Journal of Advanced Research in Applied Sciences and Engineering Technology 57, Issue 2 (2026) 301-315

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage: https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Enhanced Perceptual Feature Extraction for Blind Image Quality Assessment using Extreme Learning Machine

Nutveesa Verak[1], Phaklen Ehkan[1,*], Ruzelita Ngadiran[1], Suwimol Jungjit[2], Fazrul Faiz Zakaria[1], Mohd Nazri Mohd Warip[1], Mohd Zaizu Elyas[1]

[1] Department of Computer, Faculty of Electronic Engineering & Technology, Universiti Malaysia Perlis, 02060 Arau, Perlis, Malaysia
[2] Department of Computer and Information Technology, Faculty of Science, Thaksin University, Phatthalung Campus, Phatthalung 93210, Thailand

ABSTRACT

In recent years, blind image quality assessment (BIQA) has become a major research topic with the practical applications as it promises the effective outcome as compared to the others image quality metrics. In situation where the reference image is unavailable, BIQA is the most suitable approach. However, the state- of-the-art image quality assessment (IQA) metrics is specific to certain types of distortion and not consistent with human perception. These existing BIQA do not consider for human visual characteristics impact on image content and specific to certain types of distortion. To overcome the problem, the research study proposed perceptual based features during initial data collection and combine with pooling algorithm based on Extreme Learning Machine (ELM). This is to match human perception that can see noticeable difference at certain frequency range and conclude the value similar to neuron. The approach aims to develop the lifting wavelet-based feature extraction with the aid of extreme learning machine (ELM) algorithm that able to enhance the image quality, extract the significant features from the image and reduce noise to the maximum extent possible. Unlike discrete wavelet transform (DWT), lifting wavelet transform (LWT) provides low computational complexity as it does not require convolution, dilation, and translation of the traditional mother wavelets. Besides, the pooling strategy of ELM is employed to overcome the limitations of previous pooling techniques such as neural networks (NNs) and support vector regression (SVR). This is because ELM has higher learning accuracy with faster learning speed. The proposed approach is verified on several type of image database that consist of different distortion type to make it general based BIQA. Based on experimental result, it proved that the approach has good performance than other features in terms of accuracy, specificity, and sensitivity. The outcome of this work will be essential in various image processing applications such as optimizations of image enhancement in medical for tumour or cancer detection, image watermarking for security detection, image coding and compression or image forensic. In biodiversity monitoring, image enhancement plays an important role for tracking and data collection.

Keywords:

Blind image quality assessment (BIQA); Extreme machine learning (ELM); Lifting wavelet transform (LWT)

* *Corresponding author.*
E-mail address: phaklen@unimap.edu.my

https://doi.org/10.37934/araset.57.2.301315

1. Introduction

The image quality assessment (IQA) plays a significant role in many applications with a rapid proliferation of digital and communication technology for the past few years [1-3]. It is used to measure the quality degradation of an image and employed to estimate and compare the performance of different processing system. An IQA has ability to predict the perceived image without human's intervention and its metrics are in demand for measuring the perceived quality of the multimedia content. An evaluation based on subjective viewing is the most reliable way for IQA in most cases since the human visual system (HVS) is the ultimate receiver of sensory information [4-6]. IQA can be classified into subjective and objective forms [7]. However, the subjective form evaluation is too cumbersome, time consuming, and suffers from high cost to be used in real-time and automated system [8]. On the other hand, it also needs the suitable degree of correlation with the subjective result. This is because it pays no attention to the image structure and not corrected well with the human perception. Therefore, many researchers have developed numerous types of HVS-based IQA metrics in the past [25].

A general agreement of the objective quality assessment can be broadly classified into full-reference (FR), reduced-reference (RR) and no-reference (NR) or blind IQA (BIQA) [9,10]. The full-reference metric required access to the reference image to evaluate the quality of the distorted images [11,12]. However, reduced-reference metric uses some features extracted from the reference image, while no- reference method assumes completely no access to the reference [13]. The peak signal-to-noise ratio (PSNR) and mean squared error (MSE) are full-reference (FR) objective quality assessment metrics that widely used due to its simplicity in mathematical and clear physical meanings. However, these metrics seem to have good performance for low content distortions and not correlated well with human perceptual measures. There are numerous developed FR- IQA models had been introduced by researchers based on various mechanism such as image statistics, image structure and human visual system (HVS). The structural similarity (SSIM) index, and multi-scale SSIM (MS-SSIM) are the example of objective metric based on the HVS properties [14,15] that have evaluated from previous metrics. However, high correlation rate with human perception measures is existing in these IQA models. Under most circumstances, there are unavailable of partial or full reference image information in practical application. In such case, no-reference metrics is most useful technique in applications because of an original image is not available.

Additionally, the efficiency of the image quality assessment can be enhanced with the aid of the feature pooling algorithm. The distortion factors such as contrast, luminance and gradient should be pooled or weighted together for the overall image quality gauging [8]. The simple summation or weighted multiplication are the simplest distortion factors use in conventional IQA metrics such as PSNR and SSIM. Despite that, such pooling technique seems to be ad-hoc with the limited theoretical grounds and no general method available for this. Fortunately, learning-based pooling technique can overcome these limitations. It can be applied to deduce the complexity of mathematical functions to model the relationship of the distortion effect variable with the image quality. Presently, the development of the no-reference metrics in the literature mostly is based on the neural networks (NNs) that using learning algorithms. Among the neural networks, the extreme learning machine (ELM) has been used for the evaluation of image due to the ability of avoiding randomly of weights and biases before training. For instant, modified metrics to measure the quality of the JPEG image based on two schemes such as K-fold selection scheme and real-coded genetic algorithm to select the input of weights and biases value [16]. The ELM tends to have good generalization performance and higher learning speed as compared to the support vector machine

(SVM) method. When ELM is used as feature mapping, the subjective evaluation scores of human beings are integrated into the trained model, so the approach metric is more stable with the human perception as compare with the conventional feature mapping methods.

This paper aimed to formulate perceptual based features using lifting wavelet transform as input into feature extraction and implement Extreme Learning Machine as a pooling agent that generate the objective IQA values. The paper is organized as follows. In section 1, the detail of the introduction proposed BIQA method is discussed. Section 2 represented the quality assessment related works while section 3 illustrated the feature of the BIQA databases used the detail of quality calibration and statistical method used for measuring the image quality assessment. The experimental result is then compared and analysed in section 4, followed by the conclusion in last section.

2. Related Works

In this section, the algorithms or techniques that have been used in perceptual image extraction are discussed. Most research related to image quality assessment has been done in last few decades [13,30]. In previous research, the predicted distortion can be based on the analysis of the image such as natural scene statistic [18,19,26] or transform domain [9,20-23] of the images. In year 2019, Athar and Wang [27] proposed alternative complementary framework based on degradation of structural information to extract the contrast, luminance, and structural information assuming human visual system is highly sensitive to these distortions, namely as structural similarity index (SSIM). However, this application scope does not restrict to image processing and metric is mathematically more cumbersome than previous mean squared error (MSE) method. Another study proposed BLIINDS-II metric that use different set of samples DCT statistics with no statistical modelling and generalized based on NSS of block DCT coefficients and transform model parameters into feature used [28,29]. According to Moorthy and Bovik [18] recognized DIIVINE that deploys NSS model of image wavelet coefficients which divided into two distortion identification and distortion-specific quality assessment stage in year 2011.

The technique described in Mittal's [19] using in spatial domain for feature extraction in no-reference IQA which namely as blind image spatial quality evaluator (BRISQUE). It uses scene statistics of locally normalized luminance coefficients to quantify the possible losses of the "naturalness" in the LIVE [33], TID [31] and CSIQ [32] image. This is due to the presence of the distortions which leading to a holistic measure of the quality. The BRISQUE function extract the NSS features from the distorted image and using support vector regression to predict the quality score. The features employed by this model are generally invariant to distortion. When tested on the different types of image distortions in standard IQA databases, it will show the high prediction performance correlated with human perceptual measurement. Then, he improved the metric with "completely blind" natural image quality evaluator (NIQE) which did not use human opinion score for training [7].

With resurgence of the machine learning research, many learning based IQA are proposed in recent year. Huang *et al.,* [36] proposed the distortion classification and label transfer (TCLT) method following by two-stage framework used in DIIVINE IQA. The DIIVINE consists of two stages: probabilistic distortion identification and distortion-specific quality assessment [19]. The author claimed that this metric able to access the quality of distorted image across multiple distortion categories. However, DIIVINE cannot compute the specific distortion features such as blocking but extract the statistical features which resulting to abroad range of distortion measurements. Abdul Manap *et al.,* [34] proposed PATCH-IQA to overcome the limitation of using intensive training phase to optimize the regression parameters and reduce the use of human perceptual measures based on a training set of distorted images. Author used the alternating BIQA model that using the nearest neighbour methods which have virtually zero training cost to predict the image quality. This model is focusing on learning framework that needs no explicit training phase which previous BIQA models is often required. The work shows competitive result based on input from different spatial domain features. Nevertheless, PATCH-IQA facing the computational complexity problem when the new types of distortion are introducing in the metric. The increasing in the datasets size leading to the high memory and require long processing time. The latest research work by Junyong You and Jari Korhonen [35] introduced transformer in image quality assessment that extracted by convolution neural networks (CNN) using shallow Transformer encoder on the top of a feature map. However, this method only can adapt to certain type of image contents, resolutions, and distortion types.

Although previously works introduce many features for evaluation, the possibility of finding the feature that match to human vision system is still challenging research. The research effort is still required in developing suitable design of perceptual model for image assessment. An approach also required to be computationally efficient as multimedia material are extensive and real-time assessments are desired for online applications [17]. Therefore, the quality metrics that consider both lower-level features and high-level features should be modestly considered [11]. In this paper, a novel leaning- based no-reference IQA metric is proposed, and the contributions are described as follows.

i. The Wavelet based on lifting scheme is exploited as feature extraction. The property of Lifting scheme-based creates the possibilities to has fast implementation as it makes use of similarities between high pass and low pass filters as compared to DWT. Unlike others, Lifting Wavelet Transform (LWT) can directly construct from time series signal and does not require convolution, translation, or dilation of traditional mother wavelets. Moreover, it allows in place calculation that aids to save the auxiliary memory because the input image is substituted by its transform. Consequently, it provides low computational complexity and more flexible as compared to the non-lifted wavelet transform, making its ability to be more stable for IQA metric.
ii. A developing machine learning scheme is adopted as distortion effects pooling agent toward obtaining the better regression results than standard method such as weighted summation, Euclidean distance, etc [24]. With the aids of Extreme Learning Machine (ELM), the human visual system toward perceptual IQA is incorporated with the distortion pooling process to map feature vectors onto perceptual quality scores. Furthermore, ELM model has been executed as single hidden layer of neurons with randomly feature mapping, which provides fast learning accuracy with much faster learning speed than traditional Neural Networks (NNs) and Singular Vector Regression (SVR) in many applications such as image segmentation, face recognition and no-reference IQA. SVR able to handle high-dimensional data but often achieve the suboptimal solutions with high computational complexity that argued by Huang *et al.,* [36]. The low computational complexity in ELM has been focuses from research community, specifically in high dimension and large database applications. Therefore, Extreme Learning Machine (ELM) is employed in this research study as pooling agent to balance the performance and improve computational complexity to evaluate the quality score of images.

3. Methodology

The proposed methodology of the work is shown in Figure 1. It is divided into three phases which consist of derivation of perceptual features extraction using LWT, followed by training for ELM pooling, and lastly, the statistical analysis of the algorithm and performance comparison is carried out between the approach with existing IQA.

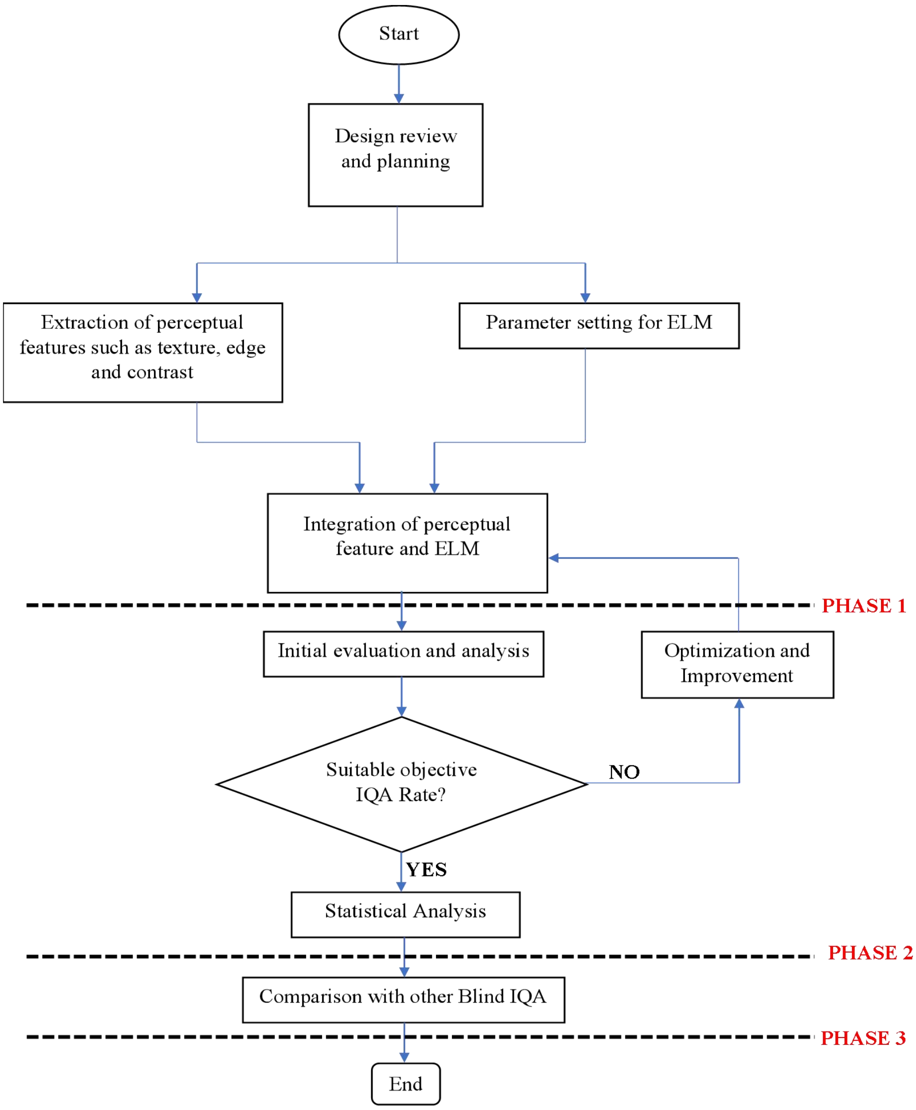

Fig. 1. Methodology of proposed metric

In the first phase, the main contribution is featuring extraction. The feature extraction is a key process where the two-dimensional (2D) image is converted to a set of mathematical parameters. In

this study, the lifting wavelet scheme is proposed for feature extraction. The wavelet was used to analyse various frequencies of image by using different scales. The LWT scheme can be divided into three main steps to decompose signal such as Split, Predict and Update signal as revealed in Figure 2. The split step formula can be written as:

$$s_e[n] = s[2n] \tag{1}$$

$$s_o[n] = s[2n+1] \tag{2}$$

After that, signal will undergo prediction process:

$$d[n] = s_o[n] - P(s_e)[n] \tag{3}$$

The signal will then be updated in last step:

$$c[n] = s_e[n] + U(d)[n] \tag{4}$$

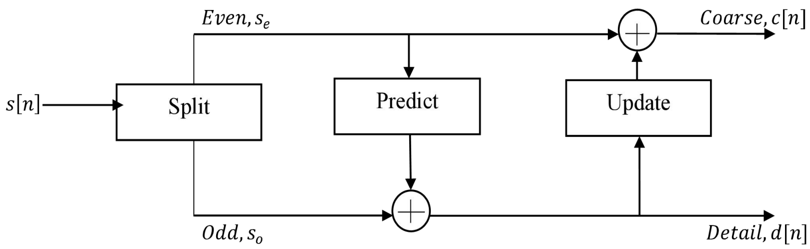

Fig. 2. Block diagram of lifting wavelet transform

Based on the Figure 3, the LWT is applied to the input image as feature extraction that resulted in four sub-bands: LL (low-low), LH (low-high), HH (high-high) and HL (high-low) with four-level wavelet decomposition. The LL1 (low-level image) will be further decomposed into second-level approximation. Then, the resultant from LWT is utilized to form the feature vector. The lifting scheme able to implement the wavelet transform in faster path as it makes use of similarities between the high pass filter and low pass filter.

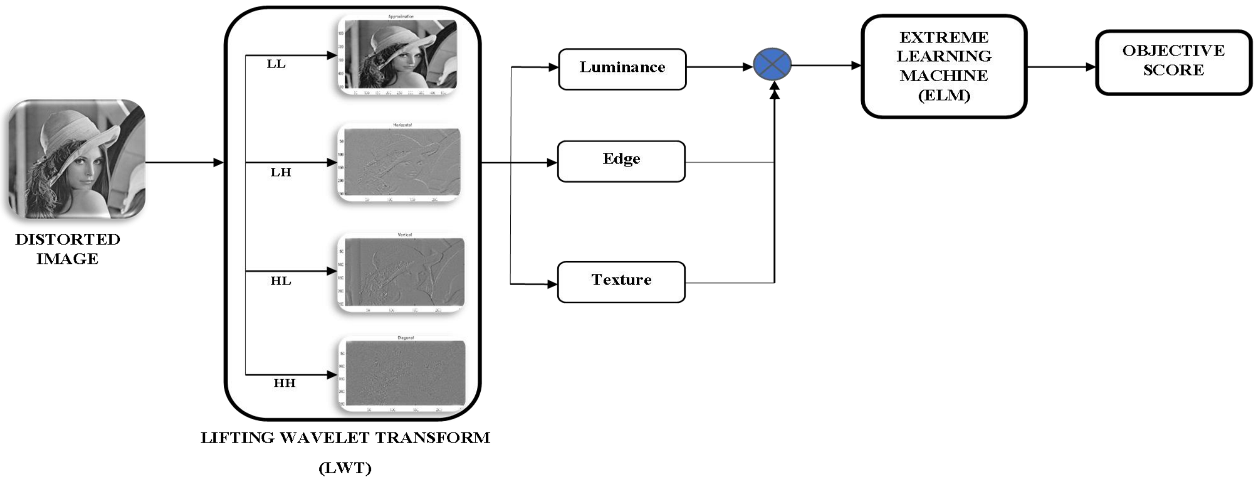

Fig. 3. Perceptual input proposed metric

In second phase, the approximation coefficients generated from feature extraction will be combined with ELM algorithm using the given subjective score to predict the overall image quality. The ELM is proposed as pooling operator to derive the classification output from evaluated image (Figure 4). ELMs are single hidden layer feed forward NNs which provide optimum and universal solution for highly complex tasks at extreme high speed. It has good scalability and runs at much higher learning speed (up to thousands of times) than NNs and Singular Vector Machine (SVM) algorithm [12]. The classification of ELM can be formulated as follow:

Given that *N* arbitrary training sample is (C_i, y_i), where the C_i is representing the feature vector of the *i*th pair of the original or distorted images and y_i is subjective quality score of its corresponding distorted image. The aim of the ELM is to determine the smallest derivation of the subjective quality score y_i of all training data [8]. The function can be represented as:

$$f(C_i) = \sum_{j=1}^{L} \beta_j g_j(C_i), i = 1, \ldots, N \tag{5}$$

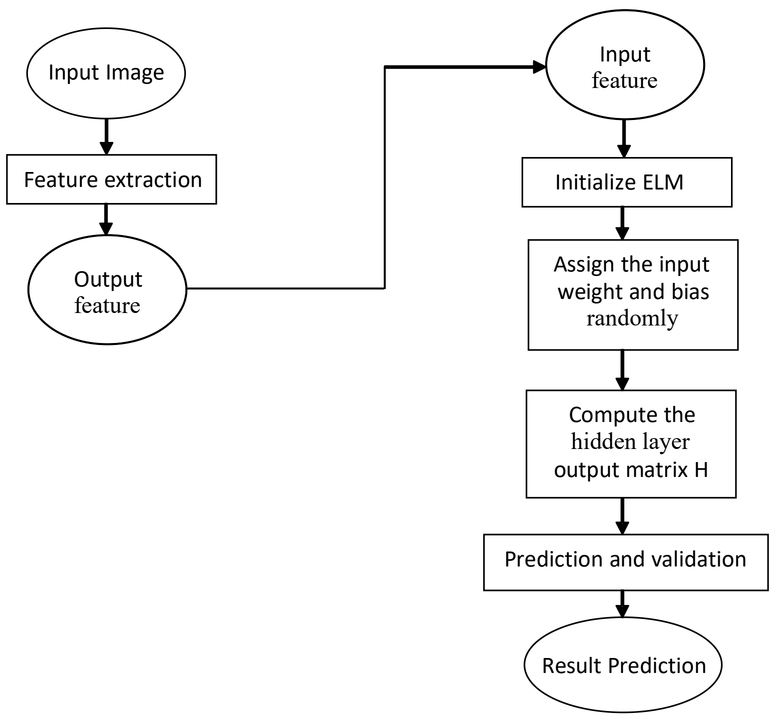

Fig. 4. Pooling process of Extreme Learning Machine (ELM)

where $\beta = [\beta_1, \ldots, \beta_l]^T$ is the output weighing vector, while $g_j(C_i)$ is the activation function which can approximate the *N* training samples with zero error, mean that $\sum_{i=1}^{N} \|f(C_i) - y_i\| = 0$. So, it can be formulating as:

$$g_j(C_i) = g\left(w_j . C_i + b_j\right) \tag{6}$$

where w_j is input weighing vector that connecting the *j*th hidden node, while the b_j (input node) is the threshold of the *j*th hidden node [7]. For $w_j . C_i$, it denotes the inner product between w_j and C_i. In theory [7], it has been proven that the w_j (input weighing vector) and bias term b_j can be randomly generated based on the continuous probability distribution. Besides, all hidden nodes are also randomly generated as well and independent each other. Hence, the β in Eq. (5) is only the parameter to be measured. This is one reason why ELM has fast learning speed than other feedforward NNs algorithms [14]. For *N* training sample in Eq. (5) can be written compactly as:

$$Y_H \beta = Y \tag{7}$$

where Y_H is called as hidden layer output matrix of the neural network as named in Huang *et al.*, [36], which can be shown as:

$$Y_H = \begin{pmatrix} g(w_1.C_1 + b_1) & \cdots & g(w_L.C_1 + b_L) \\ \vdots & \vdots & \vdots \\ g(w_1.C_N + b_1) & \cdots & g(w_L.C_N + b_L) \end{pmatrix}_{NxL} \tag{8}$$

$$\beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_L \end{pmatrix}_{Lx1} , Y = \begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix}_{Nx1} \tag{9}$$

where L is the number of hidden nodes in SLFNs [14]. The minimal norm least square method is employed in ELM algorithms instead using the standard optimization method to minimize the norm of the output weights. Therefore, β (output weights vector) can be estimated analytically as:

$$\beta = YY^{\dagger}{}_H \tag{10}$$

where $\mathrm{Y}^{\dagger}$ is the Moore-Penrose generalized inverse of matrix $\mathrm{Y_H}$; the *i*th column is the *i*th hidden node output with respect to the inputs $C_1, \ldots, C_N$. Boose *et al.,* [20] claimed that the best generalization performance of the feedforward NNs is obtained from the smallest norm weights. By adopting the Moore-Penrose generalized pseudo-inverse concept in Eq. (10), $\beta = YY^{\dagger}{}_H$ tend to have smallest norm among all the optimization solutions. This is reason why ELM algorithm has higher learning accuracy as compared to NNs and SVR.

When the ELM is called, the input weight and bias of the hidden neurons are generated randomly. The ELM is extended to Radial Basic Function (RBF) which allow the centres and the impact widths of the RBF Kernels to be randomly produced and making the output weights to be simply analytically calculated instead of iteratively tuned. The parameter of ELM such as number of hidden neurons, average value and training times is set at maximum to compute the hidden layer output matric H. The objective score produced then will be used to predict the image quality and validate in term of performance evaluations. The result prediction of image quality is then generated at the end of experiment.

4. Result and Discussion

In this section, the extensive experimental results and analysis are demonstrated to evaluate the overall accuracy of the proposed IQA metric. The most commonly used standard image processing grayscale image 'Lena' with the grayscale size of 512 x 512 as shown in Figure 5 is selected to decomposed in research. The simulation work has been carried out using MATLAB version 2021a on Window platform with computational device of 4 GB RAM, 2.0 GHz processor. The different type of wavelets was used to analyses different frequencies of an image using different scales. In this research, Discrete Wavelet Transform (DWT) and proposed Lifting Wavelet Transform (LWT) are used as feature extraction tool to extract the coefficient from image. The DWT and LWT analyses the image in Figure 5 by decorrelating the high pass in the image from the low pass.

Fig. 5. Input image (Lenna image)

Based on Figure 3, in our proposed model, the first level selected 2D wavelet transform is conducted. As analysed below, the image obtained are downsampled by the column indicated by 2 that means only even indexed columns are selected. The resultant image produced then decomposed again with the high pass and low pass filters. Now, the image is downsampled by row denoted by 1 which yields for subband that the half size of original image. Thus, the four subband image generated LL_1 (low-low), LH_1 (low-high), HL_1 (high-low) and HH_1 (high-high) that containing horizontal, vertical, and diagonal information of the image. The directional features extracted from LL_1 subbands wavelet transform gives the approximation version of the image while the details of the image are found in the rest of the subbands. LL_1 is then selected for the next round of the decomposition in the same manner as of that input image. From the next round, approximation coefficient LL_2 is extracted. Similarly, the image is decomposed by the two-dimensional wavelet decomposition up to the five levels. The performance analysis of DWT and LWT after feature extraction are shown in Figure 6 and Figure 7, respectively.

From experiment, the decomposition process only done up to five levels since at sixth level the image will loss most of its details as shown in Table 1. Based on result obtained, RMSE value at decomposition level 5 is 0.039135, whereas at level 6 of decomposition is 0.039615 of RMSE. Through the experimental result, it has been proved that the accuracy of the feature extraction was reduced starting from sixth level decomposition. The approximation coefficient obtained, that is LL_1, LL_2, LL_3, $LL_4$and LL_5 are then used to form the feature set as demonstrated in the following subsections.

Table 1
Different decomposition level in feature extraction

Decomposition Level	Root Mean Square Error, RMSE
Level 1	0.040053
Level 2	0.040220
Level 3	0.039662
Level 4	0.039609
<u>Level 5</u>	<u>0.039135</u>
Level 6	0.039615

The Table 2 shows the RMSE value of the grey-scale Lenna image (512x512) for the different type of wavelet filter. Based on theory, the increasing of the levels of wavelet will increase the complexity eventually the value of RMSE value reduces.

Table 2
Different wavelet filter used in feature extraction

Type of Wavelet filter	Root Mean Square Error, RMSE
Daubechies (db1/Haar)	0.040339
Daubechies (db2)	0.040018
Daubechies (db3)	0.039840
Daubechies (db4)	0.039428
Biorthogonal (bior3.1)	0.040125
Biorthogonal (bior3.3)	0.039667
Biorthogonal (bior3.5)	0.039598

The RMSE is used to measure the quality between the reconstructed image and input image which can defined as follows:

$$RMSE = \sqrt{MSE(input - output)} \tag{11}$$

The lower the value of RMSE, the better is the state of the reconstructed image. In this experiment, Daubechies wavelets and Biorthogonal wavelets filter are apply to the LWT during the feature extraction to compare the quality of an image. There are many Daubechies wavelet filters namely ‘'db1' or 'Haar', 'db2', ...,'db10', ..., etc but all these are nearly similar. Since the Daubechies wavelet are family of orthogonal wavelet characterized by the maximum number of the vanishing moments, thus it extracts better features as compared to the simple wavelet such as Haar wavelet and achieved similar results when compared to the complex wavelets like Gabor. Unlike complex wavelet techniques, it also takes less time to retrieve the results thus become suitable wavelet use in the experiment. Due to the used of overlapping window in these wavelets, the high frequency changes reflected by high frequency coefficient spectrum. Based on result in Table 2, the RMSE value decreases when the number of the decomposition levels of wavelet increases. It can determine that db4 (lowest RMSE value) become most preferred wavelet and db1 (highest PSNR value) become the least favoured wavelet. The Biorthogonal Wavelets is the extended family of wavelet with two scaling factors that responsible for the generation of different multi-resolutions on basis of the different wavelets. The quality of extracted image through different filter like bior3.1, bior3.3, bior3.5 are compared. The PSNR value for bior3.5 is 0.039598 is lowest in comparison to the bior3.1 and bior3.3. Overall, the Daubechies wavelets “db4” outperform other type of wavelet filters. So, it has been selected as wavelet filter parameter in this experiment for extracting coefficient. The extracted features then were used as input vectors for training and testing the performance of the Extreme Learning Machine (ELM) classifier.

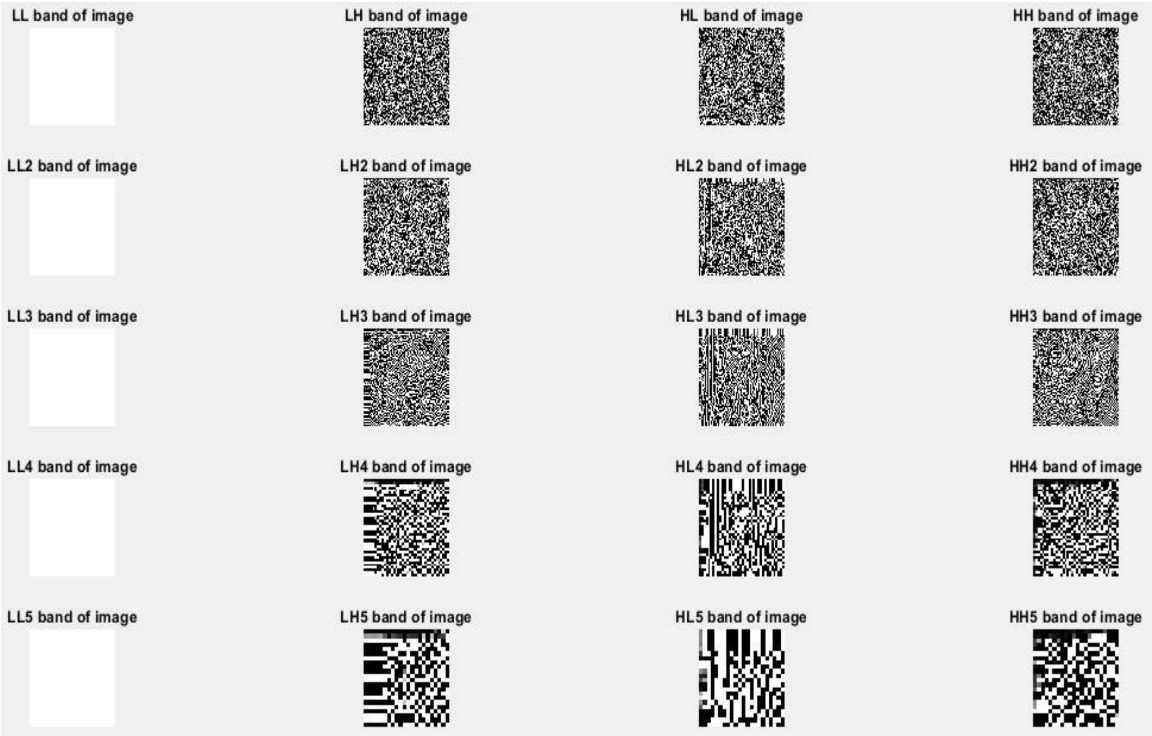

Fig. 6. Feature extraction using DWT (five-levels decomposition)

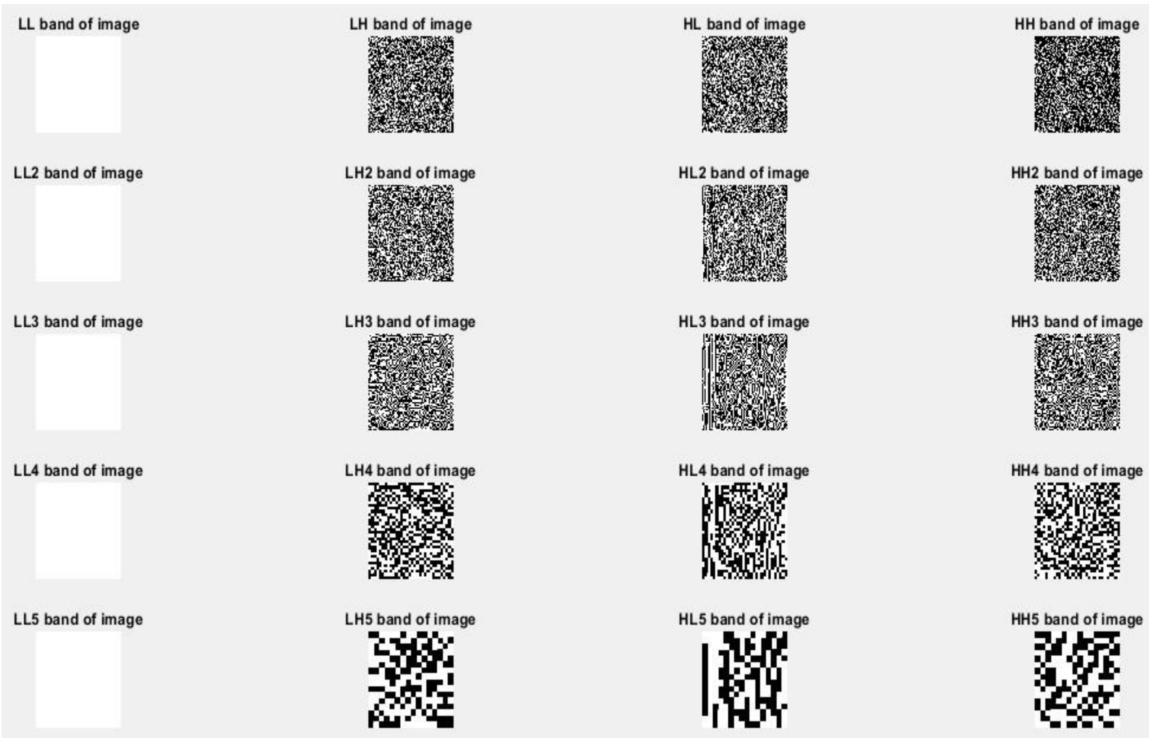

Fig. 7. Feature extraction using LWT (five-levels decomposition)

The approximation coefficients generated from feature extraction then combined with ELM algorithm to predict the overall image quality. From the simulation result obtained as shown in Figure 8 and Figure 9, the proposed approach LWT+ELM achieves exciting performance with 0.0039327of RMSE value, while the DWT+ELM achieve 0.040053 RMSE value. It shown that the LWT has lower value of RMSE indicate better performance in the extracted image. It demonstrated that the proposed approach is more effective than state-of-the-art work due to fast lifting-based wavelet and high speed of pooling ELM. The lower value of RMSE value proven that it has better segmentation performance.

Fig. 8. Regenerated result obtained (DWT+ELM)

Fig. 9. Regenerated result obtained (LWT+ELM)

From Table 3, it demonstrated that the RMSE value decreases with increasing of the number of hidden neurons in ELM. It means that the accuracy of the proposed model can be at optimum level with higher number of hidden neurons. When the hidden neurons are set at 20, the RMSE value achieve 0.066857 and decrease to 0.039897 for number of hidden neurons of 50. In this study, the number of hidden neurons of 50 is selected to conduct the experiment. In contrary, the ELM algorithm can achieve the best performance by adjusting the number of hidden layer neurons to improve the ability of the system regression and classification analysis.

Table 3
Different number of hidden neurons during pooling process

Number of Hidden Neurons	Root Mean Square Error, RMSE
20	0.066857
30	0.054497
40	0.046293
50	0.039897

5. Conclusion

Perceptual Image Quality Assessment has been an important issue in various applications. A simple but effective BIQA metric is proposed to estimate the image quality without the presence of a reference image. An ELM based IQA metric with lifting-based feature extraction is proposed to overcome the existing drawbacks in distortion description and distortion effects pooling of IQA. Due to low computation and less memory property in lifting scheme, it can be considered as the parts-based representations of image scenes, which are more stable with the human perception. Besides, the use of ELM exploits the advantages of machine learning to generate an effective mapping of distortion effects into overall quality scores. Therefore, ELM is recommended as pooling operator to derive the regression output from evaluated image due to its high leaning accuracy and faster learning speed. From the result obtained, the proposed metric is outperforming the relevant existing algorithms. It also proves that approach have good performance and robustness on the individual distortion evaluation results.

Acknowledgement
The authors would like to thank the anonymous reviewers for their insightful comments. This work is fully supported by KPT Fundamental Research Grant Scheme (FRGS/1/2020/ICT02/UNIMAP/02/5) and Universiti Malaysia Perlis (UniMAP).

References

[1] Lu, Pengluo, Yongchang Li, Longxu Jin, and Shuangli Han. "Blind image quality assessment based on wavelet power spectrum in perceptual domain." *transactions of tianjin university* 22 (2016): 596-602. https://doi.org/10.1007/s12209-016-2726-7
[2] Yang, Pei, Jordan Sturtz, and Letu Qingge. "Progress in blind image quality assessment: a brief review." *Mathematics* 11, no. 12 (2023): 2766. https://doi.org/10.3390/math11122766
[3] Wang, Lanjiang. "A survey on IQA." *arXiv preprint arXiv:2109.00347* (2021).
[4] Wang, Zhou. "Applications of objective image quality assessment methods [applications corner]." *IEEE signal processing magazine* 28, no. 6 (2011): 137-142. https://doi.org/10.1109/MSP.2011.942295
[5] Zhang, Weixia, Dingquan Li, Chao Ma, Guangtao Zhai, Xiaokang Yang, and Kede Ma. "Continual learning for blind image quality assessment." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 45, no. 3 (2022): 2864-2878.
[6] Ma, Rui, Hanxiao Luo, Qingbo Wu, King Ngi Ngan, Hongliang Li, Fanman Meng, and Linfeng Xu. "Remember and reuse: Cross-task blind image quality assessment via relevance-aware incremental learning." In *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 5248-5256. 2021. https://doi.org/10.1145/3474085.3475642
[7] Mittal, Anish, Rajiv Soundararajan, and Alan C. Bovik. "Making a “completely blind” image quality analyzer." *IEEE Signal processing letters* 20, no. 3 (2012): 209-212. https://doi.org/10.1109/LSP.2012.2227726
[8] Wang, Shuigen, Chenwei Deng, Weisi Lin, Guang-Bin Huang, and Baojun Zhao. "NMF-based image quality assessment using extreme learning machine." *IEEE transactions on cybernetics* 47, no. 1 (2016): 232-243. https://doi.org/10.1109/TCYB.2015.2512852
[9] Ahmed, ISMAIL T., Chen Soong Der, and Baraa Tareq Hammad. "A survey of recent approaches on no-reference image quality assessment with multiscale geometric analysis transforms." *International Journal of Scientific & Engineering Research* 7, no. 12 (2016): 1146-1156.
[10] Madhusudana, Pavan C., Neil Birkbeck, Yilin Wang, Balu Adsumilli, and Alan C. Bovik. "Image quality assessment using contrastive learning." *IEEE Transactions on Image Processing* 31 (2022): 4149-4161. https://doi.org/10.1109/TIP.2022.3181496
[11] Hu, Bo, Leida Li, Jinjian Wu, and Jiansheng Qian. "Subjective and objective quality assessment for image restoration: A critical survey." *Signal Processing: Image Communication* 85 (2020): 115839. https://doi.org/10.1016/j.image.2020.115839

[12] Wyllian Bezerra da Silva, Alexandre Mikowski, and Rafael Machado Casali, "No-reference Video Quality Assessment Method Based on Spatial-temporal Features using ELM Algorithm", IET Image Processing, 14(7), (2020): 1316-1326. https://doi.org/10.1049/iet-ipr.2019.0941
[13] Manap, Redzuan Abdul, and Ling Shao. "Non-distortion-specific no-reference image quality assessment: A survey." *Information Sciences* 301 (2015): 141-160. https://doi.org/10.1016/j.ins.2014.12.055
[14] Wang, Shuigen, Chenwei Deng, Weisi Lin, and Guang-Bin Huang. "A novel NMF-based image quality assessment metric using extreme learning machine." In *2013 IEEE China Summit and International Conference on Signal and Information Processing*, pp. 255-258. IEEE, 2013. https://doi.org/10.1109/ChinaSIP.2013.6625339
[15] Li, Dingquan, Tingting Jiang, and Ming Jiang. "Norm-in-norm loss with faster convergence and better performance for image quality assessment." In *Proceedings of the 28th ACM International conference on multimedia*, pp. 789-797. 2020. https://doi.org/10.1145/3394171.3413804
[16] Parikh, Nikunj, Santosh Chapaneri, and Gautam Shah. "No-reference image quality assessment using extreme learning machines." In *2016 International Conference on Inventive Computation Technologies (ICICT)*, vol. 2, pp. 1-5. IEEE, 2016. https://doi.org/10.1109/INVENTIVE.2016.7824891
[17] Wang, Shuigen, Chenwei Deng, Baojun Zhao, Guang-Bin Huang, and Baoxian Wang. "Gradient-based no-reference image blur assessment using extreme learning machine." *Neurocomputing* 174 (2016): 310-321. https://doi.org/10.1016/j.neucom.2014.12.117
[18] Moorthy, Anush Krishna, and Alan Conrad Bovik. "Blind image quality assessment: From natural scene statistics to perceptual quality." *IEEE transactions on Image Processing* 20, no. 12 (2011): 3350-3364. https://doi.org/10.1109/TIP.2011.2147325
[19] Mittal, Anish, Anush Krishna Moorthy, and Alan Conrad Bovik. "No-reference image quality assessment in the spatial domain." *IEEE Transactions on image processing* 21, no. 12 (2012): 4695-4708. https://doi.org/10.1109/TIP.2012.2214050
[20] Bosse, Sebastian, Dominique Maniry, Klaus-Robert Müller, Thomas Wiegand, and Wojciech Samek. "Deep neural networks for no-reference and full-reference image quality assessment." *IEEE Transactions on image processing* 27, no. 1 (2017): 206-219. https://doi.org/10.1109/TIP.2017.2760518
[21] Sheikhpour, R., Sheikhshoaie, M., and Kiani, K., "A novel no-reference image quality assessment model using a deep convolutional neural network." Signal, Image, and Video Processing, 14(5), (2020): 909-916.
[22] Strisciuglio, Nicola, Manuel Lopez-Antequera, and Nicolai Petkov. "Enhanced robustness of convolutional networks with a push–pull inhibition layer." *Neural Computing and Applications* 32, no. 24 (2020): 17957-17971. https://doi.org/10.1007/s00521-020-04751-8
[23] Madhusudana, Pavan C., Neil Birkbeck, Yilin Wang, Balu Adsumilli, and Alan C. Bovik. "Image quality assessment using contrastive learning." *IEEE Transactions on Image Processing* 31 (2022): 4149-4161. https://doi.org/10.1109/TIP.2022.3181496
[24] Deng, Chenwei, Shuigen Wang, Alan C. Bovik, Guang-Bin Huang, and Baojun Zhao. "Blind noisy image quality assessment using sub-band kurtosis." *IEEE transactions on cybernetics* 50, no. 3 (2019): 1146-1156. https://doi.org/10.1109/TCYB.2018.2889376
[25] Beghdadi, Azeddine, M-C. Larabi, Abdesselam Bouzerdoum, and Khan M. Iftekharuddin. "A survey of perceptual image processing methods." *Signal Processing: Image Communication* 28, no. 8 (2013): 811-831. https://doi.org/10.1016/j.image.2013.06.003
[26] Yang, Pei, Jordan Sturtz, and Letu Qingge. "Progress in blind image quality assessment: a brief review." *Mathematics* 11, no. 12 (2023): 2766. https://doi.org/10.3390/math11122766
[27] Athar, Shahrukh, and Zhou Wang. "A comprehensive performance evaluation of image quality assessment algorithms." *Ieee Access* 7 (2019): 140030-140070. https://doi.org/10.1109/ACCESS.2019.2943319
[28] Kim, Jongyoo, Anh-Duc Nguyen, and Sanghoon Lee. "Deep CNN-based blind image quality predictor." *IEEE transactions on neural networks and learning systems* 30, no. 1 (2018): 11-24. https://doi.org/10.1109/TNNLS.2018.2829819
[29] Wu, Jinjian, Jupo Ma, Fuhu Liang, Weisheng Dong, Guangming Shi, and Weisi Lin. "End-to-end blind image quality prediction with cascaded deep neural network." *IEEE Transactions on image processing* 29 (2020): 7414-7426. https://doi.org/10.1109/TIP.2020.3002478
[30] Liu, Jianzhao, Wei Zhou, Xin Li, Jiahua Xu, and Zhibo Chen. "LIQA: Lifelong blind image quality assessment." *IEEE Transactions on Multimedia* 25 (2022): 5358-5373. https://doi.org/10.1109/TMM.2022.3190700
[31] N. Ponomarenko. "Tampere Image Database 2008 TID2008." (2008). http://www.ponomarenko.info/tid2008.htm
[32] Larson, Eric C., and D. M. Chandler. "Categorical image quality (CSIQ) database." (2010).
[33] Sheikh, Hamid R. "Image and video quality assessment research at LIVE." *http://live. ece. utexas. edu/research/quality/* (2003).

[34] Manap, Redzuan Abdul, Ling Shao, and Alejandro F. Frangi. "PATCH-IQ: a patch based learning framework for blind image quality assessment." *Information Sciences* 420 (2017): 329-344. https://doi.org/10.1016/j.ins.2017.08.080
[35] You, Junyong, and Jari Korhonen. "Transformer for image quality assessment." In *2021 IEEE international conference on image processing (ICIP)*, pp. 1389-1393. IEEE, 2021. https://doi.org/10.1109/ICIP42928.2021.9506075
[36] Huang, Guang-Bin, Hongming Zhou, Xiaojian Ding, and Rui Zhang. "Extreme learning machine for regression and multiclass classification." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42, no. 2 (2011): 513-529. https://doi.org/10.1109/TSMCB.2011.2168604