LACE: Controlled Image Prompting and Iterative Refinement with GenAI for Professional Visual Art Creators

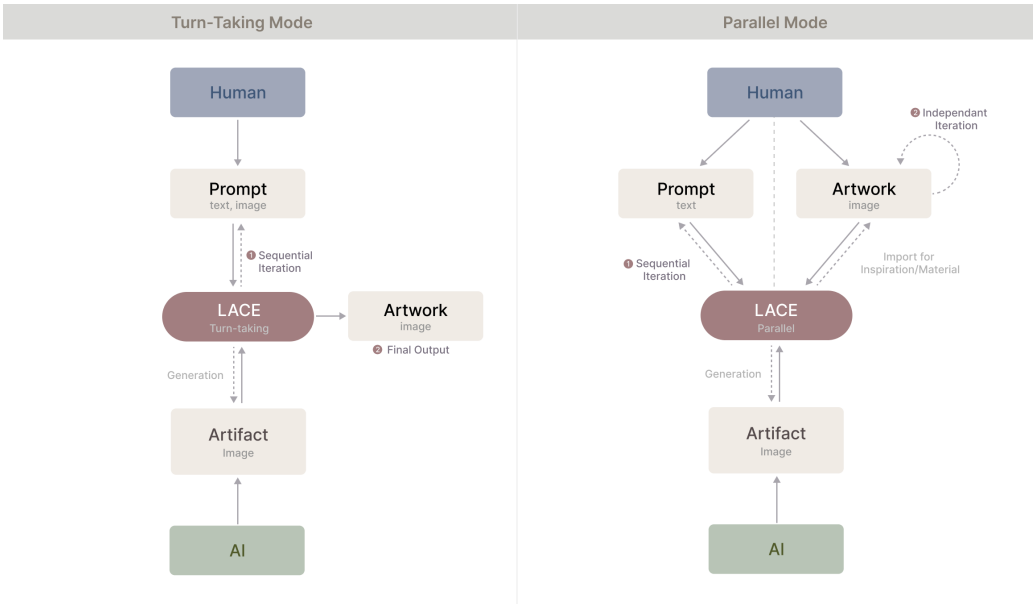


Fig. 1. Turn-Taking vs. Parallel: Turn-Taking follows a sequential workflow where human and AI iterate on prompts and outputs step by step. Parallel mode enables simultaneous refinement, with AI adapting to snapshots of the artist's evolving canvas, fostering a dynamic interplay between human and machine contributions.

Additional Key Words and Phrases: creativity support, generative AI, professional creative tools, image creation

ACM Reference Format:

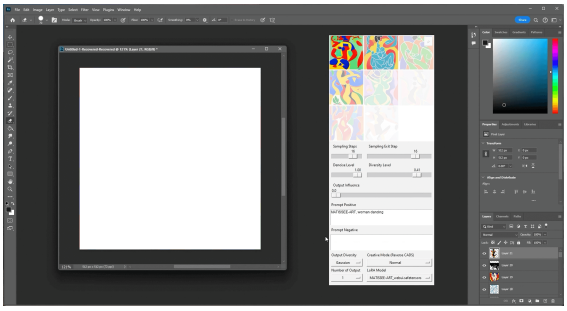
. 2025. LACE: Controlled Image Prompting and Iterative Refinement with GenAI for Professional Visual Art Creators. 1, 1 (February 2025), 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

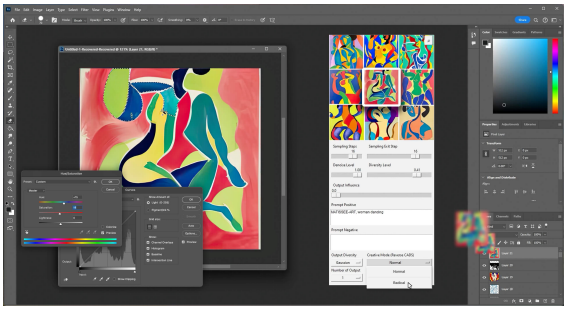
The advent of generative AI has opened exciting possibilities in professional art creation by generating high-quality outputs almost instantly. However, recent research [1, 4, 6, 7] highlights several challenges that hinder its practical adoption by professional users:

Author's Contact Information:

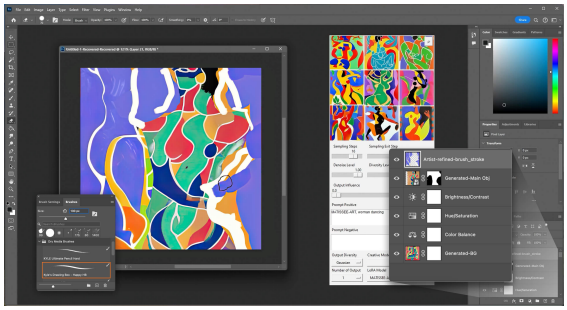
Manuscript submitted to ACM



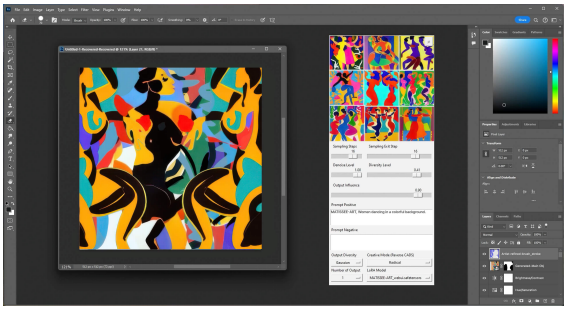
(a) LACE integrates into Photoshop via a popup GUI, linking manual edits with AI.



(b) LACE uses Photoshop's robust tools to guide precise image prompting.



(c) By layering AI outputs in Photoshop, LACE keeps artwork consistent without extra tuning.



(d) AI adapts in real time to the evolving canvas, creating a dynamic two-way creative exchange.

Fig. 2. Overview of LACE: Interface and Interaction

- **Inadequate expressiveness of text-based prompting**. Purely textual descriptions often lack the precision required to control nuanced elements such as lighting, texture patterns, subtle positioning, and overall artistic ambiance. This ambiguity makes it challenging for artists to achieve their intended vision.
- **Lack of coherence in iterative refinement**. Generative workflows are highly sensitive to input parameters [9], making it difficult for artists to progressively build on previous versions and maintain creative continuity.
- **Incompatibility with established workflows**. Many artists are reluctant to switch to unfamiliar AI tools, especially when these tools disrupt their current routines or require specialized technical knowledge [7].

To address these issues, we introduce the *Latent Auto-recursive Composition Engine* (LACE), a system that integrates generative AI into industry-standard environments. LACE provides fine-grained control through Photoshop's familiar editing features (e.g., blending modes, curves, layer-based adjustments), letting artists seamlessly refine and compose AI-generated outputs. By coupling direct manipulation with a customizable AI pipeline, LACE allows users to balance manual control and automated generation, enhancing both creative freedom and iterative coherence.

Our preliminary pilot study ($N = 21$) indicates that LACE significantly improves usability, user ownership, and overall satisfaction compared to baseline AI workflows. In the following sections, we describe LACE's design, present key findings from our user study, and discuss implications for integrating generative AI into professional art practices.

2 LACE SYSTEM

LACE bridges Photoshop and an AI pipeline (Fig 3), allowing users to edit AI-generated elements as individual layers rather than as flattened images. By leveraging Photoshop's native features (e.g., layers, masks, blending modes, and color adjustments), LACE maintains familiar workflows while integrating AI-assisted image creation [2, 3] and generative control mechanisms [5, 10]. Grounded in the COFI framework [8] by Rezwana et al., LACE supports both turn-taking and parallel collaboration for flexible co-creative interactions. Additionally, LACE's modular design allows users to configure or replace the underlying AI pipeline and connect new models directly to Photoshop.

2.1 Layer-Based Prompting

Instead of generating flattened images, LACE imports AI outputs into Photoshop as separate layers (Fig 4), enabling creators to isolate and modify foreground, mid-ground, and background elements independently without retraining or fine-tuning models. Conventional AI tools produce a single, flattened image, making it challenging to adjust individual components without extensive manual editing. In contrast, LACE allows artists to create distinct layers for different parts of their artwork, offering a level of control and precision that traditional methods cannot achieve. This layer-based approach ensures that the iterative refinement process remains flexible and artist-driven while continuously integrating AI-generated content.

2.2 Flexible Collaboration Modes

LACE integrates with Photoshop's editing timeline, enabling continuous interaction between the artist and AI (Fig 1). In this workflow, artists refine images while the AI generates suggestions, forming a feedback loop that maintains creative continuity. This structure supports both sequential turn-taking for precise adjustments and parallel interaction, where AI suggestions evolve as the artist works. Fluid transitions between these modes enhance flexibility, allowing creators to tailor AI collaboration to their process.

- **Turn-Taking mode**: Starting from scratch, users provide text prompts that generate AI outputs, forming the basis for iterative, step-by-step refinement.
- **Parallel mode**: AI adapts to the evolving canvas, producing outputs that inspire the artist while incorporating their refinements in real time.

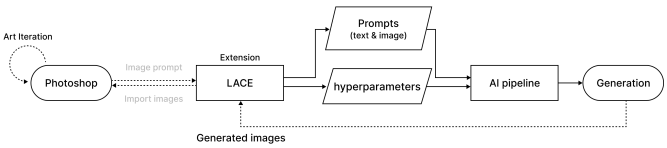


Fig. 3. LACE's modular design lets users configure or replace the underlying AI pipeline.

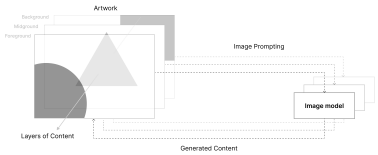


Fig. 4. Layer-based prompting offers granular control over AI-generated elements.

3 STUDY AND KEY FINDINGS

We conducted a within-subject pilot study with 21 participants (ages 20–27; 45% beginners, 25% intermediate, 30% advanced) to assess LACE's impact on creativity and workflow. They completed one of three art tasks (Representational,

Non-Representational, or Design Challenge) using three workflows: W1 (text-to-image), W2 (text-to-image with latent consistency), and W3 (LACE, image-to-image with latent consistency). Workflow order was randomized, with each session lasting 15 minutes. All generative model parameters remained consistent across conditions.

3.1 Key Findings

Quantitative findings show a strong preference for LACE (W3) over the two text-based workflows (W1 and W2). A Friedman test identified significant differences in satisfaction ($p = 0.039$), ownership ($p = 0.009$), usability ($p = 0.003$), and artistic perception ($p = 0.005$). Post-hoc comparisons indicate that LACE outperforms both text-only (W1) and text-based iterative (W2) approaches, most notably in usability and the sense of ownership.

3.1.1 ***Improved Usability and Ownership.*** Participants found it easier to position and refine image elements using LACE's layer-based editing compared to purely text-based methods. They reported higher artistic control and fewer prompt-driven uncertainties. Overall, 71.4% selected LACE as their favorite workflow, citing its ability to retain and combine favorable elements from earlier iterations. The pixel-art design challenge (T3) generated mixed feedback; although users appreciated the deeper control, some found it more time-intensive.

3.1.2 ***Correlation Between Ownership and Usability.*** Spearman correlation analyses reveal that higher usability ratings in LACE are associated with stronger ownership and satisfaction scores. By contrast, text-based workflows do not show a similar relationship—participants who found the interface straightforward did not consistently feel satisfied with the final outputs. Some noted difficulties in generating precise compositions or styles through text alone, highlighting a gap between prompt adjustments and desired visual outcomes.

3.1.3 ***Engagement in Structured Tasks.*** Time logs indicate that participants spent more effort with LACE, particularly on structured design tasks like T3. While text-based prompts required fewer steps, LACE's iterative image edits encouraged ongoing refinements. This suggests that LACE may increase both creative investment and the learning curve for users less familiar with advanced image-editing features. However, participants generally viewed the extra effort as worthwhile, citing improved alignment between their vision and the final results.

4 DISCUSSION AND FUTURE WORK

Our findings suggest that the choice of Human-AI collaboration mode depends on both the creative task and its stage. Participants generally preferred turn-taking during early ideation, where iterative text prompts facilitated concept exploration, while later, detail-oriented stages benefited from parallel interaction for real-time refinements. However, these preferences are not rigid; the interplay of timing, task complexity, and user goals remains nuanced and warrants further investigation. Future research will examine how different workflows and phases of the creative process influence user perception and behavior.

Interestingly, participants favored LACE even when its outputs were suboptimal, suggesting that artistic agency during iteration significantly impacts user experience. This highlights the importance of control and interactivity in AI-assisted creativity. Moving forward, our goal is to extend LACE's design principles to other creative domains supported by AI, such as 3D rendering, animation, and graphic design, where fine-grained control and iterative workflows are essential. Future work will also explore adapting collaboration modes for different skill levels and optimizing layer-based editing to enhance usability without overwhelming novice users.

References

[1] AC BD, Christine Bauer, and Afsaneh Doryab. 2016. Solving the Battle of First-Authorship: Using Interactive Technology to Highlight Contributions. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI EA '16)*. Association for Computing Machinery, New York, NY, USA, 609–620. https://doi.org/10.1145/2851581.2892582

[2] Katica Bozsó, András Béres, and Bálint Gyires-Tóth. 2024. Semantic segmentation mask-guided diffusion models: A pathway to enriched datasets in autonomous systems. 79–84. https://doi.org/10.3311/WINS2024-014

[3] Yu Cao, Xiangqiao Meng, PY Mok, Xueting Liu, Tong-Yee Lee, and Ping Li. 2023. AnimeDiffusion: Anime Face Line Drawing Colorization via Diffusion Models. *arXiv preprint arXiv:2303.11137* (2023).

[4] Fiona Draxler, Anna Werner, Florian Lehmann, Matthias Hoppe, Albrecht Schmidt, Daniel Buschek, and Robin Welsch. 2023. The AI Ghostwriter Effect: Users Do Not Perceive Ownership of AI-Generated Text But Self-Declare as Authors. *ArXiv* abs/2303.03283 (2023). https://doi.org/10.48550/arXiv.2303.03283

[5] Aosong Feng, Weikang Qiu, Jinbin Bai, Kaicheng Zhou, Zhen Dong, Xiao Zhang, Rex Ying, and Leandros Tassiulas. 2024. An Item is Worth a Prompt: Versatile Image Editing with Disentangled Control. *arXiv preprint arXiv:2403.04880* (2024).

[6] Steven Heller. 2024. *How Worried Should Creative Professionals Be About Artificial Intelligence?* https://www.printmag.com/web-interactive-design/how-worried-should-creative-professionals-be-about-artificial-intelligence/ Accessed: 2024-10-01.

[7] Srishti Palani and Gonzalo Ramos. 2024. Evolving Roles and Workflows of Creative Practitioners in the Age of Generative AI. In *Proceedings of the 16th Conference on Creativity & Cognition* (Chicago, IL, USA) *(C&C '24)*. Association for Computing Machinery, New York, NY, USA, 170–184. https://doi.org/10.1145/3635636.3656190

[8] Jeba Rezwana and Mary Lou Maher. 2023. Designing Creative AI Partners with COFI: A Framework for Modeling Interaction in Human-AI Co-Creative Systems. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. ACM, 1–28.

[9] Qiucheng Wu, Yujian Liu, Handong Zhao, Ajinkya Kale, Trung Bui, Tong Yu, Zhe Lin, Yang Zhang, and Shiyu Chang. 2023. Uncovering the disentanglement capability in text-to-image diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 1900–1910.

[10] Lingfeng Yang, Yueze Wang, Xiang Li, Xinlong Wang, and Jian Yang. 2024. Fine-grained visual prompting. *Advances in Neural Information Processing Systems* 36 (2024).

A SAMPLED QUALITATIVE RESULTS

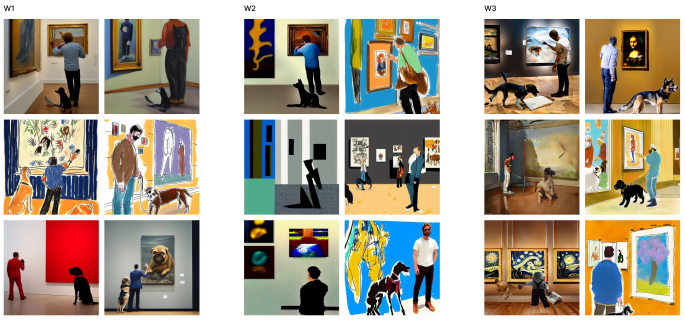


Fig. 5. Results from Task 1. Prompt: "A man reaching for a painting in an art gallery, accompanied by a dog sniffing another artwork on the floor."

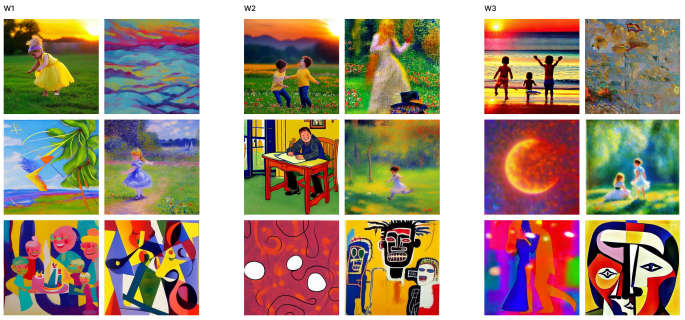


Fig. 6. Results from Task 2. Prompt: "An abstract composition that embodies the dynamics and motion associated with joy."

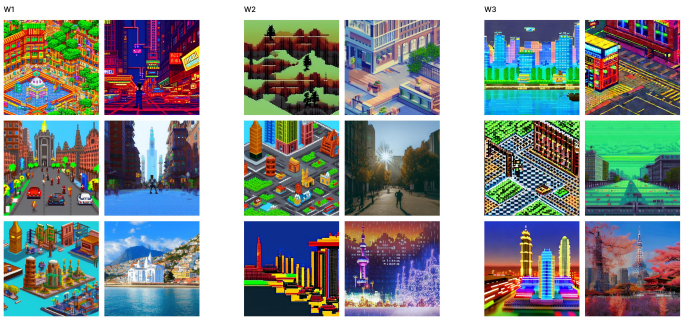


Fig. 7. Results from Task 3. Prompt: "A pixel art game scene with a bustling cityscape featuring assorted architectural styles."

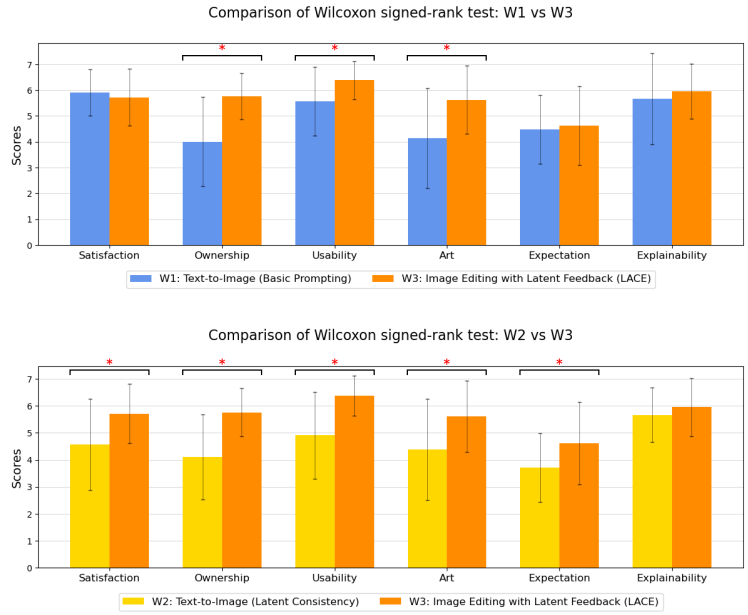


Fig. 8. The charts compare Workflow 1 and Workflow 2 against Workflow 3 (LACE) using Likert scale scores (1-7) across six categories. Red asterisks (*) indicate statistically significant differences between the workflows, while error bars represent score variability across participants.

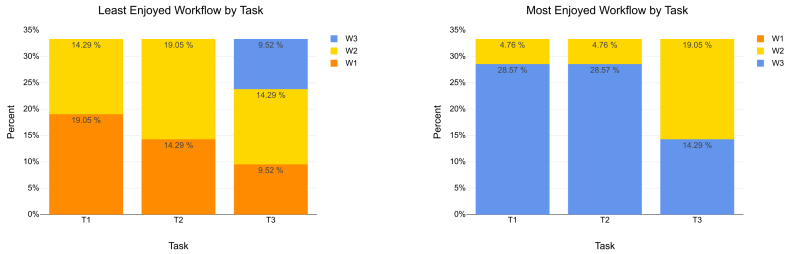


Fig. 9. Preference of Workflow by Tasks

Fig. 10. Spearman Correlation Heatmap of Likert-Scale Variables and task completion time