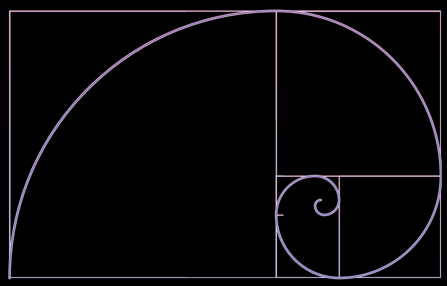
HOW MIGHT **HUMAN EXPERIENCE** INFORM **AI SYSTEMS**

Shuyue Jin
02/2022 - 07/2022
TU Delft
Design for Interaction



CHAIR Derek Lomas
MENTOR Jun Wu

AUTHOR Shuyue Jin

Faculty of Industrial Design Engineering
Delft University of Technology

PREFACE

I am a person who is curious about the world. When I first learned about the profession of a UX designer as an undergraduate, I found that it gave me unlimited space to develop this curiosity. Because what we are designing can influence people's future lives. During the two years of study in TUDelft, the rigorous study of design theory and design training enable me to explore the future world. As the learning goes on, I become more curious about the future that designers can bring to people. I am wondering - as a designer, how can we bring people a better future?

This thesis is my attempt to answer this question. Combing my design background in Human-Center Design and my passion for the AI field, I tried to link human experiences and AI systems. Through this, I think more emotionally compelling AI systems will be produced and a better future world can be built.

The process of the past 100 days has been more complicated than I imagined. Whether it is the technical exploration or the establishment of a huge system, I have felt unprecedented difficulties. I would like to take this opportunity to express my gratitude to everyone who accompanied me to complete this project.

First of all, I would like to thank my supervisor teams. Thanks to my chair, Derek. Whether it is the overall planning of the project, the revision of the project direction, or even the storyline of the presentation. He always gave me plenty of valuable feedback. More importantly, he has ignited my enthusiasm for the AI design field, and his continuous encouragement has brought me closer to the answer. Thanks to my mentor, Jun. He always put forward a lot of challenging and interesting questions that enable me to think more deeply. His perspective makes the whole project logical and fruitful. At the same time, his rigor and seriousness make me feel the charm of doing scientific research. If there is an opportunity in the future, I am still willing to continue to explore the field of research.

Second, thanks to Frederik Ueberschär. His amazing design LANDSHAPES is the starting point of this project. His explorations provide a lot of material for my designs. He is always eager to answer all kinds of questions about the project for me. He helps me get to know the technology and the background part of the project. Thanks to the team members of GANs Aesthetics, Willem van der Maden, Betul Irmak Celebi, Ton Hoang Nguyen, Moshiur Rahman, and Joseph Catlett. In my cooperation with them, I feel their professionalism. And it is the accurate data they provide me that allows me to finally build the entire output.

Thanks to my friends, it is their company that gives me the courage to move forward every time I am disappointed. My life has also become interesting and enjoyable because of meeting them.

Thanks to my family. Because of their support, I can come to TUDelft for my education and have the opportunity to explore the world.

Finally, thanks to my grandparents. It is the regret of my life that you have not been able to witness my graduation. You took care of my growth and taught me the importance of learning, and it is also because of your remindings that I have always believed in the meaning of studying hard. I hope that the fragrance of tulips on the day of my graduation can float into the sky so that you can also smell it and be proud of me.

Two years have passed, I am still curious about the world. After leaving TUDelft, I believe that I will carry more different roles in this world, and in each role, I will continue to explore - how to make a better world.

Shuyue Jin

金姝玥.

25/07/2022

EXECUTIVE SUMMARY

Artificial intelligence has brought changes to the human world. However, many aspects of human experience, like aesthetic beauty, are not accessible to AI systems. The gap between them makes AI systems unable to use the human experience to improve AI capabilities. People's experience in those systems is also damaged due to this gap. Meanwhile, as more designers dive into the AI field, AI product designers play more crucial roles in shaping AI systems. Are there any opportunities for them to become a translator between the qualitative human world and the quantitative AI world?

This graduation project focuses on the GANs(Generative Adversarial Networks), an AI system that can generate fake output for human society, sets the target users as designers working in the AI field, and identifies elements that affect aesthetic beauty. I investigate how to enable designers to efficiently inform GANs of people's feedback and then steer their output.

It originates from LANDSHAPES, designed by Frederik Ueberschär. It is an interactive exhibition piece. Its contents are AI-generated aerial landscapes generated by GANs. It proves GANs' potential influence on people's experiences and the designer's important role in AI products.

Desk research and practical operations in GANs show that GANs have the potential ability to "understand" human experience according to their plenty of parameters like latent space and discriminator. The qualitative research on the human experience of GANs' output revealed the different factors influencing aesthetics. Some factors are abstract(such as the memory evoked by images), while other factors(such as contrast and saturation) have the potential to be quantified. Also, the research about the target user - AI product designers are conducted to make the final methods more practical.

Based on those insights from three fields, two hypotheses, including retraining with the highly-rated images and building new computational models to iterate the factors, are put forward to inform the AI system of human experience.

Finally, two approaches were built. The first one is the CURATION APPROACH - Putting the beautiful GANs' outputs selected by people into the input dataset and retraining the GANs. The second is the ALGORITHMIC AESTHETICS APPROACH - transferring the factors from people's experiences to the algorithms and designing the computational models that can predict human ratings of beauty. The evaluation results show their validity in informing AI systems about the human experience.

Moreover, LANDSHAPES also shows that the videos generated by GANs can evoke people's emotions. So, an investigation in the video is conducted to investigate some future directions for designers to inform AI systems about the human experience.

By exploring the approach to inform AI systems about the human experience, the project provides simple and effective methods for AI product designers to become the professional translator between the human and the AI world. It also proves the importance of the human experience for AI and becomes a starting point to use the human experience to improve AI systems.

CONTENT

ABSTRACT

01 INTRODUCTION

Project Introduction /01
Problem Definition /03
Approach /05
Motivation /08
Reader Guide /09

02 EXPLORATION IN GANs

Why choose GANs /13
What are GANs /15
The Factors Of GANs /24
How To Measure It /29
The Experiment Of GANs /33

03 RESEARCH ON HUMAN EXPERIENCE

Exploration In Aesthetics /49
Qualitative Research Process /53
Rating results /56
Factor Analysis /57
Research on the evaluation criteria /73

04 TARGET USER

Target user in the project /79
Why we choose AI product designers /79
The characteristics of AI product designers /89

05 HYPOTHESES

The highly-rated images will improve the aesthetic quality of the output images /83

Computational models can be designed to predict human ratings of beauty. /86

06 FINAL APPROACHES

CURATION APPROACH /91

ALGORITHMIC AESTHETICS APPROACH /113

07 QUALITATIVE RESEARCH ON CHANGING SYSTEMS

The importance of changing systems /145

GANs' features influencing videos /146

Qualitative Research /149

08 CONCLUSION

Summary /163

Address the questions /165

Contribution /167

Future work /169

REFERENCE /171

ABSTRACT

The problem we want to address

AI(Artificial intelligence) is diving into people's lives as its algorithm continues to iterate. However, the algorithmic and quantitative systems do not seem to access people's experiences, which always include emotional and qualitative factors. How can an AI system understand people's feelings? How can an AI system like GANs optimize for meaningful human values, like beauty?

Our approach

In this project, I mainly build two approaches for designers to translate the human experience into GANs. Starting from the exploration of GANs, I investigate their potential chances to "understand" human experience. Then, the qualitative research on human experience reveals the aesthetic factors influencing people's feelings about GANs' output. Through the interviews, some pain points from AI product designers are concluded. Combining the exploration results in those three scopes, I hypothesize that retraining with the highly-rated images and building new computational models to iterate the factors from the human evaluation can help designers inform the AI system about the human experience. I cooperated with CSE students and proposed two approaches- CURATION APPROACH and ALGORITHMIC AESTHETIC APPROACH- based on the above assumptions. The build process demonstrates the achievability of the assumptions. In the evaluation part, the retrained images in both methods are rated higher than the original ones. Both approaches are proved to be practical and feasible.

Our result

For the curation approach, all produced models with highly-rated images outperformed the original model. The original model's score is 0.2. When replacing 500,1000,2985 highly-rated images into the input dataset, their scores increased to 0.245, 0.269, and 0.255.
For the algorithmic aesthetic approach, we select three categories of images(coastline, forest/desert, arctic). In each category of images, three factors are selected to improve. Among the nine groups whose corresponding factors are iterated by algorithms, eight groups' new images are rated higher by people than the original images. The curation and algorithmic aesthetic approaches are verified to help inform AI systems of human experience.

Potential impact

Following the curation and algorithmic approaches, designers can successfully inform GANs about people's aesthetic evaluation of their images. Using the human experience to improve AI systems is a starting point. It also proves the importance of the human experience for AI and provides a template for designers in all AI fields to inform their AI systems of human experience.

01

INTRODUCTION

1.1 PROJECT INTRODUCTION

1.2 PROBLEM DEFINITION

1.3 ASSIGNMENT

1.4 MOTIVATION

1.5 READER GUIDE

1.1 Project Introduction

Artificial intelligence has brought changes in all areas of people's lives. However, many aspects of human experience, like aesthetic beauty, are not accessible to AI systems like Generative Adversarial Networks (GANs). AI systems like GANs have no chance to consider it.

As we all know, AI systems are built by plenty of computational algorithms. At the same time, human experience always contains individual and subjective feedback. AI systems- such algorithmic and quantitative systems do not seem to have access to human experiences, which always include emotional and qualitative factors. How can an AI system understand people's feelings? How can an AI system like GANs optimize for meaningful human values, like beauty?

At the same time, as more and more designers invest in AI product design, AI product designers play an increasingly important role in shaping how AI systems develop. Unlike ordinary designers, AI product designers can combine expertise in HCD(Human-center design) and AI fields to improve AI systems from the user and technical aspects.

Therefore, how can AI designers efficiently inform AI systems of the user experience and improve their products' aesthetics? Are there any methods for each designer to become an accurate and professional translator, translating qualitative human experience into quantitative AI systems?

This project focuses on the GANs to identify elements that affect aesthetic beauty and investigate how designers can efficiently inform GANs of the feedback from people to steer their output. Approaches to help designers inform the AI system are proposed. They connect AI systems to the human world, demonstrating the importance of user experience to AI systems. In the future, people can use them as a starting point to explore methods for different AI system user experiences.

Artificial Intelligence has impacted people's lives in various fields like AI Music and AI artworks. This graduation project limits the scope to the different images generated by GANs. It aims to investigate how to accurately transform the human evaluation of aesthetic beauty to GANs and enable GANs to use people's feedback efficiently. While improving AI systems' products, the user experience's significance for AI systems will be proved.

There are mainly three aspects(Figure 1) to the problem that we want to solve:

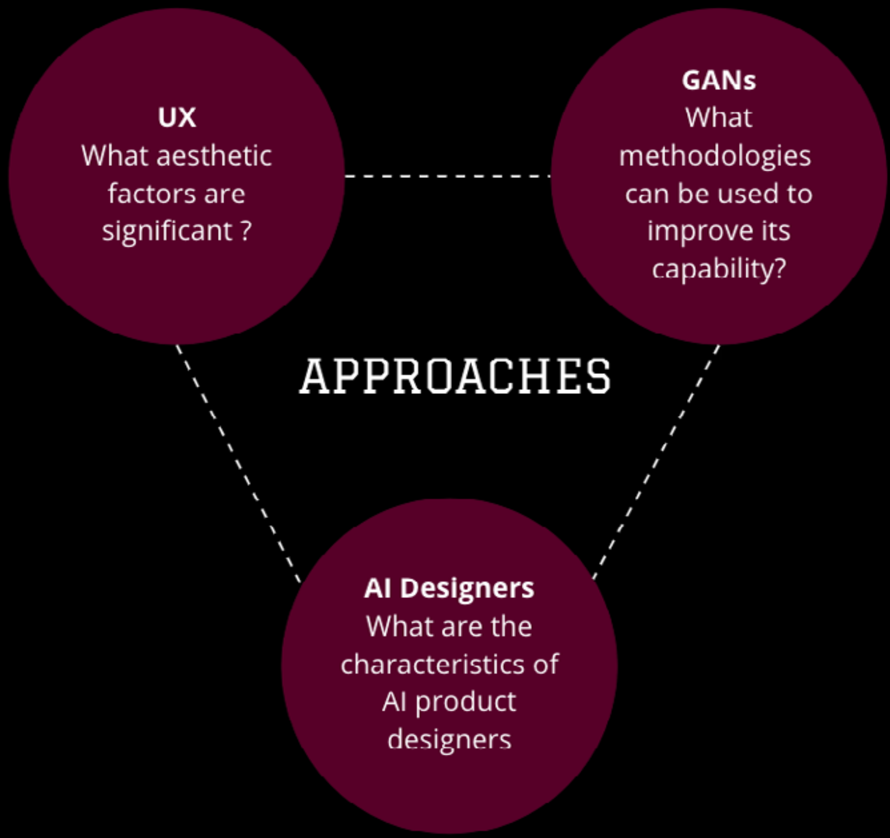


figure1. The scope of the project

1. For GANs:
What information is valuable and essential to iterate their ability and produce better works? How can the user experience influence AI systems in a more efficient way?

2. For human experience:
When people evaluate the aesthetic beauty of images generated by GANs, what factors are most significant for them? What are the key elements affecting the user experience for the aesthetics and beauty of the images from GANs?

3. For our target user - AI product designers:
What are the characteristics of AI product designers? How can an AI product designer fully inform GANs of the outcome from users? How can designers translate the qualitative evaluation from people to quantitative factors and improve GANs' work? What role should a designer play in the translation process?

To tackle this challenge, the project has four phases(Figure 2).

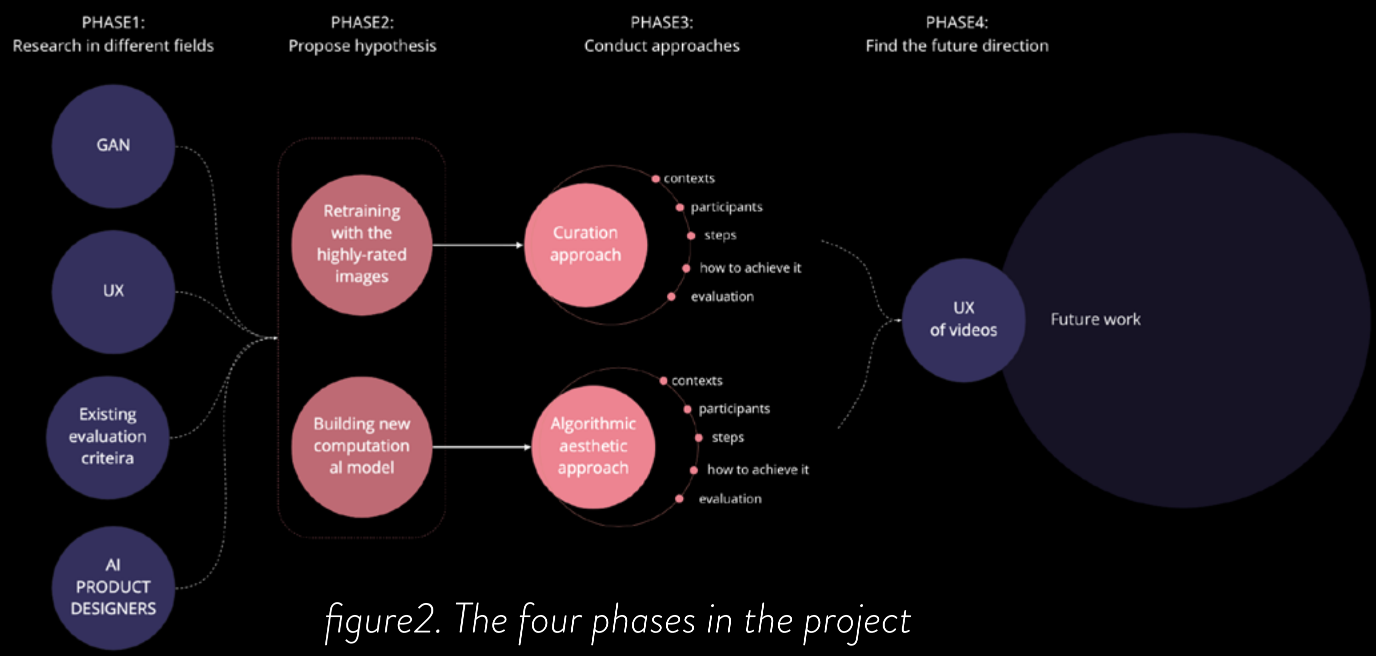


figure2. The four phases in the project

The first phase is mainly to dive into the three scopes of the projects. Investigate the features of GANs, the factors of the AI system's user experience, and the characteristics of target user groups.

This graduation project originates from LandShapes, an interactive exhibition piece to evoke people's awareness of climate change by Frederik Ueberschär. In LandShapes, he designs with GANs to generate fake images and videos for landscapes. Landscapes show us the GANs' ability to influence human experience. Also, Frederick's design process with GANs shows us the designer's deep involvement in the operation of GANs. Therefore, GANs can be an excellent beginning to explore the translations between AI systems and the human world for AI product designers.

In the first phase, desk research and practical operations were conducted to understand GANs in-depth and investigate the potential factors that can be utilized in the algorithmic systems. What is more, considering the function of GANs, generating fake images for human society, the context of human experience in this project is defined as an evaluation of image aesthetics. Further research for aesthetics helped narrow down a clear and clarified scope of aesthetics. Next, by conducting workshops for users to evaluate the images, the qualitative research revealed the different factors influencing aesthetics. There are factors like the memory behind the images, which are very subjective and individual. Some factors, such as contrast and saturation, have the potential to be quantified. Moreover, research about the target user - AI product designers- was conducted to create more practical methods.

In the second phase, combining the insights from the first phase, two hypotheses, including retraining with the highly-rated images and building a new computational model to iterate the factors, are put forward to inform the AI system's human experience.

Based on the hypothesis, two approaches were built cooperating with CSE students. The third phase is the instruction of the approaches. Following them, designers can complete the approaches quickly and confidently. This part explains approaches' originals, the context that fits them, their steps, their realization process, and their evaluation. Also, their concerns and limitation are mentioned to enable the audience to use them with a more cautious mindset.

The first is the CURATION APPROACH - by putting the beautiful GANs' outputs selected by people into the input dataset and retraining the GANs, GANs can successfully get people's feedback and steer their output. The second is the ALGORITHMIC AESTHETICS APPROACH - by transferring the factors from people's experiences to the algorithms and designing the computational models that can predict human ratings of beauty, GANs can understand people's experiences and improve their output.

Last but not least, the context of the project is the image, a static output from GANs. However, GANs' output also contains changing systems like videos. This kind of output is also an essential component in translation between humans and AI designers working in the AI field. So, in the fourth phase, an investigation in the video was conducted to show some future direction and potential for designers to inform AI system human experience.

1.4 Motivation

As an AI Designer who will explore the field of AI in the future, two years of study in the field of AI design in TUDelft allowed me to dive into this field. Moreover, the work experience in the AI design department of Alibaba also made me look forward to the changes AI will bring to people's lives.

With the development of AI technology, the products produced by AI, like AI artworks and AI music, will gradually penetrate people's lives. In the ITD course, I had the opportunity to learn about GANs. We leverage GANs to provide designers with inspirational images and mood boards that can be generated based on their ideas. This process made me feel the powerful ability and value of GANs in generating pictures and videos.

However, as a UX designer with a massive passion for AI systems, I noticed that the current exploration of AI is mainly at the algorithmic level. People's experiences and feelings about the information generated by AI seem to be ignored. Therefore, I hope to start from GANs to explore the importance of people's experience in the AI system. By enabling designers to inform AI systems of human experience, I believe it will iterate more evocative and emotionally compelling AI systems.

1.5 Reader Guide

This thesis is a comprehensive summary of the project. This subchapter introduces the structure of the thesis and guides readers through the crucial components of the following chapters.

Chapter 2 explains the knowledge of GANs, and the representation of the AI system in this project. It summarized the findings of literature research and experiments in GANs. Readers will know the fundamentals of how GANs work and their potential parameters for understanding human experience.

Chapter 3 mainly answers the question, "what influences human experience in GANs." It
presents the results of the qualitative research on humans and identifies the aesthetic factors that influence people's experience in GANs' output.

Chapter 4 shows the characteristics of our target user- AI product designers. It interprets the context for designers to translate the human experience to AI systems.

Combining the results in the above chapters, two hypotheses about how human experience might inform AI systems are shown in chapter 5. This chapter explains the content of the hypothesis and the specific reasons for making the hypothesis.

Chapter 6 introduces the outcome of this project- two approaches that enable designers to inform AI systems about the human experience. This chapter is a manual for those approaches. Following them, designers can complete the approaches quickly and confidently. This part explains approaches' originals, the context that fits them, their steps, their realization process, and their evaluation. Also, their concerns and limitation are mentioned to enable the audience to use them with a more cautious mindset.

Chapter 7 displays the results of the qualitative result of human experience in fake videos generated by GANs. Changing systems is also essential in translation between humans and AI designers working in the AI field. So, in chapter7, the investigation in the video shows the future direction and potential for designers to inform AI system human experience.

Chapter 8 gives a conclusion, including the summary of the project, how the project addresses the initial questions, contribution, and future work.

02

EXPLORATION IN GANS

2.1 WHY CHOOSE GANs

2.2 WHAT ARE GANs

2.3 THE FACTORS OF GAN

2.4 HOW TO MEASURE IT

2.5THE EXPERIMENT OF GANs

2.1 Why choose GANs

2.1.1 What are Generative Adversarial Networks

This graduation project originates from LandShapes, an interactive exhibition piece to evoke people's awareness of climate change through the changing process of images by Frederik Ueberschär. In LandShapes, all images and videos are fake earth landscapes generated by GANs.

On the one hand, this project proves that GANs can change people's lives with their output. On the other hand, Frederik's design process with GANs shows that GANs have many spaces for designers to participate in their operation. It is a good starting point to explore how to inform AI systems of human experience and steer their output.

The specific reasons are as follows:

1. GANs' output can influence people's experience
In Landshapes, when people see the images and videos generated by GANs, different experiences are brought, including "Aesthetics," "Experience Narrative," "Emotional Response," and "Negative emotions." The different experiences evoked by GANs show AI systems' ability to influence the human world. However, Landshapes does not mention whether those experiences can influence GANs. If human experience can be informed GANs, whether GANs' outputs can be changed?

2. The output from GANs is easy to explore
The outputs from GANs are images, videos, and music. Unlike the output from other AI systems, those kinds of output features are simple, and we can measure and explore them more easily. For example, NLP's(Neuro Linguistic Programming) results are the "information" based on understanding people's thinking, feeling, language and behavior. The outcome is too abstract and comprehensive to explore. However, many measurements have been put forward for images and videos to explore their human experience. So it is more practical to solve how human experience informs AI systems from GANs.

3. Designers can be deeply involved in the operation of GANs

Frederik's design process with GANs shows the deep involvement of designers in the operation of GANs. Frederik completes the process, including gathering datasets, training GANs, and exploring GANs' output individually. When using other AI systems like OCR (Optical Character Recognition), designers often stay in the design stage outside the system. For GANs, designers can directly use it as their design tool to solve problems. It means that designers have more opportunities to translate human experience into an AI system in this system.

4. The lack of instruction that leads designers to improve GANs' output

During Frederik's design process, he found some factors that influence the quality of images, such as the input dataset and the number of iterations. Nevertheless, because of the lack of appropriate methods, he can only process the data based on his experience. Meanwhile, the process is time-consuming and complicated. For example, designers must rely on their judgment to remove bad images from the input dataset. Designers waste much time, and these messy processes cannot guarantee that the final product will meet the user's preferences.

According to the characteristics of GANs and the participation process of designers, GANs are finally chosen as the starting point to explore the question - how might human experience informs AI systems.

2.2 What are GANs

2.2.1 What are Generative Adversarial Networks

A generative adversarial network(GANs) is a class of machine learning frameworks designed by Ian Goodfellow and his colleagues in June 2014. It is an unsupervised learning approach and contains three parts(Figure 3).

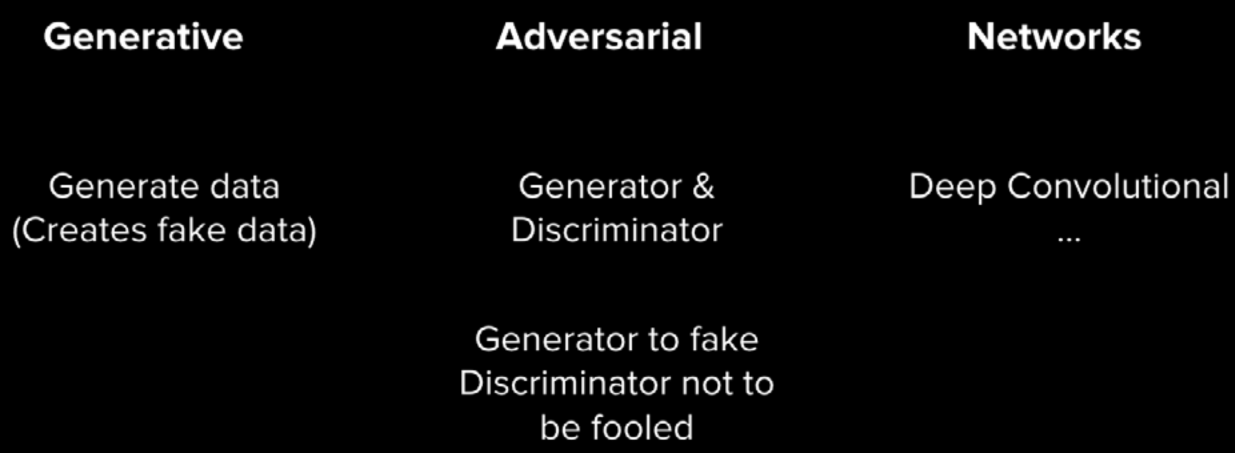


figure3. The structrue of GAN

GENERATIVE
Generative means it can generate data and create fake data.

ADVERSARIAL
The adversarial part is the key to this system. It contains two networks - generator and discriminator.
The generative network generates candidates while the discriminative network evaluates them. Based on the latent space, the generative network learns to map to a data distribution of interest. Then, the discriminative network distinguishes candidates produced by the generator from the true data distribution.

NETWORK
The network stands for network Neural Network. A neural network can be understood
as a network of hidden layers that try to mimic the working of a human brain. For the GANs, the network can be deep convolutional.

2.1.2 How does it work?

The overview architecture of GAN is shown in the figure4.

First, there will be a dataset that includes the real images. At the same time, random noise seeds, which we call latent vectors, go into the generator network. This random noise goes into this generator network, and then the network can scale this to create a two-dimensional image.
In the beginning, the fake images are not realistic, and they just look like noise. Furthermore,
it is straightforward for the discriminator to distinguish between counterfeit and real images.
As the learning goes, the generator gets better until the discriminator does not know if
this is real or fake. We are looking at discriminator and generator loss as the network

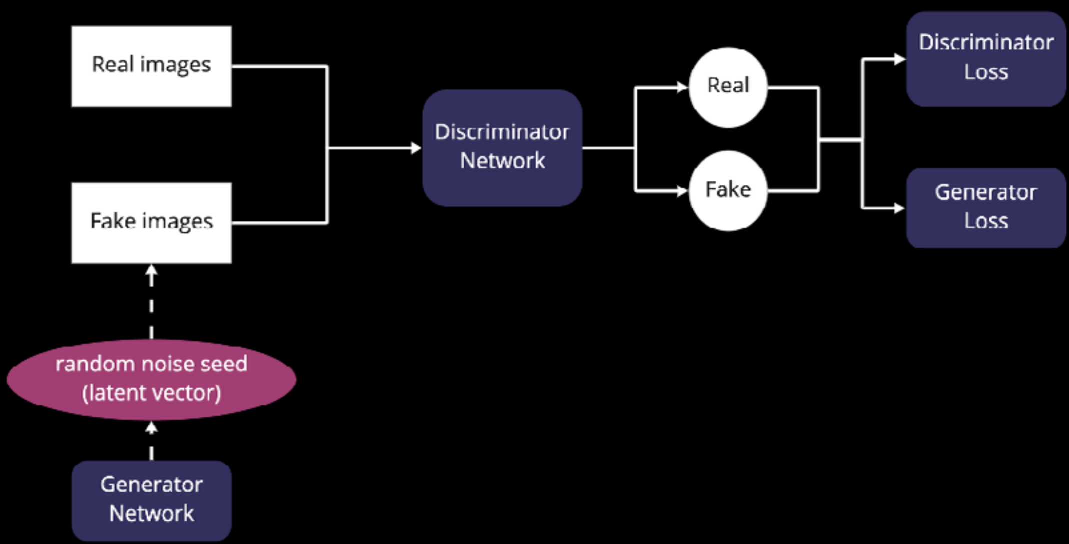


figure4. The overview architecture of GAN

2.1.3 How to train it

The training process contains two phases.

1. First phases - train the discriminator

In the first phase (figure5), we train the discriminator and freeze the generator, which means that the training set for the generator is turned false. The network will only do the forward pass, and they will not be any backpropagation. Then, it is time for the discriminator to be trained with real data and check if it can predict them correctly. And the same with fake data. It needs to identify them as fake.

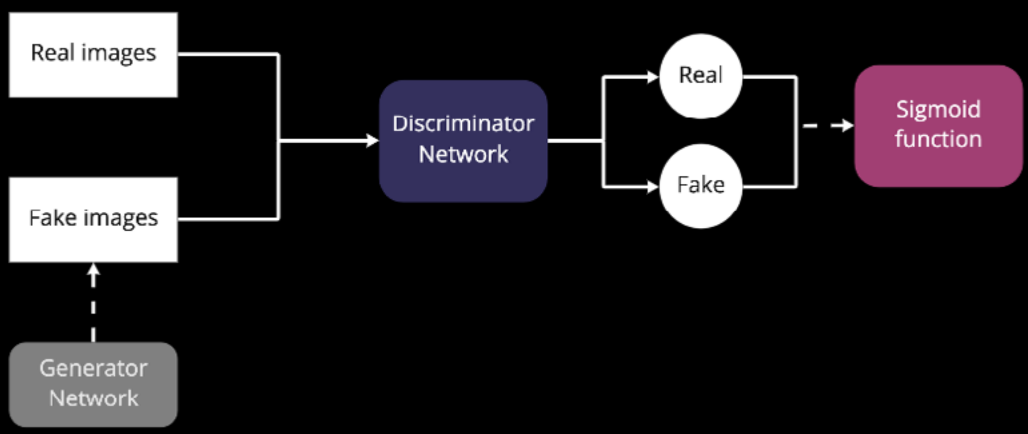


figure5. Phase 1 of the GAN

2. Second phases - train the generator

In the second phase(figure6), we train the generator and freeze the discriminator. After we get the result from the first phase, we can use them to make a better output from the previous state and fool the discriminator.

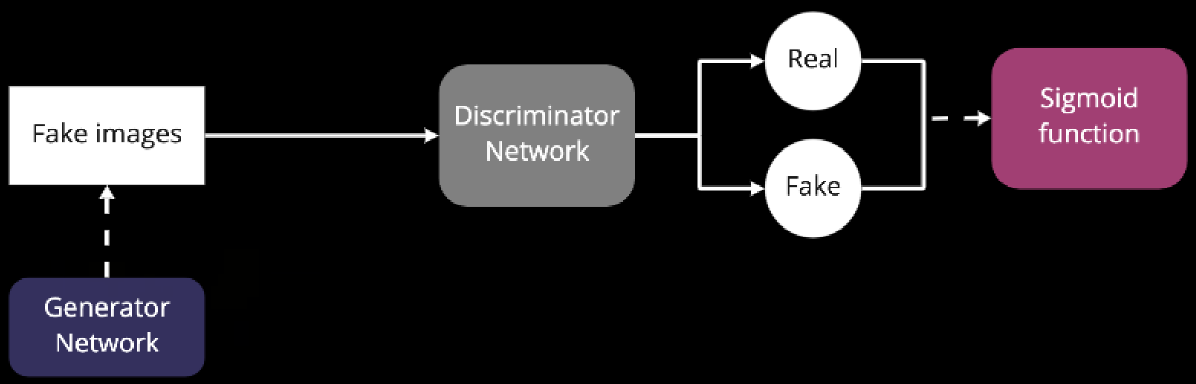


figure6. Phase 2 of the GAN

2.1.4 How to apply it in the practical context

How can we use GAN in our practical context? Usually, there will be six steps for us to train the GAN.(figure7)

1.Define the problem

In the first step, we have to define the problem that the GAN need to solve and collect the data for our context.

2.Choose Architecture of GAN

After we have selecte our context , we should define the architecture based on the application. For example, if the application is super resolution to generate higher resolution images based on low resolution images, we need to work with Khan neural networks.Usually, it will be better to take something exists like ECG and then build our GAN architecture based on them.

3.Train Discriminator on Real data

During this process, we train the discriminator with real data to predict them as real for a number of times which we call it **epoch**.

4.Train the generator to fake the data that can fool this discriminator

In this step we are going to generate the fake samples from the generator.

5.Train Discriminator on Fake Data

Then we train the discriminator on fake data. We're going to train the discriminator to predict the generated data as faith so that's how we know that discriminator is actually predicting the values as correctly.

6.Train Generator with the output of discriminator
After getting the discriminator predictions, we train the generator to fool the discriminator.

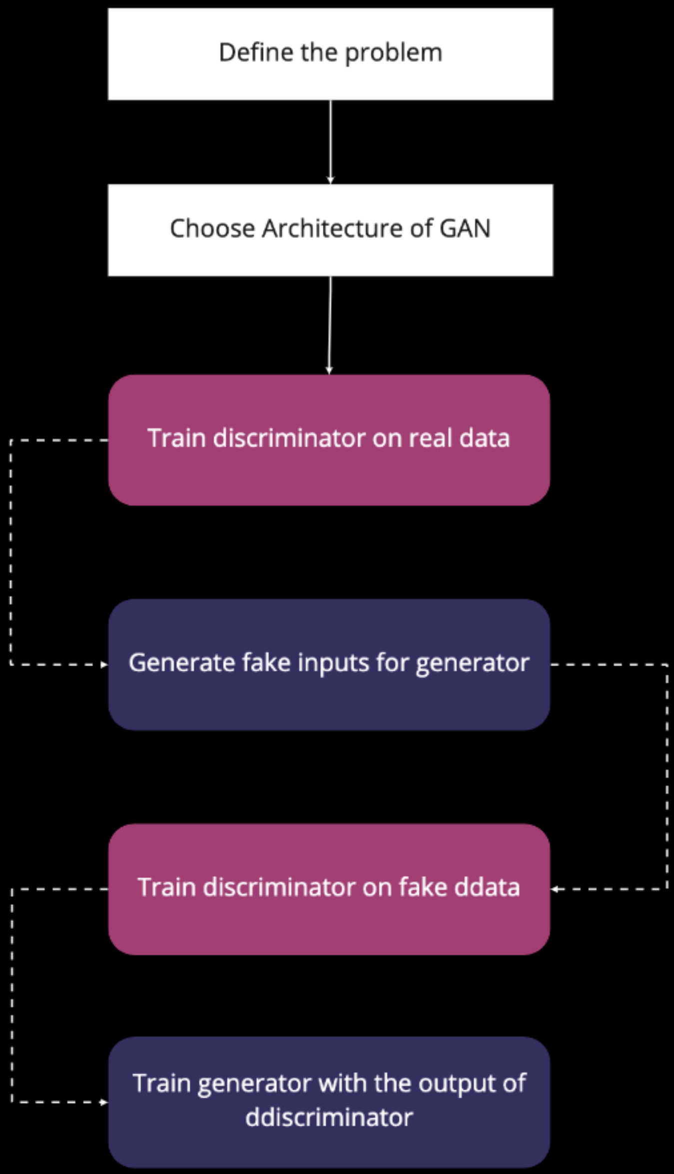


figure7. The process of the GAN's appication

TAKE AWAY

1. Input dataset is the "inspiration" for GANs to generate fake images.
Like an oil painter who needs the real landscape as inspiration, GANs also use the real image as their input dataset to generate fake images. Those real images provide GANs features to learn. So the input dataset is significant for GANs.

2.GANs are built by a lot of algorithmic components
Different networks and algorithmic frameworks in this AI system provide designers and developers many opportunities to translate the human experience.

2.3 The Factors of GANs

The following concepts are the factors that will affect the training process of GANs. I want to figure out if some factors in the system of GANs are similar to the elements in the human world. Find the potential ways to inform GANs human experience.

2.2.1 LATENT SPACE

Definition
A high-dimensional space that encapsulates all learned features from real data

Value
The positions that are close to each other tend to be more similar.

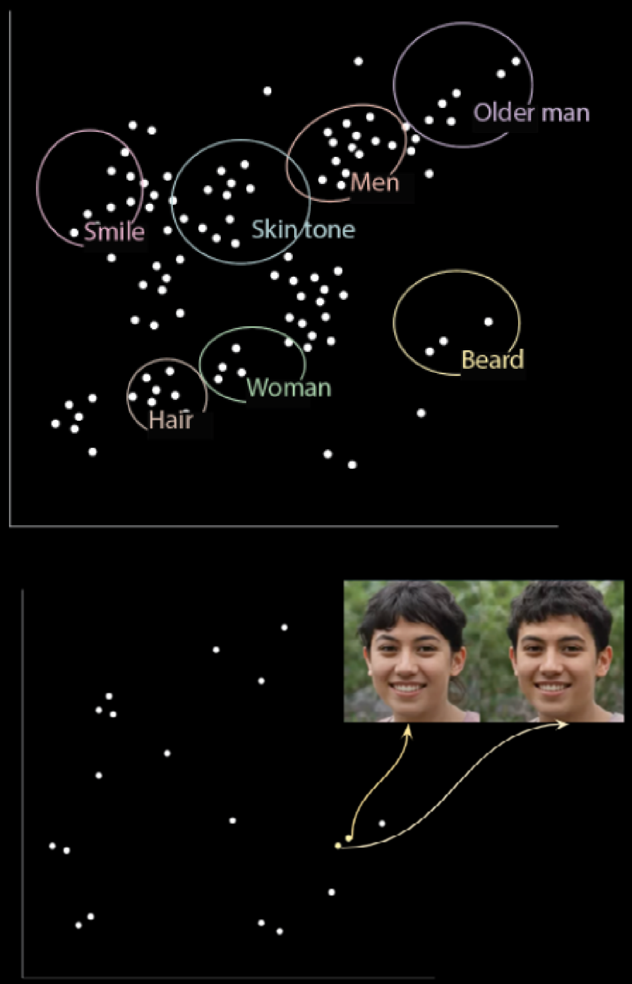


figure8. The definition of latent space

Latent space is one of the most significant concepts in GANs. It is a high-dimensional space encapsulating all learned features. Model in GANs is classifying certain features and putting them in space.

In deep learning, latent is the core because it can learn data features and simplify data representations to find patterns. In the GANs, latent space is also the central core. As we all know, GANs aim to generalize training data features to produce new similar images in the GANs. We generate images using the GANs generator, and the input to that is a latent vector. Through adversarial training, GANs learn the mapping from a latent space to real data distribution. After learning such a nonlinear mapping, GANs can produce photo-realistic images by sampling latent code from a random distribution.

For example, in StyleGANs2, the latent space is 512-dimensions. Different positions represent the various features. Each of those dimensions is a particular feature of an image. With a dataset of humans, it could mean something like a smile, skin tone, men, and older man(figure6). Also, it could be a background color.
Also, in the latent space, the positions close to each other tend to be more similar, and positions further away tend to be more different.

2.2.2 TRUNCATION

Definition
Truncate the latent space. It is a parameter that looks at real-ness.

Value
The smaller:, the more realistic & less diverse
The bigger:, the less realistic & more diverse

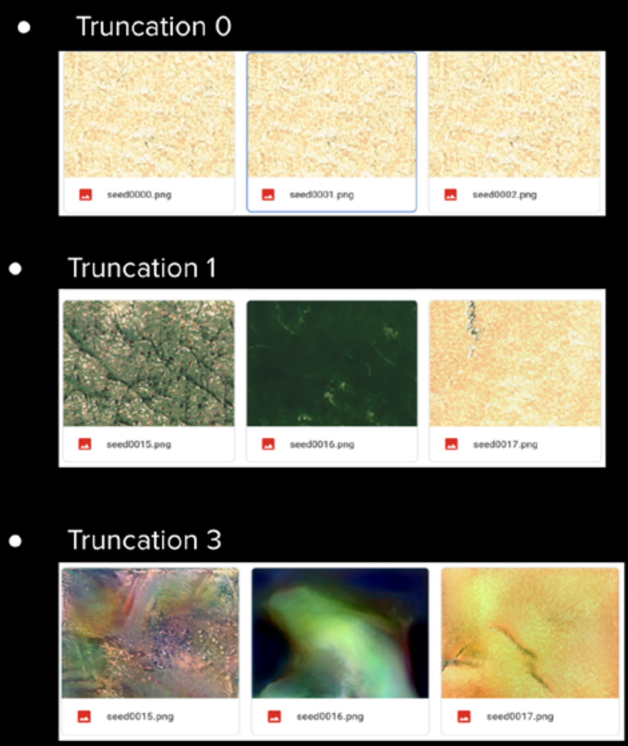


figure9. The output with different truncations

Truncation is to truncate the latent space. It is a parameter for realness. Depending on our value, it can have a subtle or dramatic effect on the images. The smaller the number, the more realistic images should appear. Nevertheless, this will also affect diversity. Most people choose between 0.5 and 1.0, but technically it is infinite.

TAKE AWAY

GANs have plenty of factors.
Like an oil painter with a set of painting equipment, GANs include various algorithmic factors like latent space, Kimg, generator, discriminator, etc. Combining its complex algorithmic frameworks, they provide designers and developers methods to inform GANs human experience. Among these factors, latent space has the most potential to be one of our methods. Because it represents the features of the image, it means that by changing it, the quality of the image can be changed directly.

2.5 How to Measure it

As an AI System with complex algorithms, it is not easy to conduct a complete and practical evaluation of the effect of the entire GANs model. By researching the measurements of GANs, I want to figure out if there are some measures related to output aesthetics that can be used to evaluate human experience?

2.3.1 Why evaluate it is so hard?

In supervised image classification, evaluation is straightforward. We have to compare the predicted output to the actual output.
However, we pass in some random noise with GANs to get this fake(generated) image. We want this generated image to look as real as possible. So, it is hard to quantify the realism of this generated image without any comparison.

Until recently, there have been various evaluation measures to measure the models in GANs. Most of them are quantitative and analyze the outcome based on computable factors. Some qualitative methods aim to get users' feedback and measure the output's quality. Here, I have listed some effective metrics currently used to measure GANs.

2.3.2 FID : Frechet Inception Distance

Definition
Score the generated images compared to your results

Value (0-300)
The smaller: better image quality and diversity
The higher:, the less realistic & less like the results dataset

Evaluation Metric
Fidelity: The high quality of the imagesthat we want our GAN to generate.
Diversity: Our GAN should generate images that are inherent in the training dataset.

Exact factors
Pixel Distance: This is a naive distance measure where we subtract two images' pixel values.
Feature Distance: We use a pre-trained image classification model and use the activation of an intermediate layer. This vector is the high-level representation of the image. Computing a distance metric with such representation gives a stable and reliable metric.

TAKE AWAY

1. The criteria for those methods are limited

The quantitative measurements mainly compare the input(real images) and output(generated images). The input dataset is primarily used as the reference standard for product evaluation. On the one hand, the aesthetic of the input has not been verified; on the other hand, the criteria lack research on the properties of the image itself.

2. Whether they are consistent with human judgments

The current measurement is studied from the perspective of fidelity and diversity. The degree of matching between these measurements and human judgments has not been verified. User experience plays almost no role in these measurements.

3.How to improve the score?

Those measurements seldom mention how to improve the scores. The metrics that come with GANs, such as FID, did not say how to improve the training process and results based on these scores, and most of them just put forward evaluation methods and standards.

2.5 The Experiment of GANs

2.5.2 THE EXPERIMENT OF SYTLEGAN2

It mainly describes the first experiment of STYLEGAN2.

Tool
StyleGAN2-ADA

Goal
- Get familiar with and understand GAN
- Analyze the three aspects - input, training process, and output
- Find whether there are factors that affect the aesthetic of the final output generated by GAN in those three aspects

Input

In this experiment, the context of the selected input dataset is landscape.

Image type: landscape
Number of original pictures: 2000
Image size: imageDimensions = ‘1024x1024’

Image quality:

I obtained the image via the Google engine, based on the geographic location randomly selected by QGIS on the map. Finally, the Google engine successfully obtained 2000 pictures.
In the process of this experiment, in order to ensure the **randomness of the final result**, no processing was performed on the original data.

Evaluation of the input

Through the analysis of the input picture, we found that the landscape pictures obtained by this method mainly have the following characteristics.

1.Content for images

1.1 Many types

Because of the randomness of the selected locations, the final images cover a wide range of types. Also, the color, contrast, etc., of every kind of pictures are very different. The elements contained in each image are pretty other.

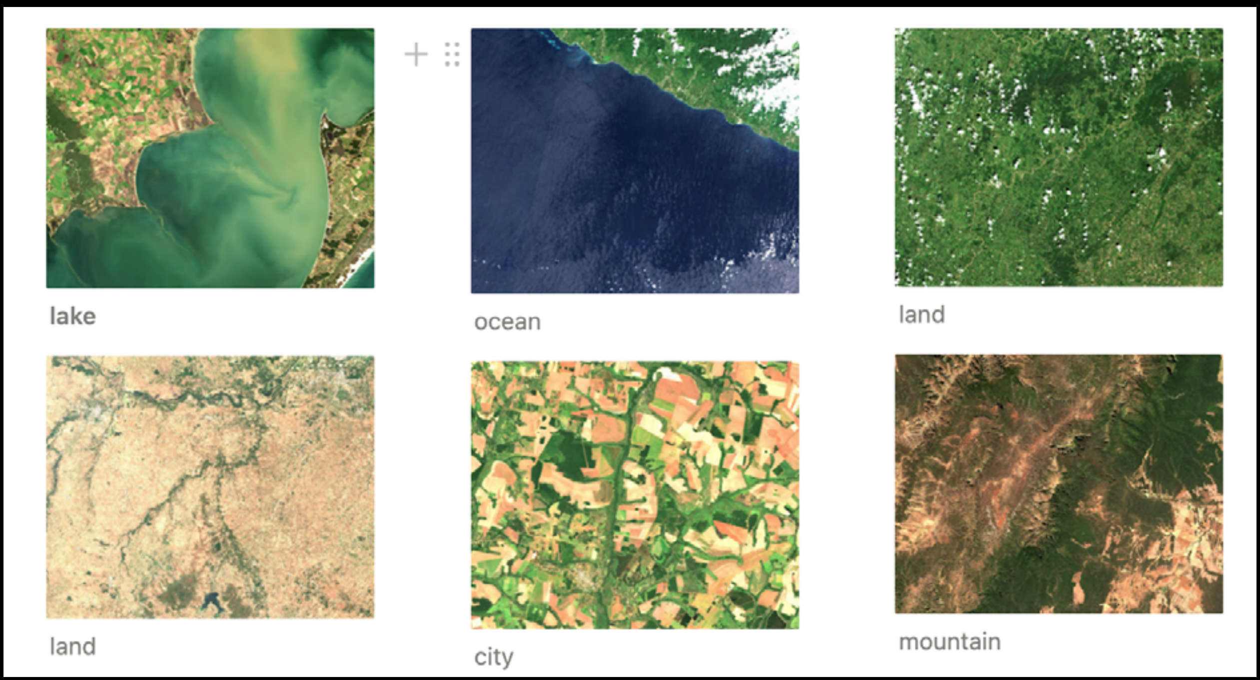


figure12. The content of the input

1.2 Similar pictures of the same type

Although the differences between different types of pictures are relatively significant, for the same kind of pictures, the differences between each image are not very large. It is mainly a simple texture change, and there is not much difference in color value and contrast.

E.g.:
The following are pictures of the type of City. The colors that make up the pictures are mainly yellow and green; the degree of the contract is not significant. The main difference is in the distribution of different colors.

Similarly, for the image of the ocean, the color of the sea is mainly dark blue; the color of the island is primarily green. The main difference is the size of the color area.

figure13. One type of images

2. Features of the picture

2.1 The color of the picture is relatively single

Since the selected dataset is the landscape that exists on the earth extracted from the earth engine, compared with datasets such as artworks and plants, the colors contained in this dataset are mainly blue, green, yellow, and white. Furthermore, it is primarily the splicing of color blocks and less processing of colors such as gradients.

figure14. The color for one image is single

2.2 The pictures are repetitive

Among 2000 randomly obtained geographic locations, if the geographic locations are relatively close, there will be many duplicate images in the finally got landscape pictures. The presence of these images reduces the diversity of the dataset. In this experiment, these same pictures were not processed to ensure the randomness of the obtained pictures.

figure15. Repetitive images

2. 3 The dataset has a lot of blank images.

Among the 2,000 randomly obtained geographic locations, statistics found that about 500 images were blank. The existence of these images will also affect the training of GANs.

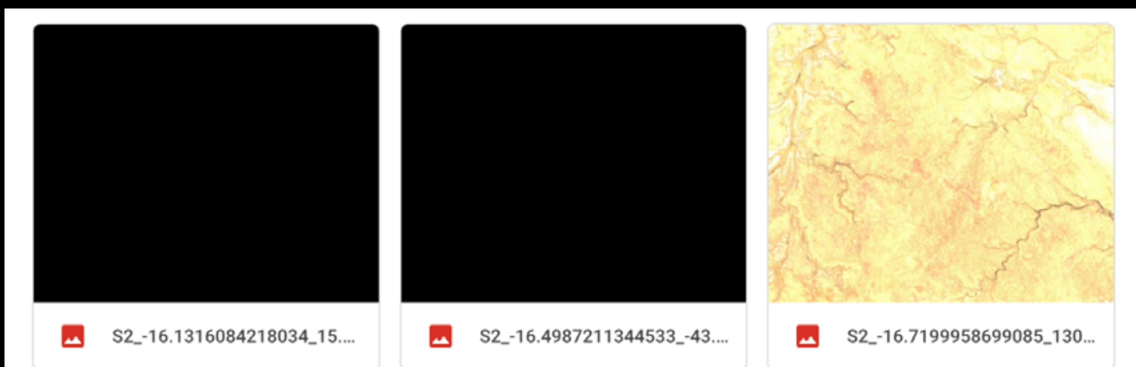


figure16 Blank images

2. 4.Overexposured.

In analyzing the pictures, although the related content can be seen in some photos, the overall light is too bright, and the content and input of higher quality cannot be presented. Further analysis found that these overexposure pictures are mainly land-based.

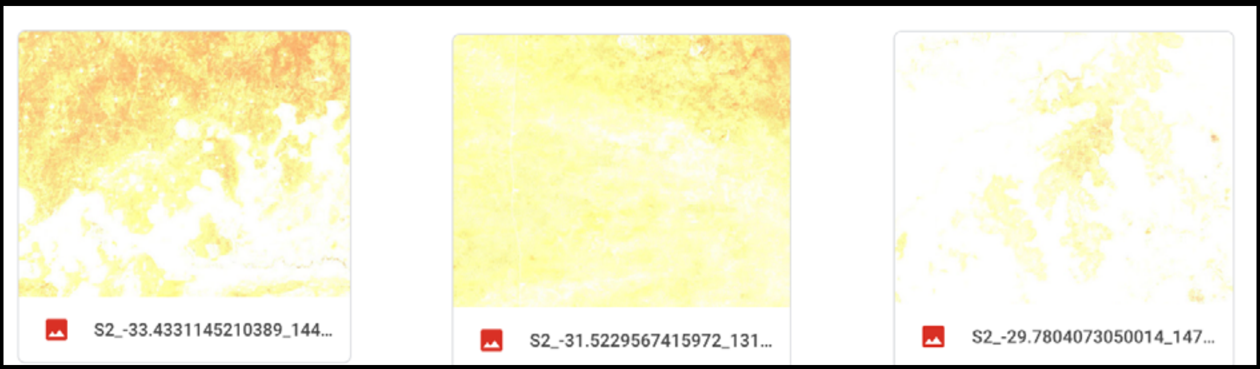


figure17. Overexposured images

Training process

In this experiment, styleGANs2 is used as the training platform.
Rounds: 2*4

In the process of this experiment, two bases were selected for training; and for each basis, four different truncation values were tried. A total of eight rounds of training were completed.

TRAINING ROUNDS

BASIS	TRUNCATION
./pretrained/wikiart.pkl	0
./pretrained/wikiart.pkl	0.8
./pretrained/wikiart.pkl	1
./pretrained/wikiart.pkl	3
ffhq5k1024x1024-apa.pkl	0
ffhq5k1024x1024-apa.pkl	0.8
ffhq5k1024x1024-apa.pkl	1
ffhq5k1024x1024-apa.pkl	3

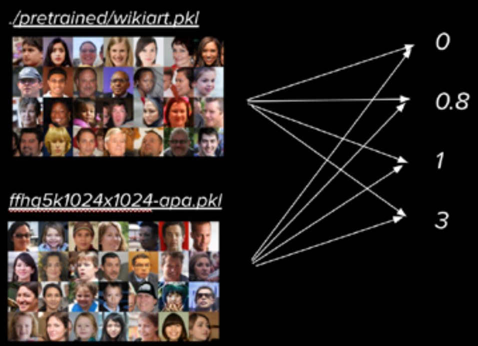


figure18. Training process

Outcome

In a total of eight rounds, using styleGANs2, there are 512 images per round. The result is as follows:

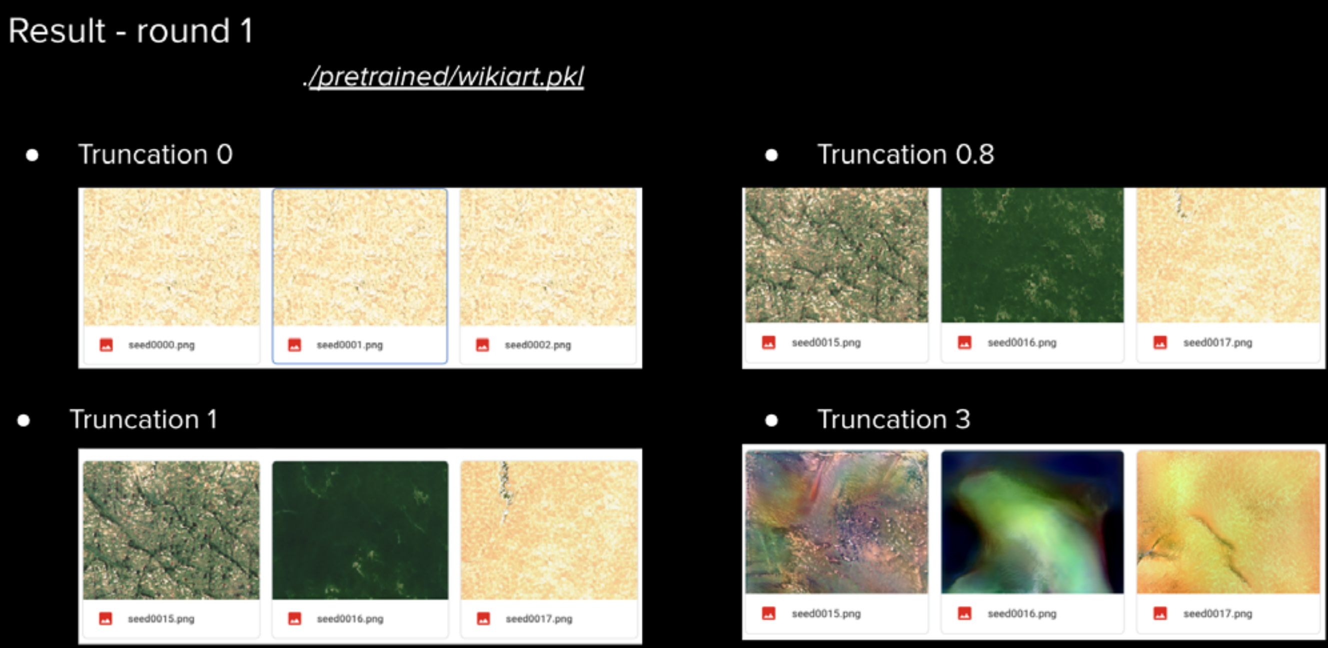
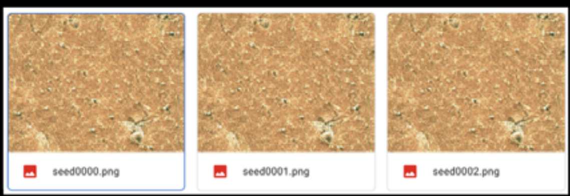


figure19. Outcome for round1

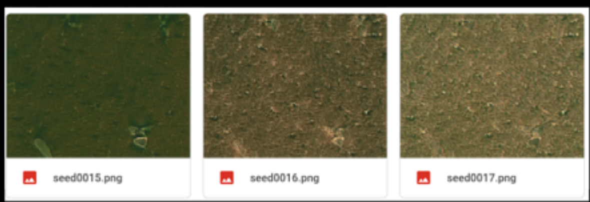
Result - round 2

ffhq5k1024x1024-apa.pkl

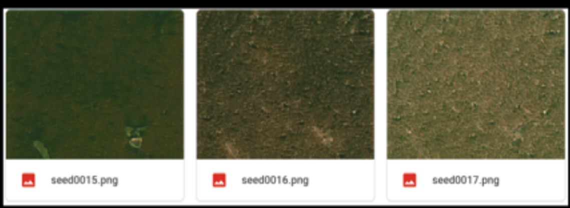
- Truncation 0



- Truncation 0.8



- Truncation 1



- Truncation 3

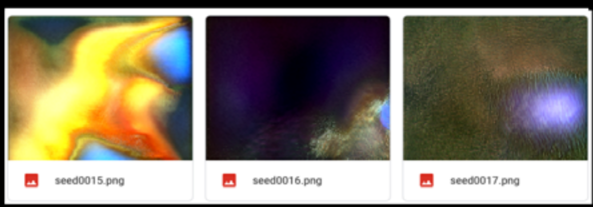


figure20. Outcome for round2

Evaluation of the outcome

During the evaluation process, seed3/seed10/seed12/seed15/seed18 in each round are randomly selected as the evaluation dataset. The overall assessment is as follows:

1.Different basis will influence its aesthetics

1.1 Round 1 generate more useful picturess

In round2, there are many blank pictures or overexposure pictures.

figure21. Outcome for round1

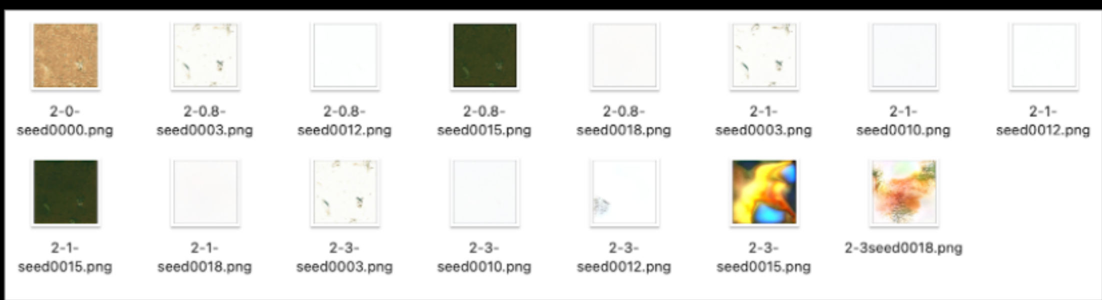


figure22. Outcome for round2

1.2 Round1's outcome is more detailed and colorful

The overall color of the round2 image is relatively simple, and the quality is lower than that of the original data-set.

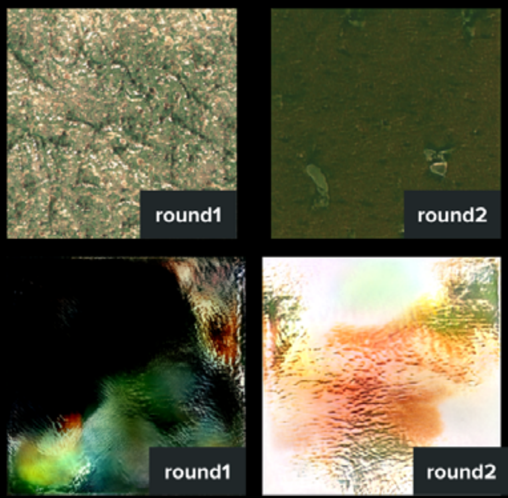


figure23. Outcome for 2 rounds

2.Different truncations will influence its aesthetics

When the truncation is smaller, the diversity is smaller, which is also more "realistic" as defined by itself. As the value of truncation increases, the pictures it produces are more abstract.

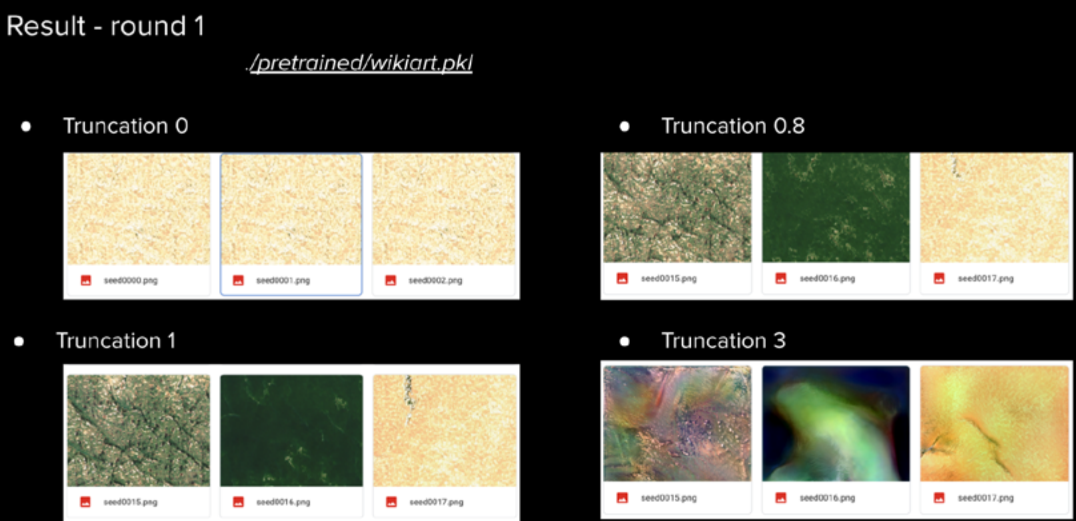


figure24. Outcome in different truncations for round1

TAKE AWAY

1. The qualities of the random input dataset are not satisfied
If the images in the input dataset are obtained randomly, the quality of the acquired images is not good. In the experiment, there are lots of overexposed and repeated. Usually, designers need to filter those images manually.
The exploration of the GANs system shows that the input dataset affects its output results. Therefore, instead of using a random input dataset directly, improving the input dataset can be an important way to translate to inform GANs of the human experience.

2. There are different parameters in GANs to iterate their output
In the experiment, when changing GANs‘ parameters like truncation, the aesthetic of its output changes. Though we do not know whether these changes meet the criteria for human evaluation, these parameters in GANs mean that we have an excellent opportunity to translate the user experience by improving the algorithm.

03

RESEARCH ON HUMAN EXPERIENCE

3.1 EXPLORATION IN AESTHETICS

3.2 QUALITATIVE RESEARCH PROCESS

3.3 RATING RESULTS

3.4 FACTORS ANALYSIS

3.5 RESEARCH ON CRITERIA

3.1 Exploration In Aesthetics

In this graduation project, I limit the scope of user experience to people's evaluation of the aesthetics of pictures. Therefore, before conducting the experiments, I explore the concept of aesthetics to provide clear evaluation criteria for the experiments.

3.1.1 What is Aesthetics

The definitions of aesthetics have different meanings in various fields. So I explored the concept of aesthetics.

The Schematic model of aesthetic experience established by Leder et al. shows that when people judge their aesthetics from an image, people's perception level will be awakened at first. After perceptual analyses, combining personal experience, people will see implicit information integration, explicit classification, cognitive mastering, and finally, finish the evaluation.

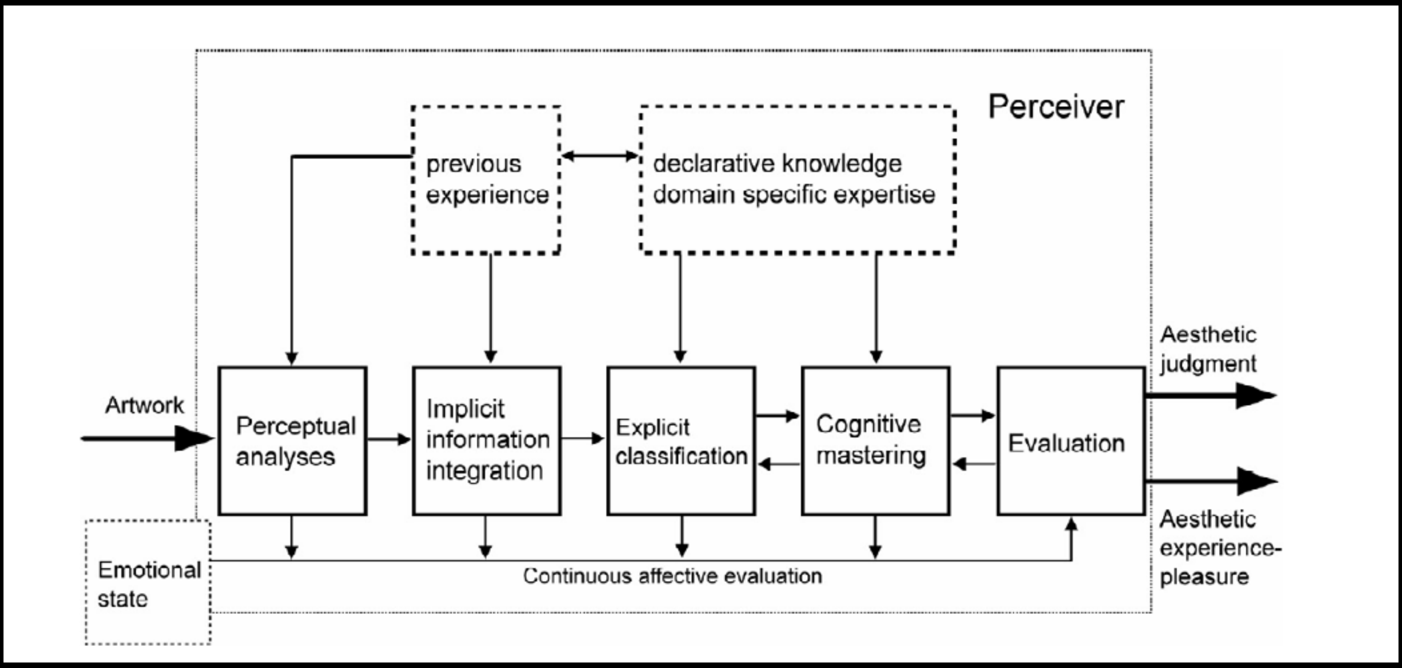


*figure25.*Schematic model of aesthetic experience

Paul Herkker's model of aesthetic preference shows that the perceptual level mainly determines people's judgment of aesthetics, cognitive, and social levels. There are safety and accomplishment needs at each level. For different contexts, the significance of aesthetic preference is not the same.

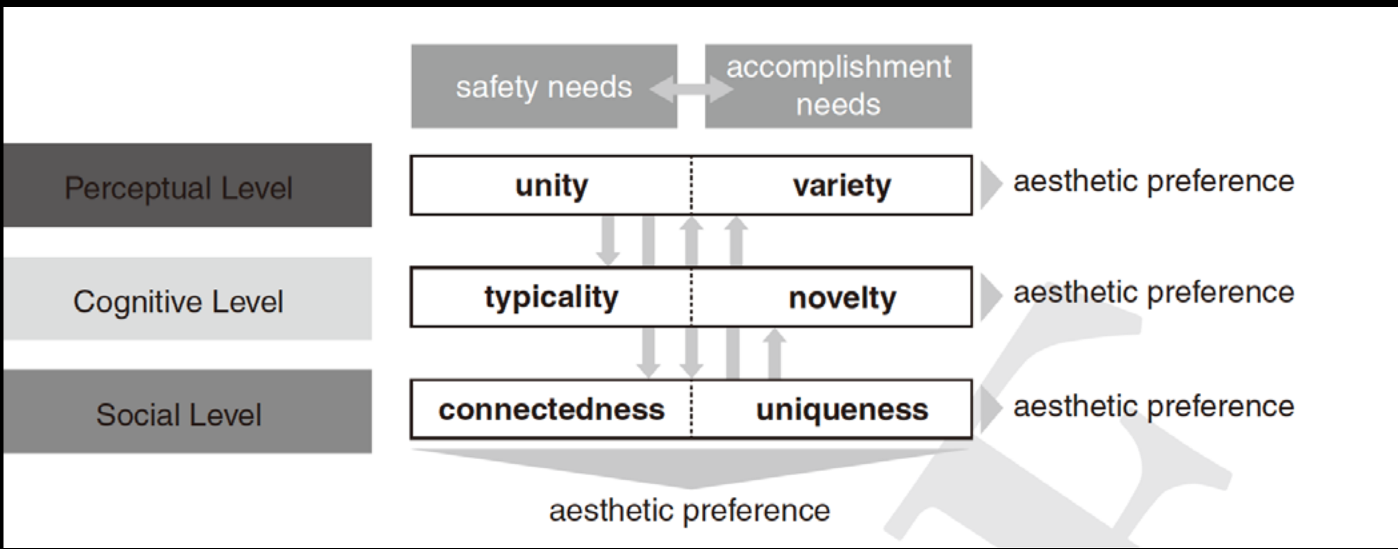


*figure26.*aesthetic preference

3.1.2 The Scope of Aesthetics

This project's definition of aesthetics is narrowed down to the perceptual level. In people's perceptual cognition, the scope is further defined as the aesthetic feeling at the visual level.

The main reasons are as follows:

1. The image type of the experiment is landscape

In this experiment, the type of image is landscape. Unlike artworks, which need to awaken people's thinking to express the content of the pictures, landscape images mainly provide people with beautiful enjoyment through the most authentic natural scenery. Therefore, rather than letting people evaluate ideas by deep thinking, this project pays more attention to the most intuitive feelings that images bring.

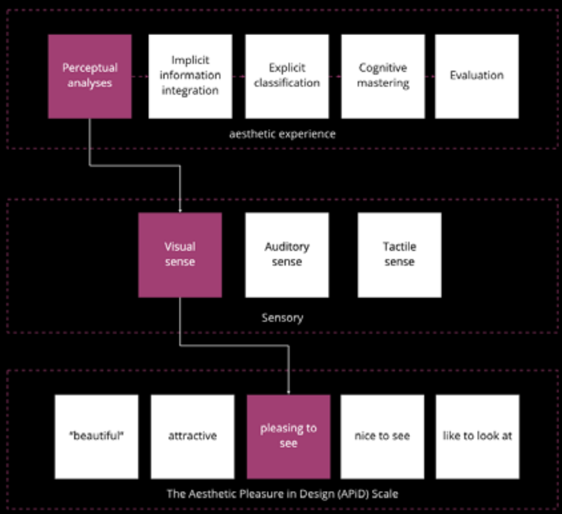


figure27.The scope of aesthetics

2. Stability

The literature shows that compared with the cognitive aspect and other aesthetic judgments related to people's personal life, the sensory level has stability. The sensory level can ignore the influence of culture and never changes. Therefore, people's evaluation criteria and scores are more stable and malleable at the sensory level. The judgments made under the same sensory system are also in line with human cognition of aesthetics.

3. Beauty lies in the "eyes of the beholder."

Although we can judge the aesthetics of objects through various orGANs, visual evaluation is still the most direct way to be recognized by the public. Compared with hearing, touch, etc., the aesthetics brought by visual senses are relatively intuitive and fast.

3.2 Qualitative Research Process

In this part, I mainly use qualitative experiments to obtain the factors that affect people's ratings of pictures.

Introduction

Through desk research and initial research, we know aesthetic judgments often have subjective evaluation criteria when obtaining user parameters. They are cognitive, perceptual, or other psychological factors. I first use qualitative research to get the user's factors in the aesthetic judgment of the picture.

figure28. The process of qualitative research

Aim

- Obtain users' for different factors in image aesthetic
- Understand why users are choosing factors
- Investigate user definitions for those factors

Research questions
The research question is:
What influences human experience in the AI system - GANs

There are some sub-questions in this research:

1. What parameters will affect the user's judgment of image aesthetics? What are the most important factors?
2. What is the meaning of these parameters in this aesthetic context?
3. For an AI system like GANs, which factors in user experience can be used directly?
4. For GANs, can the user's feedback be connected with GANs' parameters (such as outdir and truncation)? Is there a specific relationship between them? Can we improve GANs' output by adjusting its parameters to improve the factors users think affect aesthetics?

Participants

The participants will be recruited without specific criteria. A total of ten participants from various age groups and occupations were invited for the quantitative analysis.

PARTICIPANT	NAME	GENDER	Field	JOB	AGE	TIME COST
1	Zifei Li	woman	Psychology	student	23	23min
2	Mila Zhang	woman	Economics	student	24	32min
3	Melody	woman	Design	student	26	23min
4	Shuangyin	woman	Design	designer	26	40min
5	Sophia	woman	Traffic	student	28	32min
6	Anna	woman	Philosophy	phd	34	20min
7	Xiaoyang	man	Politics	officer	54	45min
8	Jane	woman	Education	teacher	46	35min
9	Ming	man	Math	teacher	45	29min
10	Asker	man	Civil enginerring	enginner	32	30min

figure31. The information about the participants

3.3 Rating results

(Table X) gives an overview of the results of two different datasets, and TableX shows the scores of each image.

1. The average value of StyleGANs2's pictures (3.229) is significantly higher than those generated by citydoesnotexsit (2.373), indicating that the images generated by sytleGANs2 are more aesthetically pleasing than thecitydoesnotexsit.

2. The maximum value of pictures in StyleGANs2 is generally higher than thecitydoesnotexsit. The maximum value of the former is 7, and three images get that score. For thecitydoesnotexist, the top score is only 5. It means StyleGANs2 has a greater chance of providing the most pleasing images.

3. The minimum score for both is 1 . 33 of 40 images get a score of 1. It shows that for the same image, people's experience is different. Most images have the potential to be rated low. At the same time, in the picture set of thecitydoesnotexsit, almost all the pictures (19/20) were given a low score of 1, indicating that the pictures generated by this GANs are more likely to produce unpleasing images.

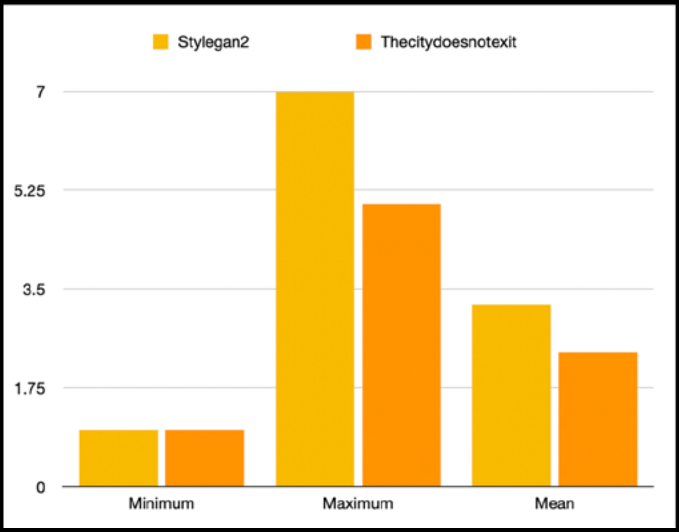


figure32. The results for the two dataset

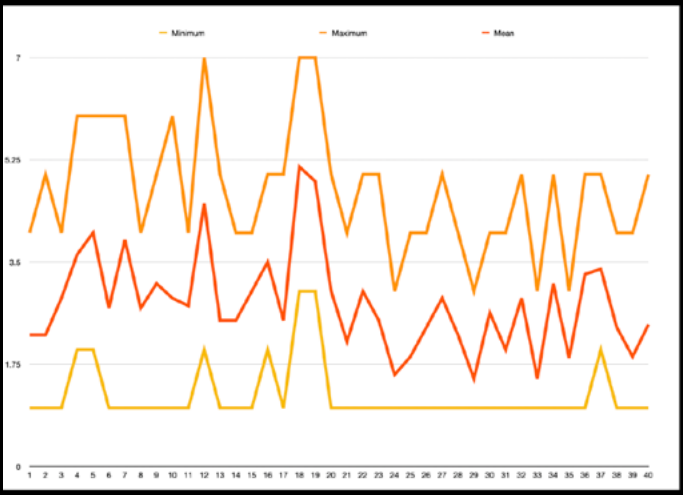


figure33. The score for each images
(1-20 images from STYLEGAN2
21-40 images from thecitydoesnotexist)

3.4 Factor Analysis

3.4.1 Process

During the research process, it is found that the factors of images can be divided into the following two types, one is related to the content of the images, and the other is related to the parameters of the images. Moreover, exploring GANs shows two opportunities for this AI system to iterate the output. The first is changing the input, and the other is improving the algorithms.
So the following analysis method was established(figure34) to analyze the qualitative factors of the human.

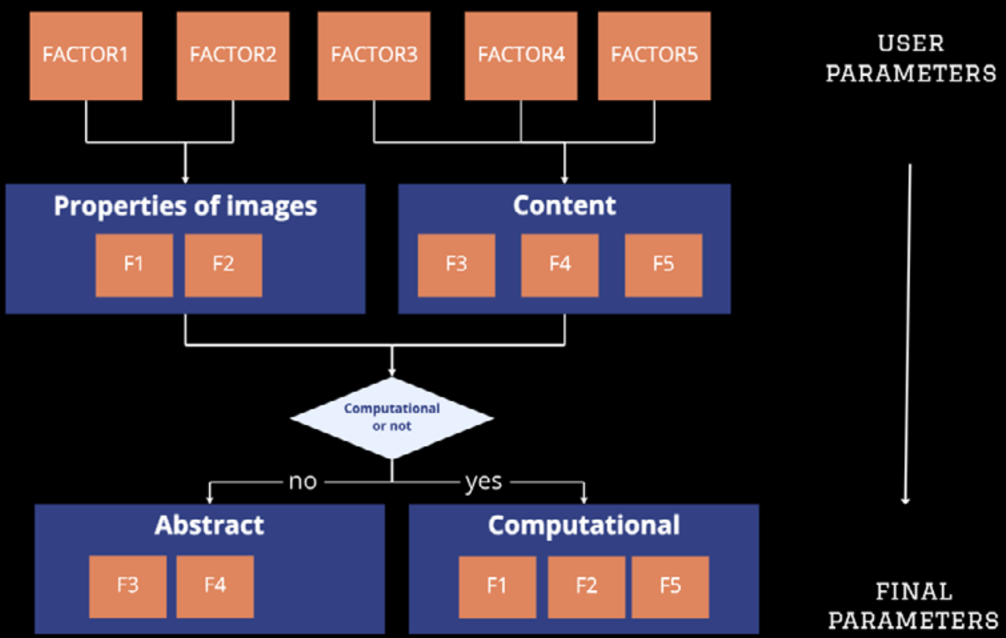


figure34. Analysis method

The analysis process of the whole process mainly includes the following steps:

Step1 Divide factors into "content" and "properties of images."

During the workshops, when people explain the factors, some factors are highly related to the images' content, and some are the parameters from the images. In our context, we choose the landscape as the only type of image. Therefore, considering that these factors will be applied to different image types in the future, we need to distinguish the above two types of factors.

Factors influenced by the content.

Those are highly related to one image's content, such as the elements that make up the picture, the layout, etc. The content of the image generates them. When the system adjusts them, it should consider the dataset type like landscape, followers, people, etc.
The characteristics of this type of factor are:
1. It has a strong relationship with the type of the picture. If the kind of images changes, those factors will vary. For different types of images, such factors may or may not exist.
2. It mainly conveys the information of the picture. According to the different reports in the picture, the standard for the factor will change. For example, for people, the layout is not necessary. Nevertheless, the glassed eyes will influence its aesthetics; the layout will be essential when it comes to landscape.

Properties of images

The second type of factor is the properties of images. Unlike content, this factor belongs to the nature of all images. No matter what information it contains, these factors will exist. For example, the color, brightness, etc., In general, they can be processed and adjusted by tools such as image processors (photoshop/illustrator. There is also a greater possibility of quantifying them.

The main factors of this type of factor are:
1. They are not affected by the content of the image itself.
2. They can be processed with image processing tools, and the possibility for AI systems to use is higher.

Step2. Divide different factors into computational and abstract.

According to the research on GANs, many algorithmic frameworks in this AI system have the chance to iterate the outputs. For these frameworks, only the computational factors can be used. Moreover, the rest of the factors that are hard to quantify are classified as abstract factors.
Therefore, we divide it into two categories: computational and abstract.

Computational factor

Those factors' properties can be computational and quantified.

Abstract images

The factors that cannot be directly converted into a statistical factor.

3.4.2 Factors

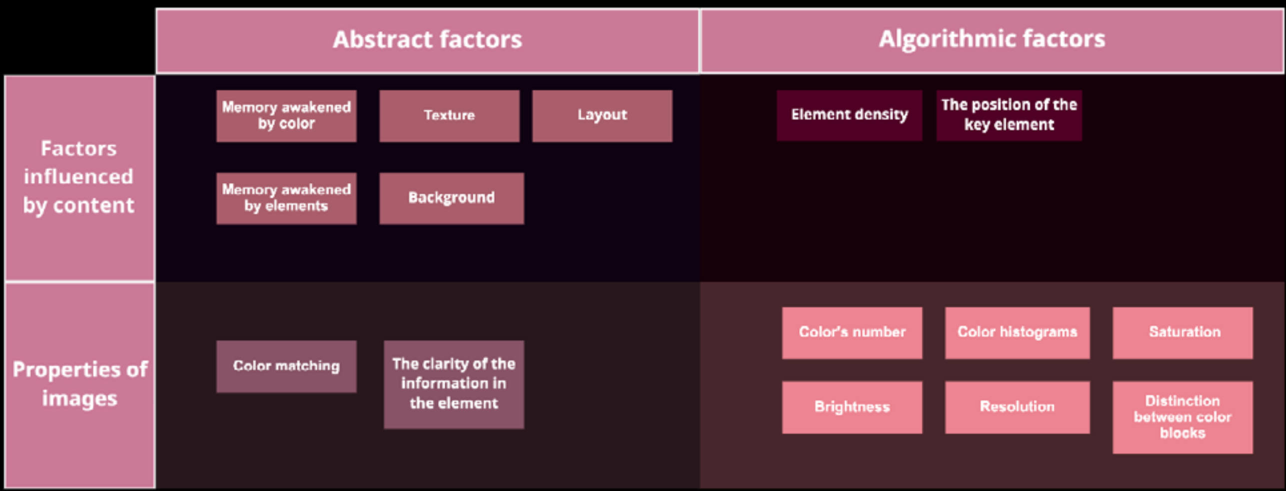


figure35. Final factors

Because the number of participants in the survey was 10, in order to ensure the comprehensiveness and diversity of the factors, the factors that are mentioned by two or more participants are all selected into the final factors.
Finally, after analysis, 13 factors that users think have a more significant impact on the aesthetics of the image are obtained. Moreover, they are classified into four categories according to people's explanations. The result is shown in figure 35.

1. Factors influenced by the content

1.1Abstract factors

1.1.1 Memory awakened by color

Different colors can awaken various associations of users. For example, grey-colored pictures(figure36) remind users of depressive scenes such as “sewers,” and they tend to give low scores. In contrast, orange-colored and yellow-colored pictures (figure37)remind users of joyous scenes such as the sun, and then they will provide high scores.

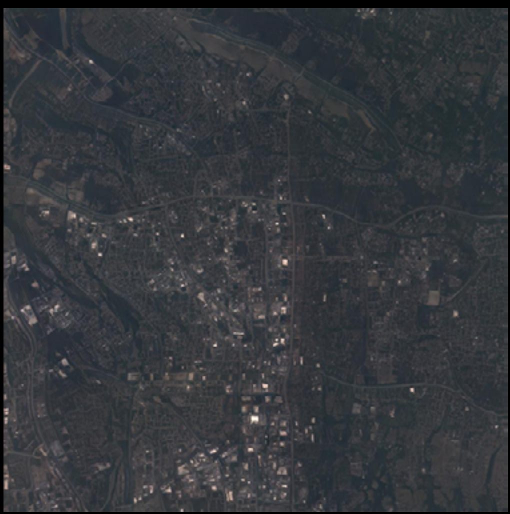


figure36.the output

figure37.the output

1.1.2 Memory awakened by content

Users also mentioned that different image content's recall of their memory also affected their scores. For example, when most users see pictures(figure38) related to the ocean, the overall score is higher because they are connected to their pleasant memories of playing at the seaside. When they see pictures of deserts(figure 39), dark cities, etc., their memories are associated with uncomfortable experiences such as moderate dryness and heat, and the score is lower.

figure 38. the output

figure39.the output

1.1.3 Texture

The texture of the picture gives people the overall feeling of the image. In experiments, graphics with relatively soft textures have a higher score than pictures with rather rough textures. Especially when it comes to landscapes, the smoothness of images such as oceans can improve people's scores.

1.1.4 Background

Whether the background in the picture can give a better atmosphere, for For example, people score higher on a background with soft and open elements like the ocean(figure 40).

figure40.the output

1.1.5 Layout

In pictures involving key elements and backgrounds, whether the level of the picture elements is precise or not will significantly affect people's ratings. For example, when the background and the elements have a clear priority, or there is white space overall will encourage people to give a high score. Even if the critical element is relatively small, it can be divided into the background and higher score.

figure41.the output

1.2 Algorithm's factors

1.2.1Element density

In the picture set for this experiment, some pictures are dominated by hous-
es,
the density of all sub-elements is relatively high, and the scores of such images are relatively low. Users also mentioned that if the density is too high, it will reduce their score.

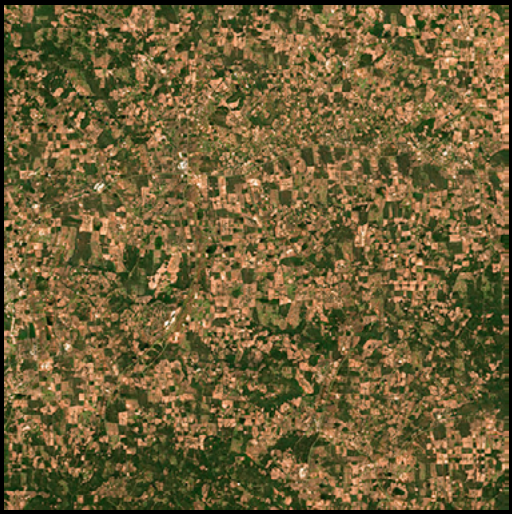


figure42.the output

1.2.2The position of the critical element

People prefer the image where they can find the critical element quickly. In general, if the key factor is in the center of the picture or four vertices, the picture has a higher score.

2. Properties of images

2.1Abstract factors

2.1.1 Color matching

The combination of colors in the picture is the most mentioned factor by users. The color combination mainly includes the following aspects:

a. Contrast between colors.

In the selected image, different elements have different colors, and the contrast among
them has a more significant visual impact on the user. The color contrast between yellow and green is relatively high-scoring in the experiment.

b. Gradient color.

When there are gradient elements in the image, the overall score is higher in the survey. Gradient colors are “reasonably matched colors” in users’ eyes, enhancing the entire screen’s high-level sense.

c. Same color.

Some users also mentioned that it would be more comfortable to see the picture if it had the exact color matching.

2.1.2 The clarity of the information in the element

Whether the information of the element itself can be effectively conveyed. It includes:
a. Quickly understanding the picture elements :
Whether the information of the picture elements can be quickly understood;
b. High degree of differentiation between elements,
Whether the distinction between the elements is high enough;
c. Match the real scene
Whether the information of the elements is consistent with the actual scene. For example, images of the ocean and coastline that match the scene in people's memory will receive a higher score. However, if the information in the picture is mixed and the elements are challenging to identify, it will reduce people's ratings.

2.2Algorithmic factors

2.2.1Color's number

The number of colors also has a specific impact on the scoring criteria. From the test, pictures with colors between 3~4 align with the nature of high-scoring images.

2.2.2Brightness

The overall brightness of the screen. In the two image sets shown, the brightness of the pictures produced by styleGANs2(figure 43) is higher than that of thecitydoes-notexsit(figure 44). The picture's brightness affects the user's perception of image recognition. Several users mentioned: "The brighter picture can ensure that they can obtain relevant information to identify the picture." Therefore, relatively moderate brightness is also one of the factors affecting the score.

figure43.the output

figure44.the output

2.2.3Color histograms

Pictures dominated by warm colors can obtain relatively high scores, while images dominated by cool colors (figure 45)have lower scores. At the same time, the user's preference for the color also has a certain degree of influence on the color's color value. Some users have higher ratings for blue-green pictures(figure46), while others have higher ratings for bright yellow images.

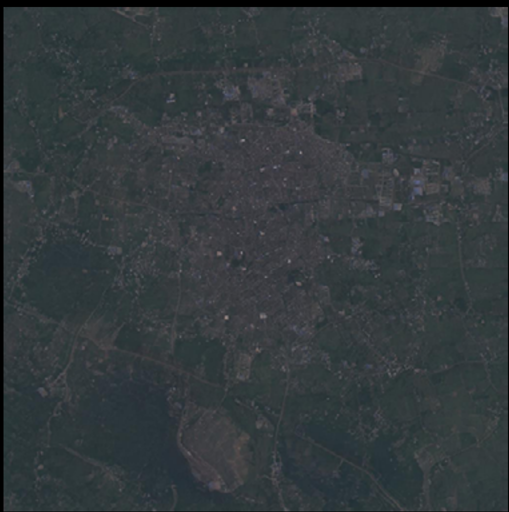


figure45.the output

figure46.the output

2.2.4 Resolution

Due to the different performances of different models, the resolution of the pictures produced by styleGANs2 is higher(figure 47). Compared with thecitydoesnotexsit(-figure 48), users can also understand the details more quickly.

figure47the output

figure48.the output

2.2.5 Saturation

Consistent with the previous direct data analysis, the image set in the thecitydoes-notexist has a low overall saturation score. The overall picture in styleGANs2 has a higher saturation and is mainly bright, with a higher score.

2.2.6 Distinction between color blocks

A clear separation between the color blocks will increase the ratings.

figure49.the output

3.4.3 Discussion

There are four kinds of aesthetic factors influencing people's evaluation

Thirteen factors influence people's evaluation of landscape images. They can be clustered into four categories: factors influenced by content, properties of images, abstract, and algorithmic factors.

Different factors have various methods to be informed by AI systems.

The classification results show that most images' properties may be quantized and turned into data parameters. It means that of the nature of the images, they can probably be solved algorithmically. It will be more accessible for GANs to iterate those factors. However, those factors generated by the image's content are primarily abstract. Changing them to GANs's statistic parameters is hard. Therefore, for different aspects, the translation methods are other.

Different datasets need different methods.

What is more, the user experience is translated differently for different types of image datasets. For example, changing the quantifiable factors can improve the user experience for datasets containing only one kind of image because all the images' elements are the same in the dataset. It is hard to change the specific aesthetic factors for datasets containing multiple types of images, which are very different for each image. Other methods should be used to iterate their output.

The correlation between highly-rated images and aesthetic factors

In addition to analyzing the overall data, it also conducts a more in-depth analysis combined with specific pictures. It shows that highly-rated images can represent abstract factors. In other words, the output with a high score always contains those good factors in the abstract category.
During the workshops, when participants are asked to cluster the images into the "factor." for the abstract factors, including Memory awakened by color/Memory awakened by elements/Texture/Background/Layout, the images that are classified as most matching these factors are the images with high scores. However, for the algorithmic factors, some images with lower scores are also classified that fit these factors. For instance, figureX with a score of 1.3 is clustered to the "Distinction be-tween color blocks," Resolution." Moreover, those pictures with the highest scores(-Figure 51) are all selected into the abstract factors groups.

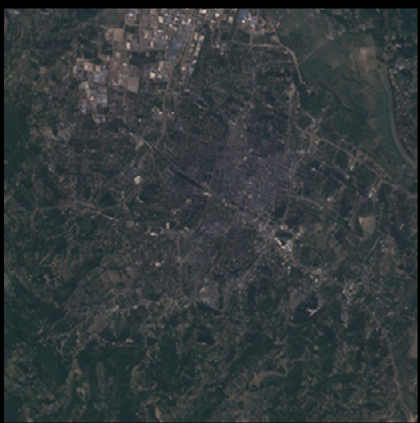


figure50.the output

figure51. The most aesthetic images in SYTLE-GAN2

3.5 Research on the evaluation criteria

After determining the scope of aesthetics in the project, I investigated some current methods of translating human experience into computers. Understanding the pros and cons of these approaches can help us better find informing AI systems human experience

Computational Aesthetics

In 1933, George David Birkhoff wrote the first quantitative theory of aesthetics in his book Aesthetic Measure. Since it involves computational methods, this work is often regarded as the beginning of Computational Aesthetics.

Computational Aesthetics is the research of computational methods that can similarly, make applicable aesthetic decisions as humans can

Moreover, there are many measurements for computational aesthetics.

3.5.1 Order/Complexity

Definition
It represents the reward one experiences when putting effort by focusing attention (complexity) but then realizing a certain pleasant harmony (order).

3.5.2Rooke - expression trees

Definition
The ability of the expression trees to make aesthetic rankings is explained by the the fact that the underlying primitives in the nodes of the trees were able to make statistical assessments of the images.

3.5.3 Sprott - global complexity measurement

Definition
For aesthetically evaluating fractal-like images

3.5.4 Baluja et al - categorizing the user rankings of imag-

Definition
These researchers attempted to train a neural net to perform this evaluation task using as training sets images that were obtained by categorizing the user rankings of images evaluated while users were running an interactive version of their generative system.

TAKE AWAY

How to inform computer - extract user parameters from human information and convert this into static parameters.

These methods show how to translate information from the human world into computer systems: extract user parameters from human information and convert this into static parameters.

1.User parameters
From the previous research process and initial experiments, we know that the human world's information is messy, cognitive, and perceptual. So the first step to translation is extracting user parameters that can represent the messy information.

2.Statistical measurements
The core of those methods is that the researchers can quantify those parameters. Because for the algorithmic world, statistical parameters are the language they can understand. So finding the statistical measurements for those user parameters is an essential step in the translation process.

04

THE TARGET USER

4.1 TARGET USER

4.2 WHY CHOOSE AI PRODUCT DESIGNER

4.3 THE CHARACTERISTICS

4.1 Target user in the project

In this project, the target user is narrowed down to AI product designers who are working in the AI field.

4.2 Why we choose AI product designers

As AI products gradually penetrate people's lives, AI product designers play an increasingly important role in shaping how AI systems develop. Moreover, my experience as an AI product designer at Alibaba made me realize the significance of translating the qualitative human experience into a quantitative AI system. Also, based on their HCD(Human-center design) mindset and technical knowledge, AI product designers can be the translator between the AI world and the human world.
So I want more AI product designers to be able to inform AI systems about the human experience.

4.3 The characteristics of AI product designers

According to some research and the interview with AI product designers, AI product designers have the following characteristics:

1. Have some basic knowledge about technology

Unlike ordinary designers focusing on interfaces or objects, AI product designers face various types of products. Furthermore, the iterative process of different products is often accompanied by the algorithm iteration. This means that designers must have a relevant understanding of technology in the design process. Therefore, AI product designers can combine expertise in HCD(Human-center design) and AI fields to improve AI systems from both user and technical aspects.

2. Work closely with the algorithm team

Due to the closeness of the design process and algorithms, AI product designers often have more opportunities to communicate directly with algorithm engineers. Compared to most UX designers, AI product designers work more closely with the algorithm team.

05

HYPOTHESIS

5.1 The highly-rated images will improve the aesthetic quality of the output images

5.2 Computational models can be designed to predict human ratings of beauty

Based on the above survey and analysis, two approaches based hypotheses are proposed on how to inform AI systems human experience.

5.1 Curation approach - The highly-rated images will improve the aesthetic quality of the output images

The curation approach hypothesizes that retrieving the highly rated images will improve the aesthetic quality of the output images.

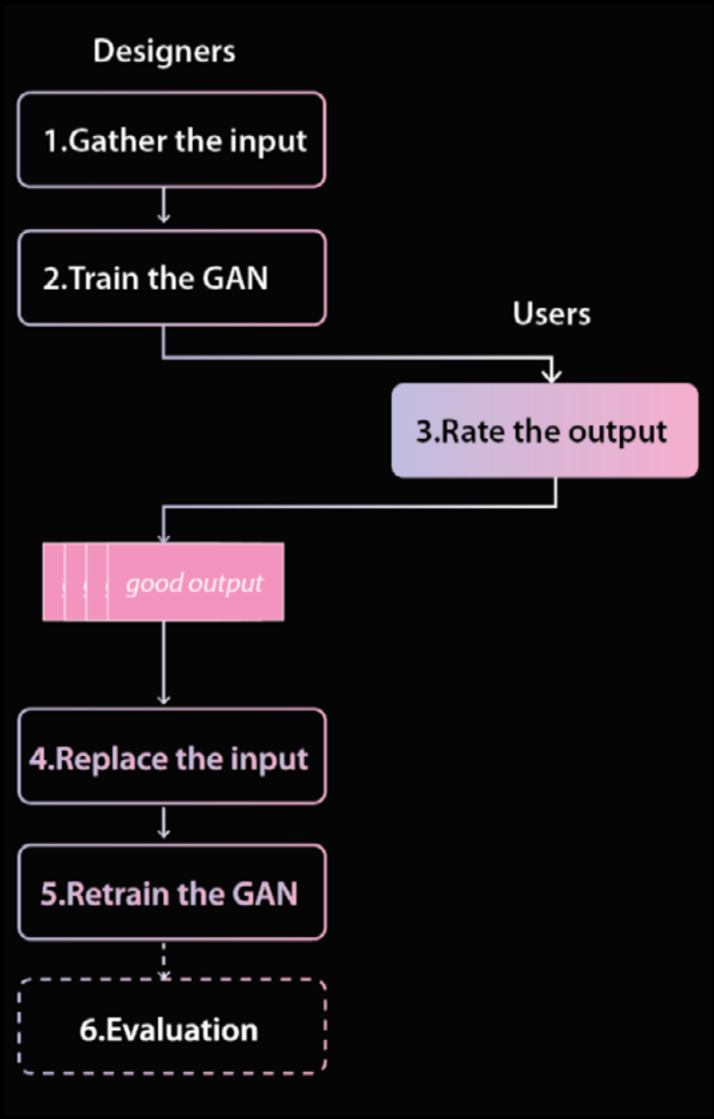


figure51.Curation approach

The reasons for proposing this approach are as follows:

Based on the qualitative research for human experiences in GANs, some factors are subjective and individual. For example, people feel excited when "It can evoke my beautiful memory, "or "The images with natural texture can absorb me." Those factors are hard to quantify directly by computer since they are connected with everyone's experience.

However, during the exploration of GANs, it is discovered that As a machine learning framework, apart from the algorithms part, which is significant for GANs' capability to generate the output, the dataset that we feed it is also essential. For GANs, the real images we provide, called input datasets, are the inspirations and materials for learning the features and generating the fake images. So, iterating the input dataset can be an excellent way to improve those factors that are difficult to compute.

Also, during the qualitative research, when users explain the factors that influence their ratings, the output with a high score always contains those good factors. So instead of collecting the new input dataset from other materials and methods, the highly-rated images selected by people can be a good resource for designers to use to represent the user experience.

What is more, considering the different contexts when we use GANs to generate the fake images, sometimes, in one set of the input dataset, there will be plenty of images that include various information. For example, if designers want to generate fake images based on flowers, landscapes, or trees, some computational factors like brightness and saturation define aesthetics differently for different images. Therefore, retraining the GANs using people's preferences is the most efficient and generative way to apply to all kinds of images.

In summary, considering the computational factors' limitations, the input dataset's significant influence on GANs, and the context when there are broad types of images in one input dataset, the curation approach can be a good guidance for designers to translate the human experience to AI systems.

5.2 Algorithmic aesthetic approach - Computational models can be designed to predict human ratings of beauty.

The algorithmic aesthetics hypothesizes that computational models can be designed to predict human ratings of beauty.

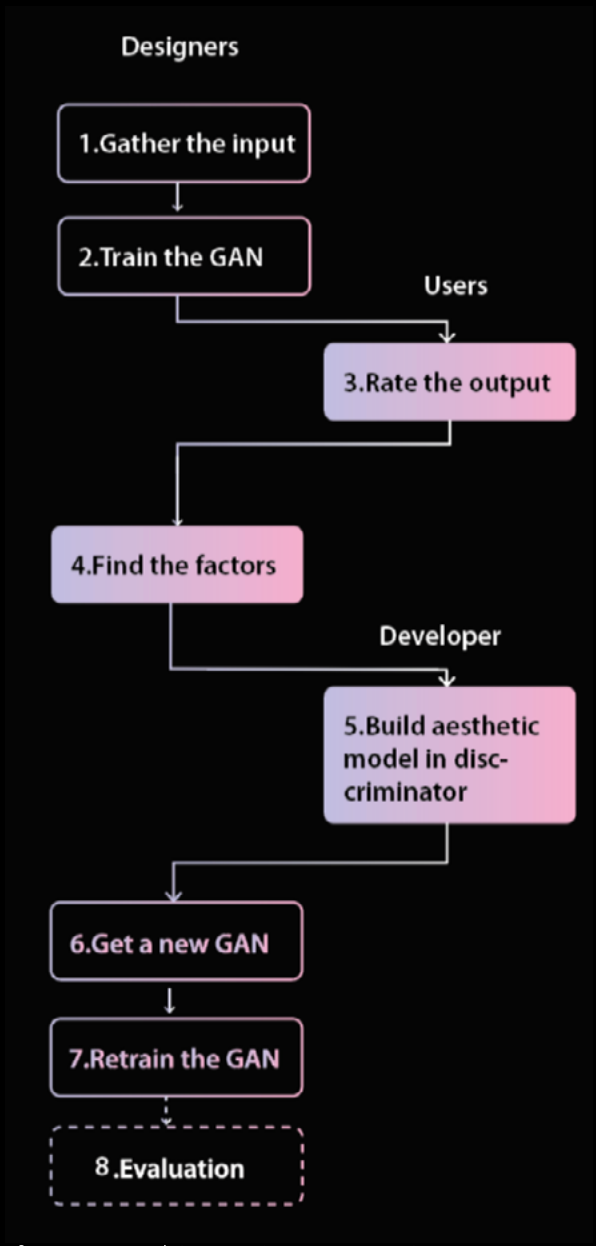


figure52.Algorithmic aesthetic approach

The reasons for proposing this approach are as follows:

Based on the qualitative research for human experiences in GANs, some factors are the images' visual and spatial features like saturation, contrast, etc. Those factors can be quantified and have the potential to be converted into statistical factors.

In the algorithmic world, some automated measures of aesthetic beauty can rate the factors from images from the perspective of AI systems. If those factors correlate highly with human ratings, they can be aesthetic predictors in AI systems. It means designers have the opportunity to find the similarities between the AI world and the human world for the evaluation of image aesthetics.

Also, during the exploration of GANs, like an oil painter with a set of painting equipment, GANs include various algorithmic components like latent space, generator, discriminator, etc. Those algorithmic frameworks provide designers and developers a lot of space and chances to iterate GANs' output by improving those aesthetic predictors.

Moreover, designers in AI often have a specific basic understanding of algorithm knowledge in the field of AI. Compared to the traditional designers who used to design only from the user's perspective, AI product designers can usually solve problems from both the view of the AI system and the user. This mindset enables those designers to analyze the factors from human experience algorithmically.

Last but not least, in the design process of AI products, the cooperation and connection between designers and developers tend to be closer. So, this close collaboration with developers gives designers more opportunities to translate the qualitative human experience to the quantitative world more scientifically and precisely.

In summary, considering the similarities in computational measurements and human experience, the rich algorithmic architectures in GANs, AI designers' technical mindset, and the close cooperation between AI designers and developers, the algorithmic approach can be good guidance for AI product designers to translate the human experience to AI systems.

06

APPROACHES

6.1 CURATION APPROACH

6.2 ALGORITHMIC AESTHETICS APPROACH

6.1 CURATION APPROACH

6.1.1 What is the curation approach

The curation approach is a method that can be useful for any kind of input dataset regardless of its content and information. Guided by the curation approach, designers can iterate the GANs' output quickly and conveniently. Without the technical theories of GANs, designers can still become good translators to inform AI systems' human experience accurately.

In the curation approach, by putting the beautiful GANs' outputs selected by humans into the input dataset and retraining the GANs, GANs can get people's feedback and steer their output.

In this process, GANs obtain the human experience in the form of "good output" selected by humans as their inspiration, and the designer "translates" the human experience through the action of "using excellent output as input and retraining the GANs."
From the perspective of the input image, this method informs the AI system of people's evaluation through retraining GANs using good output.

6.1.2 Metaphor

In this process, we provide our GANs painters with their previous landscapes loved by humans as their new inspiration and use them to draw the new paintings.

figure53. The metaphor

6.1.3 Participant

1. Designer
For the curation approach, designers do not need to know the technical theories of GANs. The only operation they should handle is training the GANs.

2. User
The curation approach relies on images selected by humans. The higher the number of users involved, the more accurate and general the high-scoring images we get. The minimum number of users is 4.

6.1.4 When do we use it

For a messy input dataset
When designers want GANs to generate a set of images that are a combination of different kinds of images, they need to provide GANs with a dataset with plenty of images. Then, for those messy input datasets, retraining the GANs with the highly-rated output in the first training process can be valuable and accurate.
For example, when designers want to generate fake images combining the artworks, landscapes, and people, it is hard to change the specific aesthetic factors, which are very different for each image. In the curation approach, designers can let humans select the beautiful images from the output and feed those good outputs to the GANs directly to get a better output.

For an independent designer
Because the curation approach does not involve knowledge of any other field, designers do not need to cooperate with experts from other fields like developers. They can complete the whole process with the user independently. So the curation approach is straightforward to use and handle for almost any designer, even if they do not know any technical knowledge of AI systems.

For a short period
In the curation approach, due to its simplicity, designers can complete the whole process in a short time. Designers do not need to change any parameters in GANs, and even the training time does not change when they replace the input dataset. It offers designers a quick method to follow when they want to iterate their output.

6.1.5 How to use it?

Facilities

1.GANs
Designers should operate the GANs expertly. Although they do not need to learn the algorithmic theory, they should know how to train the GANs and get the output.

2. 4-choice rating system
To get the images accurately and efficiently, users use the 4-choice rating system to select the images. In this system, the pictures most pleasing to people's eyes in each set of images represent the highly-rated images.

Curation Approach

In the curation approach, by putting the beautiful GANs' outputs selected by humans into the i

Participants

1.Designer

Designers do not need to know the technical theories of GANs.

2.Users

The higher the number of users involved, the more ac high-scoring images we get. The minimum number of

Designers

Gather the input dataset

Train the GANs

GitHub qualtrics XM

Make the rating survey

Gather random images without any processing

1. Connect to the drive
2. Upload the input folder
3. Generate the images

GitHub

1. Rename the images

XM

2. Upload the images to Qualtrics

GitHub

3. Copy the image ID
4. Paste the survey template
5. Run the survey creation algo

XM

6. Upload the survey.qsf file
7. Add other materials in Qualt

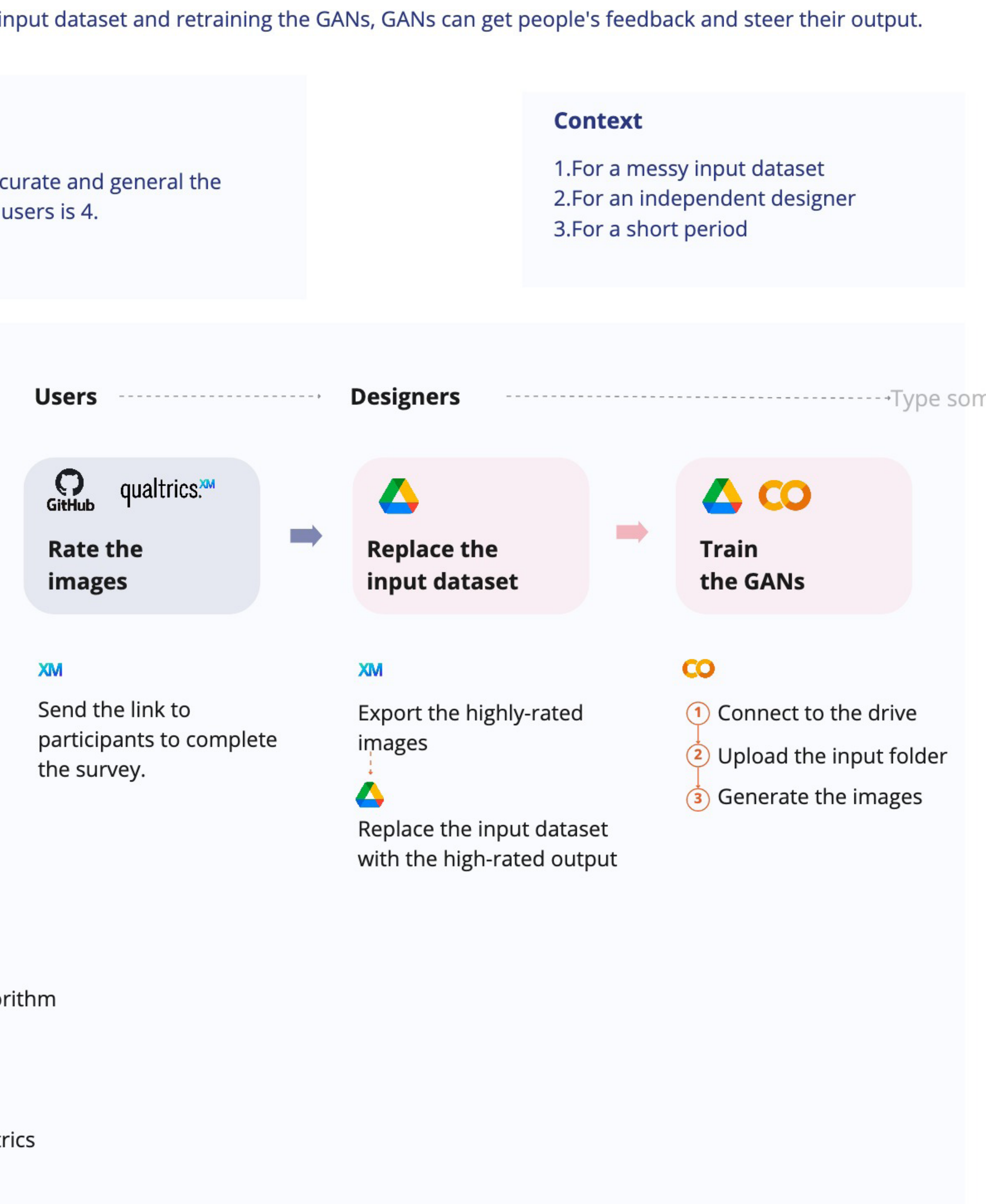


figure54. Curation approach

Steps

For the curation approach, there are six steps.

STEP1 - Get the random input dataset

In this process, the designer first gathers random images as its first input dataset. In the first round of the training, the designer does not need to do any processing on the dataset. They can collect any type of image.
For the number of the images, ideally, for GANs, ~4000 images can be enough dataset for it to generate fake images.
All the images should be compressed in a folder and uploaded to the drive.

STEP2 - Train the GANs

After getting the dataset, designers use them for training the GANs and getting the output.
2.1 Connect to the drive
Click the "play" button to view the system's GPU and connect the drive to the co-lab.

2.2 Upload the input folder.
Enter the location of the compressed file into "dataset," and "resume_from" will determine the final training result to a certain extent. Under this model, if it is a new train, you can choose 'ffhq1024' or './pretrained/wikiart.pkl'. Click the "play button" after making changes.
Click the "play button" again; the model will start training to generate images.

2.3 Generate the images
After setting the parameters like network, seeds, and truncation, designers click the play button to get the images, and finally, they can get the output in the drive's folder.

STEP3- Make the rating survey

After getting the output, designers should upload the images in an automation system and then let people rate the output in a 4-choice system. This can help humans select the best output accurately and quickly. In this process, designers will get the "good output" selected by users.

Platform

1. GitHub link
Designers must organize and select the images and then create the survey in the automation process. In this process, designers can finish such a harrowing experience in 10 minutes.

2. Qualtrics
A 4-choice system in Qualtrics is built to let designers process the images efficiently, and humans rate the images accurately.

Materials for the survey

1. Images
For an input dataset including 4000 images, getting 1000 high-rated images can be the ideal dataset to replace the previous dataset. Designers need to choose the number of good outputs given the number of their original datasets. Ideally, it is more effective if the number of new output is a quarter of the input dataset.
The size of the images is 1024*1024.

2. Question
Based on the exploration of aesthetics, the definition of aesthetics is narrowed down to the scope of the perceptual level, and the visual sense is selected as the evaluation scope. Then, from the APiD Scale, "pleasing to see" is chosen to be the question for the aesthetics of the images. So the question is, "Which image is the most pleasing to your eyes."

3. Consent form
The consent form is a document signed by persons of interest to confirm that they agree with an activity that will happen and that they are aware of the risks or costs that may come with it.

4. Example question
Some example questions can ensure the participants understand the survey process completely.

Process

3.1 Rename the images
Put the output from GANs into folders and rename them into the correct format by running the script in the correctlynameimage.py file.

3.2 Upload the images to Qualtrics
Create four different graphic folders in Qualtrics and upload the images there.

3.3 Copy the image ID
Copying the image ID from Qualtrics and paste it into the images to add.txt.fil

3.4 Paste the survey template
Paste the appropriate survey template on the Template. qsf.file

3.5 Run the survey creation algorithm
Run the survey creation algorithm from the formatqsffile.py file. This will create the edit the Template. qsf file and put the GANs images in the survey. The algorithm should output a new file named survey.

3.6 Upload the survey.qsf file
Upload the survey.qsf file to qualtrics.

3.7 Add other materials in Qualtrics
Manually add the consent form, example question, attention checking questions, and attitude questions to Qualtrics and finalize the survey.

3.8 Start the survey
Send the link to participants to complete the survey.

STEP4- Let users rate the images
Send the link to participants to complete the survey.

STEP5- Replace the input dataset
After getting the new output, designers must replace the input dataset with the high-rated output. Designers have no limitations when choosing which images in the input dataset should be replaced.

Designers will have a new folder with sound output selected by users and the previous random image as their new dataset. Furthermore, the new folder should be compressed and uploaded to the drive.

STEP6- Retrain the GANs
Finally, designers retrain the GANs based on the new input dataset, and they can generate the new output directly.

6.1 Upload the input folder.
Enter the location of the compressed file into “dataset,” and “resume_from” will determine the final training result to a certain extent. Under this model, if it is a new train, you can choose ‘ffhq1024’ or ‘./pretrained/wikiart.pkl’. Click the “play button” after making changes.
Click the “play button” again; the model will start training to generate images.

6.2 Generate the images
After setting the parameters like network, seeds, and truncation, designers click the play button to get the images, and finally, they can get the output in the drive’s folder.

6.1.6 The role that the designer plays

In this process, designers translate the human experience in the action of "making the survey" and "replacing the input dataset". There are two main roles they play in the whole process.

CATCHER

Designers have the knowledge of human-center design. So they are able to know how to get accurate and useful information from people. In the process, by making a useful survey system, designers help crowdworkers to express their feeling efficiently and accurately. This guarantee the quality of the human experience.

IMPLEMENTER

During this process, designers can complete the whole process by themselves. So they are the only implementer in this process. Because AI product designers have the knowledge from HCD(human-center design) and the technology field. They can implement the whole process easier. It means that though AI systems are technical, designers can inform AI system's of human experience by themselves.

6.1.7 How do we achieve it?

For the curation approach, starting from establishing the survey and rating system, then using them to prove the hypothesis “Retraining using the highly rated images will improve the aesthetic quality of the output images.”, three experiments were completed to establish the approach and evaluate it finally.

Establish the automatic survey creation for designers
In the curation approach, designers must upload many images to the survey. Manually organizing the images and creating the survey is very laborious. To make designers complete the approach efficiently, an automatic system is built for designers to create the survey.

Establish the four-choice rating system for users
For the curation approach, after getting the output from GANs, the most significant part is to get the high-rated images selected by users. We need an efficient system that allows designers and humans to use it easily. This system should be a low threshold for designers and provide them outcomes efficiently. Users should be able to select the most aesthetic images from many pictures accurately and quickly.

The target users in the system

1. Designers
Designers without a background in design are using the system to upload the rating materials and export the final output. They aim to get the output efficiently.
2. Users
In the rating process, the people who are selected to rate the images are the crowd-sourced workers from various backgrounds. Usually, four people are enough to provide a reliable outcome. However, the higher the number of users can produce more accurate and general high-scoring images. So, for users, this system should be easy to operate and enable humans from any background can rate the images.

The materials in the system

1. Images
The number of images in the system depends on the output from GANs and designers. For most contexts, 1000 images should be selected by humans in this system.

2. Question
"Pleasing to see" is chosen to be the question for the aesthetics of the images.

3. Consent form

4. Example questions, attention checking questions, and attitude questions

Survey formats

For the formats of the survey, cooperating with Moshiur, two formats (binary-choice and four-choice) are tested and compared to select the most efficient and useful one for the curation approach context.

Related work

Much research commended the binary choice compared the other methods like the numerical 10-point scale. It reveals that comparison can be the most efficient way for people to judge the images for evaluating images. However, when it comes to the enormous image dataset, using a binary choice will lead participants to repeat the evaluation process many times. So four-choice can be another opportunity for users to evaluate the images. Based on the comparison between binary choice and four-choice formats, I want to figure out which one is most effective in the curation approach context.

Method

Two hundred fifty-eight images are selected from 4 GANs as the materials for people to rate.
Four surveys are conducted to find the advantages and disadvantages of each format.
They are Binary-Choice-1 (BC1), Binary-Choice-2 (BC2), Four-Choice-1 (FC1), and Four-Choice-2 (FC2). The BC1 and BC2 surveys were 128 pages long and had a binary-choice design format per page. While completing the survey, the participants had to choose the image they are “the most pleasing to the eye.” The FC1 and FC2 have a four-choice design format per page, and the participants had to choose from the four images that they found “the most pleasing to the eye.”

Comparison between binary-choice and four-choice

The survey results show that both binary-choice and four-choice systems have merits and flaws.
For the people preference, people show a more positive attitude to binary choice because it is much easier for them to compare when there are only two options.
However, when it comes to the reality of the result. For binary-choice, if we change the images selected from the same GANs, they will influence its results significantly. In other words, for one GANs, using the binary-choice format cannot reveal its accuracy because the result from binary-choice will be affected by the images chosen by GANs. However, for the four-choice system, the choice of images from one GANs will not influence people’s rating for the GANs.

The choice of the format for the curation approach

For users, we want to get accurate results from people. Also, this system should be easy and efficient to operate for humans from any background.
According to the different factors in the two formats and the context for the curation approach, the four-choice format is selected in the rating step.
There are two reasons for it:
First, the four-choice format can guarantee the reliability of different GANs. The curation approach aims to enable designers to improve all kinds of images' aesthetic qualities, which means that its result should be equal for different GANs. Moreover, the four-choice format is not influenced by the GANs.
Second,four-choice can decrease the rounds to improve the efficiency when the number of images is enormous. In the curation approach, the number of images is more than 1000. Though people prefer binary choice due to its convenience for comparison, when the number of images becomes more extensive, the selection rounds in the survey will become more critical. Moreover, the four-choice format can help people save time and decrease the round they need to repeat.

Therefore, the curation approach selects the four-choice format to get people's experience with images.

6.1.7 Evaluation

The evaluation part aims to prove the validity of the curation approach. The curation approach hypothesizes that retraining using highly rated images will improve the aesthetic quality of the output images.

Related work

Some studies have mentioned the importance of the input dataset's quality for GANs' output. However, when it comes to the iteration of the input dataset, no existing literature mentions that a human's highly-rated output can be directly used. Based on the research to examine the improvement when replacing g the input dataset with highly-rated output, combining the context of the curation approach for designers and users, I want to evaluate the feasibility of the curation approach to inform AI systems user experience.

Method

The research examined the size of the dataset and the number of iterations that bring significant improvement in how pleasing the images are. Following the curation approach, the research got a set of 6,000 images and let people select the pleasing images. After getting the pleasing dataset, create subsets with different amounts of the good output (500,1000,2985) and different iterations (80kimg,200kimg,500kimg) to investigate their improvement. As a result, nine fine-tuned models and the Landshapes model, as the baseline, were used for comparative analysis.
Crowdsourced workers curated the images generated by each network to gather ratings and assess their perceived aesthetic quality.

Discussion

In the curation approach, we assume that GANs can obtain the human experience in the form of "good output" selected by humans as their inspiration, and the designer can "translate" the human experience through the action of "using excellent output as input and retraining the GANs." Based on the two assumptions and research results, the curation approach's feasibility is proved in the following perspectives.

1. Highly-rated output selected by humans contains the aesthetic factors
From the highly-rated images selected by humans, we can see that most of them have subjective and individual factors influencing people's judgment. For those factors, due to their more substantial connection with individual experience and preference, it is hard for algorithms to quantify them directly. The high qualities of factors in the highly-rated output show that they can be a valuable representation of human experience in the context of images.

figure55. Beautiful images selected by humans

In these beautiful images(figure 55) selected by humans, their factors are consistent with qualitative research's personal and subjective characteristics.

1.1 The excellent memory evoked by images
All images with the highest ratings are coastal images because, for most humans, the memories of the ocean are always connected to the beautiful sceneries and the charm of nature. In qualitative research, the memories evoked by images are a significant factor in aesthetics.

1.2 The combination of the background and the key elements
In the qualitative research, the people mentioned that a good combination of the background and the key elements would improve their ratings. Those highly-rated images have a transparent background(ocean) and the key elements(land). The layout for them is comfortable.

1.3 A clear content
In these highest-rating images, it is straightforward for people to recognize the content and information they provide. Explicit content is also an essential factor for aesthetics.

2. Retraining the GANs with excellent output is efficient in improving GANs' output

Comparing the results, we can see that the output of the three input datasets with high-rated images has higher scores than the random input dataset's output. It shows that people prefer the model with a new input dataset. Also, It proves that during the curation approach, GANs understand people's experiences successfully since, after the iteration, it is enabled to provide the output with a better human experience. So, "using excellent output as input and retraining the GANs" is a valuable method for designers to "translate" the human experience to AI systems.
Regardless of the influence from iteration, the input with 1000 good output got the highest score. In the research, the number of input images is 4000. Therefore, designers can replace the input dataset with highly-rated images from people, and the ideal number of highly-rated images is a quarter of the input dataset.

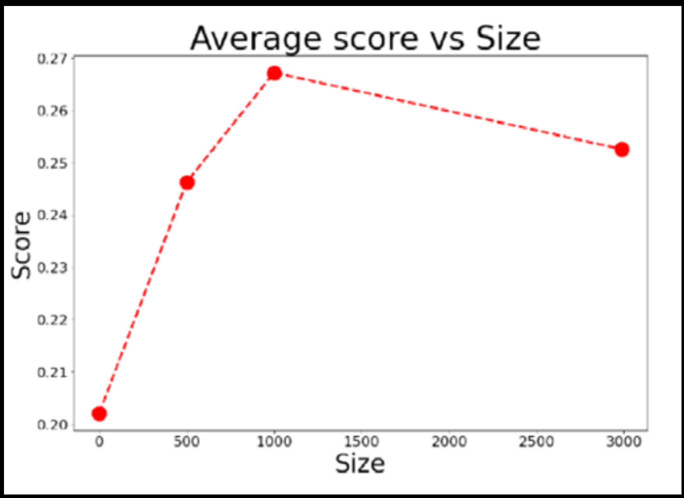


figure56. Average score vs Size

6.1.8 Limitations

There are also some limitations in the curation approach we must consider when designers use it.

The influence of iteration
In the evaluation research, we also find that apart from the size of images, the iteration the designer chooses when training the GANs will also influence its output. Though 1000 images are the best for the average score, the combination of 500 kimg and 2985 images works best considering the iteration in the training process. So when using the curation approach, to get the best output, designers should also consider the iteration in training the GANs.

The different contexts for the curation approach
In the curation approach, the scope of the human experience is narrowed down to "aesthetics ."However, in the natural human world, the human experience has broader factors like "ugly," "surprising," etc. When it comes to other dimensions, does replace the input dataset with the images having aiming factors still work? For example, will the ugliness improve when designers retrain the GANs using the ugliest output selected?

6.2 ALGORITHMIC AESTHETICS APPROACH

6.2.1 What is the algorithmic aesthetics approach?

The algorithmic aesthetics approach is a method to improve the quality of GANs' output from quantitative iteration, and it complements qualitative research and computational analysis. The algorithmic aesthetics approach can provide designers with a more nuanced and precise method to inform AI systems human experience. Guided by the algorithmic aesthetics approach, designers can iterate the GANs' output in more detail.

In the algorithmic aesthetics approach, by transferring the factors from people's experiences to the algorithms and designing the computational models that can predict human ratings of beauty, GANs can understand people's experiences and improve their output.

In this process, the system obtains the human experience in the form of "good factors," which can be predicted by aesthetic models in GANs,
Moreover, the designer "translates" the human experience through the action of "finding the factors that influence people's rating and also can be measured by automated measures in algorithms."

6.2.2 Metaphor

In this process, it is just like we replace a painter's oil color with those colors that people prefer to enable the painter to draw the painting(figure57).

figure57. The metaphor

6.2.3 Participant

1. Designer
For the algorithmic aesthetics approach, designers need some AI design experience. It will also be helpful when designers have basic knowledge about GANs system.

2. Developer
The algorithmic aesthetics approach relies on statistical and computational methods; it involves critical activities such as data analysis and iteration of GANs model. So the participation of developers is essential. Ideally, there are two types of developers in the whole process.
The first is the data analysis programmer. They should be able to find the correlation between the algorithmic measurements and human rating data.
The second is algorithm engineer. They are supposed to build the new model in GANs' architecture.

3. User
The algorithmic aesthetics approach relies on the aesthetic factors that influence people's ratings. Some workshops will take place to get insights from users. Usually, six to eight participants will be enough to get those factors.
There are no limitations on the background of the users.

6.2.4 When do we use it?

For the input dataset where the images types are single

The algorithmic aesthetics approach can improve the precious factors for the images. So, it works better when the types of images in the input dataset are single. The factors that influence aesthetics sometimes vary for different kinds of images. However, when all the photos in the dataset belong to one type, improving the aspects via algorithmic is more efficient.
For example, when designers want to generate fake images for landscapes, their aesthetic factors, such as colors, saturation, etc., can be the same. When the images are people, the factors are people's emotions, hair, etc. In the algorithmic aesthetics approach, designers can improve the specific factors for the images.

For an AI product design team

The process of the algorithmic aesthetics approach contains experts from different fields, including the algorithm and data analysis. So an AI product design team can complete the whole process, and the designer can play the role of the leader to facilitate the entire approach.

For a precious iteration

Guided by the algorithmic aesthetics approach, designers can translate the human experience in a more precise way instead of improving the output generally. Because in this process, designers translate the human experience to the quantified factors. Improving those factors via algorithm enables AI systems to understand the human experience in their whole "language" and mindset.

6.2.5 How to use it?

Facilities

1.GANs

Designers should operate the GANs expertly. Although they do not need to learn the algorithmic theory, they should know how to train the GANs and get the output.

2. Qualtrics

ALGORITHMIC AESTHETICS APPROACH

In the algorithmic aesthetics approach, by transferring the factors from people's experiences t
of beauty, GANs can understand people's experiences and improve their output.

Participants

1.Designer

Designers need some AI design experience

2.Users

The minimum number of users is 4.

3.Developer

The first is the d
The second is al

Designers — **Users** — **Data pro**

Gather the input dataset

Gather random images without any processing

Make the rating survey

1. Upload the images to Qualtrics
2. Make the rating question in Qualtrics

Rate the images

Send the link to participants to complete the survey.

Filt bes

Prog
and
imag
imag
featu
Mea
Regr
the h
filter

Train the GANs

1. Connect to the drive
2. Upload the input folder
3. Generate the images

o the algorithms and designing the computational models that can predict human ratings

ata analysis programmer.
gorithm engineer.

Context

1. For the input dataset where the images types are single
2.For an AI product design team
3.For a precious iteration

a analysis grammer **Algorithmic engineer**

ter out the the st predictors

Build the aesthetic model in GANs

Train the GANs

grammers extract the visual spatial feature data for each ge in OpenCV. Combine ges' rating scores and ure data through PCA, K-ns Clustering, and ression, the features with highest correlations will be red out.

In latent space, algorithm engineers can find the regions that represent the best predictor features and then iterate them based on the data from last step.

1. Connect to the drive
2. Upload the input folder
3. Generate the images

figure58. The algorithmic aesthetics approach

Steps
For the algorithmic aesthetics approach, there are six steps.

STEP1 - Get the random input dataset
For the algorithmic aesthetics approach, there are six steps.

In this process, the designer first gathers random images as its first input dataset. In the first round of the training, the designer does not need to do any processing on the dataset. They can collect any type of image.
For the number of the images, ideally, for GANs, ~4000 images can be enough dataset for it to generate fake images.
All the images should be compressed in a folder and uploaded to the drive.

STEP2 - Train the GANs
After getting the dataset, designers use them for training the GANs and getting the output.
2.1 Connect to the drive
Click the “play” button to view the system’s GPU and connect the drive to the co-lab.

2.2 Upload the input folder.
Enter the location of the compressed file into “dataset,” and “resume_from” will determine the final training result to a certain extent. Under this model, if it is a new train, choose ‘ffhq1024’ or ‘./pretrained/wikiart.pkl’. Click the “play button” after making changes.
Click the “play button” again; the model will start training to generate images.

2.3 Generate the images
After setting the parameters like network, seeds, and truncation, designers click the play button to get the images, and finally, they can get the output in the drive's folder.

STEP3- Let the user rate the output
After getting the output, designers should create a survey for people to rate the output in Qualtrics.

Platform

1. Qualtrics
Designers build a rating survey in Qualtrics to get human's numerical rating of the output.

Materials for the survey
1. Images
To investigate the factors that influence people's ratings, ideally, it is more effective when more images are rated by people so that people will have a more comprehensive understanding of the aesthetic judgment of pictures. One hundred images are enough for people to evaluate the images. The number of their original datasets does not influence the number of images for people to rate.
The size of the images is 1024*1024.

2. Question
Based on the exploration of aesthetics, the definition of aesthetics is narrowed down to the scope of the perceptual level, and the visual sense is selected as the evaluation scope. Then, from the APiD Scale, “pleasing to see” is chosen to be the question for the aesthetics of the images. In the survey, humans rate the aesthetic value for each image between 0 and 9. The question is, “To what extent is this picture pleasing to your eyes?”

3. Consent form
The consent form is a document signed by persons of interest to confirm that they agree with an activity that will happen and that they are aware of the risks or costs that may come with it.

4. Example question
Some example questions can ensure the participants understand the survey process completely.

Process

3.1 Rename the images
Put the output from GANs into folders and rename them by their orders.

3.2 Upload the images to Qualtrics
Create graphic folders in Qualtrics and upload the images there.

3.3 Create the survey in Qualtrics
Create a rating survey in Qualtrics.

3.4 Add other materials in Qualtrics
Manually add the consent form, example question, attention checking questions, and attitude questions to Qualtrics and finalize the survey.

3.5 Start the survey
Send the link to participants to complete the survey.

3.6 Download the CSV. File with human rating data

STEP4- Find the applicable factors with the data analysis programmer
After getting the scores from humans, designers should cooperate with programmers to find automated measures of aesthetic beauty.

4.1 Find images factors that can be automatic measured
Designers provide programmers with some factors of images, and then programmers filter out factors that can be automatically measured.
During our research, we have found some factors for people's reference:
Visual features: Saturation / Luminance / Contrast / Sharpness / Colour Histogram
Spatial features: Rule of Thirds & Diagonal Dominance / Rule of Thirds & Diagonal Dominance /Symmetry / Line Orientation Ratios

4.2 Filter out the automated measures that be the best predictors for human aesthetic ratings

Platform:OpenCV
Designers give developers the CSV file with the people's rating scores and the images. Then, the programmer can extract the visual and spatial feature data for each image in OpenCV. Developers combine images' rating scores and feature data through PCA, K-Means Clustering, and Regression, the features with the highest correlations will be filtered out.

STEP5 - Build the aesthetic model in GANs

Designers provide the algorithm engineers with those features. In the latent space, algorithm engineers can find the regions that represent those features and then iterate them based on the data from step 4.2. For algorithm engineers, the GANsAesthetic approach is a valuable method for completing the iteration and building the new model.

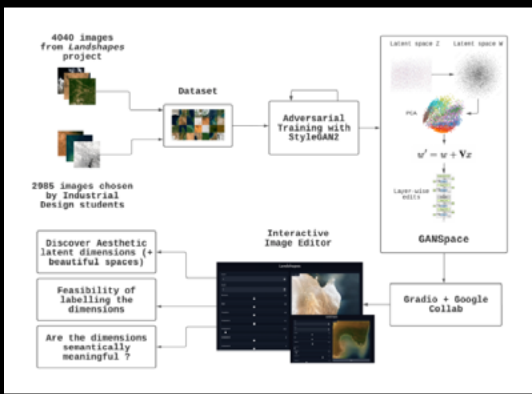


figure59. The process of GANsAesthetic approach

STEP6- Retrain the GANs
Finally, designers retrain the GANs based on the new GANs.
6.1 Upload the input folder.
Enter the location of the compressed file into “dataset,” and “resume_from” will determine the final training result to a certain extent. Under this model, if it is a new train, you can choose ‘ffhq1024’ or ‘./pretrained/wikiart.pkl’. Click the “play button” after making changes.
Click the “play button” again; the model will start training to generate images.

6.2 Generate the images
After setting the parameters like network, seeds, and truncation, designers click the play button to get the images, and finally, they can get the output in the drive’s folder.

6.2.6 The role that the designer plays

Facilitator
This process involves different kinds of people including designers, data analysis, programmers, algorithm engineers. AI product designers should facilitate the whole process. Also, since AI product designers have the knowledge in design and technology fields, they are able to communicate with programmers better when they need to transfer the information.

Catcher
In this process, based on their human-center design background, the designer should get the information from people accurately. For example, when people select the good output, designers should analyze the factors from those images accurately and efficiently. It means they should be a good "catcher" to filter out the factors that people really care about in those images.

6.2.7 How do we achieve it?

The algorithmic aesthetics approach hypothesizes that computational models can be designed to predict human ratings of beauty. In this process, designers “translate” the human experience to GANs through the action of “finding the factors that influence people’s rating and also can be iterated in algorithms.”
Two sub-questions are put forward:
1. Whether the factors from humans can be measured by algorithms?
2. Whether GANs can improve those automated measures?

Two experiments were conducted to investigate those two questions, and finally, the achievability of this method was confirmed.

Using Automated Measures of Aesthetic Beauty to Improve GANs Output of Satellite Images

Method
The research shows that some computational factors correlate with human rating. Firstly, visual and spatial features that are computational are filtered out. Then, those features are extracted from an image dataset. For the human rating data, crowd-sourced workers are asked to fill a survey attributing an aesthetic value between 0 and 9 for each image. Finally, in the form of an Ordinary Least Squares Regression model, the statistical analysis of the outputted values is performed to find the correlation between the features and human ratings. The features highly correlated with human ratings are the quantified factors in building the new model.

Take away

This research shows us that algorithms can quantify some aesthetic factors from human ratings.

1. Most properties of the image that can be quantified
During the previous qualitative research, it is assumed that some aesthetic factors can be quantified, and algorithms can measure them. In this research, those aesthetic factors are filtered out. They can be the reference for designers when selecting elements to compare with human ratings.
The features are:
Visual features: Saturation / Luminance / Contrast / Sharpness / Colour Histogram
Spatial features: Rule of Thirds & Diagonal Dominance / Rule of Thirds & Diagonal Dominance /Symmetry / Line Orientation Ratios

In those quantified features, we can see that most of the image's properties are computational. Moreover, they are not related to the content carried by the image. It means that designers can use the algorithmic approach to iterate different types of images like landscape, followers, people, etc.

2. Some factors are strongly correlated with people's ratings
In the regression results, it shows that for this dataset, saturation and color histograms,
novelty, contrast, straight to diagonal line ration, horizontal to vertical line ratio, and symmetry highly correlate with human ratings.

Therefore, it proves those features can be good predictors of human experience. Iterating those quantified factors through algorithms can improve the human experience. Designers can translate the human experience into quantified features.

In summary, for the first sub-question, the factors from humans can be measured by algorithms. Moreover, designers can confidently translate the human experience of various types of images into quantified features.

OLS Regression Results

Dep. Variable:	aesthetic	R-squared:	0.452
Model:	OLS	Adj. R-squared:	0.421
Method:	Least Squares	F-statistic:	14.22
Date:	Wed, 15 Jun 2022	Prob (F-statistic):	1.40e-24
Time:	13:43:10	Log-Likelihood:	558.43
No. Observations:	256	AIC:	-1087.
Df Residuals:	241	BIC:	-1034.
Df Model:	14		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.3580	0.060	5.964	0.000	0.240	0.476
novelty	0.1030	0.033	3.137	0.002	0.038	0.168
typicality	-0.0389	0.047	-0.826	0.410	-0.132	0.054
histograms	-0.3699	0.047	-7.939	0.000	-0.462	-0.278
entropy	-0.0068	0.004	-1.834	0.068	-0.014	0.001
straight_diagonal_line_ratio	-0.0646	0.026	-2.505	0.013	-0.115	-0.014
horizontal_vertical_line_ration	-0.0747	0.035	-2.123	0.035	-0.144	-0.005
diagonal_dominance	1.261e-06	5.21e-06	0.242	0.809	-9e-06	1.15e-05
symmetry	-0.0001	5.69e-05	-2.090	0.038	-0.000	-6.83e-06
rule_of_thirds_power_points	-4.082e-05	4.28e-05	-0.953	0.342	-0.000	4.36e-05
rule_of_thirds_gridlines	4.596e-05	4.95e-05	0.928	0.354	-5.16e-05	0.000
sharpness	-5.001e-07	1.35e-06	-0.371	0.711	-3.16e-06	2.16e-06
contrast	0.0002	6.75e-05	2.848	0.005	5.92e-05	0.000
luminance	0.0001	5.83e-05	1.856	0.065	-6.61e-06	0.000
saturation	0.0002	3.93e-05	5.222	0.000	0.000	0.000

figure60. The result for the quantified features

Exploring the best algorithmic methodologies to iterate quantified features in GANs

Method
The research shows the GANs ability to iterate the output from the algorithmic perspective. Firstly, explore different methodologies in GANs to iterate its output. Then, GANsAesthetics is the most efficient way for cognitive scientists and design engineers to study factors contributing to the aesthetic or to study semantic representations in perceptual spaces.

Take away
This research shows that GANs can improve computational features in the first experiment. By cooperating with algorithmic architecture, designers can translate the qualitative human experience to quantitative AI systems through the computational model.

Comparison among various algorithmic approaches

After getting the quantified features as the human rating predictors, designers can use the algorithmic approach when GANs can iterate those factors in their architecture. While exploring GANs, we find GANs include various algorithmic components like latent space, generator, discriminator, etc. Furthermore, many existing approaches are conducted to improve the beauty of the images generated by GANs.

A. Human-based discriminator

This method adds humans as an additional assessment in the discriminator. In the GANs, discriminators can determine whether the data generated from the generator is real or fake. People's evaluations, such as how beautiful the images are, are added to a human-based discriminator.

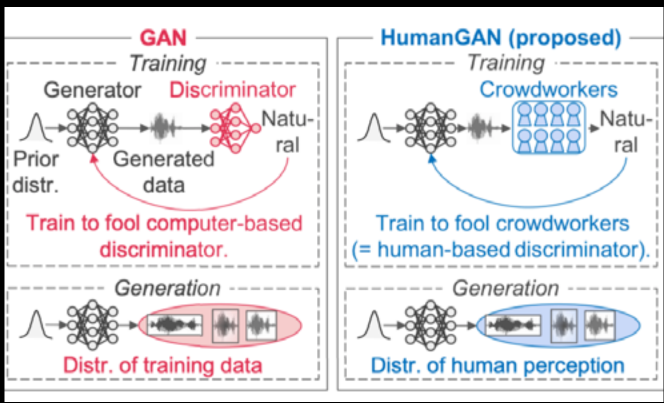


figure61. The structure of Human-based ddiscriminator

Based on the research from Japanese researchers, this method can represent a human's perception distribution.
[figure 62] shows the transition of generated data during training. A brighter grid means people's rating is higher. The figure indicates that the generated data transited from a lower probability range (darker area) to a higher probability range (brighter area) during the training iteration. It means that during the iteration of HumanGANs, its output can represent a human's perception distribution.
However, like the curation approach, they did not improve a specific factor of the picture. Still, they screened out the output that conformed to people's perception through people's overall evaluation of the output.
Therefore, it is unsuitable for the algorithmic aesthetic approach, which needs to iterate the features from images.

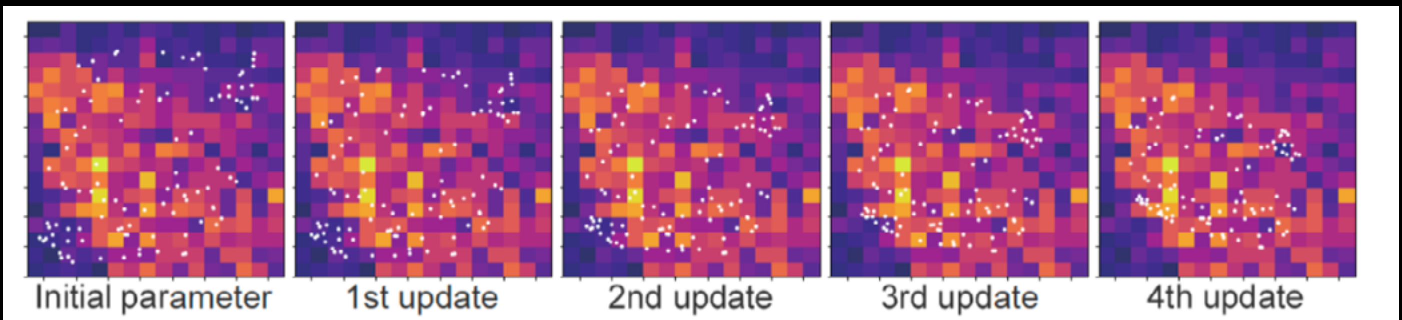


figure62. The transition of generated data during training.

B.Deep learning approach for human rating data prediction

For the deep learning approach, some neural nets are put in the new model to predict human ratings on the set of images. Few neural networks have shown significant improvement in performance, such as AlexNet, VGG-19, and resnets.
Up to now, most of these neural networks are binary classifiers that predict whether the image is aesthetic or not. It means they can not find the exact factor to improve the quality of the picture. For developers, neural networks are called the" black box. Models" that are well-known as non-identifiable models. So it is hard for developers to get the approximation function.
Thus, the deep learning approach is not helpful in the algorithmic aesthetic approach.

C.GANsAesthetic
GANsAesthetic iterates the images by adjusting the critical semantic latent directions in GANs. The procedure starts with data collection and training StyleGANs2. Afterward, using GANspace to obtain important semantic latent directions, Gradio and Google Collab construct the user interface with the sliders.
In exploring GANs, we know that the latent space represents the different features from the real images. After finding those regions that represent the quantified factors, algorithm engineers can iterate those regions based on the data from designers.

Ultimately, developers can use the GANsAesthetic approach to build the new model in the algorithmic aesthetic approach.

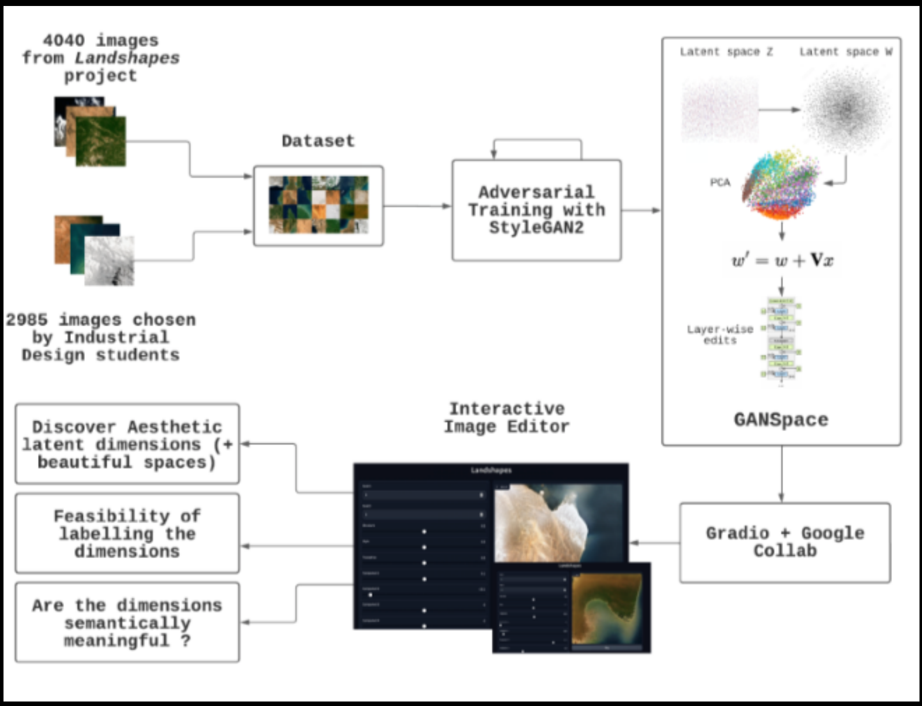


figure63. The sturcture of the GANAesthetics

6.2.7 Evaluation

The evaluation part aims to prove the validity of the algorithmic aesthetic approach. The algorithmic approach hypothesizes that computational models can be designed to predict human ratings of beauty. In cooperation with CSE students, the algorithmic aesthetic approach‘s achievability has been demonstrated. Some auto-measured factors by algorithms can be good predictors for human aesthetic evaluation. Furthermore, adjusting the regions for those factors in latent space can iterate those factors. Following the algorithmic aesthetic approach, the evaluation proves the approach’s effect on improving GANs‘ output quality.

Method
In the evaluation, to ensure the accuracy of the factors. Considering the algorithmic aesthetic approach’s context - For the input dataset where the images’ types are single, three different images of seeds generated by the StyleGAN2 are used in the experiment. The first image depicts a coastline area, the second a forest and desert area, and the third depicts an arctic area.
There are two sub-experiments in the evaluation.

1. Find the predictors
In GANAesthetic, by adjusting the different properties of the image represented by the latent space, to discover the aesthetic factors in latent dimensions. In this process, 33 participants are asked to pick three rows of each of the different images that they deem the most beautiful/pleasing to the eye.

2. Comparing the scores between the old model and the new model
After finding those factors' regions, 10 participants are asked to rate the pictures with the changed factors in the new model and the original pictures. The comparison proves the validity of the new model. The images are shown randomly. The question for each image is "whether the image is pleasing to see."This is one of the items from "The Aesthetic Pleasure in Design (APiD) Scale." Meanwhile, based on the theory of APiD, The optimal number of rating bars are '1~7'. (1 = strongly disagree, 7 = strongly disagree agree)

Result

In GANAesthetic, different components represent means the different regions in latent space. They represent various factors of the images. As mentioned above, different types of images have different predictors. The experiment's results are shown in figure 64.
People's scores for the original output and new output, whose factors are adjusted in latent space, are shown in figure 65.

	COASTLINE	FOREST/DESERT	ARCTIC
c_0	7	25	2
c_1	10	15	6
c_2	2	5	10
c_3	16	1	10
c_4	9	14	15
c_5	2	3	13
c_6	14	1	6
c_7	6	15	3
c_8	6	6	9
c_9	10	4	8
c_{10}	4	4	6
c_{11}	3	5	2
c_{12}	7	1	1
c_{13}	3	0	7

figure64. People's voting for the pictures

For coastline, c3,c6, and c9 receive the most votes. By analyzing the different images in different property values, the factors represented by the latent space can be found. C3 mainly influences the color of the ocean. C6 mainly influences the color and the size of the land. C9 mainly influences the overall color.

In summary, the coastline's color and the land's size are predictors of its aesthetics. In figure6, for all kinds of factors, the new models are higher than the baselines. For example, for C3 and C6, when the factor is adjusted to -2, the images get the highest score. For C9, after the factor is adjusted to 2, the score is the highest. It means that the coastline, by changing the latent space and adjusting the corresponding factors (the coastline's color and the land's size), the output quality generated by GANs can be improved.

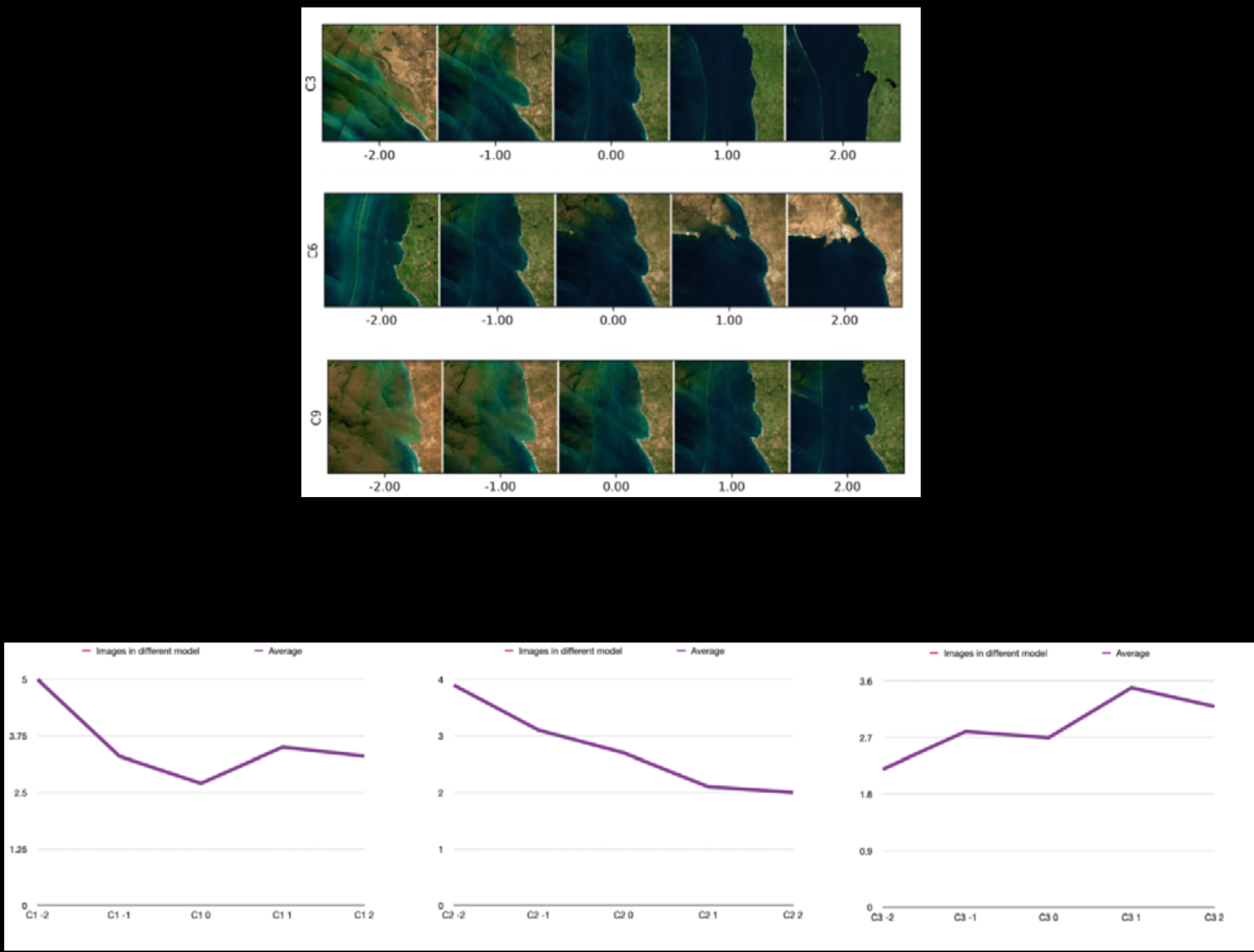
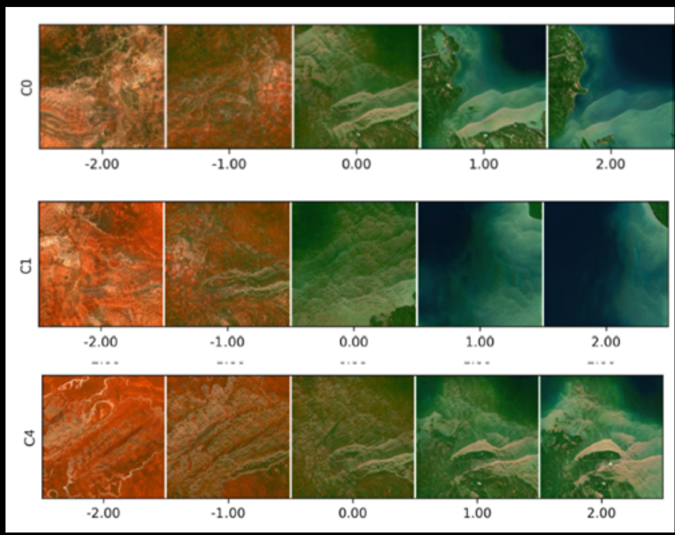


figure66. Ratings for different factors



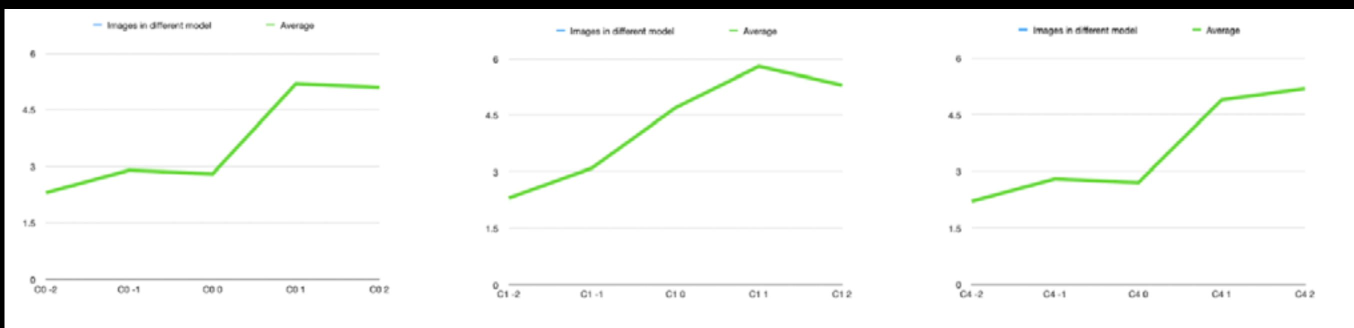
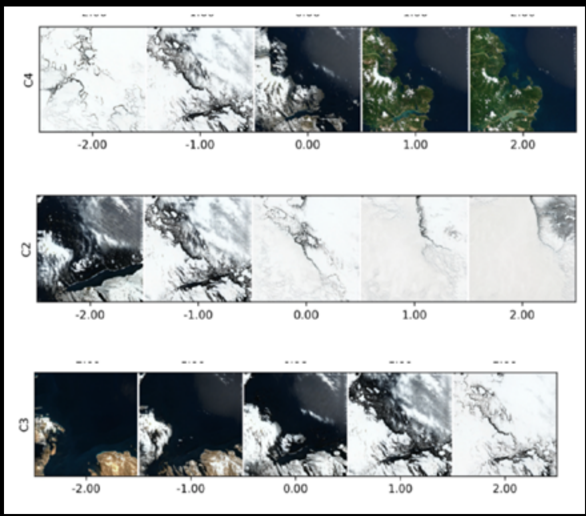


figure67. Ratings for different factors

For forest/desert, c0,c1, and c4 are the top three factors related to human evaluation. They represent that c4 represents the image's color, and c0 changes the shape and the color of the ocean. The new models generate better factors for all kinds of factors than the original ones. For c0 and c1, the value of 1 is the best. For c4, people like it most when the factor's values are adjusted to 2ost. The baselines do not always have the lowest scores, but they are still lower than some new models.



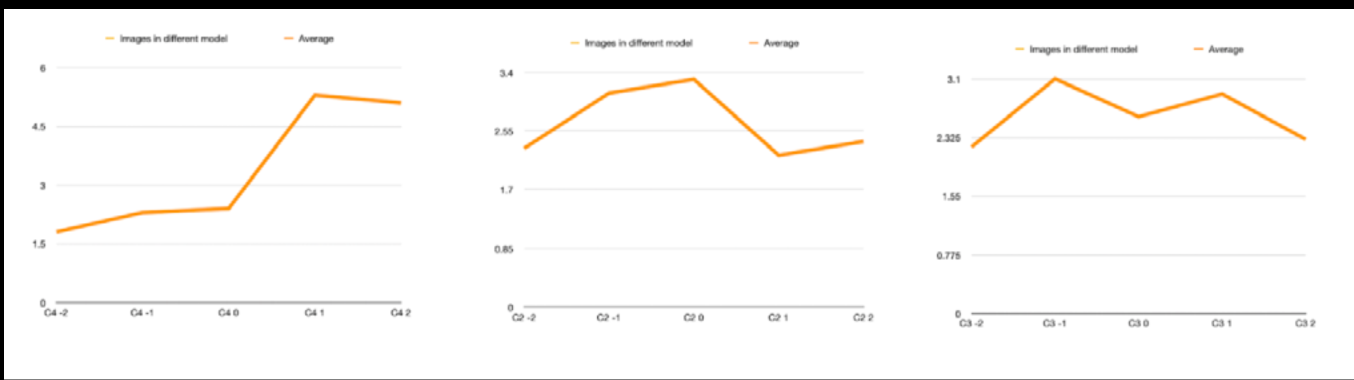


figure68. Ratings for different factors

The ratings in c4, c2, and c3 are the highest for the arctic. For c4, it changes the color of the images. c2 mainly changes the resolution of the images. c3 represents the brightness of the image. For c4 and c3, when their values are adjusted in latent space, the scores are higher than the baseline. However, for c2, the baseline is the highest score.

Discussion
According to the results from three types of images, we can conclude that the algorithmic aesthetic approach can inform GANs people's experiences. By cooperating with developers, designers can translate human experience into GANs and steer their output.

1. Algorithmic aesthetic approach is effective
The first experiment proves that latent space can find the predictors for human ratings. Furthermore, the comparison between the new and original models shows that the output generated by new models can get higher scores for most factors. Following the algorithmic aesthetic approach, people's evaluation of the images can be translated into GANs.

2. The factors of different types of images vary greatly
For different types of images, their factors vary considerably. The algorithmic aesthetic approach will be more beneficial when the input dataset includes single images. For example, when the input dataset only includes the coastline, designers can adjust the factors of the coastline more preciously in the latent space.

6.2.8 Limitations
There are also some limitations in the algorithmic aesthetic approach designers must consider when using.

The influence of different measurement difficulty of factors on the correlation

In step4.2 -Filter out the automated measures that are the best predictors for human aesthetic ratings. We can not ignore that the difficulties in measuring them are different for various features. For example, saturation only requires one parameter of the image. However, one image will have countless data for a color histogram, and many parameters like the RGB values need measuring for this factor.
Would a more easily measurable factor enhance its correlation with the rating? For those factors that are not easily measurable, will the difficulty of quantification reduce its correlation with the rating? More research to investigate this influence should be conducted in the future.

The accuracy of automated measures of aesthetic beauty

I have investigated people's definitions of aesthetic factors in qualitative research. The descriptions of the same factor usually change based on the image's content. For instance, when the image is the beach, people pay more attention to color matching when it comes to color. However, when the content is the ocean, the gradient of color will be more critical.
In the automated measures, color is only quantified by a color histogram. Whether it can represent the real definition of color for the human world? Designers and developers need to find more accurate automated measures that can represent human definition.

The accuracy of latent space for the representation of factors
Adjusting one region of the latent space for some images through the evaluation part will influence several factors. Therefore, how to find the latent space that represents the factor accurately. For designers, cooperating with developers to find the accurate regions that predict people's evaluation is essential.

07

QUALITATIVE RESEARCH ON CHANGING SYSTEMS

7.1 THE IMPORTANCE OF CHANGING

7.2 GANS' FEATURES INFLUENCING

7.3 QUALITATIVE RESEARCH

7.1 The importance of changing systems

The evaluation of LANDSHAPES showed the ability of GANs to produce media that evoke an emotional response in the viewer. Visitors mention that unidirectional interaction from the images' transitions allows them to push for change in the shown landscape. Also, In today's world, video feeds such as YouTube and Tiktok are playing an increasingly important role in people's lives. Compared with static information like images, dynamic data such as video can bring more helpful information to the human world in the future. There has been some research on enabling GANs to generate more real videos from an algorithmic perspective. However, no experiment on whether the videos satisfy humans and what factors influence the human experience in videos. Therefore, for GANs, video is another form of output worth exploring. Qualitative video research explores the human experience in the videos generated by GANs and investigates the future direction of the translation between the AI and human worlds.

7.2 GANs' features influencing videos

GANs videos can show transitions among the generated images. Through some practical operation and desk research, some features in GANs that influence the quality of videos are summarized. Those features can show GANs' potential abilities to understand the human experience of changing systems.

7.2.1 Different types of videos in GANs

GANs provide us with two kinds of videos that show the transitions of images. Videos are produced by interpolation. Interpolation is the process of generating minimal changes to a vector in order to make it appear animated from frame to frame.

Linear interpolation
The first is linear interpolation. It is the most popular type of interpolation for GANs. For linear interpolation, we need to choose various seeds as their materials. It moves from seed to seed, and there is an equal number of frames between each seed. Each image corresponds to a point. During this process, interpolation moves smoothly with latent space between different points. So, for linear interpolation, it will be helpful if you have some specific images.

Slerp Interpolation
Slerp interpolation has the same process as linear interpolation. Technically, it will be more suitable for high-dimensional GANs. Nevertheless, there is no vast difference between linear interpolation and slerp interpolation.

Noise loop interpolation
The second is noise loop interpolation. When making a noise loop video, we only pick one image. This seed will be the start point and end point. Noise latent interpolation does a random walk in latent space and then returns to where it begins. Compared to linear interpolation, the noise latent interpolation will be more smooth since it will choose the transition frames in the latent space by itself. Nevertheless, you can not get control of the specific points during the noise latent interpolation.

Circular loop interpolation
Like the noise loop interpolation, a circular loop also provides a random loop starting from one seed. Nevertheless, the loop is more smooth than the latter one.

In applying GANs' video, showing the transition of multiple images is a more common scenario. At the same time, more changes in the video can bring people a richer user experience. Therefore, linear interpolation is selected as the material for the experiment.

7.2.2 The parameter of GANs

Truncation
Like the truncation for the images, it will define the realness of a video. The smaller the number, the more realistic images should appear. Nevertheless, this will also affect diversity. Most people choose between 0.5 and 1.0, but technically it is infinite.

Frames
Frames in a video will influence the speed of transition of a video. It will define how many frames you want to produce. Use this to manage the length and the speed of the video.

7.3 Qualitative Research Process

7.3.1 Introduction
Qualitative research aims to find the factors influencing people's evaluation of videos. By finding the characteristics and the stamps people like and dislike, human experience in the transition produced by GANs can be concluded.

7.3.2 Aim
- Find the factors that influence people's evaluation of aesthetics in video
- Compare the factors between video and image

7.3.3 Research questions
The research question is:
What influences human experience in the fake videos generated by GANs

There are some sub-questions in this research:
1. What parameters will affect the user's judgment of video aesthetics? What are the most important factors?
2. What are the differences and similarities between videos and images?
3. In transition, which stamps attract people and which moments do not?

7.3.4 Materials
1. A linear loop interpolation.
Seed:1/10/20/30/40/50
Truncation:0.5

figure69. Seeds for the videos

2. Eight clips from the linear loop interpolation

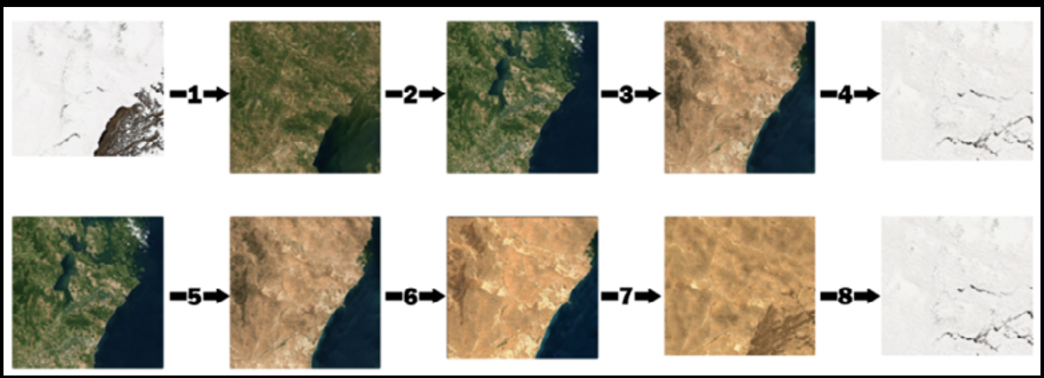


figure70. Clips from the linear loop interpolation

7.3.5 Steps

The research is conducted in a workshop where users are interviewed individually. The overall process is as follows:

Step1-Show the linear video

In this step, find the factors that affect the entire video experience through the participants' overall sense of the video.

Interview question: What factors influence the video's aesthetics?

Step2-Rate the eight clips

In this step, participants are asked to rate the eight clips. More details about the transition's factors can be found by comparing the different periods in the linear loop.

The question is "whether the clip is pleasing to see." The optimal number of rating bars is '1~7'. (1 = strongly disagree, 7 = strongly disagree agree)

Step3-Pick the appealing moments and unappealing moments

Replay the video; participants are asked to pick the appealing and unappealing moments. For each type of moment, participants are asked to choose two images. Because the video is composed of pictures, this step finds the video's factors by investigating the specific images in the video.

The task is: Please find the three most beautiful and least attractive moments in the video. During the process, if you see a relevant moment, just tell me to stop the video. I will record the pictures through video screenshots.

Step4-Compare with images

A group of images is shown to participants to let them compare the factors between the static image and dynamic video.

The interview question: what are the differences between images and videos?

7.3.6 Participants

The participants are all from previous research in image evaluation to better compare images and video.

7.4 Result and discussion

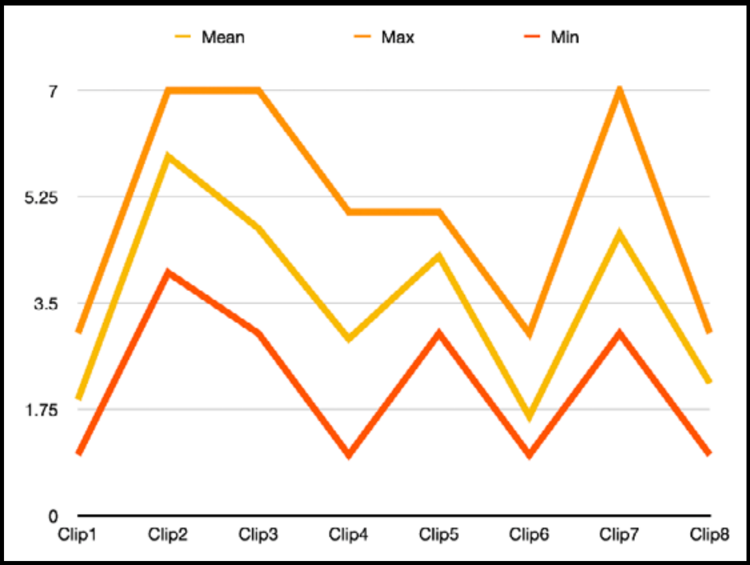


figure71. Rusults for the clips

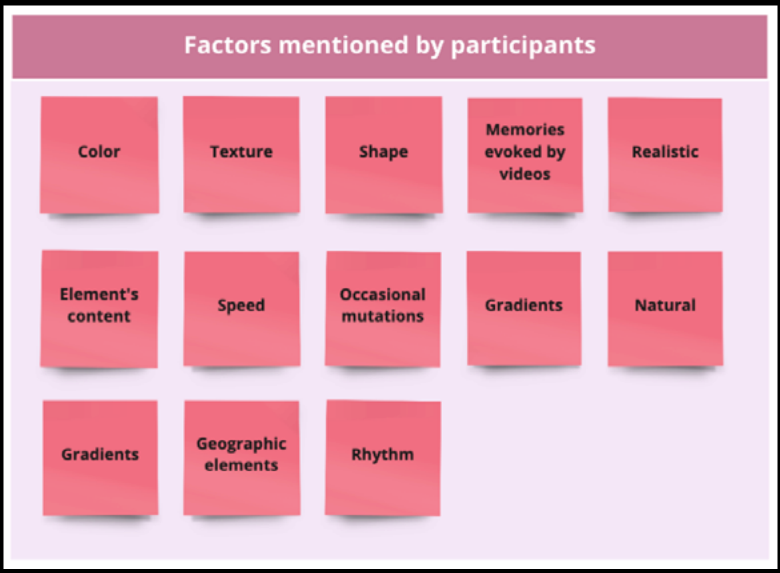


figure72. Factors for the videos

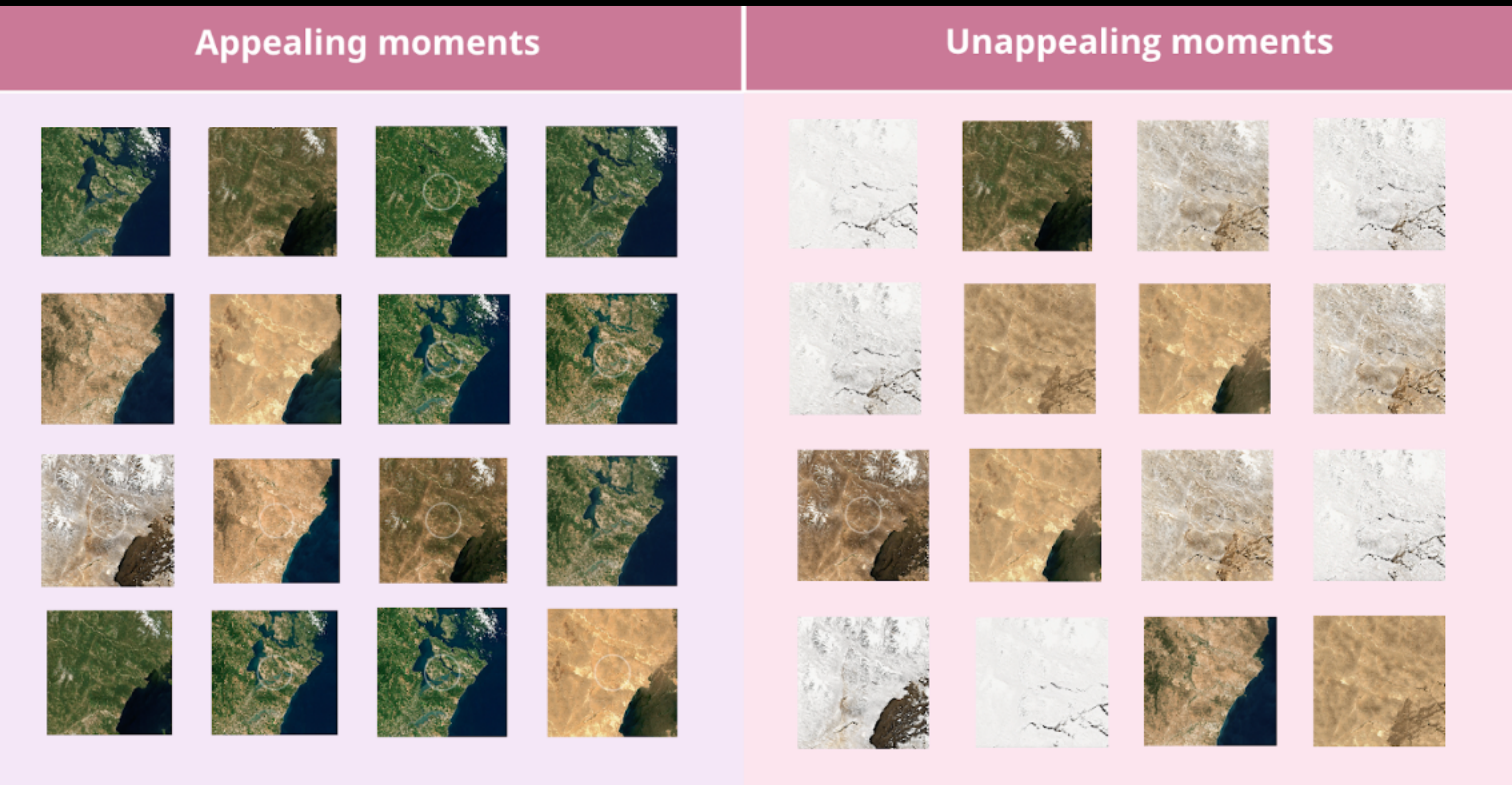


figure73. The appealing and unappealing moments selected by people

The results of rating results for the eight clips are shown in figure71. All the average, minimum and maximum scores are included in the order of appearance of clips in the video.
Figure 72 shows the factors mentioned by participants in the first steps. Because the number of participants in the survey was 10, to ensure the comprehensiveness and diversity of the factors, the factors that two or more participants mentioned are all selected into the final factors. The video's appealing and unappealing moments can be found in Table. X

7.5 Discussion

7.5.1The factors for the changing system

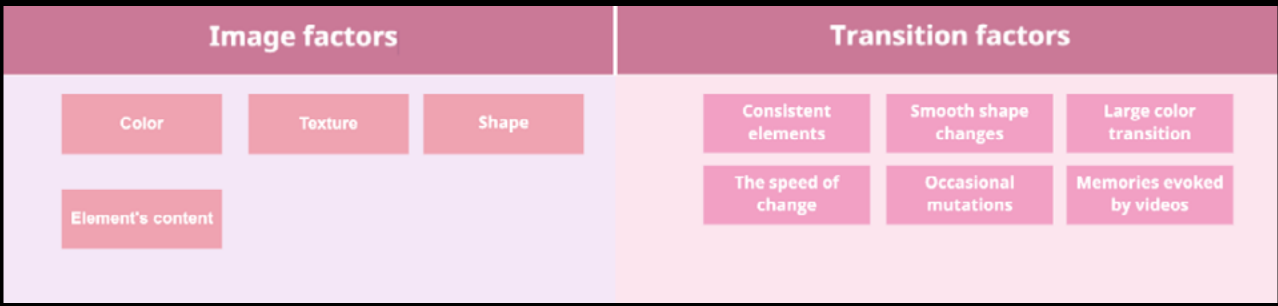


figure74. The final factors for chaing systems

After combining all the survey results, according to the nature of the factors, all the factors can be divided into the following two categories- image factors and transition factors.

1. Image factors

Image factors are related to the images in the videos. They are properties of each frame of pictures contained in videos. On the one hand, they are related to the properties of the seeds that make up the linear loop; on the other hand, they are also the properties of the frames composed of latent space in the transition of each seed.

1.1 Color

All the participants mention the frames' color. Reasonable color matching is essential for users. In the adorable moments for people(Table. X), Most of their colors consist of green and blue. What is more, in the two clips with the highest scores, clip2 and clip3, the pictures that makeup they also contain relatively rich colors.

1.2 Texture
Consistent with the qualitative findings of the picture, the texture of the frames gives people the overall feeling of the video. In experiments(Table. X), frames with relatively soft textures are more appealing than frames with relatively rough textures.

1.3 Shape
Participants also mention that the shape of the elements will influence their experience. For instance, in the unappealing frames, most are unformed shapes. It was mentioned that gentle curves like the coastline enhance their experience.

1.4 Element's content
Whether the frame's information of the element itself can be effectively conveyed is also crucial for videos. In the highly-rated clips (clip2,clip3), their frames' content is easy to understand and recognize. However, for clip4 and clip8, their content is not explicit enough. Participants need to take more time to recognize their information. It will influence their experience.

2. Transition factors

Transition factors focus on the factors in the process of image changes displayed by the video. Different from the information of each frame, these factors are the characteristics that affect the user experience in the picture's changes.

2.1 Consistent elements

In a video, if the elements in its starting and ending frames are consistent, the transition will bring people the most natural feeling. For example, people think that in clip2, only the features of the ocean and land have changed. The elements in the video are still preserved. So the overall transition is natural and harmonious.

2.2 Smooth shape changes

The shape transition of elements can affect people's judgment. For example, in clip7, the blue part gradually changes from a triangle to a point; and the land part gradually expands. This smooth shape changes improve people's experience with this video.

2.3 Large color transition

Unlike the elements which need to be consistent, people prefer the transition where the color changes a lot. For example, in clip2 and clip3, people prefer the latter change because the change from green to yellow in clip3 can make people feel the overall change more obviously.

2.4 The speed of change
Among the videos containing 480/960/1920 frames, people like 960 frames most. On the one hand, people can have enough time to see the changes in “details,” On the other hand, those speeds will not make the overall rhythm too long.

2.5 Occasional mutations
In the video, if there can be one or two relatively simple mutations, it will give people a more impressive impression.

2.6 Memories evoked by videos
If the content of videos can awaken people’s feelings about related memories, people will be more inclined to think these videos are more meaningful. For example, clip3 and clip5 can awaken people’s thinking about the problem of land desertification, while clip1, and clip8, are more like a purely artistic expressions. People do not feel the deeper content.

7.5.2 Analysis of the factors

We can draw some characteristics that affect the video human experience based on the above experimental results.

1. The abstract factors of seed are an essential part
Through the analysis of the overall factor, the quality of the seed in the video significantly impacts the user's experience. On the one hand, they are part of the video content. On the other hand, the difference between seeds can determine their transition. For the factors related to seed, except for color, the others mainly belong to the abstract factors in image qualitative research.

2. Transition factors have a more significant impact on people's human experience
From the perspective of the overall proportion, the number of transition factors is more. It means that in the video, when designers want to translate videos' human experience to AI systems, they need to consider the algorithmic properties of the changing process, such as frames, path selection of latent space, etc.

7.6 Future direction for approaches

Based on the above conclusions, several hypotheses can be proposed for transferring video user experience to AI systems.

Hypothesis one: By selecting the high-rated images as the seeds, the video's quality will be improved.
Whether in generating the video or in human experience, the images that make up the video are vital parts. What is more, most of the factors from images are abstract. Combing the curation approach for the image, electing the high-rated images as the seeds can be a potential method for designers to translate the human experience in videos to GANs.

Hypothesis two: The video's quality will be iterated by adjusting the latent space.
In the factors parts, the transition factors play an essential role. Moreover, by exploring the generation process of GANs, the latent space will define the factors in transition. So, is latent space a good tool for designers to translate those qualitative factors into quantitative features?

08

CONCLUSION

8.1 SUMMARY

8.2 ADDRESS THE QUESTIONS

8.3 CONTRIBUTION

8.4 FUTURE WORK

8.1 Summary

This graduation project aims to investigate how human experience might inform AI systems. The target user is AI product designers working in the AI field. Selecting GANs as the representative of AI systems, this project proves the importance of human experience in steering AI systems' production by enabling designers to translate the qualitative human experience into quantitative AI systems.

The desk research and experiments on GANs show GANs' features and potential to transform the user experience. For GANs, the input dataset is their "raw materials "to generate fake images. Furthermore, their abundance of algorithmic components like latent space and discriminator networks provide ample opportunities to "understand" the information from the human world. Qualitative research on the human evaluation of GANs' output reveals the characteristics of people's experience with AI systems. Different factors like abstract and algorithmic factors need various methods to be translated into GANs. Also, the research with AI product designers figures out designers' needs and the contexts for the project. Based on those insights, two methodologies are put forward to help designers inform GANs of the human experience. Cooperating with CSE students, the pipelines of the methods are built to prove their realizability. Finally, the evaluations demonstrate their effectiveness in translating the qualitative human experience into GANs.

The project's final output is two approaches that enable designers to inform AI systems with human experience. The first one is the curation approach - by putting the beautiful GANs' outputs selected by people into the input dataset and retraining the GANs, GANs can successfully get people's feedback and steer their output. The second is the algorithmic aesthetics approach - by transferring the factors from people's experiences to the algorithms and designing the computational models that can predict human ratings of beauty, GANs can understand people's experiences and improve their output.

In the process of informing AI systems human experience, AI product designers paly an important role. First, they are the FACILITATOR for the whole process. This process involves different kinds of people including designers, data analysis, programmers, algorithm engineers. AI product designers should facilitate the whole process. Also, since AI product designers have the knowledge in design and technology fields, they are able to communicate with programmers better when they need to transfer the information. Second, they are the CATCHER. In this process, based on their human-center design background, the designer should get the information from people accurately. For example, when people select the good output, designers should analyze the factors from those images accurately and efficiently. It means they should be a good “catcher” to filter out the factors that people really care about in those images.

The two approaches in the project prove the possibility for designers to translate human experience into AI systems. They are also a good starting point to explore more methods to inform AI systems of human experience in different contexts. The research process also provides compelling examples of how designers work with developers. Moreover, the research on the human experience of videos generated by GANs offers some direction for future work.

8.2 Address the questions

The graduation project research is "how might human experience inform AI systems."
To solve it, some sub-questions are put forward:
1. What methodologies can be used to understand the information from the human world?
2. What aesthetic factors are significant for the human experience of GANs' output?
3. What are the characteristics of AI product designers?

The final approaches are built on the insights from those three fields. Based on the different features of people's experience with GANs products, two methods help designers translate these qualitative and subjective experiences into GANs in the most suitable ways. The construction process of the methods and the final evaluation confirm the effectiveness of the curation and algorithmic approaches.

The project’s final output is two approaches that enable designers to inform AI systems with human experience. The first one is the curation approach - by putting the beautiful GANs’ outputs selected by people into the input dataset and retraining the GANs, GANs can successfully get people’s feedback and steer their output. The second is the algorithmic aesthetics approach - by transferring the factors from people’s experiences to the algorithms and designing the computational models that can predict human ratings of beauty, GANs can understand people’s experiences and improve their output.

The two approaches in the project prove the possibility for designers to translate human experience into AI systems. They are also a good starting point to explore more methods to inform AI systems of human experience in different contexts. The research process also provides compelling examples of how designers work with developers. Moreover, the research on the human experience of videos generated by GANs offers some direction for future work.

8.3 Contribution

The graduation project aims to find the methodologies for designers to translate qualitative human experience into quantitative AI systems. Following those steps, designers can steer AI systems' output according to human experience.

1. Provide methods to inform GANs human experience
Following the curation and algorithmic approaches, designers can successfully inform GANs about people's aesthetic evaluation of their images. For the curation approach, GANs obtain the human experience in the form of "good output" selected by humans as their inspiration, and the designer "translates" the human experience through the action of "using excellent output as input and retraining the GANs." For the algorithmic approach, the system obtains human knowledge in the form of "good factors," which can be predicted by aesthetic models in GANs. Moreover, the designer "translates" the human experience through the action of "finding the factors that influence people's rating and also can be measured by automated measures in algorithms."

2. Prove the importance of the human experience for AI systems
The evaluation of two approaches proves the significance of people's evaluation to steer the output of AI systems. After designers use curation and algorithmic approaches, the images generated by GANs get higher scores than the previous GANs. It shows that in the future, besides optimizing the algorithms, considering human feedback is also an efficient way to improve AI systems.

3. Build a general model for designers to inform AI systems about the human experience
GANs are an excellent beginning to explore how all AI systems can "understand" human experience. Based on the curation approach and algorithmic aesthetic approach, a general model to help designers translate the qualitative information from the human world into the quantitative parameter in the AI world is built. This model is a template for designers in all AI fields to inform their AI systems of human experience.

4. Give an example of cooperating with developers for AI designers
Due to the technical nature of AI products, AI product designers need to collaborate with developers frequently. The research process in this project provides an excellent example for AI designers to cooperate with developers. On the one hand, designers are supposed to learn some basic knowledge about the technique so that they can communicate with developers better. On the other hand, designers should know which step they should dive into. Combining the HCD(Human-center Design) and technical mindset, the AI product will be more beneficial for the human world.

8.4 Future work

This graduation project originates from LandShapes, an interactive exhibition piece to evoke people's awareness of climate change through the changing process of images generated by GANs. According to the advantages of GANs and the fake images, the scope of AI systems is narrowed down to the GANs; the aesthetic evaluation is chosen as the representation of the human experience. Considering the broad scope of AI systems and human experience, there is some exciting direction that people can dive into in the future.

1. Methodogies to iterate human experience of the fake video generated by GANs

The evaluation of LANDSHAPES showed the ability of GANs to produce media that
evoke an emotional response in the viewer. Visitors mention that unidirectional interaction from the images' transitions allows them to push for change in the shown landscape. Based on the research on videos, there is some direction for people to explore in the future:
The video's quality will be improved by selecting the high-rated images as the seeds. By adjusting the latent space, the video's quality will be iterated.

2. Approaches for other emotions in human experience
In the approaches, the scope of the human experience is narrowed down to "aesthetics ."However, in the natural human world, the human experience has broader factors like "ugly," "surprising," etc. When it comes to other dimensions, does replacing the input dataset with the images having aiming characteristics still work? For example, will the horror improve when designers retrain the GANs using the ugliest output selected?

3. A system that includes all the steps
This project provides two approaches for designers and the AI product teams to follow when they need to use the human experience to steer AI systems. In the future, systems like a website that can include all the steps will be more helpful for designers to complete the transition process. The system can contain all the steps and the software in the approaches. Instead of following the methods' instructions, designers can complete the whole process in a website or a system so that more designers can be accurate and professional translators between the human world and the AI world.

4.Improvement of aesthetics in different fields
Nowadays, there is some research provides a good basis for improving the aesthetic pleasure of topology optimized designs, either manually or ultimately by integrating them into the topology optimization formulation. Combing the exploration of the aesthetic pleasure in different fields, more research about the principles in different products can be concluded in the future.

REFERENCE

Frederik Ueberschär. Ai for experience: Designing with generative adversarial networks to evoke climate fascination. pages 1–156, 02 2021.

Chen, T., Guo, W., Gao, X., & Liang, Z. (2021). AI-based self-service technology in public service delivery: User experience and influencing factors. Government Information Quarterly, 38(4), 101520.

Brady, L., & Phillips, C. (2003). Aesthetics and usability: A look at color and balance. Usability News, 5(1), 2-5.

Perc M. 2020 Beauty in artistic expressions through the eyes of networks and physics. J. R. Soc. Interface 17: 20190686. http://dx.doi.org/10.1098/rsif.2019.0686

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

Luc, P., Couprie, C., Chintala, S., & Verbeek, J. (2016). Semantic segmentation using adversarial networks. arXiv preprint arXiv:1611.08408.

Harrison, P., Marjieh, R., Adolfi, F., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., ... & Jacoby, N. (2020). Gibbs sampling with people. Advances in Neural Information Processing Systems, 33.

Wang, C., Komodakis, N., & Paragios, N. (2013). Markov random field modeling, inference & learning in computer vision & image understanding: A survey. Computer Vision and Image Understanding, 117(11), 1610-1627

Harrison, P., Marjieh, R., Adolfi, F., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., ... & Jacoby, N. (2020). Gibbs sampling with people. Advances in Neural Information Processing Systems, 33, 10659-10671.

Ragot, M., Martin, N., & Cojean, S. (2020, April). Ai-generated vs. human artworks. a perception bias towards artificial intelligence?. In Extended abstracts of the 2020 CHI conference on human factors in computing systems (pp. 1-10).

Hekkert, P., & Van Wieringen, P. C. (1998). Assessment of aesthetic quality of artworks by expert observers: An empirical investigation of group decisions. Poetics, 25(5), 281-292.

Acebo, E., & Sbert, M. (2005, May). Benford's law for natural and synthetic images. In Proceedings of the First Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging (pp. 169-176).

Greenfield, G. (2005, May). Designing metrics for the purpose of aesthetically evaluating images. In Proceedings of the First Eurographics conference on Computational Aesthetics in Graphics, Visualization and Imaging (pp. 151-158)

Neumann, L., Sbert, M., Gooch, B., & Purgathofer, W. (2005). Defining computational aesthetics. Computational aesthetics in graphics, visualization and imaging, 13-18.

Ferschin, P. (2004). Perception based Illustration Methods.

Wang, Q. (2016). Emotional Architecture for Everyday Life. Architectural Design for Senior Living Oriented by the Psychological Pattern of Elderly People.

Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. Psychology science, 48(2), 157.

Hekkert, P., Snelders, D., & Van Wieringen, P. C. (2003). 'Most advanced, yet acceptable': Typicality and novelty as joint predictors of aesthetic preference in industrial design. British journal of Psychology, 94(1), 111-124.

Schifferstein, H. N., & Hekkert, P. (Eds.). (2011). Product experience. Elsevier.

Wen, S., Liu, W., Yang, Y., Huang, T., & Zeng, Z. (2018). Generating realistic videos from keyframes with concatenated GANs. IEEE Transactions on Circuits and Systems for Video Technology, 29(8), 2337-2348.

Loos, S., Wolk, S. V. D., Graaf, N. D., Hekkert, P., & Wu, J. (2022). Towards intentional aesthetics within topology optimization by applying the principle of unity-in-variety. Structural and Multidisciplinary Optimization, 65(7), 1-11.

Wang, J., Neuhauser, C., Wu, J., Gao, X., & Westermann, R. (2021). The 3D Trajectory-based Stress Visualizer. arXiv preprint arXiv:2112.09202.

Agrawal, A., Gans, J. S., & Goldfarb, A. (2018, May). Human judgment and ai pricing. In AEA Papers and Proceedings (Vol. 108, pp. 58-63).

Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. Psychology science, 48(2), 157.

Forsey, J. (2016). The aesthetics of design. Oxford University Press.

Findeli, A. (1994). Ethics, aesthetics, and design. Design issues, 10(2), 49-68.

Stolterman, E. (1994). Guidelines or aesthetics: design learning strategies. Design Studies, 15(4), 448-458.

Oh, S., Jung, Y., Lee, I., & Kang, N. (2018, August). Design automation by integrating generative adversarial networks and topology optimization. In International design engineering technical conferences and computers and information in engineering conference (Vol. 51753, p. V02AT03A008). American Society of Mechanical Engineers.

Burnap, A., Hauser, J. R., & Timoshenko, A. (2019). Design and evaluation of product aesthetics: A human-machine hybrid approach. arXiv preprint arXiv:1907.07786.

Marchesi, M. (2017). Megapixel size image creation using generative adversarial networks. arXiv preprint arXiv:1706.00082.

Shen, Y., Gu, J., Tang, X., & Zhou, B. (2020). Interpreting the latent space of gans for semantic face editing. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 9243-9252).

Rebecca K. Priestley, Zo¨e Heine, and Taciano L. Milfont. Public understanding of climate change-related sea-level rise. PLOS ONE, 16(7):1–12, 07 2021

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C.J. Burges, L. Bottou, and K.Q. Weinberger, editors, Advances in Neural Information Processing Systems, volume 25. Curran Associates, Inc., 2012.

Martin Ragot, Nicolas Martin, and Salom´e Cojean. Aigenerated vs. human artworks. a perception bias towards artificial intelligence? In Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, CHI EA ’20, page 1–10, New York, NY, USA, 2020. Association for Computing Machinery.

Xin Fu, Jia Yan, and Cien Fan. Image aesthetics assessment using composite features from off-the-shelf deep models. In 2018 25th IEEE International Conference on Image Processing (ICIP), pages 3528–3532, 2018.

Nishi Doshi, Gitam Shikhenawis, and Suman K Mitra. Image aesthetics assessment using multi channel convolutional neural networks, 2019.

Yujun Shen and Bolei Zhou. Closed-form factorization of latent semantics in gans, 2020.

Sidharth Mishra, Uttam Sarkar, Subhash Taraphder, Sanjoy Datta, Devi Swain, Reshma Saikhom, Sasmita Panda, and Menalsh Laishram. Principal component analysis. International Journal of Livestock Research, page 1, 01 2017.

J. L. Fleiss, “Measuring nominal scale agreement among many raters.” Psychological Bulletin, vol. 76, no. 5, p. 378–382, 1971.

F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, “Lsun: Construction of a large-scale image dataset using deep learning with humans in the loop,” 2015. [Online]. Available: https: //arxiv.org/abs/1506.03365

Yubin Deng, Chen Change Loy, and Xiaoou Tang. “Image aesthetic assessment: An experimental survey”. In: IEEE Signal Processing Magazine 34.4 (2017), pp. 80–106.

Elz˙e Sigut˙e Mikalonyt˙e and Markus Kneer. "Can Artificial Intelligence Make Art? Folk Intuitions as to whether AI-driven Robots Can Be Viewed as Artists and Produce Art". In: ACM Transactions on Human-Robot Interaction (2022).

Elz˙e Sigut˙e Mikalonyt˙e and Markus Kneer. "Can Artificial Intelligence Make Art? Folk Intuitions as to whether AI-driven Robots Can Be Viewed as Artists and Produce Art". In: ACM Transactions on Human-Robot Interaction (2022).

F. Snyders. Example of golden triangle method on a painting. compositional elements fall within the triangles, 2018. [Online; accessed May 27, 2022].

Ruben Rodriguez Torrado, Ahmed Khalifa, Michael Cerny Green, Niels Justesen, SebastianRisi, and Julian Togelius. Bootstrapping conditional gans for video game level generation. 2020 IEEE Conference on Games (CoG), pages 41–48, 2020.

Fei Yang, Jianfeng Ren, Zheng Lu, Jialu Zhang, and Qian Zhang. Rain-component-aware capsule-gan for single image de-raining. Pattern Recognition, 123, 2022.

H. Symons. How to create the rule of thirds to improve your art, 2022. [Online; accessed May 30, 2022].

I. Roseman and A. Evdokas. Appraisals cause experienced emotions: Experimental evidence. Review of General Psychology, 18(1):1–28, 2004.

P. J. Silvia. Emotional responses to art: From collation and arousal to cognition and emotion. Review of General Psychology, 9(4):342–357, 2005.