Cross-Domain Imagination via Persona-Achievement Matrix Transformation

Pinphong Ruangraweenukit[1],
Virach Sornlertlamvanich[2], Waranrach Viriyavit[1],
Thatsanee Charoenporn, and Ponlawat Chophuk*
[1]*Faculty of Informatics, Burapha University, Thailand.*
[2]*Asia AI Institute (AAII), Faculty of Data Science, Musashino University, Japan.*
ORCID ID: Pinphong Ruangraweenukit https://orcid.org/ 0009-0008-1055-4628
ORCID ID: Virach Sornlertlamvanich https://orcid.org/0000-0002-6918-8713
ORCID ID: Waranrach Viriyavit https://orcid.org/0000-0002-5192-8998
ORCID ID: Thatsanee Charoenporn https://orcid.org/0000-0002-9577-9082
ORCID ID: Ponlawat Chophuk https://orcid.org/ 0000-0002-7284-2675

Abstract. This research formalizes a creator's stylistic identity as a consistent geometric signature in latent space, enabling cross-domain translation. We propose a novel computational framework to extract and translate this identity, specifically from visual arts to culinary arts—by leveraging multimodal embeddings (CLIP) to construct a quantifiable "Persona-Achievement Matrix (PAM)". Using a technique termed Cross-Singular Value Decomposition (Cross-SVD), we decouple the artist's geometric structure from their material medium and bridge the semantic gap via a "Pépin Bridge," a pivot-creator strategy that derives a linear translation matrix. This transformation is injected directly into the cross-attention layers of a generative U-Net, modulating the latent space to synthesize artifacts that possess the source artist's structural composition but the target domain's material properties. Validation confirmed a 39.7% trajectory shift toward food textures while maintaining a 0.37 structural similarity to the original artworks. These results demonstrate that AI can perform grounded, intentional cross-domain imagination through linear manifold alignment, offering a pathway toward more trustworthy and semantically aware generative systems.

Keywords. Style transfer, Computational Creativity, Cross-Domain Creativity, Domain Translation, Representation Learning, Shared Vector Space, Disentangled Representation, CLIP Model

1. Introduction

Contemporary Artificial Intelligence is dominated by Generative AI, a technology capable of producing vast quantities of novel content by recognizing and recombining patterns from immense training datasets. While models like ChatGPT and DALL-E can generate compelling text and images, their reliance on statistical patterns often leads to hallucinations outputs that appear plausible but are factually or logically incoherent. This highlights a critical distinction discussed in recent literature, while these models are proficient at recalling and recombining facts, they often lack true knowledge or a deep, causal understanding of the concepts they process [1]. This lack of genuine

[1] Corresponding Author: Ponlawat Chophuk, Faculty of Informatics, Burapha University, 169 Long Had Bangsaen Road, Saensuk, Muang, Chonburi 20131, Thailand; E-mail: ponlawat.ch@go.buu.ac.th

understanding is a core limitation that reveals they create from what has been, not from what could be.

A Consider Albert Einstein’s departure from classical mechanics. He realized that existing Newtonian theory could not explain the bending of light across the universe. Instead, he synthesized a new framework based on the observation that light’s path is not fixed but is curved by the gravity of massive objects. Crucially, the trajectory of this light depends entirely on the specific object it encounters. This principle mirrors the function of our Imaginative Model. Just as the path of light is shaped by a specific gravitational lens, the translation of a creative style is shaped by the specific 'Bridge' creator we employ. The resulting artifact is not a generic average; it is a unique outcome determined by the specific 'stylistic gravity' of our expert pivot. Unlike standard Generative AI, which merely interpolates past data to find a probable means, our model calculates this specific curvature, allowing for a grounded, intentional innovation that, like Einstein’s theory, it extends beyond the limits of existing patterns.

The central challenge in verifying true AI creativity is that when a model operates within its training domain for instance, an AI creating a new painting after seeing millions of others, it is nearly impossible to prove that the output is not just a clever amalgamation of its training data. A more rigorous test of imaginative capability, therefore, is to challenge a model to translate a creator's identity into a completely new and unfamiliar domain. This provides a clear testbed to distinguish pattern recombination from genuine creative translation.

This leads us to our central thought experiment: Imagine Vincent van Gogh, not in a sun-scorched field in Arles with an easel, but in a bustling kitchen. What would he create? Perhaps a dish defined by turbulent swirls of saffron sauce, with bold, impasto-like textures and an emotional intensity that reflects his inner world. This thought experiment is the philosophical core of our research as illustrated in Figure 1. It posits a fundamental question: does the signature of genius, the unique style of a gifted individual, exist independently of their chosen medium?

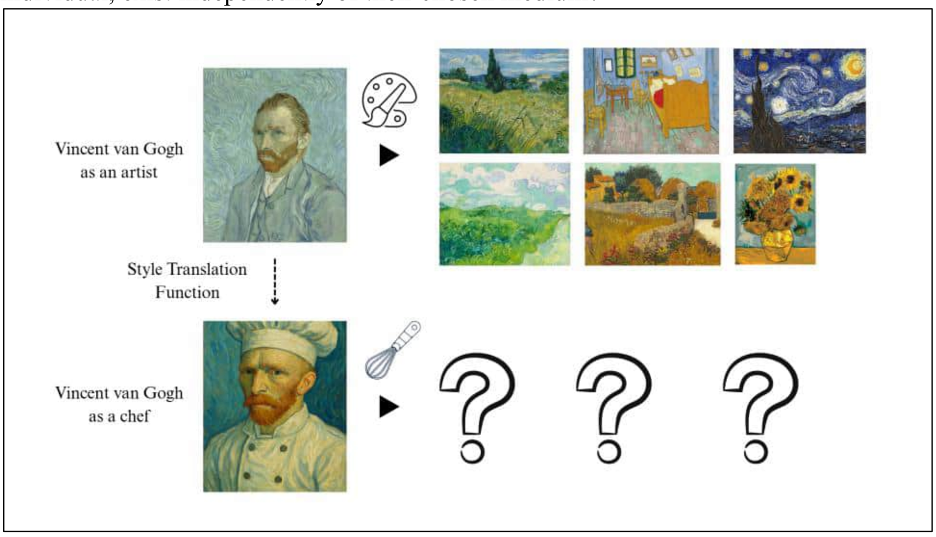

Figure 1. A conceptual illustration of the core research problem.

However, realizing this thought experiment computationally presents a fundamental geometric challenge: the Modality Misalignment Problem. In high-dimensional latent spaces, the data manifold for “Visual Art” and the manifold for “Culinary Concepts” occupy disparate regions with distinct origin points. A vector representing “texture” in an oil painting does not share a coordinate system with “texture” in a soufflé; they are mathematically incongruent (akin to comparing apples and oranges).

Furthermore, relying on a single modality to bridge this gap is insufficient. Text-based approaches suffer from semantic ambiguity (e.g., the word “rich” implies wealth in art narratives but fat content in culinary descriptions), while purely image-based approaches often confuse surface-level color distributions with structural identity (e.g., confusing a red ball with a red apple). To accurately translate a creative signature, one must align these disparate spaces without losing the unique identity of the creator.

To solve this geometric misalignment, we introduce Cross-SVD[2] Strategy. Mathematically, the vector space for Visual Art and the vector space for Culinary Concepts rely on different coordinate systems; their origin points do not match. A vector direction that signifies “boldness” in the artistic domain does not necessarily point to “boldness” in the culinary domain. To map one to the other, we need a reference point that exists validly in both spaces to serve as an anchor.

This is where Jacques Pépin serves as our “Rosetta Stone.” Our central hypothesis is that a single creator possesses a consistent internal stylistic origin, a creative soul that remains constant regardless of the medium. Therefore, the mathematical relationship between Pépin’s Art Vectors and his Food Vectors is not random; it represents the fundamental Transformation Matrix required to bridge the two disparate manifolds. By calculating the precise vector rotation needed to align Pépin’s art with his cooking, we derive universal logic for translation. We then apply this learned logic to Vincent van Gogh, effectively “rotating” his artistic identity into the culinary space.

This methodology is designed to specifically address the limitations of current AI, such as Popularity Bias [3] and systemic generative biases [4]. Instead of relying on a model's generalized and biased associations (e.g., assuming “French Art” just means “croissants”), we ground our analysis in the specific, curated stylistic logic of our bridge subject. This approach aligns with state-of-the-art research aiming to enhance base models with external, specialized knowledge to enable more robust, Semantically Aware Reasoning (SAR) [1]. In our work, this specialized knowledge is the derived Transformation Matrix, which forces the AI to follow a proven stylistic path rather than guessing.

The central problem this research addresses is the lack of a framework for learning a direct translation function for a creator's identity across disparate domains. Our methodology unfolds in three stages. First, we extract multimodal vectors and consolidate them into a Persona-Achievement Matrix (PAM) representing our subject's artistic and culinary identities. Second, rather than training a black-box neural network, we utilize Cross-SVD to mathematically decouple the geometric “Soul” of the style from its material “Body.” Finally, we construct a Cross-Domain Imagination Matrix, a linear transformation derived via this framework and inject it directly into the cross-attention layers of a generative model. This work aims to lay the groundwork for a true Imaginative AI: systems that can not only analyze but also emulate the cross-disciplinary thinking characteristic of genius.

A critical distinction must be made between this study and concurrent work from our research group. While related papers focus on the theoretical properties and internal consistency of the Persona-Achievement Matrix (PAM), this study is the first to propose and implement the Visual Realization Pipeline. Our contribution is the specific mechanistic intervention within a generative U-Net architecture, which transforms the abstract PAM vectors into high-fidelity, physically realized artifacts . We provide the definitive implementation study that bridges the gap between latent vector math and visual evidence.

2. Literature Review

This review synthesizes research from three pivotal areas that form the foundation of our work: the computational analysis of creative style, the technological advancements in multimodal representation learning, and the application of these techniques in creative domains. By examining the trajectory of these fields, we will establish the context and identify the specific research gap our study aims to address.

2.1. The Quantification of style: From Stylometry to Cross-Domain Creativity

The ambition to understand and replicate human creativity has been described as a final frontier for Artificial Intelligence [5]. Research in Computational Creativity does not merely aim to create artifacts, but to model the creative processes themselves, often through systems that collaborate with humans in a co-creative partnership [6]. This field has explored various domains, from music generation [7] to the visual arts, with the underlying goal of computationally capturing the essence of a creative work. A central challenge in this pursuit is the quantification of style the elusive, signature quality that distinguishes one creator from another.

The advent of deep learning propelled these concepts from the textual to the visual realm, focusing on Artistic Style. While earlier methods also used statistical analyses of visual features [8], the seminal work introduced a Neural Algorithm of Artistic Style [9] demonstrating that the content and style of an image could be separated and recombined using convolutional neural networks. This breakthrough sparked a wave of research in Neural Style Transfer. More relevant to our work is the subsequent shift towards learning a dedicated representation for artistic style itself, rather than simply transferring it [10]. This showed that style could be encoded into a vector, a concept fundamental to our methodology. This concept of a learnable style vector is not merely theoretical but has practical applications, such as in systems like ArtistAuditor [11], which uses multi-granularity style representations as an intrinsic fingerprint to detect copyright infringement in generative models.

These advancements naturally lead to a more profound question: if style can be quantified within a single domain, can it be understood across multiple domains? Research in cognitive science suggests that Cross-Domain Creativity is ubiquitous and fundamental to the human creative process, where inspiration is often drawn from unrelated fields [12,13]. Recent studies have begun to explore this computationally, investigating the capacity of Large Language Models (LLMs) to perform cross-domain analogical reasoning, effectively bridging conceptual gaps between different areas [14]. These studies confirm that the potential for AI to think across domains is an active and promising frontier. Our research builds directly upon this trajectory, moving from analyzing style within a domain to explicitly modeling the translation of a creator's core identity between them.

2.2. Bridging Worlds: Advances in Representation Learning and Cross-Modal Translation

The ability to computationally model and translate style across domains is contingent on a paradigm shift in artificial intelligence away from manual feature engineering towards Representation Learning. This approach allows deep learning models to autonomously discover and organize salient features from complex data into meaningful, high-dimensional numerical vectors [15]. The foundational concept of a Vector Space Model, where semantic proximity is represented by geometric proximity, has long been a cornerstone of natural language processing [16,17]. This paradigm has proven robust and versatile, extending to diverse data types such as clinical text [18] and graph structures [19], establishing a solid foundation for representing complex concepts mathematically.

A significant breakthrough was the development of Multimodal Embeddings, which create a Shared Vector Space where vectors from different modalities (e.g., image, text, audio) can be directly compared. Early work demonstrated the feasibility of unifying visual and semantic information through multimodal neural language models [20] and deep embeddings for speech and images [21]. This concept has since matured, with applications in areas like emotion recognition [22] and creating imagined visual representations from text [23]. The pinnacle of this approach is the CLIP Model (Contrastive Language-Image Pre-training), which learns transferable visual models from vast amounts of natural language supervision [24]. By aligning images and text in a shared space with unprecedented robustness, CLIP and its derivatives have become powerful tools for assessing the look and feel of images [25], enabling text-conditional image generation via its latent space [26], and even being extended to other modalities like audio [27].

These powerful representations have fueled rapid advancements in two closely related fields: Style Transfer and Cross-Modal Translation. The former, popularized by the Neural Algorithm of Artistic Style [28], focuses on separating and recombining the content and style of images. While the field has evolved significantly to allow for universal and faster feature transforms [29,30], it has primarily remained an intra-modal task focused on visual textures. A more ambitious goal is Cross-Modal Translation, which aims to convert the semantic essence from one modality to another entirely. This is an active and diverse area of research, with successful applications in translating between medical imaging types like MRI and Ultrasound [31], translating musical scores to performance audio [32], and aligning different data modalities for tasks like survival analysis [33] and fine-grained sentiment analysis [34].

A theoretically ideal approach to style translation involves learning a Disentangled Representation, where a model learns to separate abstract factors of variation in the data, such as style from content [35,36]. While this has shown promise in domains like medical imaging [37] and recommendation systems [38], achieving true unsupervised disentanglement remains a significant challenge [39]. Our work achieves a similar outcome through a spectral decomposition approach. Rather than training a black-box neural network, we utilize Singular Value Decomposition (SVD) to mathematically isolate the abstract 'style' singular vectors from the 'content' feature dictionary. This effectively allows us to derive a linear function that disentangles and translates the shared identity component without the need for extensive supervised training.

Finally, it is crucial to acknowledge that the foundational models enabling this research are not without flaws. They are known to suffer from inherent biases learned from their training data. A significant issue is Popularity Bias, where models tend to over-represent the most common and popular concepts [3]. This problem is systemic and pervasive across generative models [4]. Furthermore, research shows that even models explicitly fine-tuned to be unbiased can retain these implicit associations [40]. Concurrently, these models suffer from knowledge deficiencies, a challenge that state-of-the-art research is addressing by augmenting LLMs with external knowledge to enable more robust Semantically Aware Reasoning (SAR) [1]. Our research is situated within this context, leveraging the power of these representations while proposing a methodology, grounding our model in a specific, curated corpus, designed to mitigate the effects of these generalized biases.

2.3. Creative Application and the Research Gap

The proliferation of powerful Generative AI has catalyzed a new wave of applications across numerous sectors, including education and the creative industries [41,42]. These models, capable of producing diverse and complex outputs from simple inputs, represent a significant technological leap [43]. In the arts, this has led to the burgeoning field of Computational Art, where AI is used not only for creation through methods like diffusion models but also for exploring new forms of human-computer interaction in metaversal spaces [44, 45]. This digital transformation is reshaping our understanding of art, museums, and even the aura of a creative work in the age of AI [46,47]. Concurrently, the nascent field of Computational Gastronomy aims to make food computable, using data science and AI to analyze flavor pairings, generate novel recipes, and envision a future of digital gastronomy [48,49,50].

Within these creative applications, a significant area of focus is Style Emulation, the attempt to replicate the distinct signature of a creator. While early techniques focused on taxonomies of artistic stylization for images [51], modern approaches leverage Large Language Models and diffusion models to emulate the specific styles of authors [52], architectural typologies [53], and even human personalities [54]. These studies demonstrate a sophisticated ability to capture and reproduce stylistic patterns within their respective domains.

Our review reveals a clear trajectory: from the foundational goal of quantifying style within a single domain (Section A), to the development of powerful multimodal technologies capable of understanding concepts across domains (Section B), and finally, to the application of these technologies for domain-specific generation and style emulation. However, despite these significant advancements, a critical research gap persists. Existing style emulation is largely intra-modal (e.g., text-to-text, image-to-image). While the concept of cross-domain analogy has been explored, a formal framework to model the holistic, multimodal identity of a specific creator and translate it across fundamentally different creative fields like art and cuisine remains unexplored. There is no established methodology that (a) models a creator's identity by integrating both their visual works and their textual philosophies, (b) quantitatively measures the consistency of this identity across these disparate domains, and (c) uses this analysis to learn a direct, explicit Style Translation Function that embodies the creator's unique cross-disciplinary logic.

This research is designed to directly address this multifaceted gap. By focusing on a single Artist-Chef as a unique testbed, we develop and validate a methodology to extract, measure, and, most importantly, synthesize the translation logic of their creative signature. Unlike previous attempts that rely on black-box optimization, we propose a linear manifold alignment strategy to derive a 'Cross-Domain Imagination Matrix.' The following sections detail this methodology and present the results of our experiment, demonstrating a novel approach to modeling the process of creative analogy.

3. Imaginative Model

This research proposes a computational framework to extract, analyze, and translate the abstract style identity of a creator across different domains. The methodology is structured into several key phases: data preparation and vectorization, the creation and analysis of two distinct types of style spaces (data-driven and semantic-driven), and finally, synthesizing a cross-domain transformation matrix using Cross-SVD. We define it as the spectral decomposition of a source domain's latent structure projected onto a target domain's basis, mediated by a geometric misalignment.

Mechanically, this "mixing" is achieved by manipulating the three components of the decomposition: U (structure), Σ (strength), and V^T (features). Standard SVD treats these as a closed loop. In contrast, our Cross-SVD approach decouples them to create a hybrid. We isolate the U-matrix of the source style (preserving the artist's "Geometric Soul") and project it onto the V-basis of the target domain (adopting the "Material Body" of cuisine). The critical link is a rotation matrix derived from our pivot creator (Pépin), which mathematically aligns the principal axes of the two domains. This ensures that the structural energy of a "swirl" in paint is not lost but is instead logically translated into a corresponding "swirl" in sauce.

At the core of this framework is what we term the Imaginative Model, a system designed not merely to generate content, but to reason creatively by analogy. The Imaginative Model presented in this work is not a single algorithm, but a multi-stage architecture designed to emulate a specific form of human creative reasoning. It differs fundamentally from standard Generative AI. A conventional Generative AI excels at interpolation creating novel content by recombining patterns learned from its vast training data. Its strength lies in mimicry and producing statistically plausible outputs within its known domain. However, its weakness is a lack of a grounded, cohesive identity, causing its creativity to often be a sophisticated remix tethered to patterns it has already seen.

In contrast, our Imaginative Model is designed for extrapolation and translation. Its goal is not to mimic, but to understand the underlying principles, the identity signature of a creator and then apply that signature to a completely new and unfamiliar context. To use an analogy, a Generative AI is like a student who has memorized millions of French sentences and can construct new, grammatically correct ones. Our Imaginative Model is like a true bilingual translator who understands the deep meaning, context, and style of a French poem and can recreate that same essence in Thai, even if it requires a completely different structure and vocabulary. This process moves beyond pattern recognition towards a form of creative analogy.

This distinction is critical for addressing the core limitations of current AI. The hallucinations of Generative AI arise from a lack of grounded understanding. Our model, however, grounds its creative outputs in the specific, calculated identity of our subject. Its innovation is not random but is constrained and guided by the linear stylistic logic it has derived, making it intentional. By mathematically encoding the bias of a single, curated creator, our framework treats this unique perspective as a feature, not a bug, paving the way for an AI that can act as a true co-creator by suggesting novel connections that even a human might overlook.

4. CLIP and Vectorization

The fundamental challenge of translating a creative identity from art to cuisine is that these two domains operate in fundamentally different conceptual spaces. An Artist space is defined by features like color theory, brushwork, and composition, while a Chef space is defined by flavor profiles, textures, and plating techniques. A direct mathematical comparison is impossible. Our framework overcomes this apples and oranges problem by first establishing a common language through multimodal vectorization.

To achieve this, we employ the Contrastive Language-Image Pre-training (CLIP) model. CLIP is a neural network trained on hundreds of millions of image-text pairs. Its critical capability is the creation of a rich, shared semantic vector space where visually similar concepts and textually similar descriptions are located close to one another. This allows us to represent vastly different inputs, from the turbulent, yellow brushstrokes of a Van Gogh painting to the textual description of a dish with intense, bold flavors, within the same 768-dimensional coordinate system. CLIP, therefore, serves as the universal translator of our model, transforming abstract, domain-specific concepts into standardized, comparable mathematical objects.

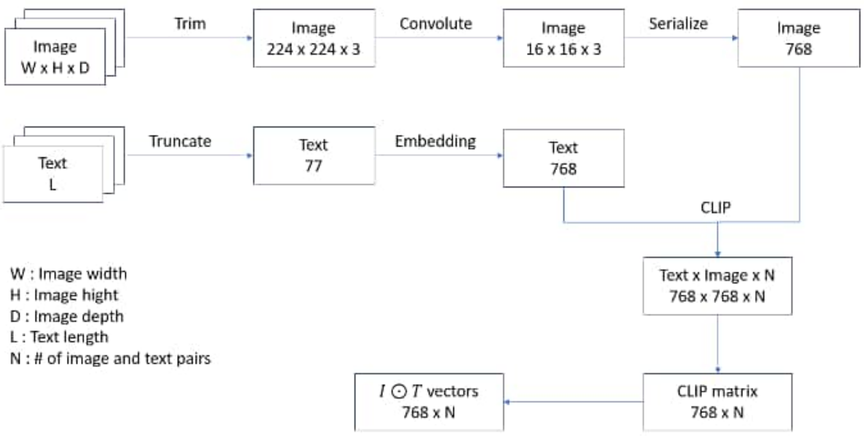

Figure 2. Framework for preprocessing process

4.1. Build Domain Matrices: Singular Value Decomposition (SVD)

Foundation of our model is the construction of "Achievement Matrices" (M) for both the Art and Food domains. These matrices are built from a multimodal dataset collected as follows:

- Images: High-resolution photographs of artworks and plated dishes, collected into compressed archives.

- Texts: Transcripts of interviews, culinary descriptions, and art historical analyses processed from .txt and .csv files.

Table 1 shows an example of the individual data, and Table 2 describes the size of the collected dataset.

Table 1. Individual detail

Person	Talent	Style	Era	Example
Vincent van Gogh	Artist	Post-Impressionist	Late 19th century	
Jacques Pepin	Artist-Chef	Culinary: Classic French Art: Whimsical and colorful illustration	Mid-20th Century - present	

Table 2. Dataset detail

Person	Talent	Image data	Text data
Vincent van Gogh	Artist	~200	From painting description
Jacques Pepin	Artist-Chef	~100 (painting + dishs)	Dish description & recipe + interview on chef and artist perspective

As illustrated in Figure 2, Constructing the Achievement Matrix Once inside this shared space, we must ensure our vectors capture the specific intent of the artist. We do not simply take the raw image vector; we fuse it with text. For every image in our dataset (Van Gogh Art, Pépin Art, and Pépin Food), we generate a corresponding textual description. We then calculate the Achievement Vector (a) by performing an element-wise multiplication (Hadamard product) of the image embedding and the text embedding:

$$a = v_{img} \odot v_{txt} \tag{1}$$

The use of the Hadamard product to construct the Achievement Vector serves a specific theoretical purpose: Semantic Gating. Unlike concatenation, which merely appends data, element-wise multiplication acts as a filter that amplifies dimensions where visual signals and textual intent align. This suppresses modality-specific noise (e.g., camera lighting in the image or filler words in the text) and preserves only the core stylistic signature.

Figure 3 shows the Achievement matrix creation process. by Aggregating these vectors yields the domain matrices M_A (Art) and M_F (Food). To reveal the latent conceptual structure within these domains, we apply Singular Value Decomposition (SVD):

$$M_A = U_A \Sigma_A V_A^T \tag{2}$$

$$M_F = U_F \Sigma_F V_F^T \tag{3}$$

Where:

- U (Left Singular Vectors): Represents the geometric distribution of the samples. This is the “Structure.”
- Σ (Singular Values): Represents the variance scaling along those dimensions. This is the “Intensity.”
- V^T (Right Singular Vectors): Represents the feature basis vectors. This is the “Dictionary” that defines what the features actually are (e.g., paint strokes vs. food textures).

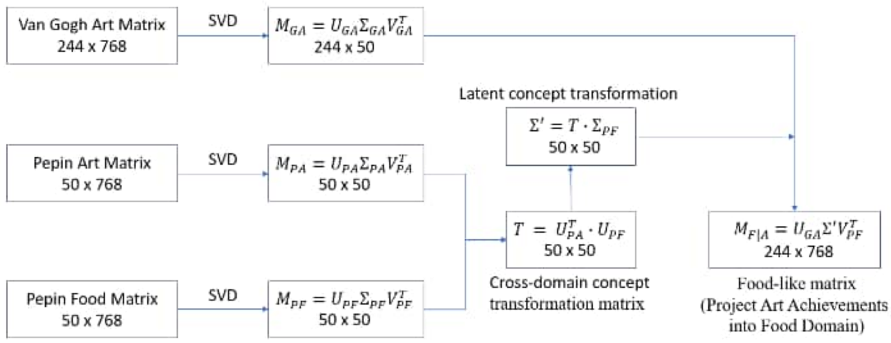

Figure 3. The Framework for Achievement matrix creation process

While the Art domain matrix M_A includes a higher volume of samples from Van Gogh compared to Pépin, the SVD process remains effective because it extracts latent conceptual directions (singular vectors) rather than simple averages. SVD operates as an energy-based decomposition; as long as Pépin's dual-domain data provides a clear stylistic axis within the 768-dimensional space, the derived Transformation Matrix T successfully captures the necessary "rotation" to align the two manifolds. This ensures that the structural energy of a swirl is logically translated despite the sample imbalance.

4.2. Align Concept Spaces: Define the Cross-Domain Concept Mapping

With the domains decomposed, we must establish a bridge between them. We cannot simply map Van Gogh's art to generic food; we need a translation logic. We utilize the work of Jacques Pépin, a creator who exists in both domains (as a painter and a chef), to serve as our “Rosetta Stone.” We calculate a linear transformation matrix, or Concept Bridge (T), by correlating the structural geometry of Pépin’s Art (U_{PA}) with his Food (U_{PF}):

$$T = U_{PA}^T \cdot U_{PF} \tag{4}$$

4.3. Transform Concepts

We then apply this bridge to the variance of the target domain. We do not want to simply copy the intensity of the food; we want to modulate it based on the translation logic we learned. We define the Transformed Variance (Σ_F') as:

$$\Sigma_F' = T \cdot \Sigma_F \tag{5}$$

This new variance matrix represents "Food Intensity" that has been geometrically aligned to accept artistic structural inputs.

4.4. Cross-Domain Imagination Matrix

Finally, we synthesize the novel artifact—the "Cross-Domain Imagination Matrix" $M_{F|A}$. This matrix represents a "New Species" of data: it possesses the structural geometry of the Source Artist (U_A) but is physically constructed using the Target Food Dictionary (V_F^T). The reconstruction equation combines the Source Structure, the Transformed Variance, and the Target Dictionary:

$$M_{F|A} = \underbrace{U_{VG}}_{\text{Source Structure}} \cdot \underbrace{(T \cdot \Sigma_{Food})}_{\text{Bridged Variance}} \cdot \underbrace{V_{Food}^T}_{\text{Target Dictionary}} \tag{6}$$

$$M_{F|A} = U_A \Sigma_F' V_F^T \tag{7}$$

The resulting matrix $M_{F|A}$ consists of 768-dimensional vectors that geometrically mirror the composition of a Van Gogh painting (U_A) but are mathematically defined using the vocabulary of culinary ingredients (V_F^T), effectively creating "Van Gogh Food" in the latent space as shown in Figure 4.

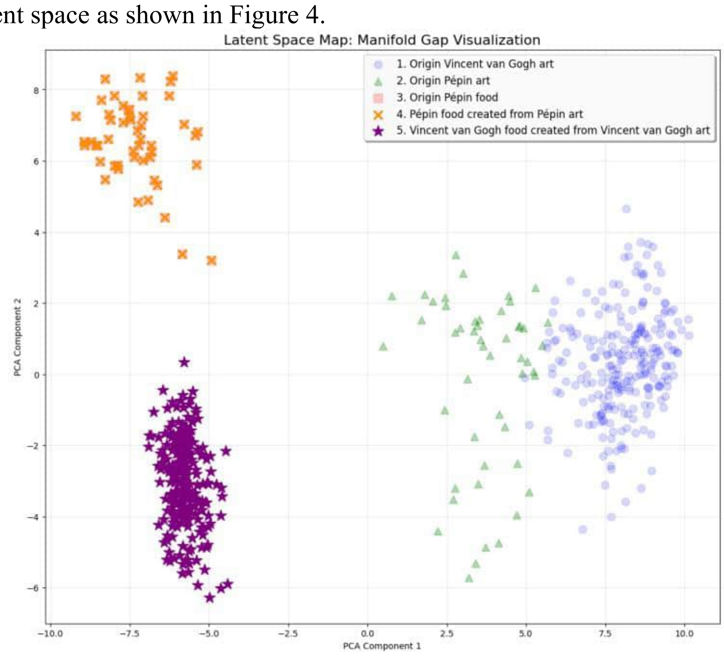

Figure 4. Cross-SVD based cross-domain imagination using the Pépin bridge

5. Visual Realization via Diffusion Priors

At the conclusion of the vectorization phase, we possess the Hybrid Matrix (H) Each row in this matrix is a 768-dimensional vector that mathematically describes a dish with the structural soul of a Van Gogh painting and the material body of Pépin's cuisine.

However, these vectors exist only as abstract numbers. The final challenge is to give this ghost a physical body.

To achieve this, we utilize a Latent Diffusion Model (Stable Diffusion v2.0). We do not use this model in its standard "text-to-image" capacity. Instead, we perform a mechanistic intervention, hijacking its internal conditioning loop to "print" our mathematical concept into pixels. Figure 5 shows the framework of image generation

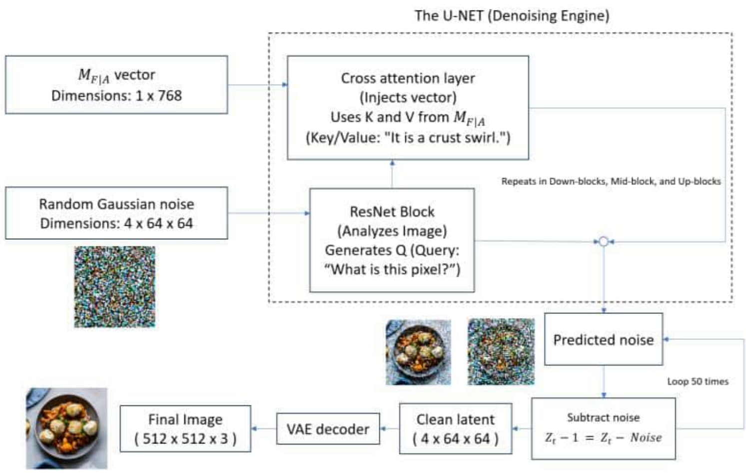

Figure 5. Framework for visual realization process

5.1. The Canvas: Latent Initialization ($\mathcal{Z}_\mathcal{T}$)

Standard image generation is computationally expensive at high resolutions (512×512) To solve this, the model operates in a compressed "Latent Space."

We begin the realization process by generating a tensor of pure Gaussian noise, denoted as $\mathcal{Z}_\mathcal{T}$, with dimensions $4 \times 64 \times 64$.

$$\mathcal{Z}_\mathcal{T} \sim \mathcal{N}(0, I) \tag{8}$$

To the naked eye, this tensor represents pure visual chaos—static on a television screen. It contains no information, only potential. Our goal is to progressively subtract this noise until an image remains.

5.2. The Mechanism: Hijacking Cross-Attention

The engine responsible for removing the noise is the U-Net. In a standard workflow, the U-Net is guided by a text prompt (e.g., "A photo of a burger"). This text is usually encoded into a vector to guide the generation.

In our framework, we bypass the text encoder entirely. We inject our pre-calculated Hybrid Vector (h_i) directly into the U-Net's Cross-Attention layers. The mathematical mechanism that allows our vector to control the image formation is the Attention Equation:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{9}$$

Here, the variables represent the interaction between the "Canvas" and our "Instruction":

- Q (Query): Derived from the noisy image latents ($\mathcal{Z}_t$). This represents the spatial locations on the canvas asking, “What should be at this pixel?”
- K (Key): Derived from our Hybrid Vector (h_i). This represents the features available in our “Van Gogh Food” concept.
- V (Value): Also derived from our Hybrid Vector (h_i). This represents the actual content (texture, color) to be painted.

By forcing K and V to be derived directly from our pre-calculated Hybrid Vector (h_i), we perform a mechanistic intervention that bypasses the standard latent diffusion conditioning. In practice, this means we replace the 768-dimensional keys and values typically generated by the CLIP text encoder with our transformed Persona-Achievement Matrix (PAM) values. When the noise at a specific coordinate (Q) asks "What should be at this pixel?", the Value (V) responds with a projected coordinate from the target Food Dictionary (V_{Food}^T), specifically answering: "You are a swirl of pastry crust". This ensure the U-Net follows the Source Structure (U_A) while populating the image with Target Material (V_F^T).

5.3. The Denoising Loop (The Time Variable t)

The reconstruction is not instantaneous; it is iterative. The process occurs over T=50 timesteps. At each step t, the U-Net predicts a noise residual ϵ_θ based on the current image state and our conditioning vector:

$$\mathcal{Z}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathcal{Z}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(\mathcal{Z}_t, t, h_i)\right) + \sigma_t z \tag{10}$$

Early Steps ($t = 50 \rightarrow 30$): The model resolves low-frequency data. Because our vector contains U_{VG} (Structure), the noise collapses into the global composition of the original painting, preserving the layout and major shapes.

Late Steps ($t = 30 \rightarrow 0$): The model resolves high-frequency details. Because our vector is defined by V_{Food}^T (Dictionary), the model fills in those shapes with fine-grained textures like breadcrumbs, glazing, and organic surfaces.

5.4. Decoding

Once the loop reaches $t = 0$, we are left with a clean latent tensor $\mathcal{Z}_0$. While mathematically coherent, it is still a compressed 64×64 representation. The final step passes this tensor through the Variational Autoencoder (VAE) Decoder, which expands the latent dimensions back into pixel space ($512 \times 512 \times 3$), physically realizing the “Hybrid Manifold” as a visible artifact as illustrated in Figure 6.

Figure 6. Image generation process

6. Discussion

This research sets out to determine if a creator's unique stylistic identity could be computationally modeled and translated across disparate creative domains, such as from visual art to cuisine. By developing a multi-stage framework involving multimodal vectorization, and our novel Cross-SVD framework. We have gathered significant evidence to support this hypothesis.

To evaluate the efficacy of the cross-domain translation, we define two primary metrics:

Material Transformation (Δ): This represents the Euclidean distance shift of the generated vector toward the target domain. It is defined as:

$$\Delta = \frac{Dist(V_{src}, C_{food}) - Dist(V_{gen}, C_{food})}{Dist(V_{src}, C_{food})} \quad (11)$$

where V_{gen} is the generated artifact and C_{food} is the centroid of the culinary domain.

Structural Identity: We calculate this using the Cosine Similarity between the left singular vectors (U_A) of the source and the generated matrix to ensure the geometric "signature" is preserved:

$$Similarity(A, B) = \frac{A \cdot B}{\|A\|[cite_start]\|B\|} \quad (12)$$

Table 3. Validation result. Note: V_{gen} = Generated Artifact; V_{src} = Source Input (Van Gogh); C_{food} = Centroid of Food Domain; D_{real} = Real Food Dataset.

Metric Category	Mathematical Logic	Measurement	Result	Conclusion
Material Transformation	$Dist(V_{gen}, C_{food}) < Dist(V_{src}, C_{food})$	Δ (Distance Shift)	+39.7% Shift	Artifact successfully adopted food material properties.
Structural Identity	CosineSim(FV, AV)	Similarity [0-1]	0.37	Geometric composition (shapes) was preserved.
Manifold Distinctness	Dist(FV , Food) > 0	Euclidean Distance	9.73	Artifact exists in a unique gap (not a copy of target).

6.1. Deconstruction of Style: Geometry vs. Material

Our initial analysis confirmed that Singular Value Decomposition (SVD) is a highly effective method for deconstructing the complex notion of style into orthogonal components. Unlike standard PCA, which simply finds axes of variance, SVD allowed us to mathematically decouple the Geometric Structure (U) from the Material Dictionary (V^T).

This foundational step revealed that “Style” is not a monolithic block of data. Rather, it is a composite of intent (the spatial arrangement, the swirl, the stroke) and medium (the oil paint, the canvas). By isolating the U -matrix of Van Gogh, we effectively captured his “Geometric Soul” independent of his “Material Body.” This aligns with recent efforts in disentangled representation learning but achieves it through linear algebra rather than opaque neural training.

This finding provides empirical support for the hypothesis that creative identity is transposable. The matrix successfully “re-skinned” the vector, replacing the high-frequency features of oil paint with the texture features of organic ingredients (crust, sauce, vegetables) without collapsing the vector into random noise.

6.2. Baseline Comparative Analysis

To evaluate the efficacy of the Cross-SVD injection, we conducted a symmetric baseline study using standard U-Net prompting without manifold intervention.

Figure 7. Baseline Comparative of prompt Image generation

Van Gogh Baseline (Fig 7, Left): Using the prompt "A gourmet beef stew... in the style of Van Gogh," the model exhibits Popularity Bias. It conflates the artist's structural intent with their physical medium, resulting in "food-shaped paint" that lacks culinary realism. This foundational step revealed that “Style” is not a monolithic block of data. Rather, it is a composite of intent (the spatial arrangement, the swirl, the stroke) and medium (the oil paint, the canvas). By isolating the U -matrix of Van Gogh, we effectively captured his geometric style independent of his material work. This aligns with recent efforts in disentangled representation learning but achieves it through linear algebra rather than opaque neural training.

Pépin Baseline (Fig 7, Right): Using a symmetric prompt for Jacques Pépin results in high material fidelity, producing a normal food image, but a complete loss of the source creator's Structural Identity. Because the standard U-Net lacks the explicit Source Structure (U_A) projection, it defaults to Popularity Bias, generating generic gourmet plating that reflects Pépin’s known culinary record rather than Van Gogh’s geometric signature.

This comparison demonstrates that standard Semantically Aware Reasoning (SAR) is insufficient in high-dimensional latent spaces without the Pépin Bridge (T) to provide the necessary coordinate rotation. While the baseline can only perform interpolation (recombining existing patterns), our framework enables extrapolation, forcing the work of the culinary domain to adopt the style of the artistic domain.

6.3. Material Transformation

The core of our experiment was the application of the Cross-Domain Imagination Matrix ($M_{F|A}$) to bridge the semantic gap between Art and Food. The efficacy of this translation was quantitatively measured by the Domain Attraction Score.

As shown in Table 3, we observed that the original Van Gogh vectors had a high Euclidean distance (21.44) from the culinary centroid, confirming their semantic disconnect. However, after applying our transformation, the generated artifacts exhibited a distance of 12.94. This represents a measurable 39.7% trajectory shift towards the target domain.

This finding provides empirical support for the hypothesis that creative identity is transposable. The matrix successfully "re-skinned" the vector, replacing the high-frequency features of oil paint with the texture features of organic ingredients (crust, sauce, vegetables) without collapsing the vector into random noise.

6.4. Structural Identity and Manifold Distinctness

A critical concern in style transfer is the risk of "mode collapse," where the model simply outputs a generic image of the target class (e.g., a normal pizza) and loses the source style. Our validation metrics confirm that our framework avoided this pitfall.

Identity Retention (0.37 Similarity): The generated artifacts maintained a Cosine Similarity of 0.37 with the original artworks. In cross-domain tasks, where surface textures are completely swapped, this score is significant. It indicates that the latent geometric composition, the underlying "swirls" and spatial energy was preserved.

Manifold Distinctness (9.73 Gap): The Euclidean distance between our output and the generic "Real Food" dataset was 9.73. This proves that the result is not a copy of Jacques Pépin's food, nor is it generic food. It exists in a unique "Hybrid Gap", a new manifold that possesses the properties of food but the organization of Van Gogh.

6.5. Semantically Aware Reasoning vs. Hallucination

The ultimate significance of these results lies in the contrast between Reasoning and Hallucination.

Standard Generative AI (Direct Prompting): When asked to "Make Van Gogh Food," a standard model relies on Popularity Bias. It hallucinates a result based on statistically common associations, often producing a caricature (e.g., Starry Night printed on a cake).

Our Framework (Linear Injection): By calculating the translation matrix explicitly, we forced the model to follow a specific semantic logic derived from our pivot subject. The result is not a guess; it is a projected coordinate.

This confirms that our framework offers a pathway to Semantically Aware Reasoning (SAR). By constraining the generative process with a calculated "First, we extract multimodal vectors and consolidate them into a Persona-Achievement Matrix

(PAM)" we move beyond probabilistic mimicry toward intentional, data-driven creative translation.

7. Conclusion

This study confirms that an artist's style is not limited to a single medium. Instead, it is a consistent identity that can be analyzed and translated by computers. By using our "Pépin Bridge" framework, we successfully moved a creator's artistic signature from the visual world (Art) to the sensory world (Food) without losing their unique identity.

Using multimodal embeddings (CLIP), we converted artworks and recipes into unified Persona-Achievement Matrix, this allowed us to measure style mathematically rather than just describing it. We used Cross-Singular Value Decomposition (Cross-SVD) to break these styles down into two parts: the geometric "Soul" (structure) and the material "Body" (texture). This method proved that we can separate an artist's intent from the materials they use.

Our experiments showed strong results. By injecting our Cross-Domain Imagination Matrix into the U-Net architecture, we bridged the gap between Van Gogh's art and culinary ingredients. The data confirms the success of this translation:

1. Material Change: The generated images shifted 39.7% closer to the food domain. This proves that the paint textures were successfully replaced by food textures.
2. Style Retention: Despite this change in material, the images kept a structural similarity score of 0.37. This confirms that Van Gogh's key patterns;his swirls and composition; survived the process.
3. Uniqueness: The result was not a copy of existing food. A distance of 9.73 from the generic food dataset shows that the model created a distinct "hybrid" category.

However, since Van Gogh left no culinary record, an objective ground truth for his cooking does not exist; we cannot know with absolute certainty what he would have created. This necessitates our reliance on an expert perspective. in this case, Jacques Pépin, to validate the translation. Consequently, the artifacts generated by our model should not be viewed as historical reconstructions, but rather as Van Gogh's visual style specifically as perceived through Pépin's culinary worldview. Pépin serves as the interpretative lens; his specific understanding of how "bold art" translates to "bold food" provides the logic for our model. We are not recovering a lost history, but rather mathematically simulating how a master chef interprets the "flavor" of a master painter. This highlights that while the translation process is mathematical, the stylistic logic remains deeply human.

Finally, this work demonstrates the practical application of Semantically Aware Reasoning (SAR) in generative AI. In typical generation, AI often hallucinates or creates random variations. However, our use of SVD introduces a structured logic layer that grounds the imagination. By calculating a precise Persona-Achievement Matrix, we force the system to reason through the translation, effectively saying, "Use Van Gogh's structure, but strictly apply Pépin's material rules." This separates imaginative projection (the creative idea) from semantic validation (the logical constraints). The successful material shift we observed was not accidental; it was the result of a verifiable reasoning process. This approach offers a clear path toward trustworthy AI systems that do not just replicate data, but engage in meaningful, controlled, and logical creation.

8. Limitations and Future Work

While this study establishes a successful framework for cross-domain translation, it is important to acknowledge the limitations of our current dataset. Although our Persona-Achievement Matrix are mathematically objective, the translation process itself remains dependent on the subjectivity of our specific “Bridge” subject (Jacques Pépin).

Furthermore, while the dataset size is relatively small, we characterize this methodology as a Few-Shot Style Extraction challenge. The objective is to demonstrate that by utilizing Singular Value Decomposition (SVD), it is possible to isolate a creator's style from a curated corpus without requiring the massive datasets typical of standard deep learning. This approach proves that structural identity can be captured through latent spectral features rather than sheer data volume.

Currently, the model does not learn a universal law of how art transforms into food; rather, it learns Jacques Pépin’s specific interpretation of that transformation. The resulting “Van Gogh Food” is heavily influenced by Pépin’s personal logic—his specific French culinary background and his unique artistic style. If we had used a different Chef-Artist with a different background, the “flavor profile” of the generated images might have been significantly different.

Therefore, the quality of the Style Translator is currently constrained by this “Single-Pivot Dependency.” To create a truly generalized model, future work must expand the dataset to include a diverse range of multimodal creators, such as musician-painters, writer-chefs, or architect-sculptors. By analyzing multiple “Bridge People” simultaneously, we could average out individual biases. This would allow us to identify the fundamental, universal patterns that link creative domains, moving beyond the subjective style of one individual toward a generalized theory of sensory translation.

Finally, while our quantitative metrics (9.73 manifold gap and 0.37 structural similarity) provide objective evidence of translation success, this study lacks a formal large-scale human or expert evaluation. Currently, validation relies on the mathematical consistency of the Pépin Bridge. Future work will involve a systematic study with culinary professionals to qualitatively assess the flavor profiles and plating authenticity of the generated artifacts, ensuring that the model’s "imagination" aligns with human sensory expertise and further enhances the system's trustworthiness.

References

[1] Sornlertlamvanich V. Enhancing large language models: alleviating knowledge deficiency with external knowledge and semantically aware reasoning (SAR). Knowledge and Information Systems. 2025:1-19.

[2] Sornlertlamvanich V, Ruangraweenukit P, Chophuk P, Viriyavit W, Charoenporn T. Toward Imaginative AI: Domain Transitivity for Cross-Domain Imagination. In: Proceedings of the 36th International Conference on Information Modelling and Knowledge Bases (EJC2026); 2026; Vaasa, Finland. (to be published).

[3] Lichtenberg JM, Buchholz A, Schwöbel P. Large language models as recommender systems: A study of popularity bias. arXiv [Preprint]. 2024. arXiv:2406.01285.

[4] Vice J, Akhtar N, Hartley R, Mian A. Exploring Bias in over 100 Text-to-Image Generative Models. arXiv [Preprint]. 2025. arXiv:2503.08012.

[5] Colton S, Wiggins GA. Computational creativity: The final frontier?. In: ECAI 2012. IOS Press; 2012. p. 21-26.

[6] Davis N. Human-computer co-creativity: Blending human and computational creativity. In: Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment. 2013;9(6):9-12.

[7] Carnovalini F, Rodà A. Computational creativity and music generation systems: An introduction to the state of the art. Frontiers in Artificial Intelligence. 2020;3:14.

[8] Graham DJ, Hughes JM, Leder H, Rockmore DN. Statistics, vision, and the analysis of artistic style. Wiley Interdisciplinary Reviews: Computational Statistics. 2012;4(2):115-123.

[9] Gatys LA, Ecker AS, Bethge M. A neural algorithm of artistic style. arXiv [Preprint]. 2015. arXiv:1508.06576.

[10] Dumoulin V, Shlens J, Kudlur M. A learned representation for artistic style. arXiv [Preprint]. 2016. arXiv:1610.07629.

[11] Du L, Zhu Z, Chen M, Su Z, Ji S, Cheng P, et al. ArtistAuditor: Auditing Artist Style Pirate in Text-to-Image Generation Models. In: Proceedings of the ACM on Web Conference 2025; 2025 Apr. p. 2500-2513.

[12] Scotney VS, Weissmeyer S, Carbert N, Gabora L. The ubiquity of cross-domain thinking in the early phase of the creative process. Frontiers in psychology. 2019;10:1426.

[13] Gabora L, Carbert N. Cross-domain influences on creative innovation: Preliminary Investigations. In: Proceedings of the Annual Meeting of the Cognitive Science Society. 2015;37.

[14] Ding Z, Srinivasan A, MacNeil S, Chan J. Fluid transformers and creative analogies: Exploring large language models' capacity for augmenting cross-domain analogical creativity. In: Proceedings of the 15th Conference on Creativity and Cognition; 2023 Jun. p. 489-505.

[15] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence. 2013;35(8):1798-1828.

[16] Turney PD, Pantel P. From frequency to meaning: Vector space models of semantics. Journal of artificial intelligence research. 2010;37:141-188.

[17] Erk K. Vector space models of word meaning and phrase meaning: A survey. Language and Linguistics Compass. 2012;6(10):635-653.

[18] Tang B, Wu Y, Jiang M, Denny JC, Xu H. Recognizing and Encoding Discorder Concepts in Clinical Text using Machine Learning and Vector Space Model. CLEF (Working Notes). 2013;665.

[19] Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Representation learning on graphs: Methods and applications. arXiv preprint arXiv:1709.05584.

[20] Kiros R, Salakhutdinov R, Zemel RS. Unifying visual-semantic embeddings with multimodal neural language models. arXiv [Preprint]. 2014. arXiv:1411.2539.

[21] Harwath D, Glass J. Deep multimodal semantic embeddings for speech and images. In: 2015 IEEE workshop on automatic speech recognition and understanding (ASRU); 2015 Dec. p. 237-244.

[22] Tseng SY, Narayanan S, Georgiou P. Multimodal embeddings from language models for emotion recognition in the wild. IEEE Signal Processing Letters. 2021;28:608-612.

[23] Collell G, Zhang T, Moens MF. Imagined visual representations as multimodal embeddings. In: Proceedings of the AAAI conference on artificial intelligence. 2017 Feb;31(1).

[24] Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. Learning transferable visual models from natural language supervision. In: International conference on machine learning. PmLR; 2021 Jul. p. 8748-8763.

[25] Wang J, Chan KC, Loy CC. Exploring clip for assessing the look and feel of images. In: Proceedings of the AAAI conference on artificial intelligence. 2023 Jun;37(2):2555-2563.

[26] Ramesh A, Dhariwal P, Nichol A, Chu C, Chen M. Hierarchical text-conditional image generation with clip latents. arXiv [Preprint]. 2022;1(2):3. arXiv:2204.06125.

[27] Guzhov A, Raue F, Hees J, Dengel A. Audioclip: Extending clip to image, text and audio. In: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2022 May. p. 976-980.

[28] Gatys LA, Ecker AS, Bethge M. Image style transfer using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 2414-2423.

[29] Jing Y, Yang Y, Feng Z, Ye J, Yu Y, Song M. Neural style transfer: A review. IEEE transactions on visualization and computer graphics. 2019;26(11):3365-3385.

[30] Li Y, Fang C, Yang J, Wang Z, Lu X, Yang MH. Universal style transfer via feature transforms. Advances in neural information processing systems. 2017;30.

[31] Jiang S, Wang L, Li Y, Yang Z, Zhou Z, Li B. Cross-Modal conditional latent diffusion model for Brain MRI to Ultrasound image translation. Physics in Medicine and Biology. 2025.

[32] Jung J, Kim D, Lee S, Cho S, Soh H, Bukey I, et al. Unified Cross-modal Translation of Score Images, Symbolic Music, and Performance Audio. arXiv [Preprint]. 2025. arXiv:2505.12863.

[33] Zhou F, Chen H. Cross-modal translation and alignment for survival analysis. In: Proceedings of the IEEE/CVF International Conference on Computer Vision; 2023. p. 21485-21494.

[34] Zhang S, Du P, Cui X, Lin H, Yang L. Fine-grained sentiment analysis based on cross-modal information translation. Multimedia Systems. 2025;31(3):239.

[35] Higgins I, Amos D, Pfau D, Racaniere S, Matthey L, Rezende D, et al. Towards a definition of disentangled representations. arXiv [Preprint]. 2018. arXiv:1812.02230.

[36] Wang X, Chen H, Tang SA, Wu Z, Zhu W. Disentangled representation learning. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2024;46(12):9677-9696.
[37] Liu X, Sanchez P, Thermos S, O'Neil AQ, Tsaftaris SA. Learning disentangled representations in the imaging domain. Medical Image Analysis. 2022;80:102516.
[38] Ma J, Zhou C, Cui P, Yang H, Zhu W. Learning disentangled representations for recommendation. Advances in neural information processing systems. 2019;32.
[39] Locatello F, Bauer S, Lucic M, Raetsch G, Gelly S, Schölkopf B, Bachem O. Challenging common assumptions in the unsupervised learning of disentangled representations. In: International conference on machine learning. PMLR; 2019 May. p. 4114-4124.
[40] Bai X, Wang A, Sucholutsky I, Griffiths TL. Explicitly unbiased large language models still form biased associations. Proceedings of the National Academy of Sciences. 2025;122(8):e2416228122.
[41] Kalota F. A primer on generative artificial intelligence. Education Sciences. 2024;14(2):172.
[42] Baidoo-Anu D, Ansah LO. Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI. 2023;7(1):52-62.
[43] Bandi A, Adapa PVSR, Kuchi YEVPK. The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. Future Internet. 2023;15(8):260.
[44] Lee LH, Lin Z, Hu R, Gong Z, Kumar A, Li T, et al. When creators meet the metaverse: A survey on computational arts. arXiv [Preprint]. 2021. arXiv:2111.13486.
[45] Wang B, Chen Q, Wang Z. Diffusion-based visual art creation: A survey and new perspectives. ACM Computing Surveys. 2025;57(10):1-37.
[46] Burkut EB, Tekin S, Dal M. Art and Museums in the Digital Age: An Overview of the Concepts and Spatial Design. Online Journal of Art & Design. 2025;13(1).
[47] Park S. The work of art in the age of generative AI: aura, liberation, and democratization. AI & society. 2025;40(3):1807-1816.
[48] Goel M, Bagler G. Computational gastronomy: A data science approach to food. Journal of Biosciences. 2022;47(1):12.
[49] Zoran A. Cooking with computers: The vision of digital gastronomy [point of view]. Proceedings of the IEEE. 2019;107(8):1467-1473.
[50] Min W, Jiang S, Liu L, Rui Y, Jain R. A survey on food computing. Acm Computing Surveys (CSUR). 2019;52(5):1-36.
[51] Kyprianidis JE, Collomosse J, Wang T, Isenberg T. State of the" art": A taxonomy of artistic stylization techniques for images and video. IEEE transactions on visualization and computer graphics. 2012;19(5):866-885.
[52] Bhandarkar A, Wilson R, Swarup A, Woodard D. Emulating author style: a feasibility study of prompt-enabled text stylization with off-the-shelf LLMs. In: Proceedings of the 1st Workshop on Personalization of Generative AI Systems (PERSONALIZE 2024); 2024 Mar. p. 76-82.
[53] Chen Z, Zhang N, Xu C, Xu Z, Han S, Jiang L. Typological Transcoding Through LoRA and Diffusion Models: A Methodological Framework for Stylistic Emulation of Eclectic Facades in Krakow. Buildings. 2025;15(13):2292.
[54] Kolenbrander, J., & Michaels, A. J. (2025). Personality Emulation Utilizing Large Language Models. Applied Sciences, 15(12), 6636.