Multi-Task Learning Approach for Natural Images' Quality Assessment

Redzuan Abdul Manap[1, 2], Alejandro F. Frangi[2], Ling Shao[3] and Abdul Majid Darsono[1]
[1]*Centre for Telecommunication Research and Innovation (CeTRI), Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia*
[2]*CISTIB, Department of Electronic and Electrical Engineering, University of Sheffield, Sheffield, UK*
[3]*School of Computing Sciences, University of East Anglia, Norwich, UK*
redzuan@utem.edu.my

***Abstract*—Blind image quality assessment (BIQA) is a method to predict the quality of a natural image without the presence of a reference image. Current BIQA models typically learn their prediction separately for different image distortions, ignoring the relationship between the learning tasks. As a result, a BIQA model may has great prediction performance for natural images affected by one particular type of distortion but is less effective when tested on others. In this paper, we propose to address this limitation by training our BIQA model simultaneously under different distortion conditions using multi-task learning (MTL) technique. Given a set of training images, our Multi-Task Learning based Image Quality assessment (MTL-IQ) model first extracts spatial domain BIQA features. The features are then used as an input to a trace-norm regularisation based MTL framework to learn prediction models for different distortion classes simultaneously. For a test image of a known distortion, MTL-IQ selects a specific trained model to predict the image's quality score. For a test image of an unknown distortion, MTL-IQ first estimates the amount of each distortion present in the image using a support vector classifier. The probability estimates are then used to weigh the image prediction scores from different trained models. The weighted scores are then pooled to obtain the final image quality score. Experimental results on standard image quality assessment (IQA) databases show that MTL-IQ is highly correlated with human perceptual measures of image quality. It also obtained higher prediction performance in both overall and individual distortion cases compared to current BIQA models.**

***Index Terms*—Blind Image Quality Assessment; Multi-Task Learning; Spatial Domain Image Features; Trace-Norm Regularization.**

I. INTRODUCTION

Image quality assessment (IQA) is a method to quantify the quality of a natural image by quality metrics. For applications where the end targets are human consumers, it is preferred to have IQA metrics that can quantify the natural image quality as perceived by human observers. Therefore, human perception based metrics are often considered as the gold standard in IQA. They are typically obtained by conducting experiments where human observe and rate the quality of a natural image presented to them. The ratings are then averaged across all participants to yield the mean opinion score (MOS), or differential mean opinion score (DMOS). The score represents the perceived quality metric for the image. However, this approach is expensive, time-consuming and unfeasible for real-time applications. An objective IQA model that can automatically provide quality measurement consistent with MOS/DMOS values is more favourable.

Objective IQA can be classified into two main categories [1]: full-reference IQA (FR-IQA) and blind IQA (BIQA). FR-IQA models predict the quality of a distorted natural image by comparing the entire information difference between the image and its reference image. A reference image refers to a similar image which is distortion-free and of perfect quality. The simplest FR-IQA metrics are mean squared error (MSE) and peak signal-to-noise ratio (PSNR). However, they have been shown to have poor correlation with human perceptual measures [2]-[3]. Many other FR-IQA models were then proposed to improve the correlation performance. They are developed based on various mechanisms such as human visual system (HVS) [4]-[5], image structure [6]-[7], or image statistics [8]-[9]. These FR-IQA models achieve a high correlation with human perceptual measures. However, full reference image information is not available in some applications thus a BIQA model is preferred.

BIQA models can be categorised into two classes [10]: distortion-specific (DS) models or general-purpose models. DS BIQA methods work by utilising a specific distortion model under the assumption that the distortion affecting the image is known beforehand. For examples, the quality of JPEG compressed images is estimated by the model in [11] while the quality of a natural image affected by motion blur is assessed blindly in [12]. The effects of blocking and noise artefacts are also investigated in [13] and [14], respectively. In contrast, no prior knowledge of the distortion affecting the image is required in general-purpose BIQA models. Instead, image quality is determined by assuming that the image is degraded by the same distortion mechanism that affects a database of image exemplars. Such image exemplars can be obtained from standard IQA databases such as the LIVE [15] and CSIQ [5]. Using such exemplars and their provided MOS / DMOS values, the models are then trained to predict the MOS / DMOS of the image.

The majority of current general purpose BIQA models focus on extracting relevant features that carry discriminative information about image quality. Most of the models employ handcrafted features that are designed based on various mechanism such as natural scene statistics (NSS) [16]-[20], natural colour statistics (NCS) [21] or free energy principle [22]-[23]. Some other models use features that are learned directly from raw image pixels [24]-[25]. The extracted features are then used as an input to regression algorithms to learn the mapping between the features' space and the image quality score space. Support vector regression (SVR) with linear/radial basis function is frequently used to this effect.

High prediction performances in correlation with human

perceptual scores are reported by these models [16]-[19], [24]. However, we noted that a BIQA model may perform well for images affected by a particular type of distortion but is less effective when tested on different distortion types. We believe one possible reason to this is current BIQA models are trained independently for each type of distortion, ignoring the relationship that may exist among the distortion classes. This scenario motivates us to look at an alternative way to learn prediction model for BIQA task. In this paper, we explore the use of multi-task learning (MTL) to learn prediction models for different image distortion classes simultaneously. MTL has been utilised in learning prediction models for web pages categorisation [26], disease prediction [27] and therapy screening [28]. Here, we extend its application to BIQA tasks.

MTL is a learning approach that utilises a shared representation to learn multiple related tasks simultaneously. Based on the assumption that the learner may find it easier to learn multiple tasks together rather than in isolation when the tasks share what they learn, MTL has been shown to increase the learning capability of each individual task [29]. Treating quality prediction for one type of distortion as one learning task, our BIQA model is developed to utilise this advantage. Given a set of training images, the model, *Multi-Task Learning based Image Quality assessment* (MTL-IQ), first extracts relevant spatial domain BIQA features. These features are then used to train regression models for different distortion conditions simultaneously. The training is performed using a trace-norm regularised MTL technique. For a test image of a known distortion, MTL-IQ simply selects a specific regression model to predict the image's quality score. For a test image of an unknown distortion, MTL-IQ estimates the amount of each distortion present in the image using an SVM classifier. The probability estimates from the classifier are then used to weigh the image prediction scores from different regression models. The weighted scores are then pooled to yield the final quality score.

Our work is motivated by promising results achieved by another multi-task based BIQA model, MRLIQ [30]. However, there are three substantial differences. First, MTL-IQ utilises spatial domain features as opposed to transform-based features in MRLIQ. Second, MTL-IQ employs different regularisation or penalty term to perform its MTL training. Third, in agreement to other established BIQA models, MTL-IQ is optimised using the training images' quality ratings (DMOS/MOS), whereas MRLIQ via pairwise quality rank. In addition, MRLIQ was tested only on the LIVE IQA database with limited results and analysis. Additional experiments and analysis are included in this paper to demonstrate our model capability further. The remainder of this paper is as follows. The framework for MTL-IQ is discussed in Section II. This includes a description of the utilised BIQA features as well as the chosen MTL technique to train MTL-IQ. The conducted experiments and their corresponding results are presented in Section III to evaluate MTL-IQ's performance. The paper is concluded in Section IV.

II. Multi-Task Learning Based IQA

The proposed framework for MTL-IQ is illustrated in Figure 1. It consists of feature extraction (FE), quality estimation (QE) and distortion identification (DI) stages.

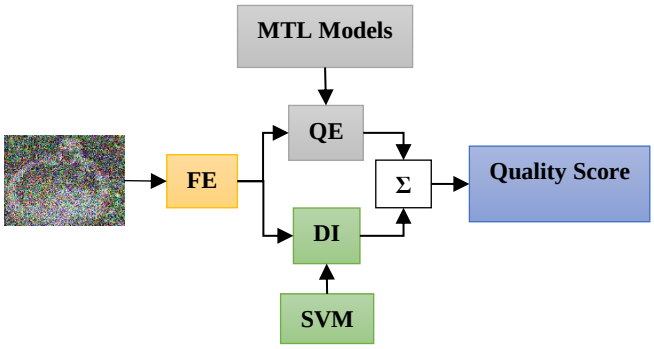


Figure 1: MTL-IQ framework

A. Feature Extraction

While most of previous BIQA models focused on discovering features that are suitably linked to perceptually relevant scores, our work focus on developing new learning strategy for a BIQA task. Therefore, rather than designing new features, MTL-IQ employs the same spatial domain features implemented by GMLOG model [19]. The features are chosen to alleviate excessive computational load often encountered by image transform based features [16]-[17]. The features consist of four statistical distributions that are derived from the image local contrast operators: gradient magnitude (GM) and Laplacian of Gaussians (LOG). GMLOG shows that the shape of these distributions will deviate from those of high-quality images when an image is distorted. As the image's distortion level increases, there are gradual changes in the distributions' shapes indicating they are predictive to image quality and can be BIQA features.

Specifically, given an image $\mathbf{I}$, its GM map $\mathbf{G_I}$ and LOG response $\mathbf{L_I}$ are defined respectively as:

$$\mathbf{G_I} = \sqrt{[\mathbf{I} \otimes \mathbf{h}_x]^2 + [\mathbf{I} \otimes \mathbf{h}_y]^2} \quad (1)$$

and

$$\mathbf{L_I} = \mathbf{I} \otimes \mathbf{h}_{\mathrm{LOG}} \, . \quad (2)$$

In Equation (1), $\mathbf{h}_x$ and $\mathbf{h}_y$ are the Gaussian partial derivative filters applied along the horizontal and the vertical direction respectively. The LOG filter in Equation (2) is represented as:

$$\mathbf{h}_{\mathrm{LOG}}(x, y|\sigma) = \frac{\partial^2}{\partial x^2}\mathbf{g}(x, y|\sigma) + \frac{\partial^2}{\partial y^2}\mathbf{g}(x, y|\sigma) \ , \quad (3)$$

where $\mathbf{g}(x, y|\sigma)$ is the isotropic Gaussian function with scale parameter σ. These GM and LOG operators are then jointly normalized to achieve stable image representations. The normalized operators are given by:

$$\bar{\mathbf{G}}_\mathbf{I} = \frac{\mathbf{G_I}}{(\mathbf{N_I}+\varepsilon)} \ , \ \bar{\mathbf{L}}_\mathbf{I} = \frac{\mathbf{L_I}}{(\mathbf{N_I}+\varepsilon)} \ , \quad (4)$$

where $\mathbf{N_I}$ is a local adaptive normalization factor while ε is a constant that prevents numerical instability.

Once both operators are normalised, MTL-IQ computes their respective marginal probability functions and use them as the first two BIQA features for the image. The marginal probability functions are defined as:

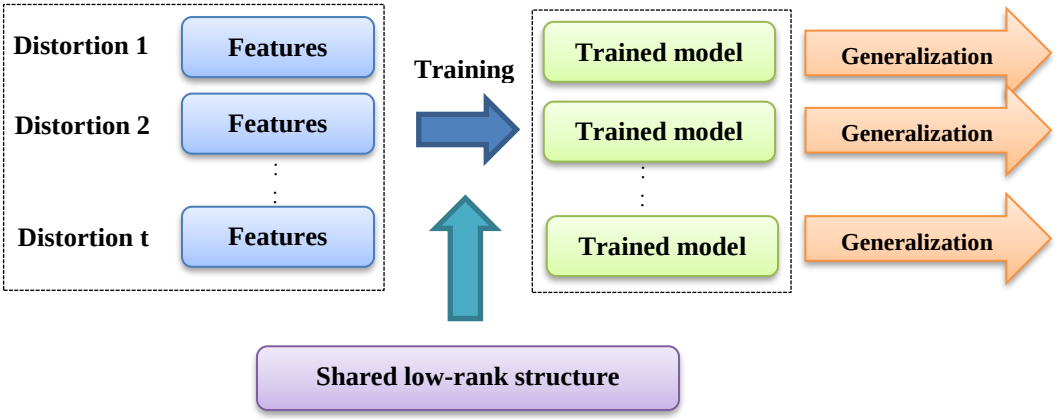


Figure 2. Trace-norm regularized MTL training structure

$$P_{\bar{\mathbf{G}}_\mathbf{I}}(\bar{\mathbf{G}}_\mathbf{I} = g_m) = \sum_{n=1}^{N} \mathbf{K}_{m,n} \quad (5)$$

and $$P_{\bar{\mathbf{L}}_\mathbf{I}}(\bar{\mathbf{L}}_\mathbf{I} = l_n) = \sum_{m=1}^{M} \mathbf{K}_{m,n} \,. \quad (6)$$

In these equations, $\mathbf{K}_{m,n} = P(\bar{\mathbf{G}}_\mathbf{I} = g_m\,, \bar{\mathbf{L}}_\mathbf{I} = l_n)$ is the joint empirical probability function of the normalized GM and LOG operators while $m = 1,2,\dots,M$ and $n = 1,2,\dots,N$ represent the quantization levels of those operators.

The remaining two BIQA features are derived based on the statistical interaction between GM and LOG operators. The features, known as independency distributions are measured by computing the dependency of each specific value $\bar{\mathbf{G}}_\mathbf{I} = g_m$ against all possible values of $\bar{\mathbf{L}}_\mathbf{I}$ and vice versa. The computations can be represented as:

$$Q_{\bar{\mathbf{G}}_\mathbf{I}}(\bar{\mathbf{G}}_\mathbf{I} = g_m) = \frac{1}{N}\sum_{n=1}^{N} P(\bar{\mathbf{G}}_\mathbf{I} = g_m | \bar{\mathbf{L}}_\mathbf{I} = l_n) \quad (7)$$

and $$Q_{\bar{\mathbf{L}}_\mathbf{I}}(\bar{\mathbf{L}}_\mathbf{I} = l_n) = \frac{1}{M}\sum_{m=1}^{M} P(\bar{\mathbf{L}}_\mathbf{I} = l_n | \bar{\mathbf{G}}_\mathbf{I} = g_m) \quad (8)$$

respectively. Equations (7) and (8) can be viewed as the sum of conditional probabilities of a specific value of $\bar{\mathbf{G}}_\mathbf{I}$ (or $\bar{\mathbf{L}}_\mathbf{I}$) over variable $\bar{\mathbf{L}}_\mathbf{I}$ (or $\bar{\mathbf{G}}_\mathbf{I}$).. The four distributions are then concatenated to produce the MTL-IQ features vector for the given image. Further details can be found in [19].

B. Quality Estimation

The extracted feature vector is then used as an input to the trained quality prediction models to estimate the quality of the given image for different distortion conditions. Previous BIQA methods train their prediction models via single-task learning (STL) approach whereby prediction model for one particular distortion is treated as single learning task and learnt independently. In contrast, MTL-IQ learns its prediction models simultaneously via an MTL approach.

Given feature vectors extracted from a set of training images, MTL-IQ aims to minimise this objective function:

$$\min_{\mathbf{W}} F(\mathbf{W}) = f(\mathbf{W}) + \Omega(\mathbf{W}) \;, \quad (9)$$

where $f(\mathbf{W})$ is the empirical loss on the training set and $\Omega(\mathbf{W})$ is the regularization term that captures the relationship among the tasks. For a BIQA case, $f(\mathbf{W})$ is represented by a loss function $\ell(\cdot,\cdot)$ as:

$$f(\mathbf{W}) = \sum_{i=1}^{n}\sum_{j=1}^{s_i} \ell\left(y_i^j, \omega_i^{\mathrm{T}} x_i^j\right), \quad (10)$$

with n is the number of distortion classes, s_i is the number of samples in the ith distortion, x_i^j and y_i^j are the jth feature vector and the associated DMOS value in the ith distortion, respectively and $\mathbf{W} = [\omega_1, \omega_2, \dots, \omega_n]$ where ω is the parameter to be estimated from the training samples.

There are many formulations for MTL. Based on the assumption that the distortion classes are related and the fact that the extracted features are in high dimension, a trace-norm regularised technique [29] is applied to train MTL-IQ. The technique captures the task relatedness through low dimensional sub-space learning whereby the models from different tasks are constrained to share a common low-rank structure. Figure 2 illustrates the trace-norm regularised training structure for MTL-IQ.

The technique treats Equation (9) as a matrix rank minimisation problem [29]:

$$\min_{\mathbf{W}} F(\mathbf{W}) = f(\mathbf{W}) + \lambda[\mathrm{Rank}(\mathbf{W})] \quad (11)$$

Since the matrix rank minimisation problem is an NP-hard problem, a convex relaxation of the rank function $\mathrm{Rank}(\mathbf{W})$ is normally required. Trace-norm relaxation method is widely used to this effect as it has been shown theoretically to be a good approximation for $\mathrm{Rank}(\mathbf{W})$ [31]. Therefore, the problem can now be approximated as a trace-norm minimization problem [29]:

$$\min_{\mathbf{W}} F(\mathbf{W}) = f(\mathbf{W}) + \lambda\|\mathbf{W}\|_* \quad (12)$$

where λ is positive regularization parameter and $\|\cdot\|_*$ denotes the trace norm defined as the sum of singular values. For faster convergence, MTL-IQ employs an accelerated gradient method [32] to solve Equation (12) to find the optimized values of $\mathbf{W}$:

$$\mathbf{W} = \arg\min_{\mathbf{W}} \frac{\gamma}{2}\left\|\mathbf{W} - \left(\mathbf{S} - \frac{1}{\gamma}\nabla f(\mathbf{W})\right)\right\|_{\mathrm{F}}^2 + \lambda\|\mathbf{W}\|_* \,, \quad (13)$$

where γ is the step size, $\nabla f(\cdot)$ is the gradient of $f(\cdot)$ and $\mathbf{S}$ is the search point. The optimized values are then used to represent the trained model for each distortion case. Further details on AGM can be found in [32].

C. Distortion Identification

The trained models are then used to predict the quality score of a test image. For a test image of unknown distortion, MTL-IQ first estimates different distortion types present in the image. The process is performed using the extracted feature vector as an input to an SVM classifier. SVM is chosen here due to its good performance in high dimensional spaces and good generalisation capabilities [33]. In this work, a multi-class SVM with a radial basis function (RBF) is employed. Note that our aim is not to perform hard classification but to estimate each distortion class present in the image. These estimates are given by the probabilities provided by the classifier. These probability values are then used to weigh the image prediction scores from different MTL models. The weighted scores are then pooled to yield the final quality score for the image.

III. Results and Discussions

A. Experimental Setup and Evaluation Protocol

Databases: MTL-IQ was tested on two well-known IQA databases: LIVE [15] and CSIQ [5]. The LIVE database contains 779 distorted images that are generated from 29 references. Each reference image is distorted at 5 or 6 degradation levels by five types of source coding or artificial artefacts: JPEG2000 compression (JP2K), JPEG compression (JPEG), additive white noise (WN), Gaussian blur (GB), and simulated fast fading (FF). The images are provided with DMOS values in the range between 0 and 100. An image of higher quality is assigned with lower DMOS. The CSIQ database is composed of 30 reference images. The reference images are subjected to six types of distortions at 4 or 5 degradation levels, yielding a total of 866 distorted images. Each image is assigned a DMOS value in the range between 0 and 1. Similar to the LIVE database, a higher DMOS value indicates a lower quality image. For the CSIQ database, only four types of distortions that contained in the LIVE database are considered: JP2K, JPEG, WN and GB.

Performance metrics: Three metrics that measure the consistency between the predicted quality scores and the subjective DMOS values were used to evaluate MTL-IQ's performance: the linear correlation coefficient (LCC), the Spearman rank order correlation coefficient (SROCC) and the root mean squared error (RMSE). The LCC and RMSE metrics are used to measure a model's prediction accuracy while the SROCC metric is used to measure the prediction monotonicity of a model. Values close to 1 for LCC and SROCC or 0 for RMSE indicates that the model has a high correlation with human subjective scores.

Benchmarked models: MTL-IQ was compared against four state-of-the-art BIQA models: BIQI [16], BRISQUE [18], GMLOG [19] and CORNIA [24], whose source codes are publicly available. MTL-IQ was also compared with two well-known FR-IQA models: SSIM [6] and FSIM [7].

Parameter and training setup: We set MTL-IQ's parameters as implemented by GMLOG. The filters' scale parameter σ to compute GM and LOG operators was set at 0.5 while the quantization level $M = N$ is 10. To train MTL-IQ and the other BIQA models, we divided the databases into two subsets: 80% of the reference images and their corresponding distorted versions were randomly selected to be a training set while the remaining 20% were used for testing. There is no overlap between the two sets. The trace-norm regularized technique to train the MTL-IQ models was implemented using the MALSAR package [34]. In the package, the loss function $\ell(\cdot,\cdot)$ is set as a least squares function. The SVM classifier for the DI stage is trained using the LIBSVM package [35]. We also used the LIBSVM package to train regression models for the competing BIQA: SVR with a RBF kernel for BIQI, BRISQUE and GMLOG as well as SVR with a linear kernel for CORNIA. For a fair comparison, their SVR parameters were determined through cross validation in accordance to their respective papers.

Table 1
Median Values Across 1000 Runs of The Overall Performance Experiment

IQA model	LIVE			CSIQ		
	LCC	SROCC	RMSE	LCC	SROCC	RMSE
SSIM	0.9464	0.9486	8.8035	0.9347	0.9362	0.0990
FSIM	**0.9612**	**0.9639**	**7.5461**	**0.9675**	**0.9629**	**0.0710**
BIQI	0.8486	0.8427	15.4068	0.8089	0.7491	0.1867
BRISQUE	0.9431	0.9421	9.3953	0.9304	0.9101	0.1073
GMLOG	0.9505	0.9503	8.8290	0.9394	0.9219	0.0997
CORNIA	0.9394	0.9416	9.9204	0.9110	0.8873	0.1254
MTL-IQ	**0.9600**	**0.9567**	**8.8060**	**0.9483**	**0.9263**	**0.0920**

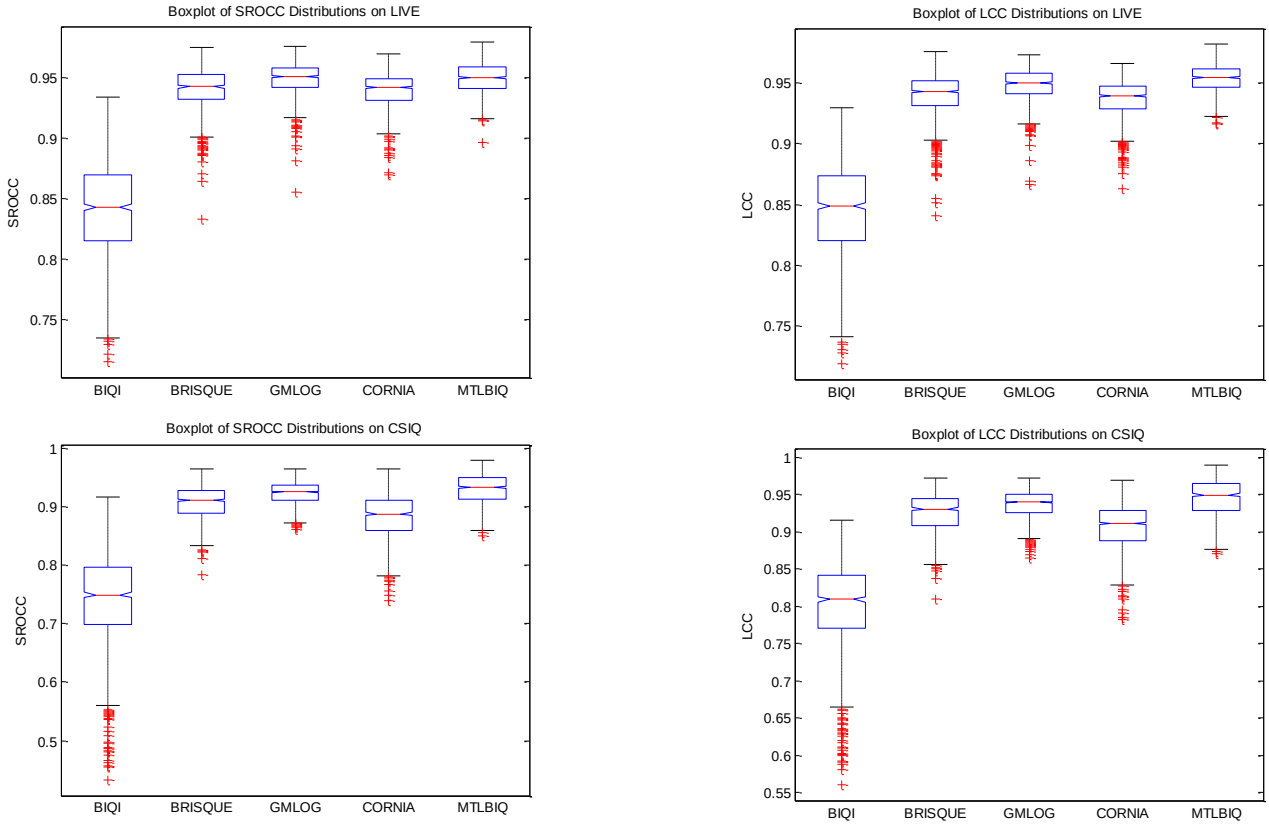
Table 2
IQR for 1000 SROCC and LCC Values for Overall Performance Experiment

BIQA model	LIVE		CSIQ	
	LCC	SROCC	LCC	SROCC
BIQI	0.0532	0.0537	0.0710	0.0960
BRISQUE	0.0197	0.0204	0.0361	0.0390
GMLOG	0.0167	0.0164	0.0237	0.0260
CORNIA	0.0184	0.0183	0.0405	0.0515
MTL-IQ	**0.0120**	**0.0154**	**0.0204**	**0.0230**

Experiments: We performed two experiments to evaluate MTL-IQ's performance: the overall performance experiment and the distortion-specific (DS) performance experiment. In the overall performance experiment, the train-test run was performed across all images regardless of their distortion type. This is to evaluate how well a BIQA model performs across all distortion classes. In the DS performance experiment, the train-test run was only conducted using images from a single distortion class. This is to evaluate how well a BIQA model performs for one specific distortion. Note that MTL-IQ contains different trained models for different distortion classes. For the DS performance experiment in which the distortion type is known, the specific trained model can directly be used for the QE stage without having to perform the DI stage. The train-test run was repeated 1000 times to ensure that the specific train-test partition does not govern the results. Due to the right-skewed distribution of the LCC and the SROCC values, the median is often used in the previous IQA works as their centre measurements. We followed the same approach to report our results.

B. Overall Performance Results

The median results for the overall experiment are tabulated in Table 1. The top FR-IQA, and BIQA models are in bold. MTL-IQ obtained the top LCC, SROCC and RMSE values among the competing BIQA models in both databases. In comparison to FR-IQA models, MTL-IQ outperformed SSIM while approaching the state-of-the-art FSIM. Given that FR-IQA models require reference images as their input, the MTL-IQ's results are promising. We then computed the inter-quartile range (IQR) value of the 1000 SROCC and LCC results obtained by each BIQA model. A model with low IQR value indicates that its' results are more consistent under



(a) (b)

Figure 3: Boxplots of (a) SROCC and (b) LCC distributions on the LIVE (top) and the CSIQ (bottom) databases

Table 3
Median SROCC Values Across 1000 Runs for the DS Performance Experiment

IQA model	LIVE					CSIQ			
	JP2K	JPEG	WN	GB	FF	JP2K	JPEG	WN	GB
SSIM	0.9614	0.9764	0.9694	0.9517	**0.9556**	0.9606	0.9546	0.8974	0.9609
FSIM	**0.9724**	**0.9840**	**0.9716**	**0.9708**	0.9519	**0.9704**	**0.9664**	**0.9359**	**0.9729**
BIQI	0.8303	0.9062	0.9328	0.8656	0.6885	0.7635	0.9102	0.5397	0.7826
BRISQUE	0.9164	0.9640	0.9791	0.9446	0.8872	0.8977	0.9212	0.9207	**0.9186**
GMLOG	0.9268	0.9630	0.9831	0.9188	0.9012	0.9161	0.9264	0.9408	0.9083
CORNIA	0.9205	0.9359	0.9608	**0.9519**	**0.9052**	0.8942	0.8820	0.7862	0.9041
MTL-IQ	**0.9356**	**0.9690**	**0.9842**	0.9288	0.9043	**0.9279**	**0.9286**	**0.9453**	0.9181

different train-test partitions. The IQR values are tabulated in Table 2 where the top model is in bold. We can see that MTLBIQ produced more consistent results than the rest.

To visualise the IQR for each model, we also generated the box-plots of the SROCC and LCC distributions in Figure 3. The central mark on each box is the median while the top edge and the bottom edge are the 25^{th} and the 75^{th} percentiles, respectively. In terms of outliers, ideally, we would like as few outliers as possible and to have them as close to the main distribution as possible. In this case, we can see that MTL-IQ has more compact outlier distributions on both databases than other competing BIQA models. Both IQR and outlier observation suggests that MTL-IQ has the best quality prediction consistency.

C. Distortion Specific Performance Results

The median results across the 1000 trials for the DS performance experiment are tabulated in Table 3. For simplicity, only the SROCC results are reported. Similar patterns can be observed from the LCC and the RMSE results. For both databases, MTL-IQ obtained the highest SROCC values for images affected by JP2K, JPEG and WN artefacts. It also gave competitive prediction performance for GB and FF images. In comparison to FR-IQA models, MTL-IQ produced a close performance to both SSIM and FSIM whereby it yielded better prediction performance for WN images.

Since MTL-IQ employs similar features as GMLOG, a direct comparison between the two models can be used to investigate whether MTL can improve GMLOG's learning capability in individual distortion classes. From Table 3, we can see that MTL-IQ achieved higher correlation values than GMLOG in all tested distortion cases. This observation indicates that MTL technique can be employed to achieve better prediction performance for BIQA task.

Table 4
Average Run-time Comparison

Model	BIQI	BRISQUE	GMLOG	CORNIA	MTL-IQ
Run times	0.05	0.10	0.07	2.43	0.08

D. Computational Complexity

The average run-time comparison between MTL-IQ models and the competing BIQA models for a typical image of 512 × 768 size is shown in Table 4. These processing times are achieved using un-optimised MATLAB R2011b code on an 8GB RAM computer with an Intel i5 3.20 GHz processor. Note that the training time is not considered here as it is assumed that the models are already trained prior to the testing stage. BIQI is the fastest model, but it has the worst prediction performance among all tested models. MTL-IQ is faster than BRISQUE and CORNIA. Although MTL-IQ uses the same feature as GMLOG, it is slightly slower due to the distortion identification requirement. Despite that, it still can process up to 12 images per second thus providing an alternative solution to real-time applications.

IV. Conclusion

This paper presents a simple yet effective BIQA model that employs a trace-norm regularised MTL technique in its learning framework. The model, dubbed as MTL-IQ, utilises a shared representation among different distortion training samples to learn prediction models for each distortion classes simultaneously. Experimental results on the LIVE and the CSIQ databases showed that MTL-IQ yields high correlation with human perceived quality measures across various types of image distortions. MTL-IQ also achieved higher prediction performance compared to some of the current BIQA models. It is worth noting that there are several steps could be taken to improve the model. For future works, different features and databases could be tested to validate MTL-IQ's performance further. Other MTL techniques could be tested for faster computation. MTL technique itself could also be employed for distortion identification in a case where the type of distortion affecting the image is unknown.

Acknowledgement

This work is partly supported by UTeM through Centre for Research & Innovation Management (CRIM) and by Machine Learning & Signal Processing (MLSP) research group under CeTRI.

References

[1] R. A. Manap, L. Shao, and A. F. Frangi, "Non-parametric quality assessment of natural images," *IEEE Multimedia*, vol. 23, no. 4, pp. 22-30, Oct. 2016.

[2] B. Girod, "What's wrong with mean-squared error?" in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993, pp. 207-220.

[3] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98-117, Jan. 2009.

[4] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Processing*, vol. 16, no. 9, pp. 2284-2298, Sept. 2007.

[5] E. C. Larson and D. M. Chandler, "Most apparent distortion: Full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011006.1-011006.21, Jan. 2010.

[6] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[7] L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Processing*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011.

[8] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Processing*, vol. 15, no. 2, pp. 430-444, Feb. 2006.

[9] C. Charrier, O. l'ezoray, and G. Lebrun, "Machine learning to design full-reference image quality assessment algorithm," *Signal Processing: Image Communication*, vol. 27, no. 3, pp. 209-219, Mar. 2012.

[10] R. A. Manap, L. Shao and A. F. Frangi, "A non-parametric framework for no-reference image quality assessment," in *Proc. of IEEE Conference on Signal and Information Processing*, Orlando, Florida, Dec. 2015, pp. 562-566.

[11] S. A. Golestaneh and D. M. Chandler, "No-reference quality assessment of JPEG images via a quality relevance map," *IEEE Signal Processing Letters*, vol. 21, no. 2, pp. 155-158, Feb. 2014.

[12] J. Zhang and T. M. Le, "A new no-reference quality metric for JPEG2000 images," *IEEE Trans. Consumer Electronics*, vol. 56, no. 2, pp. 743-750, May 2010.

[13] G. Yammine, E. Wige, and A. Kaup, "A no-reference blocking artefacts visibility estimator in images," in *Proc. of IEEE Conference on Image Processing*, Hong Kong, Sept. 2010, pp. 2497-2500.

[14] E. Cohen and Y. Yitzhaky, "No-reference assessment of blur and noise impacts on image quality," *Signal, Image and Video Processing*, vol. 4, no. 3, pp. 289-302, Sep. 2010.

[15] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, *LIVE Image Quality Assessment Database Release 2* [Online]. Available: http://live.ece.utexas.edu/research/quality

[16] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Process. Lett*ers, vol. 17, no. 5, pp. 513-516, May 2010.

[17] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: a natural scene statistics approach in the DCT domain," *IEEE Trans. Image Processing*, vol. 21, no. 8, pp. 3339-3352, Aug. 2012.

[18] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," *IEEE Trans. Image Processing*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012.

[19] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and Laplacian features," *IEEE Trans. Image Processing*, vol. 23, no. 11, pp. 4850-4862, Nov. 2014.

[20] R. A. Manap, L. Shao and A. F. Frangi, "PATCH-IQ: a patch based learning framework for blind image quality assessment," *Information Sciences*, vol. 420, no. 11, pp. 329-344, Aug. 2017.

[21] Q. Wang, J. Chu, L. Xu and Q. Chen, "A new blind image quality framework based on natural color statistic," *Neurocomputing*, vol. 172, no. 3, pp. 1798-1810, 2016.

[22] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," *IEEE Trans. Image Processing*, vol. 21, no. 1, pp. 41-52, Jan. 2012.

[23] K. Gu, G. Zhai, X. Yang, and W. Zhang, "Using free energy principle for blind image quality assessment," *IEEE Trans. Multimedia*, vol. 17, no. 1, pp. 50-63, Jan. 2015.

[24] P. Ye, J. Kumar, L. Kang, and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Providence, Rhode Island, June 2012, pp. 1098-1105.

[25] L Kang, P. Ye, Y. Li, and D. Doermann, "Convolutional neural networks for no-reference image quality assessment," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, Ohio, June 2014, pp. 1733-1740.

[26] J. Chen, L. Tang, J. Liu, and J. Ye, "A convex formulation for learning shared structures from multiple tasks," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. 35, no. 5, pp. 1025-1038, May 2013.

[27] D. Zhang and D. Shen, "Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease," *NeuroImage*, vol. 59, no. 2, pp. 895-907, Jan. 2012.

[28] S. Bickel, J. Bogojeska, T. Lengauer, and T. Scheffer, "Multi-task learning for HIV therapy screening," in *Proc. of Conference on Machine Learning*, Helsinki, Finland, July 2008, pp. 56-63.

[29] J. Zhou, J. Chen, and J. Ye, "Multi-task learning: theory, algorithms and applications," *SDM Tutorial, SIAM Conference on Data Mining*, Anaheim, CA, Apr. 2012, pp. 1-83.

[30] L. Xu, J. Li, W. Lin, Y. Zhang, L. Ma, Y. Fang, Y. Zhang, and Y. Yan, "Multi-task rank learning for image quality assessment," in *Proc. of IEEE Conference on Acoustics, Speech and Signal Processing*, Brisbane, Australia, Apr. 2015, pp. 1339-1343.

[31] M. Fazel, H. Hindi, and S. P. Boyd, "A rank minimization heuristic with application to minimum order system approximation," in *Proc. of American Control Conference*, Arlington, VA, June 2001, pp. 4734-4739.

[32] S. Ji and J. Ye, "An accelerated gradient method for trace norm minimization," in *Proc. of Conference on Machine Learning*, Montreal, Canada, June 2009, pp. 457-464.

[33] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Journal on Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121 – 167, 1998.

[34] J. Zhou, J. Chen, and J. Ye, *MALSAR: Multi-task learning via structural regularization*, Arizona State University, 2012 [Online]. Available: http://www.public.asu.edu/~jye02/software/MALSAR

[35] C. –C. Chang and C. –J. Lin, *LIBSVM: A library for support vector machines* [Online]. Available: https://www.csie.ntu.edu.tw/~cjlin/libsvm