https://doi.org/10.2352/issn.2694-118X.2020.LIM-47

CNN-based image quality assessment of consumer photographs

Luigi Celona and Raimondo Schettini
Department of Informatics, Systems and Communication, University of Milano - Bicocca, viale Sarca, 336 Milano, Italy
{***first_name.last_name@unimib.it***}

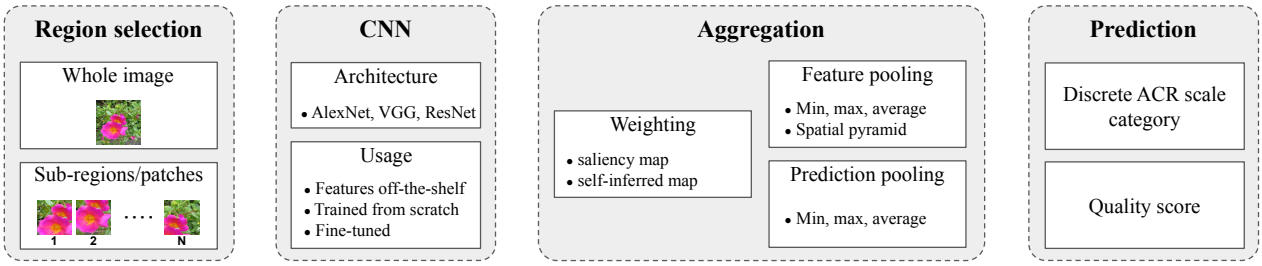

Figure 1: Graphical representation of the main building blocks of a CNN-based blind image quality assessment framework.

Abstract

Convolutional neural networks (CNNs) are used in an increasingly systematic way in a great variety of computer vision applications, including image quality assessment. However, their application to evaluate the perceived quality of images is strongly limited by the lack of adequate and consistent training data. A CNN-based framework for evaluating image quality of consumer photographs is made up of several building blocks that can be implemented in different ways. In this article, we schematically illustrate how these building blocks have been implemented and combined so far to create feasible solutions that take the most positive characteristics of CNNs while mitigating their intrinsic limitations. Some experimental results are reported to show the effectiveness of CNN-based solutions on real-world image quality datasets.

Introduction

It is hard to imagine that someone in our days does not have a smartphone in his pocket. Some statistics tell us that in the US, on average, one person takes 20 photos a day. These photos are not taken only on special occasions, often the photos are sprung to "seize the moment" and they are taken without paying particular attention to the light conditions and shooting parameters. The user often takes several photos of the same subject to increase the probability that at least one of them is the "right one" but then he has to visually select the best shot. Disregarding the image content, aesthetics [2] and emotional aspects [10] may influence the user's judgment of the overall image quality, for example, we have shown that pictures containing faces require a different image quality assessment pipeline [6]. Here we want to concentrate on an unavoidable aspect, whatever is the ultimate purpose of a photo, that is its perceived technical quality.

Convolutional neural networks (CNNs) have been successfully used in the last few years in a huge variety of applications related to image processing and computer vision, so it is reasonable to assume that even in the area of image quality these can far exceed the performance of traditional image quality methods based on hand-crafted features. Their fully exploitation is limited by the fact that neural networks require large amounts of data to be effectively trained. The currently available databases, even if acquired with a great effort, have insufficient cardinality to train architectures that have millions of parameters to be learned. The usual data augmentation techniques cannot be applied to pristine images to generate image databases having the variety and complexity of distortions typical of images acquired with common consumer devices. There is a further aspect that is no less important in our opinion, CNNs are often used as *magic boxes*, magic boxes that often worked very well, but we do not really know why. As a result, we often do not know how to improve the first promising and often exciting results.

In this short article, after briefly describing the characteristics of image databases that can be used to train a network to estimate the perceived quality of a typical consumer photo. We illustrate schematically how it is possible to build processing pipelines based on CNNs and how these are able to obtain experimental results that far exceed those obtained with the most sophisticated methods based on hand-crafted features.

Overview on image quality assessment of consumer photos

Objective image quality assessment (IQA) methods can be categorized into three groups: full-reference (FR) metrics perform a direct comparison between the image under test and a reference or original image; reduced-reference (RR) metrics which predict image quality with only partial information about the reference image; and no-reference (NR) metrics that are applied when the original image is unavailable. NR-IQA metrics assume that the image quality is determinable without a direct comparison between the pristine original and the image under test; however, most of these metrics estimate quality degradation only for specific image artifacts such as blurriness, noise, blockiness, lack of contrast, and lack of colorfulness. There are distortion-generic image quality assessment methods aiming to predict the quality of an image without any knowledge of artifacts characterizing it; these methods are called blind image quality assessment (BIQA). NR-IQA methods can be further classified into two categories: opinion-aware (OA) methods which are trained on distorted images whose quality has been rated by human subjects (e.g. CORNIA [33], HOSA [31], and FRIQUEE [15]), and opinion-unaware (OU) methods (also referred to as opinion-free) which are not trained on human-rated distorted images and measure the deviation of a distorted signal from natural scene statistics (e.g. ILNIQE [35], and dipIQ [24]). Even if OU-BIQA methods have greater potential for good generalization capability, no OU-BIQA method has shown consistently better quality prediction accuracy than the OA method yet. For this reason, in this

paper, we will focus on OA-BIQA methods.

The development of generally applicable BIQA methods requires the collection of representative image databases in terms of authenticity, scale, and diversity. The ecological validity of an IQA database [18] refers to the representativeness of the visual collection for a wide range of real-world photos such as those available on the Internet or acquired using consumer cameras; and the generalization potential of models trained on it. These databases represent a data source for training and parameter fitting of models that attempt to mimic the quality evaluation of human observers. For training OA-BIQA methods, together with the images, it is essential to carefully gather the corresponding subjective quality scores (i.e. the MOS).

The LIVE in the wild image quality challenge database (LIVE-itW) [15] and the KonIQ-10k [18] are two databases designed for blind image quality assessment. The first contains images captured using a representative variety of modern mobile device while the second consists of a selection of images from the YFCC100M [28] database. Both of them provides subjective quality ratings obtained by crowd-sourcing. Recently, for the challenge titled "quality assessment for Smartphone Camera (QA4Camera)" has been released a database composed of 1,500 pictures of 100 various scenes taken using 15 smartphones. For each scene, there are therefore 15 pictures. For each picture four quality rankings related to exposure, color, noise and texture are also provided. Finally, the CCRIQ2+VIME1 [25] database is a collection of two datasets that provide consumer images and corresponding subjective scores obtained by conducing a subjective study in a laboratory. At the moment, no methods have been tested on these two databases. Table 1 depicts the peculiarities of public available databases for BIQA in terms of number of images, subjects involved in the subjective study and annotation type.

Deep blind image quality assessment

Convolutional neural networks (CNNs) have been adopted to estimate the perceived quality of consumer photographs because they can detect and model in an holistic way the factors that influence image quality much better than hand-crafted features. But how can they do that? The process of a photographic image generation is influenced by several interacting sources of distortions that are due to the imaging device properties, and to the image processing pipeline [9]. These distortions are difficult to be individually identified and modeled even for a human observer because, as shown by several authors [27, 29, 26], image quality and image content are strongly related. However, if image content influences image quality evaluation, it is also true the opposite, at least for CNNs: image distortions *strongly* influence image recognition capability of state of the art CNNs trained for object and scene recognition, e.g. [5, 12, 13, 23]. This intrinsic weakness of CNNs is, in our intuition [3, 4], the key for solving the IQA problem. As shown in Figure 2, the resulting activation map produced by the convolving filters of the CaffeNet first convolutional layer trained on the sharp images of ImageNet database [11] is different depending on the quality of image signal. For the pristine image the result of convolution is a well defined activation map where oblique edges are emphasized, while given the noisy image the result of convolution is very different. The initial layers of the CNN see the image degradation and generate a response that can be decoupled from the image content by the subsequent layers.

Figure 1 shows the main building blocks that characterize a CNN-based framework for BIQA. In the next sections we will illustrate the possible implementations for each block.

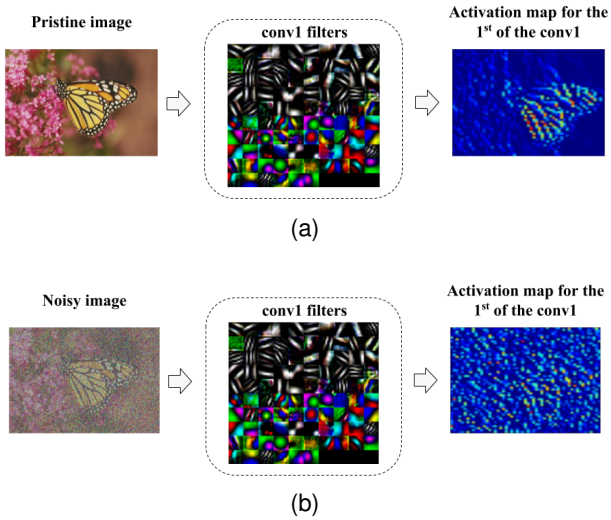

Figure 2: **Activation maps of CaffeNet first filter of the first convolutional layer trained on ImageNet.** (a) The convolution between a pristine image of a monarch and the filter. (b) The convolution between a noise version of the same monarch and the same filter.

Region selection

Depending on the region of the input image to be given to the CNN, three methods can be distinguished: whole image [30], non-overlapping sub-regions/patches [34], and overlapping sub-regions/patches [3]. In the first approach, the whole image is fed to the CNN, while in the second and in the last approach an image is divided into sub-regions and each one of them is independently processed by the CNN. The BIQA methods considering sub-regions can be further divided into those which produce an independent quality prediction for each patch, assuming that the image quality does not vary locally [19, 3], and those in which the patch features or the estimated quality predictions are aggregated to yield a single image prediction [7]. The reason behind the use of sub-regions is twofold: it is a method of data augmentation useful to mitigate the problem of model overfitting due to the reduced cardinality of available databases for BIQA; it prevents resizing operation that can mask some image artifacts for pre-trained CNNs (e.g., AlexNet) requiring a fixed input size. Bianco *et al.* [3] show how the BIQA performance increases according to the number of considered overlapping sub-regions (see Figure 3). Varga *et al.* [30] demonstrated that once the problem of the predetermined resolution has been overcome, the models trained on the whole image perform better than those trained on patches.

CNN model

Existing BIQA approaches rely on standard CNN architectures designed for image classification (e.g. ImageNet) [1]. In [18], a comparative study for BIQA of different CNN architectures is provided which demonstrates how the architectures that work better for image classification work better for BIQA as well. Considering that training a deep network is typically difficult for the small IQA database, some methods tackle the insufficient IQA database by extracting features off-the-shelf from the pre-trained CNN model [14, 21], while others by fine-tuning CNN pre-trained for example for the ImageNet recognition challenge [18, 30]. Features-off-the-shelf and fine-tuning work be-

Table 1: **Overview of the authentic distortions image quality databases.** For the column *Data format*: "MOS" stands for mean opinion score; "σ" is the standard deviation; "ACR distr." is the absolute category rank on 5 classes; "4 ranks" represents four rankings on 4 different image aspects.

Database	No. of unique contents	No. of images	Resolution	No. of subjects	No. of annotations	Data format	Range of scores
LIVE-itW (2016) [15]	1,169	1,169	500×500	$> 8,100$	350K	MOS + σ	0–100
KonIQ-10k (2020) [18]	10,073	10,073	1024×768	1,459	1.2M	MOS + σ + ACR distr.	1–5
QA4Camera (2020)	100	1,500	4032×3024	N/A	N/A	4 ranks	1–15
CCRIQ2+VIME1 (2020) [25]	11	189	1 to 20 MPs	24	N/A	MOS	N/A

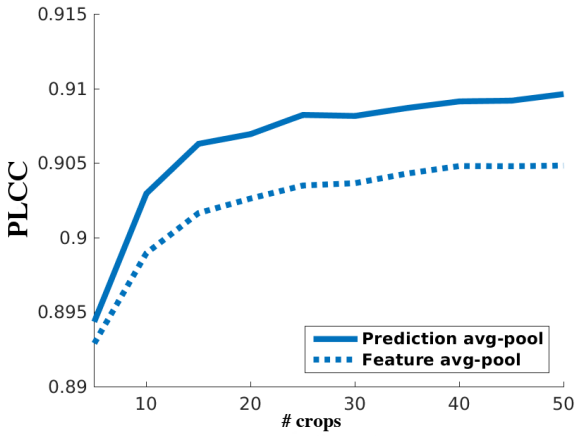

Figure 3: **How the number of selected sub-regions impacts on BIQA performance.** Median Pearson linear correlation coefficient (PLCC) on the LIVE-itW database, with respect to the number of image sub-regions given in input to the fine-tuned CaffeNet. Two aggregation schemes are considered (feature average pooling and prediction average pooling).

cause the training of the network on the good quality images, for example of ImageNet, implies the deep features to intrinsically learn natural scene statistics. Then, the fine-tuning for image quality refines deep features for measuring the deviation from these natural statistics due to the presence of image artifacts.

In DeepBIQ [3], the authors compared the IQA performance of the same CNN architecture trained on databases with different contents both in terms of type and concept cardinality. In such a way they empirically demonstrate that the more concept the CNN has been trained to recognize, the more effective are its features for modeling generic image quality.

Aggregation strategies

Aggregation is necessary to combine sub-regions quality estimates to whole image quality. The methods trained via patch-wise optimization, say methods independently treating each patch during training [19], apply aggregation only during inference. Conversely, image-wise approaches use aggregation also during training [7, 20].

Different strategies can be applied to aggregated sub-regions qualities into a global image-wise quality estimate:

- *feature pooling*: information fusion is performed element by element on the sub-region feature vectors to generate a single feature vector by considering one among minimum, average, or maximum feature pooling.
- *feature concatenation*: information fusion is performed by concatenating the sub-region feature vectors in a single longer feature vector;
- *prediction pooling*: information fusion is performed on the predicted quality scores. A quality score is predicted for each image sub-region, and these scores are then fused using one among minimum, average, or maximum pooling operators.

In [3], all of the previously mentioned strategies have been tested and it has been found that the best aggregation approach is the average prediction pooling. The previous aggregation methods equally consider the contribution of each sub-region features or predictions. Instead, some methods, trying to imitate human visual behavior and to deal with the weak assumption that the image quality does not vary locally, treat each image sub-region differently. To do so, they use a saliency map to weigh the contribution of each sub-region to the estimate of the global quality of the image [22, 8]; or let the CNN deduct the contribution of each sub-region directly by inferring a weight map [7]. There are also more sophisticated methods [30] involving new CNN layers such as the spatial pyramid pooling (SPP) [17] for aggregating features at different resolutions.

Discrete and continuous quality prediction

The output of BIQA methods is commonly expressed in terms of: 5-classes of the absolute category rating (ACR), i.e. bad, poor, fair, good and excellent [3]; distribution on subject ratings for each of the 5-classes of the absolute category rating (ACR) [30]; continuous quality score, i.e. the MOS [18, 3]. The choice of what type of output to produce depends very often on the database ground-truth and it commonly converge in predicting a continuous quality score. However, such a scheme ignores the fact that an image may receive divergent subjective scores from different subjects, which cannot be adequately represented by a single scalar number. This is particularly true on complex, real-world distorted images.

The DeepBIQ is trained for initially categorizing image sub-regions in terms of image quality according to the standard 5-point ACR scale. Given that the LIVE-itW groud-truth consists of MOS, the category labels have been obtained by crisp partitioning the MOS into five disjoint sets. Then, given that database ground-truth consists of continuous scores instead of categorical labels, the method exploits the previously learned deep features to fed a support vector regressor (SVR) machine to estimate the continuous quality score.

Experiments

In this section, we report a comparison of some CNN-based BIQA methods on two databases, namely the LIVE-itW and the KonIQ-10k. The performance of BIQA methods is commonly assessed in terms of Pearson linear correlation coefficient (PLCC) and Spearman rank order correlation coefficient (SROCC). The PLCC measures the linear correlation between the actual scores and the predicted ones and it is defined as follows:

$$PLCC = \frac{\sum_i^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^N (x_i - \bar{x})^2}\sqrt{\sum_i^N (y_i - \bar{y})^2}}, \quad (1)$$

Table 2: Comparison to state-of-the-art methods on the KonIQ-10k and the LIVE-itW databases. *DeepBIQ** indicates the the CNN model has been pretrained on ImageNet instead of Hybrid database.

Method	LIVE-itW		KonIQ-10k	
	PLCC	SROCC	PLCC	SROCC
DeepBIQ (CaffeNet)	0.91	0.89	0.91	0.89
DeepRN (ResNet-101) [30]	0.93	0.91	0.95	0.92
DeepBIQ* (VGG16) [18]	–	–	0.89	0.87
DeepBIQ* (InceptionResNet-v2) [18]	–	–	0.91	0.91
KonCept512 [18]	–	–	0.94	0.92

Table 3: Cross-database comparison among state-of-the-art methods: trained on LIVE-itW and tested on KonIQ-10k and vice versa. *DeepBIQ** indicates the the CNN model has been pretrained on ImageNet instead of Hybrid database.

Method	Train on LIVE-itW Test on KonIQ-10k		Train on KonIQ-10k Test on LIVE-itW	
	PLCC	SROCC	PLCC	SROCC
DeepBIQ (CaffeNet)	0.723	0.711	0.748	0.746
DeepRN (ResNet101) [30]	–	–	0.750	0.726
DeepBIQ* (VGG16) [18]	–	–	0.747	0.742
DeepBIQ* (InceptionResNet-v2) [18]	–	–	0.821	0.804
KonCept512 [18]	–	–	0.848	0.825

where N is the number of samples, x_i and y_i are the sample points indexed with i, finally $\bar{x}$ and $\bar{y}$ are the means of each sample distribution. Instead, the SROCC estimates the monotonic relationship between the actual scores and the predicted ones and it is calculated as follows:

$$SROCC = 1 - \frac{6\sum_i^N d_i^2}{N(N^2-1)}, \tag{2}$$

N is the number of samples, and $d_i = (\text{rank}(x_i) - \text{rank}(y_i))$ is the difference between the two ranks of each sample.

Table 2 shows the comparison of methods trained and evaluated on the same database. We want to highlight that deeper architectures perform better than shallow ones. DeepBIQ obtains almost the same performance for both databases. KonCept512 [18] method achieved the best performance on the KonIQ-10k. It takes images of 512×384 pixels as input for an ImageNet-trained InceptionResNet-v2 model fine-tuned for quality assessment by simultaneously optimizing with respect to the MOS and ACR distribution. No performance is measured on the LIVE-itW database due to the lack of ACR distribution.

We measure cross-database performance in order to evaluate generalization capabilities of the methods. Table 3 provide the performance of algorithms trained on the KonIQ-10k database and evaluated on the LIVE-itW. Compared to the other methods the KonCept512 does not suffer a drop in performance.

Final remarks

The use of CNNs in image quality is still limited by the scarcity of training data. Despite the efforts of some research groups in acquiring such datasets, such as LIVE in the wild quality challenge [15] and KonIQ-10k [18], their cardinality is still too small to train from scratch state of the art CNNs. Several authors are experimenting with different approaches of data augmentation to make up for this lack e.g. [32, 16]. Although such approaches are theoretically valid and promising experimentally their performance on images corrupted by authentic distortions is still not optimal. In the previous section, we also show how recent CNN-based BIQA methods have yielded interesting results on collections of genuinely distorted images of very different scenes. Therefore we can assume that these methods are effective for the automatic selection of photos based on the quality metric. But are these methods possibly able to guide an image enhancement process? The reason why we can say that this still does not happen is twofold: (i) the classic BIQA methods only estimate the MOS, but they do not provide any advice on the types of distortion that affect the image; (ii) there is no evidence, to the best of our knowledge, showing that CNN-based BIQA methods are sensitive to small changes in the severity of distortion on images with the same content.

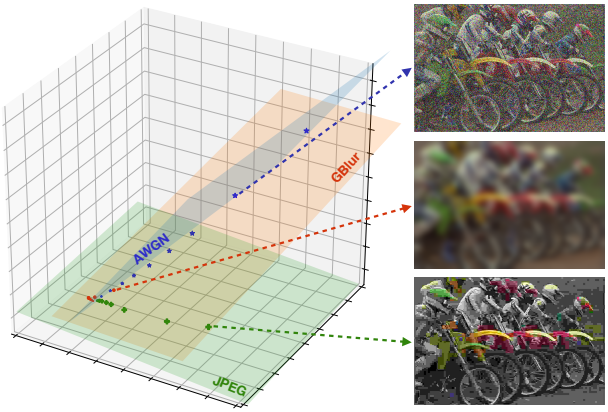

Figure 4: The deep features extracted from a specific layer of a CNN trained for a recognition task, intrinsically divide the deep feature space in such a way that visual representations corresponding to different types of distortions lay on different planes.

We think that distortion recognition is an analysis that should be conducted to explicitly justify the reason for a quality score so as not to make such CNN-based method a *magic box*. A method capable of simultaneously estimating the quality score and the type and severity of the distortions that corrupt an image on a global or local level could be a further extension of real BIQA systems.

As previously stated, pre-trained CNNs are, to some degree, already quality-aware, meaning that their learned internal features assist the performance of the task (recognition) by adapting to the presence of authentic distortions. In [4], Bianco *et al.* demonstrate how this property is effective for characterizing different types of image distortions (see Figure 4). We firmly believe that the deep representations learned under certain conditions can be even better exploited for the development of more effective opinion-unaware, distortion-unaware BIQA algorithms.

References

[1] Simone Bianco, Remi Cadene, Luigi Celona, and Paolo Napoletano. Benchmark analysis of representative deep neural network architectures. *IEEE Access*, 6:64270–64277, 2018.

[2] Simone Bianco, Luigi Celona, Paolo Napoletano, and Raimondo Schettini. Predicting image aesthetics with deep learning. In *ACIVS*, pages 117–125. Springer, 2016.

[3] Simone Bianco, Luigi Celona, Paolo Napoletano, and Raimondo Schettini. On the use of deep learning for blind image quality assessment. *Springer SIViP*, 12(2):355–362, 2018.

[4] Simone Bianco, Luigi Celona, Paolo Napoletano, and Raimondo Schettini. Dinsentangling image distortions in deep feature space. *arXiv preprint arXiv:?.?*, 2020.

[5] Simone Bianco, Luigi Celona, and Raimondo Schettini. Robust smile detection using convolutional neural networks. *Journal of Electronic Imaging*, 25(6):063002, 2016.

[6] Simone Bianco, Luigi Celona, and Raimondo Schettini. Aesthetics assessment of images containing faces. In *ICIP*, pages 2820–2824.

IEEE, 2018.

[7] Sebastian Bosse, Dominique Maniry, Thomas Wiegand, and Wojciech Samek. A deep neural network for image quality assessment. In *ICIP*, pages 3773–3777. IEEE, 2016.

[8] Aladine Chetouani. A blind image quality metric using a selection of relevant patches based on convolutional neural network. In *EUSIPCO*, pages 1452–1456. IEEE, 2018.

[9] Gianluigi Ciocca, Silvia Corchs, Francesca Gasparini, and Raimondo Schettini. How to assess image quality within a workflow chain: an overview. *Springer International Journal on Digital Libraries*, 15(1):1–25, 2014.

[10] Silvia Corchs, Gianluigi Ciocca, and Francesca Gasparini. Quality assessment of jpeg-distorted face images: Influence of affective content. In *ICETE (1)*, pages 552–559, 2018.

[11] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, pages 248–255. IEEE, 2009.

[12] Samuel Dodge and Lina Karam. Understanding how image quality affects deep neural networks. In *2016 eighth international conference on quality of multimedia experience (QoMEX)*, pages 1–6. IEEE, 2016.

[13] Samuel Dodge and Lina Karam. A study and comparison of human and deep learning recognition performance under visual distortions. In *ICCCN*, pages 1–7. IEEE, 2017.

[14] Fei Gao, Jun Yu, Suguo Zhu, Qingming Huang, and Qi Tian. Blind image quality prediction by exploiting multi-level deep representations. *Pattern Recognition*, 81:432 – 442, 2018.

[15] Deepti Ghadiyaram and Alan C Bovik. Massive online crowdsourced study of subjective and objective picture quality. *IEEE TIP*, 25(1):372–387, 2015.

[16] Xiaodi Guan, Fan Li, and Lijun He. Quality assessment on authentically distorted images by expanding proxy labels. *Electronics*, 9(2):252, 2020.

[17] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE TPAMI*, 37(9):1904–1916, 2015.

[18] V. Hosu, H. Lin, T. Sziranyi, and D. Saupe. Koniq-10k: An ecologically valid database for deep learning of blind image quality assessment. *IEEE TIP*, 29:4041–4056, 2020.

[19] Le Kang, Peng Ye, Yi Li, and David Doermann. Convolutional neural networks for no-reference image quality assessment. In *CVPR*, pages 1733–1740. IEEE, 2014.

[20] Jongyoo Kim and Sanghoon Lee. Fully deep blind image quality predictor. *IEEE Journal of selected topics in signal processing*, 11(1):206–220, 2016.

[21] Dingquan Li, Tingting Jiang, and Ming Jiang. Exploiting high-level semantics for no-reference image quality assessment of realistic blur images. In *ACM Multimedia*, pages 378–386, 2017.

[22] Jie Li and Yue Zhou. Visual saliency based blind image quality assessment via convolutional neural network. In *International Conference on Neural Information Processing*, pages 550–557. Springer, 2017.

[23] Ding Liu, Bowen Cheng, Zhangyang Wang, Haichao Zhang, and Thomas S Huang. Enhance visual recognition under adverse conditions via deep networks. *IEEE TIP*, 28(9):4401–4412, 2019.

[24] Kede Ma, Wentao Liu, Tongliang Liu, Zhou Wang, and Dacheng Tao. dipIQ: Blind image quality assessment by learning-to-rank discriminable image pairs. *IEEE TIP*, 26(8):3951–3964, 2017.

[25] Jakub Nawała, Margaret H. Pinson, Mikołaj Leszczuk, and Lucjan Janowski. Study of subjective data integrity for image quality data sets with consumer camera content. *MDPI Journal of Imaging*, 6, 2020.

[26] Kyung Hoon Oh, Sophie Triantaphillidou, and Ralph E Jacobson. Scene classification with respect to image quality measurements. In *Image Quality and System Performance VII*, volume 7529, page 752908. International Society for Optics and Photonics, 2010.

[27] Ernestasia Siahaan, Alan Hanjalic, and Judith A Redi. Semantic-aware blind image quality assessment. *Elsevier Signal Processing: Image Communication*, 60:237–252, 2018.

[28] Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. YFCC100M: The new data in multimedia research. *Communications of the ACM*, 59(2):64–73, 2016.

[29] Sophie Triantaphillidou, Elizabeth Allen, and R Jacobson. Image quality comparison between jpeg and jpeg2000. ii. scene dependency, scene analysis, and classification. *Journal of Imaging Science and Technology*, 51(3):259–270, 2007.

[30] Domonkos Varga, Dietmar Saupe, and Tamás Szirányi. Deeprn: A content preserving deep architecture for blind image quality assessment. In *ICME*, pages 1–6. IEEE, 2018.

[31] Jingtao Xu, Peng Ye, Qiaohong Li, Haiqing Du, Yong Liu, and David Doermann. Blind image quality assessment based on high order statistics aggregation. *IEEE TIP*, 25(9):4444–4457, 2016.

[32] Dan Yang, Veli-Tapani Peltoketo, and Joni-Kristian Kamarainen. CNN-based cross-dataset no-reference image quality assessment. In *ICCV-W*, pages 0–0. IEEE, 2019.

[33] Peng Ye, Jayant Kumar, Le Kang, and David Doermann. Unsupervised feature learning framework for no-reference image quality assessment. In *CVPR*, pages 1098–1105. IEEE, 2012.

[34] Chuang Zhang, Jiawei Xu, Xiaoyu Huang, and Seop Hyeong Park. No-reference image quality assessment using independent component analysis and convolutional neural network. *Springer Journal of Electrical Engineering & Technology*, 14(1):487–496, 2019.

[35] Lin Zhang, Lei Zhang, and Alan C Bovik. A feature-enriched completely blind image quality evaluator. *IEEE TIP*, 24(8):2579–2591, 2015.

Author Biography

Luigi Celona *is a Post-Doctoral Fellow with the DISCo, University of Milano - Bicocca. His current research interests focus on image analysis and classification, machine learning, and face analysis.*

Raimondo Schettini *is a professor at the University of Milano - Bicocca (Italy). Where he heads the Imaging and Vision Lab. He is a fellow of the International Association of Pattern Recognition for his contributions to pattern recognition research and color image analysis.*