

Paintings and Drawings Aesthetics Assessment with Rich Attributes for Various Artistic Categories

Xin Jin^{1,2}, Qianqian Qiao¹, Yi Lu³, Huaye Wang¹, Shan Gao⁴, Heng Huang^{5*}, Guangdong Li⁶

¹Beijing Electronic Science and Technology Institute

²Beijing Institute for General Artificial Intelligence

³Central Academy of Fine Arts

⁴Beijing University of Technology

⁵University of Science and Technology of China

⁶Beijing Institute of Fashion Technology

Abstract

Image aesthetic evaluation is a highly prominent research domain in the field of computer vision. In recent years, there has been a proliferation of datasets and corresponding evaluation methodologies for assessing the aesthetic quality of photographic works, leading to the establishment of a relatively mature research environment. However, in contrast to the extensive research in photographic aesthetics, the field of aesthetic evaluation for paintings and drawings has seen limited attention until the introduction of the BAID dataset in March 2023. This dataset solely comprises overall scores for high-quality artistic images. Our research marks the pioneering introduction of a multi-attribute, multi-category dataset specifically tailored to the field of painting: Aesthetics of Paintings and Drawings Dataset (APDD). The construction of APDD received active participation from 28 professional artists worldwide, along with dozens of students specializing in the field of art. This dataset encompasses 24 distinct artistic categories and 10 different aesthetic attributes. Each image in APDD has been evaluated by six professionally trained experts in the field of art, including assessments for both total aesthetic scores and aesthetic attribute scores. The final APDD dataset comprises a total of 4985 images, with an annotation count exceeding 31100 entries. Concurrently, we propose an innovative approach: Art Assessment Network for Specific Painting Styles (AANSPS), designed for the assessment of aesthetic attributes in mixed-attribute art datasets. Through this research, our goal is to catalyze advancements in the field of aesthetic evaluation for paintings and drawings, while enriching the available resources and methodologies for its further development and application. Dataset is available at <https://github.com/BestiVictory/APDD.git>

1 Introduction



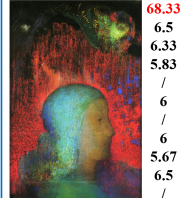
<p>Score</p> <p>Theme and Logic 41</p> <p>Creativity 5.6</p> <p>Layout and composition 4.6</p> <p>Space and Perspective 4</p> <p>Sense of Order /</p> <p>Light and Shadow 5.6</p> <p>Color /</p> <p>Details and Texture 4.4</p> <p>Overall 4</p> <p>Mood 5.2</p> <p>4.2</p>		<p>85.83</p> <p>/</p> <p>8</p> <p>8.17</p> <p>/</p> <p>8.5</p> <p>/</p> <p>8.17</p> <p>8.33</p> <p>/</p>
Artistic Category	Oil painting - Romanticism - Landscapes	Sketching - Symbolism - Portraiture
<p>77.5</p> <p>6.83</p> <p>/</p> <p>7.33</p> <p>5.83</p> <p>6.67</p> <p>7.5</p> <p>/</p> <p>7.17</p> <p>7.67</p> <p>5.67</p>		<p>22.5</p> <p>2.17</p> <p>3</p> <p>2</p> <p>/</p> <p>2</p> <p>/</p> <p>2</p> <p>1.5</p> <p>2.5</p> <p>2.17</p>
Sketching - Classicism - Still life	Traditional Chinese painting - Meticulous - Portraiture	Traditional Chinese painting - Freehand - Floral and Avian
<p>58.33</p> <p>5.5</p> <p>5.17</p> <p>5.67</p> <p>/</p> <p>5.83</p> <p>/</p> <p>5.33</p> <p>5.5</p> <p>/</p>		<p>32.5</p> <p>4.17</p> <p>3.33</p> <p>3.17</p> <p>/</p> <p>3.67</p> <p>/</p> <p>3.17</p> <p>2.83</p> <p>3.5</p> <p>3.17</p>
Sketching - Romanticism - Portraiture	Oil painting - Romanticism - Portraiture	Traditional Chinese painting - Freehand - Mountains and Water

Figure 1: Samples from the APDD dataset. APDD covers 24 artistic categories and 10 aesthetic attributes. Different artistic categories correspond to different sets of attributes.

Computational aesthetics [Datta *et al.*, 2006] aims to enable computers and robots to recognize, generate, and create beauty. In related research, computational visual aesthetics [Brachmann and Redies, 2017] primarily involves training large datasets to acquire neural network models, enabling the models to provide evaluations of aesthetic quality. Consequently, the construction of benchmark datasets for Image

*Corresponding author: Heng Huang(hecate@mail.ustc.edu.cn)

Field	Dataset	Number of images	Number of attributes	Number of categories	IAQA Task Types
Photo Field	AVA [Murray <i>et al.</i> , 2012]	255,530	-	-	Aesthetics Score
	AADB [Kong <i>et al.</i> , 2016]	10,000	11	-	Aesthetic Attributes
	PCCD [Chang <i>et al.</i> , 2017]	4,235	7	-	Aesthetic Captions
	AROD [Schwarz <i>et al.</i> , 2018]	380,000	-	-	Aesthetics Score
	AVA-Reviews [Wang <i>et al.</i> , 2019]	52,118	-	-	Aesthetic Captions
Art Field	EVA [Kang <i>et al.</i> , 2020]	4,070	4	6	Aesthetic Attributes
	BAID [Yi <i>et al.</i> , 2023]	60,337	-	-	Aesthetics Score
Art Field	VAPS [Fekete <i>et al.</i> , 2022]	999	5	5	
	JenAesthetics [Amirshahi <i>et al.</i> , 2015]	1,268	5	16	
	JenAesthetics β [Amirshahi <i>et al.</i> , 2016]	281	1 (beauty)	16	Aesthetic Attributes
	MART [Yanulevskaya <i>et al.</i> , 2012]	500	1 (emotion)	-	
	APDD (Ours)	4,985	10	24	

Table 1: A comparison between the APDD dataset and existing image datasets.

Aesthetic Quality Assessment (IAQA) has become a crucial prerequisite for advancing research in this direction. However, existing datasets predominantly focus on total aesthetic scores of images, with limited exploration in the study of image categories and aesthetic attributes. Moreover, the majority of existing datasets are concentrated in the field of photo, with sparse representation in the field of artistic images.

Existing artistic image aesthetic datasets (such as BAID [Yi *et al.*, 2023], VAPS [Fekete *et al.*, 2022], and JenAesthetics [Amirshahi *et al.*, 2015]) suffer from limitations in both image quantity and the absence of annotations for aesthetic attributes. This deficiency results in a weak interpretability of overall aesthetic scores, rendering them less persuasive. Additionally, these images predominantly represent either iconic works from the history of painting art or high-quality competition entries, lacking a representation of medium to low-quality painting works. Furthermore, these datasets predominantly focus on oil painting as the sole painting form, neglecting the diverse and extensive nature of painting as an art form, which encompasses various classifications. A singular dataset composed solely of high-quality oil paintings lacks diversity, making it unsuitable for multi-task learning and challenging to address the complexities of real-world scenarios.

To address the deficiencies in existing artistic image datasets, such as insufficient aesthetic attributes, uneven aesthetic quality, and limited artistic categories, we leveraged the knowledge and professional expertise of 28 professional artists in the field of painting. Successfully, we constructed the first-ever multi-attribute, multi-category dataset specifically tailored for the field of painting: the Aesthetics of Paintings and Drawings Dataset (APDD). APDD is structured into 24 distinct artistic categories based on different painting categories, artistic styles, and subject matter. Additionally, we selected 10 aesthetic attributes for APDD, including theme and logic, creativity, layout and composition, space and perspective, sense of order, light and shadow, color, detail and texture, overall, and mood. Based on the characteristics of different artistic categories, we selectively chose distinct sets of aesthetic attributes tailored to each artistic category. Additionally, we introduce a paintings and drawings assessment network named Art Assessment Network for Specific Painting Styles (AANSPS), designed to evaluate aesthetic attributes in mixed-attribute painting datasets.

In conclusion, the main contributions of this paper are as

follows:

Firstly, we propose a clear framework for considering aesthetic components in artistic images, providing a detailed categorization of artistic categories in the field of painting along with their corresponding aesthetic attributes. We also establish scoring criteria for these attributes within different artistic categories.

Secondly, we address the gap in the art field for the lack of a multi-attribute, multi-category painting dataset, introducing the Aesthetics of Paintings and Drawings Dataset (APDD) for the first time.

Lastly, to assess the total aesthetic scores and aesthetic attribute scores of paintings, we propose a painting image evaluation network called AANSPS. We evaluate advanced image aesthetic assessment methods and AANSPS on the APDD dataset. Our model achieves satisfactory results across all metrics, providing clear evidence of the effectiveness of our approach.

2 Related Work

2.1 Image Aesthetic Assessment Datasets

The AVA dataset [Murray *et al.*, 2012] pioneered the use of large-scale images for aesthetic analysis, thereby advancing research in computational aesthetics. In 2015, the JenAesthetics dataset [Amirshahi *et al.*, 2015] was introduced, containing 1,628 colored oil paintings exhibited in museums, with each image annotated for five attribute scores. The AADB dataset [Kong *et al.*, 2016], released in 2016, consists of 10,000 photographic images annotated by five individuals. Released in 2017, the PCCD dataset [Chang *et al.*, 2017] encompasses 4,307 photographic images with annotations providing language comments and aesthetic scores for seven aesthetic attributes. Subsequently, the 2018 release of AROD [Schwarz *et al.*, 2018] and the 2019 release of AVA-Reviews [Wang *et al.*, 2019] expanded the scale of data for image aesthetic scorings, aesthetic classification, and language comments. In 2022, the VAPS dataset [Fekete *et al.*, 2022] featured 999 representative works from the history of painting art. In 2023, the BAID dataset [Yi *et al.*, 2023] was introduced, comprising 60,337 high-quality artistic images, each associated with an overall aesthetic score.

Among all these datasets, the majority of images are concentrated in the field of photo, with relatively fewer artistic

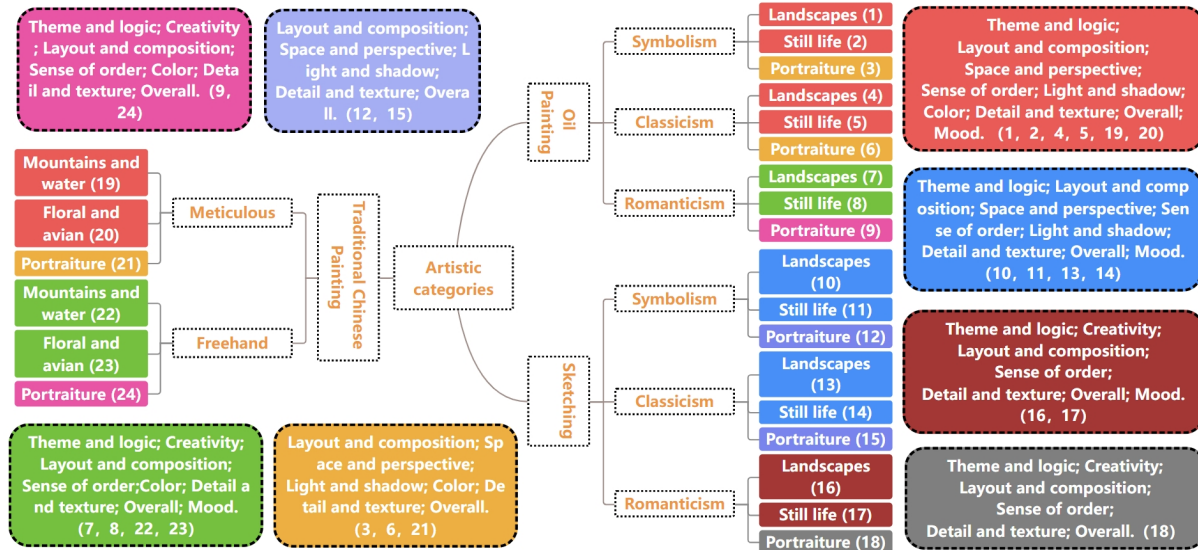


Figure 2: Correspondence between artistic categories and aesthetic attributes. The tree structure in the center represents the image categories, and the attribute sets are represented in the rectangular boxes on either side.

images. Furthermore, most datasets do not adequately classify images, and the attribute types in existing artistic image datasets cannot comprehensively reflect the fundamental characteristics of paintings. A detailed comparison of our APDD with existing datasets is provided in Table 1.

2.2 Image Aesthetic Assessment Models

Due to the powerful feature learning capability of deep networks, there has been an emergence of methods in recent years for predicting aesthetic attributes of images. Kong et al. (2016) [Kong *et al.*, 2016] introduced an attribute-adaptive deep convolutional neural network for aesthetic score prediction, providing evaluations of aesthetic attributes simultaneously with aesthetic score prediction. Malu et al. (2017) [Malu *et al.*, 2017] jointly learned aesthetic scores and aesthetic attributes using a deep convolutional network with merging layers. Building upon this, Pan et al. (2019) [Pan *et al.*, 2019] explored the inherent joint distribution of real aesthetic scores and attributes through adversarial learning to further enhance the prediction accuracy of aesthetic attributes and scores. Li et al. (2022) [Li *et al.*, 2022] proposed a deep multi-task convolutional neural network (MTCNN) model that leverages scene information to assist in predicting aesthetic attributes of images. Jin et al. (2023) [Jin *et al.*, 2023] presented a method for image aesthetic attribute assessment that achieves aesthetic classification, overall score, and scores for three attributes: light, color, and composition.

The aforementioned image aesthetic attribute evaluation methods are all proposed based on photographic images. In the field of artistic images, the AANSPS network proposed in this paper is capable of simultaneously extracting the total aesthetic scores and aesthetic attribute scores.

3 Aesthetics of Paintings and Drawings Dataset

3.1 Formation of Professional Artistic Team

In order to ensure the professionalism and authority of the developed painting dataset, we established a team of painting experts. These team members bring to the table extensive aesthetic experience and educational backgrounds garnered from art institutions worldwide. Their invaluable assistance encompassed the selection of artistic categories, identification of aesthetic attributes, the addition of paintings, and the scoring and annotation of images within the dataset. Specifically, the team is composed of artists, educators, and master’s degree candidates specializing in oil painting, sketching, and Chinese painting, spanning both domestic and international contexts. Stringent criteria were established for all team members, requiring: 1) possession of a bachelor’s degree or higher; 2) a minimum of 7 years of art education experience to ensure a profound artistic background; 3) reception of professional education in oil painting, sketching, or Chinese painting, along with the ability to evaluate artworks based on scoring standards; 4) affiliation with a professional art institute to guarantee that team members originate from institutions with robust artistic education backgrounds; 5) involvement in experiences such as exhibiting works in art exhibitions or publishing articles in professional journals.

Following these criteria, we have successfully assembled a team of specialized dataset builders, comprising 28 professional artists and 24 students with high academic qualifications. The team was segmented into a team leader (1 person), an oil painting group (19 people), a sketching group (19 people), and a Chinese painting group (13 people), with the most artistically adept member serving as the leader in each group. This team composition not only ensures a wealth of professional knowledge and extensive experience but also provides

a more nuanced artistic perspective for scoring and annotation activities.

3.2 Artistic Categories

According to different painting category (oil painting, sketching, and traditional Chinese painting), artistic styles (symbolism, classicism, romanticism, meticulous, and freehand), and subject matter (landscapes, still life, portraiture, floral and avian, mountains and water), we have categorized the APDD dataset into 24 distinct artistic categories.

Painting encompasses various types, each distinguished by its unique tools, materials, and techniques. These include oil paintings, sketches, ink drawings, traditional Chinese painting, prints, watercolors, gouache, murals, and other forms [Wang, 2010]. After in-depth research and comparison, we have chosen sketching, oil painting, and traditional Chinese painting as the fundamental categories of painting types, owing to their long-standing historical significance, widespread dissemination, and popularity among the general public.

Concerning the classification of artistic styles, we have broadly divided them into two main categories: Western art styles represented by sketching and oil painting, and Eastern art styles represented by traditional Chinese painting.

In terms of Western artistic styles, the classification is primarily based on the categories proposed by the eminent 18th-century German philosopher and aesthetician, Hegel. According to his theory, artistic styles represent different stages of idealized development, encompassing primitive symbolism, classical art (such as Greco-Roman classical art), and modern romanticism [Hegel, 2016]. Symbolism (primitive art) is characterized by a dominance of material elements, classicism (Greco-Roman) harmonizes material and spiritual factors, while romanticism (modern art) emphasizes the predominance of spiritual elements over material aspects.

In classifying Eastern art styles, the categorization is grounded in the techniques employed in traditional Chinese painting. Traditional Chinese painting techniques are primarily delineated into meticulous painting and freehand painting. Meticulous painting emphasizes exacting and precise techniques for representing objects, while freehand painting places greater emphasis on brushwork to convey the artist's emotions and artistic expression, underscoring the unity of form and spirit.

From the perspective of subject matter, painting can be categorized into various genres, including portraiture, landscape, still life, wildlife, historical themes, and genre painting. For sketching and oil painting, we have selected the three most common subject matter categories: portraiture, landscapes, and still life, as the basis for classification. Traditional Chinese painting can be broadly categorized into three major genres: portraiture, mountains and water painting, and floral and avian painting.

3.3 Aesthetic Attributes

With the assistance of the professional artistic team and guidance from art instructional materials [Yi Lu, 2021] [Dodson, 1990] [Edwards, 1997] [Gombrich and Gombrich, 1995], we

have compiled the attributes for aesthetic scoring, with the specific process outlined as follows:

Summarizing the general thought process of artistic creators:

Thinking (Creativity and Imagination) → Composition (Two-dimensional Plane) → Perspective (Three-dimensional Space) → Structure (Internal Composition) → Color (Emotional Atmosphere) → Texture (Detail).

Summarizing how art observers conduct layered observations:

Outline → Structure → Texture → Lighting.

Referring to the grading criteria for the fundamental painting exam, summarizing the evaluators' scoring approach:

Content (closely related to the theme) → Creativity (innovative design and creation) → Layout → Logic (consistent with artistic technique and theme, appropriate layout and color coordination) → Form → Color → Details → Overall.

Based on this, we summarize 10 aesthetic attributes for art images, namely, theme and logic, creativity, layout and composition, space and perspective, sense of order, light and shadow, color, detail and texture, overall, and mood.

Theme and logic require that the central idea and main content to be expressed in the work conform to the theme [Bell, 2015].

Creativity requires that the work possesses creativity and imagination, and that the design is unique and able to break the traditional rules.

Layout and composition refer to the overall formal and structural relationship of the picture and its visual effect. Layout is the underlying logic of composition, and composition is the appearance of layout.

Space and perspective indicate the layering of space and the contrast between distance and nearness. Perspective is the basic factor in painting to express the three-dimensionality of objects and create spatial effects.

The sense of order is one of the important manifestations of the elements of perfection, coordination and wholeness of visual form, the core of which is some kind of consistency existing between the elements [Hogarth, 1986].

The change of **Light and shadow** makes the picture more visually rhythmic.

Color can render emotional atmosphere, reasonable color matching can make the picture achieve better visual effect.

Details and texture require a high degree of completion of the picture, with specific and vivid details, fine texture, to enhance the sense of reality of the picture.

The overall requirements of the picture in the overall effect of coordination, clear theme [Dana, 2019].

Mood refers to the poetic space that presents the blending of scene and reality, and the rhythm of active life [Li, 2001].

3.4 Correspondence between Artistic Categories and Aesthetic Attributes

Just as sketching lacks the attribute of color, not all artistic categories encompass the 10 aesthetic attributes proposed in this paper. The correspondence between artistic categories and aesthetic attributes is shown in Figure 2.

Artistic styles of oil painting include symbolism (1, 2, 3), classicism (4, 5, 6), and romanticism (7, 8, 9). Symbolism (1,

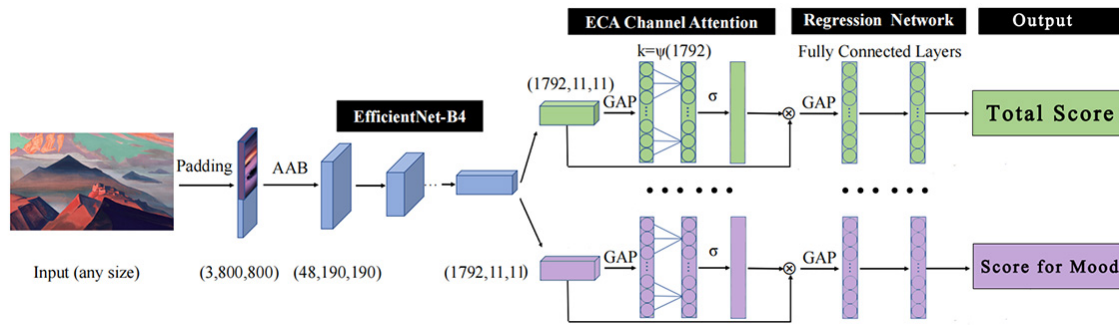


Figure 3: Overall architecture of the proposed AANSPS.

2, 3) can be understood as an objective reproduction, ignoring the attribute of *creativity*, while other aesthetic attributes are more demanding. Classicism (4, 5, 6) can be understood as the unification of material and spiritual factors, i.e., *both form and spirit*, and the attribute of *creativity* can be ignored. Among them, 3 and 6 are portraiture, in which *creativity* can be ignored along with *theme and logic*, *sense of order*, and *mood*. Romanticism (7, 8, 9) can be understood as the spiritual factor exceeds the material factor, and can ignore the attributes of *space and perspective* and *light and shadow*. Among them, 9 is portraiture, one can also ignore *mood*.

Sketching is defined as the depiction of an object’s outline, structure, light and shadow, and texture in monochromatic shades, so its *color* attribute can be ignored. Artistic styles of sketching are categorized as symbolism (10, 11, 12), classicism (13, 14, 15), and romanticism (16, 17, 18), with considerations of attributes similar to those in oil painting across different artistic categories.

Styles of Chinese painting are categorized as meticulous (19, 20, 21) and freehand (22, 23, 24). Meticulous can be understood as the use of neat and rigorous techniques to depict an object, and the attribute of *creativity* can be ignored in meticulous, while other aesthetic attributes are more demanding. Among them, 21 is portraiture, which can ignore *creativity* as well as *theme and logic*, *sense of order* and *mood*. The freehand can be interpreted as the author’s subjective expression of things, and *space and perspective*, *light and shadow* can be ignored. Among them, 24 is portraiture, and *mood* can also be ignored.

3.5 Collection of Paintings and Drawings

To construct a comprehensive and diverse collection of paintings, we carefully selected several professional art websites and institutions as data sources to ensure the breadth and diversity of artistic images. Specifically, the majority of oil paintings and sketches in the dataset are sourced from WikiArt [Phillips and Mackintosh, 2011], while most Chinese painting works come from the China Artists Association¹. Additionally, we supplemented the dataset with images from Baidu Image². The sources of all images will be documented in the dataset’s annotation files. The aesthetic quality of the paintings downloaded from these websites is

¹<https://caagov.com/>

²<https://image.baidu.com/>

Attribute	Layout and composition	Space and Perspective	Light and Shadow	Color	Details and Texture	Overall	Total Aesthetic Score
Score							
100							
80							
60							
40							
20							

Figure 4: Representative Images for the category of oil painting - symbolism - portraiture

concentrated in the intermediate and high ranges. To ensure the diversity and representativeness of the dataset, we introduced some artworks with lower aesthetic quality from the assignments of art students. Additionally, we adopted a strategy with a 3:1 ratio between works by professional artists and student assignments. This maintains the dominance of professional-level works in the dataset while also adequately considering the proportion of lower-scoring works. In the end, we successfully collected 4,985 paintings, covering 24 categories, with each category containing a minimum of 200 images.

3.6 Image Annotation Process

Recognizing the advantages of professional artists in terms of experience and objectivity, we divided the scoring process into two phases. The first phase involved 27 professional artists from the oil painting, sketching, and Chinese painting groups. Their task was to lay the foundation for the entire scoring activity. During this phase, they collaboratively established scoring criteria for each image category and aesthetic attribute, presenting them in the form of benchmark images, as shown in Figure 4, illustrating a representative image example for the oil painting - symbolism - landscape cate-

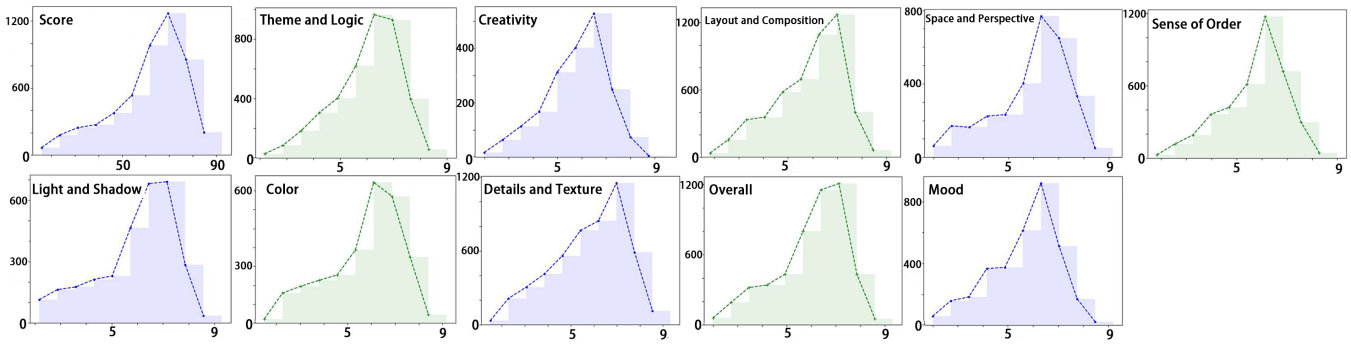


Figure 5: The score distribution of APDD. This distribution suggests a consensus on the aesthetic appeal of images, with most perceived as having a moderate level of appeal, while extreme scores are less frequent.

gory. Based on the selected benchmark images, the scoring team developed a consistent and objective scoring system, ensuring subsequent scoring work could adhere to unified standards. The first phase spanned 15 days, providing a solid theoretical foundation for the scoring activity.

The second phase involved 24 students with high academic qualifications from the oil painting, sketching, and Chinese painting groups. Leveraging the insights gained from the first phase, the second phase progressed more rapidly, completing in a total of 7 days. The division into these two phases not only maximized the use of professional artists’ experience but also allowed the student scoring team to efficiently complete the scoring task under the standards and guidance accumulated in the earlier phase.

In the end, we collected a total of 31,100 annotation records from 51 professional annotators, averaging 6.24 annotations per image. Each annotation record included the total aesthetic scores and aesthetic attribute scores of the corresponding painting work. Throughout the annotation process, we meticulously assigned tasks to each scorer, specifying the art categories and the number of images involved, ensuring that each image in the APDD dataset was evaluated by at least 6 individuals. Through a comprehensive assessment of all scores, we calculated average scores based on the evaluations of all annotators, ultimately gathering the final total scores and attribute scores for each painting work.

4 Art Assessment Network for Specific Painting Styles

4.1 Network Architecture

The entire network consists of a backbone network and eleven scoring branch networks, which include one total aesthetic scores evaluation network and ten aesthetic attribute score evaluation networks.

The backbone network extracts features from input images to obtain input feature vectors. The specific process is as follows: First, we resize the longer side of the image proportionally to 800 pixels and pad the shorter side with zeros to achieve a size of 800 pixels. Next, within the initial convolutional operation of EfficientNet-B4, the $48 \times 190 \times 190$ feature vector is reshaped using an adaptive average pooling layer, with 48 representing the number of feature chan-

nels. Lastly, the transformed feature vector undergoes further processing through the subsequent layers of the EfficientNet-B4 network, yielding an input feature vector of dimensions $1792 \times 11 \times 11$.

The scoring branch network comprises a channel attention module and regression network. In order to reduce the model’s complexity, this paper cites the Efficient Channel Attention (ECA) module [Wang *et al.*, 2020]. It employs a local cross-channel attention mechanism and adaptive 1D convolutional kernels, which can significantly enhance performance. The ECA module aggregates the input feature vector using global average pooling (GAP), resulting in 1D vector where each channel’s value represents the global average of features in that channel. Subsequently, the ECA module conducts rapid 1D convolution operations on this vector with the kernel size of k , which dynamically adapts to the channel dimension C to accommodate varying channel counts. Following this, the resulting channel weights undergo Sigmoid activation, transforming them into values ranging from 0 to 1, denoting the importance or activation level of each channel. These activated channel weights are then utilized to fuse with the input feature vector, ultimately producing the fused feature vector.

The specific size k mentioned above is formulated as follows:

$$k = \psi(c) = \left\lfloor \frac{\log_2 C}{\gamma} + \frac{b}{\gamma} \right\rfloor_{\text{odd}}$$

$\lfloor \cdot \rfloor_{\text{odd}}$ indicates the operation of rounding to the nearest odd number. The value of γ is 2, and the value of b is 1. We can get $k = 7$ if $C = 1792$.

The Regression Network consists of a GAP and three linear layers. Specifically, the fused feature vector is input into the GAP to obtain dimensionality-reduced feature. Subsequently, the dimensionality-reduced feature is flattened into a 1D vector. This 1D vector is then fed into three linear layers for linear mapping and non-linear transformation, ultimately yielding a corresponding score.

The loss function for the regression network is defined as the mean squared error between the predicted output scores and the ground-truth scores.

Score Type	MTCNN			DCNN			AANSPS		
	MSE ↓	MAE ↓	SROCC ↑	MSE ↓	MAE ↓	SROCC ↑	MSE ↓	MAE ↓	SROCC ↑
Total Aesthetic Score	0.0248	0.1230	0.0561	0.2240	0.1147	-0.1309	0.0149	0.0891	0.6085
Theme and Logic	0.0138	0.0919	-0.0264	0.0135	0.0914	0.0215	0.0143	0.0938	0.3896
Creativity	0.0153	0.1000	-0.0419	0.0154	0.0994	0.0410	0.0164	0.0996	0.4624
Layout and Composition	0.0188	0.1116	0.0453	0.0174	0.1052	0.0947	0.0505	0.1139	0.4508
Space and Perspective	0.0235	0.1266	-0.1146	0.0198	0.1126	0.0628	0.0109	0.0828	0.5288
Sense of Order	0.0163	0.1023	-0.0864	0.0159	0.1011	-0.0212	0.0111	0.0851	0.5433
Light and Shadow	0.0238	0.1249	-0.1038	0.0189	0.1070	-0.0273	0.0118	0.0826	0.6721
Color	0.02373	0.1256	-0.0158	0.0212	0.1198	0.1731	0.0139	0.0890	0.4816
Details and Texture	0.0247	0.1285	0.0110	0.0216	0.1208	-0.0636	0.0139	0.0890	0.5381
Overall	0.0203	0.1141	-0.0319	0.0168	0.1036	-0.0102	0.0122	0.0862	0.4961
Mood	0.0194	0.1133	-0.4448	0.0184	0.1096	0.0578	0.0103	0.0818	0.6284

Table 2: Comparison of MTCNN, DCNN and AANSPS on APDD.

Type of score	Predicted	GT	Predicted	GT	Predicted	GT	Predicted	GT
Total Score	75.04	73.57	73.02	77.86	62.59	65.00	59.06	60.00
Theme and Logic	/	/	/	/	6.21	6.20	5.84	5.67
Creativity	/	/	/	/	5.85	5.60	5.81	5.67
Layout and composition	7.27	7.29	7.49	7.57	6.62	6.20	6.20	5.67
Space and Perspective	7.16	7.29	7.52	7.29	/	/	/	/
Sense of Order	7.08	7.29	7.21	7.14	6.35	6.80	5.80	6.00
Light and Shadow	7.62	7.29	7.46	7.43	/	/	/	/
Color	6.89	6.86	7.32	7.43	6.69	6.60	4.97	5.83
Details and Texture	6.64	7.29	7.18	7.57	5.89	6.40	4.75	5.33
Overall	7.33	7.00	7.17	7.57	6.31	6.60	6.15	5.83
Mood	7.02	6.86	7.08	7.57	5.75	6.40	5.53	5.67

Figure 6: Test samples. *Predicted* represents the predicted score of the AANSPS output. *GT* represents the ground-truth score.

4.2 Training Process

We set the batch size to 64, with a learning rate of 0.0001. Adam optimizer is employed, with beta parameters set to (0.98, 0.999), and weight decay set to 0.0001. If the regression loss does not decrease for two consecutive rounds, the learning rate is multiplied by 0.5. Our experiments were conducted using PyTorch 1.5.0 and Nvidia TITAN XPs. The dataset used in this study is APDD, which is divided into two parts: the training set and the validation set in a 9:1 ratio. We utilized the [Jin *et al.*, 2022] model as the pretraining model.

During training, we load the pretraining model into the total aesthetic score branch network and train this branch network based on the training set of APDD to obtain the first scoring model. Then, we use the first scoring model as the pretrained model and individually train the scoring branch networks for each attribute. When training any attribute branch network, the parameters of other scoring branch networks need to be frozen. After each branch network is trained, it will include the previously trained attribute branch networks. Once all attribute branch networks have been trained, we obtain the final scoring model.

5 Experiments

For comparative experiments, we selected MTCNN [Li *et al.*, 2022] and DCNN [Malu *et al.*, 2017], which are capable of simultaneously outputting both aesthetic total scores and aesthetic attribute scores. In the case of MTCNN and DCNN, we duplicated its attribute branches for a total of 10 branches. Following the same training protocol as AANSPS, we used the original model as the pretraining model. After training each branch, the newly generated model parameters from that

branch were used to update the pretraining model. Table 2 presents the comparison of the three models in terms of MSE, MAE, and SROCC.

6 Conclusion

The construction of a multi-attribute, multi-category dataset in the field of painting aesthetics represents a pioneering new task. Through close collaboration with approximately 60 professional artists and students with high academic qualifications, we have successfully established a clear system for considering the aesthetic components of art images. We categorized the field of painting into 24 artistic categories and 10 aesthetic attributes, meticulously defining scoring criteria for corresponding attributes within different artistic categories. We have created the APDD dataset, the most broadly categorized and attribute-rich dataset in the field of painting, comprising 4,985 images of various painting types, accompanied by over 31,100 annotations for both total aesthetic scores and aesthetic attribute scores. Furthermore, based on the APDD dataset, we introduced the AANSPS painting image evaluation network, demonstrating its effectiveness in assessing total and attribute scores for artistic images.

However, given the vast scope of painting as an art field, the categorizations and attributes proposed in our study remain relatively limited. Future work necessitates continuous expansion of artistic categories and aesthetic attributes. Additionally, we plan to increase the number of images in the APDD dataset, supplementing these with more detailed language comments to enhance the interpretability of the scoring results.

Acknowledgments

We thank the ACs and reviewers. This work is partially supported by the Natural Science Foundation of China (62072014&62106118), the Fundamental Research Funds for the Central Universities (3282023014), and the Project of Philosophy and Social Science Research, Ministry of Education of China (20YJC760115).

References

- [Amirshahi *et al.*, 2015] Seyed Ali Amirshahi, Gregor Uwe Hayn-Leichsenring, Joachim Denzler, and Christoph Redies. Jenaesthetics subjective dataset: analyzing paintings by subjective scores. In *Computer Vision-ECCV 2014 Workshops: Zurich, Switzerland, September 6-7 and 12, 2014, Proceedings, Part I 13*, pages 3–19. Springer, 2015.
- [Amirshahi *et al.*, 2016] Seyed Ali Amirshahi, Gregor Uwe Hayn-Leichsenring, Joachim Denzler, and Christoph Redies. Color: A crucial factor for aesthetic quality assessment in a subjective dataset of paintings. *arXiv preprint arXiv:1609.05583*, 2016.
- [Bell, 2015] Clive Bell. *ART*. China Federation of Literature and Art Publishers (CFLP), 2015.
- [Brachmann and Redies, 2017] Anselm Brachmann and Christoph Redies. Computational and experimental approaches to visual aesthetics. *Frontiers in computational neuroscience*, 11:102, 2017.
- [Chang *et al.*, 2017] Kuang-Yu Chang, Kung-Hung Lu, and Chu-Song Chen. Aesthetic critiques generation for photos. In *Proceedings of the IEEE international conference on computer vision*, pages 3514–3523, 2017.
- [Dana, 2019] Dana. *Philosophy of Art*. Zhejiang People’s Fine Arts Press, 2019.
- [Datta *et al.*, 2006] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z Wang. Studying aesthetics in photographic images using a computational approach. In *Computer Vision-ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7-13, 2006, Proceedings, Part III 9*, pages 288–301. Springer, 2006.
- [Dodson, 1990] Bert Dodson. *Keys to drawing*. Penguin, 1990.
- [Edwards, 1997] Betty Edwards. Drawing on the right side of the brain. In *CHI’97 Extended Abstracts on Human Factors in Computing Systems*, pages 188–189. 1997.
- [Fekete *et al.*, 2022] Anna Fekete, Matthew Pelowski, Eva Specker, David Brieber, Raphael Rosenberg, and Helmut Leder. The vienna art picture system (vaps): A data set of 999 paintings and subjective ratings for art and aesthetics research. *Psychology of Aesthetics, Creativity, and the Arts*, 2022.
- [Gombrich and Gombrich, 1995] Ernst Hans Gombrich and EH Gombrich. *The story of art*, volume 12. Phaidon London, 1995.
- [Hegel, 2016] G.W.F. Hegel. *Philosophy of Fine Arts*. Chongqing Publishing House, 2016.
- [Hogarth, 1986] William Hogarth. *Analysis of Beauty*. People’s Fine Arts Press, 1986.
- [Jin *et al.*, 2022] Xin Jin, Hao Lou, Heng Huang, Xinning Li, Xiaodong Li, Shuai Cui, Xiaokun Zhang, and Xiqiao Li. Pseudo-labeling and meta reweighting learning for image aesthetic quality assessment. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):25226–25235, 2022.
- [Jin *et al.*, 2023] Xin Jin, Xinning Li, Hao Lou, Chenyu Fan, Qiang Deng, Chaoen Xiao, Shuai Cui, and Amit Kumar Singh. Aesthetic attribute assessment of images numerically on mixed multi-attribute datasets. *ACM Transactions on Multimedia Computing, Communications and Applications*, 18(3s):1–16, 2023.
- [Kang *et al.*, 2020] Chen Kang, Giuseppe Valenzise, and Frédéric Dufaux. Eva: An explainable visual aesthetics dataset. In *Joint workshop on aesthetic and technical quality assessment of multimedia and media analytics for societal trends*, pages 5–13, 2020.
- [Kong *et al.*, 2016] Shu Kong, Xiaohui Shen, Zhe Lin, Radomir Mech, and Charless Fowlkes. Photo aesthetics ranking network with attributes and content adaptation. In *Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, pages 662–679. Springer, 2016.
- [Li *et al.*, 2022] Leida Li, Jiachen Duan, Yuzhe Yang, and Yaqian Li. Scene-assisted image aesthetic attribute assessment. *Journal of Image and Graphics*, 27(11):3199–3209, 2022.
- [Li, 2001] Zehou Li. *The Journey of Beauty*. Tianjin Academy of Social Science Press, 2001.
- [Malu *et al.*, 2017] Gautam Malu, Raju S Bapi, and Bipin Indurkha. Learning photography aesthetics with deep cnns. *arXiv preprint arXiv:1707.03981*, 2017.
- [Murray *et al.*, 2012] Naila Murray, Luca Marchesotti, and Florent Perronnin. Ava: A large-scale database for aesthetic visual analysis. In *2012 IEEE conference on computer vision and pattern recognition*, pages 2408–2415. IEEE, 2012.
- [Pan *et al.*, 2019] Bowen Pan, Shangfei Wang, and Qisheng Jiang. Image aesthetic assessment assisted by attributes through adversarial learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 679–686, 2019.
- [Phillips and Mackintosh, 2011] Fred Phillips and Brandy Mackintosh. Wiki art gallery, inc.: A case for critical thinking. *Issues in Accounting Education*, 26(3):593–608, 2011.
- [Schwarz *et al.*, 2018] Katharina Schwarz, Patrick Wieschollek, and Hendrik PA Lensch. Will people like your image? learning the aesthetic space. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 2048–2057. IEEE, 2018.

- [Wang *et al.*, 2019] Wenshan Wang, Su Yang, Weishan Zhang, and Jiulong Zhang. Neural aesthetic image reviewer. *IET Computer Vision*, 13(8):749–758, 2019.
- [Wang *et al.*, 2020] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. Eca-net: Efficient channel attention for deep convolutional neural networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11534–11542, 2020.
- [Wang, 2010] Jianhong Wang. *Introduction to Art*. Culture and Arts Press, 2010.
- [Yanulevskaya *et al.*, 2012] Victoria Yanulevskaya, Jasper Uijlings, Elia Bruni, Andreza Sartori, Elisa Zamboni, Francesca Bacci, David Melcher, and Nicu Sebe. In the eye of the beholder: employing statistical analysis and eye tracking for analyzing abstract paintings. In *Proceedings of the 20th ACM international conference on multimedia*, pages 349–358, 2012.
- [Yi *et al.*, 2023] Ran Yi, Haoyuan Tian, Zhihao Gu, Yukun Lai, and Paul L Rosin. Towards artistic image aesthetics assessment: a large-scale dataset and a new method. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22388–22397, 2023.
- [Yi Lu, 2021] Donghui Cui Yi Lu. *Fundamentals of mold-making*. People’s Fine Arts Press, 2021.