AESTHETICS

Beauty is in the eye of the machine

Ansel Adams said, "There are no rules for good photographs, there are only good photographs." Is it possible to predict our fickle and subjective appraisal of 'aesthetically pleasing' visual art? Iigaya et al. used an artificial intelligence approach to show how human aesthetic preference can be partially explained as an integration of hierarchical constituent image features.

Mengmi Zhang and Gabriel Kreiman

Artificial intelligence (AI) has made rapid strides in a wide range of visual tasks, including recognition of objects and faces, automatic diagnosis of clinical images, and answering questions about images. More recently, AI has also started penetrating the arts. For example, in October 2018, the first piece of AI-generated art came to auction, with an initial estimate of US$ 10,000, and strikingly garnered a final bid of US$ 432,500 (Fig. 1). The portrait depicts a portly gentleman with a seemingly fuzzy facial expression, dressed in a black frockcoat with a white collar. Appreciating and creating a piece of art requires a general understanding of aesthetics. What are the nuances, structures, and semantics embedded in a painting that can provide us with an aesthetically pleasing sense?

Appraisal of aesthetic value by humans is fickle and subjective. However, empirical investigations have long shown that there exists some level of universality in art appreciation across cultures and history[1]. Such correlations lead to the question of whether there are any rules governing our preferences for visual arts. Previous studies have hinted at the use of feature integration frameworks to predict aesthetic value preferences[2]. However, previous work has focused on the effect of either one specific visual feature or a pool of complex features, without a clear delineation of how these features weigh towards generating value judgements. The high visual complexity of each piece of art and the large heterogeneity among different pieces of art make the selection of relevant features for value judgements challenging.

Now, a new study by Iigaya et al. in *Nature Human Behaviour*[3] represents a critical step forward in elucidating the features underlying aesthetic value preference. Iigaya and colleagues asked both in-lab and online participants to report how much they liked a piece of artwork on a four-point scale. In a first attempt, the authors used a set of 13 hand-crafted visual characteristics, such as hue, contrast, and presence of people, to successfully assess aesthetic values. Next, to circumvent the need for human intervention in the selection of features, Iigaya et al. used a deep convolutional neural network (DCNN)[4]. DCNNs are at the heart of the recent revolution in AI. They consist of neuron-like units organized into multiple layers that sequentially process an image to extract an increasingly richer and more robust set of visual features[5]. Each unit receives inputs from a myriad of other units, and the connection strengths are governed by weights that can be learned from examples via training.

Here the authors utilized a network that had been pre-trained on an object-recognition task involving labels of about 1,000,000 images from a dataset known as ImageNet[6]. The resulting DCNN model constitutes an initial approximation to the cascade of computations that take place along the ventral visual cortex. The authors then fine-tuned the DCNN model by adjusting the weights of only the last fully connected layers, using only a subset of the images ranked by human participants. This approach allowed the model to learn what is aesthetically pleasing without imposing any human wisdom or biases about specific visual features. Thus, the DCNN model learned from examples to automatically distinguish images with higher versus lower value. The experimental results by Iigaya et al. showed, surprisingly, that the DCNN model not only succeeded in predicting subjective values, but was also able to implicitly capture the 13 hand-crafted features chosen in the preliminary analyses. As observed in other studies, higher layers in the network, which represent more complex visual features, yielded a better predictability of subjective value.

In summary, this DCNN model represents a breakthrough in our understanding of how humans might make aesthetic value judgements. The results suggest a host of intriguing avenues for further studies. Sceptics may argue that computational models lack sentience, that machines do not have their own preferences, and that ultimately computers cannot understand human values. After all, our aesthetic values might represent a complex combination of evolutionary learning, cultural influences, and individuality. As in many other cases, the answer to the question of whether aesthetic value preferences are dictated by nature or nurture probably involves a mixture of both. Iigaya et al. show that a model based purely on visual features can capture aspects of universal aesthetic judgments. Yet, the model had to be trained with human-provided examples in a supervised fashion.

Maybe one day it will be possible to develop a model that is subject to the same types of visual experiences that humans go through in a largely unsupervised manner. These models may 'grow up' playing with

Fig. 1 | "Edmond de Belamy, from La Famille de Belamy". The first piece of artwork created by artificial intelligence garnered US$ 432,500 at Christie's auction in October, 2018.

toys and watching Disney movies. Perhaps then these models will spontaneously reveal aesthetic preferences without the need to learn them from examples. On the other hand, it is likely that human aesthetic judgements are also learned through examples provided by parents, teachers, and social media.

Finally, while this study was limited to visual arts, the rankings and modelling approaches could well be extended to the study of other forms of art, including music, culinary arts, or literature; or even to other forms of value judgments beyond art, including ethics, trust, or politics. Looking ahead, AI systems that can assess and create value will advance many applications, such as, exhibition planning in art galleries, recommendation systems based on individualized art preferences, and evaluation of artwork values in financial markets. Imagine the near future when you visit the Louvre: an AI system may take you on an individualized tour of the best human and computer art targeted to your own taste. ❐

Mengmi Zhang[1,2] and Gabriel Kreiman[1,2]✉
[1]Children's Hospital, Harvard Medical School, Boston, MA, USA. [2]Center for Brains, Minds and Machines, Cambridge, MA, USA.
✉*e-mail: gabriel.kreiman@tch.harvard.edu*

Published online: 20 May 2021
https://doi.org/10.1038/s41562-021-01125-5

References
1. Kant, I. & Bernard, J. H. *Kant's Critique of Judgement*. 2d ed. (Macmillan, 1931).
2. Kong, S., Shen, X., Lin, Z., Mech, R. & Fowlkes, C. Photo aesthetics ranking network with attributes and content adaptation. in *European Conference on Computer Vision* (eds Leibe B., Matas J., Sebe N. & Welling M.) 662–679 (2016).
3. Iigaya, K., Yi, S., Wahle, I., Tanwisuth, K. & O'Doherty, J. Aesthetic preference for art can be predicted from a mixture of low- and high-level visual features *Nat. Hum. Behav*. https://doi.org/10.1038/s41562-021-01124-6 (2021).
4. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. Preprint at *arXiv* https://arxiv.org/abs/1409.1556 (2015).
5. Kreiman, G. *Biological and Computer Vision* (Cambridge Univ. Press, 2021).
6. Russakovsky, O., Li, L.-J. & Li, F.-F. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2121–2131 (Computer Vision Foundation, 2014).

Competing interests
The authors declare that they have no competing interests.