

Creating an Emotionally Aware Portrait System Prototype using Aesthetic Emotion Evaluations of AI Art Portraits

Nouf Abukhodair^{1,2*} Steve DiPaola^{1†}

¹School of Interactive Arts and Technology, Simon Fraser University, Surrey, BC V3T0A3, Canada; ²Department of Computer Science, King Abdulaziz University, Jeddah, Saudi Arabia
E-mail: nouf_abukhodair_2@sfu.ca, nabukhadair@kau.edu.sa

Abstract. Visual content has the ability to convey and impact human emotions. It is crucial to understanding the emotions being communicated and the ways in which they are implied by the visual elements in images. This study evaluates the aesthetic emotion of portrait art generated by our Generative AI Portraiture System. Using the Visual Aesthetic Wheel of Emotion (VAWE), aesthetic responses were documented and subsequently analyzed using heatmaps and circular histograms with the aim of identifying the emotions evoked by the generated portrait art. The data from 160 participants were used to categorize and validate VAWE's 20 emotions with selected AI portrait styles. The data were then used in a smaller self-portrait qualitative study to validate the developed prototype for an Emotionally Aware Portrait System, capable of generating a personalized stylization of a user's self-portrait, expressing a particular aesthetic emotional state from VAWE. The findings bring forth a new vision towards blending affective computing with computational creativity and enabling generative systems with awareness in terms of the emotions they wish their output to elicit.

Keywords: computational creativity, artificial intelligence, visual art, generative art, aesthetic emotions, Geneva Emotion Wheel
This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.
[DOI: 10.2352/J.Percept.Imaging.2025.8.000503]

1. INTRODUCTION

Artists have always used different colors and styles to convey emotions in their paintings. Moreover, researchers have found that colors and properties of lines affect users' emotions [8, 15, 17, 31]. The impact of art on individuals' emotions underscores the significance of visual style in shaping their perceptions and emotional reactions. Previous psychological research has indicated connections between rendering style and how individuals perceive and feel about an object. Furthermore, other studies have investigated the association between colors and emotions [8, 17, 31]. Due to the wide range of different functions of art, many different genres have emerged, such as landscapes, still life, and portraits, all defined by varying content. Many researchers have

attempted to identify the relationship between visual and sentimental elements in images using statistical approaches. In many cases, the range of sentiment that affects humans varies with image categories. Therefore, an individual model must be established for each image category to enhance the performance of sentiment prediction. However, collecting data on this phenomenon is one of the obstacles encountered by studies in this field [29].

Automatically detecting emotions evoked by art is of considerable importance and is challenging. This can be used to organize paintings by the emotions they evoke, to recommend paintings that accentuate or balance a particular mood, to suggest paintings of a specific style or genre that depict user-determined content in a specific affective state, and to enhance computational understanding and generation of visual art. This research aims to categorize portrait art generated from our system based on aesthetic emotion and discuss the development of a system capable of generating art that elicits desired emotional responses. By conducting a study that utilizes the Visual Aesthetic Wheel of Emotion (VAWE) [1], it was possible to determine the affective response to the portrait art generated using a Generative AI Portraiture System developed by the authors.

In the study, portraits were restricted to one sitter to reduce variability. The observer derives at least two different hedonic values from such artwork: the attractiveness of the depicted person and the artistic beauty of the image that relates to the way of presentation [28]. Leder et al. [20] conducted a series of studies demonstrating that both content and appreciation of the artist's style are crucial aspects of the aesthetic experience. Aesthetic appreciation of art should depend on how content and style are evaluated. In portraits, there is a variation as the depicted persons differ according to the context such as age, gender, viewpoint, facial expressions, accessories, and clothing, which all influence likability and elicited emotions. Style refers to the way artists depict the portrayed person and is deemed by some researchers as the most relevant feature of art [20].

2. BACKGROUND AND RELATED WORK

Research surrounding the perception of emotions in relation to artistic stimuli has gained traction over the past three

*<https://orcid.org/0009-0002-8790-0927>

†<https://orcid.org/0000-0001-7549-9545>

Received Mar. 8, 2024; accepted for publication Apr. 3, 2025; published online July 04, 2025. Associate Editor: Claus-Christian Carbon.

2575-8144/2025/8/000503/14/\$00.00

decades. In the field of art aesthetics, Scherer defines “aesthetic emotions” as “emotions elicited by an appreciation of the intrinsic qualities of natural beauty or the qualities of a work of art or artistic performance” [27]. Many interesting and relevant research initiatives are emerging at the intersection of AI and art. However, the comprehension and appreciation of art are still considered an exclusive human capability. The variety of activities and research initiatives related to AI and art can generally be divided into two categories: (1) how AI is used in the process of analyzing existing art and (2) how AI is used in the process of creating new art.

Developing quantitative methods for analyzing subjective aspects of perception and aesthetic emotions is particularly challenging in the context of art images. One of the major challenges in studying perceptual characteristics of art images is the development of annotated large-scale datasets with evaluation scores obtained through experimental studies. Cetinic et al. [6] used deep-learning-based quantitative methods to extract features not only related to aesthetic evaluation but also to the sentiment and the memorability of fine art images. Besides aesthetic evaluation, sentiment analysis is the most commonly addressed task in this domain. Mohammad and Kiritchenko [23] introduced WikiArt Emotions, which is a dataset of paintings that has annotations for various emotions evoked in the observer. In a similar vein, Alameda-Pineda et al. [3] introduced an approach to automatically recognize emotions elicited by abstract paintings using the MART dataset [33]—a collection of 500 abstract paintings labeled as evoking positive or negative sentiment. Most recently, Achlioptas et al. [2] introduced ArtEmis, a large-scale dataset of emotional reactions to visual artwork, including explanations of these emotions in language, and developed machine learning models for dominant emotion prediction from images or text.

Another area of research that is pertinent in this field is developing computational approaches for quantifying and predicting values of concepts, such as aesthetic emotion, which is especially difficult for art images. The importance of perception of a particular artwork does not only emerge from its visual properties but also greatly depends on the art historical context. For that reason, it is apparent that current approaches are limited as they only take into account visual image features. This also indicates that future research in this field has to strive towards a more holistic approach if we ought to build systems that can achieve a human-like understanding of art. There is much debate within the computational creativity field regarding the affect and emotion that are conveyed in the artwork produced by creative and generative AI systems. This challenge is due to aesthetic emotions being considered higher-level semantics and based on human subjectivity, which are difficult to model computationally [18].

Our research aims to evaluate and categorize generated AI art by integrating the concepts of aesthetic emotion that the artist intends to portray in a produced artwork into our system. To achieve this objective, two studies were conducted: first, to categorize our AI portrait artworks

using VAWE based on the aesthetic emotion they convey to the viewer and second, to understand how well VAWE categorized the AI portrait artwork and potential application areas our work can open up. This allows the growth of the knowledge base and cognitive model of our Generative AI Portraiture System [9–11, 14], which currently has aesthetic reasoning based on artists’ creative processes. The integration of emotion mapping into our system amplifies the potential of interactive visual art systems, enabling them to elicit a diverse spectrum of emotions in viewers.

The VAWE was developed as a domain-specific tool to measure aesthetic emotions elicited by visual art. The VAWE is an adaptation of the Geneva Emotion Wheel (GEW) [27], which is a graphical self-report measuring tool. It consists of a circular shape divided into 20 slices, each representing a distinct emotion family. These slices form clusters of five items, organized into four quadrants with similar emotion families. Each slice further subdivides into segments based on emotion intensity, with segments closer to the center indicating low intensity and those towards the periphery indicating high intensity. The VAWE has two dimensions: valence (horizontal axis) and arousal (vertical axis). Unlike the GEW, it omits the “other emotion” option to ensure measurement reliability, dedicating a circular section in the center to neutral emotional states [1].

While Russell’s circumplex model theoretically aims to represent the full range of core emotions within a continuous two-dimensional space, the VAWE, drawing from the GEW, utilizes a selected set of 20 emotion terms. This deliberate reduction in scope was intended to enhance the practicality and efficiency of VAWE for studies requiring rapid emotion ratings, enabling participants to quickly and easily categorize their aesthetic experiences. This structure allows participants to quickly orient themselves and select specific emotions with minimal cognitive effort. Unlike traditional linear rating scales, which can overwhelm participants with lengthy lists of options, the VAWE facilitates a more focused and efficient emotional reporting process. The VAWE is not intended to represent all subtle emotional nuances but rather to capture the most salient and relevant emotions in response to visual artworks.

The development of VAWE began with a comprehensive review of the literature on aesthetic emotions evoked by visual art and music. A set of candidate emotion terms was identified, refined through expert feedback, and validated through a field study with 60 participants. This study adapted the GEW framework to place the selected terms on the wheel using a systematic clustering approach. Unlike traditional tools that present a list of terms for rating, the VAWE organizes emotions in a visually intuitive way, making it well-suited for assessing complex aesthetic experiences [1].

3. GENERATIVE AI PORTRAITURE SYSTEM

The stylistic rendering of a portrait goes through many AI and computational stages and modules that we have written, curated, and refined over multiple years for our research. This is based on cognitive modeling of both creativity



Figure 1. Process flow of our Generative AI Portraiture System from raw source to final portrait.

and the fine art painting process [9–11, 14]. However, for overview and clarity, there are three main stages of processing (Figure 1). In the first phase, the original portrait (Fig. 1(a)) of the sitter image is preprocessed by using the modified Deep Dream (mDD) system [13, 22], which conducts successive passes on the image to create the baseline art style (Fig. 1(b)). While most DD systems use pretrained networks with object recognition data such as ImageNet, the researchers have implemented an mDD system and trained new models with cognitive-based creativity art generation, not object recognition, in mind, using paintings and drawings as training data. For this phase, we have amassed a fine art painting dataset of 160,000 labeled/categorized paintings from 3000 labeled artists for a total size of 67 gigabytes of artistic visual data (one of the largest in an AI research group). However, despite its size, our AI training experiments revealed a challenge: most fine artists produce fewer than 200 paintings in their lifetime except for outliers such as Picasso. Consequently, our dataset might not have been rigorous or extensive enough for advanced convolutional neural network training for art styles. To address this limitation, we used a method developed by one of our authors called hierarchical tight style and tile [12]. This system is considerably different from Stable Diffusion or Midjourney [4], which uses prompts for an emotion under the assumption that it is creating “good art” and further validating the emotion metric, which could result in neither assumption being realized. Our system is unique as it uses a series of AI and non-AI techniques based on the cognition model of art making that was curated by artists. This was paired with a significant emotion metric that was validated first with a large user base (non-self-portraits) and later with a smaller group on self-portraits. It should be noted that our mDD database was specifically curated for a correlated and complete example set related to our cognitive-based creativity goals and has no copyrighted material. This is a departure from modern datasets used for generative art (viz., Midjourney and Stable Diffusion) that sometimes have copyrighted materials without the permission of the artists involved and are not curated for a specific goal or use.

In the last phase, the source image created from the previous phase is further refined using our ePainterly system with a combination of Deep Style [13, 22] techniques as a

surface texture manipulator as well as a series of non-AI techniques in the field of non-photorealistic rendering (NPR) techniques such as particle systems, color palette manipulation, and stroke engine techniques (see Fig. 1(c)). This iterative process refines and completes the finished portrait style. The ePainterly module is an extension of the cognitive painting system Painterly [9], and it models the cognitive processes of artists based on years of research in this area. The NPR subclass of stroke-based rendering is used as the final part of our process to realize the internal mDD models with stroke-based output informed by historic art making. Specifically, in this example, the aesthetic advantages of this additional system include reducing noisy artifacts of the generated mDD output via cohesive stroke-based clustering and a better distributed color space.

4. GENERATIVE AI FOR VISUALS

The rapid progress in deep learning systems has accelerated the development of generative AI models that create intricate visual content. Midjourney and Stable Diffusion [4] are exemplary instances of this phenomenon. Midjourney employs a conditional generative adversarial network to generate visually captivating landscapes, leveraging its capacity to synthesize novel scenes from input data. On the other hand, Stable Diffusion employs diffusion processes to create high-resolution images with impressive detail and realism. These systems showcase the potential of generative AI to produce visually compelling content. In addition to our systems, our modified Stable Diffusion system was utilized as the last step on 33% of our portrait styles while preserving the likeness to the source image, which were then included in the dataset, as demonstrated in Figure 2. This step was thoroughly applied to each portrait to generate art closer to fine art portraits. This hand-curated process was heavily based on years of research and refinement in working with our AI system. It should be noted that while many artists and researchers are simply utilizing systems (such as Stable Diffusion, Dalle, and Midjourney) to both make fine art portraits and possibly use emotional keywords in prompts, this is a brute force method where they utilize a prompt on a large dataset and assume this will (1) create good art and (2) be valid as an emotion metric. However, this



Figure 2. Examples of generated portrait art (a1, b1, c1, and d1) augmented using Stable Diffusion (a2, b2, c2, and d2).

would possibly not yield the desired outputs. Our system does something drastically different—we use a series of AI and non-AI techniques based on the cognition model of art making curated by artists (as described above). Subsequently, we created a significant emotion metric that we validated first with a large user base (non-self-portrait) and then with a smaller group on self-portraits. There are many issues with only using diffusion-based systems for emotional portraits from ethical [26] to non-explainable AI and non-repeatability. Therefore, our approach uses Stable Diffusion systems only as a refiner for the series of tools we have created to emulate fine art portraiture.

5. OBJECTIVE: AI-GENERATED PORTRAIT STUDY

Artificial intelligence has increasingly been integrated into literary and artistic expressions with the ongoing advancements in science and technology. This integration has led to more AI-generated artworks, sparking nuanced discussions regarding their evaluation. Currently, the primary debate surrounding AI-generated art revolves around its ability to convey emotions that are comparable to those of human-created art. Despite limited textual research on this matter, several studies have examined the emotional responses elicited by AI artworks, albeit with a predominant focus on comparing AI and human artists [32]. In this study, portrait artworks created by the AI Portraiture System were categorized based on the aesthetic emotion they convey using VAWE [1] as the basic measurement tool, allowing the study participants to choose from 20 emotion types and 5 intensity levels.

The objective of this study was to categorize a large corpus of portrait art styles generated from our system based on aesthetic emotions, using VAWE to capture the felt emotions when presented with artwork specific to portraiture. All portrait artworks for the study were generated using the

Generative AI Portraiture System described in Section 2. These were then rated by the study participants as follows: 1. elicited emotions from the portrait art; 2. the aesthetic liking that the style of the artwork evoked as opposed to the content. A dataset with the artworks was then created.

The online study was developed using PsychoPy, a Python platform proposed by Peirce [24]. When the study design was completed, it was exported to a JavaScript-based platform Pavlovio. One of the main reasons for choosing PsychoPy/Pavlovio over other study design applications was the ability to create customized backend layouts and interactions for the multimodal setup to fit the circular design of VAWE. Using PsychoPy v2021.2.3, we designed and implemented an interactive tool to collect participants' annotations on each artwork using an interactive interface of the emotion wheel.

5.1 Methodology

Participants, consisting of 160 adults (52.5% women, 47.5% men; age range 19–75 years; $M = 42.6$, $SD = 13.2$), were recruited using Prolific, a crowdsourcing platform [25]. The prerequisites outlined in the online consent form were that participants be fluent in English, have at least two years of college education, and have normal or corrected-to-normal vision with no color blindness. Each participant annotated 30 pieces of artwork during an average study time of 15 minutes.

Each participant was required to use a desktop computer to run the study to ensure that they were viewing the artwork on an optimal screen size. A stable internet connection was required, and due to the size of the images, the artwork was preloaded while viewing the consent page. Next, a series of instruction screens were presented. The core of the study was the VAWE presented in Figure 3. It consists of 20 options of closely related emotion terms and a neutral option at the center. The terms were arranged in four corresponding group

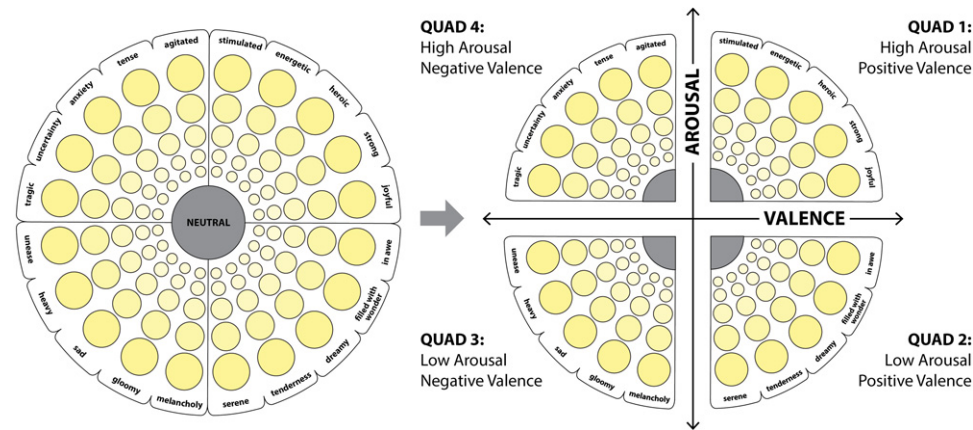


Figure 3. Visual Aesthetic Wheel of Emotion (VAWE).



Figure 4. Examples of two study images.

quadrants around the wheel based on valence and arousal to facilitate ease of annotation as shown in Fig. 3.

One hundred and twenty styles were chosen from the AI-generated styles. In the selection, we attempted to choose styles that represent different emotion categories. Each style was rendered using four different source images and presented as a set in different order to minimize the content effect/bias and the profile effect as shown in Figure 4.

The work focuses on examining emotions felt by the observer while viewing the portrait artwork set. Participants assessed the style (color, texture, feeling) of each portrait rather than interpreting the emotions expressed by the sitters in the set of four portraits created in the same style. To ensure that participants evaluate the style and not the emotion of the sitter's expression, the study specifically instructs them to evaluate the style (color, texture, feeling) as opposed to the content. This is also made clear thrice in the pre-instruction screens (see Figure 5, bottom right). They then rated the aesthetic likability of each portrait using a 5-point Likert scale. Emotional responses were measured using VAWE, where larger circles indicated more intense emotions. The

scales for each portrait set were presented consecutively on the right side of the screen in the following order: primary emotion, secondary emotion, and aesthetic likability (Fig. 5). Example instances were provided in advance to ensure clarity.

Prolific's crowdsourcing platform has been designed specifically to run academic studies. Upon reviewing other platforms, we decided to integrate Pavlovia with Prolific. Participants chose to do the task based on interest and the compensation provided. The compensation rate was 13 CAD (Canadian dollar) per hour. The research adhered closely to the guidelines and regulations outlined by Simon Fraser University's ethics board, with particular attention to protecting participant anonymity. A mechanism to avoid inattentive or malicious annotations (such as ignoring requirements: using a device without color or being colorblind. . .) was incorporated through two screens interspersed in the study to test users' cognitive attention. One screen tests whether the user can identify the primary color in an object while the other screen asks the user to specifically identify objects in a painting, both as multichoice and in a time-sensitive manner.

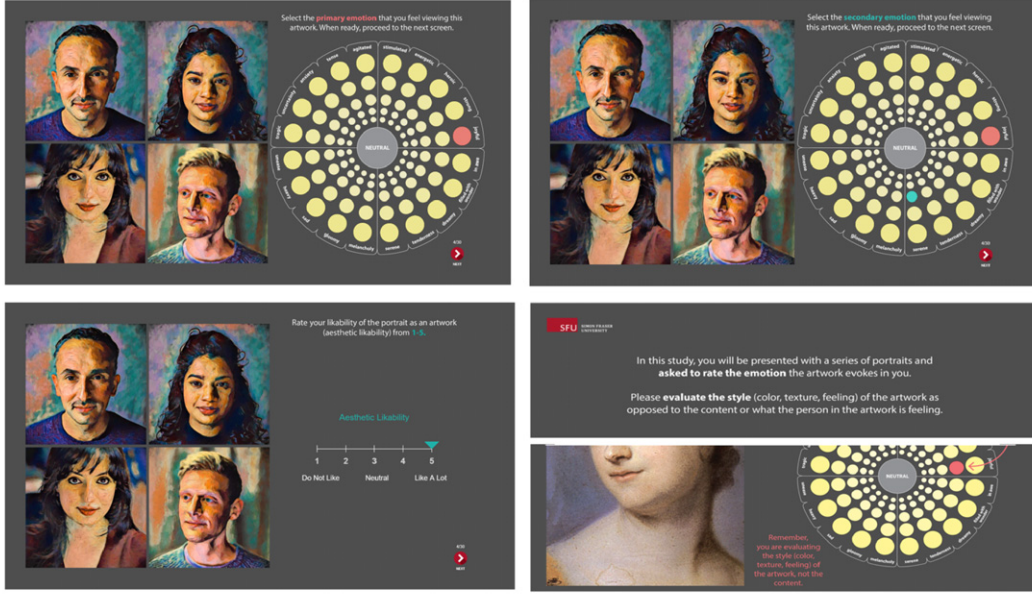


Figure 5. Examples of various instructional study screens for rating a portrait set. The bottom right shows parts of two screens displaying pre-instructions on evaluating painterly style rather than facial expression and content.

If a participant answered both questions incorrectly, their collected data was discarded. In total, 160 participants were split into four groups, with each group annotating 30 pieces of art. Forty participants evaluated each portrait set. A total of 4800 responses (for primary emotion, secondary emotion, and aesthetic likability) were obtained for the 120 pieces of portrait art sets included in the study.

6. DATA ANALYSIS

6.1 Quantifying Findings from VAWE

Previous research utilizing the GEW often analyzed emotion data with limited statistical depth, focusing on discrete emotion slices or isolated emotion intensity [5, 16, 19, 30]. For this study, a statistical measure of the goodness of fit is required for data collected with the VAWE tool; finding variances of perceived emotions on the wheel will allow us to compare emotion distinctness. The VAWE, as described by Abukhodair et al. [1], is a circumplex model of affect; therefore, circular statistics can be used to interpret its data. For the analysis, the researchers adapted the statistical method developed by Coyne et al. [7], which applies Mardia's vector method [21] for profile averaging. For each emotion (e) being tested, vector arithmetic is performed to sum all datapoints together. For n_e datapoints, with θ_i being the angular location of a datapoint and I_i its intensity, we can first calculate the resulting coordinates (X_e, Y_e) of the averaged emotion profile [7]:

$$x_e = \frac{1}{n_e} \sum (I_i \cdot \cos \theta_i) \quad (1)$$

$$Y_e = \frac{1}{n_e} \sum (I_i \sin \theta_i). \quad (2)$$

The resulting dimensional coordinates express the average valence (X_e) and arousal (Y_e) of the emotion according to the sample data. Expressing this vector in its polar coordinates yields the angular direction of the emotion profile (θ_e) and its intensity (I_e):

$$\theta_e = \tan^{-1} \left(\frac{Y_e}{X_e} \right) \quad (3)$$

$$I_e = \sqrt{X_e^2 + Y_e^2}. \quad (4)$$

This vector serves as the center of gravity of the data, where θ_e is the circular mean of the emotion profile, pointing to the predominantly perceived emotion, and I_e is the mean intensity, also indicating emotion distinctness. To determine the variance of the distribution, the *unweighted* versions of the above equations are applied to calculate the circular mean and its corresponding variance. This involves removing the intensity factor (I_i) and scaling ($1/n$) from Eqs. (1) and (2) and continuing to find the unweighted circular mean (θ'_e). The corresponding *unweighted circular variance* (V'_e) would then be expressed as follows:

$$V'_e = 1 - \frac{\sum \cos(\theta'_e - \theta_i)}{n_e}. \quad (5)$$

Considering that emotion intensity is the defining feature of the GEW compared to other circumplex emotion models, this formula is adapted to express the variance weighted by intensity or weighted circular variance (V_e):

$$V_e = 1 - \frac{\sum (I_i \cos(\theta_e - \theta_i))}{\sum (I_i)}. \quad (6)$$

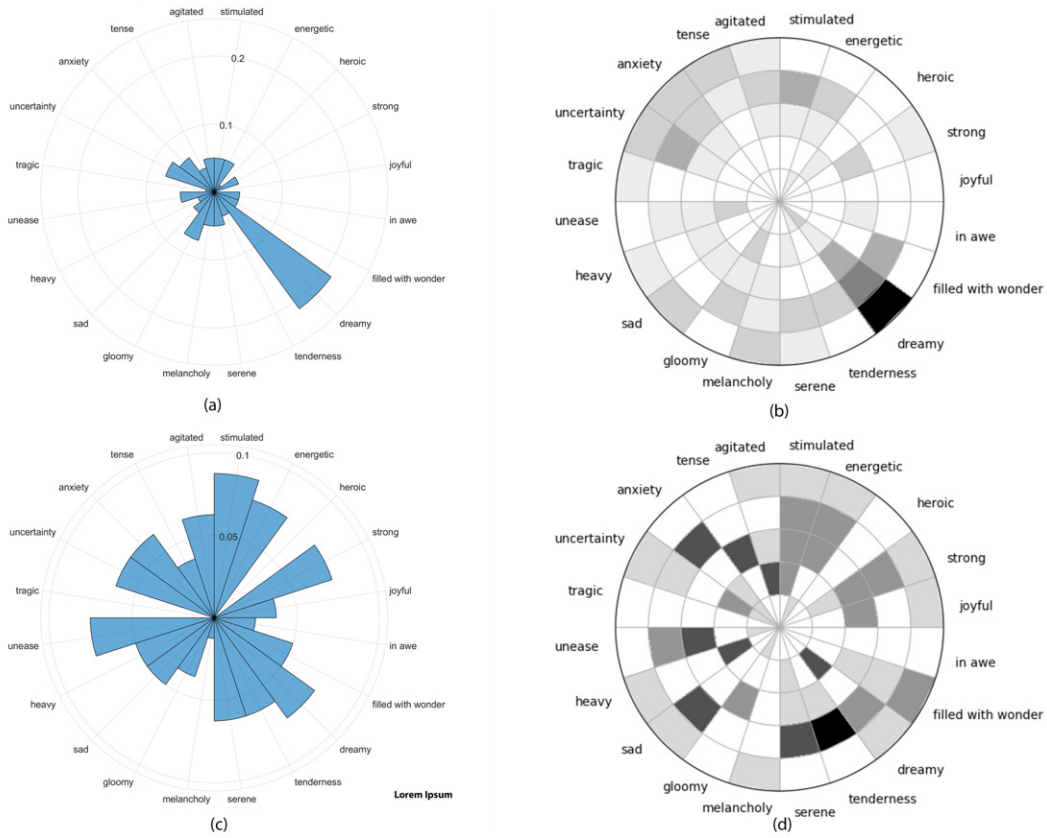


Figure 6. Visualizations of two perceived artworks. Artwork 1 (a, b) is observed to be quite distinct. Artwork 2 (c, d) is visibly less distinct as both the histogram and the heatmap show a balanced distribution.

The variance in degrees is found using the arccosine of the second term in Eq. (6):

$$V_e^o = \cos^{-1} \left(\frac{\sum (I_i \cos(\theta'_e - \theta_i))}{\sum (I_i)} \right). \quad (7)$$

To make the results more intuitive, this can be expressed as the *variance of emotions* by scaling to the total number of emotion slices on the wheel:

$$V_e^e = V_e^o \left(\frac{N}{360} \right), \quad (8)$$

where N is the total number of slices (20 for the standard GEW). The intensity of the emotion profile (I_e) and the weighted (V_e) and unweighted circular variances (V_e^o) all provide an indication of emotion distinctness. The lower the variances or the higher the intensity, the more distinct an emotion appears. Scaling variance by emotion slices provides an intuitive understanding of its value. A variance of 3 or less indicates that the emotion is distinctly defined on the wheel, insofar as that the closest wheel slice to the circular mean (θ_e) can be considered an accurate descriptor of the emotion. For example, an emotion with a variance of 2.5 and a circular mean of 0° can be said to be “joyful”—the label of the slice at 0° .

6.2 Visualizing Findings from the VAWE

The VAWE is an adaptation of the GEW; therefore, circular statistics can be used to interpret its data. To visualize reported emotions in a manner that would optimally account both qualitatively for the type of emotions and quantitatively for the intensity of emotions, the study adapted the statistical method developed by Coyne et al. [7] using heatmaps and circular histograms (or rose plots). Both methods enable better quantification of findings from the GEW format and therefore from the VAWE. The code was further adapted using MATLAB and Python. These methods were then superimposed onto the emotion wheel to represent the types of emotions reported by the participants as well as the intensities of these emotions for each artwork. Circular histograms help identify emotion families and quadrants most frequently associated with artworks, distinguishing distinct (Figure 6(a)) from less distinct emotions (Fig. 6(c)). Heatmaps offer an additional perspective on the entire emotional space and indicate the intensity levels. Here, a distinct emotion would be one where a few segments, closely spaced together, are selected most frequently (Fig. 6(b)). In contrast, an indistinct emotion would be one where there are multiple segments, widely distributed across the emotion wheel, commonly selected (Fig. 6(d)).

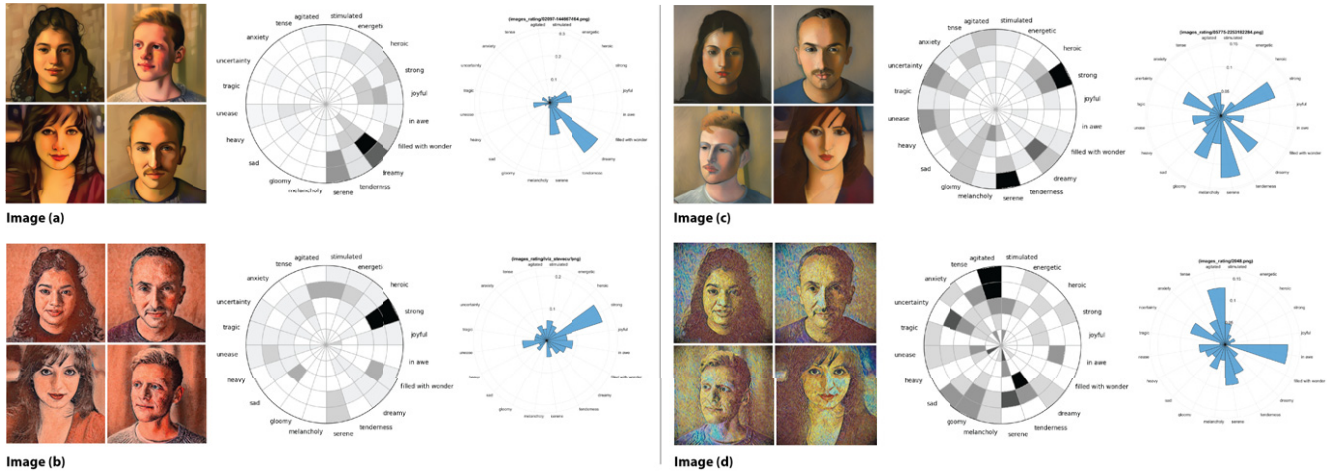


Figure 7. VAVE heatmaps (middle) and histograms (right) for sample images (a–d).

Table I. Summary of some AI portrait test results.

Portrait	θ_e (Emotion)	Closest label	I_e (Intensity)	V_e^e (Variance)
a (02097-1446674164)	17.9952	Dreamy	2.0322	3.1609
b (iviz_stevecu1)	2.3563	Strong	0.97151	4.067
c (05775-2253182284)	14.0786	Melancholy	0.50272	4.5727
d (2048)	9.5083	Tragic	0.46994	4.4768

7. RESULTS

The performance of each portrait artwork set representing a style is measured by how distinct it is, namely, how well it clusters to a local region on the VAVE tool. The observed variance provides essential insight into the clarity of the perceived emotion. High variance indicates that not all observers agreed on the perceived valence and arousal of the affective state, suggesting that it could be ambiguous. In contrast, low variance indicates that most observers perceived the artwork similarly. A sample of the findings from the AI portrait emotion study is presented graphically in Figure 7, and the associated statistics expressions are provided in Table I.

The results (from the 120 portrait sets) show a wide range of variances for each portrait style from 2.744 to 4.8718 for the weighted variance (V_e^e). There is an inverse correlation between variance and intensity in this dataset: portrait sets with high variance and low intensity exhibit indistinct emotions (e.g., c and d) while those with low variance represent distinct emotions (e.g., a and b). This is evident in the heatmaps in Fig. 7; (a) and (b) are visually tightly focused in certain regions, whereas (c) and (d) show chosen emotions more dispersed over VAVE. The highest variances are reported for (d). The closest label on the VAVE indicated by the circular mean θ_e is “Tragic,” which we consider unintuitive for the corresponding emotion shown in Fig. 7(d). The heatmap for d (Fig. 7(d)) shows that very few participants assign “Tragic” to this artwork. This observation

indicates that when variance levels are high, the circular mean (θ_e) is a poor indicator of trend and may not correctly identify the dominant label selected on the wheel. The variances of (a) were one of the lowest of the artworks tested. This indicates that according to the population sample, it was one of the most unambiguous artworks in the dataset. The artwork heatmap and histogram in Fig. 7(a), showing a single focused region, confirm that the emotion is highly defined. The parameter V_e^e is the weighted variance scaled to the 20 emotion slices. As the variance is low, emotion θ_e fits the data strongly, and the closest label on VAVE (“Dreamy”) can be considered an accurate, unambiguous descriptor of artwork (a) (Fig. 7(a)).

It is important to note that these statistics consider that angle defines emotion. This is typically accurate except when significant numbers of participants label the state as being “Neutral,” which has zero associated angle and intensity, and therefore has no impact on the calculation of the weighted circular variance (V_e^e) and only decreases the final vector intensity (I_e). This case was observed in the data for state (c); similar to (d), the variance is high and the intensity is low compared to other artworks, and although it would be intuitive to dismiss it as an indistinct emotion, its heatmap (Fig. 7(d)) shows the strong prevalence of “Neutral” in the data unlike (c) (Fig. 7(c)). This shows how vital graphical tools are for visually understanding the data and overcoming the limitations of statistics. Each emotion rating was followed by a 5-point Likert question regarding the participants’ aesthetic likability of the portrait as fine art. The findings reveal that the average likability of all artworks falls within the range 1.4–4.1. Likability demonstrates a positive correlation with positive emotions and a negative correlation with negative emotions as depicted in Figure 8.

Another research finding that contributes to the validation of the VAVE tool is that the results showed an even distribution of the emotion labeled by participants to all portrait artworks presented, indicating that the VAVE emotion term selection was comprehensive (Figure 9).

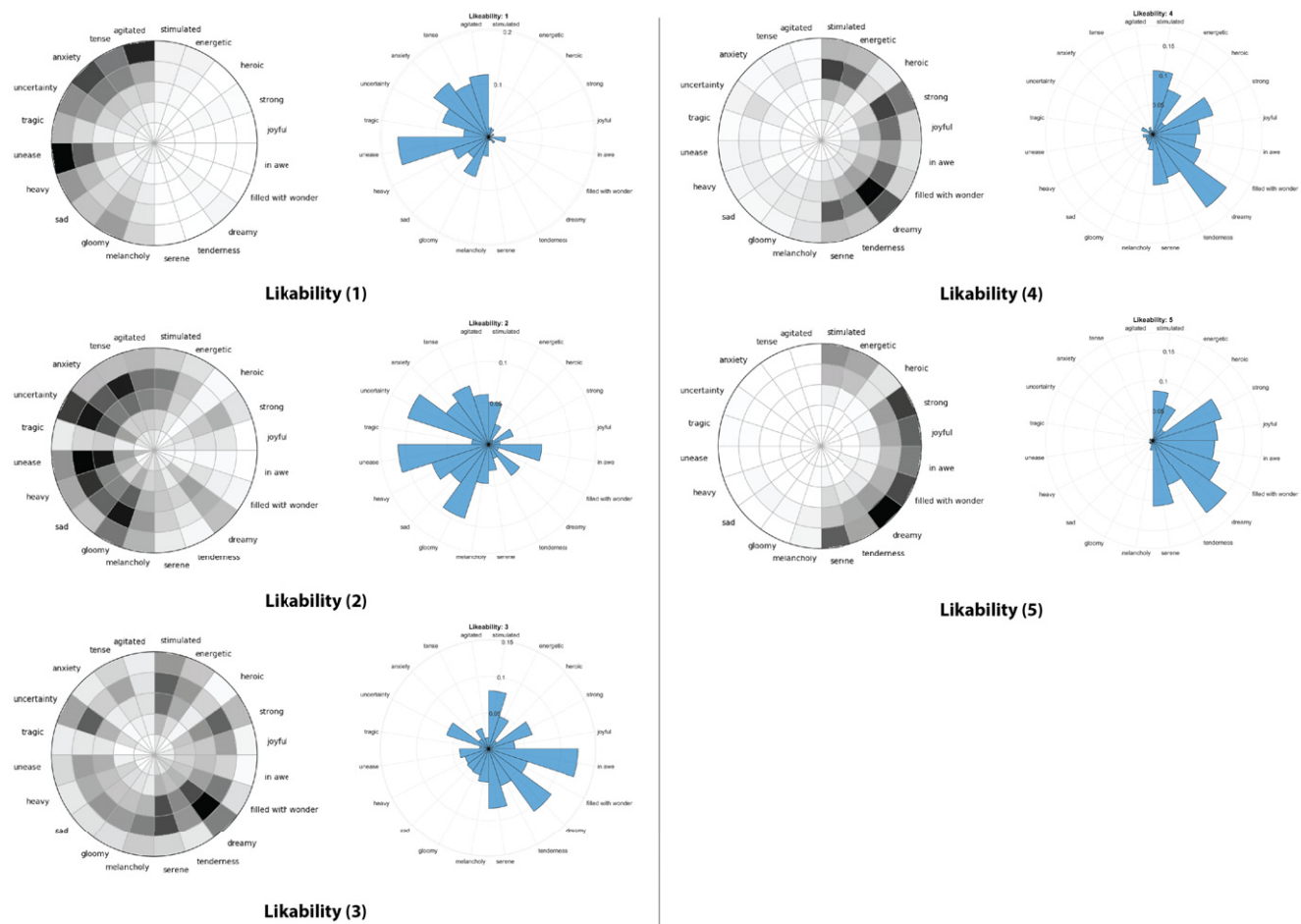


Figure 8. Likability heatmaps and histograms.

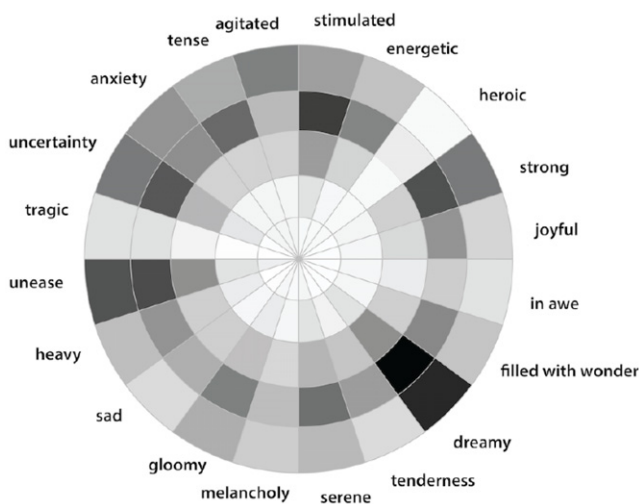


Figure 9. The heatmap for 120 portrait artwork styles.

8. TOWARDS AN EMOTIONALLY AWARE PORTRAIT SYSTEM

The large number of participants in the study appears to validate VAWE as a tool for identifying aesthetic emotions

related to art. That said, in this study, there is a strong measure to categorize our multiyear AI portrait system via matching specific emotions to given style modules. This can be seen in the results as there was an even distribution of the emotion labeled by participants to all portrait artworks presented, indicating that the VAWE emotion term selection was comprehensive. This finding contributes to the validation of the VAWE tool as an aesthetic emotion measurement tool (Fig. 9), matching our first validation of VAWE on historical portrait work. This shows that both the Generative AI Portrait System and its correlated styles with VAWE can be used by future researchers in many fields such as the arts and from art therapy to health in general. With this in mind and to further validate both VAWE and our portrait system's styles, a limited qualitative study was carried out to evaluate a prototype system for self-portraits. It should be noted that the VAWE system has already been validated with a large user population but in a non-self-portrait setup.

8.1 System Prototype/Description

A notable challenge that was missing in the prior work was the lack of a robust and valid metric for translating emotions into distinct portrait styles. In response to this, the current study incorporates the VAWE tool as a metric system,



Figure 10. Photo gallery—set of 20 personalized portrait styles of the sitter.

specifically, a transference system, utilized in mapping portrait styles to corresponding emotions. In addition, art systems were updated to build better portraits throughout, including using Stable Diffusion. We enhanced the art system, continually refining our portrait styles by integrating advanced techniques such as Stable Diffusion, aiming for a more sophisticated and improved visual experience. We believe that VAWES's emotional mapping, with our computational portrait system styles, can be used in many areas by other researchers in entertainment, art, and health, and we see this as our main contribution. However, we wanted to present an example of usage to further validate the application of VAWES. A self-portrait study was conducted using a prototype application for an “Emotionally Aware Portrait System.” It is essential to emphasize that the innovation in this study primarily lies in the development of the internal mapping system rather than a comprehensive overhaul of the entire system.

8.2 Study

A limited qualitative study involving six participants (P1–P6) was conducted with the following steps. First, a photo portrait of each participant was taken under a controlled light condition prior to the study. Second, each participant was presented with a set of 20 personalized portrait styles (Figure 10). These images were generated using our Generative AI Portraiture System, which optimally mapped the VAWES's 20 emotions based on the results of the

preceding study. The methodological approach included both quantitative and qualitative questions regarding the participants' generated portraits, covering assessments of likability as fine art, considered use, and evaluations of the system's efficacy in mapping styles to corresponding emotions. In the first part of the study, participants were presented with a photo gallery showcasing various portrait styles (Fig. 10). Subsequently, they were prompted to select their most preferred and least preferred styles, providing reasons for their choices. Participants were then asked about their likability of the portrait arts as fine art and how they would use it. The findings showed that participants consistently gave high ratings to one or more portraits as fine art, yielding a notable 86% consensus. Feedback from participants included the following:

“Some of them you can definitely consider fine art.” (P2)

“I think some of these are interesting stylistically and could be displayed in galleries (edited).” (P1)

This is quite substantial given the complexities surrounding self-portraiture and individuals' feelings observing their own faces. This agreement is significant as it shows that participants view their self-portraits as authentic works of art. This is particularly important at a time when AI is widely used to produce art of varying qualities. When asked to indicate their preferences among the 20 portraits, some comments were as follows:



Figure 11. Examples of participants' favored styles.

"Wow, that one's really nice . . . there's a lot of nice ones. Wow, they're also unique. It's very interesting." (P3)

When asked why, some comments were as follows:

"It brings out like some characteristic in my features." (P4)

"I like the more oil pastel style of it." (P5)

"The way the eyebrows accentuate and the nice contrast between some of the colors, I really like how this one turned out." (P3)

Notably, four out of the six participants favored a specific style named *iviz_port12* (Fig. 10(a)) while two out of six expressed preferences for styles named 21546-78727403 and 2017, respectively (Figure 11(b) & (c)) as their favorite or second favorite. The second part of the study included presenting the VAWE tool to the participants and explaining the dimensional model emotions and quadrants. The Self-Portrait System prototype was presented, featuring the mapped styles to the emotions. The Self-Portrait System interface is shown in Figure 12 with the sitter's original photo at the center and the 20 emotion buttons around it following the layout of the VAWE design. Clicking on one of the emotion buttons displays the stylized portrait that evokes the specific emotion. We reviewed each emotion and its corresponding style, asking if they thought the emotions matched. We obtained a high rating of 5.6 out of 7. In general, most participants were surprised by the style range of the images and how they matched the emotions. Many also found that there is a potential application to using the images on social media and marketing platforms. Some notable feedback was as follows:

"I'm surprised by like the range it's able to display and portray; I think probably down the line it gets more accurate, but for where the technology is right now, I think it's already like pretty accurate of the emotions it's trying to convey." (P2)

"A surprising amount of that matched up really well." (P1)

When we asked the participants how they would utilize the system, some responses were as follows:

"For a website in terms of like a designer, if I was trying to make a website that has like depicts emotion in this way, and to make these images cohesive hundred percent would use it. Yeah, it makes a lot of sense." (P1)

"I could see a portrait system like this (a) feeling more authentic to a person, and then (b) the utilization being higher because I have chosen my gallery if you will and so I can emote and what I am conveying emotionally is feels authentic and it is like quite literally a true reflection of myself." (P1)

"I mean, what I've seen it do and convey, it's like, very good at what it does to like a powerful degree." (P2)

In summary, this study served as an initial validation effort, ensuring the efficacy of the system before broader implementation. Despite potential challenges related to self-portraiture and individual variations in sentiments towards one's own facial features, our results show a mostly positive response. Participants showed a strong tendency to view the generated artwork as fine art, highlighting the potential usefulness and acceptance of this innovative system.

9. DISCUSSION

This research highlights the advantages of using VAWE for capturing aesthetic experiences. The wheel format enhances speed, accuracy, and intuitiveness of emotional reporting by clustering related emotions in close proximity. However, one potential drawback is that the labels are aligned along the curve of the circle, which can pose minor readability challenges compared to linear scales. Despite this, the overall design of the wheel supports a more faithful capture of participants' aesthetic experiences compared to traditional scales. Future work could explore adjustments to label alignment or alternative visualization techniques to further improve readability while maintaining the benefits of the wheel structure.

Analyzing VAWE data presented unique challenges due to its circular structure, where emotions are arranged based on valence and arousal. Some may question why we chose circular statistics over simpler linear methods. However, linear statistics are inappropriate for VAWE data because they fail to account for its cyclical nature; the wheel's endpoints connect



Figure 12. Self-Portrait System prototype—the UX matches the VAWE layout, where each button when clicked fades up from the photo to the fine art painted portrait of that emotional style.

as exemplified by the transition from “agitated” to “joyful.” Calculating a simple average of angles, for instance, would yield illogical results when averaging emotions near opposite ends of the wheel. Circular statistics, which involve converting angles to vectors, averaging the vectors, and then converting back to an angle, are essential to preserve the data’s underlying structure, treating emotions as related points on a circle. Furthermore, the circular visualization format reflects the semantic similarity and affective relationships between emotions on the VAWE. Emotions that are close together are perceived as more similar. Visualizing data in the same circular format allows us to leverage the wheel’s inherent organization and immediately understand the distribution of emotions. We avoided a frequency-based approach because it could misrepresent intensity variations. For example, “happy-3” receiving ten responses while “sad-1” through “sad-4” each receive nine would be misleadingly prioritized. Instead, averaging using circular statistics better reflects the overall emotional landscape. Therefore, we employed circular statistics to analyze the data and preserve its integrity.

Even though this research provides valuable insights into the emotional responses evoked by AI-generated portraits, its reliance on descriptive analysis limits the strength of the conclusions regarding the precise relationships between portrait styles and specific emotions (how many emotions were relatively well matched). While the current study primarily employs descriptive methods, such as heatmaps and circular histograms, to illustrate the emotional variability evoked by different portrait styles, we acknowledge that cluster analysis or multidimensional scaling could provide a more comprehensive quantitative perspective. Due to the exploratory nature of this initial study and the focus on establishing a foundational understanding of the system’s capabilities, we prioritized a broad descriptive analysis to identify key trends and patterns. As can be seen in Fig. 9, which presents the heatmaps of emotional responses across all portrait styles, our system yielded a relatively even distribution of emotions, suggesting its ability to evoke a diverse range of aesthetic experiences. We agree that future work will benefit from incorporating these more rigorous statistical

approaches to strengthen the validation of the relationship between portrait styles and the emotions they evoke.

The use of VAWE, with its focused set of 20 emotion terms, provided a practical framework for categorizing and analyzing emotional responses to AI-generated portraits. While we do not claim that VAWE comprehensively captures all possible emotions, our findings highlight its practical utility in effectively categorizing aesthetic emotions. The relatively even distribution of selected terms across portrait styles suggests that VAWE’s terms are both cognitively accessible and capable of capturing a broad emotional range.

We acknowledge that emotions not included in VAWE were not measured, and therefore, our study does not determine whether all emotions evoked by artworks were comprehensively captured. Rather, our approach builds on the GEW and aligns with the pleasure-arousal circumplex model (Russell), which theoretically accounts for all emotions within its dimensional framework. Our goal was to assess how well VAWE supports the categorization of evoked aesthetic emotions rather than to assert exhaustive emotional coverage.

Balancing a manageable number of terms without causing cognitive overload while maintaining sufficient resolution remains an ongoing challenge in aesthetic emotion research. Future work will include additional studies to refine VAWE’s affect terms, ensuring broader applicability and greater precision in capturing aesthetic emotions. Additionally, there are a number of limitations in this study. A key consideration that should be investigated further is the identification of alternative quantitative metrics that could further support our findings. The examples in portrait sets (c) and (d) in Fig. 7 were included to illustrate that portrait sets with high variance and low intensity tend to exhibit indistinct emotions. This aligns with our findings, which highlight that the highest variances were observed for set (d). The circular mean θ_e identified “Tragic” as the closest label on the VAWE (see Fig. 7(d)), which we agree may feel unintuitive for the corresponding emotion. This observation emphasizes the challenges of interpreting results in cases with high variance and low intensity (Comment 1.4.5). Future research could explore alternative visualizations of the VAWE data. For example, representing the results on the wheel as a

circle with a radius proportional to variance, rather than a single point, could provide a more nuanced understanding of the agreement or disagreement among participants.

Furthermore, facial expressions and gaze direction may contribute to the perceived emotional content of the portraits. We acknowledge that the generative AI system (Stable Diffusion) may alter portrait characteristics to varying degrees during the stylization process and could impact perceived emotions. Although participants were instructed to focus on style elements rather than content, we recognize that such alterations might still influence their responses. We plan to address these issues in greater depth and propose methodological refinements to control these variables in future studies.

A fundamental challenge in measuring aesthetic emotions is ensuring clarity in how participants interpret the rating task. Even though our study aimed to assess perceived aesthetic emotions evoked by visual art, the wording of “feel” in the instructions may have introduced some ambiguity. Specifically, although participants were instructed to rate the emotions they felt in response to the style of the portraits, some may have instead reported their own emotional responses rather than the emotions they believed the artwork conveyed.

This distinction is a well-documented issue in aesthetic emotion research, as individuals can recognize and assess emotions in art without necessarily experiencing those emotions themselves. To mitigate this, our study design incorporated VAWE’s structured emotion wheel and explicit instructions emphasizing style over content. However, we acknowledge that further refinements could improve clarity. Future research could enhance instruction wording and introduce follow-up questions to better differentiate between personally experienced emotions and perceived aesthetic emotions, ensuring a more precise understanding of participants’ responses.

10. CONCLUSION AND FUTURE WORK

The main contribution of the study involved labeling 120 generated styles utilizing VAWE to measure the aesthetic emotions they elicit. The study appears to validate VAWE as a tool for measuring aesthetic art emotions and for categorizing our multiyear AI portrait system by matching emotion to given style modules. The results showed an even distribution of emotions labeled by participants to all portrait artworks presented, which contributes to the validation of the VAWE tool as an aesthetic emotion measurement tool. The analysis revealed a trend where portrait artworks evoking positive emotions received higher likability ratings compared to those evoking negative emotions. This matches our first validation of VAWE on historical portrait work.

For the research’s second contribution to further validate both VAWE and our portrait system styles, we prototyped a self-portrait system and carried out a small qualitative study. The results from the first study were utilized to build the prototype for an Emotionally Aware Portrait System capable of creating individualized custom portraits with

styles depicting the user’s affective states. A qualitative study was conducted using individual self-portraits, yielding good results and high likability ratings as fine art pieces according to the participants. The portraits generated seemed to depict the 20 emotion categories quite well. We believe that these styles could be useful in various research fields such as health, education, and entertainment.

It should also be noted that while many artists and researchers simply use systems such as Stable Diffusion to both create fine art portraits and use emotional keywords in the prompts, our proposed system is considerably different from Stable Diffusion or large language models. The quality of the generated art and the association with emotions cannot be validated in such systems. This research gap is where our system steps in. In our multidimensional research, we first use a series of AI and non-AI techniques based on the cognition model of art making, which was curated by artists. We also developed and created VAWE, a measurement tool to measure aesthetic emotions capable of robust statistical analysis.

Future work will involve further development of the Emotionally Aware Portrait System created by the authors by automating the aspects of the process that currently have human intervention, incorporating emotion styles into a text conversational system able to extract keywords to match the 20 emotion categories to generate artwork. The next phase of the project will include ongoing research utilizing the annotations from portrait artwork emotion studies to build a classifier capable of predicting aesthetic emotions evoked by artwork, which could automatically categorize new art generated from the AI Portraiture System.

REFERENCES

- ¹ N. Abukhodair, M. Song, S. Pekçetin, and S. DiPaola, “Designing a wheel-based assessment tool to measure visual aesthetic emotions,” *Cogn. Syst. Res.* **84**, 101196 (2024).
- ² P. Achlioptas, M. Ovsjanikov, K. Haydarov, M. Elhoseiny, and L. J. Guibas, “ArtEmis: Affective language for visual art,” *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2021), pp. 11569–11579.
- ³ X. Alameda-Pineda, E. Ricci, Y. Yan, and N. Sebe, “Recognizing emotions from abstract paintings using non-linear matrix completion,” *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* (IEEE, Piscataway, NJ, 2016), pp. 5240–5248.
- ⁴ Y. Balaji, S. Nah, X. Huang, A. Vahdat, J. Song, K. Kreis, M. Aittala, T. Aila, S. Laine, B. Catanzaro, and T. Karras, “eDiff-I: Text-to-image diffusion models with an ensemble of expert denoisers,” Preprint arXiv:2211.01324 (2022).
- ⁵ C. Becker-Asano and H. Ishiguro, “Laughter in social robotics—no laughing matter,” *Proc. Int’l. Workshop on Social Intelligence Design* (Springer, 2009), pp. 287–300.
- ⁶ E. Cetinic, T. Lipic, and S. Grgic, “A deep learning perspective on beauty, sentiment, and remembrance of art,” *IEEE Access* **7**, 73694–73710 (2019).
- ⁷ A. Coyne, A. Murtagh, and C. McGinn, “Using the Geneva emotion wheel to measure perceived affect in human-robot interaction,” *Proc. 2020 ACM/IEEE Int’l. Conf. on Human-Robot Interaction* (ACM/IEEE, Cambridge, UK, 2020), pp. 491–498.
- ⁸ O. da Pos and P. Green-Armytage, “Facial expressions, colours and basic emotions,” *Colour: Design and Creativity* **1**, 2 (2007).
- ⁹ S. DiPaola, “Exploring a parameterized portrait painting space,” *Int. J. Art Technol.* **2**, 82–93 (2009).
- ¹⁰ S. DiPaola, “Exploring the cognitive correlates of artistic practice using a parameterized non-photorealistic toolkit,” *Leonardo* **50**, 452–453 (2017).

- 11 S. DiPaola and L. Gabora, "Incorporating characteristics of human creativity into an evolutionary art algorithm," *Proc. 9th Annual Conf. Companion on Genetic and Evolutionary Computation* (ACM, New York, NY, 2007), pp. 2450–2456.
- 12 S. DiPaola, L. Gabora, and G. McCaig, "Informing artificial intelligence generative techniques using cognitive theories of human creativity," *Procedia Comput. Sci.* **145**, 158–168 (2018).
- 13 S. DiPaola and G. McCaig, "Using artificial intelligence techniques to emulate the creativity of a portrait painter," *Electronic Visualisation and the Arts (EVA 2016)* (BCS, Swindon, UK, 2016), pp. 158–165.
- 14 S. DiPaola, G. McCaig, K. Carson, S. Salevati, and N. Sorenson, "Adaptation of an autonomous creative evolutionary system for real-world design application based on creative cognition," *Proc. Fourth Int'l. Conf. on Computational Creativity (ICCC)* (University of Sydney, 2013), pp. 40–47.
- 15 D. J. Duke, P. J. Barnard, N. Halper, and M. Mellin, "Rendering and affect," *Comput. Graph. Forum* **22**, 359–368 (2003).
- 16 P. Gendall, J. Hoek, and K. Gendall, "Evaluating the emotional impact of warning images on young adult smokers and susceptible non-smokers," *J. Health Commun.* **23**, 291–298 (2018).
- 17 K. Hevner, "Experimental studies of the affective value of colors and lines," *J. Appl. Psychol.* **19**, 385–398 (1935).
- 18 D. Joshi, R. Datta, E. Fedorovskaya, Q. T. Luong, J. Z. Wang, J. Li, and J. Luo, "Aesthetics and emotions in images," *IEEE Sig. Process. Mag.* **28**, 94–115 (2011).
- 19 O. Korovina, F. Casati, R. Nielek, M. Baez, and O. Berestneva, "Investigating crowdsourcing as a method to collect emotion labels for images," *Extended Abstracts of the 2018 CHI Conf. on Human Factors in Computing Systems* (ACM, New York, NY, 2018), pp. 1–6.
- 20 H. Leder, A. Ring, and S. Dressler, "See me, feel me! Aesthetic evaluations of art portraits," *Psychol. Aesthet. Creat. Arts* **7**, 358–369 (2013).
- 21 K. V. Mardia, "Statistics of directional data," *J. R. Statist. Soc. Ser. B: Statist. Method.* **37**, 349–371 (1975).
- 22 G. McCaig, S. DiPaola, and L. Gabora, "Deep convolutional networks as models of generalization and blending within visual creativity," Preprint arXiv:1610.02478 (2016).
- 23 S. Mohammad and S. Kiritchenko, "WikiArt Emotions: An annotated dataset of emotions evoked by art," *Proc. Eleventh Int'l. Conf. on Language Resources and Evaluation (LREC 2018)* (European Language Resources Association (ELRA), 2018).
- 24 J. W. Peirce, "Generating stimuli for neuroscience using PsychoPy," *Front. Neuroinform.* **2**, 343 (2009).
- 25 Prolific website, <https://prolific.co>, accessed February 2024.
- 26 P. Samuelson, "Generative AI meets copyright," *Science* **381**, 158–161 (2023).
- 27 K. R. Scherer, "What are emotions? And how can they be measured?," *Soc. Sci. Inform.* **44**, 695–729 (2005).
- 28 K. Schulz and G. U. Hayn-Leichsenring, "Face attractiveness versus artistic beauty in art portraits: A behavioral study," *Front. Psychol.* **8**, 1–9 (2017).
- 29 S. Seo and D. Kang, "Study on predicting sentiment from images using categorical and sentimental keyword-based image retrieval," *J. Supercomput.* **72**, 3478–3488 (2016).
- 30 N. Tschöpe, J. E. Reiser, and M. Oehl, "Exploring the uncanny valley effect in social robotics," *Proc. Companion of the 2017 ACM/IEEE Int'l. Conf. on Human-Robot Interaction* (ACM, New York, NY, 2017), pp. 307–308.
- 31 P. Valdez and A. Mehrabian, "Effects of color on emotions," *J. Exp. Psychol.* **123**, 394–409 (1994).
- 32 R. Xu and Y. Hsu, "Discussion on the aesthetic experience of artificial intelligence creation and human art creation," *Proc. 8th Int'l. Conf. on Kansei Engineering and Emotion Research: KEER 2020* (Springer Singapore, 2020), pp. 340–348.
- 33 V. Yanulevskaya, J. Uijlings, E. Bruni, A. Sartori, E. Zamboni, F. Bacci, D. Melcher, and N. Sebe, "In the eye of the beholder: Employing statistical analysis and eye tracking for analyzing abstract paintings," *Proc. 20th ACM Int'l. Conf. on Multimedia* (ACM, New York, NY, 2012), pp. 349–358.