

Composite Reflections

Abhishek Dangeti, Pavan Gajula, and Vikram Jamwal
Computational Creativity Group, TCS Research
INDIA

{abhishek.dangeti, pavanbhargav.gajula, vikram.jamwal}@tcs.com

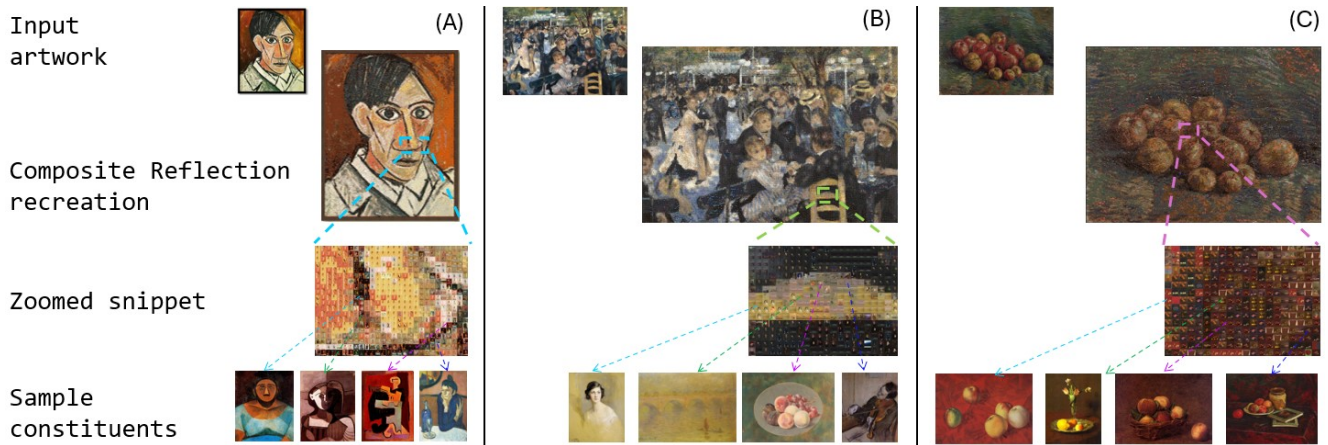


Figure 1. Composite Reflections demonstrating thematically coherent mosaic generation. (A) Pablo Picasso’s self-portrait reconstructed using 485 constituent works by the artist himself, creating an intimate artistic dialogue. (B) Pierre-Auguste Renoir’s *Bal du moulin de la Galette* recreated from 1,910 Impressionist artworks, preserving the movement’s stylistic coherence. (C) Van Gogh’s *Apples* reconstructed using 306 apple-themed artworks, maintaining both visual and semantic consistency between the subject matter and constituent pieces.

Abstract

We introduce *Composite Reflections*, a novel approach to mosaic creation where composite artworks maintain intimate thematic and visual relationships with their constituent pieces. Unlike traditional mosaics that use source images as decorative elements, our method ensures meaningful dialogue between individual artworks and their collective representation. We present an end-to-end pipeline employing color-based retrieval and structural similarity analysis to reconstruct target artworks using semantically coherent constituent collections. Our approach transcends conventional tile-based methods by preserving both the artistic integrity of individual components and the narrative coherence of the overall composition. This work opens new avenues for computational creativity, enabling artists to explore expressive relationships within their image collections through algorithmically-guided mosaic generation.

1. Introduction

Mosaic art represents a profound intersection of individual expression and collective narrative, where artists assemble disparate elements into cohesive visual statements. From the ancient tessellated floors of Pompeii to Salvador Dalí’s groundbreaking 1976 photomosaic of Abraham Lincoln, this medium has evolved to challenge traditional boundaries between composition and deconstruction, unity and fragmentation. Contemporary artists like Charis Tsevis have elevated photomosaics beyond mere technical exercises, creating works where constituent images maintain thematic coherence with their composite whole—transforming the medium from algorithmic puzzle-solving into meaningful artistic discourse.

In computer vision, photomosaic generation has been extensively studied as an optimization problem focused primarily on visual fidelity [6, 11]. Existing approaches typically segment input images and replace each region with the most visually similar candidate from a pre-defined dataset,

emphasizing computational efficiency and pixel-level accuracy. However, these methods often sacrifice artistic integrity by arbitrarily stretching, cropping, or decontextualizing constituent images, reducing the mosaic to a mere technical demonstration rather than a coherent artistic statement. Furthermore, they neglect the semantic relationships between the target image and constituent elements, missing the conceptual depth that distinguishes artistic mosaics from algorithmic reconstructions.

We introduce *Composite Reflection*, a novel approach that prioritizes thematic coherence alongside visual fidelity in photomosaic generation. Our method leverages Large Vision-Language Models (LVLMs) to curate semantically relevant image collections based on extracted themes from the target artwork. Rather than relying on arbitrary image databases, we construct purpose-built datasets that maintain conceptual dialogue with the input image’s subject matter, artistic style, or cultural context.

Our approach operates through a multi-stage pipeline that balances artistic intent with technical precision. First, we employ LVLMs to analyze the input artwork and curate a thematically coherent dataset of constituent images. Second, we perform adaptive image segmentation that respects both compositional structure and semantic boundaries. Finally, we match segments to constituent images using a dual-criteria framework that weighs color histogram similarity against structural coherence measured through the Structural Similarity Index (SSIM), ensuring both perceptual accuracy and preservation of individual artwork integrity.

The key contributions of this work are:

1. **Thematic coherence framework:** We establish a methodology for creating semantically meaningful photomosaics where constituent images maintain conceptual relationships with the target artwork, moving beyond purely visual similarity metrics.
2. **LVLM-driven dataset curation:** We demonstrate how Large Vision-Language Models can be effectively employed to automatically generate thematically relevant image collections, eliminating the need for manual dataset preparation.
3. **Dual-criteria matching with integrity preservation:** We propose a novel matching algorithm that optimizes both color fidelity and structural similarity while preserving the aspect ratios and compositional integrity of constituent artworks, avoiding common distortions in existing tile-based approaches.
4. **End-to-end automated pipeline:** We present a complete system that transforms artistic concepts into executable processes, requiring only an input artwork, target theme, and source collection to generate contextually rich photomosaics.

Through extensive evaluation on diverse artistic inputs

and comparison with state-of-the-art methods, we demonstrate that *Composite Reflection* produces photomosaics that achieve superior thematic coherence while maintaining competitive visual quality, opening new possibilities for computationally-assisted artistic creation.

2. Related Work

The concept of photomosaics has its roots in traditional art, with Salvador Dalí pioneering photo mosaics in the 1970s. Contemporary digital artists like Charis Tsevis continue this tradition, creating compelling mosaic artworks through sophisticated manipulation of constituent image collections. However, computational approaches to mosaic generation remain relatively underexplored in computer vision, despite drawing from well-established techniques in image retrieval and segmentation. Our work builds upon foundations in color-based and structure-based image retrieval to create semantically coherent digital mosaics.

2.1. Image Retrieval

Content-based image retrieval forms the backbone of computational mosaic generation. Early approaches focused primarily on color-based retrieval using features such as color histograms [19], color moments [14], and color textures [7]. While these methods effectively capture color distribution, they lack structural information crucial for creating visually coherent mosaics. We adopt color histograms as our primary retrieval mechanism due to their effectiveness in preserving color fidelity—a critical requirement for maintaining visual consistency across mosaic segments.

More sophisticated approaches combine color and structural information for improved retrieval accuracy. Recent work [25] employs cumulative histograms for color similarity and Hu invariant moments for shape similarity, enabling retrieval based on both chromatic and geometric properties. Similarly, Kumar et al. [10] integrate edge-histogram features for shape, Tamura texture features [20] for texture, and quantized color histograms for comprehensive content-based retrieval.

Our approach extends these methods by incorporating Structural Similarity Index Measure (SSIM) as a secondary filtering criterion. SSIM’s focus on luminance, contrast, and structural coherence makes it particularly suitable for mosaic applications, where preserving subtle details like shadows and shading is essential for maintaining the visual integrity of the reconstructed image.

Structure-based retrieval methods, while less directly applicable to our work, provide relevant insights into shape matching. Sketch-based retrieval approaches [16] utilize learned key shapes and contour detection for structure matching, while recent deep learning methods [17] leverage CLIP [15] with data augmentation to learn fine-grained shape representations.

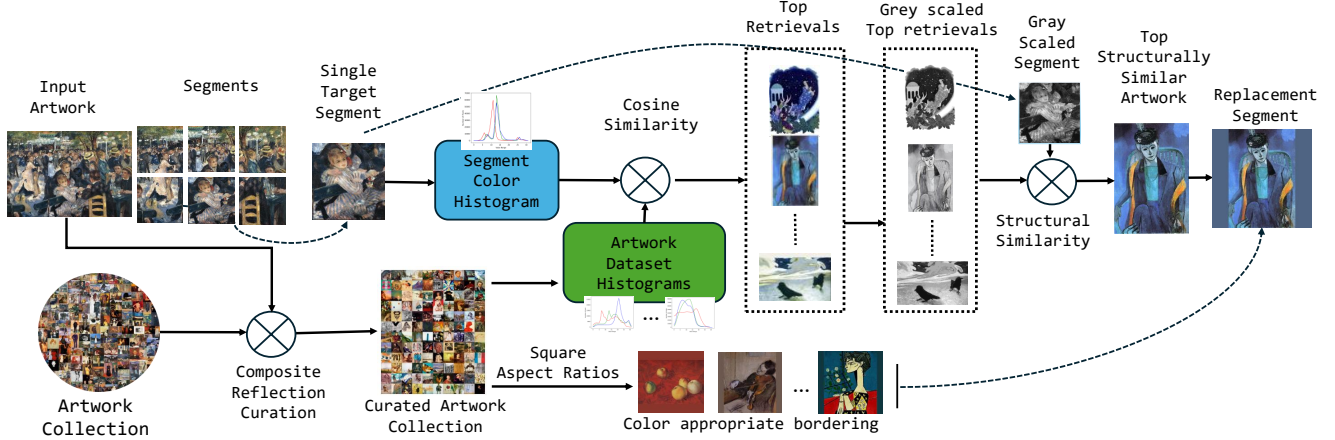


Figure 2. Overall architecture of the Composite Reflections pipeline. The curated artwork collection is extracted from a larger artwork collection and the artworks are encoded into color histograms. The input is segmented into smaller segments which are then encoded into color histograms. For each segment, the segment color histogram is matched with the curated collection histograms using cosine similarity to obtain the first set of k candidate artworks based on color similarity. The segment and the candidate artworks are then converted into grayscale to calculate structural similarity between the segment and the artworks using SSIM. The artwork with the highest SSIM score is used to replace the segment while applying color-appropriate bordering to preserve its aspect ratio. This process is repeated for each segment until the final *Composite Reflection* is obtained.

2.2. Digital Mosaics

Computational mosaic generation has evolved from manual, user-guided approaches to fully automated systems. Early work by Hausner [8] required user-defined edges for creating tiled mosaics from input artworks. Di Blasi and Gallo [5] advanced this by introducing automatic segmentation using statistical region merging [13]. In contrast, our approach employs fixed-size segmentation to preserve the aspect ratio and orientation of constituent artworks, maintaining their artistic integrity within the composite.

Photo mosaic generation has been extensively studied, with Blasi et al. [6] proposing automatic segmentation into equal-sized regions and replacement based on average RGB values. Subsequent work [11] incorporated color correction to improve visual coherence. However, we deliberately avoid color correction to preserve the authentic appearance of constituent artworks, instead relying on improved retrieval mechanisms for coherence.

Our approach differs fundamentally from existing methods by utilizing color histograms rather than simple RGB averages, providing richer color distribution information for more accurate matching. The integration of SSIM for structural similarity further distinguishes our work from purely color-based approaches. While some recent methods [24] employ edge-aware algorithms for large segment processing, direct comparison is limited by code availability constraints.

The concept of Composite Reflections extends beyond traditional mosaic generation by emphasizing the thematic relationship between constituent and composite artworks,

positioning our contribution at the intersection of computational creativity and content-based image analysis.

3. Composite Reflection Pipeline

Our Composite Reflection framework generates thematically coherent photomosaics through a five-stage pipeline that preserves both visual fidelity and semantic meaning. Given an input artwork and a target theme, our method automatically curates a relevant dataset, processes constituent images to maintain aspect ratio integrity, segments the input image, and performs dual-criteria matching to construct the final mosaic. Figure 2 illustrates the complete pipeline architecture.

3.1. Thematic Dataset Curation

The foundation of Composite Reflection photomosaics lies in establishing semantic coherence between the target artwork and constituent images. We support two primary thematic categories: *artistic themes* (art movements, specific artists, stylistic periods) and *semantic themes* (objects, entities, conceptual motifs). Our curation process adapts to the availability of metadata within existing artwork collections.

For collections with comprehensive metadata (e.g., curated museum databases), we directly filter artworks based on predefined categorical labels. However, many large-scale collections lack detailed thematic annotations. To address this limitation, we employ Large Vision-Language Models (LVLMs) for automated content analysis and annotation.

Specifically, we utilize InternVL2 [2] to generate the-

matic annotations for the WikiArt collection. For each artwork, we query the LVLM with targeted prompts such as “Describe the food items present in this artwork” or “Identify the artistic movement and style of this painting.” The resulting annotations enable fine-grained thematic filtering—for instance, extracting all artworks containing “apple” references to construct an apple-themed dataset.

This automated curation approach ensures that constituent images maintain conceptual relationships with the target artwork while scaling to large collections without manual intervention.

3.2. Constituent Image Processing

To preserve artistic integrity while enabling flexible mosaic composition, we apply two preprocessing steps to each constituent artwork in our curated dataset.

3.2.1. Aspect Ratio Preservation via Adaptive Bordering

Traditional mosaic algorithms often distort constituent images through stretching or cropping to fit predetermined segment dimensions. To maintain the original aspect ratios, we transform each artwork into a designated aspect ratio format (e.g., a square format) by adding contextually appropriate color borders.

Border colors are determined by extracting the dominant color from each artwork histogram. This approach ensures that added borders blend harmoniously with the original artwork rather than introducing jarring visual discontinuities. The bordered artworks can then be uniformly scaled to match segment dimensions without geometric distortion.

3.2.2. Color Distribution Encoding

We pre-compute color histograms for all constituent artworks to enable efficient retrieval during mosaic generation. Unlike average color representations, which reduce complex color relationships to single values, histograms capture the full distributional characteristics of an artwork’s color palette.

Our implementation uses OpenCV’s `calcHist` function with 32 bins per channel in RGB space, creating 32,768-dimensional feature vectors that encode color frequency distributions. These histograms are indexed using FAISS [9] for rapid similarity search during the matching phase. The choice of 32 bins balances discriminative power with computational efficiency - fewer bins lose important color distinctions, while more bins introduce noise from quantization artifacts.

3.3. Input Image Segmentation

We segment the input artwork into a regular grid of equally-sized rectangular regions. While adaptive segmentation based on semantic boundaries could potentially improve results, we find that uniform segmentation provides several

advantages: computational simplicity, predictable output resolution, and consistent constituent image scaling.

For each segment, we extract the same 32-bin RGB color histogram used for constituent image encoding. This creates a direct basis for similarity comparison in the subsequent matching phase.

3.4. Dual-Criteria Segment Matching

Our matching algorithm balances color fidelity with structural coherence through a two-stage retrieval process that progressively narrows the candidate pool based on complementary similarity metrics.

3.4.1. Color-Based Candidate Retrieval

For each image segment, we retrieve the top- k most similar constituent artworks based on cosine similarity between their color histograms. Cosine similarity is particularly suitable for histogram comparison as it is invariant to scaling differences and focuses on distributional shape rather than absolute magnitude.

The choice of k candidates represents a carefully tuned balance. Fewer candidates overly constrain structural matching options, while more candidates introduce colorimetrically poor matches that compromise visual coherence.

3.4.2. Structure-Aware Final Selection

Among the top- k color-matched candidates, we select the final replacement based on structural similarity using the Structural Similarity Index Measure (SSIM) [21]. Both the image segment and candidate artworks are converted to grayscale to isolate structural information from color characteristics.

SSIM quantifies perceptual similarity by comparing luminance (μ), contrast (σ), and structural correlation (σ_{xy}) between images:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1)$$

preserving structural relationships through SSIM-based selection, creating artworks that function both as faithful reproductions at a distance and as thematically meaningful collections when viewed closely.

The complete pipeline operates automatically given only an input artwork, target theme, and source collection, enabling scalable generation of contextually rich photomosaics with minimal manual intervention.

4. Experimental Setup and Evaluation

We evaluate Composite Reflection across diverse thematic contexts to demonstrate its versatility and effectiveness in generating semantically coherent photomosaics. Our experimental framework encompasses three distinct thematic categories, comprehensive evaluation metrics, and systematic analysis of key design choices.

4.1. Dataset Construction and Thematic Categories

We construct nine specialized datasets spanning three thematic categories to evaluate our method’s ability to maintain semantic coherence across different artistic and conceptual domains.

4.1.1. Artist-Specific Collections

Artist-based datasets enable evaluation of stylistic coherence within individual artistic oeuvres. We curate collections for five influential artists representing diverse artistic movements and techniques:

- **Vincent van Gogh:** 1,749 artworks from WikiArt [22], spanning his complete career from early works to his distinctive post-impressionist masterpieces.
- **Jan Brueghel the Elder:** 1,587 artworks from the specialized Brueghel dataset [18], focusing on his detailed landscape and genre paintings.
- **Pablo Picasso:** 743 artworks from WikiArt covering his blue period through cubist innovations.
- **Edvard Munch:** 5,461 artworks from the MUNCH museum foto-web archive [12], providing comprehensive coverage of his expressionist works.
- **Salvador Dalí:** 600 artworks from WikiArt representing his surrealist period and technical precision.

For artist-specific evaluations, we use self-portraits or portraits of each artist as input images, testing our method’s ability to recreate iconic representations using the artist’s own stylistic vocabulary.

4.1.2. Art Movement Collections

Movement-based datasets evaluate our method’s capacity to maintain historical and stylistic coherence within broader artistic contexts:

- **Impressionism:** 13,059 artworks capturing the movement’s emphasis on light, color, and atmospheric effects.

- **Post-Impressionism:** 5,669 artworks representing the diverse reactions to Impressionism’s limitations.
- **Early Renaissance:** 1,391 artworks showcasing the period’s technical innovations and humanistic themes.

Input images for movement-based experiments are selected from representative works within each movement, ensuring thematic consistency between target and constituent images.

4.1.3. Semantic Theme Collections

Object-based datasets demonstrate our method’s effectiveness in conceptual coherence beyond purely artistic categories. Using InternVL2 [2], we automatically annotate the complete WikiArt collection (79,469 artworks) with detailed content descriptions.

- **Apple-themed collection:** 2,000 artworks containing apple motifs, selected from automated annotations using the keyword “apple.” We limit the collection size to 2,000 as manual verification reveals diminishing annotation quality beyond this threshold.

This approach demonstrates the scalability of our automated curation process and the potential for creating highly specific thematic collections from large, unstructured datasets.

4.2. Evaluation Metrics

We employ complementary evaluation metrics to assess both technical fidelity and perceptual quality of generated mosaics.

4.2.1. Structural Fidelity Assessment

We measure structural preservation using the Structural Similarity Index Measure (SSIM) [21] applied to grayscale versions of input and output images. This approach isolates geometric and textural similarity from color information, providing a pure assessment of structural preservation.

SSIM values range from -1 (complete dissimilarity) to +1 (perfect similarity), with higher values indicating better preservation of the original image’s structural characteristics. By converting to grayscale, we ensure that structural assessment is not confounded by color matching performance.

4.2.2. Perceptual

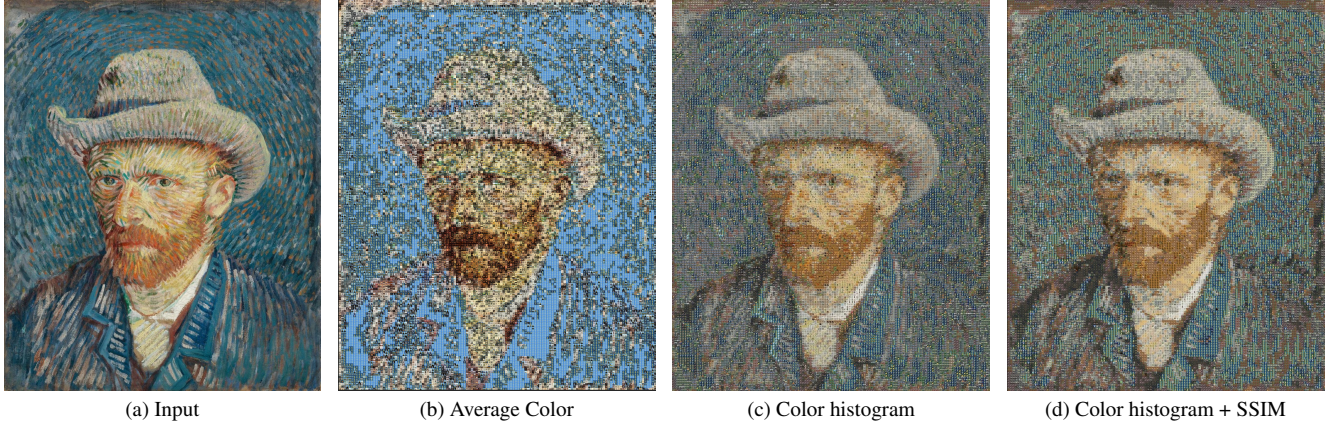


Figure 3. Qualitative comparison of Composite Reflections for Van Gogh’s self-portrait using different matching criteria: average color, color histogram, and color histogram + SSIM. Results shown for 320×260 pixel segments.

4.3. Ablation Studies

We conduct comprehensive ablation studies to validate our key design choices and demonstrate the contribution of each pipeline component.

4.3.1. Color vs. Structure Matching

We compare our dual-criteria approach against single-criterion baselines: color-only matching using histogram similarity, and structure-only matching using SSIM. Results consistently show that our combined approach achieves superior performance on both structural and perceptual metrics, validating the importance of balancing these complementary aspects.

4.3.2. Candidate Pool Size Analysis

We systematically vary the number of color-based candidates ($k = 1, 5, 10, 20, 50$) used for structural matching. Our analysis reveals that for our maximum segment granularity $k = 10$ provides the optimal balance: smaller values overly constrain structural options, while larger values introduce colorimetrically inappropriate matches that degrade overall quality.

4.3.3. Thematic Coherence Impact

We compare mosaics generated using our curated thematic datasets against those created with random image collections of equivalent size. Results demonstrate significant improvements in both quantitative metrics and qualitative assessment when using thematically coherent constituent collections, supporting our central hypothesis about the importance of semantic relationships in artistic mosaic creation.

5. Results

In this section, we showcase and discuss the results obtained through our Composite Reflection architecture. Composite Reflection samples created using different themed datasets

are available at [3] and Section 4 in the Supplementary. A video showcase can be found at [4].

5.1. Comparative Analysis

Figure 3 compares our approach with color histogram-based and color average-based mosaic approaches [1]. The color average-based mosaic fails to retrieve appropriate artworks for individual segments, as single average values cannot capture the complete chromatic complexity of image regions. This fundamental limitation results in retrieved artworks that poorly match the overall color characteristics of their corresponding segments.

The color histogram-based approach achieves better color correspondence but lacks structural coherence. While retrieved artworks exhibit appropriate color distributions, the absence of structural similarity creates mosaics that fail to preserve essential visual features such as shadows and depth information. This structural deficiency causes the final composition to lose critical definitional elements of the original subject matter.

Our SSIM-enhanced approach demonstrates significant improvement by incorporating structural similarity alongside color matching. The structural correspondence between segments and retrieved artworks enables preservation of fine details including shadows, depth gradients, and geometric features. This alignment allows retrieved components to collectively maintain the complete visual definition present in the input artwork.

5.2. Structural Impact Analysis

The importance of structural similarity is further demonstrated in Figure 4 and Table 2, where we vary the number of color retrievals used for structural filtering to modulate SSIM impact. As we increase the pool of color-based candidates for structural selection, both qualitative and quantitative improvements emerge in the resulting Composite

Segment Size Theme	5x5	20x20	100x100
Artist			
Van Gogh	0.2256	0.1509	0.1376
Picasso	0.2038	0.09479	0.0699
Munch	0.4698	0.3014	0.2555
Dali	0.3023	0.2491	0.2054
Brueghel	0.4231	0.2784	0.216
Art Movement			
Impressionism	0.3561	0.1809	0.154
Post Impressionism	0.2437	0.14083	0.1411
Early Renaissance	0.3194	0.1825	0.1965
WikiArt Dataset	0.2897	0.1659	0.1592
Thematic			
Apple	0.2837	0.213	0.2492

Segment Size Theme	5x5	20x20	100x100
Artist			
Van Gogh	0.3238	0.3769	0.4546
Picasso	0.2667	0.3005	0.4258
Munch	0.3018	0.3384	0.4406
Dali	0.2603	0.2801	0.3964
Brueghel	0.3664	0.3983	0.4957
Art Movement			
Impressionism	0.2093	0.3233	0.3773
Post Impressionism	0.2161	0.3258	0.3861
Early Renaissance	0.2516	0.2933	0.3566
WikiArt Dataset	0.1917	0.3307	0.3992
Thematic			
Apple	0.1756	0.3301	0.3755

Table 1. Quantitative evaluation using **SSIM** (left) and **LPIPS** (right) metrics comparing our Composite Reflections with original input images across different segmentation granularities. Higher SSIM scores ($\uparrow$) indicate better structural similarity, while lower LPIPS scores ($\downarrow$) indicate better perceptual similarity between the original and generated mosaic. Segment sizes are specified in pixels (height, width). SSIM is calculated between the complete input image and the Composite Reflection output.

Reflections. This progressive enhancement confirms that structural similarity is crucial for maintaining visual coherence, particularly when dealing with larger segment sizes.

Our analysis reveals that constituent artworks successfully highlight different features through coherent pattern formation that emphasizes specific characteristics of the target artwork. Notably, well-defined Composite Reflections emerge even with limited artwork collections, demonstrating the robustness of our approach.

5.3. Quantitative Validation

Table 1 presents comprehensive SSIM and LPIPS metrics comparing input artworks with their reconstructions across different granularities. These quantitative results validate our approach’s effectiveness in both structural preservation and perceptual similarity. The metrics demonstrate that our reconstructions maintain geometric fidelity while preserving human-perceived visual similarity to the original artworks, confirming the success of our dual color-structure retrieval strategy.

5.4. Limitations

Our approach has a few limitations: **1. Segmentation constraints:** The current implementation is restricted to rectangular

Number of Segments	Metric	Top 5 Color Retrievals (Color Histogram + SSIM)	Top 10 Color Retrievals (Color Histogram + SSIM)	Top 25 Color Retrievals (Color Histogram + SSIM)	Edge Aware Retrieval [23]
12x10	SSIM ($\uparrow$)	0.0974	0.106	0.1179	0.1139
	LPIPS ($\downarrow$)	0.2711	0.2686	0.2649	0.3284
48x40	SSIM ($\uparrow$)	0.1123	0.1182	0.1219	0.1289
	LPIPS ($\downarrow$)	0.2553	0.2505	0.2487	0.2532
144x120	SSIM ($\uparrow$)	0.1231	0.1355	0.1552	0.1623
	LPIPS ($\downarrow$)	0.2323	0.2316	0.2365	0.2481

Table 2. Quantitative results for varying the number of color retrievals utilized for structural retrieval. Emphasizing structure for lower granularity provides better results whereas reducing the emphasis on structure for lower granularities provides better results. SSIM is calculated between the input and the complete Composite Reflection output.

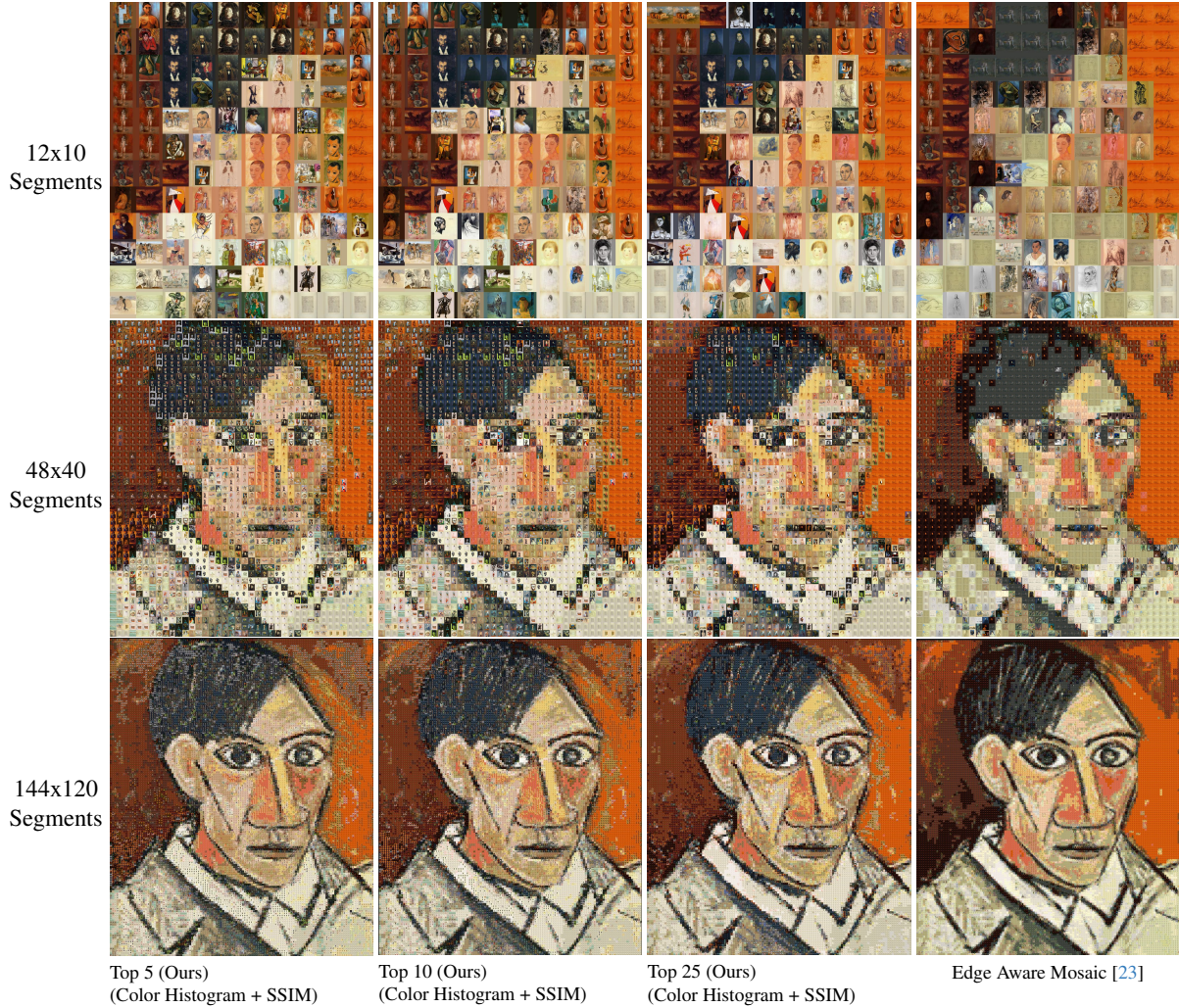


Figure 4. Composite Reflections of *Self-portrait of Picasso* with varying numbers of color histogram-based retrievals and segmentation granularity. The figure showcases qualitative comparisons between Ours and Edge Aware mosaic approaches, where segmentation granularity varies vertically and the number of color-based retrievals varies horizontally.

7. Acknowledgments

We are grateful to TCS Research management for supporting this work. We would also like to thank Partha Sarkar for his valuable guidance and assistance in developing the web application that automates the Composite Reflections pipeline.

References

- [1] Hassan Abedi. Creating mosaic art using ai. <https://github.com/habedi/mosaic-art-maker>, 2025. 6
- [2] Zhe Chen, Jiannan Wu, Wenhai Wang, Weijie Su, Guo Chen, Sen Xing, Muyan Zhong, Qinglong Zhang, Xizhou Zhu, Lewei Lu, et al. Internvl: Scaling up vision foundation models and aligning for generic visual-linguistic tasks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 24185–24198, 2024. 3, 5
- [3] Abhishek Dangeti, Pavan Gajula, and Vikram Jamwal. Composite reflections examples drive link. <https://tinyurl.com/fem38a8p>, 2025. 6
- [4] Abhishek Dangeti, Pavan Gajula, and Vikram Jamwal. Composite reflection video showcase. <https://youtu.be/0ZE-SroXLws>, 2025. 6
- [5] Gianpiero Di Blasi and Giovanni Gallo. Artificial mosaics. *The Visual Computer*, 21:373–383, 2005. 3
- [6] Gianpiero di Blasi and Maria P. Petralia. Fast photomosaic. 2005. 1, 3
- [7] Myron Flickner, Harpreet Sawhney, Wayne Niblack, Jonathan Ashley, Qian Huang, Byron Dom, Monika Gorkani, Jim Hafner, Denis Lee, Dragutin Petkovic, et al. Query by image and video content: The qbic system. *computer*, 28(9):23–32, 1995. 2
- [8] Alejo Hausner. Simulating decorative mosaics. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 573–580, 2001. 3
- [9] Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-scale similarity search with GPUs. *IEEE Transactions on Big Data*, 7(3):535–547, 2019. 4
- [10] Sumit Kumar, Arup Kumar Pal, Naushad Varish, Irfan Nurhidayat, Sayed M. Eldin, and Soubhagya Kumar Sahoo. A hierarchical approach based cbir scheme using shape, texture, and color for accelerating retrieval process. *Journal of King Saud University - Computer and Information Sciences*, 35(7):101609, 2023. 2
- [11] Hae-Yeoun Lee. Automatic photomosaic algorithm through adaptive tiling and block matching. *Multimedia Tools Appl.*, 76(22):24281–24297, 2017. 1, 3
- [12] MUNCH-Museum. Munch Museum Archives- Munch-museet Fotoarkiv, 2021. 5
- [13] Richard Nock and Frank Nielsen. Statistical region merging. *IEEE Transactions on pattern analysis and machine intelligence*, 26(11):1452–1458, 2004. 3
- [14] Alexander P Pentland, Rosalind W Picard, and Stan Sclaroff. Photobook: Tools for content-based manipulation of image databases. In *23rd AIPR Workshop: Image and Information Systems: Applications and Opportunities*, pages 37–50. SPIE, 1995. 2
- [15] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision, 2021. 2
- [16] Jose M Saavedra, Juan Manuel Barrios, and S Orand. Sketch based image retrieval using learned keyshapes (lks). In *BMVC*, page 7, 2015. 2
- [17] Aneeshan Sain, Ayan Kumar Bhunia, Pinaki Nath Chowdhury, Subhadeep Koley, Tao Xiang, and Yi-Zhe Song. Clip for all things zero-shot sketch-based image retrieval, fine-grained or not. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2765–2775, 2023. 2
- [18] Xi Shen, Alexei A Efros, and Mathieu Aubry. Discovering visual patterns in art collections with spatially-consistent feature learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9278–9287, 2019. 5
- [19] Michael J Swain and Dana H Ballard. Color indexing. *International journal of computer vision*, 7(1):11–32, 1991. 2
- [20] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki. Textural features corresponding to visual perception. *IEEE Transactions on Systems, man, and cybernetics*, 8(6):460–473, 1978