APDDv2: Aesthetics of Paintings and Drawings Dataset with Artist Labeled Scores and Comments

Xin Jin[1,2], Qianqian Qiao[1], Yi Lu[3], Huaye Wang[1],
Heng Huang[4], Shan Gao[5], Jianfei Liu[3], Rui Li[3]

[1]Beijing Electronic Science and Technology Institute, 100070, Beijing, China.
[2]Beijing Institute for General Artificial Intelligence (BIGAI), 100085, Beijing, China.
[3]Central Academy of Fine Arts, 100102, Beijing, China.
[4]University of Science and Technology of China, 230026, Fei He, China.
[5]Beijing University of Technology, 100029, Beijing, China.

Corresponding Author: hecate@mail.ustc.edu.cn

Background

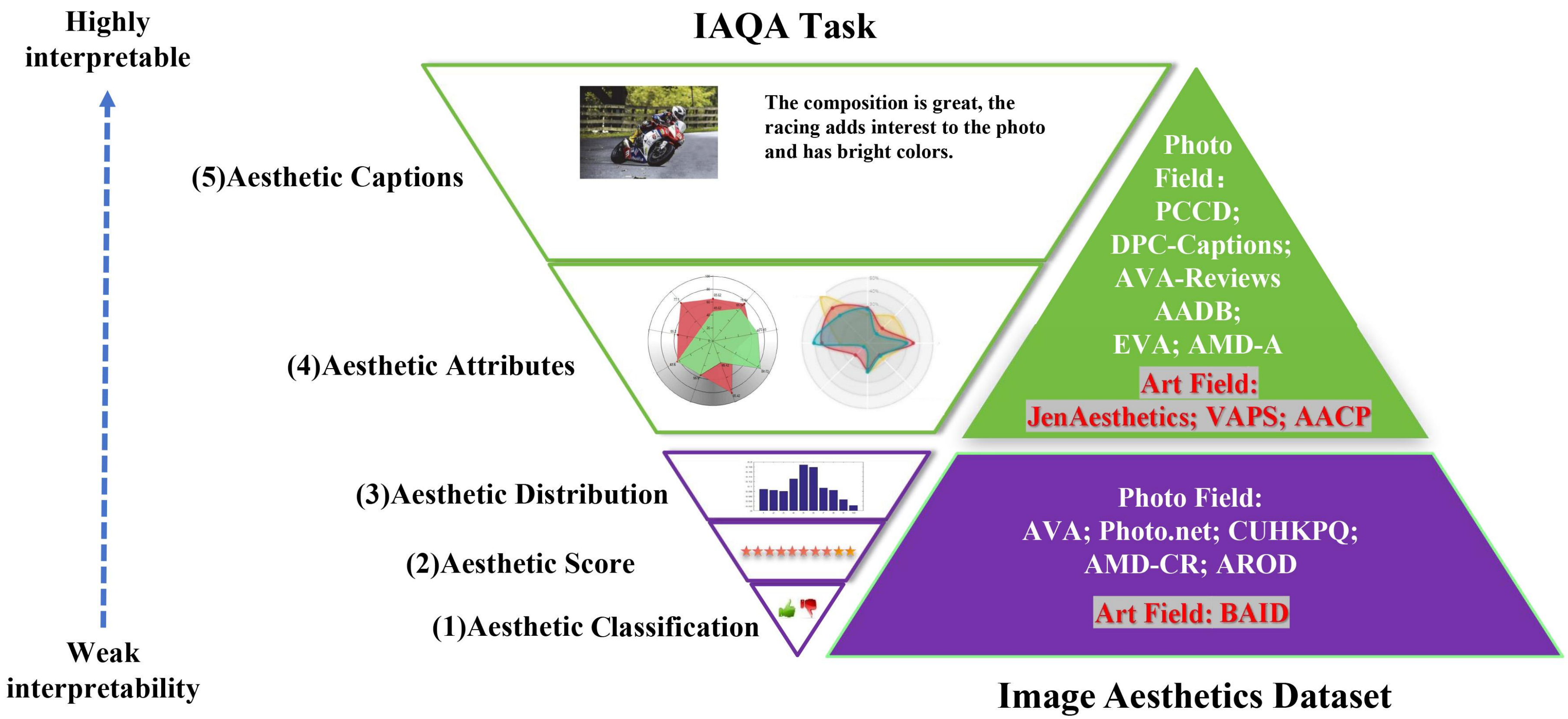


Figure 1: The five-layered tasks of IAQA.

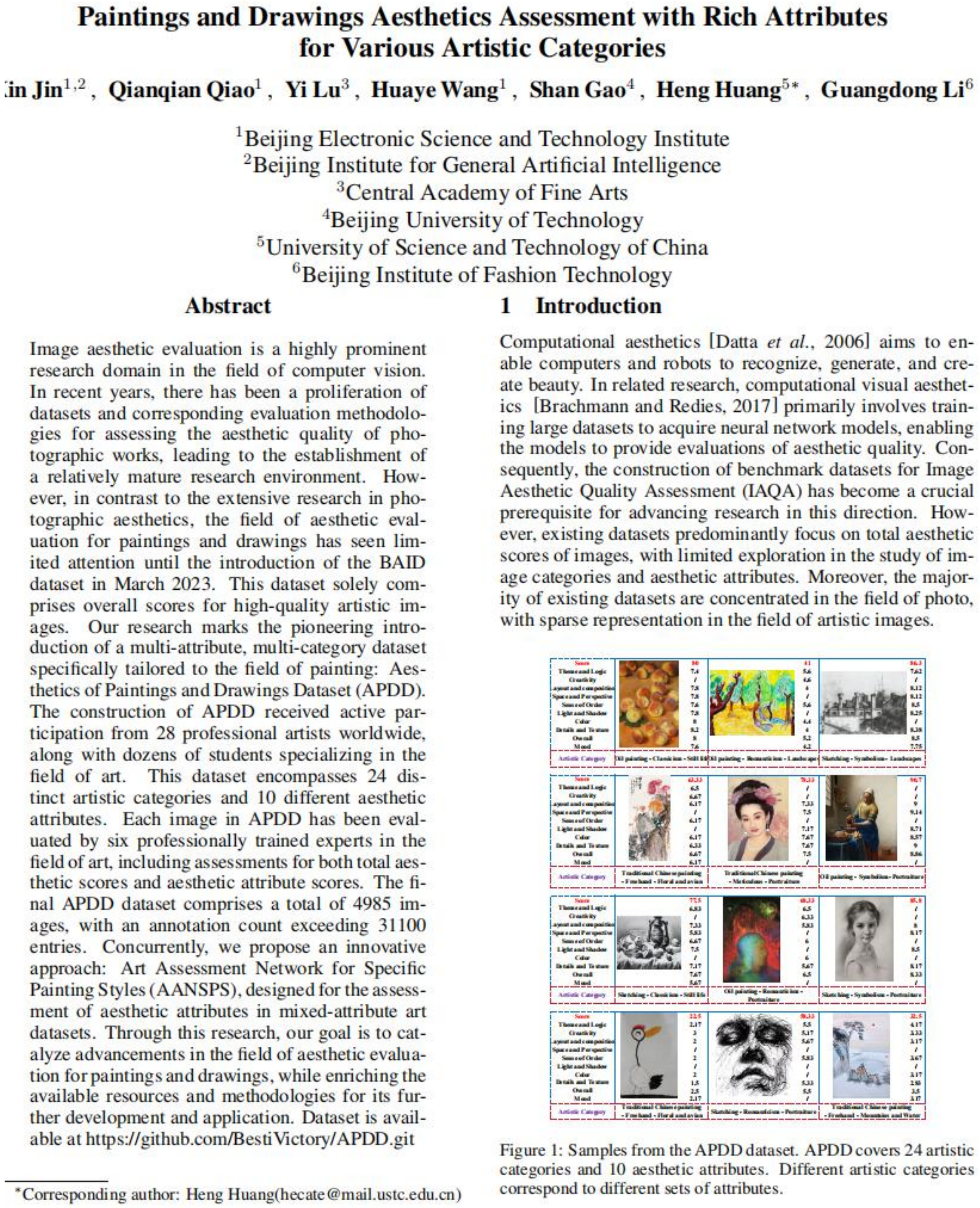
Image Aesthetic Quality Assessment (IAQA) is an important field in computer vision. While photo aesthetic quality has seen extensive research, painting aesthetics evaluation has received much less attention.

Background

Previously, we created the painting dataset APDDv1 [Jin et al., 2024] and the art image scoring model AANSPS.

This work was presented at the IJCAI 2024 conference.

However, both the dataset and the scoring model still have limitations.



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

in Jin[1,2], 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)

1 Introduction

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 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.

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.

Figure 2: Cover of the APDDv1 paper presented at IJCAI 2024.

Contributions

- Evaluation criteria: Developed tailored criteria for assessing artistic images and a scalable dataset expansion method.
- APDDv2 dataset: Introduced a dataset with 10,023 images, 85,191 annotations, and 6,249 comments.
- ArtCLIP network: Developed ArtCLIP using multi-attribute contrastive learning.

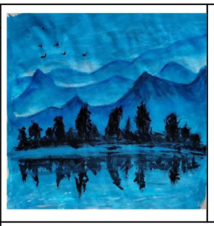
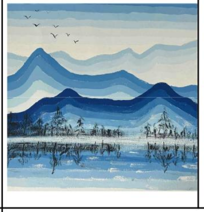
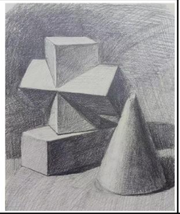
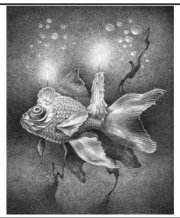
Total Aesthetic Score			38.33		45.83		56.67
Aesthetic Attribute Score	T&L Cre L&C S&P SO L&S Col D&T TO M		4.5 / 4.5 3.83 4.17 4.33 3.67 3.67 4 4.17		4.83 / 4.67 4.5 4.83 4.17 4.17 4 4.67 4.33		/ / 5.5 5.67 / 5.83 5.5 4.67 4.67 /
Language Comment		The picture is not neat enough, the reflection color is too solid and somewhat abrupt, and the colors of the lake, water, and sky are differentiated.		The overall theme is not clear, the picture is too messy, the color is a bit gray, and there are some details of trees and leaves added.		The combination of white and green colors makes the entire picture more clear, but the color connection is poor.	
Artistic Category		Oil Painting-Classicism-Landscapes		Oil Painting-Classicism-Landscapes		Oil Painting-Classicism-Portraiture	
	55 6 / 5.67 5.83 5.67 5.17 / 5.5 5.5 5.5		65 7.33 / 6.67 5.67 6.17 6.67 / 5.67 6.17 5.83		73.33 7.17 / 7 7.33 6.5 7.33 / 7.17 7.33 5.83		89.17 7.83 / 8.5 8.17 8.33 8.33 / 8.5 8 7.83
The composition of the picture is relatively reasonable, but the details are slightly insufficient, lacking light and shadow details, and the picture lacks a primary and secondary relationship.		The geometric shape of the screen performs well, with proper handling of light and dark transitions and a strong sense of three dimensionality. However, the processing of image details is relatively lacking.		The picture has a certain light and shadow effect, the overall architectural drawing is accurate, and spatial perspective can be carefully considered.		The visuals are rich in details, with a certain degree of thematic and imaginative elements, and the visuals are good on both sides.	
Sketching-Symbolism-Landscapes		Sketching-Symbolism-Still Life		Sketching-Classicism-Still Life		Sketching-Classicism-Still Life	
	35.83 4.5 4.67 4 / 4.33 / 4.17 3 4.17 3.67		72.5 7.5 / 7 6.83 6.83 7.17 7.17 7.17 6.83 6.5		80 / / 8 7.67 / 7.67 7.83 7.5 7.5 /		87.5 / / 8.67 8.5 / 8.17 8.33 8.67 8.67 /
The picture is very simple, and the brushstrokes are also very beginner like, with a very immature way of shaping. Lack of detail, lack of depth, and the picture seems to be unfinished.		The bird's posture is not agile enough, the color depiction is somewhat weakened, the plant lines are not smooth enough, and the color gradient depiction in the picture is proficient.		A very delicate painting with delicately crafted characters that are very interesting, and the ink used is very clever.		A very interesting composition method, with delicate and interesting character portrayal, and very clever ink use.	
Traditional Chinese Painting -Freehand-Mountains and Water		Traditional Chinese Painting -Meticulous-Floral and Avian		Traditional Chinese Painting -Meticulous-Portraiture		Traditional Chinese Painting -Meticulous-Portraiture	

Figure 3: Samples from the APDDv2 dataset.

Related Work（Image Aesthetic Assessment Datasets)

Dataset	Number of Images	Number of Attributes	Number of Categories	Any Comment?
BAID [Yi et al., 2023]	60,337	-	-	NO
AACP [Jiang et al., 2024]	21,200	-	-	NO
VAPS [Fekete et al., 2022]	999	5	5	NO
JenAesthetics [Amirshahi et al., 2015]	1,268	5	16	NO
JenAestheticsβ [Amirshahi et al., 2016]	281	1 (beauty)	16	NO
MART [Yanulevskaya et al., 2012]	500	1 (emotion)	-	NO
APDDv1 [Jin et al., 2024]	4,985	10	24	NO
APDDv2	**10,023**	**10**	**24**	**YES**

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

Existing artistic image datasets generally exhibit several limitations:

- Scores rely on user votes, lacking professional expertise;
- Limited variety of painting styles;
- Constrained aesthetic attributes;
- Small dataset size.

Related Work（Artistic Image Aesthetic Assessment Models)

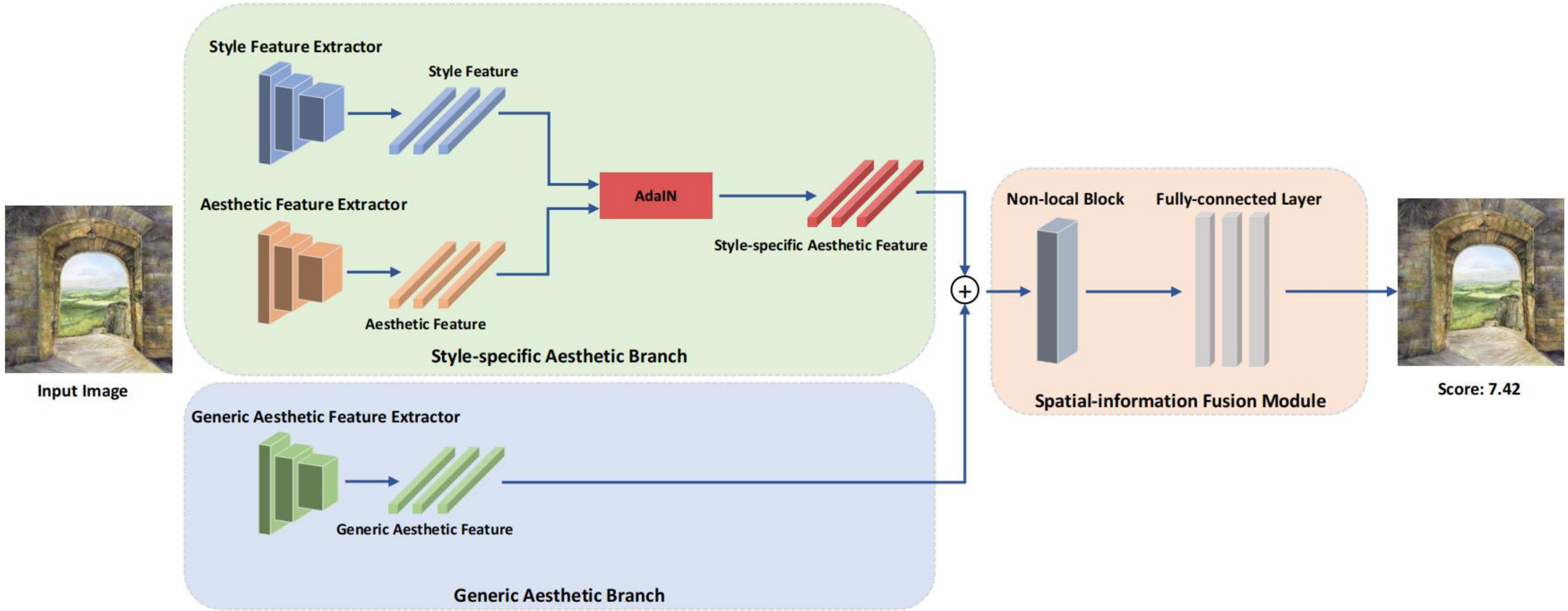


Figure 4: Overall architecture of the SAAN proposed by Yi et al. [Yi et al., 2023].

Related Work（Artistic Image Aesthetic Assessment Models)

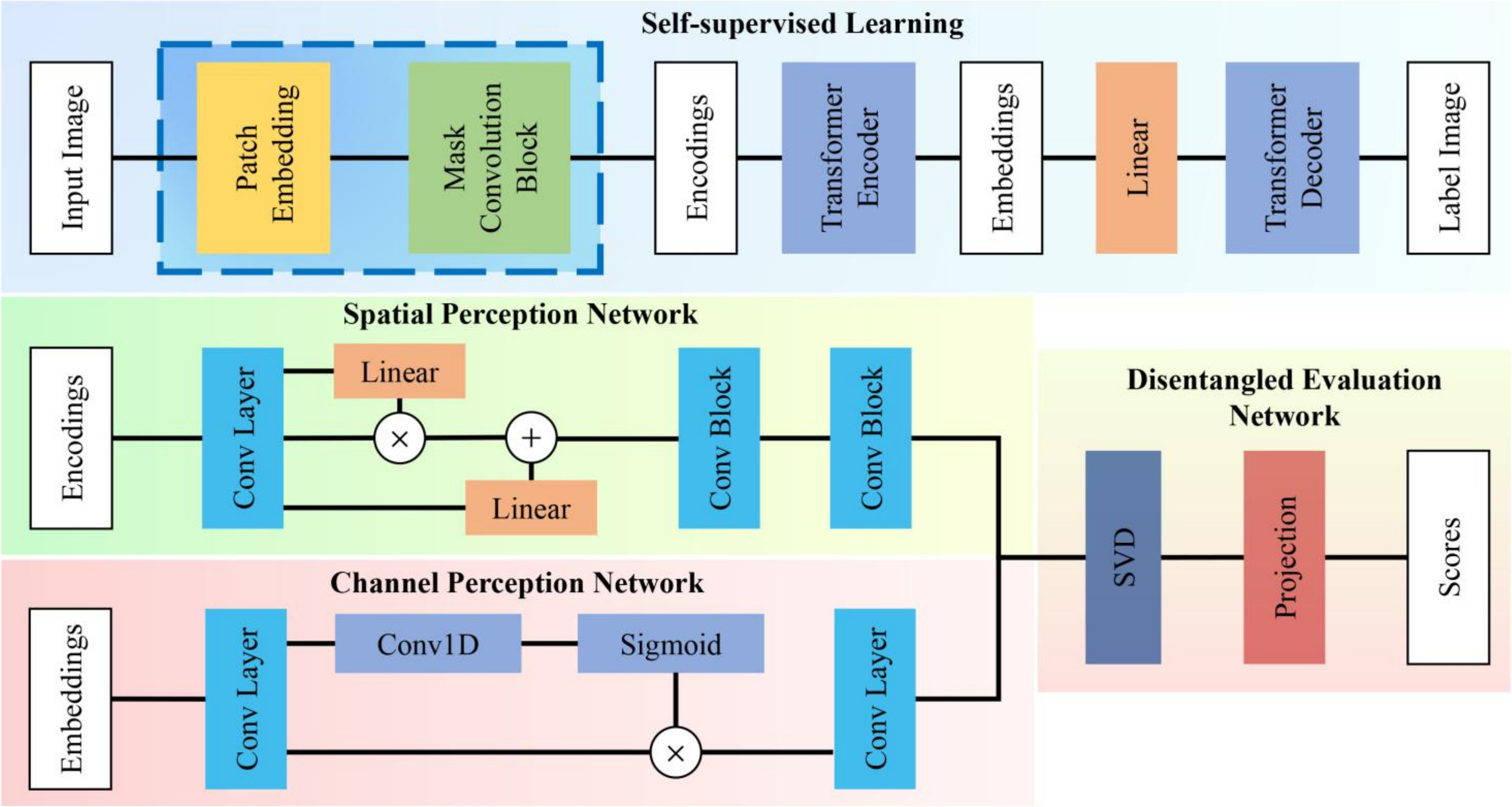


Figure 5: The network architecture for calculating scores on AACP proposed by Jiang et al. [Jiang et al., 2024].

Related Work（Artistic Image Aesthetic Assessment Models)

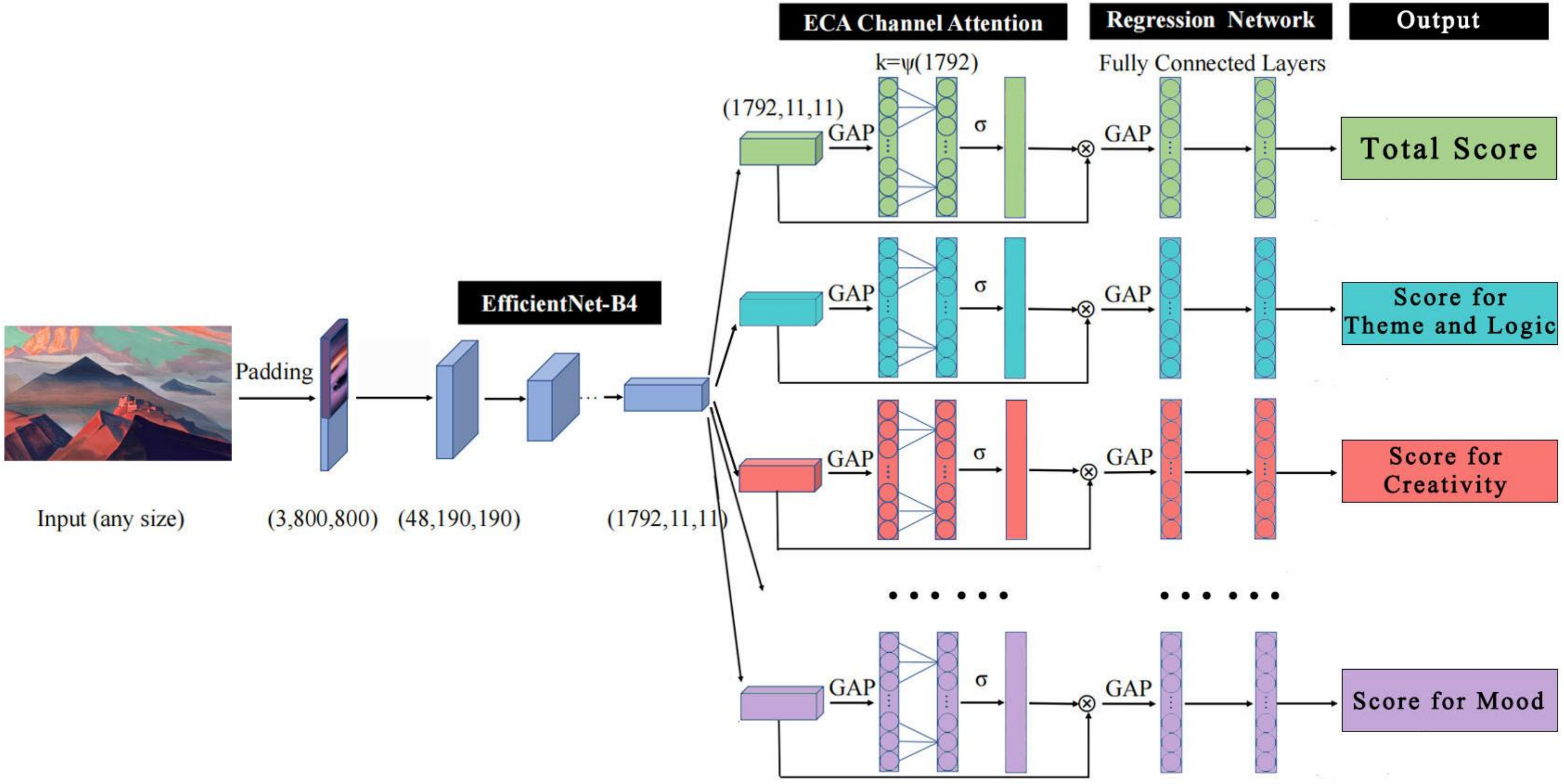


Figure 6: The network architecture of AANSPS proposed by Jin et al. [Jin et al., 2024].

APDDv2 -- 24 Artistic Categories

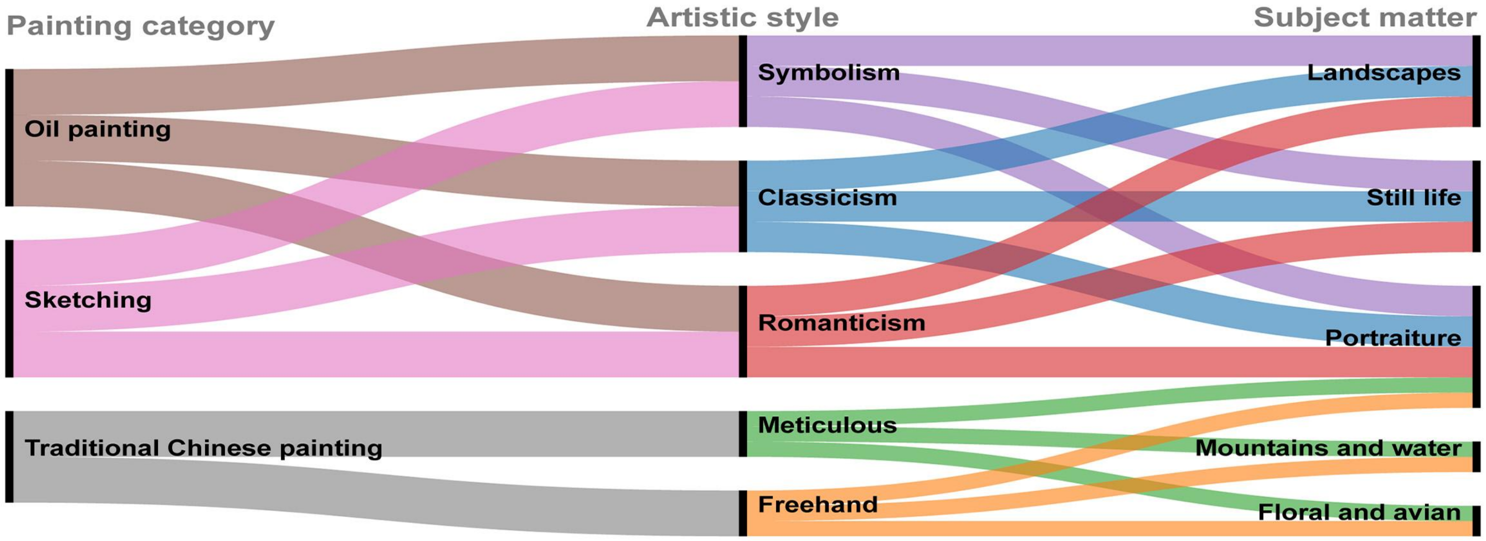


Figure 7: 24 Artistic Categories.

The images in APDD are classified into 24 categories based on:

- **painting category** (Oil Painting, sketching, and Traditional Chinese Painting)
- **artistic styles** (Symbolism, Classicism, Romanticism, Meticulous, and Freehand)
- **subject matter** (Landscapes, Still life, Portraiture, Floral and Avian, Mountains and Water)

APDDv2 -- 10 Aesthetic Attributes

Artistic Category Collection				Aesthetic Attributes									
				T&L	Cre	L&C	S&P	SO	L&S	Col	D&T	TO	M
1	Traditional Chinese painting Oil painting	Meticulous Symbolism Classicism	Floral and avian Portraiture Landscapes Still life Landscapes Still life	√		√	√	√	√	√	√	√	√
2	Oil painting Traditional Chinese painting	Symbolism Classicism Meticulous	Portraiture Portraiture Portraiture			√	√		√	√	√	√	
3	Oil painting Traditional Chinese painting	Romanticism Freehand	Landscapes Still life Mountains and water Floral and avian	√	√	√		√		√	√	√	√
4	Sketching	Symbolism Classicism	Landscapes Still life Landscapes Still life	√		√	√	√	√		√	√	√
5	Sketching	Symbolism Classicism	Portraiture Portraiture			√	√		√		√	√	
6	Oil painting Traditional Chinese painting	Romanticism Freehand	Portraiture Portraiture	√	√	√		√		√	√	√	
7	Sketching	Romanticism	Landscapes Still life	√	√	√		√			√	√	√
8	Sketching	Romanticism	Portraiture	√	√	√		√			√	√	

Figure 8: Correspondence between artistic categories and aesthetic attributes.

10 aesthetic attributes:

- **T&L**: Theme and Logic
- **Cre**: Creativity
- **L&C**: Layout and Composition
- **S&P**: Space and Perspective
- **SO**:Sense of Order
- **L&S**: Light and Shadow
- **Col**: Color
- **D&T**: Detail and Texture
- **TO**: The Overall
- **M**: Mood

APDDv2 -- Construction Process

1. Assemble a team of professional artists.

2. Collect artworks and student drawings.

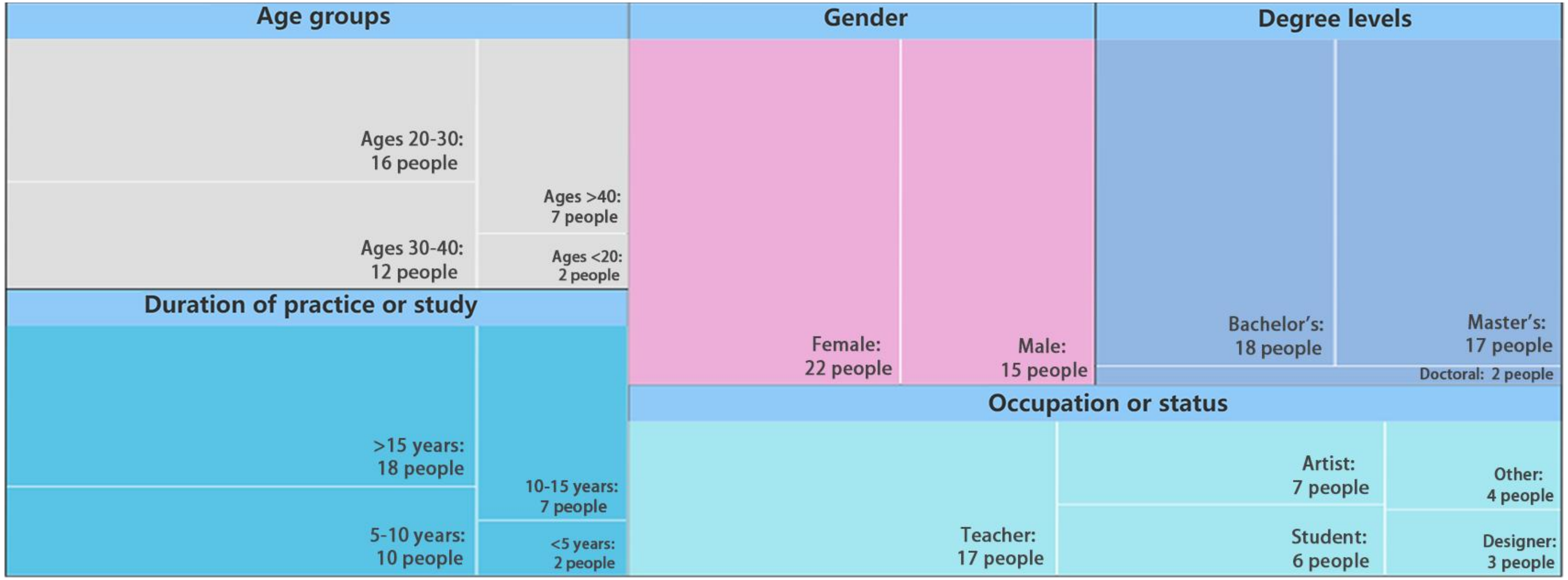


Figure 9: Labeling Team Composition.

APDDv2 -- Construction Process

3. Design the scoring criteria.

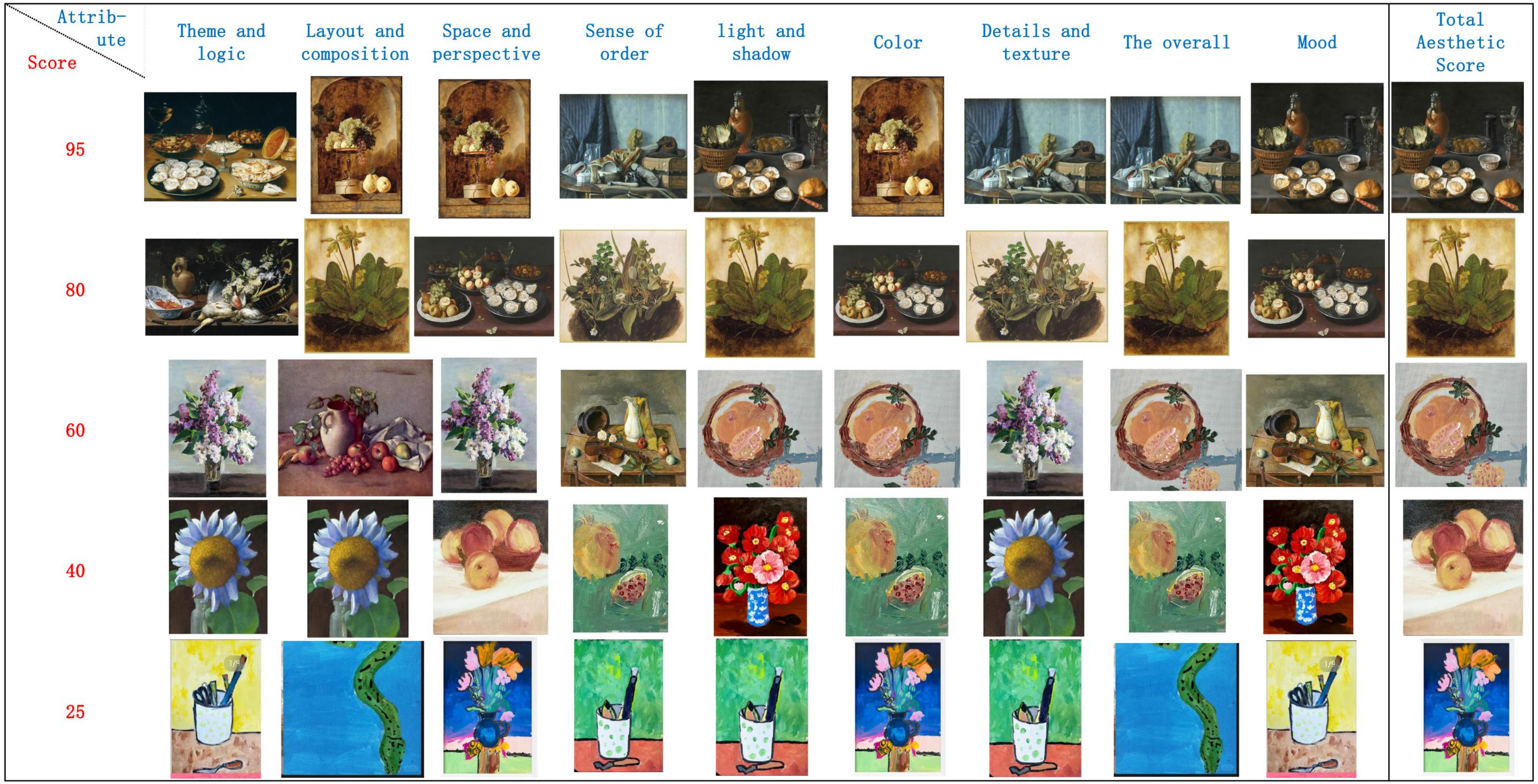


Figure 10: Scoring benchmark table for "Oil Painting - Symbolism - Still Life" category

APDDv2 -- Construction Process

3. Design the scoring criteria.

Score	95	80	60	40	25
Reference Image					
Language Comment	The theme and logic are clear, depicting a corner of the dining table. Under dim lighting, the delectable food on the table appears even more enticing against the dark background and contrasting tableware. The portrayal of items is detailed, evoking a sense of immersion. The rustic colors add to the scene's simplicity. The items on the table are arranged in an orderly yet natural manner, exuding a vibrant sense of life.	The theme is relatively clear, with the arrangement on the dining table appearing somewhat mechanical and rigid. The contrast between the dark tones and the bright colors in the scene is not sufficiently pronounced, resulting in a lack of freshness in the food. The overall composition of the scene is refined.	The purple flowers complement the deep purple tabletop nicely, while the green leaves add a touch of playfulness. However, the overall composition appears somewhat rigid due to the slightly centered arrangement. Although the scene is rich in detail, the handling of the edges seems a bit perfunctory, resulting in a lack of overall cohesion.	The contrast between the blue and yellow is striking, and the green leaves enrich the scene. However, there is a lack of depth in the layers of the leaves. The relationship between the black background and the green leaves seems somewhat forced. The composition and perspective lack aesthetic appeal, and the detailing is inadequate.	The entire scene is filled with childish charm, but the expression appears somewhat clumsy. The colors lack harmony, and there is insufficient depth in the spatial arrangement. The detailing is also lacking.

Figure 11: Benchmark table for language comments in "Oil Painting - Symbolism - Still Life" category

APDDv2 -- Construction Process

4. Develop online labeling system.

5. Divide up the work and then score online.

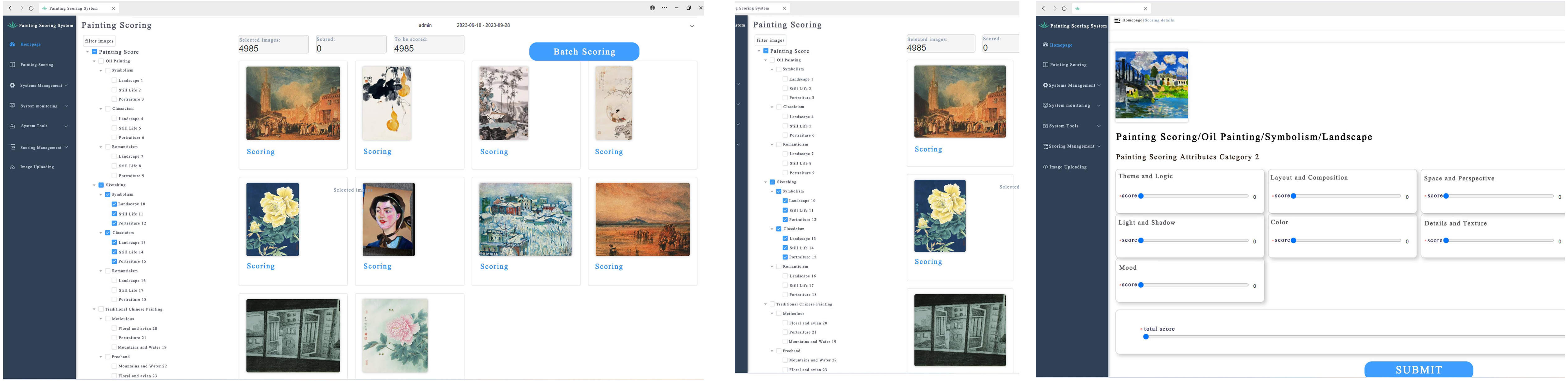


Figure 12: Labeling System Interface

APDDv2 -- Statistical Information

Score type	Total Aesthetic Score	Theme and Logic	Creativity	Layout and composition	Space and Perspective	
pre-averaging	62,790	49,967	24,122	62,790	38,668	
after averaging	10,023	7,965	3,820	10,023	6,205	
Score type	**Sense of Order**	**Light and Shadow**	**Color**	**Details and Texture**	**The Overall**	**Mood**
pre-averaging	49,967	38,644	38,870	62,790	62,790	42,115
after averaging	7,965	6,205	6,202	10,023	10,023	6,737

Table 2: Number of labels for each score type of APDDv2

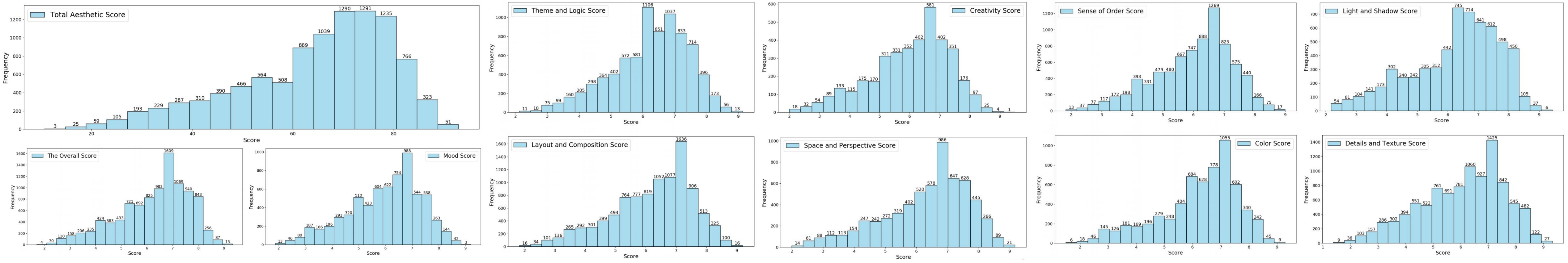


Figure 13: The score distribution of APDDv2

ArtCLIP

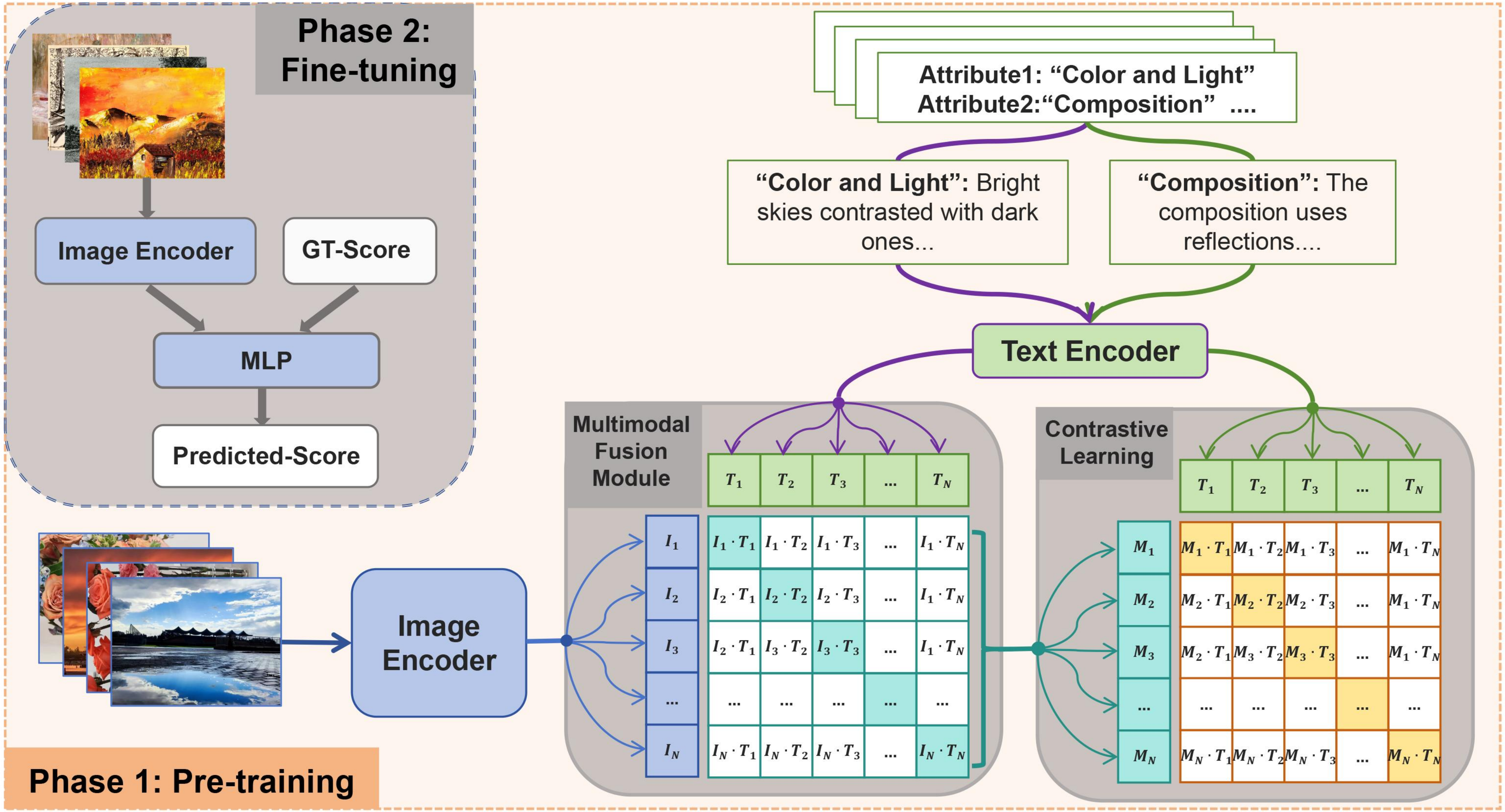


Figure 14: Network architecture of ArtCLIP

Experiments

Score Type	AANSPS					SAAN					ArtCLIP				
	MSE ↓	MAE ↓	SROCC ↑	PLCC ↑	ACC ↑	MSE ↓	MAE ↓	SROCC ↑	PLCC ↑	ACC ↑	MSE ↓	MAE ↓	SROCC ↑	PLCC ↑	ACC ↑
TAS	0.88	0.73	0.76	0.79	0.89	1.79	0.99	0.78	0.61	0.86	**0.68**	**0.63**	**0.81**	**0.84**	**0.89**
T&L	0.73	0.68	0.70	0.72	0.87	1.98	1.07	0.48	0.49	0.83	**0.60**	**0.60**	**0.74**	**0.77**	**0.87**
C	0.81	0.72	**0.71**	0.72	0.85	1.84	1.05	0.48	0.49	0.78	**0.71**	**0.67**	0.74	**0.74**	**0.85**
L&C	0.74	0.68	0.74	0.77	**0.89**	1.49	0.93	0.56	0.58	0.82	**0.63**	**0.61**	**0.77**	**0.80**	0.88
S&P	0.76	0.70	0.72	0.79	0.91	1.60	0.95	0.60	0.63	0.85	**0.60**	**0.61**	**0.79**	**0.83**	**0.91**
SO	0.75	0.68	0.73	0.75	0.87	1.60	0.94	0.52	0.52	0.81	**0.62**	**0.62**	**0.75**	**0.78**	**0.87**
L$S	0.83	0.72	0.73	0.79	0.90	1.67	1.02	0.61	0.65	0.84	**0.65**	**0.63**	**0.79**	**0.84**	**0.91**
Col	0.80	0.70	**0.79**	0.78	0.91	1.76	1.00	0.54	0.60	0.89	**0.59**	**0.59**	0.75	**0.84**	**0.92**
D&T	0.90	0.74	0.76	0.78	0.86	1.62	0.97	0.62	0.62	0.82	**0.70**	**0.65**	**0.81**	**0.83**	**0.88**
O	0.79	0.70	0.73	0.77	0.89	1.35	0.89	0.58	0.62	0.85	**0.63**	**0.62**	**0.78**	**0.81**	**0.89**
M	0.88	0.74	0.71	0.73	0.86	1.83	1.02	0.52	0.53	0.80	**0.71**	**0.67**	**0.75**	**0.78**	**0.85**

Table 3: Comparison of AANSPS [Jin et al., 2024], SAAN [Yi et al., 2023] and ArtCLIP on APDDv2.

Test Samples

Type of score		Predicted	GT		Predicted	GT		Predicted	GT
Total Score		82.10	82.50		75.50	75.80		61.30	61.70
T&L		/	/		6.85	6.83		6.43	6.50
Cre		/	/		/	/		6.00	/
L&C		7.37	7.33		7.03	7.00		6.23	6.17
S&P		7.48	7.50		6.89	7.00		5.82	6.17
SO		/	/		6.97	6.50		6.10	6.17
L&S		7.68	7.67		6.93	6.83		5.83	5.83
Col		/	/		/	/		6.46	6.33
D&T		7.50	7.83		7.09	7.00		6.28	6.17
TO		7.76	7.67		7.13	7.17		6.28	6.00
M		/	/		6.85	6.83		6.30	6.17
Total Score		60.00	54.20		33.00	34.20		21.90	20.60
T&L		/	/		4.28	4.17		/	/
Cre		/	/		3.99	4.17		/	/
L&C		5.69	6.33		3.43	4.00		2.45	3.25
S&P		5.34	5.00		/	/		2.46	2.62
SO		/	/		3.77	3.38		/	/
L&S		5.27	5.17		/	/		2.76	2.50
Col		/	/		3.58	3.67		3.51	3.38
D&T		5.17	5.17		2.72	3.17		2.44	2.50
TO		5.33	5.50		3.95	4.00		2.87	2.62
M		/	/		3.31	3.17		/	/
Total Score		82.30	84.20		79.00	78.30		70.60	67.50
T&L		/	/		/	/		6.68	6.83
Cre		/	/		/	/		6.81	6.67
L&C		8.12	8.00		7.58	7.67		6.77	6.67
S&P		7.86	8.33		7.44	7.50		/	/
SO		/	/		/	/		6.58	6.83
L&S		8.16	8.17		7.45	7.50		/	/
Col		/	/		7.67	7.67		6.79	6.67
D&T		7.90	7.83		7.39	7.17		6.88	6.83
TO		7.82	8.17		7.79	7.17		7.01	7.00
M		/	/		/	/		/	/
Total Score		63.20	65.80		50.00	47.90		45.50	41.70
T&L		6.66	6.67		/	/		/	/
Cre		6.49	6.67		/	/		/	/
L&C		6.47	6.50		5.18	5.71		4.86	4.67
S&P		/	/		4.39	4.43		4.16	4.00
SO		6.29	6.50		/	/		/	/
L&S		/	/		4.40	4.57		4.06	4.00
Col		6.12	6.50		5.29	5.43		/	/
D&T		6.59	6.83		4.12	4.43		3.78	3.83
TO		6.35	6.33		5.12	4.86		4.58	4.33
M		/	/		/	/		/	/

Figure 15: Test samples. *Predicted* represents the predicted score of the ArtCLIP output. *GT* represents the ground-truth score.

References

[Yi et al., 2023] Ran Yi, Haoyuan Tian, Zhihao Gu, Yu_x0002_Kun 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.

[Jiang et al., 2024] Shiqi Jiang, Ning Li, Chen Shi, Liping Guo, Changbo Wang, and Chenhui Li. Aacp: Aesthetics assessment of children's paintings based on self-supervised learning. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 2534–2542, 2024. 1, 2.1, 2.2.

[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 aesthet_x0002_ics research. Psychology of Aesthetics, Creativity, and the Arts, 2022.

[Amirshahi et al., 2015] Seyed Ali Amirshahi, Gregor Uwe Hayn-Leichsenring, Joachim Denzler, and Christoph Redies. Jenaesthetics subjective dataset: analyzing paint_x0002_ings 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 assess_x0002_ment in a subjective dataset of paintings. arXiv preprint arXiv:1609.05583, 2016.

[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.

[Jin et al., 2024] Xin Jin, Qianqian Qiao, Yi Lu, Huaye Wang, Shan Gao, Heng Huang, and Guangdong Li. Paintings and drawings aesthetics assessment with rich attributes for various artistic categories. In Proceedings of the 2024 International Joint Conference on Artificial Intelligence (IJCAI), August 2024. (document), 1, 1, 2.2, 3.1, 3.1, 4

NEURAL INFORMATION PROCESSING SYSTEMS

北京電子科技学院
Beijing Electronics Science & Technology Institute

中央美術學院
Central Academy of Fine Arts

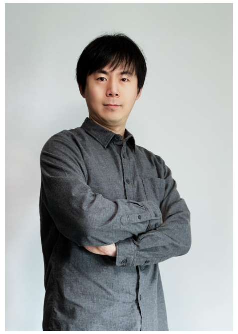
C A F A

VICTORY
Besti

THANKS!


Xin Jin


Qianqian Qiao


Yi Lu


Huaye Wang


Shan Gao


Heng Huang


Jianfei Liu


Rui Li