_{V_i} are set to a value within [100.0, 1000.0] [23].

Experimental data. We use four kinds of datasets for the experiment: the NWPU-occlusion image datasets, the ancient Chinese character occlusion (ACC-occlusion) datasets, the speech datasets, and the EEG datasets, with more details given later. We select a total of 2000 original sources, i.e. each dataset includes 500 sources. The mixtures are generated by using randomly generated mixing matrices and four sources with different weights. Each dataset includes two types: (1) instantaneous mixtures and corresponding sources; (2) convolutive mixtures and corresponding sources.

In the experiment, we randomly select 80% mixed signals and the corresponding original signals for training, the remaining 20% mixed signals for testing.

[NWPU-occlusion image datasets] This dataset is formed by the original images selected from the NWPU-RESISC dataset[2] containing 45 categories, and cloud and fog occlusion images found with Google and Baidu search engines. Each mixed image is generated by mixing two NWPU-RESISC images and two cloud and fog occlusion images in random weights.

[ACC-occlusion datasets] This dataset includes the ancient Chinese character images and occlusion images. Each mixed image is obtained by mixing one ancient Chinese character image and three occlusion images in different random weights.

[Speech datasets] The speech signals are selected from the THCHS-30 dataset[3] with each sentence containing 16384 samples randomly clipped and then normalized. Each mixture is obtained by mixing four randomly selected speech signals in different random weights. The THCHS-30 dataset was recorded by a single carbon microphone in a quiet office environment, where the signals were sampled at 16 kHz and quantized in 16 bits. Most of the speakers are college students who can speak fluent Mandarin. In total, 1000 recordings are taken, which have a total duration of more than 30 hours.

[EEG datasets] The EEG acquisition equipment used in the experiment is the EEG and evoked potential meter (model NCERP-T-240) of Shanghai Nuocheng Electric Co., Ltd. We use 24-channel silver-plated electrodes, and the electrode placement position adopts the international standard 10/20. An attention device using the ThinkGear^TM Asic Module (TGAM) chip was developed to test the attention of the participants, where a pre-defined threshold p can be used to detect the attention of the participant. If it is higher than the pre-defined threshold p (e.g. $p > 60$), it indicates that the participant has a good focus on the auditory stimuli, and at this time, EEG signals can be collected by the device.

A total of 50 participants is included in the EEG signal collection, of which 25 are males and 25 are females, aged between 20-40. All participants are either teachers or students from Taiyuan University of Science and Technology who are in good health and meet the requirements for EEG collection experiments.

We select the EEG signals according to the positions of the EEG electrodes of the temporal and frontal lobes and pre-processing them to remove artifacts such as ECG and oculus, then we randomly clipped 16384 points in each channel and normalize and convert the selected EEG signals into a multi-dimensional matrix to construct an EEG dataset. Each mixture is obtained by mixing four original EEG signals from the collected dataset in different random weights.

The four kinds of datasets are used to train our network. All the datasets are pre-processed, and each mixture corresponds to four sources, which are used as the input data of the PDualGAN network. Finally, different mixtures of four sources of different weights are selected as test signals to demonstrate the effectiveness of the proposed algorithm.

Baseline method. We compare our PDualGAN algorithm with baseline algorithms for both instantaneous and convolutive mixing model based algorithms in separating four sources from a mixture. Instantaneous mixing model based algorithms include ICA[4] [46], NMF[5] [47], EMD-ICA[6] (empirical mode decomposition and

---

2   https://hyper.ai/datasets/5449.
3   http://www.openslr.org/resources/18/data.
4   https://so.csdn.net/so/search?spm=1000.2115.3001.7499&q=FastICA_25&t=&u=&urw.
5   https://download.csdn.net/download/weixin_38688371/19076985.
6   https://so.csdn.net/so/search?q=CEEMDAN_V00&t=doc&u=&urw.

Table 3

Part of implementation details of the compared algorithms.

Algorithms	Running file	Format	Sampling rate	Number of output sources
ICA	FastICA	16384×4 matrices	-	4
NMF	nmf.m	16384×4 non-negative matrices	-	4
EMD-ICA	emd.m, FastICA	16384×1 matrices	-	4
EEMD-PCA-ICA	eemd.m, FastICA	16384×1 matrices	-	4
CEEDMAN-ICA	ceemdan.m, FastICA	16384×1 mat	-	4
E-SSA-ICA	main_SSA_ICA.m	1×16384 matrices	-	4
Gaussian WGAN	main_timit.py	16384×1 wav	16000	2
CG	demo.m	128×128 mat	-	2
IRLS	demo.m	128×128 mat	-	2
CNMF	cnmf.m	16384×4 non-negative matrices	-	4
E-S-D	main.py, deconvolve_separate.py	128×128 jpg	-	4
SCBDC	main.py, deconvolve_separate.py	128×128 jpg	-	4
E-MRP-CNN	dsd2_train.py, dsd2_eval.py	262144×2 wav	44100	2

independent component analysis) [15], EEMD-PCA-ICA[7] (ensemble empirical mode decomposition and principal component analysis and independent component analysis) [16], CEEDMAN-ICA[8] (complete ensemble empirical mode decomposition with adaptive noise and independent component analysis) [32], SSA-ICA[9] (singular spectrum analysis and independent component analysis) [33] and Gaussian WGAN[10] [20] algorithms. Convolutive mixing model based algorithms including CG[11] [36] (Conjugate Gradient), IRLS[12] [36], CNMF[13] (convolutive NMF) [48], S-D (synthesis-decomposition)[14] [21], SCBDC (single-channel blind deconvolution algorithm based on optimized deep convolutional generative adversarial networks) [47] and E-MRP-CNN[15] algorithms [40].

The ICA, NMF, CNMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, SSA-ICA, CG, and IRLS algorithms are implemented in Matlab, The facilities applied to perform the Gaussian WGAN, S-D, SCBDC, E-MRP-CNN and PDualGAN algorithms include Intel I9-10900X 13.7 GHz CPU, 2*NVIDIA RTX 8000 Graphics Card and 6*32 GB memory. Part of the implementation details of these algorithms is shown in Table 3. For the SSA-ICA and S-D algorithms, we adjusted part of the program code so that it can process four sources, therefore, we

renamed these two algorithms as extended SSA-ICA (E-SSA-ICA) and extended S-D (E-S-D) respectively. The codes of ICA, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, and E-SSA-ICA algorithms cannot be used directly to process color images, so we process the R, G, and B channels separately, and then merged the three channels. Original Gaussian WGAN and E-MRP-CNN algorithms are designed for audio and singing voice and are not applicable for processing images, so the two algorithms are only applied to one-dimensional signals. For the result of the E-MRP-CNN algorithm, the vocal represents the first source, and the accompaniment may be a mixture of the remaining sources.

Evaluation index. For the experiments on the NWPU-occlusion and ACC-occlusion image datasets, the performance of the proposed PDualGAN algorithm can be evaluated by the peak signal-to-noise ratio (PSNR) [34] structural similarity index (SSIM) [48] and correlation [35]. For the experiments on the speech datasets, the performance of the proposed PDualGAN algorithm can be evaluated by the source-to-distortion ratio (SDR) [49], source-to-interferences ratio (SIR) [49] and correlation. For the experiments on the EEG datasets, the performance of the proposed PDualGAN algorithm can be evaluated by the relative root mean squared error (RRMSE) [16][25] and correlation.

The PSNR can be defined as

$$\text{PSNR} = 10\log_{10}\left(\frac{\text{MAX}_I{}^2}{\text{MSE}}\right), \tag{14}$$

where MSE is the mean square error between two images with size $w \times v$. MAX represents the maximum value of an image without noise. The MSE of images M and H can be written as

---

7 https://so.csdn.net/so/search?q=CEEMDAN_V00&t=doc&u=&urw.
8 https://so.csdn.net/so/search?q=CEEMDAN_V00&t=doc&u=&urw.
9 https://download.csdn.net/download/qq_39065549/12318104.
10 https://github.com/ycemsubakan/sourceseparation_misc.
11 http://groups.csail.mit.edu/graphics/CodedAperture.
12 http://groups.csail.mit.edu/graphics/CodedAperture.
13 https://download.csdn.net/download/weixin_38688371/19076985.
14 https://github.com/qiuqiangkong/gan separation deconvolution.
15 https://github.com/tuxzz/emrpcnn_pub.

Table 4

The performance of our algorithm compared with instantaneous model based algorithms on the NWPU-occlusion image dataset.

Algorithms	NWPU-occlusion Dataset											
	PSNR				SSIM				Correlation			
	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4
ICA	6.71	5.49	5.26	4.98	0.37	0.29	0.17	0.02	0.32	0.26	0.10	0.12
NMF	7.92	6.88	7.43	6.05	0.29	0.21	0.19	0.11	0.33	0.23	0.28	0.12
EMD-ICA	9.64	7.88	6.13	5.69	0.30	0.25	0.15	0.04	0.41	0.26	0.17	0.11
EEMD-PCA-ICA	11.45	7.62	6.91	6.44	0.33	0.22	0.19	0.16	0.55	0.32	0.18	0.19
CEEDMAN-ICA	8.26	5.23	6.87	7.83	0.25	0.13	0.15	0.16	0.37	0.14	0.15	0.22
E-SSA-ICA	9.10	10.24	8.80	7.41	0.21	0.38	0.19	0.18	0.42	0.53	0.29	0.27
PDualGAN	22.15	17.10	21.87	16.22	0.76	0.67	0.72	0.65	0.77	0.69	0.73	0.65

Table 5

The performance of our algorithm compared with convolutive model based algorithms on the ACC-occlusion dataset.

Algorithms	ACC-occlusion Dataset											
	PSNR				SSIM				Correlation			
	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4
CG	4.55	5.48	-	-	0.18	0.12	-	-	0.17	0.14	-	-
IRLS	6.39	2.63	-	-	0.24	0.07	-	-	0.23	0.09	-	-
CNMF	8.12	7.49	7.05	5.8	0.26	0.22	0.17	0.15	0.28	0.23	0.18	0.13
E-S-D	12.50	14.46	13.41	13.09	0.52	0.61	0.57	0.53	0.53	0.60	0.64	0.57
SCBDC	14.70	13.54	14.88	14.92	0.64	0.52	0.56	0.65	0.55	0.59	0.65	0.66
PDualGAN	26.23	22.03	19.71	16.56	0.78	0.73	0.77	0.65	0.80	0.72	0.76	0.67

Table 6

The performance of our algorithm compared with convolutive model based algorithms on the Speech dataset.

Algorithms	Speech Dataset											
	SDR				SIR				Correlation			
	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4
CG	4.85	2.18	-	-	5.09	2.33	-	-	0.40	0.22	-	-
IRLS	3.38	2.13	-	-	3.31	1.89	-	-	0.26	0.13	-	-
CNMF	4.31	3.89	3.79	3.54	5.01	4.22	2.76	3.54	0.33	0.28	0.29	0.21
E-S-D	8.50	9.46	7.41	7.09	9.19	9.41	8.25	7.13	0.53	0.61	0.54	0.48
SCBDC	9.14	9.48	8.26	7.60	9.78	9.88	8.45	8.03	0.63	0.64	0.58	0.57
E-MRP-CNN	8.16	9.23	-	-	9.22	10.59	-	-	0.69	0.59	-	-
PDualGAN	11.03	12.63	17.71	18.36	12.80	12.02	18.67	17.76	0.65	0.62	0.79	0.78

Table 7

The performance of our algorithm compared with instantaneous model based algorithms on the EEG dataset.

Algorithms	EEG Dataset							
	RRMSE [%]				Correlation			
	s_1	s_2	s_3	s_4	s_1	s_2	s_3	s_4
ICA	62.27	66.14	68.26	72.11	0.36	0.31	0.23	0.11
NMF	64.33	68,17	70.45	72.29	0.34	0.24	0.19	0.14
EMD-ICA	66.12	69.24	73.54	77.34	0.21	0.14	0.06	0.03
EEMD-PCA-ICA	65.14	68.12	74.45	78.02	0.23	0.19	0.03	0.01
CEEDMAN-ICA	60.44	65.98	69.13	74.25	0.19	0.15	0.05	0.03
E-SSA-ICA	63.18	66.72	73.87	75.73	0.28	0.11	0.04	0.02
Gaussian WGAN	56.48	67.26	-	-	0.54	0.22	-	-
PDualGAN	37.12	42.59	44.32	51.14	0.74	0.66	0.69	0.67

Table 8

The running time complexity of the PDualGAN algorithm compared with state-of-the-art algorithms.

Instantaneous model based algorithms	Running time (s)	Convolutive model based algorithms	Running time (s)
ICA	0.42	CG	19.05
NMF	0.57	IRLS	19.05
EMD-ICA	0.61	CNMF	0.59
EEMD-PCA-ICA	17.34	E-S-D	30
CEEDMAN-ICA	35.44	SCBDC	29
E-SSA-ICA	2.78	E-MRP-CNN	55
Gaussian WGAN	10.92	PDualGAN	8.41
PDualGAN	8.41		
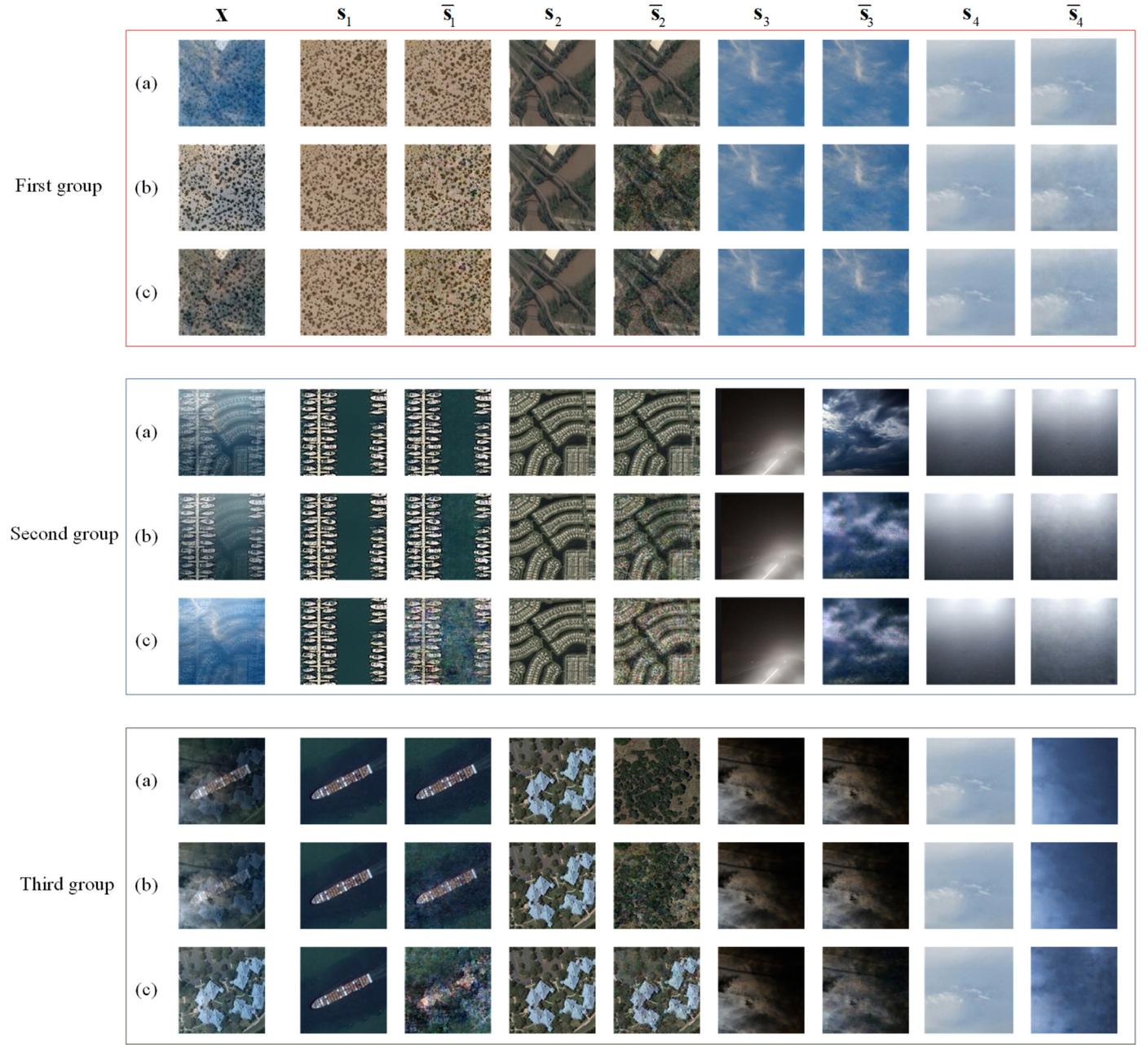
Fig. 4. Testing results of the PDualGAN approach for instantaneous mixing of the NWPU-occlusion images. Each group includes three different mixtures ((a), (b) and (c)) of four images of different random weights. The **X** represents the mixture, s_1, s_2, s_3, and s_4 represent the four original ground-truth sources, $\bar{s}_1$, $\bar{s}_2$, $\bar{s}_3$, and $\bar{s}_4$ represent the corresponding estimated sources.

Table 9
The average correlation between s_i and $\bar{s}_i$ of the proposed PDualGAN compared with the state-of-the-art instantaneous mixing algorithms.

Algorithm	First group			Second group			Third group		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
ICA	0.31	0.27	0.29	0.21	0.23	0.22	0.18	0.19	0.20
NMF	0.28	0.25	0.23	0.24	0.22	0.23	0.24	0.27	0.22
EMD-ICA	0.21	0.22	0.28	0.25	0.24	0.29	0.28	0.35	0.33
EEMD-PCA-ICA	0.28	0.27	0.28	0.24	0.22	0.25	0.33	0.31	0.30
CEEDMAN-ICA	0.37	0.27	0.25	0.35	0.29	0.32	0.28	0.18	0.21
E-SSA-ICA	0.31	0.25	0.37	0.21	0.34	0.22	0.47	0.31	0.39
PDualGAN	0.98	0.92	0.95	0.84	0.74	0.75	0.62	0.65	0.64

$$\text{MSE}(M, H) = \frac{1}{wv} \sum_{i=0}^{w-1} \sum_{j=0}^{v-1} (M(i, j) - H(i, j))^2. \tag{15}$$

The SSIM can be defined as

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \tag{16}$$

(a) Average PSNR(dB)

(b) Average SSIM
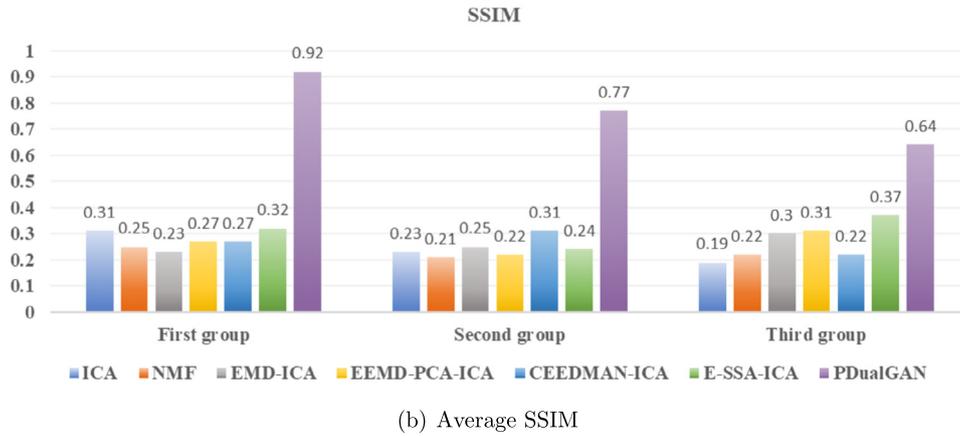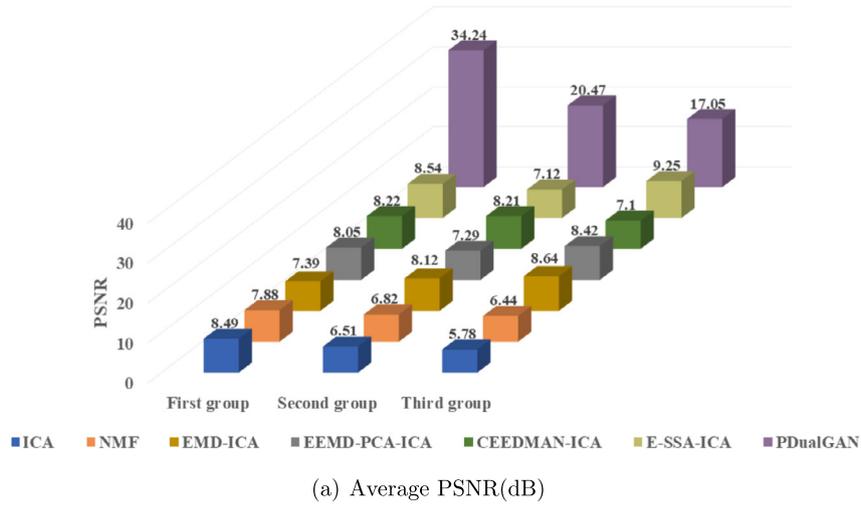
Fig. 5. Average PSNR and SSIM of the proposed PDualGAN compared with the state-of-the-art instantaneous mixing model based algorithms. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

(a) Average PSNR varying with SNR.

(b) Average SSIM varying with SNR.

(c) Average correlation varying with SNR.
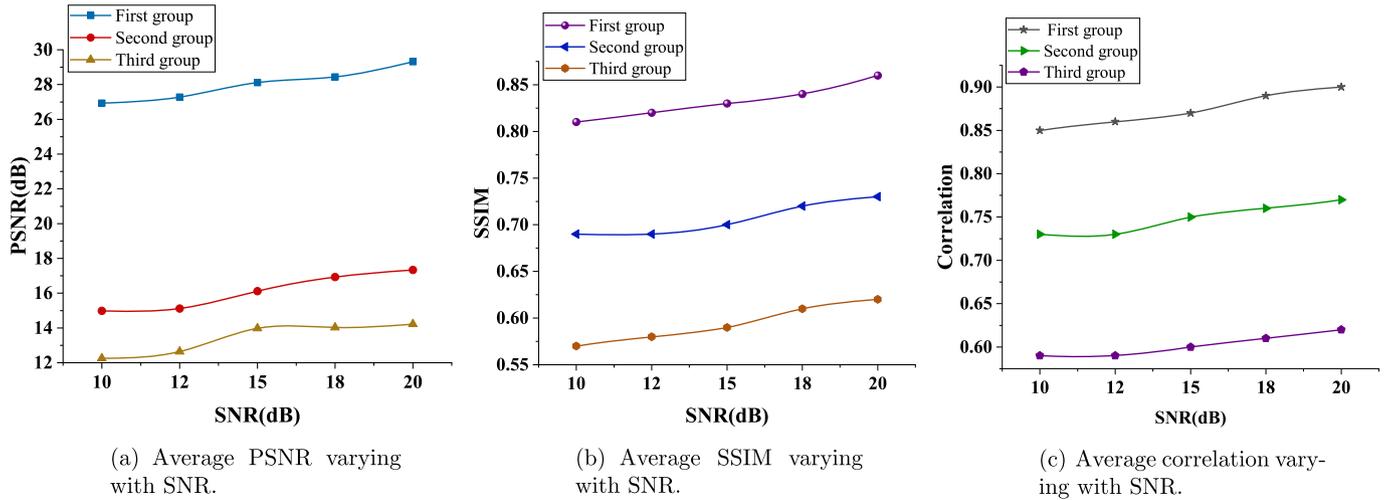
Fig. 6. Average PSNR, SSIM and correlation of the proposed PDualGAN varying with SNR.

where μ_x and μ_y are the mean of x and y, μ_x^2 and μ_y^2 are the variance of x and y, σ_{xy} is the covariance, $c_1 = (K_1L)^2$ and $c_2 = (K_2L)^2$ are constants for maintaining stability, L is the dynamic range of the pixel values, K_1 is 0.01 and K_2 is 0.03.

The SDR and SIR can be defined as

$$SDR = 20\log_{10}\frac{\|s_{target}\|}{\|e_{interf} + e_{noise} + e_{artif}\|}, \tag{17}$$

$$SIR = 20\log_{10}\frac{\|s_{target}\|}{\|e_{interf}\|}, \tag{18}$$

(a) Average PSNR(dB)

(b) Average SSIM
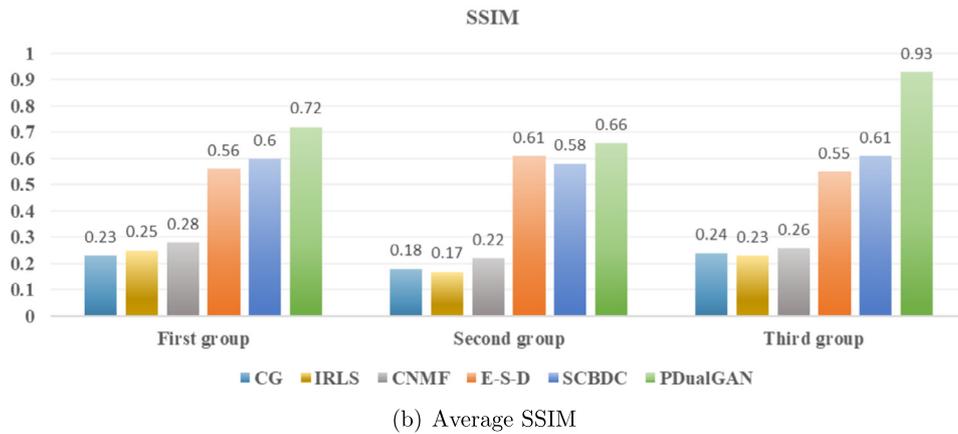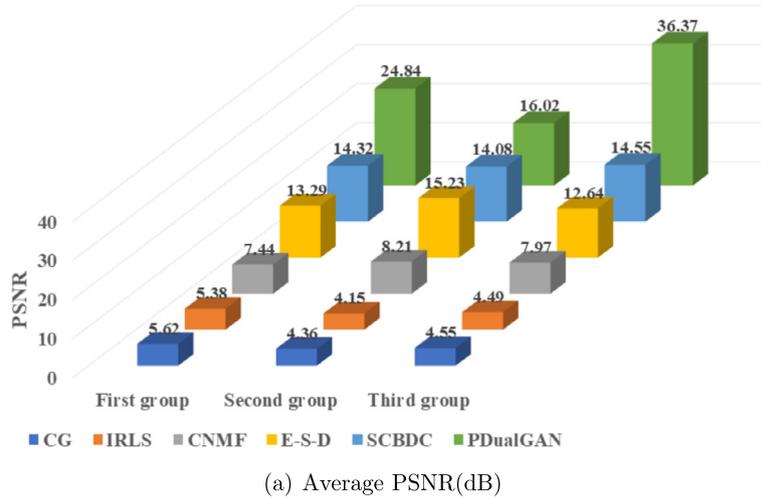
Fig. 7. Average PSNR and SSIM of the proposed PDualGAN compared with the state-of-the-art convolutive mixing model based algorithms.

where the $\mathbf{s}_{target}$ represents the target source, $\mathbf{e}_{interf}$ is an interference of the unwanted source, $\mathbf{e}_{noise}$ is a noise source, $\mathbf{e}_{artif}$ is an interference source generated by algorithm.

The RRMSE is given by

$$RRMSE = \frac{RMS\left(s\left(t\right) - \hat{s}\left(t\right)\right)}{RMS\left(s\left(t\right)\right)} \times 100[\%], \qquad (19)$$

where $\hat{s}(t)$ is the estimate signal and $s(t)$ is the original signal, RMS(.) is the root mean square.

The correlation coefficient of signals X and Y can be expressed as

$$r = \frac{\sum_{i=1}^{N}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sqrt{\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2}}. \qquad (20)$$

We conducted experiments on four datasets according to the two mixing modes. For two-dimensional signals, we show the separation result of the instantaneous mixtures on the NWPU-occlusion image datasets, and the separation result of the convolutive mixtures on the ACC-occlusion datasets; for one-dimensional signals, we show the separation result of the instantaneous mixtures on the EEG datasets, and the separation result of the convolutive mixtures on the Speech datasets. Table 4, Table 5, Table 6 and Table 7 show the comparison results of average values of our algorithm as compared with the state-of-the-art algorithms for separating four sources.

As demonstrated, in separating four sources, the traditional ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA, CG, IRLS, and CNMF algorithms have limited performance, the E-S-D and SCBDC algorithm have relatively better performance in separating four sources on ACC-occlusion datasets and speech datasets, and the E-MRP-CNN algorithm can achieve higher SDR and SIR on the Speech datasets. In general, the proposed PDualGAN algorithm has better performance than the ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA, Gaussian WGAN, CG, IRLS, CNMF, E-S-D, SCBDC and E-MRP-CNN algorithms.

Besides, Table 8 demonstrated the running time complexity of the instantaneous and convolutive model based algorithms which are approximately calculated. As demonstrated, The traditional ICA, NMF, CNMF, EMD-ICA, and E-SSA-ICA algorithms have low computational complexity, and the EEMD-PCA-ICA, CEEDMAN-ICA, CG and IRLS algorithms have relatively high complexity. Although DualGAN has a long training time of 80.50 h, it runs for less than 9 s, which is lower than other Gaussian WGAN, S-D, SCBDC, and E-MRP-CNN deep learning algorithms.

5.1. PDualGAN for NWPU-occlusion images

In the first set of simulations, we evaluate the separation performance of the proposed PDualGAN algorithm described in Algorithm 1 for instantaneous mixtures on the NWPU-occlusion image datasets.

Fig. 4 shows part of the test results of separating four images from instantaneous mixtures, three groups of instantaneous mix-
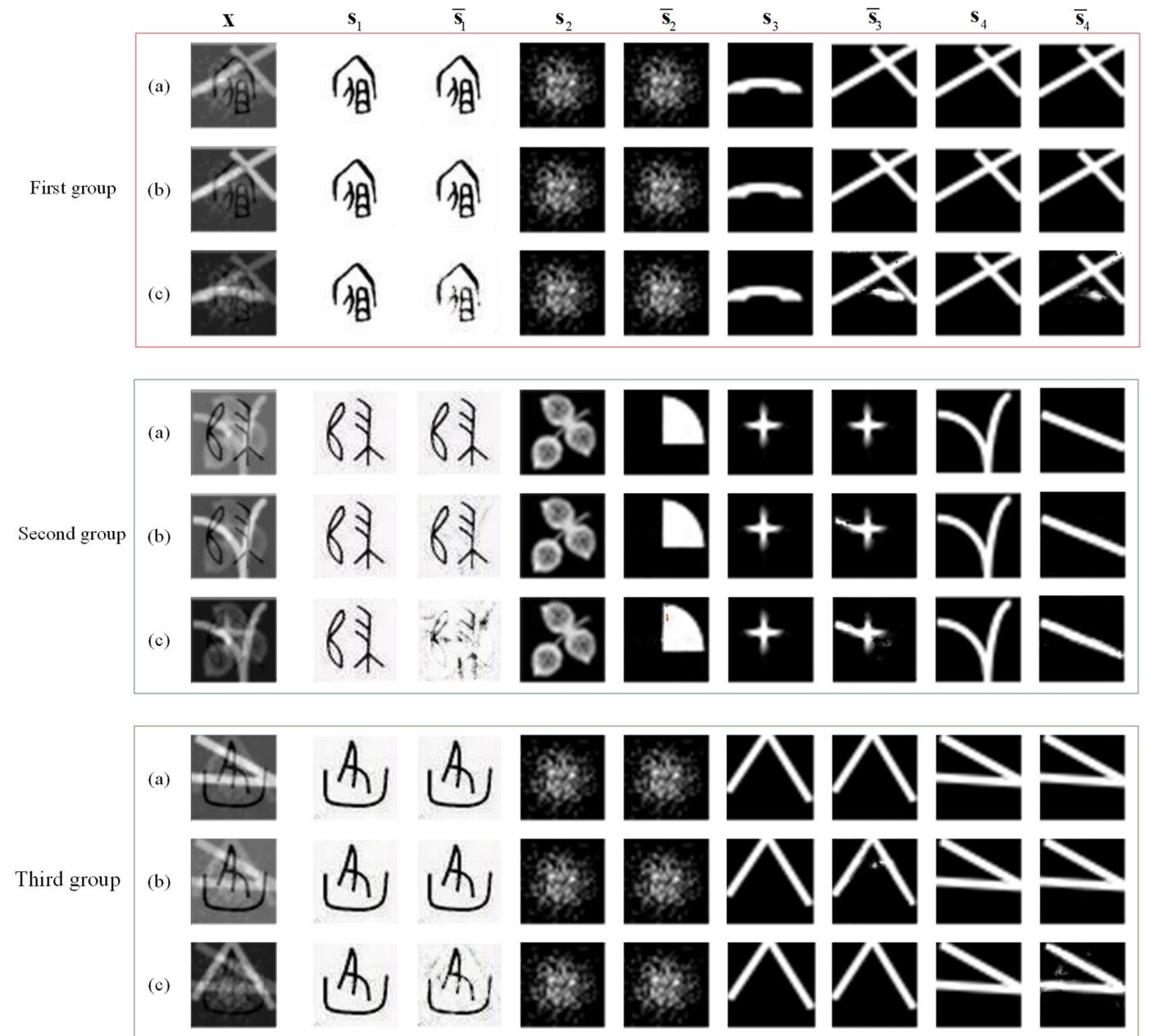
Fig. 8. Testing results of the PDualGAN approach for convolutive mixing ACC-occlusion images. Each group includes three different mixtures ((a), (b) and (c)) of four images of different random weights. The **x** represents the mixture, s_1, s_2, s_3, and s_4 represent the four original ground-truth sources, $\bar{s}_1$, $\bar{s}_2$, $\bar{s}_3$, and $\bar{s}_4$ represent the corresponding estimated sources.

tures are selected from the NWPU-occlusion datasets, where each group shows the separating results from three different mixtures ((a), (b), (c) in Fig. 4), and each mixture is generated by randomly selected mixing matrices and four images with random weights. Fig. 5(a) and Fig. 5(b) show the comparison results of average PSNR and SSIM obtained by the proposed PDualGAN algorithm and ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA instantaneous model based algorithms. Table 9 shows the comparison results of average correlation obtained by the proposed PDualGAN algorithm and these state-of-the-art algorithms.

As observed from Fig. 4, Fig. 5 and Table 9, the proposed PDualGAN algorithm can obtain high PSNR, SSIM and correlation when the four sources can be separated correctly (such as the first group in Fig. 4), otherwise, the performance will be degraded (such as the second group and the third group in Fig. 4). In general, the

average PSNR of the proposed algorithm can achieve 23.92 dB, outperforming the baseline algorithms ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, and E-SSA-ICA of 6.93 dB, 7.05 dB, 8.05 dB, 7.92 dB, 7.84 dB, and 8.30 dB, the average SSIM can achieve 0.77 which is significantly higher than these algorithms of 0.24, 0.23, 0.26, 0.27, 0.27, 0.31. The average correlation obtained by the PDualGAN algorithm can reach 0.78 which outperforms the baseline algorithms of 0.23, 0.24, 0.27, 0.28, 0.28, and 0.32. Therefore, the PDualGAN algorithm achieves better results than the baseline instantaneous mixing model based algorithms.

In addition, considering the effect of noise on the proposed PDualGAN algorithm, the white Gaussian noise is added to the instantaneous mixtures in Fig. 4 to test the separation performance. The average PSNR, SSIM and correlation are measured with the source-to-noise ratio (SNR) ranging from 10 dB to 20 dB, the re-
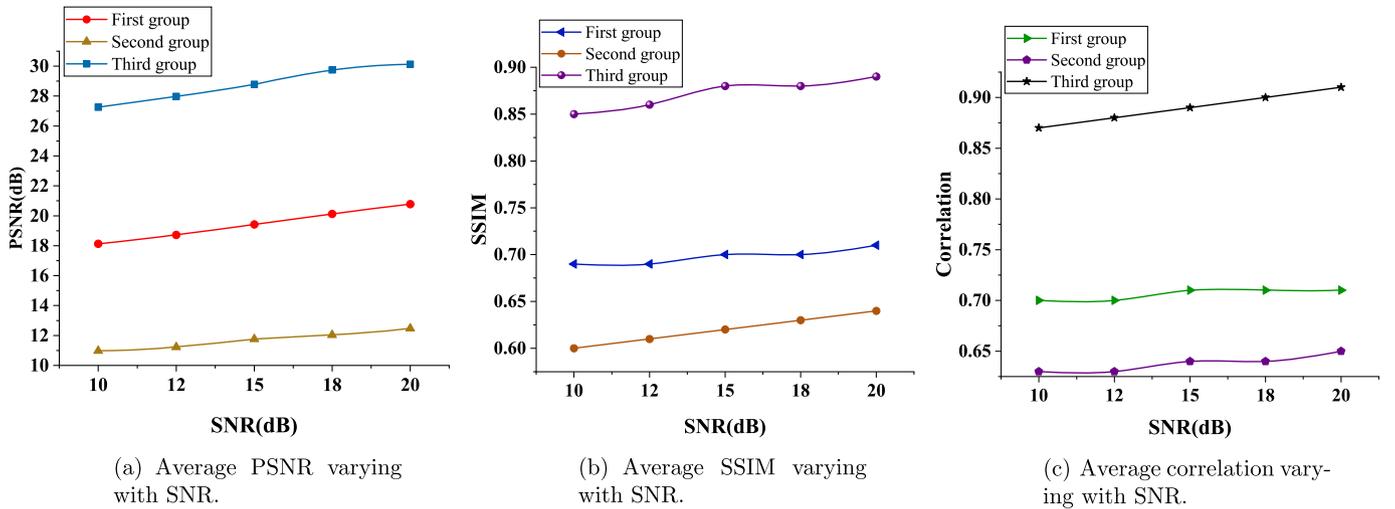
(a) Average PSNR varying with SNR.

(b) Average SSIM varying with SNR.

(c) Average correlation varying with SNR.

Fig. 9. Average PSNR, SSIM and correlation of the proposed PDualGAN algorithm varying with SNR.

(a) Average SDR.
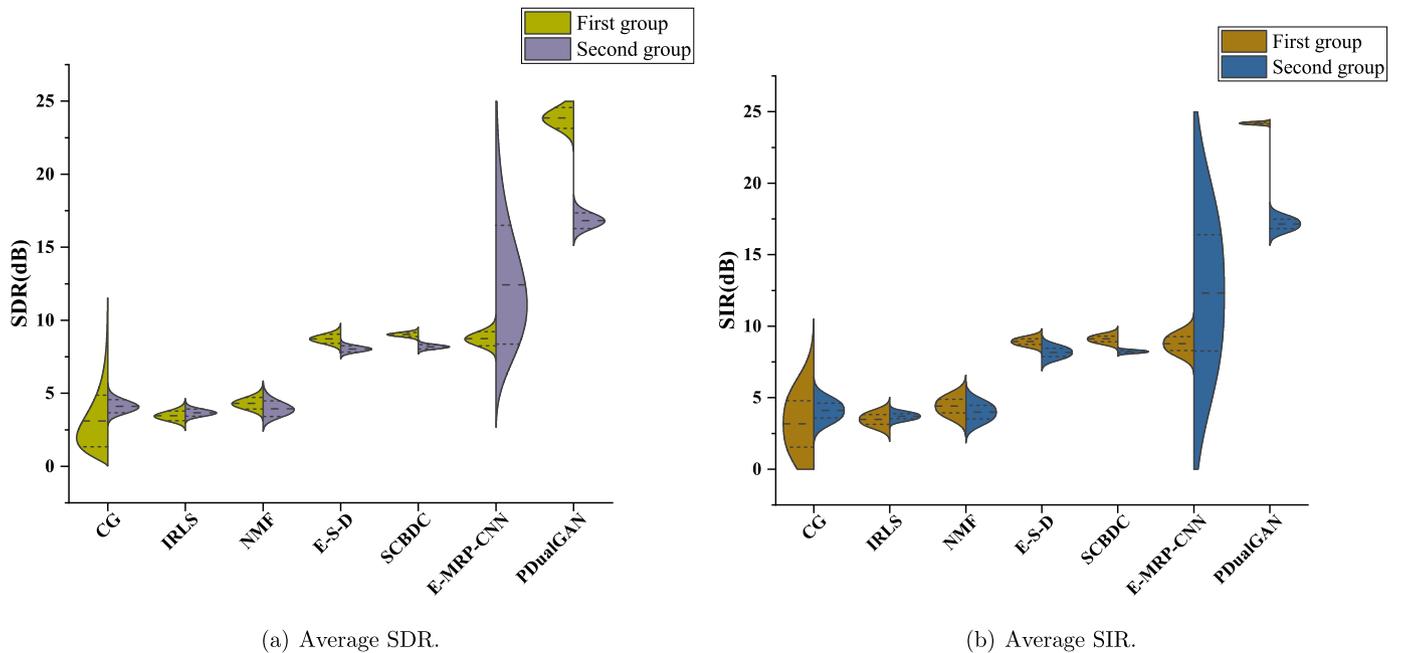
(b) Average SIR.

Fig. 10. Average SDR and SIR of the proposed PDualGAN compared with the state-of-the-art convolutive mixing model based algorithms.

Table 10
The results of average, standard deviation and confidence interval of PSNR of the instantaneous mixing model based algorithms on the NWPU-occlusion dataset.

Algorithm	Average PSNR	Standard deviation	Confidence interval
ICA	5.61	0.43	[5.52, 5.69]
NMF	7.07	0.48	[6.98, 7.16]
EMD-ICA	7.34	2.44	[6.86, 7.81]
EEMD-PCA-ICA	8.11	3.90	[7.34, 8.87]
CEEDMAN-ICA	7.05	1.35	[6.78, 7.31]
E-SSA-ICA	8.89	1.02	[8.69, 9.09]
PDualGAN	19.34	6.82	[18.01, 20.67]

sults are shown in Fig. 6. As observed, the proposed PDualGAN algorithm can still have good performance and is robust.

Considering the statistical error, we provide the average, standard deviation and confidence interval of PSNR to analyze the dispersion of the results, confidence intervals include the effect information about precision and magnitude [24]. We set the confidence level to 95%, and the test sample size to 100. Table 10 given

the results of the ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA and the proposed PDualGAN algorithms.

5.2. PDualGAN for ACC-occlusion images

In the second set of simulations, we evaluate the separation performance of the proposed PDualGAN algorithm described in Algorithm 1 for convolutive mixtures on the ACC-occlusion datasets.

Fig. 8 shows part of the test results of three groups of convolutive mixtures which are selected from the ACC-occlusion image datasets, each group shows the separating results from three different mixtures ((a), (b), (c) in Fig. 8), and each mixture is a convolutive mixture generated by random mixing matrices and four images with different random weights. We compared the PDualGAN algorithm with CG, IRLS, CNMF, E-S-D, and SCBDC convolutive model based algorithms according to the different mixtures of Fig. 8. The comparison results of average PSNR, SSIM and correlation obtained by these state-of-the-art algorithms are shown in Fig. 7(a), Fig. 7(b) and Table 11 respectively.
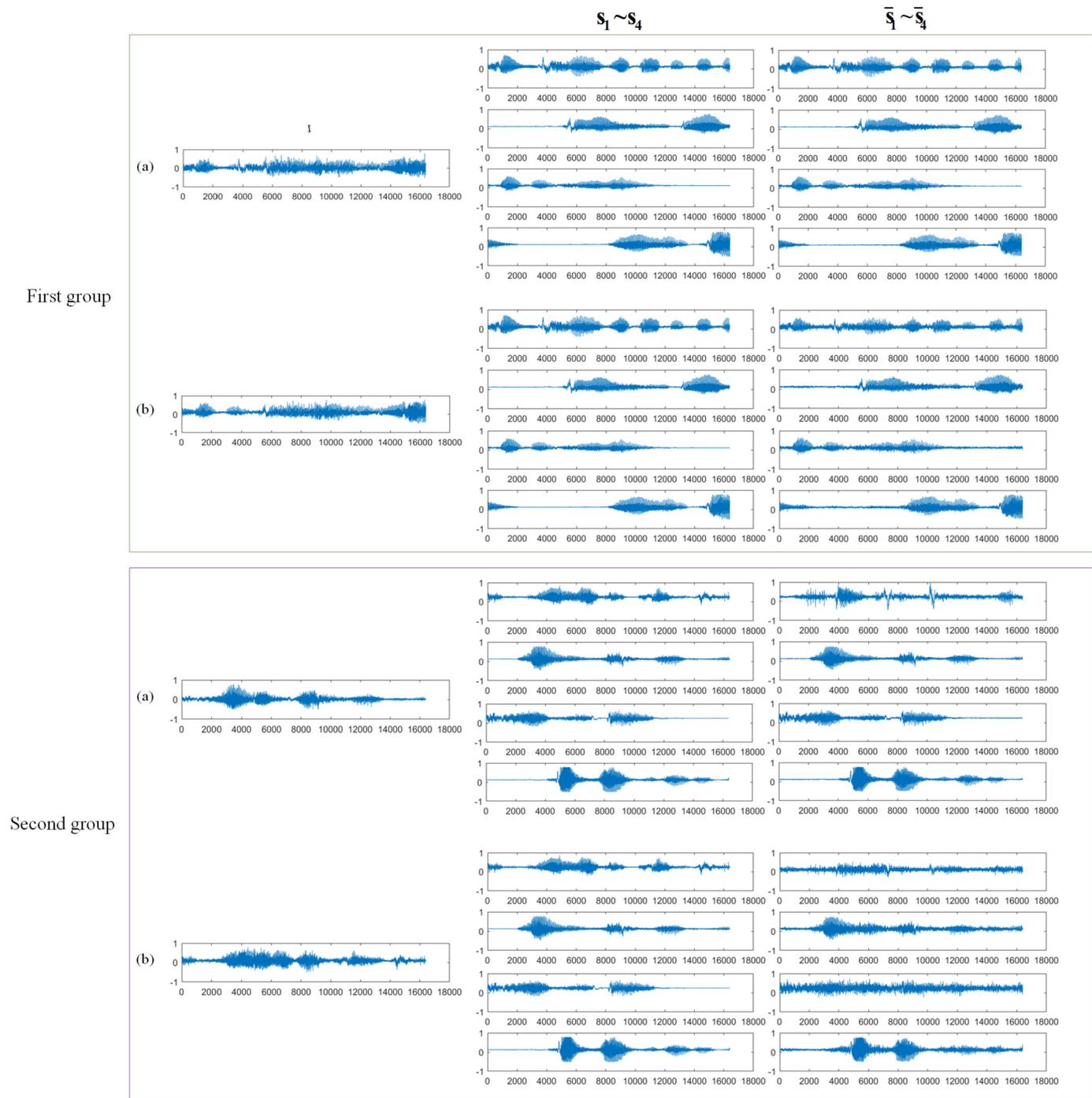
Fig. 11. Testing results of the PDualGAN approach for convolutive mixing speech signals. Each group includes two different mixtures ((a), (b)) of four signals of different random weights. The first column is the mixtures, the second column is the corresponding original sources s_1, s_2, s_3, and s_4, and the third column is the corresponding estimated sources $\bar{s}_1$, $\bar{s}_2$, $\bar{s}_3$, and $\bar{s}_4$.

Table 11
The average correlation between s_i and $\bar{s}_i$ of the proposed PDualGAN compared with the state-of-the-art convolutive mixing model based algorithms.

Algorithm	First group			Second group			Third group		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
CG	0.24	0.25	0.24	0.11	0.16	0.25	0.27	0.25	0.22
IRLS	0.15	0.27	0.24	0.13	0.21	0.15	0.23	0.26	0.22
CNMF	0.28	0.35	0.23	0.31	0.38	0.33	0.27	0.25	0.26
E-S-D	0.54	0.62	0.59	0.64	0.59	0.68	0.58	0.49	0.66
SCBDC	0.64	0.62	0.59	0.57	0.62	0.58	0.65	0.62	0.63
PDualGAN	0.78	0.77	0.75	0.69	0.68	0.55	0.99	0.97	0.97

(a) Average SDR varying with SNR.

(b) Average SIR varying with SNR.

(c) Average correlation varying with SNR.
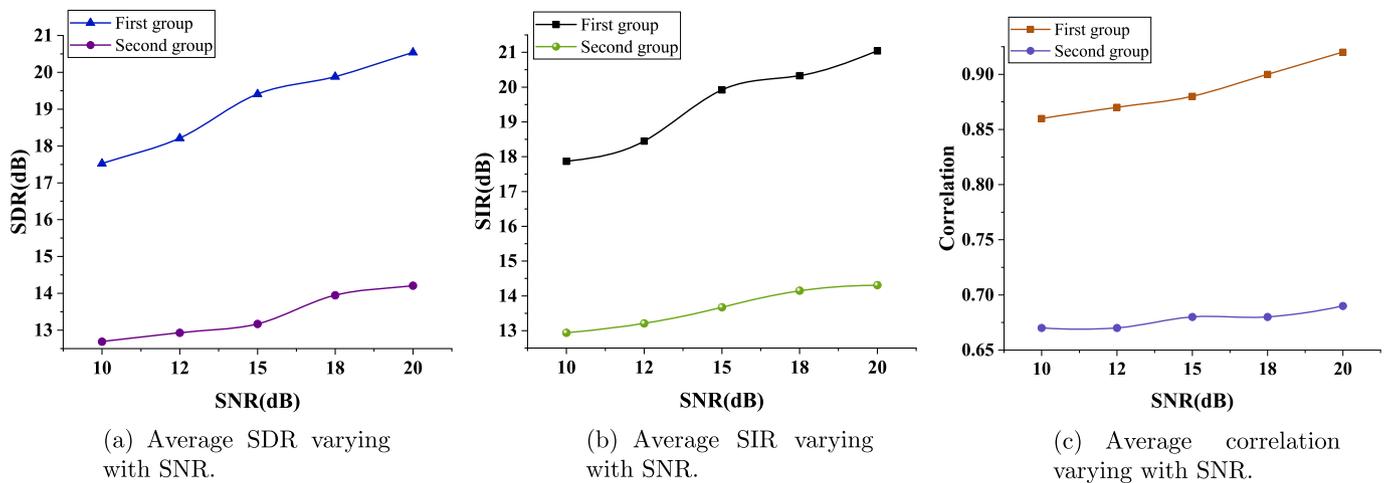
Fig. 12. Average SDR, SIR and correlation of the proposed PDualGAN algorithm varying with SNR.

Table 12
The results of average, standard deviation and confidence interval of PSNR of the convolutive mixing model based algorithms on the ACC-occlusion dataset.

Algorithm	Average PSNR	Standard deviation	Confidence interval
CG	5.02	0.22	[4.98, 5.06]
IRLS	4.51	0.77	[4.36, 4.66]
CNMF	7.12	0.72	[6.97, 7.26]
E-S-D	13.37	0.51	[13.27, 13.46]
SCBDC	14.51	0.32	[14.45, 14.57]
PDualGAN	21.13	12.43	[18.70, 23.57]

Obviously, the average PSNR of the proposed algorithm can reach 25.74 dB, outperforming the baseline algorithms CG, IRLS, CNMF, E-S-D and SCBDC of 4.84 dB, 4.67 dB, 7.87 dB, 13.72 dB and 14.32 dB, and the average SSIM can achieve 0.77 which is significantly higher than these algorithms of 0.22, 0.22, 0.25, 0.57, and 0.60. Table 11 demonstrates the average correlation that the PDualGAN algorithm can achieve 0.79 which outperforms the baseline algorithms of 0.22, 0.21, 0.30, 0.59, and 0.61.

Similar to section 5.1, the results of average PSNR, SSIM and correlation of the proposed PDualGAN algorithm measured with the source-to-noise ratio (SNR) are shown in Fig. 9.

We calculated the average, standard deviation and confidence interval of PSNR on the ACC-occlusion dataset, the results of the state-of-the-art convolutive mixing model based algorithms are shown in Table 12.

5.3. PDualGAN for speech signals

In the third set of simulations, we evaluate the separation performance of the proposed PDualGAN algorithm described in Algorithm 1 for convolutive mixtures on the Speech datasets.

As shown in Fig. 11, two groups of the convolutive mixtures are selected from the Speech datasets, and each group shows the separating results from two different mixtures ((a), (b) in Fig. 11), and each mixture is a convolutive mixture generated by four speech signals with different random weights. Fig. 10 (a) and Fig. 10 (b) show the comparison results of the SDR and SIR obtained by the PDualGAN, CG, IRLS, CNMF, E-S-D, SCBDC and E-MRP-CNN algorithms according to Fig. 11. Table 13 demonstrates the average correlation of these state-of-the-art algorithms.

The performance of CG, IRLS, and CNMF algorithms is relatively limited which can obtain the average SDR of 3.61 dB, 3.56 dB, and 4.13 dB, the SIR of 3.63 dB, 3.58 dB, and 4.17 dB, and the correlation of 0.25, 0.23, and 0.30. The performance of the E-S-D, SCBDC and E-MRP-CNN algorithms performs better and has achieved an average SDR of 8.38 dB, 8.60 dB and 14.11 dB, the SIR of 8.54 dB,

Table 13
The average correlation of the proposed PDualGAN compared with the state-of-the-art convolutive mixing algorithms.

Algorithm	First group		Second group	
	(a) s_i and $\bar{s}_i$	(b) s_i and $\bar{s}_i$	(a) s_i and $\bar{s}_i$	(b) s_i and $\bar{s}_i$
CG	0.37	0.07	0.29	0.27
IRLS	0.26	0.20	0.22	0.25
CNMF	0.34	0.30	0.28	0.27
E-S-D	0.55	0.60	0.56	0.53
SCBDC	0.63	0.59	0.58	0.55
E-MRP-CNN	0.67	0.52	0.58	0.77
PDualGAN	0.94	0.96	0.71	0.70

Table 14
The results of average, standard deviation and confidence interval of SDR of the convolutive mixing model based algorithms on the Speech dataset.

Algorithm	Average SDR	Standard deviation	Confidence interval
CG	3.52	1.78	[3.17, 3.87]
IRLS	2.76	0.39	[2.68, 2.84]
CNMF	3.88	0.08	[3.87, 3.90]
E-S-D	8.12	0.88	[7.94, 9.29]
SCBDC	8.62	0.55	[8.51, 8.73]
E-MRP-CNN	8.76	0.22	[8.72, 8.80]
PDualGAN	14.93	9.99	[12.97, 16.89]

8.66 dB and 10.58 dB, and the correlation of 0.57, 0.59 and 0.65, but are still lower than the proposed PDualGAN algorithm which can reach an SDR, SIR and correlation of 20.33 dB, 20.83 dB and 0.82.

Fig. 12 shows the performance of average SDR, SIR and correlation of the proposed PDualGAN measured with the source-to-noise ratio (SNR).

Similarly, we calculated the average value, standard deviation and confidence interval of SDR and SIR on the Speech dataset, the results of the state-of-the-art convolutive mixing model based algorithms are shown in Table 14 and Table 15.

5.4. PDualGAN for EEG signals

In the fourth set of simulations, we evaluate the separation performance of the proposed PDualGAN algorithm described in Algorithm 1 for instantaneous mixtures on the EEG signal datasets.

Fig. 13 shows part of the test results for instantaneous mixing signals, two groups are randomly selected, each group shows the separating results from two different mixtures ((a), (b) in Fig. 13), and each mixture is generated by four EEG signals with different random weights. Similarly, we compared the PDualGAN algorithm
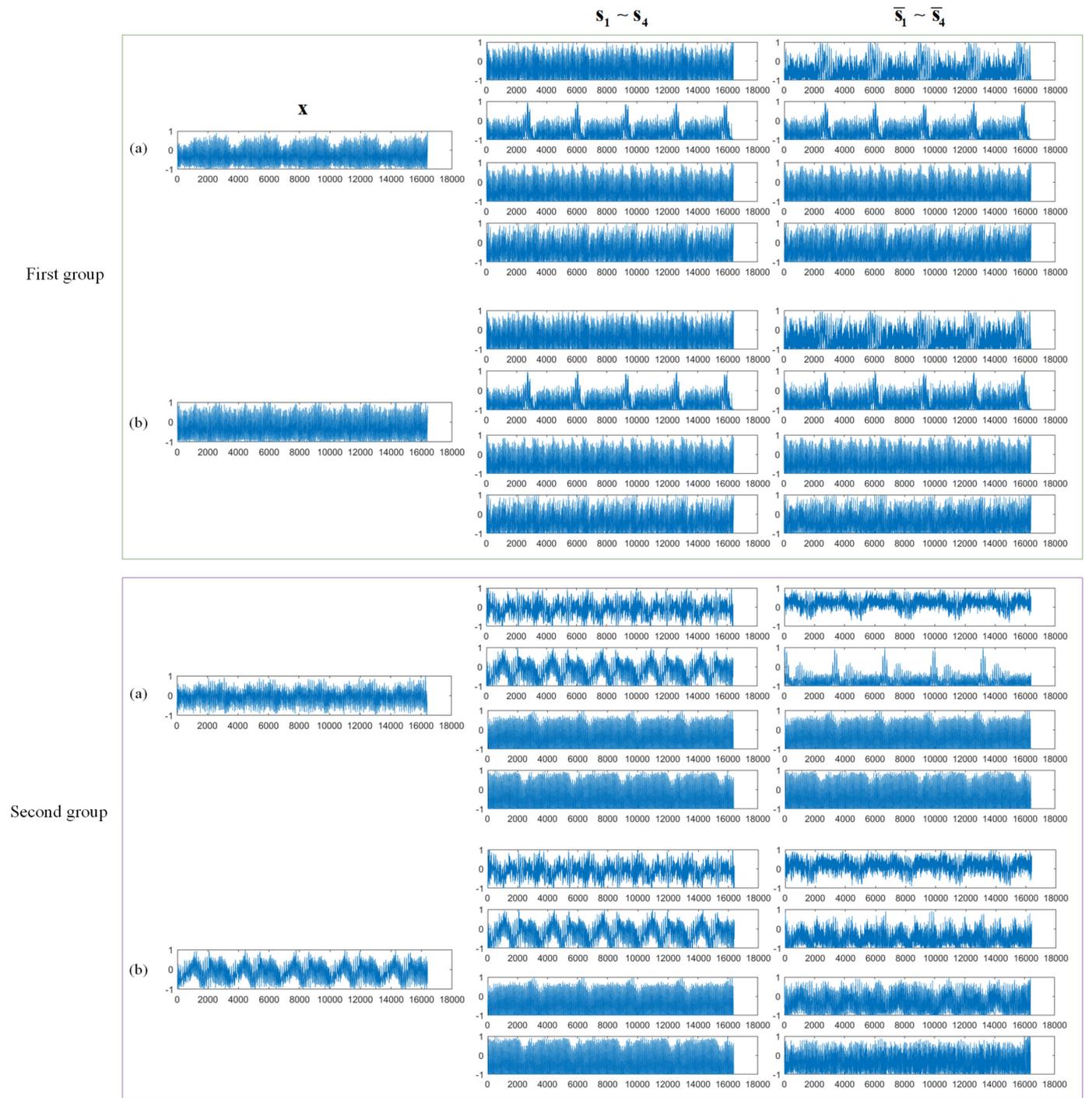
Fig. 13. Testing results of the PDualGAN approach for instantaneous mixing EEG signals. Each group includes two different mixtures((a), (b)) of four signals of different random weights. The first column corresponds to the mixtures, the second column shows the corresponding original sources s_1, s_2, s_3, and s_4, and the third column shows the corresponding estimated sources $\bar{s}_1$, $\bar{s}_2$, $\bar{s}_3$, and $\bar{s}_4$.

Table 15
The results of average, standard deviation and confidence interval of SIR of the convolutive mixing model based algorithms on the Speech dataset.

Algorithm	Average SIR	Standard deviation	Confidence interval
CG	3.71	1.90	[3.34, 4.08]
IRLS	2.60	0.50	[2.50, 2.69]
CNMF	3.88	0.69	[3.75, 4.02]
E-S-D	8.50	0.81	[8.34, 8.65]
SCBDC	9.04	0.66	[8.91, 9.16]
E-MRP-CNN	9.41	0.66	[9.28, 9.54]
PDualGAN	15.31	8.60	[13.63, 16.99]

with ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA and Gaussian WGAN algorithms according to Fig. 13. Fig. 14 and Table 16 demonstrates the comparison results of the average RRMSE and correlation obtained by these state-of-the-art algorithms.

As observed, the proposed PDualGAN algorithm has a lower RRMSE of 40.55% and a higher correlation of 0.70 which outperforms the baseline algorithms ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA, and Gaussian WGAN with an RRMSE of 69.22%, 70.63%, 69.84%, 68.52%, 69.23%, 68.27%, 59.69%, a correlation of 0.22, 0.21, 0.15, 0.21, 0.19, 0.22 and 0.46.
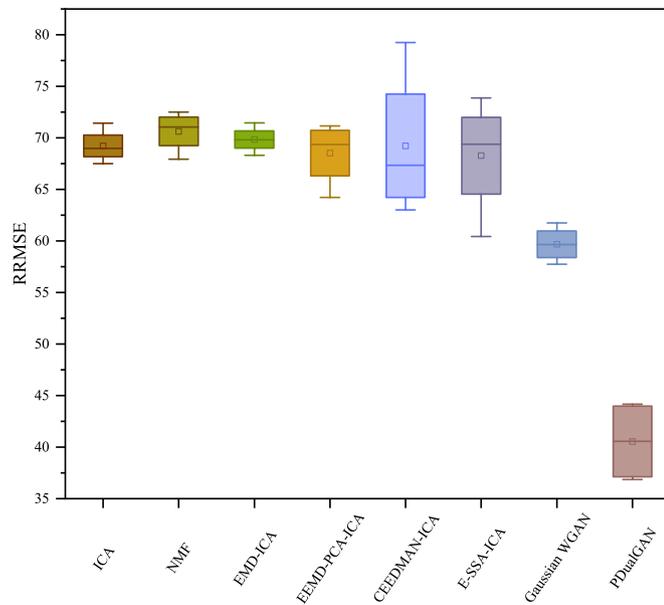
Fig. 14. Average RRMSE of the proposed PDualGAN compared with the state-of-the-art instantaneous mixing algorithms.

Table 16
The average correlation of the proposed PDualGAN compared with the state-of-the-art instantaneous mixing algorithms.

Algorithm	First group		Second group	
	(a) s_i and $\bar{s}_i$	(b) s_i and $\bar{s}_i$	(a) s_i and $\bar{s}_i$	(b) s_i and $\bar{s}_i$
ICA	0.31	0.24	0.18	0.16
NMF	0.27	0.24	0.19	0.17
EMD-ICA	0.18	0.12	0.16	0.13
EEMD-PCA-ICA	0.16	0.24	0.36	0.09
CEEDMAN-ICA	0.29	0.08	0.23	0.15
E-SSA-ICA	0.31	0.16	0.26	0.14
Gaussian WGAN	0.48	0.52	0.43	0.40
PDualGAN	0.74	0.72	0.68	0.66

Table 17
The results of average, standard deviation and confidence interval of RRMSE of the instantaneous mixing model based algorithms on the EEG dataset.

Algorithm	Average RRMSE	Standard deviation	Confidence interval
ICA	67.20	12.67	[64.71, 69.68]
NMF	68.81	8.82	[67.08, 70.54]
EMD-ICA	71.56	18.07	[68.02, 75.10]
EEMD-PCA-ICA	71.43	25.76	[66.38, 76.48]
CEEDMAN-ICA	67.45	25.09	[62.53, 72.37]
E-SSA-ICA	69.87	26.25	[64.73, 75.02]
Gaussian WGAN	61.87	29.05	[56.18, 67.56]
PDualGAN	43.79	25.06	[38.88, 48.70]

Fig. 15 shows the performance of average RRMSE and correlation of the proposed PDualGAN measured with the source-to-noise ratio (SNR).

Similar to section 5.1, we calculated the average value, standard deviation and confidence interval of RRMSE on the EEG dataset, the comparison results of the state-of-the-art instantaneous mixing model based algorithms are shown in Table 17.

The experimental results show that the proposed algorithm can obtain higher PSNR, SSIM, SDR, SIR and correlation and lower RRMSE when the corresponding sources can be estimated correctly. The comparison shows that the PDualGAN algorithm outperforms the state-of-the-art algorithms on the four different datasets for the instantaneous mixtures and the convolutive mixtures, which shows the effectiveness of the proposed algorithm for both one-dimensional and two-dimensional signals.

5.5. Minimum sample size

As we know, training on datasets of smaller size while maintaining nearly the same performance would be very beneficial, so we perform an experiment to check the minimum amount of sample of the proposed PDualGAN for separate four sources on the NWPU-occlusion dataset as an example. We still randomly select 80% mixed signals and the corresponding original signals for training and the remaining 20% mixed signals with different weights for testing. Fig. 16 shows the testing results of PSNR, SSIM and correlation varying with different amount of sample.

As demonstrated, when the amount of sample is lower than 200, the value of PSNR, SSIM and correlation increase gradually, however, with the amount of sample is more than 200, the performance of the proposed PDualGAN is stable which indicate that the data is adequately trained and the PDualGAN can work properly for separating four sources from the mixtures with different weights.

6. Conclusion

In this paper, a new algorithm for the problem of single channel blind source separation (SCBSS) has been presented. Our contributions are as follows:

Model. We have formulated a unified model for the instantaneous mixing model and convolutive mixing model for the research.

Algorithm. Based on the instantaneous mixing model and convolutive mixing model, we proposed a PDualGAN algorithm. The N DualGANs are trained simultaneously with mixtures and corresponding sources to realize one-to-multiple mapping, and the Wasserstein generative adversarial networks gradient penalty (WGAN-GP) loss function is applied in the network.

The proposed algorithm can be used to both one-dimensional and two-dimensional signals, and different mixtures are applied to test the effectiveness and generalization performance.

Datasets. We build the one-to-multiple datasets in the experiment which are composed of two parts: the mixtures and corresponding sources. Each mixture is generated by using randomly generated mixing matrices and multiple sources with different weights. These datasets can be used for related research.

Numerical experiments show that the proposed PDualGAN algorithm performs well in separating four sources from instantaneous mixtures and convolutive mixtures which outperforms the ICA, NMF, EMD-ICA, EEMD-PCA-ICA, CEEDMAN-ICA, E-SSA-ICA, Gaussian WGAN, CG, IRLS, CNMF, E-S-D, SCBDC and E-MRP-CNN algorithms. However, with the increase in the number of sources, the performance of the algorithm will decline. In the future, we will consider improving the performance of separating multiple sources.

CRediT authorship contribution statement

Ting Liu: Conceptualization, Data curation, Methodology, Software, Visualization, Writing – original draft. **Wen-wu Wang:** Conceptualization, Writing – review & editing. **Xiao-fei Zhang:** Investigation, Validation. **Yi-na Guo:** Funding acquisition, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
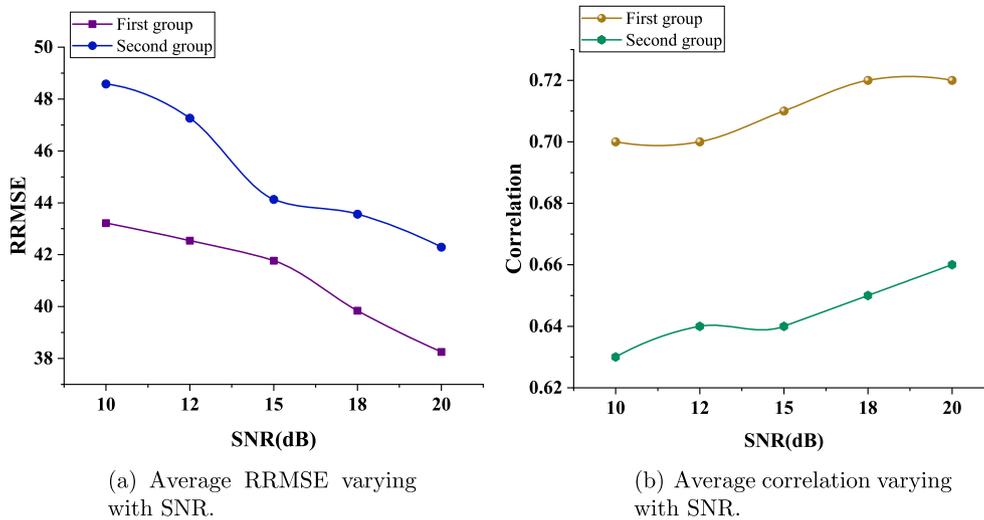
(a) Average RRMSE varying with SNR.

(b) Average correlation varying with SNR.

Fig. 15. Average RRMSE and correlation of the proposed PDualGAN algorithm varying with SNR.

(a) Average PSNR varying with amount of sample.

(b) Average SSIM varying with amount of sample.

(c) Average correlation varying with amount of sample.
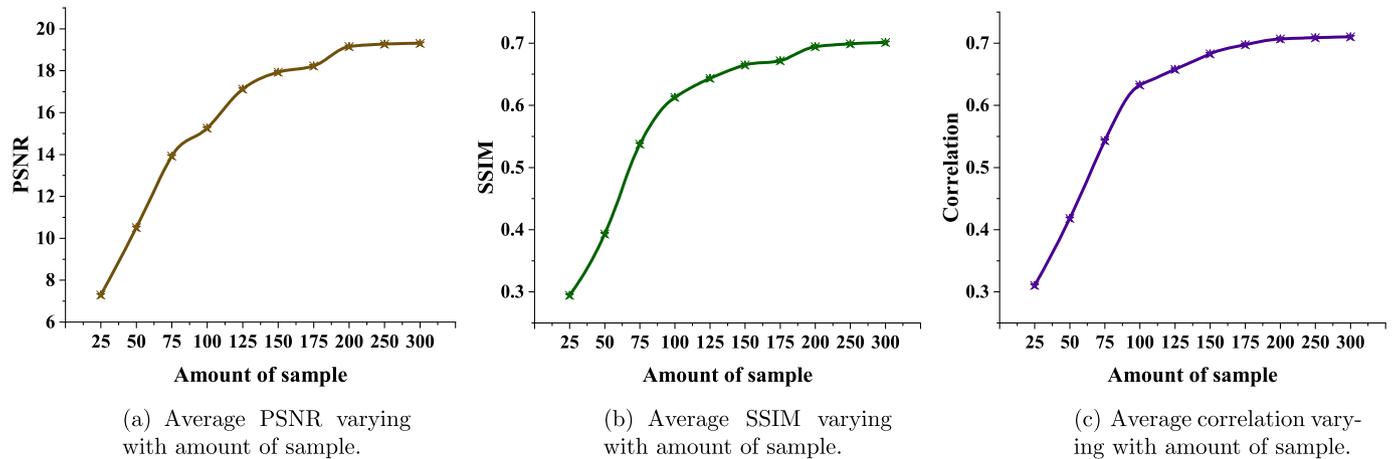
Fig. 16. Average PSNR, SSIM and correlation of the proposed PDualGAN algorithm varying with the amount of sample on the NWPU-occlusion dataset.

Acknowledgment

References

[1] P.K. Krishna, K. Ramaswamy, Single channel speech separation based on empirical mode decomposition and Hilbert transform, IET Signal Process. 11 (5) (2017) 579–586.

[2] F. Fan, K. Yang, M. Xia, W. Li, B. Fu, W. Zhang, Underwater image restoration by means of blind deconvolution approach, Front. Optoelectron. China 3 (2) (2010) 169–178.

[3] C.J. James, S. Wang, Blind source separation in single-channel EEG analysis: an application to BCI, in: Annual International Conference of the IEEE Engineering in Medicine and Biology, New York, NY, USA, 2006, pp. 6544–6547.

[4] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Z. Yu, Size-scalable content-based histopathological image retrieval from database that consists of WSIs, IEEE J. Biomed. Health Inform. 22 (4) (2018) 1278–1287.

[5] W.P. Segars, B. Tsui, J. Cai, F.F. Yin, G. Fung, E. Samei, Application of the 4-D XCAT phantoms in biomedical imaging and beyond, IEEE Trans. Med. Imaging 37 (3) (2018) 680–692.

[6] D.V. Sorokin, I. Peterlik, M. Tektonidis, K. Rohr, P. Matula, Non-rigid contour-based registration of cell nuclei in 2-D live cell microscopy images using a dynamic elasticity model, IEEE Trans. Med. Imaging 37 (1) (2018) 173–184.

[7] Y. Guo, A. Wang, W. Wang, Multi-source phase retrieval from multi-channel phaseless STFT measurements, Signal Process. 144 (2018) 36–40.

[8] Y. Guo, T. Wang, J. Li, A. Wang, W. Wang, Multiple input single output phase retrieval, Circuits Syst. Signal Process. 38 (8) (2019) 3818–3840.

[9] Y. Guo, X. Zhao, J. Li, A. Wang, W. Wang, Blind multiple input multiple output image phase retrieval, IEEE Trans. Ind. Electron. 67 (3) (2020) 2220–2230.

[10] B. Gao, L. Bai, W.L. Woo, G.Y. Tian, Y. Cheng, Automatic defect identification of Eddy current pulsed thermography using single channel blind source separation, IEEE Trans. Instrum. Meas. 63 (4) (2014) 913–922.

[11] M.P. Hobson, A.W. Jones, A.N. Lasenby, F.R. Bouchet, Foreground separation methods for satellite observations of the cosmic microwave background, Mon. Not. R. Astron. Soc. 300 (1) (1998) 1–29.

[12] L.S. Zhan, D.S. Huang, C.H. Zheng, L. Shang, Blind inversion of Wiener system for single source using nonlinear blind source separation, in: Proceedings of the International Joint Conference on Neural Networks, Piscataway, NJ, USA, 2005, pp. 1235–1238.

[13] H. Maki, T. Toda, S. Sakti, G. Neubig, S. Nakamura, EEG signal enhancement using multi-channel Wiener filter with a spatial correlation prior, in: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), South Brisbane, QLD, Australia, 2015, pp. 2639–2643.

[14] R. Sharma, S.M. Prasanna, A better decomposition of speech obtained using modified empirical mode decomposition, Digit. Signal Process. 58 (2016) 26–39.

[15] B. Mijovic, M. Vos, I. Gligorijevic, J. Taelman, S. Huffel, Source separation from single-channel recordings by combining empirical-mode decomposition

and independent component analysis, IEEE Trans. Biomed. Eng. 57 (9) (2010) 2188–2196.

[16] Y. Guo, S. Huang, Y. Li, Single-mixture source separation using dimensionality reduction of ensemble empirical mode decomposition and independent component analysis, Circuits Syst. Signal Process. 31 (6) (2012) 2047–2060.

[17] Y. Guo, S. Huang, Y. Li, G.R. Naik, Edge effect elimination in single-mixture blind source separation, Circuits Syst. Signal Process. 32 (5) (2013) 2317–2334.

[18] D.D. Lee, H.S. Seung, Learning the parts of objects by non-negative matrix factorization, Nature 401 (6755) (1999) 788–791.

[19] B. Gao, W.L. Woo, S.S. Dlay, Single-channel source separation using EMD-subband variable regularized sparse features, IEEE Trans. Audio Speech Lang. Process. 19 (4) (2011) 961–976.

[20] Y.C. Subakan, P. Smaragdis, Generative adversarial source separation, arXiv: 1710.10779, 2017.

[21] Q. Kong, Y. Xu, W. Wang, P. Jackson, M. Plumbley, Single-channel signal separation and deconvolution with generative adversarial networks, in: Proceedings of the 28th International Joint Conference on Artificial Intelligence, Macao, China, 2019, pp. 2747–2753.

[22] C. Jutten, J. Herault, Blind separation of sources, part i: an adaptive algorithm based on neuromimetic architecture, Signal Process. 24 (1) (1991) 1–10.

[23] Z. Yi, H. Zhang, P. Tan, M. Gong, DualGAN: unsupervised dual learning for image-to-image translation, in: IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2868–2876.

[24] R.G. Newcombe, Two-sided confidence intervals for the single proportion: comparison of seven methods, Stat. Med. 17 (8) (1998) 857–872.

[25] A.K. Maddirala, R.A. Shaik, Separation of sources from single-channel EEG signals using independent component analysis, IEEE Trans. Instrum. Meas. 67 (2) (2017) 382–393.

[26] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, Generative adversarial nets, in: Advances in Neural Information Processing Systems (NIPS), Montreal, Quebec, Canada, 2014, pp. 2672–2680.

[27] D. He, Y. Xia, T. Qin, L. Wang, N. Yu, T.Y. Liu, W.Y. Ma, Dual learning for machine translation, in: Advances in Neural Information Processing Systems (NIPS), Barcelona, Spain, 2016, pp. 820–828.

[28] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein gan, arXiv:1701.07875, 2017.

[29] P. Isola, J.Y. Zhu, T. Zhou, A.A. Efros, Image-to-image translation with conditional adversarial networks, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5967–5976.

[30] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 2015, pp. 234–241.

[31] M. Liu, T. Breuel, J. Kautz, Unsupervised image-to-image translation networks, in: Advances in Neural Information Processing Systems, 2017, pp. 700–708.

[32] Z.Z. Luo, Z.H. Yan, W.D. Fu, Electroencephalogram artifact filtering method of single channel EEG based on CEEMDAN-ICA, Chin. J. Sens. Actuators 31 (8) (2018) 1211–1216.

[33] S. Jabbari, Source separation from single-channel abdominal phonocardiographic signals based on independent component analysis, Biomed. Eng. Lett. 11 (1) (2021) 55–67.

[34] Q. Huynh-Thu, M. Ghanbari, Scope of validity of psnr in image/video quality assessment, Electron. Lett. 44 (13) (2008) 800–801.

[35] D.R. Fuhrmann, G.S. Antonio, Transmit beamforming for MIMO radar systems using signal cross-correlation, IEEE Trans. Aerosp. Electron. Syst. 44 (1) (2008) 1–16.

[36] A. Levin, R. Fergus, F. Durand, W.T. Freeman, Deconvolution using natural image priors, ACM Trans. Graph. 26 (2007).

[37] E.M. Grais, M.D. Plumbley, Single channel audio source separation using convolutional denoising autoencoders, in: IEEE Global Conference on Signal and Information Processing, Montreal, QC, Canada, 2017, pp. 1265–1269.

[38] Z. Fan, Y. Lai, J.R. Jang, SVSGAN: singing voice separation via generative adversarial network, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Calgary, AB, Canada, 2018, pp. 726–730.

[39] Y. Sun, W. Wang, J. Chambers, S.M. Naqvi, Two-stage monaural source separation in reverberant room environments using deep neural networks, IEEE/ACM Trans. Audio Speech Lang. Process. 27 (1) (2019) 125–139.

[40] W. Yuan, B. Dong, S. Wang, M. Unoki, W. Wang, Evolving multi-resolution pooling CNN for monaural singing voice separation, IEEE/ACM Trans. Audio Speech Lang. Process. 29 (2021) 807–822.

[41] J.E. Huang, D. Xu, Y.L. Wang, Y. Zhang, S. Wang, The research on the methodology of diagnosing the fault of bearing in warships based on NS-EMD, in: Advanced Materials Research, Hong Kong, China, 2013, pp. 352–357.

[42] R. Jenssen, T. Eltoft, Independent component analysis for texture segmentation, Pattern Recognit. 36 (10) (2003) 2301–2315.

[43] F.N. Yuan, L. Zhang, J.T. Shi, X. Xia, G. Li, Theories and applications of auto-encoder neural networks: a literature survey, Chinese J. Comput. 42 (1) (2019) 203–230.

[44] L. Cohen, What is a multicomponent signal, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), San Francisco, CA, USA, 1992, pp. 113–116.

[45] J.R. Hopgood, P.J. Rayner, Single channel separation using linear time varying filters: separability of non-stationary stochastic signals, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Phoenix, AZ, USA, 1999, pp. 1449–1452.

[46] A. Hyvarinen, E. Oja, Independent component analysis: algorithms and applications, Neural Netw. 13 (4) (2000) 411–430.

[47] T. Liu, T. Yin, Z. Gong, Y. Guo, Research on single-channel blind deconvolution algorithm for multi-source signals, J. Electron. Inf. Technol. 44 (1) (2022) 230–236.

[48] A. Mahajan, G. Birajdar, Analysis of blind separation of noisy mixed images based on wavelet thresholding and independent component analysis, Int. J. Eng. Technol. 3 (5) (2011) 560–564.

[49] E. Vincent, R. Gribonval, C. Fevotte, Performance measurement in blind audio source separation, IEEE Trans. Audio Speech Lang. Process. 14 (4) (2006) 1462–1469.

Ting Liu was born in 1991. She received the M.S. degree in computer technology from Taiyuan University of Science and Technology, Taiyuan, China, in 2017. She is currently working toward the Ph.D. degree in Control Science and Engineering with the Taiyuan University of Science and Technology, Taiyuan. Her research interests include advanced control theory and signal processing.

Wenwu Wang received the Ph.D. degree in navigation, guidance, and control from Harbin Engineering University in 2002. He is currently a Reader in Signal Processing with the University of Surrey, Surrey, U.K., and a Co-Director of the Machine Audition Laboratory within the Centre for Vision, Speech and Signal Processing. He has coauthored more than 250 publications. His research including signal processing and machine learning.

He has been a Senior Area Editor and an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is a Publication Co-Chair for the International Conference on Acoustics, Speech, and Signal Processing 2019, Brighton, U.K.

Xiaofei Zhang received the B.E. degree in Electronic Information Engineering from Hebei University of Science and Technology, Hebei, China, in 2018. He is currently working toward the master's degree in electronics and communication engineering with the Taiyuan University of Science and Technology, Taiyuan. His research interests include Brain-computer interface and intelligent information processing and pattern recognition.

Yina Guo 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. She is currently a Professor with Taiyuan University of Science and Technology, China. She has authored/coauthored more than 40 papers and two books, and was granted six patents and four software copyrights in China. Her research interests include blind source separation, bio-signal processing and phase retrieval.

She is a Senior Member of China Institute of Communications and a Member of IEEE Signal Processing Society.