Survey on Application of Generative Adversarial Networks (GANs) for Style Transfer

KAVYA KAVURI, Rutgers University-New Brunswick, USA
LIKHIT GARIMELLA, Rutgers University-New Brunswick, USA

Image style transfer is an essential field in computer vision. Style transfer is a prevalent image-to-image translation task that aims to transform an image into the style of another image while preserving its content. It enables the creation of new, visually appealing images by combining one image's style with another's content. It has numerous applications in art, design, and entertainment, allowing for the generation of unique and creative visual content. Additionally, image style transfer can be used in fields such as fashion, advertising, and marketing, where the visual presentation of products is crucial for attracting customers. Moreover, image style transfer can aid in image restoration, enhancement, and editing tasks. Therefore, the ability to transfer image styles has significant implications for both artistic and practical purposes, making it a valuable area of research. In this survey paper, we study various techniques that are being used in the field of style transfer, and we also dabble into the area of image-to-image translation that is relevant to image style transfer.

Additional Key Words and Phrases: datasets, neural networks, GANs, style transfer

ACM Reference Format:
Kavya Kavuri and Likhit Garimella. 2018. Survey on Application of Generative Adversarial Networks (GANs) for Style Transfer. *ACM Trans. Graph.* 37, 4, Article 111 (August 2018), 12 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Style transfer involves reimagining an image to mimic the characteristics of another artistic style. Essentially, this process involves transforming the appearance of an input image to appear as if it were redrawn using different stroke patterns, color schemes, perceptual representations, or artistic interpretations based on a reference style. While manual style transfer by a professional artist can be time-consuming, automated style transfer is a valuable technique with many practical applications. For instance, it can be used to quickly generate cartoon scenes from landscape or city photographs or provide novice artists with guidelines for painting. Enhancing the efficiency of style transfer techniques is a worthwhile pursuit.

Generative Adversarial Networks (GANs) are deep learning models that consist of two neural networks - a generator and a discriminator - that work in a game-like setting to produce new data. The generator is responsible for creating new data that is similar to the input data. It takes a random noise vector as input and learns to transform it into realistic data samples that match the input data distribution.

On the other hand, the discriminator is responsible for distinguishing between real data and the data produced by the generator. It takes both real and generated data samples as input and learns to classify them as either real or fake. During training, the generator tries to produce data that can fool the discriminator into thinking that it is real. In contrast, the discriminator tries to distinguish between real and fake data correctly. The two networks compete against each other in a zero-sum game, intending to improve their performance in each iteration of training.

As training progresses, the generator learns to produce increasingly realistic data samples. At the same time, the discriminator learns to become better at distinguishing real data from generated data. The ultimate goal is to reach a state where the generated data is indistinguishable from actual data. At this point, the generator can be used to create new data that is similar to the input data. The majority of style transfer architectures utilize Generative Adversarial Networks (GANs) with minor modifications. These modifications typically involve adjusting the loss function and the blocks utilized within the generator and/or discriminator networks.

This survey paper focuses on those GANs utilized to perform style transfer. There are mainly two types of models in our paper. The first one is single image style transfer. These models transfer the style onto the target image from a single-style content image. The second type is collection-style transfer. As the name suggests, these models take a collection of inputs and use their style characteristics to produce a styled target image.

Finally, this paper is organized in the following manner: first, we introduce the workings of all the models and explain the architecture. In the next section, we discuss the results of these models and illustrate what they do. Later we conclude the paper by discussing the current direction in which the research is going in the style transfer domain.

2 DATASETS

The GAN models that we have surveyed on, have used a multitude of datasets. Many of them being unpaired image sets. That is, the style content image, belongs to one domain, example, Van Gogh, Monet, Ukiyo-e etc and its pair belongs to a completely different domain. For example, a cat, a busy street etc. These models are supposed to extract the features of the style image and apply them to the target image. Thus they produce image of a cat in the style of Van Gogh or Monet etc.

A popular dataset used for style transfer is the COCO (Common Objects in Context) dataset, which contains over 200,000 images of everyday objects in various contexts. Another commonly used

Authors' addresses: Kavya Kavuri, kk1069@scarletmail.rutgers.edu, Rutgers University-New Brunswick, 57 US Highway 1, New Brunswick, New Jersey, USA, 08901; Likhit Garimella, lg836@scarletmail.rutgers.edu, Rutgers University-New Brunswick, 47 Bevier Road, Piscataway, New Jersey, USA, 08854.

0730-0301/2018/8-ART111 $15.00
https://doi.org/XXXXXXX.XXXXXXX

dataset is the ImageNet dataset, which consists of over 1 million images across 1,000 object categories.

Typically these models are trained on large datasets of images that are diverse in style and content. Most of the models that we present in section 3 have created their own datasets.

3 THE GAN MODELS

In this section, we discuss the core principles of each model that we are studying for the survey.

3.1 A Style-Based Generator Architecture for Generative Adversarial Networks

The article [Karras et al. 2019] by Nvidia proposes an alternative architecture to the traditional GAN architecture for style transfer. They do not change loss function or regularization, or hyper-parameters. Two images are provided to the network, and the resulting image is obtained by mixing the styles of both images.

Let's talk about the model architecture and how it works, The source A image is converted into a vector in latent space w ∈ W and put into a feed-forward network to obtain a representative vector. This is the input for the adaptive instance normalization (AdaIN) block, which mixes the style from the feature map and the input image(source B) from the Const 4x4x512 block above. This process repetitively forms the Synthesis network *g*.

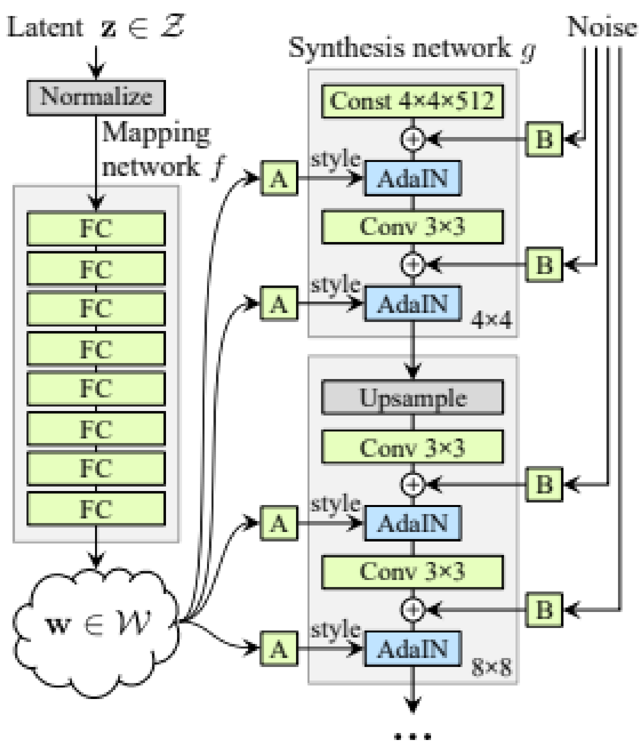

Fig. 1. Model architecture. source - [Karras et al. 2019]

3.2 Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

[Zhu et al. 2017]'s CycleGAN is a fascinating model that translates one image to another when both belong to entirely different domains. This model is primarily built to perform image-to-image translation but can achieve style transfer excellently.

The network architecture consists of a typical GAN architecture, with the generator having multiple convolutions and several residual blocks. In particular, three convolutions, six or nine residual blocks depending on the image resolution, and three convolutions that process the feature map to an RGB image. For the discriminator, they use 70 × 70 PatchGANs[[Isola et al. 2017] [Li and Wand 2016] [Ledig et al. 2017]], which aim to classify whether 70 × 70 overlapping image patches are real or fake.

They utilize a *cycle consistency loss*, which incentivizes forward or backward image reconstruction. This means forward(F) or backward(G) mappings should be able to return the original image when either function's output image is given as an input to the other.

The cycle loss is defined as:

$$\mathcal{L}_{cyc} = E_{x\sim p_{data}(x)}[||F(G(x))x||_1] + E_{y\sim p_{data}(y)}[||F(G(y))y||_1]$$

This cycle loss is combined with an adversarial loss to obtain the final objective function. Finally, the model is trained using a dataset where the source and target are unpaired because images do not belong to the same domain.

3.3 Unpaired High-Resolution and Scalable Style Transfer Using Generative Adversarial Networks

In this paper [Junginger et al. 2018], Their focus is on tackling the issue of unpaired domain translation, especially when it comes to processing high-resolution images. However, existing methods encounter a problem of high peak memory usage during both the training and translation stages, which limits the maximum image size that can be processed on a given GPU. To overcome this challenge, they have developed a scalable approach that can handle images of arbitrary-high resolutions without increasing the Neural Network's peak memory consumption.

To attain their objective, they adopt a strategy of not processing the entire image at a single time. Instead, they train and implement the domain translation on small, overlapping image subsamples.

Now, we will examine the functioning of their model.

In Fig. 2, random positions and sizes are used to extract training samples from the high-resolution image. These samples are then scaled down to a standard resolution of $x_{batch}y_{batch}$ and merged to create a training batch.

As seen in Fig.3, with the training batch from above, an iteration of training is performed on GAN with frameworks adopted from CycleGAN[[Zhu et al. 2017]] or Unit GAN.[[Liu et al. 2017]].

After the training phase, as shown in Fig. 4, the entire high-resolution image is transformed by extracting individual samples, which are then translated by the generator individually. Finally, the translated images are combined to produce the complete transformed image.

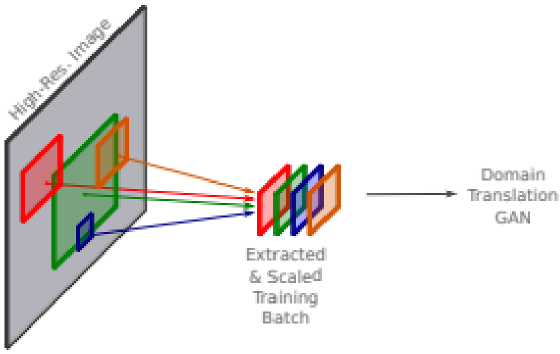

Fig. 2. First step - preprocessing. source - [Junginger et al. 2018]

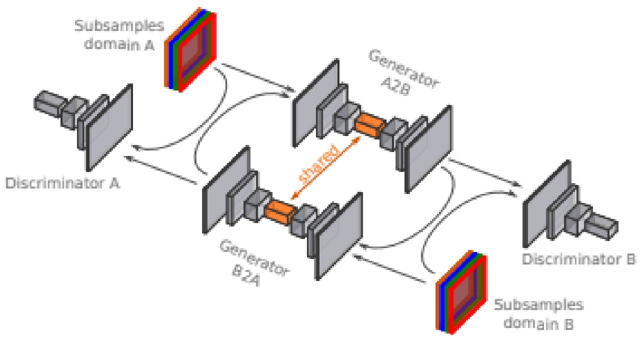

Fig. 3. Second step - training. source - [Junginger et al. 2018]

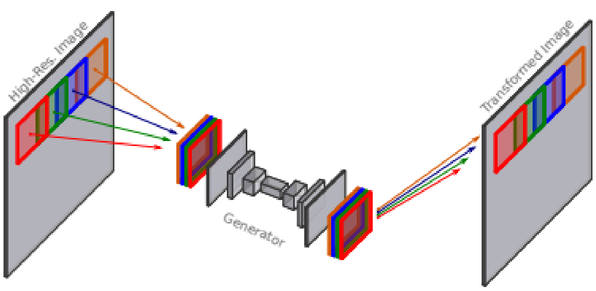

Fig. 4. Third step - producing result. source - [Junginger et al. 2018]

3.4 DRB-GAN: A Dynamic ResBlock Generative Adversarial Network for Artistic Style Transfer

The *Dynamic ResBlock Generative Adversarial Network* proposed by [Xu et al. 2021] is a model which is capable of accomplishing artistic style transfer onto the target from a single image as well as a collection of styles. They call it Collection Style Transfer. In this style transfer network, there are multiple *Dynamic ResBlocks* that are designed to integrate style code and the CNN feature map of the source image to produce a high quality synthetic, artistic version of the content image.

Each DRB consists of a Convolution Layer, a Dynamic Convolution layer [Chen et al. 2020], a ReLU layer, an AdaIN layer [Huang and Belongie 2017], and an instance normalization layer with a residual connection. This can be better understood after we get the full picture of the model.

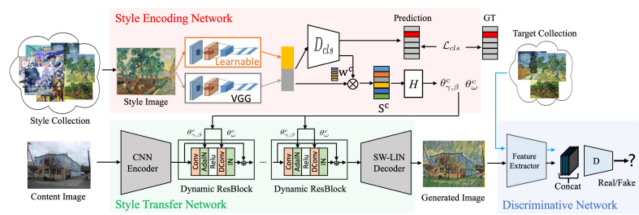

Fig. 5. Model Architecture. source - [Xu et al. 2021]

The DRB-GAN architecture, as shown in Fig. 5, can be divided into three main parts, Style Encoding Network, Style Transfer Network and Discriminative Network. The Style Encoding Network is used to generate style code from the style image, to be given to the Style Transfer Network which applies it on the content/source image. An Attention guided feature extractor - which is a combination of fixed-weight VGG network trained on COCO dataset and a learnable encoder network, is used to create a *style code* of the style image. This parameter is transferred to Style Transfer Network.

Their Collection Style Transfer is a weighted average of the *style codes* from all images of the style collection. In the Style Transfer Network, the CNN Encoder is used to down-sample the input, multiple dynamic residual blocks are given the style codes from Style Encoding Network, and a spatial window Layer-Instance Normalization (SWLIN) decoder is used to up-sample the output.

3.5 Gated-GAN: Adversarial Gated Networks for Multi-Collection Style Transfer

Gated-GAN, explained in [Chen et al. 2018b] is a model that can be used to transfer multiple styles onto target image using a single model.

There are mainly three modules in Gated-GAN: an encoder, a gated transformer, and a decoder. The generative network implicitly learns the target style from adversarial loss, aiming to fool the discriminator. We can better understand the generative network from the image below:

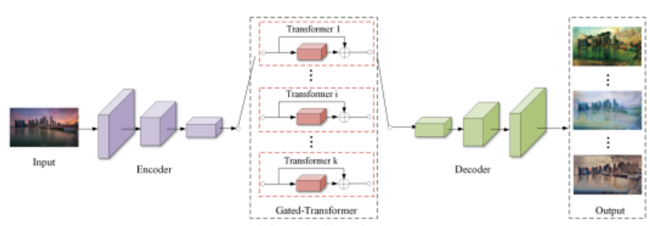

Fig. 6. Generative Network Architecture. source - [Chen et al. 2018b]

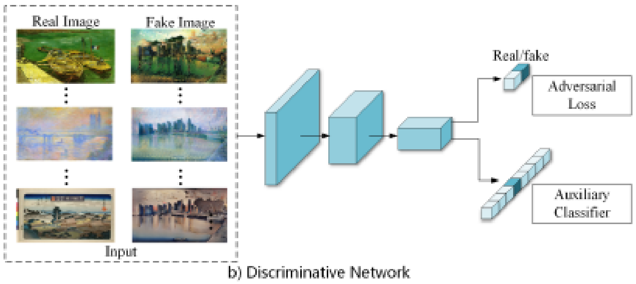

Fig. 7. Discriminative Network Architecture. source - [Chen et al. 2018b]

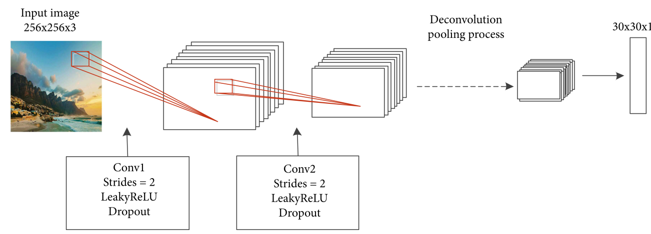

Fig. 8. Discriminator Network. source - [Liao and Huang 2022]

From the generative network figure above, we can see that The encoder consists of a series of convolutional layers that transform input image into feature space *Enc(x)* and the decoder is a series of fractionally-strided convolutional networks that decode the transformed feature into output images *Dec(T(Enc(x)))*. The style learning process is done in the gated-transformer block. It is a series of residual networks denoted by *T(·)*. An *auto-encoder reconstruction loss* is introduced for the training of Gated-GAN.

$$\mathcal{L}_R = E_{x \in X}[||Dec(Enc(x)) - x||_1]$$

This loss encourages the generator to train encoder and decoder to be able to create the original image if passed through these mappings. In the multi-collection style transfer, the gated-transformer (red blocks in Figure above) transforms the input from encoded space into different styles by switching trigger to different branches.

Finally, in the Discriminative network, the discriminator calculates the adversarial loss and also predicts the class of the input image. The goal here is to get the predictions of the classes of the real style source that is supposed to be transferred onto the target image.

3.6 Deep Learning-Based Application of Image Style Transfer

In this paper [Liao and Huang 2022], the author provides an improved version of CycleGAN [Zhu et al. 2017]. The main idea is that, in the original CycleGAN Unet network is used to downsize and up-size the image sample, whereas in this paper, they use PatchGAN to improve its efficiency greatly.The generator of the CycleGAN remains the same but the discriminator is improved. The Markov discriminator of PatchGAN has high-resolution and high-detail characteristics in the image style transfer. So it is employed to achieve better results for style transfer.

PatchGAN is used in the discriminative network. It is completely composed of convolutional layers. Unlike ordinary GAN discriminator, PatchGAN provides a patch as the same size of the input image where the confidence value in each cell of the patch indicates whether that pixel is real or fake. This proves to be a better indicator of the realness or fakeness for high quality images. It is worth noting that the loss function is again, same as the original CycleGAN.

3.7 A Domain Gap Aware Generative Adversarial Network for Multi-domain Image Translation

The model proposed by [Xu and Wang 2021] has a unified model called UMIT which can accomplish image-to-image translation without requiring the cycle consistency check of other typical GAN networks like CycleGan [Zhu et al. 2017], Gated-GAN [Chen et al. 2018b] etc. This is a multi-domain image translation model, i.e., this model is a kind of unification of all the models we have discussed. The Unified Generator, illustrated below, is the main unit that translates the image from one domain to another. Let's take a look at its functionality.

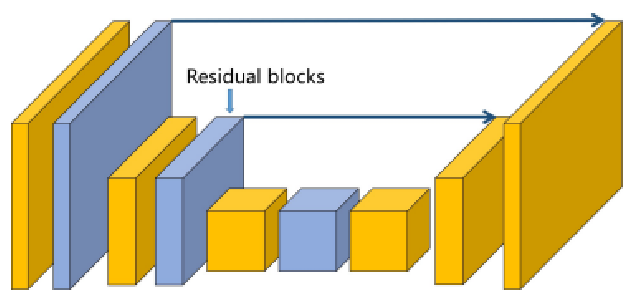

Fig. 9. Generator Network. source - [Xu and Wang 2021]

Unified Generator is developed by incorporating residual blocks at various layers within the Unet architecture; the network can gather information across multiple scales to perform transformations on features with varying spatial resolutions. Moreover, this structure takes both the images and the domain labels as inputs. A unique input-output drawer mechanism disentangles local textures at a pixel level.

The Input-Output drawer is a domain container that preserves domain-specific local textures. Each input image from the collection is assigned to one input drawer. Then the input of generator is formed as a concatenation of each input drawer. A similar operation is performed to receive output images from the output drawer.

Multi-scale Discriminator. This would require a sophisticated architecture to capture the global features and differentiate the features according to the domain information. The patch discriminator is very useful for local texture translations. However, that itself is not enough. Hence, dilated convolutional layers were added to encourage global shape changes with a large receptive field.

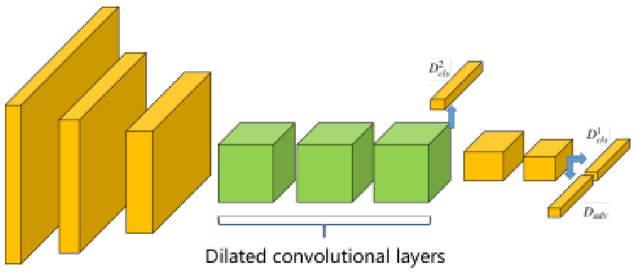

Fig. 10. Discriminator Network. source - [Xu and Wang 2021]

The objective is to minimize adversarial loss, multi-scale domain classification loss, identity loss, and perceptual self-regularization loss.

3.8 Style Transfer Using Generative Adversarial Networks for Multi-Site MRI Harmonization

In this paper, [Liu et al. 2021] proposes that the MR(Magnetic Resonance) images across multiple scanners and demographics can be converted into the target domain by applying style transfer rather than domain transfer. The need for retrospective harmonization is often inevitable when considering the scans from decade-long projects like ADNI. ADNI is a project which collects MRI scans of various demographics and geographies, so the scans range in variety. There must have been scanner upgrades, standard changes in usage etc., which introduces variability in the scans. To overcome these issues, the author proposes Style-Encoding GAN.

The architecture is very similar to [Karras et al. 2019], which uses adversarial loss, cycle consistency loss, style reconstruction loss, and style diversification loss. An encoding network takes the input style image and produces a *style code (s)*. The generator takes this s and feeds it into the synthesis network (refer to Fig. 1) to obtain a style-transferred image. The new loss functions that are introduced are style reconstruction loss:

$$\mathcal{L}_{sty} = E_{x,z}[||s - E(G(x,s))||_1]$$

and style diversification loss:

$$\mathcal{L}_{div} = E_{x,z_1,z_2}[||G(x,s_1) - G(x,s_2)||_1]$$

3.9 Style Fader Generative Adversarial Networks for Style Degree Controllable Artistic Style Transfer

[Zuo et al. 2022] proposes additions to the existing GAN networks that allow adjusting the style degree for existing GAN-based artistic style transfer frameworks in real time after training. The additions are a Style Scaling Injection (SSI) module and a Style Degree Interpretation (SDI) module. The SSI module decides the degree to which the style is transferred to the target image. The SDI module interprets the output probabilities of a multi-scale content-style binary classifier as the style degrees, providing a mechanism to parameterize the style degree of the styling. We can study these modules with the help of the diagram below:

The addition is the blue box, which says Style Degree Factor, and TB, which is the SSI module. Style Scaling Injection Module: they

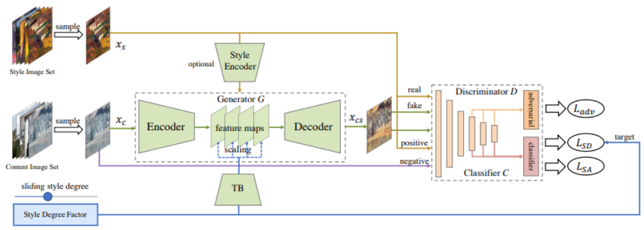

Fig. 11. Model architecture. source - [Zuo et al. 2022]

design the SSI module that scales the feature activations in existing models. As shown above, the SSI module first takes the Style Degree Factor (SDF) value, denoted as a random variable F, as the input. The SSI module then utilizes a learnable non-linear transformation block (TB) which transforms the value of f to parameters that scale the feature activations in existing GAN-based models. Style Degree Interpretation Module: The SSI module offers control signals to influence the generation process in existing models. But To make f aware of the smooth and continuous changes in style degree, they additionally introduce an SDI module. It provides a mechanism to parameterize the style degree. The SDF module provides f to the Style Degree loss:

$$\mathcal{L}_{SD} = min_G E_f[-flogC(G(.,f)) - (1-f)log(1-C(G(.,f))], f \sim U[0,1]$$

Here G denoted generator. f is the input provided by SDF module. The adversarial loss is added to this Style Degree Loss to obtain the final loss function.

3.10 MISS GAN: A Multi-IlluStrator Style Generative Adversarial Network for image to illustration translation

The author of [Barzilay et al. 2021] proposes a Multi-IlluStrator Style Generative Adversarial Network (MISS GAN) for image-to-illustration translation. This model allows learning multi-style transfer and performing an unsupervised image-to-illustration translation, with a novel generator that is based on the GANILLA generator [Hicsonmez et al. 2020], this model also combines StarGAN v2 model [] which allows learning multi-style transfer and performing an unsupervised image-to-illustration translation. To better understand, let's look at the generator architecture.

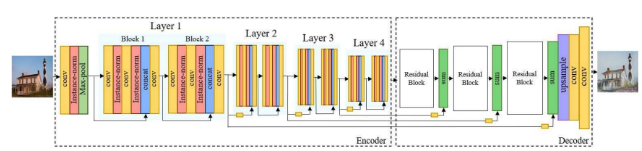

Fig. 12. Generator architecture. source - [Barzilay et al. 2021]

The generator architecture of the MISS-GAN model starts with a 7x7 convolution layer, followed by an instance normalization layer, ReLU, and max pooling layers. Then, the generator continues with four layers, where each layer consists of two residual blocks. The decoder contains three residual blocks. Each residual block starts with a convolution layer, followed by an AdaIN layer and ReLU activation. These are followed by a simple upsample layer, convolution layer, AdaIN layer, and ReLU activation.

Additionally, they added *Content features loss*, which measures distances in the feature space of a pretrained VGG16 network.

$$\mathcal{L}_{content_feat} = E[||\phi(x) - \phi(y_x)||_1]$$

Here x is the original content image, y_x is the transformed image by the network to the target domain. $\phi(x)$ and $\phi(y_x)$ are the activations of the second layer of the pretrained VGG16 network with x and y_x as inputs, respectively.They have also added a *Style aware content loss* which is defined as follows:

$$\mathcal{L}_{content_sacl} = \frac{1}{d} E[||Encoder(x) - Encoder(Decoder(Encoder(x)))||_2^2]$$

Here Encoder and Decoder are the ones used in the generator, and d is the number of the latent space domains.

3.11 CartoonGAN: Generative Adversarial Networks for Photo Cartoonization

This paper [Chen et al. 2018a] proposes a style of GAN that transforms real-world captured pictures into cartoonic-style images. Cartoon images have clear edges, smooth color shading and simple textures, and these characteristics influence the loss function in this GAN model. There are 2 kind of losses observed in this CartoonGAN's cartoonization of images, and they are, Adversarial loss and Content loss.

The loss function L(G, D) consists of two parts: (1) the adversarial loss Ladv(G, D), which drives the generator network to achieve the desired manifold transformation, and (2) the content loss Lcon(G, D), which preserves the image content during cartoon stylization. We use a simple additive form for the loss function:

$$\mathcal{L}(G, D) = \mathcal{L}_{adv}(G, D) + w\mathcal{L}_{con}(G, D)$$

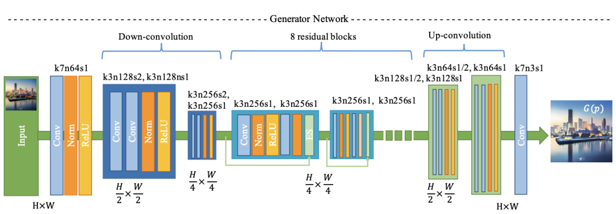

Fig. 13. Generator Network. source - [Chen et al. 2018a]

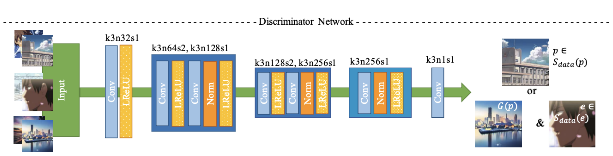

Fig. 14. Discriminator Network. source - [Chen et al. 2018a]

3.12 Weather GAN: Multi-Domain Weather Translation Using Generative Adversarial Networks

This paper [Li et al. 2021] proposes a new way of weather translation, that refers to transferring weather conditions of an image from one category to another. This is achieved by a multi-domain weather translation approach based on generative adversarial networks (GAN), called as Weather GAN, which can achieve the transferring of weather conditions like sunny, cloudy, foggy, rainy and snowy, specifically, when the weather conditions in the image are determined by various weather-cues, such as cloud, blue sky, wet ground, etc.

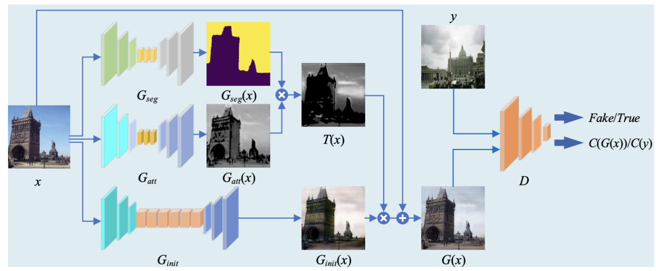

Fig. 15. Model architecture of Weather GAN. source - [Li et al. 2021]

3.13 P2-GAN: EFFICIENT STYLE TRANSFER USING SINGLE STYLE IMAGE

This paper [Zheng and Liu 2020] proposes a novel Patch Permutation GAN (P2-GAN) network, that can efficiently learn the stroke style from a single style image. For this, we use patch permutation to generate multiple training samples of images from the given style image. A patch discriminator that can simultaneously process patch-wise images and natural images seamlessly is designed.

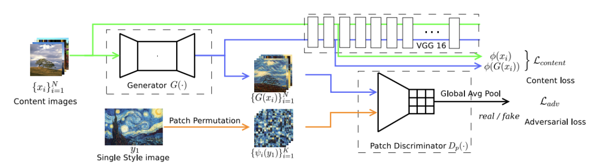

Fig. 16. Network architecture of the proposed P2-GAN. source - [Zheng and Liu 2020]

3.14 Style Transfer for Anime Sketches with Enhanced Residual U-net and Auxiliary Classifier GAN

This paper [Zhang et al. 2017] proposes an integrated residual U-net: this will not just randomly colorize sketch lines as outputs, they also do specific style transfer. This model applies the style to the grayscale sketch with auxiliary classifier generative adversarial network (AC-GAN).

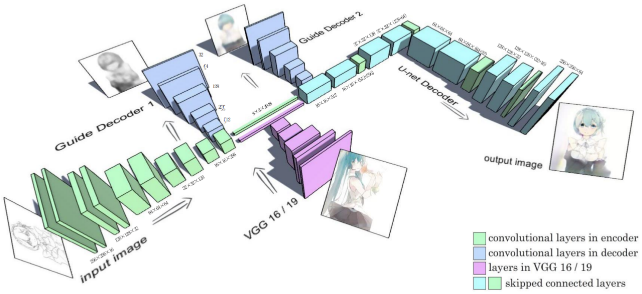

Fig. 17. The architecture of the generator. source - [Zhang et al. 2017]

3.15 CLIPstyler: Image Style Transfer with a Single Text Condition

This paper [Kwon and Ye 2022] proposes a pre-trained text-image embedding model of CLIP. A new framework that enables a style transfer 'without' a style image, but only with a text description of the desired style. This model demonstrates the modulation of the style of content images only with a single text condition. We get a successful image style transfer with realistic textures that reflect semantic query texts.

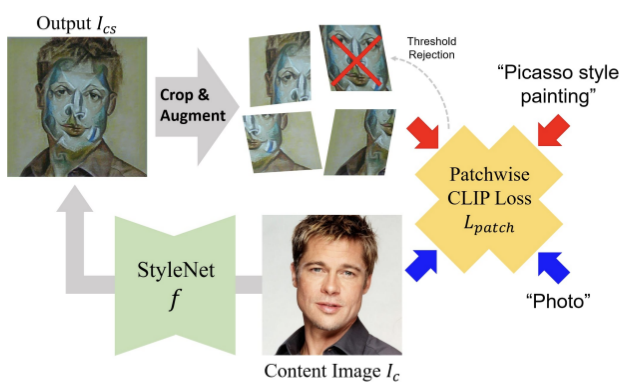

Fig. 18. Schematics of proposed patch-wise CLIP. source - [Kwon and Ye 2022]

3.16 GAN-Based Day-to-Night Image Style Transfer for Nighttime Vehicle Detection

This paper [Lin et al. 2021] proposes a type of GAN named AugGAN, a GAN-based data augmenter which could transform on-road driving images to a desired domain while image-objects would be well preserved. A structure-aware unpaired image-to-image translation network which learns the latent data transformation across different domains is designed; the domain adaptation capability of a vehicle detector is not limited by its training data; the object-preserving network provides significant performance gain in the difficult day-to-night case in terms of vehicle detection. AugGAN could generate more visually plausible images compared to competing methods on different on-road image translation tasks across domains.

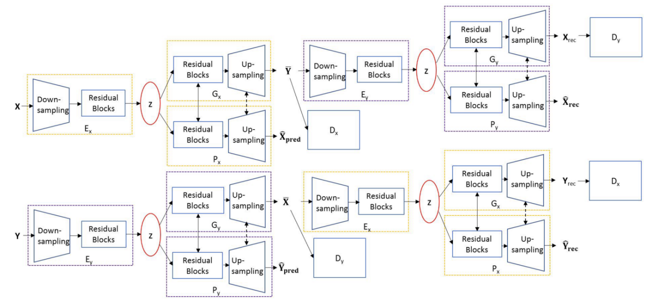

Fig. 19. Overall structure of the proposed image-to-image translation network. source - [Lin et al. 2021]

3.17 Image Generation and Style Transfer Using Conditional Generative Adversarial Networks

This paper [Murali et al. 2019] proposes the generalizability of Conditional Generative Adversarial Networks by testing their performance on a combination of image datasets, using the same type of input-output mapping. A significant consequence of using multiple datasets is that the model can transfer characteristics such as style and design from one dataset to another. In order to facilitate this, conditional GAN model must handle multiple conditional inputs, taking both the input image map and a class label.

3.18 SLGAN: Style- and Latent-guided Generative Adversarial Network for Desirable Makeup Transfer and Removal

This paper [] proposes novel, perceptual makeup loss and a style-invariant decoder that can transfer makeup styles based on histogram matching to avoid the identity-shift problem. This model is used to apply makeup to photos of the human face, and the factors taken into consideration include (1) facial components, (2) interactive color adjustments, (3) makeup variations, (4) robustness to poses and expressions, and the (5) use of multiple reference images, all these 5 features simultaneously. SLGAN is better than other comparable state-of-the-art methods.

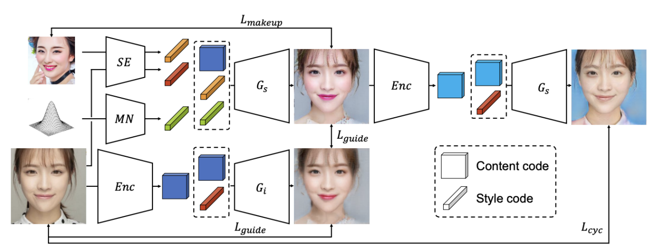

Fig. 20. SLGAN Network. source - [Horita and Aizawa 2022]

4 DISCUSSION ON RESULTS

In this section, we illustrate the results of the models that we explored in section 3.

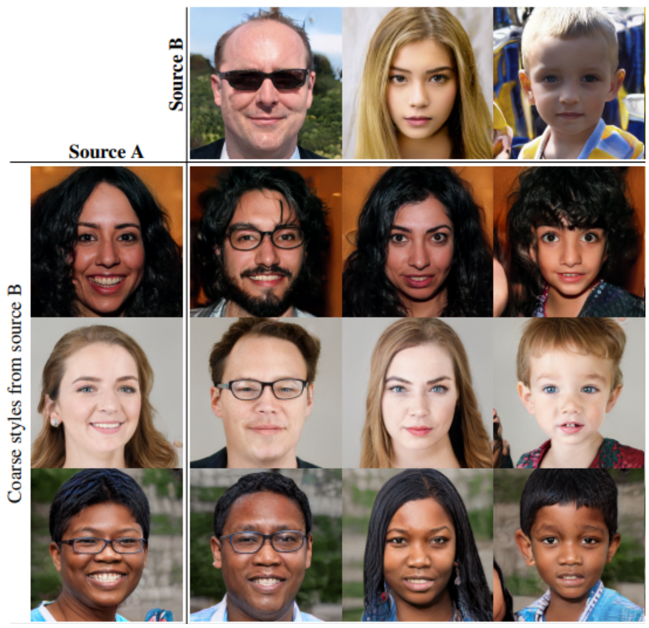

Fig. 21. Style based generator model Results. source - [Karras et al. 2019]

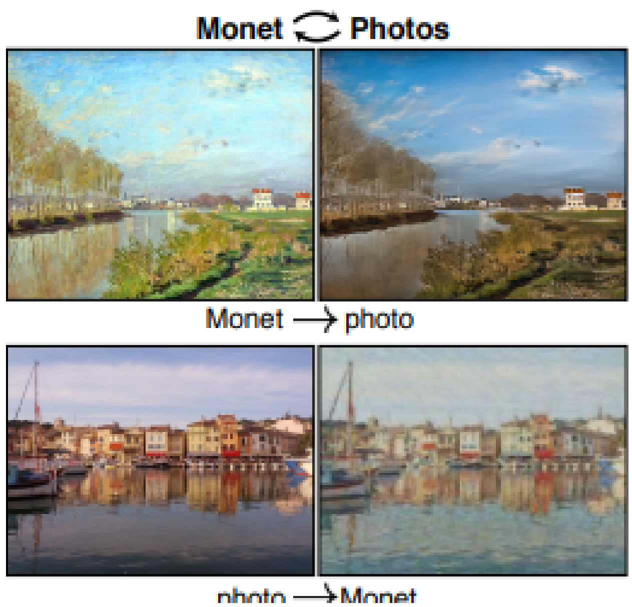

Fig. 22. Cycle GAN Results. source - [Zhu et al. 2017]

Fig. 23. Domain transfer GAN Results. source - [Junginger et al. 2018]

Fig. 24. DRB-GAN Results. source - [Xu et al. 2021]

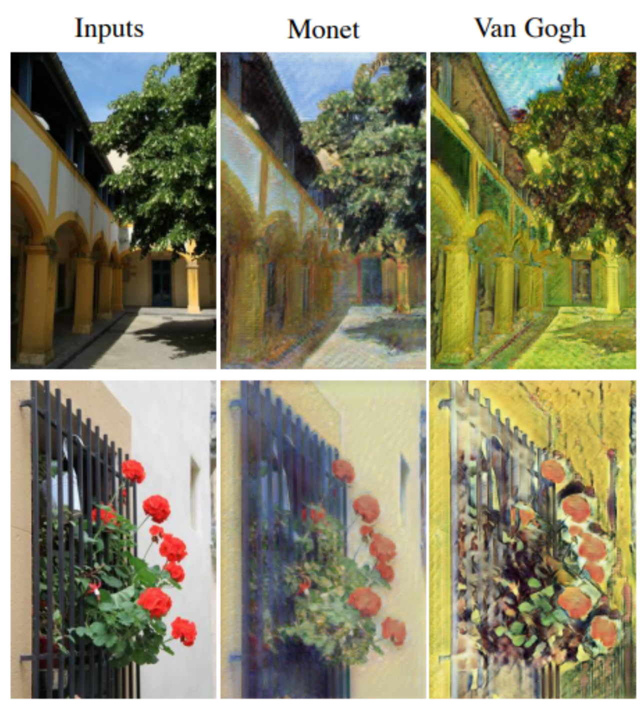

Fig. 25. Gated-GAN Results. source - [Chen et al. 2018b]

5 CONCLUSION

In conclusion, style transfer using Generative Adversarial Networks (GANs) has become an increasingly popular research topic in the deep learning community. Through the use of GANs, it is possible to create new images that combine the content of one image with the style of another.

This survey paper has reviewed various style transfer GAN architectures, including CycleGAN, DRB-GAN, GATED-GAN, MangaGAN, RPD-GAN and many more. We reviewed different loss

Fig. 26. Improved Cycle GAN Results. source - [Liao and Huang 2022]

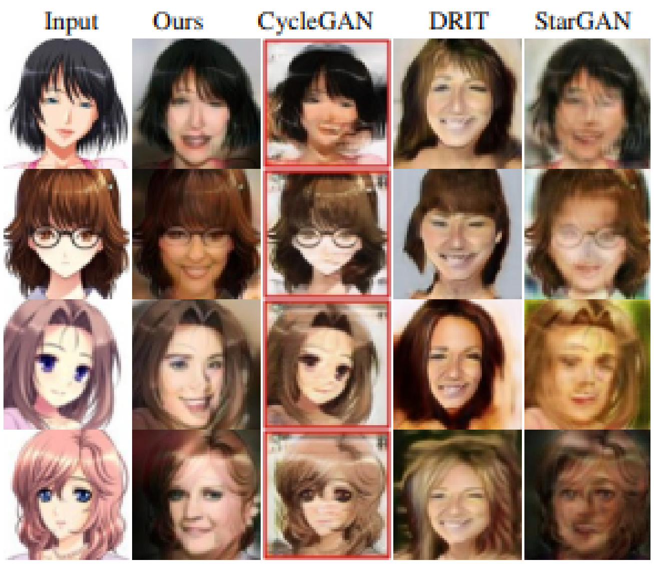

Fig. 27. Unified generator model Results - anime to human. source - [Xu and Wang 2021]

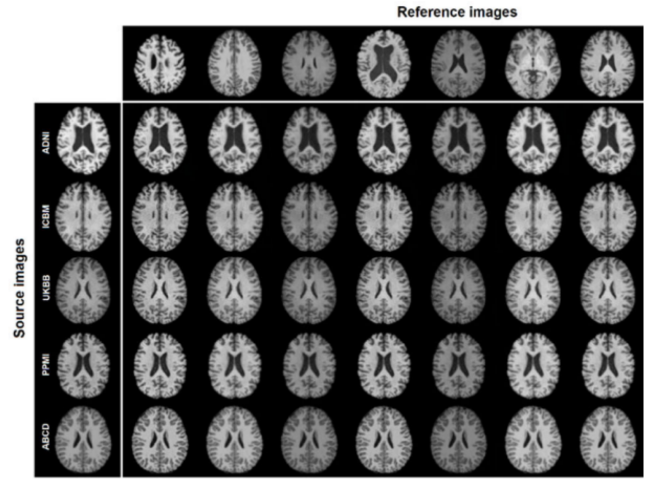

Fig. 28. Style Encoder GAN Results. source - [Liu et al. 2021]

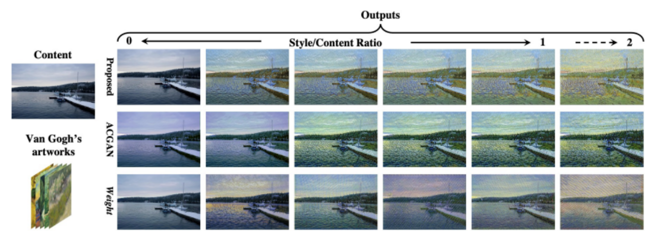

Fig. 29. Style degree controllable model Results. source - [Zuo et al. 2022]

functions as well. We have also discussed the datasets used for training style transfer GAN models, which are typically large collections of diverse images.

While GAN-based style transfer techniques have shown great promise in generating realistic and high-quality images, there are still challenges to be addressed, such as the need for more effective methods to preserve content information during style transfer. Nonetheless, with continued research and innovation in this field, it is likely that GAN-based style transfer will continue to evolve and find numerous practical applications, from creating digital art to enhancing images for use in various industries.

REFERENCES

Noa Barzilay, Tal Berkovitz Shalev, and Raja Giryes. 2021. MISS GAN: A Multi-IlluStrator style generative adversarial network for image to illustration translation. *Pattern Recognition Letters* 151 (2021), 140–147.

Xinyuan Chen, Chang Xu, Xiaokang Yang, Li Song, and Dacheng Tao. 2018b. Gated-gan: Adversarial gated networks for multi-collection style transfer. *IEEE Transactions on Image Processing* 28, 2 (2018), 546–560.

Yinpeng Chen, Xiyang Dai, Mengchen Liu, Dongdong Chen, Lu Yuan, and Zicheng Liu. 2020. Dynamic convolution: Attention over convolution kernels. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 11030–11039.

Yang Chen, Yu-Kun Lai, and Yong-Jin Liu. 2018a. CartoonGAN: Generative Adversarial Networks for Photo Cartoonization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Xiang Gao, Yingjie Tian, and Zhiquan Qi. 2020. RPD-GAN: Learning to Draw Realistic Paintings With Generative Adversarial Network. *IEEE Transactions on Image Processing* 29 (2020), 8706–8720. https://doi.org/10.1109/TIP.2020.3018856

Samet Hicsonmez, Nermin Samet, Emre Akbas, and Pinar Duygulu. 2020. GANILLA: Generative adversarial networks for image to illustration translation. *Image and Vision Computing* 95 (2020), 103886.

Daichi Horita and Kiyoharu Aizawa. 2022. SLGAN: style-and latent-guided generative adversarial network for desirable makeup transfer and removal. In *Proceedings of the 4th ACM International Conference on Multimedia in Asia*. 1–5.

Xun Huang and Serge Belongie. 2017. Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision*. 1501–1510.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1125–1134.

Andrej Junginger, Markus Hanselmann, Thilo Strauss, Sebastian Boblest, Jens Buchner, and Holger Ulmer. 2018. Unpaired high-resolution and scalable style transfer using generative adversarial networks. *arXiv preprint arXiv:1810.05724* (2018).

Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4401–4410.

Gihyun Kwon and Jong Chul Ye. 2022. CLIPstyler: Image Style Transfer With a Single Text Condition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18062–18071.

Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. 2017. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 4681–4690.

Fig. 30. MISS-GAN Results. source - [Barzilay et al. 2021]

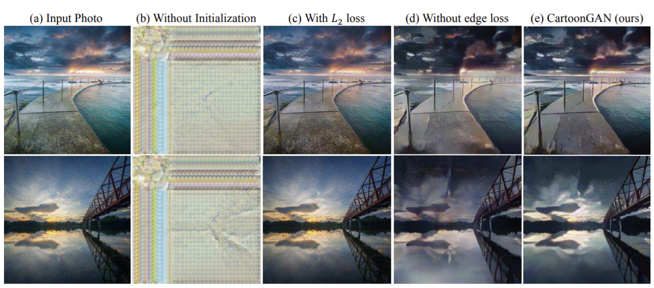

Fig. 31. Cartoon GAN Results. source - [Chen et al. 2018a]

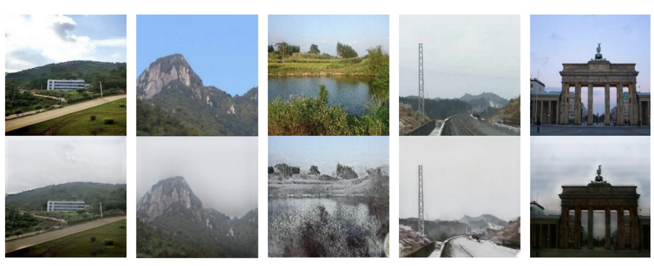

Fig. 32. Weather GAN Results. source - [Li et al. 2021]

Chuan Li and Michael Wand. 2016. Precomputed real-time texture synthesis with markovian generative adversarial networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III 14*. Springer, 702–716.

Fig. 33. Weather GAN Results. source - [Li et al. 2021]

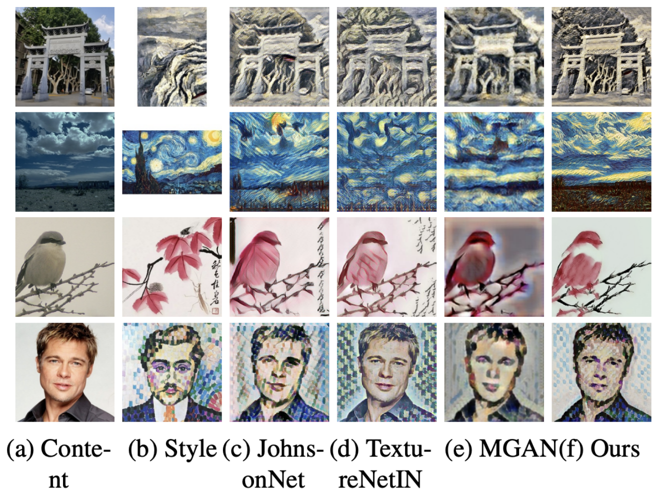

Fig. 34. P2-GAN Results. source - [Zheng and Liu 2020]

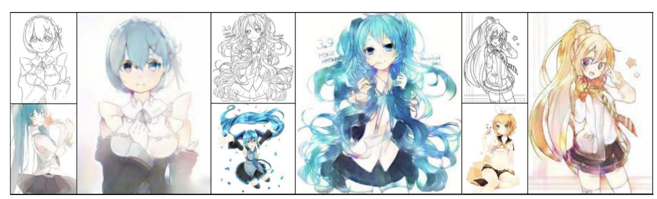

Fig. 35. AC-GAN Results. source - [Zhang et al. 2017]

Xuelong Li, Kai Kou, and Bin Zhao. 2021. Weather GAN: Multi-domain weather translation using generative adversarial networks. *arXiv preprint arXiv:2103.05422* (2021).

YiMi Liao and YouFu Huang. 2022. Deep Learning-Based Application of Image Style Transfer. *Mathematical Problems in Engineering* 2022 (2022).

Che-Tsung Lin, Sheng-Wei Huang, Yen-Yi Wu, and Shang-Hong Lai. 2021. GAN-Based Day-to-Night Image Style Transfer for Nighttime Vehicle Detection. *IEEE Transactions on Intelligent Transportation Systems* 22, 2 (2021), 951–963. https://doi.org/10.1109/TITS.2019.2961679

Mengting Liu, Piyush Maiti, Sophia Thomopoulos, Alyssa Zhu, Yaqiong Chai, Hosung Kim, and Neda Jahanshad. 2021. Style Transfer Using Generative Adversarial Networks for Multi-site MRI Harmonization. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Marleen de Bruijne, Philippe C. Cattin, Stéphane Cotin, Nicolas Padoy, Stefanie Speidel, Yefeng Zheng, and Caroline Essert (Eds.). Springer International Publishing, Cham, 313–322.

Ming-Yu Liu, Thomas Breuel, and Jan Kautz. 2017. Unsupervised image-to-image translation networks. *Advances in neural information processing systems* 30 (2017).

Sharada Murali, Mohammad Reza Rajati, and Somasekhar Suryadevara. 2019. Image Generation and Style Transfer Using Conditional Generative Adversarial Networks.

Fig. 37. Aug GAN Results. source - [Lin et al. 2021]

Fig. 38. Conditional GAN Results. source - [Murali et al. 2019]

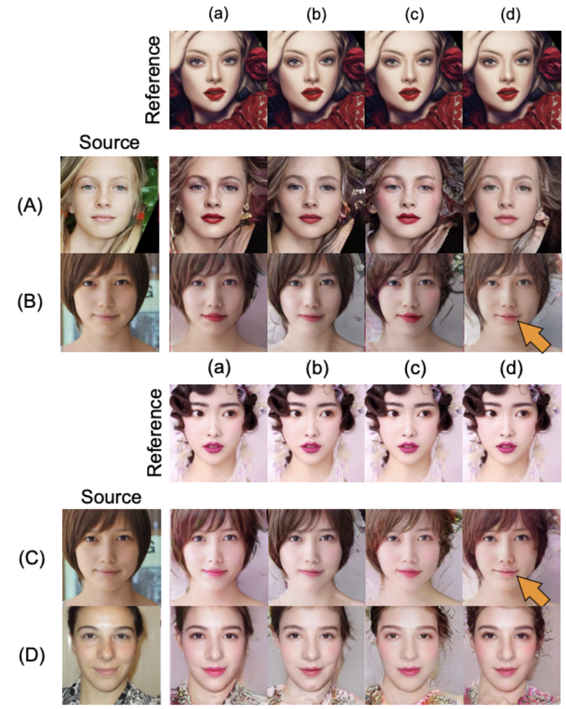

Fig. 39. SL GAN Results. source - [Horita and Aizawa 2022]

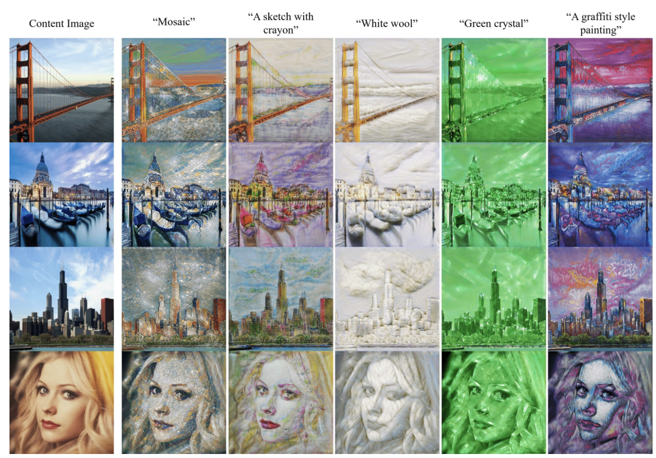

Fig. 36. CLIP Styling Results. source - [Kwon and Ye 2022]

Fig. 41. RPD-GAN Results. source - [Gao et al. 2020]

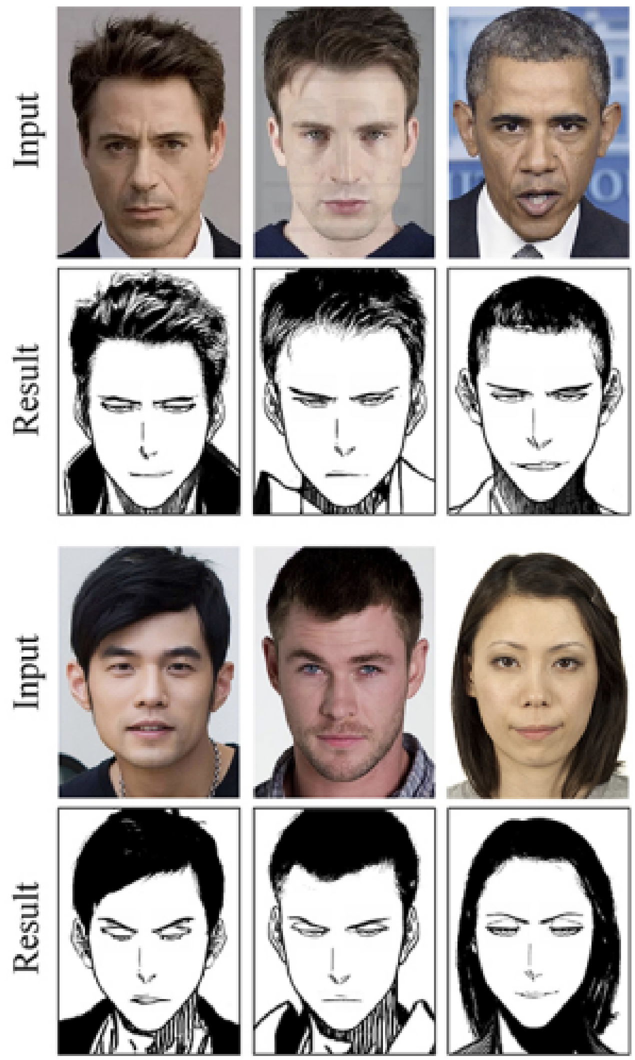

Fig. 40. Manga GAN Results. source - [Su et al. 2021]

In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)*. 1415–1419. https://doi.org/10.1109/ICMLA.2019.00231

Hao Su, Jianwei Niu, Xuefeng Liu, Qingfeng Li, Jiahe Cui, and Ji Wan. 2021. Mangagan: Unpaired photo-to-manga translation based on the methodology of manga drawing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 35. 2611–2619.

Wenju Xu, Chengjiang Long, Ruisheng Wang, and Guanghui Wang. 2021. Drb-gan: A dynamic resblock generative adversarial network for artistic style transfer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 6383–6392.

Wenju Xu and Guanghui Wang. 2021. A domain gap aware generative adversarial network for multi-domain image translation. *IEEE Transactions on Image Processing* 31 (2021), 72–84.

Lvmin Zhang, Yi Ji, Xin Lin, and Chunping Liu. 2017. Style Transfer for Anime Sketches with Enhanced Residual U-net and Auxiliary Classifier GAN. In *2017 4th IAPR Asian Conference on Pattern Recognition (ACPR)*. 506–511. https://doi.org/10.1109/ACPR.2017.61

Zhentan Zheng and Jianyi Liu. 2020. P^2-GAN: efficient style transfer using single style image. *arXiv preprint arXiv:2001.07466* (2020).

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*. 2223–2232.

Zhiwen Zuo, Lei Zhao, Shuobin Lian, Haibo Chen, Zhizhong Wang, Ailin Li, Wei Xing, and Dongming Lu. 2022. Style Fader Generative Adversarial Networks for Style Degree Controllable Artistic Style Transfer. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, Lud De Raedt (Ed.). International Joint Conferences on Artificial Intelligence Organization, 5002–5009. https://doi.org/10.24963/ijcai.2022/693 AI and Arts.