

NO-REFERENCE IMAGE QUALITY ASSESSMENT BASED ON DISTORTION SPECIFIC AND NATURAL SCENE STATISTICS BASED PARAMETERS: A HYBRID APPROACH

Jayashri V. Bagade¹, Kulbir Singh², Y. H. Dandawate³

¹Research Scholar, Department of Computer Science, Thapar University, Patiala Punjab, India.

²Associate Professor, Department of Electronics and Telecommunication,
Thapar University, Patiala, Punjab, India.

³Professor, Department of Electronics and Telecommunication,
Vishwakarma Institute of Information Technology, Pune, Maharashtra, India.

Email: jayashrihedao@rediffmail.com¹, ksingh@thapar.edu², yhdandawate@gmail.com³

DOI: <https://doi.org/10.22452/mjcs.vol32no1.3>

ABSTRACT

In web applications, for efficient use of bandwidth and storage requirement, images are compressed with lossy techniques. Quality of images degrades due to the compression artifacts. Transmission channels also contribute in degradation by adding noise. Image quality assessment plays a vital role to address this issue. In majority of cases, reference image remains unavailable for the assessment. Thus, No-Reference quality assessment techniques are widely used. In this paper, machine learning based hybrid approach for NR quality assessment is proposed. Blockiness based parameters, other statistical parameters and NSS based features are provided as input to the feed forward neural network. The back propagation training algorithm predicts a quality score. This score is correlated with differential mean opinion score (DMOS). Here, input can be images of any type from best to worst quality, and the approach used exploits the nonlinearity in the behaviors of parameters. It has been noticed that the predicted score correlates well with the DMOS with 93% accuracy. The above parameters are also used as input to Support Vector Machine for classification. The observed accuracy of this classifier is 89%.

Keywords: *No-Reference image quality assessment, Objective image quality assessment, Artificial Neural Network, Mean Opinion Score, Blockiness, Statistical Parameters, Natural Scene Statistics, Support Vector Machine.*

1.0 INTRODUCTION

In today's world of internet, huge amount of data is transmitted over communication networks of all kind. Images and videos share a considerable volume of this digital data. Efficient use of transmission channel bandwidth and storage will always remain a major concern, whatever kind of data is being stored, transmitted and reproduced. To make this possible, images and videos are compressed using lossy and lossless techniques.

Image compressions, such as lossy, result in quality degradation of an image as it introduces various artifacts such as blockiness, ringing and blurring. Channel noise also contributes to the quality degradation. Image Quality Assessment (IQA) thus plays a vital role in the evaluation of various image processing algorithms and systems. IQA is one of the critical and challenging issues in the area of image and video processing.

There are many processes like image acquisition, storage, retrieval, watermarking, authentication, enhancement, compression and synthesis where quantification of image quality is needed [1]. Image quality evaluation is carried out in two approaches: subjective and objective.

In the Subjective approach, experts and non-experts rate the image quality and the score is obtained by taking mean of total subjects ratings. Since the humans are ultimate receivers in many image processing applications, Mean Opinion Score is extensively accepted for IQA. Although the subjective evaluation Mean Opinion Score (MOS) is reliable, it is slow, inconvenient and time consuming [2]. Many times, Differential MOS (DMOS) is also used where difference of reference and test is rated in the range of 0 to 100. For example, if the difference in rating is very small the score is close to zero [3].

Objective evaluation, which is supposed to correlate with Subjective assessment, estimates perceived image quality through use of various computational models. Objective image quality assessment is categorized into Full Reference (FR), Reduced Reference (RR) and No Reference (NR) [4]. In FR, the quality of degraded image is assessed with the reference which is original or pristine image. The FR based quality metrics are mainly used in assessing the performance of different compression schemes. The normally used FR measures are peak signal to noise ratio (PSNR), image fidelity (IF), correlation coefficient (CC), and structural content (SC) [3] [4]. Though PSNR is simple and widely accepted, it is also criticized for not being associated with image quality measurement in some image processing algorithms [2] [6]. HVS based assessment metric, called as Mean Structural Similarity Index Metrics (MSSIM), is used to overcome this drawback. MSSIM is developed with the fact that the structural content in an image changes according to degradation. It takes luminance, contrast and structural content in terms of image statistics into account [2][6].

In RR, instead of original image, certain features of an image are extracted and used to compute the quality of reconstructed images. These features are normally used for image repair or impairment [7]. The wavelet coefficients of deformed images are calculated and fitted with Gaussian scale mixture (GSM) distributions to define a feature. Difference between scaled entropies of wavelet coefficients of original and deformed images is computed [8]. In another approach, which uses structural similarity, statistical features are extracted from multiscale, multi-orientation divisive normalization transform. The quality measure is developed by pursuing the concept of SSIM for FR [9].

There are plenty of applications in which reference images or features cannot be made available for assessment. For example, the images stored in digital camera are JPEG compressed by internal CODEC or compressed images transmitted over social media. The quality assessment of such images falls under NR category. Such assessments are done well by human cognition. For large amount of images, the assessment can be automated using machine learning approach [10] [11] [12]. The basic idea in NR image quality is the exploration of information fidelity or quantification of loss of information due to distortion/degradation process (compression, noise etc).

Existing NR algorithms mainly follow three trends - (1) Distortion specific approach - Presents kinds of distortions such as blockiness, ringing, blur, Just Noticeable Distortion (JND), contrast etc, (2) Feature extraction and machine learning approach - Training based prediction models or use of classifiers like Support Vector Machine (SVM), Artificial Neural Network (ANN) to determine the quality score. (3) Natural scene statistics (NSS) approach - There is assumption that natural images are a subset of the complete set of possible images. The distance between the distorted images and the subset of natural images is computed. Idea behind this approach is to determine how the statistics of images is affected by introducing distortions in them [13]. Moorthy and Bovik defined the statistical description of distorted images in distortion specific signature. In this algorithm, authors computed likelihood of each distortion in image and applied Multiclass Support Vector Machine (SVM) with radial-basis function kernel classifier to categorize test image into defined distortions [14].

From literature review, it is observed that above approaches handle individual distortion. Images may contain combination of distortions. Quality of an image cannot be assessed by considering any individual artifact as image may have any other or combination of blur, blockiness and ringing artifacts. There is need to define generalized framework. This paper presents the machine learning based hybrid approach as a new technique for IQA. Parameters from above three approaches are used in combination to determine score accurately, and images of quality from high to low are used as input. The LIVE dataset having all variants of distortions along with original images are used for complete training and testing. The accuracy of prediction obtained is around 90%. Results of this prediction model are also compared with Support Vector Machine (SVM). The accuracy obtained by this classifier is 89%.

Section 2.0 presents the related work done in the area of NR. Section 3.0 explains Feature extraction. Section 4.0 describes the technique proposed for quality score estimation and experimentation. Experimental results are presented in section 5.0. Finally, section 6.0 concludes the proposal and future directions.

2.0 PREVIOUS WORK

Z. Wang et al. [15] addressed the blocking artifacts by estimating the activities in images using following two factors: (1) Average absolute difference between in-block image samples and (2) Zero-crossing rate. The relational predictor evaluates the image quality score on the scale of 0 to 10 and only used for JPEG images. S. Suresh et al. [11] presented neural network based classifier using extreme machine learning algorithm for IQA. Estimation of quality is done as the functional relationship between HVS based features and MOS. In this paper, four blockiness based parameters of

JPEG images namely Edge Length, Edge Amplitude, Background activity, and Background luminance are calculated. Extreme machine learning classifier computes the quality score. These features are strong enough to address only blockiness. The accuracy of classifier is 70% for testing JPEG images. Other artifacts like white noise are not handled. More features are needed to be defined to handle these types of artifacts.

Z. M. Parvez Sazzad et al. [16] proposed computation of spatial features for JPEG2000 images based on the fact that human beings are very sensitive to edge information. The presented NR metric is defined using edge information and pixel distortion based features. Pixel distortion for a central pixel is evaluated as Standard Deviation (SD) that is calculated inside 5×5 neighborhood pixels. The SD is also calculated for all central pixels in the image. Another pixel distortion feature is defined as an absolute difference between neighborhoods of 5×5 block with every pixel as a center. Edge information is calculated as zero-crossing (ZC) rate which is computed row wise and column wise for the image. Another feature is histogram-based edge measure which is computed with and without edge preserving filter. Since the human visual system and the physical features of an image exhibits non-linear properties, the predictor based on Particle Swarm Optimization is used to anticipate the quality score. Authors tested the predictor with their own dataset compressed with six compression ratios and obtained accuracy of 93%. The predictor is also tested with Texas LIVE dataset and claimed same accuracy. Results obtained are only for JPEG2000 images.

P. Gastaldo et al. [10] proposed NR technique using non-parametric features. Practically, natural images are complex and difficult to analyze at global level. Objective features of an image are thus computed block by block. The proposed features are categorized in three groups, namely (1) First order histogram of image blocks, (2) Co-occurrence matrix and (3) Frequency-based representation. Estimation of quality score is done using circular back propagation model. Images are preprocessed using contrast enhancement followed by feature extraction. The reported accuracy is 86%.

R. Ferzil et al. [17] proposed the perceptual-based sharpness/blurriness metric. In this metric Just Noticeable Blur (JNB) concept is incorporated in probability summation predictor. On the basis of the probability of blur detection, the blurriness in an image is quantified. Probability summation model is used for blur estimation. Accuracy for Gaussian blur with blur variance ranging from 0.8 to 2.4 is 93% and 83% for JPEG. This metric address JPEG and Gaussian blur images; whereas it does not test high blur variance images.

L. Liang et al. [18] presented better approach to explore potential artifacts present in an image. Gradient profiles along strong edges are modeled. Statistical information on these profiles is used to determine the presence of artifacts in an image. Accuracy obtained is 94% with LIVE JPEG2000, 92% for LIVE Gblur and 86% for TID dataset. Only the blur and ringing artifacts that are present in JPEG2000 are considered.

S. Suthaharan [19] made use of blocking artifacts which determines the perceptibility of distortion which measured in terms of undistorted edges and blocking artifacts in the image. Quality score is determined by the parameterized mathematical model. This approach worked well only for LIVE JPEG images with accuracy of 92%. H. Liu et al. [20] measured the ringing artifacts in images. In his proposed work, NR metric is developed based on two-step approach. In the first step, the ringing detection method is defined, and in second step ringing annoyance is quantified. NR metric is tested for JPEG and JPEG2000 images with accuracy of 80% and 85% respectively. Some models demonstrated special features like local spatial and spectral entropies [21].

Above literature is related to specific distortions. Some researchers concentrated their work on the features that are based on Discrete Cosine Transform (DCT) statistics. M. A. Saad et al. [13] presented an excellent methodology to define the quality index that is free from any particular type of distortion. IQA is done by monitoring the information of local DCT coefficients. Authors extracted two HVS based features from an image, viz the structure and contrast. The probabilistic prediction model is proposed for determination of quality score prediction which resulted in 79% accuracy.

T. Brandao et al. [22] presented a novel approach to estimate the original DCT coefficients of an image from its quantized values available at decoder. To achieve this estimation, two techniques are used. The relationship between distribution parameter at adjoining DCT frequencies is estimated by linear prediction scheme. The predicted results are combined with maximum likelihood (ML) parameter estimation. For NR image quality assessment, the algorithm computed PSNR estimation. The mathematical model based on Laplacian parameter estimation is used to evaluate the

PSNR of the image. This technique works well for LIVE JPEG images with accuracy of 97%. Results are based on parameter estimation.

G. Zhai et al. [23] developed an algorithm to evaluate noticeable blockiness for block-based DCT images. The variations on block boundaries are computed and then converted into the block discontinuity map. Luminance variation and texture masking is applied to the blocks. All block maps are combined by applying a nonlinear operator which produces overall masking map. Lastly, integration of these maps is the discontinuity map that results in Noticeable Blockiness Map (NBM). This map is used for IQA. Accuracy is noted to be 96% for JPEG images only.

In the third trend, Natural Scene Statistics (NSS) is used. In NSS, undistorted images lie in a subset of the entire set of possible images. The distance between the distorted image and subset of natural images is determined to monitor the change in the statistics of an image as an effect of distortions.

H. Sheikh et al. [24] proposed NR metric using statistical model of natural images in the Wavelet domain. They evaluated the Wavelet coefficient of natural images in a given subband and their relationship with other wavelet coefficients across scales and orientations. Each subband vector is nonlinearly transformed and linearly combined. Linear predicate predicts the quality score. Around 93% accuracy is achieved for JPEG2000 images. It is observed that, spatial details and texture over the image affects the performance of algorithm. A. K. Moorthy and A. C. Bovik [14] presented two stage framework for NR image quality assessment based on NSS concept. In the first step, image distortion classification based on NSS modification is computed, followed by quality assessment. Five types of distortion are addressed viz JPEG, JPEG2000, white noise (WN), Gaussian Blur, and Fast Fading. Multiclass Support Vector Machine (SVM) with radial basis function kernel classifier classifies a test image into one of the above mention distortions with 82% accuracy for LIVE dataset.

M. J. Chen et al. [25] have given good insight into the natural scenes statistics. They defined reliable features for quality assessment through multi-resolution decomposition method. Natural scene statistics features are extracted from the gradient histogram. They used a Support Vector Machine (SVM) classifier. Using LIVE Gblur, 96% image accuracy is obtained. Great efforts has been put forth in this area by developing renowned algorithms like DIIVINE index, BLINDS-II, C-DIIVINE BRISQUE [26][27][28][29]. Islam et al. [30] defined the NSS model contrast distorted images using moments and entropy features. Authors [31] extracted features from log-histogram of curvelet coefficient values and the energy distribution of both orientation and scale. Zhou et al. [32] worked with noise, blur and blockiness. Quality prediction is the weighted sum of modified wavelet medium estimation, count of edge pixels points and characteristics of image pixel. The computational cost and design of the suitable threshold function, which explores special characteristics of curvelet, can be applied in area of edge detection.

The work done so far was focused on specific distortion types and images. The parameters were developed for the IQA of specific image types such as JPEG, JPEG2000, blurry, etc. Thus, IQA was restricted for specific type of images only. If automated IQA framework is required to be developed, it should assess all types of images having combination of different distortions and also images of good quality. Applications dealing with image processing use compressed images. Compression introduces blocking and ringing artifacts. To address these artifacts, blockiness based parameters are needed to be extracted. Other artifacts like noise and blurriness can also be present in images. These artifacts can be handled through statistical parameters. Web applications are supposed to handle natural images along with other images. These images possess different types of statistics which change due to distortions. This change is obtained using NSS based parameters. In this paper, we propose a framework which takes into account the features developed so far and applies the multilayered backpropagation artificial neural network based prediction model to estimate the image quality score. The network is trained with these features as inputs and DMOS as target.

3.0 FEATURE EXTRACTION

For image quality assessment of the gray images, we have considered blockiness, second ordered statistical parameters and NSS based parameters. Details of the parameters are described below. Blockiness-based parameters address blocking artifacts where statistical parameters handle blurring effect and to some extent noise. Frequency based parameters able to handle artifacts generated as effect quantization. NSS based parameters are used to identify the artifacts in natural images.

3.1 Blockiness-based Parameters

JPEG images are used in many web-based applications. Compression introduces blocking, ringing and blurring artifacts in an image. To identify blocking artifacts, Z. Wang *et al.* [15] proposed parameters which are evaluated based on pixel variations across rows and columns.

Following expression represents the image.

$$s(x, y) \text{ for } x \in [1, X] \text{ and } y \in [1, Y]$$

A differencing signal along each horizontal line is calculated as

$$p_h(x, y) = s(x, y + 1) - s(x, y), y \in [1, Y - 1] \quad (1)$$

The features are computed row wise and column wise.

- a. The average difference across block boundaries is calculated to obtain blockiness as,

$$bl_h = \frac{1}{X(\lfloor Y/8 \rfloor - 1)} \sum_{i=1}^X \sum_{j=1}^{\lfloor Y/8 \rfloor - 1} |p_h(i, 8j)| \quad (2)$$

- b. Image signal activity is calculated as two factors. First factor, average absolute difference between in-block image samples is calculated as

$$avg_h = \frac{1}{7} \left[\frac{8}{X(Y-1)} \sum_{i=1}^X \sum_{j=1}^{Y-1} |p_h(i, j)| - bl_h \right] \quad (3)$$

- c. The other factor, zero crossing rate(ZC) is calculated as

$$\text{For } y \in [1, Y - 2],$$

$$zc_h(x, y) = \begin{cases} 1 & \text{horizontal ZC at } p_h(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

ZC rate is estimated horizontally as

$$zc_h = \frac{1}{X(Y-2)} \sum_{i=1}^X \sum_{j=1}^{Y-2} zc_h(x, y) \quad (5)$$

Row wise and column wise computed features are averaged as below

$$bl = \frac{bl_h + bl_v}{2}, avg = \frac{avg_h + avg_v}{2}, zc = \frac{zc_h + zc_v}{2} \quad (6)$$

The value of bl ranges from 0.50 to 91.91, avg ranges from 0.51 to 91.48 and zc ranges from 0.0007 to 0.65. These values are ranges as per the amount of distortion present in the images. bl increases with increase in blockiness while avg and zc decreases from original to worst.

3.2 Statistical Parameters:

P. Gastaldo *et al.* [10] proposed parameters to model textual characteristics. The features are described as statistical information of an image such as gray value distribution, spatial orientation information and frequency energy distribution. In the proposed work, images are preprocessed before feature extraction. On the other hand, in our proposal parameters are extracted without preprocessing. In image processing, applications distortion quantification needs to be done without preprocessing as it can help in reducing the time complexity of algorithms. Also, preprocessing of images can lead to the wrong prediction about distortion quantification.

- a. Second Order Statistical Parameters: Second order histogram features are obtained from co-occurrence matrix. $C(g_i, g_j, r, \omega)$ represents co-occurrence matrix. It is defined as probability distribution of pair of pixels with gray levels g_i and g_j , which is apart by radial units at angle ω to x axis.
- b. Frequency based Parameters: Features are derived from energy of frequency content. It is computed as discrete cosine transform of each block.

All above features are listed in Appendix A.

3.3 NSS based Parameters

NSS based parameters are used to deal with natural scene images which are affected due to various artifacts. M. A. Saad *et al.* [33] proposed DCT based NSS features. These features are computed block wise. Image is divided into 5×5 blocks. Features are calculated for all blocks in the image. Features are then combined by averaging the values.

- a. Coefficient of Frequency Variation: Let Y is a random variable, which describes the DCT coefficients. The coefficient of frequency variation feature is

$$\psi = \frac{\sigma|Y|}{\mu|Y|} \quad (7)$$

Where, μ and σ represents mean and standard deviation respectively.

Its values range from -9.92 to 262.66

- b. Energy Subband Ratio Measure: As depicted in Fig. 1, DCT coefficients are represented increasingly in higher radial spatial function from top-left corner to bottom-right corner. We defined three frequency bands. The average energy is square of deviation. In frequency, band x it is defined as

$$AE_x = \sigma_x$$

The ratio is defined as difference between the average energy in frequency band x and average energy up to frequency band x to sum of these measures.

$$RE_x = \frac{\left| AE_x - \frac{1}{n-1} \sum_{i \in x} AE_x \right|}{AE_x + \frac{1}{x-1} \sum_{i \in x} AE_x} \quad (8)$$

Their value ranges from 0.22 to 1.47.

DC	P ₁₂	P ₁₃	P ₁₄	P ₁₅
P ₂₁	P ₂₂	P ₂₃	P ₂₄	P ₂₅
P ₃₁	P ₃₂	P ₃₃	P ₃₄	P ₃₅
P ₄₁	P ₄₂	P ₄₃	P ₄₄	P ₄₅
P ₅₁	P ₅₂	P ₅₃	P ₅₄	P ₅₅

Fig. 1. DCT coefficient with three energy bands

- c. Orientation: As an effect of distortion, local orientation energy changes abnormally. The HVS is highly sensitive to local orientation energy. This directional information is represented as block DCT coefficients

along with three orientations as depicted in Fig. 2. The values of orientations range from 0.14 to 2.18, 0.11 to 2.16 and 0.19 to 2.32 respectively.

DC	P ₁₂	P ₁₃	P ₁₄	P ₁₅
P ₂₁	P ₂₂	P ₂₃	P ₂₄	P ₂₅
P ₃₁	P ₃₂	P ₃₃	P ₃₄	P ₃₅
P ₄₁	P ₄₂	P ₄₃	P ₄₄	P ₄₅
P ₅₁	P ₅₂	P ₅₃	P ₅₄	P ₅₅

Fig. 2. DCT coefficients defined along three orientations

4.0 METHODOLOGY

Fig. 3 shows blocking parameters for original JPEG image along with five nonlinear distortions. Fig. 4 and Fig 5 shows the deviation in second order statistical and NSS based parameters for images with white noise and fast fading distortion types, respectively. If the framework is required to be proposed, it should consider all images, varying from original to distorted, for determination of score. Model based approach leads to wrong prediction due to model tuning parameters. In the model based approach, though MOS shows that quality of images is good, predictor predicts bad quality. From data analysis, it is observed that the nature of extracted features is nonlinear. In order to overcome this, we are proposing neural network based modeling as ANN is a good solution to establish linear relationship between nonlinear feature values and DMOS [34][35]. Bagade et al. [34] [35] proposed blocking and frequency domain statistical features in their earlier work with 95% accuracy for JPEG images. They concluded that sometimes, quality score predicted by mathematical predictors do not correlate with MOS which leads to incorrect prediction. To characterize the nonlinear relationship between extracted features and MOS, the machine learning approach is used.

Fig. 6 depicts the methodology of experimentation. In the hybrid approach, blocking artifacts are computed using blocking parameters. Second order statistical parameters model textual characteristics. Features are described as correlation between gray scales and spatial energy distribution. NSS based parameters are computed to model natural scene statistics. These parameters are DCT based parameters and computed for each block of size 5×5. All the extracted features form feature vectors. The back-propagation neural network is trained with the feature vector. During training, the DMOS is used as a target to the network. ANN depicts nonlinear and complex relation among the features from feature vector and DMOS. The trained neural network computes the quality score for a test image.

Proposed framework uses multilayer feedforward network with back-propagation learning algorithm. The network comprises an input layer with twenty input neurons and a hidden layer with nine neurons. Logarithmic sigmoid transfer function is used for hidden neurons. Subsequently, an output layer neuron has linear transfer function. The nonlinear transfer functions of neurons permit the network to gain knowledge of nonlinear and linear correlation between the input and the output.

In neural network, output depends on initialization of the weights. In each run, the decision boundaries change with the initial weights. Indirect nature of discrimination function in ANN is unable to handle the boundary conditions. In SVM, decision boundaries are determined to minimize classification error for training as well as testing data. SVM has high generalization ability due to the capability of maximizing margins. SVM is more robust than ANN. To check the sturdiness of framework we also tested prediction with SVM.

Training of SVM is done through feature vectors. The vectors are categorized in four classes, namely best, good, average, and worst. Trained SVM predicts the class for test vectors.

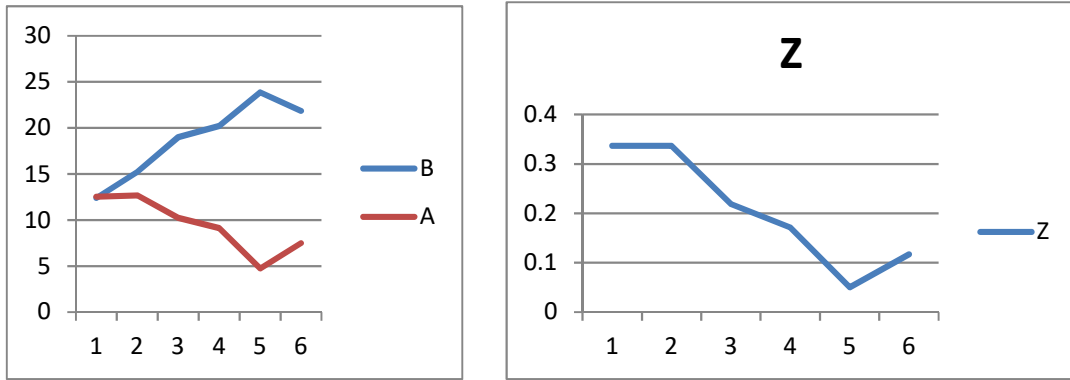


Fig. 3. Plot for blocking parameters for JPEG type of image

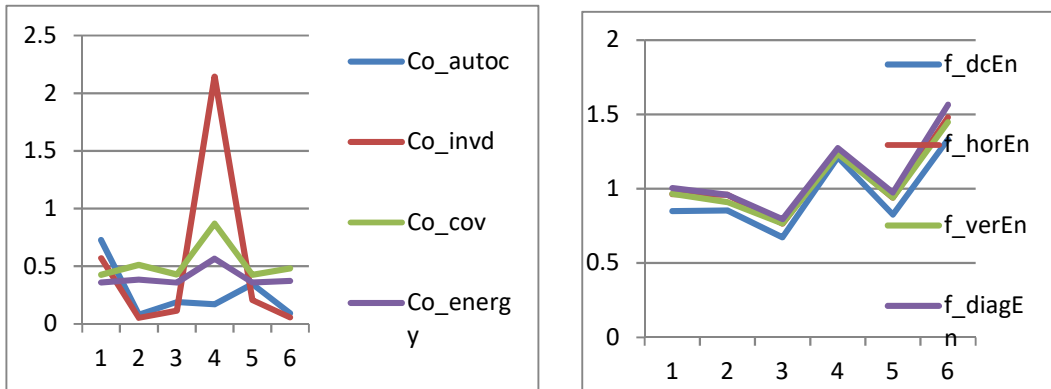


Fig. 4. Plot for statistical parameters for white noise images

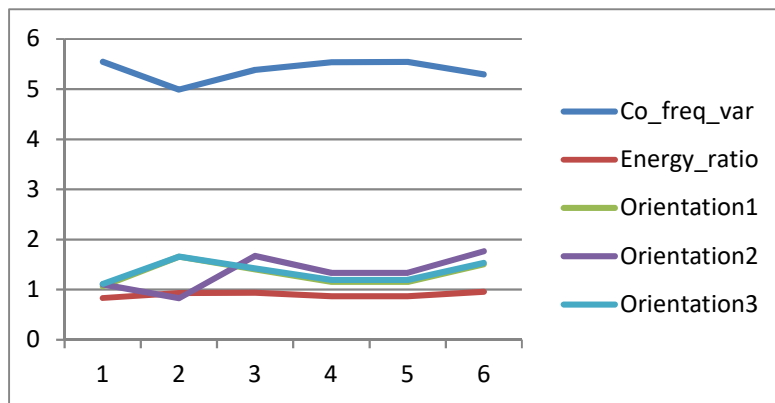


Fig. 5. Plot for NSS based parametrs for Fast Fadding images

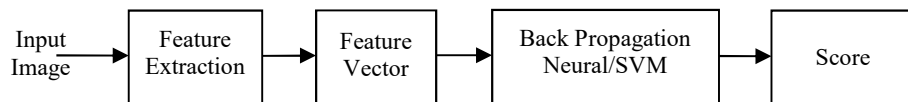


Fig. 6. Block diagram for methodology

5.0 EXPERIMENTATION

To model the framework, LIVE database is used [36]. Training set comprises of 20 original images and its five distortion variants *viz.* JPEG, JPEG2000, Gaussian blur, White noise, and Fast Fading. 9 original images along with

its variants are tested with trained network. In both training and testing phases of ANN and SVM, 29 original images are used along with 3 original images from training dataset, each of which has five distortions. Each image has five variants which comes out to be 960 images in total. The network is trained with extracted features from training dataset. Quality score estimation is done by trained network for the test dataset. The sample images are shown in appendix B.

This model is also tested for TID2008 [37] dataset which comprises of 25 original images and its four distortion variants viz. JPEG, JPEG2000, Gaussian blur, and White Noise. 18 original images with its four distortion variants are used to train the ANN and SVM. 7 original images with its distortion variants are used to test the classifiers. Image size and intensity of the image inputted to this framework is used as mentioned in the dataset. This framework works with any image size and intensity. Memory storage requirement is limited to the available physical memory.

6.0 RESULTS

In the first experiment, quality score is calculated by network and compared with DMOS. MATLAB version 7.11.0(R2010b) is used. Table 1 depicts performance of neural network for blockiness based parameters and combination of statistical parameters. It works well for specific individual distortions. In the second experiment, Support Vector Machine is trained using above extracted features. Feature vectors are divided in four classes based on DMOS. DMOS values ranging from 0 to 25 is defined as class 1, values greater than 25 and less than 50 are defined as class 2, 50 to 75 as class 3 and all above 75 are defined as class 4. Trained SVM predicts the class for test vector. For both the experiments, framework is tested for LIVE data Set and TID2008 dataset. In Table 2, we have presented comparative study of ANN and SVM by combining blockiness based parameters, statistical parameters and NSS based parameters for LIVE dataset. In Table 3, performance of classifiers for TID dataset is presented. ANN shows overall accuracy of 0.93 for LIVE dataset and 0.92 for TID dataset. The performance of ANN for JPEG2000 and Fast Fading is slight low. JPEG2000 images have ringing effect and in FF frequencies in the image varies randomly. The regression plot of ANN for overall dataset is depicted in Fig 7. The error plot of DMOS and estimated quality score by network is shown in Fig 8. This error plot is plotted for test dataset containing 9 original images from LIVE dataset. It also includes distorted and compressed versions of original images like JPEG, JPEG2000, Gaussian Blur, White Noise, and Fast Fading along with its five variants. The mean error is 0.02. From the error plot, it can be observed that Fast Fading images have more errors as the extracted parameters are not enough to handle the artifact generated due to the distortions.

Table 4 depicts performance comparison between quantification of blockiness, statistical parameters, NSS based parameters, and combined approach. Combined approach shows edge in the results. Wang tested his approach for JPEG images only and though MOS of an image is higher, predicted score found to be less [26][27]. In CBP approach, images are pre-processed before feature extraction. In combined approach, images are not pre-processed. Though observed results for ANN shows the edge over SVM, results of ANN do change with every execution because of weight initialisation. SVM results do not show such change in the execution.

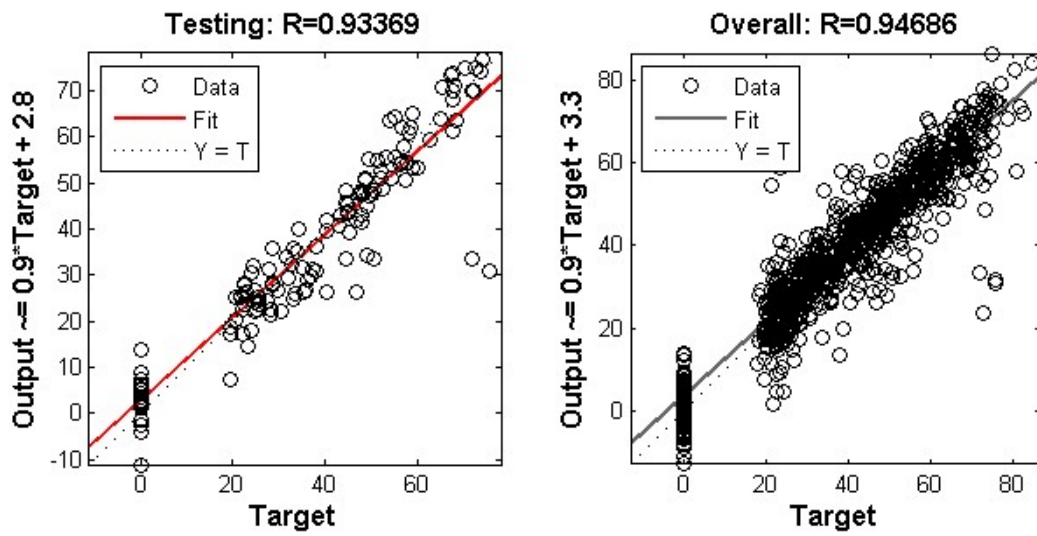


Fig. 7. Regression plot for Back Propagation Neural Network

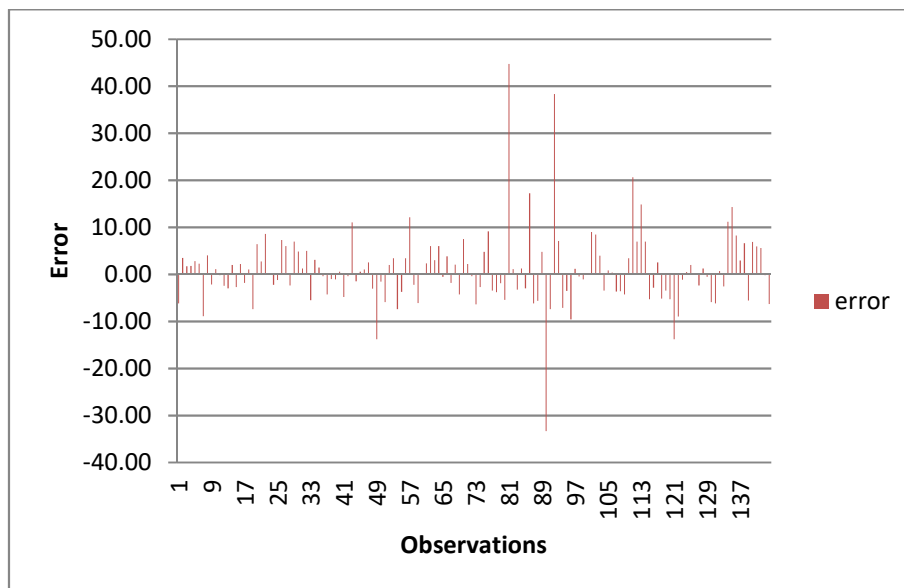


Fig. 8. Error Plot: Error between DMOS and predicted score for test images

Table 1. Performance of network for particular image type

Image type/ Parameters	Linear Correlation
Blocking Parameters	
JPEG	0.93
JPEG2k	0.83
Gblur	0.88
WN	0.98
FF	0.92
Blocking +Second order statistical parameters	
JPEG	0.92
JPEG2k	0.81
Gblur	0.90
WN	0.97
FF	0.95

Table 2. Comparative performance of ANN and SVM for LIVE dataset

Image type/ Parameters	Linear Correlation Coefficient	
	ANN	SVM
JPEG	0.89	0.92
JPEG2k	0.87	0.91
Gaussian blur	0.89	0.93
White Noise	0.98	0.95
Fast Fading	0.89	0.87
Overall	0.93	0.89

Table 3. Comparative performance of ANN and SVM for TID dataset

Image type/ Parameters	Linear Correlation Coefficient	
	ANN	SVM
JPEG	0.93	0.94
JPEG2k	0.84	0.92
Gaussian blur	0.89	0.90
White Noise	0.96	0.95
Overall	0.92	0.90

Table 4. Comparison of different approaches with proposed method for LIVE dataset

Approaches	Linear Correlation Coefficient
Proposed approach	0.93
Wang[15]	0.75
CBP[10]	0.86
BLIINDS-II[33]	0.82

7.0 CONCLUSION

Images usually have combination of distortions than specific individual distortion. Considering blockiness, statistical and NSS based parameters, quality assessment is done irrespective of image type and distortions. Original and distorted images are given as input to the system and neural network is trained using DMOS as a target. It is observed that predicted score of 93% images are correlated with its DMOS. As various parameters are used to find out different types of distortions, image specific parameters are better used in combination with neural networks. This is possible due to the characteristics of neural network to evaluate nonlinear relationship between extracted features and DMOS. These parameters are tested with SVM and observed accuracy is 89%. SVM is performing well with all the distortions. In case of Fast Fading images, its performance is found slightly lower. To improve the accuracy, NSS based and other features are needed to be extracted. Other classifiers can be used to attain better accuracy. The study of neuro-fuzzy classifier for robustness will be the focus for our future work.

REFERENCES

- [1] H. R. Sheikh, A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, Vol. 15, No. 2, February 2006, pp. 430–444.
- [2] Z. Wang, A. C. Bovik, "Modern Image quality assessment", Morgan & Claypool ISBN: 1598290231, March 2006.
- [3] I. Avcibas, B. Sankur, K. Sayood, "Static evaluation of image quality measures", *Journal of Electronic Imaging*, Vol. 11, April 2002, pp 206-223.
- [4] A. M. Eskicioglu, P. S. Fisher, "Image quality measures and their performances," *IEEE Transaction Communications*, Vol.43, No. 12, December 1995, pp. 2959-2965.
- [5] Z. Wang, A. C. Bovik, "A universal image quality index", *IEEE Signal Processing Letters*, Vol.9, No.3, March 2002, pp. 81-84.
- [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 Processing*, Vol. 13, No.4, April 2004, pp. 600–612.
- [7] Q. Li, Z. Wang, "Reduce-reference image quality assessment using divisive normalization based image representation," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 3, No. 2, April 2009, pp. 202-211.
- [8] R. Soundararajan, A. C. Bovik, "RRED Indices: Reduce reference entropic differencing for image quality assessment," *IEEE Transactions on Image Processing*, Vol. 21, No. 2, February 2012, pp. 517-526.
- [9] A. Rehman, Z. Wang, "Reduced-reference image quality assessment by structural similarity estimation", *IEEE Transactions on Image Processing*, Vol. 21, No.8, May 2012, pp. 3378-3389.
- [10] P. Gastaldo, R. Zunino, I. Heynderickx, E. Vicario, "Objective quality assessment of displayed images by using neural networks", *Signal Processing: Image Communication*, Vol. 20, No.7, August 2005, pp. 643-661.
- [11] S. Suresh, R. Venkatesh Babu, H. J. Kim, "No-reference image quality assessment using modified extreme learning machine classifier", *Applied Soft Computing*, Vol. 9, No. 2, March 2009, pp. 541-552.
- [12] K. T. Islam, and R. G. Raj. "Real-Time (Vision-Based) Road Sign Recognition Using an Artificial Neural Network", *Sensors* 2017, Vol. 17, 853.
- [13] M. A. Saad, A. C Bovik, C. Charier, "A DCT statistics-based blind image quality index", *IEEE Single Processing Letters*, Vol. 17, No. 6, June 2010, pp. 583-586.
- [14] A. K. Moorthy, A. C. Bovik, "A Two-step framework for constructing blind image quality indices", *IEEE Signal Processing Letter*, Vol. 17, No. 5, May 2010, pp. 513-516.
- [15] Z. Wang, H. R. Sheikh, A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *Proc. IEEE International Conference on Image Processing*, Vol. 1, September 2002, pp. 477–480.
- [16] Z. M. Parvez Sazzad, Y. Kawayoke, Y. Horita, "No reference image quality assessment for JPEG2000 based on spatial features," *Signal Processing: Image Communication*, Vol. 23, No. 4, April 2008, pp. 257-268.
- [17] R. Ferzil, L. J. Karam, "A No-reference objective image sharpness metric based on the notion of just noticeable blur (JNB)", *IEEE Trans. Image Processing*, Vol. 18, No. 4, April 2009, pp. 718-728.
- [18] L. Liang, S. Wang, J. Chen, S. Ma, , D. Zhao, W. Gao, "No reference perceptual image quality metric using gradient profile for JPEG 2000", *Signal Processing: Image Communication*, Vol. 25, No.7, August 2010, pp. 502-516.
- [19] S. Suthaharan, "No-reference visually significant blocking artifact metric for natural scene images", *Signal Processing*, Vol. 89, No.8, August 2009, pp. 1647-1652.
- [20] H. Liu, N. Klomp, I. Henynderickx, "A No-reference metric for perceived ringing artifacts in images", *IEEE Tran. Circuit and Systems for Video Technology*, Vol. 20, No. 4, April 2010, pp. 529-539.
- [21] L. Liu, B. Liu, H. Huang, A. Bovik, "No reference image quality assessment based on spatial and spectral entropies", *Signal processing: Image communication*, Vol. 29, No. 7, June 2014, pp. 856-863.

- [22] T. Brandao, M. P. Queluz, "No-reference image quality assessment based on DCT domain statistics", *Signal Processing*, Vol. 88, No. 4, April 2008, pp. 822-833.
- [23] G. Zhai, W. Zhang, X. Yang, W. Lin, Y. Xu, "No-reference noticeable blockiness estimation in images", *Signal Processing*, Vol. 23, No. 6, May 2008, pp. 417-432.
- [24] H. R. Sheikh, A. C. Bovik, L. K. Cormack, "No-reference quality assessment using natural scene statistics: JPEG2000," *IEEE Trans. Image Processing*, Vol. 14, No. 11, November 2005, pp. 1918–1927.
- [25] M. J. Chen, A. C. Bovik, "No. reference image blur assessment using multiscale gradient", in *Proc. IEEE Quality of Multimedia Experience*, July 2009, pp. 70-74.
- [26] A. K. Moorthy, A. C. Bovik, "Blind image quality assessment: From natural scene statistics to perceptual quality", *IEEE Transaction image processing*, Vol. 20, No. 12, December 2011, pp. 3350-3364.
- [27] A. Mittal, A. K. Moorthy, A. C. Bovik, "No reference image quality assessment in the spatial domain", *IEEE Transactions on image processing*, Vol. 21, No. 12, December 2012, pp. 4695-4707.
- [28] Y. Zhang, A. Moorthy, D. Chandler, A. Bovik, "C-DIIVINE: No reference image quality assessment based on local magnitude and phase statistics of natural scenes", *Signal processing: image communication*, Vol. 29, No. 10, May 2014, pp. 725-747.
- [29] A. Mittal, R. Soundararajan, A. Bovik, "Making a completely blind image quality analyzer", *IEEE Signal processing letters*, Vol. 20, No. 3, March 2013, pp. 209-213.
- [30] K. T. Islam, R. G. Raj, G. Mujtaba, "Recognition of Traffic Sign Based on Bag-of-Words and Artificial Neural Network", *Symmetry*, 2017, Vol. 9, No. 8, pp. 138.
- [31] M. A. Shayegan, S. Aghabozorgi, R. G. Raj, "A Novel Two-Stage Spectrum-Based Approach for Dimensionality Reduction: A Case Study on the Recognition of Handwritten Numerals," *Journal of Applied Mathematics*, Vol. 2014, Article ID 654787, 14 pages, 2014. doi:10.1155/2014/654787.
- [32] L. Zhou, Z. Zhang, "No reference image quality assessment based on noise, blurring and blocking effect", *Optik*, Vol. 125, May 2014, pp. 5677-5680.
- [33] M. A. Saad, A. C. Bovik, C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain", *IEEE Transactions on image processing*, Vol. 21, No. 8, August 2012, pp. 3339-3351.
- [34] J. V. Bagade, K. Singh., Y. H. Dandawate, "No reference image quality assessment using block-based and frequency domain statistical features: a machine learning approach", *Int. Journal of Communication Network and Distributed System*, Vol. 12, No. 1, November 2014, pp. 95-112.
- [35] J. V. Bagade, Y. H. Dandawate, K. Singh, "No reference image quality assessment using block based features and artificial neural network." in *proc. 4th International conference on communication in computer and information science*, Springer, Vol. 270, November 2011, pp. 128-138.
- [36] H. R. Sheikh, Z. Wang, L. Cormack, A. C. Bovik, Live image quality assessment database, <http://www.live.ece.utexas.edu/research/quality>.
- [37] [41] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, ID2008—a database for evaluation of full-reference visual quality assessment metrics," *Adv. Modern Radioelectron.*, 2009, vol. 10, no. 10, pp. 30–45.

APPENDIX A: STATISTICAL FEATURES.

1. Second Order Parameters

$$co_autoc = \sum_{g_i, g_j} g_i g_j c_q(g_i, g_j, r, \omega)$$

Values ranges from 0.05 to 3.7

$$co_invd = \sum_{g_i, g_j} \frac{c_q(g_i, g_j, r, \omega)}{1 + (g_i - g_j)^2}$$

Values ranges from 0.02 to 323

$$co_cov = \sum_{g_i, g_j} (g_i - \mu_i)(g_j - \mu_j) c_q(g_i, g_j, r, \omega)$$

Values ranges from 0.21 to 1.88

$$co_energy = \sum_{g_i, g_j} [c_q(g_i, g_j, r, \omega)]^2$$

Values ranges from 0.21 to 1.51

2. DCT based Parameters

$$f_dcEn = B_p[0,0] / \sum_{x,y} B_p[x,y]$$

$$f_horEn = \sum_y B_p[0,y] / \sum_{x,y} B_p[x,y]$$

$$f_verEn = \sum_x B_p[x,0] / \sum_{x,y} B_p[x,y]$$

$$f_diagEn = \sum_{\substack{x,y \\ x=y}} B_p[x,y] / \sum_{x,y} B_p[x,y]$$

where $B_p[x, y]$ is the DCT component at the angular frequencies m, n .

Values for f_dcEn range from -2040.98 to 737.28, Values for f_horEn range from -588.96 to 152.98, Values for f_verEn range from -231.24 to 107.85, Values for f_diagEn range from -2048.16 to 698.82

APPENDIX B:

