Neural Style Transfer With Adaptive Auto-Correlation Alignment Loss

Zheng Zhou, Yue Wu, Xiaofei Yang, and Yicong Zhou, *Senior Member, IEEE*

***Abstract*—The neural style transfer has achieved a significant improvement with deep learning methods. However, the existing methods are susceptible to lack the ability for handling the texture style transfer because of their less consideration of the textural structure from style images. To overcome this drawback, this letter presents a simple method to capture the textural structure by using an adaptive auto-correlation alignment loss function. Furthermore, we also introduce three metrics to quantitatively evaluate the performance. We qualitatively and quantitatively evaluate the proposed methods. The experimental results demonstrate the superiority of the proposed method and our method can synthesize the stylized images with rich texture style patterns.**

***Index Terms*—Neural style transfer, auto-correlation, structure preserving, adaptive periodicity.**

Fig. 1. Comparative results of style transfer methods. Richer style texture patterns (repeated red woven structure) are shown in the proposed method than Gatys [8]. Gatys' result loses the repeated texture style patterns.

I. Introduction

The goal of style transfer [1]–[3] is to synthesize a new image based on the content image and style image. In recent years, lots of style transfer methods based on the neural network have been proposed and performed very promising results [4]–[7]. For example, Gatys *et al.* [8] first proposed the gram matrices to represent the style features and showed that the correlation between feature layers in the pre-trained neural network can describe the style pattern well. It is a single model and iteratively minimizes the style loss between the synthesized images and input style images, resulting in synthesizing a new image.

Since the single style transfer methods [9]–[13] are capable of capturing the correlation between features, they have achieved very exciting results. However, these methods cost a lot of computing resources and time during the training step. Therefore, lots of feed-forward methods have been presented to reduce the computational time and efficiently synthesize the stylization images [14]–[17]. These methods that called universal style transfer learn a transformation function using a deep learning model and synthesize the stylized image based on the features extracted from the content-style pair inputs. For instance, Huang *et al.* [18] propose the AdaIN algorithm to learn the style pattern with the mean and variance of the deep features and synthesize the stylization image by minimizing the style loss with style image and content loss with content image. Li *et al.* [19] present a new style transform method called WCT using the whitening and coloring transform to transfer the style feature into content. The universal style transfer methods are encoder-decoder-based approaches, in which the encoder is applied to extract features and the symmetric decoder is adapted to synthesize the stylized image. However, these methods directly compute the transform matrices from the deep features, which could not offer a general solution for style transfer. To overcome these problems, Li *et al.* [20] propose an efficient style transfer method and achieve high-quality arbitrary style transfer using a linear feed-forward module. Park *et al.* [21] adjust the content features to match the style features through self-attention mapping and preserve the content structure by using identity loss. Kalischek *et al.* [22] develop a dual central moment discrepancy (CMD) to match the feature distribution between the generated image and style image. The style transfer includes color and texture style transfer. However, the above-mentioned methods fail to fully consider the texture style in the style images, resulting in generating the unpleasant synthesized images and destroying the complex content structure.

To address these issues, we propose a 2D adaptive auto-correlation alignment (ACA) loss function and apply it in style transfer to capture the structural pattern from the style image without corruption of complex content spatial layout (as shown in Fig. 1). Furthermore, we adopt the deception score to evaluate the performance quantitatively. The main contributions of our work are as follows:

- We propose a new and simple component called adaptive auto-correlation alignment loss function to capture more abundant style features, and apply it into the style transfer methods.
- We further introduce three metrics into the style transfer, which to our knowledge, it is the first time to quantitatively evaluate the performance of the style transfer methods by using these metrics.

Manuscript received December 20, 2021; revised March 27, 2022; accepted April 2, 2022. Date of publication April 7, 2022; date of current version May 2, 2022. This work was supported by University of Macau under Grant MYRG2018-00136-FST. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Minglun Gong. *(Corresponding author: Yicong Zhou.)*

Zheng Zhou, Xiaofei Yang, and Yicong Zhou are with the Department of Computer and Information Science, University of Macau, Macau 999078, China (e-mail: alexzhouzhengz@gmail.com; xiaofeiyang@um.edu.mo; yicongzhou@um.edu.mo).

Yue Wu is with the Amazon Alexa Natural Understanding, Manhattan Beach, CA 90266 USA (e-mail: wuayue@amazon.com).

Digital Object Identifier 10.1109/LSP.2022.3165758

- We conduct experiments to quantitatively and qualitatively evaluate the style transfer methods, which indicates the superiority of our proposed method.

II. Related Work

Since the deep learning models could extract many abundant features, many deep learning-based methods are proposed for neural style transfer [23]–[25]. VGG [26] and ResNet [27] are the famous basic networks that are always used to extract deep features. For example, Gatys *et al.* [8] firstly applied the VGG network into the style transfer to extract the deep features and presented the gram matrix to compute the style loss between the generated image and the input style image. This approach stylized the content image per style per model. To improve the transfer ability, some researchers proposed arbitrary style transfer methods to transfer any style to generate new images. For instance, Huang *et al.* [18] proposed a universal style transfer method called AdaIN by using the adaptive instance normalization. In AdaIN, the pre-trained VGG network was utilized to extract the deep features, while the instance normalization layer was used to capture the statistical information. However, the AdaIN could not explore the inner relationship between features with simple statistical information. In order to explore the inner relationship between the features, Li *et al.* [19] employed the whitening and coloring transform and presented a new style transfer method called WCT. It employed the matrix manipulation procedure to process the style features, resulting in the satisfactory stylized result. However, matrix computation is often highly expensive in computation. Thus, Li *et al.* [20] utilized an auto-encoder module to process the style features rather than the matrix computation and proposed a style transfer method, i.e., LST. Park *et al.* [21] introduced a style-attention module into the auto-encoder network and proposed a new style-attention network (SAN) for extracting the features both from the style and content images.

However, all these above-mentioned methods fail to make full use of the structural pattern while handling the style features. In this work, we propose an adaptive auto-correlation alignment loss function to extract the texture style, including shapes, periodicity, and structure pattern. It is noted that the proposed loss can be adapted into the existing style transfer models. We further introduce three metrics to evaluate the performance of the style transfer methods.

III. Methodology

A. Overview

The proposed method (as shown in Fig. 2) is established based on the style transfer model and an adaptive auto-correlation alignment loss function. We adopt the same setting with SAN [21] in the content-style (CS) feature transfer module. We will introduce the auto-correlation alignment (ACA) loss function.

B. Auto-Correlation Alignment (ACA) Loss

The goal of auto-correlation [28]–[30] is to capture the co-occurrence of the inputs, which is usually used in 1D signal processing to predict future growth. For a time series $s(t)$, $t = 0, 1, \ldots, N$, the auto-correlation with lag τ is defined as:

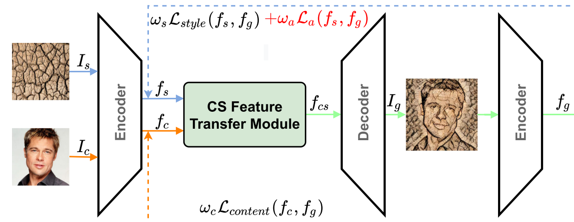

Fig. 2. Structure of style transfer with adaptive auto-correlation alignment loss. CS feature transfer module means the content-style feature transfer module which converts the style feature into content. ω_a and $\mathcal{L}_a$ denote the weight and adaptive auto-correlation alignment loss defined in (4).

$$a_\tau(s) = \frac{1}{N}\sum_{t=0}^{N} s(t)\cdot s(t+\tau), \tag{1}$$

where the a_τ denotes the auto-correlation of time series signal with the lag of τ. For a matrix M that has a size of $W \times H$, the 2D auto-correlation is formulated by

$$\mathcal{A}_{\delta x,\delta y}(M) = \frac{1}{WH}\sum_{x=0}^{W-1}\sum_{y=0}^{H-1} M(x,y)\cdot M(x+\delta x, y+\delta y), \tag{2}$$

where $M(x,y)$ denotes the value at position (x,y), δx and δy denote the positional offsets in the matrix.

We introduce the ACA loss function into the style transfer model to align the tensor-based features for minimizing the difference between the style and generated images. The ACA loss function is calculated by:

$$\mathcal{L}_a(f_s, f_g) = ||\mathcal{A}(f_s) - \mathcal{A}(f_g)||_2, \tag{3}$$

where the objective function $\mathcal{L}_a$ is to calculate the difference between auto-correlation of style and generated features denoted by $f_s \in \mathbb{R}^{C\times N\times M}$ and $f_g \in \mathbb{R}^{C\times N\times M}$. $\mathcal{A}(\cdot)$ denotes the 2D auto-correlation function. With this equation, we could capture the structural pattern from the style images.

C. Adaptive Auto-Correlation Alignment Loss

In order to capture the subtle periodic structure patterns from the images, we propose an adaptive minimum periodicity search algorithm. The goal of the proposed algorithm is to capture the minimum periodicity area. We employ spatial auto-correlation to learn the strong response for the structure in the feature layers and accordingly find the attendance area of the minimum periodicity. We analyze the minimum periodicity pattern in the restricted area across the feature dimension. It is defined as:

$$\mathcal{L}_a(f_s, f_g) = ||\mathcal{A}(\mathcal{P}(f_s)) - \mathcal{A}(\mathcal{P}(f_g))||_2, \tag{4}$$

$$\mathcal{P}(x)_k = \mathcal{I}(top_k(x)) \tag{5}$$

where the $\mathcal{P}$ denotes the adaptive search function of finding minimum periodicity area in the feature spaces, $top_k(\cdot)$ denotes the function to find the top k ($k = 9$) peaks in the given feature $x \in \mathbb{R}^{H\times W}$, $\mathcal{I}(\cdot)$ returns the found periodicity area which includes the positions of the top k peaks, $\mathcal{A}(\cdot)$ denotes the auto-correlation function defined in (2). The top k peaks represent the maximum response and the texture patterns in the feature layout. Searching top k peaks in the feature layout will help simplify the complexity of calculation and obtain significant

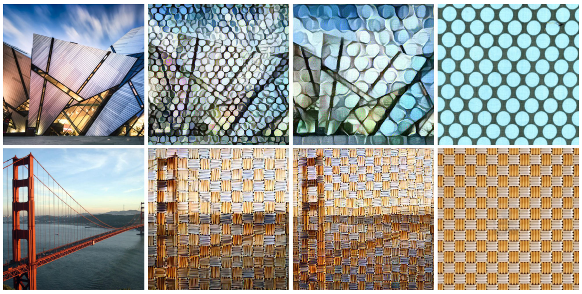

Fig. 3. The comparison results for adaptive alignment auto-correlation. w/o $\mathcal{P}$ and w/ $\mathcal{P}$ denote the auto-correlation alignment loss without and with the adaptive periodicity search function, respectively. The proposed adaptive periodicity search function can attain the authentic periodic texture patterns (repeated polka dot and woven).

style patterns from images. Finally, we can obtain the minimum adaptive periodicity $\mathbf{P}$ to narrow the scope of the area. The auto-correlation $\mathcal{A}$ is applied into the corresponding adaptive periodic area to preserve the structure pattern in the style image.

D. Adaptive Auto-Correlation Alignment Loss in Style Transfer

The structure of style transfer with the proposed loss is shown in Fig. 2. I_c denotes the content image, I_s represents the style image, and I_g denotes the generated image. Thus, the content loss function $\mathcal{L}_c$ is calculated by:

$$\mathcal{L}_c = \sum_{l \in l_c} ||f_c^l - f_g^l||_2, \tag{6}$$

where l_c denotes the selected content layers in the VGG network, f_g and f_c denote the extracted features of the generated and content images, f^l denotes the features in the corresponding layer l. Style loss $\mathcal{L}_s$ is calculated by the square euclidean distance between the Gram matrices based on the style and generated VGG features f_s, f_g in the layer l:

$$\mathcal{L}_s = \sum_{l \in l_s} ||G(f_s^l) - G(f_g^l)||_2, \tag{7}$$

where $G(\cdot)$ is the Gram matrix function [8] which represents the style in the corresponding VGG style layers l_s. The final objective function with the additional proposed loss is defined as follows:

$$\mathcal{L} = \omega_c \mathcal{L}_c + \omega_s \mathcal{L}_s + \omega_a \sum_{l \in l_a} ||\mathcal{A}(\mathcal{P}(f_s^l)) - \mathcal{A}(\mathcal{P}(f_g^l))||_2. \tag{8}$$

where $\mathcal{A}(\cdot)$ is the auto-correlation of the VGG layer to encode the structural pattern in the image, ω_c, ω_s, and ω_a denote the weight of content loss, style loss and adaptive auto-correlation alignment loss, l_a denotes the selected VGG layers to compute the adaptive auto-correlation alignment loss.

The visual comparison of the proposed adaptive periodicity search function $\mathcal{P}$ is presented in the Fig. 3. As it can be seen, the generated images from the adaptive periodicity alignment version focus on the style pattern with minimum periodicity than the original auto-correlation. The authentic periodic texture patterns (repeated polka dot and woven in Fig. 3) are extracted out from the style images.

IV. Experimental Results

A. Implementation Details

To conduct stylization and assess the proposed method, we present a neural style transfer model with the proposed adaptive auto-correlation alignment loss. For a fair comparison, we use the pre-trained VGG-19 as the feature extractor, and the same content ($ReLU\_4\_1$ and $ReLU\_5\_1$), style ($ReLU\_4\_1$), and auto-correlation ($ReLU\_4\_1$) layers to compute the content, style and proposed loss, respectively. There are also some different settings, such as that we replace the statistical information (mean and variance) as the style loss like AdaIN [18]. Inspired by [21], we have experimentally tested different weights ratio to balance the factors of content, style, and ACA loss to find the best one. We set the weight of content loss ω_c as 1 and the weight of style loss ω_s as 100. Moreover, we set the weight of adaptive auto-correlation alignment loss ω_a as 10. We train the different models on the platform tensorflow-1.8 [31] and NVIDIA Quadro P6000.

B. Periodicity Evaluation on Dataset

In our experiments, the Microsoft COCO 2014 (MS-COCO) dataset [32] and Describable Textures Dataset (DTD) [33] are used for content and style images, respectively. The MS-COCO dataset consists of 82,783 images and DTD contains 5,640 images with 47 categories labeled.

We split the dataset into training and testing subsets for both the content and style images. 22 kinds of texture style images are selected from the 47 categories in DTD dataset. In the selected 22 kinds of classes, we randomly choose 5 inference texture style images as the testing images and the others are the training images for each class. As to the MS-COCO dataset, 100 content images are randomly selected as the testing and the others are training images. It is noted that the testing style images are not shown in training procedures.

In order to test the impact of periodicity on the method, we split the DTD dataset into four categories datasets: Low, Normal, High, and Overall by calculating the angular second moment (ASM) [34], [35]. The higher for the ASM means more similarity for the image.

C. Comparison of the Neural Style Transfer

We evaluate the proposed method qualitatively and quantitatively. The qualitative comparison results are presented in Fig. 4. We can clearly observe that our proposed method outperforms other comparison methods. In detail, our proposed method preserves abundant color, shape and texture style patterns, including bubbles style of different sizes and better periodic shape of honeycombed. As shown in the "Ours" column of the first row, the proposed method can construct the human face with different bubbly shape of blue color to make the synthesized image recognizable. The SAN is followed, which captures the color and shape but misses the texture style patterns such as various bubble sizes and periodic arrangement of the honeycombed. CMD [22] well matches the color style into content structure, but changes the input texture style shape (such as bubbly style of different size) and periodicity patterns (such as repeated hexagon) according to the content image. The WCT and LST methods retain the color and shape style patterns, however, they blurred the structure both of the style and content images. Finally, the AdaIN produce the worst results, which could not retain the

Fig. 4. Stylized images by using different neural style transfer methods. We train the model in the adaptive auto-correlation alignment layers as same as the style layers. Ours can well generate the bubbles style of different sizes and better periodic shape of honeycombed as shown in the last column.

TABLE I
EVALUATION METRICS OF DIFFERENT NEURAL STYLE TRANSFER METHODS ON SYNTHESIZED IMAGES

Methods	**Performance**											
Texture Style	**Low**			**Normal**			**High**			**Overall**		
Metric	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
WCT [19]	0.402	0.234	0.179	0.359	0.160	0.190	0.237	0.274	0.230	0.337	0.224	0.199
Gatys [8]	0.213	0.169	0.281	0.183	0.236	0.261	0.150	0.184	0.186	0.275	0.240	0.195
LST [20]	0.028	0.123	0.039	0.114	0.044	0.059	0.116	0.096	0.091	0.082	0.090	0.062
SAN [21]	**0.495**	0.343	0.259	0.454	0.277	0.297	**0.444**	**0.347**	0.293	**0.466**	0.323	0.283
CMD [22]	0.455	0.300	0.255	0.377	0.304	0.306	0.434	0.340	0.294	0.422	0.327	0.284
AdaIN [18]	0.015	0.085	0.013	0.006	0.041	0.010	0.001	0.001	0.001	0.007	0.044	0.008
Ours	0.490	**0.368**	**0.284**	**0.491**	**0.313**	**0.340**	0.317	0.343	**0.300**	0.435	**0.342**	**0.307**

Note: The **boldface** denotes the best results. "Low," "Normal," "High" refer to the categories with low, normal, high ASM values.

integral color attribute and the texture pattern. It is due to the inherent quality of the oversimplified adaptive instance normalization. These results demonstrate that our proposed ACA loss function could improve the performance of the style transfer method. However, our method still has some limitations. As shown in Fig. 4, even if the texture shape is well transferred without corruptions in the human face, the color style transfer has some defects such as hair color from content image. It is due to the unbalanced weight of content, color style, and texture patterns. The quantitative comparison results are presented in Table I. We can observe that our proposed method has achieved the highest F1-score with the four different samples (such as Low, Normal, High, and Overall), which is consistent with the above qualitative analysis.

We also apply the proposed ACA loss function to the other styles transfer methods, such as Gatys, AdaIN and LST, resulting in **OurGatys**, **OurAdaIN**, and **OurLST**. The results are presented in Table II. Compared with the original style transfer methods (e.g., Gatys, AdaIN, and LST), the new methods with ACA loss function have achieved better results. In detail, **OurGatys** achieves improvement of 0.146, 0.101, and 0.096 in terms of precision, recall, and F1-score, respectively. **OurAdaIN** obtains improvement of 0.261, 0.072, and 0.081 in terms of precision, recall, and F1-score, when we apply the ACA loss function to the AdaIN. **OurLST** outperforms the LST over 0.034, 0.026, and 0.024 in terms of precision, recall, and F1-score, respectively.

TABLE II
EVALUATION METRICS OF DIFFERENT NEURAL STYLE TRANSFER METHODS ON SYNTHESIZED IMAGES

Methods	**Performance**		
Metric	**Precision**	**Recall**	**F1**
Gatys [8]	0.275	0.240	0.195
OurGatys	**0.421**	**0.341**	**0.291**
AdaIN [18]	0.007	0.044	0.008
OurAdaIN	**0.268**	**0.116**	**0.089**
LST [20]	0.082	0.090	0.062
OurLST	**0.116**	**0.116**	**0.086**

Note: The **boldface** denotes our method.

V. CONCLUSION

In this letter, we proposed an adaptive auto-correlation alignment (ACA) loss function to capture the structural pattern from the input images. We then applied the ACA loss into the style transfer task to extract much more structural patterns from the style images and proposed a new style transfer method based on SAN. We further introduced three metrics into the style transfer tasks to quantitatively evaluate the performance. The experimental results obtained on the public datasets: MS-COCO content image dataset and DTD style image dataset, demonstrate that our proposed methods outperform other neural style transfer methods.

ACKNOWLEDGMENT

This work was done prior to Amazon involvement of the authors.

REFERENCES

[1] Y. Yao, J. Ren, X. Xie, W. Liu, Y.-J. Liu, and J. Wang, "Attention-aware multi-stroke style transfer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 1467–1475.

[2] J. Cheng, Z. Han, Z. Wang, and L. Chen, 'One-shot' super-resolution via backward style transfer for fast high-resolution style transfer," *IEEE Signal Process. Lett.*, vol. 28, pp. 1485–1489, 2021.

[3] J. Yoo, Y. Uh, S. Chun, B. Kang, and J.-W. Ha, "Photorealistic style transfer via wavelet transforms," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 9035–9044.

[4] Y. Jing *et al.*, "Dynamic instance normalization for arbitrary style transfer," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, pp. 4369–4376.

[5] Z. Hu, J. Jia, B. Liu, Y. Bu, and J. Fu, "Aesthetic-aware image style transfer," in *Proc. 28th ACM Int. Conf. Multimedia*, 2020, pp. 3320–3329.

[6] M. Lu, H. Zhao, A. Yao, Y. Chen, F. Xu, and L. Zhang, "A closed-form solution to universal style transfer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 5951–5960.

[7] X. Chen, Y. Zhang, Y. Wang, H. Shu, C. Xu, and C. Xu, "Optical flow distillation: Towards efficient and stable video style transfer," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 614–630.

[8] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2414–2423.

[9] L. A. Gatys, A. S. Ecker, M. Bethge, A. Hertzmann, and E. Shechtman, "Controlling perceptual factors in neural style transfer," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 3730–3738.

[10] A. Sanakoyeu, D. Kotovenko, S. Lang, and B. Ommer, "A style-aware content loss for real-time HD style transfer," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 698–714.

[11] X. Deng, R. Yang, M. Xu, and P. L. Dragotti, "Wavelet domain style transfer for an effective perception-distortion tradeoff in single image super-resolution," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 3076–3085.

[12] D. Kotovenko, A. Sanakoyeu, S. Lang, and B. Ommer, "Content and style disentanglement for artistic style transfer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 4421–4430.

[13] S. Yang, Z. Wang, Z. Wang, N. Xu, J. Liu, and Z. Guo, "Controllable artistic text style transfer via shape-matching GAN," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 4441–4450.

[14] V. Dumoulin, J. Shlens, and M. Kudlur, "A learned representation for artistic style," in *Proc. Int. Conf. Learn. Representations*, 2017, pp. 1–26.

[15] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Diversified texture synthesis with feed-forward networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 266–274.

[16] D. Chen, L. Yuan, J. Liao, N. Yu, and G. Hua, "StyleBank: An explicit representation for neural image style transfer," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2770–2779.

[17] X. Pan, M. Zhang, D. Ding, and M. Yang, "A geometrical perspective on image style transfer with adversarial learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 63–75, Jan. 2022.

[18] X. Huang and S. Belongie, "Arbitrary style transfer in real-time with adaptive instance normalization," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 1510–1519.

[19] Y. Li, C. Fang, J. Yang, Z. Wang, X. Lu, and M.-H. Yang, "Universal style transfer via feature transforms," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 386–396.

[20] X. Li, S. Liu, J. Kautz, and M.-H. Yang, "Learning linear transformations for fast image and video style transfer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 3804–3812.

[21] D. Y. Park and K. H. Lee, "Arbitrary style transfer with style-attentional networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5873–5881.

[22] N. Kalischek, J. D. Wegner, and K. Schindler, "In the light of feature distributions: Moment matching for neural style transfer," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 9377–9386.

[23] T.-Y. Chiu, "Understanding generalized whitening and coloring transform for universal style transfer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 4451–4459.

[24] J. Yim, J. Yoo, W.-J. Do, B. Kim, and J. Choe, "Filter style transfer between photos," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 103–119.

[25] W. Gao, Y. Li, Y. Yin, and M.-H. Yang, "Fast video multi-style transfer," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2020, pp. 3211–3219.

[26] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Representations*, 2015, pp. 1–14.

[27] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.

[28] Z. Li, D. Mahapatra, J. A. W. Tielbeek, J. Stoker, L. J. van Vliet, and F. M. Vos, "Image registration based on autocorrelation of local structure," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 63–75, Jan. 2016.

[29] M. Dimitrov, T. Baitcheva, and N. Nikolov, "Efficient generation of low autocorrelation binary sequences," *IEEE Signal Process. Lett.*, vol. 27, pp. 341–345, 2020.

[30] D. Shi, W.-S. Gan, B. Lam, and K. Ooi, "Fast adaptive active noise control based on modified model-agnostic meta-learning algorithm," *IEEE Signal Process. Lett.*, vol. 28, pp. 593–597, 2021.

[31] M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Des. Implementation*, 2016, pp. 265–283.

[32] T.-Y. Lin *et al.*, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 740–755.

[33] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, and A. Vedaldi, "Describing textures in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3606–3613.

[34] M. Hall-Beyer, "GLCM texture: A tutorial," *Nat. Council Geographic Inf. Anal. Remote Sens. Core Curriculum*, vol. 3, pp. 4–24, 2000.

[35] P. Mohanaiah, P. Sathyanarayana, and L. GuruKumar, "Image texture feature extraction using GLCM approach," *Int. J. Sci. Res. Pub.*, vol. 3, no. 5, pp. 1–5, 2013.