IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 4 February 2024, accepted 27 February 2024, date of publication 14 March 2024, date of current version 21 March 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3375940

RESEARCH ARTICLE

Blind Image Quality Evaluation Method Based on Cyclic Generative Adversarial Network

XIAOYING LI AND SHOUWU HE
College of Computer Application, Guilin University of Technology at Nanning, Nanning 530001, China
Corresponding author: Xiaoying Li (lxygliten@126.com)

This work was supported by Guangxi Education Department Basic Research Ability Promotion Project for young and middle-aged teachers in universities, China, under Grant 2019KY0270.

ABSTRACT The mission of blind image quality evaluation is currently a challenging computer vision problem. Due to the shortage of reference images, it is hard for blind image quality evaluation methods to achieve the same performance as full reference image quality evaluation methods. In addition, current quality evaluation methods are difficult to effectively forecast the quality scores of synthesized distorted images as well as real distorted images. To address such issues, this study proposed a cyclic generative adversarial network composed of a quality perception network and a quality regression network on the grounds of generative adversarial networks. For further enhancing the predictive performance of quality aware networks, this study proposed using attention blocks for adaptively fusing high-level semantic features and low-level semantic features. It extracted content and distortion information from images through an image quality evaluation method on the grounds of content perception and distortion inference. And according to the different properties of the extracted features, adaptive fusion blocks were used for adaptively fusing content features and distortion features. Experiments showcased that the Spearman order correlation coefficient and Pearson linear correlation coefficient obtained by the proposed method on multiple datasets were higher than other similar methods. At the same time, the proposed method has achieved good prediction results on various types of distorted images, and has surpassed other methods. The prediction accuracy of the proposed method on five types of distortion was 0.971, 0.963, 0.984, 0.971, and 0.926, respectively. The proposed method achieved the highest predictive accuracy on all distortion types in the LIVE dataset, with predicted accuracy values of 0.973, 0.965, 0.984, 0.963, and 0.944, respectively. In summary, the proposed method not only achieved good prediction accuracy, but also had strong generalization performance in cross dataset testing. This provides a scientific and effective research direction for blind image quality evaluation.

INDEX TERMS Blind images, deep learning, attention block, quality score, evaluation.

I. INTRODUCTION

The image quality assessment aims to make the predicted score as reflective of human visual perception as possible. The human vision system is an excellent image processing system, which is good at object recognition and quality judgment, and provides simulated targets for machine vision algorithms to enhance their accuracy [1], [2]. When people perceive images, the visual system is combined with the Internal Generative Mechanism (IGM) to work together [3].

The associate editor coordinating the review of this manuscript and approving it for publication was Daniel Augusto Ribeiro Chaves.

Current blind image quality assessment techniques use end-to-end deep learning models to identify image distortion and predict quality scores. However, when faced with the diverse types of distortion in images, a single model often struggles to adapt to this complexity. In contrast, the human brain IGM can deduce the main content of the image when processing chaotic distorted images, demonstrating its high efficiency of human brain in image analysis [4]. When reasoning about the content of an image becomes more difficult, it usually means that the image has a higher degree of distortion. On the other hand, if the reasoning is simpler, the distortion degree is lower. Driven by deep learning technology, remarkable

progress has been made in the field of image quality evaluation. The main advantage of deep learning is its ability to build end-to-end models that automatically extract and generalize the features of training samples to form effective quality prediction models. This allows for deeper mining and exploitation of complex patterns in image data, effectively simulating how the human visual perception system works [5]. In order to deal with the limitations of a single model in dealing with complex image distortion, this paper adopts deep learning technology, combines the quality perception network and quality regression network of Generative Adversarial Network (GAN), and innovatively simulates the process of image quality degradation. This method not only focuses on synthetic distorted images, but also can predict the quality of real distorted images. This paper proposes an image quality evaluation method based on content perception and distortion inference. The method imitates the perception process of the human visual system to accurately evaluate image quality. The proposed method is innovative in that it integrates a deep understanding of image content with accurate recognition of distortion types. By optimizing antagonism and cyclic consistency loss, it generates evaluation results that are more consistent with human visual perception.

The proposed method simulates the whole process of image quality degradation and breaks through the limitation of dealing with various image distortions. The proposed method takes into account the complexity of images and the diversity of distortions, so that the multi-dimensional quality characteristics of images can be understood more deeply. This achievement represents significant progress in the field of image quality evaluation. In addition, it is proposed that attention blocks come from adaptive fusion of higher semantic features and lower semantic features. On this basis, the hierarchical recovery features of quality perception network are extracted as the basis of quality prediction. In the data set of the pre-trained quality perception network, the perception ability and prediction accuracy of the real distortion type are improved by adding the real distortion image. This study is to further enhance the predictive accuracy of Image Quality Evaluation (IQE) methods by integrating human visual system and DL technology.

The research mainly includes five parts, and in the first part of the article, the background and significance of IQE are mainly introduced. The content of Part 2 provides a comprehensive overview of IQE. The third part is the proposed method, mainly divided into two sections. In section III-A, a model of image quality degradation process on the grounds of cyclic GANs was studied and constructed. In section III-B, a blind IQE improvement model integrating content perception and distortion inference was constructed. The fourth part is about verifying the effectiveness of the proposed model. The fifth part is a summary of the proposed methods and an analysis of the experimental results. At the same time, the shortcomings of proposed methods and future research directions are proposed.

II. RELATED WORK

Currently, multimedia digital images, as carriers of information transmission, are possessing an essential influence on people's lives. Domestic and foreign researchers have also made many studies on IQE. Researchers are attempting to further enhance the quality evaluation through the powerful extraction of Convolutional Neural Network (CNN). To reduce the size of image files while ensuring image quality, Ma D et al. proposed a CNN-based image partitioning quality evaluation method. This study shows that this method has strong advantages in image compression time and compression efficiency [6]. Yang X et al. proposed a novel reference-free image quality evaluation method using transfer learning technology. In this method, a network called TTL is introduced, which focuses on the transmission of semantic features, so as to optimize the feature sharing process. Experiments on multiple datasets show that the method exhibits excellent generalization ability, indicating that it works effectively on different types of image data [7].Wu J et al. presented a lightweight quality evaluation method to extract hierarchical features of images. This study indicates that this method greatly reduces the parameter complexity of the network while ensuring prediction accuracy [8]. Fei L et al. proposed a CNN-basednon-reference IQE algorithm to address the issues of non universality in high dynamic range IQE and excessive reliance on original reference images. This method uses the Salience Detection By Self assembly (SDSR) algorithm to extract salient regions of the image. This study indicates that this method has high consistency with the subjective perception of the human body [9].

Lu Y et al. proposed an image quality evaluation algorithm that integrates multi-scale and dual-domain features. The algorithm uses two branches to process the original image and its phase consistency graph. It also employs feature pyramid and Attention Mechanisms (AM) to extract multi-level features. Bilinear pooling technology is used to integrate spatial and frequency domain features, and support different scale image input to extract deep quality features. The results show that the algorithm is both robust to different types and cross-databases and sensitive to scale [10].To address the instability and stability of GAN in the relevant process, Tang Y et al. proposed a gradient penalty based single image super-resolution acquisition method on the grounds of GAN. This study indicates that this method outperforms traditional methods in terms of accuracy in image resolution acquisition and perceptual quality [11]. JinT et al. proposed a GAN based method for establishing a composite crack image dataset with pixel level annotations. The results indicate that this method provides a new approach for traditional data augmentation methods [12]. Zhang L et al. proposed a fusion particle swarm optimization algorithm to solve the instability and pattern collapse problems of GAN algorithm training. This is to improve the stability of training. This study indicates that it can enhance the facial image generation and has excellent robustness [13].

In summary, existing IQE methods are difficult to accurately predict the process of image quality degradation. Therefore, it is difficult to effectively extract content and distortion information from images. Therefore, to improve the predictive accuracy of IQE methods, this study proposes to study IQE methods on the grounds of cyclic GANs and IQE methods on the grounds of content perception and distortion inference.

III. PROPOSED METHOD

CycleGAN is a deep learning model for image-to-image conversion. Based on CycleGAN, the research introduces deep learning techniques and innovatively combines a Quality Prediction Network (QPN) and a Quality Recovery Network (QRN). It simulates the whole process of image quality degradation and breaks through the limitations of dealing with various image distortions. In addition, an image quality evaluation method based on content perception and distortion inference is proposed. The image's complexity and distortion diversity are taken into account to extract content and distortion information. This leads to a deeper understanding of the image's multidimensional quality characteristics.

A. RESEARCH ON THE CONSTRUCTION OF IMAGE QUALITY DEGRADATION PROCESS MODEL ON THE GROUNDS OF RECURRENT GANS

GANs can better model data distribution and theoretically train any generator network. And GANs is not necessary for utilizing the Markov chain for repeated sampling and inference during the learning. It avoids the difficulty of approximate calculation of probabilities without a complex variational lower bound [14]. GANs are composed of two important parts, namely generators and discriminators. The generator network produces more realistic data samples by learning the distribution of real sample data in the training set [15]. The discriminator network will make endeavor to distinguish in the generated data and the real data samples. The structure of the GAN is shown in Figure 1.

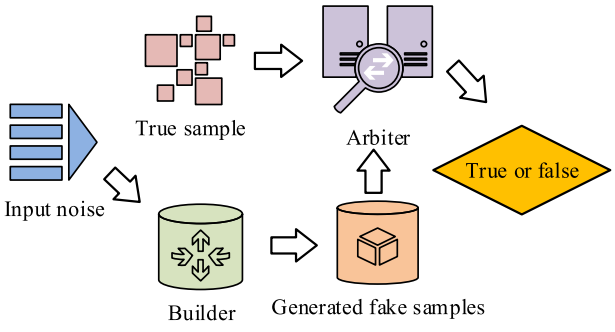


FIGURE 1. Workflow of GAN.

In Figure 1, the generator and discriminator form a dynamic game process. The entire network will continuously update the network parameters of both over time, and optimization will stop at the minimum value relative to the generator and the maximum value relative to the discriminator. The training goal of GANs is to achieve Nash Equilibrium, and its objective function is shown in equation (1) [16].

$$\min\max L(D,G) = E_{x\sim Pdata(x)}\left[\log D(x)\right] + E_{Z\sim Pz(z)}\left[\log\left(1-D\left(G\left(z\right)\right)\right)\right] \quad (1)$$

In equation (1), $G()$, $D()$ represent the generator network and discriminator network. x serves as input noise, and z serves as real data. P_{data}, P_z serve as the distribution of real data and input noise. The generator network minimizes the value of $\log(1-D(G(z)))$ and deceives the discriminator by generating more "real" samples. The discriminator network assigns the correct labels to the real data and the generated data through training iterations. However, during the trainingof GANs, unstable situations such as gradient disappearance and pattern collapse may occur. To stabilize the training process of the network, scholars such as Arjovsky proposed using Earth Mover to measure the closeness between the generated data distribution as well as the real data distribution [17]. Traditional GANs using Multi-Layer Perception (MLP) are difficult to perform satisfactorily when faced with image samples containing complex information. To effectively address this issue, GANs use CNNs to replace traditional multi-layer perceptron structures to process image samples. Although the use of CNNs for image quality assessment has achieved good accuracy, the following challenges still exist. Firstly, distorted images can suffer from various types of distortion. Secondly, due to the lack of reference information, blind IQE methods find it difficult to achieve predictive accuracy comparable to full reference IQE methods. To address these two issues, a cyclic GAN was established to perceive the process of image quality degradation. To train the QPN, the reference image I_r is first fed into the generator network of the loss recovery module to obtain a false loss image I'_d. Then, it is fed into the lossless restoration module generator network for learning how to remove distortion information and obtain image I'_d again. In the loss recovery module and lossless recovery module, the function of the discriminator network is to distinguish the generated images, thereby generating more realistic distorted images as well as reference images. Figure 2 shows the network structure of the discriminator.

In Figure 2, the VGG-16 network uses two $3*3$ Convolutional Layers (CL) instead of the $5*5$ CL, and replaces the $7*7$ CL with three $3*3$ CL. Its purpose is for decreasing the number of network parameters while ensuring that the perception domain remains unchanged. Because of the superiority of U-Net network in image processing tasks, its unique encoder decoder structure combines low-level and advanced semantic information, improving the network's feature extraction ability. Therefore, to enhance the feature extraction of the overall network, this structure is studied as the backbone network of the generator network in the loss

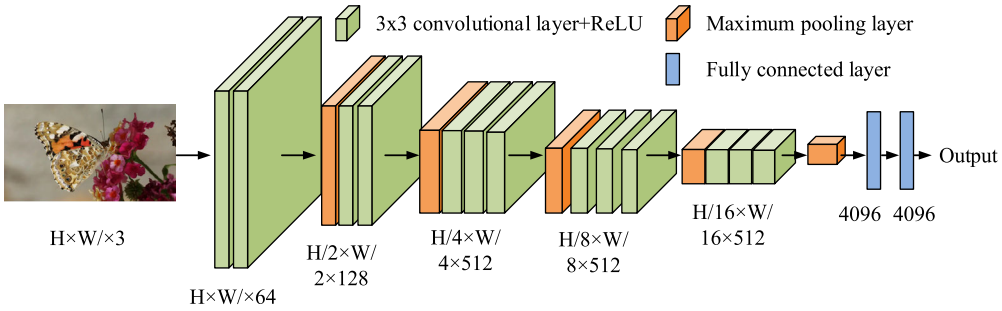


FIGURE 2. VGG-16 network structure of discriminator.

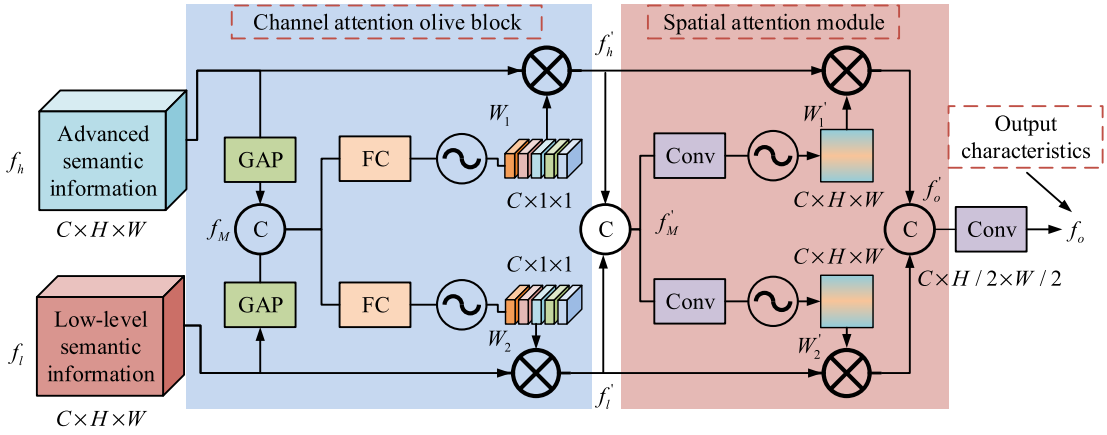


FIGURE 3. Network structure of attention blocks.

recovery module and the lossless recovery module. After giving an image, first extract low-level features using a 3 ∗ 3 CL. Then, an encoder model composed of 5 residual blocks and 4 maximum pooling layers extracts features of different scales. In the U-Net network, the encoder and decoder contain low-level semantic information and advanced semantic information, respectively. For fully and effectively utilizing different types of features, this study proposes using attention blocks to fuse the most useful features while suppressing invalid features. Figure 3 shows the network structure of attention blocks.

In Figure 3, the attention block includes a channel Attention Module (AM) and a spatial AM. The purpose of the channel AM is to learn which type of feature is more essential, while the focus of the spatial AM is to learn where important regions are. The channel AM aims to calculate the weight values between channels by learning two one-dimensional channel attention maps. Firstly, it performs Global Average Pooling Layer (GAPL) processing on the features in the image, and then performs stitching. Finally, two fully connected layers with 2C, 16, and C nodes are utilized for calculating the weight score of the input features. Finally, a one-dimensional channel attention map is produced through the Sigmoid activation function. For effectively generating spatial attention maps, the spatial AM uses CL for learning the spatial relationships of features. The relevant expression of this module is showcased in equation (2).

$$\begin{cases} f'_M = f'_l \otimes f'_h \\ W'_1 = \sigma\left(Conv_1\left(f'_M\right)\right) \\ W'_2 = \sigma\left(Conv_2\left(f'_M\right)\right) \\ f'_o = \left(f'_l \times W'_1\right) \otimes \left(f'_h \times W'_2\right) \end{cases} \tag{2}$$

In equation (2), both f'_l, f'_h represent the input features. $Conv_1()$, $Conv_2()$ represent two separate 3 ∗ 3 CL. $\sigma()$ represents the sigmoid activation function. $\otimes$ represents concatenation operations, respectively. In the pre-training stage of the quality aware network, the generators and discriminators of the loss recovery module and the lossless recovery module optimize the overall network through a maximum minimum game. In the game, the relevant network parameters the loss recovery module are updated by minimizing adversarial losses. The process is shown in equation (3).

$$L_{R2D} = E\left[\log D_{R2D}\left(I_d\right)\right] + E\left[\log\left(1 - D_{R2D}\left(G_{R2D}I_r\right)\right)\right] \tag{3}$$

In equation (3), I_d, I_r represent distorted images and reference images, respectively. D_{R2D}, G_{R2D} represent the discriminator and generator networks of the loss recovery module, respectively. In the stage of learning lossless image restoration in the lossless restoration module, the relevant network

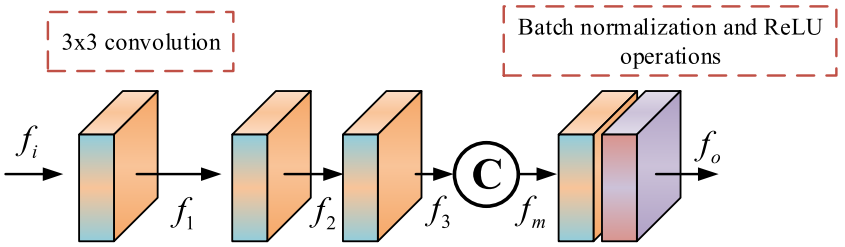


FIGURE 4. Network structure of multi-scale blocks.

parameters are updated by minimizing adversarial losses. The process is shown in equation (4).

$$L_{D2R} = E\left[\log D_{D2R}\left(I_r\right)\right] + E\left[\log\left(1 - D_{D2R}\left(G_{D2R}I'_d\right)\right)\right] \tag{4}$$

In equation (4), I'_d represents the false distorted image generated by the loss recovery module. D_{D2R}, G_{D2R} represent discriminator and generator networks for lossless restoration modules. The adversarial loss of the overall network is shown in equation (5).

$$L_{adv} = L_{D2R} + L_{R2D} \tag{5}$$

To further ensure that the generated image and the target image are the same image content, the study also adopted cyclic consistency loss [18]. The purpose is to convert the generated outcomes back into the source domain again. The calculation of cyclic consistency loss is shown in equation (6).

$$L^m_{cycle} = \left\| G_{D2R}\left(I'_d\right) - I_r \right\|^2 \tag{6}$$

To further ensure that the generated images contain rich details and texture information, the study also adopted a perceptual loss function. The calculation is shown in equation (7) [19].

$$L^p_{cycle} = \frac{1}{H_n W_n} \sum_{h=1}^{H_n} \sum_{w=1}^{W_n} \left\| \phi_n\left(G_{D2R}\left(I'_d\right)\right) - \phi_n\left(I_r\right) \right\| \tag{7}$$

In equation (7), H_n, W_n represent the height and width of the input feature image. It represents the ϕ_n ()-th layer of the VGG-19 network. The calculation expression of the comprehensive cycle consistency loss is shown in equation (8) [20].

$$L_{cycle} = L^m_{cycle} + L^p_{cycle} \tag{8}$$

The objective loss function for optimizing quality aware networks is shown in equation (9).

$$L = L_{adv} + L_{cycle} \tag{9}$$

For markedly establishing the relation in the extracted hierarchical features and the predicted quality score, the QRN uses four multi-scale blocks for feature extraction. The multi-scale blocks used are shown in Figure 4 [21].

In Figure 4, this method utilizes three CL to improve the expression ability of multi-scale features. The operation of multi-scale blocks is showcased in equation (10).

$$\begin{cases} f_1 = Conv\left(f_i\right) \\ f_2 = Conv\left(f_1\right) \\ f_3 = Conv\left(f_2\right) \\ f_m = f_1 \otimes f_2 \otimes f_3 \\ f_o = \delta\left(f_m\right) \end{cases} \tag{10}$$

In equation (10), δ () represents the batch normalization operation and the correction linear unit operation. Finally, the output features of each multi-scale block are regressed into quality scores through fully connected layers with 480, 512, and 512 nodes.

B. RESEARCH ON AN IMPROVED MODEL FOR BLIND IQE INTEGRATING CONTENT PERCEPTION AND DISTORTION REASONING

The proposed QPN models the process of introducing distorted information into images. However, distorted images contain both rich content and distorted information. The overall network lacks a focus on extracting content features, which to some extent limits the improvement of network prediction accuracy. The prediction results of IQE algorithms should be as consistent as possible with human ratings. To effectively predict the quality score of distorted images, this study starts from three aspects: the content change of distorted images, the distortion change of distorted images, and the fusion of content features and distorted features. For achieving a universal blind IQE method, a method for IQE on the grounds of content perception and distortion inference was proposed. It is used to predict and evaluate the quality scores of synthesized distorted images and real distorted images. The network structure is shown in Figure 5.

In Figure 5, the network framework contains four parts, the content Feature Extractor (CFE) and Distortion Feature Extractor (DFE) are for extracting content and distortion features in the image. The Fusion Feature Block (FFB)is for adaptively fusing extracted content features and distorted features. The quality prediction block is used for mapping the fused features into predicted quality scores. Firstly, the distorted image is input into the CFE for perceiving features at multiple scales. Secondly, to compensate for the insensitivity

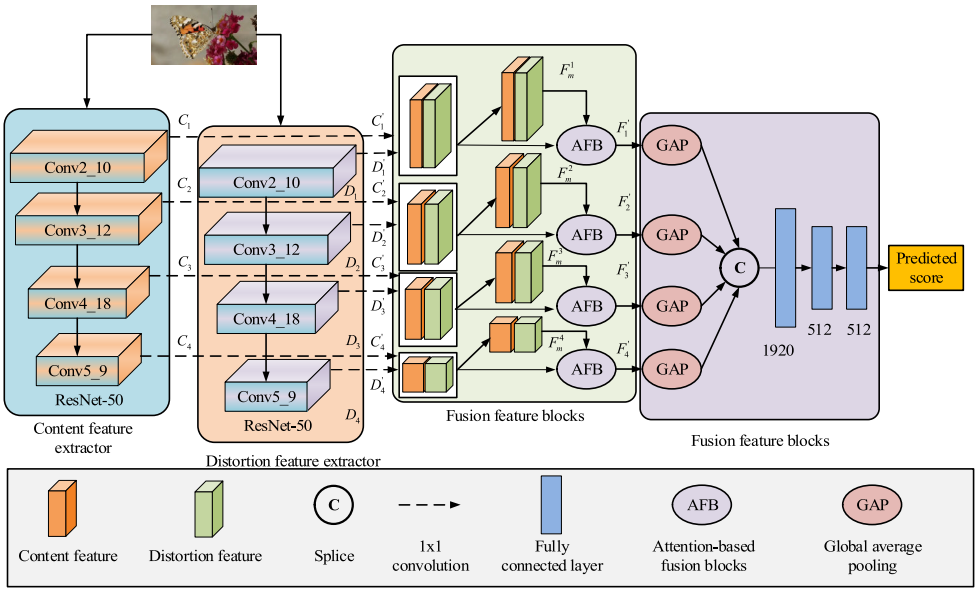


FIGURE 5. Quality evaluation network framework on the grounds of content perception and distortion reasoning.

of the network to distorted information, the distorted image is input into the distorted feature extractor and the distorted perceptual features are extracted at multiple scales. The process of extracting Content Features and Distortion Features (CFDF) is shown in equation (11).

$$\begin{cases} C_i = G_C\left(I_d\right) \\ D_i = G_D\left(I_d\right) \end{cases} \tag{11}$$

In equation (11), C_i, D_i represent multi-scale CFDF, respectively, where $i \in \{1, 2, 3, 4\}$. $G_C\left(\right), G_D\left(\right)$ represent extractors for CFDF, respectively. The function of attention based fusion blocks is for calculating the weights of C_i, D_i [22]. Among them, pooling operations will implement average pooling and maximum pooling operations along the channel [23]. $1 * 1$ CL could markedly extract non sparse global features. Therefore, to effectively compress feature dimensions and retain most relevant features, this study used 8 of $1 * 1$ CL to extract more valuable features [24]. The process is shown in equation (12).

$$\begin{cases} C_i' = f_s\left(C_i\right) \\ D_i' = f_s\left(D_d\right) \end{cases} \tag{12}$$

In equation (12), $f_s\left(\right)$ represents a $1 * 1$ convolution operation. Due to the differences between the CFDF after convolution processing, directly fusing these features will diminish the learning capability. Consequently, the study proposes for adaptively fusingvarious features through feature fusion, as shown in equation (13).

$$\begin{cases} F_m^i = C_i' \otimes D_i' \\ F_i' = G_F\left(C_i', D_i', F_m'\right) \end{cases} \tag{13}$$

In equation (13), F_m^i, F_i' respectively represent the concatenated features and the output features of the fused feature blocks. $G_F\left(\right)$ represents fused feature blocks. Finally, the quality prediction block will regress the fused features into quality scores. The expression for calculating the predicted mass fraction is shown in equation (14).

$$y_d = G_R\left(F_1', F_2', F_3', F_4'\right) \tag{14}$$

In equation (14), $G_R\left(\right)$ represents the quality prediction block. Although content extractors can obtain content information from images, they lack the ability to extract synthetic distortion information. Therefore, the study proposes the use of distorted feature extractors to solve the problem of image distortion information diversity [25]. To enhance its sensitivity to distorted information, it was studied to pre-train it on a large synthetic distorted dataset LSQA to learn rich distorted features [26]. The pre-training process is shown in Figure 6.

In Figure 6, the DFE consists of three parts, namely ResNet-50 backbone network module, quality perception module, and Quality Regression Module (QRM). The backbone network module extracts hierarchical features from distorted images, ranging from low-level to high-level. The process of extracting multi-level features from ResNet-50 is shown in equation (15) [27].

$$f_i = F_D\left(I_d\right) \tag{15}$$

In equation (15), $F_D\left(\right)$ serves as the ResNet-50 backbone network module. The quality perception module will extract corresponding feature information from multi-level features. For obtaining the predicted quality score, the QRM will regress the fused features to the quality score. The process

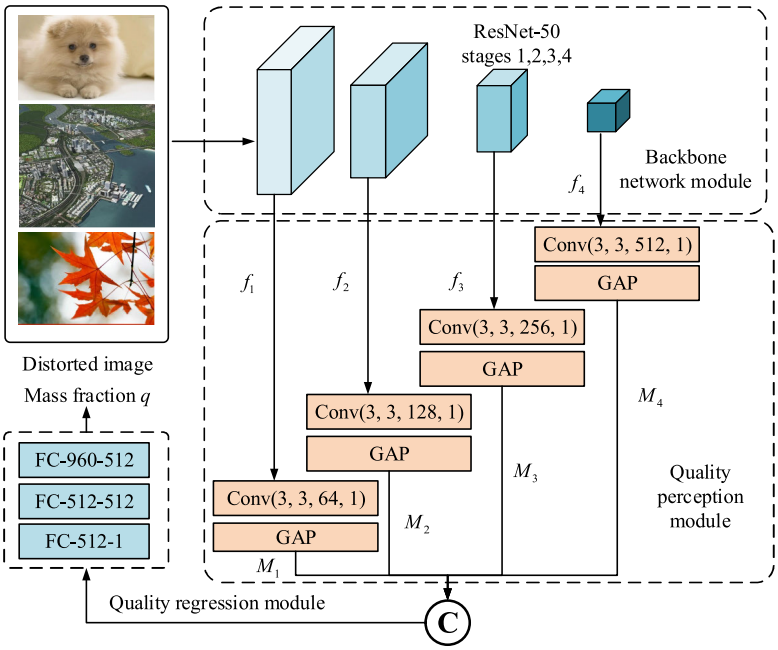


FIGURE 6. Pre-training of DFE.

of quality prediction is shown in equation (16).

$$\begin{cases} M_f = M_1 \otimes M_2 \otimes M_3 \otimes M_4 \\ q = F_R\left(M_f\right) \end{cases} \tag{16}$$

In equation (16), M_f, q represent the quality scores of fused features and predictions, respectively. $F_R()$ represents the QRM. For learning the relation in distortion features and quality scores, research is conducted to enhance prediction performance by minimizing the loss function. The calculation is showcased in equation (17).

$$l_d = \frac{1}{N}\left\|G_D\left(I_D;\theta\right) - y_g\right\| \tag{17}$$

In equation (17), y_g represents the label value of the distorted image. After completing the pre-training of the DFE, only the backbone network module will be retained for extracting distortion features. The CFE could markedly derive content perception features, while the DFE could markedly derive rich distortion features. Directly integrating such content features with various characteristics and distortion features would decrease the learning capability. To address this issue, an adaptive fusion block is proposed, which adaptively fuses the extracted content features with distorted features through weight values. It contains four attention based fusion blocks. This fusion block can adaptively fuse and extract content and distorted features on the grounds of the contributions of different features. And it can make the network concentrate on important features. The network structure is shown in Figure 7.

In Figure 7, the input of the attention based fusion block has three types of features, namely content features, distortion features, and fusion features. The purpose is to calculate the weight score of fused features through multi-scale spatial attention blocks, as shown in equation (18).

$$\begin{cases} F_m^i = C_i' \otimes D_i' \\ W_1^d = \psi_1\left(F_m^i\right) \\ W_2^c = \psi_2\left(F_i'\right) \\ F_i' = \left(W_1^d \times D_i'\right) \otimes \left(W_2^c \times C_i'\right) \end{cases} \tag{18}$$

In equation (18), W_1^d, W_2^c represent the weight scores of CFDF, respectively. F_m^i, F_i' represent the concatenated features and the adaptive fused features, respectively. $\psi_1()$, $\psi_2()$ represent two independent multi-scale spatial AMs, respectively. To solve the problem of the lack of CL for extracting features of different sizes in a single receptive field, this study proposes the use of multi-scale spatial attention blocks with various receptive fields. This is for calculating the weight scores of CFDF. Firstly, the maximum pooling layer and the average pooling layer are utilized for compressing the feature map with a size of $C \times H \times W$ into a spatial feature map with a size of $2 \times H \times W$. The features after splicing are shown in equation (19).

$$f_i' = MaxPool\left(f_i\right) \otimes AvgPool\left(f_i\right) \tag{19}$$

In equation (19), $AvgPool()$ represents the average pooling layer. Then, a set of CL with different sizes of receptive fields is used for extracting multi-scale features from f_i'. For reducing computational complexity and achieving the same receptive field, this study used two $3 * 3$ CL instead of a $5 * 5$ CL, and three $3 * 3$ CL instead of a $7 * 7$ CL. Secondly, the output features of various receptive fields are fused through splicing operations. It then performs channel compression on

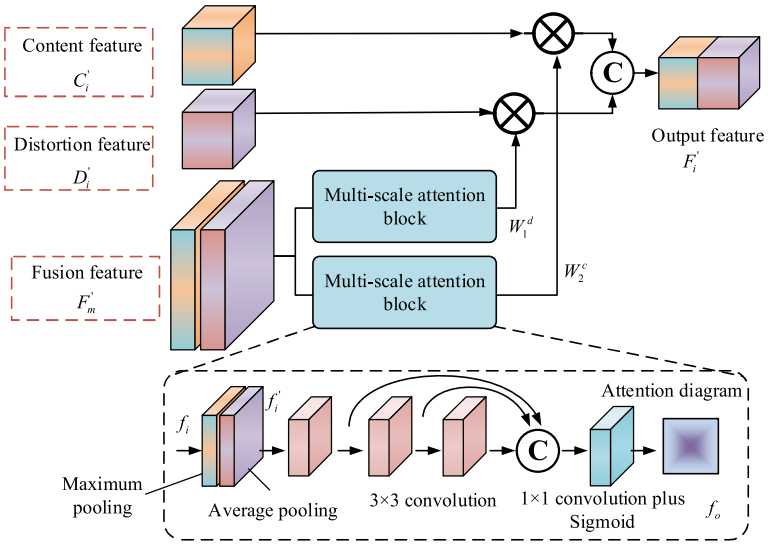


FIGURE 7. Attention-based Fusion block structure diagram.

it through a $1 * 1$ convolution operation. Finally, the Sigmoid function is utilizedfor generating a 2D attention map [28]. The multi-scale calculation process is shown in equation (20).

$$\begin{cases} f_{3\times3} = \alpha_1\left(f_i'\right) \\ f_{5\times5} = \alpha_1\left(f_{3\times3}\right) \\ f_{7\times7} = \alpha_1\left(f_{5\times5}\right) \\ f_o = Sig\left(\alpha_2\left(f_{3\times3} \otimes f_{5\times5} \otimes f_{7\times7}\right)\right) \end{cases} \tag{20}$$

In equation (20), $f_{3\times3}, f_{5\times5}, f_{7\times7}$ represent the features after $3 * 3$ CL. f_o represents the generated attention map. α_1 () , α_2 () represent convolution operations of $3 * 3$ and $1 * 1$, respectively. After getting multi-scale CFDF adjusted by adaptive fusion blocks, the mapping relation in extracted features and quality scores is learned through quality prediction blocks. It performs quality regression through fully connected layers. It assumes that the fused feature map is $\left\{F_1', F_2', F_3', F_4'\right\}$, and first applies the GAPL to such feature maps for getting the channel feature map $\{l_1, l_2, l_3, l_4\}$. The process is shown in equation (21).

$$\begin{cases} l_i = GAP\left(F_i'\right) \\ l_{concate} = l_1 \otimes l_2 \otimes l_3 \otimes l_4 \end{cases} \tag{21}$$

In equation (21), l_i represents the feature map processed by the GAPL. $l_{concate}$ represents the feature map after stitching.Finally, the subsequent feature map is mapped to the quality score by a fully connected layer. The study used three fully connected layers with 1920, 512, and 512 nodes, respectively.For growing the nonlinear relation of the network layer, the first two fully connected layers in the network structure are followed by ReLU operations. In addition, the study introduces the dropout operation into the proposed network. However, excessive dropouts will grow the convergence time of the network, so this operation is only utilized after the first fully connected layer [29], [30]. To compare the predictive accuracy of IQE methods from an objective perspective, the commonly used evaluation indicators currently include Pearson Linear Correlation Coefficient (PLCC) and Spearman Order Correlation Coefficient (SROCC) [31], [32]. PLCC is utilized for evaluating the linear correlation in image label values and predicted values, and its calculation expression is shown in equation (22) [33].

$$PLCC = \frac{\sum_{i=1}^{N}\left(y_l^i - \overline{y_l}\right)\left(y_p^i - \overline{y_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(y_i^i - \overline{y_l}\right)^2 \sum_{i=1}^{N}\left(y_p^i - \overline{y_p}\right)^2}} \tag{22}$$

In equation (22), N represents the number of test images. y_l^i, y_p^i represent the label value and predicted value of the i-th image. $\overline{y_l}, \overline{y_p}$ serve as the average of the label values and predicted values of the image, respectively. SROCC is utilized for evaluating the monotonic relation in image label values and predicted values, and its calculation expression is shown in equation (23) [34].

$$SROCC = 1 - \frac{6\sum_{i=1}^{N}\left(rank\left(y_l^i\right) - rank\left(y_p^i\right)\right)^2}{N\left(N^2 - 1\right)} \tag{23}$$

In equation (23), $rank\left(y_l^i\right)$, $rank\left(y_p^i\right)$ are the positions of the i-th image arranged in order of size, respectively. The absolute value range of these two types of evaluation indices is 0-1, and the closer their values are to 1, the more excellent the prediction accuracy.

IV. RESULT AND DISCUSSION

In the experiment, the experimental setup and its experimental results on the standard quality evaluation dataset will be introduced. The experimental setup includes experimental environment construction, network training, dataset

selection, and evaluation indicators. The experimental process includes single dataset performance evaluation, single distortion type performance evaluation, cross dataset performance evaluation, and ablation experiments.The proposed algorithm is implemented using TensorFlow framework.All experimentsare completed under Windows 11 and 64 bit operating systems. It uses Adam optimization algorithm and back-propagation algorithm to train parameters in the network.

A. EXPERIMENTAL ANALYSIS OF THE EFFECTIVENESS OF IMAGE QUALITY DEGRADATION PROCESS MODEL

To verify the effectiveness of the image quality degradation process model on the grounds of cyclic GANs studied, six representative IQE methods were selected for comparison in this experiment. These six quality evaluation methods are HOSA, CNN, TS-CNN, DeepIQA, CaHDC, and RAN4IQA, respectively. For evaluating the predictive accuracy, 7 public IQE datasets were selected for comparison. 80% of the reference images and corresponding distorted images were randomly selected as the training set, while the rest were selected as the test set.

Among them, the distorted images are obtained by adding distortion information to the reference images in the corresponding data set. The distorted image type is the format of the distorted image, such as JPEG, JP2K, etc.The number of image distortion types depends on the respective data set.Table 1 showcases the specific information.

To eliminate sample bias caused by the dataset, it is divided into 10 repetitions. The SROCC and PLCC results obtained by each method on five datasets are shown in Figure 8.

Figure 8 showcases that the SROCC value and PLCC value of the proposed method in the 5 data sets of LIVE, CSIQ, TID2013, LIVEMD and LIVEC are higher, both exceeding 0.8. The SROCC outcomesgotten by the proposed method on the five datasets of LIVE, CSIQ, TID2013, LIVEMD, and LIVEC are 0.972, 0.928, 0.835, 0.938, and 0.886, respectively. The PLCC results obtained on five datasets are 0.973, 0.930, 0.841, 0.952, and 0.806, respectively. For verifying the predictive accuracy in this section under different distortion types, three synthesized distortion data LIVE and CSIQ are selected for single distortion type testing in this experiment. The results are shown in Figure 9.

Figure 9 (a) demonstrates thatthe proposed method has a slightly lower prediction accuracy for FF image distortion types on the LIVE dataset than HOSA. The accuracy of the proposed method for JPEG, JP2K, WN and GB image distortion is higher than that of CNN, DeepIQA, TS-CNN and RAN4IQA. The predictive accuracy of the proposed method for JPEG, JP2K, WN and GB image distortion types are 0.970, 0.961, 0.982 and 0.970, respectively.Figure 9 (b) demonstrates that on the CSIQ data set, the detection SROCC results of the proposed method for JPEG, JP2K, WN, GB and PN of five distortion types reach 0.925, 0.938, 0.947, 0.928 and 0.931, respectively. The results of this method are better than other methods for the detection of these five image types. To verify the generalization of the proposed method, CSIQ, LIVE, TID2013, and LIVEC are selected for cross dataset training testing in this experiment.Each data set is divided into a training set and a test set. The training set consists of 80% of the reference images and their corresponding distorted images, randomly selected. The test set consists of the remaining 20% of the reference images and their corresponding distorted images. All models are trained on a single data set and then tested on the remaining three data sets.The results are shown in Figure 10.

TABLE 1. Details of the data set.

Data set	Tag value	Distortion type	Reference image quantity	Distorted image quantity
LIVE	[0，100]	5	29	779
LIVEC	[0，100]	/	/	1162
CSIQ	[0，1]	6	30	866
TID2013	[1，9]	24	25	3000
KonIQ-10k	[1，100]	/	/	10073
SPAQ	[0，100]	/	/	11125
LIVE MD	[0，100]	2	15	450

Note: "/" indicates no data.

Figure 10 shows that the proposed method achieves the highest prediction accuracy in 12 cross tests of 4×3 across 4 datasets. In Figure 10(a), the proposed method has higher SROCC values on the three datasets than other methods. Similarly, in Figure 10(b), (c), and (d), the proposed method continues to have higher SROCC values than other methods on the same dataset. When training and testing on synthetic distorted datasets, most methods have achieved good generalization. As trained on LIVE and tested on CSIQ and TID2013, the proposed methods exceeded the sub-optimal methods by 15.21% and 11.33%, respectively. As trained on CSIQ and tested on LIVE and TID2013, the proposed method exceeded the sub-optimal method by approximately 0.90% and 1.23%, respectively. As trained on TID2013 and tested on LIVE and CSIQ, the method in this section exceeded the sub-optimal method by approximately 0.73% and 3.54%, respectively. As trained on synthetic distorted datasets and tested on real distorted datasets, the generalization performance of most methods significantly decreases. However, the proposed method still exhibits the best generalization performance among many methods.The base-line model (BL) uses a lossless restoration module and VGG-16 as a quality regression network for image restoration and quality assessment. BL models are pre-trained with adversarial losses only. The BL+CCL model is formed by adding the loss recovery module and pre-training the adversarial loss and cyclic consistency loss. Finally, VGG-16 is replaced by a new mass regression network to get the BL+CCL+QRN model.The results of ablation experiments conducted on 5 datasets are shown in Figure 11.

Figure 11(a) shows that the extraction of multi-level features using the BL+CCL+QRN model used in the study can improve the SROCC by approximately 1.12%, 1.01%, 1.35%, and 2.59%. This verifies that the extraction of layered

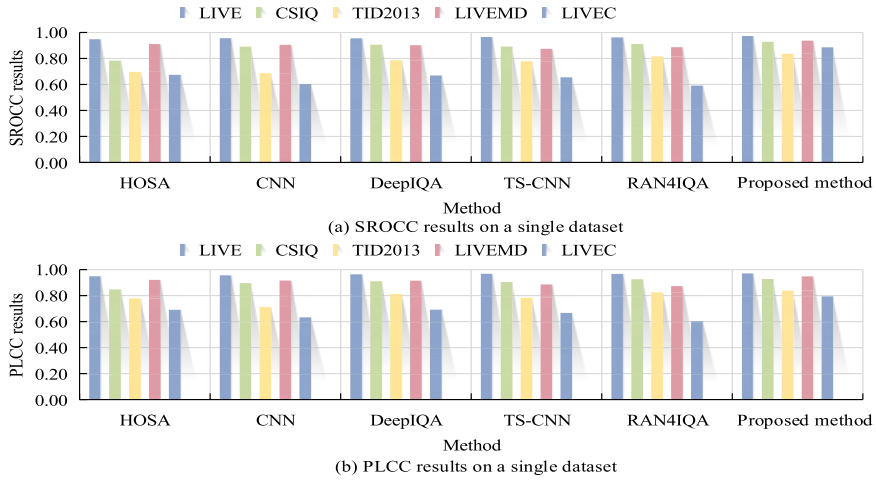


FIGURE 8. SROCC versus PLCC results on a single dataset.

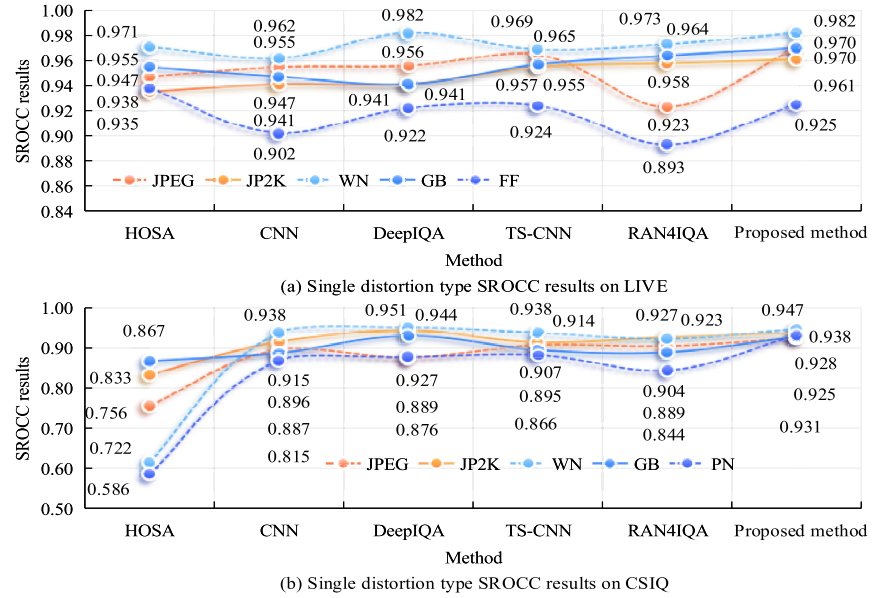


FIGURE 9. SROCC results for a single distortion type on LIVE and CSIQ.

features is superior to using only output predicted images as scoring criteria.Figure 11(b) demonstrates that after combining EF and DEF, SROCC scores on LIVE, CSIQ, TID2013 and LIVEC data sets reached 0.969, 0.927, 0.831 and 0.849, respectively, showing an overall improvement. After the addition of the AM, the SROCC on the four data sets were 0.973, 0.938, 0.846 and 0.879, respectively, indicating that the AM can effectively extract key features and improve the performance by 0.22% $\sim$ 3.66%.

B. EXPERIMENTAL ANALYSIS OF THE EFFECTIVENESS OF AN IMPROVED MODEL FOR BLIND IQE

For verifying the improved model for blind IQE, the input distorted image was randomly cropped into 5 image blocks with a size of 224 $*$ 224 pixels. It sets the learning rate to 10^-5 and multiplies it by a coefficient of 0.9 for every fifty training sessions. The DFE undergoes a total of 500 iterative training sessions until the loss function converges and the training is completed. To verify the predictive accuracy of different IQE methods on a single dataset, six common IQE datasets were chosen for validation in the experiment. Figure 12 showcases the SROCC and PLCC results on a single dataset.

Figure 12(a) demonstrates thatexcept that the SROCC value of the proposed method on the dataset CSIQ is slightly lower than that of DB-CNN, while the SROCC value of the proposed method on the other 5 datasets is higher than that of other methods. The SROCC values of the proposed method on the LIVE, TID2013, LIVEC, KonIQ-10k and SPAQ data sets are 0.971, 0.835, 0.862, 0.912 and 0.919,

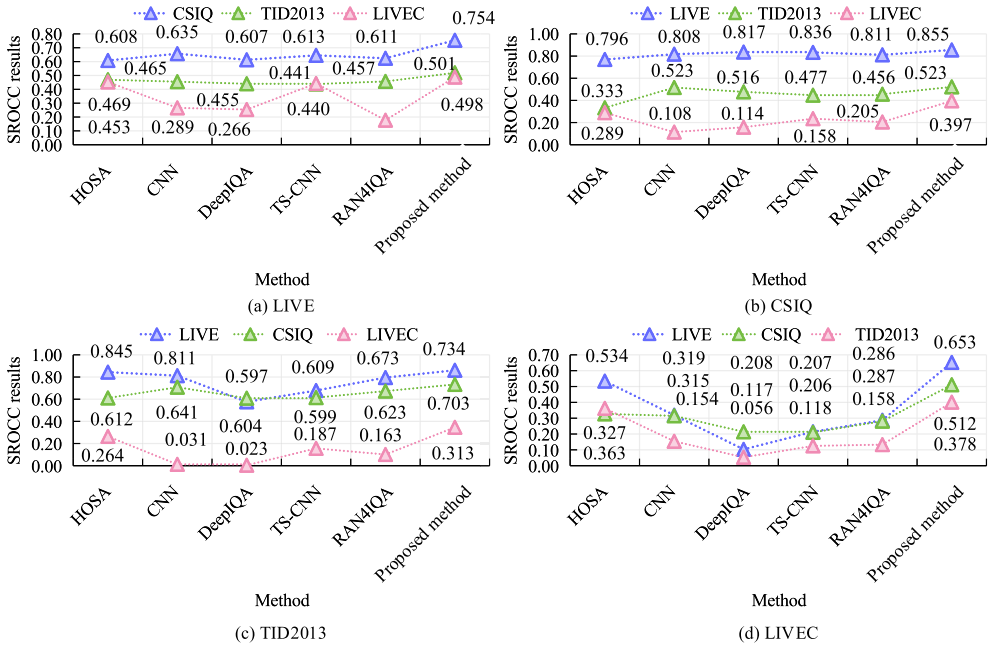


FIGURE 10. Cross-data set SROCC results.

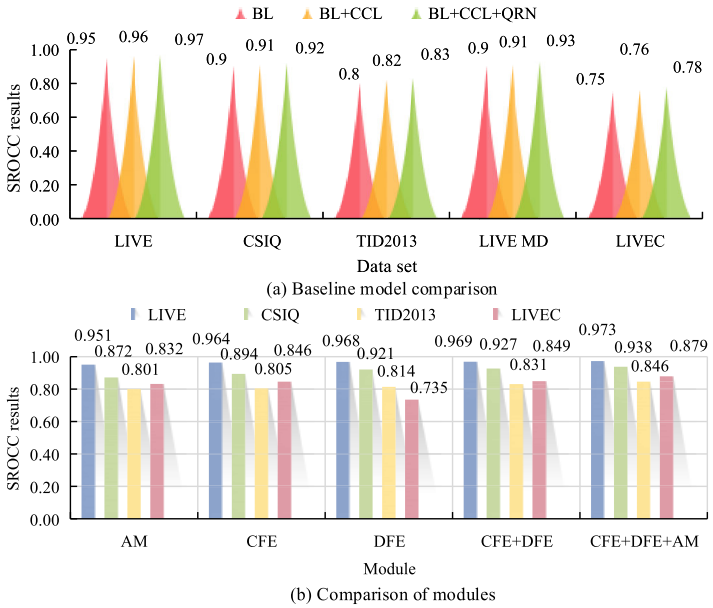
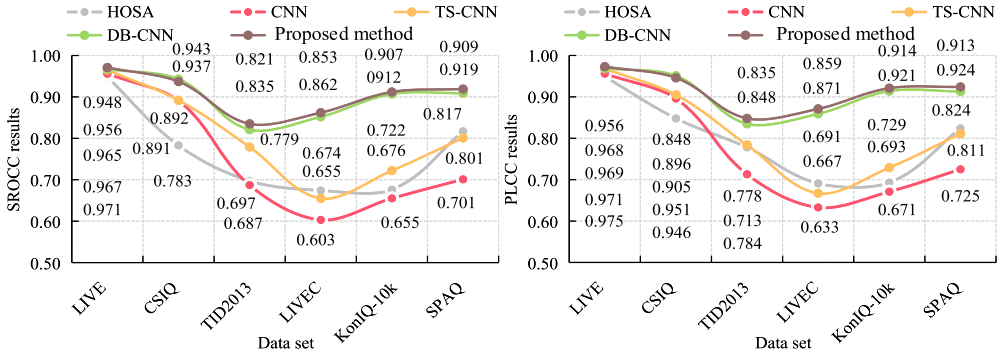


FIGURE 11. Results of ablation experiments on LIVE, CSIQ, TID2013, LIVE MD and LIVEC.

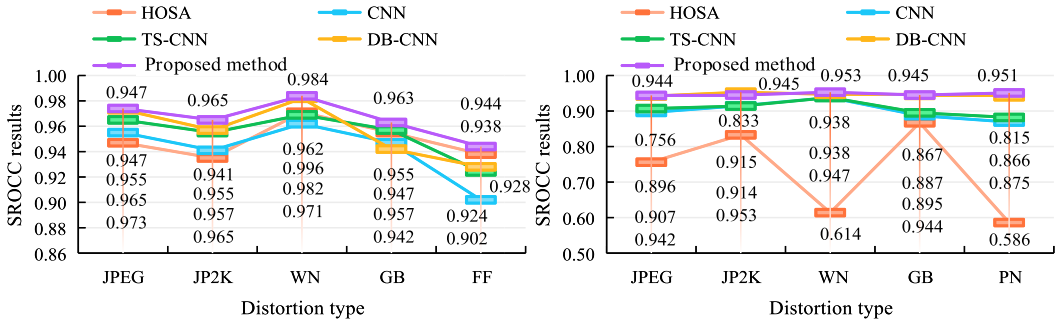
respectively. Figure 12(b) shows that the proposed method has a slightly lower PLCC value on dataset CSIQ compared to DB-CNN, but higher PLCC values on all other datasets compared to other methods. The PLCC values of the proposed method on LIVE, TID2013, LIVEC, KonIQ-10k and SPAQ are 0.975, 0.848, 0.871, 0.921 and 0.924, respectively.In summary, the prediction accuracy of proposed methods on all datasets is far superior to similar methods. For verifying the predictive accuracy for various distortion types, experiments were implemented ontwosynthetic distortion datasets LIVE and CSIQ in this section. The results are shown in Figure 13.

Figure 13 (a) shows that in the LIVE dataset, the proposed method has higher prediction accuracy for all distortion types than other methods. For JPEG, JP2K, WN, GB and FF, the SROCC results are 0.947, 0.965, 0.984, 0.963 and 0.944, respectively. Figure 13 (b) shows that the performance of



(a) SROCC results on a single dataset (b) PLCC results on a single dataset

FIGURE 12. SROCC and PLCC results on a single dataset.



(a) Single distortion type SROCC results on LIVE (b) Single distortion type SROCC results on CSIQ

FIGURE 13. SROCC results for a single distortion type on LIVE and CSIQ.

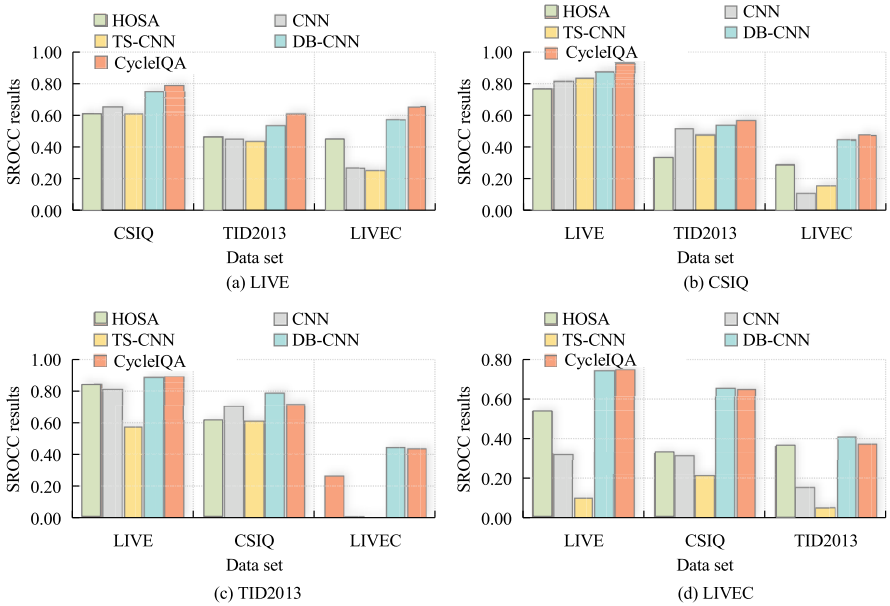


FIGURE 14. Cross-data set SROCC results.

the proposed method is about 0.19%, 0.62%, 0.14% and 1.20% higher than DB-CNN in JPEG, WN, GB and PN, respectively. And the prediction accuracy is higher than that of other methods for JPEG, JP2K, WN, GB, FF five kinds of distorted image types.In summary, the DFE proposed by the research institute can learn more abundant distortion

TABLE 2. Comparison of image quality evaluation methods.

Methods	Accuracy/%	RMSE	Prediction F1
References [6]	78.5%	0.415	82.6
References [7]	84.3%	0.401	86.7
References [8]	86.7%	0.387	87.1
References [9]	88.9%	0.363	89.6
References [10]	91.7%	0.341	87.1
Proposed method	93.2%	0.314	92.4

information, strengthening the network's capability of deriving various types of distortion. To verify the generalization performance of the proposed method, cross-experiments are carried out between the research model and other models on four datasets: LIVE, CSIQ, TID2013 and LIVEC. That is, all models are trained on one dataset and then tested for SROCC evaluation on the other three datasets. The average SROCC results are analyzed statistically. The SROCC results as shown in Figure 14.

Figure 14(a) demonstrates that in the data set LIVE, the SROCC results of the proposed method in the three cross-data sets are all higher than other algorithms, and their values on CSIQ, TID2013 and LIVEC are 0.788, 0.611 and 0.657, respectively.Figure 14(b) demonstrates that on the data set CSIQ, the SROCC results of the proposed method in the three cross-data sets are all higher than other algorithms, and their values on LIVE, TID2013 and LIVEC are 0.933, 0.569 and 0.473, respectively. Figure14(b) and (d) demonstrate that the SROCC results of the proposed method on several other cross-data sets are still higher than those of other algorithms.The proposed method achieved a total of 8 good prediction results. As trained on synthetic distorted datasets and tested on real distorted datasets, the proposed method can easily exceed 0.45. It achieved an SROCC result of 0.661 when trained on LIVE, 0.474 when trained on CSIQ, and 0.442 when trained on TID2013. When training and testing on synthetic distorted datasets, the proposed method largely outperforms other methods. The proposed method is compared with the image quality evaluation methods proposed in literatures [6], [7], [8], [9], [10]. The tests were performed on the KonIQ-10K dataset. The accuracy of image quality evaluation, root mean square error and F1 score prediction results of each method are shown in Table 2.

Table 2 shows that in terms of image quality evaluation accuracy, the proposed method has a higher evaluation accuracy than other advanced methods, and its accuracy reaches 93.2%. In the root-mean-square error of image quality evaluation, the RMSE of the proposed method is smaller than that of other methods, which is 0.314. In the prediction of F1 score, the F1 value of the proposed method is higher than that of other methods, which is 92.4. In summary, compared with other advanced methods, the proposed method has a significant improvement in the accuracy of image quality evaluation.

V. CONCLUSION

The blind IQE method is hard for achieving the performance equivalent to the full reference IQE method, and it is hard for accurately predicting the image quality score through a single model. A quality perception method on the grounds of GANs was proposed to model the process of image quality degradation. And a special IQE method was used for extracting content information and distortion information from the image. The experiment showed that on the five single datasets of LIVE, CSIQ, TID2013, LIVEMD, and LIVEC, the SROCC results and PLCC prediction accuracy results obtained by the proposed method were higher than those of the other five methods. The prediction accuracy of the proposed method for distortion types in JPEG, JP2K, WN, GB, and FF was 0.971, 0.963, 0.984, 0.971, and 0.926, respectively. The SROCC results and PLCC prediction accuracy results obtained by the proposed methods on six datasets: LIVE, CSIQ, TID2013, LIVEC, KonIQ-10k, and SPAQ were higher than those of the other four methods. The proposed method achieved the highest predictive accuracy on all distortion types in the LIVE dataset, with SROCC results of 0.973, 0.965, 0.984, 0.963, and 0.944, respectively. As trained on synthetic distorted datasets and tested on real distorted datasets, the SROCC results of the proposed method can easily surpass those of other methods. In summary, the improved model for blind IQE studied achieved optimal generalization performance. Although research has achieved good results, its main extraction ability relied on the pre-trained ResNet-50 network, making it difficult to reduce most of the overall network complexity. Therefore, in the future, it is necessary to consider using lighter networks as CFEs and DFEs. It ensures accurate prediction while also reducing computational complexity.

REFERENCES

[1] I. A. Tsalafoutas, V. Tsapaki, and I. Triantopoulou, "Evaluation of image quality and patient exposure in fluoroscopy using a phantom: Is there any clinical relevance?" *Eur. J. Radiol.*, vol. 138, May 2021, Art. no. 109607, doi: 10.1016/j.ejrad.2021.109607.

[2] D. Lee, S. W. Jeong, S. J. Kim, H. Cho, W. Park, and Y. Han, "Improvement of megavoltage computed tomography image quality for adaptive helical tomotherapy using cycleGAN-based image synthesis with small datasets," *Med. Phys.*, vol. 48, no. 10, pp. 5593–5610, Oct. 2021, doi: 10.1002/mp.15182.

[3] Y. Liu, H. Al-Shehari, and H. Zhang, "Attention mechanism enhancement algorithm based on cycle consistent generative adversarial networks for single image dehazing," *J. Vis. Commun. Image Represent.*, vol. 83, Feb. 2022, Art. no. 103434, doi: 10.1016/j.jvcir.2021.103434.

[4] Z. Chen, S. Zhang, J. Zhang, Z. Hu, X. Han, and M. Xu, "A novel artificial intelligence model for color image quality assessment for security enhanement weighted by visual saliency," *J. Intell. Fuzzy Syst.*, vol. 40, no. 4, pp. 8091–8100, Apr. 2021, doi: 10.3233/jifs-189632.

[5] S. Choudhuri, S. Adeniye, and A. Sen, "Distribution alignment using complement entropy objective and adaptive consensus-based label refinement for partial domain adaptation," *Artif. Intell. Appl.*, vol. 1, no. 1, pp. 43–51, Jan. 2023, doi: 10.47852/bonviewaia2202524.

[6] D. Ma, H. Wen, X. Li, T. Xie, and X. Li, "A sub-regional compression method for greenhouse images based on CNN image quality assessment," *J. Food Process. Preservation*, vol. 46, no. 11, pp. 1–6, Nov. 2022, doi: 10.1111/jfpp.16992.

[7] X. Yang, F. Li, and H. Liu, "TTL-IQA: Transitive transfer learning based no-reference image quality assessment," *IEEE Trans. Multimedia*, vol. 23, pp. 4326–4340, Nov. 2021, doi: 10.1109/TMM.2020.3040529.

[8] 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, doi: 10.1109/TIP.2020.3002478.

[9] F. Kefeng, L. Jiyun, L. Fei, and Q. Puye, "CNN based no-reference HDR image quality assessment," *Chin. J. Electron.*, vol. 30, no. 2, pp. 282–288, Mar. 2021, doi: 10.1049/cje.2021.01.008.

[10] Y. Lu, W. Li, X. Ning, X. Dong, L. Zhang, L. Sun, and C. Cheng, "Blind image quality assessment based on the multiscale and dual-domains features fusion," *Concurrency Comput., Pract. Exper.*, vol. 35, no. 18, pp. 4326–4340, Aug. 2023, doi: 10.1002/cpe.6177.

[11] Y. Tang, C. Liu, and X. Zhang, "Single image super-resolution using Wasserstein generative adversarial network with gradient penalty," *Pattern Recognit. Lett.*, vol. 163, pp. 32–39, Nov. 2022, doi: 10.1016/j.patrec.2022.09.012.

[12] T. Jin, X. W. Ye, and Z. X. Li, "Establishment and evaluation of conditional GAN-based image dataset for semantic segmentation of structural cracks," *Eng. Struct.*, vol. 285, Jun. 2023, Art. no. 116058, doi: 10.1016/j.engstruct.2023.116058.

[13] L. Zhang and L. Zhao, "High-quality face image generation using particle swarm optimization-based generative adversarial networks," *Future Gener. Comput. Syst.*, vol. 122, pp. 98–104, Sep. 2021, doi: 10.1016/j.future.2021.03.022.

[14] L. Goetschalckx, A. Andonian, and J. Wagemans, "Generative adversarial networks unlock new methods for cognitive science," *Trends Cogn. Sci.*, vol. 25, no. 9, pp. 788–801, Aug. 2021, doi: 10.1016/j.tics.2021.06.006.

[15] W. Qian, Y. Xu, W. Zuo, and H. Li, "Self-sparse generative adversarial networks," *CAAI Artif. Intell. Res.*, vol. 1, no. 1, pp. 68–78, Sep. 2022, doi: 10.26599/air.2022.9150005.

[16] Z. Li, Z. Dong, W. Chen, and Z. Ding, "On the game-theoretic analysis of distributed generative adversarial networks," *Int. J. Intell. Syst.*, vol. 37, no. 1, pp. 516–534, Jan. 2022, doi: 10.1002/int.22637.

[17] Z. Yao, Y. Wang, F. Yang, Y. Cheng, and B. Ran, "Map matching for travel route identification based on Earth Mover's distance algorithm using wireless cell trajectory data," *J. Intell. Transp. Syst.*, vol. 25, no. 6, pp. 644–656, Nov. 2021, doi: 10.1080/15472450.2021.1955209.

[18] B. S. N. V. Chaitanya and S. Mukherjee, "Single image dehazing using improved cycleGAN," *J. Vis. Commun. Image Represent.*, vol. 74, Jan. 2021, Art. no. 103014, doi: 10.1016/j.jvcir.2020.103014.

[19] K.-J. Chen, M. Wu, Y. Zhang, and Z. Chen, "SR-AFU: Super-resolution network using adaptive frequency component upsampling and multi-resolution features," *Frontiers Comput. Sci.*, vol. 17, no. 1, pp. 123–132, Aug. 2022, doi: 10.1007/s11704-021-0562-y.

[20] A. V. Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of pneumonia from chest X-ray images," *Global Transitions Proc.*, vol. 2, no. 2, pp. 375–381, Nov. 2021, doi: 10.1016/j.gltp.2021.08.027.

[21] X. Chen, W. Zhang, Y. Hou, and L. Yang, "Improved stereo matching algorithm based on multi-scale fusion," *J. Northwest. Polytechnical Univ.*, vol. 39, no. 4, pp. 876–882, Sep. 2021, doi: 10.1051/jnwpu/20213940876.

[22] S. Zhang, N. Zheng, and D.-L. Wang, "A novel attention-based global and local information fusion neural network for group recommendation," *Mach. Intell. Res.*, vol. 19, no. 4, pp. 331–346, Aug. 2022, doi: 10.1007/s11633-022-1336-1.

[23] Y. Zeng, R. Liu, and X. Liu, "A novel approach to tool condition monitoring based on multi-sensor data fusion imaging and an attention mechanism," *Meas. Sci. Technol.*, vol. 32, no. 5, Mar. 2021, Art. no. 055601, doi: 10.1088/1361-6501/abea3f.

[24] L. Liu and S.-B. Tsai, "Intelligent recognition and teaching of English fuzzy texts based on fuzzy computing and big data," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–10, Jul. 2021, doi: 10.1155/2021/1170622.

[25] X. Yuan, E. Pang, K. Lin, and J. Hu, "Deep protein subcellular localization predictor enhanced with transfer learning of GO annotation," *IEEJ Trans. Electr. Electron. Eng.*, vol. 16, no. 4, pp. 559–567, Apr. 2021, doi: 10.1002/tee.23330.

[26] A. Kamli, R. Saouli, H. Batatia, M. B. Naceur, and I. Youkana, "Synthetic medical image generator for data augmentation and anonymisation based on generative adversarial network for glioblastoma tumors growth prediction," *IET Image Process.*, vol. 14, no. 16, pp. 4248–4257, Dec. 2020, doi: 10.1049/iet-ipr.2020.1141.

[27] D. McNeely-White, J. R. Beveridge, and B. A. Draper, "Inception and ResNet features are (almost) equivalent," *Cogn. Syst. Res.*, vol. 59, pp. 312–318, Jan. 2020, doi: 10.1016/j.cogsys.2019.10.004.

[28] S. Singh and S. Harsha, "Analysis of porosity effect on free vibration and buckling responses for sandwich sigmoid function based functionally graded material plate resting on Pasternak foundation using Galerkin Vlasov's method," *J. Sandwich Struct. Mater.*, vol. 23, no. 5, pp. 1717–1760, Jun. 2021, doi: 10.1177/1099636220904340.

[29] Z. Hu and J. Zhou, "Convergence of city relational network, production sector structure, and regional development," *Discrete Dyn. Nature Soc.*, vol. 2021, pp. 1–13, Aug. 2021, doi: 10.1155/2021/8913858.

[30] G. Plonka, Y. Riebe, and Y. Kolomoitsev, "Spline representation and redundancies of one-dimensional ReLU neural network models," *Anal. Appl.*, vol. 21, no. 1, pp. 127–163, Jan. 2023, doi: 10.1142/s0219530522400103.

[31] M. R. Faraji, "Full-reference tone-mapped images quality assessment," *IET Image Process.*, vol. 15, no. 3, pp. 763–773, Feb. 2021, doi: 10.1049/ipr2.12060.

[32] S. Kiruthika and V. Masilamani, "Goal oriented image quality assessment," *IET Image Process.*, vol. 16, no. 4, pp. 1054–1066, Mar. 2022, doi: 10.1049/ipr2.12209.

[33] T. Hao, "Universal correlation of the superconducting transition temperature with the linear-in-T coefficient, electron packing parameter, and the numbers of valence and conduction electrons," *Phys. Chem. Chem. Phys.*, vol. 25, no. 17, pp. 12443–12449, May 2023, doi: 10.1039/d3cp00706e.

[34] Y. Dagtekin, S. Kaya, and N. Besli, "Distributed energy system selection for a commercial building by using multi criteria decision making methods," *Int. J. Hydrogen Energy*, vol. 47, no. 86, pp. 36672–36692, Oct. 2022, doi: 10.1016/j.ijhydene.2022.08.208.

XIAOYING LI was born in Fenyang, Shanxi, in July 1981. She received the bachelor's degree in computer science and technology from Xinzhou Normal University, in 2004, and the master's degree in computer software and theory from Jiangxi Normal University, in 2007.

From 2007 to 2021, she was a full-time Teacher with Guilin University of Technology at Nanning. From 2021 to 2023, she was the Deputy Director of the Teaching and Research Office, Guilin University of Technology at Nanning. In recent years, she has led and participated in three municipal-level scientific research projects and four provincial and ministerial-level educational reform projects. She has published 13 academic articles, including two indexed by EI. She has been granted one utility model patent and one software copyright. Her research interests include machine learning, image processing, and computer vision.

SHOUWU HE was born in Hequ, Shanxi, China, in July 1979. He received the bachelor's degree in computer science and technology from Xinzhou Normal University, in 2004, and the master's degree in computer application technology from Jiangxi Normal University, in 2007.

From 2007 to 2016, he was a full-time Teacher with the College of Computer Application, Guilin University of Technology at Nanning. Since 2017, he has been the Deputy Director of the College of Computer Application, Guilin University of Technology at Nanning. He led and completed one research project with the department level and participated in two research projects at or above the department level in other cities. He has published five scientific research articles, including two indexed by EI, obtained one utility model patent, and one software copyright. His research interests include big data processing and machine learning.