# Hair Shading Style Transfer for Manga with cGAN

Masashi Aizawa, Ryohei Orihara[a], Yuichi Sei[b], Yasuyuki Tahara[c] and Akihiko Ohsuga
*The University of Electro-Communications, Tokyo, Japan*

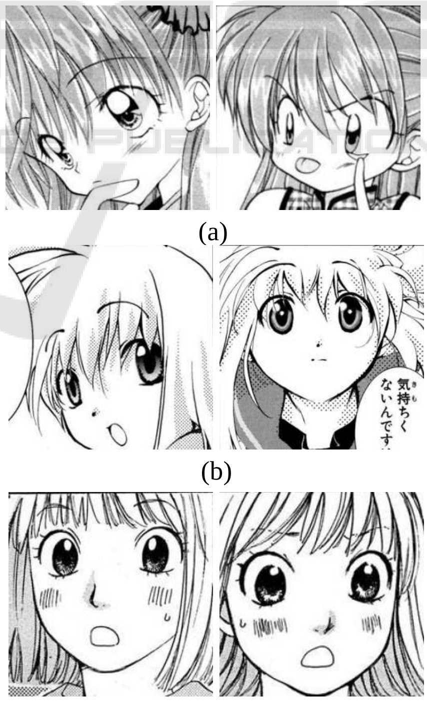
Keywords: Line Drawing, cGAN, Shading, Style Transfer.

Abstract: Coloring line drawings is an important process in creating artwork. In coloring, a shading style is where an artist's style is the most noticeable. Artists spend a great deal of time and effort creating art. It is thus difficult for beginners to draw specific shading styles, and even experienced people have a hard time trying to draw many different styles. Features of a shading styles appear prominently in the hair region of human characters. In many cases, hair is drawn using a combination of the base color, the highlights, and the shadows. In this study, we propose a method for transferring the shading style used on hair in one drawing to another drawing. This method uses a single reference image for training and does not need a large data set. This paper describes the framework, transfer results, and discussions. The transfer results show the following: when transferring the shading style to the line drawing by the same artist, the method can detect the hair region relatively well, and the transfer result is indistinguishable from the transfer target in some shading styles. In addition, the evaluation results show that our method has higher scores than an existing automatic colorizing service.

## 1 INTRODUCTION

Coloring line drawings is an important process in creating artwork. Many artists uniquely express shading and texture when coloring. This process requires special skills, experience, and knowledge, which makes such work difficult for beginners to imitate, and proficient artists spend considerable time and effort creating these pieces of art.

In particular, there are many types of styles for shading hair, as shown in figure 1. Highlight and shadow textures overlaid on the base color are expressed in various shapes, positions, and sizes. Their combination greatly influences the impression the art gives. The hair shading styles differ not only by artist but by character and even by scene. However, artists tend to use a particular style defined by a set of common features to draw an original character. The images in figure 1(a) show the character Seri drawn by Miki Ueda. In these images, the base color and the highlights represent the shading style used for this character's hair: the base color is gray with a screen tone, and the highlights are wide white areas with jagged outlines. The images in (b) show Satori Senjuin, drawn by Minene Sakurano. In these images, the base color and the shadows represent the style used for the character: the base color is white, and the




(a)
(b)
(c)

Figure 1: Examples of various hair shading styles. ©Miki Ueda (a), Minene Sakurano (b), Kaori Saki (c).

[a] https://orcid.org/0000-0002-9039-7704
[b] https://orcid.org/0000-0002-2552-6717
[c] https://orcid.org/0000-0002-1939-4455



Aizawa, M., Orihara, R., Sei, Y., Tahara, Y. and Ohsuga, A.
Hair Shading Style Transfer for Manga with cGAN.
DOI: 10.5220/0008961405870594
In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence (ICAART 2020) - Volume 2*, pages 587-594
ISBN: 978-989-758-395-7; ISSN: 2184-433X

shadows are shaped following the tufts of hair. The images in (c) show Nadeshiko Toba, drawn by Kaori Saki. In these images, the artist shades using the base color and the shadows: the base color is white, and the back of the head is shaded. In an automatic drawing, a reference shading style is expected to enable beginners to learn how to draw in that style, as well as skilled artists to easily try combinations of line drawings and shading styles. As a result, a transfer method for a specific hair shading style has significant market demand. Although many researches on automated line drawing colorization, most of them did recently conducted, were not consider transferring a hair shading style (figure 2).

In this paper, we propose a method to transfer the hair shading style of the training data (transfer source) to another line drawing (transfer target). We perform deep learning using a pair of images: an image with hair shading and another whose entire hair region is painted white. We also adopt a machine learning method using a single pair of images due to the difficulty of collecting data sets. The model is based on the conditional generative adversarial network (cGAN). This paper describes the framework for transferring a hair shading style and the discussions on the transfer results. The framework consists of three parts: data set creation (figure 3), the training model (figure 4), and transfer.

# 2 RELATED WORK

## 2.1 Manga and Sketch Processing

Studies on manga object detection have been conducted. Tolle and Arai (Tolle and Arai, 2011) proposed a method for text extraction from balloons in an e-comic. Sun et al. (Sun et al., 2013) detected comic characters from comic panels using local feature matching. These object detection methods did not directly focus on to draw shading textures, while we aimed to transfer the texture of hair shading styles to line drawings. Although not yet applied, these methods help to improve the accuracy of our method in detecting a hair region. In addition to object detection, studies on semantic manga understanding, screen pattern analyzing, sketch simplification have been conducted. Chu and Li (Chu and Li, 2017) proposed a method for detecting the bounding box of a manga character's face; this method used a deep neural network model. A study of comic style analysis (Chu and Cheng, 2016) categorized manga by considering screen pattern, panel shape, and size as potential manga styles. Those studies aiming at semantic understanding of a line drawing in a manga were dealing with a challenging issue. Yao et al. (Yao et al., 2017) vectorized manga by detecting screen tone regions and classifying screen patterns. This method handled specific patterns such as dots, stripes, and grids. Li et al. (Li et al., 2017) extracted structural lines from manga pasted with screen tones. The methods based on the screen pattern analysis hardly lead to semantically understanding of line drawings. A study of sketch simplification (Simo-Serra et al., 2016) generated a simple line drawing from a complex rough sketch in a raster image. This method consumed heavy cost for creating a data set. Further studies of sketch simplification proposed a discriminator model that offered unsupervised training data (Simo-Serra et al., 2018a) and suggested a user interaction method (Simo-Serra et al., 2018b).

## 2.2 Line Drawing Colorization

Many studies have been conducted on line drawing colorization. Qu et al. (Qu et al., 2006) proposed a manga colorization method based on segmentation. This method extracted similar pattern regions from grayscale textures. Lazybrush (Sýkora et al., 2009) proposed a region-based colorizing method with smart segmentation based on edge extraction. In these optimization-based approaches that propagate colors based on regions, the shading style cannot be transferred because the output depends on the input line drawing.

Machine-learning-based approaches also achieved line drawing colorization. Scribbler (Sangkloy et al., 2017) generated various realistic scenes from sketches and color hints. The sketches range from bedrooms to human faces. Auto-painter (Liu et al., 2017) proposed the new training model that adjusts the adversarial loss to adapt color collocations. Frans (Frans, 2017) generated color and shading predictions and synthesized them for line drawing colorization. In the study, colorization is carried out based on the line drawing, and shading is done based on both the line drawing and color scheme. These studies could hardly generate authentic illustrations. Aizawa et al. (Aizawa et al., 2019) generated a colorized image considering a body part. Their study detected the sclera regions of a human character using semantic segmentation and synthesized it with an automated colorization result. However, in subjective evaluation, most results of colorizing sclera were indistinguishable from not colorizing. Ci et al. (Ci et al., 2018) achieved high-quality colorizing using semantic feature maps from a line drawing. The study of two-stage training (Zhang et al., 2018) prevented color bleeding and blurring. This framework had two separate training phases: the rough colorization stage and the elaborate colorization stage. The user could hardly handle shading styles of colorization results, while these methods produced more plausible colorization results. PaintsChainer (Taizan, 2016) is an online service that automates line drawing colorization. Although a user can colorize the three predefined shading style versions, the shape and position of highlights and shadows are untruthful.

## 2.3 Style Transfer

Comicolorization (Furusawa et al., 2017) proposed a method of painting each character in a manga with a distinct color. Their study transferred the color of a reference image to a manga page. Although the method performed the transfer across the entire manga page, it was difficult to produce precise colorization for small regions. This approach generated colorization results affected by textures of manga, while our method aimed to transfer the texture of hair shading styles to line drawings without explicitly providing texture. The shading style transfer approach (Zhang et al., 2017) colorized and shaded an anime character line drawing with a color reference image. Their study was aimed at a color image, not a manga. Their study also used a large data set, while our method works with a single-pair data set. Furthermore, their study could hardly preserve the shading style of the reference image.

## 2.4 Single-pair Training

Hensman and Aizawa (Hensman and Aizawa, 2017) colorized a grayscale manga. They combined a learning method and a segmentation approach. Because manga images are protected by copyright as artwork and it is difficult to collect data sets, they used a single pair for training. The study used a deep learning network similar to our method, but it differed from our method in that it was colorized with reference to the grayscale texture of the target image. Our method is aimed at generating an image with a hair shading style transferred from another image without explicitly providing a hair shading style. Their study also generated segmentation-based colorization results, while our method performed no segmentation. SinGAN (Shaham et al., 2019) proposed a method that could be trained from a single natural image. This method could be applied to various tasks. It was not confirmed whether this method could be applied to our method.

Figure 2: An example of the existing colorization approach (Taizan, 2016). A user can not transfer specific shading style.

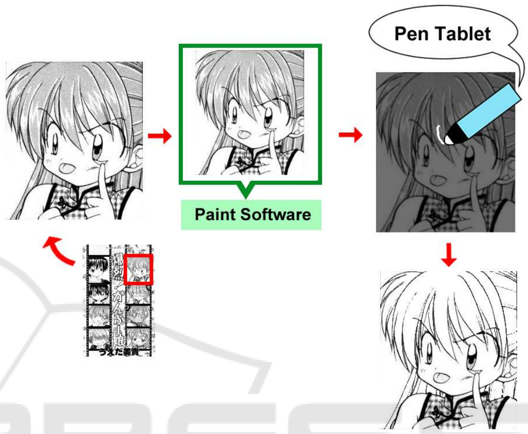



Figure 3: Data set creation. ©Miki Ueda.

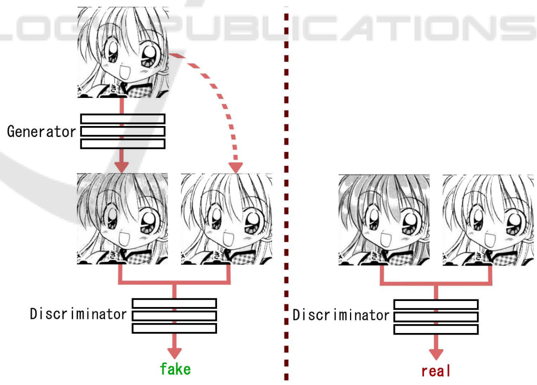



Figure 4: An example of training cGAN to map an image from a line drawing in which the hair region is painted white into an image with shaded hair. ©Miki Ueda.

# 3 FRAMEWORK

This study consists of three parts: data set creation, the training model, and hair shading style transfer. Figure 5 shows our framework.

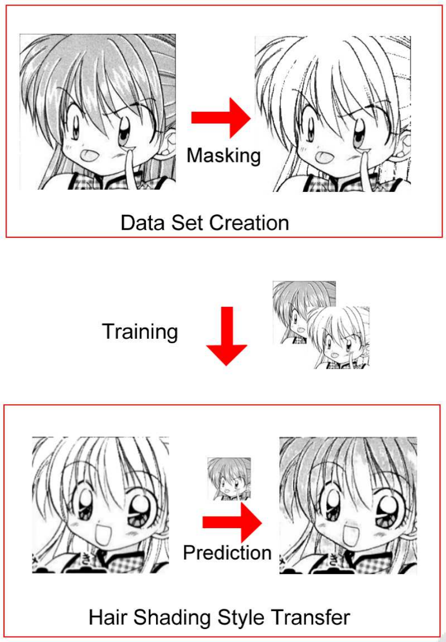



Figure 5: Our framework. ©Miki Ueda.

### 3.1 Data Set Creation

We used Manga109 (Matsui et al., 2017) to create the data set. The Manga109 data set consists of manga works of various genres and of varying release dates. An image in Manga109 represents a facing-page of a manga book. In data set creation, first, we cut out a human character's face from the Manga109 data set. The shading style of this image is the transfer source style in our method. Manga109 had several types of annotations, and the facial area is marked by its bounding box. However, this rectangle often misses out the hair region. Therefore, in this study, it was cut out manually. Next, we filled the hair region of the face image with white. As a result, we created a pair of images; one with shading and the other without shading. Figure 3 shows an example of a created data set. In filling a hair region when creating the data set, an experienced illustrator manually masked it using a paint software and a pen tablet. We used Clip Studio Paint EX (CELSYS, Inc., Tokyo, Japan), a professional illustration software package optimized for pen tablets, and an Intuos comic medium CTH-680/S3 pen tablet (Wacom Co., Ltd., Saitama, Japan).

### 3.2 Training

For training, we used the pix2pix network (Isola et al., 2016). Pix2pix is a cGAN-based image-generation neural network technology. The cGAN is a framework consisting of two networks: a generator and a discriminator network. The loss function is a combination of adversarial loss and task-dependent label data. Pix2pix learns the relationship between two image pairs. This makes it possible to generate an image belonging to one domain from an image belonging to another domain; for example, it can generate map images from aerial photos, a colorized image from a black and white image, and so on.

In this study, we used a pair of images; one with a hair shading style and another whose entire hair region was painted white, as a data set. Because we manually prepared masked images, it was difficult to collect a large-scale data set. Therefore, we performed training with a single pair. We expanded the data set by rotating and flipping it.

### 3.3 Transfer

In transfer, a style used in the transfer source was transferred to the transfer result using the trained model. Style transfer results are shown in the next section.

## 4 RESULTS

Figures 6, 7, 8 and 9 show examples of hair shading style transfer results. In (a) of each figure, the image before masking the hair region of the transfer source (top), and the transfer source (bottom) are shown. In (b) to (f) of each figure, the transfer targets (top), the transfer results (middle), and the images before masking the hair region of the transfer targets (bottom) are shown.

Figures 6 and 7 show Seri's style transfer: the style is defined by a large region filled with screen tone. Figures 8 and 9 show Satori's and Nadeshiko's style transfer: the style is defined by a large white area.

Figure 6 shows the results of a style transfer for the characters in the manga book, including the same characters as the transfer sources.

In many cases, the hair regions roughly filled with a screen pattern, which characterizes the artist's style used in the drawings of the same and other characters, are transferred.

Figure 7 shows the results of transferring to a piece of line artwork drawn by a different artist. A hair shading style in the training data is transferred to a hair region of another image. However, it is also transferred to a background and a face. It is difficult to detect a complete hair region using a single-pair data set.

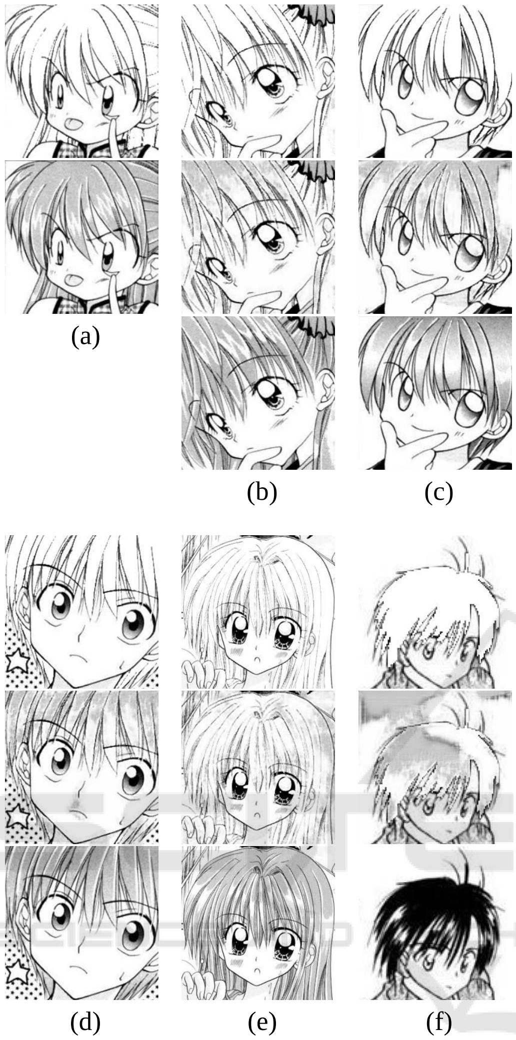


Figure 6: Examples of transferring hair shading styles with screen pattern base color (Seri's style). (a) Training data. (b) Transfer result for the target by the same character as training data. (c) to (f) Transfer results for the target by the other characters drawn by the same artist. ©Miki Ueda.

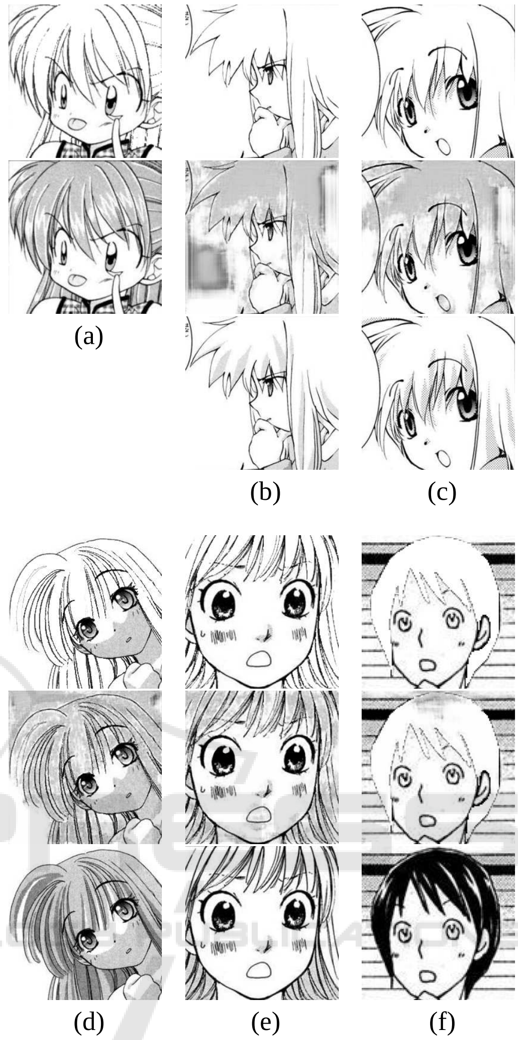


Figure 7: Examples of Seri's style transfer for the target by various artists. (a) Training data. (b) to (e) Transfer results for the targets by various artists. ©Miki Ueda (a), Minene Sakurano (b) to (d), Kaori Saki (e), (f).

In figures 8 and 9, the results are relatively consistent when transferring between the same characters. In the other cases, there is no significant difference between the transfer targets and transfer results.

## 5 EVALUATION

We evaluated the hair region shading style transfer. We performed evaluations in the following method.

- Extracting two faces of the same character from the same manga work. Choosing two similar hair shading styles.
- Painting the hair region of a face image white and training using the proposed method.
- Filling the hair region of the other face image in with white and transferring the hair shading style using the trained model.
- Assuming the transfer target before painting the hair region to be the ground truth and calculating the peak signal-to-noise ratio (PSNR), and the structural similarity index (SSIM) between the transfer result and the ground truth.

We calculated the scores for two regions: the entire image and the hair region only. For this procedure, two face images were extracted from each of the two manga works in Manga109, and the training data and target were swapped. That is, we evaluated a total of four combinations. All the image sizes were 256

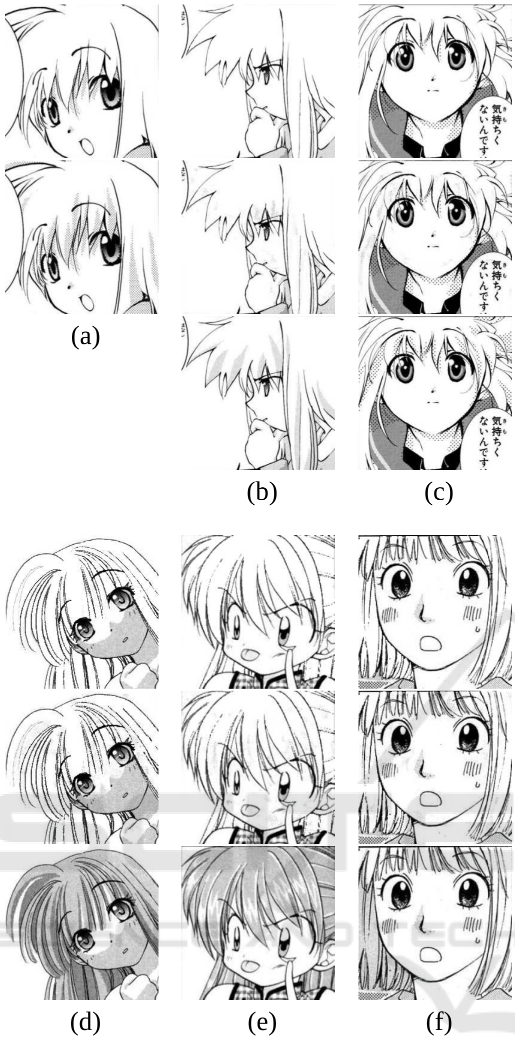



Figure 8: Examples of transferring hair shading styles with a white base (Satori's style). (a) Training data. (b), (c) Transfer results for the same character. (d) Transfer result for another character drawn by the same artist. (e), (f) Transfer results for the target by various artists. ©Minene Sakurano (a) to (d), Miki Ueda (e), Kaori Saki (f).

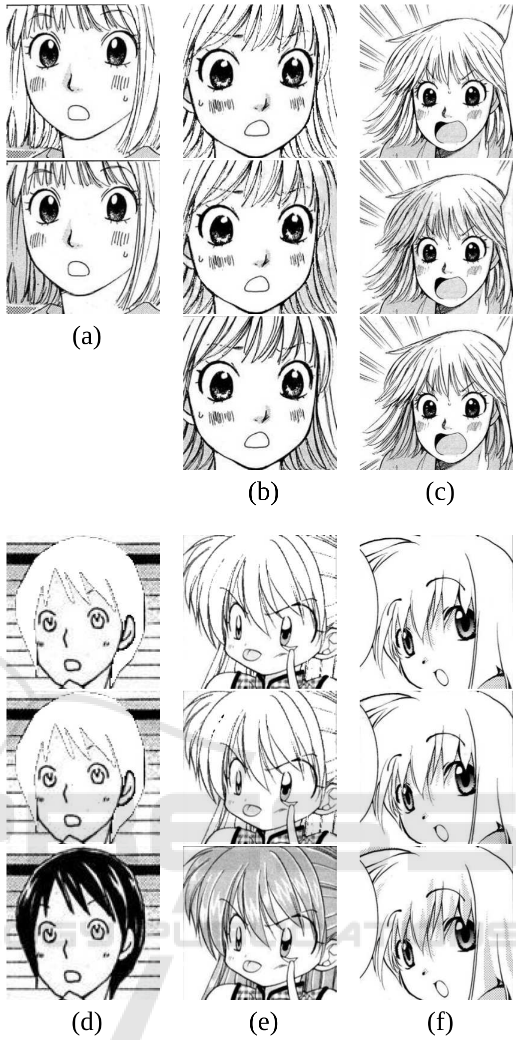



Figure 9: Examples of transferring hair shading styles with a white base (Nadeshiko's style). (a) Training data. (b), (c) Transfer results for the same character. (d) Transfer result for another character drawn by the same artist. (e), (f) Transfer results for the target by various artists. ©Kaori Saki (a) to (d), Miki Ueda (e), Minene Sakurano (f).

x 256 pixels, and each pixel was assigned a grayscale value from 0 to 255 in 256 steps. In training, an epoch was 500, and a batch size was 1.

Although the hair shading styles of the transfer source and the ground truth have similar shapes, it is unclear whether the author drew two styles in the same position and size. Figure 10 shows the images used for evaluation. In addition, for comparison with the existing automatic colorization service, we evaluated the automatic colorization results by using PaintsChainer. The method follows:

- Automatically colorizing the transfer target with PaintsChainer.
- Calculating PSNR and SSIM of the hair region between the colorization results and the ground truth.

PaintsChainer has three colorization versions: Tanpopo (watercolor-like colorization), Satsuki (slightly sharper colorization), and Canna (somewhat clearer highlight and shadow colorization). We evaluated each version. Table 1 shows the evaluation results. The values of our method show the average of ten trials. The headers (a) to (d) correspond to those in the figure 10. PC T, S and C stand for each PaintsChainer version: Tanpopo, Satsuki and Canna, respectively. For two metrics, the higher the values are, the more accurate the transfer results are. Our method achieves the best accuracy in most cases. Ta-

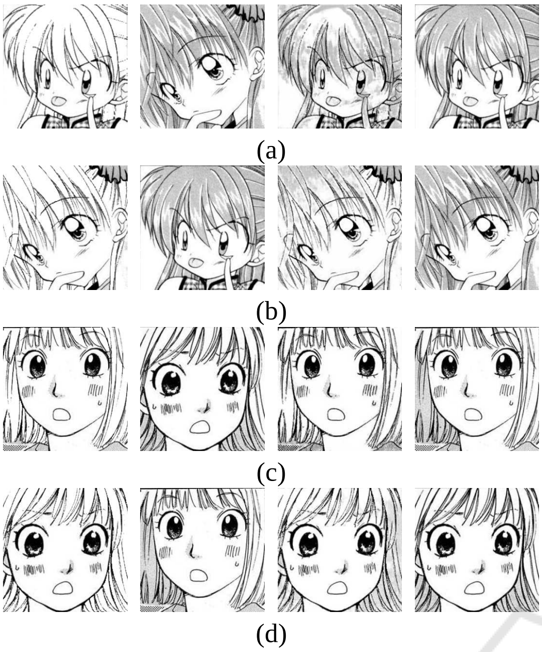



Figure 10: Images for evaluation. From left to right; transfer targets, transfer sources, transfer results, and images before masking the hair region of the transfer targets. ©Miki Ueda (a), (b), Kaori Saki (c), (d).

ble 2 shows the values of the population standard deviation of the PSNR and SSIM score of iterating our training method 10 times.

## 6 CONCLUSION

We proposed a hair shading style transfer method using a single-pair data set. Based on the transfer results, the transfer to line drawings with similar features, such as works by the same artist, managed to draw screen patterns approximately in the hair region. In styles with large white areas, such as the case where the base color was white, there was not much difference between images before and after the transfer. In the evaluation using PSNR and SSIM metrics, our method showed more accurate results than existing automatic colorization applications in both styles with a large screen pattern region and with a large white region.

There are two future works. The first is shading style classification. There are various shading styles, and analyzing the indicators that classify them helps researchers improve the appropriate visual evaluation and framework to be used. The next is framework improvement. Our method could not fully express hair shading styles: the jagged highlight in Seri's style, the shadow along the hair tufts in Satori's style and the shadow at the back of the head as seen in Nadeshiko's style are not properly transferred. Such features were unclear although the shape of highlight and shadow should not be blurred. Trying to apply the object detection methods and the other model, for example, the latest single image training network SinGAN (Shaham et al., 2019), should help us find a suitable framework for our goal.

Table 1: Evaluation results in PSNR and SSIM.

| PSNR | (a) | (b) | (c) | (d) |
|---|---|---|---|---|
| Ours (entire image) | 18.130 | 18.260 | 21.348 | 19.337 |
| Ours (hair region) | 17.344 | 15.707 | 18.731 | 15.074 |
| PC T (hair region) | 14.849 | 16.253 | 14.789 | 13.625 |
| PC S (hair region) | 15.131 | 14.432 | 15.170 | 13.414 |
| PC C (hair region) | 15.507 | 15.577 | 15.442 | 13.728 |
| **SSIM** | (a) | (b) | (c) | (d) |
| Ours (entire image) | 0.6901 | 0.7510 | 0.9162 | 0.8888 |
| Ours (hair region) | 0.4366 | 0.4412 | 0.6270 | 0.5752 |
| PC T (hair region) | 0.4451 | 0.3183 | 0.3910 | 0.3339 |
| PC S (hair region) | 0.4648 | 0.3326 | 0.4238 | 0.4915 |
| PC C (hair region) | 0.4564 | 0.3432 | 0.4828 | 0.5031 |

Table 2: The values of the population standard deviation of PSNR and SSIM.

| | (a) | (b) | (c) | (d) |
|---|---|---|---|---|
| PSNR (entire image) | 0.2165 | 0.1671 | 0.2324 | 0.1975 |
| PSNR (hair region) | 0.5393 | 0.2998 | 0.1051 | 0.2284 |
| SSIM (entire region) | 0.0197 | 0.0154 | 0.0028 | 0.0032 |
| SSIM (hair region) | 0.0078 | 0.0084 | 0.0089 | 0.0117 |

## ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Numbers JP17H04705, JP18H03229, JP18H03340, 18K19835, JP19H04113, JP19K12107.

## REFERENCES

Aizawa, M., Sei, Y., Tahara, Y., Orihara, R., and Ohsuga, A. (2019). Do you like sclera? sclera-region detection and colorization for anime character line drawings. *International Journal of Networked and Distributed Computing*, 7(3):113–120.

Chu, W. and Cheng, W. (2016). Manga-specific features and latent style model for manga style analysis. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1332–1336.

Chu, W.-T. and Li, W.-W. (2017). Manga facenet: Face detection in manga based on deep neural network. pages 412–415.

Ci, Y., Ma, X., Wang, Z., Li, Z., and Luo, Z. (2018). User-guided deep anime line art colorization with conditional adversarial networks. In *Proceedings of the 26th ACM International Conference on Multimedia*, pages 1536–1544.

Frans, K. (2017). Outline colorization through tandem adversarial networks.

Furusawa, C., Hiroshiba, K., Ogaki, K., and Odagiri, Y. (2017). Comicolorization: Semi-automatic manga colorization. In *SIGGRAPH Asia 2017 Technical Briefs*, number 12, pages 12:1–12:4.

Hensman, P. and Aizawa, K. (2017). cgan-based manga colorization using a single training image. In *2017 14th IAPR International Conference on Document Analysis and Recognition*, volume 3, pages 72–77.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2016). Image-to-image translation with conditional adversarial networks. *arXiv:1611.07004*.

Li, C., Liu, X., and Wong, T.-T. (2017). Deep extraction of manga structural lines. *ACM Transactions on Graphics*, 36(4):117:1–117:12.

Liu, Y., Qin, Z., Luo, Z., and Wang, H. (2017). Auto-painter: Cartoon image generation from sketch by using conditional generative adversarial networks.

Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., and Aizawa, K. (2017). Sketch-based manga retrieval using manga109 dataset.

Qu, Y., Wong, T.-T., and Heng, P.-A. (2006). Manga colorization. *ACM Transactions on Graphics*, 25(3):1214–1220.

Sangkloy, P., Lu, J., Fang, C., Yu, F., and Hays, J. (2017). Scribbler: Controlling deep image synthesis with sketch and color. In *The IEEE Conference on Computer Vision and Pattern Recognition*, pages 6836–6845.

Shaham, T. R., Dekel, T., and Michaeli, T. (2019). Singan: Learning a generative model from a single natural image. In *The IEEE International Conference on Computer Vision*, pages 4570–4580.

Simo-Serra, E., Iizuka, S., and Ishikawa, H. (2018a). Mastering Sketching: Adversarial Augmentation for Structured Prediction. *ACM Transactions on Graphics*, 37(1):11:1–11:13.

Simo-Serra, E., Iizuka, S., and Ishikawa, H. (2018b). Real-Time Data-Driven Interactive Rough Sketch Inking. *ACM Transactions on Graphics*, 37(4):98:1–98:14.

Simo-Serra, E., Iizuka and K. Sasaki, S., and Ishikawa, H. (2016). Learning to Simplify: Fully Convolutional Networks for Rough Sketch Cleanup. *ACM Transactions on Graphics*, 35(4):121:1–121:11.

Sun, W., Burie, J., Ogier, J., and Kise, K. (2013). Specific comic character detection using local feature matching. In *2013 12th International Conference on Document Analysis and Recognition*, pages 275–279.

Sýkora, D., Dingliana, J., and Collins, S. (2009). Lazybrush: Flexible painting tool for hand-drawn cartoons. *Computer Graphics Forum*, 28(2):599–608.

Taizan (2016). PaintsChainer. *Preferred Networks*.

Tolle, H. and Arai, K. (2011). Method for real time text extraction of digital manga comic. *International Journal of Image Processing*, 4(6):669–676.

Yao, C., Hung, S., Li, G., Chen, I., Adhitya, R., and Lai, Y. (2017). Manga vectorization and manipulation with procedural simple screentone. *IEEE Transactions on Visualization and Computer Graphics*, 23(2):1070–1084.

Zhang, L., Ji, Y., Lin, X., and Liu, C. (2017). Style transfer for anime sketches with enhanced residual u-net and auxiliary classifier gan. In *2017 4th IAPR Asian Conference on Pattern Recognition*, pages 506–511.

Zhang, L., Li, C., Wong, T.-T., Ji, Y., and Liu, C. (2018). Two-stage sketch colorization. *ACM Transactions on Graphics*, 37(6):261:1–261:14.