

A COMPARATIVE STUDY OF DNN-BASED MODELS FOR BLIND IMAGE QUALITY PREDICTION

Xiaohan Yang¹, Fan Li¹, Hantao Liu²

¹ School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, 710049, China

² School of Computer Science and Informatics, Cardiff University, Cardiff, CF243AA, UK

ABSTRACT

Recently, deep learning methods have gained substantial attention in the research community and have proven useful for blind image quality assessment (BIQA). Although previous study of deep neural networks (DNN) methods is presented, some novelty methods, which are recently proposed, are not summarized. In this paper, we provide a comparative study on the application of DNN methods for BIQA. First, we systematically analyze the existing DNN-based quality assessment methods. Then, we compare the predictive performance of various methods in synthetic and authentic databases, providing important information that can help understand the underlying properties between different methods. Finally, we describe some emerging challenges in designing and training DNN-based BIQA, along with few directions that are worth further investigations in the future.

Index Terms— deep learning, blind image quality assessment (BIQA), deep neural networks (DNN)

1. INTRODUCTION

With the development of social media and the increasing demand for imaging services, an enormous amount of visual data is making its way to consumers. Meanwhile, the acquisition, processing, compression, transmission, or storage of images is subject to various distortions, resulting degradation in visual quality. Therefore, methods for image quality assessment (IQA) have been extensively studied for the purpose of maintain, control and enhance the perceived image quality.

In principal, subjective assessment is the most reliable way to evaluate the visual quality of images [1]-[2]. But this method is time-consuming, expensive, and impossible to implement in real-world systems. Therefore, objective assessment of image quality has gained growing attention in recent years. Depending on to what extent a reference image is used for quality assessment, existing objective IQA methods can be classified into three categories: full-reference (FR), reduced-reference (RR) and no-reference/blind (NR/B) methods. Ac-

cessing all or part of the reference image information is unrealistic in many circumstances [3]-[4], hence it has become increasingly important to develop effective blind IQA (BIQA) methods.

Most BIQA methods were developed based on classical regression models [37]-[40]. The obvious limitation of those BIQA methods is that the hand-crafted features may not be able to adequately represent complex image structures and distortions. Recently, attempts have been made to adopt deep learning methods, which provide a very promising option for addressing the challenging BIQA task.

Although previous study introduces classical IQA methods or a part of end-to-end deep neural networks (DNN) methods [5]-[7], some methods that have been appeared recently are not summarized [25], [41]-[42]. In this paper, we intend to systematically discuss various deep learning methods for BIQA. First, we divide the DNN methods into two categories, which could distinguish different DNN-based methods easily. One is to extract deep features from DNN model, because it can capture high-level semantics of image and is highly related to quality degradation. The other is to directly predict quality from DNN model, which takes full advantage of back-propagated capability of DNN to optimize prediction accuracy. Moreover, we summarize useful findings and discuss the challenges of deep learning based BIQA methods.

The rest of this paper is organized as follows. In Sec. 2, we introduce and categorize different deep learning methods for BIQA. The performance of DNN-based models are analyzed in Sec. 3. The challenges of DNN-based BIQA methods are discussed in Sec. 4. The conclusion is provided in Sec. 5.

2. DNN-BASED IMAGE QUALITY PREDICTION

Deep learning techniques have achieved great success in solving various images recognition and processing problems [8]-[9]. Researchers in the image quality community have started to explore the use of deep network in IQA.

This research work was supported in part by National Science Foundation of China (61671365), and Joint Foundation of Ministry of Education of China (6141A02022344).

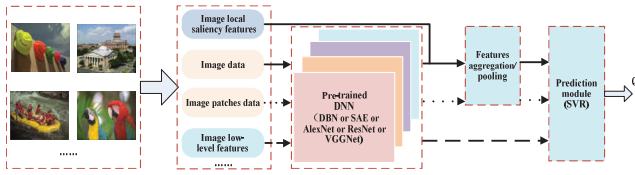


Fig. 1: The flowchart of extracting deep features methods

2.1. Extracting deep features from DNN models

The straightforward method to employing DNN models is to extract discriminative features for various distorted images, and then evaluate the image quality using conventional regression machine, such as support vector regression (SVR). It can be classified into two major schemes: 1) extracting from low-level features of images and 2) extracting from data of images or image patches. Fig.1 shows the flowchart of extracting deep feature methods from DNN models.

2.1.1. Deep features extracted from low-level features

This method aims to feed low-level image features relevant to quality perception into a pre-trained DNN model, such as the deep belief network (DBN) or stacked auto-encoder (SAE) [10]-[12], which can be used to generate high-level features. Then, the resulting deep features extracted by such pre-trained DNN model are used to train an SVR model to predict image quality.

Compared with traditional regression methods, the major advantage is deep features extracted from low-level features is highly related to quality degradation. But the limitation is hand-crafted low-level features need to be carefully designed as the input to DNN, which does not take full advantage of DNN.

2.1.2. Deep features extracted from image/image patches

Image or image patches can be directly fed into a pre-trained DNN model, which produces meaningful features automatically from the input data. The image quality can be predicted by regression methods from output features of the pre-trained DNN model [13]-[15], such as AlexNet [8], RestNet [9], VGGNet [16].

In [13], the authors adopted the AlexNet architecture [8] to extract the semantic features implied in global image content and used saliency detection and Gabor filters to capture low-level features relevant to local image content. The overall image quality was estimated by combining these features. This is inspired by the human visual perception of image quality that involves the integrated analysis of global high-level semantics and local low-level characteristics [17]. In [14], the RestNet [9] was used to represent deep features from each overlapping patches. These patch features were aggregated in a statistical way and the image quality was evaluated by a lin-

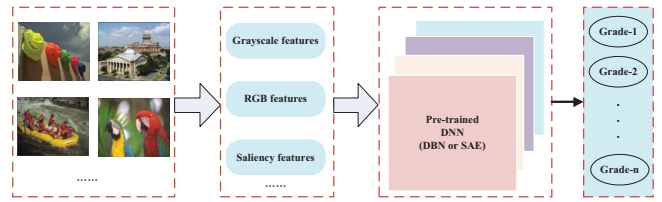


Fig. 2: The flowchart of qualitative methods

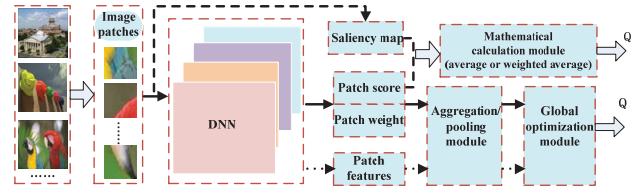


Fig. 3: The flowchart of patch-input methods

ear regression method. In [15], a DNN model formed by the VGGNet [16] was used to extract image features over each layer, the image quality was estimated by averaging layer-wise predicted scores.

The essence of above mentioned methods is to extract deep features to evaluate image quality by using classification or recognition network. However, the limitation is the features of IQA are different from those of classification or recognition tasks.

2.2. Predicting image quality from DNN models

According to different evaluation metrics for quality prediction, two classification methods are discussed based on either qualitative metric or quantitative metric.

2.2.1. Qualitative metric to predict quality grade

Extensive psychological evidence has shown that humans prefer to conduct evaluations qualitatively rather than numerically [18]. This metric treats BIQA as classification problem and the general flowchart is shown in Fig. 2. The low-level features of image were fed into the DBN or SAE and consequently classify image grades [19]-[21].

Grade quality as a qualitative adjective is usually consistent with human perception, such as excellent, good or bad. However, compared with the methods of predicting numerical scores, the drawback is that different definitions of grades of subjective opinions can significantly impact the performance of methods.

2.2.2. Quantitative metric to predict quality score

Quantitative methods based on DNN can complete the whole prediction process from image input to quality score output. Therefore, it can be regarded as a typical regression problem. Although the previous classification method is useful

[7], some DNN methods that have been appeared recently are not analyzed [25], [41]-[42]. Therefore, according to different input in DNN, we propose a classification method: patch-input and image-input to discuss quantitative methods.

In patch-input methods, the image is split into multiple patches. In [22]-[23], each patch with image subjective score as network input to train DNN, the image quality is estimated by the average score of all image patches. However, it ignores that the visual quality of different local regions is often different and humans tend to concentrate on the regions of interest when evaluating an image. Instead of using the average of all patches' scores, the weight of each patch is assigned by saliency map and image score is calculated by weighted average over all patch scores[24]. In [29]-[30], instead of using image subjective score as each patch label, they adopt proxy scores using FR methods as patch label to pre-train DNN. The goal is to obtain patch scores and weights or patch features. Then, the pre-trained DNN model can be fine-tuned to capture image scores by aggregating all image patches' features or scores and weights with subjective scores. Fig. 3 shows the flowchart of patch-input methods.

Due to the insufficient image data, the patch-input as data augmentation method is used to increase the training samples volume in order to avoid overfitting of deep network. However, these methods are problematic because the ground truth quality of each patch does not exist.

Recently, the image-input methods are developed [25],[41]-[42]. The novelty is that, despite a lack of image databases, the DNN based on image as input can also evaluate image quality very well. According to the different extended objects, we classify these methods into two sub-categories: expanding distorted images and expanding reference images.

As shown in Fig.4, expanding distorted images method [25],[41] aims to generate different distorted level images, which are similar to existing IQA databases. Then, they pre-train a DNN model to rank these images. Finally, they fine-tune the pre-trained DNN to estimate image quality scores by using IQA database. Expanding reference images method [42] is to train generative adversarial network (GAN) to generate reference images of distorted images. Then, each distorted image and corresponding discrepancy map are used to train DNN to evaluate image score. this method are presented in Fig.5.

3. THE PERFORMANCE OF DNN MODELS

We conduct the experiments on five public IQA databases, including synthetic databases (LIVE[31], TID2013[32], CSIQ[33], LIVE multiply distorted (MD)[34]) and authentic database (LIVE In the Wild Image Quality Challenge Database (LIVEC) [35]). In each database, we randomly select 80% of the distorted images as the training set and the remaining 20% of images as the testing set. There is no overlap in image contents between these training and test-

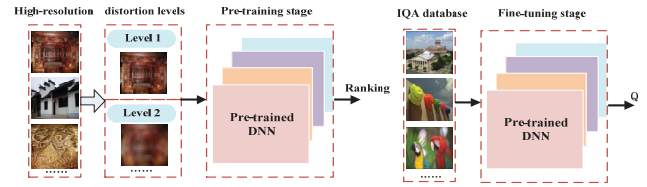


Fig. 4: The flowchart of expanding distorted images' methods

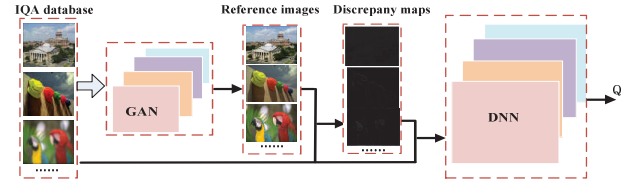


Fig. 5: The flowchart of expanding reference images' methods

ing sets. The Spearman Rank-Order Correlation Coefficient (SROCC) and Person Linear Correlation Coefficient (PLCC) are adopted for the evaluation of IQA metrics.

We compare the state-of-the-art BIQA and FR-IQA methods, including: FR-IQA methods (PSNR, SSIM [3], FSIMc [36]) and classic BIQA methods (BRISQUE [37], CORNIA [38], GMLOG [39], ILNIQE [40] and BWS [43]), current leading BIQA methods based on DNN (MGDNN[12], FRIQUEE [11], DLIQA[19], SESANIN[21], BLNDER [15], CNN [22], RANKIQA [25], DIQaM [28], BIECON [29], DIQA [30] , DB-CNN [41], HIQA [42]). However, it is difficult to exactly reproduce the IQA methods based on DNN. We therefore first adopted the results reported in the literature. Especially, for the cases where experimental results were not given, we used the released code to conduct experiments and generate results.

As shown in Table 1, the SROCC and PLCC values are reported to various methods. The best three performance among the BIQA methods are shown in bold. The weighted average of the SROCC and PLCC over the five databases is shown in the last column. The weight of each database is proportional to the number of distorted images in the database.

We find that the DNN methods are superior over all classical NR methods. This is because DNN methods can automatically extract network features instead of handcrafted features in classical NR methods. In addition, DNN methods are highly competitive with the FR methods. However, DNN methods do not use any prior information of reference for image quality assessment.

We compared with extracting deep features methods from DNN models. Although the results are not incomplete, the methods extracted from image/image patches are superior to the methods extracted from low-level features. It is easy to understand the limited low-level features can not fully represent image distortions and contents.

Compared with predicting image quality methods from

Table 1: The SROCC and PLCC comparison on the four databases

Types	Algorithms	LIVE		TID2013		CSIQ		LIVEMD		LIVEC		Weighted Average	
		SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC	SROCC	PLCC
FR	PSNR	0.876	0.872	0.636	0.706	0.806	0.800	0.725	0.815	N/A	N/A	N/A	N/A
	SSIM[3]	0.913	0.945	0.775	0.691	0.834	0.861	0.845	0.882	N/A	N/A	N/A	N/A
	FSIMc[36]	0.963	0.960	0.802	0.877	0.913	0.919	0.863	0.818	N/A	N/A	N/A	N/A
NR	BRISQUE[37]	0.939	0.942	0.572	0.651	0.775	0.817	0.897	0.921	0.607	0.645	0.676	0.729
	CORNIA[38]	0.942	0.943	0.549	0.613	0.714	0.781	0.900	0.915	0.618	0.662	0.659	0.708
	GMLOG[39]	0.950	0.954	0.675	0.683	0.803	0.812	0.824	0.863	0.543	0.571	0.713	0.727
	IL-NIQE[40]	0.902	0.908	0.521	0.648	0.821	0.865	0.902	0.914	0.594	0.589	0.651	0.719
	BWS[41]	0.934	0.943	0.597	0.622	0.786	0.820	0.901	0.922	0.482	0.526	0.666	0.693
	MGDNN[12]	0.951	0.949	-	-	-	-	-	-	-	-	-	-
	FRIQUEE[11]	-	-	-	-	-	-	-	-	0.672	0.705	-	-
	BLNDER[15]	0.966	0.959	0.819	0.838	0.961	0.968	0.944	0.964	0.945	0.953	0.890	0.902
	DLIQA[19]	0.929	0.934	-	-	-	-	-	-	-	-	-	-
	SESANIN[21]	0.934	0.948	-	-	-	-	0.836	0.838	-	-	-	-
	CNN[22]	0.956	0.953	0.558	0.653	0.683	0.754	0.933	0.927	0.516	0.536	0.604	0.702
	BIECON[29]	0.961	0.960	0.717	0.762	0.815	0.823	0.909	0.933	0.663	0.705	0.765	0.797
	DIQaM[28]	0.960	0.972	0.835	0.855	0.869	0.894	0.906	0.931	0.606	0.601	0.817	0.832
	DIQA[30]	0.970	0.972	0.843	0.868	0.844	0.880	0.920	0.933	0.703	0.704	0.839	0.857
	RankIQA[25]	0.981	0.982	0.780	0.793	0.892	0.912	0.908	0.929	0.641	0.675	0.800	0.818
DB-CNN[41]	0.968	0.971	0.816	0.865	0.946	0.959	0.927	0.934	0.851	0.869	0.868	0.897	
HIQA[42]	0.982	0.982	0.879	0.880	0.885	0.901	-	-	-	-	-	-	

Red: the highest. Blue: the second. Green: the third.

Table 2: The SROCC comparison of the cross dataset test

Train	Test	[37]	[38]	[39]	[15]	[25]	[30]
LIVE	CSIQ subset	0.890	0.898	0.897	0.700	0.797	0.906
	TID2013 subset	0.878	0.879	0.907	0.652	0.873	0.918
CSIQ subset	LIVE	0.919	0.920	0.903	0.825	0.564	0.923
	TID2013 subset	0.874	0.852	0.879	0.661	0.777	0.915
TID2013 subset	LIVE	0.877	0.907	0.889	0.751	0.769	0.905
	CSIQ subset	0.861	0.859	0.794	0.782	0.735	0.871

DNN models, the methods of predicting quality scores are better than those of predicting quality grade. This is because different definitions of grades of subjective opinions may impact prediction performance. Meanwhile, we compare the patch-input methods and image-input methods in predicting quality scores. It can be seen that there is a competitive relationship between the two types of methods. The main reason is that both of these methods solve the problem of insufficient IQA databases. However, we find that these two types of methods also have shortcomings. In patch-input methods, the main challenge is the acquisition of relevant labels for these patches of an image. In image-input methods, the main shortcoming is how to accurately simulate distortion or reference images similar to IQA databases.

Table 2 specifies the results of generalizability of the classic BIQA methods and DNN methods. More specifically, BIQA methods are trained using all the images from one database, and then tested on another database. In the CSIQ and TID2013 databases, four overlapping distortion types (WN, GB, JPEG, JP2K) are used. We find that the generalization ability of DNN model is better than the classic BIQA. Furthermore, the generalizability of patch-input method is

superior to other DNN methods.

4. CHALLENGES OF DNN METHODS

Although DNN-based IQA methods can achieve outstanding performance due to their strong representation capability, there are several challenges at the same time. The lack of large training data sets is often mentioned as a challenge. Therefore, understanding how to successfully create reliable, very large-scale databases is still an open question. Moreover, we also expect that training an end-to-end DNN model in a completely unsupervised manner is worth further investigations in the future. This is because obtaining large amounts of unlabeled data is generally much easier than labeled data and human learning is largely unsupervised.

5. CONCLUSION

This paper presents a comparative study of various DNN-based methods for BIQA. We discussed and analyzed the state-of-the-art DNN methods according to different strategies of DNN models: deep features learning and mechanisms of quality prediction. Especially, some methods, which are not present in previous study, are also discussed. Then, models with competitive performance due to use of deep learning have been observed and summarized, yet the state of research in this field is far from mature. Meanwhile, challenging issues should be noticed. We hope this comparative study of DNN methods can serve as a useful reference towards a better understanding of this research field.

6. REFERENCES

- [1] F. Li, S. Fu, Z. Li and X. Qian, "A cost-constrained video quality satisfaction study on mobile device," *IEEE Trans. Multimedia*, vol. 20, no. 5, pp. 1154–1168, May. 2018.
- [2] X. Zhang, W. Lin, S. Wang, J. Liu, S. Ma and W. Gao, "Fine-grained quality assessment for compressed image," *IEEE Trans. Image Process.*, vol. 28, no. 3, pp. 1163–1175, Apr. 2004.
- [3] 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.
- [4] S. A. Golestaneh and L. J. Karam, "Reduced-reference quality assessment based on the entropy of DWT coefficients of locally weighted gradient magnitudes," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5293–5303, Nov. 2016.
- [5] R. Manap and L. Shao, "Non-distortion-specific no-reference image quality assessment: a survey," *Information Sciences*, vol. 301, pp. 141–160, 2015.
- [6] A. G. George and K. Prabavathy, "A survey on different approaches used in image quality assessment," *Int. J. Computer Sci. and Network Security*, vol. 14, no. 2, pp. 197–203, 2014.
- [7] J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang and A. C. Bovik, "Deep convolutional neural models for picture-quality prediction," *IEEE Signal processing magazine*, pp. 130–141, 2017.
- [8] A. Krizhevsky, I. Sutskever, and H. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. NIPS*, vol. 25, no. 2, pp. 1097–1105, 2012.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. CVPR*, pp. 770–778, 2016.
- [10] H. Tang, N. Joshi, and A. Kapoor, "Blind image quality assessment using semi-supervised rectifier networks," in *Proc. IEEE Conf. CVPR*, pp. 2877–2884, 2014.
- [11] D. Ghadiyaram and A. C. Bovik, "Blind image quality assessment on real distorted images using deep belief nets," in *Proc. IEEE Global Conf. Signal Inf. Process.*, pp. 946–950, 2014.
- [12] Y. Lv, G. Jiang, M. Yu, H. Xu, F. Shao, and S. Liu, "Difference of Gaussian statistical features based blind image quality assessment: A deep learning approach," in *Proc. IEEE Conf. ICIP*, pp. 2344–2348, 2015.
- [13] C. Sun, H. Li and W. Li, "No-reference image quality assessment based on global and local content perception," in *Proc. IEEE Conf. VCIP*, pp. 27–30, 2016.
- [14] D. Li, N. T. Jiang, and M. Jiang, "Exploiting high-level semantics for no-reference image quality assessment of realistic blur images," in *Proc. ACM MM*, pp. 378–386, Oct. 2017.
- [15] F. Gao, J. Yu, S. Zhu, Q. Huang and Q. Tian, "Blind image quality prediction by exploiting multi-level deep representations," *Pattern Recognition*, vol. 81, pp. 432–442, 2018.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. ImageNet Challenge*, pp. 1–14, 2014.
- [17] W. Zhang, R. Martin and H. Liu, "A saliency dispersion measure for improving saliency-based image quality metrics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1462–1466, Jul. 2018.
- [18] E. Hutchins and G. Lintern, *Cognition in the Wild*. Cambridge, MA, USA: MIT Press, 1995.
- [19] W. Hou, X. Gao, D. Tao, and X. Li, "Blind image quality assessment via deep learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 6, pp. 1275–1286, 2015.
- [20] W. Hou and X. Gao, "Saliency-guided deep framework for image quality assessment," *IEEE Multimedia Mag.*, vol. 22, no. 2, pp. 46–55, 2015.
- [21] Y. Li, L.-M. Po, X. Xu, L. Feng, F. Yuan, C.-H. Cheung, and K.-W. Cheung, "No-reference image quality assessment with Shearlet transform and deep neural networks," *Neurocomputing*, vol. 154, pp. 94–109, 2015.
- [22] L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proc. IEEE Conf. CVPR*, pp. 1733–1740, 2014.
- [23] Y. Li, L. M. Po, L. Feng, and F. Yuan, "No-reference image quality assessment with deep convolutional neural networks," in *Proc. IEEE Int. Conf. Digital Signal Processing*, pp. 685–689, 2016.
- [24] Z. Cheng, M. Takeuchi, and J. Katto, "A pre-saliency map based blind image quality assessment via convolutional neural networks," *IEEE Intl. Symposium on Multimedia*, pp. 77–82, 2017.
- [25] X. Liu, J. Weijer, and A. Bagdanov, "RankIQ: Learning from ranking for no-reference image quality assessment," in *Proc. IEEE Conf. ICCV*, pp. 1040–1049, 2017.
- [26] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Conf. CVPR*, pp. 349–356, Jun. 2005.
- [27] J. Guan, S. Yi, X. Zeng, W. Cham and X. Wang, "Visual importance and distortion guided deep image quality assessment framework," *IEEE Trans. Multimedia*, vol. 19, no. 11, pp. 2505–2521, 2017.
- [28] S. Bosse, D. Maniry, K. Muller, 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, 2018.
- [29] J. Kim and S. Lee, "Fully deep blind image quality predictor," *IEEE J. Sel. Topics Signal Process.*, vol. 11, no. 1, pp. 206–220, Feb. 2017.
- [30] J. Kim, A. Nguyen, and S. Lee, "Deep CNN-based blind image quality predictor," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–14, 2018.
- [31] H. Sheikh, M. Sabir, and A. 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.
- [32] N. Ponomarenko et al, "Image database TID2013: Peculiarities, results and perspectives," *Signal Process.: Image Commun.*, vol. 30, pp. 57–77, Jan. 2015.
- [33] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *J. Electron. Imag.*, vol. 19, no. 1, pp. 19–21, 2010.
- [34] D. Jayaraman, A. Mittal, A. K. Moorthy, and A. C. Bovik, "Objective quality assessment of multiply distorted images," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, pp. 1693–1697, 2012.
- [35] D. Ghadiyaram and A. C. Bovik, "Massive online crowdsourced study of subjective and objective picture quality," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 372–387, Jan. 2016.
- [36] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [37] 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, 2012.
- [38] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. IEEE Conf. CVPR*, pp. 1098–1105, Jun. 2012.
- [39] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4850–4862, Nov. 2014.
- [40] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2579–2591, Aug. 2015.
- [41] W. Zhang, K. Ma, J. Yan, D. Deng and Z. Wang, "Blind image quality assessment using a deep bilinear convolutional neural network," *IEEE Trans. Circuits and Systems for Video Technology*, 2018.
- [42] K. Lin and G. Wang, "Hallucinated-IQA: no-reference image quality assessment via adversarial learning," in *Proc. IEEE Conf. CVPR*, pp. 732–741, Aug. 2018.
- [43] X. Yang, F. Li, W. Zhang and L. He, "Blind image quality assessment of natural scenes based on entropy differences in the DCT domain," *Entropy*, vol. 20, no. 12, pp. 885–906, 2018.