

中图法分类号: TP18; TP391 文献标识码: A 文章编号: 1006-8961(2025)12-3884-16

论文引用格式: Zhang Y R, Liu J T, Xu G Q and Hao D Y. 2025. Self-supervised coal mine image denoising with adaptive masking. Journal of Image and Graphics, 30(12):3884-3899(张亚茹, 刘健涛, 许国庆, 郝定溢. 2025. 基于自适应掩码的自监督矿井图像去噪. 中国图象图形学报, 30(12):3884-3899)[DOI:10.11834/jig.240624]

基于自适应掩码的自监督矿井图像去噪

张亚茹^{1*}, 刘健涛¹, 许国庆¹, 郝定溢²

1. 安徽理工大学人工智能学院, 淮南 232001; 2. 安徽理工大学安全科学与工程学院, 淮南 232001

摘要: 目的 受煤矿井下粉尘浓度高和光照不均匀等因素影响, 矿井图像容易产生不同水平的噪声、降低图像的局部对比度、丢失细节信息和边缘纹理特征, 从而影响矿井图像的信息提取精度和视觉质量。针对上述问题, 提出一种基于自适应掩码的矿井图像自监督去噪算法, 主要包含自适应掩码、掩码集成以及自适应集成损失3部分。**方法** 首先, 设计自适应掩码对矿井图像进行分块以减少后续计算消耗, 逐次对各块中心像素外的边缘像素及角点像素进行掩码, 避免恒等映射的同时增加数据多样性; 然后, 设计掩码集成对神经网络的输出和掩码区域进行重新组合, 计算两者之间的Hadamard积以增强网络对噪声与信号边界的准确识别, 综合考虑矿井图像的局部结构和全局特征, 从而提升去噪后矿井图像的完整性和连贯性; 最后, 设计自适应集成损失, 将集成图像作为训练标签, 帮助模型更好地理解矿井图像中局部特征和全局特征之间的关系, 加入原始噪声图像增强模型对信号变化的敏感性, 适应不同场景下的去噪任务。**结果** 在煤矿井下图像数据集和4个公共数据集进行实验, 去噪后的图像质量在主观感受和客观指标上均优于其他对比算法。在高斯噪声水平为50的巷道场景下, 相比B2U(blind2unblind)和NBR2NBR(neighbor2neighbor), 峰值信噪比(peak signal-to-noise ratio, PSNR)分别提高4.2 dB和2.99 dB, 结构相似性指数(structural similarity index, SSIM)分别提高0.055和0.077。在5至50的高斯噪声范围内, 本文方法计算的PSNR相较TBSN(transformer-based blind-spot network)在Kodak24(Kodak lossless true color image suite 24)数据集上提升1.09%, 在BSD300(Berkeley segmentation dataset 300)数据集上提升0.72%, 相较NBR2NBR在BSD500数据集上提升0.68%。**结论** 所提算法能够处理含有不同程度和类型噪声的矿井图像, 有效去除噪声的同时保留图像的细节信息, 展现出优越的鲁棒性和广泛的适用性。代码获取地址: <https://www.sciclb.cn/anonymous/skpswk56>。**关键词:** 自监督学习; 矿井图像去噪; 自适应掩码; Hadamard积; 自适应集成损失

Self-supervised coal mine image denoising with adaptive masking

Zhang Yaru^{1*}, Liu Jiantao¹, Xu Guoqing¹, Hao Dingyi²

1. School of Artificial Intelligence, Anhui University of Science and Technology, Huainan 232001, China;

2. School of Safety Science and Engineering, Anhui University of Science and Technology, Huainan 232001, China

Abstract: Objective The objective of this research is to enhance the quality and accuracy of information extracted from coal mine images, which are often degraded by high dust concentrations and uneven lighting conditions. These challenging environmental conditions introduce noise, reduce local contrast, and lead to the loss of fine details and edge textures, ulti-

收稿日期: 2024-11-18; 修回日期: 2025-04-18; 预印本日期: 2025-04-25

* 通信作者: 张亚茹 yrzhang1014@163.com

基金项目: 安徽省高等学校科学研究项目(2022AH050830); 安徽理工大学引进人才科研启动基金项目(2021yjrc37); 国家自然科学基金项目(52204123)

Supported by: Natural Science Research Projects of Anhui Universities(2022AH050830); Scientific Research Foundation for High-Level Talents of Anhui University of Science and Technology(2021yjrc37); National Natural Science Foundation of China(52204123)

mately compromising the visual quality and the reliability of information extraction. Aiming to address these challenges, this study proposes a self-supervised coal mine image denoising algorithm based on adaptive masking. Designed to handle a wide range of noise levels and types, this algorithm aims to restore the original integrity of the image while preserving critical visual features. The proposed algorithm is divided into three main components: adaptive masking, mask integration, and an adaptive integrated loss function. Each component plays a vital role in enhancing the denoising process, ensuring that the final output is accurate and visually appealing.

Method The adaptive masking component is the cornerstone of the proposed algorithm, enabling segmented processing of coal mine images. This segmentation not only reduces computational overhead but also allows for more targeted and effective denoising. By dividing each image into smaller blocks, the algorithm can analyze and process each section independently, thereby improving the overall efficiency of the denoising process. The module operates by sequentially applying a mask to the edge and corner pixels of each block, while deliberately excluding the central pixels. This method prevents the network from performing a trivial identity mapping that fails to enhance image quality. Instead, this approach introduces data variability that boosts the generalization capability and robustness of the neural network model, making it adaptable to previously unknown images. The adaptive nature of the mask ensures that the module responds dynamically to varying noise levels and image features. By analyzing local variance and texture complexity, the mask can adaptively determine the optimal masking strategy for each block. This tailored approach ensures that the denoising process is responsive to the specific characteristics of each image, substantially improving its effectiveness. Subsequently, once the masking process is complete, the mask integration module is employed. This module is responsible for fusing the neural network's output with the masked areas to reconstruct a coherent and denoised image. The integration involves calculating the Hadamard product (element-wise multiplication) between the network's output and the masked image. This strategic operation enhances the network's capability to distinguish between actual image content and noise, especially around edges and texture boundaries. In this stage, considering local and global features of the coal mine images is crucial. Effective integration of these features allows the algorithm to effectively interpret image context, leading to denoised outputs that are coherent and structurally complete. The mask integration module also ensures that denoised areas seamlessly blend into the rest of the image, preserving the overall visual flow and structural integrity. Furthermore, this module incorporates a quality evaluation mechanism to assess the effectiveness of the integration. The feedback from these evaluations is used to iteratively refine the integration process. The final component of the algorithm is an adaptive integrated loss function, which guides the model during training. This loss function is specifically designed to address the unique challenges of coal mine image denoising, including complex noise patterns and the need to preserve subtle image details. The adaptive integrated loss uses the integrated image as a training label, allowing the model to learn effectively from the differences between the noisy input images and the denoised outputs. Additionally, by incorporating the original noisy image, the loss function increases the model's sensitivity to signal changes, enhancing its adaptability across various denoising scenarios and noise conditions.

Result The proposed algorithm was rigorously tested using an underground coal mine image dataset alongside four additional public datasets, including Kodak24 (Kodak lossless true color image suite), BSD300 (Berkeley segmentation dataset 300), and BSDS500 (Berkeley segmentation dataset 500). The experiments were specifically designed to simulate real-world conditions, with a particular emphasis on dimly lit environments commonly encountered in coal mines. The results of these experiments demonstrated that the algorithm substantially outperformed other comparative denoising algorithms, in terms of subjective evaluations and objective metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM). In tunnel scenes with a high level of Gaussian noise (level 50), the algorithm achieved substantial improvements in PSNR/SSIM values compared to existing methods such as B2U and NBR2NBR, with increases of 4.2 dB/0.055 and 2.99 dB/0.077, respectively. Furthermore, when tested on images corrupted with Gaussian noise levels ranging from 5 to 50 on the public datasets, the algorithm consistently demonstrated substantial PSNR improvements over the second-best method, with increases of 1.09%, 0.72%, and 0.68% for Kodak24, BSD300, and BSDS500, respectively.

Conclusion The proposed self-supervised denoising algorithm has demonstrated a strong capability to remove noise while preserving overall image information from single coal mine images, across various noise levels and types. This finding highlights the algorithm's robustness and generalization capabilities, making it a promising tool for real-world applications in coal mine monitoring and safety systems. The effectiveness

of the algorithm in enhancing image quality and improving the accuracy of information extraction, even under challenging conditions, underscores its potential to make a substantial contribution to the field of coal mine image processing and analysis. The code in this paper can be obtained by <https://www.sciclb.cn/anonymous/skpswk56>.

Key words: self-supervised learning; mine image denoising; adaptive masking; Hadamard product; adaptive integrated loss

0 引言

煤炭工业是国民经济的重要支柱产业,其智能化转型升级是推动经济高质量发展的重要战略举措(刘峰等,2021;王国法,2022)。随着煤炭开采深度和强度的不断增加,煤矿安全生产面临着越来越严峻的挑战,亟需通过智能化手段提升生产效率、保障安全生产(毛善君等,2020)。高质量矿井图像作为煤矿智能化的重要信息来源,可以提供更加丰富、准确的信息,在井下视频监控(樊红卫等,2024)、井下轨道检测(卫星等,2020)以及煤矸洗选(王鹏等,2020)等智能化应用中发挥着至关重要的作用。然而,由于煤矿井下成像系统限制和环境条件影响,矿井图像存在大量噪声,严重干扰基于井下图像数据的下游任务精准分析(操晓春等,2025)。因此,去除矿井图像噪声可以有效解决图像质量低下的问题,为煤矿智能化提供更加高效、安全和可靠的信息。

图像去噪任务是从含噪声的图像中恢复原始信号,是图像预处理的基本任务之一(李潇瑶等,2022)。其中常见噪声信号(Jebur等,2024)有高斯噪声、椒盐噪声、泊松噪声以及乘性噪声。井下环境光照条件差,导致图像中出现明暗对比强烈的区域,形成光照不均匀噪声;粉尘弥漫的环境使得图像中出现随机的颗粒状噪声;成像设备本身的电子噪声和信号干扰,导致图像中出现高斯噪声和椒盐噪声。传统去噪算法针对特定噪声分布分为均值滤波、中值滤波、高斯滤波和双边滤波等,均需根据具体场景和需求来调整相应参数。非局部均值(Buades等,2005)和基于块匹配的3D协同滤波(block-matching and 3D filtering, BM3D)(Dabov等,2007a)利用像素级相似性比较以去除噪声,但其有效性依赖预先定义的自然性概念,易导致过拟合和图像细节丢失。

神经网络的发展解决了传统算法计算效率问题,其中有监督的神经网络去噪方式需要大量配对

的噪声图像和干净图像作为训练数据和标注数据完成去噪任务(Chen和Pock,2017;Zhang等,2017)。由于煤矿井下环境昏暗、潮湿以及内部结构复杂,矿井图像数据的采集和标注具有极高的挑战性(Zhang等,2018;Guo等,2019)。Lempitsky等人(2018)利用深度图像先验知识展示了神经网络参数化固有的抗噪声能力和信号保持能力,缓解了配对数据集的限制。Lehtinen等人(2018)成功消除了对干净图像的需求,但依赖于相同分布的噪声,并限制在静态场景中。然而矿井作业处于动态变化的环境中,需要能够及时应对新的噪声模式和图像特征,为此,Krull等人(2019)提出一种更为灵活的去噪策略,利用在单幅噪声图像上训练的盲点网络,避免恒等映射的同时能够适应更加广泛的动态环境和噪声模式。Batson和Royer(2019)通过模拟噪声与信号的内在关系,利用条件独立的噪声特征图像集,从而能够从单个物体的测量中自主学习去噪功能。

此外,矿井环境光照不均、粉尘弥漫,导致矿井图像中的噪声特性并不单一,需要处理类型多变的噪声,且上述方法中的手动掩码方式会导致去噪图像中的信息丢失,因此,Laine等人(2019)将盲点数据与神经网络的输出进行结合,但需要盲点数据与训练数据在噪声分布上具有高度相似性。Quan等人(2020)提供了一种基于神经网络的解决方案,通过在输入图像的Bernoulli采样实例上进行dropout训练,完成在单幅噪声图像上的去噪任务,为矿井图像去噪领域提供了更为深入的解决策略。Huang等人(2021)通过子采样生成噪声图像对在去噪网络上进行训练,避免了对噪声分布假设的过度依赖,同时解决了重复获取标签数据的需求。Wang等人(2022)使用原始去噪图像训练去噪器,并将掩码图像作为反向传播时的中介,有效减轻了由盲点引起的信息丢失。

上述方法在矿井图像去噪中拥有独立的去噪器,并与盲点数据结合能有效应对单一噪声分布的图像去噪任务,但仍存在下述问题:矿井作业环境动

态变化频繁,增加了去噪的难度;传统掩码方式在处理矿井图像时,难以精准区分噪声和有效信号,导致噪声与信号的边界模糊不清,从而在去噪过程中容易丢失图像的关键细节,造成图像失真;井下环境复杂多变,且需专业人员参与标注,导致高质量标签数据采集耗时且成本高;矿井设备计算能力有限,难以支持复杂算法高算力需求,限制了传统掩码方式的实际应用。

因此,本文提出一种基于自适应掩码的矿井图像自监督去噪算法,分为自适应掩码、掩码集成和自适应集成损失3个部分。首先,针对动态变化的场景,通过准确划分图像为具有相似视觉特征的区域,从而有效保留图像边缘细节,自适应掩码部分单独对分块处理后的边缘像素和角点像素掩码,增强图像边缘特征,避免恒等映射问题即网络输出与输入完全相同,防止学习过程无效,同时不掩码每块的中心像素,保留图像整体结构。然后,掩码集成部分对神经网络的输出与掩码区域重新组合并进行特征融合,并设计自适应集成损失,将集成图像作为训练标签,使网络模型可以从无标注数据中学习有效特征表示,并进一步分析损失函数的收敛性,证明模型仅通过单幅噪声图像即可完成图像去噪任务,提高算法在矿井环境下的实用性。最后,本文在煤矿井下

图像数据集(coal mine underground image dataset, CMUID)(程德强等,2022)上对网络模型进行训练和测试,同时,为进一步验证模型泛化能力,对合成噪声进行测试,为煤矿智能化提供更加可靠和高效的图像去噪模型。

1 算法框架

本文提出一种基于自适应掩码的矿井图像自监督去噪算法,算法总体框架如图1所示。首先,将输入的噪声图像 y 分块为 y^i ,其中 i 用于标识图像块, j 用于标识掩码位置,将分块处理后的各个图像块输入自适应掩码部分 $\Omega(\cdot)$,独立掩码各个图像块的边缘像素和角点像素,同时在给定范围内,不掩码中心像素以保证图像整体结构,经去噪网络后产生 $f_\theta(\Omega^j)$,并将原始噪声图像 y 输入去噪网络生成 $f_\theta(y)$ 。然后,经掩码集成部分重新组合为原图尺寸,对 $f_\theta(\Omega^j)$ 及掩码区域进一步集成产生集成图像 g 。最后,自适应掩码损失将集成图像作为自监督学习标签参与网络训练,保留矿井图像中关键特征,同时减轻模型复杂性,防止过拟合噪声信息。自监督去噪算法使用U-Net作为去噪网络,避免在去噪过程引入伪影。

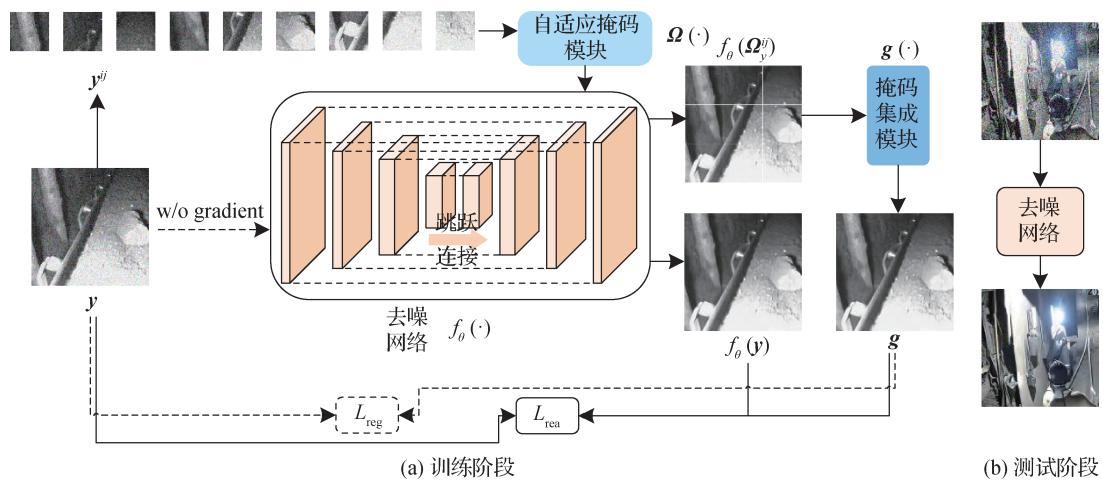


图1 去噪算法总框架

Fig. 1 Overall framework of denoising algorithm ((a) train stage; (b) test stage)

1.1 自适应掩码

考虑到矿井图像本身的边缘特征不明显,部分信息会在下采样过程中丢失,导致去噪后的矿井图像纹理细节不清晰。因此,设计针对边缘像素和角

点像素独立掩码的方法,用于增强网络模型对这些区域的关注力度,更有效捕捉图像细节变化。

为进一步保证图像空间结构,对各个图像块独立掩码后再将其重新组合。如图2所示,以 3×3 的

图像块为例,为防止网络仅学习图像整体亮度或颜色等低频信息,而忽略高频细节,除中心像素,计算各个图像块 y^{ij} 与对应掩码块 Ω_{ij} 的叉积。具体为

$$y_k = y^{ij} \times \Omega_{ij} \quad (1)$$

式中, i 表示图像块的索引, j 表示图像块内每个像素的索引。 y_k 表示各个图像块边缘像素或角点像素的掩码结果, k 表示划分的图像块数, $k \in \{0, \dots, 8\}$ 且 $j \neq 4$ 。

同时,为了使矿井图像视觉效果更加柔和自然,将各个图像块与高斯核进行卷积,保证去噪后的图像边缘与内部区域之间没有明显分界线。将其结果 y_c 与活动区域 $1 - \Omega_{ij}$ 做叉积,具体为

$$y_m = y_c \times (1 - \Omega_{ij}) \quad (2)$$

式中, $1 - \Omega_{ij}$ 表示活动区域,即边缘像素和角点像素未掩码的结果。 y_m 表示经过高斯平滑处理后的图像 y_c 与活动区域叉积后的结果, c, m 表示划分的图像块数, $c, m \in \{0, \dots, 8\}$ 且 $j \neq 4$ 。

然后,整合两者对应图像块,加大噪声信号与干

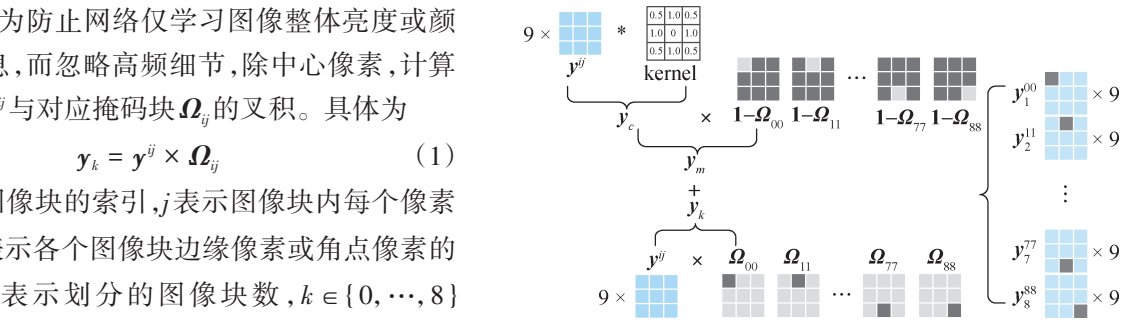


图2 自适应掩码

Fig. 2 Adaptive masking

净信号之间的差距,增强网络模型区分噪声与信号的能力,即

$$y_k^{ij} = y_k + y_m \quad (3)$$

式中, y_k^{ij} 表示掩码区域与未掩码区域相加的结果,且 $j \neq 4$ 。

最后,如图3所示,重新组合各个掩码的图像块,恢复成原始图像的尺寸 Ω_y ,并将其输入去噪网络中。

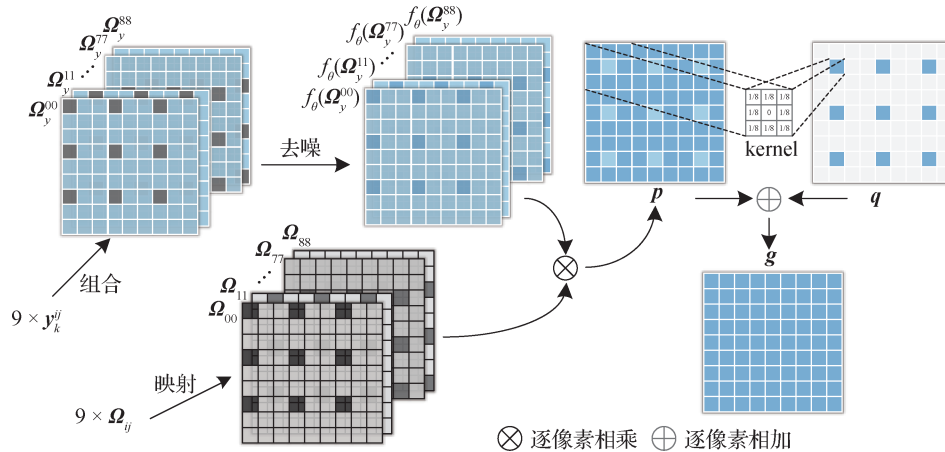


图3 掩码集成

Fig. 3 Mask integration

1.2 掩码集成

掩码的引入虽然会增加数据多样性,但也阻碍了网络模型对图像的整体理解,导致部分信息的丢失以及信号边界模糊。因此,如图3所示,设计一种掩码集成方式,通过网络对噪声与信号之间结构差异的准确识别,综合考虑局部特征与全局特征之间的相互作用,使图像纹理细节更加真实精细。

如图3所示,将各个掩码块按照对应图像块位置进行映射产生 Ω_{ij} ,并对掩码图像经去噪网络获得的图像 $f_\theta(\Omega_y^{ij})$ 与相应的掩码区域做Hadamard积。

$$p = f_\theta(\Omega_y^{ij}) \otimes \Omega_y^{ij} \quad (4)$$

式中, p 表示去噪图像与掩码图像集成之后的结果,且 $j \neq 4$ 。

为进一步增强图像整体结构,计算集成后各块区域的边缘像素和角点像素的平均值,以此作为中心像素的值。进而由 $g = p \oplus q$ 得到集成图像,其中, q 表示中心像素的恢复结果, g 表示与掩码区域相结合的最终结果。

1.3 自适应集成损失

通常将噪声图像 y 建模为 $y = x + n$,其中, x 表

示干净信号, \mathbf{n} 表示噪声信号。在假设噪声信号是零均值且独立同分布的情况下, 噪声图像的均值等同于干净信号, 即达到了去噪目的。

$$\mathbf{E}(\mathbf{y}) = \mathbf{E}(\mathbf{x} + \mathbf{n}) = \mathbf{E}(\mathbf{x}) \quad (5)$$

式中, $\mathbf{E}(\cdot)$ 表示期望。

然而, 矿井环境中的粉尘浓度不同以及光照不均匀, 难以保证噪声信号均是以零均值存在的, 因此本文将集成图像 \mathbf{g} 作为自监督学习标签, 能够保证训练过程一致性和准确性, 提高模型在不同条件下的泛化能力, 同时将原始噪声图像也加入设定的损失函数中

$$L = L_{\text{rea}} + \varphi L_{\text{reg}} = \mathbf{E} \left\| \mu \mathbf{g} + f_{\theta}(\mathbf{y}) - \mu \mathbf{y} - \mathbf{y} \right\|_2^2 + \varphi \mathbf{E} \left\| \mu \mathbf{g} - \mu \mathbf{y} \right\|_2^2 \quad (6)$$

式中, 超参数 φ 用于确定在初始阶段, 集成图像与噪声图像之间的差距对整个训练过程的影响, 即衡量掩码区域对整个图像的影响程度。超参数 μ 用于调整集成图像与噪声图像之间的集成比例, 即用于平衡网络对局部特征和整体图像结构的关注程度, 调节噪声和信号之间的边界。其正则化项缩小解的空间, 控制在 $(\mathbf{g}^*, f_{\theta}^*(\mathbf{y}))$ 之间, \mathbf{g}^* 表示最优情况下的集成图像, $f_{\theta}^*(\mathbf{y})$ 表示最优情况下的去噪图像, 稳定训练过程和约束网络复杂性, 使网络学习更加复杂的非线性关系。

针对损失函数的收敛性需明确训练的最终目标是为了寻找一个最优参数 θ , 使去噪效果最佳, 定义的目标函数为

$$\arg \min_{\theta} \mathbf{E} \left\| \mu \mathbf{g} + f_{\theta}(\mathbf{y}) - \mu \mathbf{y} - \mathbf{y} \right\|_2^2 \quad (7)$$

式中, $\mu \rightarrow 0$, 则 $\mu(\mathbf{g} - \mathbf{y}) \rightarrow 0$, 当 $f_{\theta}(\mathbf{y}) \rightarrow f_{\theta}^*(\mathbf{y})$ 时, $\mathbf{E}(f_{\theta}^*(\mathbf{y})) = \mathbf{E}(\mathbf{y}) = \mathbf{x}$ 。 $\mathbf{E}(f_{\theta}^*(\mathbf{y}))$ 表示去噪图像的均值等于噪声图像的均值, 继而等于干净图像, 避免了噪声的零均值假设。

同时, 经过对目标函数的分析可以得到, 原始噪声图像的加入并不影响最终的去噪结果, 反之有助于模型找到最优解。最终测试阶段如图 1(b) 所示, 仅使用单幅噪声矿井图像即可得到干净矿井图像。

2 实验结果与分析

2.1 数据集

本文方法采用煤矿井下图像数据集(CMUID)

(程德强等, 2022) 进行训练。该数据集是使用 KBA12B 矿用相机在多样化的煤矿环境中获取的, 为矿井图像去噪提供真实和具有挑战性的试验数据。从该数据集中选择了 5 000 幅涵盖不同矿区的 3 种不同地下场景的图像进行测试, 为确保一致性并便于分析, 所有测试图像都统一裁剪为 500×375 像素。

此外, 为确保网络模型对矿井图像去噪能力的准确性和稳定性, 使用 ILSVRC2012 (imagenet large scale visual recognition challenge 2012) (Deng 等, 2009) 数据集作为合成噪声和真实噪声的训练集, 并将图像尺寸控制在 256×256 像素到 512×512 像素之间, 随机裁剪尺寸为 128×128 像素的图像块用于训练。

为进一步全面评估本文算法, 采用 3 个公共图像数据集 Kodak24 (Kodak lossless true color image suite 24) (Jao 等, 1999)、BSD300 (Berkeley segmentation dataset 300) (Martin 等, 2001) 和 BSDS500 (Berkeley segmentation dataset 500) (Arbeláez 等, 2011)。这些数据集提供了多样化的图像特征和噪声挑战, 确保了算法的有效性和可扩展性。其中, Kodak24 (Jao 等, 1999) 提供了 24 幅高分辨率图像, BSD300 (Martin 等, 2001) 包含 100 幅在各种场景和拍摄条件下拍摄的多样化测试图像, BSDS500 (Arbeláez 等, 2011) 包含 100 幅涵盖城市环境、建筑、风景和自然场景的图像。评价指标采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性指数 (structural similarity index measure, SSIM), 用于评估在合成噪声和真实噪声上的性能。

对合成噪声, 本文在 Kodak24 (Jao 等, 1999)、BSD300 (Martin 等, 2001) 和 BSDS500 (Arbeláez 等, 2011) 上分别测试 10、15 和 20 次。考虑在 sRGB (standard RGB) 空间中的 4 种噪声类型: $\sigma = 25$ 的高斯噪声、 $\sigma \in [5, 50]$ 的高斯噪声、 $\lambda = 30$ 的泊松噪声、 $\lambda \in [5, 50]$ 的泊松噪声。其中, σ 代表高斯分布的标准差, 较大的标准差表示更强的噪声; λ 表示泊松分布的均值, 较大的均值表示更强的噪声。

为全面评估对真实噪声的去噪能力, 本文使用 PolyU (real world noisy image denoising: new benchmark) (Xu 等, 2018) 数据集, 包括 35 幅涵盖各种室内和室外场景的多样化图像, 为保持一致性和计算效率, 将测试图像按比例调整为 500×375 像素。

2.2 实验细节

本文采用 U-Net 网络结构作为骨干网络,训练过程使用 Adam 优化器,对合成噪声,初始学习率为 0.000 3,对 PolyU (Xu 等, 2018) 和 CMUID (程德强等, 2022) 等真实世界噪声数据集,初始学习率为 0.000 1。批量大小设置为 4,共训练 100 个周期。每 20 个周期后,学习率减半以减轻潜在过拟合。此外,权重衰减设置为 10^{-8} ,以进一步防止过拟合。为提升模型泛化能力,每 10 个周期后都会在对验证集上进行验证。超参数 μ 和 φ 根据经验分别设置为 5 和 3。整个实验设置使用 Python 3.8.5、PyTorch 1.8.1 以及单个 NVIDIA GTX 3080Ti GPU 进行训练。

为确保实验评估的公平性和可靠性,将本文算法与传统去噪算法,如 CBM3D (color image denoising via sparse 3D collaborative filtering) (Dabov 等, 2007b)、有监督去噪算法 DnCNN (denoising convolutional neural network) (Zhang 等, 2017)、Self2Self (self to self with dropout) (Quan 等, 2020)、Ner2N (noisier to noise) (Moran 等, 2019) 以及自监督去噪算法 N2N (noise to noise) (Lehtinen 等, 2018)、Laine19 (Laine 等, 2019)、NBR2NBR (Huang 等, 2021)、B2U (Wang 等, 2022)、B2S (blind to sound) (Wang 等, 2023)、TBSN (Li 等, 2024) 进行比较。其中, DnCNN (Zhang 等, 2017) 仅用于固定的高斯噪声水平去噪,而

Ner2N (Moran 等, 2019) 和 N2N (Lehtinen 等, 2018) 无法处理变化的噪声水平。在真实世界去噪实验中,上述方法均在相同高斯噪声条件,即 $\sigma = 25$ 下进行评估。在矿井图像去噪实验中,将会扩展到不同水平的高斯噪声,从而进行更全面的评估。

2.3 CMUID 数据集的去噪结果

本文在煤矿井下数据集中随机选择了巷道、机电硐室和带式输送机 3 个典型的场景共 100 幅图像作为测试集。

表 1 展示了不同去噪方法在 $\sigma = 25$ 下各场景的 PSNR 和 SSIM,通过对比发现,本文算法始终优于其他方法。例如,对于巷道场景的去噪任务,SSIM 的测试结果比 B2S (Wang 等, 2023) 高 0.3%,在机电硐室中,PSNR 的测试结果比 B2S (Wang 等, 2023) 高 0.6%,且在带式输送机的场景下,两者的测试结果比 B2U (Wang 等, 2022) 提升 0.07 dB/0.004。在该高斯噪声水平下,与其他算法测试结果相比,各场景的 PSNR/SSIM 均保持在一个稳定区间。

表 2 展示了不同去噪方法在 $\sigma = 15$ 、 $\sigma = 25$ 、 $\sigma = 50$ 下各场景的 PSNR/SSIM。在 $\sigma = 25$ 情况下,PSNR 比 B2S (Wang 等, 2023) 增加 0.1%,SSIM 相对于 B2U (Wang 等, 2022) 提升 0.3%;在 $\sigma = 50$ 下,SSIM 比 B2U (Wang 等, 2022)、B2S (Wang 等, 2023) 高 0.005。实验结果表明无论噪声水平高低,本文算法均能自适应噪声信号呈现不均匀分布的场景,并保

表 1 3 个不同场景的矿井图像去噪结果

Table 1 Denoising results of mine images in three different scenarios

方法	巷道		机电硐室		带式输送机	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
CBM3D (Dabov 等, 2007b)	32.54	0.875	30.77	0.883	32.09	0.862
DnCNN (Zhang 等, 2017)	33.23	0.892	31.80	0.825	34.14	0.877
Ner2N (Moran 等, 2019)	32.36	0.845	29.73	0.796	32.01	0.844
Laine19 (Laine 等, 2019)	33.34	0.889	31.87	0.887	34.00	0.863
N2N (Lehtinen 等, 2018)	33.76	0.894	32.30	0.876	32.65	0.837
NBR2NBR (Huang 等, 2021)	33.17	0.893	31.76	0.928	34.33	0.882
B2U (Wang 等, 2022)	33.36	0.893	32.00	0.929	34.44	0.884
B2S (Wang 等, 2023)	33.25	0.899	32.63	0.915	34.15	0.879
TBSN (Li 等, 2024)	32.97	0.886	31.88	0.909	33.86	0.879
本文	33.65	0.902	32.85	0.936	34.51	0.888

注:加粗字体表示各列最优结果。

证最终的图像完整。

图4和图5分别展示了在高斯噪声水平 $\sigma = 25$ 和 $\sigma = 50$ 下矿井图像的视觉去噪效果。可以发现,本文算法去噪后的矿井图像在细节保留和清晰度方面优于其他对比算法。在巷道和机电硐室这类光照对比明显的场景中,本文算法图像质量明显优于对

比算法,说明本文算法能够有效应对矿井图像中如对比度和亮度等特定环境中存在的问题。在带式输送机这种昏暗的场景中,去噪效果仍然保持在对比较算法的前列,进一步表明本文算法具有很强的鲁棒性和适应性,能在昏暗环境下处理矿井图像中存在的特定噪声模式。

表2 3个不同高斯噪声水平下的矿井图像去噪结果

Table 2 Denoising results of mine images at three Gaussian levels

方法	Gaussian $\sigma = 15$		Gaussian $\sigma = 25$		Gaussian $\sigma = 50$	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
CBM3D(Dabov等,2007b)	33.86	0.868	32.94	0.859	28.42	0.814
DnCNN(Zhang等,2017)	32.92	0.833	30.36	0.863	29.82	0.837
Ner2N(Moran等,2019)	33.92	0.871	32.15	0.833	29.31	0.812
Laine19(Laine等,2019)	34.97	0.913	32.54	0.871	29.89	0.836
N2N(Lehtinen等,2018)	33.23	0.887	31.90	0.859	30.43	0.826
NBR2NBR(Huang等,2021)	33.35	0.899	33.20	0.911	29.82	0.845
B2U(Wang等,2022)	35.64	0.941	33.27	0.909	29.96	0.847
B2S(Wang等,2023)	34.69	0.862	32.98	0.900	29.73	0.847
TBSN(Li等,2024)	33.79	0.901	33.41	0.894	29.75	0.836
本文	35.81	0.922	33.45	0.912	31.08	0.852

注:加粗字体表示各列最优结果。

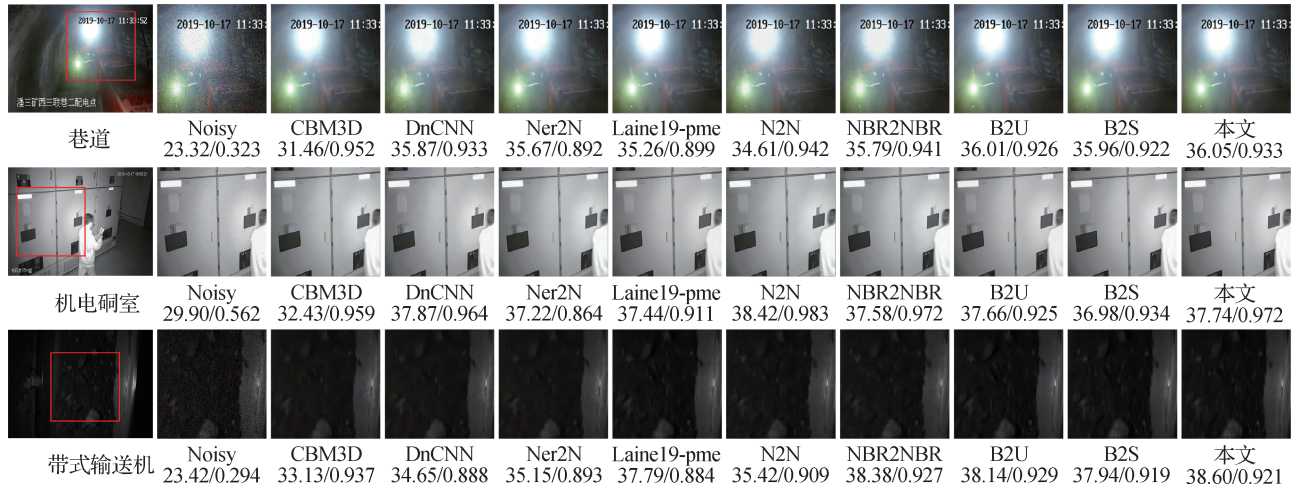


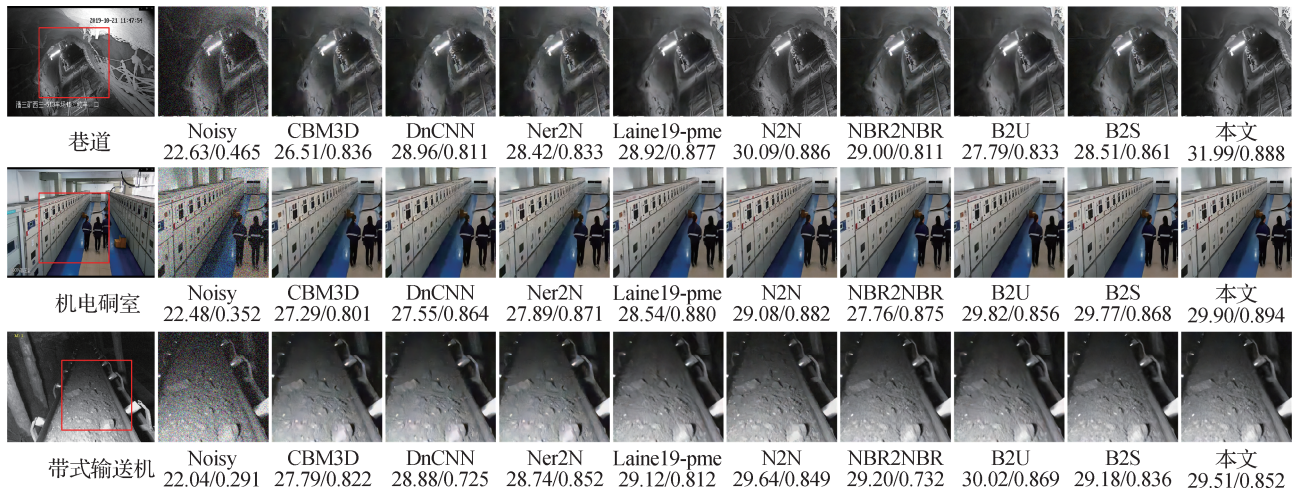
图4 在 $\sigma = 25$ 下的矿井图像可视化去噪结果

Fig. 4 Visual denoising results of mine images at $\sigma = 25$

2.4 合成噪声的去噪结果

合成噪声实验模拟理想情况下的噪声干扰,为算法提供一个可控的测试环境,帮助进一步评估和优化算法,以适应实际矿井环境中遇到的噪声条件,从而有助于深入理解算法在不同噪声水平下的表现

和稳定性。表3—表6展示了在不同高斯和泊松噪声分布下合成图像去噪的定量比较。可以看出,本文算法无论在固定或变化噪声水平下,均具有较好的稳定性。对于Kodak(Jao等,1999)数据集,在 $\sigma = 25$ 时,SSIM比B2U(Wang等,2022)提升0.34%,在

图5 在 $\sigma = 50$ 下的矿井图像可视化去噪结果Fig. 5 Visual denoising results of mine images at $\sigma = 50$ 表3 在高斯噪声 $\sigma = 25$ 的合成噪声去噪结果Table 3 Denoising results of synthesized noise at Gaussian $\sigma = 25$

算法	Kodak24		BSD300		BSD500	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
CBM3D(Dabov等,2007b)	31.87	0.868	30.48	0.861	31.83	0.837
DnCNN(Zhang等,2017)	30.48	0.838	29.91	0.837	30.27	0.846
Self2Self(Quan等,2020)	31.28	0.864	29.86	0.849	30.45	0.837
Ner2N(Moran等,2019)	30.70	0.845	29.32	0.833	30.32	0.832
Laine19-mu(Laine等,2019)	30.32	0.840	28.62	0.803	30.36	0.849
Laine19-pme(Laine等,2019)	32.40	0.883	30.99	0.877	31.45	0.866
N2N(Lehtinen等,2018)	31.80	0.870	31.72	0.874	31.53	0.865
NBR2NBR(Huang等,2021)	32.08	0.879	30.79	0.873	32.09	0.895
B2U(Wang等,2022)	32.27	0.880	30.87	0.872	31.81	0.894
B2S(Wang等,2023)	31.96	0.862	30.96	0.887	32.14	0.884
TBSN(Li等,2024)	32.54	0.874	30.94	0.890	31.87	0.891
本文	32.22	0.883	31.94	0.892	32.26	0.897

注:加粗字体表示各列最优结果。

$\sigma \in [5, 50]$ 时, PSNR和SSIM分别比TBSN(Li等, 2024)高0.36 dB和0.004,表明本文算法在对高分辨率图像进行去噪时,高斯噪声水平的变化并不影响去噪效果,在对比算法中仍处于优势地位。在BSD300(Martin等,2001)数据集上,无论对固定水平的高斯和泊松噪声,还是对变化水平的高斯和泊松噪声,均在数值上取得最优,即面对多样化的场景测试,本文算法均能适应场景变化,保证去噪效果的稳定。对BSD500(Arbeláez等,2011)数据集而言,是对

多样化场景的进一步扩充测试,在4种噪声类型下的PSNR均达到最优,进一步说明了本文算法较高的鲁棒性能。上述测试结果为矿井去噪领域提供了参考标准,确保在实际应用中能有效恢复图像细节,增强图像质量。

为进一步直观分析本文算法的有效性及其泛化能力,图6和图7分别展示了3个数据集在高斯噪声 $\sigma = 25$ 和泊松噪声 $\lambda = 30$ 的定性去噪结果。在整体纹理质量和视觉保真度方面表现出较好的效果,

表4 在高斯噪声 $\sigma \in [5, 50]$ 的合成噪声去噪结果Table 4 Denoising results of synthesized noise at Gaussian $\sigma \in [5, 50]$

算法	Kodak24		BSD300		BSD500	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
CBM3D(Dabov等,2007b)	32.02	0.860	30.56	0.847	29.12	0.913
Self2Self(Quan等,2020)	31.37	0.863	29.87	0.841	30.62	0.857
Laine19-mu(Laine等,2019)	30.52	0.833	28.43	0.794	29.75	0.833
Laine19-pme(Laine等,2019)	32.40	0.854	30.95	0.859	31.24	0.883
NBR2NBR(Huang等,2021)	32.10	0.870	30.73	0.861	32.27	0.886
B2U(Wang等,2022)	32.34	0.872	30.86	0.866	31.72	0.873
B2S(Wang等,2023)	32.19	0.884	31.29	0.879	31.91	0.884
TBSN(Li等,2024)	33.01	0.889	31.94	0.881	31.97	0.887
本文	33.37	0.893	32.17	0.883	32.49	0.891

注:加粗字体表示各列最优结果。

表5 在泊松噪声 $\lambda = 30$ 的合成噪声去噪结果Table 5 Denoising results of synthesized noise at Poisson $\lambda = 30$

算法	Kodak24		BSD300		BSD500	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Self2Self(Quan等,2020)	30.31	0.857	28.93	0.840	30.51	0.857
Ner2N(Moran等,2019)	30.12	0.844	27.36	0.798	28.32	0.806
Laine19-mu(Laine等,2019)	30.19	0.833	28.25	0.794	30.14	0.817
Laine19-pme(Laine等,2019)	31.67	0.874	30.25	0.866	30.47	0.858
N2N(Lehtinen等,2018)	26.68	0.855	27.21	0.860	27.50	0.859
NBR2NBR(Huang等,2021)	31.44	0.870	30.10	0.863	31.47	0.888
B2U(Wang等,2022)	31.64	0.871	30.25	0.862	31.14	0.886
B2S(Wang等,2023)	31.02	0.868	30.89	0.879	30.96	0.876
TBSN(Li等,2024)	31.39	0.870	31.06	0.877	31.44	0.879
本文	31.52	0.873	31.27	0.882	31.63	0.889

注:加粗字体表示各列最优结果。

表明本文算法在有效恢复细粒度细节的同时能够抑制噪声伪影。同时本文算法与其他算法相比能有效保证在固定噪声水平下的去噪图像质量,并对合成噪声具有较强的泛化能力。

2.5 PolyU数据集的去噪结果

由于真实噪声往往更加复杂和不可预测,本文除了采用煤矿井下图像数据之外,额外采用真实噪声数据评估算法在处理实际矿井图像时的鲁棒性和泛化能力。通过真实世界的去噪实验以验证算法在实际矿井环境中的适应性,评估算法是否能够有效

应对这些复杂的噪声模式,确保算法在实际应用中的可靠性。

表7定量比较了不同去噪方法在PolyU(Xu等,2018)数据集上的去噪性能,根据拍摄时使用的相机型号将图像分类为不同的子集:Sony A7II、Nikon 800、Canon 600、Canon 80D和Canon 5D2。算法能在不同成像系统上进行更细致的分析,在Nikon 800、Canon 600、Canon 80D相机型号下的去噪效果在对比算法中均达到最优,尤其是对由Canon 600和Canon 80D相机拍摄的图像,相比次优结果,PSNR分

表6 在泊松噪声 $\lambda \in [5, 50]$ 的合成噪声去噪结果Table 6 Denoising results of synthesized noise at Poisson $\lambda \in [5, 50]$

算法	Kodak24		BSD300		BSD500	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Self2Self(Quan等,2020)	29.06	0.834	28.15	0.817	29.43	0.837
Laine19-mu(Laine等,2019)	29.76	0.820	27.89	0.778	29.81	0.826
Laine19-pme(Laine等,2019)	30.88	0.850	29.57	0.841	30.96	0.853
NBR2NBR(Huang等,2021)	30.86	0.855	29.54	0.843	30.86	0.867
B2U(Wang等,2022)	31.07	0.857	29.92	0.852	30.63	0.874
B2S(Wang等,2023)	30.77	0.851	29.97	0.847	31.05	0.865
TBSN(Li等,2024)	30.79	0.847	30.41	0.861	30.89	0.861
本文	30.81	0.854	30.62	0.866	31.14	0.877

注:加粗字体表示各列最优结果。

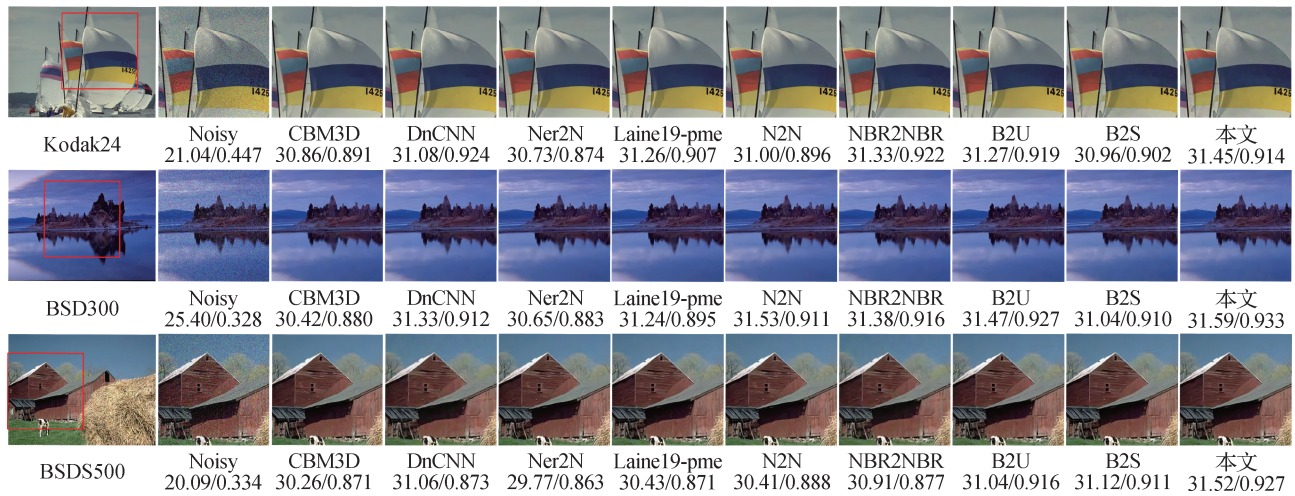
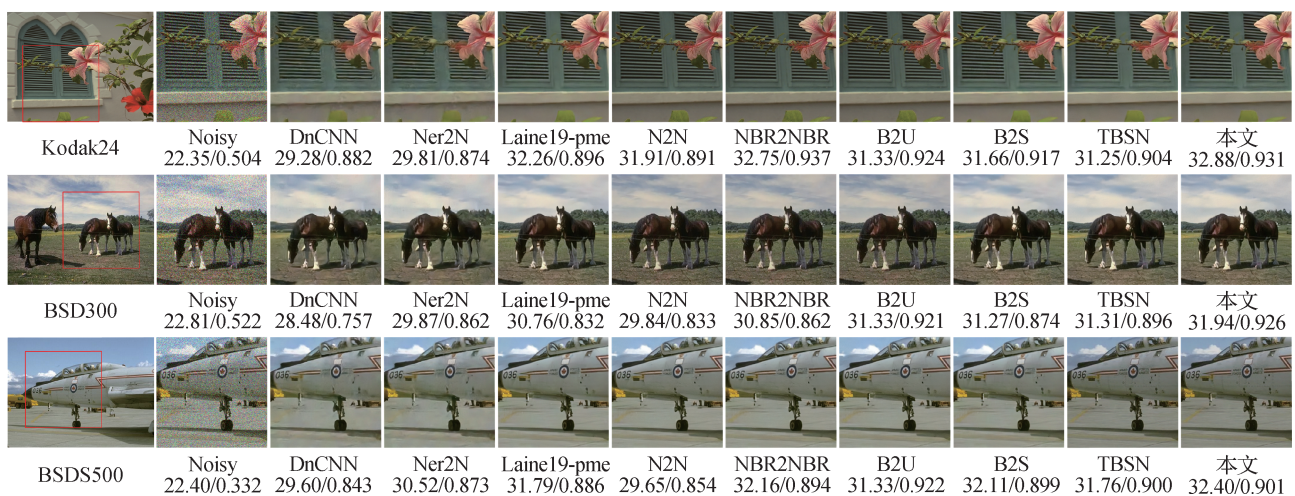
图6 在 $\sigma = 25$ 下的合成噪声可视化去噪结果Fig. 6 Visual denoising results of synthesized noise at $\sigma = 25$ 图7 在 $\lambda = 30$ 下的合成噪声可视化去噪结果Fig. 7 Visual denoising results of synthesized noise at $\lambda = 30$

表7 真实噪声的去噪结果

Table 7 Denoising results of real-world noise

方法	Sony A7II 相机		Nikon 800 相机		Canon 600 相机		Canon 80D 相机		Canon 5D2 相机	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
CBM3D(Dabov 等,2007b)	33.17	0.869	33.23	0.862	33.96	0.867	32.47	0.867	32.24	0.867
DnCNN(Zhang 等,2017)	34.69	0.932	33.32	0.920	34.75	0.926	33.52	0.925	33.72	0.932
Ner2N(Moran 等,2019)	33.52	0.883	32.46	0.857	33.44	0.892	32.73	0.889	32.87	0.889
Laine19(Laine 等,2019)	34.42	0.916	33.96	0.914	34.52	0.913	33.51	0.911	33.14	0.926
N2N(Lehtinen 等,2018)	32.50	0.885	32.43	0.888	32.41	0.894	31.79	0.883	32.67	0.915
NBR2NBR(Huang 等,2021)	34.77	0.935	33.39	0.925	34.80	0.928	33.57	0.929	33.69	0.933
B2U(Wang 等,2022)	34.61	0.932	34.85	0.931	35.09	0.932	34.16	0.930	34.92	0.933
B2S(Wang 等,2023)	34.29	0.935	32.49	0.905	35.91	0.934	33.74	0.929	34.69	0.915
TBSN(Li 等,2024)	35.61	0.932	34.71	0.931	35.27	0.911	34.19	0.930	34.55	0.904
本文	34.82	0.938	34.91	0.933	36.25	0.942	34.73	0.924	34.97	0.936

注:加粗字体表示各列最优结果。

别提升0.94%和1.57%,对由Canon 600和Canon 5D2相机拍摄的图像,相比次优结果,SSIM增加0.85%和0.32%,这不仅凸显出本文算法在降低噪声水平方面的高效性,而且证明在提取图像语义信息方面表现出较强能力。

图8展示了在PolyU(Xu等,2018)数据集上的可视化去噪结果。在Canon 5D2拍摄的机器芯片图像中,本文算法能够更清晰地恢复细节;在Canon 600拍摄的玩具图像中,相较于其他算法,本文算法能更好地恢复颜色对比度;在Canon 80D拍摄的排球图

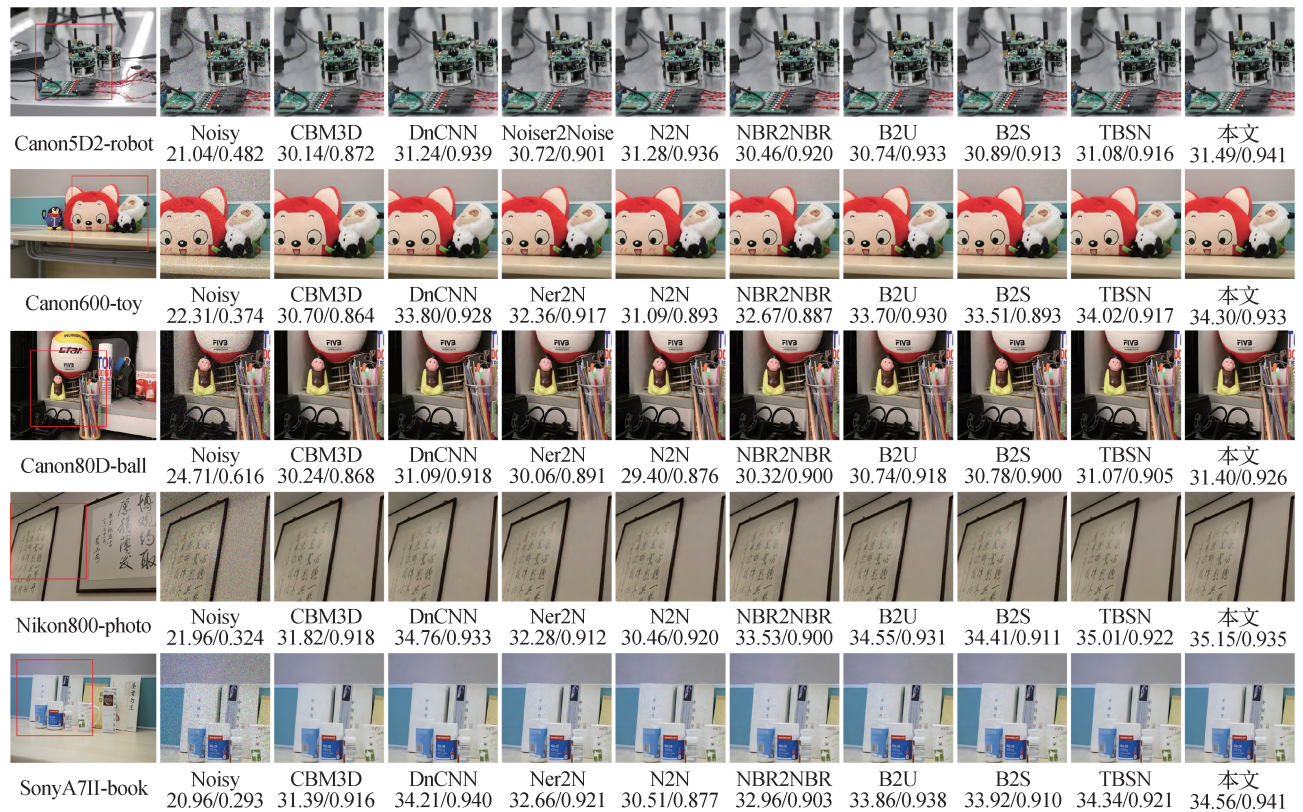


图8 真实噪声的可视化去噪结果

Fig. 8 Visual denoising results of real-world noise

像中,本文算法能更好地恢复排球上的字母;在 Nikon 800 拍摄的字画中,本文算法能更清晰地恢复字画中的细节信息;在 Sony A7II 拍摄的药品名称中,相较于其他算法,本文算法能更好地恢复药品名称的整体结构和细节信息。通过对比分析,本文算法在真实噪声去噪任务中表现出色,能有效减少图像噪声,同时保持色彩自然和细节清晰,不仅在整体图像质量方面超越了其他对比算法,并能有效完成由多样化噪声特性主导的矿井场景去噪任务。

2.6 消融实验

为进一步证明集成后的图像对网络模型的去噪影响,本文采用未集成的图像作为训练标签,改变原有的损失函数,具体为

$$L' = E \|f_{\theta}(\mathbf{y}) - \mathbf{y}\|_2^2 + \varphi E \|f_{\theta}(\Omega_{\mathbf{y}}^{\mu}) - \mathbf{y}\|_2^2 \quad (8)$$

通过对煤矿井下数据集重新训练,并对 Kodak24(Jao 等,1999)数据集进行测试,可得到两种不同损失函数的定量比较结果,如表 8 所示。可以看出无论在固定或是可变的噪声水平下,使用 L 的去噪效果均比 L' 的去噪效果好,例如在 $\sigma \in [5, 50]$ 下,PSNR 和 SSIM 比 L' 高出 4.00 dB/0.081;在 $\lambda \in [5, 50]$ 下,比 L' 高出 2.05 dB/0.107。说明选择性处理掩码区域在生成多样化训练样本的同时也

会导致部分信息丢失,而未掩码区域可以保留图像的细节信息,集成掩码与未掩码区域的图像恰好可以弥补两者的不足,有利于引导网络更有效去除噪声,恢复图像真实性。

此外,为进一步探究掩码对最终去噪图像的影响,本文通过改变超参数 φ 的值来调节掩码区域的比例,进而评估掩码对于整个去噪过程的影响。表 9 展示了在 Kodak24(Jao 等,1999)数据集上的测试结果。可以看出,合适比例的掩码能够在有限的图像数据上生成更多样化的训练样本,防止过多像素掩码造成的噪声结构变化和信丢失,引发图像失真。综合分析,本文设置 $\varphi = 3$,用于调节掩码区域的比例,可以更好地平衡掩码区域对去噪效果的影响。

表 10 展示了在 Kodak24(Jao 等,1999)数据集上不同超参数 μ 的消融实验结果。无论固定或变化的噪声水平、无论高斯噪声或泊松噪声,实验结果表明,当 $\mu = 5$ 时,算法的去噪性能达到最优。具体而言,在 $\sigma = 25$ 条件下, $\mu = 5$ 时的 PSNR 和 SSIM 分别比 $\mu = 2$ 时提升 4.21 dB 和 0.088。在 $\lambda = 30$ 条件下, $\mu = 5$ 时的 PSNR 和 SSIM 分别比 $\mu = 2$ 时提升 3.76 dB 和 0.085。通过综合对比分析,本文选择 $\mu = 5$ 作为调整集成图像与噪声图像比例的最佳参数,以确保算法在不同噪声条件下的最优去噪效果。

表 8 不同损失函数的消融实验结果

Table 8 Results of ablation with different loss functions

损失函数	$\sigma = 25$		$\sigma \in [5, 50]$		$\lambda = 30$		$\lambda \in [5, 50]$	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
L	32.22	0.883	33.37	0.893	31.52	0.873	30.81	0.854
L'	30.04	0.833	29.37	0.812	30.01	0.809	28.76	0.747

注:加粗字体表示各列最优结果。

表 9 不同超参数 φ 的消融实验结果

Table 9 Results of ablation with different hyperparameter φ

φ	$\sigma = 25$		$\sigma \in [5, 50]$		$\lambda = 30$		$\lambda \in [5, 50]$	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
0	29.44	0.818	29.70	0.825	28.65	0.807	27.92	0.814
1	30.07	0.852	29.03	0.836	30.05	0.825	30.21	0.805
3	32.22	0.883	33.37	0.893	31.52	0.873	30.81	0.854
5	32.01	0.868	32.41	0.875	32.09	0.862	30.89	0.854

注:加粗字体表示各列最优结果。

表10 不同超参数 μ 的消融实验结果Table 10 Results of ablation with different hyperparameter μ

μ	$\sigma = 25$		$\sigma \in [5, 50]$		$\lambda = 30$		$\lambda \in [5, 50]$	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
2	28.01	0.795	28.05	0.782	27.76	0.788	27.43	0.762
3	30.26	0.822	29.74	0.810	30.56	0.845	30.16	0.823
4	31.52	0.843	32.41	0.824	31.09	0.852	30.44	0.837
5	32.22	0.883	33.37	0.875	31.52	0.873	30.81	0.854

注:加粗字体表示各列最优结果。

表11展示了在Kodak24(Jao等,1999)数据集上不同掩码方式的消融实验结果。在所测试的噪声条件下,自适应掩码方式的去噪效果始终优于边缘掩码和角点掩码方式。在 $\sigma \in [5, 50]$ 噪声水平下,通过自适应掩码去噪后的图像PSNR和SSIM分别比边缘掩码方式高3.15 dB和0.094,比角点掩码方式

高3.18 dB和0.088。在 $\lambda \in [5, 50]$ 噪声水平下,自适应掩码方式的PSNR和SSIM分别比边缘掩码方式提升1.67 dB和0.047,比角点掩码方式提升1.04 dB和0.054。通过对比分析,本文提出的自适应掩码方法能够有效去除图像噪声,同时确保图像结构的完整性和细节的保留。

表11 不同掩码方式的消融实验结果

Table 11 Results of ablation with different masking methods

掩码方式	$\sigma = 25$		$\sigma \in [5, 50]$		$\lambda = 30$		$\lambda \in [5, 50]$	
	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
边缘掩码	30.19	0.796	30.22	0.799	31.03	0.811	29.14	0.807
角点掩码	30.11	0.799	30.19	0.805	29.97	0.806	29.77	0.800
自适应掩码	32.22	0.883	33.37	0.893	31.52	0.873	30.81	0.854

注:加粗字体表示各列最优结果。

3 结论

针对受粉尘和光照影响导致图像产生不同水平噪声的问题,本文提出一种基于自适应掩码的矿井图像自监督去噪算法,主要由自适应掩码、掩码集成和自适应集成损失3部分组成。1)设计自适应掩码,对各块中心像素外的边缘像素及角点像素实施掩码操作,增加了数据多样性,同时避免了恒等映射;2)设计掩码集成,对掩码区域重新组合成原始图像尺寸,集成网络输出与掩码区域,增强了网络对细节信息的关注度,同时综合考虑了局部信息和全局信息之间的相互作用;3)设计自适应集成损失训练网络模型,仅用单幅噪声图像即可实现神经网络在无标注数据中的有效特征表示学习,完成不同噪声水平甚至是不同场景下的去噪任务;4)实验结果表明,在煤矿井下图像数据集和公共数据集均取得较

好的去噪效果,能够恢复图像的整体信息并保证细节纹理信息的完整,不仅为矿井图像去噪算法的开发和优化提供了经验和数据支持,而且能提高算法在实际矿井环境中的实用性和可靠性,使其更好地服务于煤矿安全生产。

然而,本文算法难以同时处理不同尺度的特征,在高分辨率图像中,小而重要的结构可能会被忽略或误分类为噪声。此外,为确保掩码策略的准确性和有效性,需要大量数据集进行预训练,这一要求对计算资源提出很高的要求。未来将通过使用轻量级网络架构或快速近似算法减少不必要的计算负担,确保掩码策略能够在移动设备或其他嵌入式平台上快速运行。

参考文献(References)

Arbeláez P, Maire M, Fowlkes C and Malik J. 2011. Contour detection

- and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5): 898-916 [DOI: 10.1109/TPAMI.2010.161]
- Batson J and Royer L. 2019. Noise2self: blind denoising by self-supervision//*Proceedings of the 36th International Conference on Machine Learning*. Long Beach, USA: PMLR: 524-533
- Buades A, Coll B and Morel J M. 2005. A non-local algorithm for image denoising//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. San Diego, USA: IEEE: 60-65 [DOI: 10.1109/CVPR.2005.38]
- Cao X C, Zhao S C, Wu A M, Liang S Y and Wang L Y. 2025. Challenges and progress of object detection technology in open environments. *Journal of Image and Graphics*, 30(6): 1616-1637 (操晓春, 赵思成, 武阿明, 梁思源, 王立元. 2025. 目标检测技术在开放环境中的挑战与进展. *中国图象图形学报*, 30(6): 1616-1637) [DOI: 10.11834/jig.250004]
- Chen Y J and Pock T. 2017. Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1256-1272 [DOI: 10.1109/tpami.2016.2596743]
- Cheng D Q, Chen J, Kou Q Q, Nie S J and Zhang J Y. 2022. Lightweight super-resolution reconstruction method based on hierarchical features fusion and attention mechanism for mine image. *Chinese Journal of Scientific Instrument*, 43(8): 73-84 (程德强, 陈杰, 寇旗旗, 聂帅杰, 张剑英. 2022. 融合层次特征和注意力机制的轻量化矿井图像超分辨率重建方法. *仪器仪表学报*, 43(8): 73-84) [DOI: 10.19650/j.cnki.cjsi.J2209774]
- Dabov K, Foi A, Katkovnik V and Egiazarian K. 2007a. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8): 2080-2095 [DOI: 10.1109/tip.2007.901238]
- Dabov K, Foi A, Katkovnik V and Egiazarian K. 2007b. Color image denoising via sparse 3D collaborative filtering with grouping constraint in luminance-chrominance space//*Proceedings of 2007 IEEE International Conference on Image Processing*. San Antonio, USA: IEEE: I-313-I-316 [DOI: 10.1109/ICIP.2007.4378954]
- Deng J, Dong W, Socher R, Li L J, Li K and Li F F. 2009. ImageNet: a large-scale hierarchical image database//*Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami, USA: IEEE: 248-255 [DOI: 10.1109/CVPR. 2009.5206848]
- Fan H W, Zhang C, Cao X G, Liu J P, Zhang X H and Zhao H. 2024. An image dust removal and enhancement method in low illumination environment based on dark-bright channel segmentation and fusion. *Journal of China Coal Society*, 49(4): 2167-2178 (樊红卫, 张超, 曹现刚, 刘金鹏, 张旭辉, 赵寒. 2024. 一种基于暗亮通道分割融合的低照度环境图像去尘雾及增强方法. *煤炭学报*, 49(4): 2167-2178) [DOI: 10.13225/j.cnki.jccs.2023.0576]
- Guo S, Yan Z F, Zhang K, Zuo W M and Zhang L. 2019. Toward convolutional blind denoising of real photographs//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE: 1712-1722 [DOI: 10.1109/CVPR.2019.00181]
- Huang T, Li S J, Jia X, Lu H C and Liu J Z. 2021. Neighbor2neighbor: self-supervised denoising from single noisy images//*Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Nashville, USA: IEEE: 14776-14785 [DOI: 10.1109/CVPR46437.2021.01454]
- Jao C S, Hier D B and Brint S U. 1999. The display of photographic-quality images on the web: a comparison of two technologies. *IEEE Transactions on Information Technology in Biomedicine*, 3(1): 70-73 [DOI: 10.1109/4233.748977]
- Jebur R S, Zabil M H B M, Hammood D A and Cheng L K. 2024. A comprehensive review of image denoising in deep learning. *Multimedia Tools and Applications*, 83(20): 58181-58199 [DOI: 10.1007/s11042-023-17468-2]
- Krull A, Buchholz T O and Jug F. 2019. Noise2void-learning denoising from single noisy images//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Long Beach, USA: IEEE: 2124-2132 [DOI: 10.1109/CVPR. 2019.00223]
- Laine S, Karras T, Lehtinen J and Aila T. 2019. High-quality self-supervised deep image denoising//*Proceedings of 2019 Annual Conference on Neural Information Processing Systems*. Vancouver, Canada: NeurIPS: 6968-6978
- Lehtinen J, Munkberg J, Hasselgren J, Laine S, Karras T, Aittala A and Aila T. 2018. Noise2Noise: learning image restoration without clean data//*Proceedings of the 35th International Conference on Machine Learning*. Stockholm, Sweden: PMLR: 2965-2974
- Lempitsky V, Vedaldi A and Ulyanov D. 2018. Deep image prior//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake, USA: IEEE: 9446-9454 [DOI: 10.1109/CVPR.2018.00984]
- Li J Y, Zhang Z L and Zuo W M. 2024. TBSN: transformer-based blind-spot network for self-supervised image denoising [EB/OL]. [2024-11-01]. <https://arxiv.org/pdf/2404.07846.pdf>
- Li X Y, Wang L H, Zhou Y C and Zhang J. 2022. Color image denoising using adaptive non-local 3D total variation. *Journal of image and graphics*, 27(12): 3450-3460 (李潇瑶, 王炼红, 周怡聪, 章兢. 2022. 自适应非局部3维全变分彩色图像去噪. *中国图象图形学报*, 27(12): 3450-3460) [DOI: 10.11834/jig.210579]
- Liu F, Cao W J, Zhang J M, Cao G M and Guo L F. 2021. Current technological innovation and development direction of the 14th Five-Year Plan period in China coal industry. *Journal of China Coal Society*, 46(1): 1-15 (刘峰, 曹文君, 张建明, 曹光明, 郭林峰. 2021. 我国煤炭工业科技创新进展及“十四五”发展方向. *煤炭学报*, 46(1): 1-15) [DOI: 10.13225/j.cnki.jccs.2021.0042]
- Mao S J, Cui J J and Wang S B. 2020. Construction of information shar-

- ing platform of mine safe production for intelligent mining. *Journal of China Coal Society*, 45(6): 1937-1948 (毛善君, 崔建军, 王世斌. 2020. 煤矿智能开采信息共享管理平台构建研究. *煤炭学报*, 45(6): 1937-1948) [DOI: 10.13225/j.cnki.jccs.ZN20.0341]
- Martin D, Fowlkes C, Tal D and Malik J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics//Proceedings of the 8th IEEE International Conference on Computer Vision. ICCV 2001. Vancouver, Canada: IEEE: 416-423 [DOI: 10.1109/ICCV.2001.937655]
- Moran N, Schmidt D, Zhong Y and Coady P. 2019. Noisier2Noise: learning to denoise from unpaired noisy data//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 12061-12069 [DOI: 10.1109/CVPR42600.2020.01208]
- Quan Y H, Chen M Q, Pang T Y and Ji H. 2020. Self2Self with dropout: learning self-supervised denoising from single image//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE: 1887-1895 [DOI: 10.1109/CVPR42600.2020.00196]
- Wei X, Yang G Q, Li J, Lu Y and Shi L. 2020. Multiscale information-based conditional generative adversarial nets for downhole track detection. *Journal of Image and Graphics*, 25(2): 0282-0293 (卫星, 杨国强, 李佳, 陆阳, 石雷. 2020. 结合多尺度条件生成对抗网络的井下轨道检测. *中国图象图形学报*, 25(2): 0282-0293). [DOI: 10.11834/jig.190119]
- Wang G F. 2022. New technological progress of coal mine intelligence and its problems. *Coal Science and Technology*, 50(1): 1-27 (王国法. 2022. 煤矿智能化最新技术进展与问题探讨. *煤炭科学技术*, 50(1): 1-27) [DOI: 10.13199/j.cnki.cst.2022.01.001]
- Wang P, Song X N, Wu X J and Yu D J. 2020. Multi-type features network for person re-identification. *Pattern Recognition and Artificial Intelligence*, 33(10): 879-888 (王鹏, 宋晓宁, 吴小俊, 於东军. 2020. 用于行人重识别的多类型特征网络. *模式识别与人工智能*, 33(10): 879-888) [DOI: 10.16451/j.cnki.issn1003-6059.202010002]
- Wang Z J, Liu J Z, Li G Q and Han H. 2023. Blind2Unblind: self-supervised image denoising with visible blind spots//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, USA: IEEE: 2017-2026 [DOI: 10.1109/CVPR52688.2022.00207]
- Wang Z J, Liu J Z, Zhai H and Han H. 2023. Blind2Sound: self-supervised image denoising without residual noise [EB/OL]. [2024-11-01]. <https://arxiv.org/pdf/2303.05183.pdf>
- Xu J, Li H, Liang Z T, Zhang D and Zhang L. 2018. Real-world noisy image denoising: a new benchmark [EB/OL]. [2024-11-01]. <https://arxiv.org/pdf/1804.02603.pdf>
- Zhang K, Zuo W M and Zhang L. 2018. FFDNet: toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, 27(9): 4608-4622 [DOI: 10.1109/tip.2018.2839891]
- Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7): 3142-3155 [DOI: 10.1109/tip.2017.2662206]

作者简介

张亚茹,女,讲师,主要研究方向为人工智能、计算机视觉和计算机图形学。E-mail:yrzhang1014@163.com

刘健涛,男,硕士研究生,主要研究方向为矿井工程图像去噪。E-mail:ljt08252022@163.com

许国庆,男,硕士研究生,主要研究方向为图像识别与小目标检测。E-mail:15955506334@163.com

郝定溢,男,讲师,主要研究方向为煤矿智能开采。E-mail:hdycomt@163.com