AI Art Neural Constellation: Revealing the Collective and Contrastive State of AI-Generated and Human Art

Faizan Farooq Khan* KAUST
Diana Kim* KAUST
Divyansh Jha KAUST
Youssef Mohamed KAUST
Hanna H Chang KAUST
Ahmed Elgammal Rutgers University
Luba Elliott ELLUBA
Mohamed Elhoseiny KAUST

Abstract

Discovering the creative potentials of a random signal to various artistic expressions in aesthetic and conceptual richness is a ground for the recent success of generative machine learning as a way of art creation. To understand the new artistic medium better, in this work, we comprehensively analyze AI-generated art within the context of human art heritage using our dataset, "ArtConstellation," comprising annotations for 6,000 WikiArt and 3,200 AI-generated artworks. After training various generative models, we compare the produced art samples with WikiArt data using the last hidden layer of a deep-CNN trained for style classification. By interpreting neural representations with important artistic concepts like Wölfflin's principles, we find that AI-generated artworks align with modern period art concepts (1800 - 2000). Out-Of-Distribution (OOD) and In-Distribution (ID) detection in CLIP space reveal that AI-generated art is ID to human art with landscapes and geometric abstract figures but OOD with deformed and twisted figures, showcasing unique characteristics. A human survey on emotional experience indicates color composition and familiar subjects as key factors in likability and emotions. We introduce our methodologies and dataset, "ArtNeural-Constellation," as a framework for contrasting human and AI-generated art. Code and data are available here.

1. Introduction

Recent advancements in machine learning have transformed problem-solving across various domains, including art. Generative ML algorithms like GAN (Generative Adversarial Network) [6] offer artists a scientific tool to explore creativity from pure randomness to deliberate authorship. While generative art has historical roots dating back to projects like Mozart's *Musikalisches Würfelspiel* in 1792 and Harold Cohen's autonomous art systems in the 1970s, modern generative AI models are gaining attention despite controversies surrounding machine-generated art. Despite concerns about authorship and copyright, the aesthetic and conceptual richness of machine learning art is undeniable. Notable instances include the sale of the first AI artwork (*Portrait of Edmond de Belamy*) for over $\$400K$ in 2018, an AI-generated piece (*Théâtre D'opéra Spatial*) winning first place in the 2022 Colorado State Fair Fine Art competition, and a MoMA exhibition in 2023 featuring a large-scale AI-generated work by Refik Anadol titled "*Unsupervised*."

At this pivotal moment in the intersection of AI and art, understanding the distinctions between human-created and AI-generated artworks is crucial [1]. In this paper, we delve into various art principles to highlight the differences and similarities between the two, aiming to deepen our comprehension of AI as a new artistic medium. By focusing on foundational art principle concepts, we lay the groundwork for a comparative study that can inform future analyses and the creation of next-generation generative art.

Art styles, or movements, have long been essential in art history, reflecting cultural shifts, technological advancements, and periods of innovation. For instance, inventions like paint tubes and photography were pivotal for the emergence of movements like Impressionism and Cubism, respectively. Heinrich Wölfflin's theory [21], a cornerstone in art historical analysis, provides a framework for understanding stylistic patterns and changes. His comparative method, based on five fundamental visual principles—linear vs. painterly, planar vs. recessional, closed-form vs. open-form, multiplicity vs. unity, and absolute clarity vs. relative clarity—has become standard in art history training [17], offering insights into distinguishing between artistic periods such as the Renaissance and Baroque.

Recent studies [4, 12, 13] have analyzed deep neural net representations of artwork, revealing correlations with established artistic principles. Elgammal et al.[4] demon-

* indicates equal contribution.

[1] In this paper, the terms "human art" and "AI-generated art" refer to the two end-poles indicating zero and 100% of the intervention degrees of artificial automaton in art creation.

strated high correlations between the last hidden space of a deep CNN and Wölfflin's principles of art history, while [12, 13] developed a deep neural network aligning dimensions with 58 artistic concepts. While Wölfflin's principles were initially for distinguishing Baroque and Renaissance styles[21], the broader visual elements serve as universal semantic structures in art analysis [15, 20]; we term these "general art principles" in this paper.

Inspired by the studies [4, 12, 13], we conducted statistical analysis for the art principles after quantifying the art samples on the neural net's hidden layers to answer the following questions: (1) how can we represent individual human and AI-generated art samples by the art principles?, (2) what are their collective states, so how does AI-generated art differ from human artwork?, and (3) can we detect generative art instances visually outlying from the general human art population? Through Out-Of-Distribution(OOD) detection, we specify distinctive visual features that AI-generated art uniquely has. Pre-trained CLIP [16] is used for the analysis as a base neural net representation, motivated by its strong capacity to encode visual concepts in images.

Our contrastive evaluation is not limited to the formal analysis. Two other dimensions of time and emotions/likability are also investigated to identify a dominant period for human art samples that are visually closer to machine art. And, we also check if the machine generative art influences emotional experiences in art appreciation. The main philosophy of this work is to focus on the collective state of machine art from the five dimensions after conducting individual analyses for massive human and machine art samples generated by various state-of-the-art models. Finally, we evaluate the difference between human and AI-generated art, highlighting the key findings below.

- **Wölfflin's art principles**: AI-generated artworks are related more to the concepts which are originally formulated to characterize the Baroque style by Wölfflin.
- **General art principles**: compared to human art, AI artworks are visually related more to the art principle concepts for modern art.
- **Out-Of-Distribution (OOD) analysis**: human and AI-generated artworks are ID when they depict landscapes and geometric abstract figures, but when deformed and twisted figuration is generated by machines, the art samples are detected as OOD to human art.
- **Time analysis**: The majority (70%) of the AI-generated art is visually closest to the modern period art made from 1850 to 2000.
- **Likability and emotional experience**: AI-generated artworks have higher likability when they are visually closer to or resemble human art. And, like human art, they also evoke a diverse range of emotions.

Along with the findings, we propose massive data collected for this study as a whole analytical framework to contrast human and AI-generated art. For 6,000 human art from the eleventh to twentieth century and 3,200 generated art pieces from eight different generative models, we collected annotations about Wölfflin's art principles, likability, and emotions resulting in 262,000 annotations for human and AI-generated art. The evaluated AI-generated artworks were generated using StyleGAN-based models with different creative loss functions—(1) the adversarial objective of the original StyleGAN [10, 11], (2) of Creative Adversarial Network (CAN) [3], and (3) of Creative Walk Adversarial Network (CWAN) [9], and two state-of-the-art generative models: Vector-Quantized GAN (VQ-GAN) [5] and Denoising Diffusion Probabilistic Models (DDPM) [8].

2. Related Work

2.1. Art Generation Models

Our study relies on StyleGAN [10, 11], an advanced Generative Adversarial Network (GAN) [6], for art generation, allowing for hierarchical control over global style and local details. GAN algorithms have evolved from imitative to creative, as demonstrated in the literature, by modifying loss functions to encourage the production of novel and semantically meaningful content [3, 7, 9, 14]. We employ these networks to foster machine creativity by training StyleGAN architectures with creative objectives[3, 9]. Additionally, we examine two state-of-the-art image synthesis models: VQ-GAN [5], leveraging transformer architecture for high-resolution image synthesis, and DDPM [8], which progressively converts noise distributions into target data distributions to generate high-quality images.

2.2. Formal Analysis of Art in Neural Nets

In our study, art samples undergo encoding in a deep neural network, yielding quantitative representations interpreted with art principles to glean semantic knowledge about AI-generated art. Inspired by prior works like [4, 13], which showcased deep-CNNs' ability to capture smooth visual transitions over time in their last hidden layers during style classification training, we train deep neural networks for 21 human art movements. Leveraging Wölfflin's and general art principles, these approaches demonstrate the potential of deep neural nets as computational frameworks for formal art analysis. Notably, our work marks the first exploration utilizing art principles to compare human and AI-generated pieces in a neural net space.

Expanding beyond art principles, our neural net analysis delves into general semantics by considering CLIP space [16]. We observed a clear separation between human and AI-generated art in CLIP vision space, prompting an Out-Of-Distribution (OOD) analysis. OOD methods are crucial for ensuring safe and reliable machine learning. In this paper, the nearest neighbor OOD detection method [19] employed to identify AI-generated samples significantly distant from their nearest neighbors in the WikiArt dataset.

2.3. Emotional Analysis

Emotional response plays a pivotal role in art appreciation [18], intriguing both deep learning and computational art communities in understanding the relationship between visual features and emotion [2]. Recent efforts, such as ArtEmis by Achlioptas et al. [1], provide a large-scale dataset of emotional reactions to visual artwork, facilitating machine learning models for predicting dominant emotions from images or text. In our study, we conduct a human survey to explore how individuals perceive emotions in AI-generated art compared to human art, extending the ArtEmis dataset with new annotations for AI-generated art. These annotations adhere to the emotional categories and survey questions in ArtEmis, providing insights into emotional construction from both AI-generated and human art.

3. Experimental Setup

We test various neural net models and semantic spaces. For image generation, we train variants of StyleGANs, VQ-GAN, and DDPM from scratch on the WikiArt dataset. For art comparison, we consider Wölfflin's and general art principles, time, and emotions. All representative CNNs, including VGG-Nets, ResNets, and vision transformers, are trained for style classification, and CLIP (ViT-B/32) is adopted for deep nearest neighbor OOD analysis [19].

In our study, we construct six GAN-type models for art generation—based on the two versions of StyleGAN-1 and StyleGAN-2 [10, 11], combining three objectives (1) the adversarial objective of the original StyleGAN, (2) Creative Adversarial Network (CAN) [3], and (3) Creative Walk Adversarial Network (CWAN) [9]. Additionally, we include DDPM [8], which uses diffusion processes for high-quality image synthesis, and VQ-GAN [5], combining a Variational Autoencoder (VAE) with a Vector Quantization layer. In the later sections, we will bold the original Style**GAN**-1&2 to avoid confusion with StyleCAN-1&2. All models are trained on the WikiArt dataset.

We prepare two sets of art data: machine-generated art and digital human artworks. For machine-generated art, we sample 400 artworks from each of the eight trained models. We select four groups of 100 images each based on criteria such as nearest and farthest neighbor distance, shape entropy, and randomness from a total of 10,000 generated samples. For digital human art heritage, we select 6,000 art samples from the WikiArt dataset spanning from the 11th to the 20th century. They are western art except for Japanese Ukiyo-e.

3.1. Semantic Spaces for Contrastive Analysis

We analyze art samples using Wölfflin's principles, initially devised to differentiate Renaissance and Baroque styles. These principles include five pairs of opposing concepts:

Linearly vs. Painterly: Linear paintings feature clear boundaries and isolated objects, while painterly paintings depict blurry outlines and swift brushstrokes.

Planar vs. Recessional: Planar paintings arrange objects parallel to the canvas plane, while recessional paintings depict objects at angles, emphasizing spatial depth.

Closed-form vs. Open-form: Closed-form paintings have balanced figures within the frame, open-form paintings feature figures cut off, suggesting space beyond the frame.

Multiplicity vs. Unity: Multiplicity paintings have distinct parts with independent features, while unity paintings feature elements blending together as a coherent whole.

Absolute Clarity vs. Relative Clarity: Absolute clarity offers realistic representation for clear object forms, while relative clarity enhances visual effects in a holistic view.

These principles aid in capturing visual differences between human and AI-generated art, guiding our analysis.

We gathered data on Wölfflin's principles by training annotators to identify paintings based on these concepts. An interface presented descriptions of each principle, and annotators rated paintings on a scale of 1 to 5 (1: clear Linearly, 5: clear Painterly). We collected five ratings per painting to ensure accuracy. Ratings were averaged and normalized between 0 and 1, with lower scores indicating linear characteristics and higher indicating painterly characteristics.

General Art Principles: Along with Wölfflin's five principles, general art principles such as shape, color, texture, and space are also considered. Our analysis framework is based on proxy-space [12, 13], a deep neural network trained to quantify the relatedness of input paintings to 58 art principle concepts in its last hidden layer. Each dimension in proxy-space is aligned with a semantic, enabling direct statistical analysis for human and generated artworks. We use a subset of 15 visual concepts with an AUC performance of more than 0.75, including non-representational, representational, geometric, abstract, planar, closed, open, rough, perspective, broken, thin, flat, distorted, linear, and ambiguous.

Out-Of-Distribution (OOD) Analysis: Machine learning models, particularly GANs, aim to learn the distribution of real art but often produce AI-generated art that differs noticeably from human art. Machine spaces can help identify outlier samples and highlight visual similarities or differences between machine and human art. We conduct a qualitative and comparative examination based on Out-Of-Distribution (OOD) analysis, distinguishing between OOD and In-Distribution (ID) instances of machine-generated art compared to human art. Using CLIP vision space for its robust visual encoding capacity, we analyze the features determining ID or OOD.

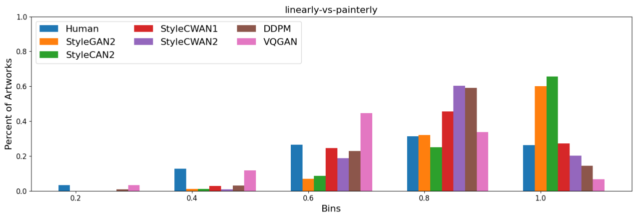

Figure 1. The distribution example of the first principle: linearly vs. painterly. The bins represent the values from the left value to the right value. We can see that AI-generated art is less uniform than human art for this Wölfflin's principle. Machine artworks are highly populated on the right side concepts more in general. Figures for other principles are in supplementary section C.

Time Analysis: Deep-CNN style classifiers can encode the smooth visual transition of art over time in its hidden layers. Even though time information is not used in training, the hidden representation learned a positive correlation with the year of art creation [4]. We estimate the time period of AI-generated art by checking the year of the human art which is spatially closest to each of AI artworks in a neural space.
Emotion and Likability: In our experiment, we investigate likability and emotion in art appreciation. Specifically, we aim to determine if AI-generated art is more liked or disliked, identify visual features prominent in highly liked art, and assess the range of emotions evoked by generative artworks compared to human art. Emotional data for human art was obtained from the ArtEmis dataset, while Mechanical Turk participants provided emotional responses and likability ratings for all 3,200 AI-generated artworks.
Deep Neural Nets: [4] is seminal in computational art analysis, showing how a deep-CNN style classifier's hidden space captures visual transitions across art movements. Inspired by this, we trained various neural networks to classify style and analyzed Wölfflin's principles and time from the hidden space. Extracting representations from the last hidden layer after ReLU activation, we applied PCA to examine principal components covering 95% of the variance. ResNet50 showed the highest style accuracy and maximum correlation with time, thus was chosen for our analysis.

4. Experimental Results

4.1. Wölfflin's Principles-based Analysis

We reaffirmed that Wölfflin's principles are implicitly learned in the last hidden layer of CNN style classifiers. For example, the Pearson correlation coefficients (PCCs) between Wölfflin's concepts and ResNet50 features for WikiArt art are: linearly-painterly: -0.21, planar-recessional: -0.19, closed-open: 0.2, multiplicity-unity: 0.25, absolute-relative clarity: 0.33. These are the maximum absolute PCCs among the top 30 principal components (95% variance). Correlation diminishes after the 30th

Table 1. The table displays the average absolute maximum PCCs of the first 30 PCA components for human and AI-generated art across various architectures. The highest PCC for each art type is bolded, with the second highest in red. Human art consistently exhibits higher PCC values for all of Wölfflin's principles across all architectures. Models labeled with "+2" denote extended ConvNet models with additional hidden layers.

Architecture	Human Art	SG2	SC2	CW1	CW2	DDPM	VQ-GAN
ResNet50	**.236**	.114	.126	.128	.124	.126	.122
ResNet50+2	**.186**	.114	.084	.130	.118	.096	.122
ResNet101	**.262**	.122	.106	.108	.128	.100	.116
ResNet101+2	**.170**	.096	.094	.134	.116	.092	.118
VGG16	**.300**	.144	.110	.112	.134	.108	.126
VGG16+2	**.344**	.118	.112	.116	.124	.108	.122
ViT-S	**.242**	.110	.108	.142	.102	.118	.146
ViT-B	**.256**	.136	.122	.120	.110	.102	.116
ViT-L	**.224**	.142	.110	.124	.132	.116	.136

component. Human and AI-generated art are compared using Wölfflin's principles. In Table 1, we averaged the maximum absolute Pearson correlation coefficients (PCCs) in the top 30 principal components for the five principles. Human artworks show a higher correlation than AI-generated images consistently. The lower PCC values for AI art suggest it varies within smaller ranges across Wölfflin's conceptual poles compared to human art. Fig. 1 illustrates that AI art is more concentrated around higher values, particularly around 0.4. Concepts associated with Renaissance style are less prevalent in AI art, with mean values biased towards concepts like "painterly, relative clarity, unity, open". This bias, coupled with smaller deviations across principles, suggests AI-generated art shares visual characteristics with modern period human art (1800 - 2000).

4.2. General Art Principles-based Analysis

To compare human and AI-generated artworks based on general art principles, we utilized the same neural network as the original framework [12], to quantify paintings' relatedness to each of the 15 visual concepts. We normalized the AI and human embeddings using human artworks' means and standard deviations to establish a unified scale, referred to as "standardized values by human art." Kernel Density Estimation (KDE) and hypothesis testing were conducted on the normalized space to assess the statistical differences between AI and human art samples for each concept. Additional details on the experimental procedure are provided in supplementary section D.1.

The densities in Fig. 2 show the differences between human and generated artworks across various visual concepts and their significance. We combined samples from all eight models to obtain a single representative density of AI art for each concept, using a Gaussian kernel with $\sigma = 0.5$ for KDE. Hypothesis testing revealed significant differences between the centers of generated and human art for most visual concepts, except for "geometric", "open", and "linear".

One interesting observation is that the generated arts are

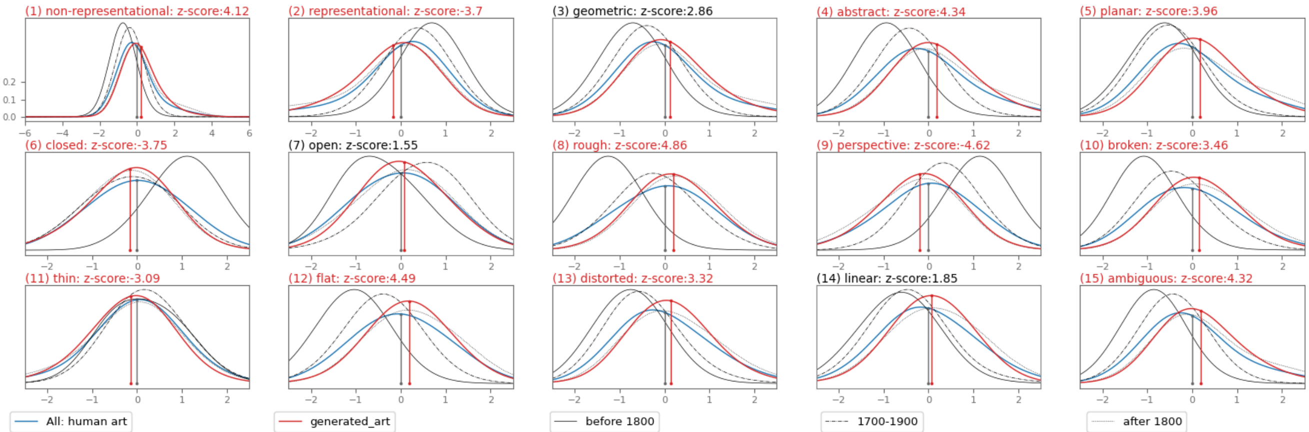

Figure 2. For each visual concept, the densities for human (in blue) and generated art (in red) are estimated to contrast their centers. Dotted densities are also added to show how human artworks visually changed from classical to modern times and to see where the generated art falls on the historical transitions. The plots' ranges are magnified around centers, and the names of the concepts are colored red when the two population means are significantly different. We found AI-generated art centers different from the human centers for most of the visual concepts except for "geometric", "open", and "linear"; in general, generated artworks are visually biased toward modern periods artworks.

visually biased toward modern periods' human art. We computed KDE after dividing the human samples into three groups by the periods: (1) before 1800, (2) 1700 – 1800, and (3) after 1800. We observed definite linear transitions of human art for visual concepts across time (dotted and black lines in Fig. 2). And, the visual bias between AI and human art is also aligned with the direction of the changes. For example, as human art is getting less "representational" closer to the modern period, the generated art is less "representational" compared to human art; the same phenomenon is observed for all the visual concepts. To confirm this observation, we tested the same hypothesis by dividing human samples into two groups: before 1800 and (3) after 1800. The center of generated samples is significantly different from the human arts as human art samples were drawn before the 1800s while the center of modern art is not much different from generated art. The results show that generated art is visually biased to modern period art from the perspectives of the principle concepts. The tendency is observed again for individual AI models and all hypothesis testing results are presented in supplementary section D.2.

To clarify the relation between human and generated art further, we examined the following questions on proxy space: (1) what are the nearest and farthest generated samples to human art, (2) how do they look depending on the distances, and (3) lastly, we check whether the generated arts are valued within the range of human arts or beyond on each axis of 15 visual concepts. Based on standardized values by human art, the most insignificant and significant AI-generated artworks from human art are sorted and analyzed. From the insignificant generative arts, we observed a pattern that AI-generated samples represent the typical subjects in human art: portraits and landscapes, but not in delicate or detailed expressions as much as in human art. Another result to note is that for all visual concepts, significant generated samples are observed on the same side of the bias toward modern arts, but AI-generated art samples are not valued beyond the extremes of human samples in both negative and positive sides. This indicates that AI-generated art is within the distribution of human art at least from the aspects of the concepts of proxy-space, but its visual characteristic is biased toward modern styles. In supplementary section H, the art examples are presented.

4.3. Out-Of-Distribution (OOD) Analysis on CLIP

Using CLIP's [16] robust representation of visual concepts, we conducted an OOD analysis between human and generated art. Initial trials showed a clear separation between human and generated samples in CLIP space, with one-third of generated samples identified as OOD to human art at 95% PCA space. To further explore the space, we conducted OOD detection across different PCA spaces (20%, 30%, 50%, 70%, 95%, 100%). OOD samples were detected based on kNN distances computed between human vs. human and AI-generated vs. human samples [19]. OOD samples were identified if their kNN distances exceeded a threshold determined by a 5% false OOD probability. Fig. 3 illustrates the variation in OOD counts and the changing t-SNE arrangements of generated and human arts with PCA. Across all models, similar patterns emerged, with the lowest number of OOD samples observed in lower dimensions (20%-50% PCA), followed by an increase in OOD samples as more principal axes were considered. This trend was reflected in t-SNE plots for 50%, 70%, and 95% PCA, indicating that top principal axes encode shared semantics, while differentiation occurs in later PCA axes.

Different PCA axes in CLIP represent various visual semantics, leading to variations in the spatial arrangements

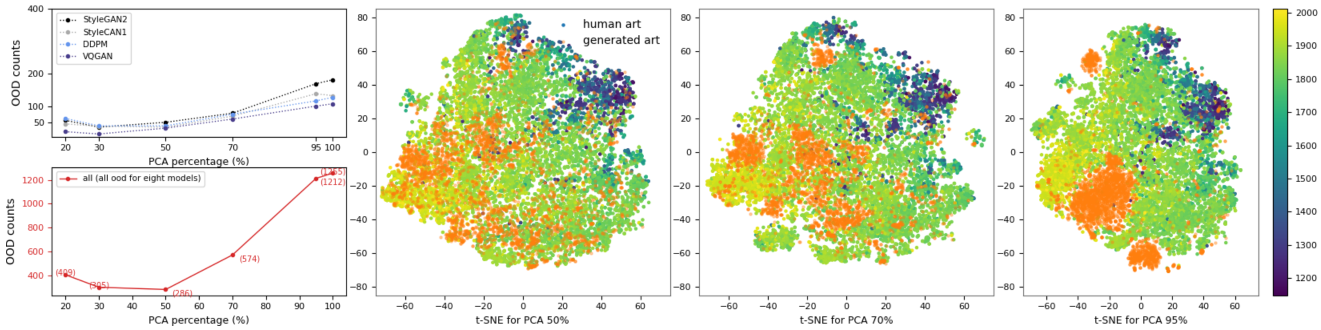

Figure 3. For the various CLIP spaces by different PCA dimensional reductions, OOD states are compared along with t-SNE representations (perplexity 30): human (greenish dots) and generated arts (orange dots). We found that for the top principal axes from 1 to 17 (50% data variance), OOD is relatively few, but the number of OOD is increased as more additional marginal dimensions are considered.

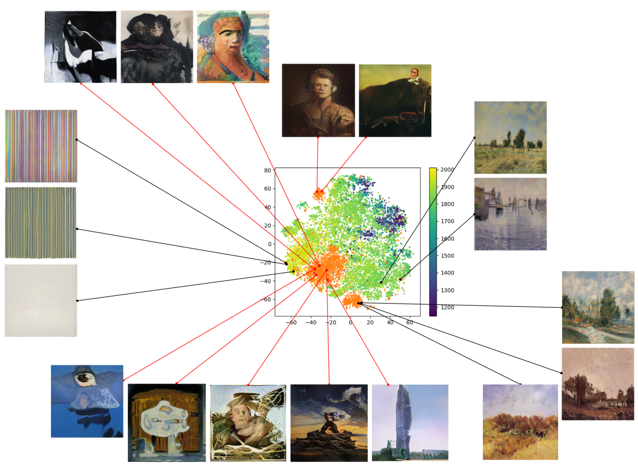

Figure 4. The most ten OOD (red lines) and eight ID samples (black lines) in 95% PCA are presented along with their t-SNE locations. All ID samples show landscape or abstract and geometric patterns, but OOD samples show irregular and twisted figures.

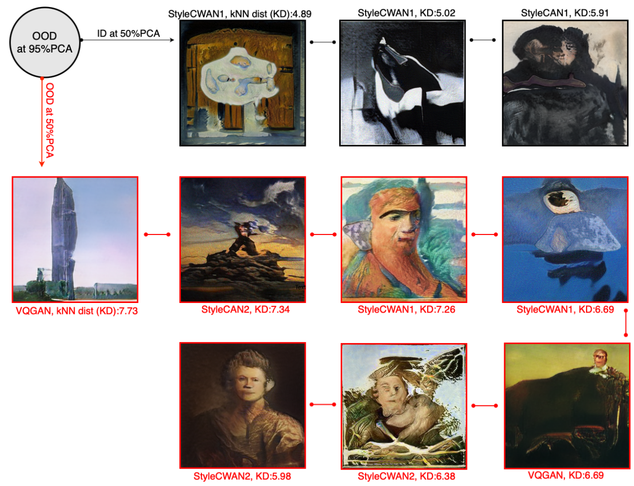

Figure 5. OODness in 50% PCA space is rechecked for the AI-generated samples detected as most OOD in 95% PCA space. The three abstract paintings in the first row are newly detected as ID, while the other generated images in the second and third rows are still OOD. The incomplete and unbalanced figuration in the seven samples is what is hardly seen in human art in general.

of arts across PCA spaces. With larger PCA spaces containing more profound semantics, we sought to identify (1) what generated samples remain close to human art when encoded and compared by the various semantics in the large PCA space (95% variance: $1-256$ axes). and (2) what generated samples are distant to human art when they are compared by a few but essential semantics consisting of the top PCA space (50% variance: $1-17$ axes)? We aimed to characterize shared and distinctive features between human and AI-generated art by addressing these questions.

To identify the generated samples closest to human art in a large PCA space (95% variance), we collected the most In-Distribution (ID) samples and Out-of-Distribution (OOD) samples. These samples were then analyzed in a smaller PCA space (50% variance) to observe changes in their OODness. The t-SNE plot in Fig. 4 illustrates ten OOD samples (red lines) and eight ID-generated samples (black lines) on the t-SNE plot in 95% PCA. Notably, ID samples predominantly depict landscapes or abstract patterns, while OOD samples feature irregular and somewhat twisted figures. When analyzed in 50% PCA, all ID samples remained ID, while some OOD samples transitioned to ID in the new space, as shown in Fig. 5. The ten OOD samples in 95% PCA become three ID and seven OOD samples in 50% PCA. This shift indicates visual distinctions between the two groups: the first group, initially OOD in 95% PCA but ID in 50% PCA, appeared too abstract to discern any figurative objects. Conversely, the second group contained recognizable figures such as towers, portraits, and landscapes, albeit incomplete and malformed.

The features observed in the sub-samples were consistent for all other OOD and ID samples. All ID-generated images in 95% PCA depict one of the subjects: landscape, portrait, and geometric abstract. All OOD-generated images in 50% PCA contain recognizable objects, but they are somewhat unbalanced, abnormal, and incomplete. Based on the observation, we confirmed two things: (1) human and

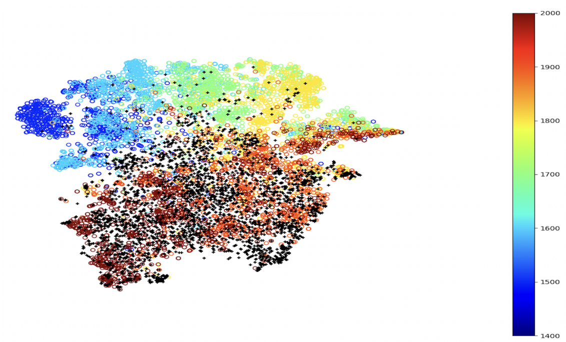

Figure 6. t-SNE visualizations of ResNet50 features of AI art (in black) are overlayed with human art.

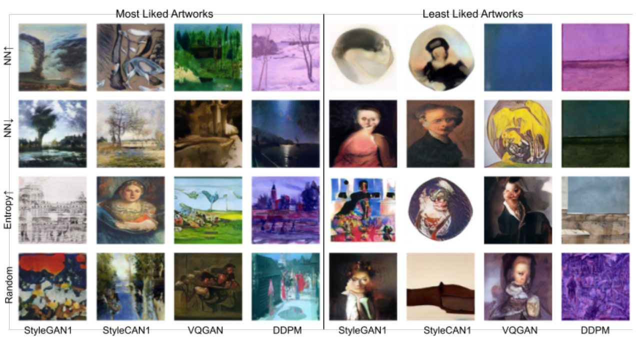

Figure 7. The most and least liked AI-generated artworks. There was a constant feature that affected likability throughout all models: the highly likable images are mostly the typical landscape composed of horizontally lined blue sky, water, and trees whereas the low-scored images are deformed and distorted portraits.

generated art have subjects in common, but (2) some generated artworks get visually distinct from human art when they try to reproduce the subjects but fail to have the delicacy and variants as much as human art. It generates incomplete and reduced forms but presents a new artistic style.

4.4. Time Analysis

The t-SNE visualization in Fig. 6 shows that AI-generated art(black dots) tends to cluster with modern human art, supporting the observation in section 4.2 of visual proximity between them. To figure out the specific period of modern art closest to our AI-generated art, we computed the following. We first find the five nearest neighbors from human art in the ResNet50 space for each of the generated artworks and then assign the average year of these five nearest human art pieces of art to the generated artwork. We observed that the average year of human art from the WikiArt dataset was 1869, but it was 1891 for the generated art. We notice that the majority (70%) of the AI-generated art was estimated to be in the period from 1850 to 2000. The average year of generated art was 1882 for Style**GAN**-2, 1894 for StyleCAN-2, 1887 for StyleCWAN-1, 1893 for StyleCWAN-2, 1885 for VQ-GAN, and 1902 from DDPM.

Table 2. Turing and likability results across different groups and models reveal interesting trends. StyleCWAN consistently showed higher mean likability compared to other StyleGAN variants, with the LowestNN group, close to human art in ResNet50, exhibiting the highest likability. DDPM artworks garnered the highest mean likability across all groups. In the Turing test, over 50% of participants believed that the machine artworks were created by human artists. For reference, [3] reported Turing test rates of 85%, 41%, and 62% for three human art sets: Abstract Expressionist, Art Basel, and Artists Sets Combined. Qualitative AI art examples and Turing results are provided in supplementary section E.

	Likability					Truing test
models	Q1-mean/std	NN↑	NN↓	Entropy↑	Random	Q2-human artists (%)
Style**GAN**-1	3.12/0.58	3.07	3.36	3.00	3.06	55.53
Style**GAN**-2	3.02/0.67	2.89	3.31	2.79	3.09	53.80
StyleCAN-1	3.20/1.14	3.01	3.61	3.05	3.11	56.55
StyleCAN-2	3.23/0.61	3.27	3.34	3.11	3.21	57.70
StyleCWAN-1	3.29/1.12	3.15	3.67	3.15	3.17	58.63
StyleCWAN-2	3.40/1.10	3.30	3.61	3.33	3.35	64.00
VQ-GAN	3.57/1.03	3.55	3.65	3.57	3.52	65.90
DDPM	3.85/0.91	3.77	3.90	3.81	3.93	63.55

4.5. Beholder's Visual Experience

Likability and emotion are the last elements for the comparison of the artworks. We conducted a human survey and found that AI-generated artworks are more likable when they depict landscapes and portraits, typical genres in human art, and they evoke a diverse range of emotions, including anger, awe, contentment, disgust, excitement, fear, sadness, and amusement. Participants identified various visual features that contributed to the likability of generated art, such as unique color combinations, imaginative compositions, intricate details, and captivating narratives. These findings highlight the potential of AI-generative art in offering engaging and emotional impact on artistic experiences.

4.5.1 Likability

We surveyed likability for all 3,200 AI-generated artworks, asking participants two questions: Q1, "How much do you like this image?" and Q2, "Was the art created by a human artist or a machine?" Likability ratings ranged from 1 to 5, with mean scores above 3 indicating a neutral to positive response. In Table 2 we report mean values for the first question (likability) and the percentage of responses that believed human artists created the AI-generated samples. Fig. 7 illustrates the most and least liked machine art samples categorized by the four generative model groups.

We have three points to remark from Table 2. First, StyleCAN-1&2 and StyleCWAN-1&2 scored higher than Style**GAN**-1&2; in the survey, the models trained with creative losses resulted in 38% and 18% more people assigning a full score of 5 over the vanilla models. Second, DDPM is liked the most (score 3.81) than any other model. DDPM images are all figurative and representational. We observed that some samples quite resemble specific paintings by hu-

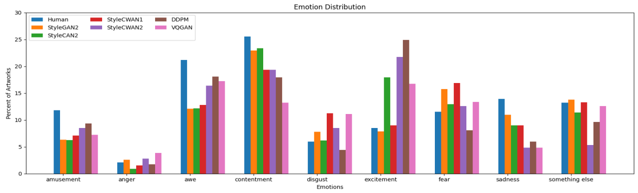

Figure 8. Emotional distribution results of the human survey on human and AI-generated artwork. Both AI-generated and human art evoke diverse emotions with negative emotions relatively rare.

man artists. For example, the most liked HighestNN in Fig. 7 reminds us of the winter landscapes of Alfred Sisley. We also found multiple DDPM variations of Marcel Duchamp's *Sad Young Man on a Train* of different tones in color. Lastly, in comparison of the two NN groups (Highest NN vs. Lowest NN), Lowest NN—visually closer to WikiArt—got higher likable scores than other groups for all the generative models. In our empirical observation, we found a correlation between the closeness to human art and likability. AI-generated artworks were more likable when they successfully depict the popular and familiar topics of landscape and portrait in human art, but they are disliked much when they fail to be drawn as a sound shape, like the deformed portraits in Fig. 7. According to groups and likability, we characterize our generative models empirically and provide detailed analyses in supplementary section F.

4.6. Emotional Distribution

We examined the collected data for emotion and investigated the diversity of emotions elicited by AI-generated art. As shown in Fig. 8, AI-generated art can construct diverse sets of emotions. To measure the diversity of these emotions, we calculated the entropy of emotions evoked by AI-generated art and compared it with human art. Human art has an entropy of 0.916, whereas the generated art from all models combined has an entropy of 0.938, indicating that generative art is capable of evoking diverse emotions. However, some differences exist in the distribution. For example, excitement accounts for around 17% of AI-generated art compared to 7.47% in human art. DDPM has the highest percentage of art evoking excitement of 24.90%. Awe accounts for 18.63% of human art, which is much higher than machine art, but contentment has the highest share of any emotion for both human and generated artworks.

Qualitative Analysis of Emotion: Based on the responses of emotion participants, we derived and analyzed common factors and elements of AI-generated artworks that constructed the different emotions. We find that color and composition are the key factors for emotions in the emotional narratives of participants. Also, some familiar and ordinary subjects make people comfortable and content by bringing a good memory related to them. Detailed analyses are provided in supplementary section G.

5. Conclusion

AI is reshaping both daily life and the landscape of art creation, sparking debates in the art community regarding the definition of art, authorship, and creativity. As AI-generated art gains traction through museum exhibitions and auctions, understanding its role as a new artistic medium becomes increasingly important. In this work, we delve into generative AI models, exploring their representation within the broader context of human art heritage. Through deep neural net representations, we aim to analyze, visualize, and interpret the constellations of AI-generated art, considering fundamental art principles and emotional experiences. Ultimately, our goal is to shed light on the current state of machine art.

From our analysis, AI-generated art simulates the most typical subjects in training data—in our case, landscape, portrait, and geometric abstracts often drawn from the 1800-2000 period samples in WikiArt—but the outcomes often consist of bizarre shapes, distorted human faces and bodies, and unrecognizable nature scenes especially when generated by GAN models. The incomplete figuration causes the bias toward the visual concepts related to modern period art in abstract and open forms. In OOD analysis, the deformed images were detected as significant outliers compared to human art images. The phenomenon might be caused by the GAN models' architectures and learning algorithms in practice, but the constraint comes to form a unique and distinctive visual style of AI-generative art. DDPM, on the other hand, showed better ability in producing art with realistic and delicate human and natural figures and they are liked the most by the participants in our survey.

Our results and empirical observations reveal current generative art's unique visual features. And, we understand that historically new art mediums have initiated new art movements or genres like the camera, which was introduced as a scientific instrument, gradually revealed its artistic potential. However, this study did not aim at positioning AI-generated art into a new abstract category. Rather than that, we demonstrated how AI-generated art differs from human art heritage based on various fundamental aspects of art history and aesthetics in fine neural net representations.

The AI-generated art samples in this work represent the case where art creation is solely based on the machine's random generation without artists' curation. However, the practical use of generative models in the art world is beyond that; existing middle of somewhere between pure randomness and the full realization of the author's free will. Also, in the developmental process of generative models, we have many choices for datasets and ML algorithms. Given all the possibilities and new algorithms, ML-based generative models possess endless potential for visual art. This work was about the starting point of the infinite spectrum, offering a baseline framework for both the art and ML communities to understand current and next-generation machine art.

References

[1] Panos Achlioptas, Maks Ovsjanikov, Kilichbek Haydarov, Mohamed Elhoseiny, and Leonidas Guibas. Artemis: Affective language for visual art. *CoRR*, abs/2101.07396, 2021. 3

[2] Eva Cetinic, Tomislav Lipic, and Sonja Grgic. A deep learning perspective on beauty, sentiment, and remembrance of art. *IEEE Access*, 7:73694–73710, 2019. 3

[3] Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. Can: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms, 2017. 2, 3, 7

[4] Ahmed Elgammal, Bingchen Liu, Diana Kim, Mohamed Elhoseiny, and Marian Mazzone. The shape of art history in the eyes of the machine. In *AAAI*, 2018. 1, 2, 4

[5] Patrick Esser, Robin Rombach, and Bjorn Ommer. Taming transformers for high-resolution image synthesis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 12873–12883, June 2021. 2, 3

[6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014. 1, 2

[7] Aaron Hertzmann. Toward modeling creative processes for algorithmic painting, 2022. 2

[8] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pages 6840–6851. Curran Associates, Inc., 2020. 2, 3

[9] Divyansh Jha, Kai Yi, Ivan Skorokhodov, and Mohamed Elhoseiny. Creative walk adversarial networks: Novel art generation with probabilistic random walk deviation from style norms. In *International Conference on Computational Creativity*, 2022. 2, 3

[10] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, 2019. 2, 3

[11] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Analyzing and improving the image quality of stylegan, 2020. 2, 3

[12] Diana Kim, Ahmed Elgammal, and Marian Mazzone. Proxy learning of visual concepts of fine art paintings from styles through language models. In *AAAI*, 2022. 1, 2, 3, 4

[13] Diana Kim, Bingchen Liu, Ahmed Elgammal, and Marian Mazzone. Finding principal semantics of style in art. In *2018 IEEE 12th International Conference on Semantic Computing (ICSC)*, 2018. 1, 2, 3

[14] Adeel Mufti, Biagio Antonelli, and Julius Monello. Conditional GANs for painting generation. In Wolfgang Osten and Dmitry P. Nikolaev, editors, *Twelfth International Conference on Machine Vision (ICMV 2019)*, volume 11433, pages 857 – 864. International Society for Optics and Photonics, SPIE, 2020. 2

[15] Otto G. Ocvirk. *Art fundamentals: Theory and practice*. McGraw-Hill, 1998. 2

[16] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision, 2021. 2, 5

[17] Higor Y. D. Sigaki, Matjaž Perc, and Haroldo V. Ribeiro. History of art paintings through the lens of entropy and complexity. *Proceedings of the National Academy of Sciences*, 115(37):E8585–E8594, 2018. 1

[18] Paul J Silvia. Emotional responses to art: From collation and arousal to cognition and emotion. *Review of general psychology*, 9(4):342–357, 2005. 3

[19] Yiyou Sun, Yifei Ming, Xiaojin Zhu, and Yixuan Li. Out-of-distribution detection with deep nearest neighbors. *ICML*, 2022. 2, 3, 5

[20] Joshua Charles Taylor. *Learning to Look: A Handbook for the Visual Arts*. The University of Chicago Press, 1957. 2

[21] Heinrich Wölfflin. *Principles of art history*. na, 1915. 1, 2