

Article

Not peer-reviewed version

TOLGAN: An End-to-end framework for producing Traditional Orient Landscape

Booyong Kim , [Heekyung Yang](#) ^{*} , [Kyungha Min](#) ^{*}

Posted Date: 19 October 2023

doi: 10.20944/preprints202310.1290.v1

Keywords: traditional oriental landscape (TOL); GAN; SPADE; generator; discriminator



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

TOLGAN: An End-to-End Framework for Producing Traditional Orient Landscape

Booyong Kim ¹, Heekyung Yang ^{2,*},[†] and Kyungha Min ^{1,*},[†]

¹ Department of Software, Sangmyung University, Chonan, Korea.

² Department of Computer Science, Sangmyung University, Seoul, Korea;

* Correspondence: yanghk | minkh@smu.ac.kr; Tel.: +82-41-550-5351 | 2-2287-5377 (F.L.)

[†] These authors contributed equally to this work.

Abstract: We present TOLGAN that generates traditional oriental landscape (TOL) image from a map that specifies the locations and shapes of the elements composing TOL. Users can create a TOL map by using a user interface or a segmentation scheme from a photograph. We design the generator of TOLGAN as a series of decoding layers where the map is applied between the layers. The generated TOL image is further enhanced through an AdaIN architecture. The discriminator of TOLGAN processes a generated image and its groundtruth TOL artwork image. TOLGAN is trained through a dataset composed of paired TOL artwork images and their TOL maps. We present a tool through which users can produce a TOL map by specifying and organizing the elements of TOL artworks. TOLGAN successfully generates a series of TOL images from the TOL map. We evaluate our approach using a quantitative way by estimating FID and ArtFID scores and a qualitative way by executing two user studies. Through these studies, we prove the excellence of our approach by comparing our results with those from several important existing works.

Keywords: traditional oriental landscape (TOL); GAN; SPADE; generator; discriminator

1. Introduction

Traditional orient landscape (TOL) is a fine art technique beloved in North-east Asian countries including Korea, China and Japan. The artists in those countries have created various landscapes using brushes with black ink. They express objects in a landscape such as mountains, trees, rocks, cloud, water and buildings in their own abstract shape. The strokes that express salient shape of the objects spread black ink on the paper. Since many well-known TOL artworks are created hundreds years ago, most of the artworks turn gray and yellow.

In computer graphics and computer vision society, many researchers have presented various computational models that produce stroke patterns of TOL brushes [1–10]. They build the skeleton of the strokes using controllable curve and the shape of the strokes using offset from the skeleton curves. Some of them exploit physically-based models for simulating the spread of ink on the surface of the paper. However, they encounter an important limitation in expressing the abstraction and distortion of the objects in TOL.

The progressive of deep learning has presented various schemes that produce images with various styles from the input photographs. These schemes analyze the style from the sampled artworks and apply them to the target photos. The styles are transferred through a texture structure such as Gram matrix or a set of loss functions. These schemes successfully transfer the styles from many famous artworks such as *Starry Nights* by Gogh, *Scream* by Munch and *Portraits* by Picasso to an arbitrary input photograph. These schemes, however, have limitations in transferring TOL style to a photograph. Since TOL artworks are expressed by strokes and spaces, the schemes that transfer texture-based styles are not effective for TOL artworks.

We present a deep learning-based approach for producing TOL artwork images from user-initiated map whose components are collected from many TOL artwork images. We collect many TOL artwork images and analyze the components of the images. Then, we segment the images into the components.

Afterwards, we devise a deep learning-based framework including GAN architecture and AdaIN network that can produce TOL image from a user-initiated map. The GAN architecture is trained using the dataset composed of the pairs of TOL image and its map. The result of GAN architecture can successfully mimic the style of TOL artwork, but it suffers from unclear strokes. To address this limitation, we further apply AdaIN network to enhance the strokes. Our framework can produce visually-convincing TOL artwork images from user-initiated map or an input landscape photograph.

2. Related work

In computer graphics society, several studies have been proposed for simulating traditional oriental paintings [1–10]. Recently, the progress of deep learning techniques presents various schemes for producing traditional oriental paintings.

2.1. Deep learning-based general approach

Gatys et al. [11] presented a pioneering work that transfers a texture-based style captured from a source image to the content of a target image. This scheme can be applied to transfer TOL style to a photograph. Ulyanov et al. [12] and Huang et al. [13] improved Gatys et al.'s work to present a stable and efficient style transfer. These texture-based style transfer schemes are very effective for preserving the content of the target image, but they suffer from improper transfer of the TOL stroke patterns, which are very unique characteristics of TOL artworks.

Generative adversarial network (GAN) [14] presents a very influential framework for synthesizing various styles on images. Mizra and Osindero [15] presented a conditional GAN that considers user-specified conditions for controlling styles and contents in the produced images. Isola et al. [16] presents pix2pix, one of effective schemes that produce style transfer between two domains. Later, Zhu et al. [17] presented CycleGAN that introduces cycle consistency loss to resolve the constraint of paired training dataset of pix2pix framework. Park et al. [18] continued this approach by introducing SPADE that controls the constraints of style transfer using a region map that specifies the contents to produce. Another direction is MUNIT, which was presented by Huang et al. [24], that translates images in multi domains. These schemes present a general-purpose style transfer framework between images of multiple domains. Therefore, they can be employed for producing images of TOL style. However, they have a limitation that the style they synthesize is produced only mimicking stroke patterns of TOL artworks. They cannot synthesize the deformation of objects, which is frequently observed in many TOL artworks.

The general approaches applied for producing traditional oriental paintings show a serious limitation in producing visually pleasing quality, since they does not consider the unique stroke patterns of oriental paintbrush observed in the traditional oriental paintings. They also do not consider the difference of the stroke patterns on different objects in the traditional oriental paintings.

2.2. Deep learning-based TOL generation approach

Li et al. [19] presented a framework that transfers input photograph into Chinese traditional painting style. The abstraction of Chinese traditional painting is trained and expressed through modified xDoG filters, which further process the result images of the generator. The results of the modified xDoG filters are processed through a discriminator with three loss functions including a morphological filter loss. Lin et al. [20] presented a multi-scale GAN model that transfers sketches to Chinese paintings. The generator of this model is trained with L1 and adversarial losses. By adding an edge detector, this model can operate a style generation similar to that of neural style transfer. This model successfully abstracts thick lines of the Chinese paintings, but they cannot properly express the spread of ink.

He et al. [21] presented ChipGAN that produces Chinese ink wash painting style from a photograph. They apply a brush stroke constraint on a generative model to extract an edge-based representation, which is further transferred to a Chinese ink wash painting style using an adversarial

loss. They also apply a cycle consistence loss for the input photograph and the reconstructed photograph from the result of the generator. The result image of the generator is further eroded and blurred with ink wash loss for the final result. Zhou et al. [22] presented Shanshui-DaDA, an interactive and generative model that produces Chinese shanshui painting style using CycleGAN. Users of this model are presented with a web-based interface through which users draw sketches for the Shanshui-DaDA. Xue [23] presented SAPGAN that combines a couple of GAN structures including sketch GAN and paint GAN. Sketch GAN produces an edge map from an input image through either Relative Average Square GAN or StyleGAN2 model, and paint GAN executes edge-to-paint translation using either pix2pix, pix2pixHD or SPADE. The paint GAN is trained with a paired dataset of real Chinese paintings and their corresponding edge maps.

Hung et al. [25] presented UTGAN, which extends CycleGAN for transforming images from very different two image sets. They employ UTGAN for transforming two different image sets such as portraits and Shan-Shui paintings. They devised a special loss term for their UTGAN in order to transform very different image sets. However, they do not support the transformation for the landscape photos and traditional oriental landscapes.

Chung and Huang [26] presented a traditional Chinese painting using a cycle-consistent GAN with pix2pix. They presented a border enhancing scheme in order to enhance the details of border images, which improves the quality of resulting images. However, they suffer from obscure stroke patterns inside an object whose border is rendered using enhanced strokes by their scheme.

The deep learning-based TOL models preserve the stroke patterns of traditional oriental paintings to their input photographs. However, these models have a limitation in applying different stroke patterns to the objects in a scene. We resolve this limitation by sampling the stroke patterns for the objects and applying them to the appropriate objects, which presents very visually pleasing quality.

3. TOL generation framework

3.1. Overview

The overall structure of our TOL generation framework is presented in Figure 1. The input to our model can be either user initiated TOL map or an input photograph, which can be converted to TOL map through an image segmentation model. The input to the generator is a latent vector Z , and the TOL map, which is inserted to each SPADE layer of the architecture. We employ the SPADE module from Park et al.'s work [18]. The result of the generator mimics the style of TOL artwork image, which is further processed through AdaIN architecture to emphasize and to enhance the salient strokes of TOL brushes.

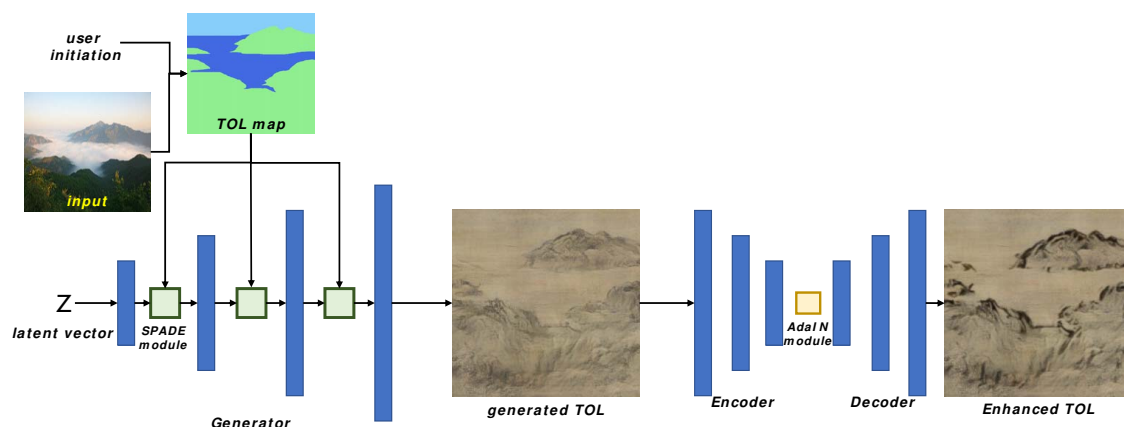


Figure 1. An overview of our TOL generation framework.

3.2. Model architecture

The structure of our model is presented in Figure 2.

Generator		Discriminator
Initial block	Linear(256, 16384)	Conv(64, 4) Stride 2 LeakyReLU
	Reshape(1024, 4, 4)	Conv(128, 4) Stride 2 Instance normalization LeakyReLU
Each block information	SPADE ResBlk(1024) Upsample(2)	Conv(256, 4) Stride 2 Instance normalization LeakyReLU
	Conv(64, 3) with Stride 2 Instance normalization LeakyReLU	Conv(512, 4) Instance normalization LeakyReLU
	SPADE ResBlk(1024) Upsample(2)	
	Conv(128, 3) with Stride 2 Instance normalization LeakyReLU	
	SPADE ResBlk(1024) Upsample(2)	
	Conv(256, 3) with Stride 2 Instance normalization LeakyReLU	
Final Block	3x3Conv-3, Tanh	
	Reshape(8192, 1, 1)	

Figure 2. The structure of our model.

3.2.1. Generator

Our generator has two stages: a TOL generator and an AdaIN module. The TOL generator generates TOL image from a latent vector and a TOL map, which specifies the content of the generated TOL image. The TOL map is processed through SPADE modules that control the content of an image. Between the SPADE modules, a deconvolutional layer with instance normalization is inserted to control the resolution and the channel size of the feature maps. After the processing of TOL generator, an auto-encoder with an AdaIN module processes the generated TOL image to enhance the strokes of TOL. The AdaIN module is pre-trained to enhance the edges of an image and to reinforce the contrast of a scene. The overview of the generator is presented in Figure 1.

3.2.2. Discriminator

For an effective training of our discriminator, we prepare a paired set of real TOL image and generated TOL image. Our discriminator processes a real TOL image and its paired TOL image for an effective training. As a first step, we produce a TOL map from a real TOL image. The map is processed through our generator to a TOL image. This generated TOL image shares the TOL map with the real TOL image. The real TOL image and the generated TOL image are processed through our discriminator.

3.2.3. Loss function

The loss functions we employed for TOLGAN are feature matching loss and hinge loss. The feature matching loss L_{FM} is defined as follows:

$$L_{FM} = \sum_i w_i ||VGG19_i(y) - VGG19_i(G(z))||^2,$$

where y belongs to TOL artwork images and $G(z)$ belongs to generated TOL images. i denotes the i -th layer of the VGG19 network.

The hinge loss L_H is defined as follows:

$$L_H = -E_{z \sim p_z, y \sim p_{data}} [\min(0, -1 + D(D(z), y))] - E_{z \sim p_z, y \sim p_{data}} [D(G(z), y)].$$

4. Training

4.1. Building TOL image dataset

The first step of our building dataset is to collect TOL images. We collect traditional oriental artwork images from various websites and museum archives as many as possible. As a result, we have collected 18K TOL artwork images at this step. At the end of the first step, we select landscape artwork images by excluding portrait, still-life and other inadequate low-quality artwork images. Finally, we select 4,627 traditional oriental landscape (TOL) artwork images for our purpose.

The second step is preprocessing where we apply color correction and resizing for the selected TOL artwork images. The images in our dataset are organized to have 256×256 resolution.

We illustrate some of the collected TOL artwork images in Figure 3.

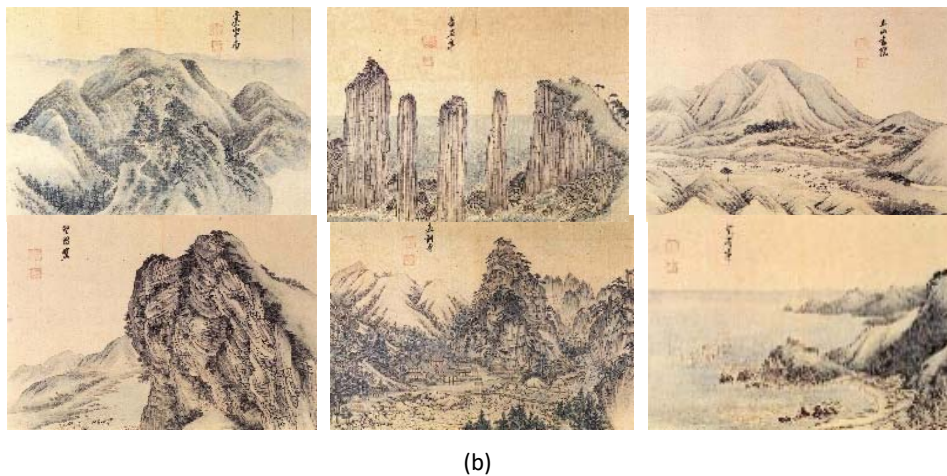
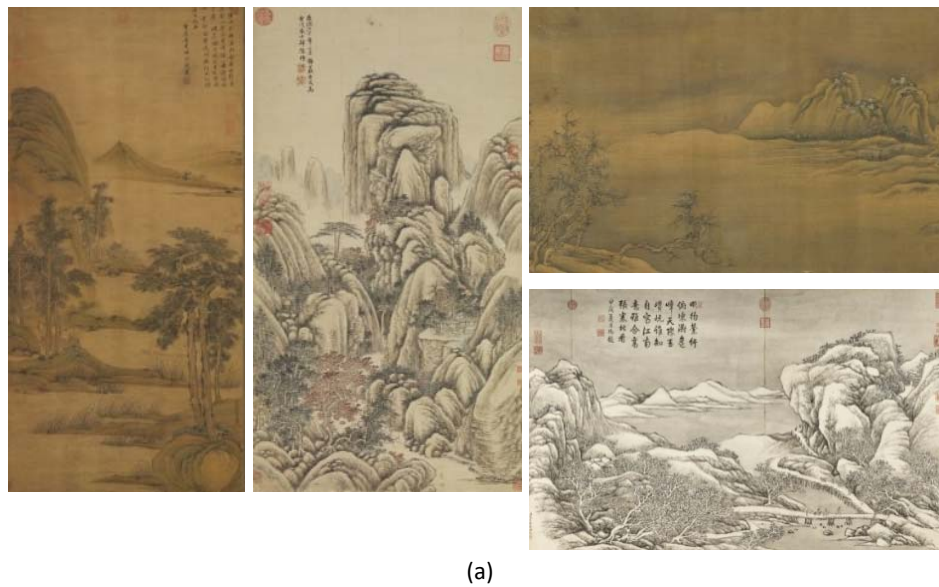


Figure 3. TOL image dataset: (a) The images belong to 18K TOL artwork images, (b) The selected TOL images.

The third step is TOL map generation. We select 253 images among the TOL artwork images and build TOL maps. For this purpose, we implement a TOL map generation tool that specifies regions corresponding to the components of TOL artwork. The components we segment include sky, close mountain, far mountain, ground, rock, grass, water, close tree and far tree.

Finally, the TOL artwork images with TOL maps are augmented by flipping, tone control and three-way rotation. Finally, we train our model with 11,418 images.

The process of building TOL image dataset is illustrated in Figure 4.

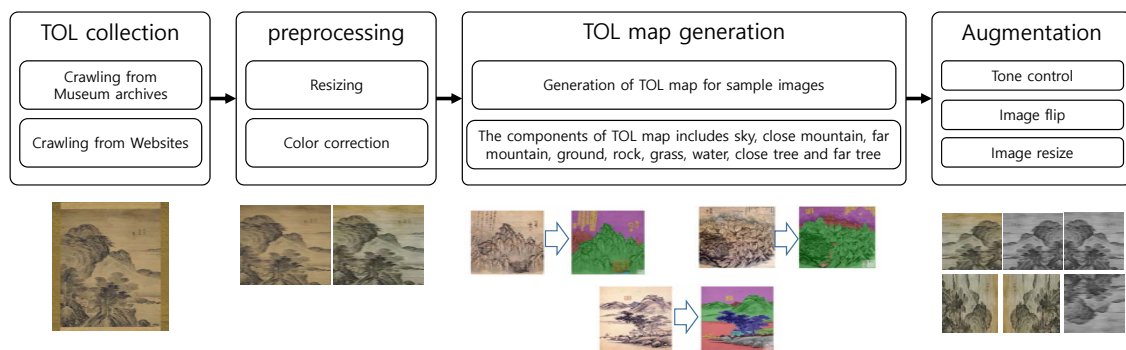


Figure 4. The process of building dataset.

4.2. Training our model

The training process of our model is explained in Figure 5. The paired set of real TOL artwork image with their matching TOL map is employed for the training. The discriminator compares the generated TOL image from the TOL map with its matching real TOL artwork image. This training process is executed for 120 epochs. The training process of 120 epochs takes 100 hours.

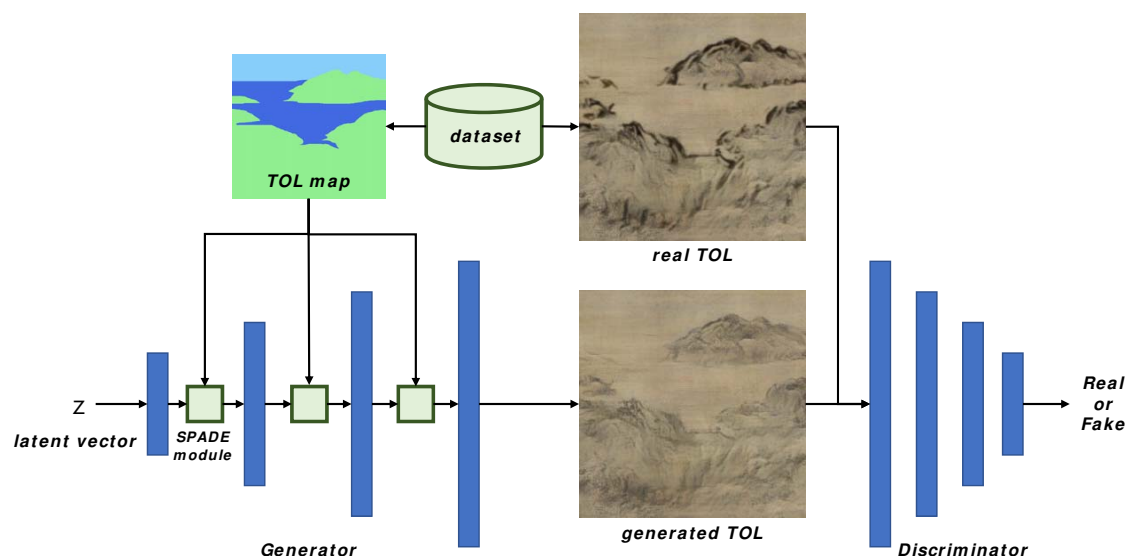


Figure 5. Training of GAN architecture.

5. Implementation and results

We implemented our TOL generation framework on a personal computer with double nVidia RTX3090 GPUs. The software environment is Python 3.9 with PyTorch 1.12.1 and torchvision 0.13.1. We employed Adam optimizer whose β_1 is 0.1 and β_2 is 0.9. The learning rate is 0.0002.

After the training we implement a tool with a user interface for generating TOL images (See Figure 6). This interface presents nine components of TOL artwork images. The components cover sky, close mountain, far mountain, ground, rock, grass, water, close tree and far tree. Users select the component to insert and draw the shape of components in the left canvas of the interface. After finishing the drawing, they select the resolution of the image to generate using the job shuttle and generate TOL image.

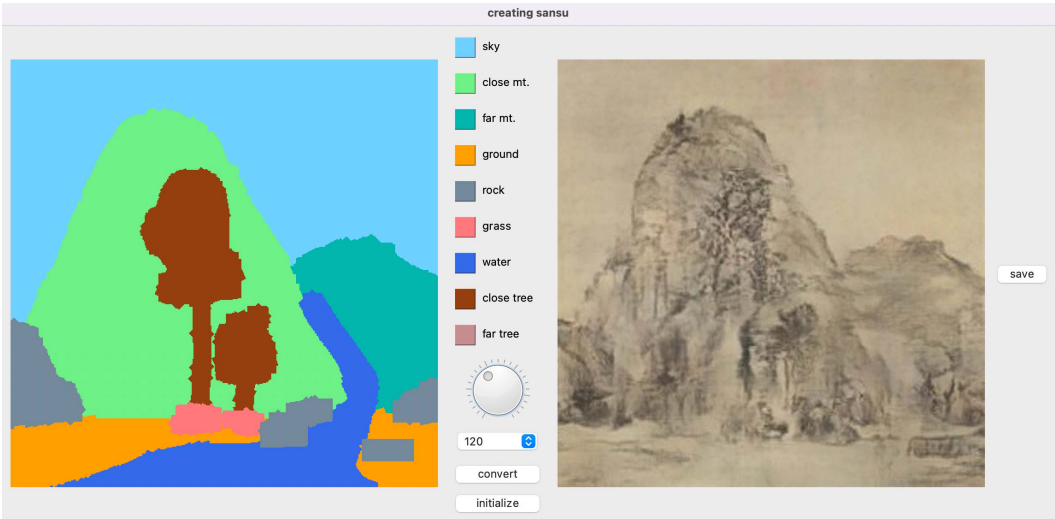


Figure 6. The user interface of our model.

We present twelve TOL map images in Figure 7 and their corresponding TOL images in Figure 8. For the comparison, some of them are generated from photographs.

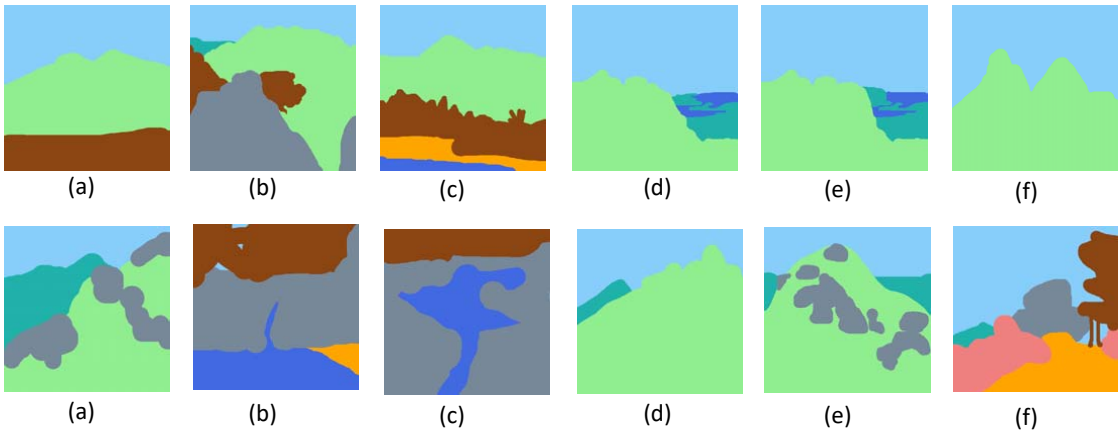


Figure 7. Input semantic maps for TOL generation.

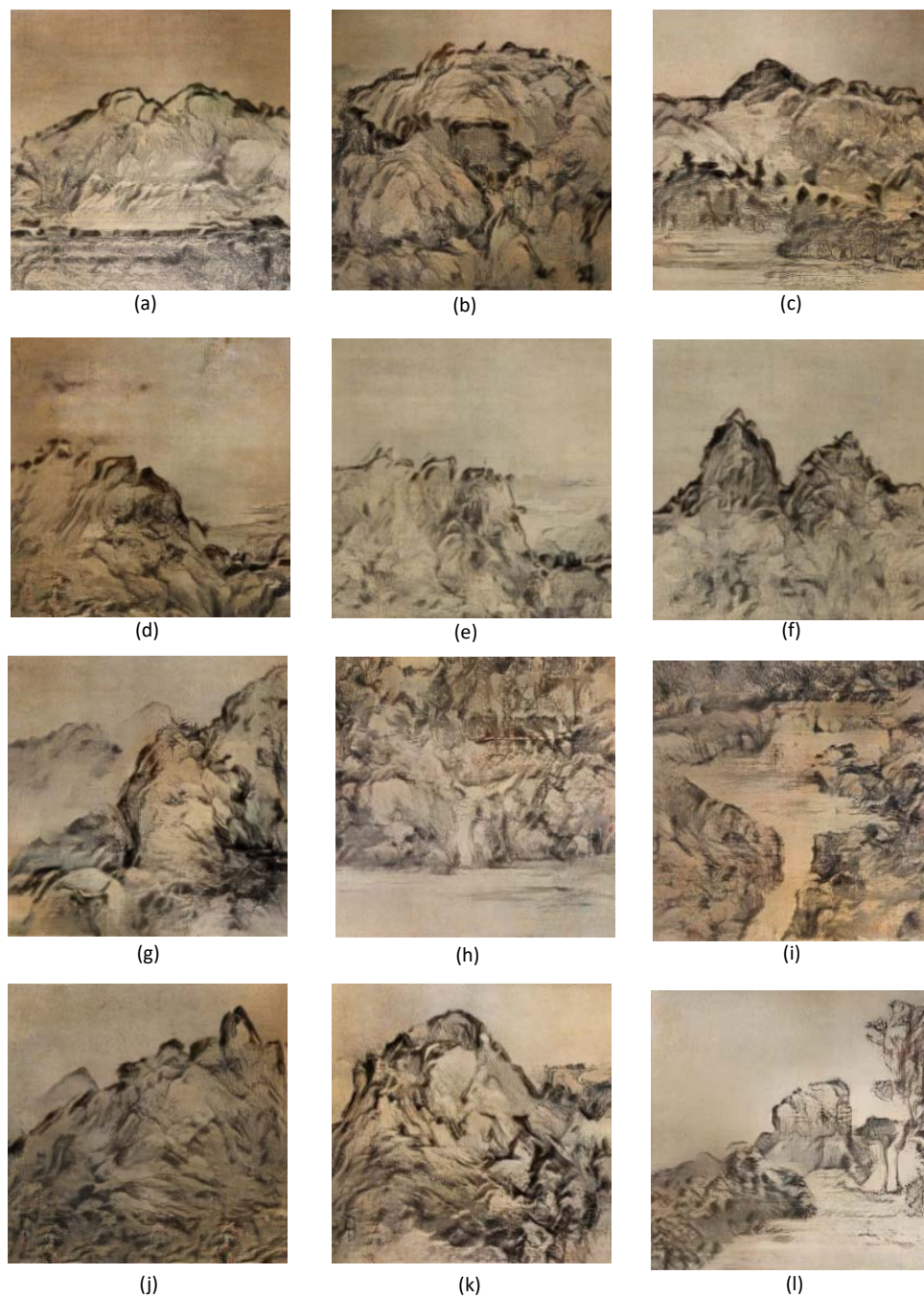


Figure 8. The generated TOL images.

6. Analysis

6.1. Comparison

We compare the results from our model with those from several important existing works including MUNIT [24], ChipGAN [21], CycleGAN [17] and neural style transfer (NST) [11]. Among them, MUNIT and CycleGAN are general-purposed GAN-based model that transfers styles between various domains of images. We re-train MUNIT and CycleGAN using the collected TOL images and landscape photographs. ChipGAN is a designated GAN-based model for TOL image generation. Neural style transfer is one of the most pioneering works for applying the styles captured from TOL to a photograph. The results are presented in Figure 9.

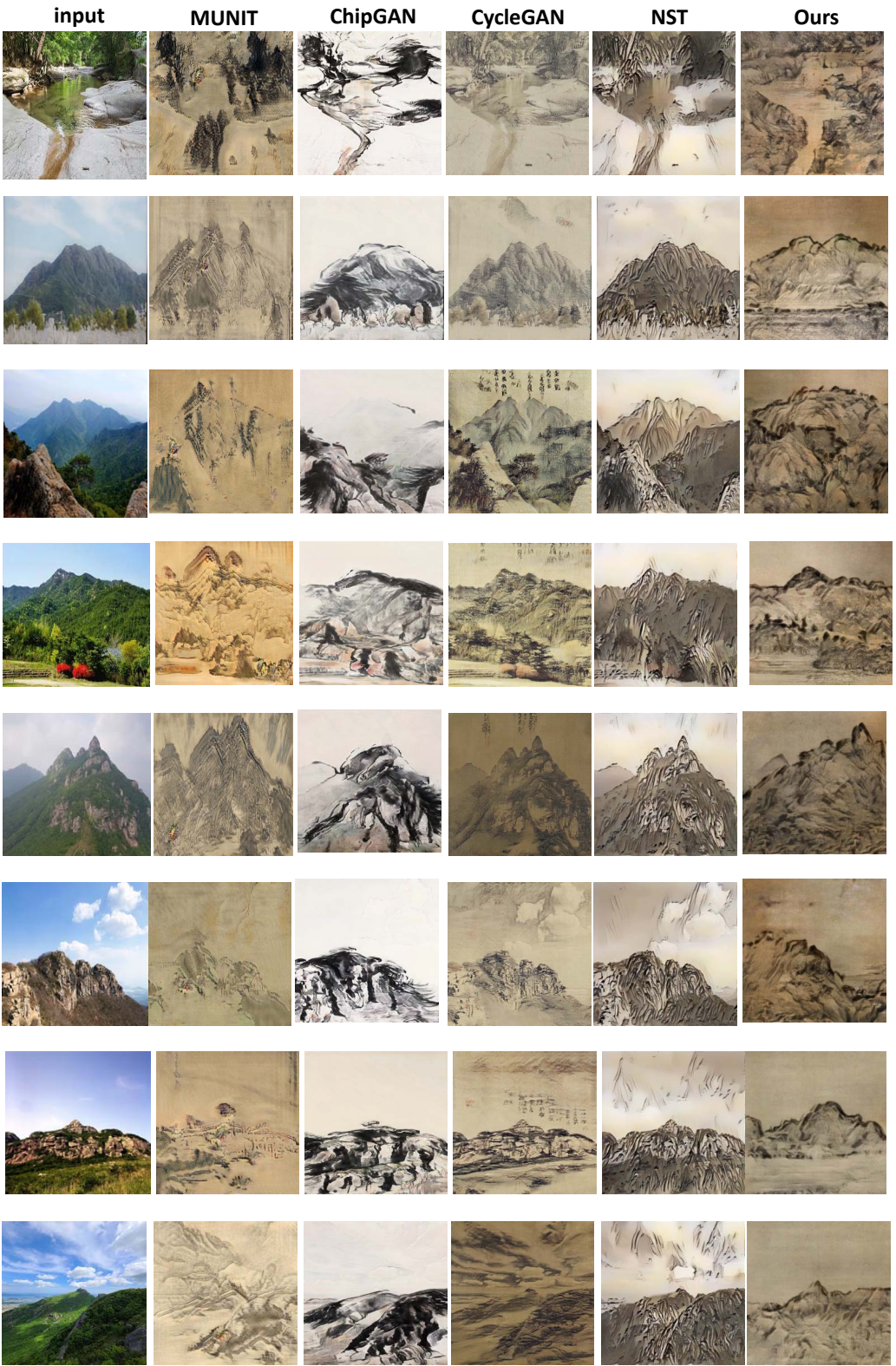


Figure 9. Comparison of our results with those from existing works.

6.2. Quantitative evaluation

For the quantitative evaluation of our study, we estimate Frechet Inception Distance (FID) and ArtFID [27] for the TOL images in Figure 9. We estimate both FID and ArtFID for the results from four

existing schemes including MUNIT, CycleGAN, ChipGAN and NST are compared with ours for the eight input landscape images. The FID and ArtFID values from the results of four existing schemes and ours are suggested in Table 1.

Table 1. The FID and ArtFID scores for the four existing studies and ours. The red figures denote the smallest score.

	MUNIT	ChipGAN	CycleGAN	NST	Ours
FID	213.5	235.7	198.4	187.6	154.5
ArtFID	185.7	176.3	156.3	163.3	131.5

6.3. Qualitative evaluation

We evaluate our results by comparing them with those from the existing works in two experiments: a user study and a focus group evaluation. The user study is executed for ordinary people who do not have special backgrounds on TOL artworks, and a focus group evaluation is executed for those who have some backgrounds on TOL artworks.

6.3.1. User study

We hire thirty volunteer participants for a user study. nineteen of them are in their twenties and eleven of them are in their thirties. seventeen of them are female and thirteen are male. They are required to have not special backgrounds on TOL painting artworks.

We have two questions for the images in Figure 9.

1. Question 1 (Quality of the generated image): Among the five result images, select one image that most resembles TOL style.
2. Question 2 (Preservation of input photograph): Among the five result images, select one image that most resembles the input image.

Each participant marks one model for the eight images in Figure 9. We summarize the marks from the participants and present the results in Table 2. For question 1, the scores of our model are higher than those of the other models for all sample images. For question 2, our model records higher scores for five images among eight sample images. Even though our model outperforms for generating TOL style, some of the models show better results for preserving shapes of the objects.

Table 2. The results of a user study for two questions: The red figure denotes best result.

question	models	01	02	03	04	05	06	07	08	sum
1	MUNIT	0	0	0	0	0	0	0	0	0
	ChipGAN	4	6	0	9	3	12	10	12	56
	CycleGAN	0	0	5	5	0	0	4	0	14
	NST	0	4	4	1	2	4	3	2	20
	ours	26	20	21	15	25	14	13	16	150
2	MUNIT	0	0	0	0	0	0	0	0	0
	ChipGAN	3	2	0	8	1	10	7	9	40
	CycleGAN	0	5	5	8	2	3	5	0	28
	NST	6	12	11	3	14	8	8	10	72
	ours	21	11	14	11	13	9	10	16	150

6.3.2. Focus group evaluation

We hire ten experts who have an experience of painting TOL artworks. All of them major in oriental paintings. Three of them are senior undergraduate students, four of them are graduate students, and three of them are professional oriental painting artists. Four of them are female and six of them are male.

We ask the participants to evaluate the completeness of the TOL images generated by our model in ten-point metric. The points range in (1 ~ 10). They are instructed to mark the generated TOL images by assuming the real TOL artworks have 10 points. The results are presented in Table 3. Among the eight sample images, our results record highest scores for six images.

Table 3. The results of a focus group evaluation: The red figure denotes best result.

models	01	02	03	04	05	06	07	08	average
MUNIT	2.5	1.8	2.2	2.1	2.2	2.1	2.2	1.9	2.13
ChipGAN	4.2	4.7	3.7	7.2	6.3	7.2	7.0	6.4	5.84
CycleGAN	3.3	4.0	5.8	5.7	2.9	5.4	5.7	1.7	4.31
NST	2.9	5.8	5.2	4.2	6.1	6.3	6.9	4.5	5.24
ours	6.9	7.4	7.6	7.7	7.6	6.5	7.1	6.3	7.14

6.4. Ablation study

For an ablation study of our model, we compare two sets of generated TOL images in Figure 10. The images in the upper row are generated without AdaIN module that enhances the generated TOL images. The images in the lower row are produced by processing the images in the upper row through the AdaIN module. In comparing the generated TOL images in both upper and lower row, we can conclude that TOL generator using SPADE module has a limitation in producing thick stroke patterns of TOL artwork images. By adding AdaIN module for the enhancement, we can generate TOL images that mimic the thick stroke patterns of TOL artwork.

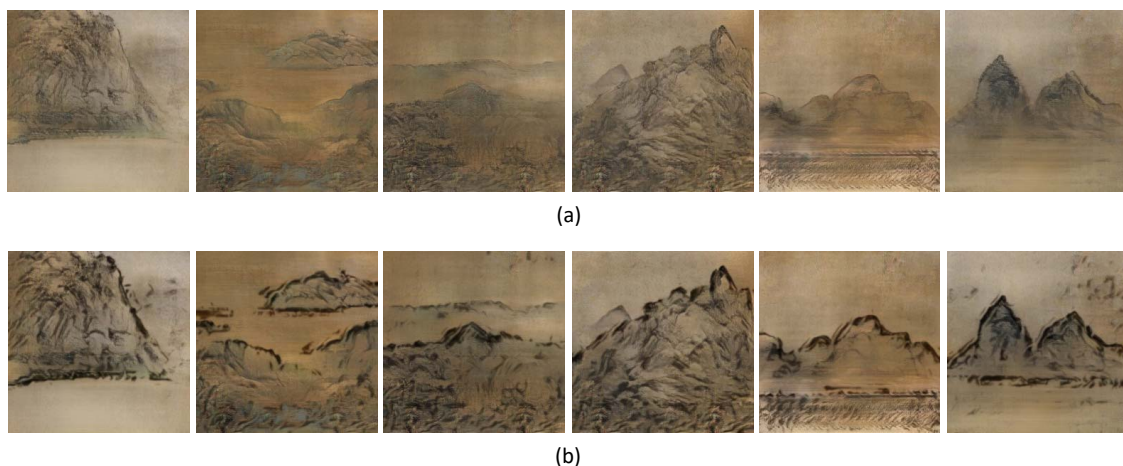


Figure 10. Ablation study of our results.

6.5. Limitation

We analyze the most serious limitation of our model is that our model depends on the segmentation map produced by users. Even though we apply a segmentation scheme for an input photograph for an automatic production of segmentation map, users are required to modify the produced map for the TOL map, the input of our model. This process requires a burden for an automatic TOL image generation.

Another limitation of our model is that the generated TOL images from our model depends on the components of the TOL map. Since we have selected components of frequently appearing components of real TOL artwork images, some components such as person or animals are neglected.

The other limitation of our model is that the result images show yellow-paper effect, which is frequently observed for papers more than hundred years. We have collected real TOL artwork images from many museums. Therefore, the collected TOL artwork images are painted on yellow papers. Some contemporary TOL artwork images are required to be collected for produce diverse TOL images.

7. Conclusion and future work

We have presented TOLGAN that generates TOL images from TOL map that specifies the locations and shapes of the elements composing TOL. TOL map is produced by a user interface or a segmentation scheme from a photograph. The generator of TOLGAN is developed as a decoder structure where SPADE module is inserted between the layers. The SPADE module processes TOL map to control the result of TOL image. The generated TOL image is further enhanced through AdaIN structure to complete a visually-pleasing TOL image.

The results of our approach are evaluated through both quantitative and qualitative evaluations. For the quantitative evaluation, we estimate FID and ArtFID scores for our results and the results from important existing studies and show that ours show smallest FID and ArtFID scores. For the qualitative evaluation, we execute two user studies including volunteers and focus group. These user studies show our framework produces better TOL images than the compared studies.

For a further work, we aim to enhance our model that can produce TOL artwork images without TOL maps. In order to produce scholastic TOL artwork images, TOL map should be produced automatically. We aim to design a model that can produce TOL maps from scratches. Another direction is to extent our TOL generation model for various applications such as TOL animation. TOLGAN can automatically generate background images of TOL animation.

References

1. Strassmann, S., "Hairy Brushes," ACM Computer Graphics Vol. 20, No. 4, pp. 225-232, 1986.
2. Guo, Q. and Kunii, T., "Modeling the Diffuse Paintings of 'Sumie'," Proc. of Modeling in Computer Graphics 1991, pp. 329-338, 1991.
3. Lee, J., "Simulating Oriental Black-ink Painting," IEEE Computer Graphics and Applications, Vol. 19, No. 3, pp. 74-81, 1999.
4. Lee, J., "Diffusion Rendering of Black Ink Paintings using new Paper and Ink Models," Computer & Graphics, Vol. 25, No. 2, pp. 295-308, 2001.
5. Way, D., Lin, Y., and Shin, Z., "The Synthesis of Trees in Chinese Landscape Painting using Silhouette and Texture Strokes," Journal of WASC, Vol. 10, pp. 499-506, 2002.
6. Huang, S., Way, D., and Shih, Z., "Physical-based Model of Ink Diffusion in Chinese Ink Paintings," Proc. of WSCG 2003, pp. 33-40, 2003.
7. Yu, J., Luo, G., and Peng, Q., "Image-based Synthesis of Chinese Landscape Painting," Journal of Computer Science and Technology, Vol. 18, No. 1, pp. 22-28, 2003.
8. Xu, S., Xu, Y., Kang, S., Salesin, D., Pan, Y., and Shum, H., "Animating Chinese Paintings through Stroke-based Decomposition," ACM Transactions on Graphics Vol. 25, No. 2, pp. 239-267, 2006.
9. Zhang, S., Chen, T., Zhang, Y., Hu, S., and Martin, R., "Video-based Running Water Animation in Chinese Painting Style," Science in China Series F: Information Sciences, Vol. 52, No. 2, pp. 162-171, 2009.
10. Shi, W., "Shan Shui in the World: A Generative Approach to Traditional Chinese Landscape Painting," Proc. of IEEE VIS 2016 Arts Program, pp. 41-47, 2016.
11. Gatys, L., Ecker, A., and Bethge, M., "Image Style Transfer Using Convolutional Neural Networks," Proc. of CVPR 2016, pp. 2414-2423, 2016.
12. Ulyanov, D., Lebedev, V., Vedaldi, A., and Lempitsky, V., "Texture Networks: Feed-forward Synthesis of Textures and Stylized Images," Proc. of ICML 2016, pp. 1349-1357, 2016.
13. Huang, X., and Belongie, S., "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization," Proc. of ICCV 2017, pp. 1501-1510, 2017.
14. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., "Generative Adversarial Nets," Proc. of NIPS 2014, pp. 2672-2680, 2014.
15. Mirza, M., and Osindero, S., "Conditional Generative Adversarial Nets," arXiv preprint arXiv:1411.1784, 2014.
16. Isola, P., Zhu, J. Y., Zhou, T., and Efros, A., "Image-to-image Translation with Conditional Adversarial Networks," Proc. of CVPR 2017, pp. 1125-1134, 2017.

17. Zhu, J., Park, T., Isola, P., and Efros, A., "Unpaired Image-to-image Translation using Cycle-Consistent Adversarial Networks," *Proceedings of ICCV 2017*, pp. 2223-2232, 2017.
18. Park, T., Liu, M., Wang, T. and Zhu, J., "Semantic Image Synthesis with Spatially-Adaptive Normalization," *Proc. of CVPR 2019*, pp. 2337-2346, 2019.
19. Li, B., Xiong, C., Wu, T., Zhou, Y., Zhang, L., and Chu, R., "Neural Abstract Style Transfer for Chinese Traditional Painting," *Proc. of ACCV 2018*, pp. 212-227, 2018.
20. Lin, D., Wang, Y., Xu, G., Li, J., and Fu, K., "Transform a Simple Sketch to a Chinese Painting by a Multiscale Deep Neural Network," *MDPI Algorithm*, Vol. 11, No. 1, Article No. 4, 2018.
21. He, B., Gao, F., Ma, D., Shi, B., and Duan, L., "Chipgan: A Generative Adversarial Network for Chinese Ink Wash Painting Style Transfer," *Proc. of ACM Multimedia 2018*, pp. 1172-1180, 2018.
22. Zhou, L., Wang, Q.-F., Huang, K., and Lo, C.-H.Q., "ShanshuiDaDA: An Interactive and Generative Approach to Chinese Shanshui Painting Document," *Proc. International Conference on Document Analysis and Recognition 2019*, pp. 819-824, 2019.
23. Xue, A., "End-to-End Chinese Landscape Painting Creation Using Generative Adversarial Networks," *Proc. of WACV 2021*, pp. 3863-3871, 2021.
24. Huang, X., Liu, M.-Y., Belongie, S., and Kautz, J., "Multimodal Unsupervised Image-to-Image Translation," *Proc. of ECCV 2018*, pp. 172-189, 2018.
25. M. Hung, M. Trang, R. Nakatsu, and N. Tosa, "Unusual Transformation: A Deep Learning Approach to Create Art," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 422, pp. 309-320, 2022.
26. C. Chung and H. Huang, "Interactively transforming chinese ink paintings into realistic images using a border enhance generative adversarial network" *Multimedia tools and applications*, Vol. 82, pp. 11663-11696, 2023.
27. Wright, M. and Ommer, B., "ArtFID: Quantitative Evaluation of Neural Style Transfer," *Proc. of German Conference on Pattern Recognition 2022*, pp. 560-576, 2022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.