# Unusual Transformation: A Deep Learning Approach to Create Art

Mai Cong Hung[1](✉), Mai Xuan Trang[2], Ryohei Nakatsu[3], and Naoko Tosa[3]

[1] Osaka University, Osaka, Japan
[2] Phenikaa University, Hanoi, Vietnam
trang.maixuan@phenikaa-uni.edu.vn
[3] Kyoto University, Kyoto, Japan
ryohei.nakatsu@design.kyoto-u.ac.jp, tosa.naoko.5c@kyoto-u.ac.jp

**Abstract.** In this research, we proposed the concept of "unusual transformation," which realizes transformation between two very different image sets as an extension of CycleGAN. CycleGAN is a new deep-learning-based AI technology that can realize transformation between two image sets. Although conventional CycleGAN researchers have tried transformation between two similar image sets, we applied CycleGAN to the transformation of two very different image sets such as between portraits photos and Ikebana or Shan-Shui paintings. Then to obtain a better result, we improved CycleGAN by adding a new loss function and developed "UTGAN (Unusual Transformation GAN)." We found that by using UTGAN, portrait photos and animal photos are transformed into Ikabana-like and Shan-Shui-like images. Then we carried out an analysis of the obtained result and made a hypothesis that the unusual transformation works well because both Ikebana and Shan-Shui are fundamental and abstracted expressions of nature. Also, we carried out various considerations to justify the hypothesis.

**Keywords:** GANs · CycleGAN · Image transformation · Ikebana · Shan-Shui · UTGAN

## 1 Introduction

In recent years, the rapid development of AI and Deep Learning raises the questions about the impact of these advance technologies in the way we create and study art. In the analyzing topic, machine learning techniques were used in the artwork clustering task [1] and the art evaluation task [2]. However, for the fundamental question such as whether AI can create artworks or not, an answer has not yet been obtained.

Style transfer is widely considered as one basic approach of AI toward such a direction. One might use generative models in Deep Learning to transform normal photos or sketches into images that have similar visual effects to artworks with a specific style.

M. C. Hung and M. X. Trang—Contributed equally to this work as first authors.


M. Wölfel et al. (Eds.): ArtsIT 2021, LNICST 422, pp. 309–320, 2022.
https://doi.org/10.1007/978-3-030-95531-1_21

Recently, the appearance of GANs (Generative Adversarial Networks [3]) has made a breakthrough regarding the topic of style transfer. In the training of GANs, a generator network G learns to generate new data while a discriminator network D tries to identify the generated data whether it is real or fake. In game theory terms, this training process can be interpreted as a minimax game. With this interesting mechanism, the training proces of GANs networks can converge provided a relatively small number of training data.

Based on the minimax game of generator and discriminator in the basic configuration of GANs, a large number of variations has been developed by modifying the network structure and the objective loss function. CycleGAN [4] is an elegant variation of GANs which study the mutual transformation between two sets of photos. CycleGAN is effective for art style transfer because of the unpaired training mechanism. It realizes set-to-set level transformation to learn the distribution of the target sets, or art styles.

Classic examples of CycleGAN and other style transfer techniques were developed by achieving the transformation between two sets of data of relatively similar size, themes, or categories. On the other hand, in this paper, we propose the idea of "Unusual Transformation," which achieves a mutual transformation between two image sets with different sizes and themes. In our previous research [5], we gave several examples of portraits and animal photos transformed into Ikebana (Japanese flower arrangement) via CycleGAN. At the same time, however, as there were problems of under transformation and over transformation, we found it necessary to improve CycleGAN [6].

By combining these previous research results, in this paper, we propose "Unusual Transformation" by explaining its concept and also by giving various examples. We also discuss the underlying connection of this concept to other art-related topics.

## 2 Generative Models in Deep Learning

In the last decade, GANs (Generative Adversarial Networks) [3] have become one of the most essential topics in Deep Learning. The generative model in GANs provides impressive performance on art style transfer even with a small number of training data. The architecture of GANs could be described as in Fig. 1 with the basic configuration of two networks, a generator network (G) and a discriminator network (D). The training of GANs is based on a minimax mechanism in the sense that the generator G learns to generate fake data from random noise while the discriminator D tries to classify the generated data into categories of "real" or "fake." In other words, the training on G tries to maximize the probability of the generated data to lie on the targeted distribution and the training process of D tries to minimize it.

Among the variations of GANs, CycleGAN [4] is an effective approach to set-to-set level learning to study the mutual transformation between two sets of photos. The architecture of CycleGAN consists of two generators and two discriminators as can be seen in Fig. 2.

To perform the mutual transformation of two image sets A and B, the training of CycleGAN learns two mappings $G_{AB} : A \rightarrow B$ and $G_{BA} : B \rightarrow A$ given the training samples: $\{a_i\}_{i=1}^{N} \in A$ and $\left\{b_j\right\}_{j=1}^{M} \in B$ with the data distributions $a \sim p_A(a)$ and

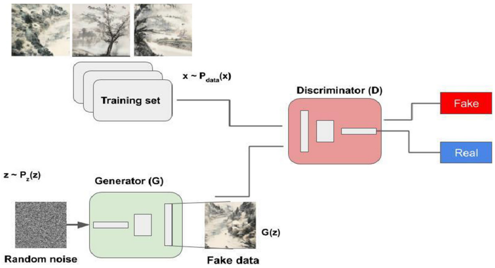


**Fig. 1.** The basic configuration of GANs

$b \sim p_B(b)$. The respected discriminators are $D_A$ and $D_B$ aim to distinguish between real photos and generated fake photos.

To emphasize the mutual transformation, the objective loss function of CycleGAN includes two components: adversarial losses for matching the generated images to the target set, and cycle consistency loss for preventing the mappings $G_{AB}$ and $G_{BA}$ from contradicting each other.

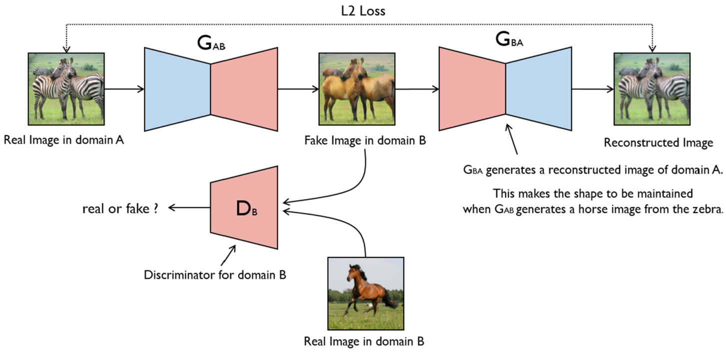


**Fig. 2.** Basic configuration of CycleGAN ([4])

**Adversarial Loss:** The adversarial loss applies to both generators.

- For $G_{AB} : A \rightarrow B$ and the respected discriminator $D_B$:

$$\begin{aligned} &\mathcal{L}_{GAN}(G_{AB}, D_B, A, B) \\ &\quad = \mathbb{E}_{b\sim p_B(b)}\big[\log D_B(b)\big] \\ &\quad + \mathbb{E}_{a\sim p_A(a)}[\log(1 - D_B(G_{AB}(a)))] \end{aligned} \tag{1}$$

- For $G_{BA} : B \rightarrow A$ and the respected discriminator $D_A$:

$$\begin{aligned} &\mathcal{L}_{GAN}(G_{BA}, D_A, B, A) \\ &\quad = \mathbb{E}_{a\sim p_A(a)}\big[\log D_A(a)\big] \\ &\quad + \mathbb{E}_{b\sim p_B(b)}[\log(1 - D_A(G_{BA}(b)))] \end{aligned} \tag{2}$$

**Cycle Consistency Loss:** For each image $a$ from domain $A$, the generated image after applying two transformations $G_{AB}$ and $G_{BA}$ should be similar to $a$: $a \rightarrow G_{AB}(a) \rightarrow G_{BA}(G_{AB}(a)) \approx a$. We call it forward cycle consistency. We also have backward cycle consistency in the reverse direction: $b \rightarrow G_{BA}(b) \rightarrow G_{AB}(G_{BA}(b)) \approx b$. The cycle consistency loss is a combination of both forward and backward cycle consistency losses:

$$\begin{aligned} \mathcal{L}_{cyc}(G_{AB}, G_{BA}) = \mathbb{E}_{a\sim p_A(a)}\left[\|\mathrm{G_{BA}}(G_{AB}(a)) - a\|_1\right] \\ + \mathbb{E}_{b\sim p_B(b)}\left[\|\mathrm{G_{AB}}(G_{BA}(b)) - b\|_1\right] \end{aligned} \tag{3}$$

The total objective loss function of CycleGAN consists of the adversarial losses and the cycle consistency loss:

$$\begin{aligned} \mathcal{L}(G_{AB}, G_{BA}, D_A, D_B) = \mathcal{L}_{GAN}(G_{AB}, D_B, A, B) \\ + \mathcal{L}_{GAN}(G_{BA}, D_A, B, A) \\ + \lambda\mathcal{L}_{cyc}(G_{AB}, G_{BA}) \end{aligned} \tag{4}$$

where constant $\lambda$ is the weight of the cycle consistency loss.

We note that the generative models in CycleGAN learn the set-to-set level of transformation while the original GANs learn to generate data to fit in a target set. In the task of art style transfer, CycleGAN could learn the mutual conversion between normal photos and art styles as well as more general transformation between two sets of data.

# 3 Unusual Transformation

## 3.1 Concept of Unusual Transformation

In classic examples of CycleGAN in [4], the generative models were used to make a mutual transformation between horse and zebra images, landscape photos and Monet paintings, etc. This means that the transformation was made between images of relatively similar size, theme, and category. Because of such similarities between two image sets, obtained results are interesting but not impressive enough. For example, in the case of the transformation from a landscape to a Monet-like image, the obtained image only looks like a Monet-like image and not more than that. This means that, at this stage, AI does not have the capability of art creation.

Here, we should understand that creation can be achieved based on the connection of different things. As has always been indicated, ideas and inventions often come from the connection of two different things [7].

A good example is Surrealism. In artworks of Surrealism such as Dali's artworks, we find that things that never co-exist in our real-world appear together such as the co-existent of day and night scenes, co-existent of a real-world, and a dream world, etc. These artworks inspire our imagination and therefore have been highly evaluated. If two different things could be connected by AI, it may be possible that AI can create art. Although CycleGAN has the capability of connecting two different things, so far what it can achieve is the transformation between two similar image sets.

Based on this, we propose the idea of "Unusual Transformation," a high-abstracted transformation where transformation is achieved between relatively different domains of objects such as macro and micro-size worlds of plants and animals. This concept of unusual transformation is believed to be a key idea to create new art. For example, portraits or animal photos could be unusually transformed into Ikebana-like or Shan-Shui-like artistic images.

Although the definition of unusual transformation is very simple, it contains several fundamentally difficult problems. One is that the unusual transformation is a naturally difficult task with a low rate of successful transfer. To increase the success rate of conversion, we think the original CycleGAN is not enough and we have to develop new style transfer technology by improving CycleGAN.

Another is that if we want to achieve a good transformation, the target data set B should have a conceptual meaning. If we use concrete images such as Monet's drawings as the target image set B, the obtained image only looks like a Monet-like image and not more than that. What kind of image set should be used as B is a crucial issue. Ikebana would be a good example as a target data set because of its minimality and flexibility. Shan-Shui paintings would be another good painting tool as well because in Shan-Shui natural elements such as rocks, streams, mountains would be put in flexible positions. This will be further discussed in Chapter 4.

### 3.2 Ikebana and Shan-Shui as Target Data Sets

**(1) Ikebana**

Ikebana (Japanese flower arrangement) is one of the most important art forms in Japanese culture. Ikebana is the art of flower arrangement where the flowers are given life under the conceptual arrangements of Ikebana artists [8].

Ikebana has a deep root in the Japanese philosophy of art under the strong influence of Zen Buddhism. The tradition of arranging flowers on Buddha was brought to Japan from China in the Heian period (794–1185) by Zen Buddhist monks. Then Ikebana grew to be an important art form along with the development of Zen.

Ikebana has a long history of development and has continued to be a great source of inspiration in modern art. For instance, Naoko Tosa, created a video artwork named "Sound of Ikebana" by applying fluid dynamics to the art creation process [9, 10].

The "minimality" and the "flexibility" are the two important properties of Ikebana which support our idea of the unusual transformation. Under the influence of Zen philosophy, "emptiness" plays an essential role in the art of Ikebana. The emptiness appearing in Ikebana artworks is believed to provide meaning and be harmonic to the whole scene. We consider the minimality of Ikebana as the appearance of the emptiness in an Ikebana artwork. At the same time, we call Ikebana flexible as the materials (flowers, leaves, branches, and so on) can be placed in various shapes and arrangements.

**(2) Shan-Shui Painting**

Shan-Shui refers to a style of traditional Chinese painting that involves or depicts scenery or natural landscapes, using a brush and ink rather than more conventional paints [11]. Mountains, rivers, and waterfalls are common subjects of Shan-Shui paintings.

Shan-Shui painting first began to develop in the 5th century in China, in the Liu Song dynasty. It was later characterized by a group of landscape painters such as Zhang

Zeduan, most of them already famous, who produced large-scale landscape paintings. These landscape paintings usually centered on mountains. Mountains had long been seen as sacred places in China, which were viewed as the homes of immortals and thus, close to the heavens. Philosophical interest in nature, or mystical connotations of naturalism, could also have contributed to the rise of landscape painting. The art of Shan-Shui, like many other styles of Chinese painting, has a strong reference to Taoism/Daoism imagery and motifs, as symbolisms of Taoism strongly influenced "Chinese landscape painting". Some authors have suggested that Daoist stress how minor the human presence is in the vastness of the cosmos, or Neo-Confucian interest in the patterns or principles that underlie all phenomena, natural and social lead to the highly structuralized nature of Shan-Shui.

Shan-Shui painting was first introduced to Japan from China along with Zen as an ink painting during the Kamakura period (1185–1333). At first, many paintings expressed Zen thought, but gradually the form of ink painting changed and Shan-Shui paintings began to be drawn. In the latter half of the 15th century, the famous Shan-Shui painter Sesshu (1420–1506) appeared and completed the Japanese Shan-Shui painting.

## 4 Preliminary Experiment and Development of UTGAN

### 4.1 Preliminary Experiment: Transformation to Ikebana

In this section, we introduce some experimental examples of unusual transformation via CycleGAN. We performed the unusual transformation via CycleGAN with the sets A1, A2, and the target set B as follows:

- Dataset A1: Portrait photos in Flickr
- Dataset A2: Kaggle Animal-10 dataset. (https://www.kaggle.com/alessiocorrado99/animals10)
- Dataset B: Ikebana photos in Google Image Search

Figure 3 shows the result of the transformation by CycleGAN. As can be seen from the results, portraits and horse photos turned into Ikebana images while keeping the original shape. This "unusual transformation" concept would inspire a new method to create art via Deep Learning. However, there are some limitations. In some cases of photos with complex background, the experiments failed to transform them into abstract Ikebana. Some photos were over-transformed so that we could not recognize the original shape. We consider the reason is that the structure of the CycleGAN was not designed to learn specific high abstract representation such as the unusual transformation in this experiment.

### 4.2 Improvement of CycleGAN

In the classic examples, CycleGAN worked well on mutual transform two sets of similar photos in terms of size, themes, or categories. However, as shown in Fig. 3, our experiments showed the over-transformation limitation of CycleGAN in unusual transformation with Ikebana. To improve the performance, we introduce the combination of

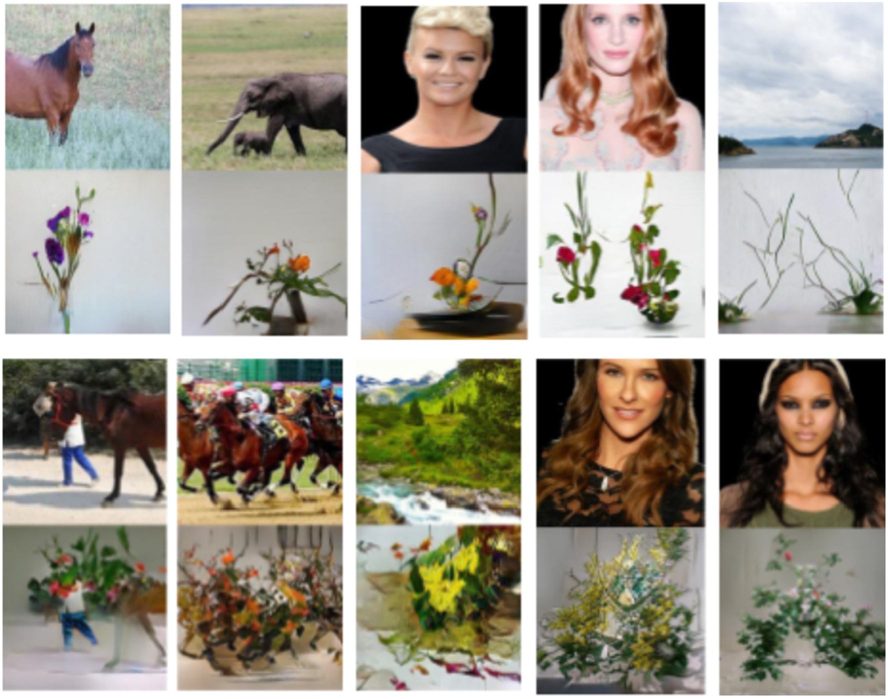

**Fig. 3.** Results of transformation into Ikebana including several failed transformations.

CycleGAN with several computer vision techniques in an improved modification called "UTGAN." To preserve the original shape of the main objects to a certain level, we apply object recognition techniques to detect the main object, and then an edge detection is used to emphasize the main shape. In the last step, we include an object edge-promoting loss to enforce the model to emphasize the original shapes of the main objects under the transformation. The calculation of the object edge-promoting loss is described in the next paragraph.

From the training photos set $A$, we generate a set of edge photos $E = \{e_i\}_{i=1}^{N}$ by removing clear edges of the main object in $\{a_i\}_{i=1}^{N}$. For each photo $a_i \in A$, we operate the following steps:

(1) To recognize objects in the photo by using a pre-trained object detector (Mobile_Net_SSD),
(2) To detect edge pixels of objects by the Canny edge detector [12],
(3) To dilate the edge regions, and
(4) To apply Gaussian smoothing in the dilated edge regions.

In our proposed UTGAN, the training target of the discriminator $D_A$ is to maximize the probability of assigning the correct label to $G_{BA}(b)$, the real photos without clear edges of the photos' main objects (i.e., $e_j \in E$) and the real photos (i.e., $a_i \in A$).

Finally, we add object edge-promoting loss to the adversarial loss:

$$\mathcal{L}_{GAN}(G_{BA}, D_A, B, A) = \mathbb{E}_{a\sim p_A(a)}\left[\log D_A(a)\right] + \mathbb{E}_{b\sim p_B(b)}[\log(1 - D_A(G_{BA}(b)))] + \gamma\mathbb{E}_{e\sim p_E(e)}[\log(1 - D_A(G_{BA}(e)))] \tag{5}$$

where constant $\gamma$ controls the weight of the object edge-promoting loss.

## 5 Examples of Unusual Transformation

### 5.1 Transformation to Ikebana

We performed the unusual transformation via UTGAN with the style sets A1, A2 and the set B as follows:

- Dataset A1: Portrait photos in Flickr (https://github.com/NVlabs/ffhq-dataset)
- Dataset A2: Animal photos from Google Image Search
- Dataset B: Ikebana photos in Google Image Search

Figure 4 shows several results of A1 to B transformation and Fig. 5 shows several results of A2 to B transformation.

**Fig. 4.** Experiment result A1-B: the first row is the original photos, the second row is the result by CycleGAN, the last row is the results by UTGAN

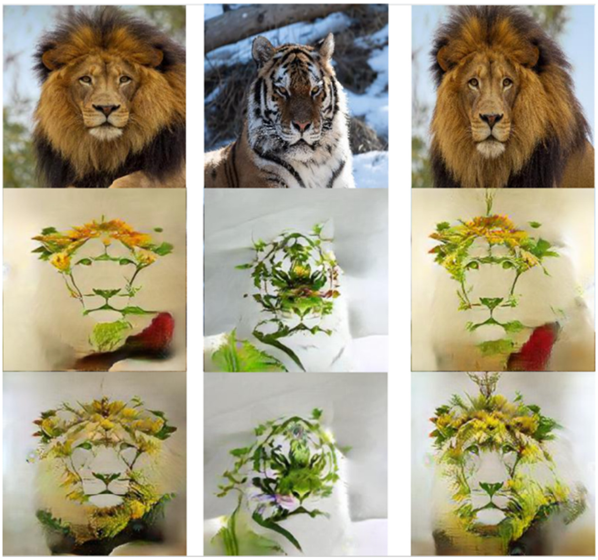

**Fig. 5.** Experiment result A2-B: the first row is the original photo, the second row is the results by CycleGAN, the last row is the results by UTGAN

### 5.2 Transformation to Shan-Shui Paintings

We applied the unusual transformation to transfer photos of human faces in Flickr to Chinese Shan-Shui paintings via UTGAN. Below is the data set we used.

- Dataset A: Portrait photos in Flickr
- Dataset B: Shan-Shui artworks in Google Image Search

Some of the obtained results are shown in Fig. 6. Portraits turned into Shan-Shui-like images while one can still recognize the original shape of human faces.

**Fig. 6.** Experimental results of portrait to Shan-Shui transformation.

## 6 Discussion

In this chapter, we will propose hypotheses regarding art by considering the functions of CycleGAN and its improved UTGAN. We will also discuss the possibility of clarifying the essence of art by using UTGAN.

In this paper, beyond the scope of transformations so far achieved by CycleGAN, we have attempted unusual transformations by carrying out transformations between image sets that seem to have no similarity at all. We tried to convert between image sets of animals and portraits and image sets of Ikebana photos and Shan-Shui paintings, which are completely different in appearance. As a result, the portraits or animal photos were converted into Ikebana-like images and Shan-Shui-like images while retaining the characteristics of the original image. Rather, obtained images may be unprecedented Ikebana images or unprecedented Shan-Shui paintings. In other words, our unusual transformation has produced paintings that have never been seen before. What does this mean? We think that the following hypotheses can be made.

- Hypothesis 1: Portraits and animal photos are successfully converted into Ikebana and Shan-Shui images because both portraits and animals are natural objects.
- Hypothesis 2: The conversion into Ikebana and Shan-Shui was successful because Ikebana and Shan-Shui paintings contain the essentials of natural objects.

There is a famous Aristotle words that "art imitates nature [13]". As expressed in these words, art represented by paintings used to express nature. The so-called realism paintings are typical examples. For the impressionism that was born after realism represented by the artworks of Monet, Cezanne, etc., artworks created by the Impressionism artists are abstract in the sense that they did not draw nature as it is but drew the impressions of the artists. However, although they have drawn the impression they received, what they tried to draw is clearly understood and not very abstract. After that, however, paintings with a higher degree of abstraction such as Cubism and Surrealism appeared, and it continued to the present extremely high degree of abstraction. This is the brief history of Western art.

Based on CycleGAN's idea to carry out transformation between two image sets, what the Western paintings have tried to express can be shown in Fig. 7. In other words, there is a conversion from the actual landscape to the landscape paintings. (The process of converting a landscape painting into a landscape photograph doesn't make much sense for our discussion, so it's enough to consider only the transformation function G here.)

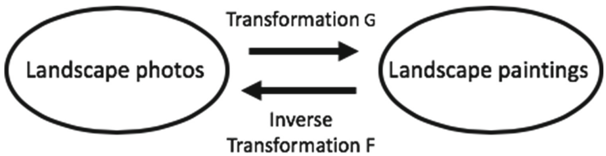


**Fig. 7.** Relationship between landscape photos and landscape paintings.

To make it even more abstract, Fig. 7 can be expressed as Fig. 8. In other words, art extracts the essential things from natural objects and phenomena.

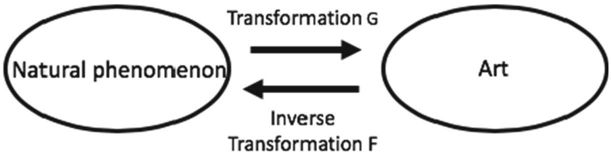


**Fig. 8.** Relationship between natural phenomenon and art.

If we think in this way, we may find that hypotheses 1 and 2 mentioned above are correct. Furthermore, when examining the characteristics of Ikebana and Shan-Shui paintings, they have the following characteristics and are appropriate for expressing the essence of natural objects and phenomena.

**(1) Minimality**

Ikebana and Shan-Shui paintings try to remove unnecessary things from natural objects and phenomena and express them with the minimum expression. For example, Ikebana tries to express the scenery of nature with a very small number of flowers and vegetation. Also, it can be said that Shan-Shui paintings express nature by decomposing what constitutes nature into the minimum basic elements (mountains, rocks, water streams, etc.) and by reconstructing them.

**(2) Flexibility**

As mentioned above, both Ikebana and Shan-Shui paintings are trying to reconstruct nature by breaking down what construct nature into the minimum elements and reconstructing them. In addition, when reconstructing nature, the individual elements have flexibility in their placement. For example, in the case of Ikebana, the arrangement of a small number of flowers and vegetation greatly differs depending on each artist. In other words, the degree of freedom of arrangement itself may lead to the diversity of Ikebana. Also, in the case of Shan-Shui paintings, the individual components such as rocks and water streams can be freely placed in the painting.

## 7 Conclusion

CycleGAN, which is one of the variations of GANs, enables mutual conversion between datasets without the need for a one-to-one correspondence of data. For example, it is possible to convert landscape photographs into Monet-like images. However, this means that AI merely produces a Monet-like image. At this stage, AI is not yet capable of creating art. The main reason for this is that the style transfer in previous studies only involves conversions between similar datasets, such as between horses and zebras, between landscape photographs and Monet's landscape paintings, etc.

Art and inventions have been a creation based on the connection of different things. Based on this basic principle, this paper proposes the transformation between different types of datasets called "Unusual Transformation." Then, as an example, we tried to convert portraits and animal photographs into Ikebana using CycleGAN. However, it has been shown that under transformation and over transformation often occur. To solve this problem, we proposed UTGAN, in which a new element is added to the loss function, to give CycleGAN a new function to keep the original structure of portraits or animal photos. It was shown that by applying UTGAN, portraits and animal photos can be successfully converted into Ikebana and Shan-Shui.

Based on these results, we considered why portraits and animal photos can be converted into Ikebana and Shan-Shui images. As a result, it became clear that these even seemingly different types of image sets are connected at the root. In other words, since human faces and animals are natural objects, and Ikebana and Shan-Shui paintings are the essences of nature, conversion is successful when there is such a relationship between two image sets. Extending this further, we may be able to approach art more deeply from the science and technology aspect.

## References

1. Gultepe, E., Conturo, T.E., Makrehchi, M.: Predicting and grouping digitized paintings by style using unsupervised feature learning. J. Cult. Herit. **31**, 13–23 (2018)
2. Mai, C.H., Nakatsu, R., Tosa, N., Kusumi, T., Koyamada, K.: Learning of art style using AI and its evaluation based on psychological experiments. In: Nunes, N.J., Ma, L., Wang, M., Correia, N., Pan, Z. (eds.) ICEC 2020. LNCS, vol. 12523, pp. 308–316. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65736-9_28
3. Creswell, A., et al.: Generative adversarial networks: an overview. IEEE Sig. Process. Mag. **35**(1), 53–65 (2018)
4. Zhu, J., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2242–2251 (2017)
5. Mai, C.H., Nakatsu, R., Tosa, N.: Developing Japanese Ikebana as a digital painting tool via AI. In: Nunes, N.J., Ma, L., Wang, M., Correia, N., Pan, Z. (eds.) ICEC 2020. LNCS, vol. 12523, pp. 297–307. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-65736-9_27
6. Hung, M.C., Trang, M.X., Tosa, N., Nakatsu, R.: IkebanaGAN: new GANs technique for digital Ikebana art. In: Rauterberg, M. (ed.) HCII 2021. LNCS, vol. 12794, pp. 88–99. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-77411-0_7
7. Jewkes, J., Sawers, D., Stillerman, R.: The Sources of Invention. W. W. Norton & Company (1971)
8. Luu, A., Matsuba, I.: Ikebana Unbound: A Modern Approach to the Ancient Japanese Art of Flower Arrangement, Artisan (2020)
9. Tosa, N., Nakatsu, R., Yunian, P.: Creation of media art utilizing fluid dynamics. In: 2017 International Conference on Culture and Computing, pp.129–135, 10–12 September 2017
10. Pang, Y., Zhao, L., Nakatsu, R., Tosa, N.: A study of variable control of Sound Vibration Form (SVF) for media art creation. In: 2017 International Conference on Culture and Computing, pp.136–142, 10–12 September 2017
11. Law, S.S.-M.: Being in traditional Chinese landscape painting. J. Intercult. Stud. **32**(4), 369–382 (2011)
12. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-8**(6), 679–698 (1986)
13. Aristotle: The Art of Rhetoric. Oxford University Press (2018)