Arab J Sci Eng (2018) 43:4057–4070
https://doi.org/10.1007/s13369-017-2803-9

RESEARCH ARTICLE - SPECIAL ISSUE - COMPUTER ENGINEERING AND COMPUTER SCIENCE

Impact of Feature Selection Algorithms on Blind Image Quality Assessment

Imran Fareed Nizami[1] · **Muhammad Majid**[2] · **Hammad Afzal**[3] · **Khawar Khurshid**[1]

Received: 5 April 2017 / Accepted: 6 August 2017 / Published online: 21 August 2017

Abstract Blind image quality assessment (BIQA) is a challenging task in real-world problems due to unavailability of reference images. The performance of BIQA techniques is highly dependent on features used to assess the image quality. In the literature, different BIQA techniques have been proposed using a two-step approach, i.e., feature extraction in different domains and prediction of quality score using extracted features. However, optimum feature selection for these techniques has not been explored. This paper investigates the impact of feature selection algorithms on the performance of BIQA techniques. In contrast to existing techniques, the proposed methodology follows a three-step approach. Firstly, features are extracted using existing BIQA techniques. In the second step, feature selection algorithm is applied on the extracted features to reduce the number of features. The selected features are then utilized for prediction of a quality score in the third step. The proposed approach is evaluated for six BIQA techniques using five commonly used feature selection algorithms. Experimental results show that the feature selection algorithms not only reduces the number of features but also improves the performance of the state-of-the-art BIQA techniques.

Keywords Image quality assessment · Feature selection · Support vector regression · Perceptual quality · Machine learning · Computational intelligence

✉ Imran Fareed Nizami
12phdinizami@seecs.edu.pk

Muhammad Majid
m.majid@uettaxila.edu.pk

Hammad Afzal
hammad.afzal@mcs.edu.pk

Khawar Khurshid
khawar.khurshid@seecs.edu.pk

1 School of Electrical Engineering and Computer Science, National University of Sciences and Technology, Islamabad, Pakistan

2 Department of Computer Engineering, University of Engineering and Technology Taxila, Taxila, Pakistan

3 Military College of Signals, National University of Sciences and Technology, Islamabad, Pakistan

1 Introduction

With the development of communication systems and widespread use of handheld devices, exchange of images has become a common practice. Distortion in images is introduced due to imperfections in acquisition, compression for data reduction and transmission over lossy channels [1]. More than one billion images are shared over the Internet everyday; therefore, evaluating the quality of images has gained importance. Image quality assessment (IQA) aims to evaluate the quality of the image, which correlates with human perception. IQAs are broadly divided into two categories, i.e., subjective and objective quality assessment. Assessment performed by human observers is termed as subjective image quality assessment [2]. Hence, objective IQA is required that mimic the behavior of human observer to predict the quality score of images, which correlates with the mean observer score.

Objective image quality assessment techniques are broadly categorized into three groups, namely blind image quality assessment (BIQA), reduced reference (RR) and full reference (FR) IQA [3]. FR-IQA techniques are purely mathematical and measure the error pixel-wise that does not offer a proper contemplation of visual perception [4]. Peak signal-to-noise ratio, mean square error, structural similarity index and feature similarity index are few of the commonly used quality measures for FR-IQA [5]. In FR-IQA, both the reference and distorted images are required, thus limiting its application [6–11]. RR-IQA techniques attempt to estimate the perceived quality of an image using limited information about the reference and the distorted image. Similar to FR-IQA, RR-IQA techniques do not account the human perception while measuring the image quality [12–15]. The aim of BIQA techniques is to develop an insight into factors that can be used to model image quality perceptually without the reference image [1,16–21].

Natural images are highly structured and possess certain properties that are affected in the presence of distortion. Such properties are known as natural scene statistics (NSS) [22]. NSS- based IQA techniques assess the quality of images by measuring the deviation of NSS features of distorted images from that of the natural images. Various BIQA techniques have been proposed in the literature that assess the quality score of images using natural scene statistics (NSS)-based features from different domains [16,18,19,22–29]. Most of these BIQA techniques convert the color images to gray scale and then compute the quality score of the image. The conversion from color to gray scale makes the BIQA techniques computationally efficient and also applicable for grayscale images.

In [23], wavelet transform is used to extract features over three scales and three orientations and regression model to determine the quality score of images. In [24], discrete cosine transform (DCT)-based features are extracted for the assessment of image quality. In [25], local normalization of luminance features and their products in spatial domain are used over two scales for BIQA. In [27], spatial and spectral entropies are extracted by dividing the image into patches of 8×8. These features are given as an input to support vector regression (SVR) for the prediction of image quality score. Statistical features in curvelet domain are extracted based on the maximum value of log-histograms and the energies of scale and orientation [1]. In [21], NSS-based features using shearlet transform are used for BIQA, which localize distributed discontinuities and makes it suitable for assessing the quality score. The quality score is computed by taking the difference between the mean and standard deviation of natural and distorted images.

Recently, few BIQA techniques have been proposed that do not require training. In [30], spatial domain features are extracted, which are mapped to a multivariate gaussian distribution (MVG). The quality score is predicted based on the distance between distorted and natural images MVG. In [17], quality aware features are extracted in a two-step framework. Firstly, the distortion type is identified, and then, the quality score is assessed based on label transfer using annotated images.

The performance of BIQA techniques deteriorates in terms of low correlation with the subjective mean opinion score if redundant and irrelevant NSS features are extracted. To the best of our knowledge, feature selection that removes redundant and irrelevant features for BIQA has not been explored. Therefore, feature selection algorithm is required that reduces the number of features and improves the performance of BIQA techniques. This paper proposes a three-step approach for BIQA. In the first step, existing BIQA techniques are used for feature extraction. In the second step, feature selection is performed by five commonly used feature selection algorithms, which reduces the number of features for all distortion types for a particular BIQA technique. Lastly, selected features are used for the assessment of image quality. The contributions of this paper are

1. Impact of feature selection algorithms on state-of-the-art BIQA techniques are explored.
2. Feature selection step used in BIQA shows better performance in terms of higher correlation with mean observer scores and lower root-mean-square error values.
3. Feature selection algorithm reduces the number of features for existing BIQA techniques, which result in less computational expense and execution time.

The rest of the paper is organized as follows. Section 2 discuses the proposed methodology in detail. Section 3 presents the experimental results of five feature selection algorithms on six BIQA techniques. Finally, conclusion is given in Sect. 4.

2 Proposed Methodology

Figure 1 shows the proposed methodology for BIQA, which follows a three-step approach. In the first step, N number of features, i.e., F_E, are extracted from color or grayscale image using existing BIQA techniques. Feature selection is performed in the second step that selects M number of features, i.e., F_S from F_E, where $M \leq N$. In the last step, features are employed for the computation of image quality score. The detail of each step is as follows.

2.1 Feature Extraction

In the first step, features are extracted using BIQA techniques. Six state-of-the-art BIQA techniques are used in this paper. All the techniques used in this work are based on NSS. These techniques assess the quality score of images based on deviation of NSS properties from the natural images, represented by the respective features. In this work, the images used for quality assessment are color as well as gray scale. These images are distorted by different types of distortions, namely JPEG, JPEG2000, white noise and Gaussian blur. The distortion in each image is introduced for four different

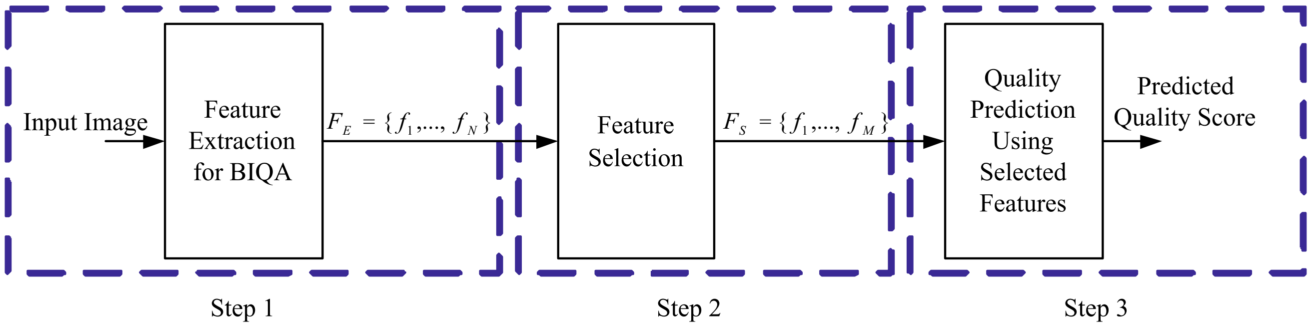

Fig. 1 Proposed methodology for BIQA using feature selection algorithm

levels, i.e., from low distorted to highly distorted image. The details of each BIQA technique used are as follows.

2.1.1 Distortion Identification-Based Image Verity and INtegrity Evaluation (DIIVINE)

DIIVINE extracts image features for BIQA in the first step and utilizes a regression model for predicting the quality score of image in the second step [22]. The technique uses a loose wavelet transform based on steerable pyramids to extract NSS features from images. The wavelet transform is considered across two scales and six orientations. Five group of features, namely spatial correlation, correlations across scales, orientation selective statistics, scale and orientation selective statistics and across orientation statistics, constitute a feature vector of length 88. DIIVINE uses these 88 features as input to SVR for the estimation of quality score.

2.1.2 Spatial–Spectral Entropy-Based Quality (SSEQ)

In [27], a hybrid technique for BIQA is introduced. The images are analyzed at three scales, i.e, low, middle and high. Bicubic interpolation is used during downsampling to avoid aliasing. SSEQ uses spatial and spectral entropies, which are computed by dividing the image into 8×8 patches. Then, features are sorted in ascending order and 60% of the central elements are selected for BIQA. In SSEQ, first the distortion type affecting the image is identified using 12 features, and then, these features are given as input to SVR for the computation of quality score.

2.1.3 Gradient Magnitude and Laplacian of Gaussian-Based IQA (GM-LOG)

In [26], Gaussian magnitude (GM) and Laplacian of Gaussian (LOG) marginal distributions are jointly normalized in an adaptive procedure called joint adaptive normalization (JAN) for assessing the quality of images. The technique follows a two-step approach, i.e., in the first step, 40 features are extracted, and in the second step SVR is used to estimate the quality score. The technique introduces a dependency index to assess and refine the relationship between GM and LOG statistics. The luminance discontinuity is used to describe the structural information by applying GM and LOG operator on the image.

2.1.4 Oriented Gradients Image Quality Assessment (OG-IQA)

In [16], gradient statistics are deeply explored relative to the surrounding and adaboosting neural network is used for computing the quality of images. Distortion in images usually results in blurring of gradient information. Gradient orientation information can augment gradient magnitude for BIQA. Anisotropic information present in images is modified by distortion to produce unnatural local anisotropic properties. OG-IQA exploits these properties to assess the quality of images. The features extracted over two scales are based on variances of gradient magnitude histogram, variances of relative gradient orientation histogram and variances of relative gradient magnitude histogram. The adaboosting neural network maps the image features to the quality score.

2.1.5 Distortion-Type Classification and Label Transfer (TCLT)

In [17], a two-step approach is adopted, i.e., distortion-type classification and label transfer. DCT, wavelet and spatial domain are utilized for feature extraction in each YCbCr channel. In DCT domain, intra-block skewness and global frequency band entropy are used as features. Exponential decay based on entropy and self-similarity across all scales are used as features in wavelet domain. In spatial domain, local binary patterns are used as features with 16 neighbors and radius of 2. Once the features are extracted, the quality score of the image is estimated by using label transfer from subjective quality labels of annotated images to the test image. TCLT does not require training as quality score prediction is based on label transfer.

2.1.6 Natural Image Quality Evaluator (NIQE)

In [30], quality aware statistical features are extracted in the spatial domain. NIQE is opinion and distortion unaware technique that does not require training for the estimation of quality score. The images are divided into $P \times P$ patches. The patches with sharp regions giving detail information are selected to construct a model of natural image statistics. Selected patches are used to extract 36 features. The extracted features are modeled by fitting them on a MVG distribution. The quality score is predicted by measuring a distance between the MVG of the distorted image with that of the undistorted pristine image.

2.2 Feature Selection

The features extracted by different BIQA techniques are subjected to feature selection. Five commonly used feature selection algorithms are employed for comparison. Feature selection algorithms select optimum features for BIQA that are most affected by the presence of distortion in the image. Therefore, features selection is performed based on the deviation of NSS properties of distorted image from natural images. The details of each feature selection algorithms are as follows.

2.2.1 Random Search (RS)

Random feature selection uses Las Vegas algorithm with randomness. It results in guiding the feature selection rapidly toward a correct solution [31]. From a total of N features, a subset of M features are selected. An inconsistency criterion is applied to check whether desired performance can be achieved by M features, which are less than currently selected features M_{best}. The inconsistency rate of M and M_{best} is compared and interchanged when $M < M_{best}$. The above- mentioned procedure is repeated $\max_{\text{tries}}$ times.

The inconsistency criterion of a set of features is determined as follows. Firstly, if the label of two instances does not match but otherwise they are identical, then they are considered as inconsistent. Secondly, the inconsistency count can be calculated by subtracting the total number of identical instances from the total number of classes. Lastly, the inconsistency rate is computed by adding up the inconsistency count and dividing it by total instances.

2.2.2 Incremental Wrapper Feature Subset Selection with Naive Bayes Classifier (IWSSENB)

In [32], IWSSENB uses ranking with the naive Bayesian classifier for feature selection. A filter based on Symmetric uncertainty (SU) is used to evaluate the predictive ability of each attribute. Attributes are ranked in descending order of SU and is given in [32] as.

$$\mathrm{SU}(x_i, Z) = 2\left(\frac{H(x_i) - H(x|Z)}{H(x_i) + H(Z)}\right), \tag{1}$$

where x_i is the ith instance of input, Z is the class and H is the entropy. IWSSENB validates the training set T using naive Bayesian classifier and fivefold cross-validation after dividing the training data into subsets. A relevance criterion based on $ttest$ value is computed over the accuracy for each fold and is used to decide whether a new attribute should be included in the selected subset.

2.2.3 Linear Forward Selection (LFS)

A commonly used sequential forward selection technique is used for linear selection. The technique starts by performing a simple hill climbing search [33]. Linear forward selection technique reduces the number of features in each forward step. It starts with an empty subset and evaluates each attribute for inclusion to the current feature subset. The attribute that helps achieve the top score is selected permanently. The search is terminated when no single attribute expansion can help in improving the current top score. As the number of evaluations increase exponentially with each step, the attributes which are considered in each step are limited by a user-defined constant. The user-defined constant is determined by adopting ORDERED-FS search algorithm using $P - fold$ cross-validation. ORDERED-FS randomly divides the data into training and testing sets.

Attributes for the algorithm are selected based on training data. The method performs P forward selections, one for each training set. The training data are utilized so that a decision can be drawn to select the attributes which can be incrementally included in each iteration for forward selection. The assessment of $P - best$ subsets is performed using the test data for a particular size of subset. The subset with highest average computed over P scores is selected as the optimal feature subset. Finally, forward selection is performed on the dataset to find a subset of features of attribute size P.

2.2.4 Particle Swarm Optimization (PSO)

Particle swarm optimization is a random population searching algorithm, which is being commonly used because it can perform nonlinear optimization effectively [34]. Each single candidate in PSO is known as particle. Each particle uses the memory gained by it in the swarm to find the best solution. A fitness value and velocity are associated with each particle. The fitness value is evaluated by a fitness function, whereas the velocity defines the direction, in which the particle moves.

The particle moves in the problem space over a wave of optimum particles. The initial swarm is selected randomly. On each iteration, every particle is updated by using *pbest* and *gbest*. For the purpose of classification, the population is taken as binary, i.e., {0, 1}. A bit value represented as 0 indicates that the feature is not selected, whereas a 1 in the bit value represents that the feature is selected. The population data are given as input to the support vector classifier (SVC) to obtain classification accuracy [35]. The search for best solution continues until the terminating criterion is reached.

2.2.5 Genetic Search (GS)

Genetic search is an optimization problem, which analyzes and optimizes a population, i.e., a set of solutions [36]. In genetic algorithm, the solution is represented by a sequence of 0ś and 1ś in the search space that are called chromosomes. The genetic algorithm allows these chromosomes to crossover or mutate. Crossover is a process through which offspring pair of chromosomes is produced. Mutation produces a chromosome in which some components are altered, but overall it is identical to the original chromosome. The optimization is done in several iterations called generations. In each generation, a new set of chromosomes are created. The next generation constitutes only the best chromosomes. Least costly feature subset is used by the genetic search that does not deteriorate the performance of the system below a certain level. The classifier error is used to evaluate the performance of feature subset.

A feasible feature subset is defined as one whose error rate falls below a feasible threshold (t). Subsets of feature with error rate below the feasible threshold are given a small reward. Feature subsets with error rate equal to the feasible threshold are evaluated according to the error rate. Subsets of feature with error rate above t but below $t + m$ are given a small penalty. Feature subsets with error rate above $t + m$ receive high penalty, where m is the error margin. The genetic function converges to optimized feature selection based on the above properties.

2.3 Quality Prediction Using Selected Features

In the third step, selected features are used for quality score prediction. The selected features are given as an input to SVR to predict the image quality for DIIVINE [22], SSEQ [27], GM-LOG [26] and OG-IQA [16]. The SVR model is given in [37] as

$$\psi(x) = \alpha\beta(x) + c, \tag{2}$$

where x is the input feature vector, β is the feature space and α is the weight corresponding to ith instance of x, i.e., β relates the input features to the feature space. Let the input features be denoted by x_i and the associated target value y_i. Then, the objective of SVR is to estimate support vector machine function such that the difference between the target value and the predicted value is minimized. Kernel function is used for all the computations of SVR because it offers the advantage of taking inner product without the need to construct the vector space explicitly. A radial function of order Q is given in [37] as

$$k(x) = \sum_{i=1}^{K} \alpha_i \frac{1}{(2\pi)^{\frac{Q}{2}} \sigma_i^{Q}} \exp\left(-\frac{||x - c_i||}{2\sigma_i^2}\right) + b, \tag{3}$$

where c_i denotes the center of the ith Gaussian basis function with standard deviation σ_i, α_i is the weight of the ith Gaussian basis function and b is the bias value.

TCLT [17] does not require training, and label transfer using annotated images is employed for quality score evaluation. Similarly, NIQE [30] does not require training and distance between the MVG distribution of distorted and natural images is used to evaluate the quality score. The distance is measured as [30]

$$D(v_1, v_2, \Sigma_1, \Sigma_2) = \sqrt{\left((v_1 - v_2)^{\mathrm{T}} \left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1} (v_1 - v_2)\right)}, \tag{4}$$

where v_1, v_2 are mean vectors and Σ_1, Σ_2 are covariance matrices of multivariate gaussian distribution.

3 Experimental Results

3.1 IQA Databases

Many subjective databases are available for the purpose of BIQA. Four commonly used IQA databases are selected to evaluate the performance of the proposed methodology, i.e., LIVE [38], TID2008 [39], CSIQ [40] and A57 [41]. There are 29 reference images in the LIVE database, which are affected by five types of distortions, namely, fast fading (FF), Gaussian blur (GB), JPEG2000 (JP2K), JPEG and white noise (WN). A total of 779 images with varying degree of distortions is present in the LIVE database. The TID2008 database consists of 25 reference images. There are totally 17 types of distortions in the TID2008 database, and each reference image is degraded at four different levels of each distortion type. The evaluators selected to assess the quality of images were selected from different countries, i.e., Ukraine, Italy and Finland and from diverse social levels, i.e., students, tutors and researchers. CSIQ database comprises of 30 reference images and five types of distortion types, i.e., JPEG, JP2K,

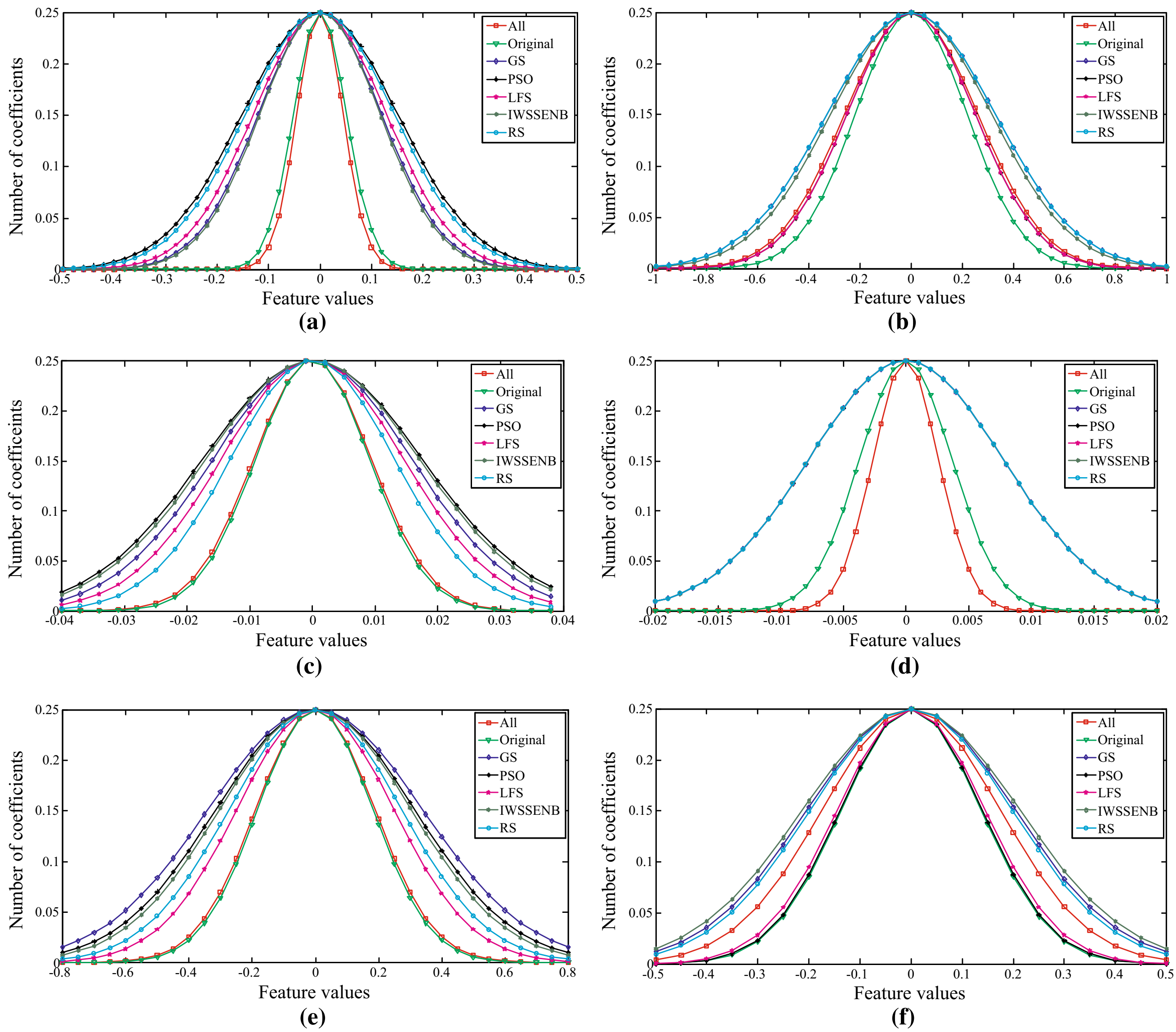

Fig. 2 Impact of feature selection algorithms on normalized feature histogram for BIQA techniques **a** DIIVINE, **b** SSEQ, **c** GM-LOG, **d** OG-IQA, **e** TCLT, **f** NIQE

contrast, WN and GB. Image quality assessment was performed by 35 observers on a LCD screen of 1920 × 1200 resolution. The A57 database consists of 3 reference images with 6 distortion types. Each distortion is degraded at 3 levels. There are 54 total images in the A57 database that have been evaluated by 7 observers. Four common distortions from the TID2008, CSIQ and A57 database are used for performance evaluation in this work, i.e., GB, JP2K, JPEG, WN.

3.2 Evaluation Criteria and Parameters

The quality score of an image affected by distortion is computed using SVR for DIIVINE [22], SSEQ [27], GM-LOG [26] and OG-IQA [16]. SVR requires training to assess the quality score of an image. Therefore, the dataset is divided into two non-overlapping groups, i.e., training and testing. The training and testing are repeated 1000 times by randomly selecting non-overlapping sets of training and testing images to nullify any bias due to selection of images. SVR parameters c and γ are selected by applying grid test for each BIQA technique, i.e., the parameters are optimized for each BIQA technique. LibSVM package is used for the implementation of SVR and SVC [42]. WEKA [43] is used for the application of five commonly used feature selection algorithms. Feature selection is performed by considering all the image features across four IQA databases for each BIQA technique, i.e., same features are used for all the databases for a particular BIQA technique.

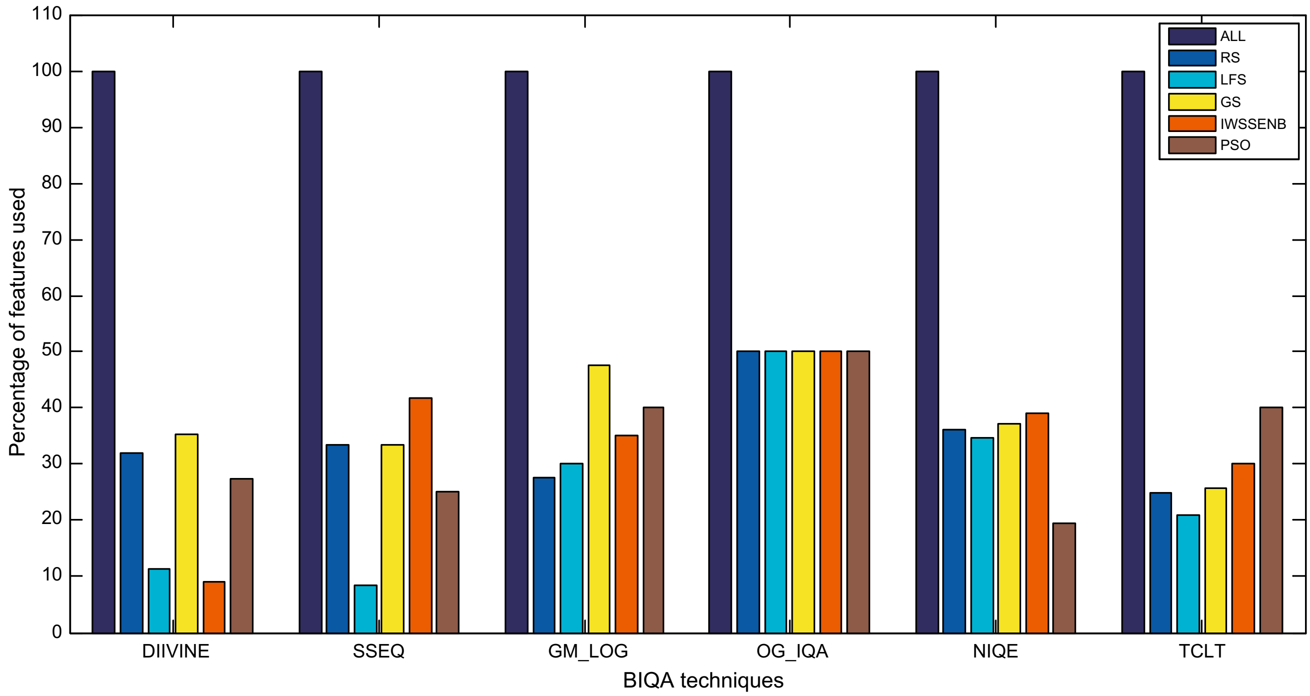

Fig. 3 Percentage of features selected using different feature selection algorithms for six BIQA techniques

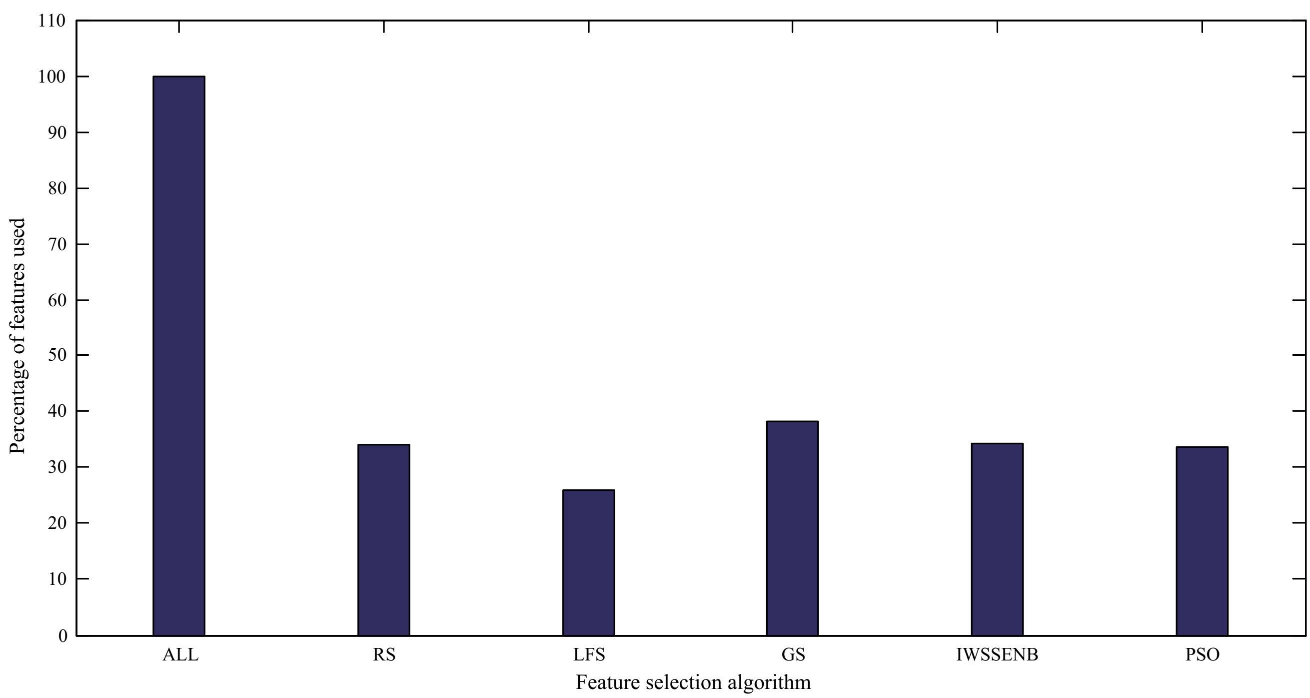

Fig. 4 Aggregate percentage of features selected using different feature selection algorithms for six BIQA techniques

For the performance comparison Spearman's ranked ordered correlation constant (SROCC), linear correlation constant (LCC), kendall correlation constant (KCC) and root-mean-squared error (RMSE) are computed. SROCC is calculated as [44]

$$\mathrm{SROCC} = 1 - \frac{6\Sigma d_i^2}{n(n^2-1)}, \tag{5}$$

where d_i is the difference between paired ranks and n is the total number of cases. The LCC is given as [45]

Table 1 Performance comparison of the proposed scheme with state-of-the-art schemes in terms of median value of SROCC for each individual distortion on LIVE, TID2008, CSIQ and A57 Databases

BIQA technique	Database	Distortion type	Feature selection algorithm					
			All	RS [31]	LFS [33]	GS [36]	IWSSENB [32]	PSO [34]
DIIVINE [22]	LIVE [38]	FF	**0.9022**	0.8695	0.8446	0.8446	0.8278	0.8783
		GB	**0.9631**	0.9172	0.8670	0.8670	0.8002	0.9360
		JP2K	0.6217	**0.8414**	**0.8234**	**0.8234**	**0.8127**	**0.8628**
		JPEG	**0.9037**	0.8766	0.7871	0.7871	0.7921	0.8742
		WN	**0.9828**	0.9818	0.9648	0.9648	0.9591	**0.9828**
	CSIQ [40]	GB	0.7139	**0.7838**	**0.7826**	**0.7826**	**0.7669**	**0.7893**
		JP2K	0.7801	**0.8185**	**0.8209**	**0.8209**	**0.8073**	**0.8140**
		JPEG	0.6274	**0.7526**	**0.6476**	**0.6476**	**0.6423**	**0.8216**
		WN	0.7227	**0.9306**	**0.8874**	**0.8874**	**0.8897**	**0.9259**
	TID2008 [39]	GB	0.5068	**0.8156**	**0.7023**	**0.7023**	**0.5982**	**0.8316**
		JP2K	0.7598	**0.8436**	**0.8378**	**0.8378**	**0.8045**	**0.8526**
		JPEG	0.4536	**0.7444**	**0.7203**	**0.7203**	**0.6534**	**0.8180**
		WN	0.5353	**0.8687**	**0.8656**	**0.8656**	**0.8511**	**0.8827**
	A57 [41]	GB	0.8223	**0.8346**	**0.8250**	**0.8292**	**0.8263**	**0.82221**
		JP2K	0.8547	0.8321	0.8252	**0.8671**	**0.8692**	**0.8562**
		JPEG	0.8135	0.8032	**0.8563**	**0.8321**	**0.8190**	0.8083
		WN	0.8351	**0.8621**	0.8210	**0.8589**	**0.84311**	**0.8421**
SSEQ [27]	LIVE [38]	FF	0.8384	**0.8778**	0.7862	**0.8778**	**0.8569**	0.7862
		GB	**0.9340**	0.9281	0.9030	0.9281	0.9261	0.9030
		JP2K	0.8692	**0.8956**	0.8349	**0.9365**	**0.8824**	0.8349
		JPEG2	0.8756	0.8745	0.8253	0.8745	**0.8949**	0.8253
		WN	0.9064	**0.9532**	**0.9133**	**0.9532**	**0.9591**	**0.9133**
	CSIQ [40]	GB	**0.8547**	0.8412	0.8238	0.8412	0.8414	0.8238
		JP2K	**0.8356**	0.8283	0.8216	0.8283	0.8300	0.8216
		JPEG	0.8283	**0.8616**	**0.8716**	**0.8616**	**0.8625**	**0.8716**
		WN	0.9123	**0.9257**	0.7437	**0.9257**	**0.9224**	0.7437
	TID2008 [39]	GB	**0.8421**	0.7925	0.8150	0.7925	0.7701	0.8150
		JP2K	0.8722	**0.8857**	**0.8782**	**0.8857**	**0.8872**	**0.8782**
		JPEG	**0.8075**	0.7459	0.7704	0.7459	0.7355	0.7704
		WN	0.8256	0.8238	0.3877	0.8238	**0.8607**	0.3877
	A57 [41]	GB	**0.8792**	0.8621	0.8560	0.8621	0.8851	0.8856
		JP2K	0.8821	**0.8863**	**0.8829**	**0.8863**	**0.8876**	**0.8829**
		JPEG	0.8463	**0.8521**	**0.8492**	**0.8521**	**0.8596**	**0.8492**
		WN	0.8369	**0.8531**	**0.8456**	**0.8531**	**0.8370**	**0.8456**
GM-LOG [26]	LIVE [38]	FF	0.8867	**0.8874**	**0.8872**	**0.8990**	**0.8966**	**0.9010**
		GB	0.8739	**0.9126**	**0.9133**	**0.8941**	**0.9222**	**0.9034**
		JP2K	0.8752	**0.8847**	**0.8834**	**0.8964**	**0.8887**	**0.8961**
		JPEG	0.8981	0.8826	0.8820	**0.9039**	0.8898	0.8862
		WN	0.9788	**0.9818**	**0.9837**	**0.9813**	**0.9808**	**0.9823**
	CSIQ [40]	GB	0.8220	**0.8616**	**0.8634**	**0.8690**	**0.8701**	**0.8638**
		JP2K	0.8683	**0.8843**	**0.8908**	**0.9146**	**0.9008**	**0.9186**
		JPEG	0.8723	**0.8832**	**0.8848**	**0.8852**	**0.8843**	**0.8928**
		WN	**0.9542**	0.9493	0.9499	0.9497	0.9515	0.9488
	TID2008	GB	0.8301	**0.8481**	0.7940	**0.8376**	0.8120	0.8263
		JP2K	0.8737	**0.9098**	**0.9053**	**0.8947**	**0.9038**	**0.8977**

Table 1 continued

BIQA technique	Database	Distortion type	Feature selection algorithm					
			All	RS [31]	LFS [33]	GS [36]	IWSSENB [32]	PSO [34]
		JPEG	0.8842	0.8767	**0.8851**	**0.9023**	**0.9008**	**0.9005**
		WN	**0.9054**	0.8876	0.8797	0.8947	0.9058	0.8980
	A57 [41]	GB	0.8651	**0.8663**	0.8492	**0.8722**	**0.8652**	**0.8695**
		JP2K	0.8231	**0.8451**	**0.8391**	**0.8321**	**0.8369**	**0.8692**
		JPEG	0.7945	**0.8021**	0.7900	**0.8331**	**0.8365**	**0.8120**
		WN	0.8523	0.8421	**0.8732**	**0.8763**	**0.8591**	0.8342
OG-IQA [16]	LIVE [38]	FF	**0.7830**	0.7783	0.7783	0.7783	0.7783	0.7783
		GB	0.8586	**0.8687**	**0.8687**	**0.8687**	**0.8687**	**0.8687**
		JP2K	0.8929	**0.8984**	**0.8984**	**0.8984**	**0.8984**	**0.8984**
		JPEG	0.7358	**0.7872**	**0.7872**	**0.7872**	**0.7872**	**0.7872**
		WN	0.9153	**0.9665**	**0.9665**	**0.9665**	**0.9665**	**0.9665**
	CSIQ [40]	GB	0.8590	**0.8705**	**0.8705**	**0.8705**	**0.8705**	**0.8705**
		JP2K	0.7580	**0.7717**	**0.7717**	**0.7717**	**0.7717**	**0.7717**
		JPEG	0.7566	**0.7976**	**0.7976**	**0.7976**	**0.7976**	**0.7976**
		WN	0.6934	**0.8185**	**0.8185**	**0.8185**	**0.8185**	**0.8185**
	TID2008 [39]	GB	**0.7820**	0.7759	0.7759	0.7759	0.7759	0.7759
		JP2K	**0.8872**	0.8857	0.8857	0.8857	0.8857	0.8857
		JPEG	0.7353	**0.7474**	**0.7474**	**0.7474**	**0.7474**	**0.7474**
		WN	0.5647	**0.7504**	**0.7504**	**0.7504**	**0.7504**	**0.7504**
	A57 [41]	GB	0.8663	**0.8736**	**0.8736**	**0.8736**	**0.8736**	**0.8736**
		JP2K	0.8121	**0.8223**	**0.8223**	**0.8223**	**0.8223**	**0.8223**
		JPEG	0.8871	**0.8893**	**0.8893**	**0.8893**	**0.8893**	**0.8893**
		WN	0.8331	**0.8421**	**0.8421**	**0.8421**	**0.8421**	**0.8421**
TCLT [17]	LIVE [38]	FF	0.9152	**0.9180**	0.9038	**0.9177**	**0.9533**	0.9038
		GB	0.9560	**0.9562**	0.9463	**0.9568**	**0.9787**	0.9537
		JP2K	0.9495	**0.9530**	**0.9560**	**0.9541**	**0.9580**	**0.9507**
		JPEG	**0.9710**	0.9620	0.9630	0.9690	0.9622	0.9627
		WN	**0.9871**	0.9860	0.9860	0.9863	0.9820	0.9848
	CSIQ [40]	GB	**0.8076**	**0.8180**	0.7851	**0.8076**	0.7921	0.7969
		JP2K	0.8206	**0.8251**	0.7981	0.8041	0.7951	0.8037
		JPEG	0.9072	**0.9174**	0.9020	**0.9214**	0.9014	**0.9152**
		WN	0.9047	**0.9310**	**0.9281**	**0.9299**	**0.9081**	**0.9339**
	TID2008 [39]	GB	0.9053	**0.9060**	0.8972	**0.9090**	**0.9092**	**0.9075**
		JP2K	0.8647	**0.8691**	0.8114	**0.8730**	**0.8752**	0.8278
		JPEG	0.8707	**0.8717**	0.8594	**0.8789**	**0.8735**	0.8677
		WN	0.8302	**0.8320**	**0.8593**	**0.8308**	**0.8353**	**0.8306**
NIQE [30]	LIVE [38]	FF	0.8473	**0.8660**	0.8231	0.8261	**0.9200**	0.8201
		GB	0.9340	**0.9394**	0.9335	**0.9365**	**0.9423**	0.9305
		JP2K	0.8914	**0.8920**	0.8860	0.8890	**0.9321**	0.8830
		JPEG	0.8960	0.8925	0.8740	0.8770	**0.9872**	0.8710
		WN	**0.9823**	**0.9823**	0.9798	**0.9828**	0.9793	0.9768
	CSIQ [40]	GB	0.9066	**0.9132**	0.8980	0.9310	**0.9150**	0.8950
		JP2K	0.8839	**0.8870**	0.8818	**0.9595**	**0.9313**	0.8788
		JPEG	0.8714	0.8590	0.8362	0.9792	**0.9220**	0.8332
		WN	**0.9608**	0.9560	0.9543	0.9573	0.9364	0.9513

Table 1 continued

BIQA technique	Database	Distortion type	Feature selection algorithm					
			All	RS [31]	LFS [33]	GS [36]	IWSSENB [32]	PSO [34]
	TID2008 [39]	GB	0.8475	**0.8647**	**0.8793**	**0.8823**	**0.8737**	**0.8763**
		JP2K	0.8737	**0.8902**	**0.8857**	**0.8887**	0.8562	0.8827
		JPEG	**0.8962**	0.8872	0.8740	0.8770	0.7677	0.8710
		WN	0.9090	**0.9097**	**0.9098**	**0.9128**	0.8563	0.9068
	A57 [41]	GB	0.8221	**0.8311**	**0.8290**	**0.8361**	**0.8356**	**0.8396**
		JP2K	0.8351	**0.8362**	**0.8436**	**0.8398**	0.8422	0.8363
		JPEG	0.8469	**0.8522**	**0.8563**	**0.8596**	**0.8566**	**0.8488**
		WN	0.8132	**0.8261**	**0.8222**	**0.8361**	**0.8356**	**0.8161**
Hit count			21	70	52	71	70	54

Bold value indicates better or at par performance of features selected using feature selection algorithms as compared to using all the features for a particular blind image quality assessment technique

$$\mathrm{LCC} = \frac{\sum_{i=1}^{n}(a_i - \bar{a})(l_i - \bar{l}))}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2}\sqrt{\sum_{i=1}^{n}(l_i - \bar{l})^2}}, \quad (6)$$

where a_i and l_i are the values in first dataset and second dataset, respectively, and $\bar{a}$ and $\bar{l}$ are mean values of a_i and l_i, respectively. KCC is given as [46]

$$\mathrm{KCC} = \frac{n_c - n_d}{n(n-1)/2}, \quad (7)$$

where n is the total number of observations, n_c is the number of concordant pairs and n_d is the number of discordant pairs. RMSE is given by [47]

$$\mathrm{RMSE} = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(x_{\mathrm{dmos}} - x_{\mathrm{score}})^2}, \quad (8)$$

where n is the total number of images, x_{dmos} is the mean observer score and x_{score} is the predicted quality score for image x.

3.3 Performance Comparison

A major portion of the literature for BIQA is based on NSS features. NSS- based BIQA techniques perform under the assumption that natural images possess certain properties that can be represented by NSS features. The presence of distortion affects these NSS properties. Image quality assessment can be performed by measuring the divergence of NSS properties between the distorted and natural image. All the BIQA techniques employed in this work are based on NSS, and each extracted feature represents NSS properties of the image as feature selection algorithms select a subset of NSS features that belong to existing BIQA techniques. Therefore, impact of selected features on NSS needs to be addressed. For this purpose normalized histogram for features of each BIQA technique over all the distortion types, all the databases and feature selection algorithms is shown in Fig. 2.

It can be observed that the NSS properties are preserved after feature selection. Moreover, the NSS properties of the undistorted and distorted images are not only different, but also the difference between the NSS properties of natural and distorted images increases after feature selection. Feature selection algorithms select optimum features for BIQA that are most affected by the presence of distortion in the image. The selected NSS features can more effectively represent the deviation of NSS properties of distorted image from natural images that is validated in Fig. 2. The impact of feature selection on BIQA can be evaluated by the considering that the feature selection algorithm with the highest deviation in NSS properties from the natural images gives the best performance for the respective BIQA technique.

In order to illustrate that the feature selection algorithms reduce the number of features for different BIQA techniques, the percentage of features used by different feature selection algorithms as compared to using all the NSS features is shown in Fig. 3. It can be observed that all feature selection algorithms reduces the number of features for each BIQA technique. The largest reduction in features is achieved by incremental wrapper feature subset selection with naive Bayes classifier for DIIVINE, linear forward selection for SSEQ, random search for GM-LOG, all the five feature selection techniques for OG-IQA, particle swarm optimization for NIQE and linear forward selection for TCLT. The largest reduction of 91.67% is achieved for SSEQ using linear forward selection algorithm. Figure 4 shows the overall percentage of features selected by different feature selection algorithms. It can be observed that highest amount of feature reduction is achieved by linear forward search algorithm. Thus, linear forward selection algorithm ranks top feature

Table 2 Overall performance comparison of feature selection algorithms on different BIQA techniques for LIVE, TID2008, CSIQ and A57 databases

BIQA technique	Feature selection technique	SROCC	LCC	KCC	RMSE
DIIVINE [22]	All	0.7491	0.7321	0.5953	10.072
	RS [31]	0.8521	0.8627	0.6807	8.1890
	LFS [33]	0.8370	0.8498	0.6654	9.2077
	GS [36]	0.8150	0.8246	0.6422	9.4155
	IWSSENB [32]	0.7927	0.8012	0.6178	9.5863
	PSO [34]	0.8676	0.8785	0.6950	7.7902
SSEQ [27]	All	0.8643	0.8840	0.6833	7.6782
	RS [31]	0.8698	0.8952	0.7078	7.1303
	LFS [33]	0.8095	0.8290	0.6484	8.7321
	GS [36]	0.8698	0.8852	0.6978	7.1303
	IWSSENB [32]	0.8695	0.8943	0.7075	7.1787
	PSO [34]	0.8095	0.8290	0.6484	8.7321
GM-LOG [26]	All	0.8870	0.8947	0.7184	7.1671
	RS [31]	0.8931	0.9113	0.7370	6.5506
	LFS [33]	0.8932	0.9112	0.7334	6.0210
	GS [36]	0.9000	0.9185	0.7426	5.9813
	IWSSENB [32]	0.9006	0.9268	0.7506	5.6388
	PSO [34]	0.9013	0.9186	0.7440	5.2578
OG-IQA [16]	All	0.7928	0.7970	0.6190	11.398
	RS [31]	0.8232	0.8203	0.6494	10.415
	LFS [33]	0.8232	0.8203	0.6494	10.415
	GS [36]	0.8232	0.8203	0.6494	10.415
	IWSSENB [32]	0.8232	0.8203	0.6494	10.415
	PSO [34]	0.8232	0.8203	0.6494	10.415
TCLT [17]	All	0.8730	0.8856	0.7381	7.0719
	RS [31]	0.8862	0.8882	0.7457	6.5321
	LFS [33]	0.8742	0.8879	0.7407	7.0658
	GS [36]	0.8767	0.8794	0.7347	6.8446
	IWSSENB [32]	0.8816	0.8860	0.7392	6.6521
	PSO [34]	0.8759	0.8890	0.7465	6.9939
NIQE [30]	All	0.8999	0.8992	0.7437	7.7754
	RS [31]	0.9022	0.9151	0.7517	7.6276
	LFS [33]	0.8920	0.8992	0.7368	7.8142
	GS [36]	0.9092	0.9239	0.7581	7.5079
	IWSSENB [32]	0.9159	0.9266	0.7603	7.1371
	PSO [34]	0.8890	0.8964	0.7344	7.9275

selection algorithm when feature reduction is considered in all of the BIQA techniques.

In this set of experiment, the proposed three-step approach is evaluated in terms of statistical parameters. For this purpose, all the BIQA techniques are trained separately using features selected from feature selection algorithms. The SROCC score for each distortion type over four databases, three color, i.e., LIVE, TID2008 and CSIQ and one gray scale, i.e., A57 and five feature selection algorithms for DIIVINE, SSEQ, GM-LOG, OG-IQA and NIQE techniques is presented in Table 1. TCLT does not convert the color images into gray scale for the prediction of quality score; therefore, TCLT results are only reported for color databases in Table 1. It is evident that all feature selection algorithms improve the performance of the selected BIQA techniques over majority of the distortion types. Hit count in Table 1 indicates how many times each feature selection algorithm performs better or equal than the state-of-the-art techniques using all features. The highest hit count of 71 is achieved for genetic search feature selection algorithm. The lowest hit count of 52 is achieved for linear forward selection algorithm, which is still much higher than the hit count of 21 when using

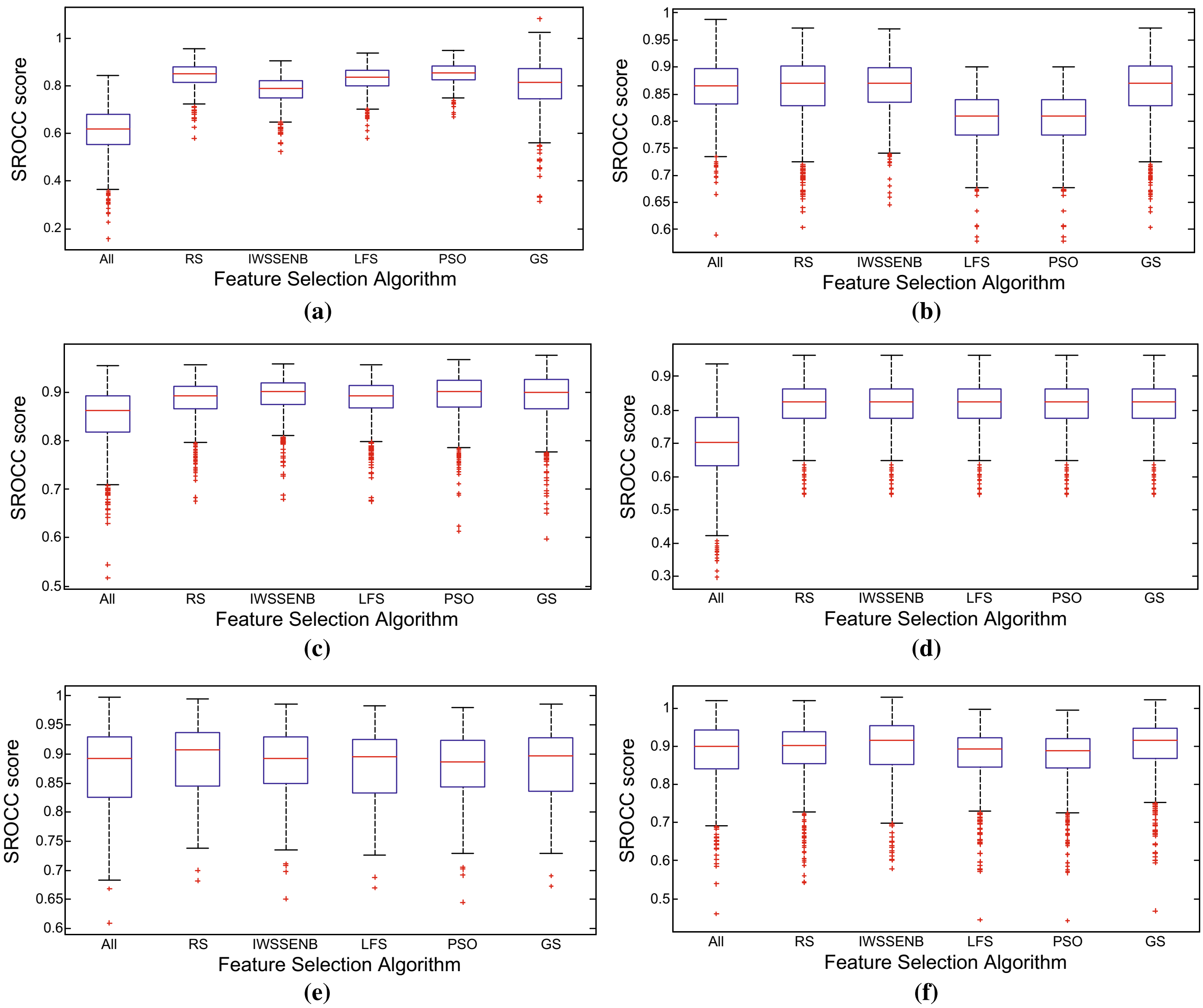

Fig. 5 Box plot of SROCC using feature selection algorithms for different BIQA techniques averaged over all the databases **a** DIIVINE, **b** SSEQ, **c** GM-LOG, **d** OG-IQA, **e** TCLT, **f** NIQE

all the features. It is evident from Table 1 that the proposed methodology improves the performance of BIQA techniques on color as well as grayscale images.

Table 2 shows an overall performance comparison of the proposed method with state-of-the-art BIQA techniques. All the feature selection algorithms give better results for each BIQA technique except linear forward selection and particle swarm optimization on SSEQ and NIQE. As all the BIQA techniques used are based on NSS; therefore, the feature selection algorithm, which selects features with the highest deviation from natural images, gives the best performance. Table 2 validates the improvement in performance using feature selection algorithms in a similar way as depicted in Fig. 2.

Figure 5 shows a box plot of averaged SROCC score for feature selection algorithms over different BIQA techniques for four IQA databases. It can be observed that the feature selection algorithms improves the performance of BIQA techniques using less number of NSS features. It is also seen that the interquartile range that represents the deviation in SROCC scores is also reduced by using feature selection algorithms as compared to using all features. The quality score prediction of BIQA techniques shows higher correlation with MOS after feature selection algorithm is applied.

Computational time of BIQA technique has a significant importance. Therefore, execution time of each BIQA technique using all features and the execution time when features are selected by feature selection algorithms is shown in Table 3. It can be observed that the execution time is decreased for all the BIQA techniques. IWSSENB performs the best for DIIVINE, linear forward selection for SSEQ, random search for GM-LOG, particle swarm optimization for NIQE and linear forward selection for TCLT. The maximum reduction in time is 91.66% for SSEQ using linear forward

Table 3 Impact of feature selection algorithm on different BIQA techniques in terms of execution time in seconds

Feature selection algorithm	BIQA technique					
	DIIVINE [22]	SSEQ [27]	GM-LOG [26]	OG-IQA [16]	TCLT [17]	NIQE [30]
All	28.2	1.71	0.43	0.32	0.1724	4.89
RS [31]	8.97	0.57	0.12	0.16	0.0427	1.76
LFS [33]	3.20	0.15	0.13	0.16	0.0359	1.69
GS [36]	9.93	0.57	0.20	0.16	0.0441	1.81
IWSSENB [32]	2.56	0.74	0.15	0.16	0.0516	1.9
PSO [34]	7.69	0.44	0.17	0.16	0.0690	1.53

selection. It can be concluded from Tables 2 and 3 that the feature selection algorithms which take the least amount of time in execution show the least improvement in performance in terms of evaluation parameters, i.e., SROCC, LCC, KCC and RMSE. Furthermore, it can be observed from Fig. 3 and Table 3 that the feature selection algorithm which selects fewer number of features requires lesser number of computations resulting in reduced execution time of the model. The experimental results show that feature selection improves the performance of existing BIQA techniques with higher correlation scores, lower RMSE values and less execution time as compared to using all features.

4 Conclusion

Blind image quality assessment is a difficult task due to the absence of the reference image. The performance of BIQA techniques is directly related to features used to assess the quality of the image. Redundant and irrelevant features can degrade the performance of BIQA techniques in terms of low correlation with the mean observer score and large execution time. In this paper, three-step approach is proposed that incorporates feature selection in the existing BIQA framework. The performance is evaluated on three-color and one grayscale IQA database for six BIQA techniques. The experimental results show that the performance of BIQA techniques improves in terms of higher correlation with the mean observer score and less execution time when feature selection is performed. It is observed that feature selection algorithm improves the overall performance of all BIQA techniques. But the least amount of improvement in overall performance in terms of SROCC, LCC, KCC and RMSE is observed for those feature selection algorithms, which reduces largest percentage of features and have least computational time. The proposed three-step approach is applicable to and performs well for color, grayscale images and BIQA techniques that require training for prediction of image quality score, as well as for quality aware feature-based BIQA techniques.

References

1. Liu, L.; Dong, H.; Huang, H.; Bovik, A.C.: No-reference image quality assessment in curvelet domain. Signal Process. Image Commun. **29**(4), 494–505 (2014)
2. Hong, R.; Pan, J.; Hao, S.; Wang, M.; Xue, F.; Wu, X.: Image quality assessment based on matching pursuit. Inf. Sci. **273**, 196–211 (2014)
3. Chetouani, A.; Beghdadi, A.; Deriche, M.: A hybrid system for distortion classification and image quality evaluation. Signal Process. Image Commun. **27**(9), 948–960 (2012)
4. He, L.; Gao, F.; Hou, W.; Hao, L.: Objective image quality assessment: a survey. Int. J. Comput. Math. **91**(11), 2374–2388 (2014)
5. Wang, Z.; Bovik, A.: Mean squared error: love it or leave it? A new look at signal fidelity measures. IEEE Signal Process. Mag. **26**(1), 98–117 (2009)
6. Bampis, C.G.; Goodall, T.R.; Bovik, A.C.: Sampled efficient full-reference image quality assessment models. In: 2016 50th Asilomar Conference on Signals, Systems and Computers, pp. 561–565. IEEE (2016)
7. Oszust, M.: Full-reference image quality assessment with linear combination of genetically selected quality measures. PLoS ONE **11**(6), e0158333 (2016)
8. Saha, A.; Wu, Q.J.: Full-reference image quality assessment by combining global and local distortion measures. Signal Process. **128**, 186–197 (2016)
9. Wen, Y.; Li, Y.; Zhang, X.; Shi, W.; Wang, L.; Chen, J.: A weighted full-reference image quality assessment based on visual saliency. J. Vis. Commun. Image Represent. **43**, 119–126 (2017)
10. Deriche, M.: An image quality index based on mutual information and neural networks. Arab. J. Sci. Eng. **39**(3), 1983–1993 (2014)
11. Lahoulou, A.; Bouridane, A.; Viennet, E.; Haddadi, M.: Full-reference image quality metrics performance evaluation over image quality databases. Arab. J. Sci. Eng. **38**(9), 2327–2356 (2013)
12. Lai, M.J.: Reduced reference image quality assessment, Ph.D. thesis, UTAR (2016)
13. Golestaneh, S.; Karam, L.J.: Reduced-reference quality assessment based on the entropy of DWT coefficients of locally weighted gradient magnitudes. IEEE Trans. Image Process. **25**(11), 5293–5303 (2016)
14. Wu, J.; Lin, W.; Shi, G.; Li, L.; Fang, Y.: Orientation selectivity based visual pattern for reduced-reference image quality assessment. Inf. Sci. **351**, 18–29 (2016)
15. Balanov, A.; Schwartz, A.; Moshe, Y.: Reduced-reference image quality assessment based on DCT subband similarity. In: 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX), pp. 1–6. IEEE (2016)
16. Liu, L.; Hua, Y.; Zhao, Q.; Huang, H.; Bovik, A.C.: Blind image quality assessment by relative gradient statistics and adaboosting neural network. Signal Process. Image Commun. **40**, 1–15 (2016)

17. Wu, Q.; Li, H.; Meng, F.; Ngan, K.N.; Luo, B.; Huang, C.; Zeng, B.: Blind image quality assessment based on multichannel feature fusion and label transfer. IEEE Trans. Circuits Syst. Video Technol. **26**(3), 425–440 (2016)
18. Xu, J.; Ye, P.; Li, Q.; Du, H.; Liu, Y.; Doermann, D.: Blind image quality assessment based on high order statistics aggregation. IEEE Trans. Image Process. **25**(9), 4444–4457 (2016)
19. Li, Q.; Lin, W.; Xu, J.; Fang, Y.: Blind image quality assessment using statistical structural and luminance features. IEEE Trans. Multimed. **18**(12), 2457–2469 (2016)
20. Lu, P.; Li, Y.; Jin, L.; Han, S.: Blind image quality assessment based on wavelet power spectrum in perceptual domain. Trans. Tianjin Univ. **22**(6), 596–602 (2016)
21. Lu, W.; Xu, T.; Ren, Y.; He, L.: Statistical modeling in the shearlet domain for blind image quality assessment. Multimed. Tools Appl. **75**(22), 14417–14431 (2016)
22. Moorthy, A.K.; Bovik, A.C.: Blind image quality assessment: from natural scene statistics to perceptual quality. IEEE Trans. Image Process. **20**(12), 3350–3364 (2011)
23. Moorthy, A.K.; Bovik, A.C.: A two-step framework for constructing blind image quality indices. IEEE Signal Process. Lett. **17**(5), 513–516 (2010)
24. Saad, M.A.; Bovik, A.C.; Charrier, C.: Blind image quality assessment: a natural scene statistics approach in the DCT domain. IEEE Trans. Image Process. **21**(8), 3339–3352 (2012)
25. Mittal, A.; Moorthy, A.K.; Bovik, A.C.: No-reference image quality assessment in the spatial domain. IEEE Trans. Image Process. **21**(12), 4695–4708 (2012)
26. Xue, W.; Mou, X.; Zhang, L.; Bovik, A.C.; Feng, X.: Blind image quality assessment using joint statistics of gradient magnitude and laplacian features. IEEE Trans. Image Process. **23**(11), 4850–4862 (2014)
27. Liu, L.; Liu, B.; Huang, H.; Bovik, A.C.: No-reference image quality assessment based on spatial and spectral entropies. Signal Process. Image Commun. **29**(8), 856–863 (2014)
28. Zhou, W.; Yu, L.; Qiu, W.; Zhou, Y.; Wu, M.: Local gradient patterns (LGP): an effective local-statistical-feature extraction scheme for no-reference image quality assessment. Inf. Sci. **397**, 1–14 (2017)
29. Nizami, I.F.; Majid, M.; Khurshid, K.: Efficient feature selection for blind image quality assessment based on natural scene statistics. In: 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), pp. 318–322. IEEE (2017)
30. Mittal, A.; Soundararajan, R.; Bovik, A.C.: Making a completely blind image quality analyzer. IEEE Signal Process. Lett. **20**(3), 209–212 (2013)
31. Liu, H.; Setiono, R. et al.: A probabilistic approach to feature selection-a filter solution. In: ICML, Citeseer, vol. 96, pp. 319–327 (1996)
32. Bermejo, P.; Gámez, J.A.; Puerta, J.M.: Speeding up incremental wrapper feature subset selection with naive Bayes classifier. Knowl. Based Syst. **55**, 140–147 (2014)
33. Gutlein, M.; Frank, E.; Hall, M.; Karwath, A.: Large-scale attribute selection using wrappers. In: IEEE Symposium on Computational Intelligence and Data Mining, CIDM'09, pp. 332–339. IEEE (2009)
34. Kennedy, J.: Particle swarm optimization. In: Encyclopedia of Machine Learning, pp. 760–766. Springer (2011)
35. Sharkawy, R.; Ibrahim, K.; Salama, M.; Bartnikas, R.: Particle swarm optimization feature selection for the classification of conducting particles in transformer oil. IEEE Trans. Dielectr. Electr. Insul. **18**(6), 1897–1907 (2011)
36. Siedlecki, W.; Sklansky, J.: A note on genetic algorithms for large-scale feature selection. Pattern Recogn. Lett. **10**(5), 335–347 (1989)
37. Smola, A.J.; Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**(3), 199–222 (2004)
38. Sheikh, H.R.; Sabir, M.F.; Bovik, A.C.: A statistical evaluation of recent full reference image quality assessment algorithms. IEEE Trans. Image Process. **15**(11), 3440–3451 (2006)
39. Ponomarenko, N.; Lukin, V.; Zelensky, A.; Egiazarian, K.; Carli, M.; Battisti, F.: TID 2008-a database for evaluation of full-reference visual quality assessment metrics. Adv. Mod. Radioelectron. **10**(4), 30–45 (2009)
40. Larson, E.C.; Chandler, D.M.: Most apparent distortion: full-reference image quality assessment and the role of strategy. J. Electron. Imaging **19**(1), 011006–011006 (2010)
41. Chandler, D.; Hemami, S.: A57 database (2007)
42. Chang, C.C.; Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. (TIST) **2**(3), 1–27 (2011)
43. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explor. Newslett. **11**(1), 10–18 (2009)
44. Zar, J. H.: Spearman rank correlation. In: Encyclopedia of Biostatistics, Wiley (1998)
45. Benesty, J.; Chen, J.; Huang, Y.; Cohen, I.: Pearson correlation coefficient. In: Noise Reduction in Speech Processing, pp. 1–4. Springer (2009)
46. Abdi, H.: The Kendall rank correlation coefficient. In: Encyclopedia of Measurement and Statistics, pp. 508–510. Sage, Thousand Oaks (2007)
47. Willmott, C.J.; Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. **30**(1), 79–82 (2005)