ArtStyleNet: Enhancing Artistic Style Similarity through Deep Learning

Introduction:

Art has been a fundamental component of human culture, serving as a medium for expression, storytelling, and societal reflection. From classical paintings to contemporary installations, art captures the essence of human creativity and emotion. The study of art encompasses a wide range of disciplines, from historical analysis to aesthetic evaluation, each aiming to uncover the meaning and significance behind artistic works.

The art work of one artist can be little bit similar to another artists work in a degree to which artworks share visual, structural, or conceptual attributes. These attributes can include elements like color schemes, brushstroke patterns, composition, and texture. Traditionally, artistic similarity is analyzed through subjective judgments by art experts, focusing on theoretical classifications such as movements or periods. However, for the general audience, the perception of similarity often arises from an intuitive "feeling" of connection between artworks. This project aims to quantify these intuitive judgments and analyze them systematically.

ArtStyleNet is an exploration into the realm of artistic styles, aiming to uncover hidden patterns and similarities that go beyond traditional art theory. Unlike established definitions in art history, this project leverages deep learning to investigate whether computational methods can mimic the human perception of style similarity—the intuitive "feeling" that connects various artworks. By taking a computational approach, ArtStyleNet seeks to bridge the gap between theoretical knowledge and practical application, enabling broader accessibility and engagement with art. Traditionally, only people who are knowledgeable about art theory have been able to appreciate art, allowing casual viewers to rely on their own subjective perceptions.

However, the growing intersection of technology and the arts offers an opportunity to democratize the understanding of artistic styles. With the development of deep learning, it becomes possible to quantify and analyze visual elements in ways that were previously unimaginable. This project aims to explore whether artificial intelligence can replicate and even enhance the intuitive judgments of style similarity made by human viewers. By identifying hidden styles through data-driven methodologies, ArtStyleNet opens new avenues for art recommendation systems and broader appreciation among diverse audiences.

Deep learning techniques, such as convolutional neural networks (CNNs), excel at identifying intricate patterns in visual data, making them particularly suitable for analyzing paintings. By clustering these features, machine learning can uncover hidden patterns and group artworks based on shared characteristics, providing novel insights into artistic styles and facilitating applications such as personalized art recommendations.

The recognition and classification of artistic styles have historically been grounded in subjective interpretations by art experts. While these interpretations are invaluable, they are inaccessible to most art viewers who lack deep theoretical knowledge. Also traditional methods struggle to identify nuanced, hidden styles that might not conform to established categorizations. Factors such as individual biases, limited exposure to diverse artworks, and the sheer complexity of artistic elements often hinder comprehensive analysis. These challenges underscore the need for a systematic approach that can complement and enhance traditional methods.

ArtStyleNet proposes a data-driven approach, utilizing deep learning to analyze and cluster paintings based on visual similarities perceived by general audiences

Related work

Numerous studies have explored artistic style classification using machine learning. Convolutional Neural Networks (CNNs), such as ResNet, have been widely employed for feature extraction due to their ability to capture intricate visual patterns. For instance, (Gatys et al, 2016) utilized CNNs to transfer artistic styles between images, demonstrating the versatility of neural networks in capturing stylistic elements. Dimension reduction techniques like PCA and t-SNE have been used to simplify high-dimensional feature spaces, allowing for more interpretable visualizations of data (Van Der Maaten & Hinton, 2008). Meanwhile, clustering algorithms such as K-means and LDA have been instrumental in identifying latent groupings in datasets, as shown in the work of (Achlioptas et al, 2002).

In addition to these methods, numerous studies have directly tackled the challenge of detecting artistic similarity. For example, (Saleh and Elgammal, 2015) proposed a method for visual style analysis by using computer vision techniques to quantify the stylistic features in artworks. Their work demonstrated the potential of machine learning models in detecting similarities and differences in artistic styles across centuries. Similarly, (Karayev et al, 2014) developed an approach to predict and classify artistic styles using supervised learning models trained on features extracted from artworks, providing a baseline for style recognition tasks.

Recent advancements have focused on combining feature extraction with novel clustering approaches. (Tan et al, 2020) proposed an unsupervised framework that integrates self-supervised learning with clustering to better capture nuanced visual similarities in artistic datasets. Additionally, (Elgammal et al, 2018) introduced "creative adversarial networks" (CANs) to distinguish between traditional artistic styles and those that deviate from historical norms, offering a mechanism for evaluating both stylistic similarity and novelty. Such approaches emphasize the growing role of AI not only in understanding artistic styles but also in pushing the boundaries of creative expression.

ArtStyleNet builds on these methodologies to uncover and visualize latent styles, extending the scope of artistic analysis through innovative use of deep learning frameworks.

Method

Analyzing artistic style similarity requires a systematic and multi-step approach to extract meaningful patterns and insights. The methodology adopted in ArtStyleNet leverages advanced deep learning and machine learning techniques, which include feature extraction, dimensionality reduction, clustering, and visualization. Each step plays a pivotal role in transforming raw image data into interpretable clusters of artistic styles, enabling a deeper understanding of the hidden patterns in the dataset.

Step 1: Feature Extraction with ResNet

ResNet (Residual Network) is employed as the foundation for feature extraction due to its robust architecture and proven performance in image analysis tasks. ResNet leverages skip connections to mitigate the vanishing gradient problem, allowing for the effective training of very deep networks. In this project, a pre-trained ResNet model is used to extract high-dimensional features from the paintings. These features encapsulate critical aspects such as color composition, brushstroke patterns, and texture, which are essential for understanding artistic styles. The extracted features serve as the input for subsequent steps, forming the basis for dimensionality reduction and clustering.

ResNet (Residual Network) is utilized to extract visual features from paintings. By leveraging pre-trained weights on large image datasets, ResNet captures detailed attributes such as color composition, texture, and brushstroke patterns. These extracted features serve as a high-dimensional representation of each artwork.

Step 2: Dimension Reduction with PCA

Once features are extracted, the next step involves reducing their dimensionality to make the data more manageable and interpretable. Principal Component Analysis (PCA) is employed for this purpose. PCA identifies the directions (principal components) along which the data varies the most, and projects the high-dimensional data onto these components. By retaining only the most significant components, PCA helps reduce computational complexity and ensures that the resulting feature set highlights the most critical visual patterns. This step is crucial for preparing the data for clustering while minimizing noise and redundancies.

Principal Component Analysis (PCA) is applied to reduce the dimensionality of features extracted by ResNet. This step mitigates computational complexity and emphasizes the most significant visual patterns. By retaining only the principal components, PCA ensures that the data remains interpretable while discarding noise and redundancies.

Step 3: Clustering with LDA

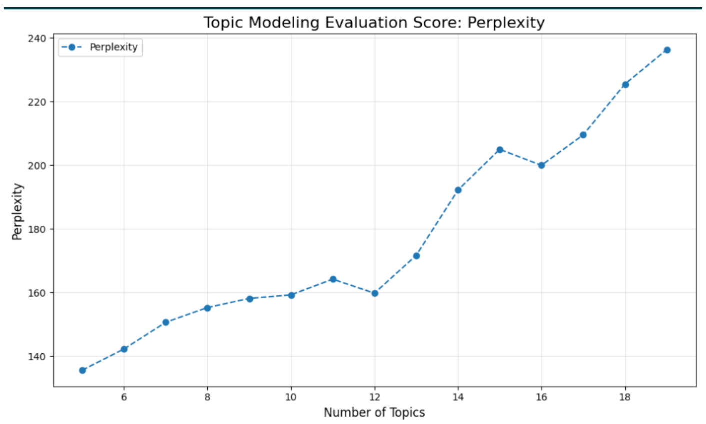
Latent Dirichlet Allocation (LDA) is used to cluster the paintings into distinct style groups. Unlike traditional clustering methods that assign each item to a single cluster, LDA models each painting as a mixture of multiple clusters, reflecting the complex and multifaceted nature of artistic expression. The range of clusters explored is between 5 and 20, balancing dataset size and cluster granularity. Evaluation metrics include:

- **Perplexity Score:** Used to measure how well the clustering fits the dataset. Lower scores indicate better fit.
- **Coherence Score:** Not utilized in this project, as it is more suitable for text-based similarity rather than image clustering.
- **Elbow Point Method:** The best number of clusters is identified by visualizing the changes in perplexity scores and locating the elbow point, which represents the optimal trade-off between model complexity and performance.

The identified clusters represent latent artistic styles, which are then analyzed and interpreted to provide meaningful insights into the dataset.

Latent Dirichlet Allocation (LDA) is used to cluster the paintings into distinct style groups. Unlike traditional clustering methods that assign each item to a single cluster, LDA models each painting as a mixture of multiple clusters, reflecting the complex and multifaceted nature of artistic expression. The optimal number of clusters is determined by evaluating coherence and perplexity scores. The range of clusters is explored between 5 and 10 due to the relatively small size of the dataset (51 painters). The identified clusters represent latent artistic styles, which are then analyzed and interpreted to provide meaningful insights into the dataset.

Latent Dirichlet Allocation (LDA) is employed to cluster paintings into hidden style groups. Unlike traditional clustering methods, LDA treats each painting as a mixture of multiple styles, reflecting the multifaceted nature of artistic expression. The optimal number of clusters is determined by evaluating coherence and perplexity scores within a range of 5 to 10 clusters, given the dataset's relatively small size.

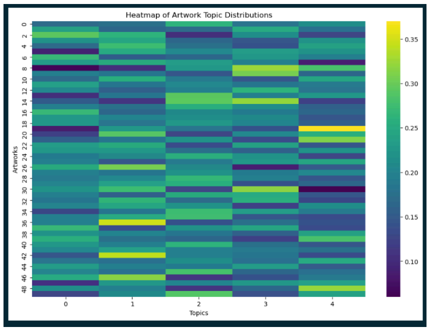


Step 4: Visualization with HeatMap

The final step in the methodology involves visualizing the clustering results using heatmaps. The heatmap visualizes the topic distributions of the first 50 artworks, where each row represents an artwork and each column corresponds to a specific topic. The intensity of color indicates the proportion of a topic in a given artwork:

- **Yellow areas** signify a higher contribution of a topic.
- **Darker areas** indicate a lower contribution of a topic.

Insights from the heatmap include the observation that topics are evenly distributed across most artworks, with notable spikes in certain topics for specific pieces. This suggests variability in stylistic influences across the dataset.



By analyzing these heatmaps, researchers and art enthusiasts can gain valuable insights into the hidden relationships between paintings, artists, and styles, paving the way for more informed discussions and applications in art analysis**.**

Dataset

Source

The dataset originates from the Monthly Dacon Artist Classification AI Competition, a prominent platform for applying artificial intelligence in creative and artistic domains. The competition provided a rich dataset that serves as a benchmark for exploring the capabilities of machine learning in art analysis.

The dataset comprises a diverse collection of artworks from 51 painters, totaling 5,910 paintings. Each painting is accompanied by metadata that includes information such as the painter's name, birth and death years, genres, and nationalities. However, this project focuses exclusively on analyzing the painting images to uncover stylistic patterns. The following key details highlight the structure and scope of the dataset:

- **Training Set Only is used:** The training set is utilized for this project to ensure consistent and complete data. The evaluation dataset, which contains cropped paintings, is excluded as it does not align with the objective of analyzing complete artworks.
- **Number of Artists:** The dataset includes works from 51 painters, providing a comprehensive sample of varied artistic styles.
- **Number of Paintings:** With 5,910 artworks, the dataset offers sufficient depth for analyzing trends and patterns across different styles and artists.
- **Metadata:** While metadata such as genres and nationalities is available, this project prioritizes visual features over textual data, focusing solely on the images of paintings.
- **Focus:** Painting images only.

Experiment

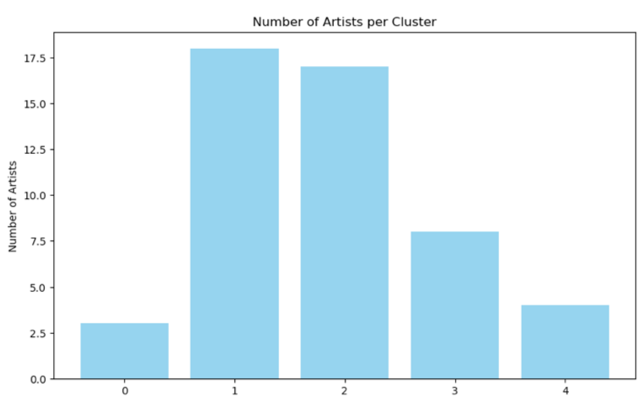
The dataset is processed through the methodology outlined above, with a focus on feature extraction, dimension reduction, clustering, and visualization. The following analyses are conducted:

1. Identification of the elbow point for optimal clustering.
2. Generation of heatmaps showing style probabilities for each painter and painting.
3. Compilation of the top 10 painters and paintings for each style cluster, based on the highest probabilities.

Results

- **Elbow Point Exploration Results:** To be documented.
- **HeatMap Visualization:** To be documented.
- **Top Painters and Paintings:** To be documented.

Top 5 artworks per cluster



Number of Artist per each cluster

Limitations

PCA assumes linear relationships between high-dimensional and low-dimensional spaces. Future work could explore nonlinear techniques, such as AutoEncoders, to capture complex patterns more effectively.

This project does not segment visual elements like color, texture, and composition for separate analysis. Incorporating metadata, such as genres and artist backgrounds, could provide richer insights into styles aligned with art theory.

Conclusion

ArtStyleNet demonstrates the potential of leveraging deep learning techniques to explore artistic style similarities in paintings. By employing a structured methodology that includes feature extraction, dimensionality reduction, clustering, and visualization, the project uncovers hidden stylistic patterns that extend beyond traditional art theory. The integration of methods such as ResNet, PCA, and LDA enables a nuanced analysis of artistic expressions, highlighting the diverse influences and unique characteristics within the dataset.

Through this exploration, ArtStyleNet bridges the gap between art enthusiasts and computational methods, making the appreciation and understanding of art more accessible. The insights gained from this project provide a foundation for future advancements, such as personalized art recommendations and enhanced artistic style discovery tools. However, challenges such as the use of linear reduction techniques and simplified feature extraction underscore the need for further refinement and innovation in this domain.

By continuing to innovate and address these limitations, ArtStyleNet and similar projects hold the promise of reshaping how art is studied, appreciated, and analyzed in the digital age.

References

- Saleh, B., & Elgammal, A. (2015). Large-scale classification of fine-art paintings: Learning the right metric on the right feature. arXiv preprint arXiv:1505.00855.
- Karayev, S., Hertzmann, A., Winnemoeller, H., Agarwala, A., & Darrell, T. (2014). Recognizing image style. In Proceedings of the British Machine Vision Conference (BMVC) (pp. 1-12).
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2414-2423).
- Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. Journal of Machine Learning Research, 9(Nov), 2579-2605.
- Achlioptas, D., McSherry, F., & Schoelkopf, B. (2002). Sampling techniques for kernel methods. In Advances in Neural Information Processing Systems (pp. 335-342).
- Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2018). Creative adversarial networks: Generating "art" by learning about styles and deviating from style norms. In Proceedings of the International Conference on Computational Creativity.
- Tan, C., Yu, Z., & Huang, J. (2020). Self-supervised learning for clustering-based analysis of visual patterns. In IEEE Transactions on Image Processing, 29, 7895-7906.