Blind Image Quality Assessment Using a CNN and Edge Distortion

Rajesh Babu Movva*, Raja Kumar Kontham

Department of Computer Science & Systems Engineering, College of Engineering(A), Andhra University, Visakhapatnam 530003, A.P, India

Corresponding Author Email: mrb.csebec@gmail.com

ABSTRACT

The present paper introduces a Convolutional Neural Network (CNN) for the assessment of image quality without a reference image, which comes under the category of Blind Image Quality Assessment models. Edge distortions in the image are characterized as input feature vectors. This approach is in justification of the fact that subjective assessment focusses on image features that emanate from the edges and the boundaries present in the image. The earlier methods were found to use complex transformations on the image to extract the features before training or as a part of the training. The present work uses Prewitt kernel approach to extract the horizontal and vertical edge maps of the training images. These maps are then input to a simple CNN for extracting higher level features using non-linear transformations. The resultant features are mapped to image quality score by regression. The network uses Spatial Pyramid Pooling (SPP) layer to accommodate input images of varying sizes. The present proposed model was tested on popular datasets used in the domain of Image Quality Assessment (IQA). The experimental results have shown that the model competes with the earlier proposed models with simplicity of feature extraction and involvement of minimal complexity.

1. INTRODUCTION

It is a well-known fact that a picture is worth thousand words. Communicating and sharing of images is an integral part of present-day life especially in the backdrop of social networking. It is equally true that image capturing, processing and transmission using low bandwidths is often vulnerable to various distortions. The judgement of image quality performs a crucial role within the design of several image processing applications. Human judgement of image quality is subject to several limitations because of which automatic image quality estimation has been an important research area in image processing and computer vision [1]. Automatic image quality assessment of distorted images is classified as Full Reference Image Quality Assessment (FR-IQA), Reduced-Reference Image Quality Assessment (RR-IQA) and No-Reference Image Quality Assessment (NR-IQA) or Blind Image Quality Assessment (BIQA) depending on the evidence obtainable regarding the original image. The evolution of CNN led to a new methodology for Blind Image Quality Assessment. A CNN is trained with a dataset of distorted images and their corresponding quality scores and is used to evaluate the image quality in BIQA. The quality scores are normally the Difference Mean Opinion Scores (DMOS) obtained from the categorical judgement (like good, poor, bad) of the distorted images by humans.

FR-IQA models need the original image to assess the image quality [2-4]. RR-IQA models have access to partial information of the original image for judging the quality of its distorted version [5-8]. BIQA models assume no information regarding the original image to estimate the quality of its distorted version. These models can be broadly classified as:

conventional models, which are based on natural scene statistics, machine learning models, which are based on image and/or hand-crafted features of the image to train the model and CNN based Deep Learning models, which extract the representational features of the image using the earlier layers of the network and map them to the quality scores with fully connected layers during training.

Mittal et al. [9] developed a blind image special quality assessment model called (BRISQUE), which quantifies costs of spontaneity within the image supported locally normalized luminance coefficients as image quality. The model claims implementation enhancements over the state-of-the-art models. Moorthy and Bovik [10] introduced a BIQA model as a two-stage framework that clusters wavelet coefficients in Inter and Intra subbands of various scale and orientation by modelling them using Gaussian scale mixtures (GSM) model. Lixiong Liu et al. [11] proposed a curvelet based model that extracts features from the strength distribution of together scale and orientation within the curvelet field and therefore the directs of the maxima of the log histograms of the curvelet coefficient values. The extracted energy features from curvelet domain are highly appropriate to real image quality among multiple non-natural categories. Support vector machine (SVM) is employed for linear classification of image quality with human subjective opinions. Saad et al. [12] developed a NR-IQA model that extracts features from the real scene statistics model of the picture discrete cosine transform measurements. The resulting elements are used in a plain Bayesian inference method to predict image quality scores. Liu et al. [13] developed oriented gradients image quality assessment model, which maps the extracted relative gradient magnitude features of the image to the image quality by using Ada boosting back

propagation neural network. The estimated accuracy of the experiment is superior to the performance of the state-of-the-art models. Ye et al. [14] designed a NR-IQA structure that uses raw image pieces separated from a set of un-labelled images to realize a vocabulary of image features in an unsupervised manner. The authors deployed a sensitive task enciphering with max pooling to acquire features that represent the image for quality estimation.

Liu et al. [15] proposed a NR-IQA approach that learns ranking of images by using synthetically generated datasets and a Siamese Network. The trained Siamese Network shares the knowledge with a traditional CNN and the CNN is trained to fine tune the transferred weights with a batch of images obtained from all sets of images in the group. During the testing phase, the distorted image is randomly sampled for thirty sub images and the image quality is computed as the median of the quality scores of the sub images. Kang et al. [16] developed a CNN by extracting features from locally normalized pixel intensities of image patches with different window sizes. The learned features are regressed to the image quality score by using two fully connected layers and an output node. For a given image, the anticipated patch scores are be a mean of to seek out the image level quality score. Bosse et al. [17] introduced an end-to-end image quality assessment CNN with ten convolutional layers followed by five pooling layers intended for feature extraction and finally introduce two fully connected layers for regression. The method incorporates joint optimization of weighted average patch accumulation for pooling local patch qualities to global image quality. Bianco et al. [18] used Caffe network architecture that has a feature extractor on top of which a SVR machine with a linear kernel was used to map the extracted features to the subjective quality scores. Li et al. [19] proposed an image quality assessment method based on ResNet [20] architecture claiming that the ReLu activation allows nonlinear changes for isolating high-level image features resulting in consistent measurement of image quality better than linear filters.

Most of the methods in the literature used low level and complex transformations to extract image features for training a CNN. This observation inspired the authors to design a method, which extracts higher-level features of an image to quantify the image quality, as the subjects are able to extract only higher-level representative features to judge the quality of an image. The proposed method extracts horizontal and vertical edge maps of the images in the training dataset and uses them for training a CNN to assess the image quality.

The present article is organized as follows. In Section 2, we describe the proposed approach. In Section 3, we describe the training, testing process, experimental results and conclude the paper in Section 4.

2. PROPOSED APPROACH

Let X be the distorted image and y be the DMOS of the distorted image present in the dataset. The value of y is normally in the range 0 to 100. A value of zero indicates that the image is not subjected to any distortion, which means that X is an original image. On the other hand, a value of 100 indicates that the image is completely distorted. However, conventions (0 to 1, 0 to 9) and their interpretation may differ from one dataset to another dataset. In BIQA, a CNN is trained with pairs of (X, y) available in the dataset. The initial convolutional layers extract the representational features of the

image and the subsequent fully connected layers map these features to the image quality score y. Such a trained network can be employed to review the quality of an unseen distorted image X'. A CNN can effectively learn the representational features of an image when it is fed with a larger number of samples. A notable problem in the domain of the image quality estimation is that the datasets are small in size and contains less than two thousand images on an average. Present work effectively solves this problem by obtaining the horizontal edge maps and the vertical edge maps in each image using Prewitt [21] edge detector and are fed to the network. The transformed input helps the convolutional layers to quickly extract the higher-level image features though the dataset which is smaller in size. Further, the input is augmented by dividing the feature maps into patches. The number of patches into which a feature map is to be divided is a critical design decision. Large number of patches will result in the loss of feature relationships and results in semantically unconnected patches. If the number is too small, then the purpose of the augmentation may not be served. In general, humans focus on four quadrants of an image during perception. Hence, in the present work, augmentation is done by dividing each feature map of image X into four patches and ascertaining the image score y to each of the patches of the feature map.

2.1 Input to the network

The Input dataset to the network is created from the original public dataset with the following steps: Each gray scale version of the image X is transformed to horizontal edge map X^h and vertical edge map X^v using Prewitt kernel.

Prewitt is a gradient-based kernel to detect the direction and degree of an image. It adds the gradient estimate of image intensity function for image power detection. At the pixels of an image, the Prewitt kernel computes either the normal to a vector or the alike gradient vector. It uses two 3 x 3 kernels which are convoluted to calculate estimates of the derivatives, one for flat changes, and the one for perpendicular. The kernels are:

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Each feature map is divided into four equal sized patches resembling the four quadrants of a plane. Thus, each image X of size (w,h) results in eight feature maps of size (w⁄4,h⁄4). The DMOS y of the image X is paired with each of these patches as the label. So, the size of the training dataset is multiplied by eight times. Figure 1 illustrate the creation of the input dataset.

2.2 Network architecture of CNN

Let $w * h$ be the width and height of the image input to the network. The proposed CNN comprises of the following convolutional layers and fully connected layers as shown in the Figure 2.

1) A Convolutional Layer with (3, 3) filter size and Rectified Linear unit (ReLu) activation function with 16 filters. The output of this layer is an image of size $w * h * 16$.
2) Average pooling layer with a window size of (3, 3). The output of this layer is an image of size $\frac{w}{2} * \frac{h}{2} * 16$.

3) Another Convolutional Layer with (3, 3) filter size and Rectified Linear unit (ReLu) activation function with 16 filters. The output of this layer is an image of size $\frac{w}{2} * \frac{h}{2} *16$.

4) Average pooling layer with a window size of (3, 3). The output of this layer is an image of size $\frac{w}{4} * \frac{h}{4} *16$.

5) A Spatial Pyramid Pooling Layer with (8, 8) blocks. The output is a vector of size 1024.

6) Fully Connected regression layer with single unit, and sigmoid activation function.

The proposed model is prototyped using TensorFlow 2.0 Python library. The network is trained for 50 epochs using Mean Squared Error (MSE) [22] as the loss function, and Adam [23] as the optimizer. MSE computes the mean of squared difference between the actual DMOS score (y_i) and predicted DOMS score ($\hat{y_i}$) of each image in a batch of images of the dataset.

$$MSE = \frac{1}{N}\sum_{i=0}^{N}(y_i - \hat{y_i})^2 \qquad (1)$$

Adam Optimizer receives the strengths of the RMSProp and AdaGrad and builds upon them to give a more improved gradient descent. The proportion of gradient descent is measured in such a way that there is smallest fluctuation when it reaches the global minimum while taking large enough steps so as to pass the local minima encountered.

TensorFlow [24] is a free and open-source software library for machine learning, based on dataflow architecture. The experiments are conducted in Colaboratory, or "Colab" for short, which is an open platform for research, offered by google. The GPU offered by the Colab facilitates to speed-up the training and testing process on the larger datasets with more accuracy.
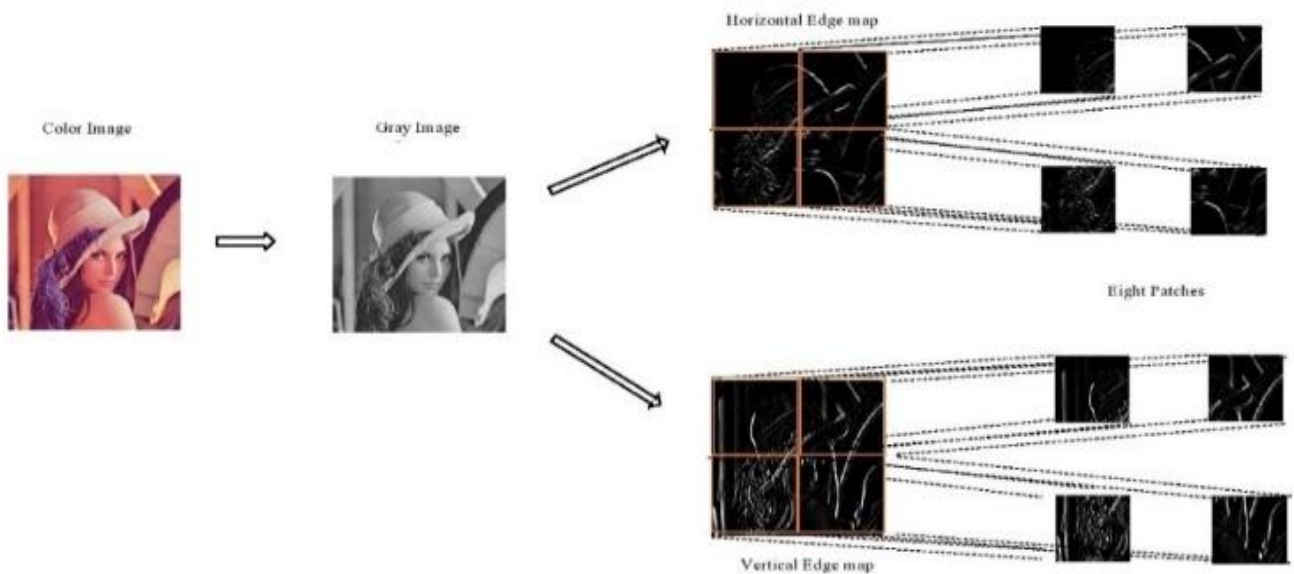

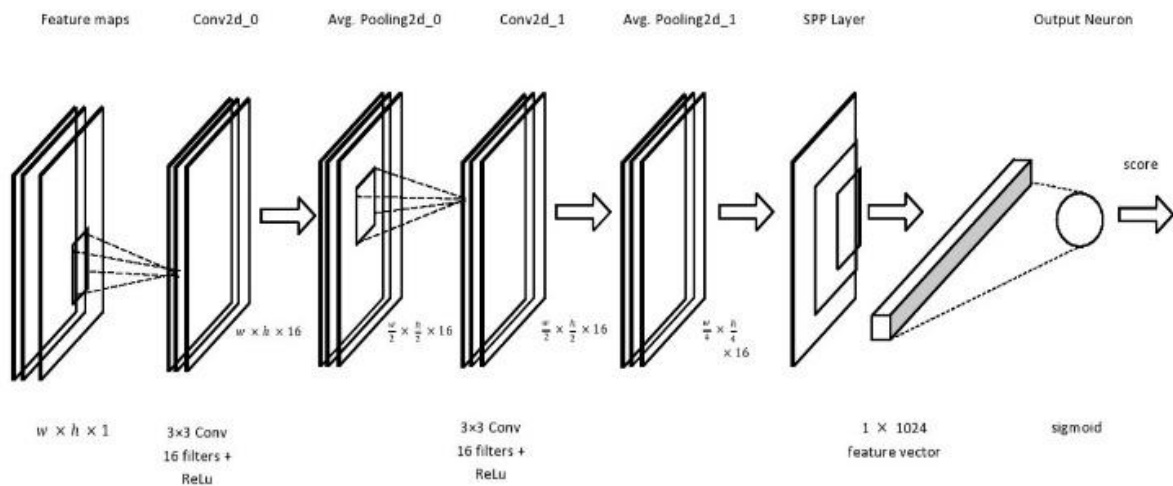**Figure 1.** Dataset preparation

Figure 2. Network architecture

3. EXPERIMENTS

The proposed method is applied on the following datasets.

3.1 Datasets

1) Laboratory for Image & Video Engineering Image Quality Assessment (LIVE IQA) [25] database: LIVE IQA consists of 29 reference images, and five types of distortions are applied with 7-8 degradation quantities. The distortion types include White Gaussian - WN, Gaussian Blur - GBLUR, JPEG compression - JPEG, and FastFading - FF. The total number of images in the dataset are 982 among which 799 are distortion images. For every image in LIVE IQA, a DMOS falling in the range [0, 100) is provided and a lower DMOS indicates higher quality. The original (a) and five levels of white noise distorted versions (b-f) of "coinsinfountain" image

of the LIVE IQA dataset are shown in Figure 8.

2) Computational and Subjective Image Quality (CSIQ) [26] database: It consists of 30 reference images, 866 distorted images corrupted by Additive White Gaussian Noise (AWGN), JP2K compression (JPEG2000), Blurring (BLUR), JPEG compression (JPEG), global contrast decrements (GCD) and additive pink Gaussian noise (APGN), with 4-5 levels for all distortion types. For every image in CSIQ, a DMOS score is provided between [0, 1], where a lower DMOS indicates higher quality. The present work considered only those distortions, which are common in both the datasets mentioned for fair comparison of results. The original (a) and five levels of JPEG distorted versions (b-f) of "sunsetcolor" image of the CSIQ dataset are shown in Figure 9.

3.2 Training and testing

Table 1. Number of train and test images in the independent validation framework

	WN		JP2K		GBLUR		JPEG		FF		All	
	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE	TR	TE
LIVE	140	34	182	45	140	34	187	46	140	34	786	196
CSIQ	144	36	144	36	144	36	144	36	-	-	504	126

Table 2. Number of train and test images in the cross-validation framework

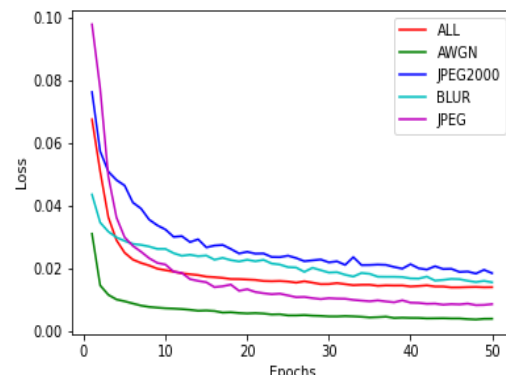|	WN		JP2K		GBLUR		JPEG		All	
	LIVE	CSIQ	LIVE	CSIQ	LIVE	CSIQ	LIVE	CSIQ	LIVE	CSIQ
TR/TE	174	180	227	180	174	180	233	180	808	630

Two different experimental frameworks were designed to evaluate the execution of the proposed model. In the first framework which is named as the independent validation framework, experiments were conducted considering all the distortions once and each individual distortion separately to obtain the performance metrics for each of the datasets mentioned in the above section. In each of the experiments, the proposed model was trained with 80% of the images under consideration in the dataset and was tested with the remaining 20% of the images. In the second experimental setup which is named as the cross validation framework, the proposed model was trained with all the images of one dataset and by testing with all the images in the other dataset for all the distortions once and for each of the individual distortions. Table 1 shows the details of number of images in the training phase (TR) and test phase (TE) for each of the datasets in the first framework of experiments. Table 2 shows the details for the second framework of experiments.
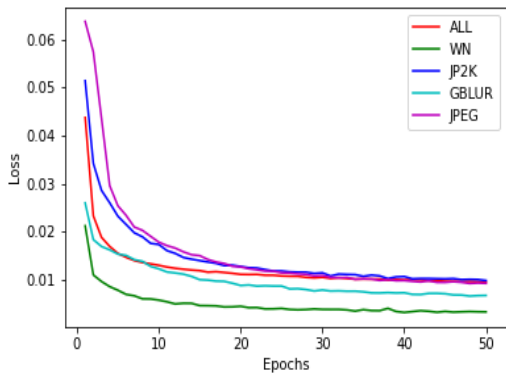
3.3 Results and discussion

The sub figures in Figure 3 show the plots of loss function during training phase for all the experiments conducted in the independent validation framework for the LIVE IQA dataset (a) and the CSIQ dataset (b). Sub figures (c) and (d) show similar plots for the experiments conducted in the cross-validation framework. The monotonic decrease of loss function in Eq. (1) during the training phase for different distortions and all distortions testify the generic capability of the IQA of the proposed model. The efficiency and computational economy of the model is further substantiated by the fact that 50 epochs only were used for training the model in both the experimental setups. The plots also show that the learning capability of the model is independent of the datasets and the distortions considered.
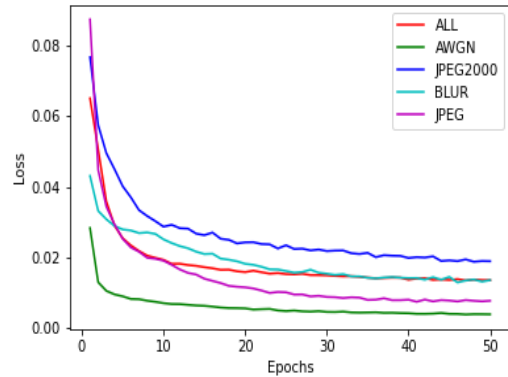

(a). Trained and tested with LIVEIQA

(b). Trained and tested with CSIQ

(c). Trained with LIVE IQA and tested with CSIQ   (d). Trained with CSIQ and tested with LIVE IQA
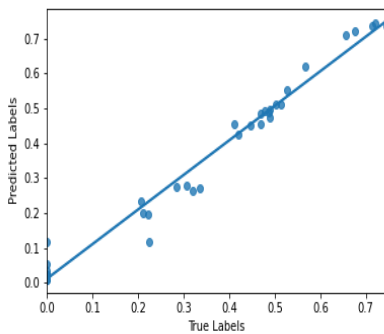
Figure 3. Independent and Cross validation loss (MSE) function plots in the training phase for all distortions and individual distortions present in LIVE IQA and CSIQ datasets

The scatter plots of predicted and actual DMOS scores during the testing phase in the independent validation framework of all distortions and individual distortions are shown in Figure 4 and Figure 5 for the LIVE IQA and CSIQ dataset respectively. The plots indicate clearly that the prediction capability of the model correlates very well with subjective assessment of image quality of distorted images. It can also be observed that the model is equally effective with all the distortions considered. This truth is further strengthened with the scatter plots of the predicted and actual DMOS scores in cross validation framework of all distortions and the individual distortions present in both the datasets mentioned, as shown in Figure 6 and Figure 7 respectively.
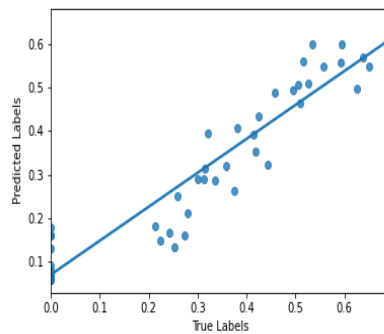
The performance of the model was compared to the other state-of-the-art models [9-14, 16] by computing SROCC and PLCC values between the predicted and actual image quality scores in the testing phase for all kinds of experiments conducted in the independent and cross validation frameworks using LIVE IQA and CSIQ datasets. The models considered for comparison include FR-IQA and NR-IQA models.
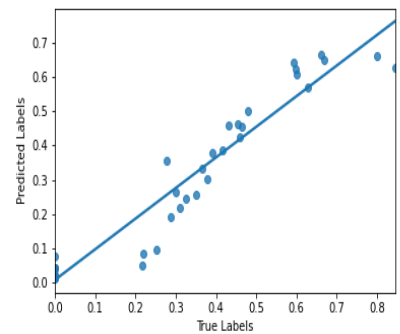
The Figure 8 shows the original (a) and five levels of white noise distorted versions (b-f) of "coinsinfountain" image of LIVE IQA dataset with actual and predicted DMOS scores at the top and bottom of each image, respectively. Figure 9 shows similar results for JPEG compressed "sunsetcolor" image of CSIQ dataset. Analogous results were obtained in the independent and cross validation of all distortions and individual distortions for both the datasets considered.
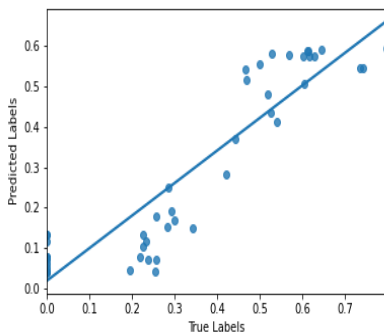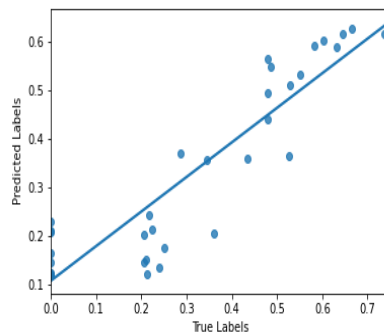

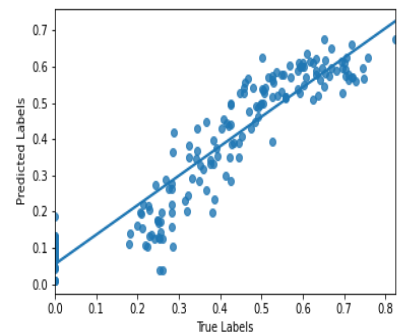(a) White noise   (b) Jp2k compression   (c) Gaussian blur

(d) Jpeg compression   (e) Fast fading   (f) All distortions

Figure 4. Scatter plots of predicted and actual DMOS scores for independent validation of all distortions and individual distortions of LIVE IQA (a-f)

(a) White noise (b) Jp2k compression (c) Blurring

(d) Jpeg compression (e) All distortions

Figure 5. Scatter plots of predicted and actual DMOS scores for independent validation of all distortions and individual distortions of CSIQ (a-e)

(a) White noise (b) Jp2k compression (c) Blurring

(d) Jpeg compression (e) All distortions

Figure 6. Scatter plots of predicted and actual DMOS scores for cross validation of all distortions and individual distortion with LIVE IQA as training dataset and CSIQ as test dataset (a-e)

(a) White noise   (b) Jp2k compression   (c) Blurring

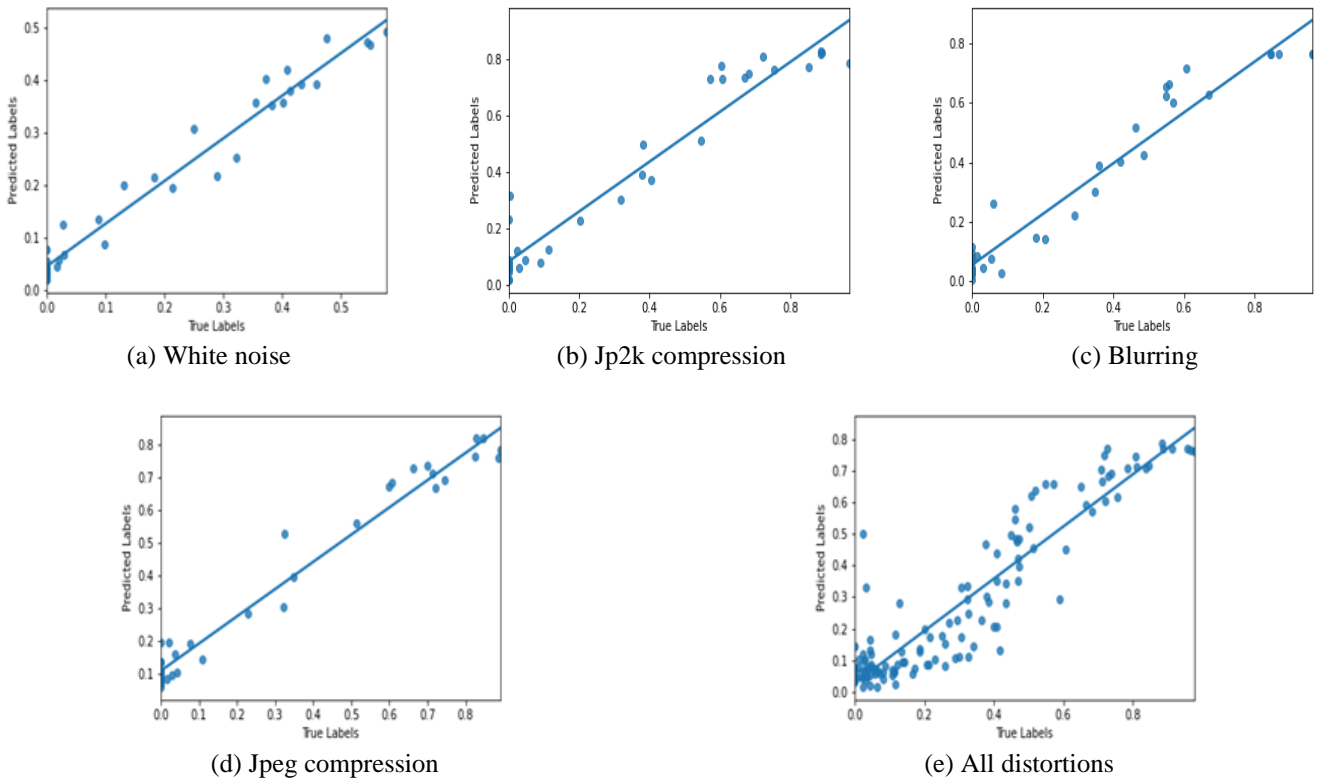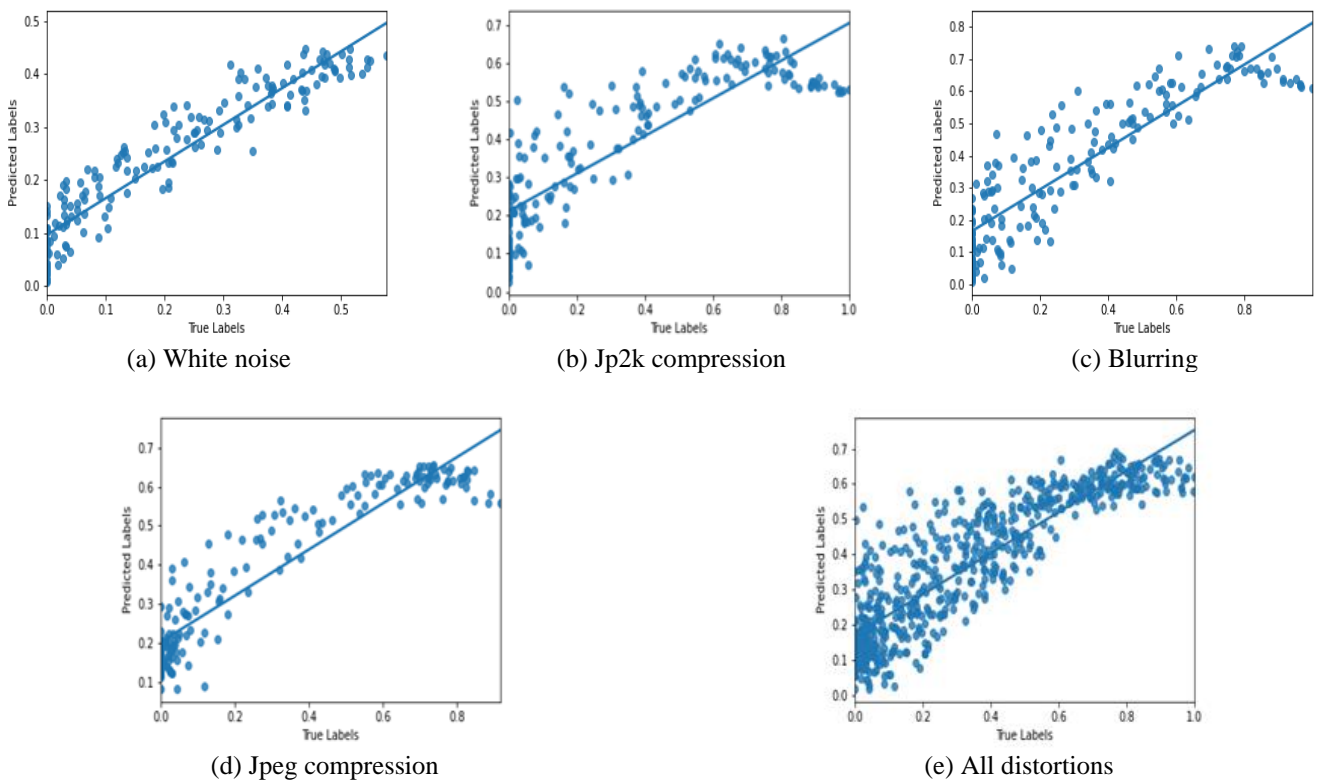(d) Jpeg compression   (e) All distortions
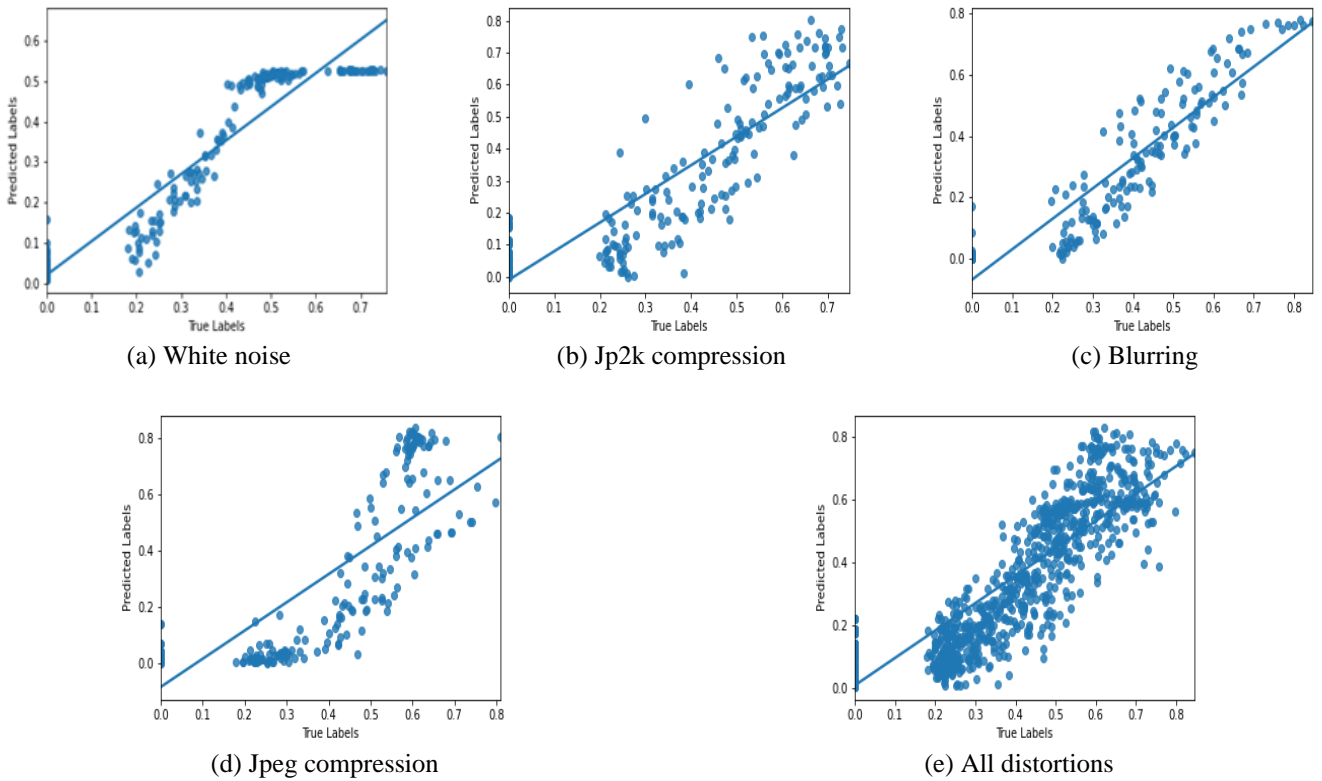
Figure 7. Scatter plots of predicted and actual DMOS scores for cross validation of all distortions and individual distortion with CSIQ as training dataset and LIVE IQA as test dataset

Original Image: WN (img146) DMOS = 0.0

WN (img8) with DMOS = 20.7114160011119

WN (img31) with DMOS = 27.4915734648744

Predicted DMOS = 0.03641214   (a)

Predicted DMOS = 0.2349273   (b)

Predicted DMOS = 0.3146033   (c)

WN (img124) with DMOS = 46.9103473254394

WN (img92) with DMOS = 51.3184224119651

WN (img24) with DMOS = 65.5358355849048

Predicted DMOS = 0.48280033   (d)

Predicted DMOS = 0.53661203   (e)

Predicted DMOS = 0.70794284   (f)

Figure 8. LIVE IQA original image, distorted images and their predicted DMOS scores by the model

Original Image: JPEG (sunsetcolor)
DMOS = 0.0

JPEG (sunsetcolor.JPEG.1) with DMOS
= 0.031

JPEG (sunsetcolor.JPEG.2) with DMOS
= 0.436

Predicted DMOS = 0.1930024
(a)

Predicted DMOS = 0.41772422
(b)

Predicted DMOS = 0.60261774
(c)

JPEG (sunsetcolor.JPEG.3) with DMOS
= 0.693

JPEG (sunsetcolor.JPEG.4) with DMOS
= 0.885

JPEG (sunsetcolor.JPEG.5) with DMOS
= 0.918

Predicted DMOS = 0.7051542
(d)

Predicted DMOS = 0.75876117
(e)

Predicted DMOS = 0.75876117
(f)

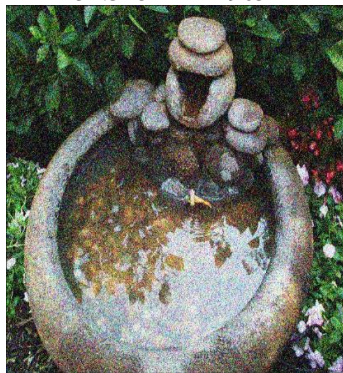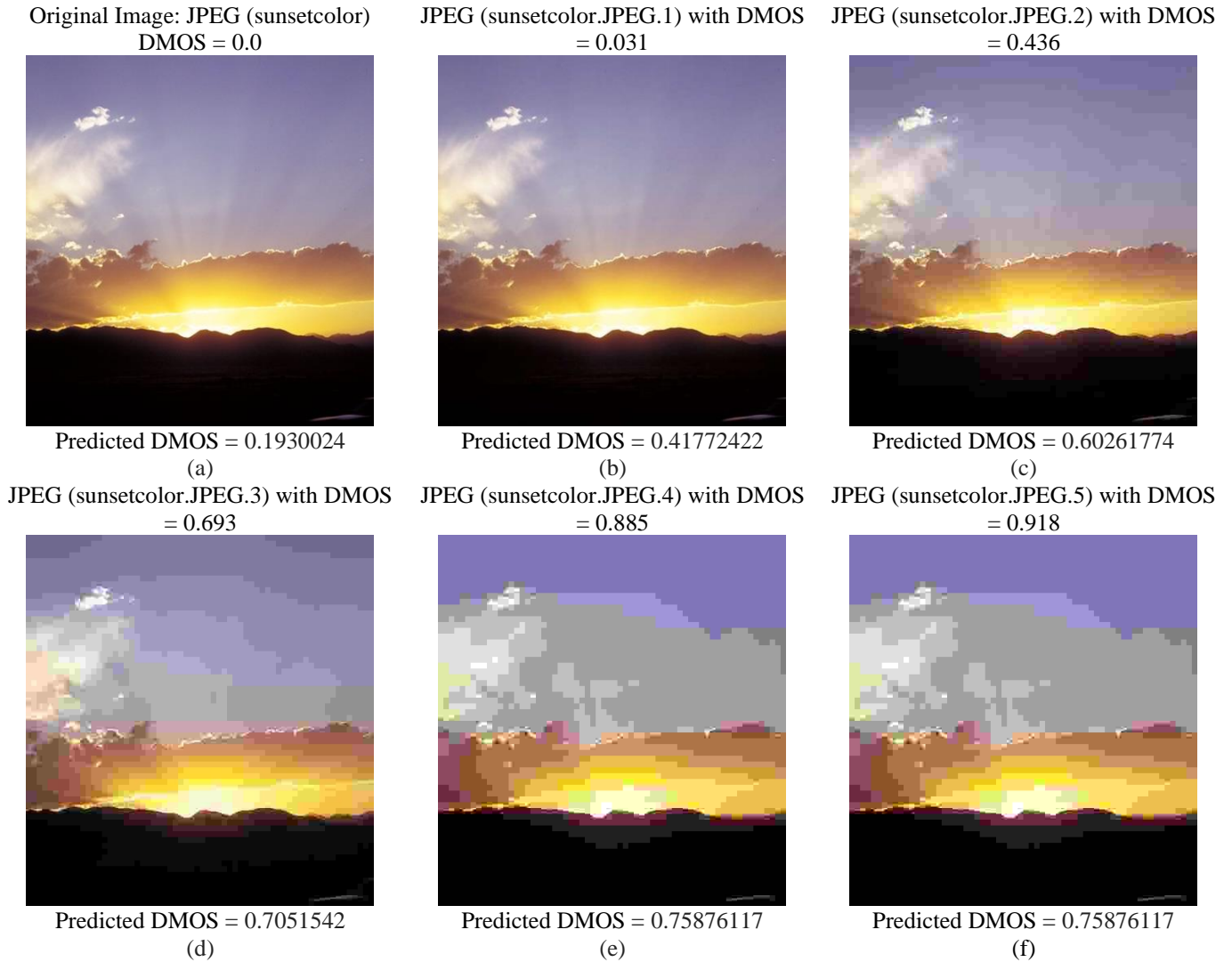**Figure 9.** CSIQ original image, distorted images and their predicted DMOS scores by the model

Table 3 and Table 4 compare the values of SROCC and PLCC respectively between the predicted and actual image quality scores computed with the proposed model in the testing phase with the other state-of-the-art models. The values are listed for all the experiments conducted with all and individual distortions using LIVE IQA dataset in the independent validation framework. The values show that the proposed model competes with the performance of the other models mentioned despite its simplicity and less complexity. Table 5 and Table 6 show similar results for the CSIQ dataset. The performance metrics of the proposed model in both cases generalize its capability of image quality assessment independent of the distortions and the datasets.

Table 3. SROCC for individual and all distortions in independent validation framework using LIVE IQA dataset. FR-IQA algorithms are mentioned in font Italic and others are NR-IQA algorithms

SROCC	WN	JP2K	GBLUR	JPEG	FF	ALL
PSNR	0.982	0.904	0.807	0.894	0.894	0.883
SSIM [27]	0.970	0.960	0.951	0.973	0.956	0.948
VIF [28]	0.984	0.968	0.971	0.982	0.962	0.963
DIVINE [10]	0.984	0.913	0.921	0.910	0.863	0.916
BLINDS-II SVM [12]	0.969	0.928	0.923	0.942	0.889	0.930
BLINDS-II Prob. [12]	0.978	0.950	0.943	0.941	0.862	0.920
BRISQUE [9]	0.978	0.913	0.951	0.964	0.876	0.939
CNN [16]	0.978	0.952	0.962	0.977	0.908	0.956
OG-IQA [13]	0.986	0.937	0.961	0.964	0.898	0.950
CurveletQA [11]	0.987	0.937	0.965	0.911	0.900	0.930
CORNIA [14]	0.976	0.943	0.969	0.955	0.906	0.942
Proposed model	0.981	0.959	0.970	0.894	0.886	0.946

Table 4. PLCC for individual and all distortions in independent validation framework using LIVE IQA dataset. FR-IQA algorithms are mentioned in font Italic and others are NR-IQA algorithms

PLCC	WN	JP2K	GBLUR	JPEG	FF	ALL
PSNR	0.982	0.885	0.803	0.878	0.892	0.864
SSIM [27]	0.986	0.971	0.955	0.981	0.962	0.946
VIF [28]	0.992	0.980	0.977	0.989	0.968	0.961
DIVINE [10]	0.988	0.922	0.923	0.921	0.888	0.917
BLINDS-II SVM [12]	0.979	0.934	0.938	0.967	0.895	0.930
BLINDS-II Prob. [12]	0.985	0.963	0.948	0.979	0.863	0.923
BRISQUE [9]	0.985	0.922	0.950	0.973	0.903	0.942
CNN [16]	0.984	0.953	0.953	0.981	0.933	0.953
OG-IQA [13]	0.990	0.945	0.967	0.982	0.911	0.952
CurveletQA [11]	0.985	0.946	0.969	0.928	0.918	0.932
CORNIA [14]	0.987	0.951	0.968	0.965	0.917	0.935
Proposed model	0.986	0.943	0.954	0.914	0.910	0.940

Table 5. SROCC for individual and all distortions in independent validation framework using CSIQ dataset. All are NR-IQA methods

SROCC	AWGN	JPEG2000	BLUR	JPEG	ALL
SFOSR [29]	0.918	0.923	0.877	0.938	0.887
PRLIQM-I [30]	0.856	0.872	0.886	0.887	0.863
PRLIQM-II [30]	0.868	0.884	0.902	0.901	0.872
Proposed model	0.962	0.923	0.922	0.911	0.872

Table 6. PLCC for individual and all distortions in independent validation framework using CSIQ dataset. All are NR-IQA methods

PLCC	AWGN	JPEG2000	BLUR	JPEG	ALL
SFOSR [29]	0.892	0.925	0.889	0.941	0.883
PRLIQM-I [30]	0.879	0.916	0.902	0.901	0.873
PRLIQM-II [30]	0.883	0.925	0.905	0.911	0.907
Proposed model	0.983	0.967	0.970	0.984	0.924

Table 7 presents the values of SROCC and PLCC between the predicted and actual image quality scores computed with the proposed model in the testing phase of cross validation framework. The values are listed for all the experiments conducted with all and individual distortions using LIVE IQA dataset for training CSIQA dataset for testing. Table 8 shows similar results where CSIQ dataset was applied for training and LIVE IQA dataset for testing. The performance metrics of the proposed model further consolidates the superior capability of the image quality assessment, which is independent of the distortions and the datasets.

Table 7. SROCC, and PLCC for individual and all distortions in cross validation framework using LIVE IQA for training and CSIQ testing

Proposed model	WN	JP2K	GBLUR	JPEG	ALL
SROCC	0.956	0.883	0.904	0.907	0.884
PLCC	0.940	0.854	0.891	0.929	0.877

Table 8. SROCC, and PLCC for individual and all distortions in cross validation framework using CSIQ for training and LIVE IQA testing

Proposed model	WN	JP2K	GBLUR	JPEG	ALL
SROCC	0.978	0.897	0.945	0.894	0.898
PLCC	0.933	0.880	0.914	0.817	0.869

4. CONCLUSIONS

We proposed a CNN for BIQA, which comes under the category of NR-IQA models. The proposed model uses the edge maps of the distorted images, which are computed using Prewitt kernel as an input to the CNN, which extracts the higher-level characteristics of the input image. The distortions in the high-level features are quantified as the image quality score by regression. The results conclude that the proposed model is generic and its capability of image quality estimation is independent of the datasets and the distortions present in the dataset. The performance metrics in terms of SROCC and PLCC prove that the model competes well with the state-of-the-art models.

REFERENCES

[1] Zhai, G.T., Min, X.K. (2020). Perceptual image quality assessment: A survey. Science China Information Sciences, 63(11): 211301. https://doi.org/10.1007/s11432-019-2757-1

[2] Rezazadeh, S., Coulombe, S. (2013). A novel discrete wavelet transform framework for full reference image quality assessment. Signal, Image, and Video Processing, 7(3): 559-573. https://doi.org/10.1007/s11760-011-0260-6

[3] Charrier, C., Lézoray, O., Lebrun, G. (2012). Machine learning to design full-reference image quality assessment algorithm. Signal Processing: Image

Communication, 27(3): 209-219. https://doi.org/10.1016/j.image.2012.01.002

[4] Tong, Y., Konik, H., Cheikh, F.A., Tremeau, A. (2010). Full reference image quality assessment based on saliency map analysis. Journal of Imaging Science and Technology, 54(3): 30503-1. https://doi.org/10.2352/j.imagingsci.technol.2010.54.3.030

[5] Tao, D., Li, X., Lu, W., Gao, X. (2009). Reduced-reference IQA in contourlet domain. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 39(6): 1623-1627. https://doi.org/10.1109/TSMCB.2009.2021951

[6] Wang, Z., Simoncelli, E.P. (2005). Reduced-reference image quality assessment using a wavelet-domain natural image statistic model. Human Vision and Electronic Imaging X, 5666: 149-159. https://doi.org/10.1117/12.597306

[7] Ma, L., Li, S., Zhang, F., Ngan, K.N. (2011). Reduced-reference image quality assessment using reorganized DCT-based image representation. IEEE Transactions on Multimedia, 13(4): 824-829. https://doi.org/10.1109/tmm.2011.2109701

[8] Xue, W., Mou, X. (2010). Reduced reference image quality assessment based on Weibull statistics. Second International Workshop on Quality of Multimedia Experience (QoMEX), pp. 1-6. https://doi.org/10.1109/QOMEX.2010.5518131

[9] Mittal, A., Moorthy, A.K., Bovik, A.C. (2012). No-reference image quality assessment in the spatial domain. IEEE Transactions on Image Processing, 21(12): 4695-4708. https://doi.org/10.1109/TIP.2012.2214050

[10] Moorthy, A.K., Bovik, A.C. (2011). Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE Transactions on Image Processing, 20(12): 3350-3364. https://doi.org/10.1109/TIP.2011.2147325

[11] Liu, L., Dong, H., Huang, H., Bovik, A.C. (2014). No-reference image quality assessment in curvelet domain. Signal Processing: Image Communication, 29(4): 494-505. https://doi.org/10.1016/j.image.2014.02.004

[12] Saad, M.A., Bovik, A.C., Charrier, C. (2012). Blind image quality assessment: A natural scene statistics approach in the DCT domain. IEEE Transactions on Image Processing, 21(8): 3339-3352. https://doi.org/10.1109/tip.2012.2191563

[13] Liu, L., Hua, Y., Zhao, Q., Huang, H., Bovik, A.C. (2016). Blind image quality assessment by relative gradient statistics and adaboosting neural network. Signal Processing: Image Communication, 40: 1-15. https://doi.org/10.1016/j.image.2015.10.005

[14] Ye, P., Kumar, J., Kang, L., Doermann, D. (2012). Unsupervised feature learning framework for no-reference image quality assessment. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1098-1105. https://doi.org/10.1109/CVPR.2012.6247789

[15] Liu, X., De Weijer, J.V., Bagdanov, A.D. (2017). RankIQA: Learning from rankings for no-reference image quality assessment. IEEE International Conference on Computer Vision (ICCV), pp. 1040-1049. https://doi.org/10.1109/iccv.2017.118

[16] Kang, L., Ye, P., Li, Y., Doermann, D. (2014). Convolutional neural networks for no-reference image quality assessment. IEEE Conference on Computer Vision and Pattern Recognition, pp. 1733-1740. https://doi.org/10.1109/cvpr.2014.224

[17] Bosse, S., Maniry, D., Muller, K.R., Wiegand, T., Samek, W. (2018). Deep neural networks for no-reference and full-reference image quality assessment. IEEE Transactions on Image Processing, 27(1): 206-219. https://doi.org/10.1109/TIP.2017.2760518

[18] Bianco, S., Celona, L., Napoletano, P., Schettini, R. (2017). On the use of deep learning for blind image quality assessment. Signal, Image and Video Processing, 12(2): 355-362. https://doi.org/10.1007/s11760-017-1166-8

[19] Li, Y., Ye, X., Li, Y. (2017). Image quality assessment using deep convolutional networks. AIP Advances, 7(12): 125324. https://doi.org/10.1063/1.5010804

[20] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778. https://doi.org/10.1109/cvpr.2016.90

[21] Zhang, L., Zhang, L., Mou, X.Q., Zhang, D. (2011). FSIM: A feature similarity index for image quality assessment. IEEE Transactions on Image Processing, 20(8): 2378-2386. https://doi.org/10.1109/TIP.2011.2109730

[22] Botchkarev, A. (2019). A new typology design of performance metrics to measure errors in machine learning regression algorithms. Interdisciplinary Journal of Information, Knowledge, and Management, 14: 045-079. https://doi.org/10.28945/4184

[23] Kingma, D.P., Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint, arXiv:1412.6980.

[24] Abadi, M., Agarwal, A., Barham, P. et al. (2016). Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467.

[25] Ghadiyaram, D., Bovik, A.C. (2016). Massive online crowdsourced study of subjective and objective picture quality. IEEE Transactions on Image Processing, 25(1): 372-387. https://doi.org/10.1109/TIP.2015.2500021

[26] Larson, E.C., Chandler, D.M. (2010). Most apparent distortion: Full-reference image quality assessment and the role of strategy. Journal of Electronic Imaging, 19(1): 011006. https://doi.org/10.1117/1.3267105

[27] Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4): 600-612. https://doi.org/10.1109/TIP.2003.819861

[28] Sheikh, H.R., Bovik, A.C. (2006). Image information and visual quality. IEEE Transactions on Image Processing, 15(2): 430-444. https://doi.org/10.1109/TIP.2005.859378

[29] Feng, T., Deng, D., Yan, J., Zhang, W., Shi, W., Zou, L. (2016). Sparse representation of salient regions for no-reference image quality assessment. International Journal of Advanced Robotic Systems, 13(5): 1-11.

[30] Xu, L., Li, J., Lin, W.S., Zhang, Y., Zhang, Y.B., Yan, Y.H. (2016). Pairwise comparison and rank learning for image quality assessment. Displays, 44: 21-26. https://doi.org/10.1016/j.displa.2016.06.002