

An aerial image data cleaning of power lines with multi-information fusion perception

Yanjun Fan^a, Qian Yang^b, Yangming Guo^{*a}, Lujuan Jiang^c, Jiang Long^a, Zhuqing Wang^a, Guo Li^a

^aSchool of Computer Science, Northwestern Polytechnical University, Xi'an, Shaanxi, China;

^bWenzhou Science and Technology Branch of Zhejiang Tusheng Power Transmission and Transformation Engineering Co., Ltd., Wenzhou, Zhejiang, China; ^cSchool of Computer Science and Engineering, Northwest Normal University, Lanzhou, Gansu, China

ABSTRACT

Recently, unmanned aerial vehicles (UAVs) have been used to conduct the task of power line inspection. Unfortunately, most of the images acquired by UAVs are invalid for visual inspection. The main reason is that the collected images are low quality and repetitive. To obtain a set of valid images with high quality, a novel multi-information fusion perception (MIFP) model is proposed to automatically clean the large-scale aerial image data. Firstly, the image quality features and image content features are extracted, in which the weights of different features are used to evaluate the effect on the final image quality. Secondly, the image spatial features are exploited to aggregate spatial information in weight maps of different sizes. Then, the image quality and content features are merged into the multi-information features that characterize the quality of the final image. As a result, the image is picked out according to the quality score. Finally, experimental results show that the proposed multi information fusion perception model has excellent performance on real databases.

Keywords: Data cleaning, power line inspection, deep learning, image quality assessment

1. INTRODUCTION

The rapid growth of power line inspection has produced large-scale aerial images. However, the accuracy rate is only 10%-30%¹ if they are directly used for power line detection without evaluating and cleaning. Therefore, it is significantly crucial for pre-process the aerial images to improve the accuracy of recognition and detection. The purpose of blind image quality assessment (BIQA)² is to enable computers to imitate human beings that assess images quality in the absence of a reference image. A typical application of the BIQA is the wholesale image data cleaning, which the purpose is to select high-quality images from large-scale images containing invalid images. Hence, designing a good IQA model plays a significant role in image data cleaning.

In the past few years, inspired by the many successful applications with deep learning methods in vision tasks³⁻⁶, several learning-based BIQA studies have been successively proposed in References^{2,7-18}. For instance, Siahaan et al.¹² use image semantic information to observe its quality prediction, which improves the performance of existing method. Zhu et al.⁸ give a deep meta-learning based BIQA approach that learns the meta-knowledge and can better deal with a variety of distortions. Norwegian Research Centre and Shenzhen University jointly launched a multi-scale image quality assessment method (i.e., MUSIQ model¹⁹), which was the first application of the transformer in the field of BIQA.

However, most of these algorithms are proposed over elaborated databases for BIQA and underperform on real-world aerial images in the wild, which is problematic for image data cleaning (DC) in practical application. The main reasons are as follows: 1) The existing BIQA models usually apply a single quality score as a cleaning indicator¹³, which cannot characterize the diversity of images quality; 2) The existing BIQA models are not well adapted to the variability of real aerial images; 3) The authentic wild image distortion databases are mixed with various distortions, which makes their application scope is limited¹⁵.

Our goal is to develop a BIQA method, in this paper, which considers the variable content and spatial information of images to accurately estimate the quality of the areal images and ensure that invalid images can be removed from original large-scale image databases. The main work is summarized as follows: 1) A novel multi-information fusion perception method is proposed, which combines the perceptual content information with network learning for enhancing the

* yangming_g@nwpu.edu.cn

adaptability in practical applications and accurately predicting the final quality of the aerial images of power line; 2) Given the fact that the spatial information of the aerial images will affect the image quality, the WSP module in Reference¹⁴ is introduced to adapt the variation of feature vectors of different sizes such that the final quality of the evaluated images can be comprehensively represented; 3) Experiments are performed on the actual aerial image data, and the results show that the MIFP method can not only accomplish the data cleaning of power line images but also achieve desirable experimental results in the field of BIQA.

2. MULTI-INFORMATION FUSION PERCEPTION METHOD

The architecture of our MIFP is shown in Figure 1. The aerial images are delivered to the network of feature extraction, called the Resnet50, to generate image quality features (IQA Features) and content-aware features (ICA Features). Another branch is the WSP which can learn the corresponding weights of spatial hierarchical information. Then, the MIFP module merges the image quality features, the image content features and image spatial information into the multi-information features which as input of image quality prediction regression network. The specific details of each component are as follows.

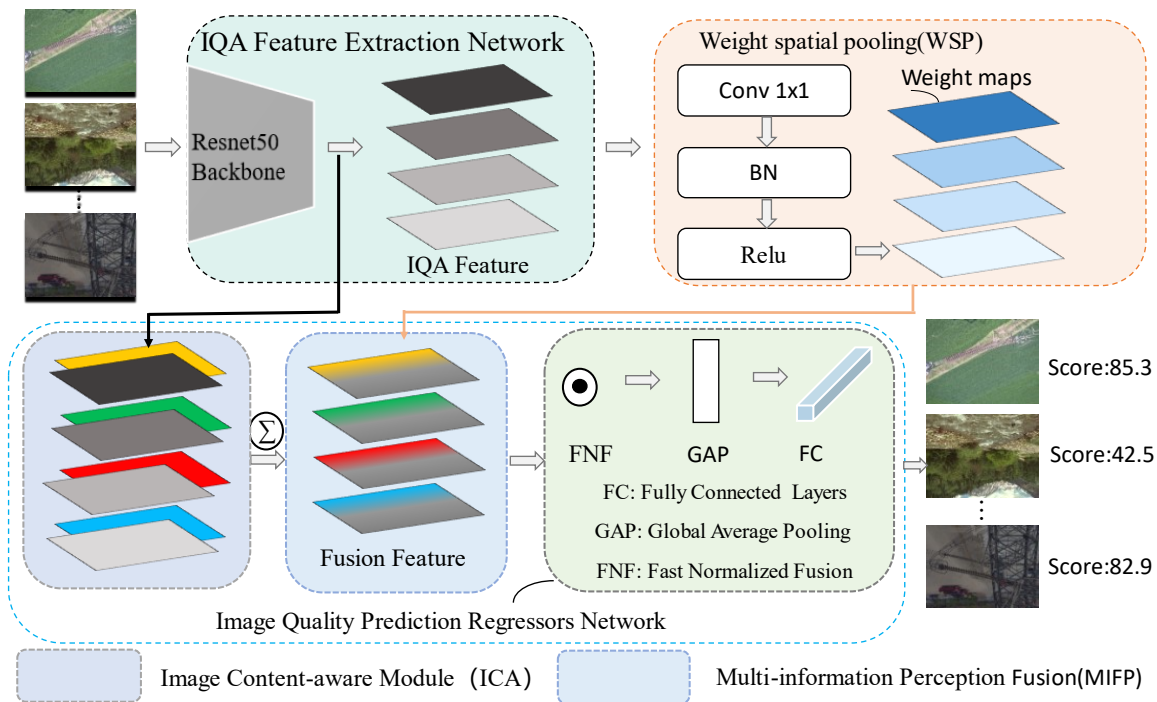


Figure 1. The overall architecture of the MIFP.

Feature extraction network. Inspired by References^{12,20}, the human visual system perceives an image mainly deriving from four aspects: the luminance and chrominance information of an image, the sharpness of the picture, the picture information and the local contrast information. The existing methods take all the above features together as the standard of image quality evaluation, without considering the impact of the division of different features on image quality in different real tasks. To enhance the robustness of the model in different tasks, the luminance and chrominance information and the sharpness of the picture are used as the basic characteristics of the image to characterize a little part of the image quality. To achieve this, we using the Resnet50 as the backbone network which was initialized on ImageNet²¹ while removing the global average pooling and fully connected layers. We naturally extract IQA features from conv1, conv2_9, and conv3_12 layers which can be described as:

$$V_{iqa}^{k+1} = g(\mathbf{w}_{iqa}^k * V_{iqa}^k * + b^{k_1}) \quad (1)$$

where V_{ica}^{k+1} stands for k -th ICA feature, w_{ica}^k is k -th convolution kernel and b^{k_1} represents the bias. The conv4_18 is used to extract local contrast information of images which outputs local content feature. Finally, the features extracted by conv5_9 can be used as the global semantic features of the image. In this paper, the local contrast information and global semantics are called image content feature. The output potentially relevant content feature is defined as:

$$V_{ica}^{k+1} = g(w_{ica}^k * V_{ica}^k * b^{k_2}) \quad (2)$$

where V_{ica}^{k+1} stands for k -th ICA feature, w_{ica}^k and b^{k_2} is convolution kernel and bias.

Multi-information fusion perception module. Intuitively speaking, the characteristics of different contents have different influences on the quality of the original image, and the quality is related to image spatial information. The prescriptive BIQA methods often ignore the influence of spatial information and different content on image quality in multi-scale information fusion. With the above two reasons, we leverage the ICA and WSP to draw into visual multi-information in our task hierarchically. As shown in Figure 1, we add a MIPF module which incorporate multi-information and WSP into IQA features to generate an objective evaluation result. In particular, we find the proportion of fusion features that best fits the task by adjusting different weight parameters. The output feature is defined as:

$$V_{fusion}^k = w_{ica}^k * f(g(g_{ica}^k)) \odot w_{ica}^k * f(g(g_{ica}^k)) \odot w_{wsp}^k * f(g(g_{wsp}^k)) \quad (3)$$

where V_{fusion}^k denotes the k -th fusion feature, w_{ica}^k , w_{ica}^k and w_{wsp}^k are the k -th weight of image content feature at different stages, and \odot denotes the cross-channel fusion.

Regression stage. The goal of regression is to map image quality features to image quality scores, which can be given by:

$$Q_p = \varphi(V_f, \gamma) \quad (4)$$

where Q_p is the image quality score, φ is the network model and γ represents the parameters. In this paper, a simple network is used for quality prediction. The function of the quality regression network maps the image quality to a specific score. As shown in Figure 2, in the quality regression stage, we use fully connected layer predict the final quality. It receives the multi-information feature vector as the input. Meanwhile, ReLU function is used for the activation function. The loss function is defined as⁸:

$$L_{loss} = \|Q_p - Q_1\|^2 \quad (5)$$

where Q_1 is the ground-truth quality score derived from subjective experiments.

Image data cleaning. After completing the above work, the cleaning work is relatively simple. The distribution of quality scores is observed through experiments. The computer can pick out the images according to the image quality score indicators and finally obtain high-quality databases.

3. EXPERIMENTAL SECTION

3.1 Experimental setup

Databases: In the experimental phase, images of power lines in different backgrounds collected from multi-rotor drones were used. In order to facilitate industry experts' subjective scoring of images, we also designed a subjective scoring crowdfunding gadget, and finally obtained the subjective scoring file corresponding to the database. At the same time, a separate flag is introduced into the subjective scoring file for faster learning of image content features. In addition, we use some authentically BIQA image databases, such as BID¹⁶, LIVEC¹⁷ and KonIQ-10k¹⁸. We also tested our MIFP method on CSIQ¹⁹.

Implementation details and performance criteria. We implemented our MIFP by PyTorch (1.10.2+cu113) and execute testing on the NVIDIA GeForce GPU (RTX 2080 Ti). Firstly, we randomly sample the training image and flip it horizontally into 25 patches of size 224×224 for data augmentation. Secondly, we use the Adam¹⁸ optimizer to train the

network and set the learning to rate 2×10^{-5} and the weight decay is 5×10^{-5} . Moreover, two performance criteria¹⁸, namely Pearson’s linear correlation coefficient (PLCC) and Spearman’s rank order correlation coefficient (SRCC) are used to evaluate the proposed MIFP.

3.2 Methods comparison

To validate images quality assessment performance of the MIFP for variable contents, we compare with the existing model¹⁵ when pre-testing specific image quality scores as shown in Figure 2. It can be seen that in the face of specific applications, the existing models cannot achieve the expected results while our model successively pick out images that don’t look great but are useful for the task. For instance, for power line detection, the images in the first row and the second column are useless images. Since the image is better than other images in terms of clarity and brightness, the existing method see it as a higher quality image, which is contrary to practical application. Although it does not look good, compared to the image in the first row and third column, it is still useful for work.

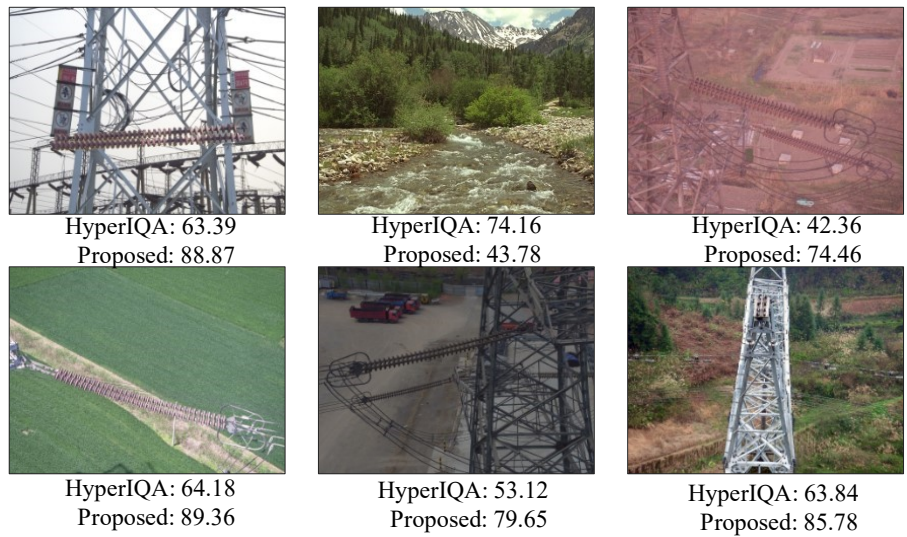


Figure 2. Comparing results with the existing methods in real task.

We also compare the MIFP approach with the latest BIQA methods based on deep learning models^{10,22-24}. From the results in Table 1, we can observe that although not specifically designed for IQA data, our method exhibits comparable performance on LIVE Challenge and KonIQ. This also just shows the robustness of our method_proposed in real applications.

Table 1. Experimental comparison on different datasets.

| BIQA methods | IQA Database | | | | | | | |
|------------------------|----------------|-------|--------------|--------------|-------|-------|-------|-------|
| | LIVE Challenge | | BID | | KonIQ | | CSIQ | |
| | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC |
| DB-CNN ⁹ | 0.851 | 0.869 | 0.848 | 0.859 | 0.875 | 0.884 | 0.946 | 0.959 |
| MetaIQA ⁸ | 0.802 | 0.835 | / | / | 0.850 | 0.887 | / | / |
| CaHDC ²⁴ | 0.738 | 0.744 | / | / | / | / | 0.903 | 0.914 |
| MMMNet ¹⁰ | 0.852 | 0.876 | / | / | / | / | 0.924 | 0.937 |
| HyperIQA ¹⁵ | 0.859 | 0.882 | 0.869 | 0.878 | 0.906 | 0.917 | 0.923 | 0.942 |
| UNIQUE ²⁵ | 0.854 | 0.890 | 0.858 | 0.873 | 0.896 | 0.901 | 0.902 | 0.927 |

| BIQA methods | IQA Database | | | | | | | |
|------------------------|----------------|--------------|-------|-------|--------------|--------------|-------|-------|
| | LIVE Challenge | | BID | | KonIQ | | CSIQ | |
| | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC | SRCC | PLCC |
| TRIQ ²⁶ | 0.779 | 0.800 | / | / | 0.882 | 0.893 | / | / |
| GraphIQA ¹¹ | 0.845 | 0.862 | / | / | 0.911 | 0.915 | 0.930 | 0.959 |
| HOSA ²² | 0.640 | 0.678 | 0.721 | 0.593 | 0.671 | 0.694 | 0.741 | 0.823 |
| BRISQUE ²³ | 0.608 | 0.629 | 0.562 | 0.593 | 0.665 | 0.681 | 0.746 | 0.829 |
| Proposed | 0.856 | 0.892 | 0.849 | 0.870 | 0.911 | 0.918 | 0.825 | 0.879 |

3.3 Ablation experiment

We perform ablation analysis to analyse the contribution level of each part in the MIFP approach. The results are shown in Figures 3 and 4. We designed four comparative experiments in this section and sequentially added different network components. The first experiment employs the ResNet50, fine-tuning the pre-trained network in Reference¹⁵. The ICA module is used in the second experiment on the basis of the ResNet. For CSIQ and LIVEC, ResNet with MIFP module and ICA module greatly improve accuracy compared with the vanilla ResNet. Particularly, compared with other ablation experiments, the improvements on LIVEC of the MIFP module and the ICA module are the highest, which suggests that the MIFP module and the ICA module are the key to the success of the proposed method.

To test the performance of the MIFP module, the MIFP module is introduced to replace the ICA module. As Figures 3 and 4, we can clearly observe that incorporating the MIFP module and WSP module outperforms IQA module. For PLCC, after adding the MIFP module and the WSP operation, its value increases from 0.85 to 0.879, The improvement ratio is 3.41%, and for SROCC, its value is improved by 4.33%; However, the performance on the CSIQ is remarkably small. SROCC is improved by 1.51% and PLCC is improved by 2.07%. This just explained that the MIFP method has better performance on real-world databases than synthetic image databases.

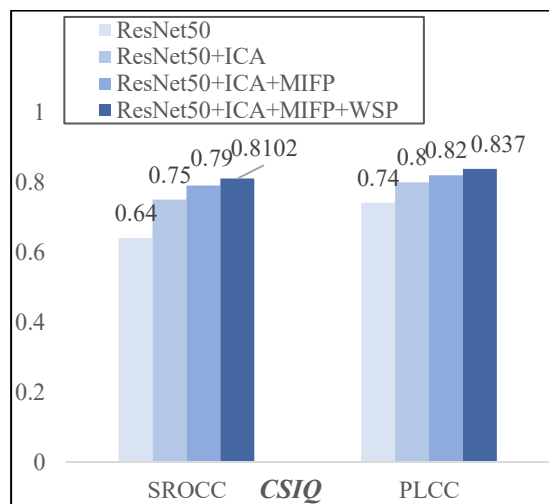


Figure 3. Experimental comparison result on CSIQ.

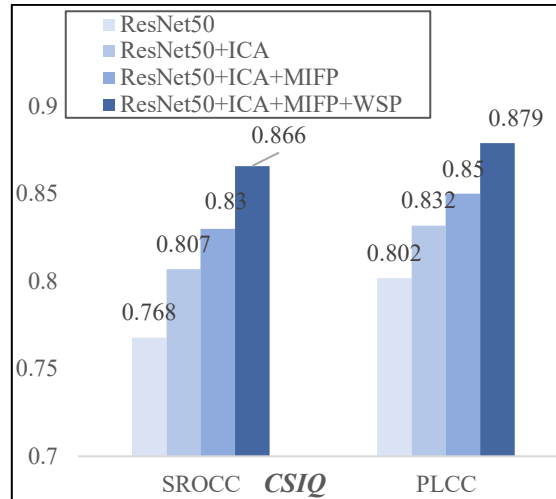


Figure 4. Experimental comparison result on LIVEC.

4. CONCLUSION

In this paper, a multi-information fusion perception method is proposed to solve two difficulties that arise in the intelligent processing task of aerial images. The MIFP method predicts the quality of the image after understanding its quality and content. Particularly, the WSP module is introduced to accommodate feature vectors of different sizes. Then, the quality of aerial images of power line can be more comprehensively characterized. Finally, the significance of the MIFP approach has been proved by experimental results.

ACKNOWLEDGMENTS

This work is supported by Open Fund of CETC Key Laboratory of Data Link Technology (No. CLDL-202022082), Opening Project of Science and Technology on Reliability Physics and Application Technology of Electronic Component Laboratory (No. 6142219200205), Fundamental Research Funds for the Central University, Shaanxi College Students Entrepreneurship Training Innovation Practice Project (No. S20211069427S).

REFERENCES

- [1] Tao, X., Zhang, D., Wang, Z., Liu, X., Zhang, H. and Xu, D., "Detection of power line insulator defects using aerial images analyzed with convolutional neural networks," *IEEE Transactions on Systems, Man, and Cybernetics Systems* 50(4), 1486-1498 (2018).
- [2] Park, K. C., Motai, Y. and Yoon, J. R., "Acoustic fault detection technique for high-power insulators," *IEEE Transactions on Industrial Electronics* 64(12), 9699-9708 (2017).
- [3] Zhao, Z., Xu, G. and Qi, Y., "Representation of binary feature pooling for detection of insulator strings in infrared images," *IEEE Transactions on Dielectrics and Electrical Insulation* 23(5), 2858-2866 (2016).
- [4] Zhao, Z., Xu, G. and Qi, Y., "Vancouver multi-patch deep features for power line insulator status classification from aerial images," *Inter. Joint Conf. on Neural Networks*, 2161-4407 (2016).
- [5] Liao, S. and An, J., "A robust insulator detection algorithm based on local features and spatial orders for aerial images," *IEEE Geoscience & Remote Sensing Letters* 12(5), 963-967 (2017).
- [6] Yang, Q., Yang, Z. and Zhang, T., "A random chemical reaction optimization algorithm based on dual containers strategy for multi-rotor UAV path planning in power line inspection," *Concurrency and Computation: Practice and Experience* 31(12), e4658 (2019).
- [7] Ghadiyaram, D. and Bovik, A. C., "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Transactions on Image Processing* 25(1), 372-387 (2015).
- [8] Zhu, H., Li, L., Wu, J., Dong, W. and Shi, G. M., "MetalQA: Deep meta-learning for no-reference image quality assessment,"

- Conf. on Computer Vision and Pattern Recognition, 14131-14140 (2020).
- [9] Zhang, W., Ma, K., Yan, J. and Deng, D., "Blind image quality assessment using a deep bilinear convolutional neural network," *IEEE Transactions on Circuits & Systems for Video Technology* 30(1), 36-47 (2019).
 - [10] Li, F., Zhang, Y. and Cosman, P. C., "MMNet: An end-to-end multi-task deep convolution neural network with multi-scale and multi-hierarchy fusion for blind image quality assessment," *IEEE Transactions on Circuits and Systems for Video Technology* 31(12), 4798-4811 (2021).
 - [11] Sun, S., Yu, T., Xu, J., Zhou, W. and Chen, Z., "GraphIQA: Learning distortion graph representations for blind image quality assessment," *IEEE Transactions on Multimedia*, (2022).
 - [12] Siahaan, E., Hanjalic, A. and Redi, J. A., "Semantic-aware blind image quality assessment," *Signal Processing Image Communication* 60, 237-252 (2017).
 - [13] Nguyen, V. N., Jenssen, R. and Roverso, D., "Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning," *International Journal of Electrical Power & Energy Systems* 99, 107-120 (2018).
 - [14] Su, Y. and Korhonen, J., "Blind natural image quality prediction using convolutional neural networks and weighted spatial pooling," *IEEE Inter. Conf. on Image Processing*, 191-195 (2020).
 - [15] Su, S., Yan, Q., Zhu, Y., Zhang, C., Ge, X., Sun, J. Q. and Zhang, Y. N., "Blindly assess image quality in the wild guided by a self-adaptive hyper network," *Conf. on Computer Vision and Pattern Recognition*, 3667-3676 (2020).
 - [16] Ciancio, A., Costa, A. and Silva, E., "No-reference blur assessment of digital pictures based on multifeature classifiers," *IEEE Transactions on Image Processing* 20(1), 64-75 (2011).
 - [17] Ghadiyaram, D. and Bovik, A. C., "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Transactions on Image Processing* 25(1), 372-387 (2015).
 - [18] Hosu, V., Lin, H., Sziranyi, T. and Saupé, D., "KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment," *IEEE Transactions on Image Processing* 29, 4041-4056 (2020).
 - [19] Ke, J., Wang, Q. and Wang, Y., "MUSIQ: Multi-scale image quality transformer," *Proc. of the Inter. Conf. on Computer Vision*, 5148-5157 (2021).
 - [20] Li, D., Jiang, T., Lin, W. and Jiang, M., "Which has better visual quality: The clear blue sky or a blurry animal?" *IEEE Transactions on Multimedia* 21(5), 1221-1234 (2018).
 - [21] Deng, J., Dong, W., Richard, S., Li, L. J., Li, K. and Li, F. F., "Imagenet: A large-scale hierarchical image database," *IEEE Conf. on Computer Vision and Pattern Recognition*, 248-255 (2009).
 - [22] Xu, J., Ye, P., Li, Q., Du, H., Liu, Y. and Doermann, D., "Blind image quality assessment based on high order statistics aggregation," *IEEE Transactions on Image Processing* 25(9), 4444-4457 (2016).
 - [23] Mittal, A., Moorthy, A. K. and Bovik, A. C., "No-reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing* 21(12), 4695-4708 (2012).
 - [24] Wu, J., Ma, J., Liang, F., Dong, W., Shi, G. and Lin, W., "End-to-end blind image quality prediction with cascaded deep neural network," *IEEE Transactions on Image Processing* 29, 414-7426 (2020).
 - [25] Zhang, W., Ma, K. and Zhai, G., "Uncertainty-aware blind image quality assessment in the laboratory and wild," *IEEE Transactions on Image Processing* 30, 3474-3486 (2020).
 - [26] Tan, M., Pang, R. and Le, Q. V., "Transformer for image quality assessment," *Inter. Conf. on Image Processing*, 1389-1393 (2021).