



Technological Development of Image Aesthetics Assessment

Ruoyu Zou¹(✉), Jiangbo Xu¹, and Ziyu Xue^{1,2}

¹ State Key Laboratory of Media Convergence and Communication, Communication University of China, Beijing, China

xujiangbo@cuc.edu.cn

² Academy of Broadcasting Science (NRTA), Beijing, China

xueziyu@abs.ac.cn

Abstract. Quantitative research on aesthetics is a classic interdisciplinary research. With the rapid development of deep learning, various approaches have been made in image aesthetics assessment (IAA). Starting from the concept of image aesthetics, this report roughly follows the chronological sequence and first introduces the manual design of image aesthetic features. We divide IAA into generic image aesthetics assessment (GIAA) and personalized image aesthetics assessment (PIAA) to introduce separately in the deep learning part. Majority of approaches are GIAA, which purpose is to simulate general aesthetics. In this section, we separately reviewed representative studies of five assessment methods (aesthetic classification, aesthetic regression, aesthetic distribution, IAA with attributes, aesthetic description). Due to the subjectivity of aesthetics, human's aesthetics will more or less deviate from the generic value. PIAA aims to model the aesthetic preferences of specific user, and the research is of great value. We introduced this novel research in the fifth section. Finally, image aesthetic datasets of different uses are summarized. We hope this comprehensive survey can be helpful to researchers in the field of image and enhance the connection between computer and art.

Keywords: Image aesthetics assessment · Deep learning · Dataset

1 Introduction

The pursuit of beauty is human instinct. IAA can serve as guidance for tasks such as image enhancement, image cropping, image retrieval etc. In daily application scenarios, e-commerce websites use automatically generated IAA results as a guide to select product posters, and smart phones can use IAA to generate photo suggestions. At the same time, as the most widely used means of information recording, aesthetics assessment for images can also be applied in ecology [34], art

This work is supported by Research on Quality Evaluation Method of UHD Video based on Hvc (GJ181901).

© Springer Nature Switzerland AG 2021

Y. Peng et al. (Eds.): ICIG 2021, LNCS 12890, pp. 341–352, 2021.

https://doi.org/10.1007/978-3-030-87361-5_28

[10] and other fields. In view of the mass media resources and people’s growing aesthetic needs, there will be more and more fields using IAA models in the future.

1.1 Image Quality Assessment (IQA)

IQA includes technical quality assessment and aesthetic quality assessment. The purpose of technical quality assessment is to simulate human eyes’ perception of image distortion. For example, TID2013 [35] dataset for technical quality assessment contains 25 types of image distortion, such as artifact, noise, and blur. Aesthetic quality assessment also aims to achieve assessment that is close to subjective feelings. Treat technical quality as fidelity, aesthetic quality is artistic attribute based on that. Technical quality assessment uses objective grading to represent the distortion degree of images, while aesthetic quality assessment uses more complex and subjective evaluation results such as “beautiful” and “ugly”.

Therefore, there is no perfect reference image for IAA, which means that IAA belongs to non-reference quality assessment. Talebi et al. [39] trained a CNN model with both technical quality dataset (TID2013 [35], LIVE [11]) and aesthetic quality dataset AVA [31], and verified that this model had good performance in both technical quality assessment and aesthetic quality assessment. To sum up, the study of IAA improves the requirement of image quality assessment from the basic technical quality to the more complex aesthetic level.

1.2 IAA Research

IAA research was launched less than 20 years, later than the development of machine learning and deep learning theory. Researchers usually adopt a data-driven method, the rise of large photography rating websites and mature subjective rating experiments provide sufficient resources for this method. IAA models mainly fall into two categories: extracting image features and inputting them into machine learning algorithms for decision making, and using neural network for end-to-end assessment (see Fig. 1)

In the first method, there are two ways of image feature extraction: manual design and feature extraction using CNN. Manual designed features often target the basic properties of the image, spatial layout, subject objects, and various photographic rules. Trained CNN can be used as feature extractor for image. Image features extracted by the deep learning model trained on other tasks are called generic deep features. The features extracted by the model trained with aesthetic data are called aesthetic deep features. In the decision making stage, machine learning algorithms such as KNN, SVM, random forest, linear regression and SVR can be used to classify or regress the aesthetic quality.

The end-to-end aesthetic assessment model benefits from the rapid development of deep learning and the establishment of large-scale image aesthetic dataset. Researchers designed different neural network models, calculated the loss between the output of last layer and training label by constructing loss function, and then iteratively updated the parameters of the model by back

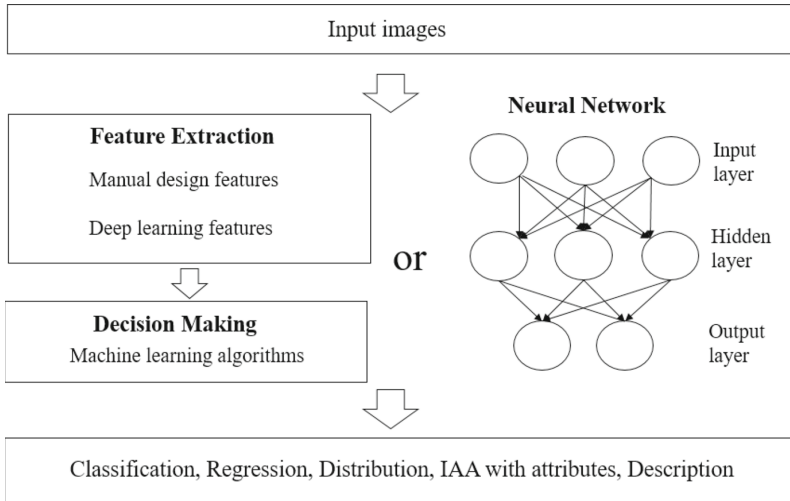


Fig. 1. Two general processes of IAA: Feature extraction and decision making, end-to-end assessment based on neural network.

propagation algorithm. The end-to-end assessment using neural network can be applied to all types of assessment: aesthetic classification, aesthetic regression, aesthetic distribution, IAA with attributes and aesthetic description.

Considering the abstraction of image aesthetics, this report intends to introduce the concept of image aesthetics in the Sect. 2 combining with the research on computational aesthetics. Then in the Sect. 3, we review the attempts of manual design features. IAA based on deep learning is divided into GIAA and PIAA, which are summarized in Sect. 4 and Sect. 5 respectively. There are various types of aesthetic datasets, and a new dataset is often accompanied with a novel research. We summarize the representative aesthetic datasets in Sect. 6. We also notice that IAA takes many forms and is developing. So this report mainly focuses on the research ideas rather than the performance of these models.

2 Inception of Image Aesthetics

Image aesthetics assessment (IAA) is a cross research direction of computational aesthetics, computer vision and psychology, which requires computer scientists to have a certain understanding of image aesthetics. The main difficulty of the study lies in the subjective and abstract aesthetic quality, as well as the variety of assessment methods.

In the field of art, artists tend to pay more attention to the emotions and ideas conveyed by their works than the aesthetic properties. Human's perception of aesthetics and emotion have something in common under some conditions. However, emotional impulses are completely subjective. Emotional responses shared between different people are hard to analyze.

The main task of computational aesthetics is to build a model to simulate human perception of aesthetics. Including aesthetic measures of vision, literature, music, cooking, etc. [33] Concept of computational aesthetics has been born as early as the 1930s, when the American mathematician George D. Burkhoff gave his own calculation method of aesthetics in his book, that is, the aesthetic quality should be the ratio of order to complexity [3].

The definition of order and complexity of image is not clear in the traditional computer vision field. The aesthetic quality formula for images was first proposed by Machado et al. [27] in 1998:

$$Aesthetic = IC^a / PC^b \quad (1)$$

IC (Image Complexity) represents the complexity of an image, and PC (Process Complexity) represents the complexity of brain in analyzing an image. a and b respectively represent the weight of the two complexity degrees. However, IC and PC are still very difficult to measure, which makes this calculation too abstract.

Lakhal et al. [21] believed that the complexity on aesthetics is different from the information entropy used in the field of communication. They defined two kinds of complexity, namely, the entropy complexity representing the image information amount and a non-monotone increasing structural complexity.

Joshi et al. [17] discussed aesthetic and emotion in images from philosophy, photography, painting and other fields. They believed that computational framework based on machine learning is an essential method of computational aesthetics and analyzed image aesthetic data on the Internet.

With the development of computational aesthetics and definition of image aesthetic quality, calculation of image aesthetics has gradually changed from rule-driven method to data-driven method. Researchers are generally committed to designing an IAA model recognized by users with different cultural backgrounds and knowledge levels. These models have a wide range of applications although may be unconvincing for complex and abstract works of art.

3 Manual Designed Features

Researchers in the early stages of IAA adopted the method of manual design of aesthetic features. The development process of manual design features can be roughly summarized as the transformation from low-level image features to high-level aesthetic features. Low-level features of an image can reflect the basic attributes and technical quality, while the high-level aesthetic features are often based on photography rules and have a stronger ability of aesthetic expression.

In 2004, Tong et al. [40] took the lead in selecting the features of texture, color, shape and other concatenation into a 846-dimensional feature vector to classify the aesthetic quality of image. Then some researchers began to study the impact of global features on image aesthetic quality. Ke et al. [19] designed the spatial distribution of image edges, color distribution, hue, degree of blur, contrast and brightness. Aydin [2] uses five global features of sharpness, depth, clarity, hue and saturation. Since global features cannot fully represent the spatial

structure and regional aesthetic properties in the image, researchers [7, 25, 43] tried to combine local features, features between regions and global features. Similarly, general image descriptors such as BOV, FV, and SIFT, can also be applied to IAA [30, 44], but these features have limited performance due to the lack of attention to image aesthetics.

Images of different content usually have different aesthetic features. Aesthetic attributes can also be subdivided using evaluation criteria in the field of photography. Therefore, more complex and advanced aesthetic features are often targeted.

Luo et al. [25] divided images into seven categories: landscape, plant, animals, night, human, static, and architecture and designed different features for different types of images. Dhar et al. [8] designed 26 features to reflect the aesthetics and interest of the image through the classification of photos. Nishiyama et al. [32] uses the Moon-Spencer model to analyze color harmony in images. Jin et al. [16] summarized four types of lighting commonly used in portrait photography: Rembrandt, Paramount, Loop and Split, and used the stepwise feature pursuit algorithm to learn the contrast characteristics of the local lighting of photographic works.

Low-level features lack the ability to express image aesthetics. Complexity and abstraction of photography rules and the various types of pictures make the design of high-level aesthetic features a very complicated work. Therefore, IAA based on deep learning has become the mainstream research method at this stage. But we still believe in this time, manual design features have many practical applications in industry and can perform better in some area that have explicit aesthetics.

4 Generic Image Aesthetic Assessment (GIAA) Based on Deep Learning

GIAA model based on deep learning use neural network as a feature extractor or perform end-to-end IAA, which purpose is to model recognized aesthetics. This section introduces the research ideas according to the five types of IAA. We intersperses the analysis of five assessment types and datasets proposed in research during introduction.

4.1 Aesthetic Classify and Aesthetic Regression

Some studies use CNNs as feature extractors. In the early stage, Dong et al. [9] used image pyramid model to convert the original image into image blocks of different scales and same size and input them into AlexNet to extract features. In 2020, Sheng et al. [37] pointed out that the manipulation of image usually causes negative aesthetic effects. For this reason, they designed a novel self-supervised learning method to identify attributes like blur, camera shake and so on in the image. Then the features extracted by the recognition task were input into the linear classifier for aesthetic classification.

In 2013, the AVA dataset [31] containing 250,000 images was constructed and open sourced by Murray et al., which promoted end-to-end IAA as the mainstream algorithm.

In order to extract aesthetic features of different scales, many researchers have adopted the method of multi-column or multi-patches CNN models. In 2014, Lu et al. [23] designed the RAPID model, using two-column CNN to extract the features of global image and local image obtained by random cropping, and then splicing the global features and local features to classify aesthetic quality. Subsequently, Lu et al. [24] designed a multi-column neural network named DMA-Net with shared parameters. In order to extract detailed information in the image, multiple image patches of same size were randomly cropped from the original image for training. They also designed two feature fusion layers based on statistics and ranking to aggregate the output from multi-column network. Ma et al. [26] used a saliency detection method [45] to extract the salient areas of the image. The salient image blocks and the overall image were taken as vertices, and the spatial information between vertices were used as edges to construct an undirected attribute graph. The undirected attribute graph is converted into an one-dimensional vector and input to the network to extract composition information. In order to preserve the original size of image, Mai et al. [28] designed MNA-CNN. They designed an adaptive spatial pooling layer (ASP) that can output fixed-dimensional features. The model is a multi-column network using an ASP layer to extract aesthetic features of different scales. In addition, they trained a scene classification network to perform feature aggregation.

Kao et al. [18] believed that image semantic recognition is the key to assessing the aesthetics. The proposed model uses semantic recognition tasks to assist aesthetic quality evaluation under the framework of multi-task learning. The experiment found that some tags such as “Seascapes” are positively related to aesthetics, and some tags such as “Candid” are negatively related to aesthetics.

In 2018, Sheng et al. [38] applied attention mechanism to IAA. They randomly cropped out several image blocks, and then designed three attention mechanisms (average, minimum, and adaptive) to adjust the weight of image blocks during training. The experimental results show that the attention mechanism plays a positive role in the classification of image aesthetic.

4.2 Aesthetic Distribution

In the aesthetic distribution method, probability distribution is used to describe the possibility that an image is considered to belong to a certain aesthetic level, which reflects the subjectivity of IAA. Besides, distribution can be easily converted to classification and aesthetic score, which has been favored by many researchers.

Jin et al. [13] used kurtosis of image score histogram to measure the reliability of photos in AVA dataset, combined with the Jensen-Shannon divergence based on cumulative distribution as the loss function of aesthetic distribution task.

Hou et al. [12] found that EMD loss performs well on dataset that has inherent sorting among different categories. Subsequently, Talebi et al. [39] removed

the last layer of MobileNet, Inception-v2 and VGG16 as the baseline model, after that they added a fully connected layer and a Softmax layer to output aesthetic distribution. They used EMD loss as loss function and made significant progress compared with other methods on AVA dataset. Cui et al. [6] combined the semantic information of image in the aesthetic distribution network, and chose a FCN to preserve the original size of input images.

4.3 IAA with Attributes

IAA with attributes means that the assessment results are generated for different aesthetic attributes. Combined with the other four decision making methods, it has a better ability to express aesthetic quality. IAA with attributes based on deep learning was first proposed in 2016 by Kong et al. [20]. They built AADB dataset containing about 10,000 pictures and open sourced. Photos in AADB have eleven aesthetic attributes (Rule of thirds, color harmony, interesting content, etc.) evaluated. Kong's model is trained using pictures in the AADB dataset and can output the quality of each attribute of the picture.

Malu et al. [29] adopted eight attributes in AADB. They used a multi-task neural network to extract features for these attributes, and used a visualization technology based on gradient back propagation to show the corresponding area of each attribute in the image. Jin et al. [15] uses a multi-task regression learning strategy to extract the general features and features of six attributes. The assessment result were displayed as an intuitive radar map.

4.4 Aesthetic Description

The research of aesthetic description is inspired by the task of image caption. Image caption is to generate a descriptive text for an image, while the task of aesthetic description is to generate aesthetic comment.

Aesthetic description is a more subjective assessment method, and often contains descriptions of one or more aesthetic attributes. Therefore, aesthetic descriptions are generally considered to be the highest level of IAA at the moment. It combines the research of computer vision and natural language processing. Limited by the scale and effectiveness of existing datasets, there are relatively few studies in this area.

The aesthetic description research started in 2017. Chang et al. [4] built PCCD containing image comments and aesthetic attributes. They proposed a novel model to generate aesthetic comments. Regarding the evaluation criteria of the generated aesthetic reviews, they pointed out that unlike the image captions datasets, the comments in PCCD have fewer synonymous sentences. Therefore, they believe that the SPICE [1] standard is more suitable for aesthetic description. In addition, Chang et al. also proposed a diversity index to measure the similarity between aesthetic reviews. Regrettably, PCCD has a small amount of data and has stopped updating.

Subsequently, Wang et al. [42] built a dataset called AVA-Reviews containing 52118 photos and 312708 reviews. Jin et al. [14] were inspired by PCCD dataset

and crawled 330,000 pictures and comments of these pictures. After screening the content of the comments, 150000 pictures that have comments with one to five aesthetic attributes were retained. These photos helped them train a CNN-LSTM model combined with attention mechanism, which can generate five comments for different aesthetic attributes.

5 Personalized Image Aesthetic Assessment (PIAA)

PIAA is a challenging job and has great application prospects. GIAA can only reflect the aesthetics of a relatively small number of people in some controversial pictures. Unlike GIAA, PIAA is dedicated to learning aesthetic preferences that belong to specific users.

Constructing a personalized recommendation model for users is a problem that has been researched in the recommendation field. Because it is difficult to obtain effective and large amounts of user feedback in the field of IAA, traditional recommendation algorithms (collaborative filtering etc.) are not effective in PIAA tasks.

In 2017, Ren et al. [36] raised the issue of Personalized Image Aesthetics Assessment (PIAA). In order to link IAA with user's identity, he downloaded 40,000 images from the photography website Flickr and asked 210 workers to mark these images with 1 to 5 points on the online crowdsourcing survey platform, and finally built FLICKR-AES. They also built a dataset called REAL-CUR, consisting of 14 photo albums of real users with aesthetic ratings. Ren et al. proposed a PAM method that uses aesthetic bias of a single user to adjust the GIAA model to make it fit the user's aesthetic preferences, and an active PIAA method (Active-PAM) in order to reduce dependence on personalized data.

Li et al. [22] used personality characteristics to assist in the completion of GIAA and PIAA learning under the framework of multi-task learning. They used the PsychoFlickr dataset proposed in the research [5] to learn personality characteristics. The personality are The Big-Five (BF): Openness (O), Conscientiousness (C), Extroversion (E), Agreeableness (A), and Neuroticism (N). Trained GIAA model is fine-tuned using the aesthetic data of a single user in FLICKR-AES to generate PIAA model.

Zhu et al. [46] and Wang et al. [41] proposed methods based on meta-learning. The idea of meta-learning is considered "learning how to learn", and the purpose is to train a model that can quickly fit new tasks. In the training process of meta-learning, each user's aesthetics is treated as a single task and the aesthetic data is divided into a support set and a query set. Then the trained model is fine-tuned and tested on the test task. Experiments proved that meta-learning strategy performs well on PIAA tasks.

6 Aesthetic Datasets

IAA based on deep learning is a data-driven model. As a result of subjective assessment, aesthetic data is often accompanied by words even emoticons in

daily life. Aesthetic data collection is much more complicated than other tasks such as image classification and saliency detection.

Looking back at the entire IAA development process, novel methods often accompanied with new datasets. A large-scale open source dataset can greatly promoted the development of IAA.

The above has briefly introduced some datasets and their built methods. This section intends to make a summary of some key information in the dataset. Table 1 is prepared for scale, assessment results, whether it contains aesthetic attributes, the identity of users, and whether it contains semantic information.

Table 1. Comparison of the properties of representative image aesthetics datasets

	CUHK-PQ [25]	AVA [31]	AADB [20]	FLICKR-AES [36]	PCCD [4]
Number of images	17690	255530	10000	40000	4235
Assessment type	Category	Distribution	Distribution	Distribution	Score
With attributes	No	No	Yes	No	Yes
Rater's ID	No	No	Yes	Yes	Yes
Semantic tags	No	Yes	No	No	No

7 Conclusion and Future Works

How does the brain perceive beauty? What are the characteristics of aesthetics? So far, IAA still has a lot of room for development. Research on aesthetic description and IAA with attributes are relatively small and not mature enough. In the near future, more advanced evaluation methods may be applied. There are also many problems need to be solved in PIAA.

This report reviews the development process of IAA roughly in chronological order, but those studies that are not yet popular are not worthless. IAA is a complex and huge subject. Different fields have different emphasis on aesthetics. For example, researches on composition and lighting can play a role in real-time shooting suggestions, and researches on color harmony can be used in the field of fashion etc.

Aesthetics datasets are complex and diverse. Many researchers choose to crawl photos from photography websites. Manipulation, technical quality and aesthetic value of photos are issues that researchers have to consider. How to value the multi-modality information on photography websites is also one of the works being researching.

This article reviews representative IAA approaches, We hope this report can help researchers who are engaged in or intend to engage in the work of IAA!

Acknowledgements. This work is supported by Training of Outstanding Talents in Beijing in 2017, Research on Quality Evaluation Method of UHD Video based on HEVC (GJ181901).

References

1. Anderson, P., Fernando, B., Johnson, M., Gould, S.: SPICE: semantic propositional image caption evaluation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) ECCV 2016. LNCS, vol. 9909, pp. 382–398. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46454-1_24
2. Aydın, T.O., Smolic, A., Gross, M.: Automated aesthetic analysis of photographic images. *IEEE Trans. Vis. Comput. Graph.* **21**(1), 31–42 (2014)
3. Birkhoff, G.D.: *Aesthetic Measure*. Harvard University Press (2013)
4. Chang, K.Y., Lu, K.H., Chen, C.S.: Aesthetic critiques generation for photos. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 3514–3523 (2017)
5. Cristani, M., Vinciarelli, A., Segalin, C., Perina, A.: Unveiling the multimedia unconscious: Implicit cognitive processes and multimedia content analysis. In: Proceedings of the 21st ACM International Conference on Multimedia, pp. 213–222 (2013)
6. Cui, C., Liu, H., Lian, T., Nie, L., Zhu, L., Yin, Y.: Distribution-oriented aesthetics assessment with semantic-aware hybrid network. *IEEE Trans. Multimedia* **21**(5), 1209–1220 (2018)
7. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Studying aesthetics in photographic images using a computational approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3953, pp. 288–301. Springer, Heidelberg (2006). https://doi.org/10.1007/11744078_23
8. Dhar, S., Ordonez, V., Berg, T.L.: High level describable attributes for predicting aesthetics and interestingness. In: CVPR 2011, pp. 1657–1664. IEEE (2011)
9. Dong, Z., Shen, X., Li, H., Tian, X.: Photo quality assessment with DCNN that understands image well. In: He, X., Luo, S., Tao, D., Xu, C., Yang, J., Hasan, M.A. (eds.) MMM 2015. LNCS, vol. 8936, pp. 524–535. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-14442-9_57
10. Elgammal, A., Liu, B., Elhoseiny, M., Mazzone, M.: CAN: Creative adversarial networks, generating “art” by learning about styles and deviating from style norms. arXiv preprint [arXiv:1706.07068](https://arxiv.org/abs/1706.07068) (2017)
11. Ghadiyaram, D., Bovik, A.C.: Massive online crowdsourced study of subjective and objective picture quality. *IEEE Trans. Image Process.* **25**(1), 372–387 (2015)
12. Hou, L., Yu, C.P., Samaras, D.: Squared earth mover’s distance-based loss for training deep neural networks. arXiv preprint [arXiv:1611.05916](https://arxiv.org/abs/1611.05916) (2016)
13. Jin, X., et al.: Predicting aesthetic score distribution through cumulative Jensen-Shannon divergence. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32 (2018)
14. Jin, X., et al.: Aesthetic attributes assessment of images. In: Proceedings of the 27th ACM International Conference on Multimedia, pp. 311–319 (2019)
15. Jin, X., et al.: Predicting aesthetic radar map using a hierarchical multi-task network. In: Lai, J.H., et al. (eds.) PRCV 2018. LNCS, vol. 11257, pp. 41–50. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-03335-4_4
16. Jin, X., Zhao, M., Chen, X., Zhao, Q., Zhu, S.-C.: Learning artistic lighting template from portrait photographs. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010. LNCS, vol. 6314, pp. 101–114. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15561-1_8
17. Joshi, D., et al.: Aesthetics and emotions in images. *IEEE Sig. Process. Mag.* **28**(5), 94–115 (2011)

18. Kao, Y., He, R., Huang, K.: Deep aesthetic quality assessment with semantic information. *IEEE Trans. Image Process.* **26**(3), 1482–1495 (2017)
19. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2006, vol. 1, pp. 419–426. IEEE (2006)
20. Kong, S., Shen, X., Lin, Z., Mech, R., Fowlkes, C.: Photo aesthetics ranking network with attributes and content adaptation. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9905, pp. 662–679. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46448-0_40
21. Lakkhal, S., Darmon, A., Bouchaud, J.P., Benzaquen, M.: Beauty and structural complexity. *Phys. Rev. Res.* **2**(2), 022058 (2020)
22. Li, L., Zhu, H., Zhao, S., Ding, G., Lin, W.: Personality-assisted multi-task learning for generic and personalized image aesthetics assessment. *IEEE Trans. Image Process.* **29**, 3898–3910 (2020)
23. Lu, X., Lin, Z., Jin, H., Yang, J., Wang, J.Z.: RAPID: rating pictorial aesthetics using deep learning. In: *Proceedings of the 22nd ACM International Conference on Multimedia*, pp. 457–466 (2014)
24. Lu, X., Lin, Z., Shen, X., Mech, R., Wang, J.Z.: Deep multi-patch aggregation network for image style, aesthetics, and quality estimation. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 990–998 (2015)
25. Luo, W., Wang, X., Tang, X.: Content-based photo quality assessment. In: *2011 International Conference on Computer Vision*, pp. 2206–2213. IEEE (2011)
26. Ma, S., Liu, J., Wen Chen, C.: A-Lamp: adaptive layout-aware multi-patch deep convolutional neural network for photo aesthetic assessment. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4535–4544 (2017)
27. Machado, P., Cardoso, A.: Computing aesthetics. In: de Oliveira, F.M. (ed.) *SBIA 1998*. LNCS (LNAI), vol. 1515, pp. 219–228. Springer, Heidelberg (1998). https://doi.org/10.1007/10692710_23
28. Mai, L., Jin, H., Liu, F.: Composition-preserving deep photo aesthetics assessment. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 497–506 (2016)
29. Malu, G., Bapi, R.S., Indurkha, B.: Learning photography aesthetics with deep CNNs. arXiv preprint [arXiv:1707.03981](https://arxiv.org/abs/1707.03981) (2017)
30. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: *2011 International Conference on Computer Vision*, pp. 1784–1791. IEEE (2011)
31. Murray, N., Marchesotti, L., Perronnin, F.: AVA: a large-scale database for aesthetic visual analysis. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2408–2415. IEEE (2012)
32. Nishiyama, M., Okabe, T., Sato, I., Sato, Y.: Aesthetic quality classification of photographs based on color harmony. In: *CVPR 2011*, pp. 33–40. IEEE (2011)
33. Perc, M.: Beauty in artistic expressions through the eyes of networks and physics. *J. R. Soc. Interface* **17**(164), 20190686 (2020)
34. Polat, A.T., Akay, A.: Relationships between the visual preferences of urban recreation area users and various landscape design elements. *Urban Forest. Urban Greening* **14**(3), 573–582 (2015)
35. Ponomarenko, N., et al.: Color image database TID2013: peculiarities and preliminary results. In: *European Workshop on Visual Information Processing (EUVIP)*, pp. 106–111. IEEE (2013)

36. Ren, J., Shen, X., Lin, Z., Mech, R., Foran, D.J.: Personalized image aesthetics. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 638–647 (2017)
37. Sheng, K., et al.: Revisiting image aesthetic assessment via self-supervised feature learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 5709–5716 (2020)
38. Sheng, K., Dong, W., Ma, C., Mei, X., Huang, F., Hu, B.G.: Attention-based multi-patch aggregation for image aesthetic assessment. In: Proceedings of the 26th ACM International Conference on Multimedia, pp. 879–886 (2018)
39. Talebi, H., Milanfar, P.: NIMA: neural image assessment. *IEEE Trans. Image Process.* **27**, 3998–4011 (2017)
40. Tong, H., Li, M., Zhang, H.-J., He, J., Zhang, C.: Classification of digital photos taken by photographers or home users. In: Aizawa, K., Nakamura, Y., Satoh, S. (eds.) *PCM 2004*. LNCS, vol. 3331, pp. 198–205. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30541-5_25
41. Wang, W., Su, J., Li, L., Xu, X., Luo, J.: Meta-learning perspective for personalized image aesthetics assessment. In: 2019 IEEE International Conference on Image Processing (ICIP), pp. 1875–1879. IEEE (2019)
42. Wang, W., Yang, S., Zhang, W., Zhang, J.: Neural aesthetic image reviewer. *IET Comput. Vis.* **13**(8), 749–758 (2019)
43. Wong, L.K., Low, K.L.: Saliency-enhanced image aesthetics class prediction. In: 2009 16th IEEE International Conference on Image Processing (ICIP), pp. 997–1000. IEEE (2009)
44. Yeh, M.C., Cheng, Y.C.: Relative features for photo quality assessment. In: 2012 19th IEEE International Conference on Image Processing, pp. 2861–2864. IEEE (2012)
45. Zhang, J., Sclaroff, S., Lin, Z., Shen, X., Price, B., Mech, R.: Unconstrained salient object detection via proposal subset optimization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5733–5742 (2016)
46. Zhu, H., Li, L., Wu, J., Zhao, S., Ding, G., Shi, G.: Personalized image aesthetics assessment via meta-learning with bilevel gradient optimization. *IEEE Trans. Cybern.* (2020)