

# Multiscale Convolutional Network for Repairing Coal Slime Foam Images

Xianwu Huang,\* Yuxiao Wang, Haili Shang, and Jinshan Zhang



Cite This: *ACS Omega* 2023, 8, 9547–9554

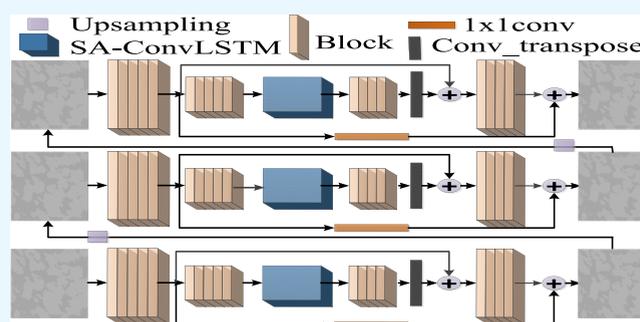


Read Online

ACCESS |

Metrics & More

Article Recommendations



**ABSTRACT:** Visual feature information regarding flotation foam is crucial for the flotation process. Owing to a large amount of noise and blur in the foam images collected in the flotation field, feature extraction and segmentation of foam images pose considerable challenges. Furthermore, the visual properties of foam are strongly correlated with current flotation conditions. Therefore, this study presents a method to repair blurred pixels in foam images. In addition to enhancing the image dataset necessary for network model training, the restored images can provide high-quality images extracting foam-feature information. In addition, this research presents a novel fifth-order residual structure that enlarges the network structure by stacking, enhancing the learning ability of complex networks. Experimental results demonstrate that the suggested method can achieve a satisfactory repair effect for foam images under various blurring conditions, laying a foundation for guiding the intelligent adjustment of flotation field parameters.

## 1. INTRODUCTION

Foam flotation is one of the most efficient processes for coal separation.<sup>1–3</sup> It is a process of mineral separation that involves adding flotation reagents to coal slurry so that hydrophobic coal particles adhere to the bubbles while hydrophilic gangue remains in the slurry.<sup>4</sup> During the flotation process, the onsite staff artificially adjust the factors affecting flotation production. These factors include the addition of flotation reagents and the amount of aeration by observing the characteristic information of the foam surface.<sup>5–8</sup> However, owing to subjective bias, production environment changes, and other factors, the findings of manual judgment are frequently inaccurate, causing frequent fluctuations in flotation production indicators and low utilization of mineral resources.<sup>7</sup> Consequently, using machine vision to capture the characteristic information of flotation foam, to monitor and identify the foam in the flotation process in real time, and to guide the adjustment of production factors, such as reagent dosage and liquid level, affecting flotation production increases the yield and economic benefits of the flotation concentrate in a coal washing plant.<sup>8–10</sup>

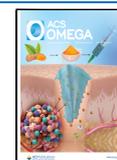
In addition to information on static features such as color and shape, flotation foam images contain information on

dynamic feature information such as foam velocity and stability.<sup>3</sup> In the flotation process, rising and floating of bubbles, collision between materials and bubbles, stirring of the flotation machine, and other factors usually cause many irregular movements of bubbles, affording considerable noise and blurring in the collected foam images.<sup>11,12</sup> This complicates the extraction of features from foam images as well as necessitates the screening of collected foam images when learning-based methods are employed to analyze flotation foam images. After removing fuzzy low-quality images, the foam dataset for model training is created, which influences the final performance of the model. In this regard, the CNN-SVM hybrid training model has been proposed for classifying and identifying coal slime flotation froth images. Herein, multiple images were gathered sequentially for each

**Received:** December 31, 2022

**Accepted:** February 20, 2023

**Published:** March 1, 2023



ash label. From these images, a subset was selected as the dataset corresponding to each ash label. Correspondingly, the accuracy rate reached 87.66% for froth classification, and it was concluded that a higher accuracy may be achieved by acquiring more image datasets.<sup>13</sup> As such, the network described herein uses the pixel information in the foam image to repair the blurred and damaged pixels in the image, yielding a more detailed image and resolving the difficulties associated with using the foam images collected onsite to construct the dataset and extract the feature information.

In recent years, multiscale Convolutional Neural Networks (multiscale CNNs) have been applied to a variety of computer vision problems and have achieved good results in a variety of image tasks.<sup>14–26</sup> Li et al. used multiscale fully CNNs for feature extraction from high-resolution remote sensing images and for change detection tasks,<sup>15</sup> Yao et al. proposed an end-to-end multiscale convolutional neural network (MSCNN) for extracting features of objects at different scales in HRS images,<sup>16</sup> and Li and Yu found that the multiscale features extracted using the CNN have high visual saliency, so they proposed a new visual saliency model, combining the CNN saliency model with the spatial consistency model and image segmentation into a complete framework for saliency detection,<sup>17</sup> and their experiments achieved good results.

Image restoration is an important application area and a challenging topic for multiscale CNNs,<sup>18–27</sup> where image restoration plays an important role in processing, understanding, and representing images. Using a multiscale CNN, when restoring the scratches on old images,<sup>18</sup> it is possible to collect image semantic information while repairing the image texture, affording a more realistic photo repair result. To fix the hues and contours of damaged murals,<sup>20</sup> the most similar relevant aspects can be selected. When repairing the occluded face image,<sup>22</sup> the high-level information of the image is used to maintain image consistency and the low-level feature information is used to fill in the texture. For the restoration of blurred coal slurry foam images, Xiao et al. screened the foam images for deblurring and then performed image enhancement, followed by deblurring of the processed images using the FS-DeblurGAN network.<sup>28</sup> Liu et al. used the MsD model to increase the extraction of image feature information without estimating the blurring kernel using a multiscale architecture for the restoration of blurred foam image restoration work.<sup>29</sup>

This study offers a coal slime flotation foam image repair network based on an analysis of the coal slime flotation process and foam image features to address the difficulties associated with precisely and effectively extracting foam features from fuzzy damaged foam images. The contributions of this study include the following. (1) The network described herein repairs the blurred and damaged pixels in the foam image obtained from the scene, yielding a more detailed image to construct the foam image dataset and extract the feature information. (2) A five-ordered residual structure is designed in the network, and  $1 \times 1$  conv. is performed on the parts connected by summation for each order. The residual structure is stacked to form an encoder–decoder structure, and a long short-term memory network with a self-attention mechanism module is connected between the two so that the entire network can be designed by stacking as well as perform complex learning in time and space dimensions to make better use of interrelated feature information to repair images. (3) This study also presents a method for repairing the blurred

pixels in a foam image while maintaining a high level of image detail. The resulting high-quality image facilitates the extraction of foam visual feature information, laying the groundwork for the intelligent adjustment of flotation field parameters.

This paper is organized as follows: in Section 2, the content related to the coal slime foam image repair network is introduced. In Section 3, the proposed method foam image repair network will be described in detail. In Section 4, the effectiveness of the image repairing method is experimentally verified from qualitative and quantitative aspects, and the repairing effect is discussed for the experiment. In Section 5, this work is summarized and provides an outlook on future research topics.

## 2. RELATED WORK

**Residual structure.** The network architecture becomes difficult to design as the number of network layers increases. To address this challenge, the ResNets network<sup>30</sup> adopts the method of stacking blocks of the same shape in the VGG network.<sup>31</sup> The stacked ResBlock structure<sup>30</sup> facilitates the network ability to achieve accuracy gains from the considerably increased network depth through skip connections. Subsequently, the ResNeXt structure,<sup>36</sup> based on the stacking technique, sums and aggregates the outputs of each module to achieve various transformations without redesigning the architecture. While preserving complexity, the aggregate transformation is found to be superior to the original ResBlock structure. When constructing a network using the Inception series architecture,<sup>32–35</sup> the number and size of filters should be carefully designed. However, it is often difficult to determine how to design for greater precision owing to factors such as network structures and hyperparameters. Through the study of networks designed using residual structures<sup>30,36</sup> and Inception series architecture<sup>32–35</sup> when dealing with image tasks such as image recognition and classification, it has been discovered that the presence of jump connections in a network can make the network produce good results in image-processing tasks.

**Long short-term memory network.** The long short-term memory network (LSTM) is a special type of recurrent neural network that may acquire long-term-dependent information. Convolutional LSTM (ConvLSTM)<sup>37</sup> comprises a coding network and a prediction network with multiple ConvLSTM layers stacked on top of one another. The final prediction is to concatenate and generate all the states in the prediction network via  $1 \times 1$  conv. To better capture the temporal and spatial correlation of feature information, the nested LSTM (NLSTM) architecture<sup>38</sup> proposes to increase the depth of LSTM by nesting. In this architecture, the external storage unit can selectively read and write related long-term information to its internal unit. Compared to the stacked LSTM, it can process events on a longer time scale and is more suited for learning long-term dependencies. Self-attention ConvLSTM (SA-ConvLSTM) is incorporated into ConvLSTM as a result of the self-attention mechanism.<sup>39</sup> The constructed memory module is capable of aggregating past-related image feature information and determining the temporal and spatial relation between image feature information.

**Multiscale convolutional neural network.** Typical multiscale convolutional neural networks define three scales in the decreasing order of image resolution. The network with the greatest size generates the coarsest level of clear images, which

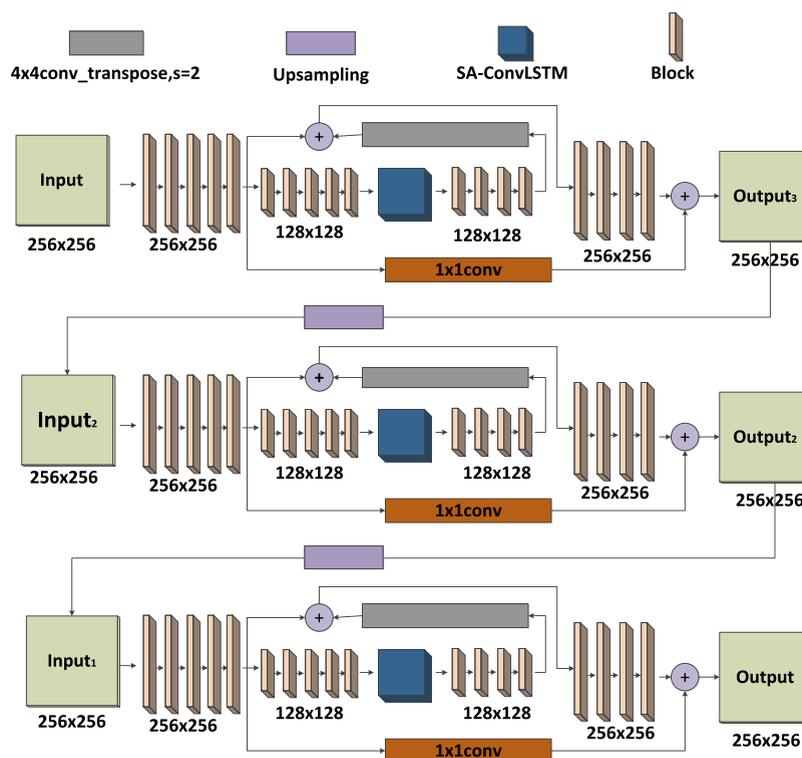


Figure 1. Architecture of the proposed multiscale convolutional network.

are then used as the input by a network with a finer scale to generate images with a finer level of clarity. By directly stacking 19 modified residual structure extension networks, a multiscale loss function is employed to converge the training loss of the multiscale CNN,<sup>24</sup> allowing the image to be repaired without estimating the blur kernel. The scale-recurrent network (SRN)<sup>26</sup> enhances the stability of the network by sharing network weights across scales. The encoder–decoder ResBlock structure can receive additional image feature information for image inpainting. MPRNet<sup>40</sup> is a network that exchanges information between multiple scales. Each scale network contains an adaptive design in which each pixel reweights local features. The encoder–decoder architecture of MPRNet<sup>40</sup> combines cross-scale and finer-grained networks after image feature learning. The network<sup>24,26</sup> primarily uses spatial image feature information for image restoration. MPRNet<sup>40</sup> demonstrates that the network can better repair images by combining temporal and spatial feature information.

### 3. NETWORK ARCHITECTURE

The architecture of the proposed multiscale convolutional network is shown in Figure 1. The network is divided into three scales according to the order of image resolution reduction. A coarse-to-fine strategy is adopted. The maximum-scale network generates the coarsest level of clear images, which are transmitted as input to the next finer-scale network to generate finer clear images. Simultaneously, the training efficiency is improved by weight sharing and recursive structures between different scales.

**3.1. Residual Structure.** The residual structure of this paper is shown in Figure 2. Based on the residual structure,<sup>27</sup> the residual function is extended to the fifth order and the original one-time jump connection is redesigned to conduct a five-time jump connection. The input portion is performed 1 ×

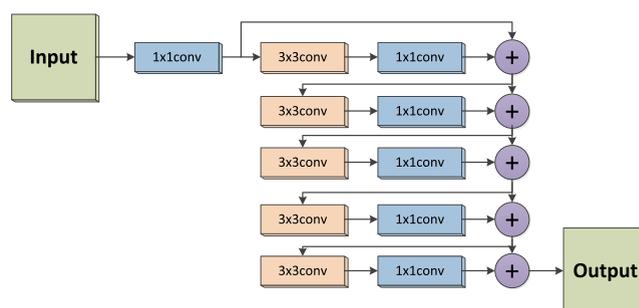


Figure 2. Residual block structure used herein.

1 conv. before the first feature summation. After each feature summation, 3 × 3 conv. features are passed, followed by 1 × 1 conv. features before the following feature summation. Redesigning the network structure using a stacking technique for residual structures enhances the ability of the network to perform complex learning. In this network, the convolution kernel size of the residual structure is set to 3 × 3.

The residual structure can be defined as

$$x_n = \begin{cases} \mathcal{P}(x_n), & n = 1 \\ x_{n+1} + \mathcal{F}_{n-1}(x_{n-1}), & n \geq 2 \end{cases} \quad (1)$$

where  $\mathcal{P}(x_n)$  is the result of  $x_n$  through a 1 × 1 conv. and 1 through a combination of 3 × 3 conv. and 1 × 1 conv. is  $\mathcal{F}(x_n)$ . The specific process of feature information through the residual structure is as follows.

(1) Input image feature information  $x$  after a 1 × 1 conv. feature information is expressed as  $\mathcal{P}(x_1)$ , recorded as  $x_1$ .

(2) The new feature information generated by summing the feature information combined with  $x_1$  through the first 3 × 3

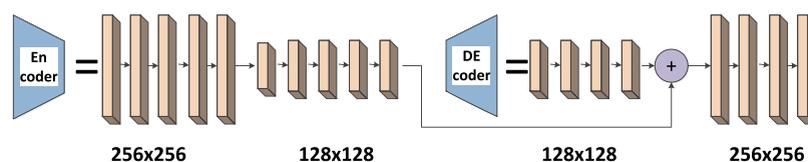


Figure 3. Encoder–decoder structure.

conv. and  $1 \times 1$  conv. is expressed as  $x_1 + \mathcal{F}(x_1)$ , denoted as  $x_2$ .

(3) The new feature information generated by summing the feature information combined with  $x_2$  through the first  $3 \times 3$  conv. and  $1 \times 1$  conv. is expressed as  $x_2 + \mathcal{F}(x_2)$ , denoted as  $x_3$ .

(4) The new feature information generated by summing the feature information combined with  $x_3$  through the first  $3 \times 3$  conv. and  $1 \times 1$  conv. is expressed as  $x_3 + \mathcal{F}(x_3)$ , denoted as  $x_4$ .

(5) The new feature information generated by summing the feature information combined with  $x_4$  through the first  $3 \times 3$  conv. and  $1 \times 1$  conv. is expressed as  $x_4 + \mathcal{F}(x_4)$ , denoted as  $x_5$ .

(6) The new feature information generated by summing the feature information combined with  $x_5$  through the first  $3 \times 3$  conv. and  $1 \times 1$  conv. is expressed as  $x_5 + \mathcal{F}(x_5)$ , denoted as  $x_6$ .

**3.2. Encoder–Decoder Structure.** Encoder–decoder<sup>29</sup> does not refer to a specific algorithm. In the process of image processing, the image feature extraction process is termed as the encoding process and the output process is termed as the decoding process. The encoder–decoder structure is shown in Figure 3. The encoder is stacked by five residual structures with a resolution of  $256 \times 256$  and five residual structures with a resolution of  $128 \times 128$ . The decoder structure comprises four residual structures with a resolution of  $128 \times 128$  and four residual structures with a resolution of  $256 \times 256$ .

The encoder–decoder arrangement creates a large receptive field that effectively utilizes feature information for image restoration. The SA-ConvLSTM network structure connected between the encoder and decoder<sup>41</sup> may connect all the states in the prediction network by stacking multiple ConvLSTM layers and using the self-attention memory module to aggregate past-related features for the final prediction. In addition, feature summation is used to establish a skip connection between the encoder and decoder to ensure image consistency.

**3.3. Loss Function.** Herein, the loss function uses a root mean square error (RMSE) to calculate the square root by determining the ratio of the square sum of the deviation between the observed value and the true value by the number of observations  $n$ . The smaller the value utilized to measure the deviation between the repaired and blurred images in the network, the greater the fitting result. The RMSE can be expressed as follows.

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{observed}_i - \text{predicted}_i)^2} \quad (2)$$

where  $\text{observed}_i$  represents the blurred image and  $\text{predicted}_i$  represents the repaired image. After the experiments, when

RMSE is used as the loss function, the network model can converge faster and achieve satisfactory repair results.

## 4. RESULTS AND DISCUSSION

**4.1. Foam Image Acquisition Process.** The design of the image-capturing process is shown in Figures 4 and 5. Industrial

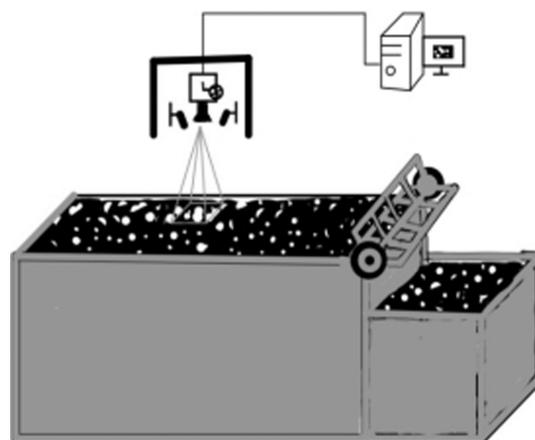


Figure 4. Acquisition device structure.

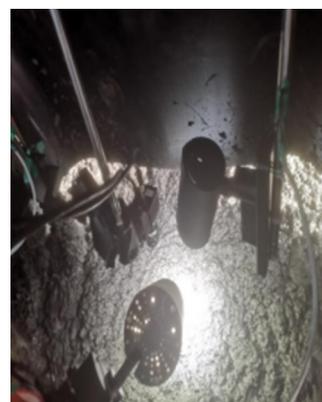


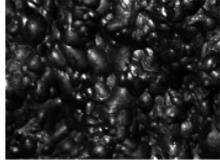
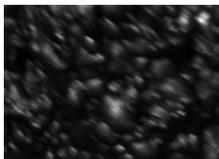
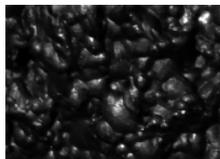
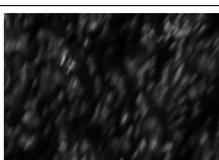
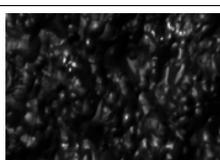
Figure 5. Field-image acquisition device.

cameras, equipment boxes, computers, and lighting equipment were used. Foam images of flotation tanks No. 1 and No. 2 at a coal washing company in the Inner Mongolia Autonomous Region were captured over the course of eight days. The daily duration of filming was 11 h. At the same time, every 20 min, the slime foam from two flotation tanks was collected. After the samples were dried, they were loaded to prepare for subsequent flotation-ash prediction based on foam characteristics.

**4.2. Experimental Results of the Foam Image Set.** We randomly selected 1000 foam images from over 20,000 foam images obtained onsite as datasets and then repaired them using our proposed network. For different degrees of blurred foam images, our proposed network has obtained good repair

results for varied levels of fuzzy bubble images, and the specific repair effect is shown in Table 1.

**Table 1. Our Proposed Network on Repairing Different Degrees of Blurred Foam Images**

	original images	images repaired by our proposed network
mildly blurred image		
moderately blurred image		
severely blurred images		

Based on the visual renderings of the repaired foam, we can find that the repaired foam images have sharp foam edges and clear textures, which is beneficial for our later segmentation and feature extraction of the foam.

The network proposed herein is compared with SRN,<sup>26</sup> DeepRFT,<sup>23</sup> MPRNet,<sup>40</sup> and the network suggested by Gao et al.,<sup>25</sup> which may be used to deblur the multiscale convolutional neural network for image restoration on foam image datasets. To meet the requirements of the image size of MPRNet,<sup>40</sup> the foam image size is reduced to  $1280 \times 968$  in equal proportions during testing.

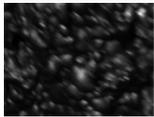
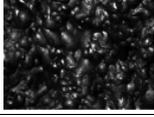
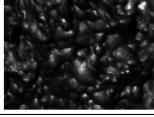
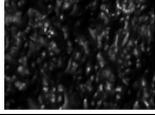
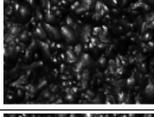
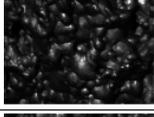
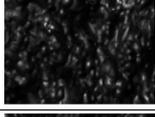
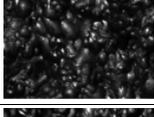
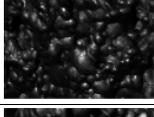
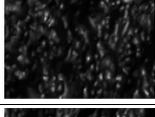
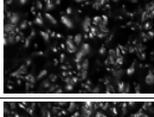
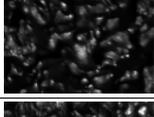
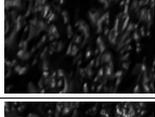
To evaluate the image quality of the restored image, the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and entropy are used. The greater the value of these three evaluation methods, the higher the quality of the image restoration. The visual effect of repairing different degrees of blurred foam images is shown in Table 2.

The evaluation results of repairing different degrees of blurred foam images are shown in Table 3.

The PSNR obtained by the network test herein is 0.216 points higher than the highest in other comparison networks, the SSIM is 0.010 points higher than the highest in other comparison networks, and the entropy is 0.015 points higher than the highest in other comparison networks, indicating that the network herein has achieved good results in the repair of foam images. The quantified results are consistent with our observation that our framework outperforms other frameworks in the remediation of slime flotation foam images.

**4.3. GOPRO Image Set Experimental Results.** The GOPRO<sup>24</sup> dataset contains 3214 pairs of blurred and clear images with an image resolution of  $1280 \times 720$ . Different intensity blurs in the dataset are generated by averaging a varied number of consecutive frames from a video captured at 240 fps using a GOPRO4 camera.

**Table 2. Different Networks in Repairing Different Degrees of Blurred Foam Images**

	mildly blurred images	mildly blurred images	severely blurred images
original images			
our proposed network			
SRN			
DeepRFT			
MPRNet			
network proposed by Gao et al.			

**Table 3. Evaluation Results of Different Networks on the Foam Image Set**

	our proposed network	SRN	DeepRFT	MPRNet	network proposed by Gao et al.
PSNR	33.657	33.422	33.371	32.908	33.441
SSIM	0.868	0.858	0.858	0.855	0.858
entropy	4.288	4.272	4.262	4.269	4.273

To ensure accuracy, all the networks herein are trained and tested in the same environment. The 2103 images in the GOPRO<sup>24</sup> dataset are used for model training, while the remaining images are used to test blurred images of varying degrees. The PSNR, SSIM, and entropy results are shown in Table 4.

**Table 4. Evaluation Results after GOPRO Dataset Repair**

	our proposed network	SRN	DeepRFT	MPRNet	network proposed by Gao et al.
PSNR	33.811	33.308	33.219	33.170	33.067
SSIM	0.861	0.829	0.815	0.813	0.813
entropy	5.192	5.195	5.193	5.189	5.198

The PSNR obtained by the network test herein is 0.503 points higher than the highest in other comparison networks, and the SSIM is 0.032 points higher than the highest in other comparison networks, although the entropy is lower than the highest reported by Gao et al., and the difference is only 0.003,

indicating that the network herein has achieved a good repair effect on the GOPRO<sup>24</sup> dataset images.

For the images in the GOPRO<sup>24</sup> dataset with varying degrees of blur, the visual effect is shown in Table 5.

**Table 5. Visual Results after GOPRO<sup>24</sup> Dataset Repair<sup>a</sup>**

original images			
our proposed network			
SRN			
DeepRFT			
MPRNet			
network proposed by Gao et al.			

<sup>a</sup>Photograph courtesy of “Nah, S.; Tae, H. K.; Kyoung, M. L.”. Copyright 2023.

**4.4. Experimental Results of the HIDE Image Set.** The HIDE<sup>42</sup> dataset includes 8422 pairs of clear and blurred images. Motion blur is developed by averaging frames of several real-world scenarios with varying numbers of persons captured using GoPro Hero cameras at 240 fps. The PSNR, SSIM, and entropy results obtained by evaluating 1200 randomly selected blurred images of varying degrees from the HIDE<sup>42</sup> dataset are shown in Table 6.

**Table 6. Evaluation Results after HIDE Dataset Repair**

	our proposed network	SRN	DeepRFT	MPRNet	network proposed by Gao et al.
PSNR	32.978	32.671	32.862	32.777	32.577
SSIM	0.832	0.804	0.794	0.788	0.789
entropy	5.195	5.196	5.185	5.174	5.198

From the results in Table 6, it can be seen that the PSNR obtained by the network proposed herein is 0.503 times higher than the highest in other comparison networks, whereas the SSIM is 0.032 times greater than the highest in other comparison networks, and the entropy is lower than the highest Gao et al., but the difference is only 0.003. This demonstrates that the network proposed herein can effectively repair images from the HIDE<sup>42</sup> dataset.

The visual effect for images in the HIDE<sup>42</sup> dataset with varying degrees of blur is shown in Table 7.

**4.5. Ablation Study.** Table 8 demonstrates that connecting SA-ConvLSTM<sup>39</sup> between the encoder and decoder may aggregate the previous relevant features of the image, obtain the connection in the spatial and temporal

**Table 7. Visual Results after HIDE Dataset Repair<sup>a</sup>**

original images			
our proposed network			
SRN			
DeepRFT			
MPRNet			
network proposed by Gao et al.			

<sup>a</sup>Photograph courtesy of “Shen, Z.; Wang, W.; Lu, X.; Shen, J.; Ling, H.; Xu, T.; Shao, L.”. Copyright 2023.

**Table 8. Influence of Networks on Image Restoration**

	froth images		GOPRO		HIDE	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
No SA-ConvLSTM	33.451	0.856	33.341	0.832	32.593	0.800
SA-ConvLSTM	33.657	0.868	33.811	0.861	32.978	0.832

dimensions, and enable the network to achieve superior outcomes in image restoration.

From the results in Table 8, it can be seen that for the froth image dataset, the PSNR for the SA-ConvLSTM network connected to the network in this paper is 0.206 higher and the SSIM is 0.012 higher than the PSNR for the unconnected SA-ConvLSTM network, for the GOPRO image dataset, the PSNR is 0.470 higher and the SSIM is 0.029 higher, for the GOPRO image dataset, the PSNR is 0.385 higher and the SSIM is 0.032 higher. This suggests that connecting the SA-ConvLSTM network can help the network achieve better results when performing image restoration.

The quantitative analysis of the restored coal slurry foam image dataset with different order residual structures is shown in Table 9.

From the results in Table 9, it can be seen that the fifth-order residual structure helps the network in this paper to achieve better restoration results in the image restoration process.

**Table 9. Quantitative Analysis of the Structure of Residuals of Different Orders**

	froth images		
	PSNR	SSIM	entropy
fifth-order residual structure	33.657	0.868	4.288
fourth-order residual structure	33.409	0.855	4.280
third-order residual structure	33.631	0.866	4.282
second-order residual structure	33.524	0.867	4.283

## 5. CONCLUSIONS

This study proposes a multiscale convolutional neural network to repair blurred images to solve the challenge of constructing image datasets owing to image blur, which hinders the feature extraction of foam images. This paper proposes a novel five-order residual structure that may be used to extend the network structure through simple stacking. It can not only extend the network structure through direct stacking but also enhance its ability for complex learning. This study demonstrates that connecting the SA-ConvLSTM network across encoder–decoder structures stacked using the residual structure enables the network to obtain good results in image restoration with varying degrees of blur. Numerous studies demonstrate that the proposed image restoration technique has a superior effect on the fuzzy restoration of foam images and has the potential to be implemented in the flotation field. Subsequently, utilizing high-quality foam image generated by the network repair described herein, we will construct a dataset, extract the feature information of the foam image, and analyze the correlation between the foam characteristics and the key process parameters of the onsite flotation to guide the intelligent adjustment of flotation site parameters.

## ■ AUTHOR INFORMATION

### Corresponding Author

Xianwu Huang – School of Information Engineering, Inner Mongolia University of Science & Technology, Baotou 014010, China; [orcid.org/0000-0002-8867-6747](https://orcid.org/0000-0002-8867-6747); Email: [huangxianwu1983@163.com](mailto:huangxianwu1983@163.com)

### Authors

Yuxiao Wang – School of Information Engineering, Inner Mongolia University of Science & Technology, Baotou 014010, China; [orcid.org/0000-0002-9313-1995](https://orcid.org/0000-0002-9313-1995)

Haili Shang – Key Laboratory of Coal Processing and Efficient Utilization, China University of Mining and Technology, Ministry of Education, Xuzhou 221116, China; [orcid.org/0000-0001-9952-0687](https://orcid.org/0000-0001-9952-0687)

Jinshan Zhang – School of Mines and Coal, Inner Mongolia University of Science & Technology, Baotou 014010, China

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsomega.2c08293>

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The study was financially supported by the Key Laboratory of Coal Processing and Efficient Utilization, (China University of Mining and Technology), Ministry of Education, supported by the Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region (NJZY22451), supported by the Fundamental Research Funds for Inner Mongolia University of Science and Technology.

## ■ ABBREVIATIONS

CCDC, Chinese control and decision conference; LSTM, long short-term memory; PSNR, peak signal-to-noise ratio; RMSE, root mean square error; SSIM, structural similarity

## ■ REFERENCES

- (1) Zarie, M.; Jahedsaravani, A.; Massinaei, M. Flotation Froth Image Classification Using Convolutional Neural Networks. *Miner. Eng.* **2020**, *155*, No. 106443.
- (2) Jian, H.; Lihui, C.; Yongfang, X. Design of Soft Sensor for Industrial Antimony Flotation Based on Deep CNN. In *Chinese control and decision conference (CCDC)*; IEEE Publications, 2020; pp 2492–2496.
- (3) Lu, M.; Liu, D.; Deng, Y.; Wu, L.; Xie, Y.; Chen, Z. RK algorithm: A Novel Dynamic Feature Matching Method of Flotation Froth. *Measurement* **2020**, *156*, No. 107581.
- (4) Pawlik, M. Fundamentals of Froth Flotation. *ChemTexts* **2022**, *8*, 1–40.
- (5) Wen, Z.; Zhou, C.; Pan, J.; Nie, T.; Jia, R.; Yang, F. Froth Image Feature Engineering-based Prediction Method for Concentrate Ash Content of Coal Flotation. *Miner. Eng.* **2021**, *170*, No. 107023.
- (6) Cao, W.; Wang, R.; Fan, M.; Fu, X.; Wang, H.; Wang, Y. Recognition of reagent dosage condition image for coal flotation system based on joint classification model of MRMR and SSGMM. *Appl. Intell.* **2022**, *52*, 732–752.
- (7) Ding, L.; Yang, J.; Zhang, Z.; Zhao, Y. A summary of extraction algorithms of froth features in slime flotation. *Coal Prep. Technol.* **2021**, *38*, 2045–2058.
- (8) Tan, J.; Liang, L.; Peng, Y.; Xie, G. Challenges of Using Froth Features to Predict Clean Coal Ash Content in Coal Flotation. *Int. J. Coal Prep. Util.* **2020**, *42*, 1–37.
- (9) Ozmak, M.; Zeki, A. Coal Froth Flotation: Effects of Reagent Adsorption on the Froth Structure. *Energy Fuels* **2006**, *20*, 1123–1130.
- (10) Zhao, L.; Peng, T.; Xie, Y.; Gui, W.; Zhao, Y. Froth Stereo Visual Feature Extraction for the Industrial Flotation Process. *Ind. Eng. Chem. Res.* **2019**, *58*, 14510–14519.
- (11) Wang, Z.; Cheng, H.; Xu, Y.; Cheng, K.; Gao, H.; Li, H. Development of intelligent controls system for slime flotation. *Coal. Process. Compr. Util.* **2022**, *8*, 6–11.
- (12) Guo, Z.; Wang, R.; Fu, X.; Wei, K.; Wang, Y. Method for extracting froth velocity of coal slime flotation based on image feature matching. *Ind. Mine Autom.* **2022**, *48*, 34–39.
- (13) Sun, Y.; Chen, C.; Yang, Z.; Wang, X. Image Recognition Method of Flotation Foam in Coal Preparation Plant Based on CNN-SVM. *Coal Process. Compr. Util.* **2021**, *2*, 8–11.
- (14) Wang, X.; Zhang, C.; Zhang, S. Multiscale Convolutional Neural Networks with Attention for Plant Species Recognition. *Comput. Intell. Neurosci.* **2021**, *2021*, No. 5529905.
- (15) Li, X.; He, M.; Li, H.; Shen, H. A Combined Loss-Based Multiscale Fully Convolutional Network for High-Resolution Remote Sensing Image Change Detection. *IEEE Geosci. Remote Sens. Lett.* **2022**, *19*, 8017505.
- (16) Yao, Q.; Hu, X.; Lei, H. Multiscale Convolutional Neural Networks for Geospatial Object Detection in VHR Satellite Images. *IEEE Geosci. Remote Sens. Lett.* **2021**, 18–27.
- (17) Li, G.; Yu, Y. Visual saliency detection based on multiscale deep CNN features. *IEEE Trans. Image Process.* **2016**, *25*, 5012–5024.
- (18) Gao, W.; Wu, S. Scratch Repairing of Old Photos Based on Multi-Scale Attention Semi-Supervised Learning. *Comput. Eng.* **2022**, *48*, 245–251.
- (19) Chen, G.; Liao, Y.; Yang, Z.; Liu, W. Image inpainting algorithm of multi-scale generative adversarial network based on multi feature fusion. *J. Comput. Appl.* **2022**, 1–10.
- (20) Zhang, H.; Xu, D.; Luo, H.; Yang, B. Multi-scale mural restoration method based on edge reconstruction. *J. Graph.* **2021**, *42*, 590–598.
- (21) Lin, X.; Zhou, Y.; Li, D.; Huang, W.; Sheng, B. Image Inpainting Using Multi-Scale Feature Joint Attention Model. *J. Comput. Aided Des. Comput. Graph.* **2022**, *34*, 1260–1271.
- (22) Bai, Z.; Yi, T.; Zhou, M.; Wei, W. Face Image Inpainting Method Based on Multi-Scale Feature Fusion. *Comput. Eng.* **2021**, *47*, 213–220.

- (23) Mao, X.; Liu, Y.; Shen, W.; Li, Q.; Wang, Y. Deep Residual Fourier Transformation for Single Image Deblurring. 2021, arXiv preprint arXiv: 2111.11745.
- (24) Nah, S.; Tae, H. K.; Kyoung, M. L. Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2017, pp 3883–3891.
- (25) Gao, H.; Tao, X.; Shen, X.; Jia, J. Dynamic Scene Deblurring with Parameter Selective Sharing and Nested Skip Connections. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2019; pp 3848–3856.
- (26) Tao, X.; Gao, H.; Shen, X.; Wang, J.; Jia, J. Scale-recurrent Network for Deep Image Deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2018; pp 8174–8182.
- (27) Su, J.; Xu, B.; Yin, H. A survey of deep learning approaches to image restoration. *Neurocomputing* **2022**, *487*, 46–65.
- (28) Xiao, W.; Tang, Z.; Luo, J.; Liu, J. FS-Deblur GAN: a spatiotemporal deblurring method for zinc froth flotation. *Eur. Phys. J.: Spec. Top.* **2022**, *231*, 1983–1993.
- (29) Liu, J.; Gao, Q.; Tang, Z.; Xie, Y.; Gui, W.; Ma, T.; Niyoyita, J. P. Online monitoring of flotation froth bubble-size distributions via multiscale deblurring and multistage jumping feature-fused full convolutional networks. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 9618–9633.
- (30) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016; pp 770–778.
- (31) Simonyan, K.; Andrew, Z. Very Deep Convolutional Networks for Large-scale Image Recognition. *Comput. Sci.* 2014, ar Xiv preprint ar Xiv: 1409.1556.
- (32) Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. A. Inception-v4, Inception-resnet and the Impact of Residual Connections on Learning. In *Thirty-first AAAI conference on artificial intelligence*; 2017.
- (33) Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2015; pp 1–9.
- (34) Ioffe, S.; Christian, S. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In *International Conference on Machine Learning*. PMLR; 2015; pp 448–456.
- (35) Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016; pp 2818–2826.
- (36) Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated Residual Transformations for Deep Neural Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2017, pp 1492–1500.
- (37) Shi, X.; Chen, Z.; Wang, H.; Yeung, D. Y.; Wong, W. K.; Woo, W. C. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. In *Adv. Neural Inf. Process. Syst.*; 2015; p 28.
- (38) Moniz, J. R. A.; David, K. Nested lstms. In *Asian Conference on Machine Learning*. PMLR; 2017, pp 530–544.
- (39) Lin, Z.; Li, M.; Zheng, Z.; Cheng, Y.; Yuan, C. Self-attention ConvLstm for Spatiotemporal Prediction. *AAAI Conf. Artif. Intell.* **2020**, *34*, 11531–11538.
- (40) Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; Yang, M. H.; Shao, L. Multi-stage Progressive Image Restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*; 2021; pp 14816–14826, DOI: 10.1109/CVPR46437.2021.01458.
- (41) Mao, X.; Chunhua, S.; Yu-Bin, Y. Image Restoration Using Very Deep Convolutional Encoder-decoder Networks with Symmetric Skip Connections. In *Advances in Neural Information Processing Systems*; 2016; p 29.
- (42) Shen, Z.; Wang, W.; Lu, X.; Shen, J.; Ling, H.; Xu, T.; Shao, L. Human-aware Motion Deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*; 2019; pp 5572–5581.