Affective and Pleasurable Design, Vol. 123, 2024, 29–36
https://doi.org/10.54941/ahfe1004673

Generating Paintings Eliciting Specific Emotions Using Machine Learning for Application in Painting Therapy

Keisuke Kisu[1], Motohiro Kozawa[2], and Keiichi Watanuki[1,3]

[1]Graduate School of Science and Engineering, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338–8570, Japan
[2]Faculty of Education Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338–8570, Japan
[3]Advanced Institute of Innovative Technology, Saitama University, 255 Shimo-okubo, Sakura-ku, Saitama-shi, Saitama 338–8570, Japan

ABSTRACT

This study aims to label paintings based on biometric information and generate paintings that elicit specific emotions using machine learning. To create the dataset, we conduct experiments with eight participants using multi physiological measurement sensors. We focus on the arousal axis of emotion and use the skin conductance response as a measure of arousal. The results suggest that machine learning may be effective in generating paintings that elicit emotions because features related to arousal, such as brightness and color, can be appropriately learned.

Keywords: Painting, Emotion, Machine learning, GAN, Skin conductance response

INTRODUCTION

In recent years, stressful societies have been regarded as a problem, and painting therapy, which is effective in relieving stress and treating psychological disorders, is expected to become widespread. Current painting therapy research and practices focus on active drawing activities. However, considering that there are some people who resist drawing, it is important to examine the viewing of paintings, which can be done passively, and to distinguish between viewing and drawing depending on the person and situation. It is important to clarify the relationship between appreciation and emotions in the appreciation of paintings; the evaluation of the relationship between appreciation and psychological stressors based on emotions such as anger and anxiety, which are the causes of stress, may not only help solve the underlying causes of stress but also motivate people to paint by evoking positive emotions through the appreciation of paintings.

Appreciation and Emotion

In previous studies that evaluated human emotions when viewing paintings, studies focusing on each element of the paintings were conducted, such as one that evaluated the relationship between the color of the paintings and emotions (Marshall et al., 2019) and another that evaluated the relationship between factors such as symmetry, color combination, saturation, and brightness of images and emotions (Taemin et al., 2020). However, considering that the components of a painting affect each other in a complex manner, it is necessary to extract the features that affect the emotions in each painting. Machine learning is a method for extracting the features of paintings, and a report by David et al., (2017) classified paintings into 10 emotions based on subjective evaluation and generated paintings based on these emotions. However, previous studies have evaluated emotions based on subjective evaluations, and there are no examples of emotion evaluations using biometric information. There is a possibility that emotions can be evaluated appropriately and quantitatively using biometric information instead of subjective evaluation, which is qualitative and easily influenced by the impression of a painting. Therefore, the purpose of this study is to evaluate the effect that viewing a painting has on emotions using biometric information and to generate a painting that induces specific emotions using machine learning. Previous studies have reported that human emotions can be evaluated along two axes: pleasant-unpleasant and arousal-non-arousal (Russell, 1980). Of these two axes, this study evaluated emotions using biometric information on arousal and non-arousal and generated paintings that evoke arousal and non-arousal emotions.

PAINTING APPRECIATION EXPERIMENT

We conducted a painting viewing experiment using biometric measurements to obtain biometric information while viewing paintings. This experiment was approved by the Ethics Committee on Research Involving Human Subjects at Saitama University (R4-E-22). Informed consent was obtained from all the collaborators.

Experimental Method

Figure 1 shows the environment used in the experiment. The skin conductance response (SCR) was measured using a Nexus-10 MARK II biometric measurement (manufactured by Mind Media). The experimenter sat 0.75 m away from the monitor and the SCR sensors were attached to the index and ring fingers of the non-dominant hand. The flow of each trial is shown in Figure 2. The participants rested for 1 min, viewed the paintings for 7.5 s, answered a questionnaire about their emotions, and rested for 5 s. The participants viewed 30 paintings by repeating the process of viewing the paintings and resting for 5 s 30 times. This was considered one trial and was repeated eight times at intervals. Ultimately, the participants observed a total of 240 paintings during the experiment. For the rest of the task, participants were asked to look at black crosses to measure the baseline biometric information. In the painting viewing task, the paintings were displayed in random order. In the emotion questionnaire, the participants answered a questionnaire based on Russell's circle model (Russell, 1980), in which the horizontal axis was pleasant-unpleasant and the vertical axis was arousing-non-arousing. The experiment was conducted on eight male participants (23.9±3.1 years old).

Two groups of four were formed, with each group viewing 240 identical paintings for a total of 480 paintings. Paintings designated as public domains were selected for the experiments.

DATA PROCESSING FOR BUILDING PAINTING GENERATION MODEL

To create a dataset for the painting generation model, biometric information was used to classify paintings into arousal and non-arousal paintings. Because autonomic nervous system activity is related to the arousal state, the SCR was used as an evaluation index: an increase in the SCR value indicates sympathetic dominance and arousal, whereas a decrease indicates parasympathetic dominance and non-arousal. For the SCR, the Z-score was calculated by standardization, as shown in Equation (1).

$$X(t)_{z\text{-}score} = \frac{X(t)_{Raw} - \mu_{All}}{\sigma_{All}} \tag{1}$$

where $X(t)_{Raw}$ is the data before standardization, μ_{All} is the mean of the entire data, and σ_{All} is the standard deviation of the entire data. After standardization, the mean SCR was calculated for the four datasets within a group that viewed the same painting. The paintings were classified into an arousal group if the average SCR was positive, and into a non-arousal group if the average SCR was negative. Consequently, a dataset was created with 216 paintings in the arousal group and 264 paintings in the non-arousal group.

In the experiment, paintings of different sizes were displayed and viewed on a 1920 × 1080 monitor. Padding was performed to align the images with the actual size as viewed by the participants. In the padding, the margins were filled with white images, and all 480 paintings were converted to 1920 × 1080 pixels. An example of padding is shown in Figure 3. In Figure 3, the background is grey to make the addition of the white images easier to understand.

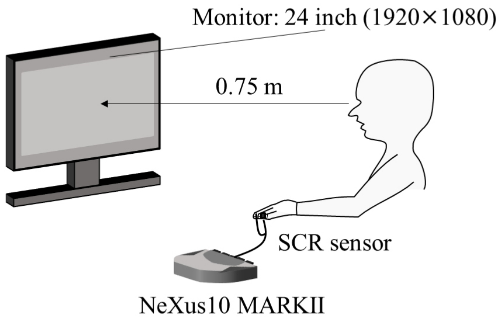


Figure 1: Experimental environment.

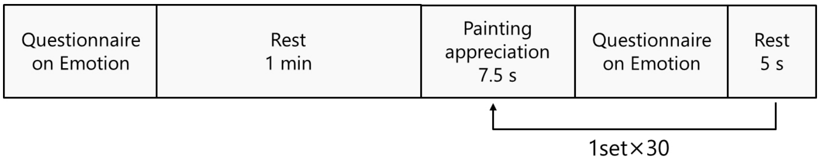
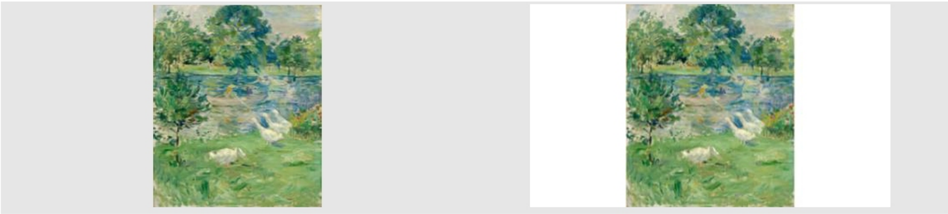


Figure 2: Experiment flow.



A. Original image B. After padding

Figure 3: Comparison before and after padding.

CONSTRUCTION OF PAINTING GENERATION MODEL

We constructed a painting generation model using a GAN (Ian et al., 2014) as shown in Figure 4. The model consists of a generator and discriminator. The generator generates fake samples from the input noise, the discriminator discriminates between the fake samples generated by the generator and the training data, and outputs the classification error. The generator consisted of an all-coupling layer and an output layer. The activation function for the all-coupling layer was a leaky rectified linear unit (leaky ReLU) to avoid gradient loss, whereas the output layer activation function was a tanh function. Using the tanh function instead of the sigmoid function, which transforms the input value from 0 to 1 and outputs it, the input value can be transformed from -1 to 1 and a clearer image can be produced. The discriminator consisted of an all-coupling layer and an output layer, with a leaky ReLU as the activation function for the all-coupling layer and a sigmoid function for the output layer. The leaky ReLU, tanh function, and sigmoid function are expressed in Equations (2), (3), and (4), respectively. For training, we set the leaky ReLU to 0.01, the loss function to binary cross-entropy, the optimization algorithm to Adam, the learning rate to 0.001, the batch size to 128, and the number of epochs to 20000.

$$Leaky\ ReLU = \begin{cases} x & (x \geq 0) \\ ax & (x < 0) \end{cases} \tag{2}$$

$$tanh = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{3}$$

$$sigmoid = \frac{1}{1 + e^{-x}} \tag{4}$$

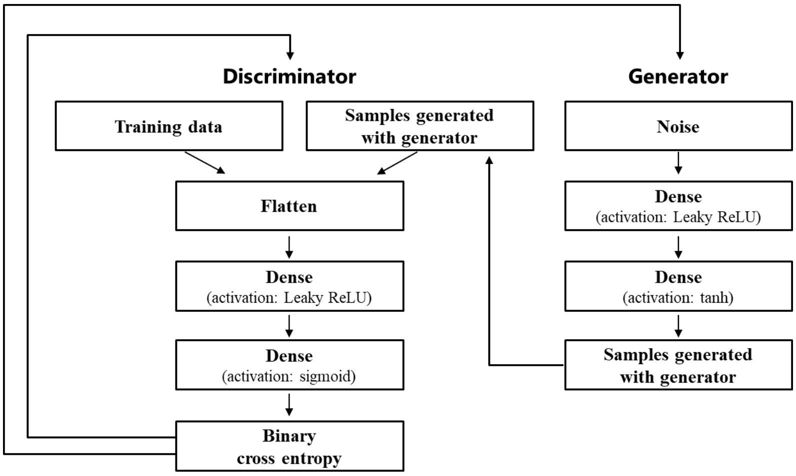


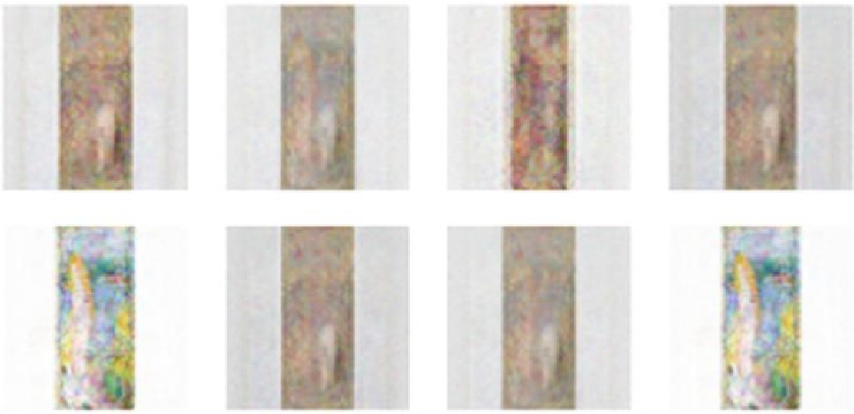
Figure 4: Model overview.

GENERATED RESULTS

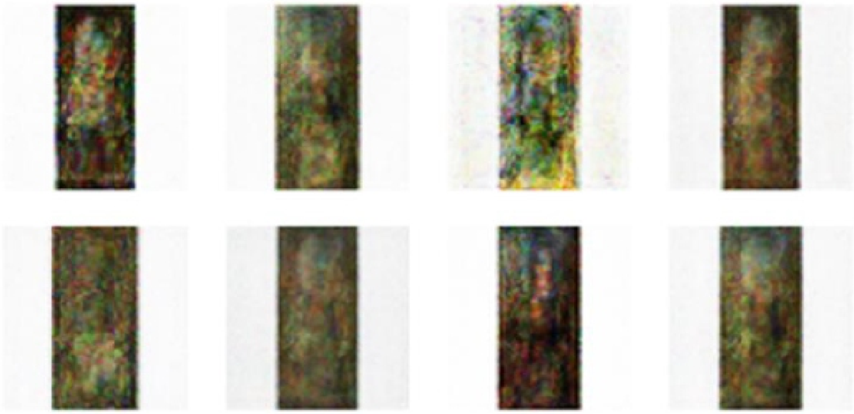
We generated paintings using the developed GAN models. The same model used 216 paintings for the arousal group and 264 paintings for the non-arousal group as the training data. The size of the generated paintings was 64 × 64 pixels, and eight paintings were generated for the arousal group and non-arousal group, respectively. Figure 5 shows the painting generation results. Different groupings based on the arousal levels produced different trends. The arousal group produced brighter paintings, whereas the non-arousal group produced darker paintings (mainly green and black). To evaluate the differences in brightness between the two groups, we compared the brightness of the paintings produced by the arousal and non-arousal groups. The results are presented as box plots in Figure 6a. The Mann-Whitney u-test showed a significant difference at the 1% level of significance. The brightness of the training data of 216 images in the arousal group and 264 images in the non-arousal group were compared between the arousal group and non-arousal group. The results are shown as a box-and-whisker diagram in Figure 6b. The Mann-Whitney U-test was conducted at a significance level of 5%, and no significant differences were found.

Subjective Evaluation of Paintings

A subjective evaluation questionnaire was administered to 13 participants to evaluate eight paintings in each group. The questionnaire asked the participants to rate all 16 paintings on a 7-point scale, from "arousal/non-arousal," "light/dark," "pleasant/unpleasant," "good/bad," and "like/dislike," and the 7-point scale was converted from 3 to -3 and compared between the arousal group and non-arousal group. The Mann-Whitney u-test showed that there was a significant difference at the 1% level for "light-dark" and a significant difference at the 5% level for "pleasant-unpleasant." The results are shown in the box-and-whisker diagram in Figure 7.

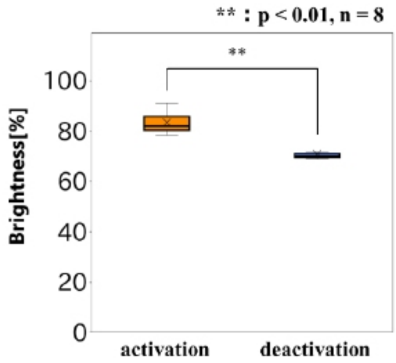


a. Generated results using the paintings of the activation group as training data

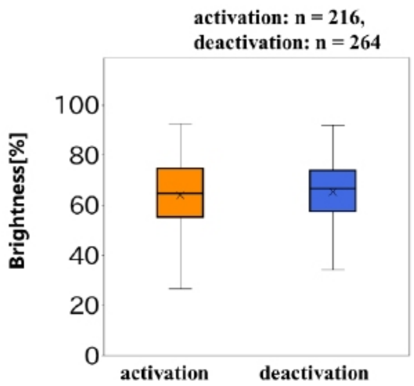


b. Generated results using the paintings of the deactivation group as training data

Figure 5: Generated results.

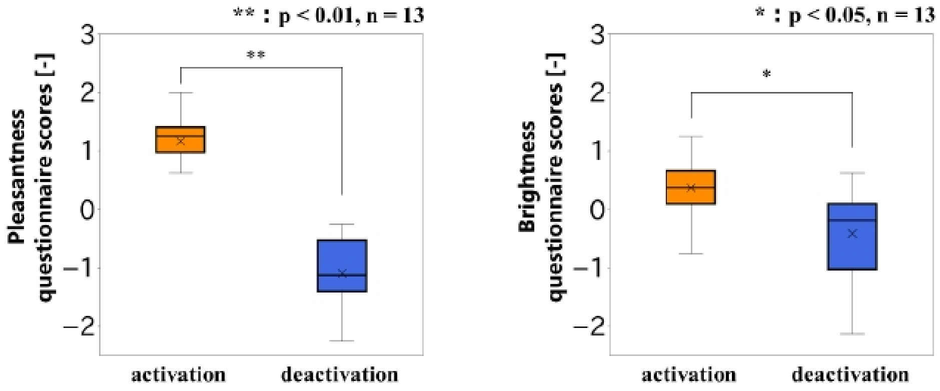


a. Comparison results of brightness of paintings generated by GAN



b. Comparison results of brightness of paintings in training data

Figure 6: Brightness comparison results.



a. Results of light-dark questionnaire

b. Results of pleasant-unpleasant questionnaire

Figure 7: Questionnaire comparison results.

DISCUSSION

Figure 6a and 7a show that when the brightness of the generated paintings was compared between the arousal and non-arousal groups, significant differences were observed in both the brightness and subjective brightness. In addition, compared to the arousal group, the non-arousal group produced paintings mainly in green and black. As the cause of these findings, previous studies have reported that colors with higher brightness have an arousing effect, with red having the highest arousing effect, followed by green and blue (Lisa et al., 2018). It is possible that the present generative model could effectively learn the features that affect arousal, resulting in the generation of paintings with a relatively high brightness in the arousal group and a relatively low brightness, mainly green, in the non-arousal group. The fact that no difference was observed in the brightness of the training data in Figure 7b suggests that there was no bias in the brightness of the training data, and that the GAN model could learn the features appropriately during the learning process.

Figure 7b shows a significant difference between the aroused and non-aroused groups in terms of pleasantness. In the questionnaire, some respondents gave reasons for their discomfort, such as "I don't understand it well and it makes me feel uncomfortable." Previous studies have shown that abstract paintings have lower scores on the evaluative factors, such as pleasantness and displeasure, than figurative paintings (Azami et al., 2021). Consequently, it is possible that the non-arousal group felt uncomfortable with the paintings produced in this study because they were more abstract than those of the arousal group, and therefore were not completed as paintings. Therefore, even if the characteristics of the non-arousal group could be learned appropriately, the evaluation of the paintings would change depending on whether the quality of the paintings was high enough to be evaluated as paintings, and it would be important to generate images that are more accurate and closer to the actual paintings.

CONCLUSION

In this study, we conducted viewing experiments, built models, and generated paintings with different arousal levels with the aim of generating paintings that elicit specific emotions using machine learning. Paintings in the training data were padded and classified into arousal and non-arousal groups by evaluating their arousal levels using standardized SCRs. The classified paintings were used to generate paintings using the GAN. The results suggested that machine learning may be effective in generating emotion-inducing paintings because it can appropriately learn features related to arousal, such as brightness and color. It was also suggested that the abstractness of generated images may influence emotions.

REFERENCES

Azami, Ritsuko and Totsuka, Mayo. (2021). The Association between Evaluations of Representational and Abstract Paintings and Attitude toward Ambiguity, Volume 31. pp. 39–48.

David, Alvarez-Melis and Judith, Amores. (2017). The Emotional GAN: Priming Adversarial Generation of Art with Emotion, NeurlIPS Machine Learning for Creativity and Design Workshop.

Ian, Goodfellow. Jean, Pouget-Abadie. Mehdi, Mirza. Bing, Xu. David, Warde-Farley. Sherjil, Ozair. Aaron, Couville and Yoshua, Bengio. (2014). Generative Adversarial Nets, Advances in Neural Information Processing Systems, Volume 27.

Lisa, Wilms and Daniel, Oberfeld. (2018). Color and emotion: effects, of hue, saturation, and brightness, Psychological Research Volume 82. pp. 896–914.

Ma Marshall, Xiaoyin. Charles, Noussair. and Luc, Renneboog. (2019). Colors, Emotions, and the Auction Value of Paintings, European Economic Review, Volume 142.

Russell, James. (1980). A Circumplex Model of Affect, Journal of Personality and Social Psychology Volume 39, No. 6, pp. 1161–1178.

Taemin, Lee. Dongwann, Kang. Kyunghyun, Yoon and Sanghyun, Seo. (2020). Emotion-Based Painting Image Display System, Intelligent Automation And Soft Computing, Volume 26, No. 1, pp. 181–192.