

Review

Will the Machine Like Your Image? Automatic Assessment of Beauty in Images with Machine Learning Techniques [†]

Matteo Bodini 

Department of Computer Science, Università degli Studi di Milano, Via Celoria 18, 20133 Milano, Italy; matteo.bodini@unimi.it

[†] This paper is an extended version of our paper published in “Bodini, M. Automatic Assessment of the Aesthetic Value of an Image with Machine Learning Techniques. In the Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering 2019 (ISMAC-CVB), Elayampalayam, India, 13–14 March 2019; Springer International Publishing: Cham, Switzerland, 2019; in press”.

Received: 7 June 2019; Accepted: 23 June 2019; Published: 28 June 2019



Abstract: Although the concept of image quality has been a subject of study for the image processing community for more than forty years (where, with the term “quality”, we are referring to the accuracy with which an image processing system captures, processes, stores, compresses, transmits, and displays the signals that compose an image), notions related to aesthetics of photographs and images have only appeared for about ten years within the community. Studies devoted to aesthetics of images are multiplying today, taking advantage of the latest machine learning techniques and mostly due to the proliferation of huge communities and websites, specialized in digital photography sharing and archiving, such as Flickr, Imgur, DeviantArt, and Instagram. In this review, we examine the latest advances of computer methods that aim at computationally distinguishing high-quality from low-quality photos and images, relying on machine learning techniques. The paper is organized as follows: First, we introduce many approaches to aesthetics, studied in philosophy, neurobiology, experimental psychology, and sociology, to see what lighting they propose to researchers. Such points of view let us explain the weakness of the current consensus on the difficult aesthetics problem and the importance of the ongoing debates on it. Then, we analyze the work done in the community of pattern recognition and artificial intelligence on the task of automatic aesthetic assessment, and we both compare and critically examine the presented results. Finally, we describe many issues that have not been addressed, and starting from these, we outline some possible future directions.

Keywords: aesthetics; beauty; machine learning; deep learning; deep neural networks

1. Introduction

The problem of assessing image quality has been faced in the image processing community for at least forty years: it refers to the level of accuracy with which an imaging system captures, processes, stores, compresses, transmits, and displays the signals that compose an image [1]. On the contrary, notions linked to aesthetic assessment of photography and images appeared in the field of computer science only in the last ten years. Image aesthetic (or beauty) assessment aims at computationally distinguishing high-quality from low-quality photos, based on different features and methods.

The automatic aesthetic assessment of images is a new challenging task for the communities of computer vision and image processing that has several applications, for instance: photo management, image retrieval, photo enhancement, and many others [2–4]. These applications have received growing attention in the last decade, because of the evolution of ICT (Information and Communication

Technology), which has had important consequences on business and societal practices (storage on very large and distributed databases, archival and retrieval functions, automatic learning processes) and because of the flourishing of photographic exchange sites, such as Flickr, Imgur, DeviantArt, and Instagram. On these websites, the filter “beauty images” is more and more chosen as the discriminating criterion to select images in retrieval operations.

The theme of automatic aesthetic assessment of images is now rich in papers. In the last ten years, many communications have proliferated in the pattern recognition and Artificial Intelligence (AI) community, offering to provide an automatic assessment of the aesthetic value of an image. A variety of approaches has been proposed in the literature to try to solve this challenging problem [5–11]. These works use Deep Neural Networks (DNNs), and they follow the works published at the beginning of the 2000s, which tackled this problem with the help of the more classical pipeline of machine learning: detection of primitives, chosen by the user (for this reason, they are usually said to be hand-crafted) and classifiers of many types [3,12–15]. As usual, DNN techniques quickly outperformed the more traditional methods, as they did in many other areas of pattern recognition. In this article, we introduce many of these machine learning approaches (based on hand-crafted features and deep features). Then, we analyze and highlight the main contributions and the novelties of such approaches.

To get immediately an idea about the articles that we are going to review, we introduce one of the most remarkable works. We take into account NIMA: Neural Image Assessment [11]. The proposed deep Convolutional Neural Network (CNN), can be used to rank photos from an aesthetic point of view. In the article, the “Large-Scale Database for Aesthetic Visual Analysis” (AVA) dataset is taken into account [14]: it is a dataset containing over than 250,000 images, along with a rich variety of metadata, including a large number of aesthetics scores for each image, semantic labels for over 60 categories, as well as labels related to photographic style. The authors had each photo scored by an average of 200 people in response to photography contests. After training, the aesthetics ranking of these photos by NIMA closely matched the average scores given by human raters. A ranking example is given in Figure 1.

The advances that we are going to review have also been made possible because they take place in a scientific context that makes it possible to approach this problem in very different ways, in particular by neurobiological approaches, as well as experiments in social psychology. We are going to analyze what is the role of these studies and what influence they have on the evolution of AI techniques. In particular, AI has had huge benefit from 25 centuries of literature on aesthetics, beauty, and art, in philosophy, in sociology and in experimental psychology. Without a doubt, AI has also taken advantage of the ongoing studies in physiology and in neurobiology that intend to analyze how our brain works, some of them directed towards explaining our aesthetic judgment, now recognized under the name of “neuro-aesthetics”. These points are very important, and they will be addressed in detail, despite them almost never being considered in other surveys that are focused on the problem of aesthetic evaluation of images. For instance, see [4].

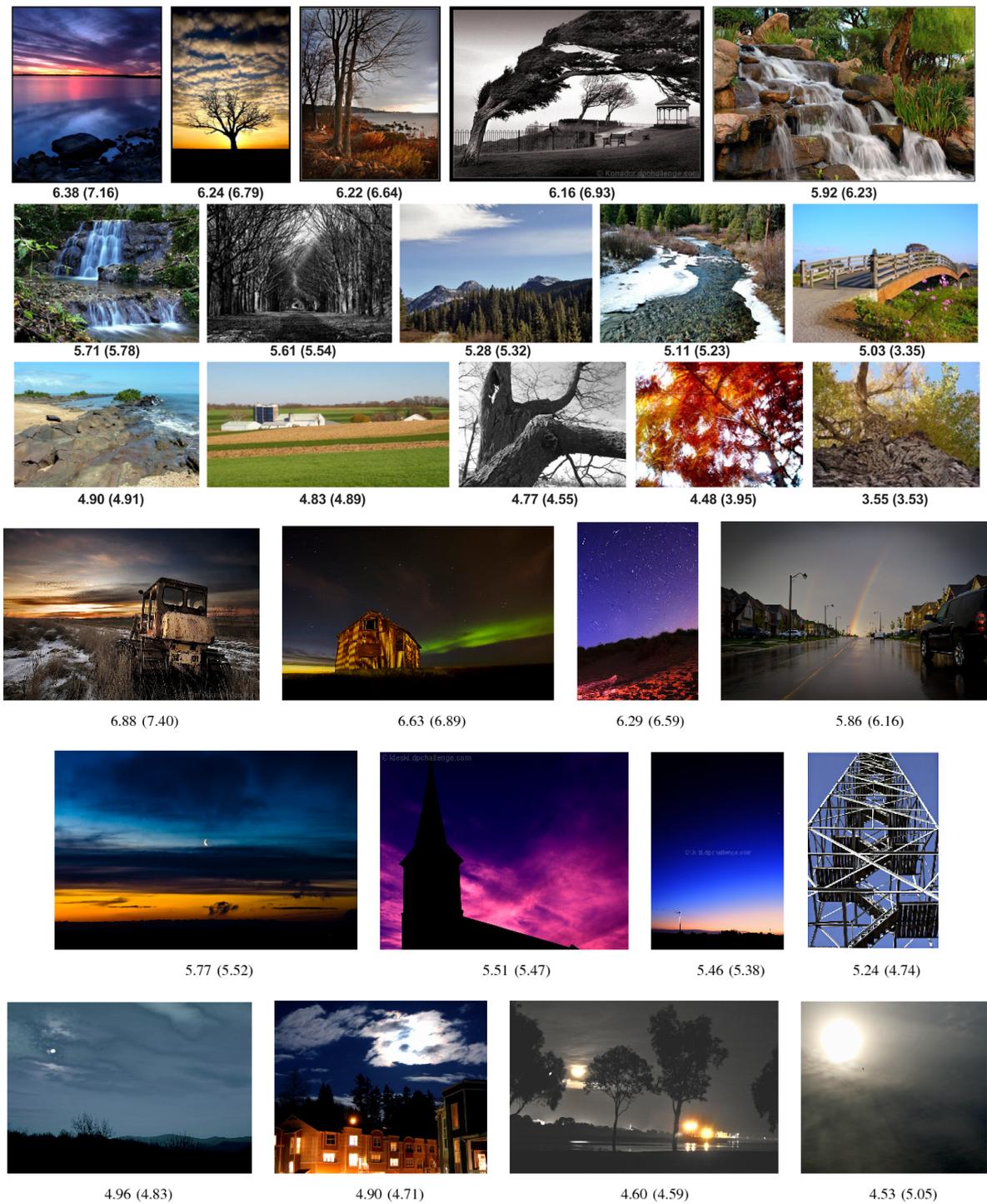


Figure 1. Ranking of some examples, labeled with the “landscape” tag, from the Aesthetic Visual Analysis (AVA) dataset using Neural Image Assessment (NIMA). Predicted NIMA (and ground truth in the brackets) scores are shown below each image.

2. How Can We Measure Beauty?

The first works that proposed a mathematical measure of beauty were due to Charles Henry [16], but it is the mathematician Birkhoff who proposed the first operational formulation [17,18]. This formulation, inspired by 25 centuries of philosophical literature on aesthetics in the visual arts in particular, was built on notions of order and simplicity at a time when these two terms had little meaning in mathematics. It has been enriched over the last century by the successful contributions

of the gestalt, the theory of information, mathematical morphology, and the theory of complexity to arrive at algorithmic and algebraic expressions [19–21], which had interesting results, but did not result in a very large agreement within the community.

Techniques based on machine learning, which appeared within this century, outperformed the previous works, as they opportunely exploited a completely new scientific context: the many images accessible on the Internet, the availability of numerous sources of expertise through specialized social networks or the general public, and finally, the advantages provided by powerful statistical techniques, which are able to build rules of classification and extend them to large unknown groups. Then, with the diffusion of CNN techniques, successively exploiting convolutional filtering and then fully-connected neural networks, we notice also in this field the “black box” approach, for which human expertise is only reduced to the constitution of indexed databases, needed in the learning phase.

This complete break in the paradigms behind the aesthetic approach should be analyzed and should be considered with respect to the consequences that can be expected from it. In the next subsections, we analyze the scientific background, provided by many fields, that contributed to this evolution.

2.1. Philosophical Approaches

It is hardly possible to treat in a few lines all the works that have gradually made aesthetics a clean and recognized part of philosophy. However, the masters have tried doing so during the centuries [22–25]. We can oppose schematically the “objectivist” school, resulting from the Greek philosophers of the classical period (which defends a universal idea of beauty, attached to the object or to the person whom it qualifies, an idea that is shared by all and in every place), with the subjective school, born from the philosophers of the Enlightenment (which relates beauty to individual experience and its experimental contingencies). The great currents of thought that, due to psychoanalysis, traversed philosophy during the last century carried this problem into more modern terms: is beauty so unanimously perceived because it selectively activates universal physiological sensations or is it the result of conjunctive and individual biochemical and environmental influences [26–29]? Far from being concluded, the debate rebounds perpetually, taking advantage of the new light of science.

This debate has a direct impact on our project of carrying out a beauty measurer. Should we analyze the most beautiful objects and try to discover and catch the beauty canons or identify the emotional springs of consciousness, in order to provide materials capable of satisfying them? Among the important results of these philosophical works, let us report some marginal contributions, but fundamental for our purpose, that clearly help us to distinguish the part of the aesthetics in art (in particular, contemporary art) [30–32], as well as to reveal multiple elements that could mask the role of aesthetics in the appeal of an artwork [33–36].

2.2. Neuro-Aesthetic Approaches

If our scientific world has now passionately devoted itself to neural networks (mainly CNN), it has done so for 30 years, and with the same enthusiasm, in the perspectives of cerebral imagery. Magnetic resonance imaging and functional magnetic resonance imaging (fMRI) provide exceptional tools for trying to understand how our brain works. They were used yet from the beginning to understand the mysterious rules of our artistic judgment, and fMRI has thus given birth to a distinct branch of neurobiology, which recognizes itself as neuro-aesthetics [37,38]. The literature contains more than 3000 publications, where most of them are devoted to visual arts. Neuro-aesthetics brings us much knowledge, much more than we can summarize in a few lines, of course, but we can find in [39] a very good synthesis of the related works.

Neuro-aesthetics allows us to dismiss the idea, considered many years ago, of hedonic areas: it was the idea that there existed specialized areas in the treatment of beauty. On the contrary, it is known today that there are many cerebral areas, involved also in other different cerebral tasks, that

contribute to aesthetic judgment. They are shown and explained in detail in Figure 2 and are briefly summed up and grouped in the points below:

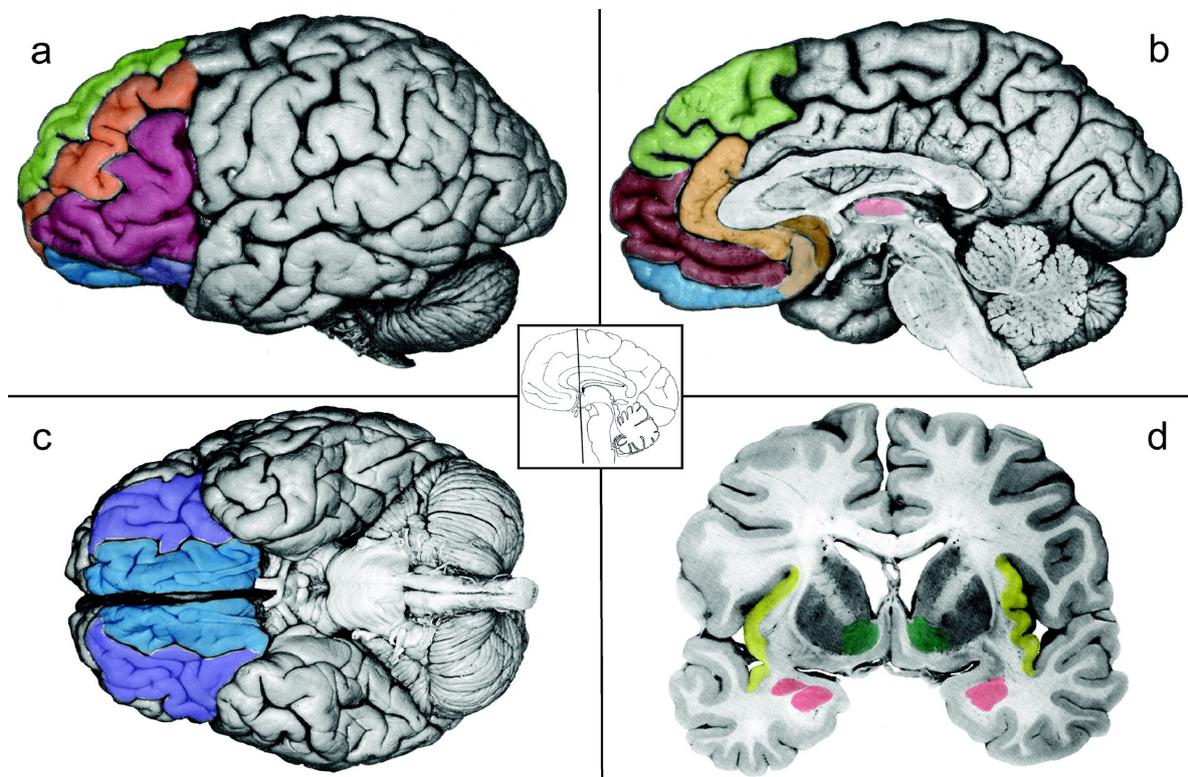


Figure 2. The areas involved during aesthetic judgment [40]. They are the areas that compose the prefrontal brain circuitry. Figure 2 highlights the cortical components [41]. The ventral system includes two closely-connected circuits that are anchored in the orbitofrontal cortex (OFC; c). The sensory system involves the lateral sector of the OFC (a,c, purple). It is closely connected to the anterior insula (d, yellow) and the basolateral complex in the amygdala (d, rose, ventral aspect). The visceromotor circuitry includes the ventral portion of the ventromedial prefrontal cortex, which lies in the medial sector of the OFC (a–c, blue) where the medial and lateral aspects of OFC connect; the ventromedial prefrontal cortex is closely connected to the amygdala (including the central nucleus, d, rose, dorsal aspect) and the subgenual parts of the anterior cingulate cortex on the medial wall of the brain (b, copper and peach). The dorsal system is associated with mental state attributions including the dorsal aspect of the ventromedial prefrontal cortex corresponding to the frontal pole (b, maroon), the anterior cingulate (2b, peach), and the dorsomedial prefrontal cortex (2a,b, green). The ventrolateral prefrontal cortex is shown in red (a). Structures in the reward circuitry include the OFC, dorsolateral prefrontal (2a, orange) and cingulate cortex (b, copper and tan), the thalamus (b, light pink), the ventral striatum d, green), the amygdala (d, rose), the hippocampus (d, gray), and the limbic brainstem.

- The visual areas, the occipital and inferior lateral zones, the insular cortex, and the superior parietal lobule are active for the task of vision and also for the extraction of shapes, colors, movements, and faces.
- The orbitofrontal cortex acts during the evaluation of risks, and also when we feel pleasure. It is clear that humans feel pleasure when they look at beautiful objects [42]. It seems also an important part in the control of our decisions.
- The insular cortex, which controls our emotions, is also an important and always involved part when observing artworks, images, and photographs.
- The areas engaged in cognition (the amygdala) and memory operations (medial parietal areas, prefrontal lobe) are often active in the task of aesthetic assessment.

- The areas in charge of the premotor control (ventral premotor cortex, temporal lobes, hippocampus) are active, specifically in situations of strong empathy and embodiment, which often occur when observing an artwork.

Regarding how the mentioned brain areas that allow aesthetic judgment capability were individuated, numerous details are given in [40,43–45] about the experiments conducted, the results obtained, and on the conclusions that can be drawn from them. In [46,47] was presented a synthetic vision of our knowledge on this subject.

However, we have to remark how fMRI, in the state of its actual development, is insufficient to understand the mechanisms actually implemented in brain circuits: the response time of instruments, the need for averaging experiments, and individuals altering the deductive abilities of this technique. In particular, it is almost impossible to trace the chronology between visual stimuli and activation of higher areas, conditions that are essential for a true causal explanation [48]. It is on these conditions, however, that the debate between objectivists and subjectivists goes on. Finally, It should also be mentioned that works based on analysis through fMRI techniques face deeper theoretical criticism [49,50].

2.3. Experimental Psychology, Psycho-Sociology, and Photography

Another important source of information on aesthetics comes from the literature on photography, and in particular from the recommendations of photographers and art photography books. The relevance of these notions can be noticed in various ways: through their high frequency in the photos/artworks (this is, for instance, the case of the rule of thirds: the rule suggests that a photo should be divided into nine equal parts by two equally-spaced horizontal lines and two equally-spaced vertical lines). Then, the most relevant subjects should be positioned along such lines or their intersections. The proponents of this method claim that aligning compositional objects with these lines creates higher interest, tension, and energy in photos, than simply centering them [51]. An example is contained in Figure 3 (original photo by Pir6mon; the edited photo of Figure 3 was from Teeks99 (<https://commons.wikimedia.org/wiki/File:RuleOfThirds-SideBySide.gif>, under CC BY-SA 3.0 license <https://creativecommons.org/licenses/by-sa/3.0>)), through the popularity of their author, and even through the rating of the artworks that make use of them on the art market (which then seek the sociological approach [52]). Eventually, many verifications can be conducted using tests of experimental psychology, as well as statistical verifications on corpora.



Figure 3. Two pictures of the Moul n'ga Cirque in the Tadrart region, Southeast Algeria, with wavy clouds above. The picture on the right is cropped with the rule of thirds; the one on the left is not.

However, we can conclude that the rules proposed in the literature of photography are not universal: very few of them resist objective verifications. First of all, we have to reject the idea that many classical features are fundamental in the definition of the beauty of an image, such as

resolution, spike, signal-to-noise ratio, and contours, because quality and beauty evolve in spaces that are subjectively different [53,54]. Many rules of composition are also not universal: the rule of thirds, the Fibonacci spiral, the golden ratio, symmetries, privileged orientations, etc. [55–57]. The rules regarding the distribution of shadows and lights seeming to validate regular $1/f^2$ decays in the power density spectrum are fairly well verified [58], while the laws on the histogram of gray levels are reduced to fairly good preferences [59]. Finally, the preferences on the chromatic palette, which seemed to be well anchored with the well-established theories of Moon and Spencer and Matsuda, collide with several refutations [60,61].

The universality of aesthetics criteria is therefore often defeated when one refers to these kinds of literature. Further, many studies on the observer's eye gaze during the examination of photographs confirm that the rules governing the analysis of a photo or an artwork are very dependent on the cultural baggage of the observer [62,63].

3. Machine Learning Approaches

As we said in Section 2, the first aesthetic measurement systems were algebraic (they did not use any machine learning technique) and did not result in a very large consensus inside the community. However, in recent years, many research efforts have been made, and various approaches have been proposed, exploiting the latest machine learning techniques and the huge available datasets. The works available in the literature can be grouped following this categorization: First, we have to divide between works that follow the classical machine learning pipeline (feature extraction of handcrafted designed features, followed by classification or regression) and works that make use of deep learning techniques, where the feature representation is learned from a huge amount of data. These methods showed promising performances in many tasks, such as recognition, localization, retrieval, and tracking, beating the capability of conventional handcrafted features [64–67].

3.1. Classical Machine Learning Approaches

If we consider the works that follow the standard machine learning pipeline, we can individuate two subgroups, according to the way the problem is formulated: we can divide between aesthetic classification and aesthetic regression, which are both embodied in the supervised learning approach. A typical pipeline for this one assumes a set of training data $\{x_i, y_i\}_{i \in [1, N]}$, from which a function $f : g(X) \rightarrow Y$ is learned, where $g(x_i)$ denotes the feature representation of the image x_i . The label y_i is represented as $\{0, 1\}$ for a binary classification problem (where the function f is considered to be a classifier) or a continuous score range for regression (where f is considered to be a regressor). Following this formulation, the said pipeline can be broken into two main components, i.e., feature extraction and decision component (which can be a regressor or a classifier).

Regarding feature extraction, it is the first component to design for an image aesthetics assessment system. The aim is to extract meaningful and robust feature representations that describe the aesthetic content of an image. Such features are assumed to model the quantity of the photographic/artistic aspect of an image to distinguish between them. Many efforts have been tried to design features catching the aesthetics rules.

Within the set of methods that face the aesthetic assessment as a binary classification problem (i.e., they distinguish between aesthetic and unaesthetic images), most of them have focused on designing features able to imitate the way people perceive the aesthetic quality of images. For instance, Datta et al. [68] designed specific visual features (colorfulness, the rule of thirds, low depth of field indicators, etc.) and made use of Support Vector Machine (SVM) and Decision Tree (DT) to classify between *beautiful* and *ugly* images. In Marchesotti et al. [13], it was demonstrated that generic image descriptors, such as the well-known GIST, Bag-of-Visual-words (BOV) encoded from Scale-Invariant Feature Transform (SIFT) information, and the Fisher Vector (FV) encoded from SIFT information, are able to capture several measures useful for aesthetic evaluation of images. Nishiyama et al. [69] proposed a method that relied on color harmony and bags of color patterns to catch color variations

in local regions. In Simond et al. [70], it was shown that the aesthetics in images depends on context, since the authors obtained more accurate predictions by selecting features for specific image categories.

In the literature, we can also notice many methods that are able to learn effective aesthetic features directly from images through deep learning methods and then make use of the classical machine learning pipeline. Kao et al. [71] exploited an SVM using features extracted from a CNN, pre-trained on the ImageNet classification task [72]. Lu et al. [5] presented in their work the RAting PICTorical aesthetics using Deep learning (RAPID) system, which made use of a CNN to learn features for aesthetic categorization automatically.

Within the approaches that consider aesthetic assessment as a regression problem, i.e., they predict an aesthetics score or rating of the images, Bhattacharya et al. [73] proposed to use saliency maps and a high-level semantic segmentation technique for extracting aesthetic features, then for training a Support Vector Regression (SVR) machine. Datta et al. [68] proposed the use of Linear Regression (LR) with polynomial terms of the features to predict the aesthetics score. In Wu et al. [74] was designed a new algorithm called Support Vector Distribution Regression (SVDR) in order to use a distribution of user ratings instead of a scalar one for model learning. More recently, Kao et al. [71] proposed a CNN regression model, which achieved state-of-the-art results on aesthetic quality assessment.

However, most of the conventional approaches [3,13,14,68,69,73,75–83], that typically adopt handcrafted features to model many photographic rules have been outperformed by the most recent works in which generic deep features [84,85] and learned deep features [5,6,71,86–94] are used.

3.2. Deep Learning Approaches

From their appearance in the field of aesthetic evaluation, the DNN-based techniques showed superior performance with respect to the more conventional approaches. The architectures adopted are those found throughout the field of recognition in images: layers of convolutions followed by totally-connected layers, or more recently, only convolutional layers. However, many refinements have been proposed to adapt these systems to the specificities of the problem:

- Several solutions have been proposed to allow treating very large images while preserving the fine structure of details: window preselection around points of interest [10,95], parallel processing of randomly-drawn windows [96], use of hierarchical structures [7,97], etc. Despite these solutions, the size of the operational DNN input layers is a limit for the works on aesthetics that handle large images.
- Taking into account additional information, very important in the choice of the criteria to be applied, led to networks with multiple flows [9,10], which exploit various knowledge: the type of image, the style of the photo, the class of the main object, etc.
- The reproduction of certain brain mechanisms led to the separation of the processing architecture in different ways [9,98] or, sometimes, in a succession of DNNs: one in charge of the low level, another in charge of the high information level [92].

The use of DNN-based techniques significantly changed the work done on the aesthetics of images. A first element of differentiation concerns the choice of databases. The need for very large learning databases led to the abandonment of the works that used original databases, composed only by a few thousand images. The community has thus focused on the AVA database, which has the merit of having images that are often very beautiful with many annotations on each image. However, for training networks, its size (it is made up of 250,000 images) is often insufficient. Then, usually, researchers perform a dataset augmentation by manipulating images [92].

A second element to consider is the almost complete disappearance (except in [10]) of the aesthetic criteria for the construction of the DNN architecture. The works that rely on information external to the image mainly use data based on the type of image: interior, portrait, sport, etc., data that seem, however, quite unrelated to the beauty of the image.

3.3. Datasets

In the assessment of aesthetics, a training and a test set containing both high-quality and low-quality images are assumed. Evaluating the aesthetic quality of a given image, i.e., the ground-truth, is, however, a completely subjective task. Hence, it is challenging to obtain a large amount of well-annotated data. Most of the earlier articles [68,76,77] on aesthetic assessment built a small amount of private image data. These datasets usually contain a few thousand images at maximum, with binary labels or aesthetics scores for each contained image. These datasets on which the performances of the models are evaluated are not publicly available. Later, a huge effort was made to contribute publicly available aesthetics datasets of a larger scale for more comparable evaluation of performances. In the following, we list the main datasets that are frequently used in benchmarking for automatic aesthetic assessment:

- The Photo.net dataset and the DP Challenge dataset [99,100]. This can be considered the earliest attempt to construct large-scale image databases for aesthetic assessment. The Photo.net dataset contains 20,278 images, with a minimum of ten score ratings for every image. The ratings range is from zero to seven, with seven assigned to the most aesthetically-pleasing photos. Typically, images uploaded to Photo.net are evaluated as somewhat pleasing [99]. The DPChallenge dataset is more challenging and provides several ratings. The DPChallenge dataset is composed of 16,509 images and was extended by the Aesthetic Visual Analysis (AVA) dataset, in which several images derived from DPChallenge.com are also included.
- The Chinese University of Hong Kong-PhotoQuality (CUHK-PQ) dataset [81,101]. It is composed of 17,690 images also collected from DPChallenge.com and many photographers. All the images come provided with binary aesthetic labels and are grouped into seven categories: architecture, landscape, humans, animals, plants, static, and night. Usually, the training and test set are selected as random partitions of a 50/50 split, or a ten-fold cross-validation, where the ratio of the positive examples and the negative examples is around 1:3. Many sample images taken from the dataset are available in Figure 4.



Figure 4. Several images contained in the Chinese University of Hong Kong-PhotoQuality (CUHK-PQ) dataset [81,101]. Many distinctive differences can be visually observed between the high-quality and low-quality photos.

- The AVA dataset contains $\sim 250,000$ photos [14]. These were obtained from DPChallenge.com and labeled with scores. Every image received hundreds of votes, in the range one to ten. The average score of an image is commonly taken to be the ground truth. The dataset contains many challenging examples. For the task of binary aesthetic classification, images with an average score higher than a threshold of $5 + \nu$ are treated as positive examples, and images with a score lower than $5 - \nu$ are treated as negative examples. Further, the AVA dataset contains 14 style attributes and 60 category attributes. There are two typical training and test splits used with this dataset: (1) a large-scale standardized partition with $\sim 230,000$ training images and $\sim 20,000$ test images, with a hard threshold of $\nu = 0$, and (2) an easier partition modeling the one of CUHK-PQ, taking those images whose score ranking is at the top 10% and bottom 10%. This results in $\sim 25,000$

images for training and ~25,000 images for testing. The ratio of the total number of positive examples to that of the negative examples is around 12:5.

To date, the AVA dataset is regarded as a standard benchmark for evaluation of the performances of aesthetic assessment, as it is the first large-scale dataset with very detailed annotations. However, we must take into account, during the evaluation of the results, that the distribution of positive and negative examples in the dataset is fundamental for the effectiveness of trained models: false-positive predictions are as bad as having a low recall in image retrieval and searching applications. This factor is crucial, as the majority of the presented datasets are not well balanced.

3.4. Evaluation Metrics and Comparison of the Methods

Across the literature, we can find different metrics for performance evaluation of aesthetic assessment:

- The classification accuracy [5,6,13,14,68,73,75,78,84,87–95,102–106], which reports the proportion of the results that are correctly classified.
- The Euclidean distance between the ground truth and aesthetics ratings [76,105,107,108] and the correlation ranking [77,82,87] are used for evaluating performances in regression tasks.
- The Precision-and-Recall (PR) curve, used for instance in [3,69,78,80,109], considers the degree of relevance of the classified items and the retrieval rate of the items.
- The mean average precision [5,6,85,91] is the average precision between multiple queries, which is often used to summarize the PR curve for the considered set of samples.
- Finally, the Receiver-Operating Characteristic (ROC) curve [79,105,109–111] and area under the curve [81,101,109] concern the performance of binary classifiers when the discrimination threshold is varied.

As always happens, it is not feasible to provide a comparison between all the methods: different datasets and evaluation methods are taken into account within literature. Hence, we compare the results considering the AVA dataset. To date, the AVA dataset (assuming the standard partition) is considered as the most challenging by the majority of the reviewed works, and it is the most used. Further, the overall accuracy metric appears to be the most popular metric. It is always computed in the considered works, and it can be written as:

$$\text{Overall Accuracy} = \frac{TP + TN}{P + N},$$

where TP stands for True Positive, TN for True Negative, P for Positive, and N for Negative. We have to consider that this metric could be easily biased when considering unbalanced datasets. In Table 1, we can see the overall accuracies obtained by the cited articles.

Table 1. The subset of the reviewed methods that use AVA as the training dataset. We can notice the sensible improvement carried by the latest deep learning techniques, in terms of overall accuracy. Further, it seems that a proper balancing of the training and test set provides classifiers that obtain better performances. RAPID, RAting Pictorial aesthetics using Deep learning.

The Reviewed Methods Evaluated on the AVA Dataset				
Method	Dataset	Employed Method	Overall Accuracy	Train/Test Split
Marchesotti et al. [110]	AVA	Elastic Net	67.89%	Standard partition
AVA handcrafted features [14]	AVA	SVM	68.00%	Standard partition
Lu et al. [6]	AVA	CNN	72.85%	Standard partition
RAPID (full method) [5]	AVA	CNN	74.46%	Standard partition
Peng et al. [88]	AVA	CNN	74.50%	Standard partition
Kao et al. [94]	AVA	CNN	74.51%	Standard partition
Lu et al. [6]	AVA	CNN	75.41%	Standard partition
RAPID (improved version) [91]	AVA	CNN	75.42%	Standard partition
Kao et al. [71]	AVA	CNN	76.15%	Standard partition
Wang et al. [89]	AVA	CNN	76.94%	Standard partition
Kong et al. [87]	AVA	CNN	77.33%	Standard partition
Wang et al. [92]	AVA	CNN	78.08%	Standard partition
Tian et al. [90]	AVA	CNN	80.38%	10% subset/20k*2
Liu et al. [112]	AVA	CNN	83.09%	Standard partition
Zhang et al. [113]	AVA	Bayesian model	83.24%	10% subset/12.5k*2
Lv et al. [85]	AVA	Ranking model	84.32%	10% subset/20k*2
Dong et al. [84]	AVA	CNN	83.52%	10% subset/19k*2
Wang et al. [89]	AVA	CNN	84.88%	10% subset/25k*2

4. Analysis of the Works

Thanks to the work done by the computer vision community, we can evaluate the beauty of photos using machine learning techniques. We discuss here some interesting points that arise from the analyzed works, and from which we can give many future directions.

4.1. Non-Exploited Features

The classical DNN architectures showed their power in recognizing and locating objects, even deformed or partially occluded. However, it seems that some important properties of the aesthetic evaluation would require evolving such architectures. We have already pointed out the importance of being able to process large images with many fine details. Notice also the importance that should be given to chromatic harmony, which is undeniably an important component of aesthetics (the work of [95] is exemplary). It is not obvious that architectures that carry out convolutions in the first layers respect the nuances. The internal construction of photography is itself an important element of the aesthetic quality of the photos (D. Diderot made it a major argument of his aesthetics approach [23]). Let us recognize that, although many works tried to take it into account, very few gave themselves the means to do so through the initially convoluted layers, and then those totally connected for the DNN. To our knowledge, only the authors of [10] considered this point of view.

4.2. The Binary Criterion: Ugly vs. Beautiful

The binary criterion is widely adopted by the community to compare the various available approaches. It can be applied quickly on very large databases; it can be easily considered for different databases; it can be confirmed with a simple visual check; it offers a good solution to some of the problems that the Internet community poses: sort very quickly large archives to keep quintessence, provide attractive examples for illustrations, assist an operator in his/her shooting, and so on.

However, this criterion suffers from being hugely simplified. It is based on the assumption that all the images come from one or the other category, a postulate of which no trace is found in the literature. Moreover, it is commonly accepted, both in philosophy and in neurobiology, that the attribute of

beauty has only a positive valence and no equivalent to a negative valence (which would be called *ugliness*), where negative valence could be supported by other attributes like “scary”, “sad”, “boring”, “banal”, “rough”, and many others.

Thus, the complexity of the information transmitted through the annotations for each image of the AVA database is currently insufficiently analyzed, even if some works tried to exploit it [11,92,114]. It would be important, however, to distinguish the annotations considering the heterogeneity of interest, attention, culture, motivation, etc., of the experts and of the intrinsic properties of the photo (what the authors of [114] attributed to be an inherent “difficulty” of interpretation).

4.3. A Continuous Ranking for Evaluation

From the beginning [68], many works had set themselves the objective of classifying photos according to a scale of beauty that was almost continuous. Although many algorithms provide a score between zero and 10, few studies report the quality of these notations [11] except to refine the binary decision [92,94]. The evaluation of a continuous ranking is very difficult today and seems to us a major issue. Let us note that in [15], a classification at five levels made it possible to refine the measurement substantially. Note especially the very original approach of [115], which proposed to compare two by two the images of the database, to reach a relative evaluation.

4.4. Which Beauty? Which Expert?

The images used for testing performances represent what we can expect from quality images from social networks. The most beautiful are undoubtedly generally superior to the *ugly* ones. However, if the qualities of beautiful images are not always obvious, we can notice that they rarely show the flaws that make ugly images so evident: poor composition, poor chromatic distribution, lack of focus, etc.

An attentive and demanding observer will often disagree with the decisions made by the system, even if these decisions are in accordance with the judgments annotated in the database. This is often explained either because the beautiful images are commonplace or, especially, because a quality image has been classified as ugly. In the latter case, it is frequently observed that the original aspects of the image have been ignored. Further, it turns out that DNNs prefer *normal* images, and this is hardly in accordance with experts’ recommendations.

Finally, let us discuss one of the most sensitive points of the DNN approach. The importance of having a database of fairly high quality has been felt since the implementation of the approaches that made use of handcrafted features, but it has become crucial for DNN approaches. The AVA database [14] provided a good answer to this request. Beyond the collection of images, AVA provides several pieces of information attached to each photo: the evaluations, the theme covered by the image (among more than 900, taken from the competitions of DPChallenge), a semantic annotation (among 66), and the photographic style (ascribed by professional photographers, among 14).

Is this sufficient? This is not certain. Certainly, for the objectivists, who place all the beauty only in the object, the object is faithfully reproduced in, AVA and on average, the expression of consensus on its appreciation is annotated. There are therefore all the elements sufficient to allow a machine to reproduce the human judgment, provided that we master the AI techniques.

If a more important place to the observer is given, the information that will be needed for the evaluation will be more important. Without adopting the extreme positions of the subjectivists, who attribute the total authority over the judgment to the moods of the observer, one can ask for other information to simulate a feeling that appeals to the sensations, on the one hand (those coming by the visual signal of the image), the conscious and unconscious mental faculties of the observer, on the other hand, and finally his/her temperament. It is unlikely that we can draw such information from the AVA database. Thus, in [87], it was considered necessary to build a database, Aesthetics with Attributes Database (AADB), different from AVA, keeping the evaluator’s mark during the evaluation. The authors indicated that such a choice made it possible to obtain a better match between the ratings obtained by the same expert. In [15], a great deal of attention was paid to the cultural context of the

experts used to build the BEAUTY database. Only users from a small number of countries with a high degree of cultural homogeneity were selected, and their opinions were subsequently screened to discard points of view were too different.

5. Conclusions

The success of methods for evaluating the beauty of photos and images is certain. Taking advantage of a very large number of images, they allow separating with reasonable performances the most beautiful from the ugly ones. There is no doubt that these performances will improve over time, as the works currently being presented still have a great deal of margin for progress.

However, let us point out that today, the interest of the presented methods resides mainly in their capacity to elaborate a first sorting on large quantities of images. If the aim is to distinguish among the most beautiful images, it is still necessary to analyze manually the automatic returned sort (provided by machine learning methods) and then to select the small number of images that subjectively surpass all the others.

Further, we regret, as we do for the majority of the other kinds of recognition problems, that DNN-based solutions are delivered to us without explicit intermediate decision steps, or rather that these intermediate results, accessible in the form of maps, are not available and readable today with our knowledge. Thus, if we know how to sort the images according to beauty, we do not really know how this sorting is done. This can be considered, for our understanding, a step back from previous approaches that made use of handcrafted features.

Finally, let us insist on the fact that the methods implemented to date have completely ignored an important part of the aesthetic judgment that the literature puts forward: the cultural and socio-educational context of the observer. This lack is understandable because, if aesthetics is a complex and poorly-understood field, culture is even more complex and poorly modeled. We do not know how to use it in the proposed architectures, but this fact allows us to reason about a hidden culture, which is a step beyond the knowledge of the expert. When evaluating the beauty of images, it is therefore a community of experts or enthusiasts in photography, distributed throughout the world, who are rather fond of social life via the Internet, often enthusiastic about technology, who serves as a reference. This is a fundamental consideration in terms of what has helped to build the latest aesthetics databases.

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Burningham, N.; Pizlo, Z.; Allebach, J.P. Image Quality Metrics. In *Encyclopedia of Imaging Science and Technology*; American Cancer Society: 2002. Available online: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471443395.img038> (accessed on 27 June 2019) doi:10.1002/0471443395.img038.
2. Datta, R.; Li, J.; Wang, J.Z. Learning the Consensus on Visual Quality for Next-generation Image Management. In Proceedings of the 15th MM '07 ACM International Conference on Multimedia, Augsburg, Germany, 25–29 September 2007; ACM: New York, NY, USA, 2007; pp. 533–536. doi:10.1145/1291233.1291364. [CrossRef]
3. Ke, Y.; Tang, X.; Jing, F. The Design of High-Level Features for Photo Quality Assessment. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; Volume 1, pp. 419–426. doi:10.1109/CVPR.2006.303. [CrossRef]
4. Deng, Y.; Loy, C.C.; Tang, X. Image Aesthetic Assessment: An experimental survey. *IEEE Signal Process. Mag.* **2017**, *34*, 80–106. doi:10.1109/msp.2017.2696576. [CrossRef]
5. Lu, X.; Lin, Z.; Jin, H.; Yang, J.; Wang, J.Z. RAPID: Rating Pictorial Aesthetics Using Deep Learning. In Proceedings of the 22nd MM '14 ACM International Conference on Multimedia; ACM: New York, NY, USA, 2014; pp. 457–466. doi:10.1145/2647868.2654927. [CrossRef]

6. 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 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 990–998. doi:10.1109/ICCV.2015.119. [[CrossRef](#)]
7. Mai, L.; Jin, H.; Liu, F. Composition-Preserving Deep Photo Aesthetics Assessment. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26–30 June 2016; pp. 497–506. doi:10.1109/CVPR.2016.60. [[CrossRef](#)]
8. Wang, Z.; Liu, D.; Chang, S.; Dolcos, F.; Beck, D.; Huang, T. Image aesthetics assessment using Deep Chatterjee’s machine. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 941–948. doi:10.1109/IJCNN.2017.7965953. [[CrossRef](#)]
9. Kao, Y.; He, R.; Huang, K. Deep Aesthetic Quality Assessment With Semantic Information. *IEEE Trans. Image Process.* **2017**, *26*, 1482–1495. doi:10.1109/tip.2017.2651399. [[CrossRef](#)] [[PubMed](#)]
10. Ma, S.; Liu, J.; Chen, C.W. A-Lamp: Adaptive Layout-Aware Multi-patch Deep Convolutional Neural Network for Photo Aesthetic Assessment. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 722–731. doi:10.1109/CVPR.2017.84. [[CrossRef](#)]
11. Talebi, H.; Milanfar, P. NIMA: Neural Image Assessment. *IEEE Trans. Image Process.* **2018**, *27*, 3998–4011. doi:10.1109/TIP.2018.2831899. [[CrossRef](#)] [[PubMed](#)]
12. Datta, R.; Wang, J.Z. ACQUINE: Aesthetic Quality Inference Engine—Real-time Automatic Rating of Photo Aesthetics. In *Proceedings of the International Conference on MIR '10 Multimedia Information Retrieval*; ACM: New York, NY, USA, 2010; pp. 421–424. doi:10.1145/1743384.1743457. [[CrossRef](#)]
13. Marchesotti, L.; Perronnin, F.; Larlus, D.; Csurka, G. Assessing the aesthetic quality of photographs using generic image descriptors. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011; pp. 1784–1791. doi:10.1109/ICCV.2011.6126444. [[CrossRef](#)]
14. Murray, N.; Marchesotti, L.; Perronnin, F. AVA: A large-scale database for aesthetic visual analysis. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 2408–2415. doi:10.1109/CVPR.2012.6247954. [[CrossRef](#)]
15. Schifanella, R.; Redi, M.; Aiello, L.M. An Image is Worth More than a Thousand Favorites: Surfacing the Hidden Beauty of Flickr Pictures. In Proceedings of the ICWSM’15: Proceedings of the 9th AAAI International Conference on Web and Social Media, Oxford, UK, 26–29 May 2015.
16. Henry, C. *Introduction à une Esthétique Scientifique*; La revue contemporaine: Paris, France, 1885.
17. Garabedian, C.A. Book Review: Aesthetic Measure. *Bull. Am. Math. Soc.* **1934**, *40*, 7–11. doi:10.1090/s0002-9904-1934-05764-1. [[CrossRef](#)]
18. Birkhoff, G.D. *Aesthetic Measure*; Harvard University Press: Cambridge, MA, USA, 1933.
19. Moles, A.A. Théorie de l’information et perception esthétique. *Revue Philosophique de la France et de l’Étranger* **1957**, *147*, 233–242.
20. Bense, M. Einführung in die informationstheoretische Ästhetik. In *Ausgewählte Schriften: Band 3 Ästhetik und Texttheorie*; J.B. Metzler: Stuttgart, Germany, 1998; pp. 251–417. doi:10.1007/978-3-476-03716-9_3.
21. Rigau, J.; Feixas, M.; Sbert, M. Informational Aesthetics Measures. *IEEE Comput. Graph. Appl.* **2008**, *28*, 24–34. doi:10.1109/MCG.2008.34. [[CrossRef](#)]
22. Cooper, J.M.; Hutchinson, D.S. *Plato: Complete Works*; Hackett Publishing: Indianapolis, IN, USA, 1997.
23. Grimm, F.M.; Diderot, D. *Correspondance Littéraire, Philosophique et Critique de Grimm et de Diderot, depuis 1753 jusqu’en 1790*; Furne: Paris, France, 1831.
24. Hegel, G.W.F. *Vorlesungen über die Ästhetik*; Duncker und Humblot: Berlin, Germany, 1842; Volume 1.
25. Kant, I. *Critique of Judgment*; Hackett Publishing: Indianapolis, IN, USA, 1987.
26. Gombrich, E.H. *Art and Illusion: A Study in the Psychology of Pictorial Representations*; Princeton University Press: Princeton, NJ, USA, 1960.
27. Berlyne, D.E. *Aesthetics and Psychobiology*; Appleton-Century-Crofts: New York, NY, USA, 1971; Volume 336.
28. Zemach, E.M. *La Beauté Réelle: Une Défense du Réalisme Esthétique*; Presses Universitaires de Rennes: Rennes, France, 2005.
29. Pouivet, R. *Le Réalisme Esthétique*; Presses universitaires de France: Paris, France, 2015.
30. Tatarkiewicz, W. *History of Aesthetics*; Harrrel, J., Ed.; PwN-Polish Scientific Publishers: Mouton, Warsaw, Poland, 1970; Volume I.
31. Danto, A. The artworld. *J. Philos.* **1964**, *61*, 571–584. [[CrossRef](#)]

32. Arnheim, R. *Art and Visual Perception: A Psychology of the Creative Eye*; University of California Press: Oakland, CA, USA, 1965.
33. Solso, R.L. *Cognition and the Visual Arts*; MIT Press: Cambridge, MA, USA, 1996.
34. Lange, C.G.; James, W. *The Emotions*; Williams & Wilkins: Philadelphia, PA, USA, 1922; Volume 1.
35. Sperber, D.; Wilson, D. Relevance theory. In *Handbook of Pragmatics*; Blackwell: Oxford, UK, 2004; pp. 607–632.
36. Dessalles, J.L. *La Pertinence et ses Origines Cognitives-Nouvelles Théories*; Hermes-Science Publications: Paris, France, 2008.
37. Zeki, S. *Inner Vision: An Exploration of Art and the Brain*; Oxford University Press: New York, NY, USA, 1999.
38. Zeki, S. Artistic creativity and the brain. *Science* **2001**, *293*, 51–52. [[CrossRef](#)]
39. Brown, S.; Gao, X.; Tisdelle, L.; Eickhoff, S.B.; Liotti, M. Naturalizing aesthetics: brain areas for aesthetic appraisal across sensory modalities. *Neuroimage* **2011**, *58*, 250–258. [[CrossRef](#)]
40. Chatterjee, A.; Vartanian, O. Neuroscience of aesthetics. *Ann. N. Y. Acad. Sci.* **2016**, *1369*, 172–194. [[CrossRef](#)]
41. Barrett, L.F.; Mesquita, B.; Ochsner, K.N.; Gross, J.J. The experience of emotion. *Annu. Rev. Psychol.* **2007**, *58*, 373–403. [[CrossRef](#)] [[PubMed](#)]
42. Kühn, S.; Gallinat, J. The neural correlates of subjective pleasantness. *NeuroImage* **2012**, *61*, 289–294. doi:10.1016/j.neuroimage.2012.02.065. [[CrossRef](#)] [[PubMed](#)]
43. Cinzia, D.D.; Vittorio, G. Neuroaesthetics: A review. *Curr. Opin. Neurobiol.* **2009**, *19*, 682–687. doi:10.1016/j.conb.2009.09.001. [[CrossRef](#)] [[PubMed](#)]
44. Ishizu, T.; Zeki, S. Toward A Brain-Based Theory of Beauty. *PLoS ONE* **2011**, *6*, e21852. doi:10.1371/journal.pone.0021852. [[CrossRef](#)] [[PubMed](#)]
45. Vartanian, O.; Skov, M. Neural correlates of viewing paintings: Evidence from a quantitative meta-analysis of functional magnetic resonance imaging data. *Brain Cogn.* **2014**, *87*, 52–56. doi:10.1016/j.bandc.2014.03.004. [[CrossRef](#)] [[PubMed](#)]
46. Changeux, J.P. *La beauté Dans le Cerveau*; Odile Jacob: Paris, France, 2016.
47. Changeux, J.P. Beauty in the brain: For a neuroscience of art. *Rendiconti Lincei* **2012**, *23*, 315–320. [[CrossRef](#)]
48. Ehtemami, A.; Scott, R.; Bernadin, S. A Survey of FMRI Data Analysis Methods. In Proceedings of the SoutheastCon 2018, St. Petersburg, FL, USA, 19–22 April 2018; pp. 1–7.
49. Vidal, F. Neuroaesthetics: Getting rid of art and beauty. *BioSocieties* **2012**, *7*, 209–213. doi:10.1057/biosoc.2012.7. [[CrossRef](#)]
50. Brown, S.; Dissanayake, E. The arts are more than aesthetics: Neuroaesthetics as narrow aesthetics. In *Neuroaesthetics*; Routledge: Abingdon-on-Thames, UK, 2018; pp. 43–57.
51. Peterson, B. *Learning to See Creatively: Design, Color, and Composition in Photography*; Amphoto Books, London, UK, 2015.
52. Bourdieu, P.; Castel, R.; Schnapper, D.; Boltanski, L.; Lagneau, G.; Chamboredon, J.C. *Un Art Moyen: Essai sur les Usages Sociaux de la Photographie*; Les éditions de Minuit: Paris, France, 1965.
53. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Process. Lett.* **2013**, *20*, 209–212. doi:10.1109/lsp.2012.2227726. [[CrossRef](#)]
54. Chandler, D.M. Seven Challenges in Image Quality Assessment: Past, Present, and Future Research. *ISRN Signal Process.* **2013**, *2013*, 905685. doi:10.1155/2013/905685. [[CrossRef](#)]
55. McManus, I.C.; Stöver, K.; Kim, D. Arnheim’s Gestalt Theory of Visual Balance: Examining the Compositional Structure of Art Photographs and Abstract Images. *i-Perception* **2011**, *2*, 615–647. doi:10.1068/i0445aap. [[CrossRef](#)]
56. Amirshahi, S.A.; Hayn-Leichsenring, G.U.; Denzler, J.; Redies, C. Evaluating the Rule of Thirds in Photographs and Paintings. *Art Percept.* **2014**, *2*, 163–182. doi:10.1163/22134913-00002024. [[CrossRef](#)]
57. Hübner, R.; Fillinger, M.G. Comparison of Objective Measures for Predicting Perceptual Balance and Visual Aesthetic Preference. *Front. Psychol.* **2016**, *7*. doi:10.3389/fpsyg.2016.00335. [[CrossRef](#)] [[PubMed](#)]
58. Schweinhardt, A.M.; Esock, E.A. Structural Content in Paintings: Artists Overregularize Oriented Content of Paintings Relative to the Typical Natural Scene Bias. *Perception* **2013**, *42*, 1311–1332. doi:10.1068/p7345. [[CrossRef](#)]
59. Attewell, D.; Baddeley, R.J. The distribution of reflectances within the visual environment. *Vis. Res.* **2007**, *47*, 548–554. doi:10.1016/j.visres.2006.11.015. [[CrossRef](#)] [[PubMed](#)]
60. Mehrabian, A. Individual differences in stimulus screening and arousability. *J. Person.* **1977**, *45*, 237–250. doi:10.1111/j.1467-6494.1977.tb00149.x. [[CrossRef](#)] [[PubMed](#)]

61. Smith, D. Color-person-environment relationships. *Color Res. Appl.* **2008**, *33*, 312–319. doi:10.1002/col.20424. [[CrossRef](#)]
62. Tatler, B.W.; Wade, N.J.; Kwan, H.; Findlay, J.M.; Velichkovsky, B.M. Yarbus, Eye Movements, and Vision. *i-Perception* **2010**, *1*, 7–27. doi:10.1068/i0382. [[CrossRef](#)] [[PubMed](#)]
63. Huston, J.P.; Nadal, M.; Mora, F.; Agnati, L.F.; Conde, C.J.C. (Eds.) *Art, Aesthetics, and the Brain*; Oxford University Press: Oxford, UK, 2015. doi:10.1093/acprof:oso/9780199670000.001.0001.
64. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436. [[CrossRef](#)] [[PubMed](#)]
65. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
66. Bodini, M. A review of facial landmark extraction in 2d images and videos using deep learning. *Big Data Cognit. Comput.* **2019**, *3*, 14. [[CrossRef](#)]
67. Boccignone, G.; Bodini, M.; Cuculo, V.; Grossi, G. Predictive Sampling of Facial Expression Dynamics Driven by a Latent Action Space. In Proceedings of the 2018 14th International Conference on Signal-Image Technology Internet-Based Systems (SITIS), Las Palmas de Gran Canaria, Spain, 26–29 November 2018; pp. 143–150. doi:10.1109/SITIS.2018.00031. [[CrossRef](#)]
68. Datta, R.; Joshi, D.; Li, J.; Wang, J.Z. Studying Aesthetics in Photographic Images Using a Computational Approach. In *Computer Vision—ECCV 2006*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 288–301. doi:10.1007/11744078_23.
69. Nishiyama, M.; Okabe, T.; Sato, I.; Sato, Y. Aesthetic quality classification of photographs based on color harmony. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011. doi:10.1109/cvpr.2011.5995539.
70. Simond, F.; Arvanitopoulos, N.; Susstrunk, S. Image aesthetics depends on context. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015. doi:10.1109/icip.2015.7351513.
71. Kao, Y.; Wang, C.; Huang, K. Visual aesthetic quality assessment with a regression model. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015. doi:10.1109/icip.2015.7351067.
72. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009. doi:10.1109/cvpr.2009.5206848.
73. Bhattacharya, S.; Sukthankar, R.; Shah, M. A framework for photo-quality assessment and enhancement based on visual aesthetics. In Proceedings of the International Conference on Multimedia—MM10, Firenze, Italy, 25–29 October 2010. doi:10.1145/1873951.1873990.
74. Wu, O.; Hu, W.; Gao, J. Learning to predict the perceived visual quality of photos. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011. doi:10.1109/iccv.2011.6126246.
75. Tong, H.; Li, M.; Zhang, H.J.; He, J.; Zhang, C. Classification of Digital Photos Taken by Photographers or Home Users. In *Advances in Multimedia Information Processing—PCM 2004*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 198–205. doi:10.1007/978-3-540-30541-5_25.
76. Sun, X.; Yao, H.; Ji, R.; Liu, S. Photo assessment based on computational visual attention model. In Proceedings of the Seventeen ACM International Conference on Multimedia—MM’09, Beijing, China, 19–24 October 2009. doi:10.1145/1631272.1631351.
77. You, J.; Perkis, A.; Hannuksela, M.M.; Gabbouj, M. Perceptual quality assessment based on visual attention analysis. In Proceedings of the Seventeen ACM International Conference on Multimedia—MM’09, Beijing, China, 19–24 October 2009. doi:10.1145/1631272.1631356.
78. Luo, Y.; Tang, X. Photo and Video Quality Evaluation: Focusing on the Subject. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 386–399. doi:10.1007/978-3-540-88690-7_29.
79. Lo, L.Y.; Chen, J.C. A statistic approach for photo quality assessment. In Proceedings of the 2012 International Conference on Information Security and Intelligent Control, Yunlin, Taiwan, 14–16 August 2012. doi:10.1109/isic.2012.6449719.
80. Dhar, S.; Ordonez, V.; Berg, T.L. High level describable attributes for predicting aesthetics and interestingness. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011. doi:10.1109/cvpr.2011.5995467.

81. Tang, X.; Luo, W.; Wang, X. Content-Based Photo Quality Assessment. *IEEE Trans. Multimed.* **2013**, *15*, 1930–1943. doi:10.1109/tmm.2013.2269899. [\[CrossRef\]](#)
82. Yeh, M.C.; Cheng, Y.C. Relative features for photo quality assessment. In Proceedings of the 2012 19th IEEE International Conference on Image Processing, Orlando, FL, USA, 30 September–3 October 2012. doi:10.1109/icip.2012.6467496.
83. Marchesotti, L.; Perronnin, F. Learning beautiful attributes. In Proceedings of the British Machine Vision Conference 2013, Bristol, UK, 9–13 September 2013. doi:10.5244/c.27.7.
84. Dong, Z.; Shen, X.; Li, H.; Tian, X. Photo Quality Assessment with DCNN that Understands Image Well. In *MultiMedia Modeling*; Springer International Publishing: Cham, Switzerland, 2015; pp. 524–535. doi:10.1007/978-3-319-14442-9_57.
85. Lv, H.; Tian, X. Learning Relative Aesthetic Quality with a Pairwise Approach. In *MultiMedia Modeling*; Springer International Publishing: Cham, Switzerland, 2016; pp. 493–504. doi:10.1007/978-3-319-27671-7_41.
86. Bodini, M.; D’Amelio, A.; Grossi, G.; Lanzarotti, R.; Lin, J. Single Sample Face Recognition by Sparse Recovery of Deep-Learned LDA Features. In *Advanced Concepts for Intelligent Vision Systems*; Springer International Publishing: Cham, Switzerland, 2018; pp. 297–308. doi:10.1007/978-3-030-01449-0_25.
87. Kong, S.; Shen, X.; Lin, Z.; Mech, R.; Fowlkes, C. Photo Aesthetics Ranking Network with Attributes and Content Adaptation. In *Computer Vision—ECCV 2016*; Springer International Publishing: Cham, Switzerland, 2016; pp. 662–679. doi:10.1007/978-3-319-46448-0_40.
88. Peng, K.C.; Chen, T. Toward correlating and solving abstract tasks using convolutional neural networks. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Placid, NY, USA, 7–9 March 2016. doi:10.1109/wacv.2016.7477616.
89. Wang, W.; Zhao, M.; Wang, L.; Huang, J.; Cai, C.; Xu, X. A multi-scene deep learning model for image aesthetic evaluation. *Signal Process. Image Commun.* **2016**, *47*, 511–518. doi:10.1016/j.image.2016.05.009. [\[CrossRef\]](#)
90. Tian, X.; Dong, Z.; Yang, K.; Mei, T. Query-Dependent Aesthetic Model With Deep Learning for Photo Quality Assessment. *IEEE Trans. Multimed.* **2015**, *17*, 2035–2048. doi:10.1109/tmm.2015.2479916. [\[CrossRef\]](#)
91. Lu, X.; Lin, Z.; Jin, H.; Yang, J.; Wang, J.Z. Rating Image Aesthetics Using Deep Learning. *IEEE Trans. Multimed.* **2015**, *17*, 2021–2034. doi:10.1109/tmm.2015.2477040. [\[CrossRef\]](#)
92. Wang, Z.; Chang, S.; Dolcos, F.; Beck, D.; Liu, D.; Huang, T.S. Brain-inspired deep networks for image aesthetics assessment. *arXiv Preprint* **2016**, arXiv:1601.04155.
93. Zhang, L. Describing human aesthetic perception by deeply-learned attributes from flickr. *arXiv Preprint* **2016**, arXiv:1605.07699.
94. Kao, Y.; Huang, K.; Maybank, S. Hierarchical aesthetic quality assessment using deep convolutional neural networks. *Signal Process. Image Commun.* **2016**, *47*, 500–510. doi:10.1016/j.image.2016.05.004. [\[CrossRef\]](#)
95. Lu, P.; Peng, X.; Li, R.; Wang, X. Towards aesthetics of image: A Bayesian framework for color harmony modeling. *Signal Process. Image Commun.* **2015**, *39*, 487–498. doi:10.1016/j.image.2015.04.003. [\[CrossRef\]](#)
96. Lu, P.; Kuang, Z.; Peng, X.; Li, R. Discovering Harmony: A Hierarchical Colour Harmony Model for Aesthetics Assessment. In *Computer Vision—ACCV 2014*; Springer International Publishing: Cham, Switzerland, 2015; pp. 452–467. doi:10.1007/978-3-319-16811-1_30.
97. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. In *Computer Vision – ECCV 2014*; Springer International Publishing: Cham, Switzerland, 2014; pp. 346–361. doi:10.1007/978-3-319-10578-9_23.
98. Yao, L.; Suryanarayan, P.; Qiao, M.; Wang, J.Z.; Li, J. OSCAR: On-Site Composition and Aesthetics Feedback Through Exemplars for Photographers. *Int. J. Comput. Vis.* **2011**, *96*, 353–383. doi:10.1007/s11263-011-0478-3. [\[CrossRef\]](#)
99. Joshi, D.; Datta, R.; Fedorovskaya, E.; Luong, Q.T.; Wang, J.; Li, J.; Luo, J. Aesthetics and Emotions in Images. *IEEE Signal Process. Mag.* **2011**, *28*, 94–115. doi:10.1109/msp.2011.941851. [\[CrossRef\]](#)
100. Datta, R.; Li, J.; Wang, J.Z. Algorithmic inferencing of aesthetics and emotion in natural images: An exposition. In Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008. doi:10.1109/icip.2008.4711702.
101. Luo, W.; Wang, X.; Tang, X. Content-based photo quality assessment. In Proceedings of the 2011 International Conference on Computer Vision, Barcelona, Spain, 6–13 November 2011. doi:10.1109/iccv.2011.6126498.

102. Wong, L.K.; Low, K.L. Saliency-enhanced image aesthetics class prediction. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009. doi:10.1109/icip.2009.5413825.
103. Bhattacharya, S.; Sukthankar, R.; Shah, M. A holistic approach to aesthetic enhancement of photographs. *ACM Trans. Multimed. Comput. Commun. Appl.* **2011**, *7S*, 1–21. doi:10.1145/2037676.2037678. [CrossRef]
104. Wu, Y.; Bauckhage, C.; Thurau, C. The Good, the Bad, and the Ugly: Predicting Aesthetic Image Labels. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010. doi:10.1109/icpr.2010.392.
105. Lienhard, A.; Ladret, P.; Caplier, A. Low Level Features for Quality Assessment of Facial Images. In Proceedings of the 10th International Conference on Computer Vision Theory and Applications, Berlin, Germany, 11–14 March 2015. doi:10.5220/0005308805450552.
106. Yin, W.; Mei, T.; Chen, C.W. Assessing photo quality with geo-context and crowdsourced photos. In Proceedings of the 2012 Visual Communications and Image Processing, San Diego, CA, USA, 27–30 November 2012. doi:10.1109/vcip.2012.6410821.
107. Li, C.; Gallagher, A.; Loui, A.C.; Chen, T. Aesthetic quality assessment of consumer photos with faces. In Proceedings of the 2010 IEEE International Conference on Image Processing, Hong Kong, China, 26–29 September 2010. doi:10.1109/icip.2010.5651833.
108. Sun, R.; Lian, Z.; Tang, Y.; Xiao, J. Aesthetic visual quality evaluation of Chinese handwritings. In Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, Buenos Aires, Argentina, 25–31 July 2015.
109. Lo, K.Y.; Liu, K.H.; Chen, C.S. Assessment of photo aesthetics with efficiency. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, Japan, 11–15 November 2012; pp. 2186–2189.
110. Marchesotti, L.; Perronnin, F.; Meylan, F. Learning beautiful (and ugly) attributes. *BMVC* **2013**, *7*, 1–11.
111. Su, H.H.; Chen, T.W.; Kao, C.C.; Hsu, W.H.; Chien, S.Y. Scenic photo quality assessment with bag of aesthetics-preserving features. In Proceedings of the 19th ACM international conference on Multimedia—MM’11, Scottsdale, AZ, USA, 28 November–1 December 2011. doi:10.1145/2072298.2071977.
112. Liu, Z.; Wang, Z.; Yao, Y.; Zhang, L.; Shao, L. Deep Active Learning with Contaminated Tags for Image Aesthetics Assessment. *IEEE Trans. Image Process.* **2018**. doi:10.1109/tip.2018.2828326. [CrossRef] [PubMed]
113. Zhang, L.; Gao, Y.; Zimmermann, R.; Tian, Q.; Li, X. Fusion of Multichannel Local and Global Structural Cues for Photo Aesthetics Evaluation. *IEEE Trans. Image Process.* **2014**, *23*, 1419–1429. doi:10.1109/tip.2014.2303650. [CrossRef] [PubMed]
114. Jin, B.; Segovia, M.V.O.; Susstrunk, S. Image aesthetic predictors based on weighted CNNs. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016. doi:10.1109/icip.2016.7532767.
115. Schwarz, K.; Wieschollek, P.; Lensch, H.P.A. Will People Like Your Image? Learning the Aesthetic Space. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018. doi:10.1109/wacv.2018.00226.



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).