SCIENCE CHINA
Information Sciences

• NEWS & VIEWS •

February 2019, Vol. 62 027101:1–027101:2
https://doi.org/10.1007/s11432-018-9692-2

Two open-source projects for image aesthetic quality assessment

Le WU[1,2,3], Xin JIN[1,2,3*], Geng ZHAO[1] & Xinghui ZHOU[1]

[1]*Beijing Electronic Science and Technology Institute, Beijing* 100070, *China;*
[2]*State Key Laboratory of Cryptology, P.O. Box* 5159, *Beijing* 100878, *China;*
[3]*China Electronics Technology Group Corporation Big Data Research Institute Co., Ltd., Guiyang* 550018, *China*

Received 8 September 2018/Accepted 14 November 2018/Published online 26 December 2018

Citation Wu L, Jin X, Zhao G, et al. Two open-source projects for image aesthetic quality assessment. Sci China Inf Sci, 2019, 62(2): 027101, https://doi.org/10.1007/s11432-018-9692-2

Here we introduce two open-source projects for image aesthetic quality assessment. The first one is ILGnet, an open-source project for the aesthetic evaluation of images based on the convolution neural network. The second is CJS-CNN, an open-source project for predicting the aesthetic score distribution of human ratings.

ILGnet. An image aesthetic classification method based on a deep convolution neural network can classify the results of the image aesthetic quality as "good" or "bad". This method combines local and global features and designs a new depth convolution neural network, ILGnet, which learns the image aesthetic classification model using approximately 230000 images from the largescale database for aesthetic visual analysis (AVA). The AVA dataset contains 255530 valid images, each of which comes from the website[1)], which is a well-known photographic community abroad. Each image is scored by members, who are human artists registered on multiple websites. Each image is scored by 78–539 people, with an average of 210 participants. The dataset is a recognized benchmark in the field of image aesthetic evaluation. The quality of the annotated data is high, and it can support the study of the aesthetic classifications, aesthetic scores, and aesthetic distributions. The ILGnet can automatically distinguish two types of images: high aesthetic quality and low aesthetic quality. It uses a new structure of a convolution neural network for image classification and fuses the information of different receptive fields to complete the aesthetic feature extraction process. Compared with previous methods, the classification accuracy is greatly improved. The accuracy we achieve on the AVA dataset is 81.68%, and the accuracy is up to 82.66% by using the Inception V4 module [1, 2].

CJS-CNN. The classification and scoring of the image aesthetic assessment usually use a scalar to express the aesthetic quality of an image and largely ignore the diversity, subjectivity, and individuation of human aesthetics in a certain consensus. In general, nearly all image recognition tasks have standard answers but few have aesthetic images. This is the biggest difference between aesthetic evaluations and general image recognition. The probability distribution of an image aesthetic score can describe the aesthetic subjectivity to a certain extent; for example, the variance can describe the human consensus for an image to a certain extent and the kurtosis can describe the popularity of an image to a certain extent. In this study, a convolution neural network (CNN) based on the cumulative distribution with Jensen-Shannon divergence (CJS-CNN) is proposed to predict the fractional distribution of the image aesthetic quality evaluation, which is different from previous single scalar evaluations. This method can give the score distribution of an image aesthetic qual-

* Corresponding author (email: jinxin@besti.edu.cn)
1) https://www.dpchallenge.com/

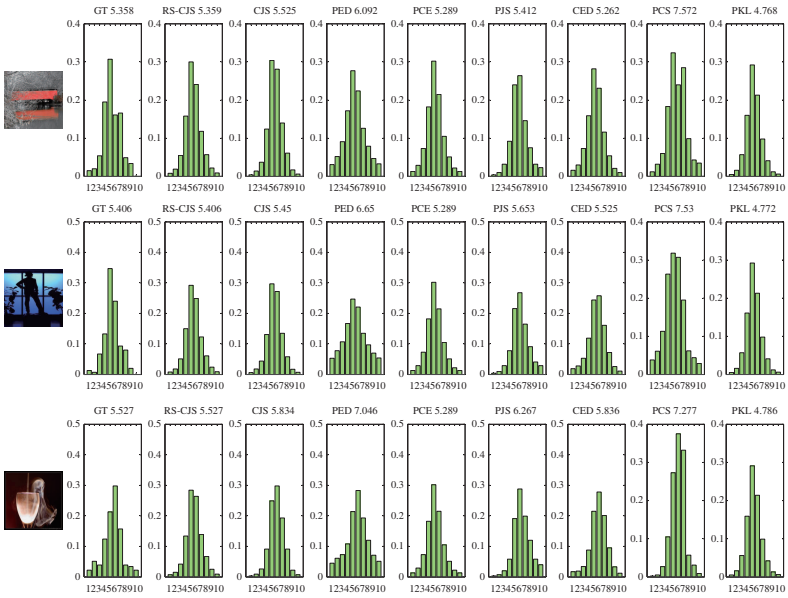

Figure 1 (Color online) The predicted aesthetic score histogram according to CJS, RS-CJS, and other loss functions. The leftmost column shows each test image, and the number at the top of each graph is the average value of the scores calculated by the histogram. The second column represents the real image fraction distribution. The third and fourth columns are the results based on the CJS method and the RS-CJS method, respectively. The other columns represent the results of other loss functions.

ity evaluation, and can better describe the diversity of human aesthetic perceptions. Further, a reliability-sensitive learning algorithm based on kurtosis statistics is proposed. This algorithm only requires the normalized evaluation score distribution data to preform reliability-sensitive learning, while the previous reliability-sensitive learning algorithm based on the number of evaluators requires the complete evaluation score distribution data prior to normalization. As shown in Figure 1, the result of CJS-CNN method is closer to the true distribution than the other methods. In the experiments, our proposed CJS-CNN method outperform other methods on various mean divergence indicators [3, 4].

Usage. ILGnet is an open-source project for the aesthetic evaluation of images based on the deep learning framework, Caffe. If you need to train the model, you should download the AVA dataset and treat the dataset according to the following method. First, you can choose a score of 5 as a criterion to divide the dataset. Second, download the pre-trained GoogleNet model on ImageNet. During training, you should fix the parameters of all the layers before the first full-connected layer and fine-tune the last layers using the training set of the AVA dataset. We suggest training and testing the model using the Caffe framework. The accuracy in the AVA dataset is 81.68% and can be up to 82.66% by using the Inception V4 module. The CJS-CNN project is very easy to use because it can be implemented on any deep learning network. If you want to train in the AVA dataset, you can copy the include file and source code of the CJS layer to the Caffe source and include the path and rebuild the Caffe framework. Next, you need to process the AVA dataset into distributed forms and make the training and test labels. Finally, you can select any neural network and change the loss function to the CJS loss function for training and testing.

Access methods. **ILGnet** can be downloaded from https://github.com/BestiVictory/ILGnet. **CJS-CNN** can be downloaded from https://github.com/BestiVictory/CJS-CNN.

References

1 Jin X, Chi J Y, Peng S W, et al. Deep image aesthetics classification using inception modules and fine-tuning connected layer. In: Proceedings of the 8th International Conference on Wireless Communications and Signal Processing (WCSP), Yangzhou, 2016

2 Jin X, Wu L, Li X D, et al. ILGNet: inception modules with connected local and global features for efficient image aesthetic quality classification using domain adaptation. IET Comput Vision, 2018. doi: 10.1049/iet-cvi.2018.5249

3 Jin X, Wu L, Li X D, et al. Predicting aesthetic score distribution through cumulative Jensen-Shannon divergence. In: Proceedings of AAAI Conference on Artificial Intelligence, New Orleans, 2018

4 Wang J L, Lu Y H, Liu J B, et al. A robust three-stage approach to large-scale urban scene recognition. Sci China Inf Sci, 2017, 60: 103101