Mathematical Biosciences and Engineering

MBE, 20(11): 19661–19685.
DOI: 10.3934/mbe.2023871
Received: 28 August 2023
Revised: 08 October 2023
Accepted: 16 October 2023
Published: 26 October 2023

http://www.aimspress.com/journal/MBE

Research article

ARU-DGAN: A dual generative adversarial network based on attention residual U-Net for magneto-acousto-electrical image denoising

Shuaiyu Bu[1,2,3], Yuanyuan Li[1,2,*], Wenting Ren[4] and Guoqiang Liu[1,2]

[1] Institute of Electrical Engineering, Chinese Academy of Sciences, Beijing 100190, China
[2] School of Electronic Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
[3] State Grid Beijing Electric Power Company, Beijing 100031, China
[4] Department of Radiation Oncology, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100730, China

* **Correspondence:** Email: lyy@mail.iee.ac.cn.

Abstract: Magneto-Acousto-Electrical Tomography (MAET) is a multi-physics coupling imaging modality that integrates the high resolution of ultrasound imaging with the high contrast of electrical impedance imaging. However, the quality of images obtained through this imaging technique can be easily compromised by environmental or experimental noise, thereby affecting the overall quality of the imaging results. Existing methods for magneto-acousto-electrical image denoising lack the capability to model local and global features of magneto-acousto-electrical images and are unable to extract the most relevant multi-scale contextual information to model the joint distribution of clean images and noise images. To address this issue, we propose a Dual Generative Adversarial Network based on Attention Residual U-Net (ARU-DGAN) for magneto-acousto-electrical image denoising. Specifically, our model approximates the joint distribution of magneto-acousto-electrical clean and noisy images from two perspectives: noise removal and noise generation. First, it transforms noisy images into clean ones through a denoiser; second, it converts clean images into noisy ones via a generator. Simultaneously, we design an Attention Residual U-Net (ARU) to serve as the backbone of the denoiser and generator in the Dual Generative Adversarial Network (DGAN). The ARU network adopts a residual mechanism and introduces a linear Self-Attention based on Cross-Normalization (CNorm-SA), which is proposed in this paper. This design allows the model to effectively extract the most relevant multi-scale contextual information while maintaining high resolution, thereby better modeling the local and global features of magneto-acousto-electrical images. Finally, extensive

experiments on a real-world magneto-acousto-electrical image dataset constructed in this paper demonstrate significant improvements in preserving image details achieved by ARU-DGAN. Furthermore, compared to the state-of-the-art competitive methods, it exhibits a 0.3 dB increase in PSNR and an improvement of 0.47% in SSIM.

Keywords: magneto-acousto-electrical tomography; image denoising; generative adversarial networks; attention mechanism

1. Introduction

Magneto-Acousto-Electrical Tomography (MAET) integrates the advantages of ultrasound imaging and electrical impedance imaging, representing a promising electro-characteristic imaging technique with broad clinical application prospects. It has demonstrated significant potential, particularly in early diagnosis of diseases such as breast cancer and liver cancer [1]. However, as a complex coupled imaging method, effectively mitigating the impact of noise on MAET images is undoubtedly a critical challenge. Interference generated by experimental equipment and environmental white noise can potentially degrade image quality [2]. Crucially, the electrical conductivity of normal biological tissue is as low as 0.2 *S/m* or even lower. To further detect signals equivalent to the conductivity of biological tissues, it is necessary to reduce the conductivity of the tissues being examined. However, when the conductivity of these tissues decreases, the signal becomes more susceptible to environmental or experimental noise. This underscores the pressing need to enhance image quality against a backdrop of weak electrical signals. Therefore, effective denoising of magneto-acousto-electrical images not only enhances image quality but also more accurately reflects the actual conditions within the patient's body, which is of significant importance for improving the application effectiveness of MAET technology in the medical field. We will explore the feasibility and efficacy of using deep learning algorithms for magneto-acousto-electrical image denoising, with the aim of providing new perspectives and methods for current medical imaging research.

In recent years, numerous scholars have dedicated their efforts to enhancing the quality of magneto-acousto-electrical images. In previous studies, researchers have attempted to improve image quality through various methods such as adjusting the signal pattern and energy of the ultrasonic excitation source [3–10], and enhancing the intensity of the static magnetic field [11–16], among others. Additionally, some studies have focused on the characteristics of MAET signals, adopting wavelet filtering methods to enhance the signal-to-noise ratio and utilizing the reciprocity theorem to derive the fluctuation equation of the detection signal. This equation, combined with the time reversal imaging algorithm, allows for more precise imaging of targets with low conductivity using the filtered signals [17]. However, despite some progress, current denoising methods still exhibit significant limitations due to the unique properties of magneto-acousto-electrical images, such as their weak electrical signals and complex noise environments. Specifically, existing denoising techniques lack effective modeling capabilities for local and global features of magneto-acousto-electrical images and are unable to extract the most relevant multi-scale contextual information to model the joint distribution of clean and noisy image pairs. However, by constructing such a joint distribution, the model can more accurately understand and depict the noise patterns in the image, enabling the denoising algorithm to more effectively distinguish between noise and useful signals. Moreover, this approach can capture the

characteristics and variations of the image at different scales, which is crucial for handling complex noise environments and preserving detailed information in the image. Therefore, finding more effective denoising strategies to further enhance the quality of magneto-acousto-electrical images remains an important goal in current research.

In the field of image denoising, methods based on deep learning have achieved significant accomplishments in recent years [18–20]. However, training these deep learning models requires a large number of pairs of clean and noisy images, the collection of which is both time-consuming and costly. To address this issue, researchers have proposed various noise generation techniques to simulate more pairs of clean and noisy images, thereby facilitating the training of deep learning models. The core idea of these methods primarily involves utilizing Generative Adversarial Networks (GAN) [21] to directly learn the noise distribution $f(v)$ [22,23]. However, these methods face challenges when simulating complex noise images, where the noise level and characteristics vary with the signal and may differ across multiple dimensions.

To address these issues, we propose a Dual Generative Adversarial Network Based on Attention Residual U-Net (ARU-DGAN), which has been successfully applied to magneto-acousto-electrical image denoising. By approximating the joint distribution of magneto-acousto-electrical clean and noisy images from both the perspectives of noise removal and noise generation, we propose an effective method that leverages attention residual U-Net to learn image information and extract the most relevant multi-scale contextual information. The major contributions of this paper are as follows:

1) We propose a novel Dual Generative Adversarial Network (DGAN) architecture that allows the model to approximate the joint distribution $f(u, v)$ of magneto-acousto-electrical clean and noisy images from two perspectives: Noise removal and noise generation, within a unified Bayesian framework as illustrated in Figure 1. More crucially, by using the additional clean and noisy image pairs simulated by our trained generator, we can expand the scale of the training set, thereby further enhancing the denoising performance of the model.

2) We design an Attention Residual U-Net (ARU) as the backbone for both the denoiser and generator within the DGAN framework. The ARU architecture employs a residual mechanism and incorporates the proposed linear Self-Attention based on Cross-Normalization (CNorm-SA). This design enables the model to maintain high resolution while efficiently extracting the most relevant multi-scale contextual information, thus enhancing its capability to model local and global features of magneto-acousto-electrical images.

3) We build a Magneto-Acousto-Electrical Imaging (MAEI) dataset, aiming to establish a standardized benchmark for the field of magneto-acousto-electrical image denoising. To validate the practical performance of ARU-DGAN, we conduct extensive experiments on the MAEI dataset. The experimental findings underscore the exceptional ability of ARU-DGAN in restoring sharp edges and detailed textures, resulting in denoised outputs that closely resemble real-world scenarios. Furthermore, ARU-DGAN outperforms the state-of-the-art competitive methods in denoising effects, with an increase of 0.3 dB in PSNR and a 0.47% improvement in SSIM.

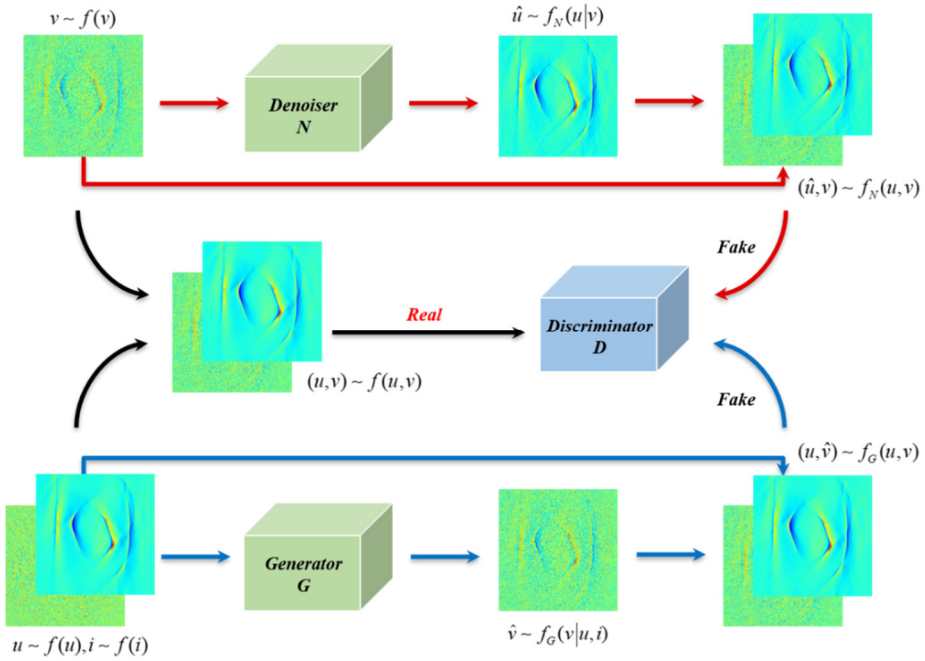

Figure 1. Overall architecture of the dual generative adversarial network. The denoiser N aims to approximate the joint distribution $\boldsymbol{f_N(u,v)}$ in Eq (2) by passing the conditional distribution $\boldsymbol{f_N(u|v)}$. Moreover, the generator G targets approximating the joint distribution $\boldsymbol{f_G(u,v)}$ in Eq (4) by passing the conditional distribution $\boldsymbol{f_G(v|u,i)}$. Through the adversarial action of the discriminator D, the denoiser N and generator G gradually optimize during the training process, making their pseudo joint distributions $\boldsymbol{f_N(u,v)}$ and $\boldsymbol{f_G(u,v)}$ progressively approach the real joint distribution $\boldsymbol{f(u,v)}$.

The remainder of this paper is organized as follows. In Section 2, a brief review of the relevant literature on magneto-acousto-electrical image processing and image denoising is provided. In Section 3, the overall architecture and individual modules of the proposed model are described in detail. Section 4 presents experiments conducted on a real-world dataset of magneto-acousto-electrical images to validate the effectiveness of the proposed model. Finally, in Section 5, we conclude the paper and discuss future directions.

2. Related work

2.1. Magneto-acousto-electrical image processing

The fundamental principle of MAET involves positioning the target object within an ultrasonic field, which induces local ions to oscillate in tandem with the propagation of ultrasonic waves. These oscillating ions are subjected to the Lorentz force within a magnetic field, leading to ion separation and the generation of a localized electric field. Concurrently, electrodes attached to the imaging body receive electrical signals, which are further utilized for the realization of electrical property imaging.

Many teams have made significant contributions to the MAET process. Initially, MAET was referred to as Hall Effect Imaging, first proposed by Han et al. [2]. Their research primarily focused on imaging the interfaces between regions of differing conductivities. Subsequently, some scholars

explored the gradient changes in conductivity by studying the correlation between ultrasound and conductivity images [24]. However, due to the interaction of multiple physical fields such as sound field, electric field and magnetic field, the signal amplitude is extremely low, only at the μV level. Moreover, interference in the experimental system and environmental noise may lead to a reduction in the signal-to-noise ratio.

In order to enhance the imaging quality of magneto-acousto-electrical images, researchers have conducted in-depth theoretical and applied explorations from multiple perspectives. These explorations include adjusting the signal pattern and energy of the ultrasonic excitation source, enhancing the intensity of the static magnetic field and employing strategies such as filtering algorithms. Specifically, the evolution of the ultrasonic excitation source includes a transition from differential frequency signal excitation [15] and linear frequency modulation [14] to the use of single pulse excitation [3,4]. This transition not only optimized the generation of ultrasound waves but also improved their propagation effects within objects. The ultrasonic transducer has also undergone a technical upgrade from planar transducers [16] to focused transducers [5–7]. This advancement allows for more precise localization and irradiation of the target area, thereby enhancing the quality of imaging. In terms of detection, we have progressed from being able to use only uniform static magnetic field stimulation to now being capable of imaging under non-uniform static magnetic field excitation [25]. As for the detection mode, researchers have shifted from using coils [8,9] to electrodes [10–12,26]. This transition has improved the efficiency of signal reception and contributed to enhancing the signal-to-noise ratio. Moreover, in accordance with the characteristics of MAET signals, researchers have also introduced wavelet filtering methods to effectively boost the signal-to-noise ratio, thereby improving image quality [17].

2.2. Image denoising

Image denoising is a significant research direction in the field of computer vision, aiming to recover clear images from those contaminated with noise. Traditional image denoising strategies largely rely on prior knowledge of the image, including but not limited to sparsity [27,28], low-rank [29], self-similarity [30,31] and smoothness [32,33]. However, these priors are severely constrained when dealing with image denoising tasks under extreme conditions, making it exceptionally challenging to denoise severely corrupted images. Furthermore, discriminative learning methods have provided new insights and directions for image denoising research. These methods mainly include Markov Random Fields (MRF) [34–36], Cascade of Shrinkage Fields (CSF) [37,38] and Trainable Nonlinear Reaction Diffusion (TNRD) [39] approaches. By learning the distribution characteristics of data, these methods attempt to establish more accurate noise models, thereby achieving more effective image denoising.

In recent years, image denoising methods based on deep learning have achieved significant breakthroughs in the field of image denoising. For instance, Jain and Seung [40] first employed a five-layer network for denoising tasks, while Burger et al. [41] used a simple Multi-Layer Perceptron (MLP) for image denoising, successfully achieving performance comparable to the BM3D algorithm. Zhang et al. [20] proposed the convolutional denoising network DnCNN, which achieved state-of-the-art performance in Gaussian denoising. In addition, numerous other network architectures have been designed and applied to image denoising tasks, including but not limited to RED [42], NLRN [43], N3Net [44], RIDNet [18], VDN [19] and DANet [45]. Moreover, in the field of image segmentation, Mu et al. [46] proposed an attention-based residual U-Net method to learn how to segment intracranial

aneurysms through various preprocessing and geometric post-processing techniques. On the other hand, Liu et al. [47] proposed a dual-branch network based on Transformer and convolution for retinal vessel segmentation in OCTA images. These image segmentation approaches provide valuable insights for the image denoising method ARU-DGAN proposed in this paper.

Moreover, for image denoising tasks affected by uncertainty and/or imprecision, some studies suggest to exploit fuzzy image preprocessors based on fuzzy logic [48]. This approach proposes a blur method based on fuzzy similarity computation, which has very low computational complexity and can effectively handle images with blurry content to approximate the real noise environment more accurately.

3. A dual generative adversarial network based on attention residual U-Net

Similar to most supervised deep learning denoising methods, our approach is constructed based on a given training dataset, which contains a large number of pairs of magneto-acousto-electrical clean and noisy images. Our learning objective is not to rigidly force the model to learn the mapping from noisy image v to clean image u, but rather to approximate the latent joint distribution $f(u,v)$ between the clean and noisy image pairs. Next, we will introduce our method from a Bayesian perspective.

3.1. Dual generative adversarial network

3.1.1. Approximate joint distribution

This section decomposes the joint distribution $f(u,v)$ of magneto-acousto-electrical clean and noisy image pairs from two different perspectives [49]. First, we approach from the perspective of noise removal, focusing on minimizing the impact of the noisy image to restore a more accurate clean image. By modeling the relationship between the noisy and clean images, we aim to enhance the performance and effectiveness of the denoiser. Second, we delve into the generation of noise, studying how to produce corresponding noisy images from clean ones. Such analysis aids in better simulating and understanding the characteristics of noise in real images. By comprehending the generation process and features of noise, we can more accurately evaluate and improve the generator.

Noise removal: In Bayesian inference, we aim to deduce the probability distribution of the corresponding clean image u through the observed magneto-acousto-electrical noise image v. The conditional distribution $f(u|v)$ represents the probability distribution of the clean image u given the noisy image v. However, since the actual distribution $f(u|v)$ is often difficult to model in practice, we have designed a denoiser N, which approximates the real distribution $f(u|v)$ by learning the mapping relationship from the magneto-acousto-electrical noise image v to the clean image u.

Through the training process, the denoiser N learns an implicit distribution $f_N(u|v)$ that approximates the real distribution $f(u|v)$ as closely as possible. This allows it to generate a corresponding clean image u when given a magneto-acousto-electrical noise image v. Therefore, the output of the denoiser N can be viewed as an image sampled from this learned implicit distribution $f_N(u|v)$.

With this understanding, we can obtain the pseudo magneto-acousto-electrical clean-noise image pair $(\hat{u},v)$ as follows:

$$\hat{u} = N(v), v \sim f(v) \Rightarrow (\hat{u},v) \tag{1}$$

This can be viewed as an instance sampled from the pseudo joint distribution $f_N(u|v)$. Clearly, the better the performance of the denoiser *N*, the higher the accuracy with which it approximates the real joint distribution $f(u|v)$.

$$f_N(u,v) = f_N(u|v)f(v) \tag{2}$$

Noise generation: In real magneto-acousto-electrical imaging systems, image noise primarily originates from interference within the experimental system and white noise in the environment. To more comprehensively describe the generation process from clean magneto-acousto-electrical images *u* to noisy images *v*, we introduce an additional latent variable *i*. This latent variable *i* represents the random noise in the magneto-acousto-electrical imaging system. Therefore, the noise generation process can be characterized by the conditional distribution $f(v|u,i)$. In this task, the role of the generator *G* is to learn the implicit distribution $f_G(v|u,i)$ to approximate the real distribution $f(v|u,i)$ as closely as possible. Hence, the output of the generator *G* can be considered as instances sampled from $f_G(v|u,i)$, that is, $G(u,i)\sim f_G(v|u,i)$. Similar to Eq (1), we can obtain the pseudo pair of magneto-acousto-electrical clean-noise image pair $(u,\hat{v})$ as follows:

$$u \sim f(u), i \sim f(i), \hat{v} = G(u,i) \Rightarrow (u,\hat{v}) \tag{3}$$

By introducing the latent variable *i* and using the generator *G* to approximate the real conditional distribution $f(v|u,i)$, we can better understand and simulate the image noise phenomenon in magneto-acousto-electrical imaging. Theoretically, the latent variable *i* can be marginalized to obtain the following pseudo joint distribution $f_G(u,v)$, which serves as an approximation to the real joint distribution $f(u,v)$:

$$f_G(u,v) = \int_i f_G(v|u,i)\, f(u)f(i)di \approx \frac{1}{M}\sum_m^M f_G(v|u,i_m)\, f(u) \tag{4}$$

As suggested in [50], we can set the number of samples *M* to 1, provided that the batch size is sufficiently large. Under this setting, the pseudo pair of clean and noisy images $(u,\hat{v})$ obtained through the generation process of Eq (3) can be approximated as a sample drawn from the pseudo joint distribution $f_G(u,v)$. Such a sampling process can effectively simulate the correlation between real images and noise, thereby providing beneficial training samples for the noise removal task. Using these samples, we can better train the generator to more accurately restore clean images and remove noise, thereby improving the model's performance and robustness.

3.1.2. Network architecture

In the preceding section, two pseudo joint distributions, namely $f_N(u,v)$ and $f_G(u,v)$, are derived from the perspectives of noise removal and noise generation. The critical question now is how to effectively train the denoiser *N* and generator *G* so that they can approximate the real joint distribution $f(u,v)$ well. We can control the training process through the sampling process defined in Eqs (1) and (3), thereby making it possible to use methods similar to GAN [21]. This approach approximates the real joint distribution by gradually optimizing the two pseudo joint distributions. Specifically, we articulate this idea as a dual generative adversarial problem inspired by Triple-GAN [51].

$$\min_{N,G}\max_{D} L_{DGAN}(N,G,D) = E_{(u,v)}[D(u,v)] - \alpha E_{(\hat{u},v)}[D(\hat{u},v)] - (1-\alpha)E_{(u,\hat{v})}[D(u,\hat{v})] \tag{5}$$

where $\hat{u} = N(v)$, $\hat{v} = G(u, i)$. D represents the discriminator, whose primary task is to distinguish between real clean-noise image pairs (u, v) and pseudo image pairs, i.e., $(\hat{u}, v)$ and $(u, \hat{v})$. The hyperparameter α is used to adjust the relative weights between the denoiser *N* and generator *G*. In DGAN, the discriminator *D* attempts to differentiate the distribution of real images from the distribution of images generated by the denoiser *N* and generator *G*. Compared to traditional distance measures such as Jensen-Shannon (JS) divergence or Kullback-Leibler (KL) divergence, the Wasserstein distance can better address the vanishing gradient problem in certain cases and it provides more stable results when there are overlapping areas between distributions [52]. Therefore, we use the Wasserstein-1 distance as the loss function to make the images generated by the denoiser *N* and generator *G* closer to the real distribution, thereby enhancing the performance and stability of DGAN.

The working mechanism of DGAN is illustrated in Figure 1. The denoiser *N* aims to approximate the joint distribution $f_N(u, v)$ in Eq (2) by passing the conditional distribution $f_N(u|v)$. Furthermore, the generator *G* targets approximating the joint distribution $f_G(u, v)$ in Eq (4) by passing the conditional distribution $f_G(v|u, i)$. Through the adversarial action of the discriminator *D*, the denoiser *N* and generator *G* gradually optimize during the training process, making their pseudo joint distributions $f_N(u, v)$ and $f_G(u, v)$ progressively approach the real joint distribution $f(u, v)$. This adversarial training mechanism allows the denoiser and generator to learn from each other's information, thereby better simulating the joint distribution between real images and noise images.

In addition, we employ a dual regularization between the denoiser *N* and the generator *G* to mutually enhance their capabilities. During the training process, for any given real noisy-clean image pair (u, v) as well as the pseudo image pair $(\hat{u}, v)$ generated by the denoiser *N* or the pseudo image pair $(u, \hat{v})$ produced by the generator *G*, the discriminator *D* is updated based on the adversarial loss. Subsequently, we fix the parameters of the discriminator *D* and simultaneously update the denoiser *N* and the generator *G*. This implies that in each iteration, the denoiser *N* and the generator *G* maintain interaction and guide each other's optimization process. Through this adversarial training approach, the denoiser and generator are able to collaborate, gradually improving the performance of image denoising and noise generation, while also enhancing the robustness of the model.

In the overall architecture of DGAN, the denoiser *N*, generator *G* and discriminator *D* are all parameterized through deep neural networks. As shown in Figure 1, the denoiser *N* takes the noisy image *v* as input and generates the denoised image $\hat{u}$. The generator *G*, on the other hand, takes the clean image *u* and latent variable *i* as inputs to generate an image $\hat{v}$ with simulated noise. To construct the denoiser *N* and generator *G*, we adopt the attention residual U-Net architecture proposed in this paper as the backbone (see Section 3.2.1 for details). In addition, both employ a residual learning strategy [20] to enhance the model's performance and training efficiency. The discriminator *D* consists of a series of stride convolutional layers and a fully connected layer, which serve to reduce image size and fuse information, aiding in distinguishing between real and generated images. Through the collaborative action of these networks, DGAN demonstrates greater effectiveness in image denoising and noise generation tasks.

3.2. Denoiser and generator

As previously mentioned, DGAN comprises three major components: The denoiser *N*, generator *G* and discriminator *D*. Starting from this section, we will provide a detailed description of the model structure for these components.

3.2.1. Attention residual U-Net

Inspired by U-Net [53], we propose a novel Attention Residual U-Net (ARU) as the backbone architecture for both the denoiser N and generator G. Figure 2 illustrates the overall structure of the ARU network, where d_{in} and d_{out} represent the number of input and output channels, respectively. "Conv (k, s, p)" denotes the convolution operation with a kernel size of (k, k), stride of s and padding of p. Similarly, "TransConv (k, s, p)" refers to the transposed convolution operation with a kernel size of (k, k), stride of s and padding of p. As can be seen from the figure, the ARU network primarily consists of three parts.

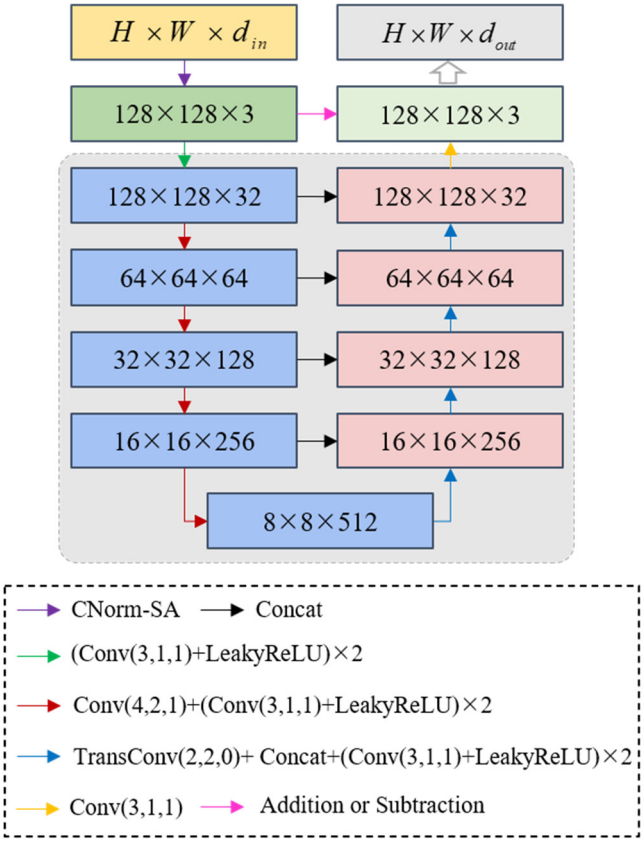

Figure 2. Overall architecture of the attention residual U-Net. The network consists of key components such as a global feature extraction layer and a symmetric encoder-decoder structure of the residual U-Net. "Conv (k, s, p)" denotes the convolution operation with a kernel size of $(\boldsymbol{k}, \boldsymbol{k})$, stride of s and padding of p. Similarly, "TransConv (k, s, p)" refers to the transposed convolution operation with a kernel size of $(\boldsymbol{k}, \boldsymbol{k})$, stride of s and padding of p.

1) Global Feature Extraction Layer. This layer transforms the input feature map x $(H \times W \times d_{in})$ into a global feature map $g_1(x)$ using the linear self-attention operation based on cross-normalization proposed in this paper (see Section 3.2.2 for details). Through the processing of the global feature extraction layer, the original input image can be converted into a more enriched and high-dimensional global feature map, providing beneficial feature representations for subsequent processing steps.

2) Based on the global feature map $g_1(x)$, we adopt a symmetric encoder-decoder structure of the residual U-Net for learning to extract and encode the most relevant multi-scale contextual information $F(g_1(x))$, as shown in Figure 2. Within this network structure, we utilize progressively downsampled feature maps to extract multi-scale features. During the step-by-step upsampling process, the output feature map is concatenated with the feature map before upsampling, and after convolution operations, it is progressively encoded into a high-resolution feature map. This process aids in mitigating the potential loss of fine details that may result from direct large-scale upsampling. Through this structure, we are able to extract the most relevant multi-scale contextual features, thereby enabling the model to better capture the rich local and global information within the image. Local features are crucial for capturing subtle structures and detail information in the image, while global features assist in understanding the overall semantics and structure of the image. In tasks of image denoising and noise generation, the extraction of such multi-scale features is particularly critical.

3) Residual connections that fuse global features and multi-scale features through addition or subtraction operations: $g_1(x) \pm F(g_1(x))$, where $g_1(x)$ represents the global features obtained from the input feature *x* through the global feature extraction layer, and $F(g_1(x))$ represents the multi-scale features processed by the symmetric encoder-decoder structure of the residual U-Net after extracting global features. Specifically, the corresponding operation processes for the denoiser *N* and generator *G* are as follows:

$$N(v) = g_1(v) - F(g_1(v)) \tag{6}$$

$$G(u,i) = g_1(u) + F(g_1(u,i)) \tag{7}$$

We name this design as attention residual U-Net and compare it with the traditional residual blocks [54] to elucidate the underlying intuition.

In the original residual block, the operation can be summarized as $RES(x) = x + g_2(g_1(x))$, where $RES(x)$ represents the desired mapping of the input feature x, and g_1 and g_2 represent the weight layers of the convolution operation. The primary design difference between ARU and traditional residual blocks is that ARU replaces the ordinary convolution weight layer with a symmetric encoder-decoder structure of the residual U-Net and replaces the original feature *x* with the global feature $g_1(x)$ transformed through the CNorm-SA: $ARU(x) = g_1(x) \pm F(g_1(x))$, where *F* represents the ARU structure shown in Figure 2. Such a design enables the network to directly extract the most relevant multi-scale contextual features from each residual block, thereby enhancing the network's understanding of the image.

3.2.2. A linear self-attention based on cross-normalization

In order to extract the most task-relevant multi-scale features within the ARU network, we propose a novel linear Self-Attention based on Cross-Normalization (CNorm-SA). The design of this mechanism aims to mitigate the excessive dependency on initial weights. Furthermore, CNorm-SA eliminates the nonlinearity by replacing the Softmax non-linear activation function and altering the operation sequence, thereby reducing the computational complexity of this module to $O(H \times W)$, where H and W represent the height and width of the original input feature map respectively. Consequently, our model can efficiently handle high-resolution inputs, leading to enhanced performance in image processing tasks.

For an input feature map, the formula for the traditional self-attention mechanism is as follows [55]:

$$Q = xW_Q, K = xW_K, V = xW_V \tag{8}$$

$$Att(x) = Softmax(\frac{QK^T}{\sqrt{d_k}})V \tag{9}$$

Where Q, K and V represent the Query, Key and Value matrices generated by linear transformations on the input feature map, respectively, and $W_q \in \mathbb{R}^{d_{in} \times d_k}$, $W_k \in \mathbb{R}^{d_{in} \times d_k}$ and $W_v \in \mathbb{R}^{d_{in} \times d_{out}}$ are weight matrices.

As shown in Figure 3, within CNorm-SA, we introduce the CNorm operation and eliminate the Softmax non-linear activation function. Simultaneously, this self-attention mechanism allows the module to compute K^TV first, then multiply it with Q. The computational and storage complexity of this operation is $O(H \times W)$, thus the process is linearly related to the size of the input feature map.

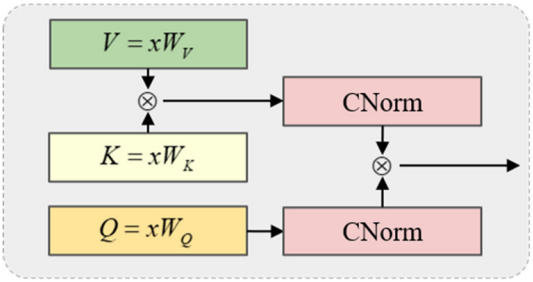

Figure 3. Linear self-attention based on cross-normalization. This self-attention mechanism allows the module to first compute $\boldsymbol{K^TV}$ and multiply it with Q after undergoing separate CNorm operations. It plays a crucial role in attention calculation, effectively capturing the correlations between input features. By utilizing cross-normalization technique, excessive dependence of the self-attention mechanism on initial weights is avoided.

Specifically, CNorm-SA is defined as follows:

$$Att(x) = CNorm_{Row}(Q)(CNorm_{Col}(K^TV)) \tag{10}$$

$$CNorm(a) = \frac{\beta a}{\sqrt{\|a\|^2 + \varepsilon}} \tag{11}$$

where a represents a vector and β is a learnable parameter. ε represents a positive number close to zero, typically chosen to be very small, such as $1e-8$. This setting ensures that the denominator does not become zero during the computation process. $CNorm_{Row}$ or $CNorm_{col}$ respectively denote the application of cross-normalization operation to a matrix row-wise or column-wise, that is:

$$Att(x) = \begin{bmatrix} \hat{q}_0 \cdot \hat{o}_0 & \cdots & \hat{q}_0 \cdot \hat{o}_{d_{out}} \\ \vdots & \ddots & \vdots \\ \hat{q}_{H \times W} \cdot \hat{o}_0 & \cdots & \hat{q}_{H \times W} \cdot \hat{o}_{d_{out}} \end{bmatrix} \tag{12}$$

$$\hat{q}_j = CNorm(Q_{j:}) \tag{13}$$

$$\hat{o}_j = CNorm((K^TV)_{:j}) \tag{14}$$

According to Eq (12), the relational feature can be defined as the cosine similarity between q and

o. To ensure the effectiveness of the relational feature, we employ the CNorm operation to constrain q and o to unit vectors and limit their magnitudes within a finite range through regularization. This treatment prevents their values from having a suppressive effect on the relational feature. Without such handling, the scope of attention would depend on initialization, leading to instability in the attention mechanism.

3.3. Discriminator

In the architecture of GAN [21], the discriminator plays a crucial role. Its primary task is to distinguish real and fake instances, providing guidance for the generator to move in the correct direction of generation. In the framework we propose, we adopt the discriminator structure widely used in literature [56,57]. This structure consists of a series of convolutional layers and a fully connected layer, which are used to gradually reduce the feature size and fuse the extracted features. Figure 4 provides a detailed depiction of the overall design of the discriminator. The numbers next to each feature map represent its spatial dimensions and depth. The input to the discriminator is a pair of concatenated magneto-acousto-electrical images with dimensions of 128 × 128 × 6. After processing by the discriminator, a scalar value is outputted. This scalar value can be interpreted as the discriminator's assessment of the authenticity of the input image pair, thereby guiding the denoiser and generator for more precise image denoising and noise generation.

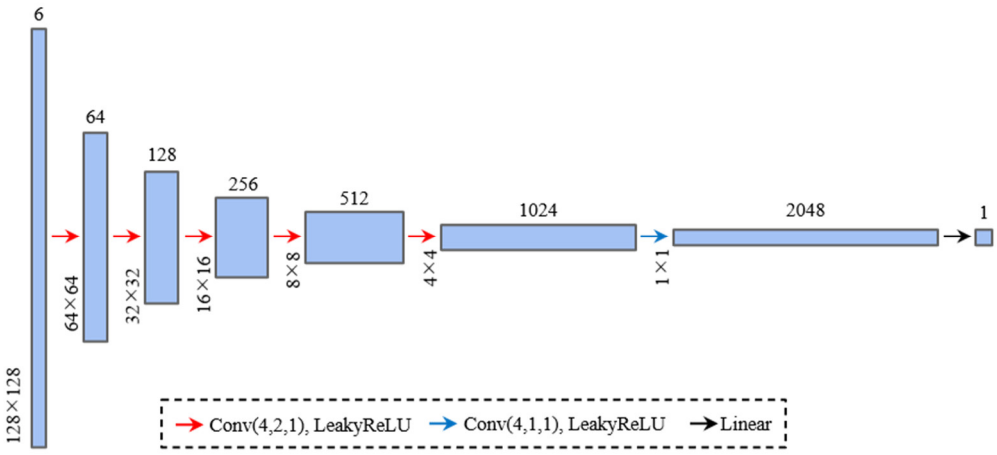

Figure 4. Overall Architecture of the Discriminator. The network consists of a series of convolutional layers and a fully connected layer, which are used to gradually reduce the feature size and fuse the extracted features. The numbers next to each feature map represent its spatial dimensions and depth. The input to the discriminator is a pair of concatenated magneto-acousto-electrical images with dimensions of 128 × 128 × 6. After being processed by the discriminator, a scalar value is outputted, which can be considered as the discriminator's evaluation of the authenticity of the input image pair. "Conv (k, s, p)" denotes the convolution operation with a kernel size of ($\boldsymbol{k}$, $\boldsymbol{k}$), stride of s and padding of p.

3.4. Model optimization and training strategy

In previous studies [58,59], scholars found that combining adversarial loss with traditional loss functions could effectively accelerate and stabilize the training process of GANs. In our image denoising task, we choose to use L_1 loss, i.e., $\|\hat{u} - u\|_1$, which can make the output of the denoiser *N* closer to the real image. However, for the generator *G*, due to the randomness of noise, directly applying L_1 loss may not bring the expected effect. Therefore, we apply the L_1 constraint to the statistical features of the noise distribution:

$$\|G(\hat{v} - u) - G(v - u)\|_1 \tag{15}$$

where $G(\cdot)$ represents the Gaussian filter used to extract the first-order statistical information of noise. By integrating these two regularization factors into the adversarial loss in Eq (5), we can obtain the final objective function:

$$L = \min_{N,G}\max_{D} L_{DGAN}(N, G, D) + \gamma\|\hat{u} - u\|_1 + \eta\|G(\hat{v} - u) - G(v - u)\|_1 \tag{16}$$

where γ and η represent the hyperparameters balancing the losses of the denoiser and generator.

In the objective function defined by Eq (16), three key components need to be optimized: The denoiser *N*, the generator *G* and the discriminator *D*. This is consistent with most methods in GAN-based research literature [21,51,52], i.e., jointly training *N*, *G* and *D*. In implementing the training process, we first fix the denoiser *N* and the generator *G*, then update the parameters of the discriminator *D*. Next, while keeping the discriminator *D* and one other component fixed, we sequentially update the parameters of the denoiser *N* and the generator *G*. This alternating update strategy aids in balancing the learning progression among different components, thereby enhancing the overall performance of the model. To ensure the stability of the training process, we draw upon the gradient penalty technique from WGAN-GP [56]. By introducing an additional gradient penalty term, the discriminator is forced to satisfy the 1-Lipschitz constraint, which in turn improves the stability and effectiveness of the model.

After training, the generator *G* has acquired the ability to simulate additional noise images given any magneto-acousto-electrical clean image. This capability allows us to augment the training dataset by incorporating a large number of magneto-acousto-electrical clean images produced by generator *G* along with their corresponding noise images, thereby retraining the denoiser *N*. This process not only enhances the diversity of the training dataset but also aids in further optimizing the performance of the denoiser *N*, enabling it to better adapt to various noise environments.

3.5. Analysis of algorithm complexity

The algorithm complexity of ARU-DGAN can be calculated by analyzing the complexities of its three major components, namely the denoiser *N*, generator *G* and discriminator *D*.

1) The denoiser *N* and the generator *G* both utilize the ARU network as the backbone architecture. The ARU network consists of key components such as a global feature extraction layer and a symmetric encoder-decoder structure of the residual U-Net. Specifically, the global feature extraction layer employs CNorm-SA operations. By replacing the softmax non-linear activation function and altering the operation sequence, CNorm-SA eliminates the non-linearity in the self-attention mechanism, thereby reducing the computational complexity of this module to $O(H \times W)$, where H

and W represent the height and width of the input feature map, respectively. In the symmetric encoder-decoder structure of the residual U-Net, the upsampling and downsampling processes are composed of a series of convolutional layers. Assuming the input feature map size is $H \times W \times d_{in}$, the kernel size is $K \times K$ and the output channel number is d_{out}, the computational complexity of a single convolution operation is given by $H \times W \times d_{in} \times K \times K \times d_{out}$, where d_{in} represents the number of input channels. Based on the settings of kernel size, input and output channel numbers in each convolutional layer as shown in Figure 2, the main computational cost of the convolution operations is determined by the size of the input feature map, $H \times W$. Therefore, the algorithm complexity of the symmetric encoder-decoder structure of the residual U-Net is also $O(H \times W)$. In summary, both the denoiser N and the generator G have an algorithm complexity of $O(H \times W)$.

2) As shown in Figure 4, the discriminator D consists of a series of convolutional layers and a fully connected layer. Based on the above analysis, the algorithm complexity of the discriminator D is mainly determined by the convolutional layers, thus being $O(H \times W)$.

In summary, the overall algorithm complexity of ARU-DGAN is $O(H \times W)$, which is directly proportional to the size of the input feature map.

4. Experiment

To comprehensively evaluate the proposed ARU-DGAN model, we conducted a series of experiments on a real-world magneto-acousto-electrical image dataset. These experiments aim to answer the following key research questions:

Q1: How does the performance of ARU-DGAN compare with current image denoising methods?

Q2: Do the core components such as the generator G, ARU network and CNorm-SA play a crucial role in enhancing the performance of ARU-DGAN?

4.1. Experiment setup

4.1.1. Datasets and preprocessing

In order to propel research and development in the field of magneto-acousto-electrical image processing and to provide a standard benchmark testing platform for academia, we conducted a large number of real-world measurements and constructed a Magneto-Acousto-Electrical Image (MAEI) dataset. This dataset not only includes magneto-acousto-electrical images from various real-world environments but also reflects the impact of environmental and experimental noise on image quality. By utilizing this dataset, we are able to validate the performance of ARU-DGAN under real-world conditions and ascertain its potential and applicability in the task of magneto-acousto-electrical image denoising.

In the MAEI dataset, it comprises 50 sets of magneto-acoustic-electric clean images and their corresponding noisy image pairs, all of which are obtained from real measurements. Each original image has a size of 3968 × 3968 pixels. To further augment the dataset and increase the diversity of samples, we adopted a cropping strategy: each original image was cropped into 100 patches of 512 × 512 pixels with a stride of 384 pixels. Through this approach, we successfully expanded the scale of the dataset to 5000 pairs of magneto-acoustic-electric images. This augmentation strategy not only elevated the order of magnitude of the dataset but also enriched the diversity of image samples, thereby aiding us

in more comprehensively and accurately evaluating and optimizing the performance of ARU-DGAN.

In the measurement experiments, each real magneto-acoustic-electric noise image was obtained through precise measurements on specific phantoms. These phantoms possess low conductivity characteristics with a conductivity value of 0.2 *S/m* and they contain anomalies that we preset internally. This measurement method enabled us to generate a series of real and representative magneto-acoustic-electric noise images. These images reflect the conductivity distribution inside the phantom and also reveal the existence and location of the anomalies.

In the preparation phase of the experiment, we first prepared the required phantoms. This process involved dissolving an appropriate amount of sodium chloride in water to form a specific solution. To accurately measure the conductivity value of this solution, we employed the Zurich Instruments MFIA device for detection. Next, we mixed agar with the aforementioned solution at a ratio of 1 *g*/100 *ml* of water. This mixture was subsequently heated to boiling and cooled to solidify, successfully creating the phantom. This phantom served as the foundation of our experiment, providing an environment that simulates real biological tissues. Within this phantom, we designed and set anomalies of different sizes and shapes, including circular, elliptical and rectangular forms. The purpose of these anomalies is to simulate potential tumor changes within biological tissues.

In the experimental process, we initially applied a single pulse signal with a center frequency of 0.5 *MHz* to the ultrasonic transducer. This procedure generates ultrasonic waves that penetrate the phantom in the designated direction of the transducer. Subsequently, under the influence of the static magnetic field, magneto-acousto-electrical signals are produced. These signals are captured and measured by electrodes and then amplified by 56 *dB* to enhance their intensity and detectability. Finally, these amplified signals are collected by the NI data acquisition card. By varying the number, position, orientation and size of the anomalies, we were able to generate multiple magneto-acousto-electrical noise images. These images not only reflect the distribution of electrical conductivity within the phantom but also reveal the presence and characteristics of the anomalies.

In the MAEI dataset, the clean images are obtained through a series of processing steps. Initially, the noise images corresponding to the clean images were preprocessed in the measurement experiment using the filtering means described in [17]. Subsequently, these processed images were further optimized and adjusted by medical professionals, resulting in the clean images present in our dataset.

4.1.2. Baselines and evaluation metrics

To validate the effectiveness of our proposed method in the task of magneto-acousto-electrical image denoising, we compared it with several advanced competing methods as follows:

1) WNNM [29]: We discuss the Weighted Nuclear Norm Minimization (WNNM) problem and applies it to image denoising.

2) CBDNet [60]: Convolutional Blind Denoising Network (CBDNet) is an advanced deep learning method that enhances denoising effects by predicting noise levels and incorporating noise uncertainty mapping.

3) RIDNet [18]: Real Image Denoising Network (RIDNet) is a single-stage blind denoising method that adopts a modular architecture and utilizes residual structures and feature attention mechanisms.

4) VDN [19]: Variational Denoising Network (VDN) is a novel blind image denoising method that integrates noise estimation and image denoising into a unique Bayesian framework, which can be flexibly used for estimating and eliminating complex non-independent and identically distributed noise

collected in real scenarios.

5) DANet [45]: Dual Adversarial Network (DANet) is a new unified framework capable of handling both noise removal and noise generation tasks by learning the joint distribution of clean-noise image pairs, rather than merely inferring the posterior distribution of the underlying clean image.

To evaluate the performance of image denoising tasks, we employed two widely used metrics: Peak Signal-to-Noise Ratio (PSNR) [61] and Structural Similarity Index (SSIM) [62]. These two metrics assess image quality from different perspectives.

PSNR is a commonly used metric for assessing image quality, primarily utilized to measure the overall error between the denoised image and the clean image. A higher PSNR indicates less impact from noise. The formula for calculating PSNR is as follows:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{17}$$

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j) - K(i,j)]^2 \tag{18}$$

where MAX_I represents the possible maximum pixel value (for an 8-bit image, $MAX_I = 255$). I and K respectively denote the denoised image and the clean image, while m and n represent the height and width of the image. MSE stands for Mean Squared Error, which is the average of the squared differences between the pixels of the denoised image and the clean image.

SSIM, on the other hand, places more emphasis on measuring the visual quality of the denoised image compared to the clean image. It takes into account the brightness, contrast and structural information of the image, aligning more closely with human subjective perception of image quality. A larger SSIM value indicates higher similarity. The formula for calculating SSIM is as follows:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{19}$$

where x and y represent the denoised image and the clean image respectively, μ_x and μ_y are the averages of x and y, σ_x and σ_y are the variances of x and y, σ_{xy} is the covariance of x and y, and C_1 and C_2 are stability constants to avoid division by zero.

Both of these metrics can be used to evaluate the effect of image denoising. However, since they focus on different aspects, they are typically used in conjunction in practical applications to obtain a more comprehensive evaluation result.

4.1.3. Experiments details

In the training process of ARU-DGAN, we adopted specific initialization and optimization strategies. Specifically, all convolutional layer weights of the denoiser N and the generator G were initialized according to the Xavier method [63]. This approach ensures that the network weights have an appropriate scale at the beginning of the training, thereby accelerating convergence. The weights of the discriminator D, on the other hand, were initialized from a zero-centered normal distribution with a standard deviation of 0.02 [57].

For network training, we selected the Adam optimizer [64], an adaptive learning rate optimization algorithm that can effectively handle sparse gradients and non-stationary objectives. Furthermore, we trained the model for 70 epochs with learning rates set to $1e-4$, $1e-4$ and $2e-4$ for N, G and D

respectively. Moreover, we stipulated that the learning rate would decay linearly by half every 10 epochs. This setting aims to gradually reduce the learning step size as training progresses, allowing the model to fine-tune parameters more delicately when approaching the optimal solution, thereby enhancing the performance of the model.

During the model training phase, we randomly selected 128 × 128 pixels patches from the input images for training. This method can enhance the robustness of the model as it forces the model to learn a more diverse set of image features. The convolutional layer's kernel size, stride and padding, as well as the dimensions and channel numbers of all input and output feature maps, are thoroughly defined for both the ARU network and the discriminator *D*. The specific values of these parameters are presented in Figures 2 and 4, respectively. In the optimization process, before each update of the denoiser *N* and generator *G*, we first updated the discriminator *D* three times. This approach helps maintain the stability of the model and prevent oscillations during the training process. In terms of experimental parameter settings, we set $\gamma = 1000$, $\eta = 10$ and $\alpha = 0.5$. The setting of α implies that the contributions of the denoiser *N* and generator *G* are considered equally important. Additionally, following the default settings of WGAN-GP [56], its penalty coefficient was set to 10. As described in Section 3.4, the trained generator *G* in ARU-DGAN can augment the original training set by generating more synthetic pairs of magneto-acousto-electrical clean and noisy images. Therefore, we retrained the denoiser *N* based on the expanded training dataset.

All models were trained using PyTorch [65]. We ensured optimal performance for all baseline models by carefully adjusting the parameters. All experiments were run independently in the same experimental environment with Intel(R) Xeon(R) Platinum 8358P host and NVIDIA Tesla A40-48GB GPU.

4.2. Overall performance comparison (Q1)

4.2.1. Quantitative comparison

To demonstrate the superiority of ARU-DGAN, we compared its results with five baselines on the MAEI dataset, with specific results shown in Table 1. The optimal and suboptimal results are highlighted in bold and underlined, respectively. From the experimental results, it can be observed that:

1) Leveraging the powerful fitting capability of Deep Neural Networks (DNN), deep learning-based methods significantly outperform traditional WNNM methods in terms of performance. This is primarily due to the self-learning and self-optimizing characteristics of deep learning, enabling it to better understand and handle complex image noise.

2) ARU-DGAN outperforms the state-of-the-art competitive methods in denoising effects, with an improvement of 0.3 *dB* in PSNR and a 0.47% increase in SSIM. The significant improvements of our model over the baseline models can be attributed to two key factors: First, ARU-DGAN employs a dual generative adversarial network, where the synthetic data generated by *G* actively promotes the training of the denoiser *N*; second, the major structures of the denoiser *N* and generator *G* adopt the attention residual U-Net structure and CNorm-SA mechanism, which help extract more relevant multi-scale features, thereby effectively enhancing the image denoising effect. These innovative designs endow ARU-DGAN with higher accuracy and stability in handling magneto-acousto-electrical image denoising tasks.

Table 1. Comparison of our method with five baselines on the MAEI dataset in terms of PSNR and SSIM. The optimal and suboptimal results are highlighted in bold and underlined, respectively.

Methods	Metrics	
	PSNR	SSIM
WNNM	15.59	0.6735
CBDNet	19.32	0.7537
RIDNet	19.54	0.7654
VDN	20.13	<u>0.7664</u>
DANet	<u>20.22</u>	0.7636
ARU-DGAN	20.52	0.7700

4.2.2. Qualitative comparison

Figure 5 presents the results of various methods in terms of visual denoising. As can be seen from the figure, WNNM does not perform well in dealing with noise in magneto-acousto-electrical images. Although CBDNet and RIDNet alleviate some noise during the denoising process, extensive noise remains. VDN and DANet often cause over-smoothing of edges and textures during denoising, resulting in loss of image detail. In particular, DANet loses important information. In contrast, ARU-DGAN excels in restoring sharp edges and detailed textures and its denoising results are closer to the real situation.

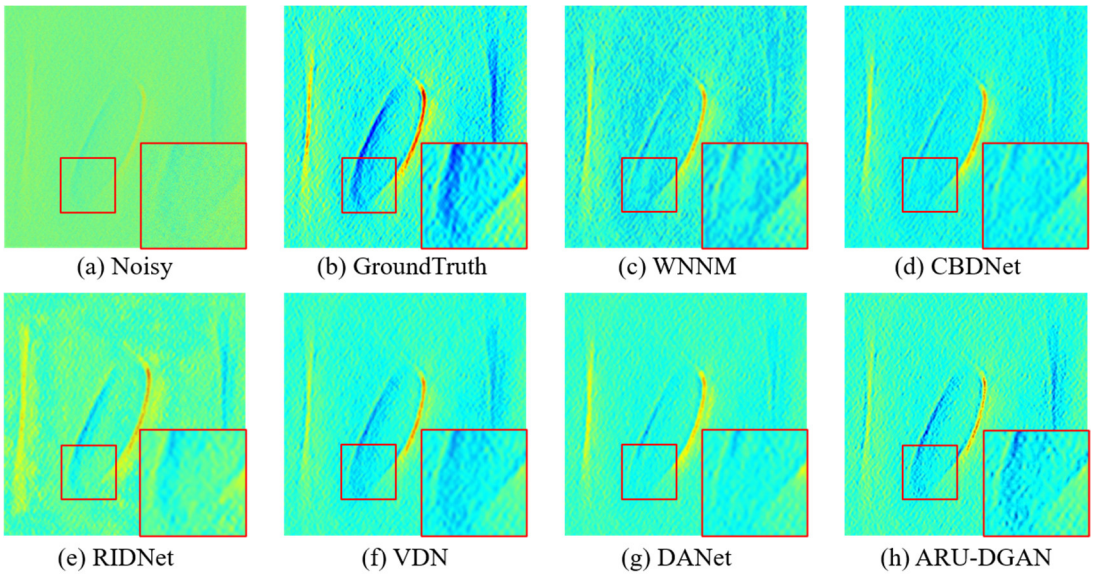

Figure 5. Visual denoising effects of different methods on the MAEI dataset. Here, (a) represents the noisy image, (b) represents the clean image, (c)–(g) represent the denoising results of five baselines, respectively, and (h) represents the denoising result of the proposed ARU-DGAN.

4.3. Detailed model analysis (Q2)

4.3.1. Impact of generator G

To further investigate the contribution of the generator *G* to the overall performance of the model, we trained ARU-DGAN under conditions without generator *G*. As shown in Table 2, ARU-DGAN achieved superior performance. This demonstrates that the interactive relationship between the denoiser *N* and the generator *G* plays a crucial role in enhancing the performance of the model.

Table 2. Performance Comparison of ARU-DGAN without generator *G*. "ARU-DGAN w/o *G*" denotes training ARU-DGAN without the generator *G*. The bold numbers indicate the best performance results.

Variants	Metrics	
	PSNR	SSIM
ARU-DGAN w/o *G*	20.38	0.7654
ARU-DGAN	**20.52**	**0.7700**

4.3.2. Impact of different U-Net networks

In this section, our primary goal is to validate the effectiveness of ARU. Specifically, we employ U-Net and Residual U-Net to respectively replace ARU as the backbone structure for the denoiser *N* and the generator *G*. Table 3 presents the quantitative results from these ablation studies.

Table 3. Quantitative results comparison of different U-Net architectures. "U-Net" and "Residual U-Net" respectively refer to the regular U-Net and the residual U-Net. The bold numbers indicate the best performance results.

Variants	Metrics	
	PSNR	SSIM
U-Net	20.03	0.7529
Residual U-Net	20.21	0.7639
Attention Residual U-Net	**20.52**	**0.7700**

The results clearly show that the performance of U-Net is the poorest, while that of Residual U-Net surpasses U-Net. However, their performances are inferior to ARU. The significant improvement of ARU can be attributed to the CNorm-SA mechanism and the residual mechanism employed in its structure. These two mechanisms allow the network to directly extract the most relevant multi-scale contextual features from each residual block for image denoising.

5. Conclusions

We propose a dual generative adversarial network based on attention residual U-Net for magneto-acousto-electrical image denoising. Specifically, our model simulates the relationship between magneto-acousto-electrical clean and noisy images from two perspectives: First, it maps the noisy

image to the clean image through a denoiser; second, it maps the clean image to the noisy image through a generator. Then, we employ a dual adversarial strategy to train both the denoiser and the generator simultaneously. After this, the trained denoiser can be directly applied to actual denoising tasks, or its performance can be further enhanced by simulating new pairs of clean and noisy images using the trained generator. Furthermore, we design an attention residual U-Net to serve as the backbone for the denoiser and generator within the dual generative adversarial network. The ARU network incorporates a residual mechanism and introduces a linear self-attention based on cross-normalization, proposed in this paper. This design allows the model to effectively extract the most relevant multi-scale contextual information while maintaining high resolution, thereby better modeling the local and global features of magneto-acousto-electrical images. Furthermore, the composite loss function and training strategy within the model can better preserve image details during denoising. Finally, extensive experiments on a real-world magneto-acousto-electrical image dataset constructed for this study demonstrate that ARU-DGAN exhibits excellent performance in both quantitative and qualitative analyses.

Despite the impressive performance of our model in experiments, there are certain potential limitations.

1) Our model heavily relies on a substantial amount of training data. The performance of the model might be affected if there is an insufficient number of training samples available in specific application scenarios.

2) While our model has demonstrated significant achievements in the task of magneto-acousto-electrical image denoising, its applicability to other types of images or datasets requires further validation.

In future research, we intend to explore additional data augmentation techniques to address the issue of limited training samples. Furthermore, we plan to test our model on a broader range of image types and datasets to verify its generality and robustness.

Use of AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant No. 52377227, 52007182 and 51937010.

Conflict of interest

The authors declare that there are no conflicts of interest.

References

1. P. Grasland-Mongrain, C. Lafon, Review on biomedical techniques for imaging electrical impedance, *IRBM*, **39** (2018), 243–250. https://doi.org/10.1016/j.irbm.2018.06.001

2. H. Wen, R. S. Balaban, The potential for hall effect breast imaging, *Breast Dis.*, **10** (1998), 191–195. https://doi.org/10.3233/BD-1998-103-418
3. Y. Zhou, Z. Yu, Q. Ma, G. Guo, J. Tu, D. Zhang, Noninvasive treatment-efficacy evaluation for HIFU therapy based on magneto-acousto-electrical tomography, *IEEE Trans. Biomed. Eng.*, **66** (2018), 666–674. https://doi.org/10.1109/TBME.2018.2853594
4. Y. Li, G. Liu, H. Xia, Z. Xia, Numerical simulations and experimental study of magneto-acousto-electrical tomography with plane transducer, *IEEE Trans. Magn.*, **54** (2017), 1–4. https://doi.org/10.1109/TMAG.2017.2771564
5. G. Guo, J. Wang, Q. Ma, J. Tu, D. Zhang, Non-invasive treatment efficacy evaluation for high-intensity focused ultrasound therapy using magnetically induced magnetoacoustic measurement, *J. Appl. Phys.*, **123** (2018), 154901. https://doi.org/10.1063/1.5024735
6. M. S. Gözü, R. Zengin, N. G. Gençer, Numerical implementation of magneto-acousto-electrical tomography (MAET) using a linear phased array transducer, *Phys. Med. Biol.*, **63** (2018), 35012. https://doi.org/10.1088/1361-6560/aa9f3b
7. P. Grasland-Mongrain, F. Destrempes, J. Mari, R. Souchon, S. Catheline, J. Chapelon, Acousto-electrical speckle pattern in electrical impedance tomography, in *2014 IEEE International Ultrasonics Symposium*, (2014), 221–223. https://doi.org/10.1109/ULTSYM.2014.0056
8. L. Guo, G. Liu, H. Xia, Magneto-acousto-electrical tomography with magnetic induction for conductivity reconstruction, *IEEE Trans. Biomed. Eng.*, **62** (2014), 2114–2124. https://doi.org/10.1109/TBME.2014.2382562
9. L. Kunyansky, C. P. Ingram, R. S. Witte, Rotational magneto-acousto-electric tomography (MAET): theory and experimental validation, *Phys. Med. Biol.*, **62** (2017), 3025. https://doi.org/10.1088/1361-6560/aa6222
10. L. Kunyansky, A mathematical model and inversion procedure for magneto-acousto-electric tomography, *Inverse Probl.*, **28** (2012), 035002. https://doi.org/10.1088/0266-5611/28/3/035002
11. H. Ammari, P. Grasland-Mongrain, P. Millien, L. Seppecher, J. Seo, A mathematical and numerical framework for ultrasonically-induced Lorentz force electrical impedance tomography, *J. Math. Pures Appl.*, **103** (2015), 1390–1409. https://doi.org/10.1016/j.matpur.2014.11.003
12. Y. Li, J. Song, H. Xia, G. Liu, The experimental study of mouse liver in magneto-acousto-electrical tomography by scan mode, *Phys. Med. Biol.*, **65** (2020), 215024. https://doi.org/10.1088/1361-6560/abb4bb
13. Z. Sun, G. Liu, H. Xia, S. Catheline, Lorentz force electrical-impedance tomography using linearly frequency-modulated ultrasound pulse, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **65** (2018), 168–177. https://doi.org/10.1109/TUFFC.2017.2781189
14. M. Dai, X. Chen, T. Sun, L. Yu, M. Chen, H. Lin, A 2D magneto-acousto-electrical tomography method to detect conductivity variation using multifocus image method, *Sensors*, **18** (2018), 2373. https://doi.org/10.3390/s18072373
15. E. Renzhiglova, V. Ivantsiv, Y. Xu, Difference frequency magneto-acousto-electrical tomography (DF-MAET): application of ultrasound-induced radiation force to imaging electrical current density, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **57** (2010), 2391–2402. https://doi.org/10.1109/TUFFC.2010.1707
16. A. Montalibet, J. Jossinet, A. Matias, Scanning electric conductivity gradients with ultrasonically-induced lorentz force, *Ultrason. Imaging*, **23** (2001), 117–132. https://doi.org/10.1177/016173460102300204

17. Y. Jin, H. Zhao, G. Liu, H. Xia, Y. Li, The application of wavelet filtering method in magneto-acousto-electrical tomography, *Phys. Med. Biol.*, **68** (2023), 145014. https://doi.org/10.1088/1361-6560/ace09c
18. S. Anwar, N. Barnes, Real image denoising with feature attention, in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, (2019), 3155–3164. https://doi.org/10.1109/ICCV.2019.00325
19. Z. Yue, H. Yong, Q. Zhao, D. Meng, L. Zhang, Variational denoising network: toward blind noise modeling and removal, in *Advances in Neural Information Processing Systems*, **32** (2019). Available from: https://proceedings.neurips.cc/paper_files/paper/2019/file/6395ebd0f4b478145ecfbaf939454fa4-Paper.pdf.
20. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, Beyond a gaussian denoiser: residual learning of deep CNN for image denoising, *IEEE Trans. Image Process.*, **26** (2017), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206
21. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative adversarial nets, in *Advances in Neural Information Processing Systems*, **27** (2014). Available from: https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.
22. J. Chen, J. Chen, H. Chao, M. Yang, Image blind denoising with generative adversarial network based noise modeling, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2018), 3155–3164. https://doi.org/10.1109/CVPR.2018.00333
23. D. W. Kim, J. R. Chung, S. W. Jung, Grdn: Grouped residual dense network for real image denoising and gan-based real-world noise modeling, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (2019), 2086–2094. https://doi.org/10.1109/CVPRW.2019.00261
24. P. Grasland-Mongrain, J. M. Mari, J. Y. Chapelon, C. Lafon, Lorentz force electrical impedance tomography, *IRBM*, **34** (2013), 357–360. https://doi.org/10.1016/j.irbm.2013.08.002
25. Y. Li, S. Bu, X. Han, H. Xia, W. Ren, G. Liu, Magneto-acousto-electrical tomography with nonuniform static magnetic field, *IEEE Trans. Instrum. Meas.*, **72** (2023), 1–12. https://doi.org/10.1109/TIM.2023.3244814
26. H. Lin, Y. Chen, S. Xie, M. Yu, D. Deng, T. Sun, et al., A dual-modal imaging method combining ultrasound and electromagnetism for simultaneous measurement of tissue elasticity and electrical conductivity, *IEEE Trans. Biomed. Eng.*, **69** (2022), 2499–2511. https://doi.org/10.1109/TBME.2022.3148120
27. M. Aharon, M. Elad, A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, *IEEE Trans. Signal Process.*, **54** (2006), 4311–4322. https://doi.org/10.1109/TSP.2006.881199
28. M. Elad, M. Aharon, Image denoising via sparse and redundant representations over learned dictionaries, *IEEE Trans. Image Process.*, **15** (2006), 3736–3745. https://doi.org/10.1109/TIP.2006.881969
29. S. Gu, L. Zhang, W. Zuo, X. Feng, Weighted nuclear norm minimization with application to image denoising, in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, (2014), 2862–2869. https://doi.org/10.1109/CVPR.2014.366

30. A. Buades, B. Coll, J. M. Morel, A non-local algorithm for image denoising, in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, **2** (2005), 60–65. https://doi.org/10.1109/CVPR.2005.38
31. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Trans. Image Process.*, **16** (2007), 2080–2095. https://doi.org/10.1109/TIP.2007.901238
32. J. Portilla, V. Strela, M. J. Wainwright, E. P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, *IEEE Trans. Image Process.*, **12** (2003), 1338–1351. https://doi.org/10.1109/TIP.2003.818640
33. L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, **60** (1992), 259–268. https://doi.org/10.1016/0167-2789(92)90242-F
34. A. Barbu, Training an active random field for real-time image denoising, *IEEE Trans. Image Process.*, **18** (2009), 2451–2462. https://doi.org/10.1109/TIP.2009.2028254
35. K. G. G. Samuel, M. F. Tappen, Learning optimized MAP estimates in continuously-valued MRF models, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (2009), 477–484. https://doi.org/10.1109/CVPR.2009.5206774
36. J. Sun, M. F. Tappen, Learning non-local range Markov Random field for image restoration, *CVPR 2011*, (2011), 2745–2752. https://doi.org/10.1109/CVPR.2011.5995520
37. U. Schmidt, *Half-quadratic Inference and Learning for Natural Images*, Ph.D thesis, Technische University, 2017. Available from: https://tuprints.ulb.tu-darmstadt.de/id/eprint/6044.
38. U. Schmidt, S. Roth, Shrinkage fields for effective image restoration, in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, (2014), 2774–2781. https://doi.org/10.1109/CVPR.2014.349
39. Y. Chen, T. Pock, Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration, *IEEE Trans. Pattern Anal. Mach. Intell.*, **39** (2017), 1256–1272. https://doi.org/10.1109/TPAMI.2016.2596743
40. V. Jain, S. Seung, Natural image denoising with convolutional networks, in *Advances in Neural Information Processing Systems*, **21** (2008). Available from: https://proceedings.neurips.cc/paper_files/paper/2008/file/c16a5320fa475530d9583c34fd356ef5-Paper.pdf.
41. H. C. Burger, C. J. Schuler, S. Harmeling, Image denoising: can plain neural networks compete with BM3D? in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, (2012), 2392–2399. https://doi.org/10.1109/CVPR.2012.6247952
42. X. Mao, C. Shen, Y. B. Yang, Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections, in *Advances in Neural Information Processing Systems*, **29** (2016). Available from: https://proceedings.neurips.cc/paper_files/paper/2016/file/0ed9422357395a0d4879191c66f4faa2-Paper.pdf.
43. D. Liu, B. Wen, Y. Fan, C. C. Loy, T. S. Huang, Non-local recurrent network for image restoration, in *Advances in Neural Information Processing Systems*, **31** (2018). Available from: https://proceedings.neurips.cc/paper_files/paper/2018/file/fc49306d97602c8ed1be1dfbf0835ead-Paper.pdf.
44. T. Plötz, S. Roth, Neural nearest neighbors networks, in *Advances in Neural Information Processing Systems*, **31** (2018). Available from: https://proceedings.neurips.cc/paper_files/paper/2018/file/f0e52b27a7a5d6a1a87373dffa53dbe5-Paper.pdf.

45. Z. Yue, Q. Zhao, L. Zhang, D. Meng, Dual adversarial network: toward real-world noise removal and noise generation, in *Computer Vision – ECCV 2020*, (2020), 41–58. https://doi.org/10.1007/978-3-030-58607-2_3
46. N. Mu, Z. Lyu, M. Rezaeitaleshmahalleh, J. Tang, J. Jiang, An attention residual u-net with differential preprocessing and geometric postprocessing: learning how to segment vasculature including intracranial aneurysms, *Med. Image Anal.*, **84** (2023), 102697. https://doi.org/10.1016/j.media.2022.102697
47. X. Liu, D. Zhang, J. Yao, J. Tang, Transformer and convolutional based dual branch network for retinal vessel segmentation in OCTA images, *Biomed. Signal Process. Control*, **83** (2023), 104604. https://doi.org/10.1016/j.bspc.2023.104604
48. M. Versaci, G. Angiulli, P. Crucitti, D. de Carlo, F. Laganà, D. Pellicanò, et al., A fuzzy similarity-based approach to classify numerically simulated and experimentally detected carbon fiber-reinforced polymer plate defects, *Sensors*, **22** (2022), 4232. https://doi.org/10.3390/s22114232
49. Z. Yue, H. Yong, D. Meng, Q. Zhao, Y. Leung, L. Zhang, Robust multiview subspace learning with nonindependently and nonidentically distributed complex noise, *IEEE Trans. Neural Networks Learn. Syst.*, **31** (2020), 1070–1083. https://doi.org/10.1109/TNNLS.2019.2917328
50. D. P. Kingma, M. Welling, Auto-encoding variational bayes, preprint, arXiv:1312.6114.
51. C. Li, T. Xu, J. Zhu, B. Zhang, Triple generative adversarial nets, in *Advances in Neural Information Processing Systems*, **30** (2017). Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/86e78499eeb33fb9cac16b7555b50767-Paper.pdf.
52. M. Arjovsky, S. Chintala, L. Bottou, Wasserstein generative adversarial networks, in *Proceedings of the 34th International Conference on Machine Learning*, **70** (2017), 214–223. Available from: https://proceedings.mlr.press/v70/arjovsky17a.html.
53. O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, (2015), 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
54. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 770–778. https://doi.org/10.1109/CVPR.2016.90
55. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, et al., Attention is all you need, in *Advances in Neural Information Processing Systems*, **30** (2017). Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.
56. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. C. Courville, Improved training of wasserstein GANs, in *Advances in Neural Information Processing Systems*, **30** (2017). Available from: https://proceedings.neurips.cc/paper_files/paper/2017/file/892c3b1c6dccd52936e27cbd0ff683d6-Paper.pdf.
57. A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, preprint, arXiv:1511.06434.
58. P. Isola, J. Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 1125–1134. https://doi.org/10.1109/CVPR.2017.632

59. J. Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, (2017), 2223–2232.
60. S. Guo, Z. Yan, K. Zhang, W. Zuo, L. Zhang, Toward convolutional blind denoising of real photographs, in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2019), 1712–1722. https://doi.org/10.1109/CVPR.2019.00181
61. J. Korhonen, J. You, Peak signal-to-noise ratio revisited: is simple beautiful? in *2012 Fourth International Workshop on Quality of Multimedia Experience*, (2012), 37–38. https://doi.org/10.1109/QoMEX.2012.6263880
62. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Trans. Image Process.*, **13** (2004), 600–612. https://doi.org/10.1109/TIP.2003.819861
63. X. Cao, Y. Chen, Q. Zhao, D. Meng, Y. Wang, D. Wang, et al., Low-rank matrix factorization under general mixture noise distributions, in *2015 IEEE International Conference on Computer Vision (ICCV)*, (2015), 1493–1501. https://doi.org/10.1109/ICCV.2015.175
64. D. P. Kingma, J. Ba, Adam: a method for stochastic optimization, preprint, arXiv:1412.6980.
65. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, et al., PyTorch: an imperative style, high-performance deep learning library, in *Advances in Neural Information Processing Systems*, **32** (2019). Available from: https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf.

AIMS Press