



香港城市大學
City University of Hong Kong

專業 創新 胸懷全球
Professional · Creative
For The World

CityU Scholars

Learning from Mixed Datasets

A Monotonic Image Quality Assessment Model

Feng, Zhaopeng; Zhang, Keyang; Jia, Shuyue; Chen, Baoliang; Wang, Shiqi

Published in:

Electronics Letters

Published: 01/02/2023

Document Version:

Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

License:

CC BY-NC-ND

Publication record in CityU Scholars:

[Go to record](#)

Published version (DOI):

[10.1049/ell2.12698](https://doi.org/10.1049/ell2.12698)

Publication details:

Feng, Z., Zhang, K., Jia, S., Chen, B., & Wang, S. (2023). Learning from Mixed Datasets: A Monotonic Image Quality Assessment Model. *Electronics Letters*, 59(3), Article e12698. <https://doi.org/10.1049/ell2.12698>

Citing this paper

Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher's definitive version for pagination and other details.

General rights

Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission

Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy

Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.

Learning from mixed datasets: A monotonic image quality assessment model

Zhaopeng Feng,¹ Keyang Zhang,² Shuyue Jia,³
Baoliang Chen,³ and Shiqi Wang³

¹School of Mechanical Engineering and Automation, Harbin Institute of Technology, Shenzhen, China

²Department of Electrical Engineering, City University of Hong Kong, Hong Kong, Hong Kong SAR

³Department of Computer Science, City University of Hong Kong, Hong Kong, Hong Kong SAR

E-mail: shiqiwan@cityu.edu.hk

Deep learning based image quality assessment models usually learn to predict image quality from a single dataset, leading the model to overfit specific scenes. To account for this, mixed datasets training can be an effective way to enhance the generalization capability of the model. However, it is nontrivial to combine different image quality assessment datasets, as their quality evaluation criteria, score ranges, view conditions, as well as subjects are usually not shared during the image quality annotation. Instead of aligning the annotations, this paper proposes a monotonic neural network for image quality assessment model learning with different datasets combined. In particular, this model consists of a dataset-shared quality regressor and several dataset-specific quality transformers. The quality regressor aims to obtain the perceptual quality of each image of each dataset and the quality transformer maps the perceptual quality to the corresponding annotation monotonically. The experimental results verify the effectiveness of the proposed learning strategy and the code is available at <https://github.com/fzp0424/MonotonicIQA>.

Introduction: Image quality assessment (IQA) model which aims to provide image quality objectively has been widely used as a key component in various computer vision tasks [1]. Popular IQA models can be classified into full-reference IQA (FR-IQA) models and no-reference (NR-IQA) models, depending on whether the reference image is available. FR-IQA models can be dated back to the 1970s when the signal fidelity measures are adopted for quality prediction, such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM) [2], and multi-scale SSIM (MS-SSIM) [3]. Compared with FR-IQA, NR-IQA is regarded as a more challenging task due to the lack of reference information. Typical NR-IQA models rely on the natural scene statistics (NSS) construction [4–6] with the assumption that the image quality can be estimated by measuring the destruction level of the NSS. In re-



Fig. 1 Different IQA datasets consist of images with different contents, distortion types, and score ranges. The distortion type and mean opinion score (MOS) value are shown below each sub-image. (a–d) Images sampled from CSIQ [20] and (a) is the reference image. (e–h) Images sampled from LIVE [21] and (e) is the reference image. (i–l) Images sampled from KonIQ-10k [26]

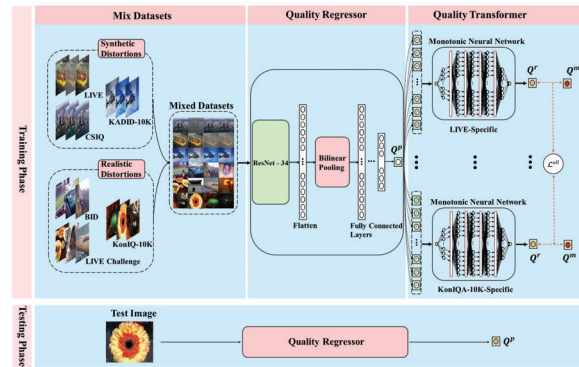


Fig. 2 The overview framework of our proposed method. It contains two main components: dataset-shared quality regressor and dataset-specific quality transformer. In the testing phase, only the quality regressor is used no matter which dataset the testing image is from

cent years, many pioneers resort to deep learning technology for both FR-IQA [7–11] and NR-IQA [8, 12–15], endowing the IQA models with a strong capability for quality-aware feature extraction. In particular, a shallow ConvNet was adopted for patch-based NR-IQA learning in [12]. Both the quality estimation and distortion identification were learned in a multi-task manner for NR-IQA in [16]. The rank information between image pairs was utilized in [17–19] which was able to enrich the training data efficiently.

Generally speaking, the performance of deep learning based methods heavily relies on the training data. An ideal training set should contain sufficiently diverse image contents, plenty of distortion types, and reliable annotations. However, it is an extremely laborious task to collect and construct such a dataset, leading to the constrained size of existing IQA datasets. The lack of training data usually brings the over-fitting problem and leads to the data-driven methods with poor generalization capability. An efficient way to enrich the training data is to combine the existing datasets together for training. As shown in Figure 1, the image contents and distortion types in different IQA datasets can be diverse, implying the promising potential for generalized IQA model learning when they are mixed together. Despite the advantages, the quality annotations of existing datasets are usually not fully aligned. As shown in Figure 1, mean opinion scores (MOSs) range of the CSIQ dataset [20] is from 0 to 1 while the range is [0,100] in the LIVE [21] dataset. Herein, one can re-scale the MOSs of each dataset non-linearly for training, while the re-scaling function is difficult to be acquired reliably. To account for this, Zhang et al. proposed to use the pairwise rank information in each dataset for mixed datasets training [22]. However, the number of rank pairs will increase explosively with the image number, causing the model convergence for a long-time. In [23], continual learning was adopted for diverse datasets learning while the catastrophic forgetting problem may not be completely avoided. Ma et al. proposed an incremental learning scheme for cross-task NR-IQA [24]. However, the pruning operator potentially causes performance degradation on the old task. In comparison, we learn the model from all the datasets simultaneously, achieving a better performance balance among different datasets. Inspired by the four-parameter logistic function, Li et al. proposed a non-linear mapping layer to map the relative quality to the perceptual quality in [25], while the fitting capability of the non-linear function might be questioned.

In this paper, we develop a new paradigm for mixed dataset training by introducing a monotonic neural network. In particular, we first adopt a dataset-shared network for perceptual quality regression of each image and subsequently map the perceptual quality monotonously to the annotated MOSs with the quality ranking in each dataset. The proposed training strategy enjoys two desired advantages. First, it gets rid of the laborious quality re-alignment to successfully combine different datasets for training. Second, the learnable monotonic neural network possesses a powerful non-linear fitting capability to map the image perceptual quality to the annotations. We evaluate our method on six IQA datasets that extensively cover both synthetic distortions and realistic distortions. Compared with state-of-the-art NR-IQA models, our method presents

Table 1. The median Spearman's rank-order correlation coefficient (SRCC) and Pearson linear correlation coefficient (PLCC) of ten splitting results. The databases used for training models are included in the bracket

Database	SRCC							PLCC						
	LIVE	CSIQ	KADID-10k	BID	LIVE Challenge	KonIQ-10k	Weighted	LIVE	CSIQ	KADID-10k	BID	LIVE Challenge	KonIQ-10k	Weighted
NIQE	0.906	0.632	0.374	0.468	0.464	0.521	0.471	0.908	0.726	0.428	0.461	0.515	0.529	0.503
ILNIQE	0.907	0.832	0.531	0.516	0.469	0.507	0.528	0.912	0.873	0.573	0.533	0.536	0.534	0.576
PQR (BID)	0.663	0.522	0.321	–	0.691	0.614	–	0.673	0.612	0.403	–	0.740	0.650	–
PQR (KADID-10k)	0.916	0.803	–	0.358	0.439	0.485	–	0.927	0.870	–	0.405	0.501	0.484	–
DB-CNN (CSIQ)	0.855	–	0.501	0.329	0.451	0.499	–	0.854	–	0.569	0.383	0.472	0.515	–
DB-CNN (LIVE Challenge)	0.723	0.691	0.488	0.809	–	0.770	–	0.754	0.685	0.529	0.832	–	0.825	–
Koncept512 (KonIQ-10k)	0.776	0.639	0.491	0.800	0.781	0.917	0.800	0.772	0.663	0.508	0.818	0.844	0.931	0.727
UNIQUE-Linear (All databases)	0.935	0.821	0.870	0.809	0.799	0.868	0.865	–	–	–	–	–	–	–
UNIQUE (All databases)	0.969	0.902	0.878	0.858	0.854	0.896	0.888	0.968	0.927	0.876	0.873	0.890	0.901	0.892
Proposed (All databases)	0.958	0.898	0.885	0.847	0.854	0.902	0.892	0.965	0.920	0.887	0.862	0.886	0.911	0.900

Table 2. The median SRCC of ten splitting results using different structure of the proposed monotonic neural network

SRCC	LIVE	CSIQ	KADID-10k	BID	LIVE Challenge	KonIQ-10k	Weighted
Three CFCLs and two ELUs	0.963	0.821	0.806	0.862	0.851	0.903	0.857
Five CFCLs and four ELUs (Proposed)	0.958	0.898	0.885	0.847	0.854	0.902	0.892
Seven CFCLs and six ELUs	0.965	0.812	0.834	0.860	0.855	0.902	0.868

superior performance for image quality prediction, revealing the effectiveness of the proposed monotonic neural network.

Proposed method: The overview framework of our method is presented in Figure 2. As shown in Figure 2, our method contains two main components: (1) the dataset-shared quality regressor which aims to obtain the perceptual quality of each image in the mixed datasets; (2) the dataset-specific quality transformer which monotonously maps the perceptual quality to the MOSs annotated in each dataset. The details of each component are elaborated.

Quality regressor: For the perceptual quality regression, we first adopt the pre-trained ResNet-34 [27] to extract the quality-aware feature of each image. Then we flatten the spatial dimensions of the feature representation after the last convolution to obtain the feature $\mathbf{F} \in \mathbb{R}^{c \times l}$, where c is the channel number and l is the spatial length. Analogous to [22], we perform the bilinear pooling on \mathbf{F} , aiming to acquire the interactions between different channels,

$$\bar{\mathbf{F}} = \mathbf{F} \times \mathbf{F}^T, \quad (1)$$

where $\bar{\mathbf{F}} \in \mathbb{R}^{c \times c}$. Based upon $\bar{\mathbf{F}}$, two fully connected layers (each followed by a ReLU activation) with a final fully connected layer (the output dimension is one) appended are utilized to regress the $\bar{\mathbf{F}}$ to the perceptual quality Q^p . The quality regressor is shared by different datasets.

Quality transformer: The connection between the perceptual quality Q^p of each image and its MOS Q^m is that their rank should be the same within the dataset. Following this principle, we propose a monotonic neural network to map the Q^p to the Q^m . The architecture of our monotonic neural network consists of five constrained fully connected layers (CFCL) whose weights are all constrained to be positive. The positive weight ensures the gradient of the mapping function to be positive otherwise higher Q^p will be mapped to the higher Q^m and vice versa. In addition, to maintain the network monotonicity and enhance the transformer non-linear fitting ability, we adopt the ELU [28] as the activation function. We provide a proof of the monotonicity of the quality transformer at https://drive.google.com/file/d/1OGwv6y2sSgvOAY_s53gaKYAL_BtNP3Gx/view?usp=share_link.

Loss function: We denote the mapping result (output) of the quality transformer as Q^r and the smooth l_1 loss \mathcal{L}^{sl} and normal-in-normal loss \mathcal{L}^{nin} [29] are adopted for training as follows,

$$\mathcal{L}_{all} = \mathcal{L}_{sl}(Q^r, Q^m) + \lambda \mathcal{L}_{nin}(Q^r, Q^m), \quad (2)$$

where λ is a hyper-parameter and

$$\mathcal{L}_{sl}(Q^r, Q^m) = \begin{cases} \frac{1}{2}(Q^r - Q^m)^2 & \text{if } |Q^r - Q^m| < 1, \\ |Q^r - Q^m| - \frac{1}{2} & \text{otherwise,} \end{cases} \quad (3)$$

and

$$\mathcal{L}_{nin}(Q^r, Q^m) = \frac{1}{2} \left(\frac{Q^r - \mu^r}{\sigma^r} - \frac{Q^m - \mu^m}{\sigma^m} \right)^2, \quad (4)$$

with the μ^r , σ^r , μ^m , and σ^m are the mean and standard deviation of the predicted quality scores and MOSs in a training batch. Our model is trained in an end-to-end manner.

Experiments: To verify the effectiveness of our method, six IQA databases are used and mixed for training, including LIVE [21], CSIQ [20], KADID-10k [30], BID [31], LIVE Challenge [32] and KonIQ-10k [26]. The synthetic distortions are involved in the first three datasets and the last three datasets contain the realistic distortions. We randomly split each database into three subsets for training, validation, and testing. There is no content overlap among the three subsets. The split proportions are 60%, 20%, and 20%, respectively. We apply Spearman's rank-order correlation coefficient (SRCC) and Pearson linear correlation coefficient (PLCC) to evaluate the monotonicity and accuracy. The larger SRCC and PLCC indicate better quality prediction results. We linearly re-scale the MOS ranges of each dataset to [0,10]. We implement our method by PyTorch [33] with Adam [34] as the optimizer. The learning rates of the quality regressor and quality transformer are 3×10^{-5} and 3×10^{-4} , respectively. The learning rates of the pre-trained ResNet-34 and the rest of the model are initialized to 3×10^{-5} and 3×10^{-4} , respectively. In the training phase, we set the batch size by 32 and the λ in Equation (2) is 1.0. We first re-scale the short length of images in the training sets and test sets to 512 and 768, respectively, with their aspect ratios maintained. Then patches with a

size of $384 \times 384 \times 3$ are cropped from each training sample for data augmentation during training.

Performance comparisons: We compare our method with six state-of-the-art NR-IQA models, including NIQE [35], ILNIQE [36], PQR [37], DB-CNN [38], Koncept512 [26] and UNIQUE [22]. The linear re-scaling strategy proposed in [22] is also included and denoted as UNIQUE-Linear [22] in Table 1. The median SRCC and PLCC of ten splitting results are reported. For better comparison, the image number weighted SRCC and PLCC of all six datasets are also presented. As shown in Table 1, we can observe that the NSS-based methods including NIQE [35] and ILNIQE [36] achieve better performance on synthetic datasets while are less effective on realistic datasets. The performance of deep learning based methods is determined by the training set. For example, when the DB-CNN [38] is trained on the CSIQ [20] dataset, higher SRCC can be achieved on the LIVE [21] and KADID-10k [30] datasets than it on the LIVE Challenge [32], BID [31], and KonIQ-10k [26] datasets. On the contrary, the performance on synthetic datasets will drop dramatically when the KonIQ-10k [26] is used for training. Compared with the models that are trained on a single dataset, significant performance improvement can be observed by UNIQUE [22] and our method, revealing the effectiveness of the mixed dataset training. Moreover, our method achieves the best performance in terms of both the weighted SRCC and PLCC, demonstrating the superiority of the monotonic neural network based learning strategy when compared with the linear re-scaling strategy (UNIQUE-Linear) or the rank-based strategy (UNIQUE).

As shown in Table 2, we further conduct extra experiments to study the best number of CFCL in our proposed monotonic neural network. From Table 2, we can observe the best layer number is five. Less or more layers will cause the performance to drop to some extent. The reason may lie in that fewer layers will bring a negative effect on the non-linear fitting ability and more layers will cause the overfitting problem. To verify our model efficiency, we further compare the floating point operations (FLOPs) that our method consumed during inference with the FLOPs of the other three deep-learning based methods, including the Koncept512, DB-CNN, and UNIQUE. The results show that our method is comparable with the UNIQUE (3.95G vs 3.68G) and presents a significant gain compared with DB-CNN (16.53G) and Koncept512 (13.22G). It reveals that our method achieves better quality prediction performance without much sacrifice of model complexity.

Conclusion: In this paper, we concentrate on the NR-IQA model learning from mixed datasets. To avoid the laborious quality annotation alignment, a monotonic neural network is proposed to map the regressed quality to the MOS values. Experimental results demonstrate that the proposed method is more effective than both the unreliable linear re-scaling strategy and pair rank strategy. We believe our method is able to shed the light on the exploration of more generalized IQA models and inspire more works for learning monotonic neural networks.

Author contributions: Zhaopeng Feng was in charge of literature reviewing, methodology development, and experiments conducting. Keyang Zhang collected the datasets. Shuyue Jia validated the experimental results. Baoliang Chen wrote the draft. Shiqi Wang was in charge of the paper revision and experiments analysis. All authors read and approved the final manuscript.

Conflict of interest: The authors declare no conflict of interest.

Data availability statement: The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

© 2023 The Authors. *Electronics Letters* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. Received: 25 September 2022 Accepted: 11 November 2022

doi: 10.1049/ell2.12698

References

- Liu, Y., Yan, J., Ouyang, W.: Quality aware network for set to set recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 5790–5799. IEEE, Piscataway, NJ (2017)
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
- Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: IEEE Asilomar Conference on Signals, Systems and Computers, vol. 2, pp. 1398–1402. IEEE, Piscataway, NJ (2003)
- Moorthy, A., Krishna, Bovik, A.C.: Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **20**(12), 3350–3364 (2011)
- 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)
- Hou, W., Gao, X., Tao, D., Li, X.: Blind image quality assessment via deep learning. *IEEE Trans. Neural Networks Learn. Syst.* **26**(6), 1275–1286 (2014)
- Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 586–595. IEEE, Piscataway, NJ (2018)
- Bosse, S., Maniry, D., Müller, K.-R., Wiegand, T., Samek, W.: Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Trans. Image Process.* **27**(1), 206–219 (2017)
- Kim, J., Lee, S.: Deep learning of human visual sensitivity in image quality assessment framework. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1676–1684. IEEE, Piscataway, NJ (2017)
- Ding, K., Ma, K., Wang, S., Simoncelli, E.P.: Image quality assessment: Unifying structure and texture similarity. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(5), 2567–2581 (2020)
- Chen, B., Zhu, L., Kong, C., Zhu, H., Wang, S., Li, Z.: No-reference image quality assessment by hallucinating pristine features. *IEEE Trans. Image Process.* **31**, 6139–6151 (2022)
- Kang, L., Ye, P., Li, Y., Doermann, D.: Convolutional neural networks for no-reference image quality assessment. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1733–1740. IEEE, Piscataway, NJ (2014)
- Gu, J., Meng, G., Xiang, S., Pan, C.: Blind image quality assessment via learnable attention-based pooling. *Pattern Recognition* **91**, 332–344 (2019)
- Fu, J., Wang, H., Zuo, L.: Blind image quality assessment for multiply distorted images via convolutional neural networks. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1075–1079. IEEE, Piscataway, NJ (2016)
- Kim, J., Nguyen, A.-D., Ahn, S., Luo, C., Lee, S.: Multiple level feature-based universal blind image quality assessment model. In: IEEE International Conference on Image Processing, pp. 291–295. IEEE, Piscataway, NJ (2018)
- Kang, L., Ye, P., Li, Y., Doermann, D.: Simultaneous estimation of image quality and distortion via multi-task convolutional neural networks. In: IEEE International Conference on Image Processing, pp. 2791–2795. IEEE, Piscataway, NJ (2015)
- Liu, X., van de Weijer, J., Bagdanov, A.D.: RankIQ: Learning from rankings for no-reference image quality assessment. In: IEEE International Conference on Computer Vision, pp. 1040–1049. IEEE, Piscataway, NJ (2017)
- Niu, Y., Huang, D., Shi, Y., Ke, X.: Siamese-network-based learning to rank for no-reference 2D and 3D image quality assessment. *IEEE Access* **7**, 101583–101595 (2019)
- Chen, B., Li, H., Fan, H., Wang, S.: No-reference screen content image quality assessment with unsupervised domain adaptation. *IEEE Trans. Image Process.* **30**, 5463–5476 (2021)
- Larson, E.C., Chandler, D.M.: Most apparent distortion: Full-reference image quality assessment and the role of strategy. *J. Electron. Imaging* **19**(1), 011006:1–011006:19 (2010)
- 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)

- 22 Zhang, W., Ma, K., Zhai, G., Yang, X.: Uncertainty-aware blind image quality assessment in the laboratory and wild. *IEEE Trans. Image Process.* **30**, 3474–3486 (2021)
- 23 Zhang, W., Li, D., Ma, C., Zhai, G., Yang, X., Ma, K.: Continual learning for blind image quality assessment. *IEEE Trans. Pattern Anal. Mach. Intell.* (2022). <https://doi.org/10.1109/TPAMI.2022.3178874>
- 24 Ma, R., Luo, H., Wu, Q., Ngan, K.N., Li, H., Meng, F., Xu, L.: Remember and reuse: Cross-task blind image quality assessment via relevance-aware incremental learning. In: ACM International Conference on Multimedia, pp. 5248–5256. IEEE, Piscataway, NJ (2021)
- 25 Li, D., Jiang, T., Jiang, M.: Unified quality assessment of in-the-wild videos with mixed datasets training. *Int. J. Comput. Vision* **129**(4), 1238–1257 (2021)
- 26 Hosu, V., Lin, H., Sziranyi, T., Saupe, D.: KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment. *IEEE Trans. Image Process.* **29**, 4041–4056 (2020)
- 27 He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778. IEEE, Piscataway, NJ (2016)
- 28 Clevert, D.A., Unterthiner, T., Hochreiter, S.: Fast and accurate deep network learning by exponential linear units (elus). arXiv:1511.07289 (2015)
- 29 Li, D., Jiang, T., Jiang, M.: Norm-in-norm loss with faster convergence and better performance for image quality assessment. In: ACM International Conference on Multimedia, pp. 789–797. Association for Computing Machinery, New York, NY (2020)
- 30 Lin, H., Hosu, V., Saupe, D.: KADID-10k: A large-scale artificially distorted IQA database. In: IEEE International Conference on Quality of Multimedia Experience, pp. 1–3. IEEE, Piscataway, NJ (2019)
- 31 Ciancio, A., Targino da Costa, A.L.N.T., da Silva, E.A.B., Said, A., Samadani, R., Obrador, P.: No-reference blur assessment of digital pictures based on multifeature classifiers. *IEEE Trans. Image Process.* **20**(1), 64–75 (2010)
- 32 Ghadiyaram, D., Bovik, A.C.: Massive online crowdsourced study of subjective and objective picture quality. *IEEE Trans. Image Process.* **25**(1), 372–387 (2015)
- 33 Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: PyTorch: An imperative style, high-performance deep learning library. In: *Neural Information Processing Systems*, pp. 8026–8037. Curran Associates Inc., Red Hook, NY (2019)
- 34 Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: 3rd International Conference for Learning Representations, pp. 1–15. Association for Computing Machinery, New York, NY (2015)
- 35 Mittal, A., Soundararajan, R., Bovik, A.C.: Making a “completely blind” image quality analyzer. *IEEE Signal Process Lett.* **20**(3), 209–212 (2012)
- 36 Zhang, L., Zhang, L., Bovik, A.C.: A feature-enriched completely blind image quality evaluator. *IEEE Trans. Image Process.* **24**(8), 2579–2591 (2015)
- 37 Zeng, H., Zhang, L., Bovik, A.C.: Blind image quality assessment with a probabilistic quality representation. In: IEEE International Conference on Image Processing, pp. 609–613. IEEE, Piscataway, NJ (2018)
- 38 Zhang, W., Ma, K., Yan, J., Deng, D., Wang, Z.: Blind image quality assessment using a deep bilinear convolutional neural network. *IEEE Trans. Circuits Syst. Video Technol.* **30**(1), 36–47 (2020)