Ukiyo-e Analysis and Creativity with Attribute and Geometry Annotation

Yingtao Tian
Google Brain
Tokyo, Japan

Tarin Clanuwat
ROIS-DS Center for
Open Data in the Humanities
NII

Chikahiko Suzuki
ROIS-DS Center for
Open Data in the Humanities
NII

Asanobu Kitamoto
ROIS-DS Center for
Open Data in the Humanities
NII

Abstract

The study of Ukiyo-e, an important genre of pre-modern Japanese art, focuses on the *object* and *style* like other artwork researches. Such study has benefited from the renewed interest by the machine learning community in culturally important topics, leading to interdisciplinary works including collections of images, quantitative approaches, and machine learning-based creativities. They, however, have several drawbacks, and it remains challenging to integrate these works into a comprehensive view. To bridge this gap, we propose a holistic approach: We first present a large-scale Ukiyo-e dataset with coherent semantic labels and geometric annotations, then show its value in a quantitative study of Ukiyo-e paintings' *object* using these labels and annotations. We further demonstrate the machine learning methods could help *style* study through soft color decomposition of Ukiyo-e, and finally provides joint insights into *object* and *style* by composing sketches and colors using colorization. We make our dataset (Tian, CODH, and ARC 2021) available online [1].

Introduction

The Edo period of Japan (16th to 19th century) has seen the prosper of Ukiyo-e (浮世絵), a genre of pre-modern Japanese artwork that consists of paintings and woodblock printings. Unlike early dominating Emakimono (絵巻物, *picture scroll*) and Ehon (絵本, *picture book*) that focus on famous figures and stories in Sinosphere culture and classic Japanese stories, the topic of Ukiyo-e extends broadly to daily subjects, such as characters like beauties and *Kabuki* (歌舞伎), landscape arts, animals and plants in everyday life, and even contemporary news. As an example, Figure 1 shows an Ukiyo-e depicting a *Kabuki* performance. The popularity of woodblock printing makes it possible to produce paintings on a larger scale at a lower cost, which contributes to the flourish of Ukiyo-e and leaves us with a vast collection of artworks in this genre (Kobayashi 1994; IUS 2008). Such an extensive and varied collection provides a valuable corpus for Japanese artwork research.

The subject of such artwork study could be multi-faceted involving several aspects, of which two crucial are the *object* in the painting, such as the outline and the shape of depicted figures, and the *style* of painting, such as textures

[1] https://github.com/rois-codh/arc-ukiyoe-faces

Attribute	Value
Title (画題)	*Kizukansuke* (「木津勘助」)
Painter (絵師)	Hirosada (広貞)
Format (判種)	Middle-size / *Nishiki-e* (中判/錦絵)
Year in AD (西暦)	1849
……	

Figure 1: An example of Ukiyo-e work in ARC Ukiyo-e Collection (Object Number arcUP2452) titled *Kizukansuke* by painter Hirosada. The painting on the left is accompanied by metadata for this work on the right. For example, metadata further indicates this work is a middle-sized *Nishiki-e* (multi-colored woodblock printing) produced in 1849.

and colors. For example, the former reveals the trend of objects depicted over time, and the latter allows the identification of artists (Suzuki, Takagishi, and Kitamoto 2018). The renewed interest by the machine learning community in the culturally essential topics has led to works addressing the traditional Japanese artworks from an interdisciplinary perspective. Along this line of research, building open collections of digitized images has been proposed for Ehon (Suzuki, Takagishi, and Kitamoto 2018) and Ukiyo-e (Art Research Center, Ritsumeikan University 2020; Pinkney 2020). Further works use quantitative approaches into the *object* for artworks, such as studying the geometry features of Buddha statues (Renoust et al. 2019) and Ukiyo-e faces (Renoust et al. 2019), Alternatively, inspired by the art nature of painting, machine learning-based creativity has been leveraged for studying *style*, such as painting process generation (Tian et al. 2020) and image synthesis across artwork and photorealistic domains (Pinkney and Adler 2020). These works provide valuable connections between machine learning and the humanities research of Japanese artwork.

We, however, also notice that these works present several drawbacks. For example, collection on digitized images may either comes with no semantic (Pinkney 2020) or is in a format not designed with machine learning-based applications in mind.Furthermore, quantitative approaches are only conducted on a small set of artworks (Murakami and

Proceedings of the 12th International Conference on Computational Creativity (ICCC '21)
ISBN: 978-989-54160-3-5

Urabe 2007) or require extensive human labor to adapt for Ukiyo-e (Renoust et al. 2019), and machine learning-based creativity works may deal more with cross-domain art expression (Pinkney and Adler 2020) than the very domain of artwork on which humanities research focuses. Finally, the art study into a particular genre requires insights into both the *object* and *style* to acquire a comprehensive understanding. Current works, however, only address one of the *object* or *style*, falling short of the expectation.

To overcome the aforementioned drawbacks and to provide deeper insight into the artistic style of Ukiyo-e, we propose a new approach that is (1) holistic in both studying the *object* and *style* through the joint use of images, labels, and annotations, and (2) powered by large scale data and state-of-the-art machine learning model than the prior works. To summarize, our main contributions are as follow:

- We present a large-scale ($11,000$ paintings and $23,000$ faces) Ukiyo-e dataset with coherent semantic labels and geometric annotations, through augmenting and organizing existing datasets with automatical detection.
- We are the first to conduct a large-scale quantitative study of Ukiyo-e paintings (on more than $11,000$ paintings), providing understanding into *object* in artworks by jointly quantifying semantic labels and geometric annotations.
- We show that machine learning-based models could provide insights into *style* by decomposing finished Ukiyo-e images into color-split woodblocks that reflect how Ukiyo-e images were possibly produced.
- We study and show machine learning-based creativity model could engage problems that arise jointly studying *object* and *style* by separating geometry shapes and artistic styles in an orthogonal and re-assemblable way.

Dataset

Art research in traditional paintings often asks questions regarding the work, like the author and production year. One focus in such research is on faces since they could help answer these questions through quantitative analysis. In this direction, Collection of Facial Expressions (Suzuki, Takagishi, and Kitamoto 2018; Tian et al. 2020) provides a large-scale (8848 images) set of coarse-grained cropped faces. Another study (Murakami and Urabe 2007) deals with facial landmarks which are more fine-grained than cropped faces to support quantitative analysis. However, its manual labeling process only allows analysis on a small set (around 50 images) of Ukiyo-e paintings.

To combine both works' advantage, we extend existing datasets through augmentation and automated annotation, resulting in a large-scale Ukiyo-e dataset with a more fine-grained facial feature. The rest of this section details the process and analysis of our new proposed dataset.

Fundamental Datasets

We build our work based on two foundation datasets. One of them is ARC Ukiyo-e Collection (Art Research Center, Ritsumeikan University 2020), a publicly available service

Facial Region	Landmarks
Left Eye	Center, Left, Right, Up, Down
Right Eye	Center, Left, Right, Up, Down
Left Eyebrow	Left, Right, Up
Right Eyebrow	Left, Right, Up
Left Pupil	Center
Right Pupil	Center
Mouth	Left, Right, Up, Down
Nose	Center, Left, Right
Jawline	Upper Left & Right, Mid Left & Right, Chin Bottom

Figure 2: An exmaple of detected landmarks and the extracted face in Figure 1's Ukiyo-e painting. On the left, the red dots show detected facial landmarks and the rectangle shows the bounding box inferred from these landmarks. The right image shows the extracted face from the bounding box. The table lists a summary of all landmark locations.

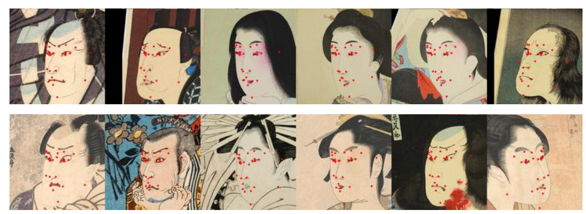


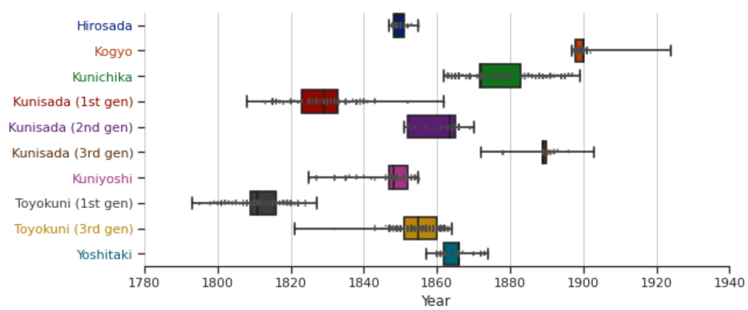
Figure 3: Faces with their landmarks. In each row, we show six examples of extracted faces annoated with their corresponding landmarks in the same format as Figure 2.

that provides access to digitized Ukiyo-e paintings primarily in the Edo period, plus metadata compiled by domain experts. It has $11,103$ entries of painting and the associated metadata, one example of which is shown in Figure 1. This service allows researchers to dive into curated metadata for comparative study for art research.

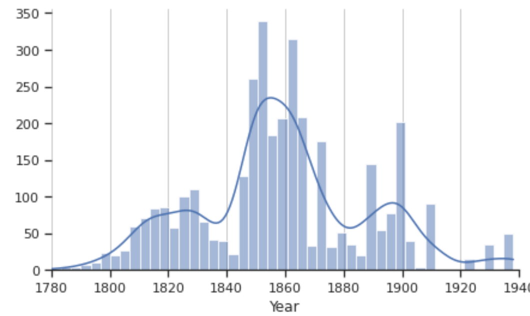
Another dataset is Ukiyo-e Faces Dataset (Pinkney 2020), a public available dataset of Ukiyo-e faces extracted from Ukiyo-e images available online. With $5,000$ high-quality faces, this dataset plays an essential role in controllable image generation across Ukiyo-e faces and photo-realistic human faces (Pinkney and Adler 2020). However, as this dataset focuses on image synthesis, it does not include metadata for Ukiyo-e paintings from which faces are extracted.

Geometric Annotation with Facial Landmark Detection

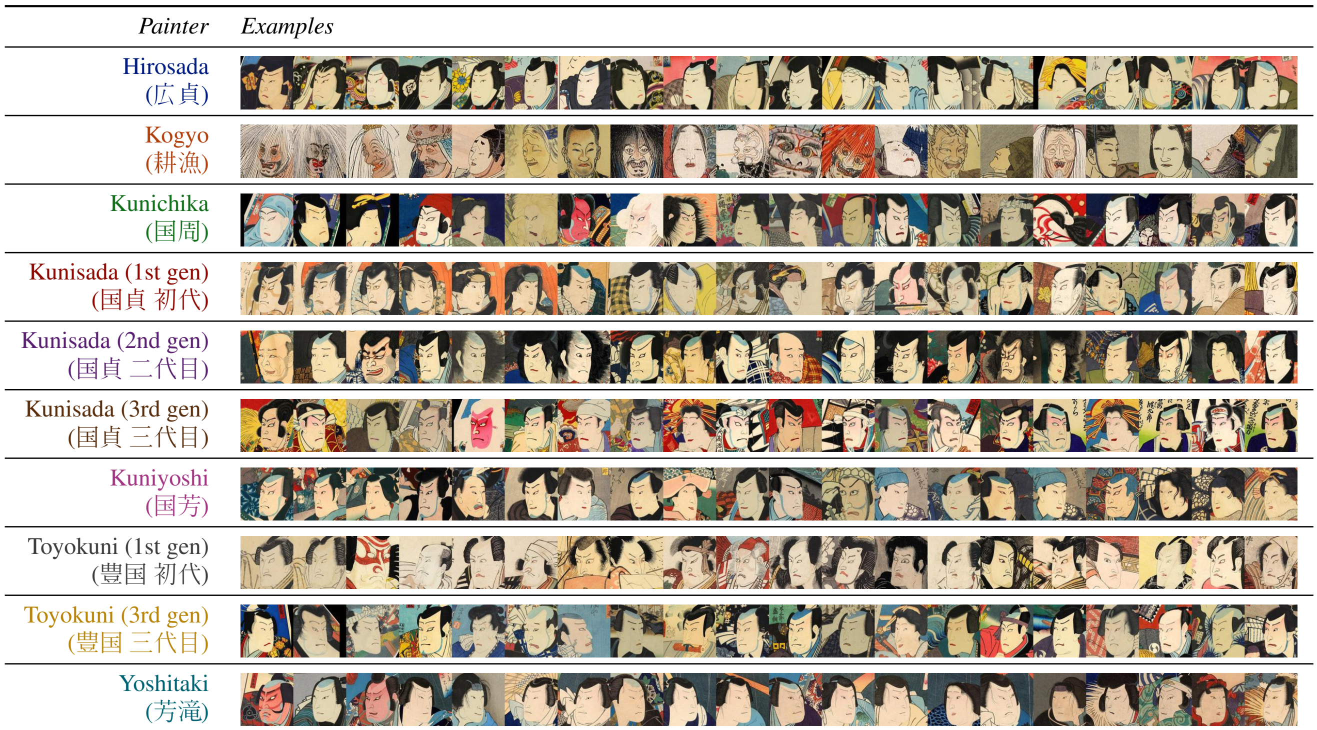
Inspired by Pinkney (2020) , we use an face recognition API, Amazon Rekognition (link), to detect facial landmarks in in Ukiyo-e Faces Dataset paintings. Despite targeting photo-realistic human face images, this API demonstrates compelling accuracy on Ukiyo-e paintings. Since the detected faces may not be well-aligned, we infer the possibly rotated bounding box of faces for cropping faces from the



(a) Distribution of years with respect to authors.



(b) Distribution of years of all works in the dataset.



(c) Example of paintings, represented by the extracted faces, by authors.

Figure 4: Metadata and their associated paintings. We jointly show two important metadata, year and author. We also show show ten authors with the most painting in the dataset. **(a)** illustrates the year distribution with respect to authors, and **(b)** shows the overall year distribution. Exemplary paintings belonging to these authors are shown in **(c)**.

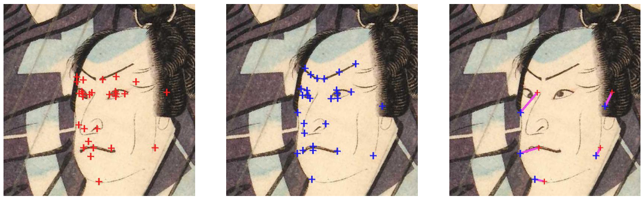
painting, inspired by the preprocessing in FFHQ (Karras, Laine, and Aila 2019). In Figure 2 we show an example of detected landmarks and the face extraction process.

A total of $18,921$ faces and their corresponding facial landmarks have been detected from paintings in ARC Ukiyo-e Collection. Furthermore, since Ukiyo-e Faces Dataset (Pinkney 2020) also follows the same preprocessing as FFHQ, its $5,000$ faces are comparable to the faces extracted from ARC Ukiyo-e Collection. Although faces in Ukiyo-e Faces Dataset lack metadata, we can still incorporate them for geometry statistics by going through the above-mentioned landmark detecting process. In doing so, we have a total of around $23,000$ Ukiyo-e faces. In Figure 3, we show examples of such faces and their landmarks.

Semantic Labels Incorporation

As our dataset is derived from ARC Ukiyo-e Collection, we can also relate faces and the corresponding landmarks with the original metadata, such as the year of creation and the author of the painting. In Figure 4 we show these two metadata jointly, as well as exemplary paintings belonging to several authors. For example, we can observe Shumei (襲名, *name succession*) system common in traditional Japanese art community where an artist takes his/her teacher's name, as the case of the lineage of Kunisada 1st gen (国貞 初代), Kunisada 2nd gen (国貞 二代目) and Kunisada 3rd gen (国貞 三代目). Furthermore, we can also notice three peaks of production of Ukiyo-e painting, occupying the early, mid, and late 19th century. The last peak is dominated by Kogyo (耕漁) who painted well into the 20th century and whose

Proceedings of the 12th International Conference on Computational Creativity (ICCC '21)
ISBN: 978-989-54160-3-5



(a) Landmarks detected automatically

(b) Landmakrs manually annotated by domain experts

(c) Detailed Study in Jawline landmarks

Facial Region	Landmarks (Mean Error in Pixel Distance)		
Left Eye	Center (10.2)		
Right Eye	Center (18.5)		
Mouth	Left (9.7)	Right (13.4)	
Nose	Center (22.1)	Left (18.6)	Right (16.5)
Left EyeBrow	Left (34.2)	Right (43.0)	Up (23.1)
Right EyeBrow	Left (17.4)	Right (41.8)	Up (53.1)
Jawline *	Upper Left (69.4)	Upper Right (57.4)	Mid Left (25.0)
	Mid Right (56.2)	Chin Bottom (57.8)	

(d) Landmarks with the mean error in pixel distance between detected and expert labeled postion. Landmarks in green are considered of high-quality and those in red of low-quality.

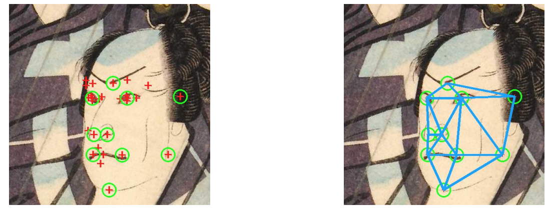
Figure 5: Study of landmark quality by comparing automatically detected positions **(a)** with expert labeled positions **(b)**. As Ukiyo-e faces are mostly towards either left or right, we normalize all paintings to face left for the geometry purpose. The study has been conducted for 69 Ukiyo-e paintings, and the mean error of pixel distances are aggregated in **(d)**, which we use to decide which landmarks are considered high-quality. The decision is based on picking landmarks with a low error of pixel distance (heuristically those < 20), except for Jawline (*) that needs special consideration: as **(c)** shows, landmarks on the the direction of facing (Upper Left, Mid Left) are useless since they are invisible in most Ukiyo-e. The others (Upper Right, Mid Right, Chin Bottom) are valuable since they still lie on the jawline, and they are far from other landmarks, allowing larger error marging when used for calculating angular features.

uniqueness is further shown in his exemplary paintings under the influence of modern painting.

Experiment

Study Ukiyo-e *Object* using Geometry Features

Regarding the content, art researches may be divided into two categories: the shape that deals with geometry features and the texture that deals with brushes and color features. To quantitatively provide insights on attributes such as the author and the painting year, either category can be used for unsupervised learning, like clustering, or supervised learning, like predicting metadata, While the texture features could help analyze attributes for a single work (Tian et al. 2020), the geometry features could also be considered since the texture may vary due to Ukiyo-e's frequent reprint (Murakami and Urabe 2007) or sculpture's preservation condition (Renoust et al. 2019), Both works propose to leverage



(a) High quality landmarks circled in green.

(b) Examples of angles formed by high quality landmarks

Figure 6: Extracted geometry features from high quality landmarks. In (a) we highlight landmarks of high quality, and in (b) we show some a subset of of angles formed by high quality landmarks for brevity.

facial landmarks to infer geometry features such as angles and iconometric proportions to quantify artwork.

However, since both works rely on manually labeled landmarks, they either suffer from being too small (only around 50 Ukiyo-e paintings are annotated with landmarks) or require extensive human effort if we ever want to apply the technique used on sculpture to Ukiyo-e. To bridge this gap, we propose to use automatically detected landmarks as geometry features. To our best knowledge, we are the first to conduct large-scale (more than 10k paintings) quantitative analysis of Ukiyo-e painting. We hope it could serve as an invitation for further quantitative study in artworks.

Geometry Features from Landmarks Inspired by Murakami and Urabe (2007) , we consider the angles formed by landmarks as they are geometry-invariant under rotation. Here we focus on rotation in two-dimensional space. As tranditional Japanese painting are not photo-realistic, the 3D perspective in viewing is a more complex issue invovling deformation and we left it for future study. To attain a clear understanding of the quality of landmarks, we conduct a study on 69 Ukiyo-e paintings, comparing landmarks that are automatically detected with positions manually annotated by domain experts, as detailed in Figure 5. We observe that, despite the general high-quality of the predicted landmarks on Ukiyo-e painting, some landmarks have systematically worse quality than others we decided not to consider. In the end, we calculate 252 angles formed by all possible triplets of high-quality landmarks as geometry features for each face, as illustrated in Figure 6.

Analysis on Authorship To illustrate the information of geometry features, we conduct unsupervised (PCA, T-SNE, UMAP) and supervised (LDA) clustering of faces using geometry features in Figure 7. All clusterings show two distinctive authors, Kogyo and Hirosada (広貞), are separated from other authors. Such separation could be supported through visual inspection into original paintings. For example, Figure 4 (c) shows Kogyo and Hirosada has visually distinctive styles compared to other painters. Furthermore, such separation could also be cross-verified with analysis leveraging other information sources, Figure 4 (a) shows that Kogyo was active well into the 20th century where in contrast Ukiyo-e paintings are mainly around the middle 19th century, Furthermore, his uniqueness of style is visu-

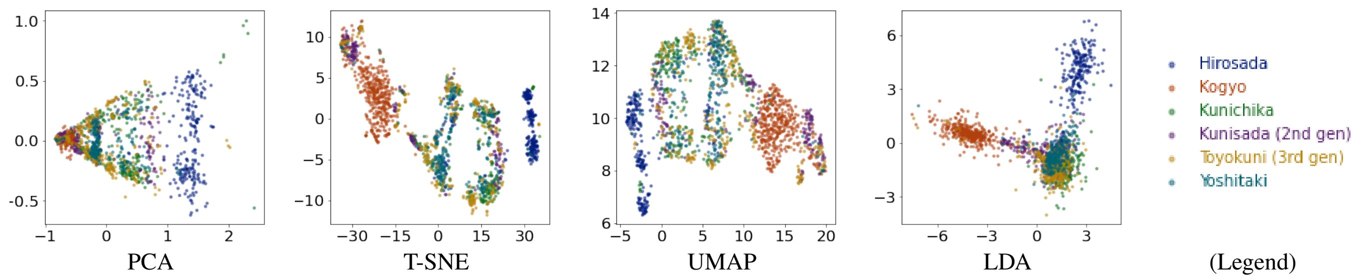


Figure 7: Unsupervised (PCA, T-SNE, UMAP) and supervised (LDA) clustering of faces' geometry features in a two-dimensional plane. We show works by six most-frequently appearing authors in the clustering. Labels are used for coloring the authors for illustrative purposes only, and are not used in the clustering except for LDA. Visually, Hirosada and Kogyo are shown with clear separation from other authors, which could be cross-verified with explanation using other information.

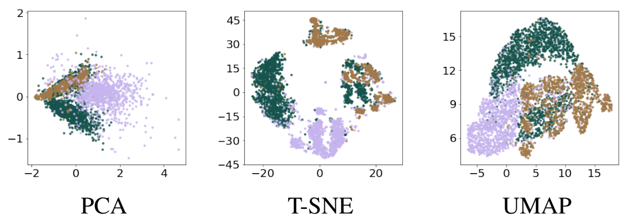


Figure 8: Unsupervised (PCA, T-SNE, UMAP) clustering of faces using geometry features for Ukiyo-e paintings, Kaokore paintings and photo-realistic human faces. Visually, we see a separation between faces of different sources.

ally illustrated in exemplary paintings. We can also observe Hirosada forms a unique style related to the geographical factors. While most of the painters analyzed at that time worked in Edo (modern Tokyo), Hirosada was active in Osaka. Comparison of culture at that time could be made between the around Edo/Tokyo region, the *de facto* capital of Japan under Tokugawa shogunate, and Kamigata (上方) region encompassing Kyoto and Osaka, the *de jure* capital of Japan and the cultural center of western Japan. The style of Hirosada and Yoshitaki (芳滝) who were active in western Japan is therefore called Kamigatae (上方絵, *Kamigata painting*) and is a subject for comparative study.

Comparing Ukiyo-e, Ehon and Human Faces As exemplary paintings in Figure 4 (c) show, Ukiyo-e paintings are characterized by their particular facial geometry, which could potentially be different from other art genres or photo-realistic human faces. To quantify such observation, we conduct unsupervised (PCA, T-SNE, UMAP) clustering of Ukiyo-e (popular in the 19$^{\text{th}}$ century) faces, Ehon / Emakimono (another Japanese artworks genre popular in the 16$^{\text{th}}$ to 17$^{\text{th}}$ century) faces, and realistic human faces.

Concretely, we use Kaokore (Tian et al. 2020) for Ehon / Emakimono faces, as well as human face photos collected in FFHQ (Karras, Laine, and Aila 2019) dataset that are published under CC BY 2.0 license. In Figure 8, we can observe that the geometry of Ukiyo-e faces is different from Kaokore, and only share similarities to a small section of realistic human faces. This observation confirms the uniqueness of Japanese artworks' way of portraying humans compared to the real-world image and shows that the development of Japanese artworks over time is a drastic one.

Study Ukiyo-e *Style* through Color Separation

Ukiyo-e printings distinguish themselves from other traditional Japanese artworks by the very manner of producing. Unlike Ehon, which is targeted at a small audience and thus painted by hand, Ukiyo-e is mass-produced using woodblock printing after the painter finishes the master version. As shown in a modern reproducing process (link), multiple woodblocks are carved, each for a portion in the image with a single color, and are printed sequentially with corresponding inks onto the final canvas. Unfortunately, such a process for a given Ukiyo-e painting is not precisely reproducible since the underlying woodblocks are vulnerable, easily worn-out, and often discarded after a certain number of prints. Thus from an art research point of view, it would be interesting to recover the above-mentioned separated portions for a given Ukiyo-e painting with only access to the image itself.

We address this challenge by framing it as a *soft color segmentation* (Aksoy et al. 2017) task, which decomposes an input image into several RGBA layers of homogeneous colors. The alpha channel ("A" in "RGBA") in each layer allows pixels to potentially belong to multiple layers, which captures ambiguity unavoidable due to imperfect woodblock carving and alignment in multi-pass printing. In detail, we use state-of-the-art Fast Soft Color Separation (FSCS) (Akimoto et al. 2020) for efficient processing. As shown in Figure 9, FSCS decomposes Ukiyo-e paintings into layers of homogeneous colors using color palette. The inferred layers could be interpreted as woodblocks with corresponding colors that could be used for making a particular artwork.

The decomposition of a painting into multiple layers of homogeneous colors allows us to explore further creativity. One example in this direction is recoloring, where we pick a new color for each of the individual layers and compose them into a recolored painting. As shown in Figure 10, the recoloring could be done either automatically using the inferred color palette from other artworks or manually in Adobe After Effects with *alpha add* mode for blending. The recoloring here serves as an example to study artworks and opens the door to reinterpret them in a new way.

Figure 9: Soft color separation takes as input Ukiyo-e paintings (left) and a color palette (middle), and produces decomposed layers of homogeneous colors (middle). These layers can be used as the inferred woodblocks for corresponding colors and composed back to a reassembled painting (right) resembling the original one. We infer the color palette by applying K-means clusttering (Lloyd 1982) on the input painting's pixels.

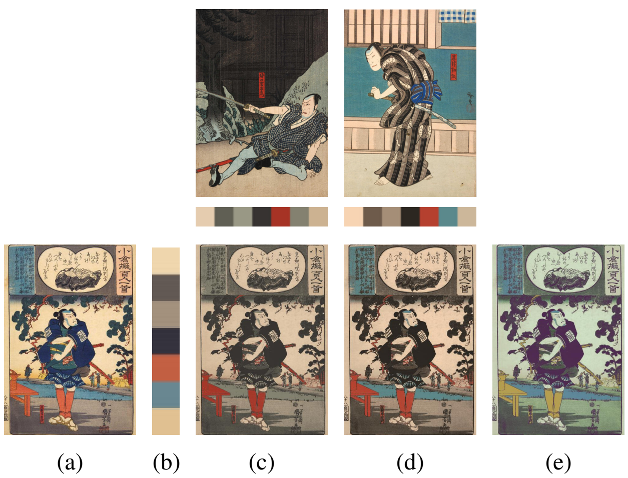


Figure 10: Decomposing an Ukiyo-e painting **(a)** with color palette **(b)** and recoloring, which could be done automatically **(c, d)** using color palettes inferred from the reference images, or manuallly with Adobe After Effects **(e)**.

Study Jointly Ukiyo-e *Object* and *Style* by Composing Sketch and Color

As we deals with a dataset focusing on artworks, it becomes natural to ask whether we could engage them with approaches invoking creativity and artistic expression. One direction is to examine whether the recent advances of machine learning models could create structurally sound, or even artistically impressive, results. In this direction, generative models has been proposed to generate faces in Japanese painting style (Tian et al. 2020) and blend generative models trained on data of different domains by swapping layers of two image generation neural networks (Pinkney and Adler 2020). However, the former lacks controllability in the generation as it can only produce images as a whole, and the latter focuses on transferring across separated, different domains by the nature of its design.

Figure 11: Pairs of original Ukiyo-e faces on the upper row and corresponding line art sketches on the lower row.

Figure 12: Colorization on Ukiyo-e faces. For a face painting **(a)**, we extract its line art sketch **(b)**. A colorization model takes both the sketch and a reference painting **(c)**, and produces a colorized painting **(d)** reflecting the sketch's geometry and the reference's style in colors and textures.

Thus we identify an unbridged gap in the *in-domain* separation of artistically essential aspects. In detail, we ask the following question: what is the (dis)entanglement between the *object* and *style* within the Ukiyo-e. Answering this question reveals the relation between Ukiyo-e's *object* and *style*. Furthermore, it also allows editing one of them while keeping another intact for creative expression. One way to separate the *object* and *style* is to represent the former with line art sketches for what person/scene is depicted, and the latter with color and texture information showing the painting style. They could be composed with a colorization process, which blends a sketch as an *object* reference and an image as a reference for instance-level painting *style*.

Face Images. We extract line art sketches from Ukiyo-e

(a) Ukiyo-e faces. (b) Whole Ukiyo-e paintings.

Figure 13: Matrices of blending line art sketches and painting style for Ukiyo-e faces **(a)** and whole Ukiyo-e paintings **(b)**. Within a single matrix, each row represents an art line sketch, each column represents the reference image for style, and images at an intersection are the blending results of the corresponding row and column.

(a) Origin (b) Sketch (c) Reference (d) Colorized

Figure 14: Top: Pairs of whole Ukiyo-e painting and line art sketches in the same format as Figure 11. Bottom: Colorization on Ukiyo-e faces in the same format as Figure 12.

images using SketchKeras (Lvming Zhang 2017), as illustrated in Figure 11. We further train image colorization (Lee et al. 2020) using a public-available implementation (Sou Hasegawa 2020). The whole pipeline is illustrated in Figure 12. As shown in Figure 13, Since the model learns to separate the *object*, indicated in the sketch image, and the *style*, indicated by reference image, as two orthogonal and composable semantics, it could blend arbitrary combination of sketch and reference style images. Such separation could enable future works to help with humanities research on combinations of Ukiyo-e color and subject. For example, in Ukiyo-e depicting Kabuki, the attributes and colors of the characters are somewhat correlated semantically. Therefore swapping colors can change the meaning of scenes and people in the painting. We envision that discoveries could be made by studying how the impression of Ukiyo-e paintings changes through the process of swapping colors.

Figure 15: Colorization in-the-wild woodblock printing using the model trained on the whole Ukiyo-e paintings. Each row represents a woodblock printing work, and each column represents the reference image for style.

Whole Painting. We go beyond faces and work on whole Ukiyo-e painting images. By employing the same pipeline

Figure 16: Comparation of conditional and the unconditional colorization method. The former uses style reference images while the latter does not. Four rows are the ground truth color image, conditional colorization, unconditional colorization, and line art sketches, respectively.

to the whole painting images, as shown in Figure 14, the model can be further leveraged to colorize in-the-wild woodblock printing images, as Figure 15 shows. However, while the resulting colorized images are reasonable, they are of lower quality than those of faces. Such observation is anticipated since the whole Ukiyo-e painting is more complex than face in many ways, like topics and topological configuration of objects, which presents a much more challenging task for colorization. This issue could be further exaggerated by the discrepancy between the Ukiyo-e domain where the model is trained and the woodblock painting domain where the model is applied. We would leave higher quality, whole Ukiyo-e painting colorization for future study.

Conditional vs. Unconditional Colorization. While we choose to use a conditional colorization method, which produces results from a sketch *and* a reference image for color and styles, it is also worth considering a simper, unconditional colorization method that directly generates the results from a sketch, such as Pix2PixHD (Wang et al. 2018). This alternation, however, suffers from the inability to control the color and style of the generated image. Moreover, as we show in Figure 16, the unconditional colorization method produces worse colorization results than the conditional colorization method (Lee et al. 2020). We argue that this is expected since the former method has to fall back to safe colors that valid for any Ukiyo-e images, while the latter could make a wiser choice based on the reference images.

Discussion We show that Ukiyo-e paintings can be studied by (1) representing *object* with line art sketches, (2) representing *style* as a color reference image, and (3) composing them using colorization. This pipeline provides a clear separation of two semantics important in the art research and allows further creativity through compositions of both in unseen ways. As it is just one possible way of studying the interaction between the *object* and the *style*, we expect further works could explore different forms of creative expression.

For example, one possible further work on computational creativity could be focused on the controlled generation of Ukiyo-eimages. Although we proposed to using sketch and color as *object* and *style* for image composition, they nonetheless could take other forms, such as categorical variables like a person's social status, gender, or the painter's style factors such as the art school. Furthermore, another research direction could be on interpretability in a cultural sense, where the association between styles and colure background can be revealed by exploring the generated images' factors. Finally, we also envision that creative work can jointly consider the information in multiple modalities. In this direction, one may consider the relation of painters as a desecrate graph with the paintings themselves as continuous images, combining graph analysis such as Graph Neural Networks (Scarselli et al. 2008) and image processing techniques.

Conclusion

In this work, we propose to bridge the machine learning and humanities research on the subject of Ukiyo-e paintings. Besides the presented dataset with coherent labels and annotations, we also show their value in the quantification approach to humanities research, Furthermore, we demonstrate that machine learning models in a creative setting could address art-style research problems.

Acknowledgement

We thank Hanjun Dai, David Ha, Yujing Tang, Neil Houlsby, Huachun Zhu, Justin Pinkney, and the anonymous reviewers for their comments and helpful suggestions.

References

[Akimoto et al. 2020] Akimoto, N.; Zhu, H.; Jin, Y.; and Aoki, Y. 2020. Fast Soft Color Segmentation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

[Aksoy et al. 2017] Aksoy, Y.; Aydin, T. O.; Smolić, A.; and Pollefeys, M. 2017. Unmixing-based Soft Color sSegmentation for Image Manipulation. *ACM Trans. Graph.*

[Art Research Center, Ritsumeikan University 2020] Art Research Center, Ritsumeikan University. 2020. ARC Ukiyo-e database (ARC所蔵浮世絵データベース), Informatics Research Data Repository, National Institute of informatics. `https://doi.org/10.32130/rdata.2.1`.

[IUS 2008] IUS. 2008. *Encyclopedia of Ukiyo-e (浮世絵大事典)*. Tokyo, Japan: Tokyodo Shuppan (東京堂出版).

[Karras, Laine, and Aila 2019] Karras, T.; Laine, S.; and Aila, T. 2019. A Style-based Generator Architecture for Generative Adversarial Networks. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

[Kobayashi 1994] Kobayashi, T. 1994. *Encyclopedia Nipponica s.v. Ukiyo-e*. Tokyo, Japan: Shogakukan.

[Lee et al. 2020] Lee, J.; Kim, E.; Lee, Y.; Kim, D.; Chang, J.; and Choo, J. 2020. Reference-Based Sketch Image Colorization Using Augmented-Self Reference and Dense Semantic Correspondence. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*.

[Lloyd 1982] Lloyd, S. 1982. Least squares quantization in PCM. *IEEE transactions on information theory*.

[Lvming Zhang 2017] Lvming Zhang. 2017. SketchKeras.

[Murakami and Urabe 2007] Murakami, M., and Urabe, J. 2007. A Quantitative Analysis of Portraits of Kabuki Actors. In *Proceedings of the Institute of Statistical Mathematics*.

[Pinkney and Adler 2020] Pinkney, J. N., and Adler, D. 2020. Resolution Dependant GAN Interpolation for Controllable Image Synthesis Between Domains. *arXiv preprint arXiv:2010.05334*.

[Pinkney 2020] Pinkney, J. N. M. 2020. Aligned Ukiyo-e Faces Dataset. Link.

[Renoust et al. 2019] Renoust, B.; Franca, M.; Chan, J.; Garcia, N.; Le, V.; Uesaka, A.; Nakashima, Y.; Nagahara, H.; Wang, J.; and Fujioka, Y. 2019. Historical and Modern Features for Buddha Statue Classification. In *SUMAC 2019*.

[Scarselli et al. 2008] Scarselli, F.; Gori, M.; Tsoi, A. C.; Hagenbuchner, M.; and Monfardini, G. 2008. The graph neural network model. *IEEE transactions on neural networks* 20(1):61–80.

[Sou Hasegawa 2020] Sou Hasegawa. 2020. Automatic Line Art Colorization. Link.

[Suzuki, Takagishi, and Kitamoto 2018] Suzuki, C.; Takagishi, A.; and Kitamoto, A. 2018. 'Collection of facial expressions' with IIIF Curation Platform - Close Reading and Distant Reading for Style Comparative Studies. *Proceedings of IPSJ SIG Computers and the Humanities Symposium*.

[Tian et al. 2020] Tian, Y.; Suzuki, C.; Clanuwat, T.; Bober-Irizar, M.; Lamb, A.; and Kitamoto, A. 2020. KaoKore: A Pre-modern Japanese Art Facial Expression Dataset. In *International Conference on Computational Creativity*.

[Tian, CODH, and ARC 2021] Tian, Y.; CODH, R.-D.; and ARC. 2021. ARC Ukiyo-e Faces Dataset. (Created by Yingtao Tian, ROIS-DS CODH; Collected from ARC) DOI: `https://doi.org/10.20676/00000394`. URL: `https://github.com/rois-codh/arc-ukiyoe-faces`.

[Wang et al. 2018] Wang, T.-C.; Liu, M.-Y.; Zhu, J.-Y.; Tao, A.; Kautz, J.; and Catanzaro, B. 2018. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*.