2019 2nd International Conference on Mechanical, Electronic and Engineering Technology (MEET 2019)

A Behavioral Study of Improved Style Transfer Network with Various Textural Styles Applied to a Female Human Face

Yiting Li

Yorba Linda High School, Yorba Linda, CA, USA 92886

E-mail: yyyitingli@gmail.com

Keywords: Improving the Neural Algo-rithm of Artistic Style, a neural network study

Abstract: Art is a required class in many countries. It helps people learn and construct ideas more visually. Recognizing patterns and textures is a skill that challenges even many art teachers when instructing their students. Based on a neural network study called Improving the Neural Algorithm of Artistic Style, the idea of artistic transfer portrays the future that computers can be used to ameliorate the hardship some teachers find. In this research, we specifically use four patterns that have distinct styles and illustrate some basic rules on how to apply these style images. We also examine the styling behavior when using grayscale content images.

1. Introduction

Traditional art education has proven to be a vital component in our everyday lives. Although it can sometimes be overlooked by many, art education is crucial not only for art students but for everyone. According to The Masters of Traditional Art Education Guide[56], the aesthetics of art not only changes people's perception of beauty and merit, but also helps students categorize information and organize materials. Students are able to do this in the form of data tables and charts. The ability to identify different textures from a style of art and apply textures to pictures, then, becomes a really important form of art because it is also the most basic form of training. However, this skill may not come easy to everyone. Very few people know how to complete it confidently and creatively. Even though it is important, mastering this kind of art requires persistent practice and dedication. Application of texture can be more than just capturing people realistically. It can be about representing the essence and the message of the scene perceived by the beholder in a particular art style. From our own experience teaching at a non-profit organization, we have found that many students have trouble drawing facial feature, such as the eyes, the jawline etc. They might also struggle transforming a style of pattern into their drawing. Another common complaint they have is using their knowledge about textures and symbols and applying it on content pictures. We understand that sometimes explaining that kind of art orally becomes a bit of a challenge, especially for younger kids from ages 2-7. Therefore, this inspired us to think that it is advantageous to have an efficient way of providing pictures that can be used as pedagogical references.

As we can see, art moves together with the development of technology. Technology has been able to give artists a chance to reveal their inspiration in new forms. Digital art has been popular for decades and still continues to grow. Many young artists, like Alberto Seveso and Evgeny Parfenov, were inspired by these new tools and were able to come up with their own astonishing pieces of art. In particular, generative art is one of the most popular and fast-growing paradigms among the technology-assisted arts. Generative art uses algorithmic and parametric modeling to mimic real-life behaviors, including static and dynamic imagery, structure, and sound. This new form of art is up and coming and affects not just the visual art paradigm, but also the industrial, architectural and urban design paradigms. It has been able to shorten the amount of time it takes to create these works of art, and it has made it easier for the designer to refine their original models. It has also been helpful in providing people ways to visualize 3D models more clearly and interactively. One of the most early neural-based generative art technique is the style net published in 2015 by Gatys et al.[1]. We are curious about its potential to provide new pedagogical approach to art teaching and learning.

The rise and popularization of deep learning changes how research in computer vision is done and brings new energy into generative art. In particular, one of the most early neural-based generative art technique is the style-transfer neural network published in 2015 by Gatys et al.[1]. Using gram matrices over extracted feature weights to quantify “artistic style” of an image, beautiful work of style-transfer is cleverly achieved where original images are modified algorithmically to acquire the visual styles of other images. Since then there have been plenty of follow-ups and improvements on the quality, speed and the conditioning of style-transfer techniques [1,2,57,58], as well as demonstrations of its infinite possibilities published online. However, we believe there is lack of effort toward a empirical behavioral study of how different kinds of visual patterns present in style image are mixed with content image and how successful results can be obtained by proper settings of hyper parameters and choices of images to start with. This work addresses this deficiency in prior works. More importantly, we are curious about its potential to provide new pedagogical approach to art teaching and learning. Therefore in this work we choose a common topic in art education - outlining and styling a human face - and investigate how neural-based style-transfer technique can support this task.

In this work, we used a wide variety of textures and symbols to style a photo of a female human face. The outcome pictures show patterns of the style picture that can be used as an outline of the face for beginners or examples that a teacher may give to a student about how to apply creatively textures or symbols on the human face. We briefly discuss related work in section II, describe our approach in section III, summarize our results in section IV, and conclude this work in section V.

2. Related Work

One of the major approaches that researchers who focus on artistic style transfer frequently discuss is the Neural Network. The overall idea is to use neural networks to extract richly expressive features from multiple source images, some as style images and others as content images, and then optimize the pixels of an output image, whose extracted features would match as close as possible the extracted features from source images. A general argument clearly offered by Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge in their article A Neural Algorithm of Artistic Style[1] is that their invention of a neural algorithm increases the knowledge of how humans produce and analyze art pieces by demonstrating a particular content present in a content image in the form of a notable style of art present in a style image. For example, one can use this technique to transfer Vincent van Gogh's acrylic painting style onto a normal selfie. The study shows that the best way to achieve this goal is to use a Convolutional Neural Network to generate the representation of the content image and the style image and then to match the style and the content separate to the final output image. Content matching would be achieved by comparing each output on the feature maps of a specific layer, usually one of the latter layers that hold high-level structural features. On the other hand, style matching can be attained by comparing a statistical summary of each gram matrix at specific layers. This practice has been proven successful at transferring stylistic patterns onto the final image. As a result, the final image would retain the structural content of the content image while adopt the stylistic characteristic of the style image.

Similarly, Novak and Nikulin[2] also discovered many improvements. One of the their key findings is their adjustment of each layer’s content and style weight with a geometric weight scheme. Also, they proposed using all 16 layers as style layers to calculate gram matrices for richer style representation. Another important contribution of their work is the one where they shifted the activation to accelerate the process of style transfer convergence and to eliminate sparsity, which makes each picture more informative. Lastly, in order to present more features, they focused on correlations of features belonging to different layers. Due to their impressive work, we use their model as our base to examine more behavioral performances of neural style transfer.

Additionally, Mao-Chuang Yeh and Shuai Tang [55] came up with some principles to improve artistic style transfers. To be exact, they demonstrated two solutions: adaptation of universal style transfer methods onto cross layer gram matrices and usage of loss multiplication. Undoubtedly, from these modifications, the final product proved to have a better texture from the previous style, and it preserved more crucial features that needed to be included from the content. One of their pictures, using Brad Pitt's face to transfer different style of art, inspired us to start this research of using female faces as content implied from different textures and symbols.

3. Approach

The only content image we use in this work is Ziyi Zhang [3], a mugshot of a famous Chinese actress. She is not only known for her movies but also for her photogenic face, which is well structured, and easy to trace using face curls. All the experiments done in this work are based on the improved style-transfer algorithm proposed in [2]. We run our experiments on Nvidia Quadro P5000 GPU, and each run consists of 20 forward passes of style-transfer neural network and result image updates that minimize the defined style-transfer loss function. More iterations only result in negligible change in the results. All input and output images are scaled to 706x700. Larger image would provide spaces for finer visual patterns to develop but whose transfer would take longer time to complete, and we found the aforementioned dimension to be a reasonable trade-off between quality and time-cost.

3.1 Face-sized Curls

One of the most basic skills that every artist requires is the ability to identify facial structure. The most frequently appearing aspects of facial structures is curved lines. Forexample, all of the jaw lines, eyebrows, lips lines and etc are constructed from curls. By using face-sized curls (see swirl_1-5[4-8]), a beginner would be able to outline the facial features from the outcome picture. Face-sized curls is a very broad term, but in this research we chose to use a specific curl, swirls, which is a twisting or spiraling pattern that is the same size as the face in the content image.

3.2 Linear Texture, Tiles and Contour

Linear texture, tiles and contour are some basic shapes and patterns that could be easily found in real life and used in art. By combining patterns and the content picture, the artist is able to restructure the entire face with the provided linear and contour textures in the style image. The output facial image will be easier for a beginner to trace and to imitate. Also, it suggests different approaches to draw human faces, thus helps the artist to conceive facial structures in completely new ways.

Four particular classes of linear texture, tiles and contours are used in this work:

- straight lines of different orientations and coherence: In this case, coherence refers to lines of different thicknesses and proximity (see image titled, line_1[9]). For instance, in image line_1, the lines are closer together and thicker compared to the lines in image line_2[10]. Also, the lines in image line_3[11] are horizontal, which is different orientation than the diagonal lines in images line_2 and line_1.
- Locally twisted lines: To be specific, the corresponding style images use straight lines following a twisted pattern (see line_4[12] and line_5[13]). However, in line_4 the density of the lines are heavier than in line5. Also, line_5 uses only diagonal lines, whereas line_4 uses horizontal and diagonal lines.
- Tiles such as origami: Using origami creates more complex and varied style images. For instance, all of the pictures that we used are composed of different colors. There are also different shapes range from simple ones to complex ones. To be exact, Origami images 1-4[14-17] are animal shapes in different sizes. Origami images 5-9[18-22] are composed of basic geometric shapes.
- Contours such as contour map and visual art works motivated by it: By definition, contour means outline, especially one representing or bounding the shape or form of something. In this research, we use artworks that are composed of a repeated motif inspired by a contour. To elaborate, image leaf[23], waves[24], line_6[25], line_7[26] and waves_2 [27] are all examples of the contour image style.

3.3 Dotted Texture

Dotted texture has been widely used in art to create different styles. It is simple to draw, but any difference in color and size can generate a completely different style. The goal is to be able to restructure the entire face with the colored dots of various sizes provided in the style image only. This helps beginners to learn and understand the style and be able to apply it to a female face.

Three particular classes of linear texture, tiles and contours are used in this work:

- Australian aboriginal dot paintings: The Australian aboriginal dot paintings use colorful and different sized dots to portray a story or a belief of the people. They usually include the shapes of animals. (see tribe 1-2[28-29])
- Halftone: This technique is used in old comic books. The dots vary in size and shape to indicate different tones (see halftone 1-5 [30-34]).
- Yayoi Kusama: Kusama is a Japanese contemporary artist who has a distinguishable artistic style. She often uses colorful dots to create abstract artworks on unconventional canvases. For instance, she's painted on mirrors, people's bodies, ceilings, and other odd objects. (see kusama 1-8 [35-42]).

3.4 Symbolic Patterns

Many symbolic patterns are used to represent specific cultures. The ideal result is to be able to restructure the entire face with symbols of various sizes, colors and geometric features provided in the style image only. This can help generate inspiration for artists who have trouble coming up with new ideas. More importantly, it helps to propagate cultural knowledge.

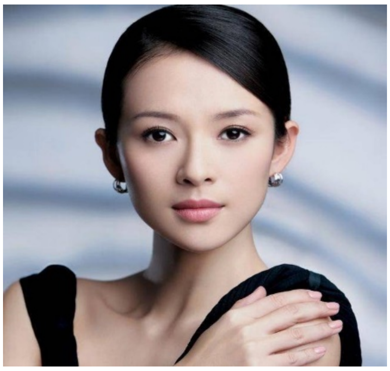
Three particular classes of linear texture, tiles and contours are used in this work:

- Chinese blue-and-white porcelain: It is often composed of willow patterns. This style appears most often on Chinese ceramics. In the research, we use picture of porcelain in various willow pattern and size of the pattern. (see blue1-4 [43-46])
- Scandinavian traditional patterns: These are symmetrical patterns based off simple illustrations of animals, geometric shapes and plants. These patterns appear in different colors and styles. (see scandi1-3 [47-49])
- Polynesian: Today, Polynesian motifs are likely appear as tattoo art. This style is composed of black symbols, plants and geometrical patterns, such as bands and triangles. (see polynesian 1-5 [50-54])

4. Results and Discussions

4.1 Face-sized Curls

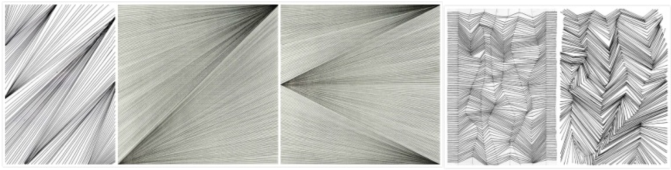
As show from fig. 1 to fig. 5, the condition for possible failures are when the size of the style picture is smaller than the content image. As a result, the outcome picture cannot outline the results, as in image four. Also, when the empty space in the style picture is more than half of the picture, it is likely that the nose, mouth, hair, ear, etc cannot be traced. For instance, in *swirl_result 1-3,* we can see that some details in the outcome picture are being ignored by the algorithm during the artistic style transfer process. In order to create a well-illustrated picture, the style picture also has to have enough detail for the computer to generate the outcome picture.



Ziyi Zhang

Figure 1. Mug-shot of Ziyi

Figure 2. Style images of face-sized



line_1 line_2 line_3 line_4 line_5

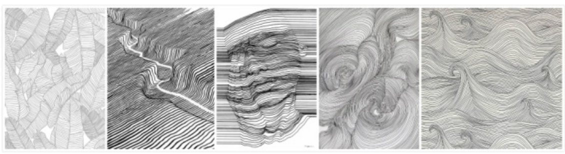
Figure 3. Style images of straight lines and locally twisted line.

origami 1,2,3

origami 4,5,6

origami 7,8,9

Figure 4. Style images of tiles.



leaf line_6 line_7 waves waves_2

Figure 5. Style images of Contour.

One particularly visually appealing result is *swirl_result 5*. In this picture, the jaw line is clearly outlined, as well as the nose, eyes, and ears. What stands out in this picture more than the other results is that the nose and mouth is portrayed by a single line instead of a shaded area, like in *swirl_result 1-3*. Most importantly, her hand and nails are distinctly projected in the outcome picture. Even though picture two also has well-traced facial features, due to the content picture having too many small, minor details and patterns as well as small tattoo-like figures on the face, prove this picture to be less visually appealing, as show from fig. 6 to fig. 9.

Figure 6. Style images of Australian aboriginal dot paintings.

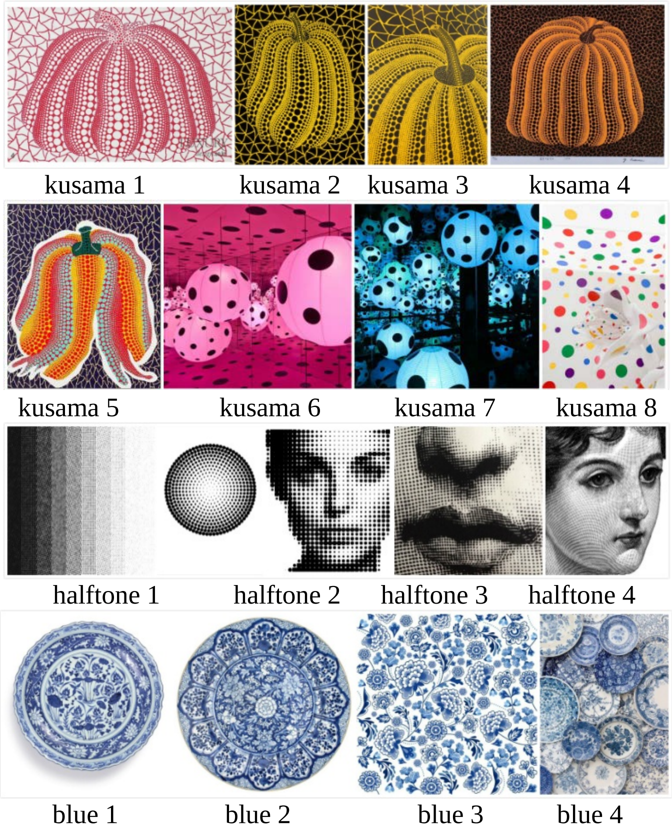


Figure 7. Style images of Chinese blue-and-white porcelain

Figure 8. Style images ofScandinavian traditional patterns.

Figure 9. Style images ofScandinavian traditional patterns.

4.2 Linear Texture, Tiles, and Contour

Some factors that greatly affect the results of these pictures are the patterns and orientation of the line, tile, and contour. One notable difference is that the lines that are twisted provide more detail to the outcome picture compared to the style picture that is only composed of straight lines. To be specific, image *line_result 1*, uses the style image that is only composed of straight lines, overlooking the minor details, such as eyes, ears, and nose. This factor affects contour style image results as well. In contour style image leaf, the picture includes straight lines and curved lines. This resulted in a more structured face than the contoured image that only included curved lines. Moreover, when using origami as tile to reconstruct the picture, the pattern in the style image cannot be as small as image o*rigami_3* and cannot be bigger than image *Origami_8*, otherwise the outcome picture would be unclear, as show from fig. 10 to fig. 12.

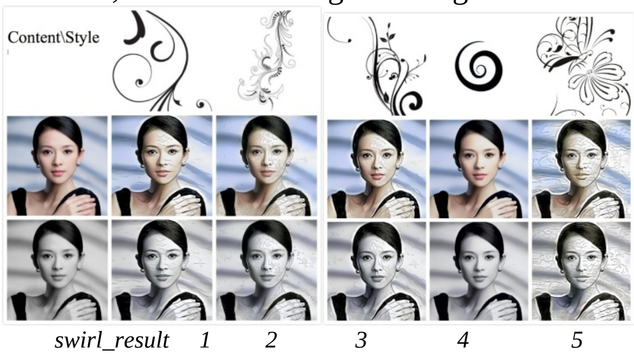


Figure 10. Output images of Swirls.

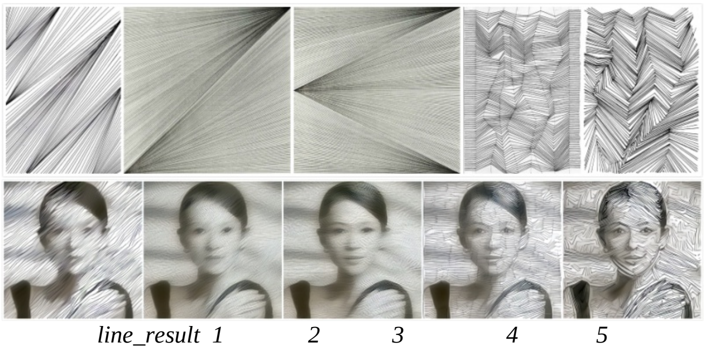
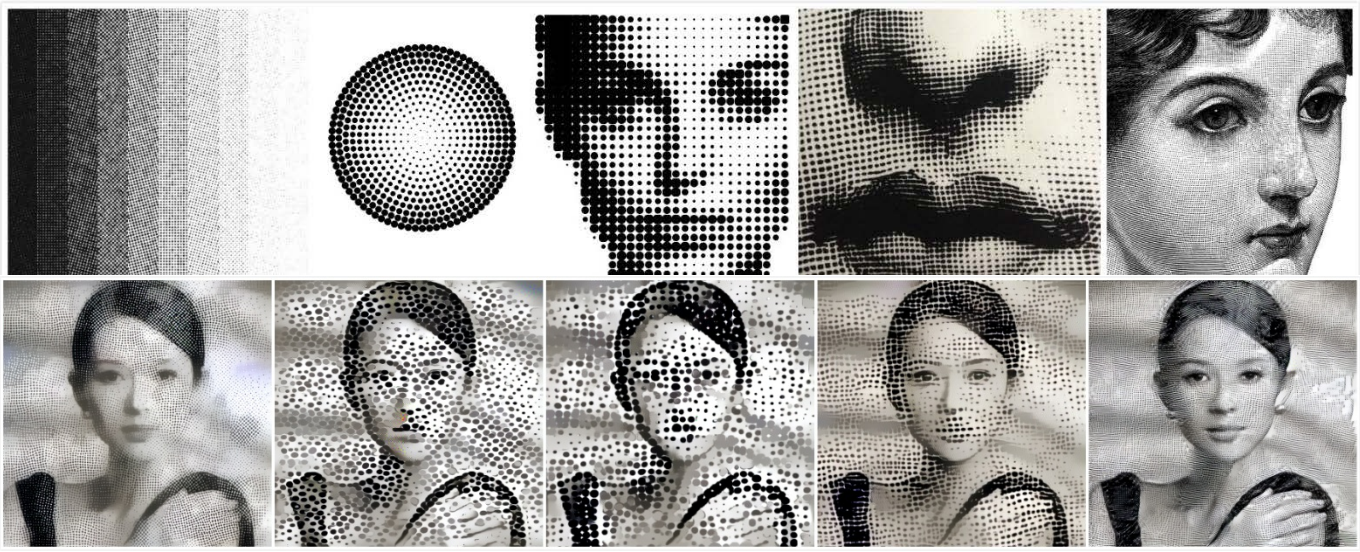


Figure 11. Output images of straight line and locally twisted line.

Figure 12. Output images of tiles.

One picture that displays high quality transformation and well expresses the linear style, is the image *line_result 5*. This image clearly resembles the face with lines and with obvious facial details. In fact, some influential factors that make this image stand out are the fact that the style image provides various detail, like twisted lines but in different orientations, and the density of the line is smaller. When there is large density of lines in the same size picture, details, such as the eyes, is hard to copy. Furthermore, for the contour image, the results of the picture using a leaf as the style is desirable because it can better illustrate the jaw line and hand. The difference with this style image is that it contains both straight and curvy lines. Image results_origami7 is an appealing picture that uses tile. It can clearly be seen that Image *origami_result 7* replaces the nose and hand with tile because the block in the origami is as small as the size of the minor details on the face, as show from fig. 13 to fig. 16.

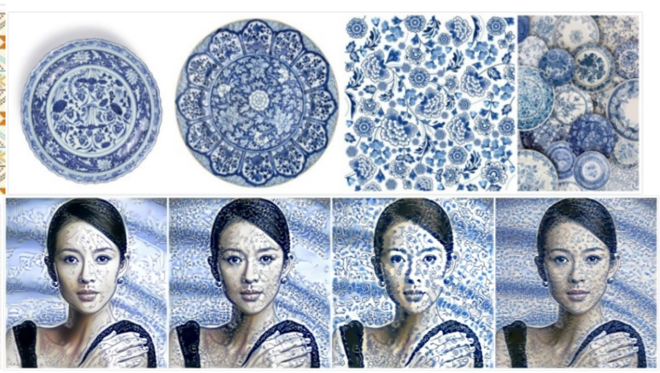


halftone_result 1 2 3 4 5

Figure 13. Output images of halftone.

Kusama_result 1 2 3 4 5 6 7 8

Figure 14. Output images of Yayoi Kusama's painting.



scandi_result1 2 3 blue_result1 2 3 4

Figure 15. Output images ofScandinavian traditional patterns.

Figure 16. Output images of Chinese blue-and-white

Another approach that we use to accomplish the goal is using the grayscale content picture. The results prove that using the grayscale content picture would display results that have different color placement. Also, pictures that have grayscale would lose some detail. For instance, by comparing image *origami_result 7* and its gray scaled image, one can see that the two different patterns appear on each image. The original picture shows more tile-like patterns, which indicates the grayscale image changes the outcome picture.

4.3 Dotted Texture

One noticeable effect that can be seen by using dots is that they change the outcome of the research. The picture with various size dots would create a clear image that outlines the facial structure. This happens because when there is different size dots, the computer has more options to choose from to illustrate the shaded regions, eyebrows, and the lip. For instance, style image *halfstone_1*, only uses same size dots, and the content image combined. With this, one can easily conclude that the final picture does not successfully transfer the style onto the face. This can be seen because many areas, such as the lip, right cheek, eyes, etc, did not have any dots on it. Moreover, another thing that can be noted is that in the style picture there should be a dot or a group of dots that is similar to the size of the pupil of the eye. Without that specific dot, the eye would not be able to be modified by the algorithm. Similar to the linear patterns, when the style image is left too empty, areas like the cheek would likely also be blank. Furthermore, when the dots are too concentrated, the hair and eyes would likely stay like the original content picture.

Image *Kusama_result 5* is one of the outstanding results that is generated in this research. In this picture, the right side of the jawline has been changed into multiple, lighter, grey dots to portray the non-shaded area. The shaded area has been changed into multiple black dots. Another reason that makes this picture extraordinary is that it uses five smaller dots to demonstrate the cracks on the lip. In order to further improve the picture, we use the grayscale content picture. From this, we discover that not only is the color placement different, but the pupil would also be more likely to transfer. For example, before using the gray scale picture, the left side of the pupil is similar to the original picture, but with a different color. After changing to grayscale image however, the result of the image illustrates a pupil that has turned into red, which is a color that can be found in the style image.

4.4 Symbolic Patterns

From the research, we see that different conditions can be analyzed as factors that impact the modification of the picture, such as using various symbols. Remarkably, when multiple patterns stack into a single picture (see image *scandi_result 2*), the pattern is more likely to be mixed. Also, when the repeated pattern in the style image is too small compared to the size of the eye in the content, the outcome picture will be blurry (see image *blue_result 4*). One notable finding in this research is that with more detail in the pattern of the style image, the outcome picture demonstrates more precision, especially for the eyes and lips. Another important aspect is that with the symbolic patterns that fill the entire image, the details such as the hair and eyebrows would more likely transfer into the expected way.

One picture, image *blue_result3*, successfully accomplishes this predicted goal. In this picture, all of the facial feature has been extracted and transformed into the willow pattern. To elaborate, the hair had transferred into the willow pattern and the left eye had transferred into another willow pattern. These are the facial characteristics that are difficult to change using symbolic patterns. Also, it changes the pearl earring into two flowery shaped earrings, which perfectly match with her shirt and face. This illustrates a potential idea that based on the simple shape of old-fashioned jewelry and artistic style transferred into a more new and popular style can help or inspire designers to produce new jewelry.

5. Conclusion and Future Work

In conclusion, the research illustrated in the study, Improving the Neural Algorithm of Artistic Style, describes essential concepts that can possibly be applied on a female face image to create a new outcome picture that has the combination of content and style. These strategies can be implemented in artistic education. Four different textures have been used in a variety of results that represent how some commonly found patterns can be applied. They also aid in developing some strategies that help when choosing style images, like density of the pattern and the orientation of the pattern. Another finding in the research is how grayscale content images can affect the patterns on the outcome image. However, the research still leaves possibility of improvement.

One improvement could be trying a wide variety of ratio between content weight and style weight and then determining the best value that helps the algorithm generate appealing results. The different styles that could be performed on animals could be another approach to improve the idea of artistic transfer. Lastly, using pictures of old objects and mixing them with style images that include new trend patterns to produce a new object could point to another interesting work.

References

[1] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, *A Neural Algorithm of Artistic Style*, arXiv:1508.06576, August 2015.

[2] Roman Novak and Yaroslav, *Improving the Neural Algorithm of Artistic Style*, arXiv:1605.04603, May 2016.

[3] Pinterest image, source: https://pin.it/oiocdjawo5raxd

[4] Pinterest image, https://pin.it/yhz2woyjugftsk

[5] Pinterest image, https://pin.it/yhz2woyjugftsk

[6] Pinterest image, https://pin.it/hxhh4dudhwsgd5

[7] Pinterest image, https://pin.it/255tr3bhd32kl6

[8] Pinterest image, https://pin.it/26eipof2gfpban

[9] Pinterest image, https://pin.it/vpw5r56x5imf4u

[10] Pinterest image, https://pin.it/36z3eipw63dhho

[11] Pinterest image, https://pin.it/ueenljapl6vsum

[12] Pinterest image, https://pin.it/73xcb6c7tm6thn

[13] Pinterest image, https://pin.it/bd62ghy46nivnq

[14] Pinterest image, https://pin.it/ygmx57wdewer2w

[15] Pinterest image, https://pin.it/twbnt2nhbsmqll

[16] Pinterest image, https://pin.it/twvswnowxwruwe

[17] Pinterest image, https://pin.it/vbvl4x3mp57nnf

[18] Pinterest image, https://pin.it/nee2sk37nw4k7w

[19] Pinterest image, https://pin.it/vxtynd4po3yzu7

[20] Pinterest image, https://pin.it/cabcjvf5rwozwy

[21] Pinterest image, https://pin.it/3prt74xw77caqf

[22] Pinterest image, https://pin.it/ibhtkqcsuwioid

[23] Pinterest image, https://pin.it/kde4xpwvqfgh64

[24] Pinterest image, https://pin.it/evgiscc2vfxzzg

[25] Pinterest image, https://pin.it/jj6qeffrdkqc3j

[26] Pinterest image, https://pin.it/udtj6xmcog5li5

[27] Pinterest image, https://pin.it/p5sdlcxzwt2iza

[28] Pinterest image, https://pin.it/oabd5qzamv7ksn

[29] Pinterest image, https://pin.it/7tftrmnmnliuxc

[30] Pinterest image, https://pin.it/5n7b3624qhh7vt

[31] Pinterest image, https://pin.it/ngoh3gxlsxqwxe

[32] Pinterest image, https://pin.it/wrq2mw5bxqauyz

[33] Pinterest image, https://pin.it/4i3bgpvvhbkajl

[34] Pinterest image, https://pin.it/zrion37nby2tav

[35] Pinterest image, https://pin.it/cl5smph2enve67

[36] Pinterest image, https://pin.it/qgb5j7agqrdqzq

[37] Pinterest image, https://pin.it/tvufv3vfqvq7lk

[38] Pinterest image, https://pin.it/kwlnakzk6kjmm6

[39] Pinterest image, https://pin.it/j375vy7cvm63zl

[40] Pinterest image, https://pin.it/73nb6jcwb7tboe

[41] Pinterest image, https://pin.it/fsgzajxkhsupgd

[42] Pinterest image, https://pin.it/oysjmhm57p2n72

[43] Pinterest image, https://pin.it/tomnh4fqz7qixd

[44] Pinterest image, https://pin.it/weh4mfp5dohvh6

[45] Pinterest image, https://pin.it/obalaunynalurr

[46] Pinterest image, https://pin.it/dkkz3lu7yh2ldd

[47] Pinterest image, https://pin.it/bbummu5cidhwbw

[48] Pinterest image, https://pin.it/yrg4wkqkaclf7q

[49] Pinterest image, https://pin.it/e6eo3rf5ozwx5x

[50] Pinterest image, https://pin.it/2m24jdmp5j6lwt

[51] Pinterest image, https://pin.it/xcjmqbgr666ano

[52] Pinterest image, https://pin.it/oe65agemcluvdy

[53] Pinterest image, https://pin.it/w762fewgurwh72

[54] Pinterest image, https://pin.it/6x3dfrnyvy5xtq

[55] Mao-Chuang Yeh and Shuai Tang, *Improved Style Transfer by Respecting Inter-layer Correlations*, arXiv:1801.01933, 5 January 2018 .

[56] Paddy Bowman, Betty Carter and Alan Glovenar. The Master of traditional Art Education, Documentary Arts, Inc., 2011.

[57] Leon A. Gatys, Matthias Bethge, Aaron Hertzmann, Eli Shechtman, *Preserving Color in Neural Style Transfer*, arXiV:1606.05897, June 2016.

[58] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge, Aaron Hertzmann, Eli Shechtman, *Controlling Perceptual Factors in Neural Style Transfer*, arXiV:1611.07865, May 2017.