

MoodAlbum: Paintings derived from emotions

Shayan Amir



Fig. 1(a)

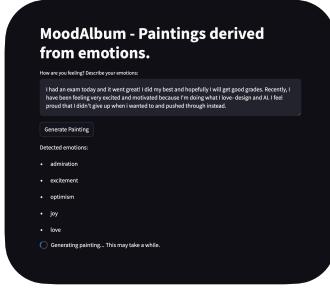


Fig. 1(b)



Fig. 1(c)

Figure 1 : (a) Initial user interface for input (b) Detected emotions after Input (c) Generated Art

Abstract

MoodAlbum is a large-language + text-to-image model based application that generates personalised art based on user emotions. By integrating emotion detection through roBERTa model and image generation using Stable Diffusion v1-5 model, MoodAlbum translates textual input reflecting the user's mood into unique visual art which is inspired by real art pieces from historic eras. This project aims to enhance human-ai interaction on an emotional level by the means of visual art therapy and self-reflection.

1 Introduction

In recent years, the field of generative AI has made significant strides in enhancing human-AI interaction, particularly in the realms of creativity and visual therapy. By leveraging advanced models, AI systems are now capable of producing art, music, and other forms of creative output that closely mimic human creativity. Research has shown that generative AI can be used effectively in therapeutic settings, helping individuals express their emotions and thoughts through visual means, thereby enhancing emotional well-being [1,2]. The potential for AI to collaborate with humans in creative processes offers exciting opportunities not only for artistic endeavours but also for therapeutic applications where visual expression is a key component of healing [3].

However, despite the advancements, there are still significant challenges in this domain. One of the primary obstacles is the accurate detection of human emotions through AI, which requires models to understand and interpret subtle emotional cues from text or speech [4]. Moreover, generating art that is genuinely reflective of these emotions and resonates on a personal level with users is a complex task. The AI must bridge the gap between understanding abstract emotional states and producing visual outputs that evoke the desired feelings in the user. The difficulty lies not only in the technical implementation but also in ensuring that the generated art remains contextually relevant and emotionally congruent with the input provided by the user [5].

MoodAlbum aims to address these challenges by integrating advanced emotion detection with state-of-the-art image generation technique. By using the large-language based model RoBERTa for precise emotion detection and text-to-image model stable diffusion for generating art, it ensures that each piece of art is not only visually compelling but also deeply connected to the user's emotional state. Unlike previous efforts that may have struggled with maintaining contextual relevance or emotional accuracy [6], MoodAlbum draws inspiration from real art pieces, incorporating styles from various artistic eras and movements. This approach not only fills the gaps left by earlier models but also enhances the therapeutic potential of AI-generated art, offering users a more personalised and meaningful creative experience.

2 MoodAlbum

2.1 Emotion Detection

Emotion detection is a crucial first step, accomplished by analysing the user's input through the use of the RoBERTa (Robustly Optimised BERT Pretraining Approach) model [7], a transformer-based language model fine-tuned for emotion classification [8]. The model is trained to recognise 28 distinct emotions; however, in our project, we focus on the top 5 emotions (based on scores given by the model) detected in the user's input to streamline the process. These emotions are then mapped to specific artists and artistic eras that historically align with those emotional states. For instance, sadness might be associated with the melancholic tones of Edvard Munch, while joy might link to the vibrant colours of the Impressionist era. This mapping allows the art generated to resonate more deeply with the user's current emotional state, creating a personalised and immersive experience.

2.2 Image Generation

The next step is generating an image that visually represents these feelings. MoodAlbum employs the Stable Diffusion v1-5 model, a state-of-the-art text-to-image generation model known for its high-quality output and flexibility. Stable Diffusion works by leveraging latent diffusion techniques [9], refining a noisy latent space input into a coherent image that matches the user-defined prompt. In our project, the prompt is carefully crafted by combining the top 5 emotions detected and referencing the corresponding artistic styles and eras. This prompt guides the diffusion model to generate an image that encapsulates the user's emotional state, providing a visual representation that resonates with their feelings. The resulting artwork is more abstract or artistic rather than containing realistic or lifelike representations, thus complementing the emotional tone.

2.3 Implementation

The system is implemented as a seamless pipeline that integrates emotion detection and image generation models, sourced from Hugging Face's Model Hub. The user begins by entering text such as a personal reflection or journal entry into the system Fig.1(a). RoBERTa processes this text to identify the top 5 emotions Fig.1(b), which are then mapped to corresponding artistic styles. This information is used to craft a prompt that combines the detected emotions, which is fed into the Stable Diffusion model to generate the final image Fig.1(c). The entire process, from input to image, is designed to be intuitive and simple using Streamlit, allowing users to engage with the system without needing to understand the underlying technology. The entire application is hosted on Hugging

Face Spaces which can be found [here](#), making it easily accessible. The result is a highly personalised piece of art that reflects the user's emotional state, offering a unique and meaningful way to explore and visualise their inner world.

3 Application Scenario

MoodAlbum offers versatile applications in art therapy, mental health, and personal creative expression. In therapeutic settings, by translating written reflections into emotionally resonant artwork, therapists can gain deeper insights into patients' mental states.

By adding a feature to store generated images, MoodAlbum can function as a creative journal, allowing users to document their daily experiences and revisit their emotional journey over a period of time. This form of self-reflection fosters emotional awareness and also provides an aesthetically rewarding way to capture personal moments.

Furthermore, MoodAlbum adds value by bridging the gap between humans and AI. Unlike traditional AI art generators that produce images based on random or arbitrary inputs, it creates art that is deeply connected to the user's emotional state, making it more relevant and personal. This not only enhances the user experience but also opens up new possibilities for human-AI interaction, where AI acts as a facilitator of personal expression and emotional exploration.

4 Conclusions

In this work, we developed MoodAlbum, a platform that transforms users' emotions into personalised artwork by integrating emotion detection with AI-driven image generation. While the system is currently hosted on a CPU, which limits its performance, especially during image generation, it could significantly benefit from being hosted on a GPU, offering faster and more efficient processing. Future work could explore optimising the pipeline for real-time interaction, expanding the range of emotions and artistic styles, and adding features such as image storage for long-term emotional tracking. This would further enhance the user experience and open up new possibilities for creative and therapeutic applications.

References

- [1] Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative Adversarial Networks, Generating "Art" by Learning About Styles and Deviating from Style Norms. *arXiv preprint arXiv:1706.07068*.
- [2] McCormack, J., Gifford, T., & Hutchings, P. (2019). Autonomy, Authenticity, Authorship and Intention in Computer Generated Art. *Computational Creativity Research: Towards Creative Machines*, 125-147.

- [3] Buechel, S., Athanasiou, N., & Gurevych, I. (2023). Harnessing AI for Human Well-Being: The Role of Generative Models in Visual Art Therapy. *IEEE Transactions on Affective Computing*.
- [4] Yang, W., Zheng, L., & Li, J. (2023). Emotion-Aware AI: Challenges and Opportunities in Affective Computing. *Journal of Artificial Intelligence Research*, 76, 35-59.
- [5] Zhou, H., Wu, Y., & Huang, M. (2022). From Emotion to Art: Affective Computation and AI-Based Artistic Creation. *Journal of Computational Aesthetics*, 19(2), 101-117).
- [6] Haque, S., & Hariri, S. (2020). Generating Emotion-Based Visual Art Using GANs. *Proceedings of the 2020 International Conference on Artificial Intelligence and Advanced Manufacturing*.
- [7] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.
- [8] Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A., Nemade, G., & Ravi, S. (2020). GoEmotions: A Dataset of Fine-Grained Emotions. *arXiv preprint arXiv:2005.00547*.
- [9] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *arXiv preprint arXiv:2112.10752*.