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

Xin Jin\(^1\), Qianqian Qiao\(^1\), Yi Lu\(^3\), Huaye Wang\(^1\), Shan Gao\(^4\), Heng Huang\(^5\)*, Guangdong Li\(^6\)

\(^1\)Beijing Electronic Science and Technology Institute
\(^2\)Beijing Institute for General Artificial Intelligence
\(^3\)Central Academy of Fine Arts
\(^4\)Beijing University of Technology
\(^5\)University of Science and Technology of China
\(^6\)Beijing 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

---

\*Corresponding author: Heng Huang(hecate@mail.ustc.edu.cn)
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 photography, 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 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 AAV-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

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

<table>
<thead>
<tr>
<th>Field</th>
<th>Dataset</th>
<th>Number of images</th>
<th>Number of attributes</th>
<th>Number of categories</th>
<th>IAQA Task Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photo Field</td>
<td>AVA [Murray et al., 2012]</td>
<td>235,390</td>
<td>-</td>
<td>-</td>
<td>Aesthetics Score</td>
</tr>
<tr>
<td></td>
<td>AADB [Kong et al., 2016]</td>
<td>10,000</td>
<td>11</td>
<td>-</td>
<td>Aesthetic Attributes</td>
</tr>
<tr>
<td></td>
<td>PCCD [Chang et al., 2017]</td>
<td>4,235</td>
<td>7</td>
<td>-</td>
<td>Aesthetic Captions</td>
</tr>
<tr>
<td></td>
<td>AROD [Schwarz et al., 2018]</td>
<td>380,000</td>
<td>-</td>
<td>-</td>
<td>Aesthetics Score</td>
</tr>
<tr>
<td></td>
<td>AVA-Reviews [Wang et al., 2019]</td>
<td>52,118</td>
<td>-</td>
<td>-</td>
<td>Aesthetic Captions</td>
</tr>
<tr>
<td></td>
<td>EVA [Kang et al., 2020]</td>
<td>4,070</td>
<td>4</td>
<td>6</td>
<td>Aesthetic Attributes</td>
</tr>
<tr>
<td>Art Field</td>
<td>BAID [Yi et al., 2023]</td>
<td>60,337</td>
<td>-</td>
<td>-</td>
<td>Aesthetics Score</td>
</tr>
<tr>
<td></td>
<td>Jenaesthetics [Amirshahi et al., 2015]</td>
<td>999</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>VAPS [Fekete et al., 2022]</td>
<td>1,268</td>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td></td>
<td>JenAestheticsβ [Amirshahi et al., 2016]</td>
<td>281</td>
<td>1 (beauty)</td>
<td>16</td>
<td>Aesthetic Attributes</td>
</tr>
<tr>
<td></td>
<td>MART [Yanulevskaya et al., 2012]</td>
<td>500</td>
<td>1 (emotion)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>APDD (Ours)</td>
<td><strong>4,985</strong></td>
<td><strong>10</strong></td>
<td><strong>24</strong></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: A comparison between the APDD dataset and existing image datasets.
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, 2, 3)
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 portraitures, 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.

Figures

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 Association. Additionally, we supplemented the dataset with images from Baidu Image. 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 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-
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 48×190×190 feature vector is reshaped using an adaptive average pooling layer, with 48 representing the number of feature channels. Lastly, the transformed feature vector undergoes further processing through the subsequent layers of the EfficientNet-B4 network, yielding an input feature vector of dimensions 1792×11×11.

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:

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

$\lfloor \cdot \rfloor$ indicates the operation of rounding to the nearest odd number. The value of $\gamma$ 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.
### 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 regression loss does not decrease for two consecutive rounds, and 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 datasets: 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 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 aesthetics represents a pioneering new task.

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 for artistic images. 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


