

Visual Practices and Algorithmic Imagination

Massimo Balestrini

massimo.balestrini@accademiabellartiverona.it

This paper articulates a theoretical and methodological framework for understanding the visual apparatus of ViaggiAccademici as a locus of negotiation between human cognition and algorithmic generativity. Rather than treating generative AI as a merely operative tool, the study conceptualises prompting as a form of visual dramaturgy through which the theatrical text is reconfigured into a sequence of iconogenic units. The integration of LLM-mediated descriptions, hand-drawn studies and diffusion-based image synthesis reveals how artistic intentionality is modulated, refracted and at times destabilised by the latent structures governing machine-learning models. By incorporating a customised painterly dataset, the research examines stylistic drift, the epistemic status of the glitch and the emergence of visual indeterminacy as constitutive aesthetic conditions rather than technical contingencies. The resulting corpus of images and videos operates not as representational support but as a critical dispositif that expands the dramaturgical field, interrogating the boundaries between authorship, perception and algorithmic inference. The paper ultimately argues that contemporary artistic practice in the context of generative AI demands a shift toward a paradigm of co-emergent imagination, in which human and computational agencies converge to produce novel regimes of visual sense-making.

Keywords: Algorithmic Aesthetics, Visual Epistemologies, Latent Image Production, Co-Emergent Creativity.

Pratiche visive e immaginazione algoritmica

Visual Practices and Algorithmic Imagination

(Engl. trans. below)

Massimo Balestrini
massimo.balestrini@accademiabelleartiverona.it

GPTHEATRE – Generative AI for Humanities: a Theatre Experiment è un progetto di ricerca dell'Università degli Studi di Milano che esplora il dialogo tra arti visive, teatro e intelligenza artificiale generativa.

Nato come esperimento transdisciplinare e presentato nel maggio 2025 al convegno internazionale *Artificial Intelligence to the Test. Creativity and Humanistic Knowledge Today*, il progetto indaga le possibilità di una creatività condivisa tra umano e macchina. In questo contesto, il mio ruolo è stato quello di curare la componente visiva dello spettacolo *ViaggiAccademici*, elaborando un metodo capace di intrecciare la ricerca artistica con la scrittura di prompt destinati ai modelli linguistici di grandi dimensioni (LLM) per la generazione automatica di immagini e video. L'obiettivo non era semplicemente illustrare il testo teatrale, ma interrogare la natura stessa dell'immagine nell'epoca dell'intelligenza artificiale: un territorio liminale in cui visione e pensiero si co-generano in un dialogo dinamico tra intelligenze eterogenee, quella umana, intuitiva e sensibile, e quella artificiale, algoritmica e probabilistica.

1. Dal testo alla scena visiva

Il punto di partenza è stato il copione, concepito come struttura aperta, suscettibile di traduzione visiva. Ho scomposto il testo in micro-sequenze narrative, ciascuna delle quali poteva essere trasformata in unità iconica. Ogni frammento è stato poi elaborato in dialogo con ChatGPT, che ha funzionato non

come strumento esecutivo, ma come interlocutore semiotico, capace di generare descrizioni evocative, ambienti e situazioni visive attraverso la formulazione di *prompt*. L'atto di prompting si è configurato come una nuova forma di scrittura scenica: una drammaturgia descrittiva che traduce l'immaginario verbale in potenzialità visive.

Accanto al linguaggio, ho impiegato schizzi a matita e composizioni preliminari, per dare corpo al ritmo visivo del racconto. Da questo intreccio fra testo, disegno e algoritmo sono nate immagini in cui la parola è diventata forme, campiture, colori e atmosfere: un processo di traduzione sinestetica in cui l'AI diventa estensore sensibile del pensiero visivo.



Figure 1 — Preliminary Sketch

Preparatory sketch establishing the visual structure and atmosphere of the scene, used as a conceptual basis for the subsequent algorithmic image generation.

2. Addestramento e deriva stilistica

I prompt generati sono stati elaborati in Midjourney e in modelli Stable Diffusion e Flux, gestiti tramite interfaccia a blocchi (principalmente ComfyUI), ai quali ho applicato un training personalizzato basato su opere tratte dal mio archivio pittorico personale. Questo innesto estetico non mirava a ottenere una semplice riconoscibilità stilistica, ma

a esplorare l'ibridazione tra la mia gestualità pittorica e la sintassi iconica della macchina. L'esito, ancora tecnicamente imperfetto, tende a costruire un linguaggio visivo instabile e poroso, in cui il segno pittorico si dissolve nella texture digitale e la forma sembra generarsi in tempo reale da un flusso cognitivo condiviso.



Figure 2 — Generated Image and Compositing

Image produced through generative models and later refined using compositing techniques to construct the final scenographic version.

3. Selezione e montaggio dell'immagine

Ogni prompt ha generato decine di varianti: scarti, errori, rivelazioni inattese. Nel processo di selezione non ho cercato la bellezza né la coerenza formale, ma l'attrito tra caos e leggibilità, quella soglia in cui l'immagine sembra smarrirsi e, proprio per questo, rivelare qualcosa di essenziale e insolito. Mi sono lasciato guidare da ciò che potrei definire “non comune”: allucinazioni visive, errori, artefatti, immagini distorte o glitchate, assenza di prospettiva, colori non armonici.

Sono queste le qualità che cerco, fragilità visive che disobbediscono alle regole della composizione e restituiscono all'immagine la sua capacità di sorprendere.

Attraverso Photoshop e tecniche di *compositing*, *photobashing* e *Generative Fill*, ho costruito immagini stratificate, in bilico tra pittura e simulacro, tra documento e sogno digitale.

Il risultato è una serie di scenografie che non rappresentano, ma evocano; non illustrano il testo, ma ne amplificano il potenziale immaginativo. Lo spazio scenico diventa così un laboratorio cognitivo, in cui le immagini non accompagnano la parola, ma la riscrivono.

4. Dalla staticità al tempo: immagini in movimento

Dalle immagini statiche sono poi nate sequenze video generate tramite piattaforme come Sora, Runway, Luma, Hailuo, Kling, Pika, Veo e Wan. In alcuni casi ho fissato il primo e l'ultimo fotogramma per controllare la transizione; in altri, ho utilizzato l'ultimo frame di una sequenza come origine della successiva, generando una continuità fluida.

Ne è scaturita una narrazione visiva priva di montaggio tradizionale, dove il tempo si espande e si piega come materia plastica: un sogno algoritmico che accompagna le parole del copione, sospeso tra flusso e contemplazione.



Figure 3 — Frame Sequence for Video Generation

Series of frames used as visual transitions for the creation of the video sequence.

5. Estetica dell'errore e spazio latente

Ritengo che l'uso dell'intelligenza artificiale per replicare immagini già note, stili e estetiche già viste sia un esercizio sterile. Il suo valore risiede altrove: nella possibilità di dischiudere nuove forme dell'immaginazione visiva e di condurre l'artista verso territori percettivi inesplorati.

Le AI generative operano all'interno di uno spazio latente, una sorta di archivio collettivo popolato da immagini archetipiche, spesso esiti di stereotipi visivi, mode effimere o linguaggi derivativi. Attraversare ed esplorare questo spazio significa confrontarsi con l'inconscio visivo della cultura digitale, con ciò che la macchina custodisce e rielabora del nostro sguardo.

Il mio lavoro tenta di oltrepassare questa soglia, di muoversi verso ciò che definisco “non comune” o, meglio, non banale: l’allucinazione, l’errore, la dissonanza cromatica, la mancata prospettiva diventano luoghi di transizione, interstizi tra gli agglomerati dell’immaginario condiviso.

In questi scarti e in queste irregolarità si apre uno spazio poetico dell’imprevisto, dove la macchina, invece di produrre perfezione, si fa generatrice di ambiguità, di senso e di stupore, un luogo in cui il difetto si fa linguaggio e l’errore si trasforma in visione.

6. Conclusione. Verso una co-creazione riflessiva

Il corpus visivo e audiovisivo di GPTHEATRE non costituisce un apparato illustrativo, bensì un dispositivo critico, un campo di tensione tra il visibile e il pensabile. In questo senso, l’AI non sostituisce l’artista, ma ne prolunga la sensibilità, traducendo l’intuizione in una forma dialogica e riflessiva. Le scenografie finali di *ViaggiAccademici* sono il risultato di questa negoziazione tra due forme d’intelligenza: l’una organica, affettiva, disordinata; l’altra computazionale, probabilistica, impersonale.

Sul palcoscenico, le immagini diventano spazi mentali condivisi, superfici di proiezione dove umano e artificiale si riconoscono l’uno nell’altro. La scena, così, non rappresenta più, ma pensa: pensa attraverso le immagini, attraverso l’errore, attraverso l’algoritmo.

Visual Practices and Algorithmic Imagination

GPTHEATRE – Generative AI for Humanities: a Theatre Experiment is a research project by the University of Milan that explores the dialogue between visual arts, theatre, and generative artificial intelligence. Born as a transdisciplinary experiment and presented in May 2025 at the international conference Artificial Intelligence to the Test. Creativity and Humanistic Knowledge Today, the project investigates the possibility of a shared creativity between humans and machines.

Within this framework, my role was to curate the visual component of the theatre performance ViaggiAccademici, developing a method that could weave together artistic research and the writing of prompts for large language models (LLMs) used to automatically generate images and videos.

The aim was not simply to illustrate the script, but to question the very nature of the image in the age of artificial intelligence: a threshold space where vision and thought co-generate each other in a dynamic dialogue between heterogeneous intelligences — human, intuitive and sensitive, and artificial, algorithmic and probabilistic.

1. From the Script to the Visual Scene

The starting point was the script, conceived as an open structure ready to be translated visually. I broke the text down into micro-narrative sequences, each of which could become an iconic unit. Each fragment was then developed in dialogue with ChatGPT, which worked not as an execution tool but as a semiotic partner capable of generating evocative descriptions, environments, and visual situations through prompt formulation. Prompting became a new form of stage writing: a descriptive dramaturgy that turns verbal imagination into visual possibilities.

Alongside language, I used pencil sketches and preliminary compositions to give shape to the visual rhythm of the story. From the interplay between text, drawing, and algorithm emerged images in which words turned into forms, colours, textures, and atmospheres — a synesthetic translation in which AI becomes a sensitive extension of visual thinking.



Figure 1 — Preliminary Sketch

Preparatory sketch establishing the visual structure and atmosphere of the scene, used as a conceptual basis for the subsequent algorithmic image generation.

2. Training and Stylistic Drift

The prompts were processed in Midjourney and in Stable Diffusion and Flux models, managed through a node-based interface (mainly ComfyUI). I applied a custom training based on artworks from my own painting archive. This aesthetic graft was not meant to achieve simple stylistic recognizability, but to explore the hybridisation between my painterly gesture and the machine's visual syntax.

The result — still technically imperfect — tends toward an unstable and porous visual language, where the painterly mark dissolves into digital texture and the form seems to emerge in real time from a shared cognitive flow.



Figure 2 — Generated Image and Compositing

Image produced through generative models and later refined using compositing techniques to construct the final scenographic version.

3. Image Selection and Assembly

Each prompt produced dozens of variations: discarded attempts, mistakes, unexpected insights. In selecting the images, I did not look for beauty or formal coherence, but for the friction between chaos and readability—the threshold where the image appears to lose itself and, for that very reason, reveals something essential and unusual. I was guided by what I would call the “uncommon”: visual hallucinations, errors, artifacts, distorted or glitched images, lack of perspective, non-harmonic colours.

These are the qualities I seek — visual fragilities that break the rules of composition and give the image back its power to surprise.

Through Photoshop and techniques such as compositing, photobashing, and Generative Fill, I built layered images that sit between painting and simulacrum, between document and digital dream. The result is a series of scenographies that do not represent but evoke; they do not illustrate the text, but amplify its imaginative potential. The stage becomes a cognitive laboratory where images do not accompany words, but rewrite them.

4. From Stillness to Time: Moving Images

From the static images I then created video sequences generated through platforms such as Sora, Runway, Luma, Hailuo, Kling, Pika, Veo, and Wan. In some cases, I fixed the first and last frame to control the transition; in others, I used the last frame of one sequence as the starting point of the next, creating a fluid continuity.

The result is a visual narrative without traditional editing, where time expands and bends like a plastic material: an algorithmic dream accompanying the script's words, suspended between flow and contemplation.



Figure 3 — Frame Sequence for Video Generation

Series of frames used as visual transitions for the creation of the video sequence.

5. The Aesthetics of Error and the Latent Space

I believe that using AI to replicate already-known images, styles, and aesthetics is a sterile exercise. Its value lies elsewhere: in its ability to open new forms of visual imagination and guide the artist toward unexplored perceptual territories.

Generative AIs operate within a latent space — a kind of collective archive populated by archetypal images, often shaped by visual stereotypes, fleeting trends, or derivative languages. Exploring this space means confronting the visual unconscious of digital culture, what the machine stores and reshapes from our gaze.

My work attempts to go beyond this threshold, moving toward what I call the “uncommon” or, more precisely, the non-banal: hallucination, error, chromatic dissonance, lack of perspective become transition zones, interstices among the clusters of shared imagery.

In these deviations and irregularities, a poetic space of the unexpected opens up: the machine, instead of producing perfection, becomes a generator of ambiguity, meaning, and wonder—a place where defect becomes language and error becomes vision.

6. Conclusion: Toward a Reflective Co-Creation

The visual and audiovisual corpus of GPTHEATRE is not an illustrative apparatus but a critical device, a field of tension between the visible and the thinkable. In this sense, AI does not replace the artist; it extends their sensitivity, translating intuition into a dialogic and reflective form. The final scenographies of ViaggiAccademici are the result of this negotiation between two forms of intelligence: one organic, affective, and disorderly; the other computational, probabilistic, and impersonal.

On stage, the images become shared mental spaces, projection surfaces where the human and the artificial recognise one another. The scene, therefore, no longer represents — it thinks: it thinks through images, through error, through the algorithm.