

Understanding and Creating Art with AI: Review and Outlook

EVA CETINIC, JAMES SHE

科技藝術書報討論

指導老師：許素朱 教授

林巖 IPHD 110003818

- **EVA CETINIC**, Rudjer Boskovic Institute and Durham University, UK
<https://scholar.google.hr/citations?user=5R4EpaEAAAAJ&hl=en>
<https://www.dvs.uzh.ch/en/Fellows/Dr.-Eva-Cetinic.html>
- **JAMES SHE**, Hamad Bin Khalifa University (HBKU) and Hong Kong University of Science and Technology <https://ece.hkust.edu.hk/eejames>

Additional Key Words and Phrases:

- AI Art, deep learning, generative art, computational creativity, generative adversarial networks, convolutional neural networks, visual arts, image understanding

ACM Reference format:

- Eva Cetinic and James She. 2022. Understanding and Creating Art with AI: Review and Outlook. ACM Trans. Multimedia Comput. Commun. Appl. 18, 2, Article 66 (February 2022), 22 pages.
- <https://doi.org/10.1145/3475799>

INTRODUCTION

- **AI and art**

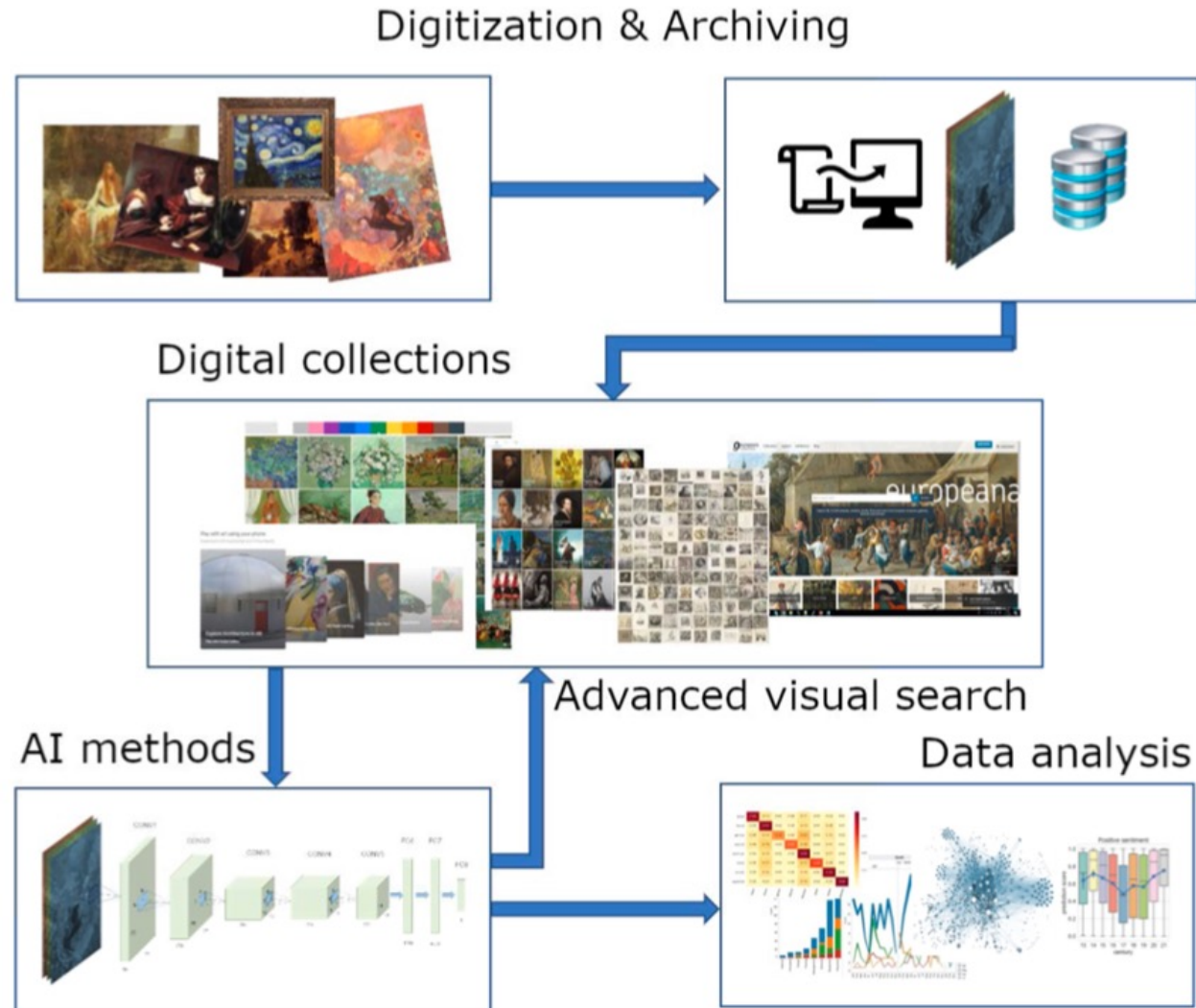
(1) AI is used in the process of analyzing existing art, or

(2) AI is used in the process of creating new art.

- The use of AI in the process of creating visual art was significantly accelerated with the emergence of generative adversarial networks (GANs)
- The use of the convolutional neural network (CNN) in AI of ART.

UNDERSTANDING ART WITH AI

- Art Collections as Data Sources



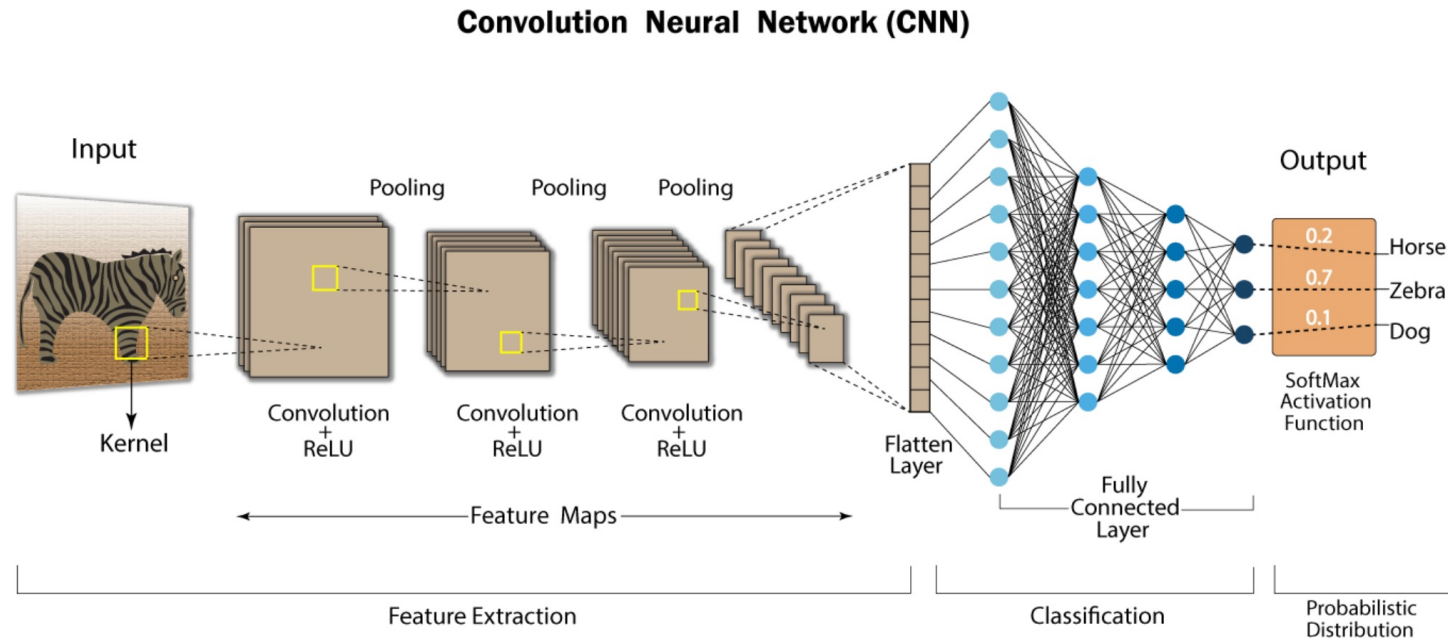
Art Dataset

Dataset	Source	No. of Images	Task
Painting-91	[79]	4k	Artist classification
Pandora	[49]	8k	Style classification
Web Gallery of Art	www.wga.hu	40k	Artist, style, period classification
WikiArt	www.wikiart.org	85k	Artist, style, genre classification
Rijksmuseum Challenge	[99]	112k	Artist, material, type classification
Art500k	[92]	550k	Artist, genre, style, event, historical figure retrieval
OmniArt	[127]	2M	Artist, style, period, type, iconography, color classification/object detection

ArtDL	[101]	42k	Iconographic classification
PRINTART	[16]	1k	Object/pose retrieval
Paintings	[34]	8.6k	Object retrieval
Face Paintings	[32]	14 k	Face retrieval
Iconart	[58]	6k	Iconographic object detection
VisualLink	[117]	38.5k	Visual link retrieval
Brueghel dataset	[120]	1.6k	Visual link retrieval
IconClass AI Test Set	[106]	87k	Iconographic classification/multimodal tasks
SemArt	[52]	21k	Multimodal retrieval
Artpedia	[125]	3k	Multimodal retrieval
BibleVSA	[8]	2.3k	Multimodal retrieval
AQUA	[53]	21k	Visual question answering
ArtEmis	[3]	81K	Multimodal sentiment analysis
WikiArt Emotions	[102]	4.1k	Sentiment analysis/emotion classification
MART	[138]	500	Sentiment analysis
JenAesthetic	[6]	1.6k	Aesthetics quality assessment

Automated Classification of Artworks

- Categories such as artist, style, or genre. These had been studied and identified art in some essential elements which become the deep learning variances to the AI in Art.
- Used CNN as the deep learning model to enable the art variable for dependence in which to construct the Art through AI or deep learning system.



Source: <https://developersbreach.com/convolution-neural-network-deep-learning/>



14,197,122 images, 21841 synsets indexed

[Home](#) [Download](#) [Challenges](#) [About](#)

Not logged in. [Login](#) | [Signup](#)

<https://www.image-net.org/index.php>

- Deep learning and AI models go through the ImageNet database to learn the graph style and identify the learning process and results in different phases for perspective identification in variance.

Object Detection and Similarity Retrieval

- The use of deep neural networks shows promising results in exploring the content of artworks and automatically recognizing objects.
- As the pioneering works in this area, CNN features from natural images can retrieve paintings containing these objects with success.
- One of the main practical goals of computational methods for automated content and style recognition in art images is to build a smart retrieval system that helps organize and analyze large collections of artwork in an efficient process.

Multimodal Tasks

- An increased number of studies have focused on analyzing both visual and textual modalities of artwork collections. Efforts to map images and their textual descriptions in a joint semantic space have mostly been made to create multimodal retrieval systems.
- In particular, Garcia and Vogiatzis [52] introduced the SemArt dataset, a collection of fine-art images associated with textual descriptions, and applied different methods of multimodal transformation with the goal to map the images and their descriptions in a joint semantic space.

<https://paperswithcode.com/dataset/semart>

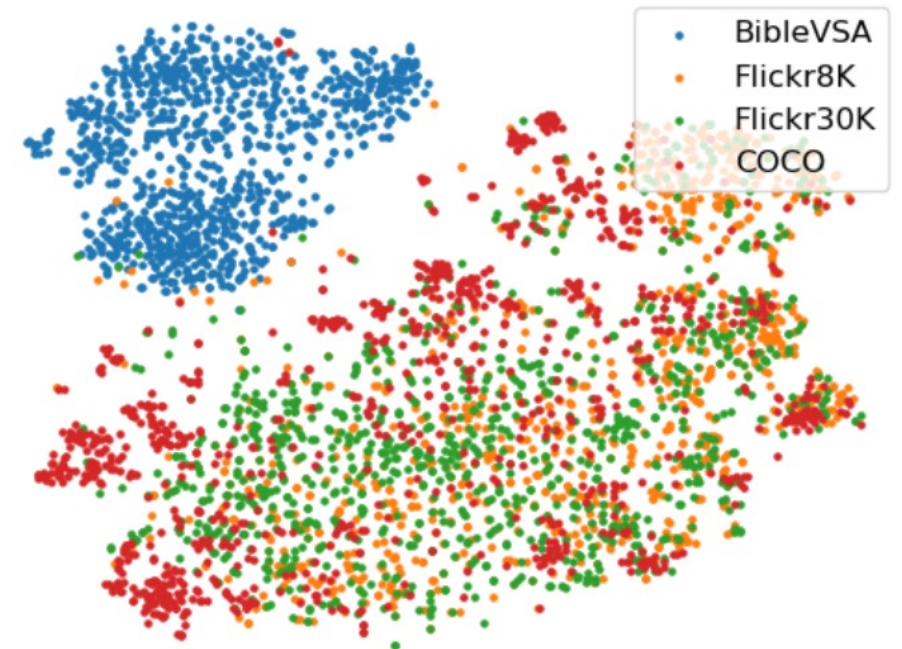
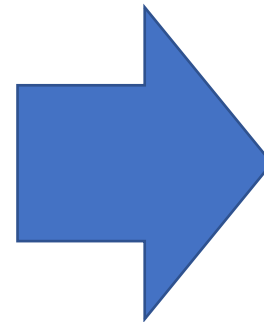


Title: Still-Life
Author: Willem van Aelst
Type: Still-Life
School: Dutch
Timeframe: 1651-1700

The painting depicts a still-life with roses, tulips and other flowers resting on a ledge. It demonstrates the elegance, refinement, and technical brilliance cultivated during the painter's formative years in Italy.

Baraldi et al. [8] introduced a new dataset named **BibleVSA**, a collection of miniature illustrations and commentary text pairs, to explore supervised and semi-supervised approaches of learning cross references between textual and visual information in documents.

https://www.researchgate.net/figure/Overview-of-the-proposed-BibleVSA-dataset-On-the-left-a-sample-page-from-the-Borso_fig1_326626330



Stefanini et al. [125] presented the Artpedia dataset where images are annotated with both visual and contextual descriptions and introduced a retrieval model that maps images and sentences in a joint embedding space and discriminates between contextual and visual sentences of the same image.

<https://aimagelab.ing.unimore.it/imagelab/page.asp?IdPage=35>

Dataset



✓ The painting depicts an idyllic, pastoral scene of a lone young woman in peasant attire posed for the artist, balancing a stick (likely her crook) across her shoulders, standing barefooted in the foreground.

✗ The title is taken from the Southern French dialect.

✗ It is currently in the permanent collection at the Philbrook Museum of Art in Tulsa, where it has become an emblematic image for the museum.



✓ In the foreground, a young man stands upon a rocky precipice with his back to the viewer.

✓ He is wrapped in a dark green overcoat, and grips a walking stick in his right hand.

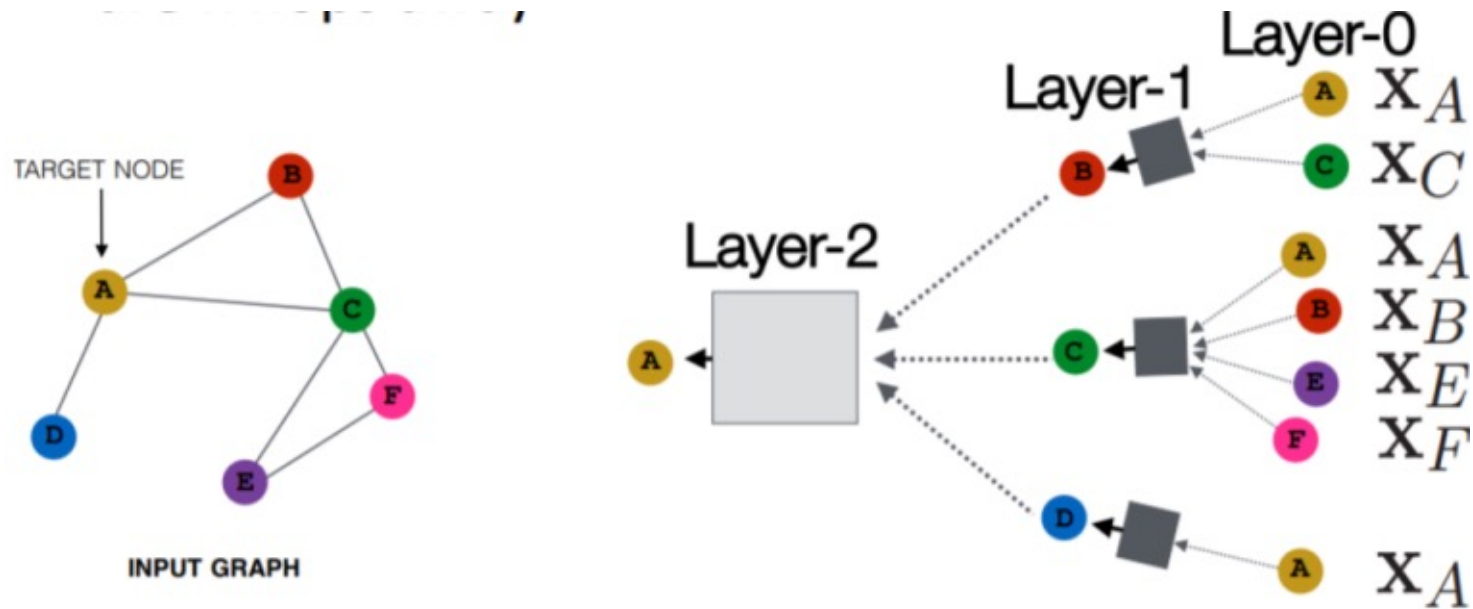
✗ It has been considered one of the masterpieces of Romanticism and one of its most representative works.

✗ It currently resides in the Kunsthalle Hamburg in Hamburg, Germany.

Efthymiou et al. proposed a novel multimodal architecture that integrates graph neural networks (GNNs) and CNNs to jointly learn visual and semantic-based artistic representations.

<https://neptune.ai/blog/graph-neural-network-and-some-of-gnn-applications>

Amal Menzli <https://arxiv.org/pdf/1812.08434.pdf>



visual question answering (VQA). VQA refers to the problem where given an image and textual question,

Garcia et al. presented a novel dataset AQUA, which consists of automatically generated visual- and knowledge-based question-answer pairs.

Apart from VQA, a few recent works addressed the task of image captioning where the goal is to automatically generate accurate textual descriptions of images.

Sheng and Moens introduced image captioning datasets referring to ancient Egyptian and Chinese art and employed an encoder-decoder framework for image captioning where the encoder is a CNN and the decoder is a long short-term memory (LSTM) network.

Knowledge Discovery in Art History

Particularly important in advanced computational analysis of art is a collaboration between different disciplines, especially computer science and art history.

Due to practical difficulties, cross-domain knowledge gaps, or animosity toward the increasing trend of quantification in humanities research.

research criteria are becoming more rigorous and computational methods are not being used only because it is fashionable but because they can provide truly novel methodological extensions

Aesthetics and Perception

WikiArt Emotions: An Annotated Dataset of Emotions Evoked by Art

Saif M. Mohammad and Svetlana Kiritchenko (saif.mohammad.svetlana.kiritchenko@nrc-crnc.gc.ca)

NRC-CRNC

National Research Council Canada

RESEARCH QUESTIONS

- what makes art evocative?
- how does art convey different emotions?
- what attributes make a painting well liked?
- how much does the title of an art impact its emotional response?
- what is the extent to which categories of art evoke consistent emotions?

THE WIKIART EMOTIONS DATASET

www.saifmohammad.com/WebPages/wikiartemotions.html

- Annotated 4,105 pieces of art for emotions evoked, amount liked, whether they depict a face
- 10 people per item, crowdsourcing
 - 4 styles: Renaissance, Post-Renaissance, Modern, Contemporary
 - 22 categories: Impressionism, Figurative art, Realism, etc.

APPLICATIONS

- search paintings evoking the desired emotional response
- automatically detect emotions evoked by paintings
- automatically transform paintings
- identify what makes paintings evocative

THE SOURCE OF THE ART

WikiArt.org: 151,151 pieces of art; 10 art styles; 168 style categories

- notable art in each category is shown in a *Featured* page
- selected ~200 items from each of the featured pages of 22 categories

Style

Contemporary Art: ~2000 pieces

Minimalism

Modern Art: ~60,000 pieces

Impressionism, Expressionism, Post-Impressionism, Surrealism, Abstract Expressionism, Cubism, Pop Art, Abstract Art, Art Informel, Color Field Painting, Neo-Expressionism, Magic Realism, Lyrical Abstraction

Post-Renaissance Art: ~35,000 pieces

Realism, Romanticism, Baroque, Neoclassicism, Rococo

Renaissance Art: ~6,000 pieces

Northern Renaissance, High Renaissance, Early Renaissance



Title: Mona Lisa
Year: 1519

EMOTION ANNOTATIONS

- 20 emotions chosen from art literature
- select all that the art brings to mind
- annotate image, title, art (image, title)

Emotion	Emotion Distribution %
Positive	
gratitude, thankfulness, or indebtedness	1.30
happiness, calmness, pleasure, or ecstasy	36.90
humility, modesty, unpretentiousness, or simplicity	12.00
love or affection	8.10
optimism, hopefulness, or confidence	5.90
trust, admiration, respect, dignity, or honor	21.50
Negative	
anger, annoyance, or rage	1.30
arrogance, vanity, hubris, or conceit	1.90
disgust, dislike, indifference, or hate	1.70
fear, anxiety, vulnerability, or terror	10.20
pessimism, cynicism, or lack of confidence	1.10
regret, guilt, or remorse	0.30
sadness, pensiveness, loneliness, or grief	9.20
shame, humiliation, or disgrace	0.70
Other or Mixed	
agreeableness, acceptance, submission, compliance	0.10
anticipation, interest, curiosity, suspicion, vigilance	15.40
disagreeableness, defiance, conflict, or strife	0.20

LABEL DISTRIBUTION

4 out of 10 say emotion present

AVERAGE ART RATINGS

-3 (dislike a lot) to 3 (like a lot)

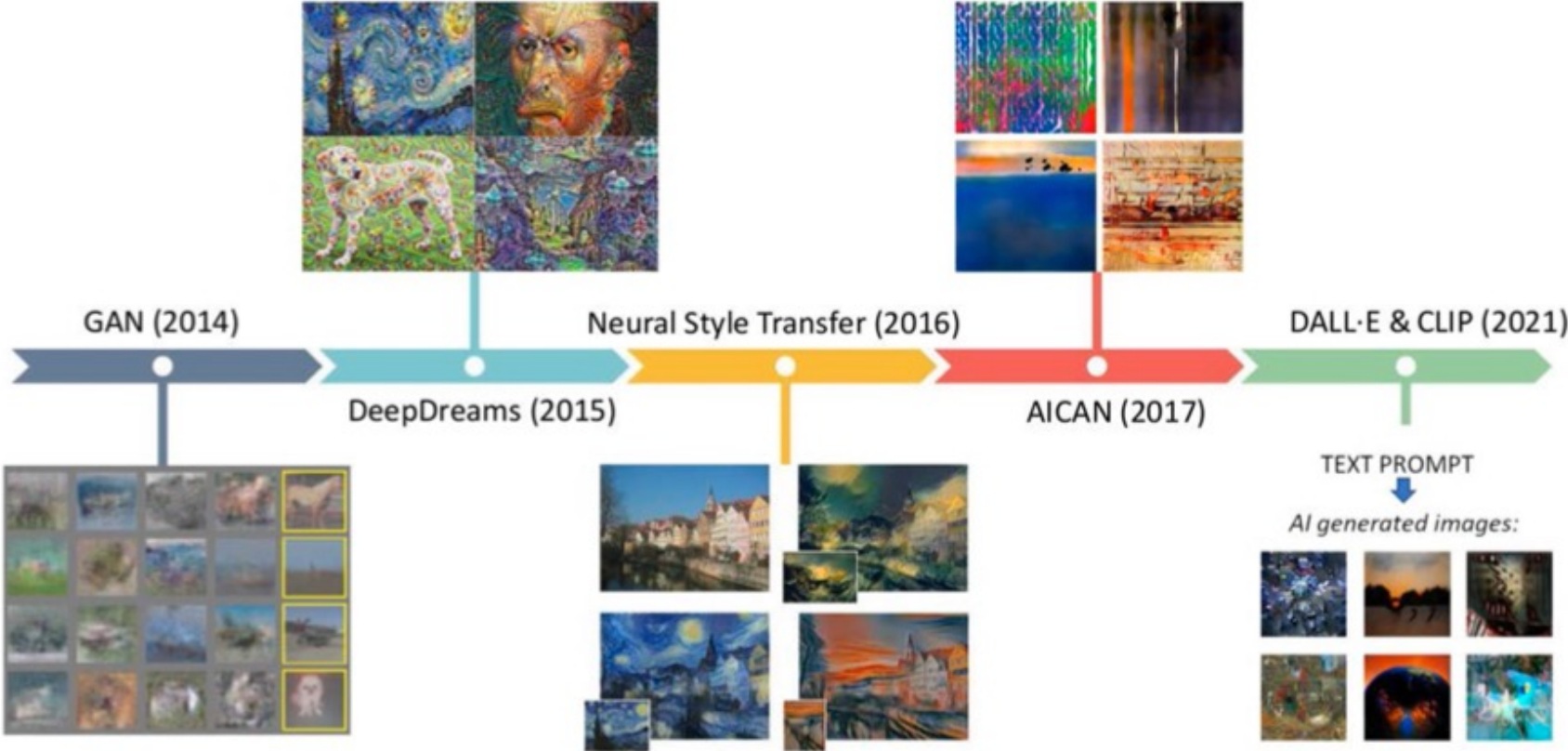
Emotion	Avg. Art Ratings (by fine emotion)
gratitude	1.87
happiness	1.79
humility	1.62
love	1.95
optimism	1.72
trust	1.78
anger	0.41
arrogance	0.80
disgust	-0.38
fear	0.27
pessimism	0.39
regret	0.89
sadness	0.79
shame	0.48
agreeableness	1.60
anticipation	0.99
disagreeableness	0.60

- Different aspects of visual perception have been studied by psychologists for a long time and have in recent years become a rising subject of interest within the computer vision and deep learning community.
- Developing quantitative methods for analyzing subjective aspects of perception is particularly challenging in the context of art images.
- Amirshahi et al. introduced the JenAesthetics dataset, a datasets of artwork images labeled with subjective scores of aesthetic evaluation. <https://github.com/Bin-ary-Li/JenAesthetics>
- Mohammad and Kiritchenko [102] introduced WikiArt Emotions, a dataset of paintings that has annotations for various emotions evoked in the observer. <http://saifmohammad.com/WebPages/wikiartemotions.html>

CREATING AI ART

Technological Milestones

Neural style transfer (NST). This method was introduced in the highly influential work of Gatys et al. that demonstrated the successful use of CNNs in creating stylized images by separating and combining the image “content” and “style.”



Elgammal et al. introduced **AICAN**—an AI creative adversarial network. In their work, they argue that if a GAN model is trained on images of paintings, it will just learn how to generate images that look like already existing art, and in a similar manner as the NST method, this will not produce anything truly artistic or novel.

In January 2021, **OpenAI** presented a 12-billion parameter network called **DALL·E** trained to generate images from text descriptions, using a large dataset of text-image pairs.

Radford et al. introduced the **CLIP (Contrastive Language-Image Pre-training)** model that was trained using 400 million image-text pairs collected from the internet.

The **CLIP model** with other generative models to guide the generative latent space search, such as **BigSleep** and **DeepDaze**.

The Contemporary AI Art Scene

Since October 2018, when the AI artwork “Portrait of Edmond Belamy” produced by the Obvious collective was sold at an auction by Christie’s for \$432,500 [27], there has been an increasing interest for AI Art but also a growing need to discuss key aspects of this new movement in the contemporary art scene.



- (1) Obvious, Portrait of Belami.
- (2) Mario Klingemann, Memories of Passersby I.
- (3) Sofia Crespo, Neural Zoo.
- (4) Robbie Barrat, Nudes.
- (5) Scott Eaton, Humanity (Fall of the Damned).
- (6) James She, Keep Running.

- **Aaron Hertzmann** presents the context of visual art history in his essay “Can Computers Create Art?”. In his article, he draws parallels between AI Art and the invention of photography which explores the evolution of collaboration between art and technology in filmmaking, 3D computer animation, and procedural artwork.
- The current AI Art works can be understood as results of sampling the “**latent space.**” Perhaps the most novel aspect of AI Art is this possibility to venture into that abstract multi-dimensional space of encoded image representations.

Machine Autonomy and the Role of the Artist

“The recent GAN-produced artworks use AI as a tool, whereas the creative process is primarily dependent on the artist’s pre- and post-curatorial actions.”

The process of the art dataset and the deep learning models combined with the signature with the texture.

Hertzmann indicates “that AI algorithms are not autonomous creators and will not be in the foreseeable future. They are still just tools, ready for artists to explore and exploit.” “Computers Do Not Make Art, People Do”

Authorship Copyright, and Ethical Issues

The case of Christie's Belamy auction revealed many issues regarding the questions of authorship and copyright, as well as raised general discussions on the ethical considerations that have to be taken into account during production, promotion, and sale of an AI artwork.

The problematic aspects of the **“Portrait of Edmond Belamy”** regarding authorship, authenticity, and other important aspects of AI Art.

CryptoArt, which led to a great expansion of the so-called crypto art market that is based on the use of blockchain technology.

The **non-fungible token (NFT)** into the digital art market. NFT is a type of cryptocurrency that differs from other cryptocurrencies by being unique and not exchangeable like-for-like.

The digital artist Beeple **“Everydays: The First 5000 Days”** sold for more than \$69 million

Perception of AI Art

One of the major arguments for labeling generative AI systems as creative was the fact that the work they produced was indistinguishable from human-made art and perceived as surprising, interesting, or aesthetically pleasing by a larger number of people.

However, even if AI systems can, or will in the future, produce convincing artworks that resemble human-made art, that does not necessarily imply that the system itself should be perceived as truly autonomous or creative.

CONCLUSION AND FUTURE OUTLOOK

From the perspective of computer vision, there are still many practical challenges that need to be solved to assist researchers working on cultural digital archives. In particular, those are problems related to annotation standards, advanced object detection and retrieval, cross-depiction, iconographic classification, multimodal alignment, and image understanding.

Furthermore, it is evident that the increasing use of AI technologies in the creation of art will have significant implications regarding the questions related to authorship, as well as on our human perception of art.

Connection

- This paper provides clear Art with AI and historical enhancement.
- With the art dataset and the data in signature with various considerations to define the ART.
- With CNN and GNN in deep learning to construct the art symbols in various methods.

Comments

With the starter of ART with AI development, this paper provided more clearly process and mindset to catch up with complex AI ART. Digitalized the art in combination and rethink the art through AI models in some successful models.

A rabbit in sky

```
File Edit Search Source Run Debug Consoles Projects Tools View Help  
pytorch-mutex-1.0 3 KB #####  
cuda-tools-11.7.1 1 KB #####  
cuda-visual-tools-11 1 KB #####  
cuda-nvtx-11.7.91 43 KB #####  
torchaudio-0.13.1 4.6 MB #####  
cuda-nvcc-11.7.99 44.0 MB #####  
cuda-toolkit-11.7.1 1 KB #####  
cuda-libraries-11.7. 1 KB #####  
pytorch-cuda-11.7 3 KB #####  
cuda-nvml-dev-11.7.9 85 KB #####  
cuda-cudart-dev-11.7 711 KB #####  
pytorch-1.13.1 1.10 GB #####  
cuda-nvrtc-11.7.99 71.9 MB #####  
cuda-command-line-to 1 KB #####  
Preparing transaction: done  
Verifying transaction: done  
Executing transaction: done  
Retrieving notices: ...working... done
```

```
(pytorch-gpu) C:\Users\Lin_Y>dream "a rabbit in sky"
```

```
setting seed of 0  
you can override this with --seed argument in the command line, or --random for a randomly chosen one  
Imagining "a_rabbit_in_sky" ...  
loss: -40.41: 100% | 1050/1050 [09:07<00:00, 1.92it/s]  
loss: -40.96: 100% | 1050/1050 [09:05<00:00, 1.92it/s]  
loss: -41.49: 100% | 1050/1050 [10:09<00:00, 1.72it/s]  
epochs: 15% | 3/20 [28:23<2:43:03, 575.48s/it]  
loss: -41.90: 63% | 664/1050 [05:46<03:19, 1.93it/s]  
image update: 18% | 76/420.0 [34:03<2:29:51, 26.14s/it]
```

Name	Date modified	Type
.anaconda	8/19/2021 10:32 PM	File folder
.cache	12/21/2022 10:09 PM	File folder
.conda	12/24/2022 5:14 PM	File folder
.continuum	8/19/2021 10:03 PM	File folder
.eclipse	11/14/2022 6:16 PM	File folder
.ipython	8/19/2021 10:33 PM	File folder
.jupyter	12/23/2022 11:15 PM	File folder
.matplotlib	8/19/2021 10:33 PM	File folder
.ms-ad	7/5/2022 5:34 PM	File folder
.oracle_jre_usage	11/14/2022 6:37 PM	File folder
.p2	11/14/2022 6:16 PM	File folder
.pytorch_pretrained_biggan	12/21/2022 10:51 PM	File folder
renesas	8/22/2021 7:38 AM	File folder

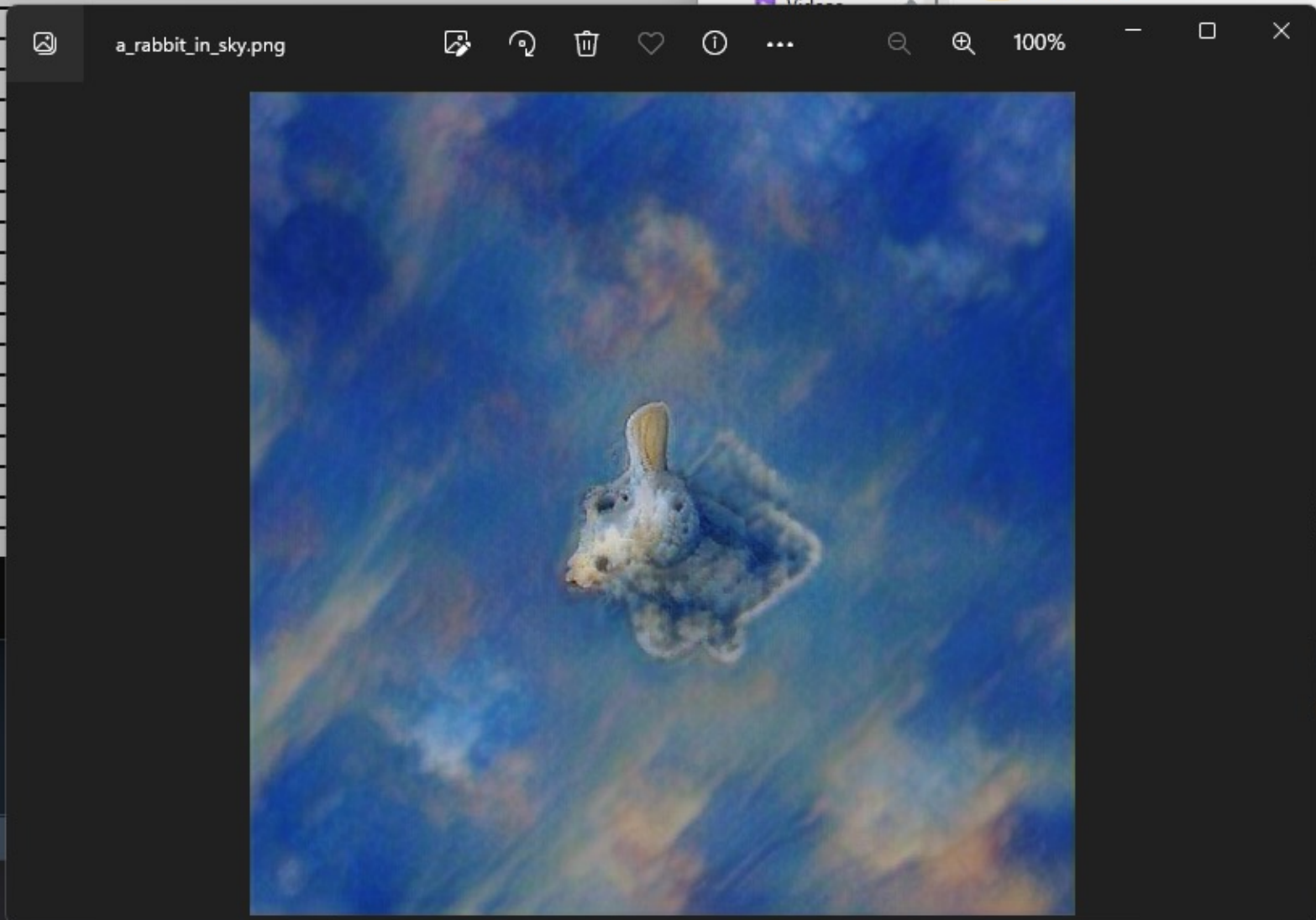
Anaconda and pip. Anaconda is the recommended pack

PyTorch dependencies in one sandboxed install, inclu

```
pytorch-gpu) C:\Users\Lin_Y>dream "a rabbit in sky"
Starting up... v0.9.1
Setting seed of 0
You can override this with --seed argument in the command line, or --random for
imagining "a_rabbit_in_sky" ...
ss: -40.41: 100%
ss: -40.96: 100%
ss: -41.49: 100%
ss: -41.46: 100%
ss: -42.07: 100%
ss: -42.20: 100%
ss: -42.11: 100%
ss: -42.23: 100%
ss: -41.55: 100%
ss: -42.10: 100%
ss: -42.49: 100%
ss: -42.45: 100%
ss: -42.06: 100%
ss: -42.89: 100%
ss: -41.91: 100%
ss: -42.84: 100%
ss: -41.92: 100%
ss: -42.48: 100%
ss: -41.96: 100%
ss: -42.21: 100%
epochs: 100%
page update: 100%
```

File Explorer window showing the contents of the 'This PC' directory. The table below lists the files and folders visible in the 'Pictures' folder.

Name	Date modified	Type
OneDrive	12/23/2022 11:04 PM	File folder
Pictures	12/19/2022 10:42 AM	File folder
pip	12/23/2022 11:28 PM	File folder
Saved Games	12/19/2022 10:42 AM	File folder
	12/19/2022 10:42 AM	File folder
	11/14/2022 9:59 PM	File folder
	12/23/2022 11:03 PM	File folder
	6/22/2022 6:27 PM	CONDARC F
	12/24/2022 7:44 PM	PNG File
	7/28/2022 4:40 PM	DAT File
	2/27/2022 9:39 AM	MATLAB Cod
	8/19/2021 11:03 PM	LOCK File
	9/5/2022 11:23 PM	Microsoft Ed



```
pytorch-gpu) C:\Users\Lin_Y>
```

...s/it]
...s/it]
... of two supported pack
...d package manager as it will provide y
... including Python and pip.
...ical installer for PyTorch 3.x. Click on t





C:\Users\Lin_Y\Downloads\serpnext-develop

s:\Lin_Y\Downloads\serpnext-develop\untitled2.py

temp.py × setup.py × untitled0.py* × untitled1.py* × untitled2.py* ×

```
# -*- coding: utf-8 -*-
"""
Created on Sat Dec 24 20:48:34 2022

@author: Lin_Y
"""

from big_sleep import Imagine

dream = Imagine(
    text = "an armchair in the form of pikachu|an armchair imitating pikachu|abstrac
    lr = 5e-2,
    save_every = 25,
    save_progress = True
)

dream()
```

Source Console Object

Usage

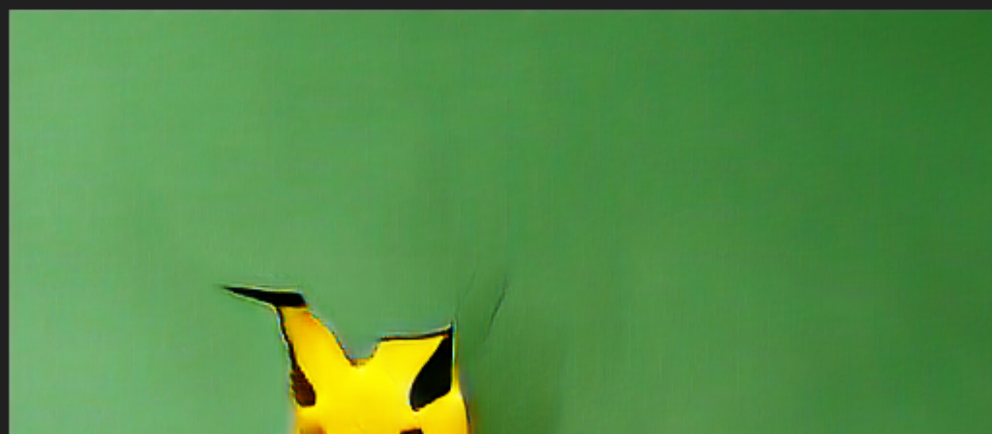
Here you can get help of any object by pressing **Ctrl+I** in front of it, either on the Editor or the Console.

Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in *Preferences > Help*.

New to Spyder? Read our [tutorial](#)

Help Variable Explorer Plots Files

an_armchair_in_the_form_o...



8:	68%	715/1050	[06:28<02:56,	1.90it/s]
8:	68%	716/1050	[06:28<02:56,	1.90it/s]
9:	68%	716/1050	[06:29<02:56,	1.90it/s]
9:	68%	717/1050	[06:29<02:56,	1.89it/s]
9:	68%	717/1050	[06:29<02:56,	1.89it/s]
9:	68%	718/1050	[06:29<02:54,	1.90it/s]
3:	68%	718/1050	[06:30<02:54,	1.90it/s]
3:	68%	719/1050	[06:30<02:53,	1.91it/s]
2:	68%	719/1050	[06:30<02:53,	1.91it/s]
2:	69%	720/1050	[06:30<02:52,	1.91it/s]
1:	69%	720/1050	[06:31<02:52,	1.91it/s]
1:	69%	721/1050	[06:31<02:51,	1.92it/s]
1:	69%	721/1050	[06:31<02:51,	1.92it/s]
1:	69%	722/1050	[06:31<02:51,	1.91it/s]
9:	69%	722/1050	[06:32<02:51,	1.91it/s]
9:	69%	723/1050	[06:32<02:50,	1.91it/s]

