The Physiology of Creativity: Autonomous Robotic Painting Driven by Synthetic Biometric Generation

Angshu Adhya*§, Eric Chen*§, Rishad Hasan*§, Hana Ichikawa*§, Mrunmayee Jere*§, Tetsu Kurumisawa*§, Tianxin Li*§, Cindy Yang*§, Longzhen Yuan*§, Abhishek Narula* and Patrícia Alves-Oliveira*

* University of Michigan

Corrspondence to robopati@umich.edu

***Abstract*—We tokenize multimodal biometric data–heart rate, electrodermal activity (EDA), emotion, and motion capture–and use it to train three generative models for stroke, image and biometric synthesis. These models drive a closed perception-action loop autonomous painting robot.**

***Index Terms*—Human-Robot Interaction, Variational Autoencoder, Embedded Biometrics, Stroke Generation, Image Synthesis, Robotic Painting**

I. Presentation Mode

The video demonstration will show our system producing paintings autonomously over 300-500 strokes ($\sim$ 60 minutes). The robot initializes with random brush strokes, and our demonstration will show how we can achieve artistic creation guided by an autonomous loop using biometrics.

II. Introduction

Existing robotic painting systems use direct mappings (text-to-image [1], image-to-image [2]) that bypass internal physiological states characterizing human creativity. While recent work has explored biometric data visualization in interactive art [3], [4] and emotion-conditioned digital image synthesis [5], [6], no prior art system uses biometrics as a generative latent representation for autonomous robotic painting.

III. System Overview & Methods

Our system demonstrates autonomous robotic painting driven by three sequential components that form a closed perception-action loop.

A. Data Collection

We captured synchronized painting sessions recording: (1) intermediate canvas states, and (2) continuous biometric signals (heart rate, EDA, motion capture, 9-category facial emotion annotations) (See Fig 1).

B. Biometrics Generator

A transformer architecture processes visual inputs, including the current canvas state and optional inspiration images, alongside current biometrics to generate predicted sequences of future biometric responses encoded as embeddings. Training on our dataset of human artist biometrics enables the model to learn temporal progression patterns in physiological responses and their relationship to visual stimuli. The resulting predictions capture how specific visual states elicit biometric changes over a given time window (See Fig 2 and 3).

§Equal Contribution (in alphabetical order)

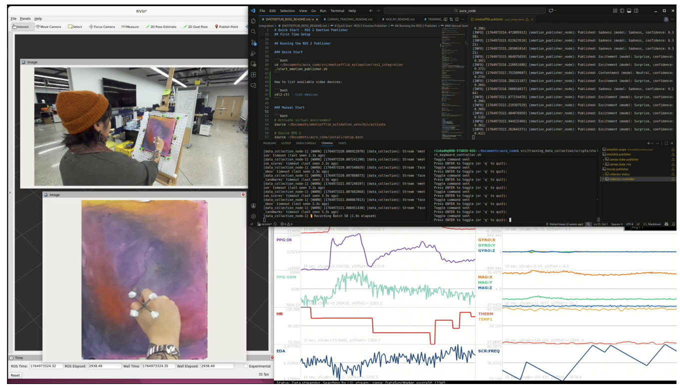

Fig. 1. Data Collection Phase: Upper Left: Artist painting on canvas, Bottom Left: Canvas state being captured, Top Right: Stream of predicted emotion, Bottom Right: Stream of heart rate and EDA

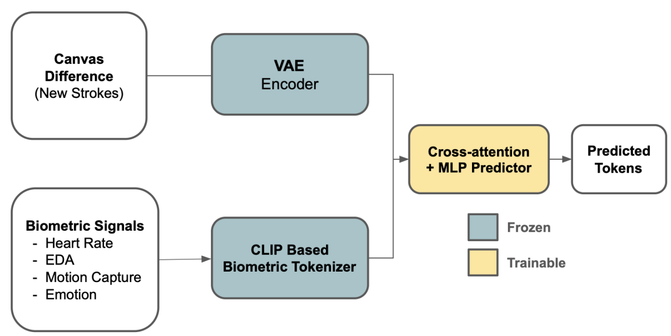

Fig. 2. Biometrics Generator (Token Prediction): Canvas differences are created by subtracting consecutive frames and thresholding the differences to isolate new strokes. Then the current biometric tokens are fused via cross-attention. The MLP predictor learns to predict the next biometric state in the CLIP token space.

C. Biometrics-Conditioned Image Generation

A diffusion model (InstructPix2Pix [7], Stable Diffusion [8]) generates the target image after the next 5-10 strokes, conditioned on both the current canvas and the synthetic biometric signals from component B. We explored multiple conditioning strategies, including cross-attention layers, latent concatenation, adaptive instance normalization (AdaIN), and time-step

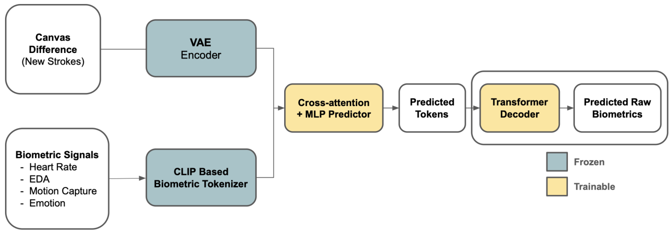

Fig. 3. Biometrics Generator (Biometrics Reconstruction): A decoder translates tokens to raw biometrics for visualization. This enables interpretation of what the robot "feels" while painting, outputting the heart rate, EDA, motion, and emotion time series.

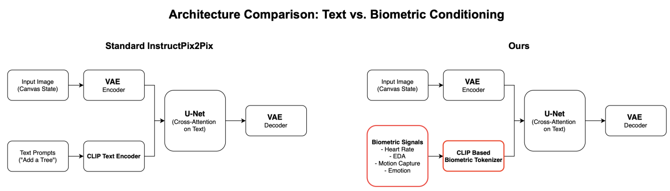

Fig. 4. Biometrics-Conditioned Image Generation: We modify the InstructPix2Pix architecture by replacing the CLIP text encoder with a biometric tokenizer. Instead of language instructions, the U-Net is conditioned on physiological signals through cross-attention layers.

embedding augmentation to integrate continuous physiological signals into the U-Net architecture (See Fig 4).

D. Emotion-Guided Stroke Optimization

Inspired by the Robot Synesthesia paradigm [9], we convert synthetic biometrics to target emotion vectors and use differentiable optimization to generate brush strokes. During stroke generation, candidate strokes are rendered and evaluated under augmentations. Emotion loss (negative cosine similarity between predicted and target emotions) backpropagates through the differentiable renderer to stroke parameters, enabling generation that is both visually and emotionally aligned with the biometric state (See Fig 5).

E. Autonomous Feedback Loop

The robot executes the three-step cycle (biometrics generation $\rightarrow$ target image synthesis $\rightarrow$ emotion-guided stroke optimization), paints the optimized strokes, and repeats for 300-500 strokes. This creates a self-sustaining creative process where the robot's paint strokes generate predicted biometric responses, which condition target image generation, which guides stroke execution—modeling the internal physiological dynamics that characterize human creative processes

IV. Demo Description

Attendees will observe:

- Real-time synthetic biometric visualization (heart rate, EDA, emotion)
- Target images generated from biometric conditioning
- Painting strokes optimized for emotional consistency
- Emergent artistic behavior from the perception-action loop

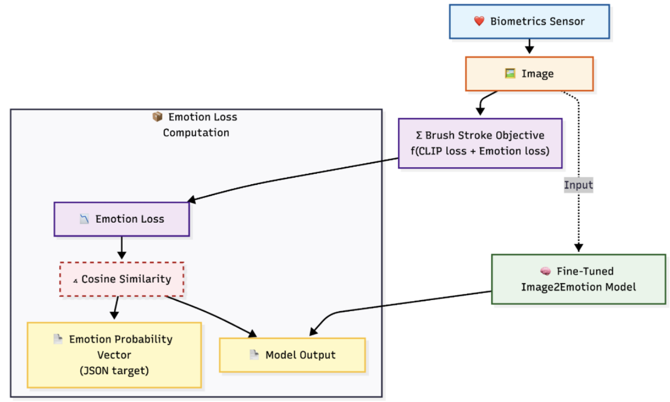

Fig. 5. Emotion-Guided Stroke Optimization: The Emotion Loss Computation subgraph translates emotional intent into a differentiable training signal for brush stroke optimization. A target emotion distribution (JSON) is compared against the emotions predicted from the generated image by a fine-tuned Image-to-Emotion model. Their cosine similarity produces an emotion loss term, which guides stroke optimization toward emotional expressiveness.

See our video for the research process and presentation concept.

V. Contributions

- **Biometrics as Creative Latent Space:** First use of synthetic biometric generation as an autonomous driver for robotic art. Unlike systems requiring collaboration [2] or explicit instructions [1], we model physiological states as a novel intermediate representation between intent and mark-making.
- **Closed-Loop Perception-Action Architecture:** Self-sustaining creative system combining biometric prediction, biometric-conditioned image synthesis, and emotion-guided stroke optimization—demonstrating emergent artistic behavior through continuous biometric modeling.

VI. Discussion

This work positions itself not only as a technical demonstration of robotic painting but as an inquiry into the quantification of human creativity. Rather than relying on high-level natural language expressions, the system uses physiological signals learned from real human painters as the latent variables driving the creative process — bringing us closer to a sci-fi notion of "cloning" the human artist from the inside out. In doing so, the project questions the assumptions embedded in contemporary AI systems that seek to formalize internal states such as emotion, attention, and intent as a means of replicating human cognition. By attempting to reduce artistic expression to measurable variables, the work exposes both the ambition and the limits of that reduction.

Through the video demonstration of the system producing an abstract painting, we invite the audience to sit with the following questions:

- Are our high-level artistic intentions distilled into our low-level physiological signals or does something essential escape the body's measurable trace?
- What aspects of the creative process can we make visible and can we see the causal relationship between biometrics and the painting?
- When the robot's generated images are optimized to mimic the correlation between biometrics and paintings from a human painter, are we "copying" the creativity of the painter that is providing the data?
- When the robot's strokes are optimized for emotional consistency with simulated emotions, who (or what) is the author of the resulting painting?

Acknowledgments

This project is funded by DARPA Young Faculty Award (YFA) #D24AP00323-00. The content of this research is solely the responsibility of the authors and does not necessarily represent the official views of DARPA.

References

[1] P. Schaldenbrand, J. McCann, and J. Oh, "Frida: A collaborative robot painter with a differentiable, real2sim2real planning environment," (London, United Kingdom), pp. 11712–11718, IEEE, 2023.

[2] P. Schaldenbrand, G. Parmar, J.-Y. Zhu, J. McCann, and J. Oh, "Cofrida: Self-supervised fine-tuning for human-robot co-painting," (Yokohama, Japan), pp. 2296–2302, IEEE, 2024.

[3] C. Livia and K. Wolf, "Emotional resonance: An interactive installation controlled by biometric data," in *Electronic Workshops in Computing*, BCS Learning Development, 2018.

[4] L. A. Sparrow, C. Galwey, B. Loveridge, S. Glasser, M. S. Osborne, and R. M. Kelly, "Heart and soul: The ethics of biometric capture in immersive artistic performance," in *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI 2024, Honolulu, HI, USA, May 11-16, 2024* (F. F. Mueller, P. Kyburz, J. R. Williamson, C. Sas, M. L. Wilson, P. O. T. Dugas, and I. Shklovski, eds.), pp. 494:1–494:23, ACM, 2024.

[5] Q. Lin, J. Zhang, Y. S. Ong, and M. Zhang, "Make me happier: Evoking emotions through image diffusion models," *CoRR*, vol. abs/2403.08255, 2024.

[6] J. Yang, J. Feng, and H. Huang, "Emogen: Emotional image content generation with text-to-image diffusion models," (Seattle, WA, USA), pp. 6358–6368, IEEE, 2024.

[7] T. Brooks, A. Holynski, and A. A. Efros, "Instructpix2pix: Learning to follow image editing instructions," (Vancouver, BC, Canada), pp. 18392–18402, IEEE, 2023.

[8] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," (New Orleans, LA, USA), pp. 10674–10685, IEEE, 2022.

[9] V. Misra, P. Schaldenbrand, and J. Oh, "Robot synesthesia: A sound and emotion guided robot painter," (Abu Dhabi, United Arab Emirates), pp. 5986–5992, IEEE, 2024.