IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 September 2022, accepted 18 October 2022, date of publication 21 October 2022, date of current version 26 October 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3216388

RESEARCH ARTICLE

SEM Image Quality Assessment Based on Intuitive Morphology and Deep Semantic Features

HAORAN WANG[1], SHIYIN LI[1], JICUN DING [2], SUYAN LI[2], LIANG DONG[3], AND ZHAOLIN LU[1,2]

[1]School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China
[2]Xuzhou First People's Hospital, Xuzhou 221116, China
[3]School of Chemical Engineering and Technology, China University of Mining and Technology, Xuzhou 221116, China

Corresponding author: Zhaolin Lu (zhaolinlu@cumt.edu.cn)

This work was supported in part by the National Natural Science Foundations of China under Grant 62101557, in part by the Xuzhou City Social Development Key Special Project under Grant KC21153, and in part by the Xuzhou Science and Technology Project under Grant KC21262.

ABSTRACT The widespread use of scanning electron microscopy (SEM) has increased the requirements for SEM image quality. SEM images obtained by electron beam feedback have more complex texture features than natural images obtained by optical imaging, and this condition results in poor performance of algorithms used for assessing natural image quality on SEM datasets,meanwhile,the field of SEM image quality assessment(IQA) is mostly aimed at specific distortion types. In order to solve the above two problems,to address the rich texture, few edges, and extreme sensitivity to the distortion degree of SEM images, we propose a texture and semantic IQA (TSIQA) method for SEM images based on sparse mask and information entropy increase. First, we construct a neural network containing sparse mask module (SMM), which is used to extract intuitive texture features in the spatial and channel domains. Simultaneously, information growth attention (IGA) is introduced into SMM to detect the difference between current and past features of the network for extracting deep semantic information. The quality assessment experiments on SEM image datasets show that compared with the state-of-the-art IQA methods, including popular no-reference techniques adapted to the SEM-IQA, the TSIQA has superiority in typical criteria.

INDEX TERMS Image quality assessment, no-reference, SEM image, texture representation, semantic information.

I. INTRODUCTION

Recently, scanning electron microscopy (SEM) has presented more realistic and three-dimensional images because of its high magnification and depth of field, which play an important role in studying the microscopic world [1]. At present, the quality assessment of SEM images largely depends on the subjective judgment of operators. Individual differences in subjective assessment and visual fatigue caused by prolonged shooting will lead to deviations in the judgment of image quality of operators. The obtained microscopic images may present distortions to a certain extent, such as defocus blur, noise blur, contrast imbalance, and brightness imbalance [2], [3]. These distortions affect the researchers' judgment of the real morphology of substances. Therefore, the effective method of SEM image quality assessment (SEM-IQA) can help obtain high-quality images, which provides guidance and reference for the operator and real and reliable information for the study of substance microstructure.

According to the amount of reference information required in the assessment process, IQA algorithms can be divided into three categories: full-reference IQA, reduced-reference IQA, and no-reference IQA (NR-IQA). However, when assessing SEM image quality, obtaining a distortion-free reference image is usually difficult and the distortion type of image cannot be determined in advance. Thus, the no-reference method is more suitable for SEM-IQA task. The effect of NR-IQA based on traditional machine learning relies on the performance of the manually designed feature extraction module, and the IQA research with convolutional neural networks

The associate editor coordinating the review of this manuscript and approving it for publication was Andrea F. Abate.

(CNNs) as the backbone network has been developed rapidly owing to the strong representation ability of CNN [4], [5]. SEM images obtained by electron beam feedback have more complex texture features than natural images obtained by optical imaging. The two types of images show great differences in the appearance of substances, which results in poor performance of the algorithms used for natural IQA on SEM datasets. As far as we know, few studies are available on the SEM image quality evaluation at present, and they only focus on specific distortion types [6], [7]. These methods have difficulty applying SEM images with unknown or mixed distortion types or need to rely on pre-tasks [8]. Therefore, the research on NR-IQA methods for SEM image distortion feature design is urgent topic to be studied deeply.

According to whether a manual design of the feature extraction method is needed, the existing natural image quality assessment methods can be divided into traditional machine learning method and deep learning method. IQA algorithm models based on traditional machine learning mainly rely on manually designed feature extraction methods, and it can be divided into supervised and unsupervised learning methods according to whether subjective image scoring is required during model training. Supervised NR-IQA is dedicated to extracting valid visual features and mapping such visual features to image quality scores using machine learning models such as support vector machine. Natural scene statistics (NSS) is a widely used method to characterize image visual features, which holds that the undistorted image visual features obey a certain distribution and uses the change in distribution as a measure of image distortion. On this basis, Wang et al. [9] used NSS to extract features for IQA regression. Lu et al. [10] introduced a contourlet transform on the basis of NSS of the image, and then, they used the characteristics of each sub-band as an indicator to assess the image quality. Moorthy et al. [11] proposed a two-stage quality evaluation framework (BIQI) for images with unknown distortion types. Moorthy et al. [12] used distortion specificity to assess image quality by identifying image distortion types in advance. Inspired by the feature extraction from different image domains, Saad et al. [13] proposed an IQA model for feature extraction from the DCT domain. Mittal et al. [14] extracted NSS features in the spatial domain (BRISQUE). Ye et al. [15] used dictionary learning to extract features from small blocks of original images. Zhang et al. [16] extracted image features (DESIQUE) in spatial and frequency domains. The abovementioned methods largely rely on the performance of the manually designed feature extraction method and perform well on small datasets, but their performance is much weaker than that of deep learning methods when applied to large datasets.

The unsupervised NR-IQA methods based on traditional machine learning aim to construct effective image quality reference models. Mittal et al. [17] proposed a completely blind image quality assessment method (NIQE). In this method, statistical data are obtained from undistorted images and a multivariate gaussian (MVG) model is obtained by fitting. Then, the image quality is assessed by the parameter difference in MVG models between distorted and clean images. Xue et al. [18] segmented the distorted image into overlapping small blocks, assessed the local quality of each small block using a percentile pooling strategy and conducted quality perception clustering. This algorithm uses the statistical law of natural undistorted images to evaluate the quality, and it can complete the training without MOS tags. However, the performance of an unsupervised algorithm is usually weaker than that of the supervised algorithm when the amount of data is sufficient.

Recently, deep learning has developed rapidly and has shown great promise in image processing [19]. A large number of IQA methods with CNN as the backbone network have emerged. Kang et al. [4] first proposed the IQA model based on CNN and then added a multi-task CNN model [20] to assess image quality and identify distortion types. Bosse et al. [21] used the VGG16 network as a regression model for the IQA task to synchronously predict the weight and quality score values of image blocks and obtain the image quality score through weighting. Zeng et al. [22] learned the probabilistic characteristics of distorted images by pre-trained ResNet [23]. Ding et al. [24] develop a full-reference image quality model with explicit tolerance to texture resampling in order to avoid image quality assessment methods are overly sensitive to resampling of texture regions. Prashnani et al. [25] proposed a image quality dataset labeled with the probability that humans will prefer one image over another,then train a deep-learning model using a pairwise-learning framework to predict the preference of one distorted image over the other.The above two methods has achieved good results in natural image dataset, but these methods is full-reference model, not apply to usually lack of high quality reference image of SEM images. Freitas et al. [26] proposed uses the statistics of the orthogonal color planes pattern (OCPP) descriptor to characterize image quality,however, the method of defining image quality by color cannot effectively assess grayscale images, such as SEM images.

With the extensive application of attention mechanisms in various image processing tasks, many deep IQA algorithms combined with attention mechanisms have been proposed. Jia et al. [27] used the self-similar attention algorithm to obtain the saliency map of the image, and then, they optimized the original predicted value of mass fraction by block weighting. Chen et al. [28] proposed using cyclic attention mechanism [29] for assisting IQA tasks and training the network for automatically learning valuable image blocks through reinforcement learning [30]. They also used it to estimate the quality score of the entire image according to the quality scores of the image blocks. The multi-task based IQA algorithm improves the accuracy of evaluation by using auxiliary tasks. Ma et al. [31] introduced a model of multi-task CNN-IQA, which assists IQA task with a distortion recognition subtask. Kim et al. [32], [33] designed a pre-train CNN model through reference images and then fine-tune the model in the training stage. Wu et al. [34] proposed a CNN-IQA method of cascade

type to extract features and hierarchical degradation splicing. Freitas et al. [35] used local binary pattern (LBP) operators to extract texture features from images, and used computational visual attention model to output saliency maps.

The aforementioned IQA methods have achieved good results in assessing the distortion degree of natural images. However, the abovementioned methods do not perform well in SEM datasets due to that these images have few edges and rich textures, which are quite different from natural image features. Recently, SEM has been widely used in scientific research and industrial production. High-quality SEM images are the key to revealing the microstructure of materials. Some methods to assess the quality of SEM images have been proposed. According to gestalt perception psychology and entropy masking characteristics, Wang et al. [6] believed that the human visual system is more sensitive to the distortion of the cartoon component than the fuzzy distortion of the redundant texture component. They also evaluated the fuzzy degree of SEM images. Li et al. [7] proposed an SEM image sharpness quality assessment method based on a dark channel prior. First, the dark channel image of the SEM image is generated to extract its edge, and then, the noise is eliminated by the filter that preserves the edge. Lu [8] used texture inpainting as the upfront task and transferred the weighting factors obtained from texture inpainting task into the IQA task for assessing SEM image quality. This method makes up for the shortcoming of the abovementioned methods that can only target specific distortion types. However, its performance depends on the performance of the texture inpainting the network. Thus, its robustness is poor. The dataset of the pre-task also needs to be prepared before executing the IQA task. In summary, some current IQA algorithms for SEM images only target specific distortion types and do not have a good generalization, and some rely on the pre-task and do not have good robustness.

The SEM images are mainly composed of repeated textures because of the nonlocal self-similarity of the microstructure of the materials. Figure 1 is a typical SEM image of mineral particles and its edge and texture decomposition graph. The edge graph is extracted by the edge detection algorithm [36], and the texture graph is extracted by the texture detection algorithm [6]. The edge information of an SEM image is relatively simple, and texture information is the main component. Figure 2 shows the influences of the four distortion types on SEM images of ants'eyes and corresponding mean opinion score (MOS) values obtained through subjective assessment. The image quality is higher when the MOS value is higher. Figure 2 shows that the compound eye texture area in the last four distorted images has different degrees of detail loss compared with the high-quality (a) image, which indicates that the image texture features are very sensitive to distortion. However, the human eye will not pay much attention to the change in the target contour edge in the image regardless of the severity of its distortion. Human's attention to observing microscopic images usually focuses on the local details because the edge information of SEM images is relatively

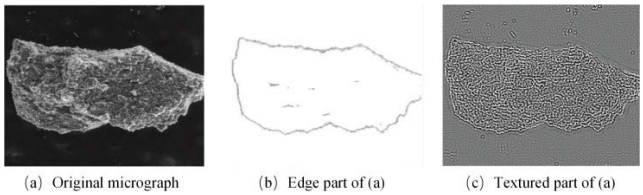

FIGURE 1. SEM image and its edge and textural parts.

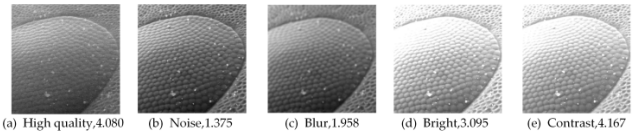

FIGURE 2. High quality SEM images and its four different distortion types, and their corresponding mean opinion score (MOS).

simple. From the above analysis, we believe that the texture features of an SEM image are the key factors that characterize its image quality, and the key to obtaining an excellent SEM-IQA method is to extract its texture features effectively. Simultaneously, IQA is also the human visual perception, which can also be classified into semantic analysis, that is, images can be described by visual words. These visual words can also be mapped to the potential semantic space to assess the image quality through the distribution of potential semantics.

Therefore, we propose a SEM-IQA network model (TSIQA) based on intuitive texture morphology and deep semantic features. The core idea is to use two sparse masks to extract texture information in the spatial and channel domains in a fine-grained way. Then, information growth attention (IGA) is used to extract semantic information. Finally, the SEM image quality is assessed. Notably, our model is designed for SEM images rich in texture and sensitive to distortion. This algorithm not only is targeted at a specific distortion type but is also a general algorithm.

Our contributions are summarized as follows: 1) A SEM-IQA method based on intuitive morphology and deep semantic features is proposed. 2) Sparse masks are used to achieve fine-grained positioning of texture information in spatial and channel domains, and this way improves performance and reduces computational cost. 3) For the first time, the increase in information entropy is considered in the IQA task and IGA is introduced to improve the deep semantics capture capability of the network.

II. PROPOSED METHOD

The proposed TSIQA is designed for SEM images with rich textures, few edges, and textures that are extremely sensitive to the degree of distortion. It uses sparse masks to screen out the intuitive morphological features of SEM images, that is, texture information. It also contains an IGA that can extract the deep semantic features of images. As shown in Figure 3, our algorithm is divided into two training sessions. Through

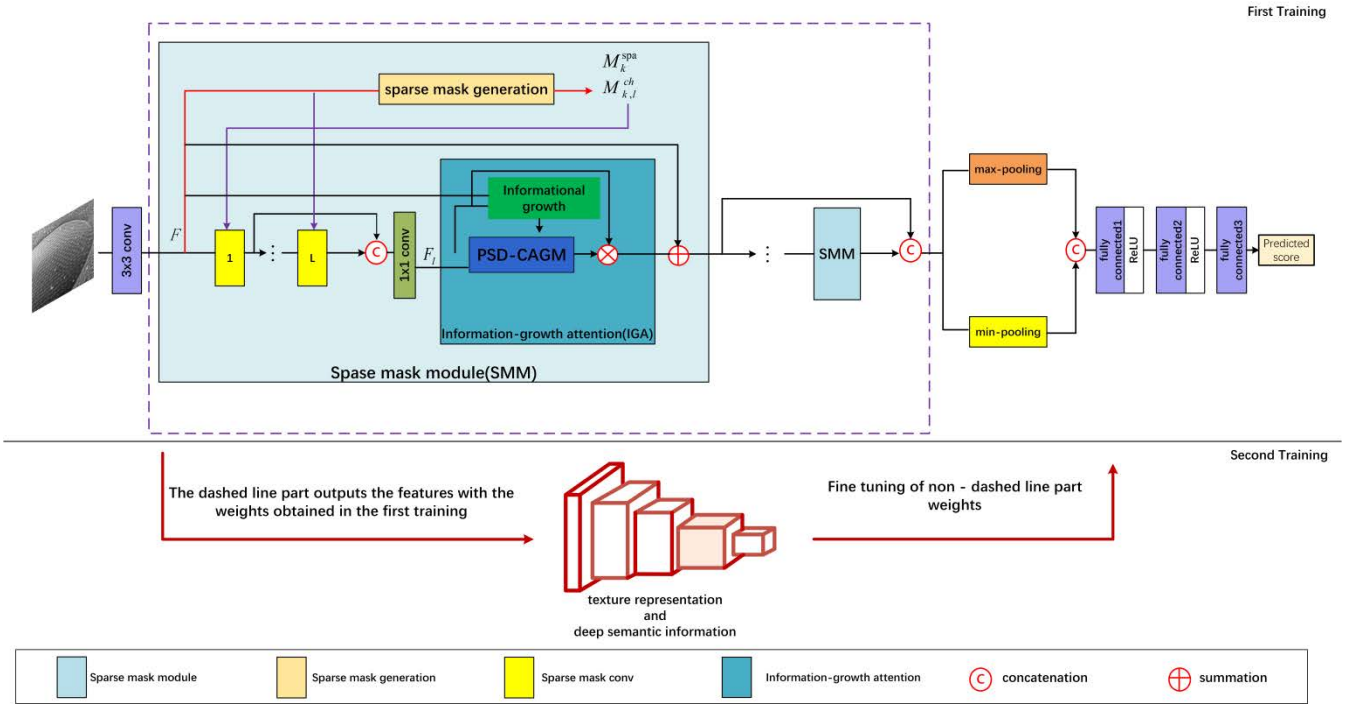

FIGURE 3. The framework of our TSIQA network.

the first training, we obtain a spatial mask that can recognize texture feature regions and a channel mask that can mark non-texture region channels. In the second training, the two sparse masks are used to sparsely convolve the feature map of the input sparse mask module (SMM) for obtaining the pure texture representation. Then, the deep semantic information of features is extracted through IGA, and the optimal weight is obtained by fine-tuning the network. Our TSIQA achieves better performance and computational efficiency given that we only extract the necessary information.

A. SPARSE MASK

1) SPATIAL MASK

For generating sparse masks, we refer to the work of Wang et al. [37]. The goal of the spatial mask is to recognize the texture region in the feature map. Its generation during the first training is shown in Figure 4(a). First, feature $F^{spa} \in R^{2\times H\times W}$ is derived from feature map $F \in R^{C\times H\times W}$ by several convolution. Then, to enable the binary spatial mask to update the gradient during training, we perform the Gumbel softmax technique to approximate a one-hot distribution [38] and obtain the softened spatial mask $M_k^{spa} \in R^{H\times W}$:

$$M_k^{spa}[x,y] = \frac{\exp((F^{spa}[1,x,y] + G_k^{spa}[1,x,y])/\tau)}{\sum_{i=1}^{2}\exp((F^{spa}[i,x,y] + G_k^{spa}[i,x,y])/\tau)} \tag{1}$$

where x, y are indices of vertical and horizontal, $G_k^{spa} \in R^{2\times H\times W}$ is a Gumbel noise tensor mapped from the Gumbel (0,1) distribution, and τ is a temperature parameter. When τ approaches zero, samples generated by Gumbel softmax distribution approach to one-hot. In other words, the Gumbel softmax distribution tends to generate binary spatial masks M_k^{spa} as τ gradually decreases.In the experiments, we anneal from high temperatures to low temperatures to obtain binary spatial masks. Specifically, we use the $\tau = \max(0.4, 1 - \frac{t}{T_{temp}})$ scheme to cool down τ, where t is the current epoch and T_{temp} is set to 500 based on experience. During the second training, we use an argmax layer to replace the Gumbel softmax layer to obtain the binary spatial mask, as shown in Figure 4(b).

2) CHANNEL MASK

The function of the channel mask is to mark redundant channels in flat areas. Its generation during the first training is shown in Figure 4(c). As before, we use the Gumbel softmax technique to obtain the binary channel mask. For the l^{th} convolutional layer in the k^{th} SMM, the random value is first extracted from the gaussian distribution N(0,1) for start to obtain the $S_{k,l}$. Then, we enter $S_{k,l} \in R^{2\times C}$ into the Gumbel softmax layer to obtain softened channel masks $M_{k,l}^{ch} \in R^{C}$:

$$M_{k,l}^{ch}[c] = \frac{\exp((S_{k,l}[1,c] + G_{k,l}^{ch}[1,c])/\tau)}{\sum_{i=1}^{2}\exp((S_{k,l}[i,c] + G_{k,l}^{ch}[i,c])/\tau)} \tag{2}$$

where $G_{k,l}^{ch} \in R^{2\times C}$ is a Gumbel noise tensor, c is the index of channel. In the first training, the annealing strategy we applied is consistent with the abovementioned spatial mask. The generation of channel mask in the second training is

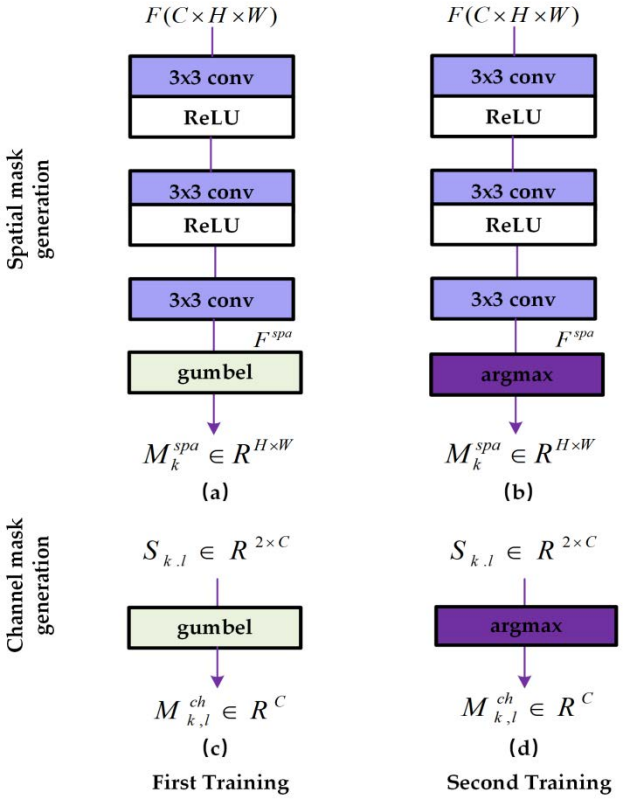

FIGURE 4. Sparse mask generation.

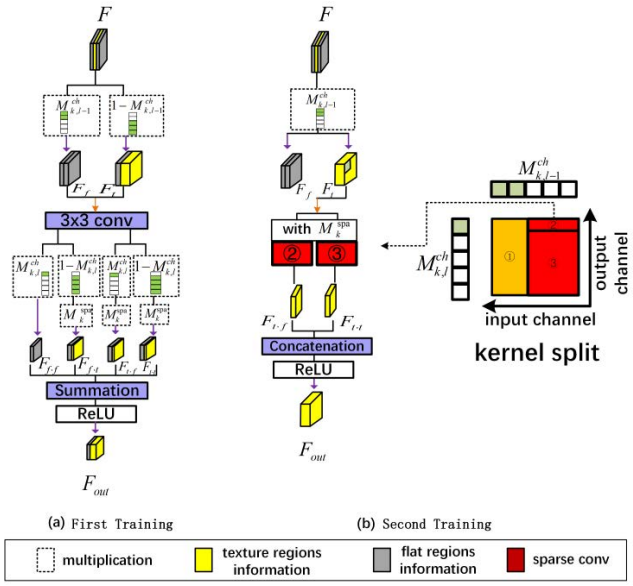

FIGURE 5. Sparse convolution.

shown in Figure 4(d). We also use the argmax layer to replace the Gumbel softmax layer to obtain the binary channel mask.

B. SPARSE MASK CONVOLUTION

1) FIRST TRAINING

During the first training, we simulate the sparse convolution by multiply the results of the ordinary convolution with the predicted spatial and channel mask to achieve the reverse update of gradients.As shown in Figure 5(a), input feature map F is first multiplied by $M^{ch}_{k,l-1}$ and $1-M^{ch}_{k,l-1}$ to produce F_f and F_t, respectively. It means the channel separating the ''flat'' and ''textured'' feature maps in feature F. Next, F_f and F_t perform convolution with the same weight to activate different portions of the feature map. Then, the obtained features are multiplied by the combination of $1-M^{ch}_{k,l}$, $M^{ch}_{k,l}$ and M^{spa}_{k}. Finally, the above mentioned results are summed to obtain the F_{out}.Given that the Gumbel softmax technique is used in sparse mask generation, the gradient of all positions can be reserved to optimize the convolution kernel weight.

2) SECOND TRAINING

In the second training stage, sparse convolution is performed according to the spatial and channel masks obtained in the first training, as shown in Figure 5(b). For example, the input feature F in layer L of the k^{th} SMM is the first decomposed into F_f and F_t based on $M^{ch}_{k,l-1}$. At this time, F_t carries the channel containing texture information and then performs sparse convolution on F_t. Notably, the sparse convolution at this time divides the convolution kernel into two sub-kernels according to $M^{ch}_{k,l-1}$ and $M^{ch}_{k,l}$, and they are used to perform convolution with F_t and cooperate with the spatial mask M^{spa}_{k}, that can locate the texture area in the space domain to obtain $F_{t\cdot f}$ and $F_{t\cdot t}$. Finally, the two obtained features are concatenated to obtain the output features F_{out}. In this way, we can extract pure texture information in channel and space domains with fine granularity.

C. INFORMATION GROWTH ATTENTION

Following Li et al. [39], an image in the IQA task is mapped to a quality fraction through the full connection layer after convolutional layers, which shows that the network learns deeper semantic information in this process. That is, the feature obtained after convolution in Figure 3 has higher information entropy than the previous one. However, the lower information entropy represents the lower uncertainty. Therefore, the generalization and stability of the network for IQA tasks are better when the information entropy is lower. Thus, we consider removing the mutual information between them to reduce the information entropy. As shown in Figure 7, IGA uses the increased information to form attention. Thus, we takein Figure 3 as the current features, regardas the former features, and subtract them. In this way, the amount of information can be increased in the convolution process and the obtained result can be input into progressive size decreasing CAGM (PSD-CAGM) to generate attention. As shown in Figure 6, the traditional self-attention adopts CAGM, which directly performs global average pooling and then forms attention through a full connection. However, IGA adopts PSD-CAGM, as shown in Figure 7, which generates attention weight through several convolution and pooling steps. This gradual operation will fully use the information

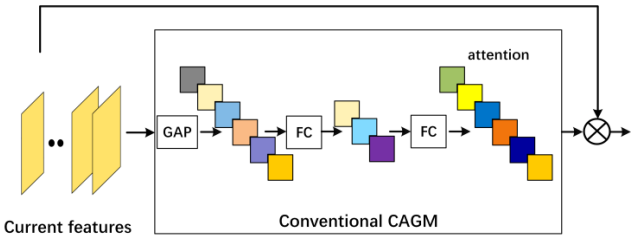

FIGURE 6. Conventional self-attention.

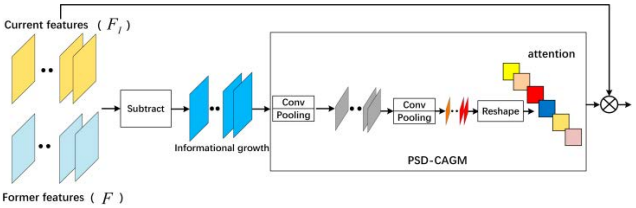

FIGURE 7. Information growth attention(IGA).

of the feature map and avoid the loss of some features and details.

D. LOSS FUNCTION

On the basis of channel and spatial masks, we introduce sparsity term $\eta_{k,l}$:

$$\eta_{k,l} = \frac{1}{C \times H \times W} \sum_{c,x,y} ((1 - M_{k,l}^{ch}[c] \times M_k^{spa}[x, y] + M_{k,l}^{ch}[c] \times I[x, y])) \quad (3)$$

Where $I \in R^{H \times W}$ is a all ones tensor, and $\eta_{k,l}$ indicates the proportion of active positions in the output feature maps. To make the output features more sparse when considering fewer active positions, we further introduce sparse regularization loss:

$$L_{reg} = \frac{1}{K \times L} \sum_{k,l} \eta_{k,l} \quad (4)$$

where K is the number of SMMs, and L is the number of sparse mask convolution layers in each SMM. Meanwhile, we express the predicted quality score and ground truth quality score of the i^{th} picture as $\hat{y}_i$ and y_i, respectively, and L1 loss is used as the score loss function:

$$L_{score} = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_i - y_i \right| \quad (5)$$

The total loss of the first training is defined as:

$$L = L_{score} + \lambda L_{reg} \quad (6)$$

We adopt the heating strategy $\lambda = \lambda_0 \times \min(\frac{t}{T_{warm}}, 1)$ to ensure the stability of the experiment, where t represents the current epochs, T_{warm} is set to 50 based on experience, and λ_0 is set to 0.1. The influence of different λ_0 values on the experiment will be discussed in detail in Section III.C.3).

The second training is a pure image quality evaluation task. Thus, we directly use the abovementioned L_{score} as the loss function.

E. OVERVIEW OF ALGORITHMS

We divide the training of the model into two steps. As shown in Algorithm 1, the input SEM image is initially trained for the first time. After a 3×3 convolution, a feature map F is generated and sent to the SMM. Each SMM contains a sparse mask generation module and L times of sparse convolution. We input F into the sparse mask generation module to generate the spatial and channel masks. Then, L sparse convolution of F is performed to obtain L feature maps and overlay them. Next, the initial F is added through 1×1 convolution and IGA. After the iteration of SMMs with a quantity of K, the feature maps are superimposed. Finally, the maximum and minimum pooling steps are conducted, and the predicted quality scores are obtained through three fully connected layers. During the training process, the weight of the sparse mask is updated with the Gumbel softmax technique until the best weight of the sparse mask is obtained.

In the second training, we can freeze the network part in front of the pool layer. Then, we can use the sparse mask obtained in the first training and the sparse convolution algorithm mentioned in Section III.B.2) to output pure texture features after K SMMs. Thereafter, training with this feature is conducted to fine-tune the network part behind the SMM. Finally, the complete network weight is obtained. At this point, our model training is completed, and the corresponding prediction scores can be obtained by inputting SEM images.

III. RESULTS

A. SEM IMAGE DATASET AND EXPERIMENTAL PROTOCOL

The SEM image dataset [40] used in this study contained 50 material samples from the fields of mineralogy, biology, and chemistry, and each sample contained 13 images of different quality and distortion categories. As shown in Table 1,one high-quality image,three defocus blur images,three noise distortion images, two astigmatism distortion images, two brightness distortion images, and two contrast distortion images were selected from each sample, these 650 images had a resolution of 1024×884 and contained the corresponding MOS value tags. An example of a class of images is shown in Figure 8. According to whether the image contains a specific background region, the image is divided into two categories,use the method of category sampling, 70% of the data are divided into the training set and 30% into the test set.

During the training process, we randomly cut each image into 16 patches of 128×128 pixels, which is the best setting found in the experiment (discussed in Section III.C.3)). A total of 7280 patches were obtained for training and 3120 patches were used for testing. This setting was found to be the best in the experiment. For TSIQA, C = 50, L = 4, and K = 5. During the training process, the Adam optimizer

Algorithm 1 TSIQA

Input: SEM images x and label score y_i
Output: Predicted quality score $\hat{y}_i$ for x
/ ※ **First Training** ※ /
$F \leftarrow 3 \times 3$ conv
for i = 1,2, ... K
    The i-th SMM:
    $(M_k^{\text{spa}}, M_{k,l}^{ch}) \leftarrow$ Sparse mask generation (use F)
    for i = 1,2, ... L
        $f_i \leftarrow$ The i-th Sparse conv
    end for
    $F_1 \leftarrow$ Concatenation, 1×1 Conv, Information-growth attention
    $F_i \leftarrow$ Summation (with F)
end for
Concatenation, Pooling, FC, ReLU
return $\hat{y}_i$
Update $(M_k^{\text{spa}}, M_{k,l}^{ch})$ with Gumble softmax
/ ※ **Second Training** ※ /
Freeze the portion of the network before the pooling layer
At this point F_i is purely textural representation
Concatenation, Pooling, FC, ReLU
return $\hat{y}_i$
Update $\theta_i = \text{Adam}(L,\theta)$ on the NR-IQA task
{L is loss function}

TABLE 1. SEM image quality dataset.

Dataset	Category	Size	Rank
SEM image and corresponding MOS	Noise	3	Weak,medium,strong
	Blur	3	Weak,medium,strong
	Contrast	2	Weak,strong
	Brightness	2	Weak,strong
	Astigmatism	2	X-axis,Y-axis
	High quality	1	\

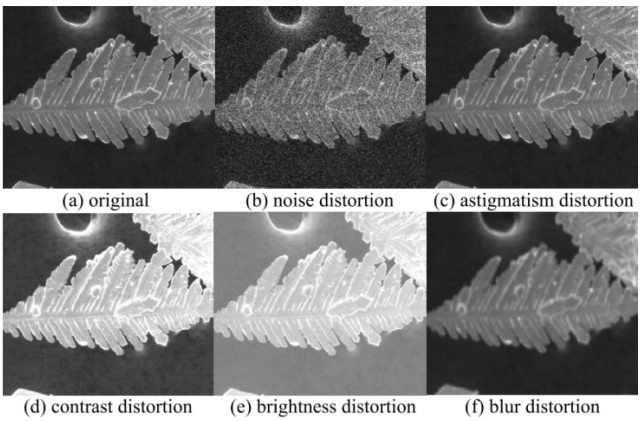

FIGURE 8. A set of sample images.

[41] with an initial learning rate of 1×10^{-3} was used, and the batch size was set to 16. The learning rate was adaptively reduced to one-tenth when no new minimum value was obtained for 10 consecutive epochs of training loss. Training was stopped when the learning rate decreased for the third time.

B. EVALUATION CRITERIA

The proposed method was evaluated using three widely used metrics to compare the IQA approaches. Spearman rank-order correlation coefficient (SRCC), Pearson linear correlation coefficient (PLCC), and root mean square error (RMSE). Before calculating the PLCC and RMSE [42], the predicted scores were input into the nonlinear logic function:

$$f(x) = l_1\left(\frac{1}{2} - \frac{1}{1 + e^{l_2(x - l_3)}}\right) + l_4 x + l_5 \tag{7}$$

where x is the predicted score of the objective assessment method,represents the mapped objective score, l_1 and l_2, l_3, l_4, l_5 are the fitting parameters.The aim of this function is to unify the objective scores calculated by different quality metrics into the same range [43].

C. MODEL ANALYSIS

First, we compared our algorithm with existing dominant algorithms. Ablation experiments were then performed to verify the effectiveness of sparse convolution, IGA, and the second training. Finally, we analyzed the influence of different parameters on the performance of our network.

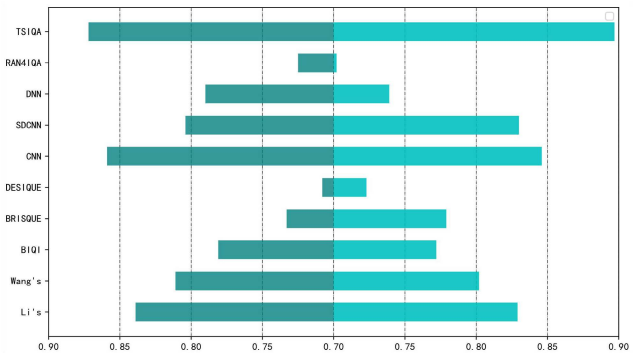

FIGURE 9. Performance comparison of defocus blur distortion.

TABLE 2. Performance comparison of various methods on the SEM image quality dataset. The best results are written in bold.

Metrics	Type	SRCC	PLCC	RMSE
MS-SSIM[44]	FR	0.6116	0.6017	0.6846
PSNR[45]	FR	0.5914	0.5587	0.5839
GMSD[46]	FR	0.5877	0.5561	0.5859
SSIM[47]	FR	0.5618	0.5409	0.5222
FSIM[48]	FR	0.4793	0.4532	0.6355
PCQI[49]	FR	0.4721	0.4462	0.6792
DESIQUE[16]	NR	0.6938	0.7038	0.5431
BIQI[11]	NR	0.7521	0.7641	0.4932
QAC[18]	NR	0.4298	0.4485	0.6833
NIQE[17]	NR	0.6316	0.6554	0.5775
BRISQUE[14]	NR	0.7494	0.7299	0.5601
CORNIA[15]	NR	0.8199	0.7933	0.3996
BLiinds2[13]	NR	0.7216	0.6965	0.4786
DIIVINE[12]	NR	0.5854	0.5608	0.5611
SSEQ[50]	NR	0.5789	0.5464	0.5672
CNNIQA[4]	NR	0.8927	0.8770	0.3674
PQR[22]	NR	0.8579	0.8279	0.4364
SDCNN[27]	NR	0.8723	0.8625	0.3870
DNNIQA[51]	NR	0.8733	0.8490	0.4040
RANK4IQA[52]	NR	0.8357	0.5494	0.6389
Hallucinated-IQA[53]	NR	0.8625	0.8328	0.4125
TSIQA	NR	**0.9172**	**0.8980**	**0.3412**

1) COMPARISON OF STATE-OF-THE-ART METHODS

The proposed method was compared with 21 state-of-the-art methods: six full-reference methods, nine no-reference methods based on traditional machine learning, and six deep learning-based methods. In Table 2, bold horizontal lines are used to distinguish between the three types of methods. To prevent overfitting, we adjusted the estimator modules in RANk4IQA to 3 for a fair comparison. The comparison results are presented in Table 1. Our proposed model is superior to other state-of-the-art methods in terms of SRCC, PLCC, and RMSE metrics. This result is consistent with our conjecture from SEM image features; that is, the key to obtaining good performance of SEM-IQA is to extract texture features effectively.

2) ABLATION STUDY

We conduct ablation research to confirm the effectiveness of our method. In the baseline method, network architecture, training settings, and training data of IQA are the same as those of TSIQA. But the difference is that the baseline method adopts the traditional dense convolution instead of sparse convolution using a sparse mask, and it does not adopt the IGA but the traditional attention mechanism and does not use the texture representation generated by sparse convolution to retrain. The SRCC and PLCC performances of the baseline algorithm throughout training are 0.8988 and 0.8786, respectively. Table 3 shows a comparison of the performance of ablation studies. After we replace the attention mechanism in the baseline with IGA, the performance of the model is improved when learning deeper semantic information. Then, on this basis, we use the sparse mask obtained in the first training to perform a sparse convolution instead of a dense convolution. As a result, the performance is improved, which also confirms the effectiveness of sparse convolution and our idea that texture can better represent the quality of SEM images. Finally, we train the network again with the texture representation generated by sparse convolution such that the network will learn to evaluate the quality with the texture features that can better represent the image. After the second training, the performance is significantly improved again and reaches the best. In addition to the overall performance, our index exceeds the baseline. As shown in Table 4, our method has also gained advantages in evaluating the performance of various distorted pictures, which proves that our method has good robustness and generalization.

We use sparse convolution to significantly skip the computation of redundant information. Thus, our method also reduces the computation cost. As shown in Table 5, our parameter quantity has increased by about 5% compared with that of the baseline algorithm, but FLOPs has decreased by about 28%. Consequently, the processing efficiency of our algorithm on CPU has also increased by about 38%. However, on GPU, sparse convolution cannot fully utilize the characteristics of GPU acceleration (such as memory merging). Thus, the advantages of our method cannot be reflected on GPU without specific optimization.

3) EFFECT OF PARAMETERS

We further study the influence of different parameters on the performance of the model. For λ_0, it is positively correlated with sparsity. The larger λ_0 corresponds to the higher sparsity we force the model to produce in the training process and the more information of the non-texture region skipped by the model. As shown in Figure10, we find that λ_0 is larger and the performance of the model is better when the number of SMMs is relatively small. Meanwhile, the smaller λ_0 can achieve better results when the number of SMM modules is relatively large. The reason may be that the complexity of the model increases when the number of SMMs is large. At this time, if λ_0 is too large and sparse convolution skips too much picture information, then over-fitting easily occurs in training for SEM datasets with relatively small number, which leads to poor performance on the test set. Our experiment shows that the performance is the best when the number of SMMs and λ_0 are set to 5 and 0.1, respectively.

TABLE 3. Ablation study results.

Model	IGA	Sparse mask	Second Training	SRCC	PLCC	RMSE
Baseline	×	×	×	0.8988	0.8786	0.3652
Baseline + IGA	√	×	×	0.9051	0.8871	0.3589
Baseline+SparseConv+IGA	√	√	×	0.9134	0.8924	0.3487
TSIQA	√	√	√	**0.9172**	**0.8980**	**0.3412**

TABLE 4. Comparison of the performance of ablation study on overall dataset and various distortion types. The first two rows represent SRCC performance, the last two rows represent PLCC performance.

Perf.	All	Noise	Cont.	Blur	Brig.	Asti.
Baseline	0.8988	0.8439	0.8682	0.8394	0.8151	0.7539
TSIQA	**0.9172**	**0.9011**	**0.8896**	**0.8721**	**0.8331**	0.7534
Baseline	0.8787	0.8254	0.8544	0.8325	0.8029	0.7964
TSIQA	**0.8980**	**0.8912**	**0.8843**	**0.8878**	**0.8274**	**0.8124**

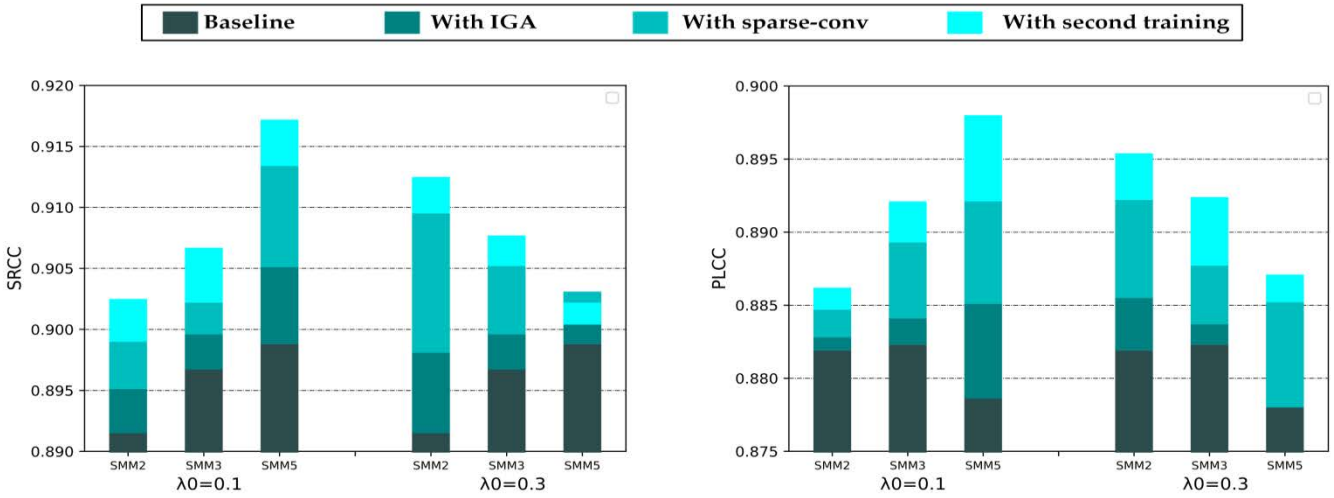

FIGURE 10. SRCC and PLCC performance of different SMM modules at $\lambda_0 = 0.1$ and $\lambda_0 = 0.3$.

TABLE 5. Comparison of computational efficiency of ablation study.

Model	Params	FLOPs	Time	
			GPU	CPU
Baseline	1582K	1.00×	1.00×	1.00×
TSIQA	1653K	0.72×	1.21×	0.62×

TABLE 6. Influence of image size on the performance of baseline algorithm.

Patch Size	32	64	128	256	512
SRCC	0.8563	0.8853	0.8988	0.8971	0.8822
PLCC	0.8324	0.8437	0.8787	0.8763	0.8549
Average	0.8443	0.8645	0.8887	0.8867	0.8685

The IQA task is a typical small sample problem. Thus, the robustness of the model is also the key factor to measure the performance of the model. Therefore, we randomly remove samples from the dataset and continuously reduce the number of training samples to observe the performance of the model. As shown in Figure 11, our model only uses about 40% of the data to achieve the same performance as the baseline algorithm on the complete dataset.

As mentioned in Section III.A, the original SEM image is cut into square image blocks for data enhancement, and different patch sizes influence the performance of the algorithm. We experiment on the baseline to determine the best input image block size. Specifically, we first use a sliding window with a fixed size to cut out the image block, in which the stride length of the sliding window is adaptively determined according to the principle of covering the entire image with minimal overlap. We set the sizes of the sliding windows to 512, 256, 128, 64, and 32, respectively, and we cut each image

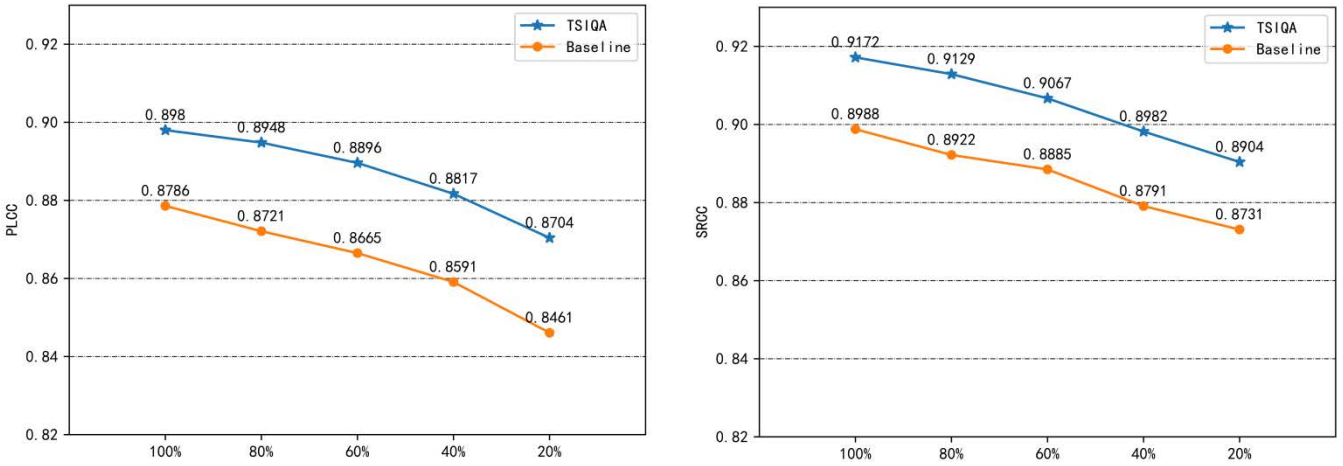

FIGURE 11. SRCC and PLCC performances of TSIQA and baseline as the number of training samples decreases. The horizontal axis indicates the sampling ratio of the training set.

TABLE 7. Performance comparison of various methods on the natural image quality dataset. The best results are written in bold.

Metrics	TYPE	CSIQ [54]		TID2013 [55]	
		SRCC	PLCC	SRCC	PLCC
MS-SSIM[44]	FR	0.906	0.889	0.786	0.830
PSNR[45]	FR	0.810	0.819	0.687	0.677
GMSD[46]	FR	**0.950**	**0.945**	0.804	0.855
SSIM[47]	FR	0.865	0.852	0.727	0.777
DESIQUE[16]	NR	0.700	0.724	0.274	0.329
BIQI[11]	NR	0.830	0.849	0.325	0.375
NIQE[17]	NR	0.247	0.303	0.025	0.098
BRISQUE[14]	NR	0.641	0.665	0.323	0.388
BLiinds2[13]	NR	0.490	0.551	0.223	0.267
DIIVINE[12]	NR	0.708	0.744	0.287	0.330
CNNIQA[4]	NR	0.842	0.836	0.803	0.817
DNNIQA[51]	NR	0.909	0.901	0.795	0.812
RANK4IQA[52]	NR	0.912	0.902	0.780	0.791
Hallucinated-IQA[53]	NR	0.938	0.925	**0.879**	**0.884**
TSIQA	NR	0.863	0.854	0.754	0.767

into small pieces of these sizes. Then, the performance of the baseline network is tested separately for each clipping method. Table 6 shows the SRCC and PLCC performance of the baseline algorithm when the original image is cut into image blocks of different sizes. The average of SRCC and PLCC performances is used as the index to select the image block size. As observed, the performance is the best when the image block size is 128. Thus, we choose the patch with the size of 128 to obtain more robust results. All experiments use a batch size of 16 to train the network for fair comparison.

4) GENERALIZATION ABILITY

Further, to verify the generalization ability of the TSIQA, the cross-database experiments are implemented on the natural image quality assessment database. We selected CSIQ [54] and TID2013 [55] datasets.The CSIQ [54] database has a total 866 test images,consisting of 30 originals and 6 types of distortions.TID2013 [55] database contains 25 originals and distorted images with 18 types of distortions.As shown in Table 7, TSIQA specially designed for SEM images can still assess the image quality normally and achieve relatively good results on natural image datasets.

IV. CONCLUSION

In this study, a new SEM image quality assessment network based on intuitive morphology and deep semantic features (TSIQA) is proposed to overcome the problems of the poor performance of traditional natural image IQA models on SEM image datasets and the lack of a robust general distortion evaluation algorithm in the SEM-IQA field. TSIQA is designed for the sensitivity of SEM image texture information to distortion,it selects the intuitive texture features and deep semantic features that can best represent SEM images quality for assessment.According to the experimental results on the SEM IQA dataset, our method achieves 0.9172 and 0.8980 values on the typical IQA indicators SRCC and PLCC, respectively,it is superior to the state-of-the-art IQA methods in terms of evaluation accuracy,and also achieved the best performance on different types of distortion subsets. At the same time, compared with the baseline algorithm, the FLOPs

is reduced by 28%, and the CPU speed is increased by 38%. In terms of network robustness,TSIQA can use 40% data to achieve the performance of baseline algorithm on complete dataset.We also maintain a relatively stable level when the training samples are reduced.

REFERENCES

[1] D. Chen, A. Miyamoto, and S. Kaneko, "Robust surface reconstruction in SEM using two BSE detectors," *IEICE Trans. Inf. Syst.*, vol. E96.D, no. 10, pp. 2224–2234, 2013.

[2] M. Lee, J. Cantone, and J. Xu, "Improving SEM image quality using pixel super resolution technique," in *Proc. 28th Metrol., Inspection, Process Control Microlithography*, vol. 16, 2014, Art. no. 90500U.

[3] Z. Ruan, S. F. Mao, and P. Zhang, "Monte Carlo simulation of realistic beam-sample interaction in SEM: Application to evaluation of sharpness measurement methods," *Scanning Microscopies.*, vol. 8729, pp. 90–103, May 2013.

[4] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1733–1740.

[5] J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang, and A. C. Bovik, "Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 130–141, Nov. 2017.

[6] H. Wang, X. Hu, H. Xu, S. Li, and Z. Lu, "No-reference quality assessment method for blurriness of SEM micrographs with multiple texture," *Scanning*, vol. 2019, pp. 1–15, Jun. 2019.

[7] Q. Li, L. Li, Z. Lu, Y. Zhou, and H. Zhu, "No-reference sharpness index for scanning electron microscopy images based on dark channel prior," *KSII Trans. Internet Inf. Systems.*, vol. 13, no. 5, pp. 2529–2543, 2017.

[8] Z. Lu, Z. Zhang, Y. Wang, L. Dong, and S. Liang, "SEM image quality assessment based on texture inpainting," *IEICE Trans. Inf. Syst.*, vol. E104.D, no. 2, pp. 341–345, 2021.

[9] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[10] W. Lu, K. Zeng, D. Tao, Y. Yuan, and X. Gao, "No-reference image quality assessment in contourlet domain," *Neurocomputing*, vol. 73, no. 4, pp. 784–794, 2010.

[11] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 513–516, May 2010.

[12] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3350–3364, Dec. 2011.

[13] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3339–3352, Aug. 2012.

[14] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012.

[15] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1098–1105.

[16] Y. Zhang and D. M. Chandler, "No-reference image quality assessment based on log-derivative statistics of natural scenes," *J. Electron. Imag.*, vol. 22, no. 4, Dec. 2013, Art. no. 043025.

[17] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[18] W. Xue, L. Zhang, and X. Mou, "Learning without human scores for blind image quality assessment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2013, pp. 995–1002.

[19] N. M. Murad, L. Rejeb, and L. B. Said, "The use of DCNN for road path detection and segmentation," *Iraqi J. Comput. Sci. Math.*, vol. 3, no. 2, pp. 119–127, 2022.

[20] L. Kang, P. Ye, Y. Li, and D. Doermann, "Simultaneous estimation of image quality and distortion via multi-task convolutional neural networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2791–2795.

[21] S. Bosse, D. Maniry, K. R. Müller, T. Wiegand, and W. Samek, "Deep neural networks for no-reference and full-reference image quality assessment," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 206–219, Jan. 2018.

[22] H. Zeng, L. Zhang, and A. C. Bovik, "A probabilistic quality representation approach to deep blind image quality prediction," 2017, *arXiv:1708.08190*.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[24] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Image quality assessment: Unifying structure and texture similarity," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, pp. 2567–2581, 2022.

[25] E. Prashnani, H. Cai, Y. Mostofi, and P. Sen, "PieAPP: Perceptual image-error assessment through pairwise preference," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1808–1817.

[26] P. G. Freitas, W. Y. L. Akamine, and M. C. Q. Farias, "No-reference image quality assessment using orthogonal color planes patterns," *IEEE Trans. Multimedia*, vol. 20, no. 12, pp. 3353–3360, Dec. 2018.

[27] S. Jia and Y. Zhang, "Saliency-based deep convolutional neural network for no-reference image quality assessment," *Multimedia Tools Appl.*, vol. 77, no. 12, pp. 14859–14872, 2017.

[28] D. Chen, Y. Wang, T. Wu, and W. Gao, "An attention-driven approach of no-reference image quality assessment," 2016, *arXiv:1612.03530*.

[29] V. Mnih, N. Heess, and A. Graves, "Recurrent models of visual attention," in *Proc. Adv. Neural Inf. Process. Systems.*, 2014, pp. 2204–2212.

[30] V. Mnih, A. Badia, and P. Mirza, "Asynchronous methods for deep reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 1928–1937.

[31] K. Ma, W. Liu, K. Zhang, Z. Duanmu, Z. Wang, and W. Zuo, "End-to-end blind image quality assessment using deep neural networks," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1202–1213, Mar. 2017.

[32] J. Kim, A.-D. Nguyen, and S. Lee, "Deep CNN-based blind image quality predictor," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 1, pp. 11–24, Jan. 2018.

[33] J. Kim and H. Zeng, "Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment," *IEEE Signal Process. Mag.*, vol. 34, pp. 130–141, 2017.

[34] J. Wu, J. Ma, F. Liang, W. Dong, G. Shi, and W. Lin, "End-to-end blind image quality prediction with cascaded deep neural network," *IEEE Trans. Image Process.*, vol. 29, pp. 7414–7426, 2020.

[35] P. G. Freitas, S. Alamgeer, W. Y. L. Akamine, and M. C. Q. Farias, "Blind image quality assessment based on multiscale salient local binary patterns," in *Proc. 9th ACM Multimedia Syst. Conf.*, Jun. 2018, pp. 52–63.

[36] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1395–1403.

[37] L. Wang, X. Dong, Y. Wang, X. Ying, Z. Lin, W. An, and Y. Guo, "Exploring sparsity in image super-resolution for efficient inference," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 4917–4926.

[38] E. Jang, S. Gu, and B. Poole, "Categorical reparameterization with Gumbel-Softmax," in *Proc. Int. Conf. Learn. Represent.*, 2017, pp. 1–13.

[39] Z. Li, G. Li, T. Li, S. Liu, and W. Gao, "Information-growth attention network for image super-resolution," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 544–552.

[40] Q. Li, Z. Lu, and L. Li, "Perceptual evaluation of scanning electron microscopy imaging," *J. Netw. Intell.*, vol. 1, pp. 10–16, Feb. 2019.

[41] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.

[42] Y. Zhou, L. Li, S. Wang, J. Wu, Y. Fang, and X. Gao, "No-reference quality assessment for view synthesis using DoG-based edge statistics and texture naturalness," *IEEE Trans. Image Process.*, vol. 28, no. 9, pp. 4566–4579, Sep. 2019.

[43] Y. Zhou, Y. Sun, L. Li, K. Gu, and Y. Fang, "Omnidirectional image quality assessment by distortion discrimination assisted multi-stream network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 4, pp. 1767–1777, Apr. 2022.

[44] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, 2003, pp. 1398–1402.

[45] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.

[46] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Trans. Image Process.*, vol. 23, no. 2, pp. 684–695, Feb. 2014.

[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[48] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 8, no. 8, pp. 2378–2386, Aug. 2011.
[49] S. Wang, K. Ma, H. Yeganeh, Z. Wang, and W. Lin, "A patch-structure representation method for quality assessment of contrast changed images," *IEEE Signal Process. Lett.*, vol. 22, no. 12, pp. 2387–2390, Dec. 2015.
[50] L. Liu, B. Liu, H. Huang, and A. C. Bovik, "No-reference image quality assessment based on spatial and spectral entropies," *Signal Process., Image Commun.*, vol. 29, no. 8, pp. 856–863, Sep. 2014.
[51] S. Bosse, D. Maniry, T. Wiegand, and W. Samek, "A deep neural network for image quality assessment," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3773–3777.
[52] H. Ren, D. Chen, and Y. Wang, "RAN4IQA: Restorative adversarial nets for no-reference image quality assessment," in *Proc. AAAI Conf. Artif. Intell.*, 2018, pp. 7308–7314.
[53] K.-Y. Lin and G. Wang, "Hallucinated-IQA: No-reference image quality assessment via adversarial learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 732–741.
[54] E. C. Larson and D. Chandler. (2010). *Categorical Image Quality (CSIQ) Database*. Accessed: Sep. 30, 2016. [Online]. Available: http://vision.okstate.edu/csiq
[55] N. Ponomarenko, L. Jin, and V. Lukin, "Image database TID2013: Peculiarities, results and perspectives," *Signal Process. Image Commun.*, vol. 30, pp. 57–77, Jan. 2015.

HAORAN WANG received the B.S. degree in communication engineering from the Hebei University of Science and Technology, Shijiazhuang, Hebei, China. He is currently pursuing the M.S. degree with the China University of Mining and Technology, Xuzhou, Jiangsu, China.

His research interest includes computer vision.

SHIYIN LI received the B.S. and Ph.D. degrees from the China University of Mining and Technology, Xuzhou, Jiangsu, China.

He is currently a Professor with the China University of Mining and Technology. He has published over 100 papers in domestic and international academic journals and conference proceedings. His research interests include wireless communication and image processing.

JICUN DING is a Young and Middle-Aged Expert with outstanding contributions in Jiangsu Province (professional and technical personnel). He is currently the President of Xuzhou First People's Hospital. He has presided over more than ten scientific research projects. His research interests include scanning electron microscopy and biomedical image processing.

SUYAN LI is currently a Professor with Xuzhou First People's Hospital. She is also a Key Medical Talent in Xuzhou. She has published over 20 papers in domestic and international academic journals and conference proceedings. Her research interest includes ophthalmology image processing.

LIANG DONG received the Ph.D. degree from the China University of Mining and Technology, Xuzhou, Jiangsu, China.

He is currently a Professor with the China University of Mining and Technology. He is also a National Young Talent Project Candidate. He has published over 30 papers in SCI journals. His research interest includes application of image processing in chemical industry.

ZHAOLIN LU received the M.S. degree from Xidian University, in 2006, and the Ph.D. degree from the China University of Mining and Technology, Xuzhou, Jiangsu, China.

He is currently a Professor with the China University of Mining and Technology. He has published over 20 papers in domestic and international academic journals and conference proceedings. Currently, he is an in charge of one project of National Natural Science Foundation of China, one project of basic scientific research of central universities. His research interest includes image processing.