

Back to Back with a Copy: A Computational Analysis of AI-Generated Visual Contemporary Art *Pastiches*

Anca Dinu, Andreiana Mihail, Andra-Maria Florescu, Claudiu Creangă, Liviu Dinu

University of Bucharest, Bucharest, Romania

anca.dinu@11s.unibuc.ro, andreiana@yahoo.com, andra-maria.florescu@s.unibuc.ro, claudiudancreanga@gmail.com, liviu.p.dinu@gmail.com

Abstract

The aim of this paper is twofold. First, it investigates whether newer generative models are getting “better” at *pastiche* contemporary artworks. Second, it explores the consistency of the multidimensional nature of stylistic evaluation across different LLMs. Building on previous work, we analyze stylistic similarity between AI generated *pastiches* and the original artworks of twelve contemporary artists. We used five complementary computer vision models to capture texture, color, semantics, composition, and perceptual features through cosine distance in high-dimensional embedding spaces. The distances obtained show that the newer image generation model that we used has produced *pastiches* with improved semantic alignment and greater diversity than the model used in previous work. However, it was slightly less performant on shallow features such as color, texture, and perceptual adherence. Our findings confirm that artistic style is inherently multidimensional, and measuring it does not depend on any spatial architecture. These quantitative findings are contextualized through feedback from human evaluators, which are the artists themselves.

Introduction

Over the past decade, Computational Creativity (CC) has experienced rapid growth, largely driven by advances in AI, in particular transformer models, which greatly increased the capacity of computer systems to create original artifacts.

In this work, we focus on a type of visual computer creativity inherently combinatorial in nature, defined in (Boden 2004) as the process of combining ideas from a given conceptual space in novel ways, which minds both convergent creativity (the capacity to identify the best or correct solution to a problem) and divergent creativity (the capacity to generate a wide range of varied original responses) (Guilford 1950), in an assisted creativity framework, defined in (Colton and Wiggins 2012) as the automatic process of supporting a human creator. In particular, we asked the Gemini Pro 3 Nano Banana model (Gemini Team, Google 2023) to generate *pastiches* after contemporary art, via prompting, by uploading the images of the original artworks.

The starting point was the work of (Dinu et al. 2026), who employed Chat GPT to create *pastiches* after the artworks of 12 contemporary artists. The study uncovered that the multi-faceted concept of style can be expressed by features like texture and color, semantics and concept, artistic

touches, visual composition, or perceptions, and that these features can be captured independently using different computer vision models which extract high-dimensional embeddings. A large cosine distance between the AI generated *pastiche* and the original human artwork in these embedding spaces means that the *pastiche* was original, but satisfies less the task requirement of creating a *pastiche*, while a low distance indicates a copy that is less original.

We kept the same framework and data, but we used the newer SOTA image generation model Gemini. We were motivated by the following research questions:

1. Are newer generative models getting “better” at producing *pastiches* after human artworks?
2. How do human and automatic evaluation correlate in this respect?
3. Are visual models consistent in extracting stylistic features across different generative AI models?

Our study directly compares the outputs of the two models and gains insight into the stability of the claims made in (Dinu et al. 2026). A vital part of our assessment was the contribution of the artists themselves through their evaluation and comparison.

Related Work

Recently, the fast growing body of work in CC spread to include a wide variety of topics, from improvement techniques and evaluation to ethical concerns and the role of humans in automatic image generation.

For example, (Morain and Ventura 2025) showed that the use of more advanced prompt engineering techniques such as Optimization as PROMPTING (OPRO) or chain-of-thought (COT) leads to no significant improvements in textual computational creative systems, while (Colloca and Rezwana 2025) analyzed how ChatGPT amplifies positive sentiment and adds descriptive language regardless of the user’s prompt style, showing no correlation between user prompt and AI-generated creative output.

The creative abilities of three artificial image-to-image generative models are formally quantified in (Ramaswamy, Chockler, and Navaratnarajah 2025). The authors propose three objective creativity criteria: requirement satisfaction, cohesion, and novelty and introduce formal measures for

Table 1: Average Cosine Distances: Gemini vs GPT. Lower values indicate higher similarity.

Model	Gemini	GPT	$\Delta\%$	Feature Type
AdaIN-Style	0.064	0.063	+1.6%	texture, color
CLIP-ViT-L	0.167	0.197	-15.2%	semantic
ResNet50-Style	0.441	0.454	-2.9%	artistic style
DINOv2	0.461	0.463	-0.4%	f.g. visual
VGG19	0.699	0.674	+3.7%	perceptual

Table 2: Model Discrimination and Consistency: Gemini vs. GPT.

Model	Variance		Consistency (Avg Diff)	
	Gemini	GPT	Gemini	GPT
DINOv2	0.0425	0.0349	0.146	0.131
VGG19	0.0082	0.0274	0.087	0.111
ResNet50-Style	0.0160	0.0249	0.089	0.115
CLIP-ViT-L	0.0072	0.0103	0.048	0.046
AdaIN-Style	0.0028	0.0014	0.029	0.034

each of them. The complexity and subjectivity of the evaluation process of CC systems, which depends mainly on the expertise and perception of the evaluator, are recognized in (Boden 1996; Boden 2004; Wiggins 2006). Moreover, (Demke and Ventura 2024; Morain and Ventura 2025) point out that the domain is extremely diverse, making objective human evaluation difficult, time consuming, and costly.

Finally, (Kaila, Holzapfel, and Jääskeläinen 2024) documented how artists negotiate authorship and ownership when experimenting with artificially generated imagery, revealing how they re-frame artistic roles and authorship through generative tools.

Data

This study uses the dataset proposed in (Dinu et al. 2026), which consists of 108 images:

- 36 original artworks by twelve contemporary artists, who generously agreed to lend 3 of their artworks each, for research purposes. The artists are: Adi Matei, Ciprian Mureșan, Ion Grigorescu, Iulia Uță, Karine Fauchard, Lazar Lyutakov, Marius Tănăsescu, Mathias Poeschl, Oana Năstăsache, Philip Patkowitsch, Răzvan Botiș, and Tom Chamberlain;
- 72 AI *pastiches* produced by AI (ChatGPT 5 model), 2 imitations/*pastiches* for each original.

We enriched this image dataset with a new series of AI generated works (e.g., Figures 3, 4, 5, 6 from the Appendix) using the latest Gemini Pro 3 Nano Banana model (Gemini Team, Google 2023). We used the same prompt as in the ChatGPT experiments conducted by (Dinu et al. 2026), so the results would be directly comparable.

Table 3: Extreme Divergence Cases: DINOv2 vs AdaIN. Divergence = $DINOv2_{norm} - AdaIN_{norm}$. Positive values indicate high texture adherence with high compositional divergence.

Artwork	DINOv2	AdaIN	Divergence
Ciprian Mureșan 1	100%	20%	+0.798
Adi Matei 2	99%	27%	+0.726
Ion Grigorescu 3	81%	11%	+0.697
Tom Chamberlain 3	50%	97%	-0.470
Karine Fauchard 3	15%	42%	-0.270
Oana Năstăsache 2	5%	27%	-0.213

Experimental Setup

In the pre-reprocessing step, we converted all images to RGB format and normalized them according to each of the five models used in (Dinu et al. 2026) to extract high-dimensional embeddings and which have been shown to capture different aspects of artistic style, as follows:

- AdaIN-Style (Huang and Belongie 2017) captures 1920-dimensions of texture/color statistics;
- ResNet50-Style (He et al. 2016) captures 2048-dimensions of artistic style features;
- CLIP-ViT-L (Radford et al. 2021) captures 768-dimensions of semantic features;
- DINOv2 (Oquab et al. 2023) captures 1024-dimensions of fine-grained visual features;
- VGG19 (Simonyan and Zisserman 2014) captures 4096-dimensions of perceptual features.

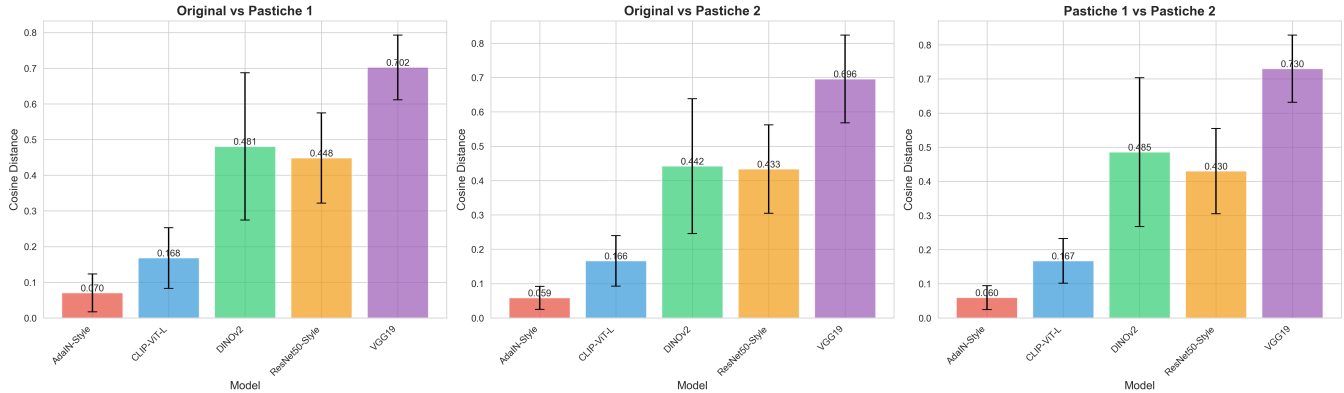
To approximate the stylistic similarity corresponding to each of the five models, for each artwork group (the original and its two Gemini-generated *pastiches*), we compute three pairwise cosine distances: original to *pastiche* 1, original to *pastiche* 2, and *pastiche* 1 to *pastiche* 2.

Automatic evaluation

The distribution of the cosine distance between the original artworks and their Gemini-generated *pastiches* using the five distinct feature embedding models is given in Table 1, in comparison to the GPT results reported in (Dinu et al. 2026). Figure 1 illustrates these distances the Gemini generated *pastiches*, for a more intuitive visual representation.

Comparing ChatGPT and Gemini reveals a consistent model distance hierarchy: AdaIN-Style remains the lowest (0.064), followed by CLIP-ViT-L (0.167), ResNet50-Style (0.441), DINOv2 (0.461), and VGG19 (0.699). This confirms that the “compositional gap” between texture matching and perceptual similarity is model-agnostic and inherent to the task itself. However, the models prioritize different elements (Table 1). Gemini favors deep semantics, better capturing concepts (15.2% closer) and style (2.9%), making it more adept at the overall *pasticheing* task based on cosine distance. In contrast, ChatGPT focuses on surface forms, slightly better imitating perceptual features (3.7% closer)

Figure 1: Average cosine distances between original artworks and Gemini-generated *pastiches* across all five embedding models. Left panel shows Original→*Pastiche 1* distances, middle shows Original→*Pastiche 2*, right shows *Pastiche 1*↔*Pastiche 2* diversity.



and texture/color (1.6%). Furthermore, Gemini generates significantly more diverse *pastiches*, as shown in Figure 10 from the Appendix. The average distance between *pastiches* of the same original is 10.1% higher for Gemini than GPT (0.374 vs. 0.340). This diversity is particularly notable in DINOv2 (+22.4%) and VGG19 (+10.3%), indicating that Gemini explores broader compositional and perceptual interpretations.

We also compared the discrimination (variance) and consistency of the two models, given in Table 2. Discrimination reflects the embedding model’s capacity to distinctly separate different artistic styles (measured by variance), while consistency indicates the stability of the generative model when producing multiple *pastiches* for the same original artwork (measured by the average distance between them).

DINOv2 remains the most discriminative model for both Gemini (0.0425) and GPT (0.0349), while AdaIN-Style remains the least discriminative. Notably, Gemini shows improved consistency on VGG19 and ResNet50-Style compared to GPT, suggesting more stable perceptual and stylistic reproduction.

The models capture complementary information, as evidenced by the weak correlation ($r = 0.293$) between DINOv2 (by composition) and AdaIN-Style (texture). This quantifies the “compositional gap”, showing that texture fidelity does not predict compositional accuracy. Table 3 highlights some extreme cases in which these metrics disagree. For example, Ciprian Muresan’s Work 1 shows the largest positive divergence (+0.798), pairing high texture adherence (AdaIN: 20th percentile) with high compositional divergence (DINOv2: 100th percentile). Conversely, Tom Chamberlain’s Work 3 displays the largest negative divergence (-0.470), combining moderate compositional adherence (50th percentile) with high texture divergence (97th percentile).

Artists’ evaluation

We asked artists to evaluate Gemini’s *pastiches* after their artwork, answering the same three questions about the *pas-*

tiches produced by ChatGPT in (Dinu et al. 2026), for a direct comparison:

- Q1. To what extent do you recognize your personal artistic language and the coherence of your visual style in this new work? (1 = not at all, 10 = completely)
- Q2. How does this new work inspire you or what thoughts does it provoke? (open answer)
- Q3. To what extent do you consider that the work generated by Gemini has aesthetic or artistic value? (1 = not at all, 10 = very high)

Similarly with ChatGPT’s *pastiches*, the grades for Gemini’s *pastiches* varied considerably by artist, as shown in Figure 2. Compared to ChatGPT, Gemini’s *pastiches* were perceived by the artists as closer to their style, with a mean grade for the first question of 4.99, representing a significant increase of 1.41 points compared to the mean grade obtained by ChatGPT. However, they were perceived as slightly less artistically valuable, with a mean grade for the third question of 4.26, which is 0.67 points lower than the one obtained by ChatGPT’s *pastiches*. Overall, the artists seem to perceive Gemini’s *pastiches* as closer to their works, in line with the computational evaluation, which constitutes a weak positive answer to our second RQ.

Answering the open-questions, the artists offered a variety of opinions about the *pastiches* of their artworks. Most of them regarded the images created by AI more like “artistic comments” than real creations, with little understanding of originality, color mixing techniques, body, and depth. Artistic comments, in this instance, refers to image that functions more as visual interpretation or reactions to the original artworks rather than independent creative works.

We reproduce here some of the answers. Oana Năstăsache says that “some colors look dirty(...) I would never associate that ochre-green combination, because it feels cold and muddy” or that “The surfaces are placed without any sense of composition”. Ion Grigorescu remarks that the new *pastiches* followed the concept and composition, but failed in

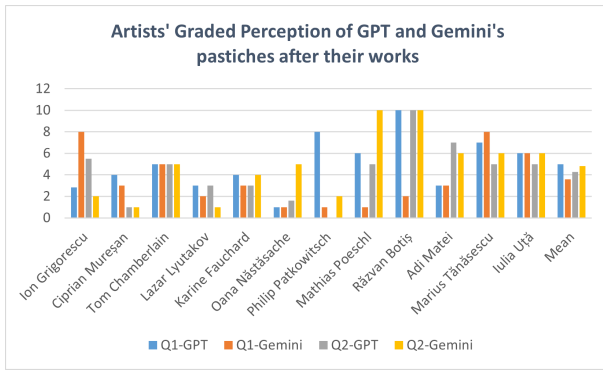


Figure 2: Comparison between artists’ graded perception of GPT and Gemini’s *pastiche*s after their works.

spirit, technique, and style: “Young man rising from a chair. Egon Schiele — closer to forgery than to *pastiche*. A theatricality cheaper than Schiele’s”. Philip Patkowitzsch has a different point of view raising the issue of intellectual property: “they actually present my works better, but just because it is using parts of the original image, so I would say copyright infringement would be an issue!”.

So, in the end, did Gemini produce artworks as perceived by the artists themselves? Most of them responded negatively, motivating that AI displayed insufficient command of color theory, which resulted in a poor chromatic atmosphere and the lack of full conceptual comprehension/integration.

Interpretation of the results

The results confirm that artistic style should be evaluated multidimensionally (Dinu et al. 2026). Both GPT and Gemini successfully match texture statistics (low AdaIN distances) but show substantial compositional divergence (high DINOv2 and VGG19 distances), showing that these dimensions are orthogonal: high texture adherence often pairs with high compositional divergence, and vice versa, as illustrated in Figures 7, 8, and 12 from the Appendix. Thus, this “compositional gap” is an intrinsic challenge for AI *pastiche* generation, as visually shown in Figure 11 from the Appendix, confirming that visual models consistently extract stylistic features across different generative AI models, positively responding to our third RQ.

The models also exhibit distinct generation strategies, pointing to a weak positive answer to our first RQ:

- **Gemini**: An “interpretive” approach prioritizing semantic alignment (15% closer to CLIP) and higher diversity (+10%).
- **GPT**: A “faithful reproduction” approach prioritizing texture matching and consistency.

Because of the weak correlation ($r = 0.293$) between texture and composition, evaluating AI art requires a multi-metric framework:

- **AdaIN-Style**: Texture/color fidelity.
- **CLIP-ViT-L**: Semantic/conceptual alignment.
- **DINOv2**: Compositional structure.

Empirical Remarks

We observed that Gemini often disregards prompt guidelines, defaulting to the styles of iconic masterpieces (e.g., Egon Schiele or Grant Wood’s *American Gothic*). Although this exploratory approach allows for originality, the model often overcompensates, producing works that compete with canonical art rather than complementing the original. Both models exhibit perfectionism, frequently completing works that artists intentionally left unfinished, while significantly altering the original image’s foreground.

While Gemini “understands” dimensionality and brushwork, it consistently fails to perceive depth, context, and conceptual irony. A prime example is the *pastiche* of Ciprian Mureșan’s *Pioneers*. Mureșan depicted communist youth as rebels and villains to mock 1980s Romanian propaganda. However, both Gemini and ChatGPT missed this irony, independently generating similar images of pioneers as standard patriotic drummers, shown in Figure 9. The striking similarity between the drummers generated by the two models suggests that they were trained on the same visual data.

Finally, Gemini performs better when creating collages and installations compared to paintings. This combinatorial technique aligns with LLM capabilities, though the results often resemble over-accumulation rather than originality.

Limitations and Future Work

Our study is constrained by a small conceptually homogeneous dataset, as most artists belong to the same gallery. Including detailed figurative artwork could provide better insight into LLM *pastiche* performance. Furthermore, comparing outputs across fundamentally different mediums (e.g., 2D paintings versus 3D conceptual installations) presents structural methodological challenges. Finally, photography was excluded because it functions as a mechanical reproduction of reality rather than an artwork.

We plan to adopt alternative metrics beyond the cosine distance that are better suited for evaluating visual similarity. To mitigate personal artist bias, we aim to incorporate formal evaluations by professional art critics. Lastly, we will apply this methodology to iconic, frequently *pasted* artworks, such as Da Vinci’s “Mona Lisa,” Van Gogh’s “Starry Night,” Grant Wood’s “American Gothic,” and Andy Warhol’s “Campbell’s Soup Cans”.

Ethical Implications

This publication raises no ethical concerns, as all software licenses and artist permissions were strictly respected. However, broader generative AI (GenAI) art introduces significant socio-ecological challenges (Marella, Erukude, and Veluru 2025). Societally, GenAI disrupts creative labor markets and raises issues regarding data exploitation, authorship, and intellectual property rights (Frosio 2023), often reinforcing existing inequalities¹. These disruptive impacts on cultural heritage and sustainability require a commitment to responsible inclusive cultural development (Roe 2025).

¹<https://www.weforum.org/stories/2025/01/the-impact-of-genai-on-the-creative-industries/>

References

- [Boden 1996] Boden, M. A. 1996. Chapter 9 - creativity. In Boden, M. A., ed., *Artificial Intelligence*, Handbook of Perception and Cognition. San Diego: Academic Press. 267–291.
- [Boden 2004] Boden, M. 2004. *The Creative Mind: Myths and Mechanisms*. Routledge.
- [Colloca and Rezwana 2025] Colloca, G., and Rezwana, J. 2025. Prompting ai in co-creation: The role of syntax and sentiment in shaping ai-generated content. In *Proceedings of the 16th International Conference on Computational Creativity*. International Conference on Computational Creativity.
- [Colton and Wiggins 2012] Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In *Proceedings of the 20th European Conference on Artificial Intelligence (ECAI 2012)*, 21–26.
- [Demke and Ventura 2024] Demke, J., and Ventura, D. 2024. Overcoming algorithmic bias as a measure of computational creativity. In *Proceedings of the 15th International Conference on Computational Creativity (ICCC'24)*. Jönköping, Sweden: Association for Computational Creativity.
- [Dinu et al. 2026] Dinu, A.; Mihail, A.; Florescu, A.-M.; and Creangă, C. 2026. The art that poses back: Assessing ai pastiches after contemporary artworks. In Machado, P.; Romero, J. J.; and Rebelo, S. M., eds., *Artificial Intelligence in Music, Sound, Art and Design*, 67–83. Cham: Springer Nature Switzerland.
- [Frosio 2023] Frosio, G. 2023. Generative ai in court. In Koutras, N., and Selvadurai, N., eds., *Recreating Creativity, Reinventing Inventiveness: International Perspectives on AI and IP Governance*. Routledge. Forthcoming chapter; available at SSRN: <https://ssrn.com/abstract=4558865>.
- [Gemini Team, Google 2023] Gemini Team, Google. 2023. Gemini: A family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*. Updated to include Gemini 3, November 2025.
- [Guilford 1950] Guilford, J. 1950. Creativity. *American Psychologist*.
- [He et al. 2016] He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- [Huang and Belongie 2017] Huang, X., and Belongie, S. 2017. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*.
- [Kaila, Holzapfel, and Jääskeläinen 2024] Kaila, A.-K.; Holzapfel, A.; and Jääskeläinen, P. 2024. Gardening frictions in creative ai: Emerging art practices and their design implications. In *International Conference on Innovative Computing and Cloud Computing*.
- [Marella, Erukude, and Veluru 2025] Marella, V. C.; Erukude, S. T.; and Veluru, S. R. 2025. The impact of artificial intelligence on traditional art forms: A disruption or enhancement. *Indian Journal of Computer Science and Engineering* 16(3):54–66.
- [Morain and Ventura 2025] Morain, R., and Ventura, D. 2025. Prompt engineering the creativity knob for large language models? In *Proceedings of the 16th International Conference on Computational Creativity*. International Conference on Computational Creativity.
- [Oquab et al. 2023] Oquab, M.; Darcet, T.; Moutakanni, T.; Vo, H.; Szafraniec, M.; Gkioxari, G.; and Bojanowski, P. 2023. Dinov2: Learning robust visual features without supervision. *arXiv preprint arXiv:2304.07193*.
- [Radford et al. 2021] Radford, A.; Kim, J. W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; Krueger, G.; and Sutskever, I. 2021. Learning transferable visual models from natural language supervision. In *Proceedings of the 38th International Conference on Machine Learning (ICML)*, volume 139 of *Proceedings of Machine Learning Research*, 8748–8763. PMLR.
- [Ramaswamy, Chockler, and Navaratnarajah 2025] Ramaswamy, A.; Chockler, H.; and Navaratnarajah, M. 2025. Defining and quantifying creative behavior in popular image generators. In Oliveira, H. G.; Spendlove, B.; Gervás, P.; and Ventura, D., eds., *Proceedings of the 16th International Conference on Computational Creativity*. Association for Computational Creativity.
- [Roe 2025] Roe, J. 2025. Generative ai as cultural artifact: Applying anthropological methods to ai literacy. *Postdigital Science and Education* 7(4):1107–1124.
- [Simonyan and Zisserman 2014] Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [Wiggins 2006] Wiggins, G. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.

Acknowledgments

Research supported by a grant of the Ministry of Research, Innovation and Digitization, CNCS - UEFIS-CDI, project SIROLA, number PN-IV-P1-PCE-2023-1701, within PNCDI IV, and by UB project.

We are grateful to all the artists who agreed to let us use their works and for their insightful feedback: Adi Matei, Ciprian Mureșan, Ion Grigorescu, Iulia Uță, Karine Fauchard, Lazar Lyutakov, Marius Tănăsescu, Mathias Poeschl, Oana Năstăsache, Philip Patkowitsch, Răzvan Botis, and Tom Chamberlain.

Appendix



Figure 3: Examples of Gemini's *pastiches* after Ion Grigorescu's work 1 "Măriuca Iosifescu", oil on canvas and plaster, 1974.

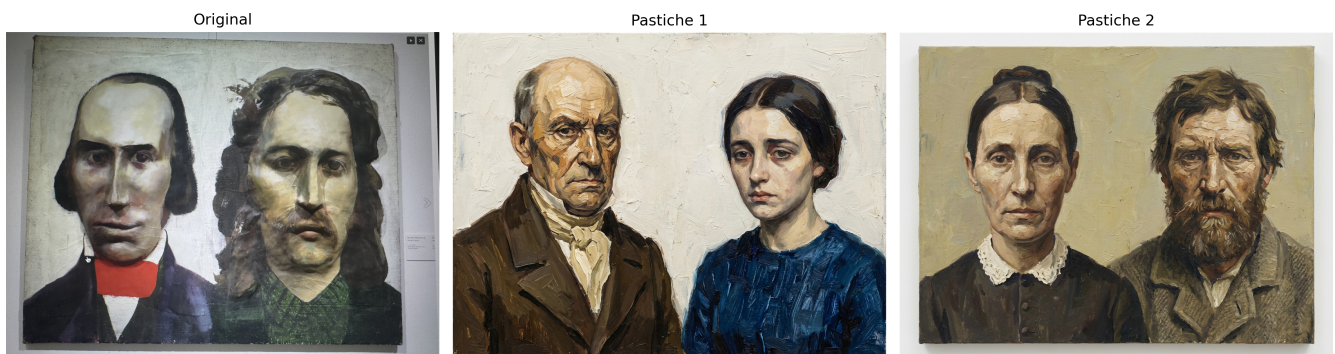


Figure 4: Examples of Gemini's *pastiches* after Ion Grigorescu's work 2 "Avram Iancu și Nicolae Bălcescu", oil on canvas and plaster, 1974.



Figure 5: Examples of Gemini's *pastiches* after Iulia Uța's work 2 Untitled, oil on canvas, 2024.



Figure 6: Examples of Gemini's *pastiches* after Lazar Lyutakov's work "24/7 the human condition", installation, 2014.

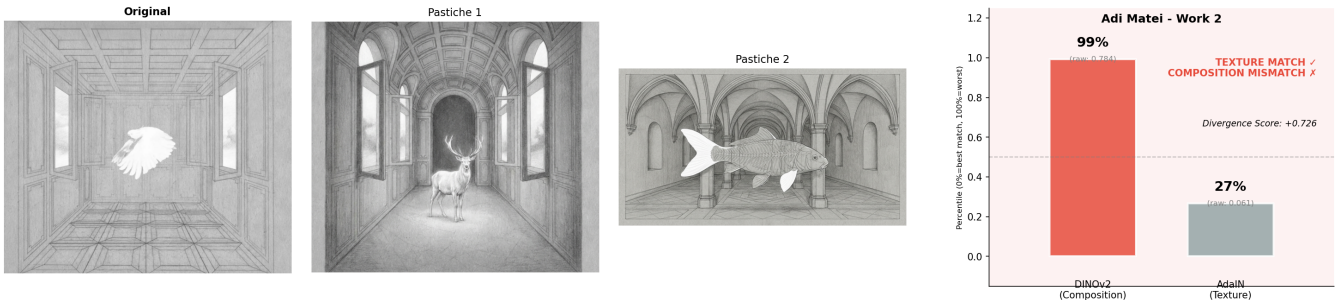


Figure 7: High texture adherence (low AdaIN), high compositional divergence (high DINOv2). These cases illustrate the core of the "compositional gap" and the orthogonality of style dimensions.

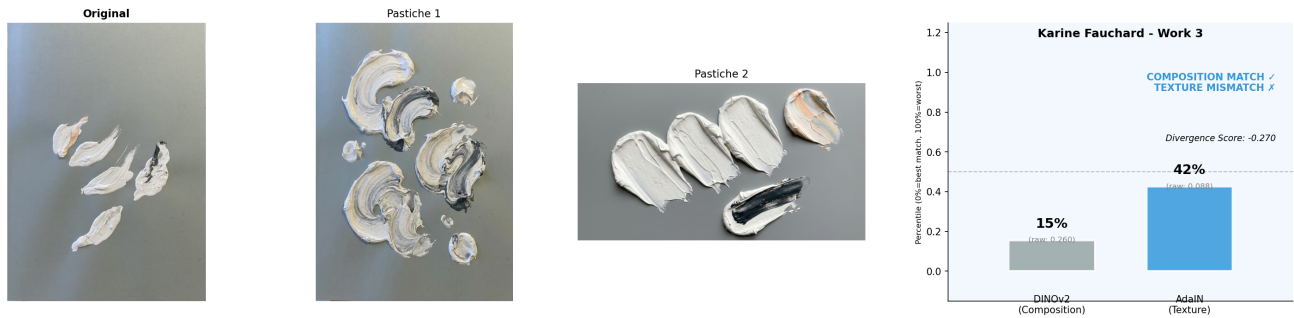


Figure 8: High compositional adherence (low DINOv2), high texture divergence (high AdaIN). Together with Figure 7, this confirms texture and composition as orthogonal dimensions.



Figure 9: GPT (middle) and Gemini (right) created similar *pastiches* of Mureșan's work (left)- suggest similar training data.

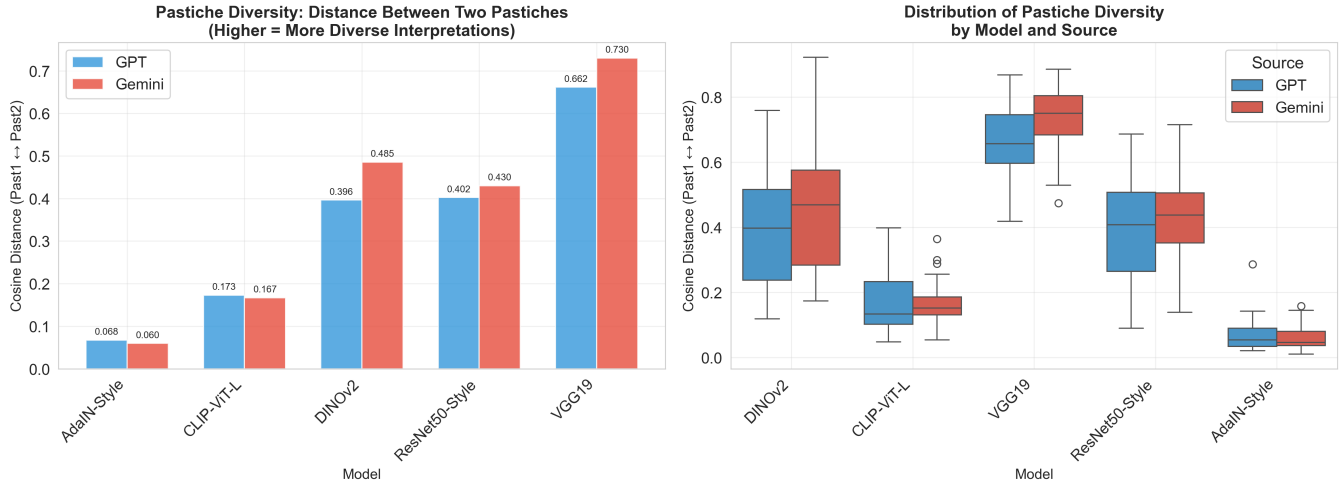


Figure 10: High *pastiche* diversity (left) indicate the model explores more varied interpretations rather than producing near-identical outputs. The distribution boxplots (right) shows that Gemini produces overall 10.1% more diverse *pastiche*s, with the largest differences in DINOv2 (+22.4%) and VGG19 (+10.3%), suggesting greater compositional and perceptual variety.

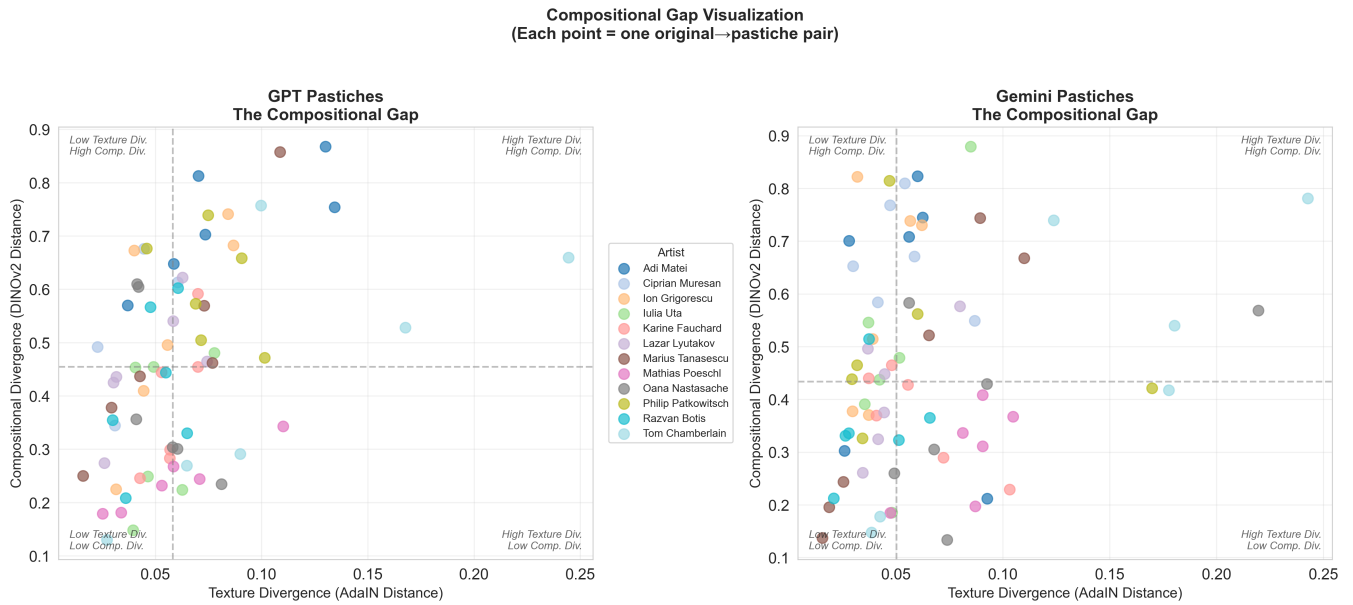


Figure 11: Texture divergence (AdaIN) vs. compositional divergence (DINOv2) for GPT (left) and Gemini (right) *pastiche*s. Each point represents one original→*pastiche* pair, colored by artist. Dashed lines indicate median values, dividing the space into four quadrants. Both models exhibit the same characteristic pattern: points cluster toward the left (low texture divergence, typically 0.03–0.10) but spread widely along the y-axis (compositional divergence ranging from 0.2 to 0.9). This asymmetric distribution—narrow horizontally, broad vertically—is the visual signature of the compositional gap: AI pastiches consistently adhere to texture/color statistics while showing high variability in compositional structure. The upper-left quadrant (“Low Texture Div. / High Comp. Div.”) is densely populated for both models, representing the dominant generation pattern. Artist-specific clustering is evident: Adi Matei (dark blue) and Ciprian Muresan (light blue) concentrate in high compositional divergence regions, while Oana Năstăsache (gray) and Mathias Poeschl (magenta) cluster toward lower compositional divergence. The similar distributions across GPT and Gemini confirm the compositional gap is model-agnostic.

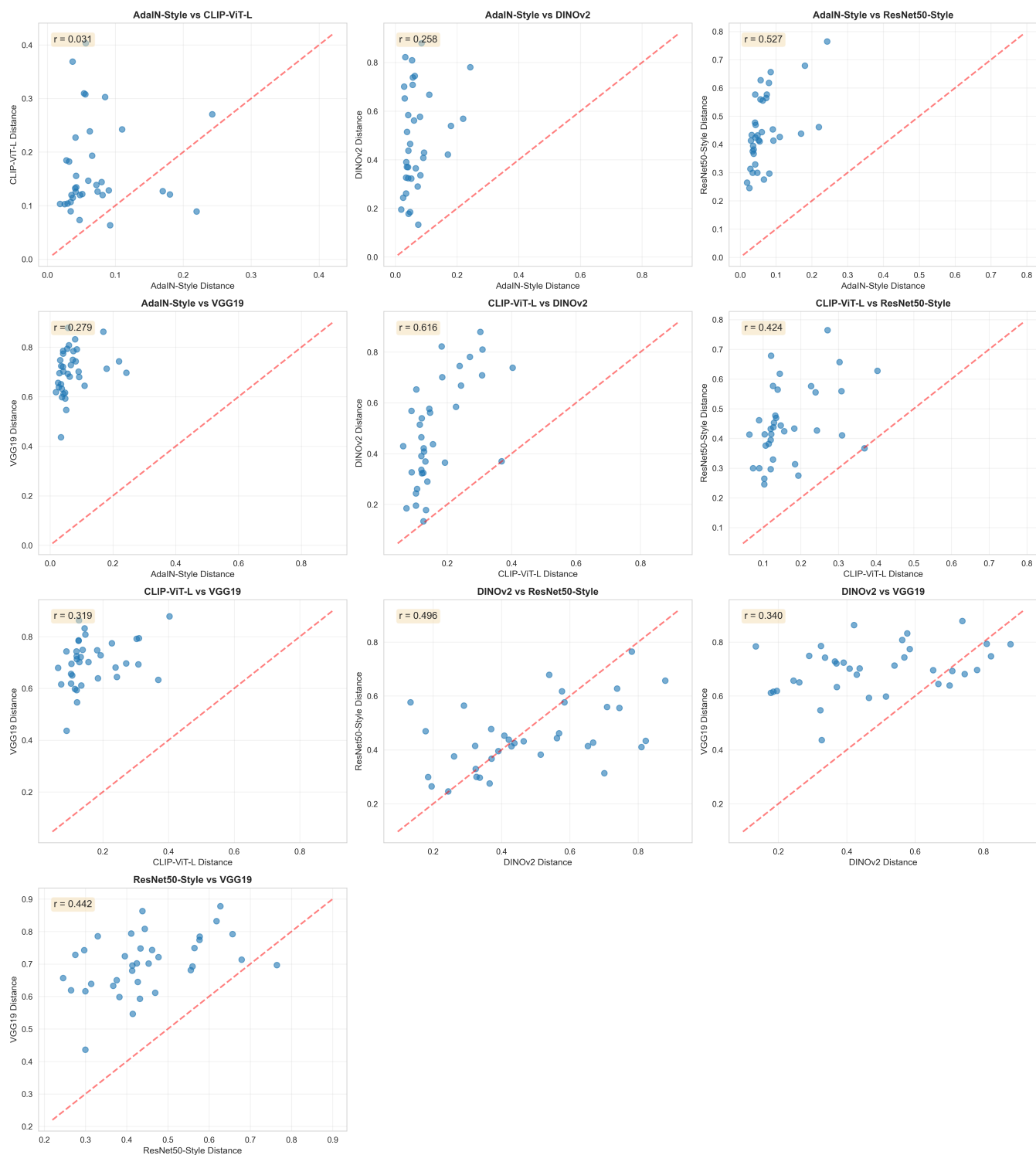


Figure 12: Pairwise correlations between embedding model distances. Each scatter plot shows the cosine distances from one model (x-axis) versus another (y-axis) across all 36 artist-*pastiche* pairs, with the Pearson correlation coefficient (r) displayed. The red dashed line indicates the identity diagonal. Key findings: (1) AdaIN-Style and CLIP-ViT-L show near-zero correlation ($r = 0.031$), confirming that texture statistics and semantic content are orthogonal dimensions; (2) The highest correlation is between CLIP-ViT-L and DINOv2 ($r = 0.616$), reflecting their shared emphasis on higher-level visual features; (3) AdaIN-Style correlates most strongly with ResNet50-Style ($r = 0.527$), as both capture style-related properties; (4) VGG19 shows weak-to-moderate correlations with all other models ($r = 0.28$ – 0.44), suggesting it captures a distinct perceptual dimension. The generally weak correlations (mean $r \approx 0.37$) confirm that the five models capture complementary rather than redundant information, justifying the multi-metric evaluation framework.