DOCTORAL THESIS

Study on Systematic Generation of Pareidolia-Inducing Stimuli using Deep Learning

Graduate School of Science and Engineering,
Iwate University
Doctor's Course,
Design & Media Technology
Yoshitaka Endo

March, 2024

Acknowledgment

In regard to my research project, I would like to deeply appreciate Associate Professor Takuya Akashi, Graduate School of Engineering Department, of Design and Media technology, Iwate University. He has supported me for 6 years and made me grow. In particular, the 3 years of the doctoral course are full of ups and downs, however, the experiences are irreplaceable and invaluable. In addition, I participate in some collaborative research of companies, and academic institutions. He gave me many opportunities to grow myself. I believe that the experience is useful in the future.

I also would like to appreciate Gertrude Baltimore Professor of Experimental Psychology Shinsuke Shimojo, Division of Biology and Biological Engineering, California Institute of Technology. I appreciate not only research advice but also an academic internship, meetings, and interesting research topics. Thank you for the international and interdisciplinary collaborative research.

Also, I would like to appreciate Senior Assistant Professor Chao Zhang, Division of Engineering Information Science, University of Fukui Faculty of Engineering, and Assistant Professor Junya Sato, Intelligent Mechanical Engineering Course,

Department of Mechanical Engineering, Faculty of Engineering, Gifu University. They are graduated from smart computer vision lab. Thank you for advice on my research and my future course.

Also, I would like to appreciate sub-chief examiners, Professor Kouichi Konno, Professor Mitsugu Saito, and in Graduate School of Engineering, Department of Design and Media Technology, Iwate University. They gave valuable advice and suggestions.

Also, I wish to thank my members in smart computer vision laboratory. The members cooperated in experiments of my doctoral thesis. In the laboratory, I enjoyed the time not only for research but also for communication, and party.

Also, I appreciate Ms. Rinka Asanuma, a master course student at Iwate University. She helped with my research data analysis and experiment.

I also would like to appreciate, my classmates from Sendai National College of Technology who graduated in 2017. They cheered my doctor course on a telephone call.

I cannot forget to thank my father and mother, sister, and grandparents. They supported me in my uncertain future, and worried about me.

Also, our lifestyle has changed because of COVID-19, and this affected to the conference style, attendance at the university, class, and research project. However, through the doctor course, I deeply realized that many people supporting and

cheered me from behind the scenes. Without this support, I might not complete doctoral course. I would like to express my gratitude to the people who supported me again.

Yoshitaka Endo

March, 2024

Abstract

Pareidolia is one of the psychological tendencies that we might experience in our daily lives. Pareidolia is the perception of a specific object from other objects or patterns. This tendency is useful for diagnosis of the patients suffering from Lewy body dementia. This diagnosis method is referred to as the “Pareidolia test”. However, the visual stimuli cannot be distributed because of their copyright. Because collecting the pareidolia stimuli is only from the Internet or generation artificially, The pareidolia stimuli generation method is required. Also, these pareidolia stimuli are difficult to measure the pareidolia-inducing power. This is because the pareidolia stimuli are difficult to compare to the other stimuli.

I aim to generate from the real face structure and natural scene images. Face pareidolia is induced by face elements such as eyes and mouth. Therefore, I use such features using preprocessing, and the features are extracted.

The face data set extracted the pareidolia elements and natural scene image are used for training of cycle-consistent adversarial networks (CycleGAN). CycleGAN can translate the image styles between two data sets bidirectionally. Face pareidolia can be considered that consist of face pareidolia elements. In this thesis, I aim to

generate pareidolia stimuli face structure and natural scene style.

In addition, one of the CycleGAN characteristics is the cycle-consistency loss. The cycle-consistency loss learns the generators of CycleGAN to be consistent with both translations. The loss function can manipulate the effects of the generation results. In this thesis, the hypothesis is that the cycle-consistency loss can manipulate the pareidolia-inducing power. Also, the cycle-consistency loss in the objective function is added to its weighted loss value. Therefore, this thesis manipulates the weight value and aims to generate the pareidolia stimuli.

This thesis consists of two-phase experiments. The first one is the face pareidolia stimuli generation experiment. In this thesis, the weight value is set to 2, 10, and 20. As a result, I confirmed the pareidolia elements are retained. The second one is the psychological experiment. This thesis enrolled some participants. The enrolled participants evaluate the generated stimuli. First, this thesis investigates whether the generated stimuli can induce pareidolia. Second, this thesis investigates the pareidolia-inducing power of the generated stimuli. The results of the evaluation experiment revealed a correlation between cycle-consistency loss and pareidolia-inducing power when the blurring process was applied excluding the pareidolia elements (the eyes and mouth) as preprocessing. In addition, this thesis confirmed that there is a significant difference (significant standard = 10%) between the weight value is set to 10 and 20 when the blurring process was applied. On the other

hand, there is no correlation between cycle-consistency loss and pareidolia-inducing power when the blurring process was applied excluding the pareidolia elements (the eyes and mouth) as preprocessing.

Also, this thesis investigates what is the cause of the pareidolia in the experiment. The detected pareidolia in the experiment can be considered for two reasons. The first one is a cause of abnormal internal criteria. The second one is this task of the experiment is too difficult to discriminate the pareidolia stimuli and others. To reveal the cause, the result data is analyzed using signal detection theory (SDT). As a result of SDT analysis, the cause of face pareidolia might be abnormal internal criteria.

The experiments suggest the face attributes which relate to face pareidolia can induce the face pareidolia although the eyes and mouth are only extracted using preprocessing. The cycle-consistency loss of CycleGAN can manipulate the pareidolia-inducing power when the blurring preprocessing is applied.

As the future work, there are many tasks to tackle. First, whether the tendency of the generated stimuli in this thesis is different from the other face pareidolia type. Most of the pareidolia stimuli are natural scene objects. In this thesis, the generated stimuli are artificial. Therefore, I need to investigate the tendencies such as perception, pareidolia-inducing power, and so on. In addition, Regarding the statistical significance test, the significant standard is set to 10%. In general, the significant

standard of statistical significance test is set to 5%. Therefore, I aim to satisfy the standard from two approaches. The first one is the increase in the number of participants. One of the participants might affect the statistical testing because of few participants. In association with the previous sentence, the second one is the investigation of the pareidolia factor. The frequency of face pareidolia varies in some factors such as specific illness, gender, and so on. In this experiment, I did not investigate whether the participants have the factors related to the face pareidolia frequency, therefore, the experiment result might be affected. In the future, this experiment is required to investigate more factors of face pareidolia frequency Finally, I will tackle manipulating pareidolia-inducing power for already existing stimuli. If the already existing stimuli can be manipulated by its pareidolia-inducing power, the research range will widen the recognition for the person, training for the artificial intelligence.

Contents

List of Figures

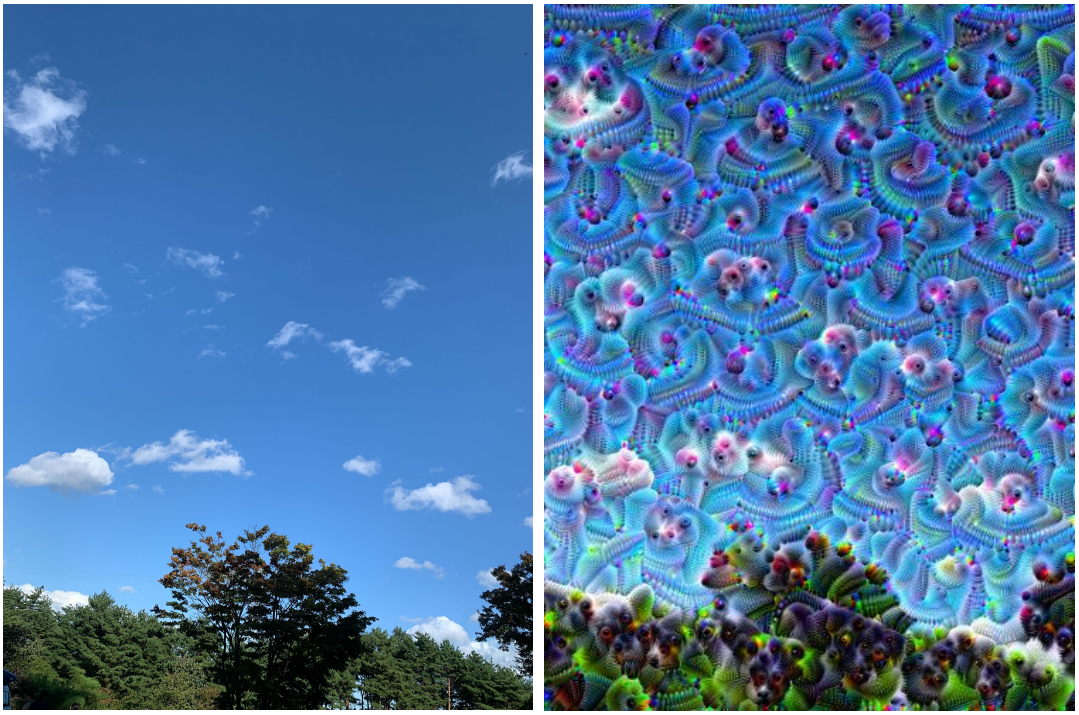
List of Tables

Chapter1 Introduction

Pareidolia, one of the psychological tendencies of human perception, is an essential mechanism for living. Pareidolia is the psychological tendency which perceive a meaningful pattern from an irregular pattern. Especially, some research papers report that the face category of pareidolia is higher frequency than the other some categories. As one of the typical examples of a pareidolia application, there is a program software, referred to as "DeepDream" [7]. The DeepDream was developed by Alexander Mordvintsev, an engineer at Google, and can create psychedelic images as shown in Fig. 1.1. The DeepDream aims to find the trained pattern from the input image and enhance the pattern of an input image.

One of the studies using the DeepDream reveals that a virtual reality video generated on DeepDream can induce a subjective experience similar to that of a realistic psychedelic [8].

The face case of pareidolia is one of the trends on the Internet (Faces In Places, FIP) [9]. The FIP images were uploaded on its blog [10]. FIP is one of the groups in which we can upload images so we can perceive face-like objects. There are about over 27000 images uploaded images on Flickr [11]. In addition, these images are



(a) Input image. (b) Generation result by DeepDream.

Figure 1.1: Generation result by DeepDream.

published in a book, called “Faces in Places: A Photographic Collection of Faces found in Everyday Places: Photos of Faces in Everyday Places” [12].

This mechanism makes us notice rapidly the face and potential facial pattern [13]. As the typical potential facial patterns, as shown in Fig.1.2, there is the front of the cars, grounded outlets, and so on.

Pareidolia is considered that there is a correlation between some specific illnesses and it, and investigated in some groups. As a typical illness example of higher pareidolia frequency, the patient suffering from Lewy body dementia tends to experience pareidolia more frequently. This is because visual-hallucination, one of the symptoms of Lewy body dementia [14], can be induced artificially by pareidolia stimuli.

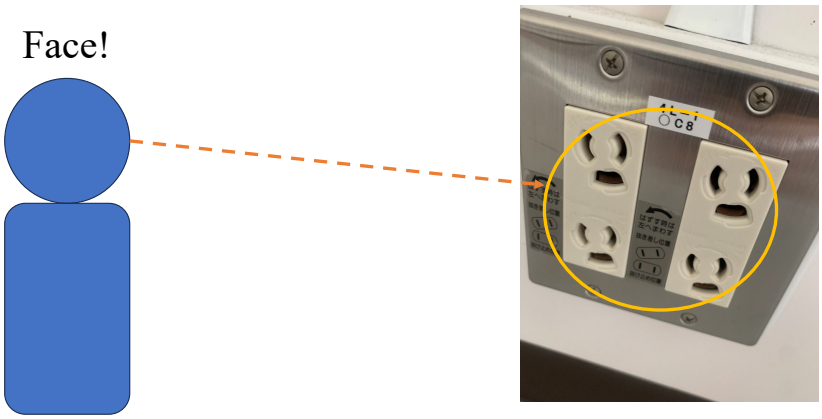


Figure 1.2: Face pareidolia example. We can perceive the grounded outlets as the face.

Based on this fact, pareidolia test [15], one of the diagnosis methods that can determine whether the patient is suffering from Lewy body dementia, has been invented. On the pareidolia test, the face pareidolia type (people and animals) is found on the presented stimuli for the patients Lewy body dementia and Alzheimer's disease. However, the stimuli used for the pareidolia test cannot be distributed because of the image copyright [16]. Vice versa, some illnesses are known to have a lower frequency of pareidolia. One of these is autism spectrum disorders [17].

In these papers, the visual face pareidolia stimuli are different from each other. For example, Alcimboldo's paintings [18], Mooney face test [19, 20] the stimuli are generated artificially [3], generated based on the human faces [21], and the pictures look like the faces [22, 23]. Comparing the results between the pareidolia papers is difficult because of some kinds of face pareidolia. Also, there is no study on the

relationship between the pareidolia categories and pareidolia-inducing power.

In a decade, there are high-quality applications of deep learning. In particular, the applications related to generation such as ChatGPT, and stable diffusion, were famous in 2023, and the words “Seisei AI (Generative AI)” became the top ten of Japan ’ s 2023 buzzword of the year. One of the famous applications of deep learning is generative adversarial networks, referred to as GANs [24]. The GANs learns the features of the images of the original existing data set and can generate images similar to the original data set. There are many applications of the GANs, and one of the typical applications is image-to-image style translation. The Cycle-consistent adversarial networks [25] (CycleGAN) is one of the famous frameworks of image-to-image style translation. The CycleGAN can translate images from one of the data sets to images applied the another data set style by learning the features of the data sets. The CycleGAN has the problem cannot learn the specific object; therefore, translation focused on the specific object is difficult. From a different perspective, the problem can be translated into the image style that can remain the original image feature.

Therefore, this thesis aims to systematically generate the visual face pareidolia-inducing stimuli using the CycleGAN. This paper focuses on the face pareidolia stimuli generated from human faces. The important elements of face detection are reported in the eyes and mouth. I consider that the face pareidolia stimuli is gener-

ated by the image style translation that remain these features.

Moreover, this thesis also focuses on pareidolia-inducing power. The pareidolia-inducing power seems to be different depending on the color, texture, shape, and edges. In this thesis, I aim to manipulate the pareidolia-inducing power by manipulation of the parameter of cycle-consistency loss, one of the loss functions of the CycleGAN.

Chapter2 Pareidolia

2.1 Abstract

Pareidolia is one of the psychological tendencies. Pareidolia is a perception of a defined pattern from an irregular pattern. However, the feature of pareidolia is the perception of not only the defined pattern but also an irregular pattern. As shown in Fig. 2.1, we can find the face pareidolia in our daily lives such as in front of cars, grounded outlets, in the cloud, and so on.

This tendency might be reflected in the social attention [26]. In the Pareidolia test, the face is one of the categories answered by the participant. Also, Taubert et al. revealed that rhesus monkeys react to the face pareidolia stimuli more than the other stimuli [27].

One of the kinds of pareidolia, "Auditory pareidolia" is known [28, 29]. Regarding auditory pareidolia, the words heard from special situations trigger perceptions that seem to make sense to the person.

Figure 2.1: One of the pareidolia example. The above holes are perceived as the eyes, and the following hole is perceived as the mouth.

2.2 Face pareidolia

Face is one of the socially important information such as personal identification, communication, emotion, face expression, and so on. For example, according to Mehrabian, the most prioritized information among language, auditory, and vision information is vision information [30] (referred to as "Mehrabian's Law").

The cause of face pareidolia detection seems to miss the face-like information. Because the face is socially important information, if we miss the real face, the loss

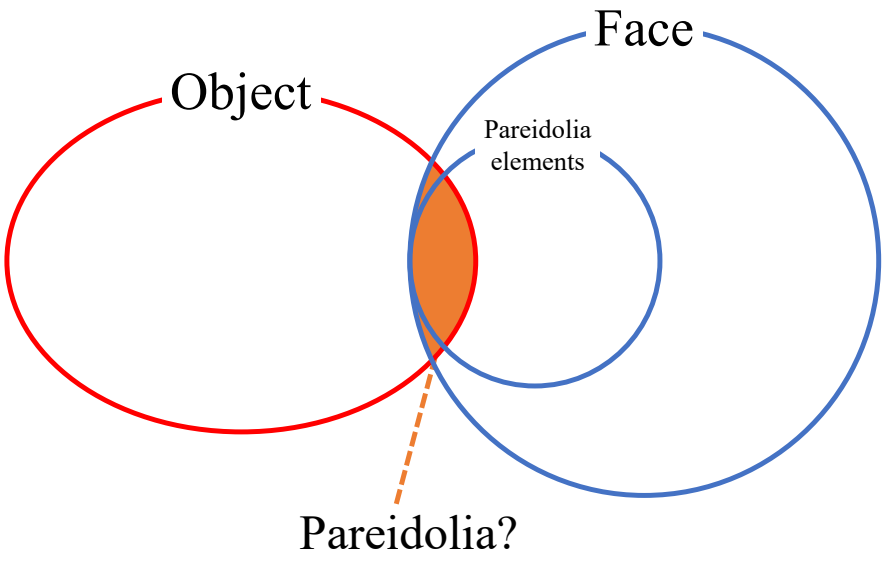


Figure 2.2: Relation between objects, face, and face pareidolia. we can experience face pareidolia not only with the inclusion of all face elements but also face pareidolia elements.

will be bigger. Therefore, we can perceive the face-like information as the face.

We often experience face pareidolia because of face importance. Face pareidolia is the stimuli that can induce face perception when the observer looks at the stimuli. As shown in Fig 2.2, I believe face pareidolia includes both face and object aspects.

For example of Fig. 2.1, we can confirm the eyes-like and mouth-like features. All face pareidolia stimuli do not include all elements of the actual face. However, we can perceive these face pareidolia as face. Therefore, the number of minimum elements for face perception can be less than the number of actual face elements.

Also, regarding the similar tendency of face pareidolia, we also experience simulacra, the forms of three points (like the form of "∴") can be perceived as the face, and may be considered one of the pareidolia. The three points of simulacra might

include the eyes-like and mouth-like features; therefore, these features can trigger the face perception.

2.2.1 Stimuli type

There are many types of face pareidolia stimuli used for the psychological experiment. One of the face pareidolia types is artificially generated. As an example, based on Alcimboldo's paintings are one of this type. As shown in Fig. 2.3, Alcimbold paintings have features that consist of multiple objects (vegetables, fruits, plants).

Based on the Alcimboldo's paintings, Pavlova et al. created "face-n-food task" [2] to investigate the gender differences of face perception. In [2], total of 10 image stimuli are created and ranked according to face recognition difficulty of the pilot group of participants. Fig. 2.4 shows the stimuli of "face-n-food task".

Pavlova et al. also used the stimuli to investigate face tuning such as Down Syndrome [31], Autistic spectrum disorders [17], Williams Syndrome [32], cultural difference [33]. In addition, the "face-n-food task" was used to investigate the factor of face tuning of gender differences [34], face sensitivity of schizophrenia [35], face sensitivity of born preterm [36]. The stimuli of the "face-n-food task" are evaluated firstly the ease of recognizing the face. Also, the stimuli are generated by the authors; therefore, the generation and evaluation of the stimuli need a lot of labor.

Figure 2.3: Spring, 1573. This figure is taken from [1].

Kato and Mugitani use the stimuli as shown in Fig. 2.5 to investigate face pareidolia perception of infants [3].

Liu et al. invented the stimuli embedded in the face and a letter in a noise image as shown in Fig. 2.6 [4]. As to the pareidolia test, the stimuli of the noise pareidolia test are also embedded in the face and object as shown in Fig. 2.7. The last one

Figure 2.4: The stimuli used for the “face-n-food”. The left image looks the worst like the face. Inversely, the right image looks the best like the face. This figure is taken from [2].

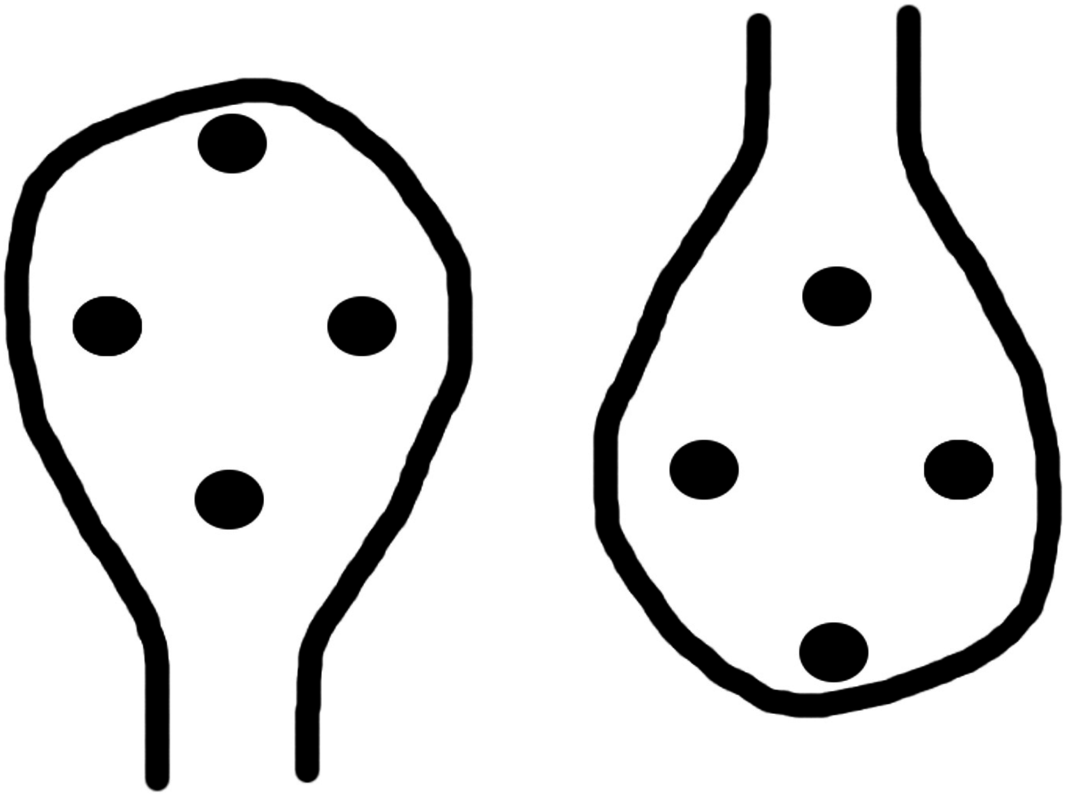


Figure 2.5: The stimuli used in [3]. This figure is taken from [3].

is based on the facial elements. Mooney face test is one of the examples of this type [19]. As shown in Fig. 2.8, the data set is a binary image with minimum face recognition information. At present, because of few images of the data set, there is

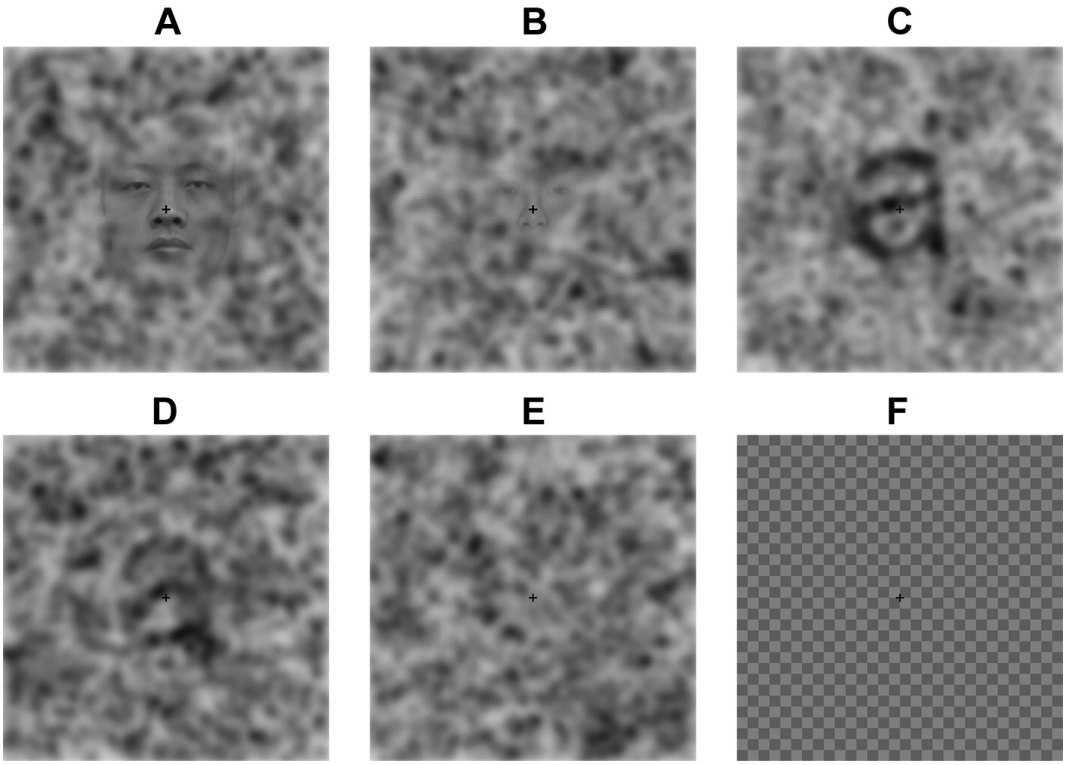


Figure 2.6: The stimuli used in [4]. The A and B images are embedded in the face. The C and D images are embedded in a letter (in this example, "a" is embedded). The E image is the pure noise image, and The F image is checkerboard image. This figure is taken from [4].

research to generate the Mooney face style images [37, 38].

These stimuli are generated either by authors artificially or naturally, There is no study to generate systematically in response to the pareidolia-inducing power.

2.2.2 Activation part of brain

The brain activation parts are measured by functional magnetic resonance imaging [39, 40], electroencephalography [23], and event-related potentials [41] for face pareidolia. Liu et al. reported that when the participants saw the faces, the right fusiform face area response specifically [4]. When we see the front of cars, we acti-

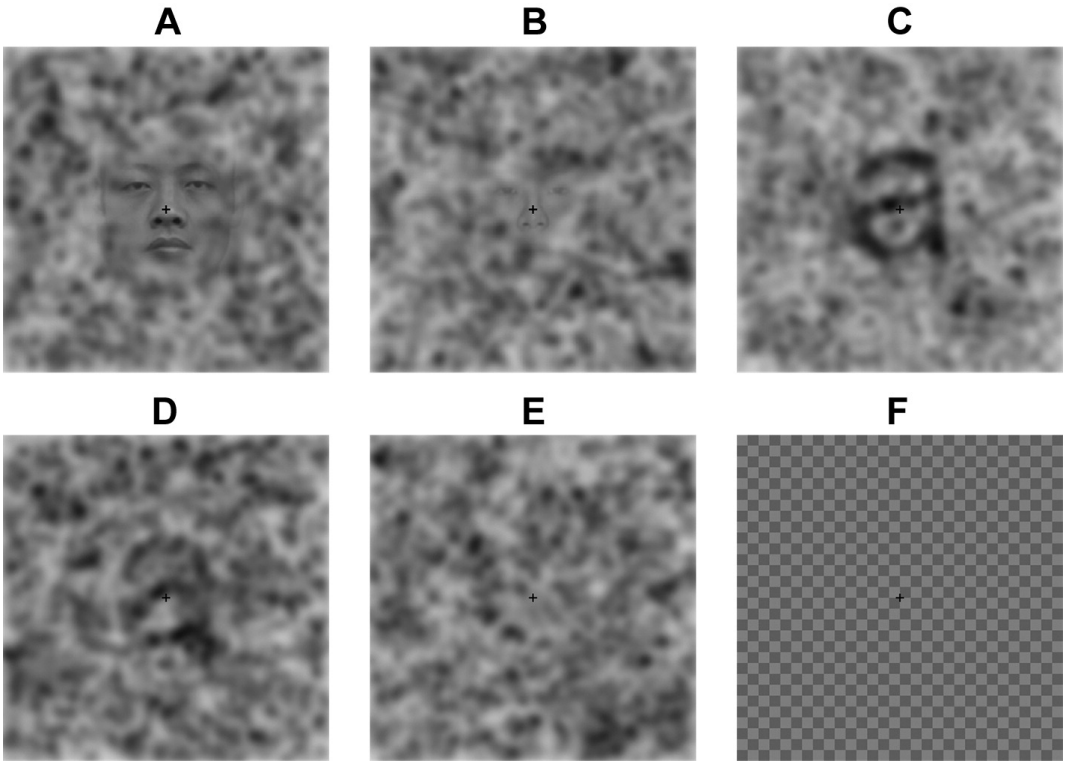


Figure 2.7: The stimuli used in [4]. The A and B images are embedded in the face. In the C and D, a letter is embedded (in this example, “a” is embedded). The E image is the pure noise image, and The F image is the checkerboard image. This figure is taken from [4].

vate the fusiform face area, temporoparietal junction, and medial prefrontal cortex, the activation part when we see the face [42]. When we see the face pareidolia, the pareidolia can evoke N170 component [43].

2.2.3 Relation to illnesses

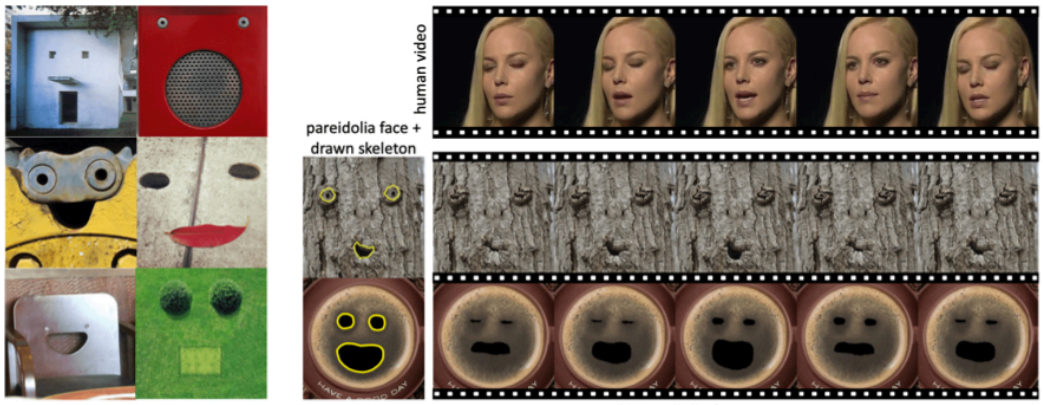
Pareidolia is reported that the experiment frequency is different by specific illnesses. Mainly, the patient with neurodevelopmental relates to the frequency of pareidolia experience. As described in Chap. 1, the patients suffering from Lewy body dementia experience more pareidolia than not only healthy control but also

(a) Facing right women image. (b) Facing left men image.

Figure 2.8: Mooney face data set. This figure is taken from [5].

the patients suffering from Alzheimer's disease. Regarding Lewy body dementia, the pareidolia-inducing stimuli have the potential to evoke visual-hallucination, a typical symptom of Lewy body dementia. Therefore, pareidolia-inducing stimuli are used when the diagnosis of Lewy body dementia.

In addition, the patient suffering from Parkinson's disease without dementia [44] and Rapid eye movement sleep behavior disorder [45] are known for the high frequency of pareidolia experience. Inversely, as the illnesses of the low frequency of pareidolia, children with Autism spectrum disorder [46], William syndrome [32], and Down syndrome [31] are known.



(a) Examples of parediolia faces

(b) Parediolia face reenactment

Figure 2.9: Manipulation of the existing stimuli to imitate the real face emotion. This figure is taken from the project page[2] of [6].

2.2.4 Relation to other research fields

Pareidolia is widely investigated not only neuroscience field but also in other research fields. In the computer vision field, there are some studies such as facial expressions of the pareidolia on house exterior design are analyzed using expression classifier [47], manipulation of the existing stimuli to imitate the real face emotion [6] (Fig. 2.9), emotional analysis of existing pareidolia stimuli [48], and analysis of products including face pareidolia [49].

In addition, convolutional neural network (VGG16 [50]) can recognize the face pareidolia by training the human face [51]. If the mechanism of training of face pareidolia for AI is revealed, it can contribute to the face pareidolia perception of human beings.

Advertisements that include face or face pareidolia attract more than advertise-

[1]https://wywu.github.io/projects/ETT/ETT.html

ments that exclude face information [52].

Studies regarding pareidolia are reported in not only the neuroscience field but also the computer vision field, advertising field, and marketing. Face is important information socially. The potential cause for the face perception of human beings will be revealed by advancing the research regarding face pareidolia.

Chapter3 Artificial intelligence

3.1 Abstract

Artificial intelligence (AI) is a new technique in a century, named by John McCarthy. In recently, the AI wins professionals such as Go [53]. There are many applications of AI such as ChatGPT[1], image generation [54], referred to as "Seisei AI". However, the information from ChatGPT is doubtful; therefore, we need to investigate whether the information is true. Also, regarding image generation, because the AI learns the image drawn by the professionals, the problem is becoming more obvious.

3.2 Perceptron

Perceptron is the algorithm invented by Rosenblatt [55]. Perceptron is the basic model of a neural network. As the mathematical description, the input signals (x_1, x_2) are inputted to the neuron firstly. The inputted signals are timed by weights (w_1, w_2), respectively. As shown in Eq. 3.1, the perceptron outputs 1 when the timed

[1] https://chat.openai.com

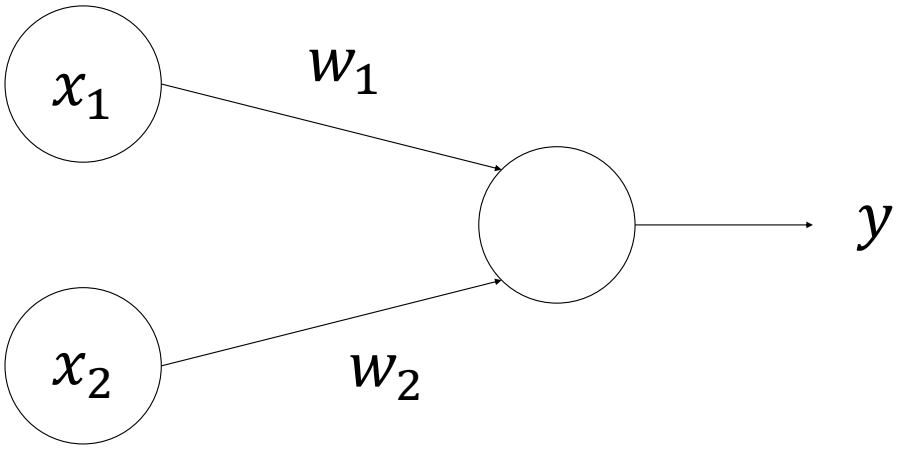


Figure 3.1: Perceptron model.

Table 3.1: OR truth table.

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	1

signals is above the threshold (θ). Also, the perceptron outputs 0 when the timed signals is less than the threshold.

$$y = \begin{cases} 0 & (w_1 \times x_1 + w_2 \times x_2 \leq \theta) \\ 1 & (w_1 \times x_1 + w_2 \times x_2 > \theta) \end{cases} \tag{3.1}$$

This model is shown in Fig. 3.1.

However, Minsky pointed out that the perceptron cannot solve the non-linear separation problem such as XOR [56]. The linear separate problem is that a simple line can separate whether the output as shown in Table. 3.1 and Fig. 3.2.

As shown in Table. 3.2 and Fig. 3.3, a simple line cannot separate the outputs.

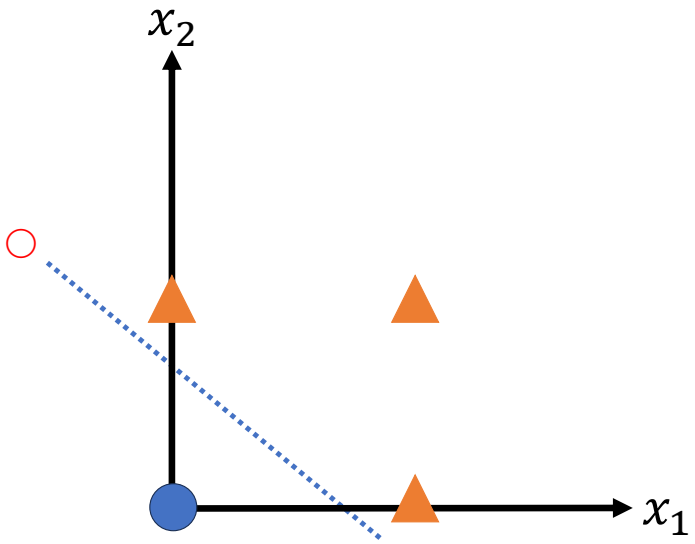


Figure 3.2: OR output. A simple line can separate the output field.

Table 3.2: XOR truth table.

x_1	x_2	y
0	0	0
0	1	1
1	0	1
1	1	0

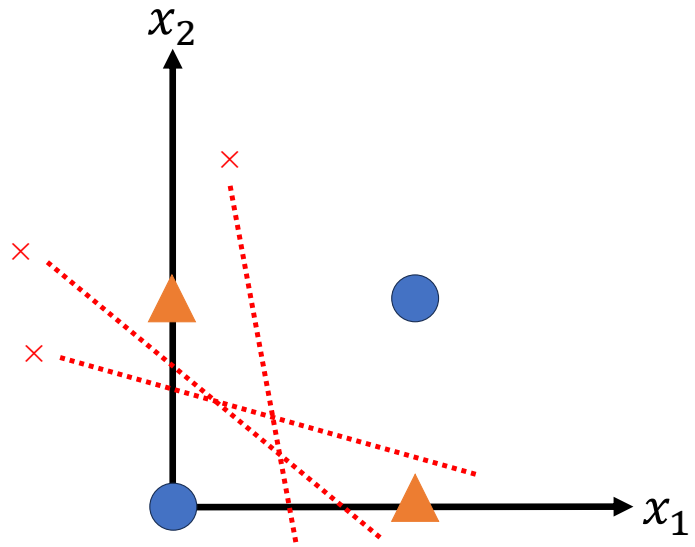


Figure 3.3: XOR output. A simple line cannot separate the output field.

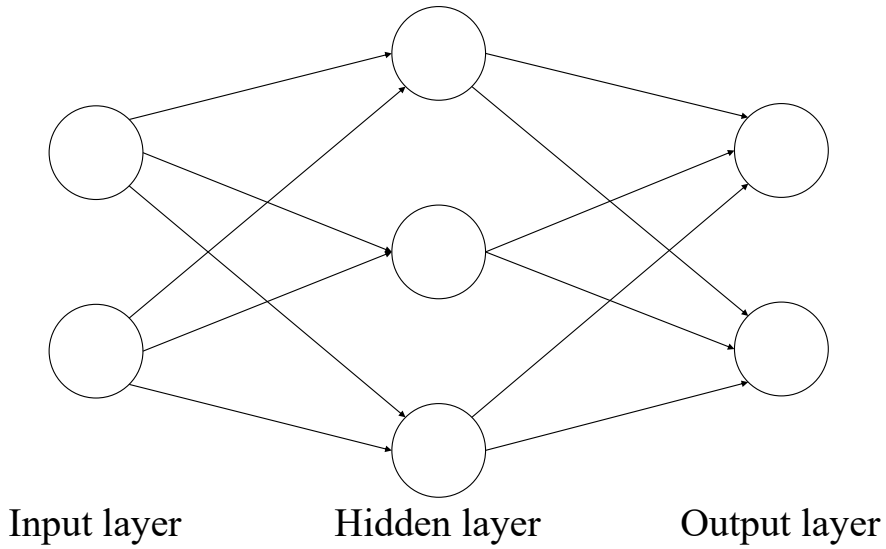


Figure 3.4: The basic model of neural network.

3.3 Neural network

Neural network is firstly invented by Romelhart et al [57]. The basic model of the neural network is shown in Fig. 3.4. The first layer of Fig. 3.4 is referred to as “Input layer”, the second is referred to as “Hidden layer”, and the third layer is referred to as “Output layer”, respectively.

3.4 Deep learning

Deep learning is first invented by Hinton et al. [58], one of the applications of neural networks.

3.4.1 Convolutional neural network

The convolutional neural network (CNN) was first invented by LeCun et al (referred as to LeNet). [59]. LeNet was applied the backpropagation algorithm [57] to solve the recognition problem of handwritten digit number. The first CNN shows high performance on the handwritten zip code recognition. The CNN can increase performance of object detection [60, 61, 62], image recognition [63, 64, 65, 66, 67], In particular, there are many CNN applications of face detection [68, 69, 70], and face recognition [71, 72]. The main features of CNN are the convolutional layer and pooling layer. Each feature is described in the next sections.

Convolutional layer

A convolutional layer function captures the locally feature target objects. In the example in Fig. 3.5, the 3×3 filter can capture the cross feature. The output in Fig. 3.5, the larger the number, the more suitable for the feature. Also, the output is referred to as a feature map and used for calculation on the pooling layer.

Pooling layer

A pooling layer function gets the maximum or average value on the feature map. A pooling layer can extract the necessary information for training and improve the training efficiency of deep learning. As shown in Fig. 3.6, the max pooling layer

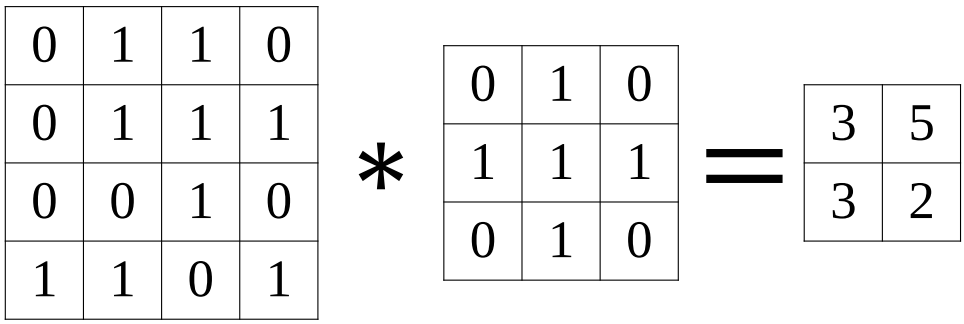


Figure 3.5: Convolutional layer process. The middle filter is applied to the left 4×4 image. Larger the value of the right 2×2 grid (referred to as "Feature map"), more like a filter feature.

can extract the maximum value of the feature map. The max pooling aims to extract the local features such as edges. As shown in Fig. 3.7, the average pooling layer can extract the average value of the feature map. The average pooling aims to extract the features of a small region. The calculation range of Figs. 3.6 and 3.7 is 2×2 grid.

3.5 Generative adversarial networks

Generative adversarial networks (GANs) is a deep learning framework invented by Ian et al [24]. The GANs is first invented in 2014. The GANs can generate the The frameworks based on the GANs are used for various applications, such as image generation [73, 74, 75], photo cartoonization [76], image super-resolution [77], and

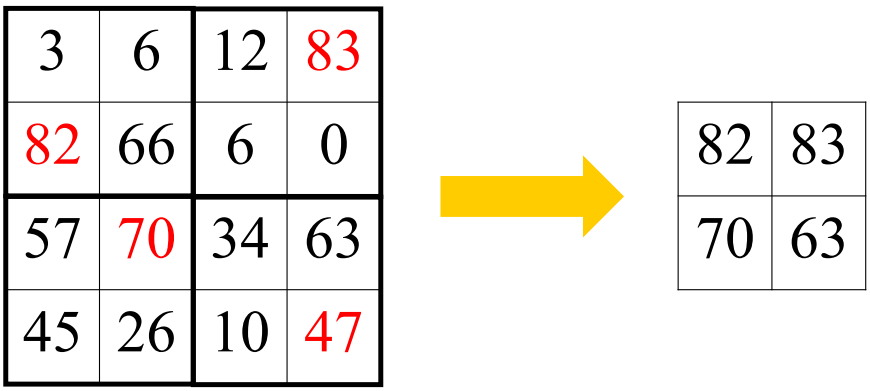


Figure 3.6: Max pooling layer process. The value of the right grid is the maximum value corresponding to the 2 × 2 grid of the left grid.

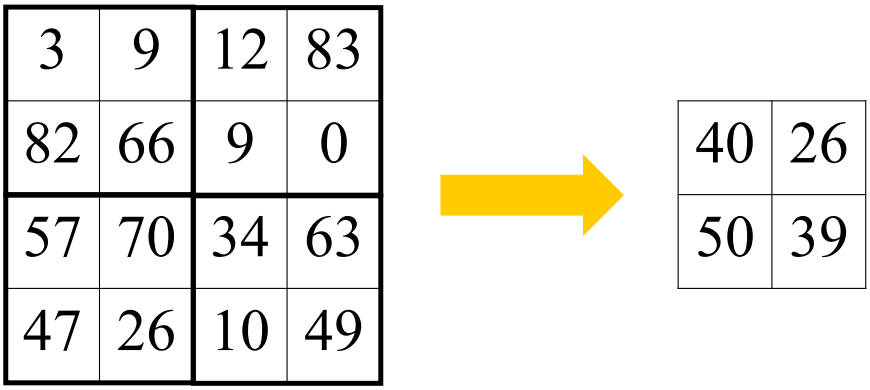


Figure 3.7: Average pooling layer process. The value of the right grid is the average value corresponding to the 2 × 2 grid of the left grid.

image-to-image style translation[25, 78, 79].

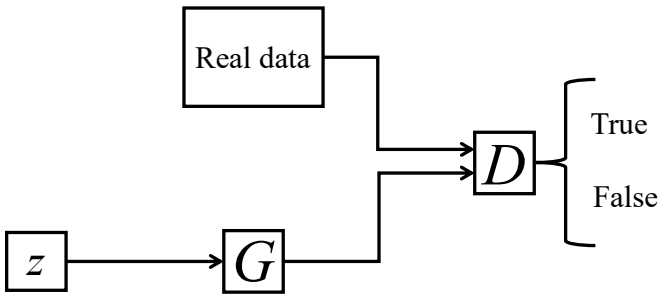


Figure 3.8: GANs framework.

3.5.1 Architecture

This framework of the GANs is shown in Fig.3.8. This framework consists of two types of networks, a generation network referred to as a "Generator" (G in Fig.3.8), and a discrimination network referred to as a "Discriminator" (D in Fig.3.8).

Here, z of Fig. 3.8 is latent variable and can generate an image like the original data set by input to the generator. The generator functions the image generation and aims to generate images that cannot be distinguished from the real data set. On the other hand, the discriminator functions the image discrimination and aims to discriminate accurately whether the image is true (from a real data set) or fake (generated from the generator). The generator and the discriminator learn alternatively,

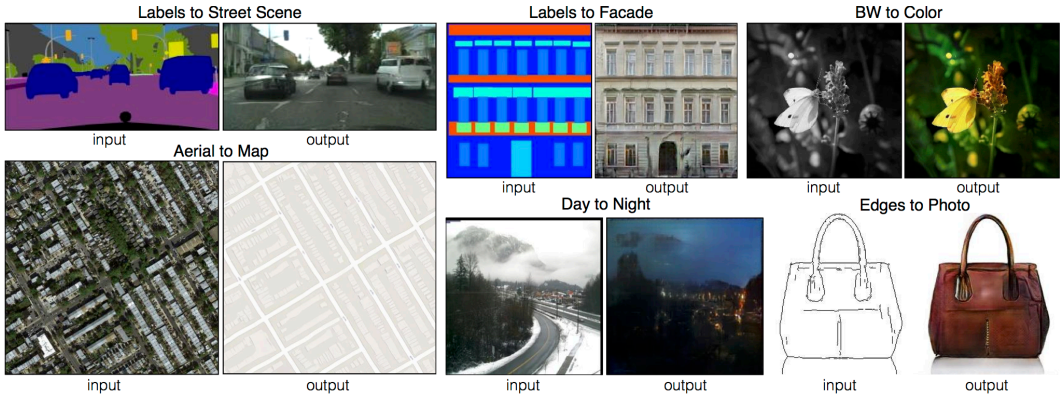


Figure 3.9: The paired data sets examples. There is an image which is corresponds to the translation from the original image. Image source from Pix2pix project page[2].

and the objective function is expressed like minimax game by Eq. 3.2:

$$\min_G \max_D V(G, D) = \mathbb{E}_{\boldsymbol{x} \sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z} \sim p_{\boldsymbol{z}(\boldsymbol{z})}}[\log(1 - D(G(\boldsymbol{z})))]. \tag{3.2}$$

3.5.2 Cycle-consistent adversarial networks

Regarding the image style translation, many approaches are proposed by researcher.

As examples of the applications of image-to-image style translation, there are a pix2pix and a cycle-consistent adversarial networks (CycleGAN). In the pix2pix method case, the methodology can only translate between the two data sets corresponding to the before-translation and after-one as shown in Fig. 3.9.

Ian et al. invented the CycleGAN to apply the translation between the two data

[2]https://github.com/phillipi/pix2pix

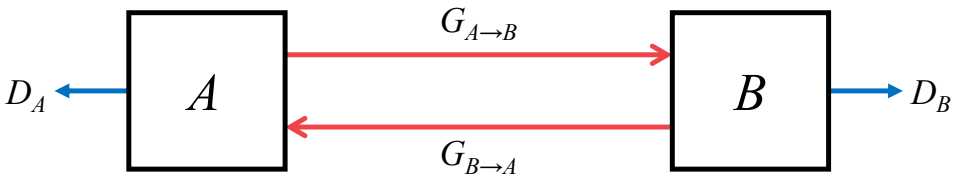


Figure 3.10: Framework of the CycleGAN. The A and B are images of the data sets of each image style, respectively. The $G_{A\to B}$ and $G_{B\to A}$ are generators. The function of the generator is the image-to-image style translator. Also, the D_A and D_B are discriminators. The function of the discriminator is discrimination of whether the input image is generated by the generator or a real data set.

sets not corresponding to the before and after. This framework is shown in Fig.3.10.

The CycleGAN consists of two generators and two discriminators. The generator of the CycleGAN is used as an image-style translation network and aims to translate to another data set image style. As shown in Fig. 3.11, CycleGAN shows high performance on various pairs of styles.

The loss function of the CycleGAN applies not only the original GANs' loss

[3]https://junyanz.github.io/CycleGAN/

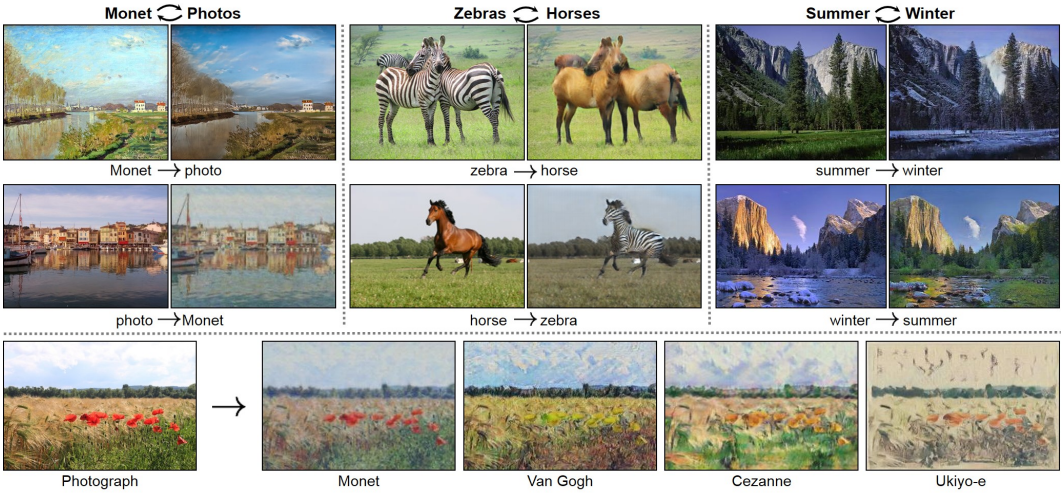


Figure 3.11: The unpaired data sets examples. There is not the image translated from the original image; however, CycleGAN can generate such images. Image source from CycleGAN project page[3].

function (like Eq.3.3 and Eq.3.4) but also cycle consistency loss.

$$\mathcal{L}_{GAN}\left(G_{A\to B}, D_B, A, B\right) = \mathbb{E}_{b\sim p_{data}(b)}[\log D_B(b)] + \mathbb{E}_{a\sim p_{data}(a)}[\log(1 - D_B(G_{A\to B}(a))]. \tag{3.3}$$

$$\mathcal{L}_{GAN}\left(G_{B\to A}, D_A, B, A\right) = \mathbb{E}_{a\sim p_{data}(a)}[\log D_A(a)] + \mathbb{E}_{b\sim p_{data}(b)}[\log(1 - D_A(G_{B\to A}(b))]. \tag{3.4}$$

Here, A and B is the each data set domain, and a and b is the images of each domain ($a \in A$, $b \in B$) In the CycleGAN, the generators ($G_{A\to B}, G_{B\to A}$) function the image translation network to translate the image of a data set to the image of another data set, and the discriminator (D_A, D_B) function the image discrimination network to discriminate the input image is whether the original image of the data

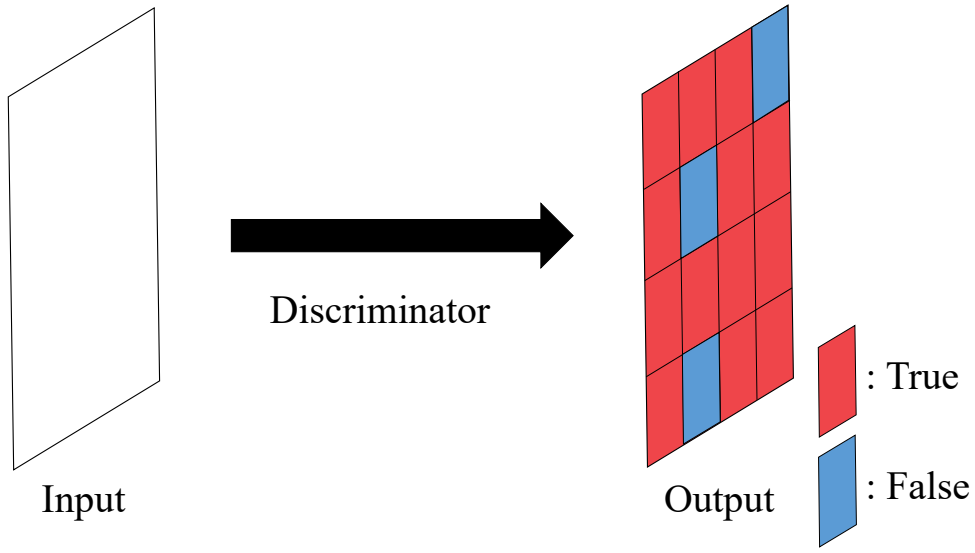


Figure 3.12: The PatchGAN process image. The discriminator does not discriminate overall the input image. The input image is discriminated on each small grid whether true or false.

set or generated from the generator. Equations 3.3 and 3.4 express the loss of data set A and B.

The generator of CycleGAN is based on [80], one of the networks of image-style-translation using deep learning. Also, the discriminator of the CycleGAN is based on PatchGAN [81]. As shown in Fig. 3.12, the PatchGAN functions discriminate the input image whether generated by the generator or from data set in $N \times N$ patch of an input image.

Herein, N is set to 70, described as a suitable value in [25, 78].

If the image is translated by one of the generator inputs to another generator, the translated image by another generator is similar to the original image. The difference between the two images is treated as the cycle consistency loss. The

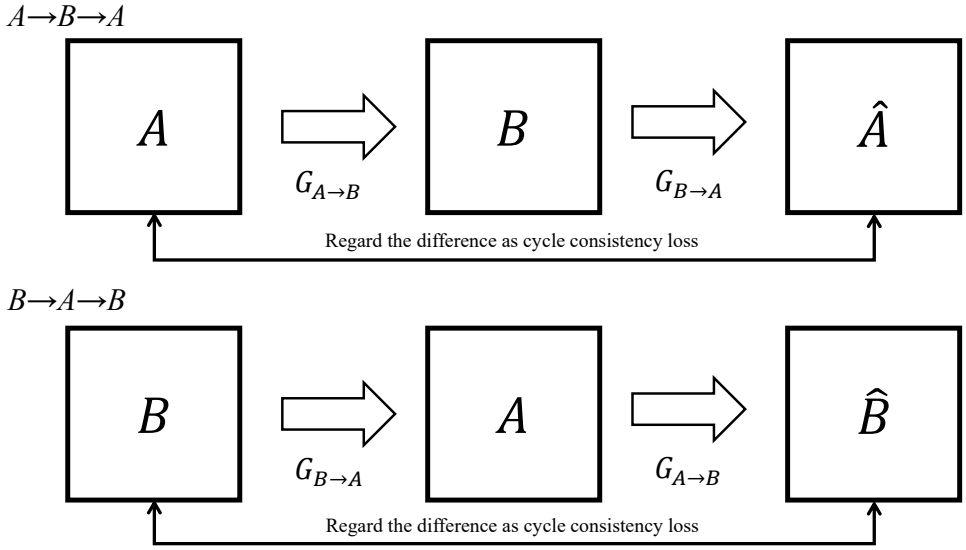


Figure 3.13: Schema of the cycle consistency loss. the A and $\hat{A}$ images are compared, and the difference between A and $\hat{A}$ is treated as the cycle-consistency loss. The difference between B and $\hat{B}$ images is treated in the same way.

cycle consistency loss is expressed as follows:

$$\begin{aligned}\mathcal{L}_{cyc}\left(G_{A\to B}, G_{B\to A}\right) = \mathbb{E}_{a\sim p_{data}(a)}[\|G_{B\to A}(G_{A\to B}(a)) - a\|_1] \\ + \mathbb{E}_{b\sim p_{data}(b)}[\|G_{A\to B}(G_{B\to A}(b)) - b\|_1].\end{aligned} \tag{3.5}$$

Figure 3.13 shows the cycle consistency loss expressed in the image.

Finally, the objective function of CycleGAN is expressed as follows:

$$\begin{aligned}\mathcal{L}\left(G_{A\to B}, G_{B\to A}, D_A, D_B\right) = \mathcal{L}_{GAN}\left(G_{A\to B}, D_B, A, B\right) \\ + \mathcal{L}_{GAN}\left(G_{B\to A}, D_A, B, A\right) \\ + \lambda\mathcal{L}_{cyc}\left(G_{A\to B}, G_{B\to A}\right).\end{aligned} \tag{3.6}$$

Here, λ is the weight of the cycle consistency loss and can manipulate the influence

of each loss function. Additionally, the example of application identity mapping loss in [25]. This identity mapping loss is defined as Eq. 3.7:

$$\begin{aligned}\mathcal{L}_{identity}\left(G_{A\to B}, G_{B\to A}\right) = \mathbb{E}_{a\sim p_{\text{data}}(a)}[\|(G_{A\to B}(a)) - a\|_1] \\ + \mathbb{E}_{b\sim p_{\text{data}}(b)}[\|(G_{B\to A}(b)) - b\|_1]\end{aligned} \tag{3.7}$$

For example, translation between the painting domain and the real picture domain needs to retain the image color information. The CycleGAN aims to learn to solve the following objectives:

$$G_{A\to B}, G_{B\to A} = \arg \min_{G_{A\to B}, G_{B\to A}} \max_{D_A, D_B} \mathcal{L}\left(G_{A\to B}, G_{B\to A}, D_A, D_B\right). \tag{3.8}$$

Chapter4 Proposed Method

4.1 Main idea

As shown in Fig. 4.1, this thesis attempts to generate the pareidolia-inducing stimuli based on the idea. The face pareidolia-inducing stimuli consist of not only without face elements but also the elements that can be perceived as the face. CycleGAN cannot learn specific objects; that is to say, CycleGAN can remain the input feature. Therefore, this thesis aims to generate face pareidolia stimuli using real face features using CycleGAN. This thesis used real-face and natural-scene data sets to generate stimuli. In this thesis, the generated pareidolia-inducing stimuli resemble the after-translation image style and include the eyes and lips, using the original feature of a face data set. The CelebAMask-HQ data set was applied herein because the objective of this thesis was to use face attribute information (such as eyes and mouth). The CelebAMask-HQ data set includes the label information of face attributes such as eyes, hair, and skin. Additionally, the training data set of natural images is collected from Flickr. In this thesis, I aim to generate the pareidolia-inducing stimuli using the eyes and mouth region.

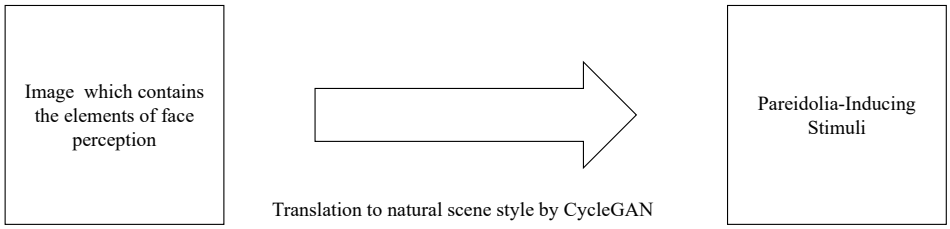


Figure 4.1: The main idea of pareidolia-inducing stimuli. In this thesis, I attempt to generate pareidolia-inducing stimuli by preservation of face elements.

4.2 Face pareidolia generation

One of the CycleGAN features, the loss function referred to as cycle-consistency loss. As shown in Fig. 3.13, the difference between the original input image and the image generated to be similar to the original image is calculated. Then, the difference is regarded as the cycle-consistency loss and makes the generators learn the after-translation image can be translated to the original input image. As the cycle-consistency loss effect, the cycle-consistency loss function can reduce the style translation-mapping solutions. In addition, the weight parameter of the cycle-consistency loss (λ, represented in 3.6) affects the performance of the generation. The pareidolia-inducing power might be associated with the weight parameter; thus, we trained the CycleGAN on the aforementioned data set with some λ values and

experimented with the generated stimuli evaluation using the stimuli generated by the model trained on these various λ values. This thesis uses the CycleGAN framework implemented on Pytorch[2] [82], one of the libraries on deep learning in Python.

4.2.1 Data set

The CycleGAN can translate the image style bidirectionally between unpaired two data sets. This thesis aims to generate natural face pareidolia; therefore, I prepare two types of the data sets, face data set and natural data set.

Face data set

There are many face data sets for studies such as face detection [83, 84, 85] and face recognition [4, 86]. In this thesis, as the face data set, this thesis uses CelebAMask-HQ [87]. The data set includes the mask images corresponding to each face attribute (eye, nose, mouth, etc.) as shown in Figs. 4.2.

I extracted the face pareidolia-inducing elements from these mask images as shown in Figs. 4.3.

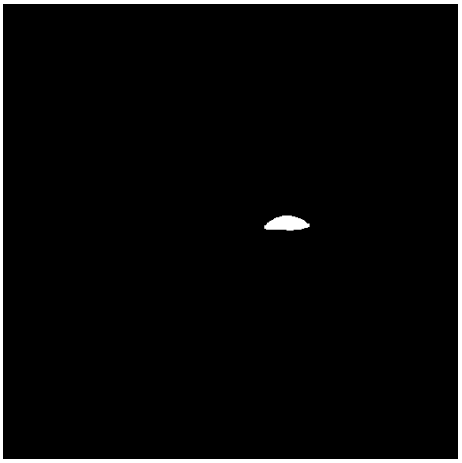
Regarding the pareidolia-elements, the eyes and mouth are important factors for face detection [88]. Therefore, I preserve the eyes and mouth features of CelebAMask-HQ (Right eye, Left eye, Upper lip, Lower lip, Mouth). In addition, I apply the pre-

[2]https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix/tree/master

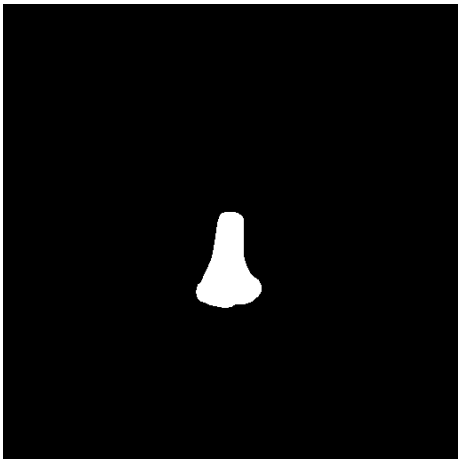
(a) Face image.

(b) Hair mask image.

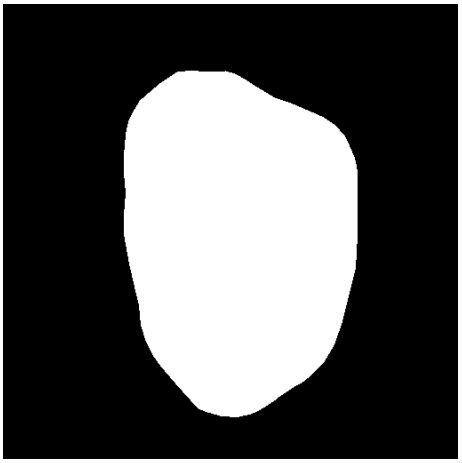
(c) Left brow mask image.



(d) Left eye mask image.

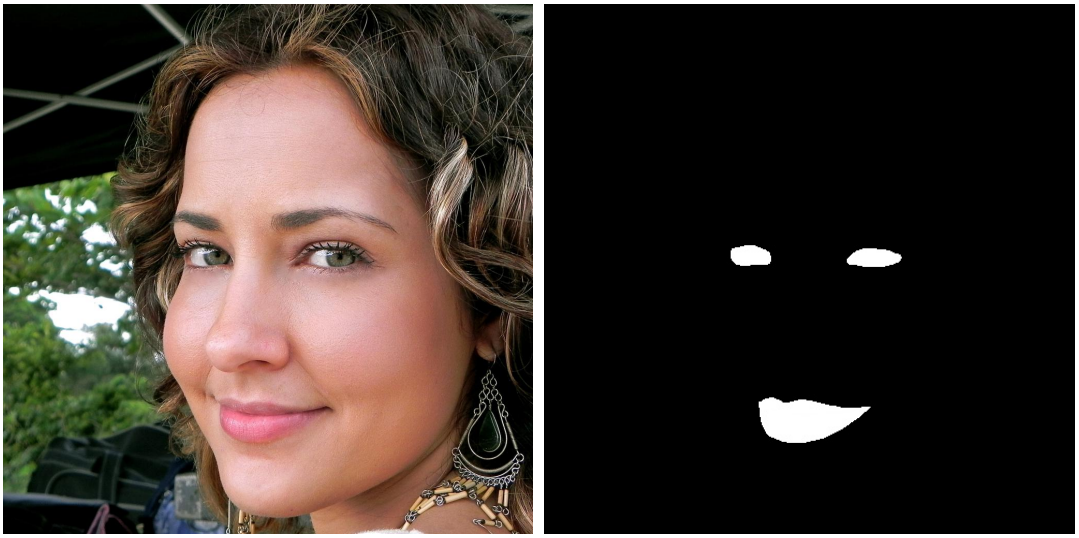


(e) Nose mask image.



(f) Skin mask image.

Figure 4.2: A face and mask image examples of CelebAMask-HQ. There are 19 classes of face attribution and accessories.



(a) Original face image. (b) Face pareidolia components mask image.

Figure 4.3: An image example of Face pareidolia elements.

processing except for the pareidolia-elements and extract the pareidolia-elements when generation of pareidolia-inducing stimuli. All labels are shown as follows:

- Hair,
- Left/Right brow,
- Left/Right eye,
- Left/Right ear,
- Upper/Lower lip,
- Mouth,
- Neck,
- Nose,

- Cloth,
- Skin,
- Earrings,
- Necklace,
- Hat,
- Eyeglass.

In addition to the above, there are 19 labels inclusion of background the other region. A total of 100 images are extracted randomly from all images and used for the training data set.

The borders of the skin, hair, and background exist in the face contours of the image data set. As described in Chap. 4, a specific object is difficult to learn for the CycleGAN. Additionally, the generated images changed significantly, with changes in the image color using the CycleGAN confirmed. Therefore, the image translated by the generator of the CycleGAN strongly retained the facial contours. To overcome this problem and preserve the pareidolia-elements features, two types of preprocessing, namely blurring, and noise processing are applied to the original face data set. An image example illustrating the preprocessing steps is shown in Fig. 4.5.

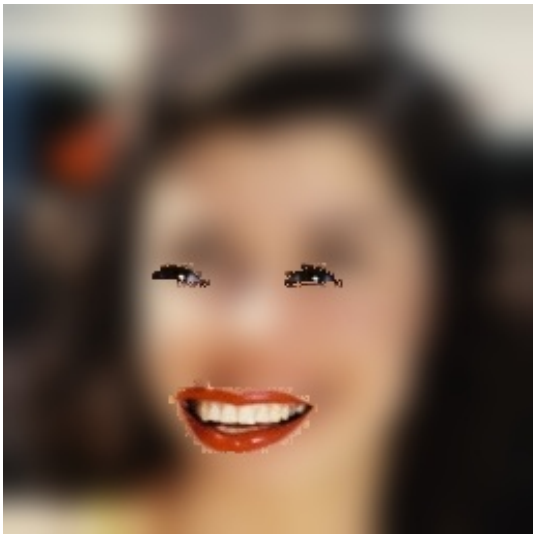
Figure 4.4: Original face image.

Also, the original face images are preprocessed including pareidolia-elements and whole image (excluding the pareidolia-elements).

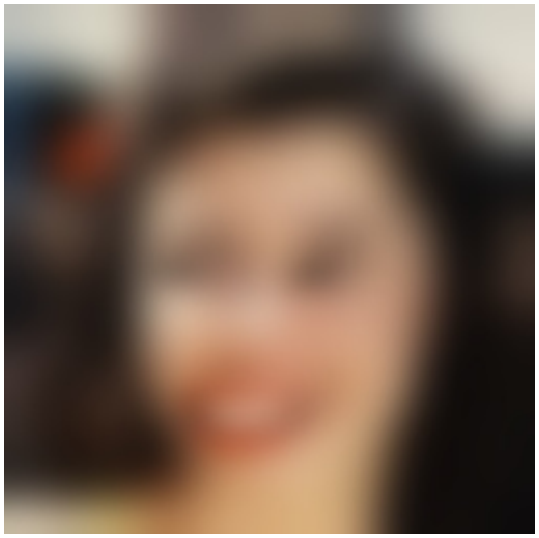
The face of the original image of the CelebAMask-HQ is placed the center of the image. Therefore, a bias toward finding the face from the center of the stimulus needs to be considered. To resolve this problem, the pareidolia-elements are moved randomly.

Natural scene data set

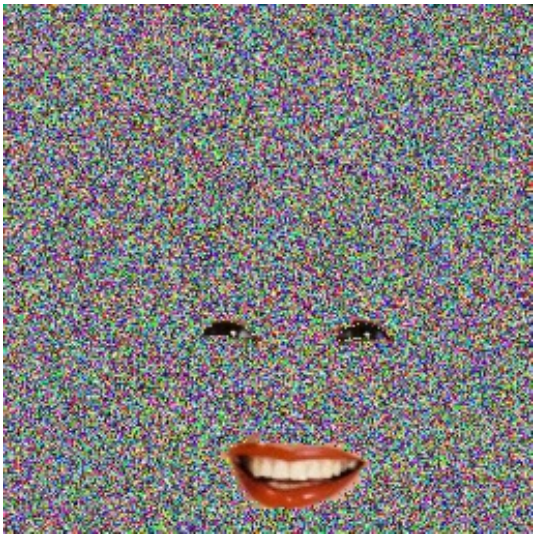
As the natural scene data set, we collect the images from Flickr [11]. The one of the exampler image is shown in Fig. 4.6.



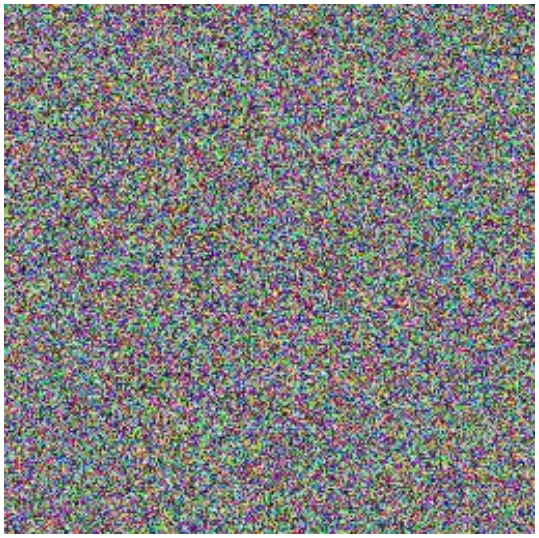
(a) Blur processing excluding the face pareidolia-elements



(b) Blur processing including the face pareidolia-elements



(c) Noise processing excluding the face pareidolia-elements



(d) Noise processing including the face pareidolia-elements

Figure 4.5: Preprocessed image example.

4.3 Evaluation of the generated stimuli

A standard psychophysical method for measurement of pareidolia-inducing power has not yet been developed. In addition, the frequency and sensitivity of pareidolia might differ between each individual. Thus, this thesis examined the measurement

Figure 4.6: Natural scene image example.

experiments conducted previously on pareidolia to design the experimental procedure. In particular, the experiment of the noise pareidolia test [21] was found to be similar to that of our research objective. The procedure for the noise pareidolia test is described. First, the investigator of the noise pareidolia test presented the experimental stimulus to the participants. Next, the participants are required to answer whether the stimulus includes the face or a specific object. When the participants answered the face in the stimulus, the participants were required to point out the facial region or specific object. In this thesis, I aim to measure the pareidolia-inducing power of the generated face pareidolia stimuli. Therefore, the participants determined whether pareidolia was included in the presented stimulus and I instructed them to evaluate the face pareidolia intensity for the stimulus. In

this thesis, I also aim to investigate the pareidolia cause; therefore, the presented stimuli include real face images, natural images, and generated images that do not include the pareidolia-elements. The experiment details are described in the next chapter.

Chapter5 Experiment

This thesis consists of two types of experiments. The first experiment is the generation of pareidolia stimuli; this thesis generated face pareidolia image stimuli using the CycleGAN. The second experiment is the evaluation of the generated stimuli. Pareidolia might differ with the sensitivity of pareidolia-inducing power between the participants; therefore, the experiment result is normalized and analyzed.

5.1 Face pareidolia generation experiment

This thesis used the CycleGAN to generate the images. As described in Sec. 4.2.1, the training dataset included a natural image dataset and a preprocessed face image dataset. Each training image dataset comprised a total of 100 images of resolution 256 $\times$ 256 pixels. The images of the training data set were randomly selected. The number of learning iterations was 1000 epochs, and the learning rate of CycleGAN was set to 0.0002 first until the 900th epoch and linearly decayed to 0 from the 901st to 1000th epochs. Moreover, the λ values used for the training were set to 2, 10, and 20. The original face images for the training data were preprocessed

(details are described in Sec. 4.2.1). The face images except for the pareidolia-elements were preprocessed in the training phase. Even if the input image used for the generation experiment did not contain elements that cause pareidolia, the effect of the output image seems to be considered slight.

5.2 Evaluation experiment

This experiment aims to evaluate the generated stimuli. In this experiment, the detection number of the intended pareidolia and its pareidolia-inducing power are investigated.

5.2.1 Evaluation procedure

This thesis verifies whether the generated visual stimuli can be generated systematically. The generated stimuli were evaluated to confirm their systematic generation. For qualitative evaluation, the participants gave a score the generated images based on all the λ values and preprocessings. This thesis evaluated face strength as face-pareidolia-inducing power. First, this thesis randomly selected 15 images that are not used for training from the CelebAMask-HQ data set as the input images. Then, four different stimuli from each input image, which depend on including/excluding the pareidolia-elements or applying the blur/noise preprocessing. In

total, a total of 60 images with respect to a value of λ are generated. Finally, a total of 180 image stimuli are evaluated because the thesis investigates three different values of λ. Additionally, this thesis extracted 15 images each from a face and natural images without using them for the training of the image-to-image style translation. Finally, a total of 210 image stimuli were evaluated in the experiment on each participant.

The procedure of the designed experiment is described as follows. First, the participant answers whether the face is contained in the presented stimuli. If the participant answered "yes", the participant selects a face region by enclosing it in an ellipse. Thereafter, the participant gives a score the face-intensity, which ranged from 1 to 99. In this thesis, I instructed to score on a 99 scale to intend the actual index. The participant answers "no" or after the scoring, the presented image is changed to the next image, and the evaluation is iterated for all the image stimuli. The experiment used a 21.5-in monitor, BenQ, G2222HDL. This thesis analyzed the result based on the evaluation by each participant.

5.2.2 Participants

There are 11 participants (age, 21-24, 2 females) who cooperated with the evaluation experiment.

All participants consented to the experiment and signed the relevant agreement.

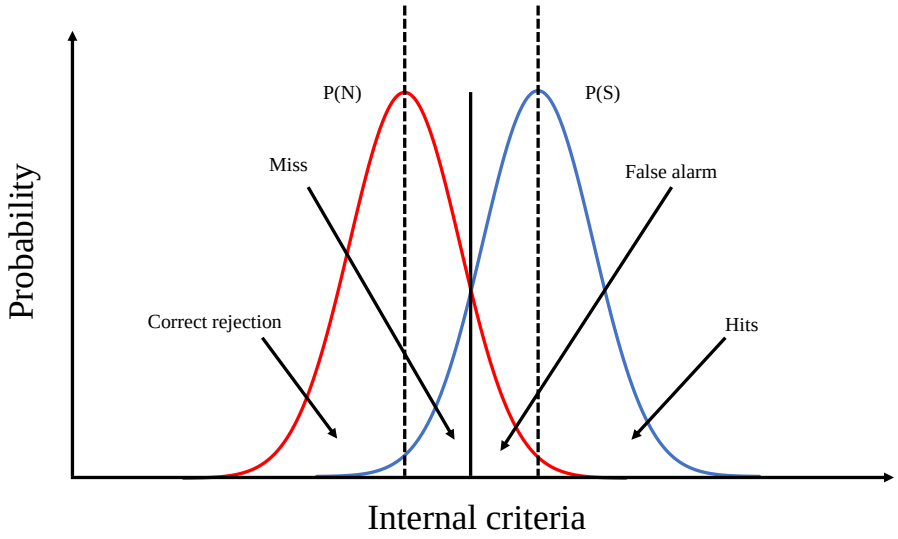


Figure 5.1: Signal detection theory model.

Also, the all participants are healthy control.

5.2.3 Data analysis

This thesis uses the signal detection theory (SDT) [89] for the analysis as shown in Fig. 5.1. An effective method for investigating whether the presented stimuli include specific information is SDT. In this thesis, the specific information is set to the face, and regarded as a "signal". Because of the exclusion of specific information, the natural image is regarded as a "no-signal". Additionally, in this thesis, when the case of inclusion of face pareidolia-elements, such case is regarded as a "signal". Invertly, when the case of exclusion is regarded as a "no-signal". In the SDT, the response can be classified into four classes: miss (M), hit (H), false alarm (FA), and correct rejection (CR). When the participant reported "yes" by observing

		Stimulus	
Response		Signal	Noise
	Yes	Hit	False alarm
	No	Miss	Correct rejection

Figure 5.2: Response classification of confusion matrix on signal detection theory. The response of each participant is classified based on the response and signal.

a stimulus inclusion of a signal, the response is classified as H, and "no" is classified as M. Similarly, if the participant reported "yes" when there is a no-signal in the present stimulus, the response is classified as FA, and "no" is classified as CR. In SDT, the ability to discriminate the stimuli is parameterized as d', and calculated by Eq 5.1:

$$d' = Z(1 - P(FA)) - Z(1 - P(H)). \tag{5.1}$$

$$P(H) = \frac{Hit}{Hit + Miss} \tag{5.2}$$

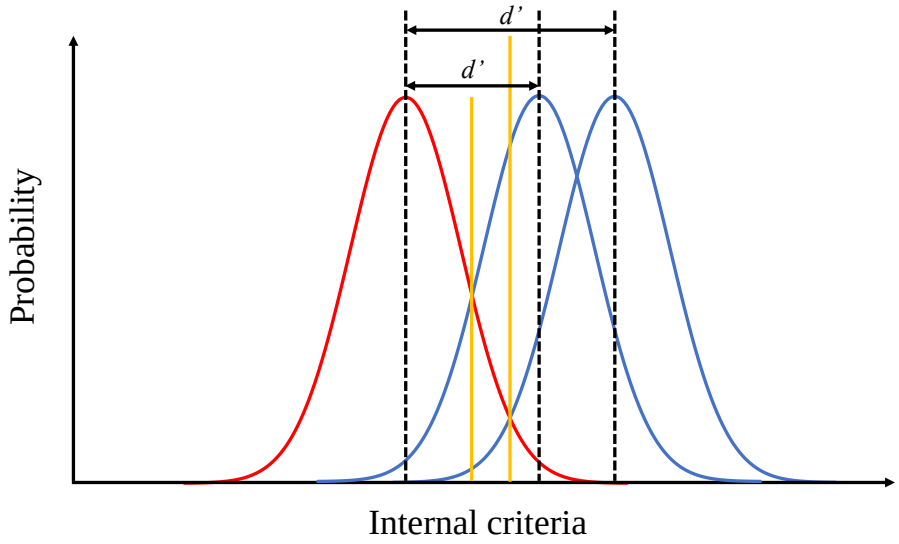


Figure 5.3: Difference of d'. The bigger the distance between red and blue probability distribution, the larger the value of d'.

$$P(FA) = \frac{FA}{FA + CR}. \tag{5.3}$$

Herein, the z-score of each probability is calculated by $Z()$ in Eq. 5.1. In addition, the probabilities of H and FA are represented by $P(H)$ and $P(FA)$ in Eqs. 5.2 and 5.3. The numbers of hit and miss can calculate the hit probability. In the same way, the numbers of false alarm and correct rejection can calculate the false alarm probability. The response time, the center of the ellipse, major axis, and minor axis are recorded. The ellipse that the participant pointed out can be reproduced using these information. Following the above SDT, each of the responses of participants is classified into four types: M, H, FA, and CR. This thesis determined whether the image was intended for face pareidolia based on the recorded ellipse information,

The mask image corresponding to the original face image to which the movement process was applied was moved by the same amount as in the preprocessing movement method. Then, the degree of overlap with the ellipse was calculated. The face pareidolia was considered acceptable when the calculated degree of overlap exceeded 90%, and the presented stimulus was treated as H. On the other hand, if the calculated degree of overlap fell short of 90% and the presented stimulus included pareidolia elements, the presented stimulus was treated as FA.

Chapter6 Results

6.1 Generation result

Figures 6.1, 6.2, 6.3, and 6.4 show the result examples of the generation experiments. The face pareidolia structure is embedded in the generated stimuli that can be perceived as faces. The generated face pareidolia stimuli can be observed to be the same shape as the face annotation data. When the input image includes face pareidolia-elements (blur preprocess: Figs. 6.1a–6.1d and noise preprocess: 6.3a–6.3d), the generated stimuli include the face pareidolia-elements. Inversely, When the input image does not include face pareidolia-elements (blur preprocess: Figs. 6.2a–6.2d and noise preprocess: 6.4a–6.4d), the generated stimuli do not include the face pareidolia-elements.

6.2 Evaluation experiment result

Pareidolia is a psychological tendency; however, the sensitivity to whether the stimuli can be perceived as a face differs between individuals. The hit transition on

(a) Input image (b) $\lambda = 2$

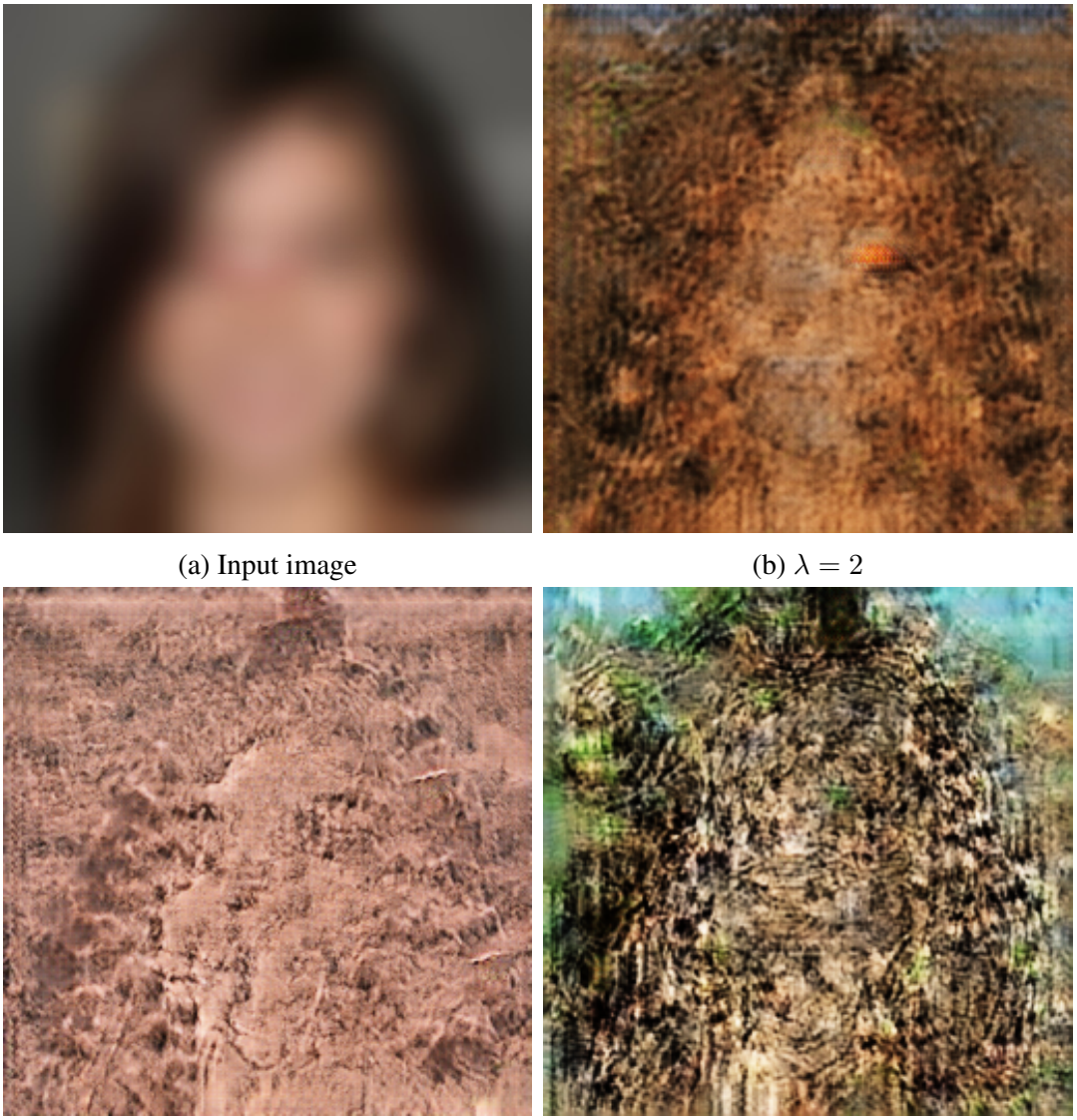
(c) $\lambda = 10$ (d) $\lambda = 20$

Figure 6.1: Generated images (blur, including the face pareidolia-elements)

each preprocessing is shown in Fig. 6.5 and Fig. 6.6.

the result in Fig. 6.5, 7 of the 11 participants increase monotonically the reported number as the value of λ on blur preprocessing. On the other hand, only 1 of the 11 participants increase monotonically the reported number as the value of λ on noise preprocessing.

In addition, when the translated image remains the face pareidolia-elements fea-



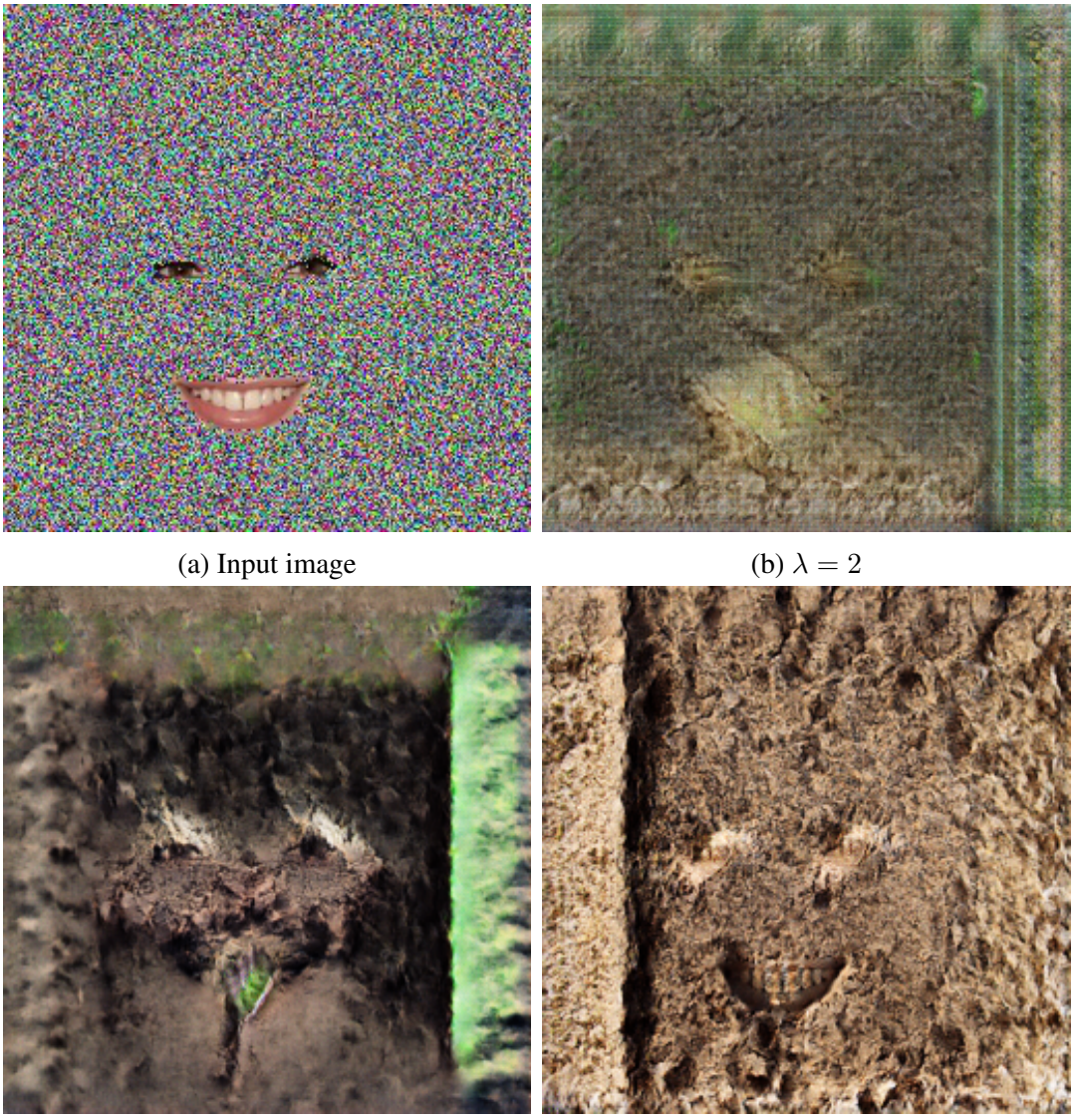
(a) Input image (b) $\lambda = 2$

(c) $\lambda = 10$ (d) $\lambda = 20$

Figure 6.2: Generated images (blur, excluding the face pareidolia-elements)

tures, these elements are perceived as the face. It reveals that the image translated to the natural image can generate pareidolia-inducing stimuli using the eyes and mouth features.

One participant could not report the real face. The cause seems the operation mistake.



(a) Input image (b) $\lambda = 2$

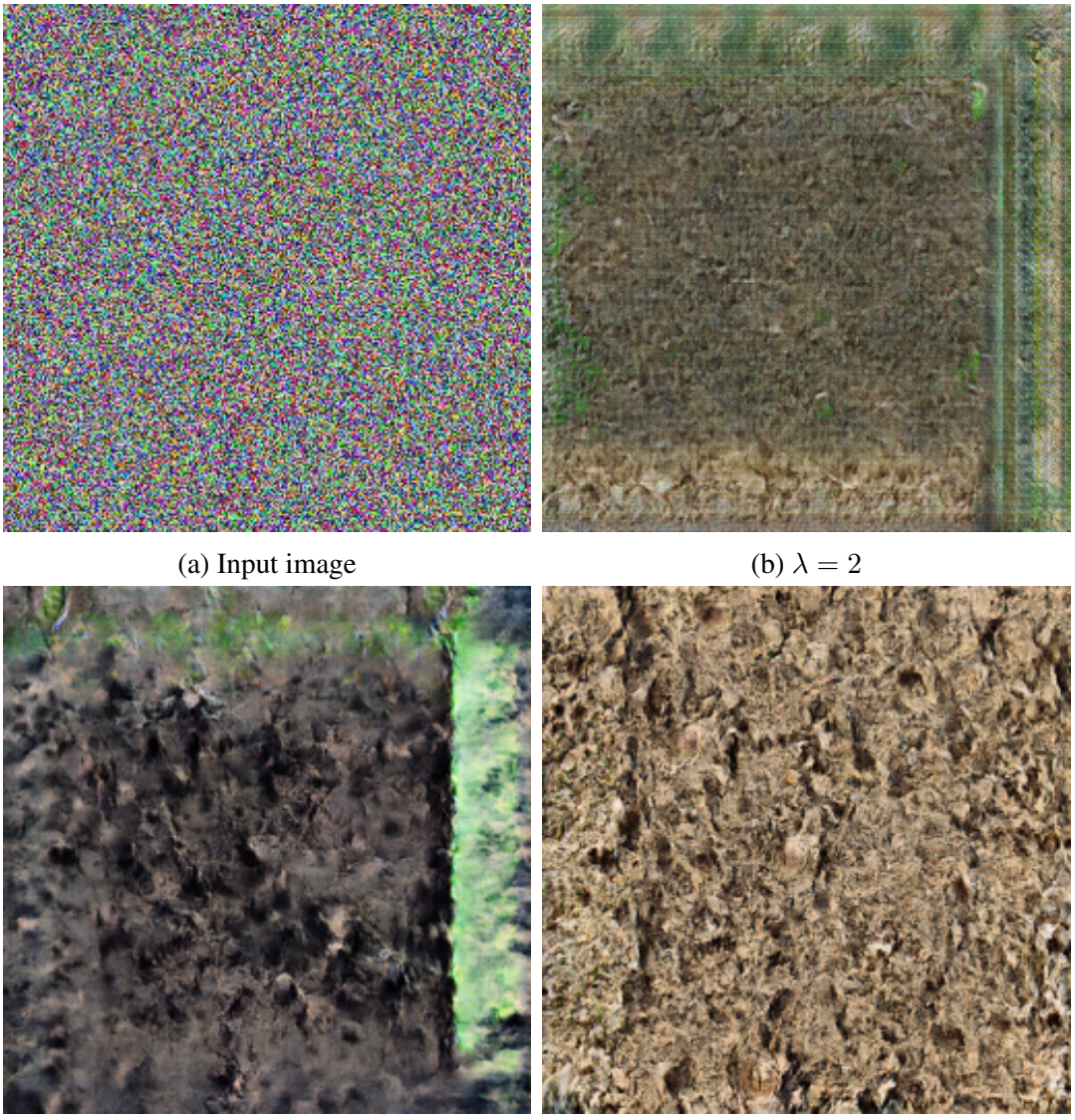
(c) $\lambda = 10$ (d) $\lambda = 20$

Figure 6.3: Generated images (noise, including the face pareidolia-elements)

6.3 Receiver operating characteristic curve

In this experiment, I instruct to find out the face. Therefore, the participant may be biased to find the face from the stimuli.

In the evaluation experiment result, “face” on the not-intended face pareidolia stimuli is sometimes reported by almost all participants. Two possible causes of factors are considered. The low ability to discriminate whether the participant can



(a) Input image (b) $\lambda = 2$

(c) $\lambda = 10$ (d) $\lambda = 20$

Figure 6.4: Generated images (noise, excluding the face pareidolia-elements)

perceive the face is the one of the reasons. This task itself is too difficult to discriminate objectively, and the d' becomes low. The abnormal setting of the internal criterion; the participants tend to set the threshold to report "face" too low, is considered another reason. To reveal the reason for pareidolia, the d' is calculated and the receiver operating characteristic (ROC) curve based on the intensity is drawn. As the result, Fig. 6.7 shows the drawn ROC curve is shown. If the participant

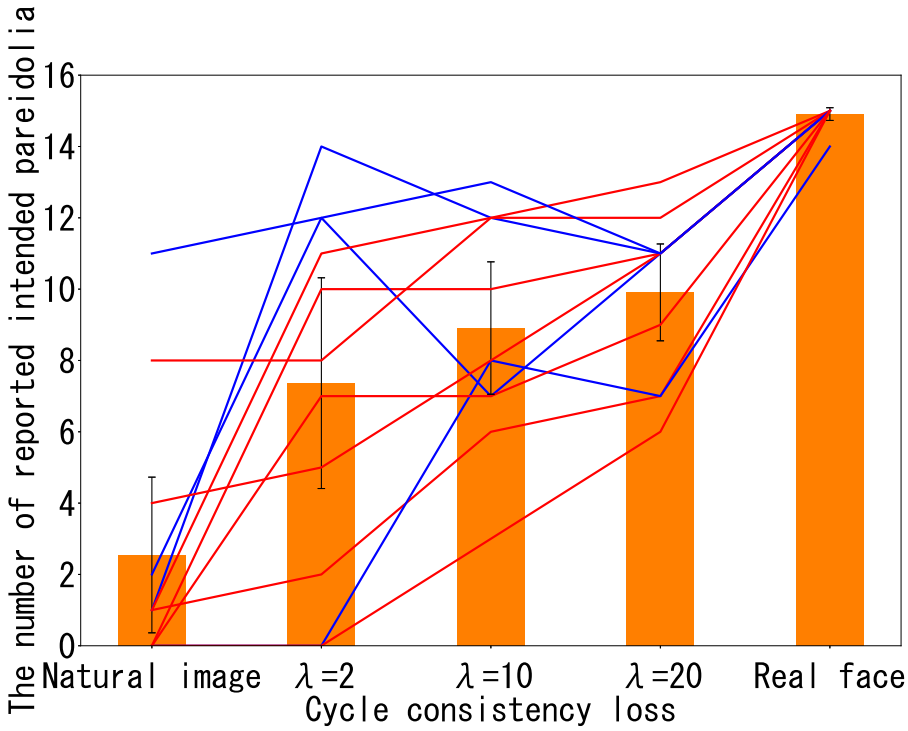


Figure 6.5: Reported number transition when blur preprocessing. The red and blue lines represent the monotonic and nonmonotonic increase of the reported number, respectively. The orange bar chart represents the average of the reported number of participants. The error bar represents 95% confidence interval of the average of the reported number.

reports "face" on the intended face pareidolia stimuli perfectly without FA, the d' is "inf". The line of the ROC curve cannot be drawn then the d' is "inf". The ability to discriminate can be observed by all participants from the ROC curve and d'. Therefore, the abnormal internal criterion for the face signal is suggested to be the pareidolia cause. Herein, Wardle reported that face pareidolia is firstly perceived as more similar to the face in the brain than consisted objects [90]. Therefrom, Wardle suggests the factors of face pareidolia is its sensitivity than its selectivity. This ROC results suggest that the factor of face pareidolia is the abnormal internal criterion for face signal; therefore, the factor of face pareidolia can be the sensitivity for the

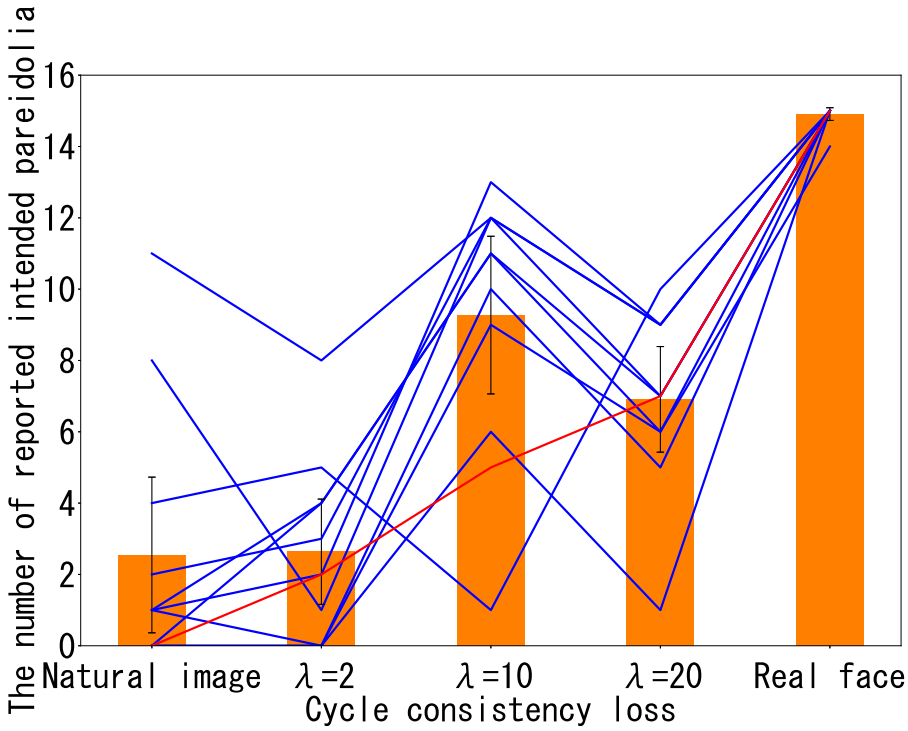


Figure 6.6: Reported number transition when noise preprocessing. The red and blue lines represent the monotonic and nonmonotonic increase of the reported number, respectively. The orange bar chart represents the average of the reported number of participants. The error bar represents 95% confidence interval of the average of the reported number.

face signal.

6.4 Data analysis result

The face sensitivity of each individuals might be different. For real face stimuli, almost all participants evaluated a 99 score. On the other hand, there are the stimuli that cannot be perceived as face. Because this thesis focuses on the perceived case (can be induced pareidolia), the stimuli that cannot be perceived as face are not considered. Also, the participant gave a different minimum score. To normalize

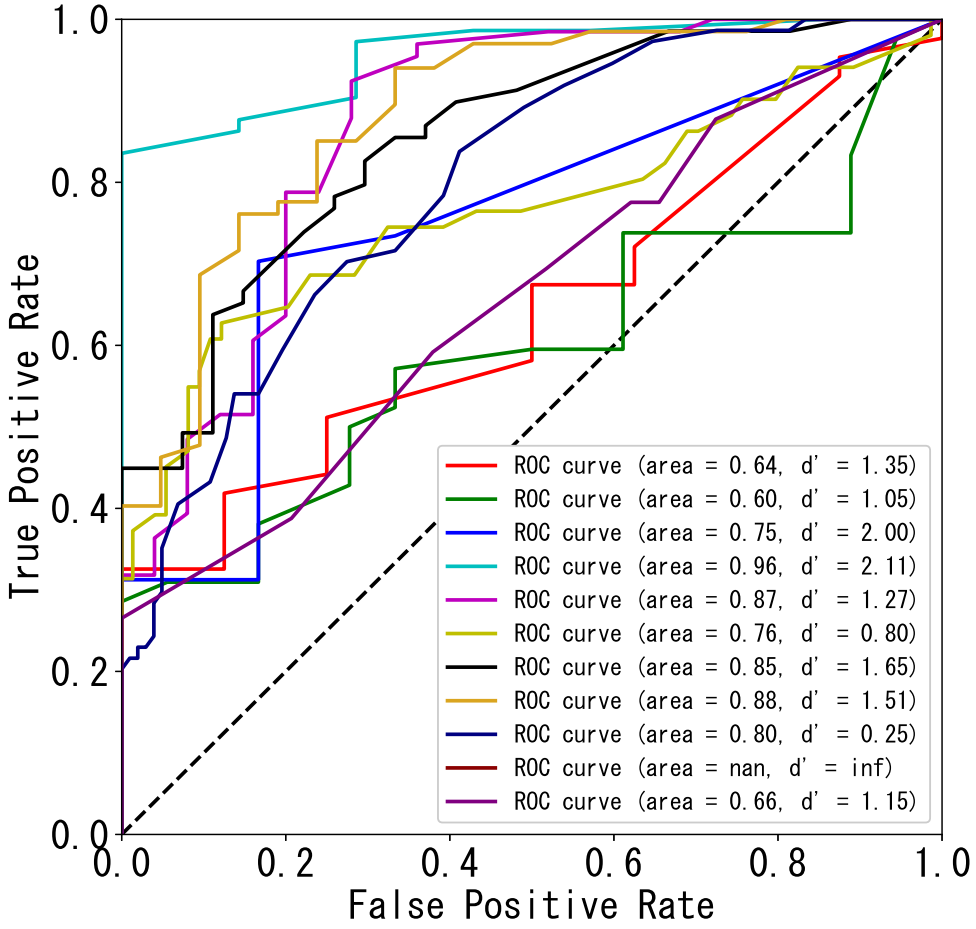


Figure 6.7: ROC curve for each participant.

these sensitivity, min–max normalization based on each minimum value to normalize the score range was applied. Eq. 6.1 represents the formula for processing the minimum value to 1 and maximum value to 99 after normalization

$$\text{Intensity} = \frac{\text{Score} - \text{Intensity}_{\text{min}}}{\text{Intensity}_{\text{max}} - \text{Intensity}_{\text{min}}} \times 98 + 1. \tag{6.1}$$

Here, $Intensity_{min}$ is the minimum value of each participant and $Intensity_{max}$ is the maximum value of each participant. The stimuli evaluation result after the

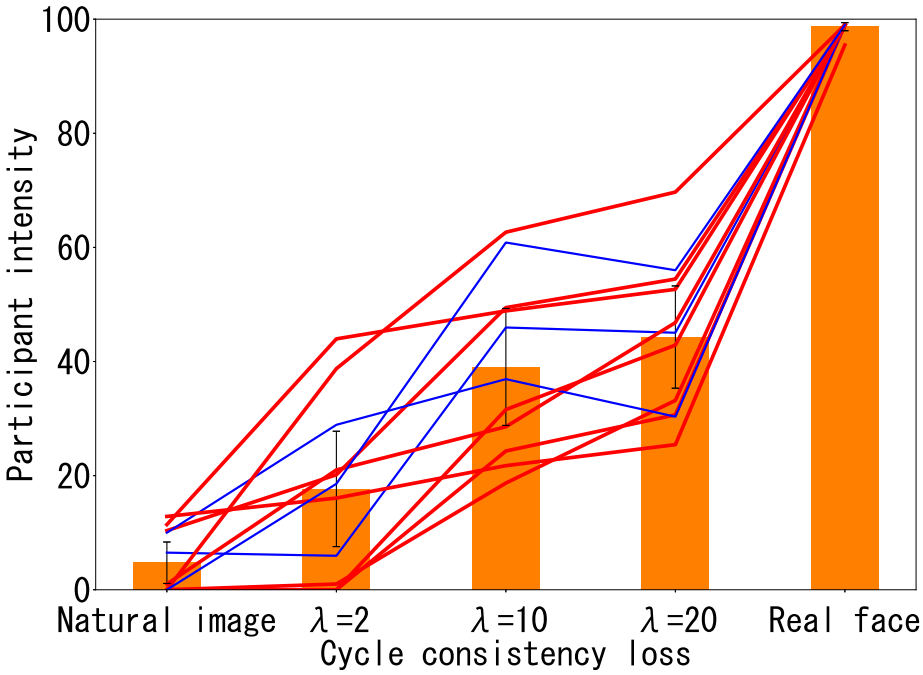


Figure 6.8: The intensity when the blur preprocessing. The red bold and blue lines represent the monotonic and nonmonotonic increase in participant intensity, respectively. The orange bar chart represents the average of the intensities of all participants. The error bar is 95% confidence interval of the average intensity.

min-max normalization process is shown in Figs. 6.8 and 6.9.

The red bold line in Fig. 6.8 and Fig. 6.9 represents the intensity of each participant monotonically increases with respect to the increase of λ. As shown in Fig. 6.8, 8 of 11 participants increase monotonically when blurring is applied as the preprocessing for stimuli generation. On the other hand, all subjects are nonmonotonically increasing when the noise process . In addition, the correlation coefficient between the λ value and the average participant ' s intensity are investigated with respect to each preprocessing method based on Fig. 6.8 and Fig. 6.9, respectively. The correlation coefficient is 0.92 when the preprocessing is blurring. On the other

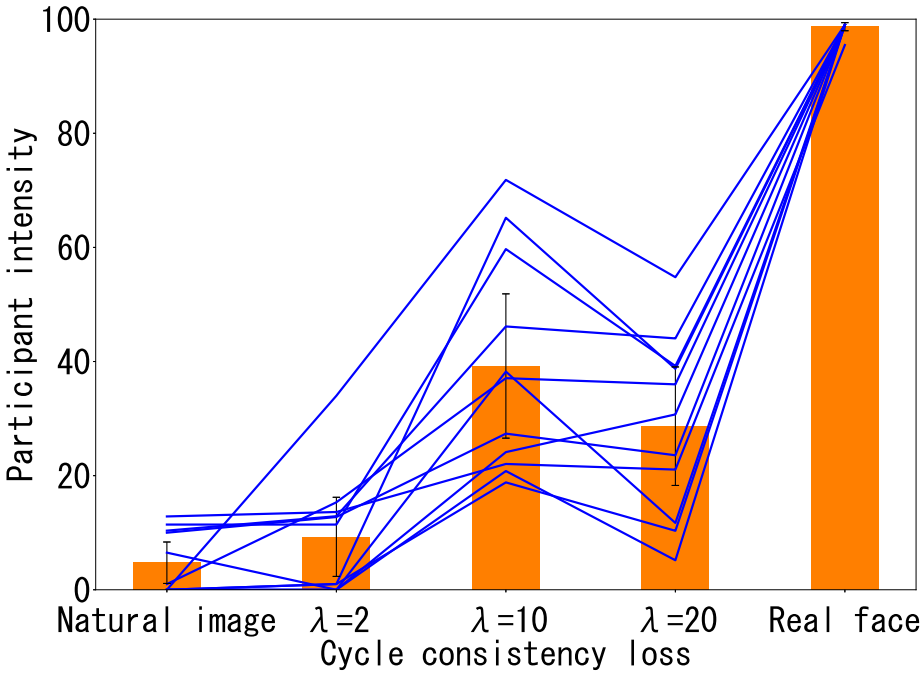


Figure 6.9: The intensity when the noise preprocessing. The blue lines represent the nonmonotonic increase in participant intensity. The orange bar chart represents the average of the intensities of all participants. The error bar represents 95% confidence interval of the average intensity.

hand, the correlation coefficient degrades to 0.59 when the preprocessing is noise. From this result, the weight (λ) was associated with the pareidolia-inducing power when the preprocess was blurring. In each category and the weight parameter, the tendency of the scores was different from that of the faces; thus, suggesting that the stimuli were generated with different characteristics from those of the faces.

In addition, Figures 6.5 and 6.8 can be observed the same intensity trend, which is the same as that in Figs. 6.6 and 6.9. The correlation coefficient between the reported pareidolia number and the average participant ' s intensity of the generated stimuli is 0.98 when the preprocessing is blur preprocessing. Furthermore, the

correlation coefficient is 1.00 when the preprocessing is noise preprocessing. From these observations, the face pareidolia report number and the pareidolia-inducing power might correlate.

6.5 Significant difference test

In Fig. 6.8, I confirmed that the correlation between pareidolia-inducing power and λ by the blur intensity transition. However, we cannot intuitively observe whether there is a significant difference between $\lambda = 10$ and $\lambda = 20$ in Fig. 6.8. Therefore, this thesis investigates whether there is a significant different using statistical significant test. Hence, the difference is investigated whether it is significant by the Wilcoxon signed-rank test [91]. The Wilcoxon can test whether there is a significant difference between the representative value of two categories obtained from the same participant.

6.5.1 Procedure

The statistics calculation of Wilcoxon signed-rank test vary depending on the data number. Procedure of Wilcoxon signed-rank test is described as follows if the data numbers is less than 25.

Step 1: Calculation of the difference between two types of categories

Step 2: Taking the absolute value of the difference values

Step 3: Rank the absolute value in descending order (If there are above two same values, the average rank is allocated)

step 4: Calculation of sum of the ranked value negative and positive, respectively

step 5: Comparison of the both sum values and the minimum value is the statistics (represented as T)

step 6: If the statistics is less than the critical value of signed-rank test, there is a statistical difference between the average of 2 types categories

Otherwise the data number (more than 25), the statistics (z) is calculated as follows formulation:

$$z = \frac{\left| T - \frac{N(N+1)}{4} \right|}{\sqrt{\left(\frac{N(N+1)(2N+1)}{24} \right)}} \tag{6.2}$$

6.5.2 Significant difference test result

In this thesis, because the number of participants is 11, the statistics is calculated in Eq. 6.2 described in Sec. 6.5.1. As the result of statistical testing, the statistics (T) is 12. This statistic is less than when the significant standard is set to 10 % (two-tailed test). From this result, there is a marginally significant between $\lambda = 10$ and $\lambda = 20$ when the preprocessing is applied blurring processing.

Chapter7 Conclusion

Herein, This thesis investigated the systematic generation of face pareidolia stimuli through the pareidolia-inducing power. In this thesis, the weight of the cycle-consistency loss was manipulated and the pareidolia-inducing stimuli were generated. The evaluation experiment revealed the pareidolia-inducing power correlated with the weight parameter of the cycle-consistency loss of CycleGAN when the blurring process was applied excluding the face pareidolia-elements (the eyes and mouth) as preprocessing. On the other hand, there is not a correlation between cycle-consistency loss and pareidolia-inducing power when the noise process was applied excluding the face pareidolia-elements (the eyes and mouth) as preprocessing. I trained the face image and natural image styles using the CycleGAN, and systematically generated face pareidolia stimuli. In particular, the generated stimuli using the human eyes and mouth can induce face perception. Therefore, the generated stimuli can induce face pareidolia. Also, the ROC curve result shows that the cause of face pareidolia is mostly due to abnormal subjective internal criterion. In addition, there is a marginally significant (significant standard = 10%) between $\lambda = 10$ and $\lambda = 20$ in Fig. 6.8. Furthermore, the generated stimuli using face attri-

bution related to the pareidolia can induce pareidolia. This suggests that preserving features such as the eyes and mouth is critical for inducing face pareidolia.

Chapter8 Future Work

In future work, there are many tasks to tackle.

First, this research expands to the other pareidolia-inducing stimuli. To manipulate the pareidolia-inducing power of existing pareidolia stimuli, the approach of this thesis aims to be applied. Furthermore, it can also be applied to more versatile applications by systematic generation in the form of existing pareidolia stimuli, except for face categories. If the already existing stimuli can be manipulated its pareidolia-inducing power, the research regarding pareidolia will be widened such as the perception of the person, and training for artificial intelligence. In addition, the opposite tendency of face pareidolia, prosopagnosia, is known. If the stimulus of pareidolia-inducing power is high, some people cannot perceive the face. In such a case, the people are suggested to be prosopagnosia. In this thesis, the scope of the investigation target is healthy control subjects. In future work, the pareidolia perception of patients suffering from not only Lewy body dementia but also prosopagnosia [92] as shown in Fig. 8.1 must be investigated the correlation

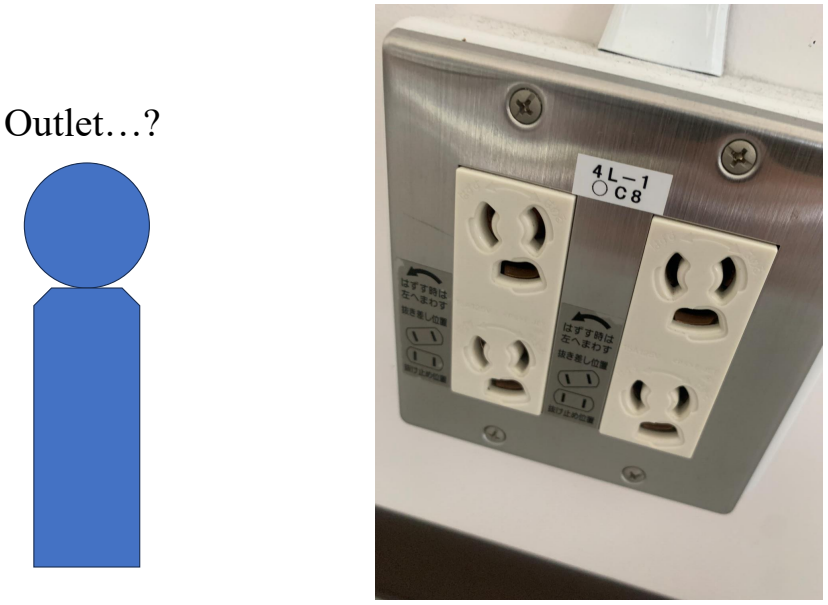


Figure 8.1: Prosopagnosia image. The person of prosopagnosia cannot perceive the face.

between pareidolia-inducing power and face pareidolia-reported number.

Also, I attempt to evaluate the generated stimuli one of each; however, this experiment is feared to take the participant a long time to measure the stimuli. In the future, the experiment procedure and method need to be improved.

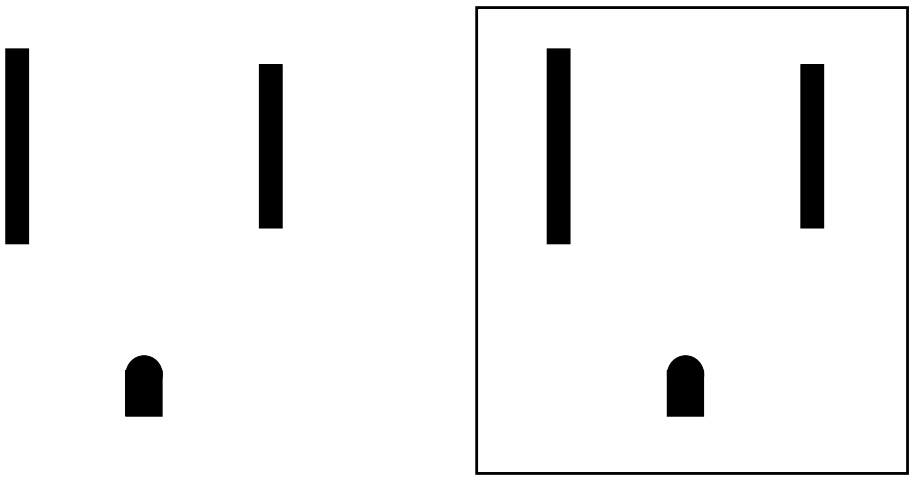
Furthermore, in this thesis, the scaling of pareidolia-inducing power is set to 99. It could be difficult to score or the small number of participants could affect the results.

In addition, Regarding the statistical significant test, the significant standard is set to 10%. In general, the significant standard of statistical significant test is set to 5%. Therefore, I aim to satisfy the standard from two approaches. The first one is the increasing of the participant. One of the participants might affect the statistical testing because of few participants. In association with the previous sentence,

the second one is the investigation of the pareidolia factor. The frequency of face pareidolia varies in some factors such as specific illness, gender, and so on. In this experiment, I did not investigate whether the participants have the factors related to the face pareidolia frequency; therefore, the experiment result might be affected. In the future, this experiment can investigate more factors of face pareidolia frequency.

Also, the studies that investigate face pareidolia have the majority of one type category. The pareidolia-inducing power is different by the pareidolia type such as the pareidolia which we can experience in everyday life, the stimuli of the noise pareidolia test, and so on. Because we can perceive these stimuli as the face, I need to investigate the tendency of each pareidolia category.

In addition, the face pareidolia-elements might relate to its contour. The pareidolia-inducing stimuli are different from each other; however, as shown in Fig. 8.2 the stimuli include the human-like contour because of the object scale, texture, design, and so on. This thesis investigates whether the face pareidolia-elements can be induced the pareidolia. As a further investigation, I need to investigate the contour and inducing face pareidolia effect. For example, we can respond more rapidly to the contour of face pareidolia-inducing stimuli than the face pareidolia-elements. Also, when we try to find the pareidolia-inducing stimuli, we might narrow down the range unconsciously because of the contour. Furthermore, there are some tasks to analyze the experiment result. First, the tendency to react to the stimulus whether



(a) Pareidolia without contour (b) Pareidolia with contour

Figure 8.2: One of the pareidolia image (outlet).

early or late might be reflected in the pareidolia-inducing power. Second, when the participant perceives the face not intended face pareidolia, it is realistic face pareidolia. It might contribute to the face pareidolia study by analysis of the natural occasion of face pareidolia. Such these tasks, the result is possible to obtain more insightful.

Bibliography

[1] Web gallery of art, searchable fine arts image database. https://www.wga.hu/frames-e.html?/html/a/arcimbol/3allegor/2seaso1.html. (Accessed on 25/10/2023).

[2] Marina A. Pavlova, Klaus Scheffler, and Alexander N. Sokolov. Face-n-food: Gender differences in tuning to faces. *PLOS ONE*, 10(7):1–12, 07 2015.

[3] Masaharu Kato and Ryoko Mugitani. Pareidolia in infants. *PLOS ONE*, 10(2):1–9, 02 2015.

[4] Jiangang Liu, Jun Li, Lu Feng, Ling Li, Jie Tian, and Kang Lee. Seeing jesus in toast: Neural and behavioral correlates of face pareidolia. *Cortex*, 53:60–77, 2014.

[5] Psychological image collection at stirling (pics). https://pics.stir.ac.uk/. (Accessed on 25/10/2023).

[6] Linsen Song, Wayne Wu, Chaoyou Fu, Chen Qian, Chen Change Loy, and Ran He. Pareidolia face reenactment. In *IEEE Conference on Computer Vision and*

Pattern Recognition (CVPR), 2021.

[7] Alexander Mordvintsev, Christopher Olah, and Mike Tyka. Inceptionism: Going deeper into neural networks. https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html, 2015.

[8] Keisuke Suzuki, Warrick Roseboom, David J Schwartzman, and Anil K Seth. A deep-dream virtual reality platform for studying altered perceptual phenomenology. *Scientific reports*, 7(1):1–11, 2017.

[9] Faces in places. https://www.flickr.com/groups/facesinplaces/. (Accessed on 31/10/2023).

[10] Faces in places (blog type). http://facesinplaces.blogspot.com/. (Accessed on 01/11/2023).

[11] Flickr. https://www.flickr.com/. (Accessed on 21/09/2023).

[12] Psychological image collection at stirling (pics). https://www.amazon.co.uk/exec/obidos/ASIN/1906672903/facinpla-21. (Accessed on 25/10/2023).

[13] N Kanwisher. Domain specificity in face perception. *Nat Neurosci*, 3(8):759–763, August 2000.

[14] Zuzana Walker, Katherine L Possin, Bradley F Boeve, and Dag Aarsland. Lewy body dementias. *The Lancet*, 386(10004):1683–1697, 2015.

[15] Yasuyuki Mamiya, Yoshiyuki Nishio, Hiroyuki Watanabe, Kayoko Yokoi, Makoto Uchiyama, Toru Baba, Osamu Iizuka, Shigenori Kanno, Naoto Kamimura, Hiroaki Kazui, Mamoru Hashimoto, Manabu Ikeda, Chieko Takeshita, Tatsuo Shimomura, and Etsuro Mori. The pareidolia test: A simple neuropsychological test measuring visual hallucination-like illusions. *PLOS ONE*, 11(5):1–13, 05 2016.

[16] Pareidolia test (in japanese). `http://www.bncn.med.tohoku.ac.jp/custom16.html`. (Accessed on 07/11/2023).

[17] Marina Pavlova, Michele Guerreschi, Lucia Tagliavento, Filippo Gitti, Alexander Sokolov, Andreas Fallgatter, and Elisa Fazzi. Social cognition in autism: Face tuning. *Scientific Reports*, 7, 05 2017.

[18] Megumi Kobayashi, Yumiko Otsuka, Emi Nakato, So Kanazawa, Masami K. Yamaguchi, and Ryusuke Kakigi. Do infants recognize the arcimboldo images as faces? behavioral and near-infrared spectroscopic study. *Journal of Experimental Child Psychology*, 111(1):22–36, 2012.

[19] Craig M. Mooney. Age in the development of closure ability in children. *Canadian journal of psychology*, 11 4:219–26, 1957.

[20] Annalisa Palmisano, Giulio Chiarantoni, Francesco Bossi, Alessio Conti, Vitiana D ' Elia, Serena Tagliente, Michael Nitsche, and Davide Rivolta. Face pareidolia is enhanced by 40 hz transcranial alternating current stimulation (tacs) of the face perception network. *Scientific Reports*, 13, 02 2023.

[21] Kayoko Yokoi, Yoshiyuki Nishio, Makoto Uchiyama, Tatsuo Shimomura, Osamu Iizuka, and Etsuro Mori. Hallucinators find meaning in noises: Pareidolic illusions in dementia with lewy bodies. *Neuropsychologia*, 56:245–254, 2014.

[22] Makoto Uchiyama, Yoshiyuki Nishio, Kayoko Yokoi, Kazumi Hirayama, Toru Imamura, Tatsuo Shimomura, and Etsuro Mori. Pareidolias: complex visual illusions in dementia with lewy bodies. *Brain*, 135(8):2458–2469, May 2012.

[23] Diane Rekow, Jean-Yves Baudouin, Renaud Brochard, Bruno Rossion, and Arnaud Leleu. Rapid neural categorization of facelike objects predicts the perceptual awareness of a face (face pareidolia). *Cognition*, 222:105016, 2022.

[24] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates, Inc., 2014.

[25] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.

[26] Colin J. Palmer and Colin W. G. Clifford. Face pareidolia recruits mechanisms for detecting human social attention. *Psychological Science*, 31(8):1001–1012, 2020. PMID: 32697673.

[27] Jessica Taubert, Susan G. Wardle, Molly Flessert, David A. Leopold, and Leslie G. Ungerleider. Face pareidolia in the rhesus monkey. *Current Biology*, 27(16):2505–2509.e2, 2017.

[28] Michael A. Nees and Charlotte Phillips. Auditory pareidolia: Effects of contextual priming on perceptions of purportedly paranormal and ambiguous auditory stimuli. *Applied Cognitive Psychology*, 29(1):129–134, 2015.

[29] Jess M. Williams, Michelle Carr, and Mark Blagrove. Sensory processing sensitivity: Associations with the detection of real degraded stimuli, and reporting of illusory stimuli and paranormal experiences. *Personality and Individual Differences*, 177:110807, 2021.

[30] Albert Mehrabian et al. *Silent messages*, volume 8. Wadsworth Belmont, CA, 1971.

[31] Marina A. Pavlova, Jessica Galli, Federica Pagani, Serena Micheletti, Michele Guerreschi, Alexander N. Sokolov, Andreas J. Fallgatter, and Elisa M. Fazzi. Social cognition in down syndrome: Face tuning in face-like non-face images. *Frontiers in Psychology*, 9, 2018.

[32] Marina A. Pavlova, Julie Heiz, Alexander N. Sokolov, and Koviljka Barisnikov. Social cognition in williams syndrome: Face tuning. *Frontiers in Psychology*, 7, 2016.

[33] Marina A. Pavlova, Julie Heiz, Alexander N. Sokolov, Andreas J. Fallgatter, and Koviljka Barisnikov. Even subtle cultural differences affect face tuning. *PLOS ONE*, 13(6):1–13, 06 2018.

[34] Marina A. Pavlova, Annika Mayer, Franziska Hösl, and Alexander N. Sokolov. Faces on her and his mind: Female and likable. *PLOS ONE*, 11(6):1–13, 06 2016.

[35] Rebecca Rolf, Alexander N. Sokolov, Tim W. Rattay, Andreas J. Fallgatter, and Marina A. Pavlova. Face pareidolia in schizophrenia. *Schizophrenia Research*, 218:138–145, 2020.

[36] Marina A Pavlova, Jessica Galli, Federica Zanetti, Federica Pagani, Serena Micheletti, Andrea Rossi, Alexander N Sokolov, Andreas J Fallgatter, and

Elisa M Fazzi. Social cognition in individuals born preterm. *Scientific Reports*, 11(1):14448, 2021.

[37] Tsung-Wei Ke, Stella X. Yu, and David Whitney. Mooney face classification and prediction by learning across tone. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 2025–2029, 2017.

[38] Caspar M. Schwiedrzik, Lucia Melloni, and Aaron Schurger. Mooney face stimuli for visual perception research. *PLOS ONE*, 13(7):1–11, 07 2018.

[39] Gulsum Akdeniz, Sila Toker, and Ibrahim Atli. Neural mechanisms underlying visual pareidolia processing: An fMRI study. *Pak J Med Sci*, 34(6):1560–1566, November 2018.

[40] Michael Scheel Robert M. Stowe Giuseppe Iaria, Christopher J. Fox and Jason J. S. Barton. A case of persistent visual hallucinations of faces following lsd abuse: A functional magnetic resonance imaging study. *Neurocase*, 16(2):106–118, 2010. PMID: 19927262.

[41] Gülsüm Akdeniz. Brain activity underlying face and face pareidolia processing: an ERP study. *Neurological Sciences*, 41(6):1557–1565, June 2020.

[42] Simone Kühn, Timothy R. Brick, Barbara C. N. Müller, and Jürgen Gallinat. Is this car looking at you? how anthropomorphism predicts fusiform face area activation when seeing cars. *PLOS ONE*, 9(12):1–14, 12 2014.

[43] Shlomo Bentin, Truett Allison, Aina Puce, Erik Perez, and Gregory McCarthy. Electrophysiological studies of face perception in humans. *J Cogn Neurosci*, 8(6):551–565, November 1996.

[44] Makoto Uchiyama, Yoshiyuki Nishio, Kayoko Yokoi, Yoshiyuki Hosokai, Atsushi Takeda, and Etsuro Mori. Pareidolia in parkinson's disease without dementia: A positron emission tomography study. *Parkinsonism & Related Disorders*, 21(6):603–609, 2015.

[45] Taeko Sasai-Sakuma, Yoshiyuki Nishio, Kayoko Yokoi, Etsuro Mori, and Yuichi Inoue. Pareidolias in REM Sleep Behavior Disorder: A Possible Predictive Marker of Lewy Body Diseases? *Sleep*, 40(2):zsw045, 01 2017.

[46] Christian Ryan, Martina Stafford, and Robert King. Brief report: Seeing the man in the moon: Do children with autism perceive pareidolic faces? a pilot study. *Journal of Autism and Developmental Disorders*, 46:3838–3843, 09 2016.

[47] Stephan K. Chalup, Kenny Hong, and Michael J. Ostwald. Simulating pareidolia of faces for architectural image analysis. 2010.

[48] Asad Abbas and Stephan Chalup. Affective analysis of visual scenes using face pareidolia and scene-context. *Neurocomputing*, 437:72–83, 2021.

[49] Andrew Wodehouse, Ross Brisco, Ed Broussard, and Alex Duffy. Pareidolia: characterising facial anthropomorphism and its implications for product design. *Journal of Design Research*, 16(2):83–98, 2018.

[50] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[51] Asad Abbas and Stephan Chalup. From face recognition to facial pareidolia: Analysing hidden neuron activations in cnns for cross-depiction recognition. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2019.

[52] Gianluigi Guido, Marco Pichierri, Giovanni Pino, and Rajan Nataraajan. Effects of face images and face pareidolia on consumers' responses to print advertising. *Journal of Advertising Research*, 59(2):219–231, 2019.

[53] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game

of go with deep neural networks and tree search. *Nature*, 529(7587):484–489, Jan 2016.

[54] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10674–10685, 2022.

[55] Frank Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.

[56] Marvin Minsky and Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry*. The MIT Press, 09 2017.

[57] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323:533–536, 1986.

[58] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural Comput*, 18(7):1527–1554, July 2006.

[59] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551, 1989.

[60] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 11 2013.

[61] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. pages 779–788, 06 2016.

[62] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. Ssd: Single shot multibox detector. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 21–37, Cham, 2016. Springer International Publishing.

[63] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90, may 2017.

[64] K Simonyan and A Zisserman. Very deep convolutional networks for large-scale image recognition. pages 1–14. Computational and Biological Learning Society, 2015.

[65] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Ra-

binovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, 2015.

[66] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2016.

[67] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger. Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269, Los Alamitos, CA, USA, jul 2017. IEEE Computer Society.

[68] Haoxiang Li, Zhe Lin, Xiaohui Shen, Jonathan Brandt, and Gang Hua. A convolutional neural network cascade for face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[69] C. Garcia and M. Delakis. Convolutional face finder: a neural architecture for fast and robust face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11):1408–1423, 2004.

[70] Huaizu Jiang and Erik Learned-Miller. Face detection with the faster r-cnn. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pages 650–657, 2017.

[71] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.

[72] S. Lawrence, C.L. Giles, Ah Chung Tsoi, and A.D. Back. Face recognition: a convolutional neural-network approach. *IEEE Transactions on Neural Networks*, 8(1):98–113, 1997.

[73] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. In Yoshua Bengio and Yann LeCun, editors, *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*, 2016.

[74] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of GANs for improved quality, stability, and variation. In *International Conference on Learning Representations*, 2018.

[75] T. Karras, S. Laine, and T. Aila. A style-based generator architecture for generative adversarial networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 43(12):4217–4228, dec 2021.

[76] Yang Chen, Yu-Kun Lai, and Yong-Jin Liu. Cartoongan: Generative adversarial networks for photo cartoonization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[77] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, September 2018.

[78] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *CVPR*, 2017.

[79] Y. Choi, M. Choi, M. Kim, J. Ha, S. Kim, and J. Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8789–8797, Los Alamitos, CA, USA, jun 2018. IEEE Computer Society.

[80] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision – ECCV 2016*, pages 694–711, Cham, 2016. Springer International Publishing.

[81] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[82] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[83] Shuo Yang, Ping Luo, Chen Change Loy, and Xiaoou Tang. Wider face: A face detection benchmark. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5525–5533, 2016.

[84] Bin Yang, Junjie Yan, Zhen Lei, and Stan Z. Li. Fine-grained evaluation on face detection in the wild. In *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, volume 1, pages

1–7, 2015.

[85] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In *2015 IEEE International Conference on Computer Vision (ICCV)*, pages 3730–3738, 2015.

[86] Gwangbin Bae, Martin de La Gorce, Tadas Baltrušaitis, Charlie Hewitt, Dong Chen, Julien Valentin, Roberto Cipolla, and Jingjing Shen. Digiface-1m: 1 million digital face images for face recognition. In *2023 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2023.

[87] Cheng-Han Lee, Ziwei Liu, Lingyun Wu, and Ping Luo. Maskgan: Towards diverse and interactive facial image manipulation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[88] Yael Omer, Roni Sapir, Yarin Hatuka, and Galit Yovel. What is a face? critical features for face detection. *Perception*, 48(5):437–446, 2019. PMID: 30939991.

[89] David Marvin Green, John A Swets, et al. *Signal detection theory and psychophysics*, volume 1. Wiley New York, 1966.

[90] Susan G Wardle, Jessica Taubert, Lina Teichmann, and Chris I Baker. Rapid and dynamic processing of face pareidolia in the human brain. *Nature communications*, 11(1):4518, 2020.

[91] Frank. Wilcoxon. Individual comparisons by ranking methods. *Biometrics*, 1:196–202, 1945.

[92] B. Bornstein and D. P. Kidron. Prosopagnosia. *Journal of Neurology, Neurosurgery & Psychiatry*, 22(2):124–131, 1959.