

## **Transfer of Control: Reconstructing the Role of Artists in AIGC Image Co-Creation**

Yichen Wei, Renmin University of China, China

The Southeast Asian Conference on Arts & Humanities 2026  
Official Conference Proceedings

### **Abstract**

This study explores visual language bias and cultural stereotypes in AIGC image creation through the AI-generated artwork *Ongoing Memories*. Addressing challenges such as AI images misaligning with artistic intent and reproducing stereotypical visual prototypes, the study proposes a “control delegation” strategy. Artists establish collaborative mechanisms by setting data boundaries and style prompts while allowing AI a degree of visual autonomy. The project reinterprets “historical” visual construction through a stylized dataset and dual-layer image structure, offering new methodological references for AIGC images within the contemporary art context.

*Keywords:* AI image generation, visual bias, cultural memory, artistic collaboration, control delegation

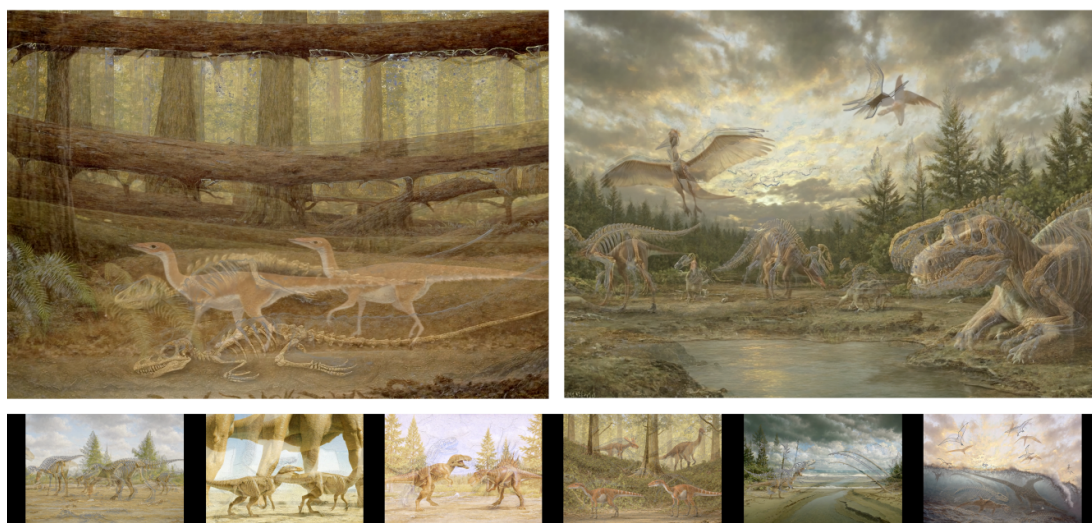
**iafor**

The International Academic Forum  
[www.iafor.org](http://www.iafor.org)

## Introduction

**Figure 1**

*Stills From Ongoing Memories*



*Note.* Selected stills from *Ongoing Memories*, showing the dual-layer structure of AI-generated imagery and fossil photographs.

In recent years, AIGC image generation has been widely applied in the field of artistic creation, greatly enriching visual expression forms. However, it has also brought about numerous methodological challenges. As contemporary artists focused on visual control and aesthetic consistency, we have long attempted to create art using AIGC technology but have repeatedly encountered failures in visual control. Specifically, these failures manifest in two aspects: first, the difficulty in aligning visual language with user expectations; second, the generated results often exhibit strong cultural biases.

In terms of visual language consistency, AI-generated models generally suffer from issues such as misaligned visual elements and stylistic deviations, failing to accurately convey the creator's aesthetic and logical intentions. AI struggles to correctly map descriptive attributes to objects (Zarei et al., 2024), while existing diffusion models exhibit limited generalization capabilities when encountering novel combinations they have not encountered before, often automatically reverting to patterns common in the training dataset, resulting in semantically confusing and illogical generated images (Park et al., 2021).

Beyond visual aesthetic issues, cultural biases and the replication of stereotypical archetypes in AI image generation constitute another major obstacle in creative practice. Extensive research has demonstrated that the outputs of AI-generated models are significantly influenced by the inherent cultural patterns in their training data, leading to generated images that tend to repeat and reinforce existing cultural biases (Grba, 2022; SIGGRAPH Digital Arts Community, 2025; University of Washington News, 2023).

**Figure 2***Example of Style Hallucination*

*Note.* Example of “style hallucination” in AI-generated output, where visual elements not present in the training prompts emerged unexpectedly.

In our personal creative attempt to use AI to generate imaginary alien creatures, even without explicit cinematic prompts, the AI still tended to output monster images similar to those in classic science fiction movies. This demonstrates that the model is not engaging in true creation but rather reproducing pre-existing “cultural visual memories.”

Furthermore, from the perspective of the overall art system, most AIGC creation practices are confined to technical demonstrations, AI art competitions, or discussions about AI as a tool for reflecting on the future, rather than being truly integrated into the theoretical framework and cultural context of contemporary art. In this context, the emphasis is often on technical prowess, lacking a deep dialogue with artistic language and theory (Grba, 2022). This state of creation suspended outside the art system constrains the in-depth development and exploration of the lasting value of AIGC image production in an artistic context.

In response to the above-mentioned challenges in the use of technology itself and the time required for its implementation, this study attempts to propose a new creative strategy: consciously ceding some visual control to algorithms and allowing the generation of AIGC images to become an integral part of the work's logic, thereby exploring a more flexible collaborative mechanism between artists and AI. Specifically, this study constructs a collaborative method based on the cultural characteristics of the training set through the creation of the video artwork “Ongoing Memories,” enabling AI to exercise autonomy within the data and style parameters provided by the artist, thereby generating visual expressions that blend human intent with algorithmic characteristics. This method not only offers a new theoretical perspective for artistic creation but also demonstrates a potential path for AIGC image generation to achieve methodological coherence within the contemporary art context planet.

### **Related Works and Key Issues**

This section will further analyze the two types of issues discussed in the previous section in the context of art, based on existing research and artistic practices. The focus of this paper is not on works that showcase AI models themselves, but rather on exploring how the role of AI images should be defined when they are used as part of a work, and proposing a creative strategy of “transfer of control” based on these companies.

## **The Heterogeneity of AI Visual Language in the Art Context**

Current AIGC image creation heavily relies on large image datasets, which often carry the cultural perspectives and biases of the creators. For example, ImageNet contains biased labels targeting people (Sivertsen et al., 2024). When AI models learn from these datasets, they inevitably internalize the implicit cultural biases into their own visual expression logic, resulting in generated images with obvious cultural tendencies (Sivertsen et al., 2024). When these culturally biased AI images enter the art domain, they produce a noticeable “heterogeneity”—that is, there is a significant asymmetry between the meaning expressed by AI images and the expectations of the artistic context.

Lev Manovich's research also points out that AI-generated images are good at replicating existing visual patterns in training sets, but they perform poorly when it comes to truly innovative, complex, or unprecedented conceptual expressions. This makes the content they generate often conform to mainstream aesthetics, but difficult to adapt to the diverse demands of artistic creation (SIGGRAPH Digital Arts Community, 2025). Additionally, existing reception studies indicate that viewers tend to lower their aesthetic evaluations of works when informed that the images were generated by AI, revealing a value recognition barrier for AI images in the realm of artistic reception (De Rooi, 2024). The above research highlights that cultural biases in AI visual language and the mismatch with artistic contexts constitute one of the core challenges facing current AIGC art practices.

### **The Shift From Tool to Collaborator**

Despite the aforementioned biases and heterogeneity issues, some artists are actively exploring how to leverage the deviations and errors in AI-generated images, attempting to transform AI from a tool into a “visual agent” or even a collaborator in artistic creation.

For example, Refik Anadol inputted a large amount of visual data into his work “Machine Hallucinations,” allowing AI to autonomously generate abstract images filled with uncertainty. He incorporated AI's uncertainty and visual hallucinations as aesthetic elements of the creation itself, a strategy that clearly leverages the inherent biases in AI visual expression (Sivertsen et al., 2024). Mario Klingemann's Neural Glitch Art, on the other hand, directly uses glitches and distortions generated during the AI process, explicitly transforming AI errors and loss of control into creative and aesthetic language, continuing the artistic tradition of Glitch Art (Sivertsen et al., 2024). These artistic practices clearly demonstrate that AI is no longer merely viewed as a tool but as an agent with partial autonomy, bias, and creative agency, forming a complex interactive relationship with artists.

However, most of the above practices emphasize AI as a creative subject with partial autonomy, and its “visual bias” is directly incorporated into the work itself. This approach focuses more on the “behavioral nature of algorithms,” that is, how AI continuously generates images, sounds, or structures during its own operation, thereby constituting a continuously evolving work site. Although this “AI as art” approach is experimental, its creative boundaries are relatively clear, and it has not been the subject of intense controversy (Putra, 2025).

The approach we focus on in this study is different. We do not regard the AI model itself as part of the artwork, nor do we make its operational process public or present it in a dynamic form. Instead, we are more concerned with how to use AI-generated images in artistic creation and embed these images as a medium of visual composition into a larger narrative and cultural

reflection framework. In this sense, AI is not the “exhibited subject” of the work but a visual collaborator—its role is akin to that of a co-creator with pre-existing biases, whose generated images, though not fully controllable, can function within the artist's framework.

### **Method**

In response to the dual challenges of visual control and cultural bias in AIGC, this study proposes a human-machine collaborative creation strategy called “control transfer” using the experimental video work *Ongoing Memories* as a practical vehicle. This method constructs a dataset with explicit cultural references, sets fictional style labels, employs the LoRA training mechanism, and combines a dual-layer narrative structure to explore how AI can generate images with shared visual characteristics within the framework defined by the artist. The following sections will elaborate on this approach from four perspectives: image structure, dataset preparation, model training, and collaborative logic.

### **Creative Objectives and Basic Concepts**

“*Ongoing Memories*” is an experimental video artwork based on AI image generation. Its objective is not to recreate the true appearance of prehistoric life but to explore, through AI's visual production mechanisms, how “our understanding of the past” is continuously constructed and imagined within the current cultural logic. The project begins with a seemingly simple question: When we talk about “dinosaurs,” where do the images that come to mind originate?

These images do not originate from direct observation or pure archaeological evidence, but rather from the continuous shaping of visual culture through films, children's picture books, museum exhibitions, and other sources. They represent a collective image memory that has been encoded and solidified. Based on this understanding, the work employs a dual-layer narrative structure: one layer consists of AI-generated dinosaur images, representing the emotionalized and anthropomorphized representations circulating in contemporary visual culture.

The other layer consists of real fossil images, symbolizing the past as “material remains.” These two types of images continuously intertwine and repeat in the footage with equal transparency, creating a visual loop and interference of time while revealing the blurred boundaries between “reality” and “fiction” on a conceptual level. Through this structure, the work aims to present an illusory mechanism of historical sense, that is, the “past” we perceive is not a stable and reproducible truth, but rather an image construct that is constantly being regenerated in the present culture.

### **Dataset Construction and Style Control Strategy**

To support the aforementioned visual narrative structure, this project constructed a specialized dinosaur image dataset, with image materials mainly sourced from contemporary visual contexts such as museum exhibitions, video games, and children's picture books. These images typically exhibit highly stylized characteristics, reflecting the mainstream visual perceptions of “dinosaur imagery” in contemporary society. The selection of these materials aims to reveal that our visual impressions of prehistoric creatures are actually the result of repeated construction and solidification by cultural media.

**Figure 3**  
*Training Dataset Overview*



*Note.* Selected training dataset images, sourced from games, artworks, children’s books, and museum displays.

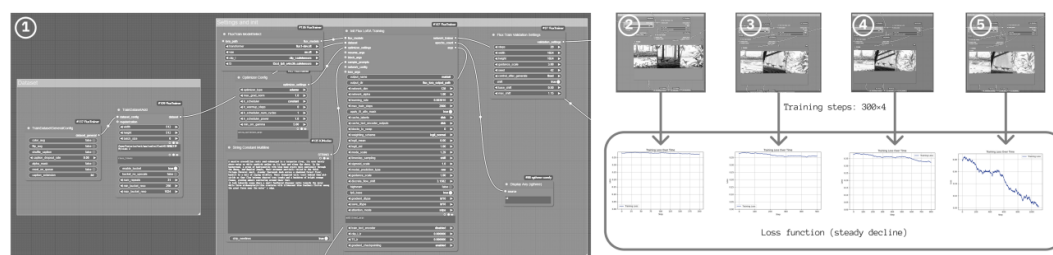
During the dataset construction process, each image is accompanied by a corresponding tag file that provides a detailed description of its content and stylistic elements, such as thematic composition, emotional tone, and color usage. In order to enable AI to better distinguish between “image content” and “style features,” this project also specifically set a fictional and unique style prompt word “dinoblend” and placed it uniformly at the beginning of all tags. This setting ensures that the model can stably identify and call pre-set styles during training and maintain consistency in style when generating images, while retaining a certain degree of content freedom.

The training process of this project also reflects the differences in objectives between rounds. From the first to the third round of training, the core objective is to improve the model's ability to fit specific image features, especially to enable the model to more effectively grasp the required stylistic elements while maintaining consistency in the prompt words. In the later stages of training, artists use prompt word combinations that are as consistent as possible with those used during training in the generation phase to improve the stability and logical consistency of the images generated by the model.

**Model Training and Technical Implementation**

In terms of technical implementation, this project adopted the LoRA training strategy. This method enables efficient style learning and image generation on small-scale datasets, effectively improving the model's ability to capture specific visual features. Compared to traditional model fine-tuning methods, LoRA is more cost-effective in terms of computing resources and training time, making it particularly suitable for targeted style transfer tasks.

**Figure 4**  
*LoRA Training Diagram*



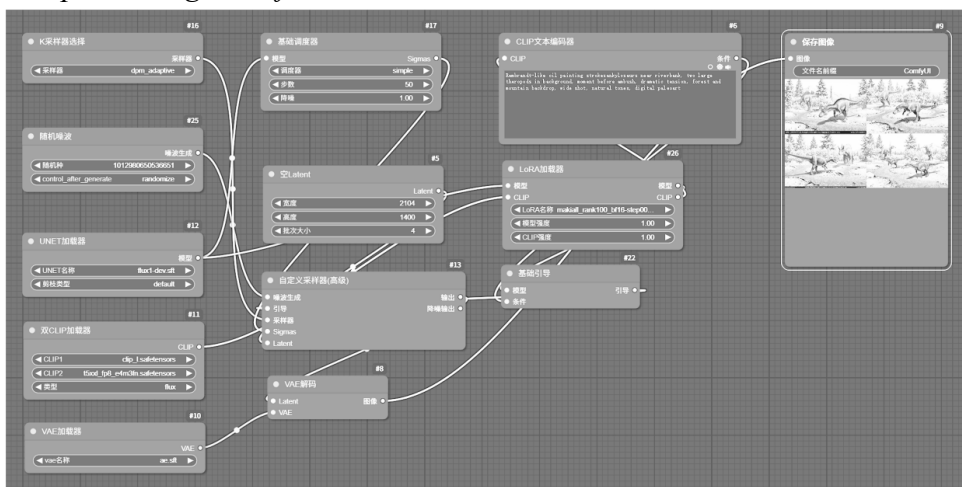
*Note.* Diagram of the LoRA training process with parameter tuning and prompt control; left shows training flow, right shows prompt semantic structure.

A total of four training rounds were conducted, with each round set to 300 training steps, for a total of 1,200 steps. During training, we continuously monitored changes in the model's loss function to assess training stability and generation performance. The overall training process showed a good convergence trend, indicating that the model was able to stably learn the target style and maintain consistent visual features in subsequent generations.

### Image Screening and Cleaning Strategy

Although this project uses an automated AI generation mechanism, this does not mean that all generated images will be directly used for image creation. In practice, artists will screen and correct the generated images as necessary, mainly in the following aspects: First, images with obvious visual logic errors, such as “sails growing out of a dinosaur's back,” which violate structural common sense, will be directly discarded. Such issues often arise because the model cannot accurately map descriptive attributes to target objects, leading to mismatches in visual elements such as color, texture, or position in unintended locations within the image. Second, some images blend multiple unrelated elements in their overall composition, creating a visual “patchwork” effect or semantic “misalignment.” These are also considered uncontrollable noise and are removed. Additionally, the project prioritizes consistency between generated images and the overall visual style of the project. For example, when the visual tone of a batch of generated images deviates significantly from the narrative tone, the artist will remove or regenerate them. The screening strategy does not pursue absolute standards, but rather a relative stability in narrative and aesthetics.

**Figure 5**  
*Post-processing Workflow*



Note. Workflow diagram of post-processing for AI-generated images.

### Visual Narrative Strategy of Double-Layer Image Structure

At the image creation level, Ongoing Memories uses a double-layer, intertwined, parallel visual narrative structure. The first layer consists of AI-generated dinosaur images with pronounced emotional and anthropomorphic characteristics, reflecting contemporary visual culture's reinterpretation of prehistoric life; the second layer is composed of real fossil images, representing fragmented, irreparable material remains. The two layers of images continuously alternate and overlap at the same transparency, interfering with and permeating each other, creating a visual tension with blurred boundaries.

This cyclical image overlap not only breaks down the traditional distinction between “reality” and “fiction” but also induces a temporal delay and dislocation in the viewing experience. As viewers struggle to discern the origin of the images, they gradually realize that our understanding of the “past” is fundamentally a visual construction of the present culture. The work thus emphasizes that history is not a static truth but a dynamic process constantly reorganized by emotions, media, and context.

To enhance the overall coherence of the dual-layer imagery, the project underwent post-production adjustments to the tone and clarity of the images, ensuring that visuals from different sources seamlessly transition in narrative rhythm and visual style. While this post-processing minimizes intervention in the generated content itself, it still enhances the fusion between images, reinforcing the emotional consistency and structural integrity of the work during viewing. The “control cession” strategy adopted in this project involves the artist not attempting to precisely control every output of the AI, but rather setting clear boundaries for its generation through the design of datasets, style tags, and prompts, while accepting a certain degree of visual autonomy within these boundaries.

### Discussion

The practical process of this project demonstrates that the key challenge of AIGC image generation in artistic creation is not simply whether the images are “beautiful” or “accurate,” but rather whether we can rethink the boundaries of artistic control and modes of expression through collaboration with AI. The creation process of *Ongoing Memories* does not seek to suppress the uncontrollable aspects of AI-generated content but actively incorporates them into the creative workflow. By setting frameworks rather than precise instructions, the model engages in a limited yet genuine visual collaboration.

From this perspective, “control transfer” is not a technical issue but a methodological choice: it requires artists to clearly define creative parameters such as datasets and style prompts, while also being willing to accept the uncertainty and heterogeneity brought by AI and transform them into part of the work's semantic meaning. The significance of this strategy lies not only in expanding the ways images are produced, but also in prompting us to rethink the role of the artist—the artist is no longer the sole visual outputter, but rather someone who defines direction, guides the process, and reconstructs meaning from the results within a collaborative system.

It is worth emphasizing that this creative strategy also raises new questions: When AI becomes a co-creator of visual results, does the subjectivity of artistic expression shift? Has AI, to some extent, gained “visual agency”? These questions have no standard answers, but they prompt us to move AI beyond its traditional roles as a “tool” or “creator,” and instead view it as a generative participant situated between the two. It possesses computational capabilities but lacks intent, has structure but lacks understanding—this ambiguity is precisely the most intriguing aspect of our current collaborative relationship with AI.

In this sense, the value of this project lies not in producing a set of images or completing a video work, but in its role as a process that opens up a new way of questioning visual collaboration. We are beginning to realize that AI is not a substitute for creativity, nor is it a fully transparent tool; it exists between the two, intervening in our visual narratives with a “biased logic.” The collaboration between artists and AI is not about efficiency, but about the redistribution of choice, control, and interpretation.

## **Conclusion**

This study proposes a creative strategy of “transfer of control” and attempts to establish a collaborative mechanism based on boundary settings, given that generative AI cannot be completely controlled. Through the practice of “Ongoing Memories,” we see that when artists no longer pursue absolute control over image generation but instead accept the visual biases and structural limitations brought about by AI, artistic creation gains a new kind of flexibility. This flexibility does not stem from AI’s “creativity,” but rather from how artists define collaborative methods and the meaning of images within an asymmetrical relationship.

Future research can further explore three areas: first, how to improve the control precision of prompt words and training labels; second, how to maintain visual style consistency in more complex multimodal structures; and third, how to understand the psychological acceptance mechanisms of audiences when faced with AIGC images, thereby enriching the interpretive pathways of their cultural contexts.

## **Acknowledgements**

The author thanks all those who provided support and feedback during the development of this project, especially in dataset preparation and image model training.

## **Declaration of Generative AI and AI-Assisted Technologies in the Writing Process**

The author declares that Grammarly, an AI-assisted writing software, was used in proofreading and refining the language used in the manuscript. The usage was limited to correcting grammatical and spelling errors and rephrasing statements for accuracy and clarity. The author further declares that, apart from Grammarly, no other AI or AI-assisted technologies have been used to generate content in writing the manuscript. The ideas, design, procedures, findings, analyses, and discussion are originally written and derived from careful and systematic conduct of the research.

## References

- De Rooij, A. (2024). *Audience perception of AI-generated artworks: Preprint*. Tilburg University.  
[https://research.tilburguniversity.edu/files/112613757/De\\_Rooij\\_2024\\_preprint\\_1.2\\_.pdf](https://research.tilburguniversity.edu/files/112613757/De_Rooij_2024_preprint_1.2_.pdf)
- Grba, D. (2022). Deep else: A critical framework for AI art. *Digital*, 2(1), 1–32.  
<https://doi.org/10.3390/digital2010001>
- Park, J., Koh, Y., Kim, H., Choi, S.-H., & Park, J. (2021). *Failure of attribute binding in novel image compositions*. OpenReview.  
<https://openreview.net/pdf?id=bKBhQhPeKaF>
- Putra, E. S. (2025). The role of AI in contemporary art: Can AI work be considered “original” art? *Cita Kara: Jurnal Penciptaan dan Pengkajian Seni Murni*, 5(1), 119–145.  
<https://doi.org/10.59997/ctkr.v5i1.5422>
- SIGGRAPH Digital Arts Community. (2025, March). *AI and artistic autonomy: Between tool and constraint*. <https://dac.siggraph.org/sparks/2025-03-ai-and-artistic-autonomy>
- Sivertsen, C., Salimbeni, G. S. G., Løvlie, A. S., Benford, S. B., & Zhu, J. (2024). Machine learning processes as sources of ambiguity: Insights from AI art. *arXiv*.  
<https://arxiv.org/abs/2403.09374>
- University of Washington News. (2023, November 29). *AI image generator Stable Diffusion perpetuates racial and gendered stereotypes, study finds*.  
<https://www.washington.edu/news/2023/11/29/ai-image-generator-stable-diffusion-perpetuates-racial-and-gendered-stereotypes-bias/>
- Zarei, N., Clinton, M. S., Cheng, S., Hua, T., & Forsyth, D. S. (2024). Text-to-image models often fail to bind attributes correctly. *arXiv*. <https://arxiv.org/abs/2406.07844>

**Contact email:** [wyc151354133@gmail.com](mailto:wyc151354133@gmail.com)