

Technical Report

TR-FSU-INF-CV-2013-07

Published in the internet on July 16th 2013

JenAesthetics—a public dataset of paintings for aesthetic research

Seyed Ali Amirshahi^{1,2}, Joachim Denzler¹, and Christoph Redies²

¹ Computer Vision Group, Friedrich Schiller University Jena, Germany
{seyed-ali.amirshahi, joachim.denzler}@uni-jena.de
<http://www.inf-cv.uni-jena.de>

² Experimental Aesthetics Group, Institute of Anatomy I, Jena University Hospital, Germany
{redies}@mti.uni-jena.de
<http://www.anatomie1.uniklinikum-jena.de>

Abstract. In recent years, aesthetic quality assessment has become a hot research topic in the image processing community. For aesthetic paintings, this research has been hampered by the lack of datasets that are available for public use and cover a wide variety of high-quality paintings. Here, we introduce the publicly available *JenAesthetics* dataset, a selection of more than 1,600 high-quality images of colored oil paintings of Western provenance by around 400 artists. Images were sampled from the Google Art Project. As a first contribution, we characterize the database in terms of art periods, styles and subject matters.

Key words: Aesthetic quality assessment, JenAesthetics dataset, aesthetic measures, color painting, art periods, subject matters in art.

1 Introduction

For a long time, psychologists and philosophers as well as art historians have explored criteria to assess the aesthetic quality of artworks and photographs. In recent years, the image processing and computer vision communities have joined this quest [1–13]. However, for a comparison of results from different studies, large common databases of aesthetic artworks are required that are freely available to the public. In the field of photography, several such databases are in general use [7–11], taking advantage of different photo-sharing websites. For paintings, the establishment of databases has been more difficult, due to the fact that art museums and art collectors generally do not allow researchers to scan or photograph their paintings, especially in large numbers. To solve this problem temporarily, some researchers have produced their own datasets by either scanning high-quality art books [1, 2, 4, 5] or by obtaining small sets of images from museums [12–14]. However, copyright laws restrict the sharing of these privately generated collections of artworks to a large degree.

Recently, museums have contributed images of their works of fine art to the Google Art (<http://www.googleartproject.com>). In this web-based project,

which is accessible to the general public at no fee, images of paintings are provided that are no longer subject to copyright restrictions in most countries. Many of the images are freely available for downloading at the Wikimedia Commons (<http://commons.wikimedia.org>) platform in high-quality jpeg format.

In the present work, we have gathered a large subsample of high-quality images from this database. The *JenAesthetics* dataset contains over 1600 images of paintings of Western provenance from around 400 painters. We are making this dataset freely available to the research community so that, in future studies, there is no need to create new databases, and results from different research groups can be compared directly.

In the following section, we will introduce the *JenAesthetics* dataset and give an overview on the different types of information available to the user.

2 Description of the JenAesthetics Dataset

The Google Art Project is an online platform for viewing artworks at a high resolution. In the present work, we take advantage of the fact that some of the more than 32,000 paintings from 46 museums in this project are available for download by the general public through the Wikimedia Commons website. Among the images available, 1625 images of colored oil paintings by 410 artist were selected and downloaded for the *JenAesthetics* dataset. Images were selected primarily based on their high resolution (image size generally higher than 3 Mbytes) and represented a large spectrum of styles, artists and subject matters from different Western countries. For each image, different types of information about the painting were then gathered and saved in a true file, as described in the following sections. A webpage with information about how to access the images and true files is dedicated to the *JenAesthetics* dataset (<http://www.inf-cv.uni-jena.de/en/jenaesthetics>).

2.1 Art periods and styles

For each painting, we list the art period/style it belongs to, according to standard textbooks on art and information available on the Wikipedia website. The dataset contains works from the 11 art periods/styles listed in Table 1. Examples are shown in Figures 1-11.

2.2 Subject matters

The content or subject matter of a painting plays a crucial role in how an observer will perceive and assess a painting. 16 different keywords were selected to cover most of the common subject matters. The subject matters with their codes in the true file shown in parentheses are: abstract (s10), nearly abstract (s11), landscapes (s21), scenes with person(s) (s22), still life (s23), flowers or vegetation (s25), animals (s26), seascape, port or coast (s27), sky (s28), portrait (one person) (s31), portrait (many person) (s32), nudes (s33), urban scene

Table 1: Initial characterization of the *JenAesthetics* dataset in comparison to other categories of images. For each image category, the art period and the number of images and painters are listed. The code representing this information in the true files is shown in parenthesis. The three numbers indicate the primary, secondary and tertiary subject matters, respectively.

	dataset/image content	number of images	number of painters
	<i>JenAesthetics</i>	1625	410
Art periods	Renaissance (p3)	113	28
	Mannerism (p4)	37	7
	Baroque (p5)	504	188
	Rococo (p6)	93	18
	Classicism (p7)	70	12
	Romanticism (p8)	142	28
	Realism (p10)	181	34
	Impressionism (p11)	206	40
	Symbolism (p13)	42	18
	Post-Impressionism (p15)	205	25
	Expressionism (p19)	32	12
Subject matters	abstract (s10)	3, 0, 0	2, 0, 0
	nearly abstract (s11)	7, 2, 2	4, 2, 2
	landscapes (s21)	179, 183, 76	66, 96, 54
	scenes with person(s) (s22)	494, 122, 62	196, 77, 39
	still life (s23)	66, 10, 1	37, 6, 1
	other (s24)	12, 2, 1	11, 2, 1
	flowers, vegetation (s25)	43, 103, 23	18, 62, 19
	animals (s26)	35, 82, 65	25, 50, 51
	seascape, port, coast (s27)	96, 111, 49	45, 65, 35
	sky (s28)	11, 42, 35	2, 23, 23
	portrait (one person) (s31)	440, 7, 1	173, 7, 1
	portrait (many person) (s32)	79, 1, 0	56, 1, 0
	nudes (s33)	43, 34, 2	34, 24, 2
	urban scene (s40)	66, 75, 47	34, 53, 35
	building (s41)	36, 80, 44	22, 49, 35
	interior scene (s42)	15, 193, 17	9, 101, 17

(s40), building (s41), interior scene (s42) and other subject matters (s51). For each painting, 1-3 subject matter keywords were assigned in the order of their impact on the visual appearance of the painting (425 paintings have 3 and 1047 paintings have 2 subject matter keywords). Figures 1-11 give the subject matter keywords assigned to each of the images shown. It is important to stress that the assessment of subject matter keywords is based on subjective grounds; exact classifications may differ for person to person.

2.3 Rule of thirds

One of the features determined for the *JenAesthetics* dataset is the presence of the Rule of thirds. The Rule of thirds is one of the most prominent features used to assess image composition in paintings and photographs [6, 7, 10, 11]. Different definitions of the rule of thirds have been given. Most, if not all, of these definitions suggest that, in order to have a photograph or painting of high aesthetic quality, the main object or focus point should be along one of the two imaginary horizontal or the two imaginary vertical lines that divide the image into nine equal parts. In the dataset, a total of 1125 images do not follow the rule of thirds while 500 images do follow the rule of thirds.

2.4 General information on the image file

For each image, the true file (Figures 12-16) provides the following information that was obtained from the Google Art Project via the Wikimedia commons platform:

- (a) file name of the image in our dataset,
- (b) download link for the image,
- (c) name of the painter,
- (d) year, in which the artwork was created,
- (e) art period the painting belongs to,
- (f) size of the image file as the number of pixels (height \times width),
- (g) information regarding the existence of the rule of thirds in painting,
- (h) subject matter(s) in the painting,

3 Conclusion

In conclusion, we introduce and characterize the *JenAesthetics* dataset of over 1.600 aesthetic colored oil paintings. This dataset is a subset of images from the Google Art Project, which is not subject to copyright restrictions in most countries. The dataset is made available to the general public for research on aesthetic artworks and covers 11 major periods of Western art and diverse subject matters. The *JenAesthetics* dataset will allow to directly compare measurements by different research groups on the same dataset.



Fig. 1: Sample images for the *JenAesthetics* dataset from the Renaissance art period. a) Paolo Veronese, 1575, b) Paris Bordone, 1550, c) Bronzino, 1533, d) Vincenzo Catena, 1531, e) Tiziano Vecelli, 1529. The numerical code indicates subject matters (see text).



Fig. 2: Sample images for the *JenAesthetics* dataset from the Mannerism art period. a) Benvenuto Tisi, 1504, b) Tintoretto, 1592, c) Tintoretto, 1543, d) El Greco, 1600, e) Alessandro Allori, 1593. The numerical code indicates subject matters (see text).



Fig. 3: Sample images for the *JenAesthetics* dataset from the Baroque art period. a) Hendrick Avercamp, 1620, b) Aelbert Cuyp, 1650, c) Giovanni Paolo Panini, 1734, d) Rembrandt, —, e) Anthony van Dyck, 1635. The numerical code indicates subject matters (see text).

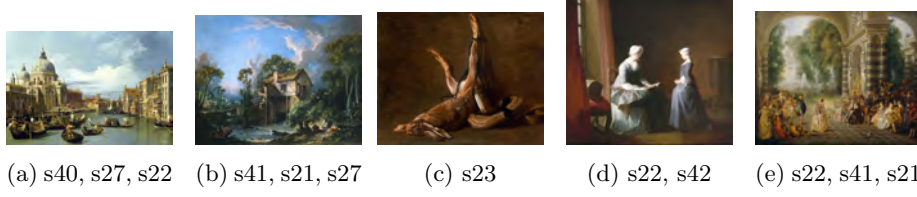


Fig. 4: Sample images for the *JenAesthetics* dataset from the Rococo art period. a) Antonio Canaletto, 1730, b) François Boucher, 1758, c) Jean-Baptiste-Siméon Chardin, 1730, d) Jean-Baptiste-Siméon Chardin, 1753, e) Antoine Watteau, 1717. The numerical code indicates subject matters (see text).



Fig. 5: Sample images for the *JenAesthetics* dataset from the Classicism art period. a) Benjamin West, 1788, b) John Singleton Copley, 1778, c) François Gérard, 1797, d) Jacques-Louis David, 1802, e) Johann Zoffany, 1777. The numerical code indicates subject matters (see text).

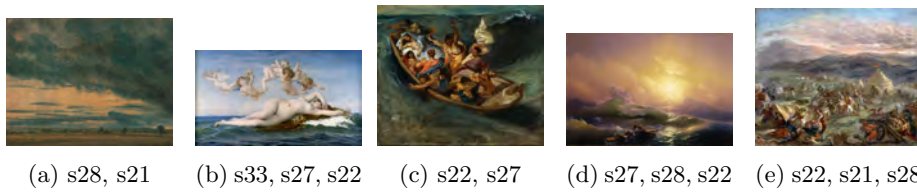


Fig. 6: Sample images for the *JenAesthetics* dataset from the Romanticism art period. a) John Constable, 1821, b) Alexandre Cabanel, 1863, c) Eugène Delacroix, 1841, d) Hovhannes Aivazovsky, 1850, e) Eugène Delacroix, 1862. The numerical code indicates subject matters (see text).



Fig. 7: Sample images for the *JenAesthetics* dataset from the Realism art period. a) Rosa Bonheur, 1849, b) Jean-Baptiste-Camille Corot, 1833, c) John Frederick Herring, 1833, d) Winslow Homer, 1884, e) Édouard Manet, 1863. The numerical code indicates subject matters (see text).



Fig. 8: Sample images for the *JenAesthetics* dataset from the Impressionists art period. a) Gustave Caillebotte, 1882, b) Mary Cassatt, 1893, c) William McTaggart, 1864, d) Monet, 1873, e) Pierre-Auguste Renoir, 1876. The numerical code indicates subject matters (see text).



Fig. 9: Sample images for the *JenAesthetics* dataset from the Symbolism art period. a) Tivadar Kosztka Csontváry, 1905, b) Pierre Puvis de Chavannes, 1881, c) Henri Fantin-Latour, 1904, d) August Malmström, 1866, e) Giovanni Segantini, 1892. The numerical code indicates subject matters (see text).

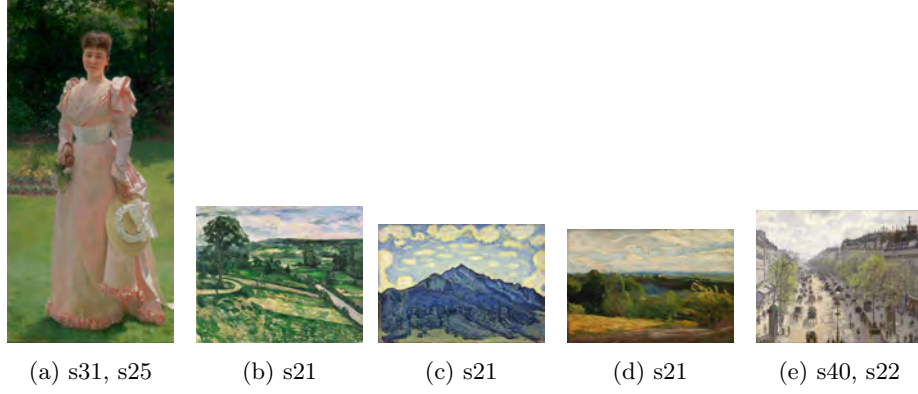


Fig. 10: Sample images for the *JenAesthetics* dataset from the Post-Impressionist art period. a) Vlaho Bukovac, 1892, b) Paul Cézanne, 1882, c) Henri-Edmond Cross, 1906, d) Martn Malharro, 1901, e) Pissarro, 1897. The numerical code indicates subject matters (see text).

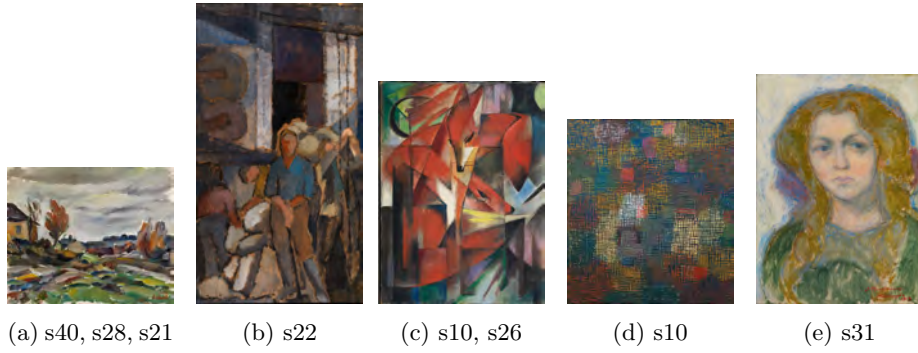


Fig. 11: Sample images for the *JenAesthetics* dataset from the Expressionism art period. a) Ilmari Aalto, 1915, b) Alvar Cawén, 1919, c) Franz Marc, 1913, d) Paul Klee, 1932, e) Jalmari Ruokokoski, 1913. The numerical code indicates subject matters (see text).



(a)

File name: Cuy6.jpg Cuy6
 Download link:
http://upload.wikimedia.org/wikipedia/commons/7/73/Aelbert_Cuyck_-_The_Passage_Boat_-_Google_Art_Project.jpg
 Artist: Aelbert Cuyck
 Year: 1650
 Period: p5
 Size: [4007 4637]
 Rule of Thirds: Y
 Subject matter 1: s27
 Subject matter 2: s22
 Subject matter 3: s28

(b) True file

Fig. 12: Sample images from the *JenAesthetics* dataset with its corresponding true file.



(a)

File name: Danby10.jpg Danby10
 Download link:
http://upload.wikimedia.org/wikipedia/commons/3/37/Francis_Danby_-_The_Avon_Gorge_at_Sunset_-_Google_Art_Project.jpg
 Artist: Francis Danby
 Year: 1850
 Period: p8
 Size: [3513 4165]
 Rule of Thirds: N
 Subject matter 1: s21
 Subject matter 2: s27
 Subject matter 3: s28

(b) True file

Fig. 13: Sample images from the *JenAesthetics* dataset with its corresponding true file.



(a)

File name: Hassan8.jpg
 Download link:
http://upload.wikimedia.org/wikipedia/commons/b/bc/Childe_Hassan_-_The_Concord_Meadow_-_Google_Art_Project.jpg
 Artist: Childe Hassam
 Year: 1891
 Period: p11
 Size: [3205 3940]
 Rule of Thirds: N
 Subject matter 1: s21
 Subject matter 2: s27
 Subject matter 3: s---

(b) True file

Fig. 14: Sample images from the *JenAesthetics* dataset with its corresponding true file.



(a)

File name: Rusin16.jpg
 Download link:
http://upload.wikimedia.org/wikipedia/commons/4/47/santiago_rusin33810752c_Spanish_-_Interior_of_a_Caf%C3%A9_-_Google_Art_Project.jpg
 Artist: Santiago Rusiñol
 Year: 1892
 Period: p15
 Size: [5234 4282]
 Rule of Thirds: N
 Subject matter 1: s22
 Subject matter 2: s42
 Subject matter 3: s---

(b) True file

Fig. 15: Sample images from the *JenAesthetics* dataset with its corresponding true file.



(a)

```

File name: Zoffany3.jpg           Zoffany3
Download Link:
http://upload.wikimedia.org/wikipedia/commons/d/d8/Johann_Joseph_Zoffany_-_The_Dr
ummond_Family_-_Google_Art_Project.jpg
Artist: Johann Zoffany
Year: 1769
Period: p7
Size: [4786 7267]
Rule of thirds: y
Subject matter 1: s32
Subject matter 2: s26
Subject matter 3: s21

```

(b) True file

Fig. 16: Sample images from the *JenAesthetics* dataset with its corresponding true file.

References

1. Redies, C., Amirshahi, S., Koch, M., Denzler, J.: PHOG-derived aesthetic measures applied to color photographs of artworks, natural scenes and objects. In: Computer Vision–ECCV. Workshops and Demonstrations, Springer (2012) 522–531
2. Amirshahi, S., Koch, M., Denzler, J., Redies, C.: PHOG analysis of self-similarity in aesthetic images. In: IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics (2012) 82911J–82911J
3. Amirshahi, S., Redies, C., Denzler, J.: How self-similar are artworks at different levels of spatial resolution? In: International Symposium on Computational Aesthetics in Graphics, Visualization, and Imaging. (2013)
4. Koch, M., Denzler, J., Redies, C.: $1/f^2$ characteristics and isotropy in the Fourier power spectra of visual art, cartoons, comics, mangas, and different categories of photographs. PloS One **5** (2010) e12268
5. Redies, C., Hasenstein, J., Denzler, J.: Fractal-like image statistics in visual art: similarity to natural scenes. Spatial Vision **21** (2007) 137–148
6. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. IEEE Journal of Selected Topics in Signal Processing **3** (2009) 236–252
7. Datta, R., Li, J., Wang, J.: Algorithmic inferencing of aesthetics and emotion in natural images: An exposition. In: IEEE International Conference on Image Processing (ICIP), IEEE (2008) 105–108
8. Datta, R., Joshi, D., Li, J., Wang, J.: Studying aesthetics in photographic images using a computational approach. Computer Vision–ECCV (2006) 288–301
9. Murray, N., Marchesotti, L., Perronnin, F.: Ava: A large-scale database for aesthetic visual analysis. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE (2012) 2408–2415
10. Bhattacharya, S., Sukthankar, R., Shah, M.: A framework for photo-quality assessment and enhancement based on visual aesthetics. In: Proceedings of the International Conference on Multimedia, ACM (2010) 271–280
11. Mai, L., Le, H., Niu, Y., Liu, F.: Rule of thirds detection from photograph. In: IEEE International Symposium on Multimedia (ISM), IEEE (2011) 91–96
12. Demetriou, M., Hardeberg, J., Adelmann, G.: Computer-aided reclamation of lost art. In: Computer Vision–ECCV. Workshops and Demonstrations, Springer (2012) 551–560

13. Oncu, A., Deger, F., Hardeberg, J.: Evaluation of digital inpainting quality in the context of artwork restoration. In: Computer Vision–ECCV. Workshops and Demonstrations, Springer (2012) 561–570
14. Graham, D., Field, D.: Statistical regularities of art images and natural scenes: Spectra, sparseness and nonlinearities. *Spatial Vision* **21** (2007) 149–164