Artificial Intelligence Research and Development
A. Cortés et al. (Eds.)

doi:10.3233/FAIA220345

Referenceless Image Quality Assessment Utilizing Deep Transfer-Learned Features

Basma AHMED[1], Osama A. OMER[3], Amal RASHED[1] and Domenec PUIG[2], Mohamed ABDEL-NASSER[2,3]
[1]*Faculty of Computers and Information, South Valley University, Qena, Egypt*
[2]*Computer Engineering and Mathematics Department, University Rovira i Virgili, Tarragona, Spain*
[3]*Electrical Engineering Department, Aswan University, Aswan, Egypt*

Abstract. Image quality assessment (IQA) algorithms are critical for determining the quality of high-resolution photographs. This work proposes a hybrid NR IQA approach that uses deep transfer learning to enhance classic NR IQA with deep learning characteristics. Firstly, we simulate a pseudo reference image (PRI) from the input image. Then, we used a pre-trained inception-v3 deep feature extractor to generate the feature maps from the input distorted image and PRI. The distance between the feature maps of the input distorted image and PRI are measured using the local structural similarity (LSS) method. A nonlinear mapping function is used to calculate the final quality scores. When compared to previous work, the proposed method has a promising performance.

Keywords. Blind image quality, Similarity measures, Pseudo-reference, Deep learning

1. Introduction

Several quality degradations are introduced in each phase of the visual communication system, e.g., capturing, transmission, compression, and display. However, high-fidelity images are needed in many fields, e.g., remote sensing image recognition, virtual reality, medical imaging, and other fields. Image quality assessment (IQA) algorithms can quantify the quality of visual content delivered to end-users, which can be adopted as the quantify criteria or optimization goal embedded in the visual communication systems [1]. IQA methods can be split into subjective and objective, according to the need for human eyes for ranking [2]. Objective assessment is more feasible and extensively applied because a machine can automatically forecast the quality of an image utilizing mathematical models. Objective assessment can be generally divided into three categories according to the presence or deficiency of a reference image: 1) full-reference IQA (FR-IQA), 2) reduced-reference IQA (RR-IQA), and 3) no-reference or blind IQA (BIQA) [14]. A reference image is often not given in practical use processes, which hinders the application remit of FR-IQA and RR-IQA. This study is focused on NR IQA. In the literature, many NR IQA has been proposed. For instance, Le et al. [6] suggested a blind technique to foretell the visual quality of multiply distorted images based on structural degradation. A structural feature is extracted as the gradient-weighted histogram of the local binary pattern (LBP) studied on the gradient map. They also utilized LBP extracted from texture and structural maps. Q. Wu et al. [7] suggested a local pattern statistics index (LPSI) mechanism by selecting statistical features extracted from binary patterns of local image structures. Min et al. [8] proposed a blind image quality assessment based on a pseudo reference image, and they used a “reference” called pseudo reference image (PRI) to assess blockiness, sharpness, and noisiness. Jinbin Hu et al. [1] proposed a deep network-based blind image quality assessment utilizing two-side pseudo reference images. In turn, deep convolutional neural networks (CNNs) are the most successful architectures for image analysis. Deep transfer learning makes it possible to use previously learned features from one learning task to another learning task as initial features. Kang et al. [9] suggested a CNN to assess image quality without a reference image.

By defining a *pseudo image* as a benchmark, the problem of NR IQA can be achieved via a pseudo reference (PR) IQA solution to bridge a practical approach between FR and NR IQAs. In such a solution, a perfect undistorted picture is not required and may not possibly be available. Besides, we can conclude that transfer learning can extract robust quality-aware features from the image without needing labelled data. This paper proposes a hybrid NR IQA method that enriches traditional NR IQA with deep learning features via deep transfer learning. The method first generates a PRI from the input distorted image. A pre-trained CNN feature extractor has been employed in our study to generate feature maps from the input distorted image and its PRI. We use the local structural similarity (LSS) to measure the distance between the feature maps extracted from the input distorted image and a PRI generated from the input distorted image. A nonlinear mapping function is used to compute the final quality scores.

2. Methodology

Figure 1 presents the proposed NR IQA method. The main components of the proposed method are the simulation of PRI, extracting deep learning features, generation of feature maps, calculation of the similarity between feature maps of the input distorted image and PRI, and applying nonlinear mapping on the similarity scores to compute the final NR IQA score for the input distorted image.

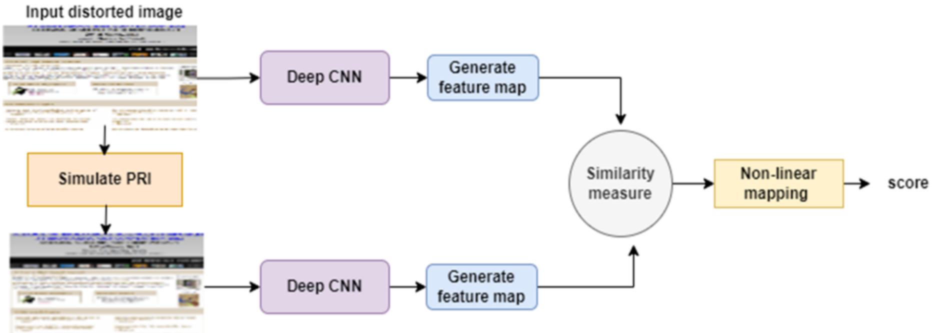

Figure 1. Proposed NR IQA method.

Following the fact that excessive blurring introduces pseudo structures that can be used to judge the quality of the blurred image. In this study, we use a 3 × 3 averaging filter to derive the PRI. Given a distorted image I, PRI is computed as follows:

$$\mathrm{PRI} = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} * \mathrm{I} \tag{1}$$

We utilize the pre-trained inception-v3 deep CNN that is 48 layers deep, trained on more than a million images from the ImageNet database to extract features from images. We use the local structural similarity (LSS) to measure the similarity between the feature map of the distorted image I and the corresponding PRI feature map. The local structure maps of the distorted image I and the PRI can be described as $K_d = \left(k_{d^{ij}}\right)_{h\times w}$, $K_m = \left(k_{m^{ij}}\right)_{h\times w}$ respectively, where h, w denote the rows and columns of the image.

We determine the overlap between $\mathrm{K_d}, \mathrm{K_m}$ as follows:

$$k_{0^{ij}} = \left(k_{d^{ij}}.k_{m^{ij}}\right)_{h\times w} \; , \tag{2}$$

We can also define the union between them as follows:

$$k_{u^{ij}} = \left(k_{d^{ij}} \setminus k_{m^{ij}}\right)_{h\times w} \tag{3}$$

Then, LSS can be defined as follows:

$$LSS = \frac{\sum_{i,j} k_{0^{ij}}}{\sum_{i,j} k_{u^{ij}} + 1} \tag{4}$$

We employ a five-parameter logistic function as a non-linear mapping function to map the quality scores, which can be defined as follows:

$$q' = \beta_1\left(\frac{1}{2} - \frac{1}{1 + \exp\left(\beta_2(q - \beta_3)\right)}\right) + \beta_4 q + \beta 5, \tag{5}$$

where $\dot{q}$ and q stand for the original and mapped quality scores, respectively; $\{\beta_j | j = 1,2, \ldots .5\}$ are five parameters determined through curve fitting. In the literature, the $\dot{q}$ values are considered for evaluation metrics computation.

Two IQA databases are adopted as testing platforms, SIQAD [13] and CSIQ [12]. All datasets consist of numerous subsets of different distortion types. In this work, we focus on Gaussian blurring. The CSIQ dataset contains 30 original screen images and 150 Gaussian blur images. The SIQAD dataset contains 20 original screen images and 140 Gaussian blur images. In this study, three performance indexes are adopted to measure the proposed method:

(1) Pearson Linear Correlation Coefficient (PLCC):

$$\mathrm{PLCC} = \frac{1}{n-1}\sum_{j=1}^{n}\left(\frac{x_j - \bar{x}}{\sigma_x}\right)\left(\frac{y_j - \bar{y}}{\sigma_y}\right) \tag{6}$$

where $\{x_1, x_2, \ldots, x_n\}, \{y_1, y_2, \ldots, y_n\}$ stand for a MOS value and a predicted one, respectively, $\bar{x}$ and $\bar{y}$ are their average scores, and σ_x and σ_y are their variances.

(2) Spearman's Rank Ordered Correlation Coefficient (SROCC):

$$\mathrm{SROCC} = 1 - \frac{6}{n(n^2-1)}\sum_{j=1}^{n}\left(r_{x_j} - r_{y_j}\right)^2, \tag{7}$$

Where r_{x_j} and r_{y_j} represent the rank of x_j and y_j in MOS values and predicted ones, respectively.

(3) Root-Mean-Square Error (RMSE):

$$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{j=1}^{n}\left(x_j - y_j\right)^2\right]^{\frac{1}{2}}, \tag{8}$$

RMSE is a metric that measures the absolute error between the subjective and objective scores.

3. Preliminary results

We compared the performance of the proposed model with NR IQA models such as DIIVINE [4], BLIINDS-II [10], BRISQUE [3], CORNIA [11], NIQE [5]. As shown in Table 1, the proposed method achieves PLCC, SROCC and RMSE values of 0.7117, 0.6035, and 8.0729 with the SIQAD dataset, respectively. From this table, we can see that the proposed model is comparable to the best-performed metrics. With the CSIQ dataset, Table 2 shows that BRISQUE achieves PLCC and SROCC values of 0.9275 and 0.9026. The proposed method achieves 0.1015 RMSE, better than DIIVINE, BLIINDS-II, BRISQUE, CORNIA, NIQE metrics. To test the proposed method in predicting the quality of images with Gaussian blur GB, we compare the predicted score and ground-truth DMOS scores in Fig.2. The deviation between the DMOS and predicted scores is relatively small, reflecting high predictors in agreement with human visual perception.

Table 1. Comparison with existing IQA algorithms for SIQAD dataset.

Algorithm	PLCC	SROCC	RMSE
DIIVINE	0.4632	0.0870	13.450
BLIINDS-II	0.4585	0.4404	13.487
BRISQUE	0.6597	0.6318	11.405
CORNIA	0.6834	0.6497	11.079
NIQE	0.6066	0.5266	12.065
Proposed	0.7117	0.6035	8.0729

Table 2. Comparison with existing IQA algorithms for CSIQ dataset.

Algorithm	PLCC	SROCC	RMSE
DIIVINE	0.8993	0.8716	0.1253
BLIINDS-II	0.8930	0.8766	0.1290
BRISQUE	0.9275	0.9026	0.1071
CORNIA	-	-	-
NIQE	0.9272	0.8925	0.1090
Proposed	0.7322	0.6651	0.1015

DMOS =84.16
Predicted score=0.9519

DMOS =72.77
Predicted score=0.9437

Figure.2. Examples of the results of the proposed method

4. Conclusion

This paper presented a hybrid NR IQA method that enriches traditional NR IQA with deep transfer learning. The method generates a PRI from the input distorted image and then uses pre-trained deep feature extractors to produce feature maps for both images. LSS is used to measure the similarity between the feature maps followed by a nonlinear mapping function to compute the final quality score of the input distorted image. The preliminary results demonstrated that our method had achieved a promising performance. Future work will focus on employing more deep feature extractors to improve the proposed method's performance further.

Acknowledgment

This research was partly supported through Project PID2019-105789RBI00.

References

[1] Hu J, Wang X, Shao F, Jiang Q. TSPR: Deep network-based blind image quality asessment using two-side pseudo reference images. Digital Signal Processing. 2020; 106:102849
[2] Yang G, Wang Y. Deep Superpixel-based Network for Blind Image Quality Assessment. arXiv preprint arXiv:211006564. 2021
[3] Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Transactions on image processing. 2012;21(12):4695-708
[4] Moorthy AK, Bovik AC. Blind image quality assessment: From natural scene statistics to perceptual quality. IEEE transactions on Image Processing. 2011;20(12):3350-64
[5] Mittal A, Soundararajan R, Bovik AC. Making a “completely blind” image quality analyzer. IEEE Signal processing letters. 2012;20(3):209-12
[6] Li Q, Lin W, Fang Y. No-reference quality assessment for multiply-distorted images in gradient domain. IEEE Signal Processing Letters. 2016;23(4):541-5
[7] Wu Q, Wang Z, Li H, editors. A highly efficient method for blind image quality assessment. 2015 IEEE International Conference on Image Processing (ICIP); 2015: IEEE
[8] Min X, Gu K, Zhai G, Liu J, Yang X, Chen CW. Blind quality assessment based on pseudo-reference image. IEEE Transactions on Multimedia. 2017;20(8):2049-62
[9] Kang L, Ye P, Li Y, Doermann D, editors. Convolutional neural networks for no-reference image quality assessment. Proceedings of the IEEE conference on computer vision and pattern recognition; 2014
[10] Saad MA, Bovik AC, Charrier C. Blind image quality assessment: A natural scene statistics approach in the DCT domain. IEEE transactions on Image Processing. 2012;21(8):3339-52.

[11] Ye P, Kumar J, Kang L, Doermann D, editors. Unsupervised feature learning framework for no-reference image quality assessment. 2012 IEEE conference on computer vision and pattern recognition; 2012: IEEE
[12] Larson EC, Chandler DM. Most apparent distortion: full-reference image quality assessment and the role of strategy. Journal of electronic imaging. 2010;19(1):011006
[13] Yang H, Fang Y, Lin W. Perceptual quality assessment of screen content images. IEEE Transactions on Image Processing. 2015;24(11):4408-21
[14] Wang Z, Bovik AC. Modern image quality assessment. Synthesis Lectures on Image, Video, and Multimedia Processing. 2006;2(1):1-156