TECHNICAL ARTICLE

Portrait Map Art Generation by Asymmetric Image-to-Image Translation

YUXIN ZHANG, FAN TANG, WEIMING DONG, THI-NGOC-HANH LE, CHANGSHENG XU, AND **TONG-YEE LEE**

ABSTRACT

The authors propose a deep neural network–based algorithm to automatically generate portrait map art (PMA), a modern art form created by British portrait artist Ed Fairburn. The authors formulate the generation of PMA as an adaptive dual-to-single image translation problem. The authors' proposed model analyzes the appearance of one portrait and one map image using two encoder networks and utilizes their hidden encodings as representations of the portrait and map image to generate new PMA using a decoder network. An adaptive style harmonization module is proposed to fuse the two hidden encodings. Optimized by cycle-consistency constraint, the model can produce new PMA images without baselines.

Portrait map art (PMA) is a modern art form created by British portrait artist Ed Fairburn. Fairburn uses maps from around the world as canvases for his interests and creativity, aiming to give viewers an abstract sense of connection and possibly stir a sense of "home" and belonging. During the creation process, Fairburn primarily produces portraits, extracting facial features from roads, rivers, and mountain contours by making gradual changes to the map. He then applies ink pens to create realistic areas of shadow and light. Sometimes he also uses an overlay technique, cutting out a specific part of the map and overlaying it upon another to reveal bursts of color [1]. Figure 1a–d shows some of Fairburn's PMA works.

Yuxin Zhang (student), NLPR, Institute of Automation, Chinese Academy of Science and School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100190, China. Email: zhangyuxin2020@ia.ac.cn. ORCID: 0000-0001-6433-2678.

Fan Tang (educator, researcher), School of Artificial Intelligence, Jilin University, Changchun, 130012, Jilin, China. Email: tangfan@jlu.edu.cn. ORCID: 0000-0002-3975-2483.

Weiming Dong (researcher, educator), NLPR, Institute of Automation, Chinese Academy of Science and School of Artificial Intelligence, Beijing, 100190, China. Email: weiming.dong@ia.ac.cn (corresponding author). ORCID: 0000-0001-6502-145X.

Thi-Ngoc-Hanh Le (student), Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan. Email: ngochanh.le1987@gmail.com. ORCID: 0000-0001-9667-9780.

Changsheng Xu (researcher), NLPR, Institute of Automation, Chinese Academy of Science and School of Artificial Intelligence, Beijing, 100190, P.R China. Email: csxu@nlpr.ia.ac.cn. ORCID: 0000-0001-8343-9665.

Tong-Yee Lee (educator, researcher), Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan. Email: tonylee@ncku.edu.tw. ORCID: 0000-0001-6699-2944.

See https://direct.mit.edu/leon/issue/56/1 for supplemental files associated with this issue.

Fairburn calls his process "topopointillism" and describes it as a "direct mix of topography and pointillism" [2]. Like a pointillistic painting, Fairburn's portraits appear abstract up close, but from afar the viewer can see the human subject emerge from the topographic patterns. Fairburn has described his process of creation: "These changes allow me to tease out the human form, resulting in a comfortable coexistence of figure and landscape. I aim to preserve the functionality of each map by feeding the composition instead of fighting it" [3]. By merging landscape and humanity, Fairburn reminds us that we are a product of our environment.

Fairburn often spends hours studying the map's terrain before beginning his artistic process. The sketches he creates are usually digital drafts and layered images created in Photoshop [4], which show the base map and placement of the subject. However, this kind of preview is inefficient because the fusion of subject and map is difficult. Some maps are paired with the subjects in a way that is elaborate but difficult to perceive. The creation of a full PMA work can take from a couple of days to a couple of months [5]. This method also inhibits a quick usable preview and adjustment of the creative scheme to create a better work.

In recent years, computational algorithms for art creation [6–8] and art analysis [9–11] have received growing interest. Especially with the development of machine learning methods such as convolutional neural networks (CNNs) and generative adversarial networks (GANs), which can help researchers with complex mathematical fitting and calculations without strict limitations (e.g. pairing, labeling) of the data, various artistic forms (e.g. paintings, poems, music) can be modeled and automatically created. Thus, to help the artist optimize the creation process of PMAs and explore the possible simulation of this artistic form by computer, we employ machine learning to develop an automatic and efficient PMA image generation approach. Previously, image-to-image translation [12,13] and image style transfer methods [14,15] were used to bridge the appearance gap of images in content, color, and texture and thus render a given natural image into an artistic style. Image

 https://doi.org/10.1162/leon_a_02323

Fig. 1. (a–d) PMA works created by Ed Fairburn [19]. (© Ed Fairburn) (e) Computer generated PMA images. (© Weiming Dong) (a) Pencil on a map. (b) Ink on an Ordnance Survey trench map. (c) Ink drawing on a street map. (d) Ink on a pocket road map. (e) PMA results generated by our model.

compositing methods [16,17] can harmonize multiple objects/contents from different sources in one target scene. However, those methods are not fit for PMA creation, because PMA is a form of compositional art in which visual elements are combined to create a new image that has a style different from both inputs. Shih et al. [18] also showed that classical deep learning-based non-photorealistic rendering style transfer does not work well for creating map art images. Therefore, in this article, we study the automatic creation of this form of digital collage art and use PMA as an experimental case.

We formulate the creation of PMA images as an adaptive dual-to-single image translation problem. Two images are given as guidance for PMA generation: one portrait image and one digital map. The proposed convolutional neural network-based approach automatically fuses these two images into one in the style of PMA works. Notably, given that the volume of existing PMA works by Ed Fairburn is limited, establishing a supervised approach by observing a baseline PMA of given portrait and map inputs is difficult. To this end, we follow CycleGAN [19] and propose two mapping workflows: dual-to-single and single-to-dual. The asymmetric cycle workflows could learn how to couple portrait and map into PMA or decouple PMA into portrait and map in an unsupervised manner. Some PMA images generated by our model are shown in Fig. 1e.

METHODOLOGY

Asymmetric Image Translation

To generate a PMA image with the visual characteristics of both samples of the source domain, we propose an asymmetric image-to-image translation model.

Denote $x_{s_1} \in \{X_{s_1}\}$ and $x_{s_2} \in \{X_{s_2}\}$ as the samples from the two source domains (portrait and map) and $x_t \in \{X_t\}$ as the sample from the target domain (PMA); the goal of the model is to generate portrait map art image $\hat{x}_t$. As shown in Fig. 2a, our model contains two mapping functions, $G : X_{s_1}, X_{s_2} \rightarrow X_t$ and $F : X_t \rightarrow X_{s_1}, X_{s_2}$, and associated adversarial discriminators D_{s_1}, D_{s_2}, and D_t. D_t encourages G to translate X_{s_1}, X_{s_2} into outputs indistinguishable from domain X_t, and vice versa for D_{s_1}, D_{s_2}, and F. We use two cycle-consistency losses to capture the intuition that if we translate from one domain to the other and back again, we should arrive at where we started. Figure 2b shows the forward cycle-consistency loss: $x_{s_1}, x_{s_2} \rightarrow G(x_{s_1}, x_{s_2}) \rightarrow F(G(x_{s_1}, x_{s_2})) \approx x_t$; Fig. 2c shows the backward cycle-consistency loss: $x_t \rightarrow F(x_t) \rightarrow G(F(x_t)) \approx x_{s_1}, x_{s_2}$.

Specifically, as shown in Fig. 3, our translation model $G_{s_1,s_2 \rightarrow t}$ consists of two encoders, E_{s_1} and E_{s_2}, a decoder, Dec_t, and an adaptive multi-style fusion (AMSF) module. Given a portrait image and a map image, we utilize the two encoder

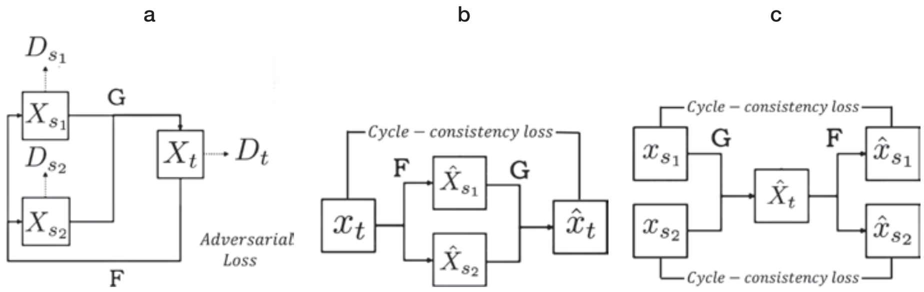


Fig. 2. Overview of our asymmetric image-to-image translation model structure. (a) The full structure. (b) Forward cycle-consistency loss. (c) Backward cycle-consistency loss. (© Weiming Dong)

networks E_{s_1} and E_{s_2} to learn their appearance representations. Then the representations are fused by the AMSF module, which guides our network to focus on the most discriminative areas of different appearances based on the attention map. Finally, a decoder network Dec_t is used to generate a PMA image from the fused representations. Vice versa, our translation model $F_{t \to s_1, s_2}$ can decompose a given portrait map art image into a portrait image and a map image. It consists of one encoder, E_t, two decoders, Dec_{s_1} and Dec_{s_2}, and a two-branch attention module. First, a given PMA image is fed into E_t to get the feature map. Next, the two-branch attention module receives the feature map and gives two attention maps to the decoders Dec_{s_1} and Dec_{s_2}. The cyclic workflow enables the neural network training in an unsupervised learning manner.

Adaptive Multi-Style Fusion Module

As shown in Fig. 3, our AMSF module consists of two pooling layers and an auxiliary classifier η_s. The output of the auxiliary classifier $\eta_s(x_1, x_2)$ represents the probability that x_{s_1} comes from X_{s_1} and x_{s_2} comes from X_{s_2}. $E^k_{s_1}(x)$ and $E^k_{s_2}(x)$ are denoted as the k-th activation maps of C_{s_1} and C_{s_1}, and $E^{k_{ij}}_{s_1}(x)$ and $E^{k_{ij}}_{s_2}(x)$ as the value at (i, j). Our AMSF module is trained to learn the weights w^k_s of the k-th feature maps for the source domains. Specifically, for the portrait domain, we use average pooling to maintain the whole structure, while for the map domain, we use max pooling to capture the details. By exploiting $w^k_{s_1}$ and $w^k_{s_2}$, we can calculate a set of domain-specific attention feature maps:

$$a_{s_1}(x) = w_{s_1} \cdot E_{s_1}(x_1) = \{w^k_{s_1} \cdot E^k_{s_1}(x_1) \,|\, 1 \le k \le n\},$$
$$a_{s_2}(x) = w_{s_2} \cdot E_{s_2}(x_2) = \{w^k_{s_2} \cdot E^k_{s_2}(x_2) \,|\, 1 \le k \le n\},$$

where n is the number of encoded feature maps. Then, our translation model $G_{s_1, s_2 \to t}$ becomes equal to $G_t(a_s(x))$. Finally, we formulate the AMSF module as

$$\eta_s(x_1, x_1) = \sigma(\Sigma_k w^k_{s_1} \Sigma_{ij} E^{k_{ij}}_{s_1}(x_1), \Sigma_k w^k_{s_2} \Sigma_{ij} E^{k_{ij}}_{s_2}(x_2)).$$

Objective

Full objective. We jointly train the encoders, decoders, discriminators, and auxiliary classifiers to optimize the adversarial loss and cycle-consistency loss:

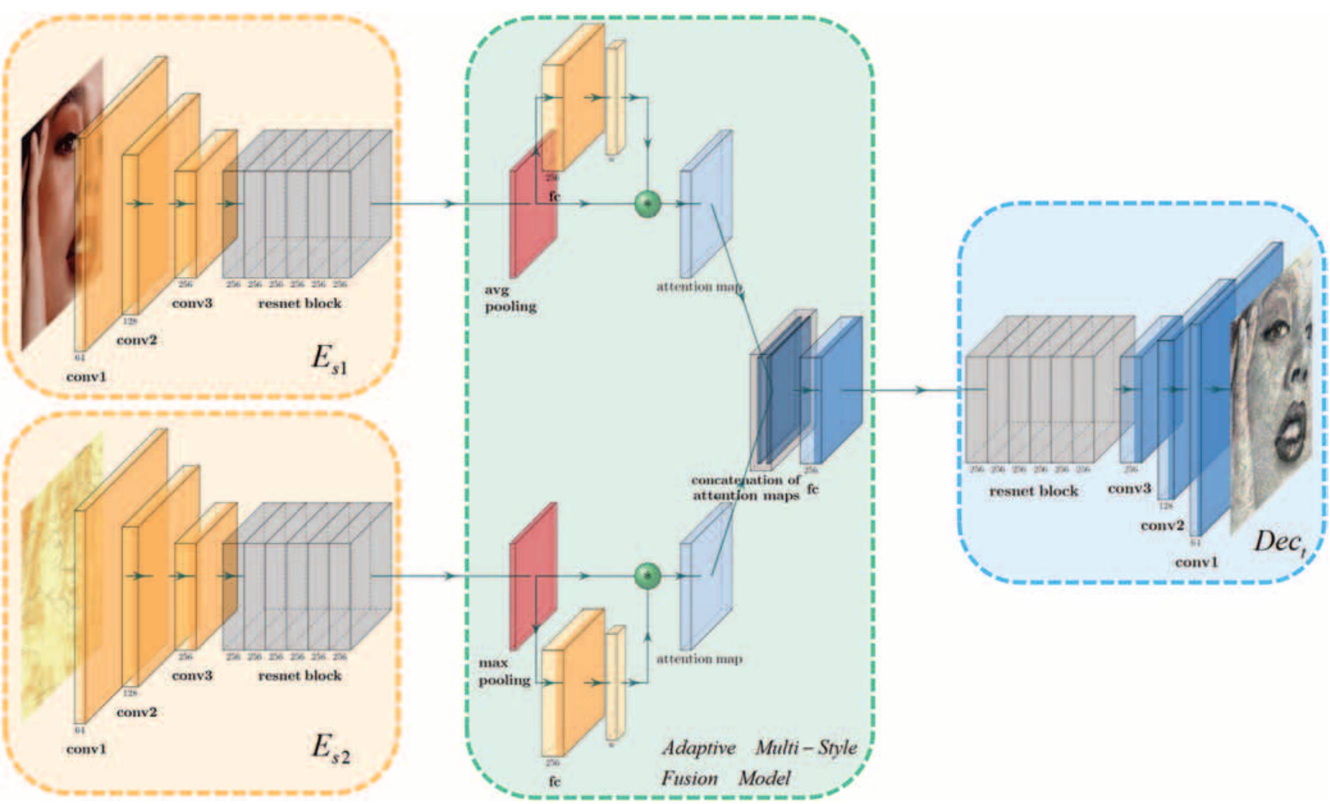


Fig. 3. Model architecture. (© Weiming Dong)

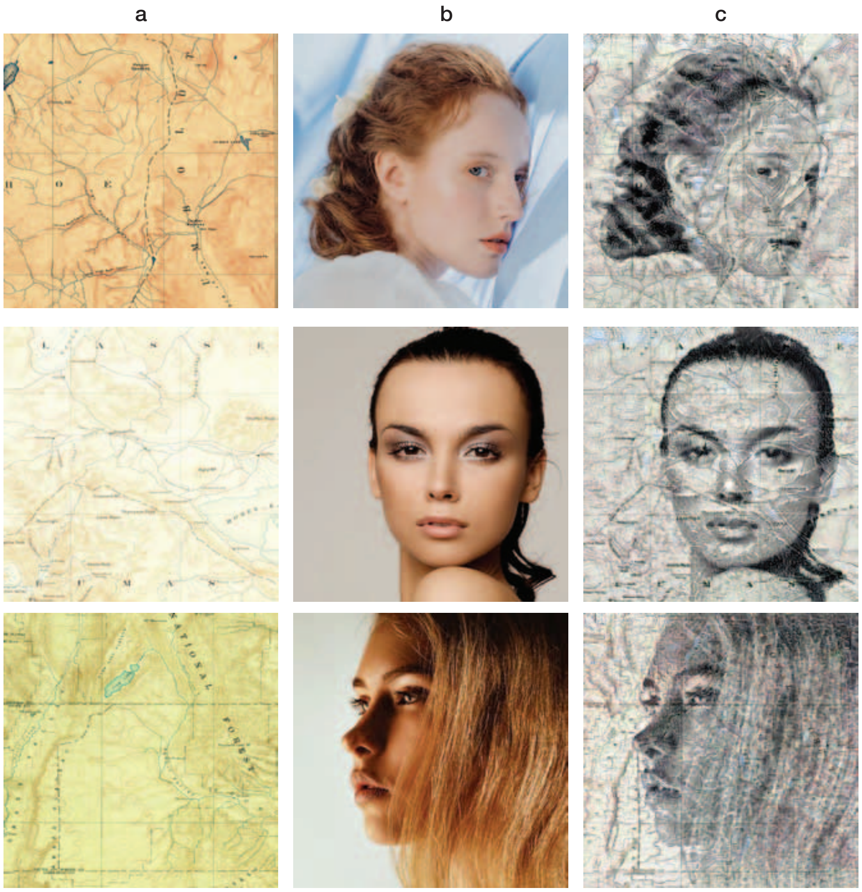


Fig. 4. Automatic PMA generation results with a style like that of Fig. 1(a).
(a) Input map images. (Public domain)
(b) Input portrait images. (Public domain)
(c) Output PMA images. (© Weiming Dong)

$$\min_{G_{s_1,s_2\to t},G_{t\to s_1,s_2}} \max_{D_{s_1},D_{s_2},D_t} \lambda_1 L_{gan}^{s_1,s_2\to t} + \lambda_2 L_{cycle}^{s_1,s_2\to t},$$

$$\min_{G_{s_1,s_2\to t},G_{t\to s_1,s_2}} \max_{D_{s_1},D_{s_2},D_t} \lambda_1 L_{gan}^{t\to s_1} + \lambda_1 L_{gan}^{t\to s_2} + \lambda_2 L_{cycle}^{t\to s_1} + \lambda_2 L_{cycle}^{t\to s_2},$$

where $\lambda_1 = 1$ and $\lambda_2 = 10$.

Adversarial loss. An adversarial loss is employed to match the distribution of the source images to the target image distribution:

$$L_{gan}^{s_1,s_2\to t} = \mathbb{E}_{x\sim X_t}[(D_t(x))^2] + \mathbb{E}_{x_1\sim X_{s_1},x_2\sim X_{s_2}}[(1 - D_t(G_{s_1,s_2\to t}(x_1, x_2)))^2],$$

$$L_{gan}^{t\to s_1} = \mathbb{E}_{x\sim X_{s_1}}[(D_{s_1}(x))^2] + \mathbb{E}_{x\sim X_t}[(1 - D_{s_1}(G_{t\to s_1}(x_t)))^2],$$

$$L_{gan}^{t\to s_2} = \mathbb{E}_{x\sim X_{s_2}}[(D_{s_2}(x))^2] + \mathbb{E}_{x\sim X_t}[(1 - D_{s_2}(G_{t\to s_2}(x_t)))^2].$$

Cycle-consistency loss. To alleviate the mode collapse problem, we apply a cycle-consistency constraint to the generator. Given images $x_{s_1} \in \{X_{s_1}\}$ and $x_{s_2} \in \{X_{s_2}\}$ after the sequential translations from X_{s_1} & X_{s_2} to X_t, and from X_t to X_{s_1} & X_{s_2}, the image should be successfully translated back to the original domain:

$$L_{cycle}^{s_1,s_2\to t} = \mathbb{E}_{x_1\sim X_{s_1},x_2\sim X_{s_2}}[|x - G_{t\to s_1,s_2}(G_{s_1,s_2\to t}(x_1, x_2))|_1],$$

$$L_{cycle}^{t\to s_1} = \mathbb{E}_{x_t\sim X_t}[|x - G_{s_1,s_2\to t}(G_{t\to s_1}(x_t))|_1],$$

$$L_{cycle}^{t\to s_2} = \mathbb{E}_{x_t\sim X_t}[|x - G_{s_1,s_2\to t}(G_{t\to s_2}(x_t))|_1].$$

EXPERIMENTS

Experimental Setup

We collect 63 PMA works by Fairburn and other artists as the training data for the target domain. We also collect 728 portrait images and 153 map images from the Internet to serve as the dual source domains. The portrait and map images are randomly cropped to 400 × 400 px in the training stage. Our model is trained on two GeForce RTX3090 GPUs, taking approximately eight hours. After training, it takes only 1.09 seconds to generate a 400-×-400-px PMA work from arbitrary given inputs.

Results

The appearance of PMA works by artists varies greatly according to the materials or the brush used for painting. Different painting tools lead to various painting styles and correspondingly convey diverse emotions. The 63 collected PMA works could be roughly divided into four categories according to their styles. The representative works of each style are shown in Fig. 1. The proposed algorithm could generate PMAs in different appearances by controlling the style of PMA works used for training.

The PMA work in Fig. 1a was created by using a pencil to draw on an original map of England's Peak District [20]. Such works usually depict shadows in portraits by filling the gaps between the contours of the map finely and are often drawn on color maps. However, when drawing this type of work, the artist only draws portraits in black. Therefore, the fusion of black portraits and color maps is very important. Figure 4 shows the automatically generated PMA works in this style.

a b c

Fig. 5. Automatic PMA generation results with a style like that of Fig. 1(b).
(a) Input map images. (Public domain)
(b) Input portrait images. (Public domain)
(c) Output PMA images. (© Weiming Dong)

a b c

Fig. 6. Automatic PMA generation results with a style like that of Fig. 1(c).
(a) Input map images. (Public domain)
(b) Input portrait images. (Public domain)
(c) Output PMA images. (© Weiming Dong)

Fig. 7. Automatic PMA generation results with a style like that of Fig. 1(d).
(a) Input map images. (Public domain)
(b) Input portrait images. (Public domain)
(c) Output PMA images. (Public domain)

The PMA work in Fig. 1b was created by using ink to draw on an OS trench map of the Western Front in World War I [21]. Such works are characterized by black and white tones and fine brushstrokes. Generating lines of the right length, number, width, and spacing at the right place is a challenge for computer-generated artworks. We selected a specific style of PMA work as the target domain for training and generating new PMA works in this style. To pursue a more refined effect, we used multiple sizes of receptive fields to analyze sketches from multiple scales, which nearly quadrupled our model parameters. Figure 5 shows the automatically generated PMA works in this style.

The PMA work in Fig. 1c was created by using ink to draw on an original street map of Norwich, U.K. [22]. Such works are different from the line drawings in that they often have large black areas and white streets. The model needs to deal with this difference selectively (the former is more like painting on white paper; the latter is more like painting on black paper). Figure 6 shows the automatically generated PMA works in this style.

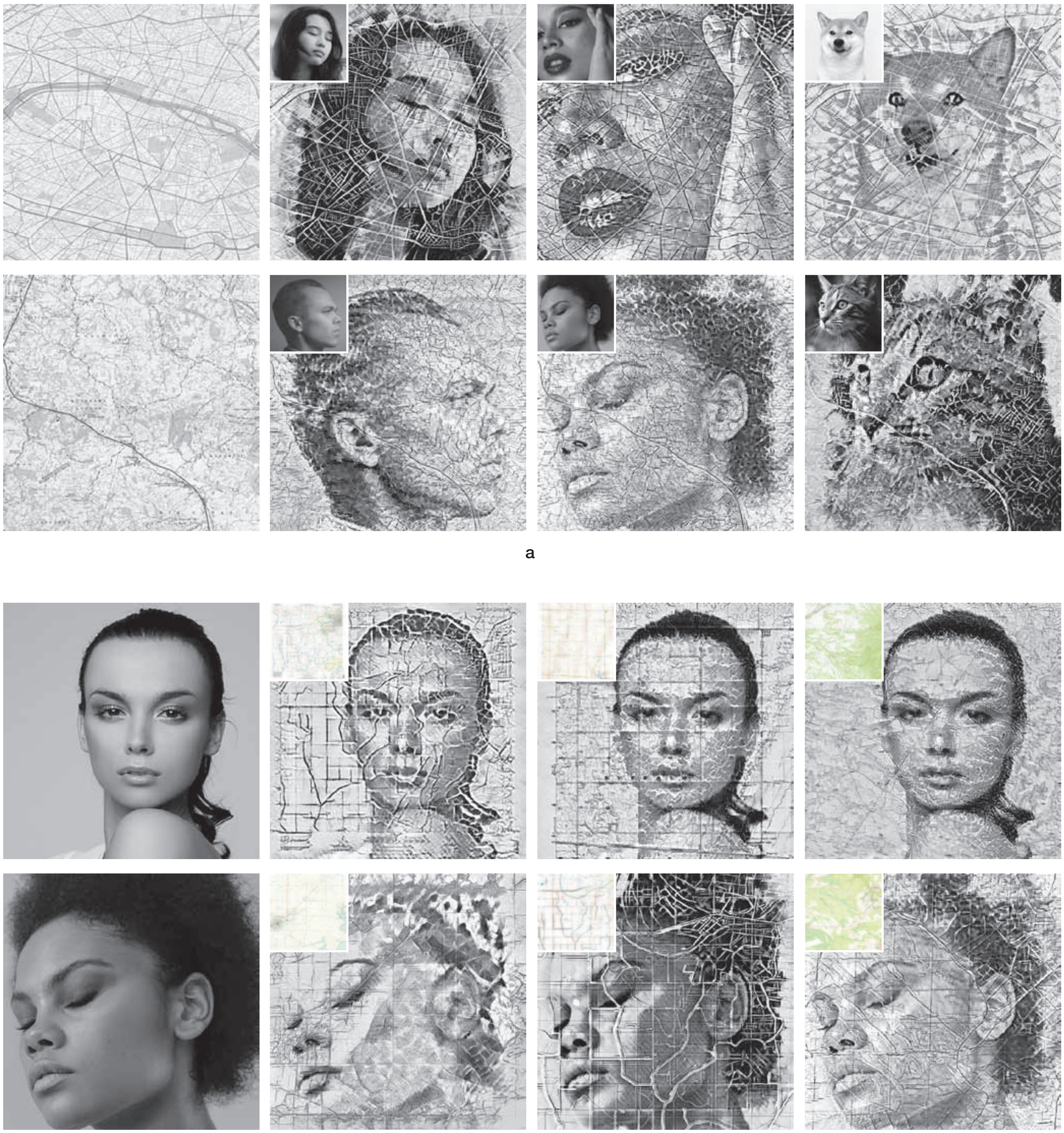
The PMA work in Fig. 1d was created by using ink to draw on a 1973 pocket road map of Germany [23]. The green part of the background and the red part of the lines have a strong visual impact. Therefore, balancing the shade and area of the portrait and the map and strengthening or weakening the colors of the map (especially the red line) are extremely important. Figure 7 shows PMA works generated in this style.

Figure 8 shows some results of using either fixed maps or fixed portraits as inputs. These experiments demonstrate the flexibility and generality of our model. Specifically, in Fig. 8a we show two results generated by using animal photos as inputs, which demonstrates that our approach can create artworks other than standard PMAs.

In all PMA images generated by our model, we can see that the important facial features of the portraits are clearly depicted on the maps in a state of natural composition. On the technical side, compared with the multi-stage map art style transfer approach [24], our method is end-to-end and can more effectively generate PMA images with styles like Fairburn's works. In fact, as orderly as Fairburn's work appears, it's the occasional disorder and unjustified choices that make his work what it is—a human approach to a human concept. Perfection through imperfection. It is through just this quality that our approach can well simulate the visual effects of PAM, since the GAN structure in our model does bring a little randomness to the results.

Limitations

Like other automatic approaches, our approach has limitations in generating specific creation styles that match the artist's interpretation. The water area is automatically vacated while depicting a portrait on the map to match the artist's creative style, which is much more challenging. Possibly this problem can be solved by adding a water area recognition module to our model to help retain the original appearance of the water areas.



a

b

Fig. 8. Automatic PMA generation results. (a) PMA images generated by using the same map but different portraits. (b) PMA images generated by using the same portrait but different maps. (Portrait and map photo: public domain. Art images: © Weiming Dong.)

Another limitation of this generation method is that it is dependent on the suitability of the data being used. The quality of image-to-image translation-based approaches is somewhat dependent on the quality of training data. Thus, we can further fine-tune our model to achieve finer local details in the results by adding more training data, especially PMA data created by the artist.

CONCLUSION

This article presents our recent attempt at the automatic generation of PMA images. Each image is created as a composite of the visual characteristics of a portrait and a map. The approach used is efficient and generic in generating a high-quality image that resembles the artist's style. The advantage of this approach is that it allows users to freely select the portrait and map images for creation according to their own preferences and then obtain the final work in real time. Artists and members of the public without professional programming skills can use our model to create their own PMA images. Artists can also use our results as drafts and create their own works through further modifications.

We believe that there are many opportunities to further refine and extend our approach. A possible immediate extension would be to combine maps and objects to generate general map art images. It might be challenging and interesting to transfer a new map art style to a certain video in which "content" comprises a map art background and animated or video objects. Moreover, we can apply this technique to generate other artistic works with artistic forms like PMAs (a composite of a background and an object in a specific style).

Acknowledgments

This work was supported by the National Key R&D Program of China under No. 2020AAA0106200; by the National Natural Science Foundation of China under Nos. 61832016, U20B2070, 6210070958, and 62102162; by the Ministry of Science and Technology under No. 110-2221-E-006-135-MY3, Taiwan; and by the Open Projects Program of the National Laboratory of Pattern Recognition.

References

1 Natasha Levy, "All Mapped Out: Topographical Portraits by Ed Fairburn": madeinshoreditch.co.uk/2021/02/08/mapped-topographical-portraits-ed-fairburn/ (accessed 8 February 2021).

2 Ed Fairburn: Original artwork and illustration: edfairburn.com/ (accessed 5 August 2021).

3 Emma Taggart, "Artist Draws Detailed Human Portraits Emerging from Map Contours and Patterns" [interview], My Modern Met (30 June 2020): mymodernmet.com/map-art-ed-fairburn/.

4 Dani Willison: "Ed Fairburn interview" (18 September 2014): daniwillison.wordpress.com/2014/09/18/ed-fairburn-interview/ (accessed 18 September 2014).

5 Fairburn [2].

6 Luis Alvarez, Nelson Monzón, and Jean-Michel Morel, "Interactive Design of Random Aesthetic Abstract Textures by Composition Principles," *Leonardo* **54**, No. 2, 179–184 (2021).

7 Thomas Metcalf, "Graphical Data Sets as Compositional Structure: Sonification of Color Graphs in RGB for Clarinet and Piano," *Leonardo* **54**, No. 3, 329–336 (2021).

8 Kang Zhang and Jinhui Yu, "The Computer-Based Generation of Fonts in the Style of Kandinsky," *Leonardo* **54**, No. 4, 437–443 (2021).

9 Prasad Gade et al., "Reactions to Imagery Generated Using Computational Aesthetic Measures," *Leonardo* **50**, No. 5, 453–460 (2017).

10 Everardo Reyes and Lev Manovich, "Cultural Viz: An Aesthetic Approach to Cultural Analytics," *Leonardo* **53**, No. 4, 408–414 (2020).

11 Yingying Deng et al., "Exploring the Representativity of Art Paintings," *IEEE Transactions on Multimedia* **23**, 2794–2805 (2021).

12 Jun-Yan Zhu et al., "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *IEEE International Conference on Computer Vision (CVPR)* (Honolulu, HI, USA, 2017) pp. 2223–2232.

13 Minxuan Lin et al., "Distribution Aligned Multimodal and Multi-Domain Image Stylization," *ACM Transactions on Multimedia Computing, Communications, and Applications* **17**, No. 3, 96:1–96:17 (2021).

14 Dongdong Chen et al., "Explicit Filterbank Learning for Neural Image Style Transfer and Image Processing," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **43**, No. 7, 2373–2387 (2021).

15 Yingying Deng et al., "Arbitrary Video Style Transfer via Multi-Channel Correlation," *AAAI Conference on Artificial Intelligence (AAAI)* (virtual, 2021) pp. 1210–1217.

16 Samaneh Azadi et al., "Compositional GAN: Learning Image-Conditional Binary Composition," *International Journal of Computer Vision*, **128**, 2570–2585 (2020).

17 Shuchen Weng et al., "MISC: Multi-Condition Injection and Spatially-Adaptive Compositing for Conditional Person Image Synthesis," *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (virtual, 2020) pp. 7738–7746.

18 Chiao-Yin Shih, Ya-Hsuan Chen, and Tong-Yee Lee, "Map Art Style Transfer with Multi-Stage Framework," *Multimedia Tools and Applications* **80**, No. 3, 4279–4293 (2021).

19 Zhu [12].

20 Fairburn [2].

21 Fairburn [2].

22 Fairburn [2].

23 Fairburn [2].

24 Taggart [3].

Manuscript received 10 September 2021.

YUXIN ZHANG *is a PhD candidate of the Institute of Automation, Chinese Academy of Sciences. Her research interests include computational visual media and computational creativity.*

FAN TANG *is an assistant professor at the School of Artificial Intelligence, Jilin University. His research interests include computational visual media and computational creativity.*

WEIMING DONG *is a professor at the Institute of Automation, Chinese Academy of Sciences. His research interests include computational visual media and computational creativity.*

THI-NGOC-HANH LE *is a PhD candidate of National Cheng Kung University. Her research interests include computer graphics and visualization.*

CHANGSHENG XU *is a professor at the Institute of Automation, Chinese Academy of Sciences. His research interests include multimedia computing and computer vision. He is a fellow of IEEE.*

TONG-YEE LEE *is a chair professor at National Cheng Kung University. His research interests include computer graphics and visualization.*