It’s all in the card!
Search Morphological Links on Tarot Card

Haotian Fang; semester project supervised by Frédéric Kaplan and Isabella Di Lenardo

École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
Emails: haotian.fang@epfl.ch

Abstract

This report presents a computational analysis of tarot card iconography across ten historically significant decks, focusing on the Major Arcana. By systematically dividing each card into forty spatial regions and generating detailed, morphologically grounded descriptions with large language models, we apply data-driven analysis for art historical interpretation. Visual and textual representations of patches are embedded in a shared vector space using CLIP, enabling fine-grained similarity comparisons and retrievals across decks and periods. We further refine these embeddings using a multi-negative triplet loss, aligning image features with textual descriptions and facilitating the discovery of recurring visual motifs independent of artistic style. Qualitative and clustering-based evaluations demonstrate the potential for this method to uncover both explicit and subtle iconographic patterns. We discuss limitations in dataset scope and annotation, and outline future work incorporating narrative structure and interactive retrieval tools to advance computational tarot studies.

Keywords: tarot cards history, cross-modal retrieval, metric learning, symbolic analysis, visual motif discovery

1. Introduction

1.1 Tarot History

Tarot cards are not just a set of beautiful images—they are also a divinatory system, a way to reveal the unknown through mystery and revelation. Tarot answers questions about the past and present and hints at the future. Many people have discovered the powerful energy within them.(Place 2005) Life is often compared to a journey, and the images on tarot cards can reflect the stages and challenges of that journey.

For those unfamiliar with tarot, the cards are simply pictures; but for a trained reader, they are a language. The reader is the translator for those seeking to understand what the cards mean. Tarot reading is deeply personal, and each reader’s interpretive style and method can vary greatly.(Pollack 1997; Greer 1984)

However, the workflow of tarot reading is almost the same. The reader typically asks the querent (person who asks a question) to shuffle the deck and then lays out the cards according to a fixed pattern, known as a “spread.” The reader then interprets the cards and points out their order.(Greer 1984)

The images on the cards are mystical symbols that can evoke responses on both the conscious and unconscious levels—these are shared elements of basic human experience. Tarot readers are trained to recognize the multiple layers of meaning in each image. Robert Place believes that the best readers always combine their knowledge of the cards with intuition. Readers must learn the symbols and the wisdom traditions behind them, but also have to make a leap of faith and trust the intuition—which, at its core, means looking at the pictures and telling a story.(Place 2005)

Each deck is different, and the meanings assigned to the cards also differ. There is no definitive tarot deck; there are hundreds to choose from, and most readers pick one or two decks to use regularly. The specific deck you use does not matter as long as the reader fully understands its symbols—the interpretations should converge in the end.(Farley 2009)

Another important aspect is the structure: a standard tarot deck usually has 78 cards, made up of two parts, or "arcana," from the Latin word for secret. The major arcana and the minor arcana. The 56 minor arcana cards are divided into four suits, much like our modern playing cards.(Pollack 1997) The major arcana cards look different; they are full of complex images representing archetypes, which Jung described as universal symbols from the collective unconscious. In the tarot system, these cards are thought to represent challenges that we all encounter in real life.(Jung 1934)

1.2 Previous Work

Prior work has explored computational analysis of visual symbolism in art, aiming to automatically detect recurring motifs across artworks. The one I'm learning through is from the Semi-supervised Clustering of Visual Signatures of Artworks by (Schaerf 2022), which introduced a human-in-the-loop pipeline to trace propagating patterns in art history. In that approach, a deep CNN (ResNeXt) was fine-tuned with a compound hinge-based metric loss on an initial set of known visual correspondences, learning visual signature embeddings that highlight shared iconographic details. The learned descriptors were then clustered using an unsupervised density-based algorithm (OPTICS), grouping artworks that contain the same recurring motif. This semi-supervised strategy (fine-tuning on known links plus unsupervised clustering) yielded high-precision clusters and outperformed purely generic clustering methods. Crucially, art historians periodically validated and annotated the motif clusters and fed these discoveries back into the training loop, markedly expanding the catalog of visual links. However, I'm facing a relatively small dataset around 200 images of tarot card without human annotation, so I have to leverage more automatic method to make it possible, and I will debrief it in the later section.

In this project, I hope to identify the symbols and structural patterns found across decks of different eras and styles to have a better understanding of tarot history.

2. Data

2.1 Basic Description

We have collected 10 different decks of tarot cards(major arcana only) suggested by humanity researchers in DH Lab, and here are the descriptions(Table 1) and the preview of the dataset(Figure 1).

The dataset are labeling as its deck number, position number, card name.

The deck number list 1-10 from earliest 1456 till 1937

The position number is listed from 1-22, 1-21 is the card from The Magician till the World, and the 22 is the zero card Fool.

Table 1. Tarot card sets we collected

No.	Name	Place and Date
1	Tarot enluminé "Colleoni"	Italy, Cremona, 1456–1458
2	Tarot enluminé "de Charles VI"	Italy, Florence, ca. 1460
3	Tarot parisien anonyme	France, Paris, 1600–1650
4	Tarocchino de Bologne	Italy, Bologna, 1600–1700
5	Tarot de Marseille	France, Lyon, 1701–1715
6	Tarot de Besançon	France, Strasbourg, 1746
7	Tarot de Marseille	France, Paris, 1930
8	Tarot d'Etteilla	France, Paris, 1850
9	Tarot occultiste – Tarot kabbalistique	France, Paris, 1889
10	Tarot occultiste – Tarot Waite-Smith	United Kingdom, 1937 (1909 for the original edition)

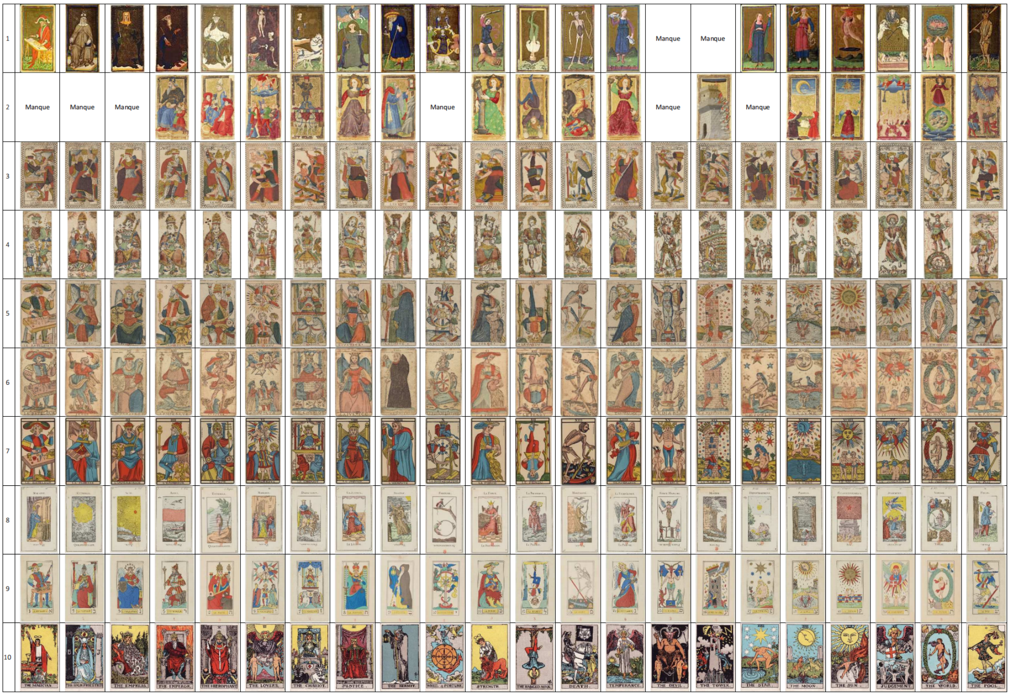

Figure 1. Preview of the dataset

2.2 Detailed Understanding

The history of tarot is vague and unclear, almost as mysterious as the images themselves. Research by Michael Dummett indicates that tarot cards originated in northern Italy in the early 15th century.(M. Dummett 1980) However, in the earliest days, tarot was not used for divination, but was simply a card game. For example, some of the earliest tarot decks, around 1450, were luxury items for the wealthy. Nobles competed to obtain the finest decks, which is why, as you can see, these cards are so richly decorated. The Visconti family even included portraits of their own relatives, such as Sister Manfreda, on the cards.(M. A. E. Dummett 1986)

See Figure 2

Figure 2. The High Priestess / Featured from Sister Manfreda

And the first two decks we collected missed some positions showing in the Figure 1 because at this stage tarot card is not for fortunetelling but more for collection.

If you pay close attention to our dataset, you will notice that the time gap between deck 2 and 3, that is, between around 1460 and 1600. This gap is historically significant. During the 15th and 16th centuries, the Church often suppressed activities or symbols perceived as contrary to Christianity.

Tarot cards, which were sometimes associated with heresy or unapproved forms of divination, were periodically banned or destroyed, and their use went underground.(M. Dummett 1980; Farley 2009)

The parlor game enjoyed by Renaissance aristocrats was only the beginning of the tarot's story. Over time, the appeal of tarot widened, and the cards were simplified, copied, and eventually printed. One of the first decks to become widely popular was the Tarot de Marseille (deck 5 in our dataset), which emerged in the early 18th century. By this time, tarot cards had evolved beyond a mere game and were increasingly used for fortune-telling—a practice that began in Italy but spread more widely about three centuries after the cards first appeared.(M. Dummett 1980; Huson 2004)

Figure 3. Tarot de Marseille

Each card came to be associated with a specific meaning, reflecting both earlier traditions and newer, often social, interpretations. In the late 18th century, a new theory arose: some believed tarot cards were evidence of ancient Egyptian magical knowledge, lost libraries, and rituals, reincarnated in the so-called "Book of Thoth." This idea was popularized by Antoine Court de Gébelin in his influential writings, although modern historians see no evidence for an Egyptian origin.(Court de Gébelin 1781; M. Dummett 1980)

Shortly after Gébelin's publications, a Frenchman named Alliette—better known as Etteilla (his name reversed)—founded the first known tarot card society of readers and developed new interpretations and spreads for the cards. He contributed to the spread of tarot as a divinatory tool, as seen in decks such as number 8 in our data (which is not included in our main analysis due to its significant differences from other decks).(Huson 2004)

Figure 4. Tarot d'Etteilla

In the 19th century, renewed interest in mysticism, magic, alchemy, and the Kabbalah brought tarot to the center of European occultism. Deck 9 in our dataset reflects the so-called "Occultist Tarot," developed by figures such as the members of the Hermetic Order of the Golden Dawn. They saw tarot as containing potent symbols that unified mystical systems like the Tree of Life in Kabbalah and the alchemical quest for the philosopher's stone.(Greer 1984)

Figure 5. Tarot occultiste - Tarot kabbalistique

This period of innovation culminated in the creation of the Waite-Smith deck (deck 10), produced by Arthur Edward Waite and Pamela Colman Smith in the early 20th century. The Waite-Smith deck is now one of the most influential and widely used tarot decks. Its imagery is much easier to read and interpret than that of the earlier Marseille deck, contributing to its lasting popularity.(Greer 1984; Place 2005)

Figure 6. Tarot occultiste - Tarot Waite-Smith

A good example is the first card, magician, for Tarot de Marseille, the shape of the hat is like infinity sign, and for Tarot Waite-Smith, it clearly shows up. Figure 12 and this is the type of pattern later we want to find.

Figure 7. Magician in Tarot de Marseille(Left) and Tarot Waite-Smith(Right)

3. Methods and Results

Here I will show you the journey of how I came up with this pipeline with understanding each step's result

3.1 Dataset Cropping

First, we want to have a localized level view of the cards we have.

I systematically divided each card into exactly 40 predefined spatial regions per image. This regions utilizes the `get_patch_regions` function, defined as follows:

- Eight primary regions covering full halves and quadrants (full top, full bottom, full left, full right, top-left, top-right, bottom-left, bottom-right).
- Each primary quadrant is further cropped into eight additional regions, including smaller quadrants and sub-regions (top, bottom, left, right, top-left, top-right, bottom-left, bottom-right).

This structured division allows systematic capture of spatially localized visual. One example is shown below Figure 8

Figure 8. Cropping One Card into 40 Regions

3.2 Patch Description Generation

Then we want to annotate the images patch of their morphological links for later triplet loss learning. However, we do not have available human expert to annotate, so I came up with the idea to leverage the capacity of LLM to replace the human expert by giving the directional prompt for reading the card.

Descriptions for each of the 40 image patches per card were generated using OpenAI's `gpt-4o-mini` model. Each patch was encoded to base64 and sent to the OpenAI API, accompanied by the following morphological-related prompt:

> *"You are an expert in describing visual patterns in symbolic artwork. This image patch is extracted from a tarot card. Please output only a single, concise English sentence (20–25 words) listing all visible objects, figures, geometric shapes, colors, textures, and their spatial relationships within the patch. Focus on literal, morphological, and compositional details that can be compared computationally across images. Do not mention any text, numbers, card names, or symbolic meanings. Avoid subjective or interpretive words. Do not start with phrases like "The patch features", "The image patch features", etc."*

This prompt is designed for consistent and detailed verbalizations of the visual part focusing on morphological details in each image patch. Examples of the result is shown below Figure 10

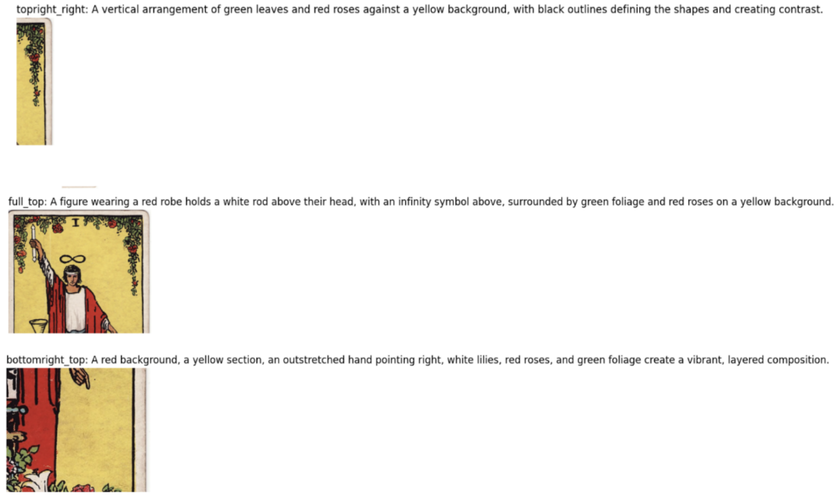

Figure 9. Prompt of the regions Example

3.3 Image Patch and Description Embedding

Then we want to embed both image and the description in the same vector space., and we are using CLIP model (`openai/clip-vit-large-patch14`), which is the most used CLIP in hugging face. Both image patches and their textual descriptions were processed to yield normalized embeddings suitable for direct cosine similarity computations. L2 normalization ensured embedding vectors were on a comparable scale.

Figure 10

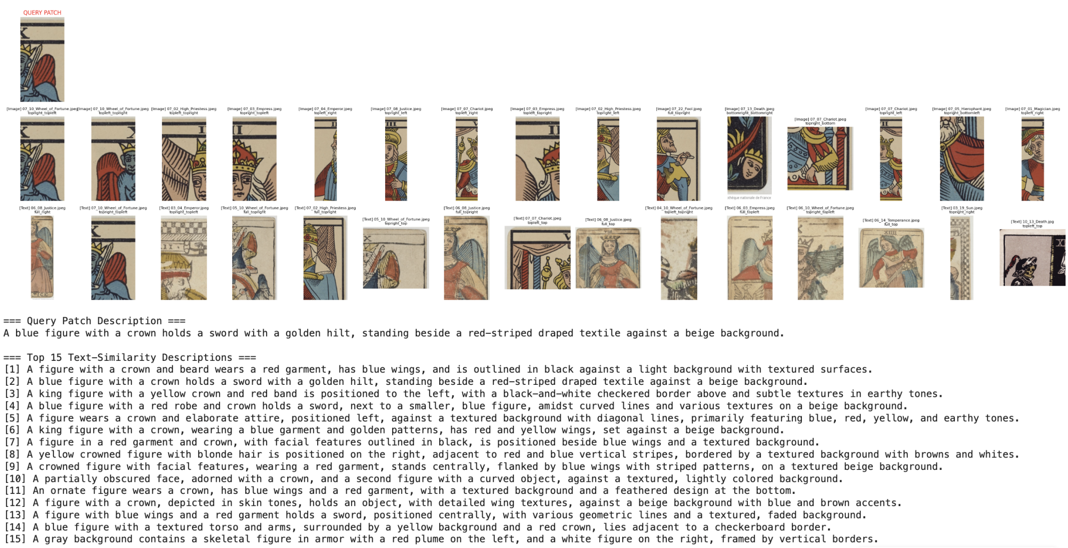

Figure 10. Prompt of the regions Example

Then we want to see how it works, as you can see the top left is our query image, the first row is the top15 similar image patch by ranking their cosine similarity of image embedding. And the second row is the top15 similar image patch and the textual description of those by ranking their cosine similarity of description embedding.

As you can see the image embedding can capture the style of the cards, the patch from same set are more similar to the query patch.

However, we want to see more interesting cross deck position similarity, and this can be shown as the similarity to the description(the second row, the corresponding image is for showing purpose, the textual description is the one being calculated). As you can observe, the wings, crown, and sword are well connected to the most similar images here.

And we obtain the following two patches which are from same position, same region but different decks, and you can see that they are from deck 5 and 6, and the query patch, which are the before and after periods.

Figure 11. Same Position and Same Region (Finding by image-text similarity)

There are other two interesting results here are more interesting pattern to be found, the shape of the red wing from the query image is alike the blue curtain behind in the justice card. And later cards will more visuaily showcase its the curtain instead of wings.

And the lady in the temperance who is pouring the water, looking like pointing a sword.

Figure 12. Example Patches for interesting patterns (Finding by image-text similarity)

3.4 Triplet Dataset Construction

So, we want to both have the image embedding's style and the description embedding's morphological information, so we decide to pull the patch image-embedding toward the closest descriptions.

And we will train a triplet loss for that.

First we have to generate a structured triplet dataset, where without human extra annotation, so we have to carefully design a way to form:

- Positive samples included patches from the same spatial positions across different decks (deck-aligned positives) and patches with topX textual embedding similarity. (X is depending on the number of the previous patches are, the non-repeated patch will be added until the total number is 20)

- Negative samples were patches identified as dissimilar based on textual embedding distance and spatial positional difference, ensuring strong discriminative power, which are those outside of positive group.

We seperated our data into training and validation dataset (80:20).

3.5 Model Training

Multi-negative Triplet Loss (Cosine Distance) Given a batch of B anchor-positive pairs (a_i, p_i), for each anchor a_i we select k hardest negatives $n_{i,j}$ $(j = 1, \ldots, k)$ within the batch based on cosine similarity. Our training objective minimizes the following multi-negative triplet loss:

$$\mathcal{L} = \frac{1}{B \cdot k} \sum_{i=1}^{B} \sum_{j=1}^{k} \max\left(0,\ 1 - \cos\left(f(a_i), f(p_i)\right) - \left[1 - \cos\left(f(a_i), f(n_{i,j})\right)\right] + m\right) \tag{1}$$

where $f(\cdot)$ denotes the normalized embedding function, $\cos(\cdot, \cdot)$ is the cosine similarity, and m is the margin. This loss encourages each anchor to be closer to its positive than to any of the k hardest negatives by at least the margin m.

- **Batch size:** 64
- **Epochs:** 10
- **Hard negatives per anchor:** $k = 20$
- **Triplet loss margin:** $m = 0.5$ (The margin controls the minimum required distance between positive and negative pairs in the embedding space. I set $m = 0.5$ to make it little harder on this task, if it is 0.2, the loss beginning with about 0.25, if it is 0.5, the loss beginning with about 0.49)
- **Optimizer:** Adam, learning rate 1×10^{-4}

Finally, we decreased average loss of last 10 from 0.49 to 0.1517

3.6 Visualization and "Evaluation"

Embeddings of test patches were evaluated by:

- Applying UMAP dimensionality reduction to visualize embedding distribution before and after training.
- Conducting DBSCAN clustering to objectively assess embedding quality, identifying visually coherent groups.

After applying UMAP we found the previous patch image-only embedding is too close to each other without clear cluster, but later after pulling those images away by textual similarity trained by triplet, we obtained the right one, which is more spread and could potentially be considered more pattern to be found in the space.

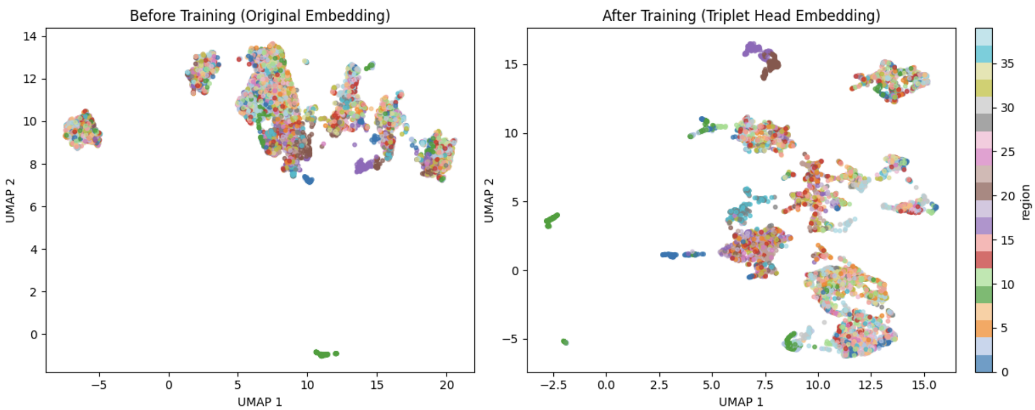

Figure 13. Visualization

To dive it deeper, we apply DBSCAN to cluster and show the top20 images closest to the centroid.

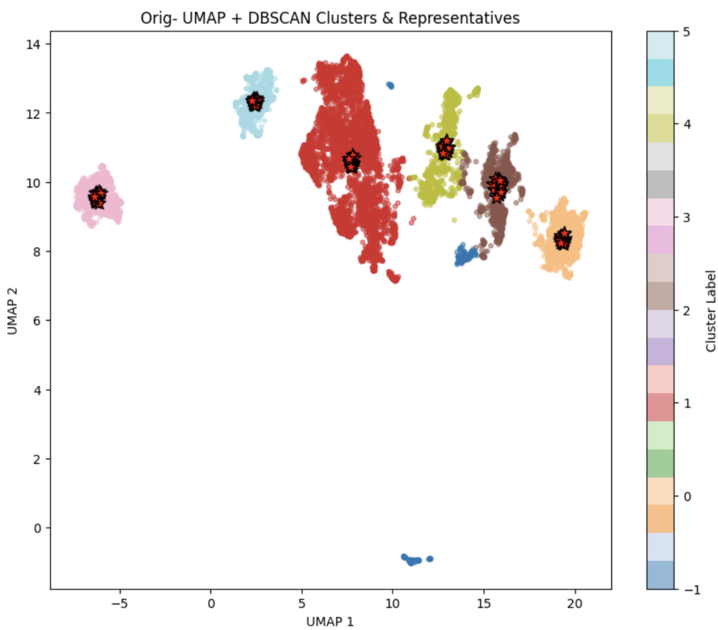

Figure 14. Visualization (before training)

Figure 15

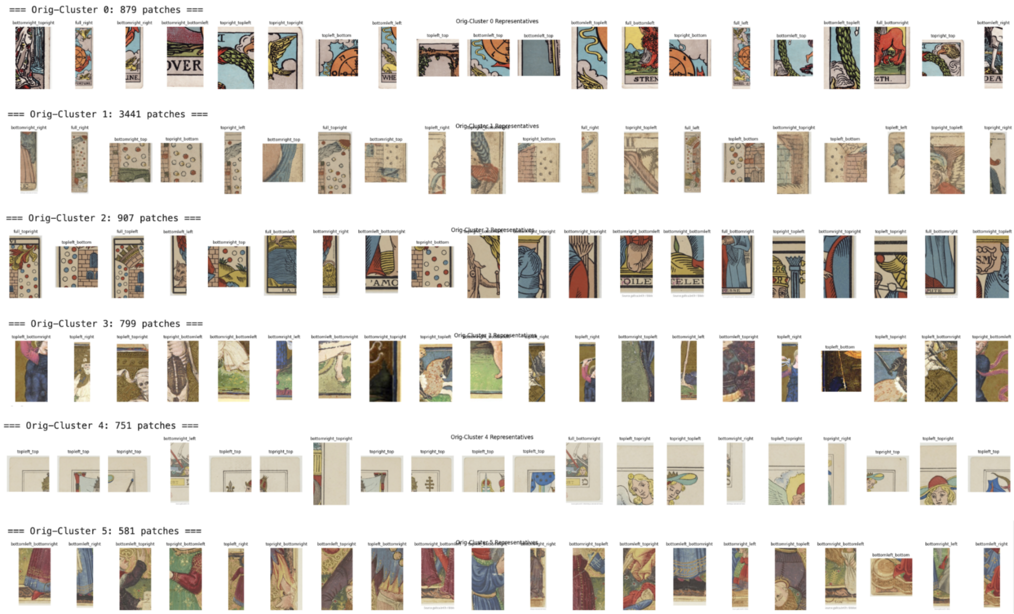

Figure 15. Top20 for each cluster (before training)

Figure 16

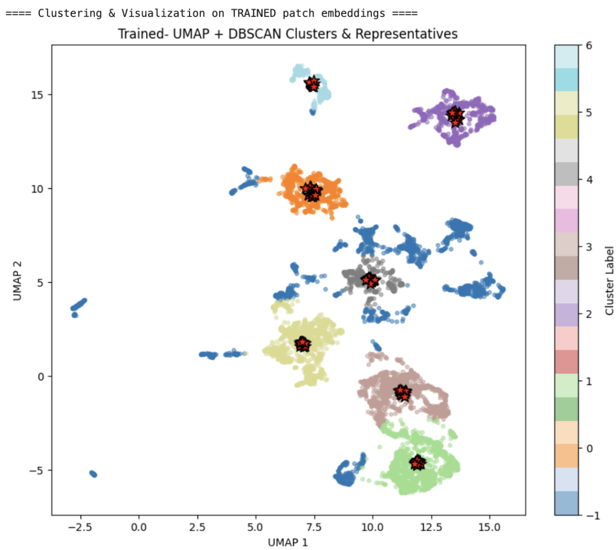

Figure 16. Visualization (after training)

Figure 17

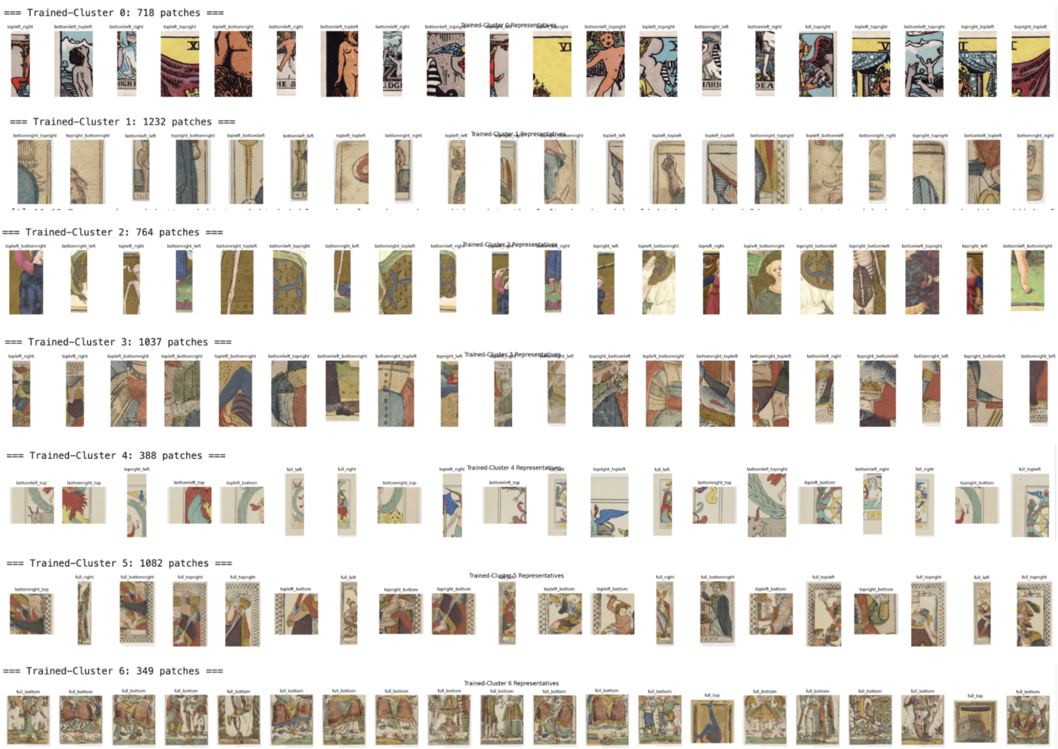

Figure 17. Top20 for each cluster (after training)

Before training, if we only rely on the image embedding, the patch in the clusters are gathering by the style which is exactly from the same deck (top 20 shown here) and hard to interpret.

After training, it has more semantic meaningful to us when looking at those clusters, for example the cluster 6 are "one leg up one leg down" pattern, and we can capture it cross the deck all as shown here.

Figure 18

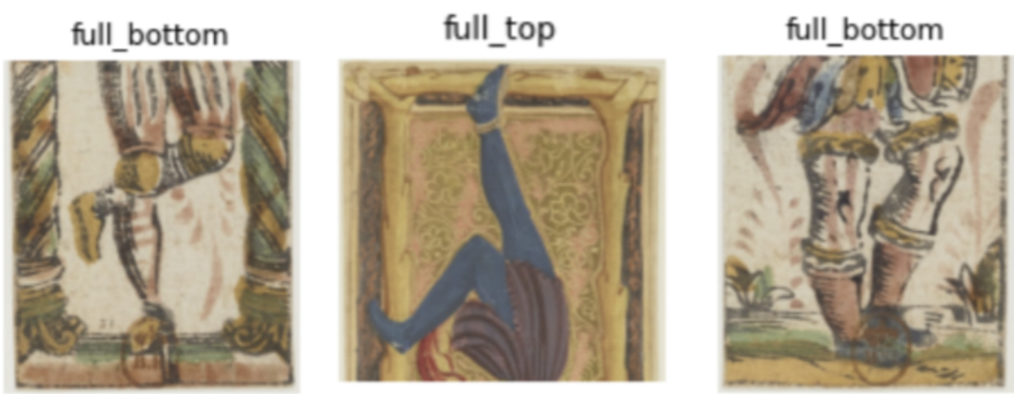

Figure 18. Pattern from Cluster 6

3.7 *Image Retrieval*

To evaluate the quality of the learned patch embeddings, we conducted a similarity-based retrieval experiment on the test set. All patch embeddings were first extracted from the test set using the trained embedding head. For each retrieval, a random patch was selected as the query, and its embedding vector was compared against all other test patch embeddings using cosine similarity

To ensure the diversity of the retrieved results, we excluded patches originating from the same source image as the query. The remaining patches were ranked by similarity, and the top 10 most similar patches, each from a different image, were selected for retrieval.

From qualitative point of view we can observe it capture a particular arm posture. as shown in Figure 19

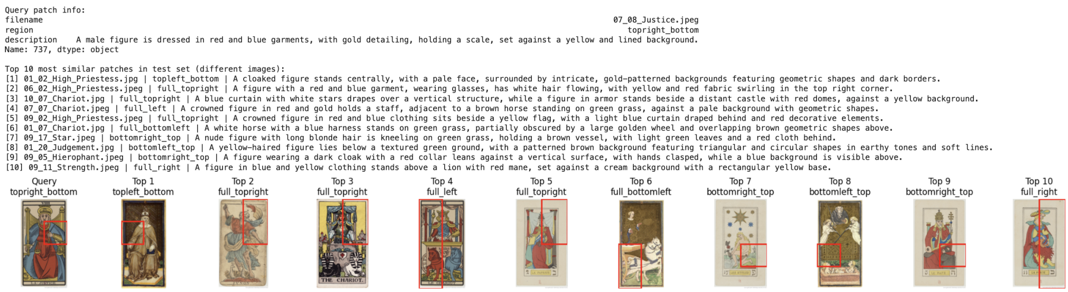

Figure 19. Image Retrieve Example

4. Limitation and Future Work

Our current dataset lacks important historical developments in tarot, particularly the inclusion of the Etteilla deck and the wave of renewed interest that followed the publication of the Waite-Smith deck. Notably, the spiritual revival of the 1960s, influenced by Eastern mysticism, brought about decks such as the J.J. Tarot. Incorporating these additional decks would provide a more comprehensive temporal and stylistic coverage of tarot's evolution.

The current prompt used for generating patch descriptions tends to produce information that, while detailed, often overemphasizes background or stylistic elements that are less relevant for morphological comparison. Refining the prompt to focus more on compositional and structural features, and less on general stylistic aspects, could yield descriptions that better serve computational analysis.

The present positive and negative grouping for metric learning does not leverage higher-level tarot knowledge. Concepts such as the "Fool's Journey"—where the Fool card (unnumbered and distinct from other major arcana) represents the archetypal journey from innocence to enlightenment—are not yet incorporated. Future work should encode such narrative structures to inform semantic grouping and hard negative mining. For example, viewing the remaining major arcana as sequential nodes in the Fool's progression could provide a more meaningful basis for defining relationships between cards.

The current cropping strategy is based on fixed grids, which may not always capture the most salient or aesthetically meaningful regions of the cards. Future work should explore data-driven or aesthetic-based cropping algorithms to automatically identify and extract visually significant patches.

Presently, clustering and retrieval performance are evaluated primarily through qualitative inspection for semester project. To benchmark our system, we plan to introduce quantitative metrics, such as cluster purity, silhouette score, and retrieval precision/recall, possibly with expert-annotated ground truth for evaluation.

Finally, an important future direction is the implementation of an interactive search engine that enables users to explore tarot patch similarities, query by image or text, and visualize relationships discovered by the model. Such a tool would facilitate broader scholarly and creative engagement with the corpus.

5. Conclusion

In this study, we demonstrate a novel framework for computationally analyzing the visual and symbolic patterns of tarot cards across several centuries of deck evolution. By combining systematic region-based image cropping, large language model-generated patch descriptions, and a unified embedding space via CLIP, we enable cross-modal and cross-deck motif discovery that transcends mere stylistic similarity. The use of triplet loss to align visual and morphological textual cues allows for more meaningful retrieval and clustering, revealing both shared and divergent iconographic traditions.

Our results show that this approach can identify nuanced, morphologically coherent patterns that would be challenging to detect manually or by style-based methods alone. While our focus has been on qualitative analysis and proof of concept, future work will expand the dataset, refine annotation prompts, and incorporate higher-level tarot knowledge—such as the narrative arc of the Major Arcana—for even deeper analysis. Furthermore, developing interactive tools for image and text-based exploration will open new possibilities for scholarly research and creative engagement. We believe that these computational methods offer promising directions for both digital art history and the broader field of symbolic visual analysis.

References

Court de Gébelin, Antoine. 1781. *Le monde primitif, analysé et comparé avec le monde moderne.* Vol. 8. chez l'auteur.

Dummett, Michael. 1980. *The game of tarot: from ferrara to salt lake city.* Duckworth.

Dummett, Michael A. E. 1986. *The visconti-sforza tarot cards* [in English]. Paperback, 141 pages. New York, NY: George Braziller. ISBN: 9780807611418.

Farley, Helen. 2009. *A cultural history of tarot: from entertainment to esotericism.* I.B. Tauris.

Greer, Mary K. 1984. *Tarot for your self: a workbook for personal transformation.* Red Wheel/Weiser.

Huson, Paul. 2004. *Mystical origins of the tarot: from ancient roots to modern usage.* Destiny Books.

Jung, Carl Gustav. 1934. *Archetypes and the collective unconscious.* Princeton University Press.

Place, Robert M. 2005. *The tarot: history, symbolism, and divination.* Tarcher/Penguin.

Pollack, Rachel. 1997. *Seventy-eight degrees of wisdom: a book of tarot.* Weiser Books.

Schaerf, Marco. 2022. Semi-supervised clustering of visual signatures of artworks. Master's thesis, École Polytechnique Fédérale de Lausanne (EPFL).