

An Edge Distortion and CNN-Based Analysis of Blind IQ

Movva. Rajesh Babu*, Kontham. Raja Kumar

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

email: mrb.csebec@gmail.com, rajakumar72@gmail.com*

Article Info

Page Number: 177 – 191

Publication Issue:

Vol 71 No. 2 (2022)

Abstract

This paper is for assessing the image quality (IQ) without using an authentic image (original image) which is a type of Blind IQ Assessment (BIQA) model by introducing a technique of Convolutional Neural Network (CNN). The distortions of edges in the image are considered as features to represent the image feature vector. This approach is justified by the evidence that the subjective evaluation concentrates on image characteristics that radiate from the boundaries and edges that exist within the image. It was identified in the prior methods that the features are extracted at the time of training or before training by applying sophisticated transformations on the image. In this work, the vertical along with horizontal edge feature maps of the training images are extracted by means of Scharr Kernel (SK) approach. These edge maps subsequently fed into a CNN, which uses non-linear transformations to bring out higher-level features. Regression is then used to link the generated features to the IQ score. To accommodate different sizes of input images, the SPP (Spatial Pyramid Pooling) layer is used in this network. The developed model was evaluated using well-known datasets in the field of IQA. The suggested model's performance reveals that it outperforms previously existing models in context of negligible complexity involvement and feature extraction simplicity.

Article History

Article Received: 30 December 2021

Revised: 28 January 2022

Accepted: 12 March 2022

Publication: 31 March 2022

Keywords: edge distortion, feature extraction, image quality assessment.

1. INTRODUCTION

A picture is equal to thousand words, is a popularly known fact. Image sharing and communication is an important component of modern life, particularly in the context of social network platforms. It's also a fact that low-bandwidth image capture, processing, and transmission are prone to a variety of distortions. The approximation of IQ is critical in the progression of numerous applications of image processing (IP). Because human picture quality judgment cannot be used everywhere, automatic IQ estimate become a well-known study topic in IP as well as computer vision (CV) [1]. Spontaneous assessment of IQ for distorted images (DI) can be labelled into several categories of Image Quality Assessment (IQA) techniques, relative on the evidence available about the authentic image. These include BIQA, also known as No-Reference IQA (NR-IQA), Reduced-Reference IQA (RR-IQA), and Full Reference IQA (FR-IQA). The advancement of CNN resulted in a new BIQA methodology. A CNN is developed for estimating the IQ in BIQA and is trained with a subset of DI and their accompanying scores of IQ. The scores of IQ are usually the DMOS (Difference Mean Opinion Scores), which are derived from human assessment of IQ such as bad, poor or good on distorted images.

2. LITERATURE REVIEW

To assess IQ, the models of FR-IQA require the authentic image [2-7]. For describing the quality of its deformed counterpart, the models of RR-IQA need access of partial information from the authentic image [8-11]. To measure the quality of altered version of an image, BIQA models require no evidence about the authentic image. These models are broadly categorized into three types such as Conventional models, Machine Learning (ML) models and CNN premised Deep Learning (CNNDL) models. Conventional models depend on demographics of native scene. ML models train the models by considering features of authentic in addition distorted forms of image and hand-crafted features. With the assistance of previous layers of a convolutional network, CNNDL models are capable to abstract representational features from an image and link these features a head of fully connected layers with quality scores in the time of training.

Mittal et al [12] proposed a BRISQUE model for assessing the special quality of blind image. This model measures the spontaneity costs that exist in the authentic and altered image's which supports locally normalized coefficients of luminance as quality of an image. The forecast model obtained better enhancements than the popular paradigms. Moorthy and Bovik [13] established the BIQA model as a two-stage outline for clustering transformation function in Intra together with Inter sub-bands of orientation and variable scale by characterizing them with a Stochastic Scale Binary mixture model. Liu et al [14] suggested a model created on curvelets to direct the log-histograms maxima of the curvelet quantity values and to bring out the scales from strength distribution of both orientation and scale exists in the curvelet field. Among a variety of non-natural categories, the generated features out-of the curvelet filed are particularly suited to quality of real image. For linear categorization of IQ using human subjective assessments, a SVM (Support Vector Machine) classifier is used. Saad et al [15] created a model of NR-IQA to extract features originating from the image discrete cosine transform measurements based genuine scene statistics paradigm. To forecast quality scores of images, the obtained features are applied in the experiment of a simple Bayesian inference approach. Yi Hua et al [16] proposed an oriented gradients IQ evaluation model that uses a back propagation neural network with Ada boosting to link the retrieved image features of relative gradient magnitude of IQ. The experiment's predicted accuracy outperforms the performance of well-known models. Ye P et al [17] developed a structure of a NR-IQA that follows unsupervised method to realize an image characteristics vocabulary by using raw image parts that are segregated from a subset of un-labeled images. To obtain features to reflect the estimation of IQ, the authors used a sensitive task encoding with max pooling.

Xialei Liu et al [18] suggested a method of NR-IQA that uses Siamese Network (SN) and synthetically created datasets to learn images ranking. The trained SN exchanges the knowledge information with a CNN, which is then trained for fine tuning the transferred weights using a group of images that are drawn from subset of images. Throughout the testing phase, thirty sub-images of the deformed image are randomly sampled, and the IQ is calculated based on the median value of those scores. Le Kang et al [19] built a CNN that uses varying window sizes to identify features from image patches where pixel intensities are locally normalized. To associate the learnt features along with the IQ score, two fully connected (FC) layers are employed in aggregation with an output node. The expected patch scores are applied to quantify the intensity of IQ score by considering the given image. Bosse et al [20] presented an end-to-end CNN approach for evaluating the IQ. This network contains a variety of layers, ten convolutional, five pooling, and two FC layers for feature abstraction

as well as regression. The combined optimization of biased average patch growth was involved in the approach for local pooling patch potentials to global IQ. Bianco et al [21] experimented with a network architecture of Caffe with an extractor for features on top of which the retrieved features were mapped to subjective quality scores using a linear kernel SVR algorithm. According to Yezhou Li et al [22], the developed approach for calculation of IQ which is built on an architecture of ResNet [23] claims that the activation function of ReLU permits non-linear modifications for separating higher level image characteristics, resulting in more reliable IQ amount than linear filters.

The developed work is planned in 5 sections. The inspiration to the work is presented in section 3 as background. The proposed technique is defined in section 4, which illustrates the input to the CNN and a CNN architecture is suggested for extracting features. The details of training, testing along with outcomes are presented in section 5. Section 6 mentions the conclusions of this work.

3. BACKGROUND

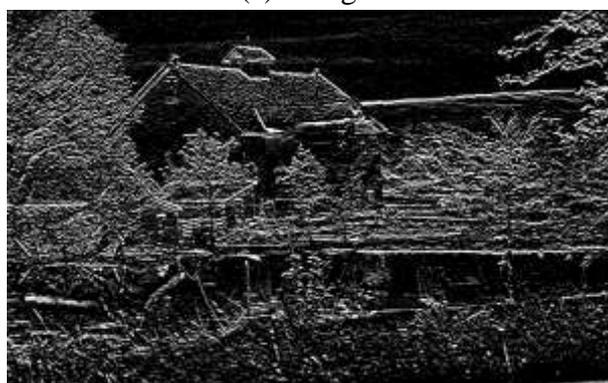
To extract the image feature maps for training, the large number of the approaches in the existing works used complex transformations at low level. The researchers are inspired with these observations to arrive a method for extracting image features at higher-level to determine the IQ [24]. The authors devised a method that identifies higher level image features to measure the IQ. The developed method identifies the vertical and horizontal edge feature maps of images from the training database and utilizes them to train a CNN to estimate IQ. The sub figures in Fig. 1 shows an input image (a), Gray Scale image (b) and its horizontal and vertical gradients (c) & (d) respectively, extracted using Scharr operators.



(a). Image.



(b). Gray Scale image.



(c). Horizontal gradient.



(d). Vertical gradient.

Fig. 1. Feature Extraction.

4. PROPOSED APPROACH

Let X is the DI in the dataset, and y be the corresponding DMOS of the DI that exist in the dataset. Normally, the y value varies from 0 to 100. A 0 value of y demonstrates that the image has not been distorted in any way and indicates that X is the authentic image. A rating of 100, on the other hand, implies that the image is fully distorted. However, the interpretations of the standards (0 to 1, 0 to 9) are differing from one dataset to the next. In the present work, we normalized the DMOS values into the range 0 to 1. Zero representing the highest quality. A CNN is trained in BIQA using (X, y) pairs that are present in the dataset. The image's representational feature maps are extracted by the first convolutional layers, which are then mapped to the IQ score y by the next FC layers.

A network which is trained like this is used to calculate the quality of unseen DI image X' . When fed with a huge number of DI's and authentic images, a CNN can successfully learn the representational image features. A significant issue in the field of estimating the IQ is the datasets are typically tiny, on average including less than few thousand images. The challenge is effectively solved in the current work by using the edge detector method of Scharr [25, 26] to acquire the vertical and horizontal edge maps in every image and feeding them to the artificial neural network.

The modified input allows extraction of higher level features of image more quickly from a smaller dataset in terms of size by make use of the convolutional layers. Furthermore, the feature maps are divided into patches, which augment the input. Every feature edge map of image X is separated to words four patches, and the score y of image is ascertained to every feature map patch to complete the augmentation process. A major design consideration is how many patches a feature map should be divided into. Due to the large count of patches, feature linkages will be lost, resulting in contextually unconnected patches. If the number is too little, the objective of the enlargement may not be accomplished. During perception, humans usually concentrate as to four quadrants of an image. Therefore, in the current study, each feature edge map of image X has been divided into 4 patches and each feature map patch is assigned the score y .

a. INPUT TO NETWORK

The following steps are used to construct the network's input set from the authentic dataset. In the first step, the Scharr kernel is applied to transform the black-and-white version of every image into vertical edge feature map X^v and horizontal edge feature map X^h . The Scharr kernel is a gradient-based kernel that learns the degree and direction of an image. Kernel Scharr does so by estimating the image's gradient intensity function. The kernels are:

$$G_x = \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ +3 & +10 & +3 \end{bmatrix} \text{ and } G_y = \begin{bmatrix} -3 & 0 & +3 \\ -10 & 0 & +10 \\ -3 & 0 & +3 \end{bmatrix}$$

In the next step, edge feature maps are separated into four same-size portions that correspond to the four quadrants of the plane. To accomplish this, for every X image of size (w, h) , generate eight edge feature maps of the same size $(w/4, h/4)$. The score (DMOS) y image is used as a label for every patch. As a result, the training data size is improved by eight times. The construction of the dataset for the input to the CNN network is depicted in Fig. 2. The horizontal feature edge map of Fig. 1 (c) is sub-divided against four patches as given in sub figures (a) to (d) of Fig. 2. The same results of Fig. 1 (d) are given in sub figures (e)-(h).

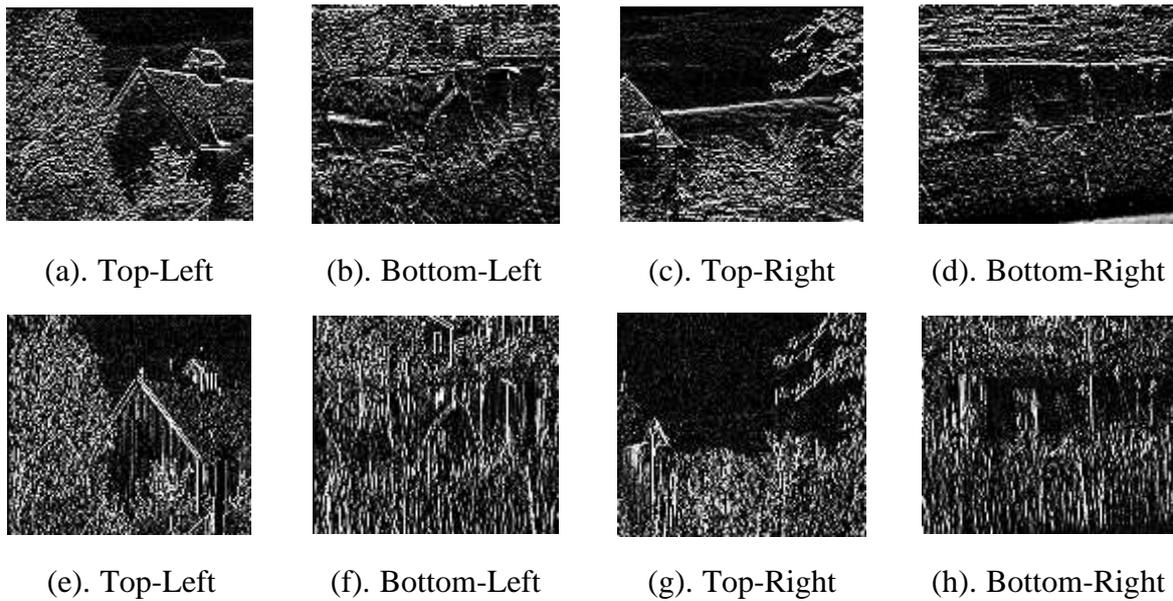


Fig. 2. Preparation of Dataset: Horizontal edge maps (a)-(d). Vertical edge maps (e)-(h).

b. CNN NETWORK ARCHITECTURE

Let $w * h$ designate the image's width and height when it is sent to the CNN network. As illustrated in Fig. 3, the anticipated CNN consists of the subsequent convolutional as well as fully connected layers. The first convolutional layer comprises of 16 filters with a filter size of (3, 3) and ReLU (Rectified Linear Unit) activation function. The indicated layer with padding produces an output with dimensions $(w * h * 16)$. The output is then passed through a layer of average pooling with (3, 3) window size whose output is $(w/2 * h/2 * 16)$. The second convolutional layer comprises of 16 filters with a window size of (3, 3) and ReLU as activation function. This layer produces an output of dimensions of $(w/2 * h/2 * 16)$. Another layer of average pooling with (3, 3) window size follows in the network, which results an output with $(w/4 * h/4 * 16)$ dimensions. An (8, 8) block SPP [20] layer reduces the input as vector size of 1024. Finally, with sigmoid activation function, a single-unit regression layer is used.

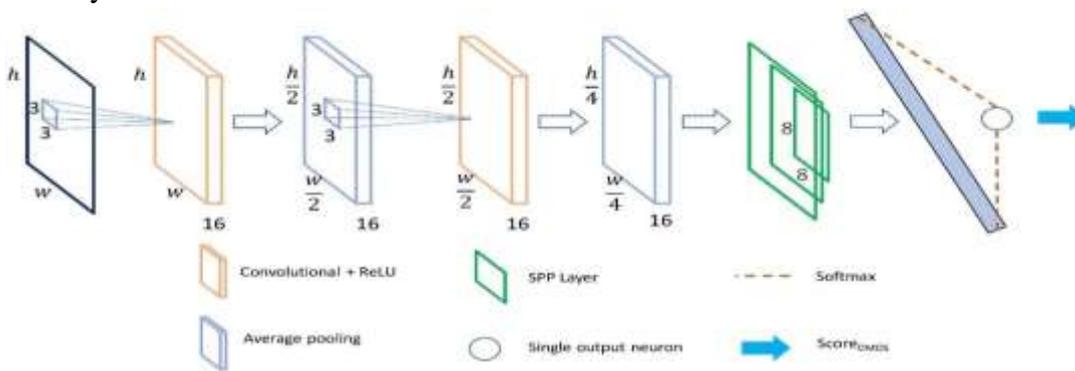


Fig. 3. Architecture of a Network

The TensorFlow 2.0 python library is employed to prototype the developed model. The CNN network is trained by using Adam [27] as the optimizer and loss function as Mean Squared Error [28] (MSE) considering 50 epochs. TensorFlow [29] is dataflow architecture-based machine learning software library that is open-source and free. The experiments are carried out in Google's Colaboratory or in short “Colab”, which is a platform for doing open research. The GPU provided in the Colab helps to

increase the speed in the process of the larger datasets training along with testing, as an outcome, the accuracy is improved. MSE calculates the average squared difference connecting each image's true and expected DOMS scores in a subset of images from the dataset.

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

The RMSProp and AdaGrad strengths are coupled in Adam Optimizer to provide a more stepped forward gradient descent. The amount of gradient succession is adjusted in such a way that there is minimum level of variation as soon as it spreads the global smallest even though captivating large adequate stages to prevent the premature convergence.

5. EXPERIMENTS

In this section, we present data used in training the model, testing the model and the results. In this work, the developed method is experimented with the two benchmark datasets TID2008 as well as LIVE-IQA.

a. Datasets

- i. TID2008 (Tampered Image Dataset 2008) database [30]: The TID2008 database comprises 25 authentic photos and 1700 distorted images divided into 17 distortion categories. In our tests, we considered the four prevalent distortions which are commonly present in the LIVE dataset, specifically JP2k, JPEG, WN, and BLUR. Each image does have a MOS (Mean Opinion Score) that varies from 0 to 9. Greater MOS, signifies higher quality. Fig. 6 demonstrates the authentic (a) and four variations of Gaussian Blur distorted forms (b-e) of the TID2008 dataset's "I09" image.
- ii. LIVE-IQA (Laboratory for Image and Video Engineering IQA) database [31]: LIVE-IQA is made up of 29 authentic images to which five varieties of distortions such as FastFading (FF), JPEG compression (JPEG), Gaussian Blur (GBLUR) and White Gaussian (WN) with 7-8 degradation levels are added. The dataset contains 982 images, with 799 of them being distorted images. Fig. 6 illustrates the concept of DMOS, which is a value between 0 and 100 that is assigned to every image in LIVE-IQA, smaller DMOS specify higher IQ. The authentic (f) as well as five grads of GBlur alternate forms (g-k) of LIVE-IQA dataset the image "churchandcapitol" have been exhibited in Fig. 6.

In the course of training as well as testing, the scores assigned to every image in the above stated datasets are normalized to a common scale of 0 to 1. For an impartial judgement of outcomes, the present work only looked at those distortions that are present in both datasets. To measure the performance of the anticipated model, two distinct experimental frameworks named as Independent Validation Framework (IVF) and Cross Validation Framework (CVF) were designed.

b. Training, Testing and Results using IVF

In IVF, the experiments were done by using each distortion separately and all the distortions combined to acquire performance measures for every dataset stated in the preceding section. The developed model used 80% of dataset during training and 20% during testing in each experiment. The images count that are employed in the TR (training) and TE (test) phases of each experiment for the couple of the datasets using the IVF method are given in Table 1.

Table 1. Count of test and train images in the IVF

	ALL		JPEG2000			GBLUR			JPEG			FF			AGN		
	TRAI	TEST	TRAI	TEST	TRAI	TEST	TRAI	TEST	TRAI	TEST	TRAI	TEST	TRAI	TEST			
TID2008	340	85	100	25	100	25	100	25	-	-	100	25					
LIVE-IQA	768	196	182	45	140	34	187	46	140	34	140	34					

The sub-figures of Fig. 4 display the loss function plots throughout the training stage in all of the trials done in IVF for the TID2008 (a) dataset as well as the LIVE-IQA (b) dataset respectively. At the training stage of IVF, the loss function used is MSE as defined in equation (1).

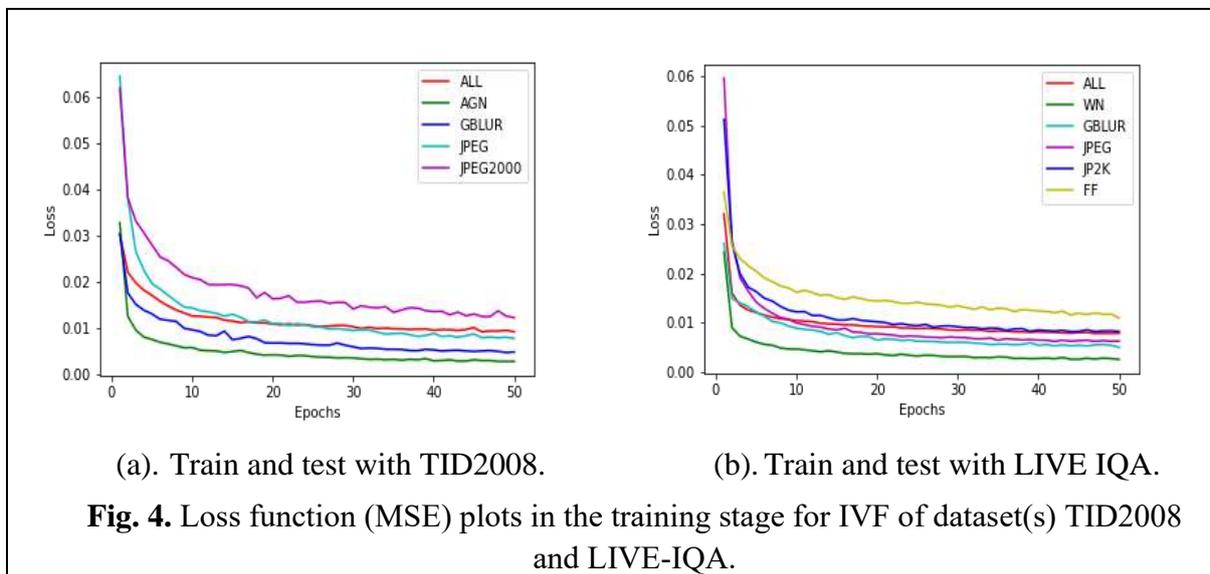
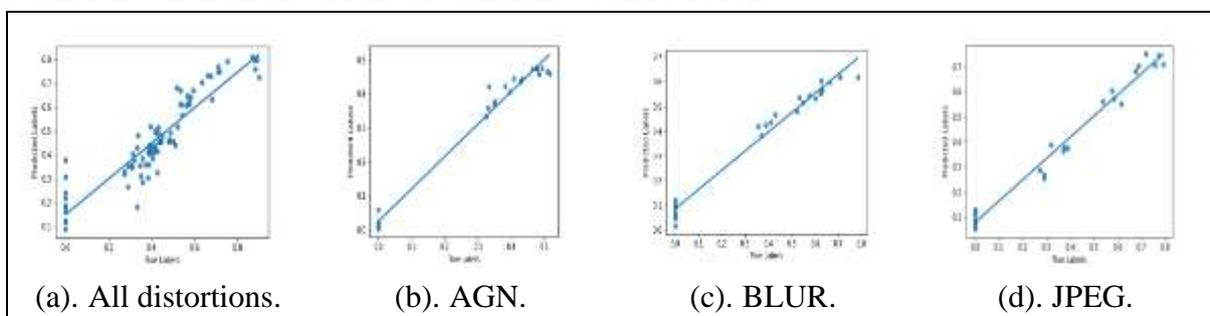
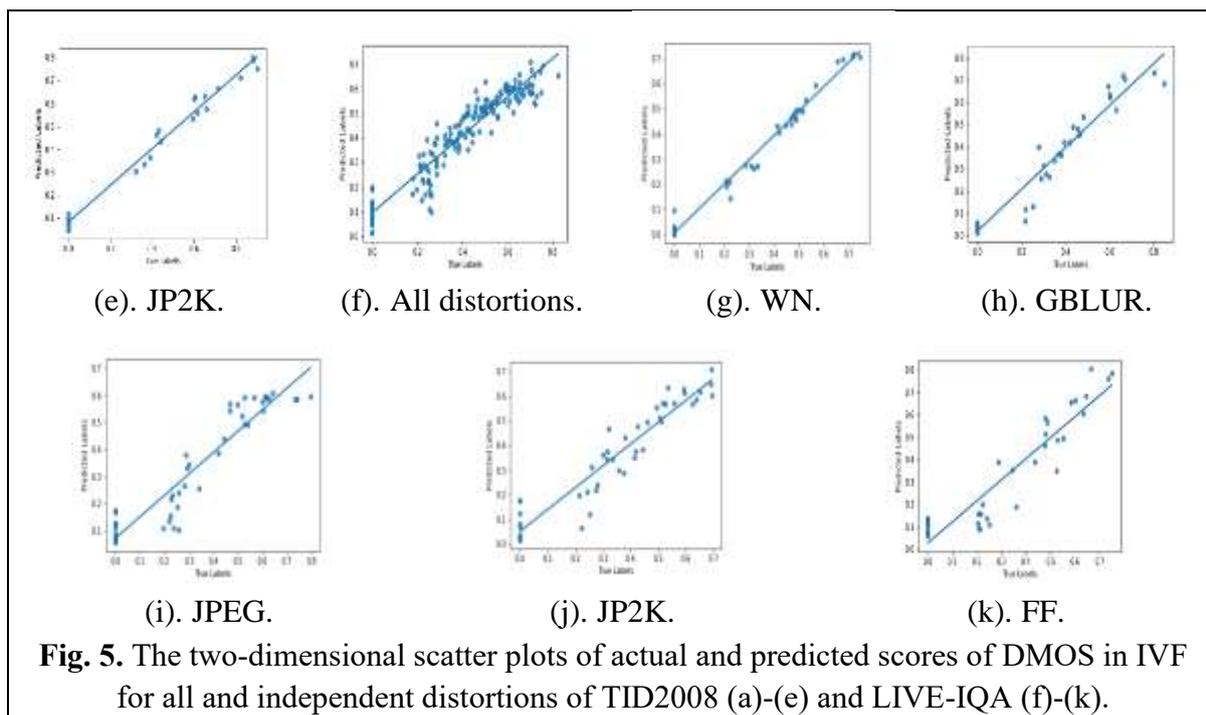


Fig. 5 displays scatter graph of true and predicted scores of DMOS in the testing stage in the IVF of all images together along with individual distortions for the datasets of TID2008 (a)-(e) and LIVE-IQA (f)-(k). The plots obviously exhibit that the model's assessments ability matches very well with subjective images quality assessments of DIs. It is also worth mentioning that the developed model is as efficient when all distortions are taken into consideration.





The PLCC and SROCC values between the real and predicted quality scores of images in the testing stage for all types of experimentations that are led in the IVF using datasets of TID2008 along with LIVE-IQA are calculated. The model's performance was compared to that of other well-known models [5-7, 12-19]. The NR-IQA and FR-IQA models were used as comparison models.

Table 2. SROCC for all and individual alterations in IVF with TID2008 dataset. Font *Italic* indicates FR-IQA algorithms, and others are BIQA algorithms

SROCC	ALL	AGN	BLUR	JPEG	JPEG2000
VSI [7]	0.898	0.922	0.952	0.961	0.984
BRISQUE [12]	0.896	0.829	0.881	0.924	0.832
DIIVINE [13]	0.889	0.851	0.862	0.866	0.924
CurveletQA [14]	0.867	0.858	0.882	0.864	0.549
CORNIA [17]	0.813	0.913	0.932	0.929	0.919
Our Model	0.934	0.942	0.973	0.964	0.969

Table 3. PLCC for all and individual alterations in IVF with TID2008 dataset

PLCC	ALL	AGN	BLUR	JPEG	JPEG2000
					0
CORNIA [17]	0.837	0.911	0.932	0.963	0.929
Our Model	0.915	0.990	0.988	0.989	0.991

Table 2 and 3 display the comparison of SROCC (Spearman Rank Ordered Correlation Coefficient) and PLCC (Pearson’s Linear Correlation Coefficient) values between the true and predicted IQ scores using dataset TID2008 that are calculated with the developed model in comparison to other well-known models during the testing phase. The values are presented for all experiments with individual and all distortions using the dataset of TID2008. Despite its simplicity and lack of complexity, the

proposed model shows best performance than other well-known models that are described. The LIVE-IQA dataset yields comparable results in Table 4 and 5.

Table 4. SROCC for all and individual alterations in IVF with LIVE-IQA. Font Italic indicates FR-IQA algorithms, and others are BIQA algorithms

SROCC	ALL	WN	GBLUR	JPEG	JP2K	FF
<i>PSNR</i>	0.883	0.982	0.807	0.894	0.904	0.894
<i>VIF</i> [5]	0.963	0.984	0.971	0.982	0.968	0.962
<i>SSIM</i> [6]	0.948	0.970	0.951	0.973	0.960	0.956
<i>VSI</i> [7]	0.952	0.983	0.952	0.976	0.960	0.943
<i>BRISQUE</i> [12]	0.939	0.978	0.951	0.964	0.913	0.876
<i>DIVINE</i> [13]	0.916	0.984	0.921	0.910	0.913	0.863
<i>CurveletQA</i> [14]	0.930	0.987	0.965	0.911	0.937	0.900
<i>BLINDS-II_{Prob.}</i> [15]	0.920	0.978	0.943	0.941	0.950	0.862
<i>BLINDS-II_{SVM}</i> [15]	0.930	0.969	0.923	0.942	0.928	0.889
<i>OG-IQA</i> [16]	0.950	0.986	0.961	0.964	0.937	0.898
<i>CORNIA</i> [17]	0.942	0.976	0.969	0.955	0.943	0.906
<i>RankIQA+FT</i> [18]	0.981	0.991	0.988	0.978	0.970	0.954
<i>CNN</i> [19]	0.956	0.978	0.962	0.977	0.952	0.908
Our Model	0.954	0.981	0.975	0.935	0.959	0.918

Table 5. PLCC for all together with individual distortions in IVF with LIVE-IQA. Font Italic indicates FR-IQA algorithms, and others are BIQA algorithms

PLCC	ALL	WN	GBLUR	JPEG	JP2K	FF
<i>PSNR</i>	0.864	0.982	0.803	0.878	0.885	0.892
<i>VIF</i> [5]	0.961	0.992	0.977	0.989	0.980	0.968
<i>SSIM</i> [6]	0.946	0.986	0.955	0.981	0.971	0.962
<i>BRISQUE</i> [12]	0.942	0.985	0.950	0.973	0.922	0.903
<i>DIVINE</i> [13]	0.917	0.988	0.923	0.921	0.922	0.888
<i>CurveletQA</i> [14]	0.932	0.985	0.969	0.928	0.946	0.918
<i>BLINDS-II_{Prob.}</i> [15]	0.923	0.985	0.948	0.979	0.963	0.863
<i>BLINDS-II_{SVM}</i> [15]	0.930	0.979	0.938	0.967	0.934	0.895
<i>OG-IQA</i> [16]	0.952	0.990	0.967	0.982	0.945	0.911
<i>CORNIA</i> [17]	0.935	0.951	0.968	0.965	0.951	0.917
<i>RankIQA+FT</i> [18]	0.982	0.994	0.988	0.986	0.975	0.960
<i>CNN</i> [19]	0.953	0.984	0.953	0.981	0.953	0.933
Our Model	0.956	0.991	0.965	0.942	0.947	0.928

Fig. 6 displays the Original Version (OV) and Gaussian Blur distorted versions at five levels (a-e) of the image of "I09" present in the dataset TID2008, with Actual Score (DMOS), and Normalized Score (NS) shown at the top of each image. The Predicted Scores (PS) of DMOS is given at the bottom of every image using IVF. The sub figures (f)-(k) display the equivalent outcomes for the LIVE-IQA dataset's GBlur (GB) image of "churchandcapitol". For both datasets, similar findings were achieved

in independent validation of individual distortions and all distortions. The proposed model's performance measures in IVF generalize its ability to judge IQ regardless of distortions and datasets.

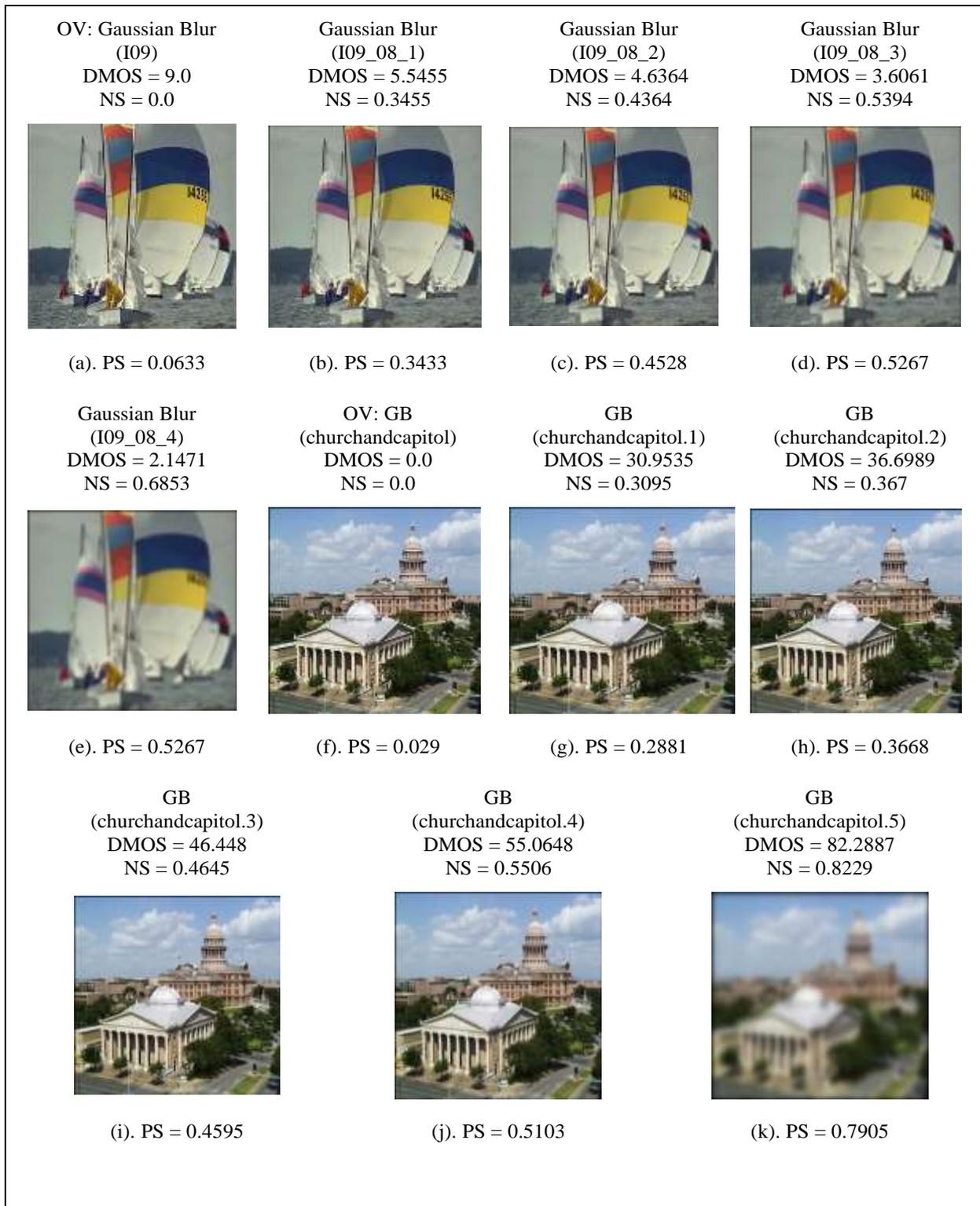


Fig. 6. The original and distorted images of TID2008 (a)-(e) and LIVE-IQA (f)-(k) with predicted scores of DMOS using IVF.

c. Training, Testing and Results using CVF

In CVF, the developed model was trained using all images of one dataset and all images of other dataset was used for testing. For each experiment, all images and individual distortion images are used once. The image counts of CVF experiments are described in Table 6. For example, in the experiment conducted with AGN 125 TID2008 dataset images were operated for training and 174 LIVE-IQA dataset images were utilized for testing. In the counterpart experiment, the roles of the images are swapped as testing and training images respectively.

Table 6. Count of test and train images in the CVF

	AGN		JPEG2000		GBLUR		JPEG		ALL	
	TID2008	LIVE-IQA	TID2008	LIVE-IQA	TID2008	LIVE-IQA	TID2008	LIVE-IQA	TID2008	LIVE-IQA
TR/TE	125	174	125	227	125	174	125	233	425	982

The sub figures (a) and (b) of Fig. 7 show loss function plots for the tests conducted in the CVF, that is, training on one dataset and testing with another dataset. The loss function's in equation (1) monotonic decline throughout the training stage for all images composed and separate distortions demonstrate the proposed model IQA's generic capabilities. The model's computational economy and efficiency are further validated by the fact that in both experimental setups, only 50 epochs were used to train the proposed CNN model. The scatter charts also illustrate that the model's capability for learning is not dependent on the considered datasets along with the distortions.

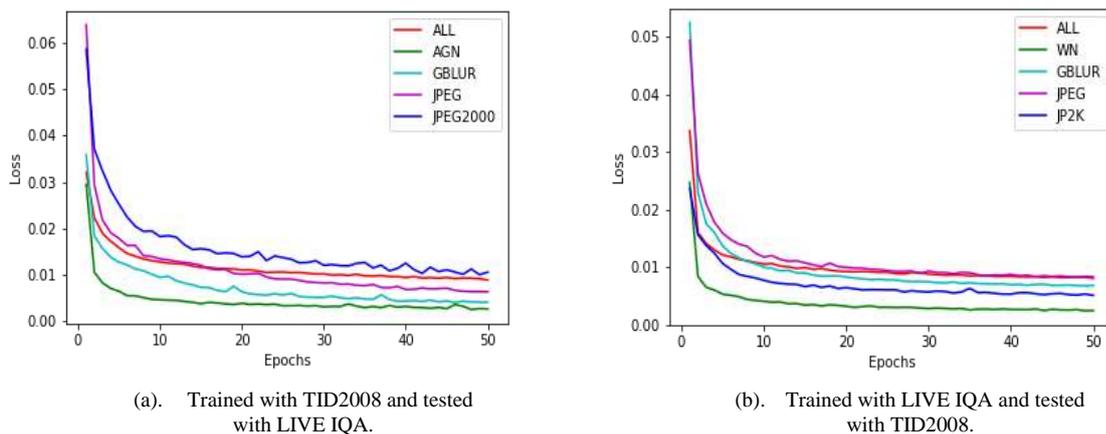


Fig. 7. Loss function (MSE) plots in the training stage for CVF.

The scatter diagrams of the real and predicted scores of DMOS in CVF of all images composed and distinct distortions that are available in both datasets are displayed in Fig. 8, training with TID2008 and testing with LIVE-IQA respectively.

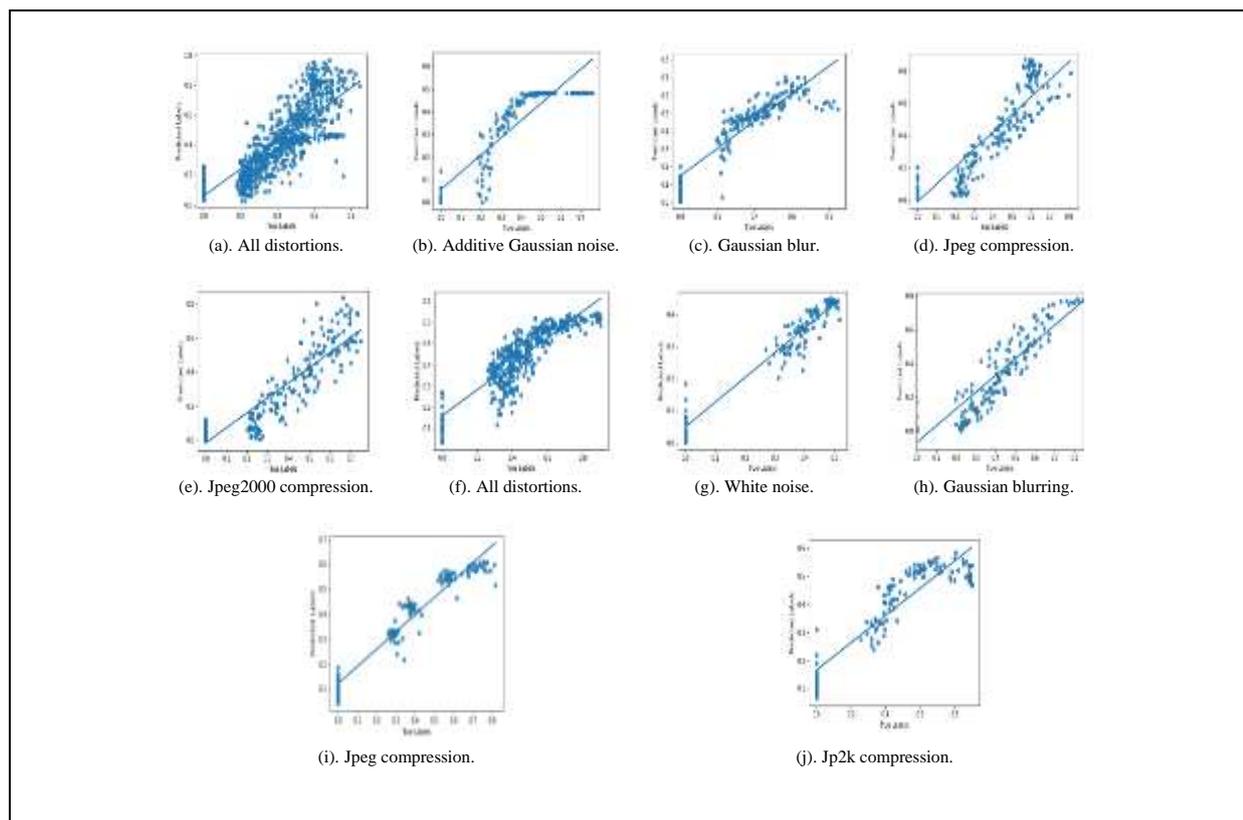


Fig. 8. Scatter plots of actual and predicted scores of DMOS in CVF with training dataset as TID2008 and test dataset as LIVE-IQA (a)-(e), and training dataset as LIVE-IQA and test dataset as TID2008 (f)-(j) respectively.

In the testing phase of the CVF, Table 7 shows the values of PLCC and SROCC among the true and predicted scores of image quality that are calculated by the developed model. The efficiency of all the trials with all distortions together as well as individual distortions make use of the dataset of TID2008 for training along with the dataset of LIVE-IQA for testing are listed. Table 8 displays comparable results using the dataset of LIVE-IQA for training and dataset of TID2008 for testing. The developed model's performance metrics further combines the superior ability of IQA, which is not dependent of distortions and datasets.

Table 7. PLCC, and SROCC for all and individual distortions in CVF with training and testing using TID2008 and LIVE-IQA respectively

Our model	ALL	AGN	BLUR	JPEG	JPEG2000
PLCC	0.857	0.907	0.897	0.915	0.904
SROCC	0.896	0.975	0.922	0.931	0.927

Table 8. PLCC, and SROCC for all and individual distortions in CVF with training and testing using LIVE-IQA and TID2008 respectively.

Our model	ALL	WN	GBLUR	JPEG	JP2K
PLCC	0.866	0.965	0.952	0.963	0.908
SROCC	0.880	0.929	0.922	0.949	0.857

Fig. 9 demonstrates the OV (a) and Jp2k distorted versions of six levels (b)-(g) of the dataset of LIVE-IQA image of "statue", with DMOS, and NS shown at the top of each given image. The Predicted Scores (PS) of DMOS is given at the bottom of each and every image using CVF, that is, trained using TID2008 and tested make use of LIVEIQA. The sub figures of Fig. 9 (h)-(l) displays the equivalent outcomes for the TID2008 dataset's Jpeg compression OV and four versions of "I22". For both datasets, similar findings were achieved in independent validation of all distortions along with individual distortions. The proposed model's performance measures in CVF generalize its ability to judge image quality regardless of distortions and datasets.

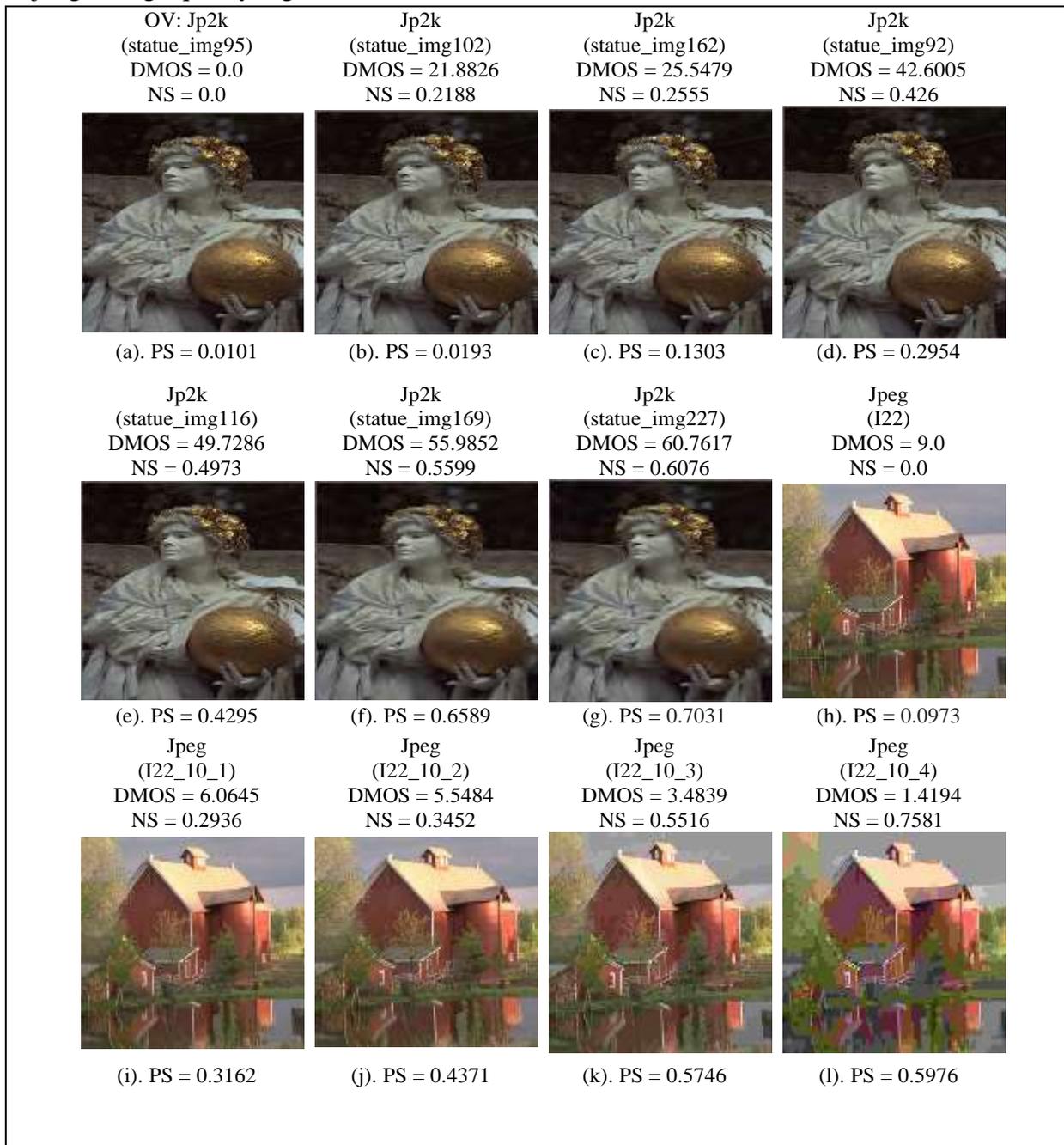


Fig. 9. The original and distorted images of LIVE-IQA (a)-(g) training with predicted scores of DMOS using CVF.

6. CONCLUSIONS

In this work, we presented a CNN framework for BIQA, which falls under the class of NR-IQA model group. The developed model feeds the altered images edge maps that are calculated with the SK, to the CNN, that extracts the input image's higher-level features. Regression technique is utilized to map the image distortions in the features at high level to IQ scores. The investigation results confine that the developed model is general, and its ability to estimate quality of image is not dependent to the distortions that are accessible in the well-known dataset(s) utilized in the experiment. The model's performance measures in terms of PLCC and SROCC values expression that it can compete with the best-in-class approaches.

REFERENCES

1. G. Zhai and X. Min, "Perceptual image quality assessment: a survey," in *Sci. China Inf. Sci.*, 2020, vol. 63, no. 11, pp. 211301: 1- 52.
2. S. Rezazadeh and S. Coulombe, "A novel discrete wavelet transform framework for full reference image quality assessment," *Signal Image Video Process*, 2013, vol. 7, no. 3, pp. 559–573.
3. C. Charrier, O. Lézoray, and G. Lebrun, "Machine learning to design full-reference image quality assessment algorithm," *Signal Process. Image Commun.*, 2012, vol. 27, no. 3, pp. 209–219.
4. Y. Tong, H. Konik, F. A. Cheikh, and A. Tremeau, "Full Reference image quality assessment based on saliency map analysis," *J Imaging Sci Technol*, 2010, vol. 54, no. 3, pp. 030503.
5. H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, 2006, vol. 15, no. 2, pp. 430–444.
6. 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.*, 2004, vol. 13, no. 4, pp. 600–612.
7. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, 2014, vol. 23, no. 10, pp. 4270–4281.
8. D. Tao, X. Li, W. Lu, and X. Gao, "Reduced-reference IQA in contourlet domain," *IEEE Trans. Syst. Man Cybern. Syst.*, 2009, vol. 39, no. 6, pp. 1623–1627.
9. Wang, Zhou, and Eero P. Simoncelli. "Reduced-reference image quality assessment using a wavelet-domain natural image statistic model." *Human vision and electronic imaging X*, 2005, vol. 5666, pp. 149-159.
10. L. Ma, S. Li, F. Zhang, and K. N. Ngan, "Reduced-reference image quality assessment using reorganized DCT-based image representation," *IEEE Trans Multimedia*, 2011, vol. 13, no. 4, pp. 824–829.
11. Xue, Wufeng & Mou, Xuanqin. "Reduced reference image quality assessment based on Weibull statistics," *Second International Workshop on Quality of Multimedia Experience (QoMEX)*, 2010, pp. 1-6.
12. A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Process.*, 2012, vol. 21, no. 12, pp. 4695–4708.
13. A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality," *IEEE Trans. Image Process.*, 2011, vol. 20, no. 12, pp. 3350–3364.
14. L. Liu, H. Dong, H. Huang, and A. C. Bovik, "No-reference image quality assessment in curvelet domain," *Signal Process. Image Commun.*, vol. 29, no. 4, 2014, pp. 494–505.
15. M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," *IEEE Trans. Image Process.*, vol. 21, no. 8, 2012, pp. 3339–3352.
16. L. Liu, Y. Hua, Q. Zhao, H. Huang, and A. C. Bovik, "Blind image quality assessment by relative gradient statistics and adaboosting neural network," *Signal Process. Image Commun.*, vol. 40, 2016, pp. 1–15.
17. Ye, P., Kumar, J., Kang, L., Doermann, D. "Unsupervised feature learning framework for no-reference image quality assessment," *IEEE conference on computer vision and pattern recognition*, 2012, pp. 1098-1105.
18. Liu, Xialei, Joost Van De Weijer, and Andrew D. Bagdanov. "Rankiq: Learning from rankings for no-reference image quality assessment." *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 1040-1049.
19. L. Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 1733–1740.

20. S. Bosse, D. Maniry, K. R. Müller, T. Wiegand, and W. Samek, "Deep neural networks for No-Reference and Full-Reference image quality assessment," *IEEE Trans. Image Process.*, 2018, vol. 27, no. 1, pp. 206–219.
21. S. Bianco, L. Celona, P. Napoletano, and R. Schettini, "On the use of deep learning for blind image quality assessment," *Signal, Image and Video Processing*, 2018, vol. 12, no. 2, pp. 355–362.
22. Y. Li, X. Ye, and Y. Li, "Image quality assessment using deep convolutional networks," *AIP Adv.*, 2017, vol. 7, no. 12, pp. 125324(1-12).
23. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770-778.
24. R. B. Movva and R. K. Kontham, "Blind image quality assessment using a CNN and edge distortion," *Rev. Intell. Artif.*, 2021, vol. 35, no. 4, pp. 315–324.
25. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, 2011, vol. 20, no. 8, pp. 2378–2386.
26. AlNouri, Mason, J. A. Saei, Manaf Younis, Fadi Bourri, M. A. Habash, Mohammed Hamza Shah, and M. A. Dosari. "Comparison of Edge Detection Algorithms for Automated Radiographic Measurement of the Carrying Angle." *J Biomed Eng Med Imaging*, 2015, vol. 2, no. 6, pp. 78–98.
27. D. P. Kingma and J. Ba, "Adam: A Method for stochastic optimization," *International Conference for Learning Representations*, 2017.
28. A. Botchkarev, "A new typology design of performance metrics to measure errors in machine learning regression algorithms," *Interdisciplinary Journal of Information, Knowledge, and Management*, 2019, vol. 14, pp. 45–76.
29. Abadi, Martín, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado et al. "Tensorflow: Large-scale machine learning on heterogeneous distributed systems." *arXiv preprint arXiv:1603.04467*, 2016.
30. N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, F. Battisti, "TID2008 - A Database for evaluation of Full-Reference visual quality assessment metrics", *Advances of Modern Radioelectronics*, 2009, Vol. 10, pp. 30-45.
31. H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, 2006, vol. 15, no. 11, pp. 3440–3451.