https://doi.org/10.2352/EI.2025.37.11.HVEI-216

Computational tools for analyses of color of costumes in large corpora of fine art paintings

***Christine Li*[1] *and David G. Stork*[2]**

[1] ***Department of Computer Science, Columbia University, New York NY 10027***
[2] ***Adjunct Professor, Electrical Engineering, Material Science and Engineering, and Program in Symbolic Systems Stanford University, Stanford CA 94305***

Abstract

Clothing is a lens through which a society expresses its culture and history. Its stylized portrayal in painting adds an immensely rich layer of cultural self introspection—how artists see themselves and their contemporaries, expressed through art. Particularly of interest in this study is color: how has color in costumes in portraiture painting changed over time, across art styles, and for different genders? In this study, we apply computational methods drawn from computer vision, machine learning, economics, and statistics to a large corpora of over 12k portrait paintings to analyze trends in color in Western art over the past 600 years. For each painting, we obtained clothing segmentation masks using a fine-tuned SegFormer model, performed gender classification using CLIP (Contrastive Language-Image Pre-Training), extracted dominant colors via clustering analysis, and computed Color Contrast Index (CI) and Diversity Index (DI). This study is, to our knowledge, the most comprehensive, large-scale analysis of colors of clothing in paintings. We share our methodology to make more widely accessible state-of-the-art computational tools for scholars studying the history and development of style in fine art paintings. Our tools empower analyses of major trends in costume colors as well as specialized domain-specific searches throughout databases of tens of thousands of paintings—far larger than can be efficiently analyzed without computer methods. These tools can reveal comparisons between different painters and trends within particular artists' careers. Our tools could be enhanced to enable refined analyses, for instance on the social status of the portrait subject, and other visual criteria.

Introduction

In the past few decades, computer vision, machine learning, pattern recognition, and artificial intelligence have been applied to an expanding range of problems in the history and interpretation of fine art, primarily two-dimensional art such as paintings and drawings. Such so-called computer-assisted connoisseurship has enriched traditional analysis methods—for instance the analysis of brushstrokes,[1] color and style,[2] portrait pose,[3, 4] lighting in tableaus,[5, 6] composition in landscape paintings,[7] and much more. Computer methods provided definitive evidence refuting the claims that some Western artists as early as Jan van Eyck (c. 1395–1441) secretly built complicated optical projectors, projected images onto their canvases or other supports, traced these images.[8, 9] Computer methods are increasingly being incorporated into otherwise traditional academic courses and scholarship, a trend that shows no signs of abating.[10]

Fine art paintings have served as an important source of information about the historical development of fashion and costume.[11] For example, the Dutch Golden Age was a particularly rich period for such analysis, given the significant changes in social, political, and especially religious concerns of the burgeoning merchant patrons that were captured in the expanding art market.[12, 13, 14, 15] Analyses of costumes in portrait and genre paintings from this historical period give insight into the costume choices in later periods.[16]

It is natural that such studies should be aided by computer-based image analysis tools. Such tools would empower studies of trends in corpora of thousands or tens of thousands of paintings over centuries—far greater than can be performed "by eye," that is, without automation. Moreover, computer-based methods can empower more refined analyses, for instance ones that include the sitters' gender,[17] social status, and material studies (matte, glossy, etc.), as is informative when estimating the type of fabric (wool, silk, etc.).[18]

A prior effort developed automated tools for analysis of costume color in portrait paintings.[19] Our work builds upon and extends this prior work and exploits state-of-the-art deep neural network methods for segmentation, provides richer analysis of color palette, and is tested on much larger corpora with finer categorization, based on metadata.

Methodology

Dataset

For our analyses, we used the *Painter by Numbers* dataset from Kaggle,[20] which contains images of over 12,926 portrait paintings (originally from `WikiArt.org`). The images are accompanied by metadata which include the date of execution, artist, and style or art movement of the painting. The dates range from 1425 to 2011, with the distribution shown in Fig. 1. We note that there are 2,881 paintings without a valid date annotation. There are 94 unique styles of the paintings represented in the dataset, and the 20 with the highest counts are presented in Fig. 2. The majority of paintings in the dataset are Western European. The paintings are authored by 1,020 unique artists in total, with the distribution for the most represented artists shown in Fig. 3.

Using this dataset of portraits, we complete a series of steps in pre-processing and feature extraction for the analysis on the clothing colors, outlined in 5. This includes (1) obtaining segmentation masks of clothing in the paintings, (2) extracting dominant colors via K-Means clustering, and (3-4) computing the Con-

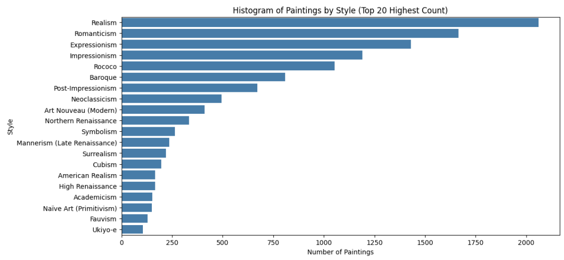

Figure 1: Histogram of Number of Works by the Top 20 Most-Represented Art Styles in the Dataset

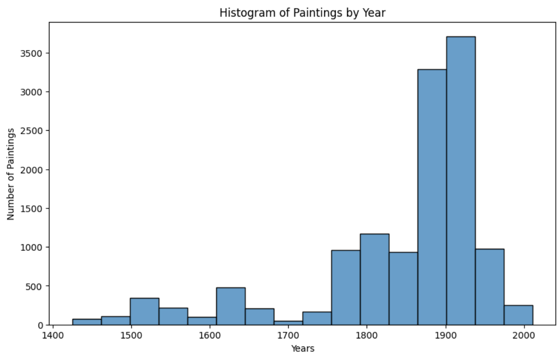

Figure 2: Distribution of Paintings by Year

trast Index and Diversity Index. Finally, we visualize and analysis these results, comparing the indices across different art styles, time periods, genders, and individual artists.

Semantic Segmentation of Clothing

First, we obtained segmentation masks for the portrait paintings using a fine-tuned SegFormer B2 model for clothing segmentation.[21] Sample segmentation masks are presented in Figs. 4 and 5.

Although we do not have ground truth for the segmentation masks, manual inspection of a random sample of segmentation results shows that the resulting masks are mostly accurate, particularly for styles that are more "realistic." We have manually filtered out masks that are clearly incorrect. While small inaccuracies in segmentation masks remain, we found that these do not detract from the data signal when taken in aggregate over the entirety of the corpora, as is shown in our later analyses.

Gender Classification of Portraits

Since the original dataset does not contain information on the gender of the portrait subjects, we used OpenAI's CLIP model for a zero-shot classification. For each portrait, we extract the joint embeddings for both the image and text prompts ("a painting of a male person," "a painting of a female person"). We then obtain the classification logits, which are converted to probabilities for the predicted gender for each portrait. Out of the total 12878 portraits paintings, 6867 (53.3%) are classified as female and 6011 (46.7%) are male.

Color Representation in CIELAB Space

As the initial step in our color analysis, we converted the pixel presentation of the paintings from the default RGB space to CIELAB space. Defined by the International Commission on Illumination (CIE), this color space uses three values: L^* represents lightness (0 to 100), a^* indicates the relative value of the red-green component of a color (-128 to 127, with positive for red and negative for green), and b^* is the yellow-blue component (-128 to 127, with positive for yellow and negative for blue). Intended to mimic the opponent color processing of the human visual system, the CIELAB representation is more suited for our numerical analysis since perceptual differences in colors can be approximated by taking the Euclidean distance between them in the CIELAB space. We visualize the distribution of raw pixels of clothing across all paintings via scatterplots of the a^* and b^* values of all pixels, grouped by half-century intervals and by gender. The figures are included in Appendix A.

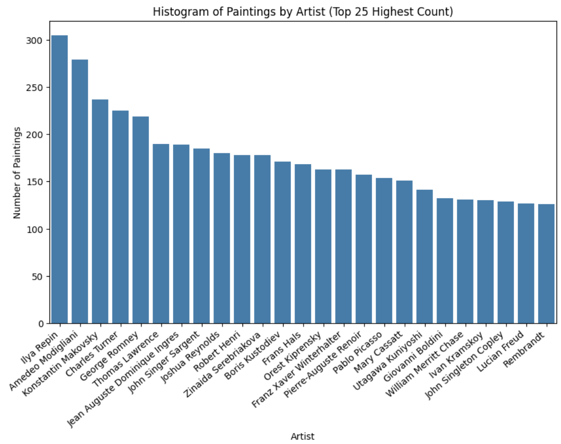

Figure 3: Histogram of Number of Works by the Top 20 Most-Represented Artists in the Dataset

Figure 4: Sample segmentation mask. Thomas Gainsborough, *John Montagu, 4th Earl of Sandwich*

Dominant Colors via Clustering

In order to better understand the overall distribution and patterns in color trends, we extracted the dominant colors in each portrait's clothing mask using K-Means clustering (with $k = 3$). Furthermore, to gain a more holistic picture of the interaction of colors both within individual paintings and across different paintings, we used key metrics to measure the color contrast and diversity. These metrics have previously been applied in the economic analyses of art in *Worth of Art*.[22] Here, we find a novel applica-

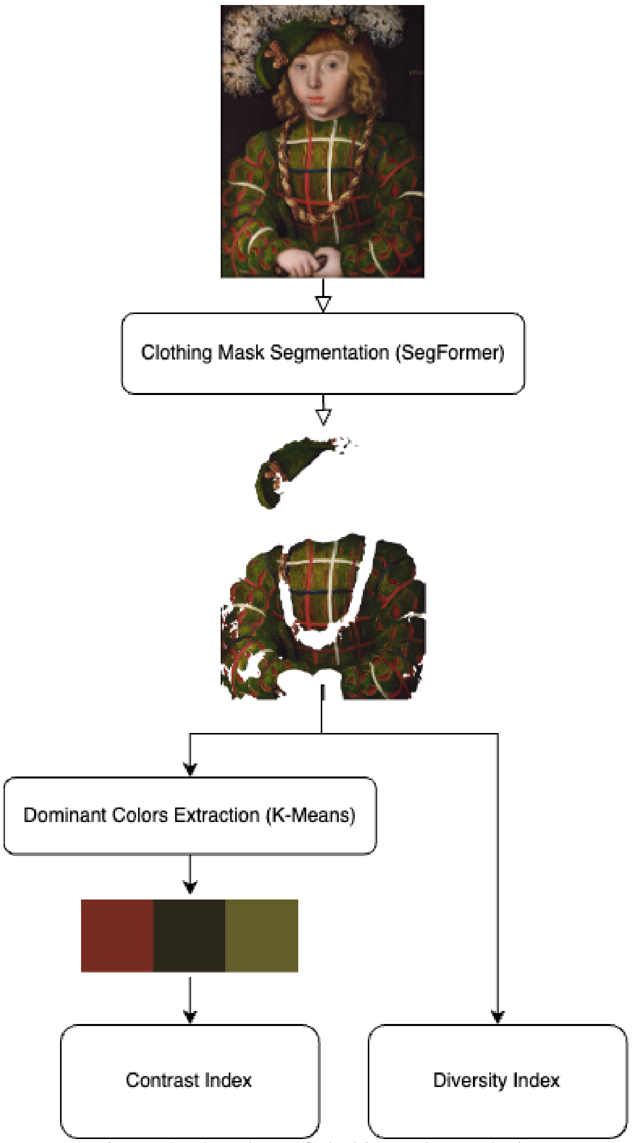

Figure 5: Flowchart of clothing color analysis.

tions for these metrics in the analysis of the portrayal of clothing from a meta-art historical perspective.

Color Contrast Index

The Color Contrast Index (CI) is defined as the log sum of the pairwise Euclidean distances between the centroids of three dominant color clusters of a painting, expressed in CIELAB space.[22] Roughly, this represents how perceptually different (far apart) the primary colors in the clothing are. The dominant colors are obtained via K-Means clustering (with $k = 3$) on all pixels in the segmented clothing mask, as described above. The centroids, C_1, C_2, C_3, are each represented as a 3-tuple (L, a^*, b^*) in CIELAB color space. Thus, the Contrast Index can be expressed as

$$\mathrm{CI} = \log(d(C_1,C_2) + d(C_1,C_3) + d(C_2,C_3)) \tag{1}$$

where each pairwise distance $d(C_i,C_j)$ is defined as

$$d(C_i,C_j) = \sqrt{(C_{i,L} - C_{j,L})^2 + (C_{i,a*} - C_{j,a*})^2 + (C_{i,b*} - C_{j,b*})^2} \tag{2}$$

Color Diversity Index

The Color Diversity Index (DI) captures, roughly speaking, how evenly distributed the pixel values are across L, a^*, and b^* ranges.[22] It is based on the normalized Herfindahl-Hirschman Index, which is a metric used in economics to measure market concentration and competition. This is adapted to analyse the "concentration" of colors. To calculate the Color Diversity Index (DI), we first partition the CIELAB evenly into 125 "bins" (5 bins for each axis) and assign each pixel of the clothing mask to its respective bin. For each bin i, we calculate σ_i, the percentage out of all pixels that fall into that bin,

$$\mathrm{DI} = \frac{1 - \sum_{i=1}^{125} \sigma_i^2}{1 - \frac{1}{125}} \tag{3}$$

A painting in which all pixels are the same color (or very similar colors that all fall into the same "bin") would have a DI of 0, indicating that the color distribution is not diversified. On the other hand, a perfectly diversified painting would have exactly 1/125 of its pixels in each bin and have a DI of 1.

CLIP Embeddings and TSNE Visualization

Additionally, we obtained CLIP (Contrastive Language-Image Pre-Training [23]) image encodings for all portraits, for both the full painting as well as the segmented clothing mask. Then, we visualized the high-dimensional (512) latent space representation by first using Principle Component Analysis for dimensionality reduction, then further projecting the resulting PCs to 2 dimensions using t-SNE (t-Distributed Stochastic Neighbor Embedding).[24] The resulting figure is included in Appendix B.

Results

Using the methods described above, we computed the dominant colors, Contrast Index (CI) and Diversity Index (DI) for all segmented clothing masks. We then visualize and analyze the distribution of the resulting indices and compare them across different art styles, half-century intervals, genders, and individual artists, with the figures and highlighted results presented in this section. We also include representative works from each of these analyses for better visualization.

First, we visualize the overall distribution of the Contrast Index and Diversity Index in Fig. 6 left and right, respectively. We note that distribution of CI across all paintings is roughly unimodal, left-skewed, with values ranging from 1.69 to 5.73 with a mean of 4.71. Similarly, the distribution of DI is left skewed, though with a smaller "peak" closer to the left tail, with values ranging from 0.000264 to 0.950915 and mean of 0.61.

We visualize the comparison of the Contrast Index across different art styles (with at least 30 examples in the dataset), ordered by increasing median value of CI, and we present the result in Fig. 7. We note that Symbolism (with 240 samples) has the lowest CI with a median of 4.54. Similarly, Social Realism

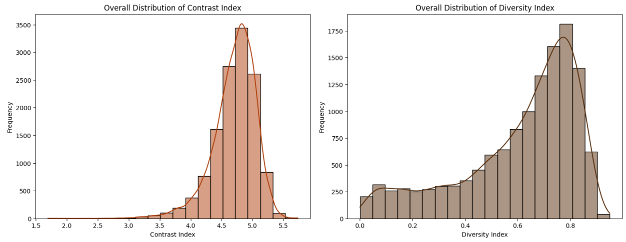

Figure 6: Overall Distributions of Contrast Index (left) and Diversity Index (right).

and Impressionism also rank low on CI. On the other hand, Pop Art (105 samples) has the highest CI with a median of 5.03, and Ukiyo-e and Fauvism trail just behind. The most representative paintings for each of these styles are displayed in Fig. 8 9 and for comparison.

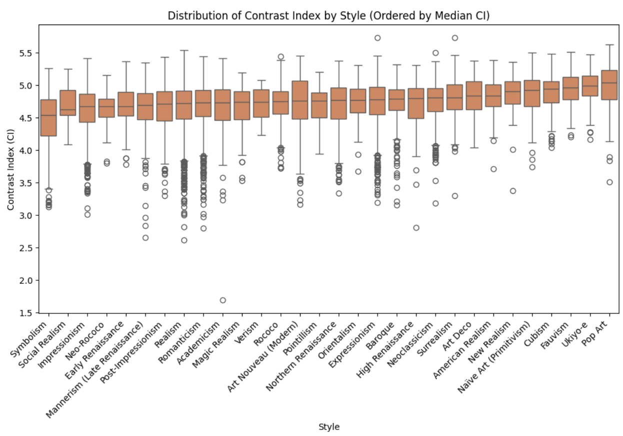

Figure 7: Box-plots of Contrast Index, Grouped by Style (in ascending order by median)

Figure 8: Representative masks from Symbolism (left), Social Realism (middle), and Impressionism (right) styles.

A similar box-whisker plot visualization of the Diversity Indices across various styles is displayed in Fig. 11. Here, Mannerism (230 samples) has the lowest DI with a median of 0.58, and Fauvism has the highest DI with a median of 0.78. We note that for DI in particular, there is a general increase in DI values for styles over time, with a notable exception of Early Renaissance having a relatively high DI (0.74 median). Representative examples are shown in Fig. 10

Observing the distributions of CI across half-century intervals (in temporal order) in Fig. 12, we note that there is a somewhat cyclic pattern in the median values of CI over time. The spread increases, with the exception of the last interval (which

Figure 9: Representative masks from Pop Art (left), Ukiyo-e (middle), and Fauvism (right) styles.

Figure 10: Representative masks from Mannerism (left), Fauvism (middle), and Early Renaissance (right) styles.

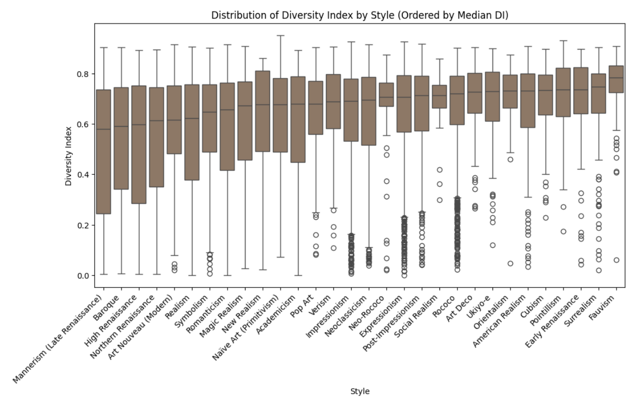

Figure 11: Box-plots of Diversity Index, Grouped by Style (in ascending order by median)

may be partly due to its relatively smaller sample size of 22 paintings).

Figure 14 shows the distributions of the Diversity Index of clothing colors over time. Here, we notice that the DI median roughly decreases from 1400 to 1699 (from 0.76 to 0.57 in 1650-1699), but jumps sharply in the interval from 1700-1749 back up to 0.75, after which it continues to decrease, then rebounds slowly after 1850. To visualize this with examples, we present the most "representative" portraits from each group (minimizes distance to median CI). Representative examples are shown in Fig. 13

Perhaps the most interesting analysis arises when comparing CI and DI values between genders, across time and art styles. In Figs. 16 and 18, we plot the distributions of CI and DI (respectively) across different styles, separately by gender. In Fig 16 for CI, we note that Art Nouveau has the largest difference in median CI, where the median CI for male-labeled portraits (124 samples) being 5.26 while for female-label portraits (178 samples) is 4.69. The style with the largest difference with the female CI being higher than male CI is High Renaissance (4.82 female and 4.78 male), followed by Mannerism and Northern Renaissance.

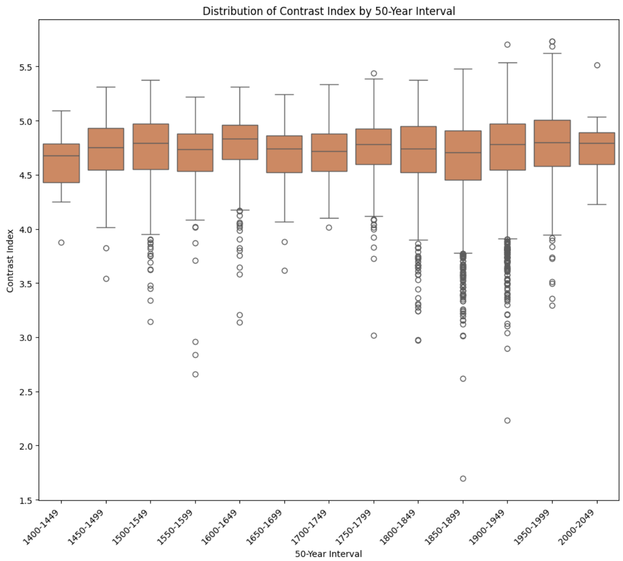

Figure 12: Box-plots of Contrast Index, Grouped by Half-Century Interval

Figure 13: Representative masks from 1440 (left), 1685 (middle), and 1739 (right) styles.

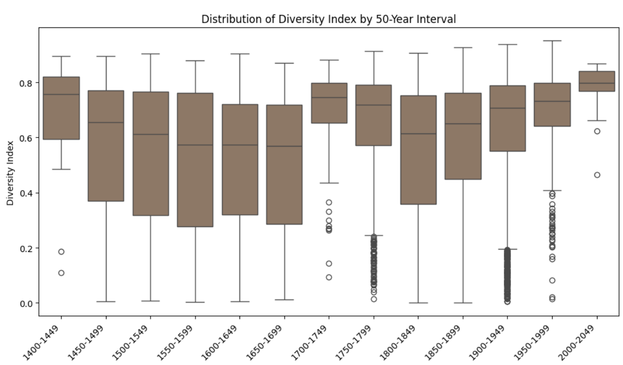

Figure 14: Box-plots of Diversity Index, Grouped by Half-Century Interval

Interestingly, Baroque has the smallest gender difference in CI, with the median male CI at 4.794575 and female CI at 4.796321. Represented portraits are presented in Fig. 15.

For the Diversity Index plot comparing styles by gender, we notice that the styles with lower DI also tend to have greater gender differences. New Realism has the greatest difference in DI between genders, with female-labeled paintings having a median DI of 0.769307 and male-labeled paintings being 0.458244. High Renaissance, Mannerism, Academism, and Northern Renaissance

Figure 15: From left to right: representative portrait of High Renaissance (female), High Renaissance (male), Baroque (female), Baroque (male)

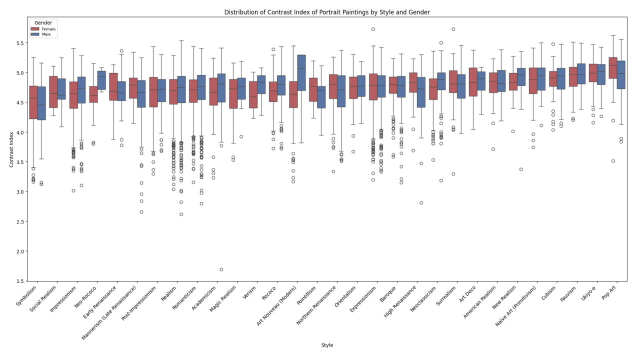

Figure 16: Box-plots of Contrast Index, Grouped by Gender and Style (in ascending order by median)

follow after with large differences in DI between genders.

With the exception of three styles—Expressionism, Surrealism, and Cloisonnism—for all other styles, the median DI of clothing in female portraits is higher than that of male portraits of the same style. We show representative female and male portraits from New Realism and Expressionsm styles to visualize the effects of varying DI values, presented in Fig. 17

Figure 17: From left to right: representative portrait of New Realism (female), New Realism (male), Expressionism (female), Expressionism (male)

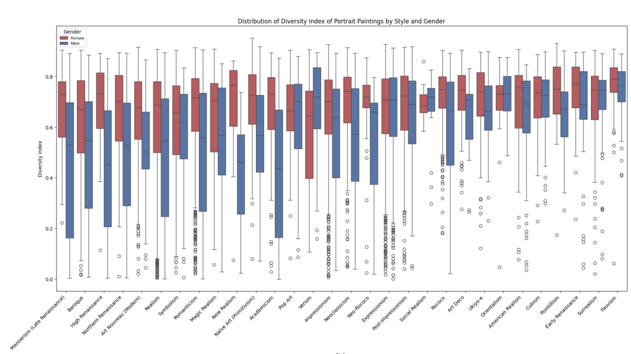

Figure 18: Box-plots of Diversity Index, Grouped by Gender and Style (in ascending order by median)

When we compare the distributions of Color Contrast Index for paintings over time (grouped by half-century intervals) by gender, as displayed in Fig. 22, we notice that from 1400 to 1599, the CI for female-labeled portraits have a higher median value than those of male portraits. In the eras thereafter, there are continual swaps between genders in terms of which has higher CI.

For a similar comparison of the Diversity Index in Fig. 23, we note that for every single time interval from 1400 to now, the DI for female portraits is always higher than that of male portraits—all except for 1950-1999. Interestingly, the magnitude of this gap varies over time, roughly increasing from 1400 to 1699 (except for a slight decrease in 1600-1649). In the 1650-1699 interval, this difference in median DI between portrait subject genders reaches 0.325976 (0.363811 for male and 0.689787 for female). However, in the interval from 1700-1749, the gender gap in DI suddenly shrinks to 0.011515 (0.736472 for male and 0.747987 for female), and the gender gap continues to decrease in the ensuing years leading up to the end of the twentieth century, to 0.00271 (0.732191 for male and 0.729481 for female). In Fig 19, we compare female and male representative portraits from 1500-1599 and from 1650-1699. Fig 20 shows the same for 1700-1749 and 1750-1799, and Fig. 21 for 1800-1849 and 1950-1999.

Figure 19: From left to right: representative portrait of 1500-1549 (female), 1500-1549 (male), 1650-1699 (female), 1650-1699 (male)

Figure 20: From left to right: representative portrait of 1700-1749 (female), 1700-1749 (male), 1750-1799 (female), 1750-1799 (male)

Figure 21: From left to right: representative portrait of 1800-1849 (female), 1800-1849 (male), 1950-1999 (female), 1950-1999 (male)

Furthermore, we analyze pairs of art styles and intervals that have statistically significant differences in the CI and DI values. We ran Tukey's HSD tests on the CI and DI of art styles and intervals, separately, and include the results below in Appendix C.

Noticing that there where many style (and interval) pairs that showed a significant differences in both CI and DI, we wondered,

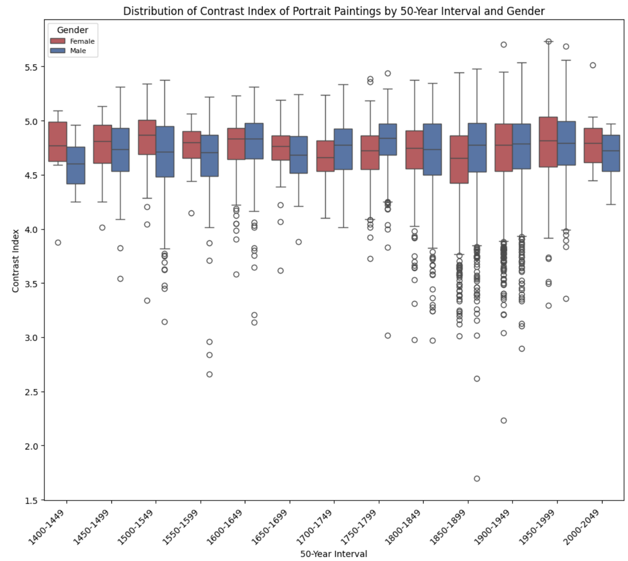

Figure 22: Box-plots of Contrast Index, Grouped by Gender and Half-Century Intervals (in ascending order by median)

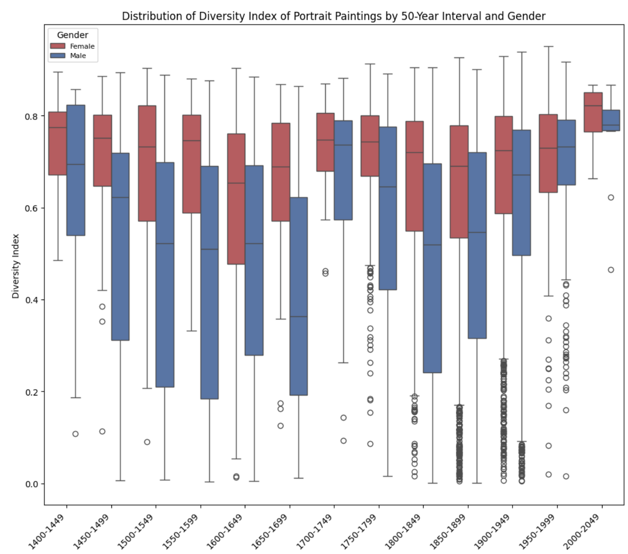

Figure 23: Box-plots of Diversity Index, Grouped by Gender and Half-Century Intervals (in ascending order by median)

do art styles tend to undergo changes in CI and DI in the same direction (i.e., is there a correlation)? And similarly for the half-century interval pairs, does CI and DI change in the same way over time? We first plotted scatter-plots for the style pairs and fitted a linear regression line, which is presented in Fig. 24.

Interestingly, we note that out of the 72 "common" pairs (pairs of styles that show a significant difference in both CI and DI), all but 6 of them see CI and DI change in the same direction. Similarly for the time-interval pairs, in Fig. 25, 7 out of 9 common pairs see changes in CI and DI in the same direction.

Lastly, we selected a few painters whose works were particularly well represented within our dataset—Ilya Repin, Amadeo Modigliani, and Pablo Picasso—to analyze how their use of color in painting clothing changes over time, through the use of Color Contrast and Diversity Indices. These results are presented in

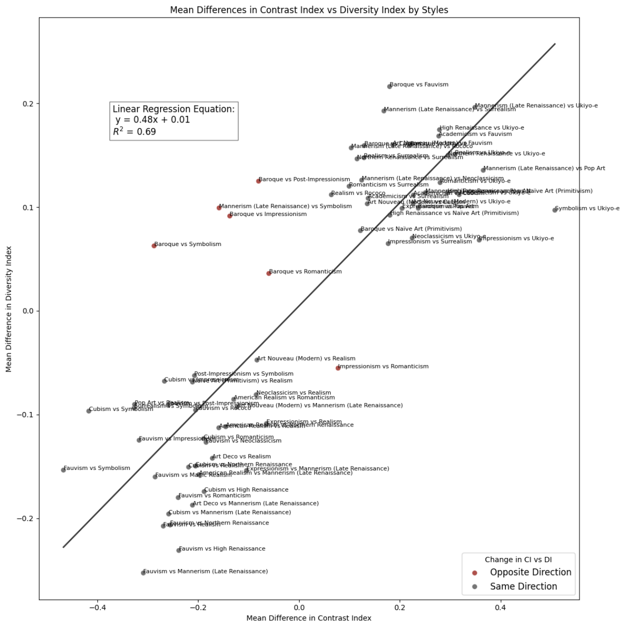

Figure 24: Linear Regression of Changes in CI vs DI for Style-Pairs with Significant Differences in both CI and DI

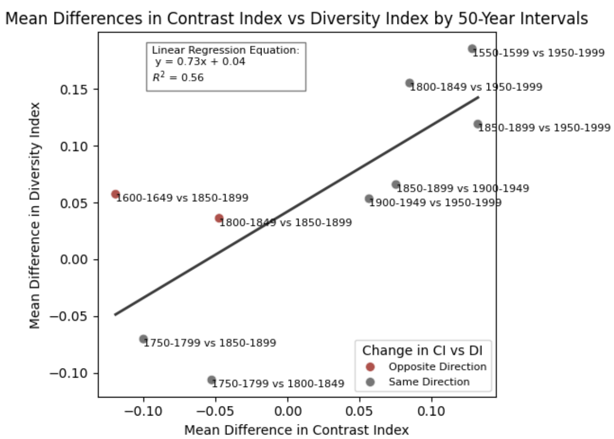

Figure 25: Linear Regression of Changes in CI vs DI for Half-Century Interval-Pairs with Statistically Differences in both CI and DI

Fig.26, 27, and 28, respectively.

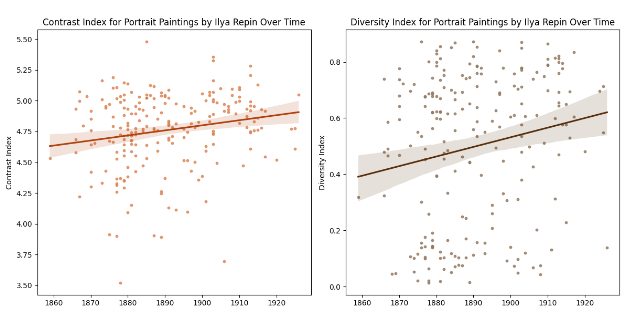

Figure 26: Scatterplots of CI and DI of Ilya Repin's Oeuvre

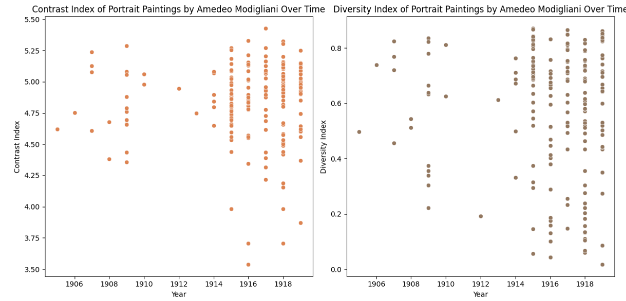

Figure 27: Scatterplots of CI and DI of Amedeo Modigliani's oeuvre.

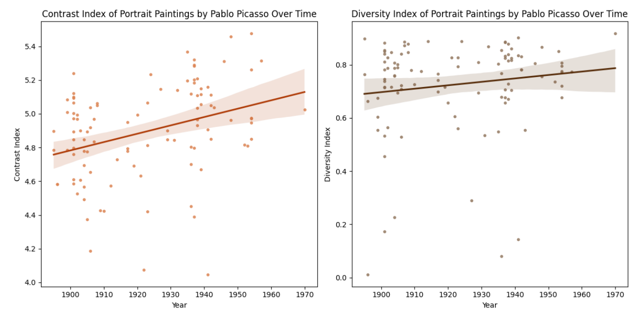

Figure 28: Scatterplots of CI and DI of Pablo Picasso's oeuvre.

Conclusion

We have presented a comprehensive, computational analysis of major trends in colors in clothing of portrait paintings over the past 600 years. With a primary goal of sharing these novel approaches from computer vision applied for fine art analysis, we hope this work can foster more cross collaboration between the two fields. Our analyses can be extended to include more fine-grained analysis of various qualities of the dominant colors across styles and time periods. Since the start-of-the-art computer vision models for segmentation are primarily trained on photographs, rather than images of fine art, there is room for improvement in segmentation and other computer vision tasks applied to paintings. Further work could include a large scale study of materials, stylistic cut, silhouette, and other attributes of clothing over time. In conjunction with the work on computational analysis of head pose in [3] and [4] as well as other prior works, these methods comprise a compendium of computational tools that can aid art scholars in the study of fine art.

Acknowledgements

The first author acknowledges the tremendous support from Professor David G. Stork throughout the research process, as all as Ali Unlu of Bilkent University for help with the dataset.

Author Biography

Christine Li is a junior at Columbia University studying Computer Science and Linguistics. She enjoys research at the intersection of computer vision and the arts.

David G. Stork is Adjunct Professor of Electrical Engineering, Symbolic Systems, and Material Science and Engineering as well as Adjunct Lecturer in Computational Mathematics and Engineering at Stanford University. He has taught many courses on computer image analysis of art.

References

[1] J. Li, L. Yao, E. Hendriks, and J. Z. Wang, "Rhythmic brushstrokes distinguish van Gogh from his contemporaries: Findings via automated brushstroke extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(6), pp. 1159–1176, 2012.

[2] A. Elgammal, B. Liu, D. Kim, M. Elhoseiny, and M. Mazzone, "The shape of art history in the eyes of the machine," in *AAAI Conference on Artificial Intelligence*, pp. 2183–2191, 2018.

[3] J.-P. Chou and D. G. Stork, "Computational tracking of head pose through 500 years of fine-art portraiture," in *Computer vision and analysis of art*, D. G. Stork and K. Heumiller, eds., SPIE, 2023.

[4] A. Unlu, J.-P. Chou, and D. G. Stork, "Computational estimation of head pose and gender through large corpora of fine-art portraits," in *Computer vision and analysis of art*, D. G. Stork and E. Spratt, eds., SPIE, 2024.

[5] D. G. Stork, "Locating illumination sources from lighting on planar surfaces in paintings: An application to Georges de la Tour and Caravaggio (abstract)," in *Optical Society of American Annual Meeting*, (Rochester, NY), 2008.

[6] D. G. Stork and M. K. Johnson, "Lighting analysis of diffusely illuminated tableaus in realist paintings: An application to detecting 'compositing' in the portraits of Garth Herrick," in *Electronic Imaging: Media forensics and security*, E. J. Delp III, J. Dittmann, N. D. Memon, and P. W. Wong, eds., **7254**, pp. 72540L1–8, SPIE/IS&T, Bellingham, WA, 2009.

[7] B. Lee, M. K. Seo, D. Kim, I. Shin, M. Schich, H. Jeong, and S. K. Han, "Dissecting landscape art history with information theory," *Proceedings of the National Academy of Science* **117**(43), pp. 26580–26590, 2020.

[8] D. Hockney, *Secret knowledge: Rediscovering the lost techniques of the old masters*, Viking Studio, New York, NY, 2001.

[9] D. G. Stork, J. Collins, M. Duarte, Y. Furuichi, D. Kale, A. Kulkarni, M. D. Robinson, C. W. Tyler, S. Schechner, and N. Williams, "Did early Renaissance painters trace optically projected images? The conclusion of independent scientists, art historians and artists," in *Digital imaging for cultural heritage preservation*, F. Stanco, S. Battiato, and G. Gallo, eds., ch. 8, pp. 379–407, CRC Press, Boca Raton, FL, 2011.

[10] D. G. Stork, *Pixels & paintings: Foundations of computer-assisted connoisseurship*, Wiley, Hoboken, NJ, 2024.

[11] S. M. Pearce, "The study of costume in painting," *Studies in Conservation* **4**(4), pp. 127–139, 1959.

[12] D. L. Chapman and L. E. Dickey, "A study of costume through art: An analysis of Dutch women's costumes from 1600 to 1650," *Dress* **16**(1), pp. 29–37, 1990.

[13] E. E. S. Gordenker, "Is the history of dress marginal? Some thoughts on costume in seventeenth-century painting," *Fashion Theory* **3**(2), pp. 219–240, 1999.

[14] J.-L. Yoo and O.-S. Cho, "A study on the civil costumes appeared on Dutch paintings in the 17th century," *Journal of Fashion Business* **3**(2), pp. 37–47, 1999.

[15] R. Hoekstra, "Images of dress in the Golden Age of Dutch painting: A methodology of research into women's clothing in the Netherlands in the seventeenth century," *Costume* **33**(1), pp. 36–45, 1999.

[16] A. Ribeiro, "Some evidence of the influence of the dress of the seventeenth century on costume in eighteenth-century female portraiture," *The Burlington Magazine* **119**(897), pp. 832–840, 1977.

[17] C. B. Ng, Y. H. Tay, and B.-M. Goi, "Recognizing human gender in computer vision: A survey," in *PRICAI 2012: Trends in Artificial Intelligence: 12th Pacific Rim International Conference on Artificial Intelligence*, pp. 335–346, (Kuching, Malaysia), 2012.

[18] A. Hollander, *Fabric of vision: Dress and drapery in painting*, Bloomsbury Visual Arts, London, UK, 2010.

[19] C. Sarı, A. A. Salah, A. Salah, and A. Almıla, "Automatic detection and visualization of garment color in Western portrait paintings," *Digital Scholarship in the Humanities* **34**(Supplement_1), pp. i156–i171, 2019.

[20] W. Kan, "Painter by numbers," 2016. https://kaggle.com/competitions/painter-by-numbers.

[21] A. C. Gallagher and T. Chen, "Using context to recognize people in consumer images," *IPSJ Transactions on Computer Vision and Applications* **1**, pp. 115–126, 2009.

[22] A. Cifuentes and V. Charlin, *The Worth of Art: Financial Tools for the Art Markets*, Columbia Business School Publishing, New York, NY, 2023.

[23] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," in *Proceedings of the 38 th International Conference on Machine Learning (PMLR)*, p. 139, 2021.

[24] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research* **9**(86), pp. 2579–2605, 2008.

Appendix A

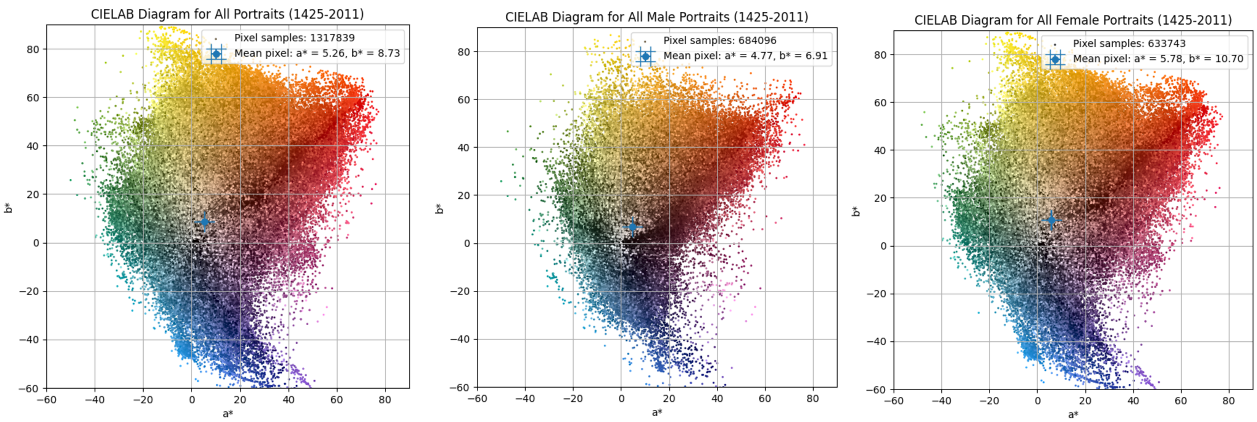

Figure 29: Aggregate CIELAB Colors for all portraits (left), all male portraits (middle), and all female portraits (right)

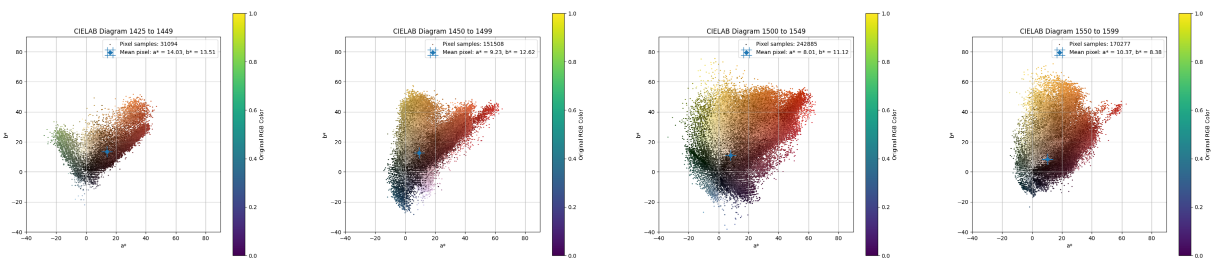

Figure 30: Aggregate CIELAB Colors over 50-year Intervals (1425-1449, 1450-1499, 1500-1549, 1550-1599)

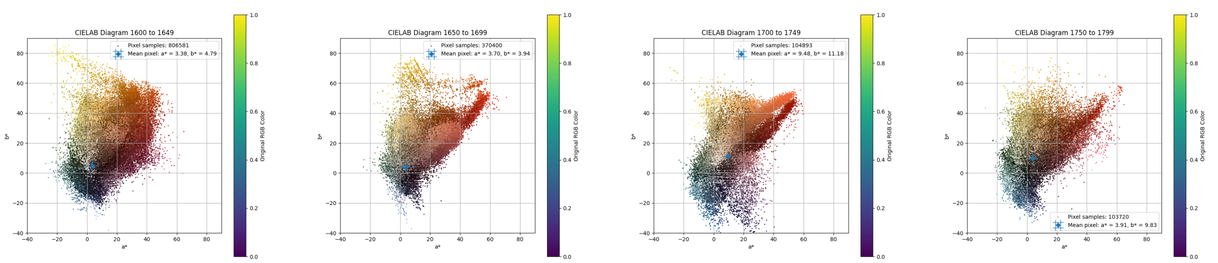

Figure 31: Aggregate CIELAB Colors over 50-year Intervals (1600-1649, 1650-1699, 1700-1749, 1750-1799)

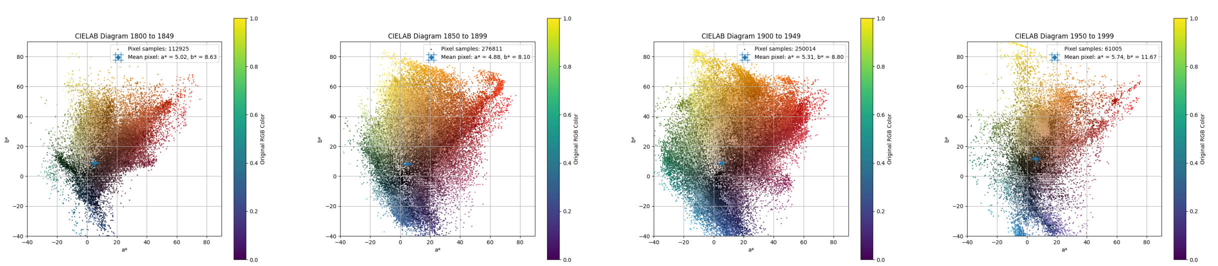

Figure 32: Aggregate CIELAB Colors over 50-year Intervals (1800-1849, 1850-1899, 1900-1949, 1950-1999)

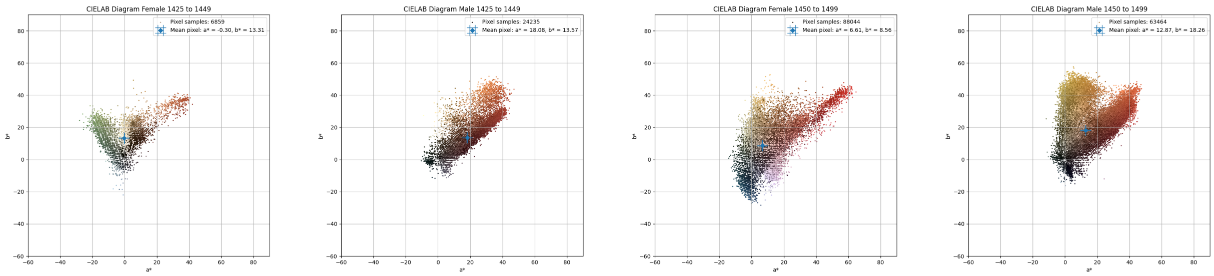

Figure 33: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1425-1449, 1450-1499)

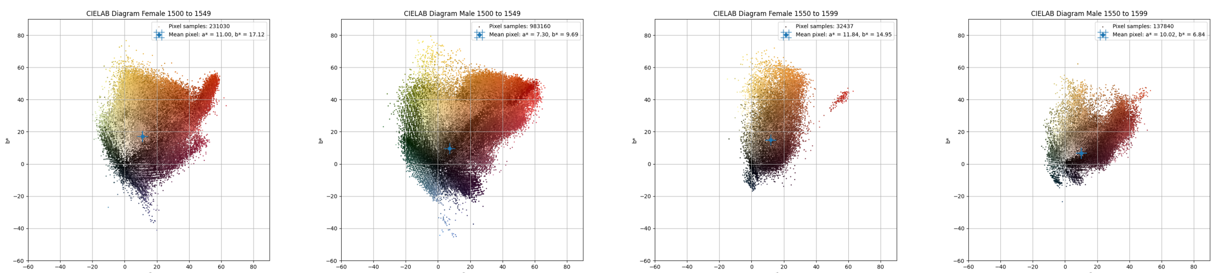

Figure 34: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1500-1549, 1550-1599)

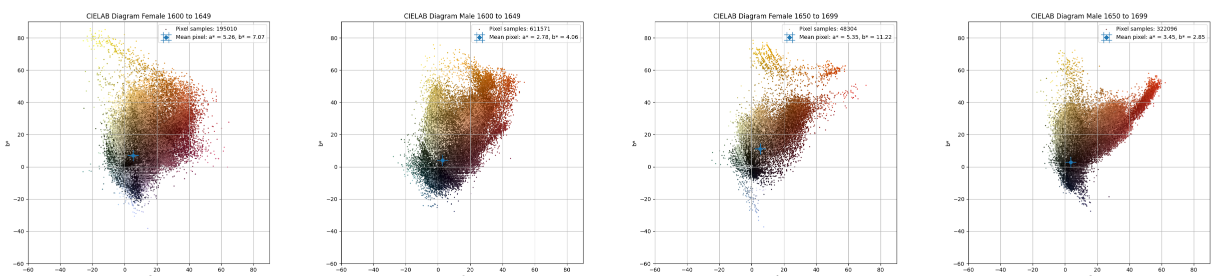

Figure 35: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1600-1649, 1659-1699)

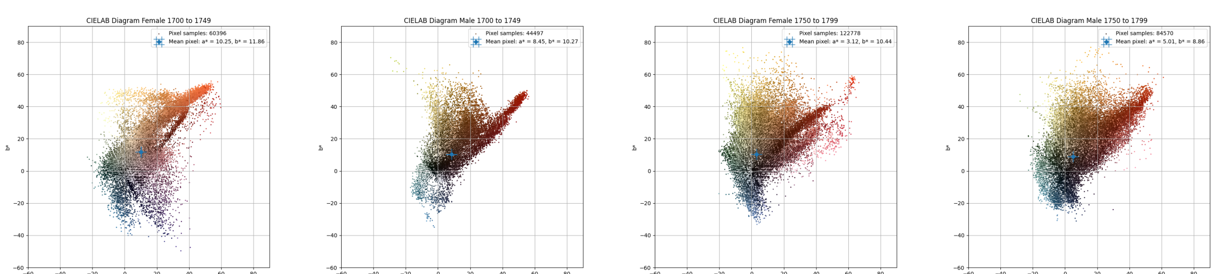

Figure 36: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1700-1749, 1750-1799)

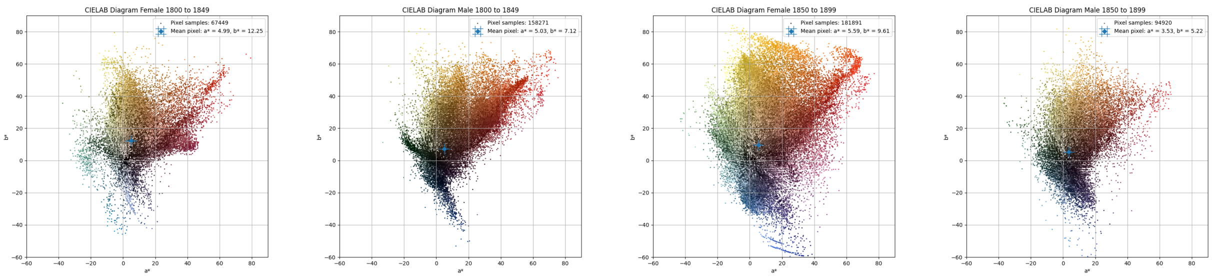

Figure 37: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1800-1849, 1850-1899)

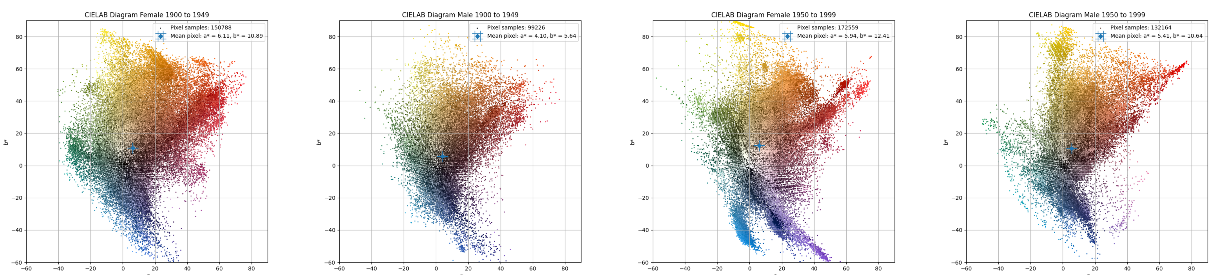

Figure 38: Aggregate CIELAB Colors over 50-year Intervals for female and male portraits (1900-1949, 1950-1999)

Appendix B

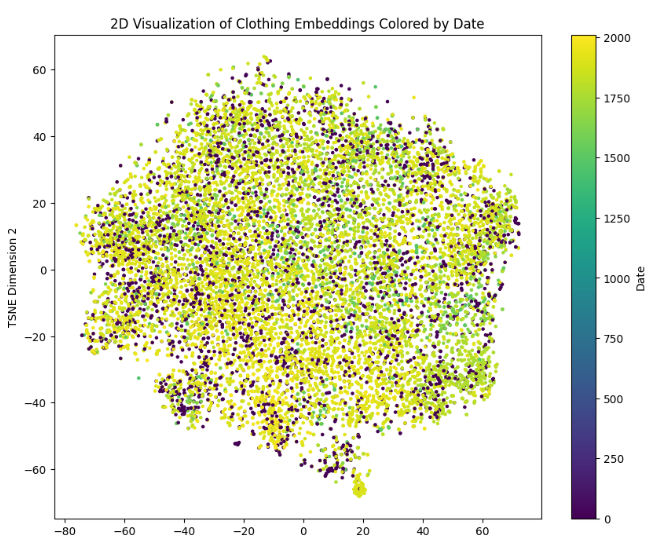

Figure 39: TSNE embeddings for all portrait paintings, colored by date

Appendix C

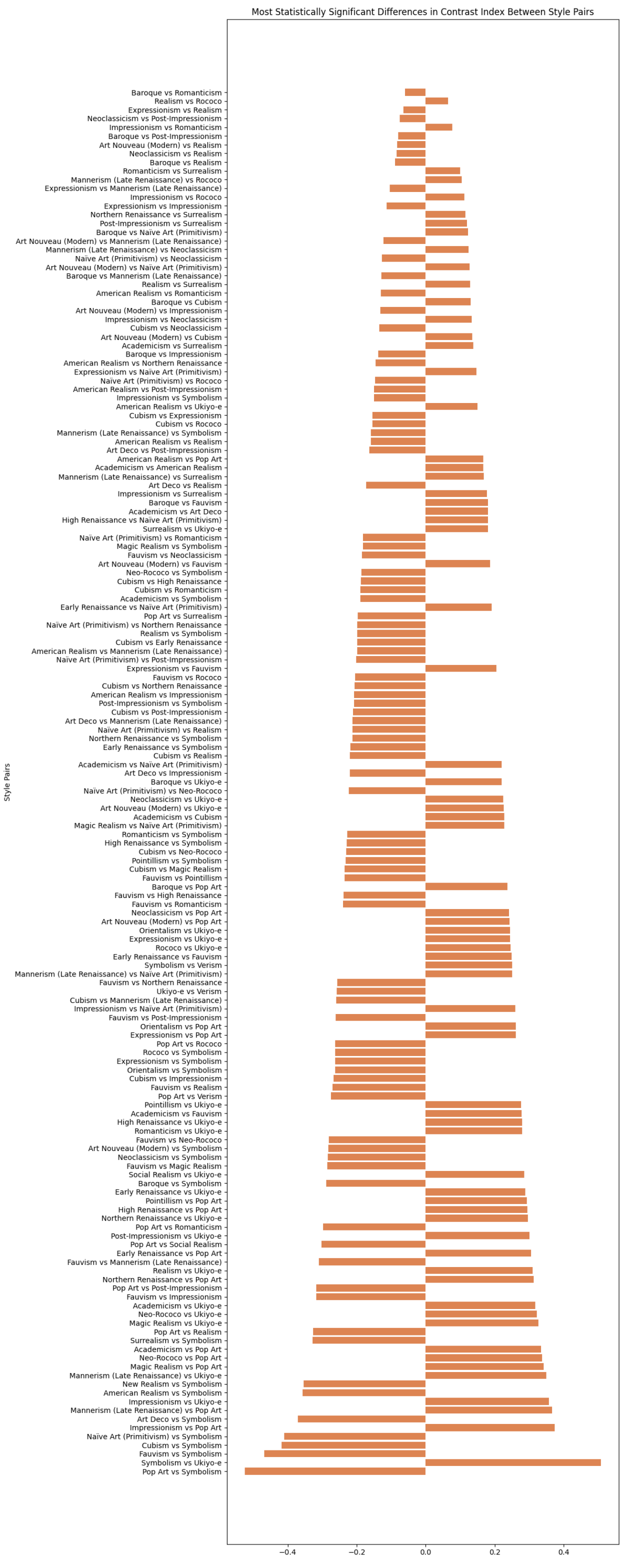

Figure 40: Pairs of Styles with Significant Difference in CI Values

Figure 41: Pairs of Styles with Significant Difference in DI Values

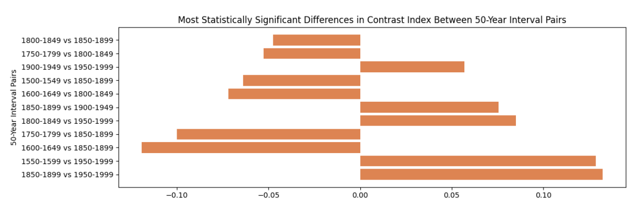

Figure 42: Pairs of Half-Century Intervals with Significant Difference in CI Values

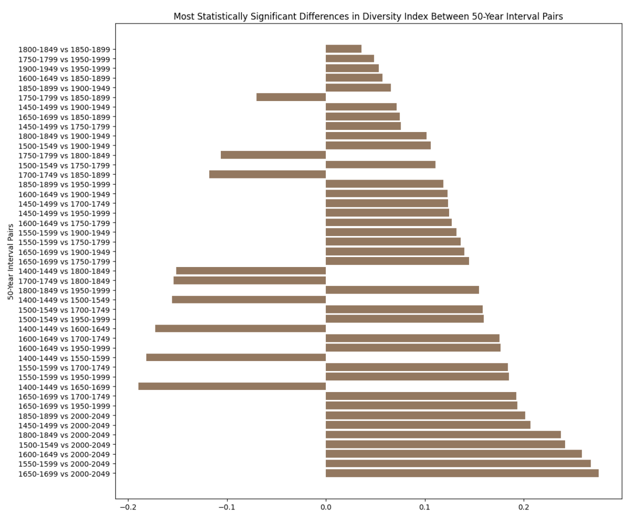

Figure 43: Pairs of Half-Century Intervals with Significant Difference in DI Values