IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Apparel Design Using Image Style Transfer

[1]Sarika Nyaramneni, [2]Budati Gayathri, [3]K. Guna Sree, [4]Shaik Rabiya, [5]Shazia Shaguftha Hussain
[1]Professor at VNRVJIET, [2]B.TECH Student at VNRVJIET, [3]B.TECH Student at VNRVJIET,[4]B.TECH Student at VNRVJIET,
[5]B.TECH Student at VNRVJIET
[1]Department of Computer Science & Engineering,
[1]VNRVJIET, Hyderabad, India

Abstract: The Style Transfer Method has recently been applied to different applications, including social communication (e.g., Ostagram, Prisma), games, clothes, and movie. The style transfer algorithms have been effectively used to create new custom apparels in the fashion industry from which we have derived interest for his model.
In this paper, we aim to provide a user interface for generating different apparels as per user's choice providing them with multiple options to customize the apparel they want to generate. Most of the works haven't focused on user preferences, hence we chose to give importance to them. The variety of user options that we are including are Style Images Scaling, Multiple Style Image Transfers, Style application on Region of Interest, Color Preservation of Content Images, Style Weighting choice in case of Multiple Style Images, Blending of Style Images and also Changing colors of style images based on user's preferences. We employed Segmentation Model(U-NET) to extract masks of the clothing so that style transfer is only applied to clothing and not to any other area. The Style Transfer algorithm we used is IFT (Iterative Feature Transforms) which provides competition to Neural Style Transform (NST).

***Index Terms* - Apparel Generation, Image Style Transfer, Multi Style Transfer, Iterative Feature Transforms, Segmentation, Mask Generation, Blending.**

I. INTRODUCTION

One of the challenging aspects faced by every general non-professional is the proficiency in aesthetics, which is also a vital aspect in Apparel design. Various steps are involved in developing apparel styles. Namely designing, application of knowledge of fabrics, texture and coordinating colors. Texture being the prerequisite element of the design potentially expresses the feel and appearance of the surface.

Apparel design involves research, idea development, design concept and pattern making. Apparel Design can be generated via many approaches such as by using Image Style Transfer algorithms, Attribute Editing Generative Networks, Virtual Try On mechanisms, etc. In this paper, we mainly focus on Apparel design generation via Image Style Transfer.

Style transfer can be defined as a technique of computer vision that enables us to modify the image contents into the styles of different type. Style transfer facilitates us to view a photo in form for a painting by any popular artist. Style can be defined as the texture of an image. Texture of an image includes the brush strokes, the angular geometric shapes, patterns and the transition between colors.

This paper has the following sections: Section A. Apparel Design using Image Style Transfer presents the detailed comparative analysis of the various deployed Style transfer algorithms in Apparel Generation. Section B. Image Style Transfer explains the details about the previous research results on Style Transfer algorithms. Section C. Clothing Segmentation explains the different segmentation methods/models for clothing image separation from background.

II. LITERATURE REVIEW

2.1 Apparel Design using Image Style Transfer

The authors of [3] developed a fully automated system that designs clothing prototypes representing fashion trends identified by exploration, detection and research making use of time-series signals generated by social media feeds. First, trendy images are discovered using competitive intelligence, social media and google trends and they are passed through pretrained VGG-19[4] network. Style transfer algorithm [5] is applied on VGG-19 representation of low- resolution trendy images. SRCNN [6] is used to apply super resolution on output of style transfer algorithm to generate high resolution image. To qualitatively evaluate designs, a questionnaire was conducted. Inspiration and manufacturability score is calculated based on designer ratings. All designs got positive score for inspiration. Some designs got low inspiration and manufacturability score due to incompatible color combination.

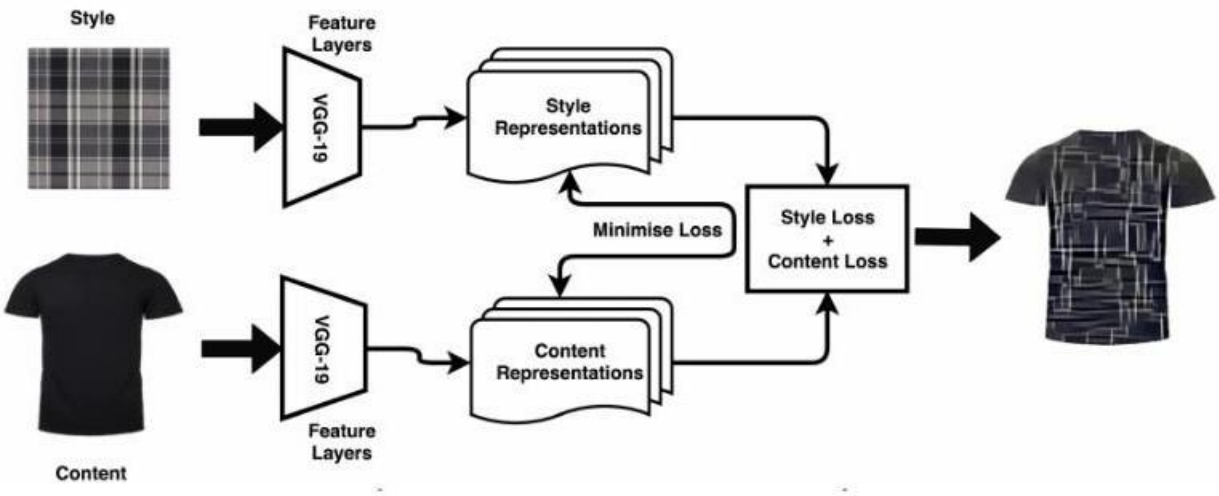


FIGURE 1. Application of Style Transfer technique

The authors of [7] suggested a method that assists designers to design apparel. DeepAttributeStyle, a new dataset is created, incorporating masks for the regions BackGround, Right Sleeve, Left Sleeve, Silhouette, Right Shoulder, Left Shoulder, Collar, Neck, Print and Hemline. DeepAttributeStyle has a wide range of information about apparel in form of annotations. They developed a system having three assistants. First, Creative Design Assistants Platform for Fashion to provide consumer insights by extracting attributes using multi-class attribute classification. Second, Apparel-Style-Merge Assistant to reconstruct new designs by segmenting multiple input apparels and positioning them appropriately. Segmentation model is trained on DeepAttributeStyle dataset using the framework Mask-RCNN. Third, Apparel- Style-Transfer Assistant to apply on the whole image,the algorithm of style transfer[1] and lastly only dress is cropped using DeepAttributeStyle segmentation model. Mask-RCNN segmentation model generated IOU score 0.90 for the class silhouette ,0.81 for the class hemline,0.78 for the class sleeves which indicated accuracy up to the mark. User survey indicated 90% users agreed on creativity and more than 80% agreed on quality.

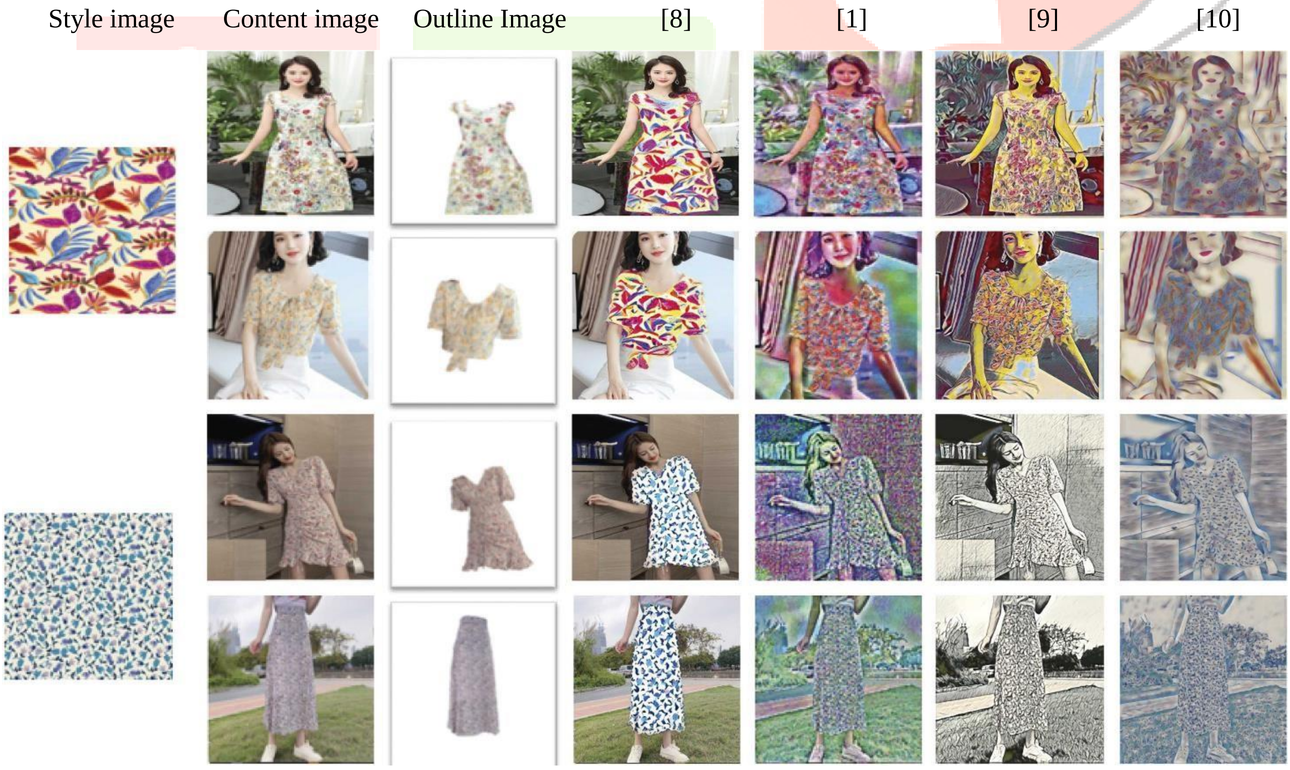


Figure 2: The fourth-left column shows results of applying method [8] given 1,2,3 columns as input. The fifth, sixth and seventh columns show the results on application of methods [1],[9],[10] taking 1,2 columns as input.

The authors of [8] introduced a method for clothing design based on interactive image localized style transfer. Users are allowed to drag a rectangle around desired region in the clothing image using. GrabCut algorithm is employed to separate the clothing from the background. A new loss function called outline loss is computed based on distance transform loss to enables perfect preservation of clothing shape. In image, noise is diminished using total variation loss. Style is applied to clothing image to create new clothing design minimizing style loss, content loss, outline loss and total variation loss. Pop-fashion.com, Taobao.com were used as sources for the experimental clothing images. On observing results of [8], [1], [9], [10] the background is kept clean by this method. A variety of artworks in diverse forms like watercolor, oil artwork are utilized as style images in experiments. Transferring large features present in style imagery such as patterns, figures to clothing images was discovered to be challenging based on experimental evidence, thus reaching the conclusion that the patterns which are smaller and distinct are transferred perfectly to clothing images.

Score	[39]	[12]	[11]
Inception Score	0.24	0.31	0.29
Human Preference Score	0.26	0.18	0.36
General Score	0.50	0.49	0.65

Table 1 The General Score of models DCGAN[4], Fashion-GAN[12],Clothing-STGAN[11]

The authors of [11] suggested a system for creating new clothing designs using style transfer technique and generative adversarial network. A collection of 52,908 images formed the dataset called Dunhuang. The images of size 64X64 are processed inunified manner with respect to the resolution. GAN module trained on Fashion MNIST datasetcontains generator using multi-layer CNNs to generate clothing images and discriminator employing a fully connected neural network and a multi-layer CNN and in order to determine the image given as input is fake or real. Style transfer algorithm[5] is used to generate clothing designs taking output image of GAN module as content image and style image from Dunhuang dataset. Table1 shows the Human Preference Score is 0.36 higher than Deep Convolutional Generative Adversarial Network[4], Fashion-Generative Adversarial Network [12]. However, the framework fails in handling complicated backdrop, missing information and information extraction from the clothing dataset Dunhuang, also in producing high resolution clothing designs.

Figure. 3. Fashion style generartors(a)FashionG (b) Spatially Constrained-FashionG

The authors of [14] introduced FashionG framework to generate clothing design from single style image and Spatially Constrained-FashionG framework to generate clothing design from multiple style images called mix-and- match based on the concept of constraining styles spatially to a particular region in clothing image using a spatial mask. FashionG and Spatially Constrained-FashionG consisting of a generator and a discriminator have end-to-end feedforward architecture. The discriminator computes the global style loss, the global content loss, the patch-based style loss and the patch-based content loss. These losses are backpropagated to the generator and optimized. Clothing form, design and patterns in style are preserved by optimization globally and locally. Fashion 144k data set [15] and Online Shopping dataset [16] are used for training FashionG and SC-FashionG frameworks. 288X288size for high resolution ,72X72 size for low resolution as content input image,128X128 size ascontent input image patch and 72x72 size as style input are used to for training. Sizes if inputsfor testing stage is arbitrary. The Chi squared test[17] is used to evaluate the user research results significance in comparison to random guessing. The Chi-squared score for designers and typical users, respectively, is 2241.554 and 808.161, with a P-value below 0.0001 in both situations.

In [18], deep learning algorithms for producing clothing patterns were centred on handloom fabric, and the accompanying challenges and applications were examined. Multiple methodologies are used in the work to study and observe the performance of style transfer algorithms and existing state of the art models for the task. Photographs of Indian sarees were used to collect data for the neural style transfer targets and image to image translation. The collection is divided into two categories which both of them are collectively referred as "Neural-Loom". Because no such dataset exists, there was a need to create a handcrafted dataset.
The most promising findings were fromNeural Style Transfer and DiscoGAN, which were both incorporated in the final application that allows users to experiment with the models. The findings demonstrate that these methods can be utilised to produce effective fusion designs, and a more specific definition of design needs to be devised and incorporated into the solution. More samples are also needed in the collection to remove any inherent bias and provide a more different and diversified data. The results were based on a

user evaluation done among 53 participants, 35 were male and 18 were female. Each of the participant was given a set of images that included both the actual and created designs. The participants were requested to label the photos "Real," "Generated," or "Not Sure," as well as grade the design as "Good," "Bad," or "Maybe". It was shown that 24.5 percent of users evaluated the "Generated" design as "Undecided", 29.7% as "Bad," and 45.8% as "Good". According to [18], more samples were needed in the collection to remove any inherent bias and provide a more colorful and diverse dataset to work with.

Figure 4: Regional Handloom" data

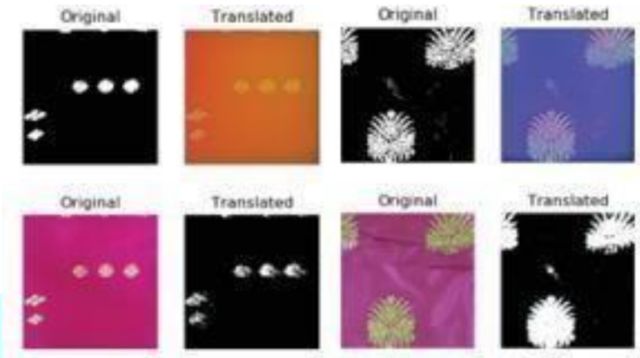


Figure 5: DiscoGAN results

Figure 6: Labeling percentages by participants. The vertical axis represents the image kind. Horizontal is a probability for the same.

[19]presents an innovative generative model-based technique for automating the creation of textile design patterns. The paper assessed the performance of image generating models such as DCGANs, CVAEs and WGANs GP across all categories and verified the methods with the inception score. A data set for such photos is needed for image generation of textile designs. The data is from a study in which the goal was to categorize textile designs using a finger-installed camera.

Checkered, solid, striped, zigzag, and flowery designs are among the six types of patterns included in the dataset. There were 2500 classified photos in the whole data set, with 400 images in each class on average. The photographs were downloaded after searching for keywords such as "checkered," "floral," and "stripes" on Google. Because of the incorrect photographs present in the dataset, they had to clean the data. Textile designs are intricate enough that many patterns can be combined. Because the dataset contains images from six classes, each of which contains only one pattern and no composite pattern, the WGAN- generated designs were not complex enough to be deemed textile designs, thus they applied Gatys' approach [20] to create complicated designs using Neural Style Transfer. Their main motive for using NST is that it is capable of producing excellent outcomes in the transformation of art styles and paintings. Because textile designs are works of art with a variety of styles underlying them, combining several styles intoone can provide very satisfactory results. Outputs created with WGANs were further selected sothat the input image class and style of the outputs remained distinct. The designs that resulted have properties of one class but a different style, such as checkered patterns stylized in a zigzag pattern, and so on.

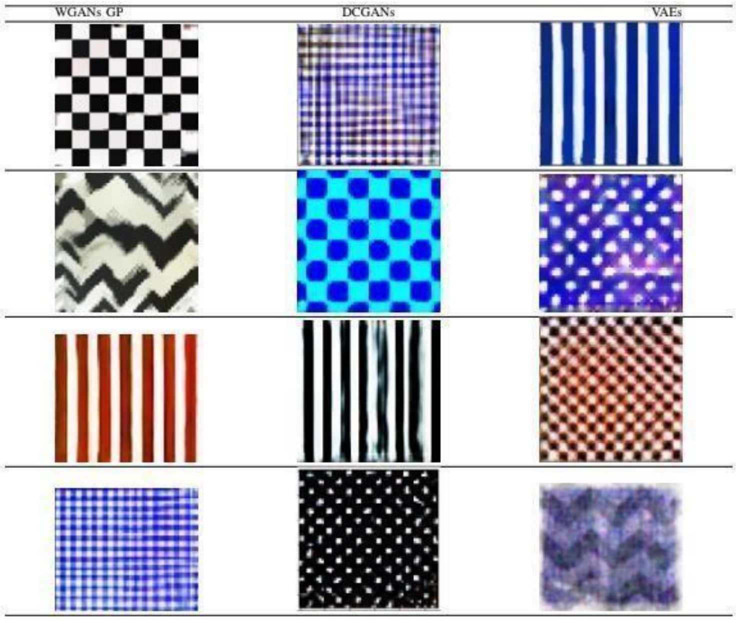


TABLE 2: Outputs of various generative algorithms for the production of textile patterns

Class Name	*DCGAN*	*WGAN*
Solid	*5.03*	*5.16*
Checkered	*3.24*	*5.03*
Dotted	*3.41*	*4.27*
Floral	*2.10*	*3.27*
Striped	*3.57*	*4.70*
Zigzag	*2.33*	*4.58*

TABLE 3: Inception score of WGAN and DCGAN model for vivid designs

2.2 Image Style Transfer

Image style transfer is the process of re-rendering the content of one image in the style of another. Most well-known algorithms combine content and style information as a black-box in their network architecture and hyper-parameters. [21] attempts to establish a new explicitly detached worldview. They proposed StyleBank, a collection of convolution filter banks that represent various styles. The related filter bank transforms a photo into a specific style using an intermediary feature provided by a single auto-encoder. Because both the auto-encoder and the StyleBank are learned at the same time, the auto-encoder doesn't encode any style features. Their method is the first to link classic texton mapping methods to style transfer networks, revealing new information about brain style transmission. [21] solved two multi-parameter image processing problems with this general idea of filterbank learning: edge-aware picture smoothing and denoising. From the Microsoft COCO data collection, the network was trained on 1,000 content shots and 50 style images at random (from the internet and existing papers). Every style image was scaled to 600 pixels on the long side and every content image was cropped to 512*512 pixels at random. The network is trained for 300,000 iterations with a batch size of four (m 14 4 in Algorithm 1). With a learning rate of 0.01 at the start and decaying by 0.8 per 30k iterations, the Adam optimization strategy is applied. They used PSNR(peak signal to noise signal) and SSIM(Structural similarity index) to compare Lo(baseline= 36.68 baseline*=36.60 our-s=36.70 our-h=36.99), RTV(baseline=38.72 baseline*=38.84 our- s=38.37 our-h=38.52), RGF (baseline= 41.34 baseline*=41.27 our-s=40.65 our-h=40.82) in TABLE 4.

They presented an efficient and effective Avatar-Net in their study [22], which enables visually believable multiscale transfer for any style. They employed a style decorator to create content features from semantically aligned style features received from any style image, which not only matched their feature distributions holistically but also kept detailed style patterns in the decorated features.The Avatar-Net enables multi-scale stylization for each style image in a single feed-forward pass by integrating this module into an image reconstruction network that fuses multiscale style abstractions. Video stylization, multiple style integration, and other applications were used to demonstrate the suggested method's advancedefficiency and effectiveness in generating high-quality stylised images. The architecture of a pretrained VGG19 is copied to the encoder using Avatar-Net. The encoder is randomly initialised, whereas the decoder is mirrored, with nearest upsampling replacing all pooling layers and reflectance padding replacing padding layers.

Around 80, 000 training samples from the MSCOCO dataset were used to build the network. With a fixed learning rate of 0.001 and a batch size of 16, the Adam optimizer is used. The training images are randomly resized and cropped to 256 256 patches during the training phase. The technique can be applied to any image size. They put their model to the test against two zero-shot stylization techniques: iterative optimization and feed-forward network approximation. Each method's method and execution time are as follows: Gatys et al. 12.18(256 × 256 (sec)) 43.25(512 × 512 (sec)), AdaIN 0.053(256 × 256 (sec)) 0.11(512 × 512 (sec)), WCT 0.62(256 × 256 (sec)) 0.93(512 × 512 (sec)), Style-Swap 0.064(256 × 256 (sec)) 0.23(512 × 512 (sec)), ZCA 0.26(256 × 256 (sec)) 0.47(512 × 512 (sec)), ZCA-Sampling 0.24(256 × 256 (sec)) 0.32(512 × 512 (sec)), AdaIN 0.071(256 × 256 (sec)) 0.28(512 × 512 (sec)).

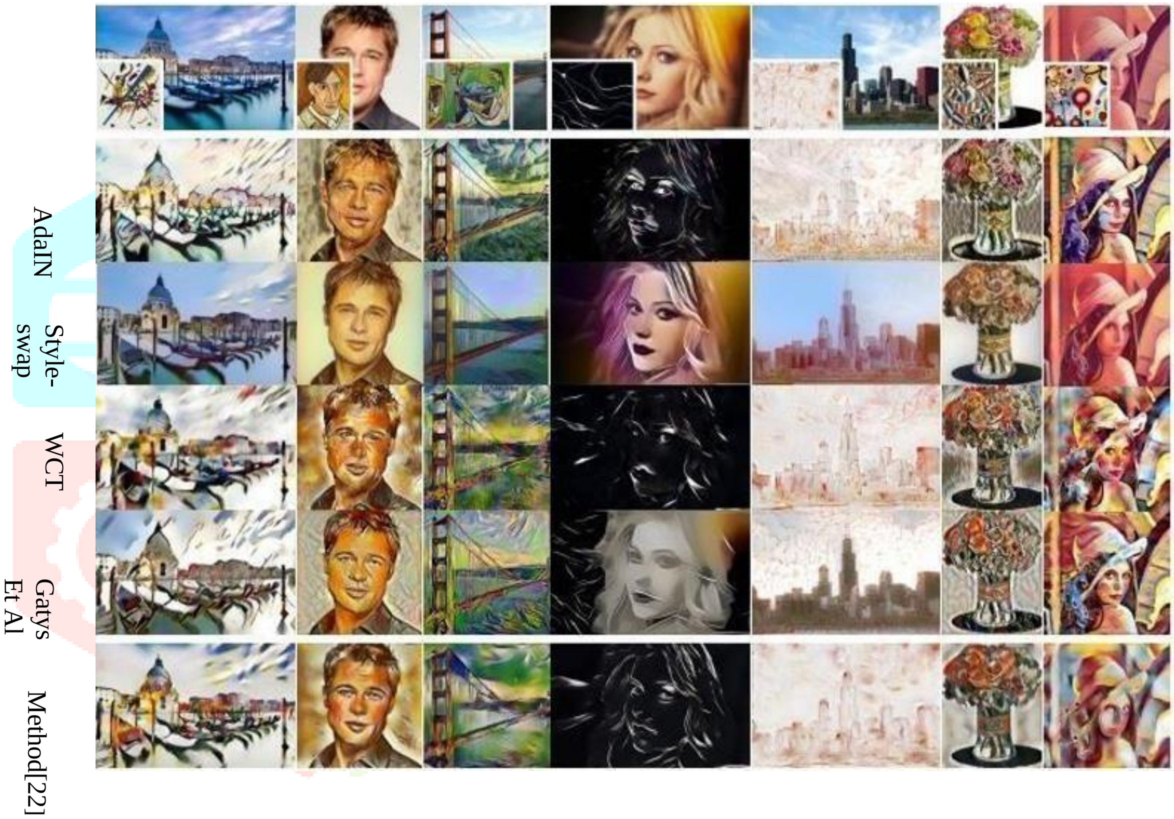


Figure 8. Some exemplar stylization outcomes are compared. The first row is the content and style image set, the other rows are as follows: [24], Style-Swap, [25], [1], and the last row is proposed method's stylized outputs

[23]compared AdaIN[24], WCT[25], SANet[26], and AAMS[27] to their technique. AdaIn [24]modifies the content image's mean and variance to match the style image worldwide. Despite being effectively preserved, the content architecture may result in poor textual patterns being reproduced in stylised graphics. Furthermore, the colour distribution of the results may alter depending on the style image. By modifying the covariance of the content picture through whitening and colouring modification operations, the WCT[25] improves AdaIN's style performance. The WCT, on the other hand, would distort the content. Style-attention is used by the SANet[26] to match style aspects to content features, resulting in beautiful styled outcomes with distinct style texture. However, it is possible that it will duplicate some style patches onto styled images, distorting the content structures. Styled images can retain the structure of the major section of content visually because the AAMS[27] is based on self-attention.

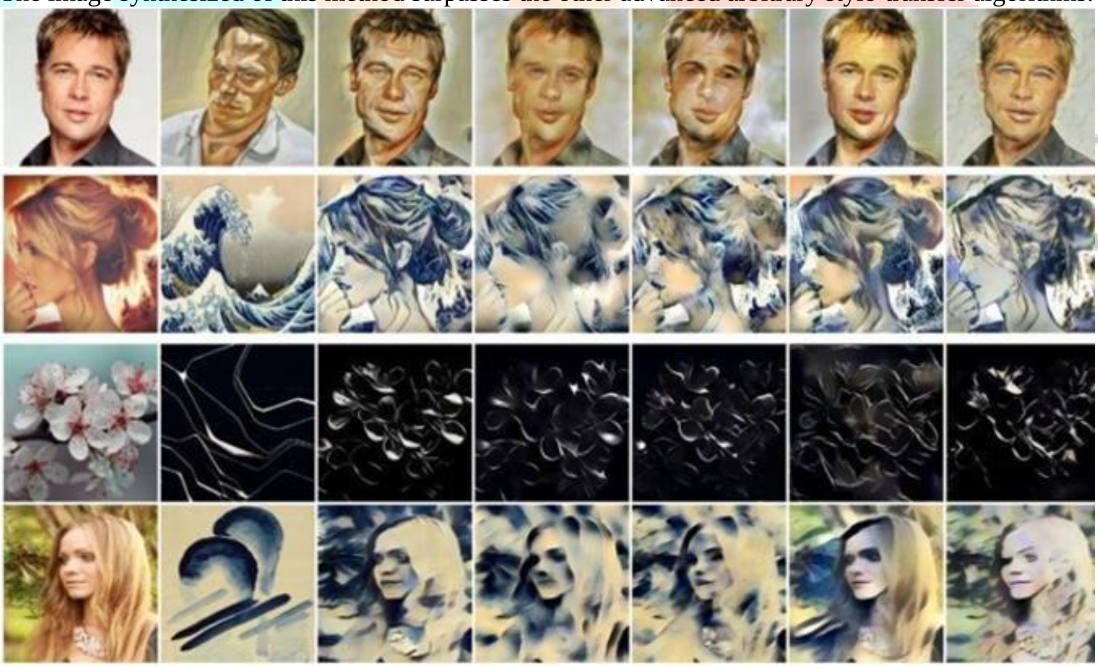
Attention map calculated by AAMS, on the other hand, has a significant impact onthe outcomes. This method may produce disappointing results in which just part of the contentstructure is preserved and the background is heavily blurred. When the content image's structures aren't evident and the style reference is monotonous, the model's fundamental weakness is that it can't handle the local association between content and style, resulting in unsatisfactory results.

Style Content[27] AdaIN[24] WCT[25] SANet[26] AAMS[27]

Figure 9: SOTA methods are compared to stylised results. The first column contains images of style, while the second has images of content. The remaining columns represent stylized AdaIN [24], WCT [25], SANet [26], and AAMS [27] results.

The author [28] put forth a style attentional network which integrates local style patterns as per semantic spatial distribution (content image). The learning approach network is composed of the two modules which is optimized with help of a new identity loss and a conventional style reconstruction loss. Various styles content-image can be essentially transferred with high quality. This network of style transfer comprises of two modules. They are encoder – decoder and style attentional. The new identity-loss function enables in retaining detailed structure while reflecting the style of the content efficiently. With both conventional style reconstruction loss and identity loss, the proposed SANet efficiently decorates the style features. This method is approximately 20 times faster than the methods that are based on matrix computation. Structure of the image and attributes of style are simultaneously maintained with the help of identity loss. Experimentally this method has been shown to be extremely efficient at producing high-quality styled images while also balancing the style pattern both globally and locally while preserving content structure.

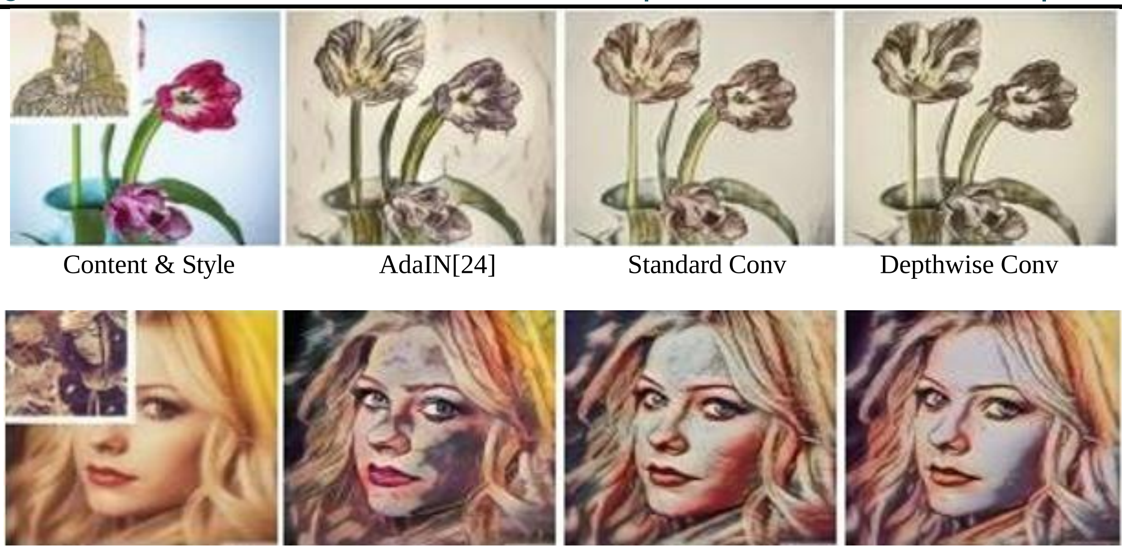
The image synthesized of this method surpasses the other advanced arbitrary style-transfer algorithms.



ContentStyle Method[28] WCT[25] Avatar-Net[22] Gatys et al.[5] AdaIN[24]

Figure 10: Comparison of stylized results

The authors of [29] introduces novel (DIN) Dynamic Instance Normalization layer which is highly pliable with greater efficiency as an arbitrary style transfer. Unlike the conventional methods utilizing shares complex encoders which necessarily help in encoding content and the style, this DIN acknowledges a style-encoder which is elaborate in order to encode a rich style pattern andthe light weight content encoder for the enhance performance, consequently resolving the predicament of shared encoder and redundant. Experimental result upheld the positive outcome against the other advanced methods precisely in transferring demanding style pattern using Mobile Net- based light weight architecture along with maintaining reduced computational cost. The DIN proposed here is capable of producing novel effects in style transfers i.e., incorporating the advanced convolutional operations for automatic spatial stroke control. There is further scop-e to research and explore regarding the usage of DIN computer vision task beyond style transfer.



Content & Style AdaIN[24] Standard Conv Depthwise Conv

Content & Style AdaIN[24] Single-level DIN Hierarchial DIN

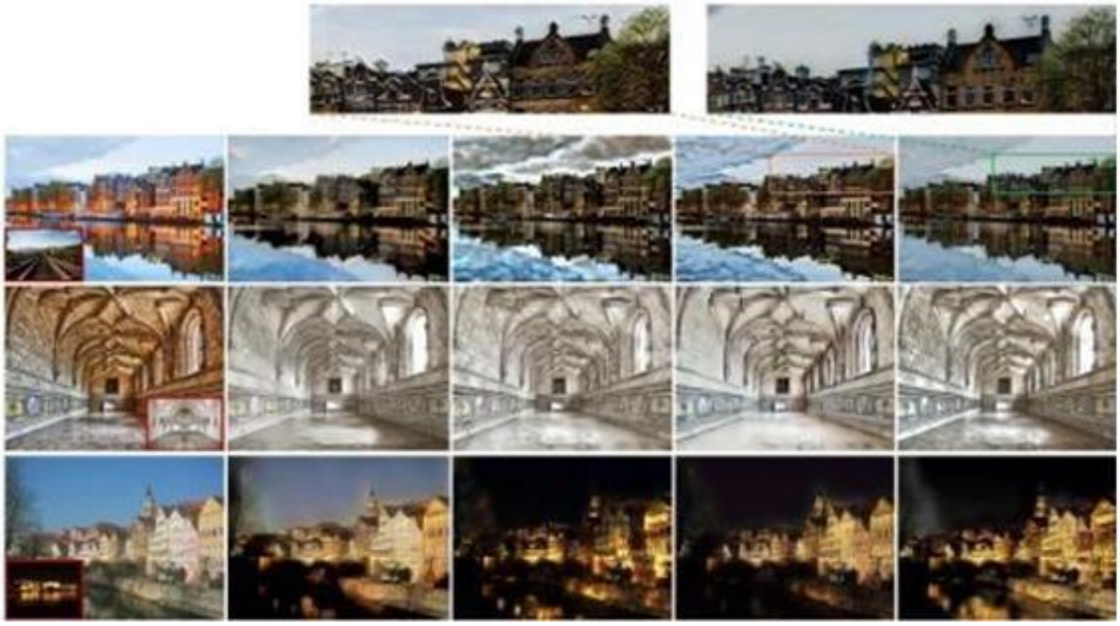
Figure 11: Results acquired by using standard convolutions, depth wise convolutions, single- level DIN, and hierarchical DIN respectively.

The authors of [30] proposed Adaptive Convolutions(AdaConv). AdaConv is an Adaptive Instance Normalization extension. This anticipates a biases and convolutional kernels from a given style embedding, which can then be used to generate layers of image decoders to change its behavior properly during the test period. AdaConv successfully transfers both global statistics as well as the spatial structures of a style image onto a content image, whereas AdaIN transfers only the style's global statistics but fails to transfer the structure of style. This is applicable in generating stye based image and practically AdaIN has been utilized everywhere. User study to check the efficiency of AdaConv a strong majority of participants nearly 71.8% shows adaptive convolutions for performing better style transfer job. This produces a new general building block for incorporating input data conditioning into convolutional neural networks-based image production and style transfer. When AdaConv is used to generative models like StyleGAN, it produces photo realistic images on a huge dataset.

Style image Content Image AdaConv Jing et al.[29] Li et al. [25] Huang.[22] John et al.[39] Belongie[24]

Figure 12: AdaConv works well for transferring style image's local structure to the content image.

The authors of [31] proposed a novel that sequentially applies style to the features using the concept of analytical gradient. The universal framework for the rapid universal style transfer comprises of an auto encoder along with feature-transformation. Experimentally this transformation has upper handas it is fast. This feature is used for transformation of universal style transfer that emerges out by estimating the solution to the objective of neural style transfer. This method offers control knob that balanced content reservation in style effect transferal. It also helps in switching between photo realistic and artistic style transfer. For adaptivity this method aids multiple style transfer and semantic control. Comparing the auto encoder-based algorithm this feature evolves maximum scores in SSIM as well as in FSIM metric showing that it succeeds in preserving the content to the finest. When compared with NST, the effectiveness of this method is shown and representative transformation for fast style transfer. Every style's integrity is well preserved than in WCT. It also shows greater extent of satisfactory result compared to that of NST.

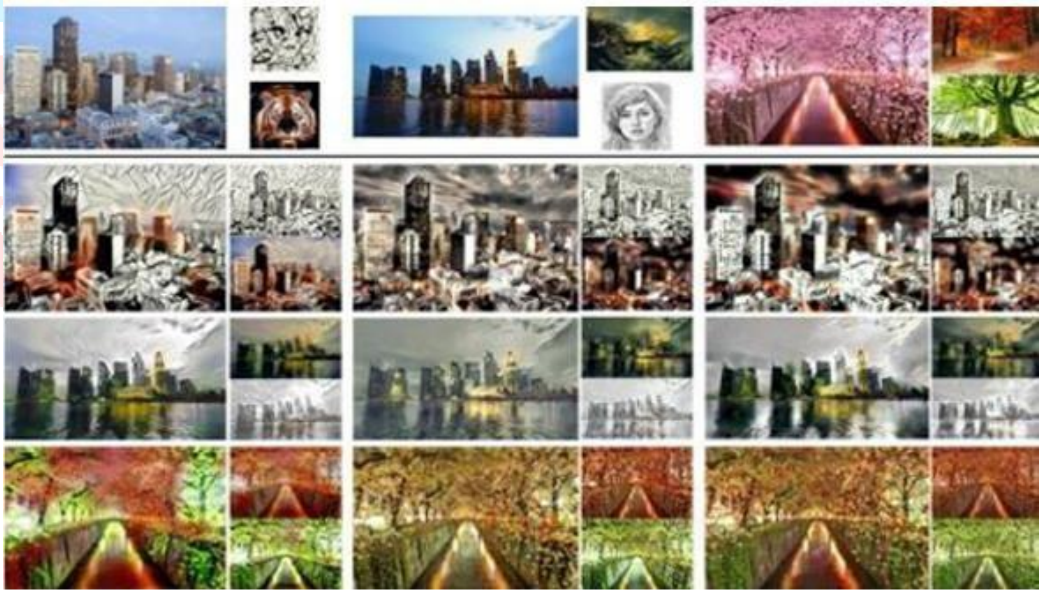


Content & Style NST[5] WCT[25] LST[40] Method[31]

Figure 13: Style transfer results for NST, WCT, LST and proposed algorithm

	NST[5]	Autoencoder-based		
		WCT[25]	LST[40]	Method[31]
SSIM	0.5644(0.7567)	0.4476(0.7265)	0.4429(0.6919)	0.5301(0.7513)
FSIM	0.7870	0.7643	0.7658	0.8197

Table 8: Evaluation of the level of distortion. Among the 3 algorithms, this feature transformation achieves the best SSIM and FSIM values while preserving the most content.



NST[5] WCT[25] Method[32]

Figure 14: Comparison of the results of methods [5],[25] & [31] on application of two styles on single content image.

Authors of [32] proposed a novel attention along with a normalization module, termed as (AdaAttN) Adaptive Attention Normalization essentially for the arbitrary style-transfer. It takes into account of shallow as well as deep features for attention score calculation & normalizing content feature in a manner that the features statistics alignment of attention weighted are in well alignment with that of attention weighted mean as well as variance-map of style features operating on the basis of every point. Local feature loss is a new optimization target that governs local features and helps not only in model training but also improves the quality of style transfer (generated image). Experiments and comparison with the complex approaches are carried out to prove efficacyof the stated method by extending this for video style transfer. This is done by the means of simple introduction of cosine- distance based attention & similarity loss resulting in a firm & satisfactory result. Potential scope of implementing AdaAttN for improving other broad scope of image-manipulation or translation-task are yet to be explored.

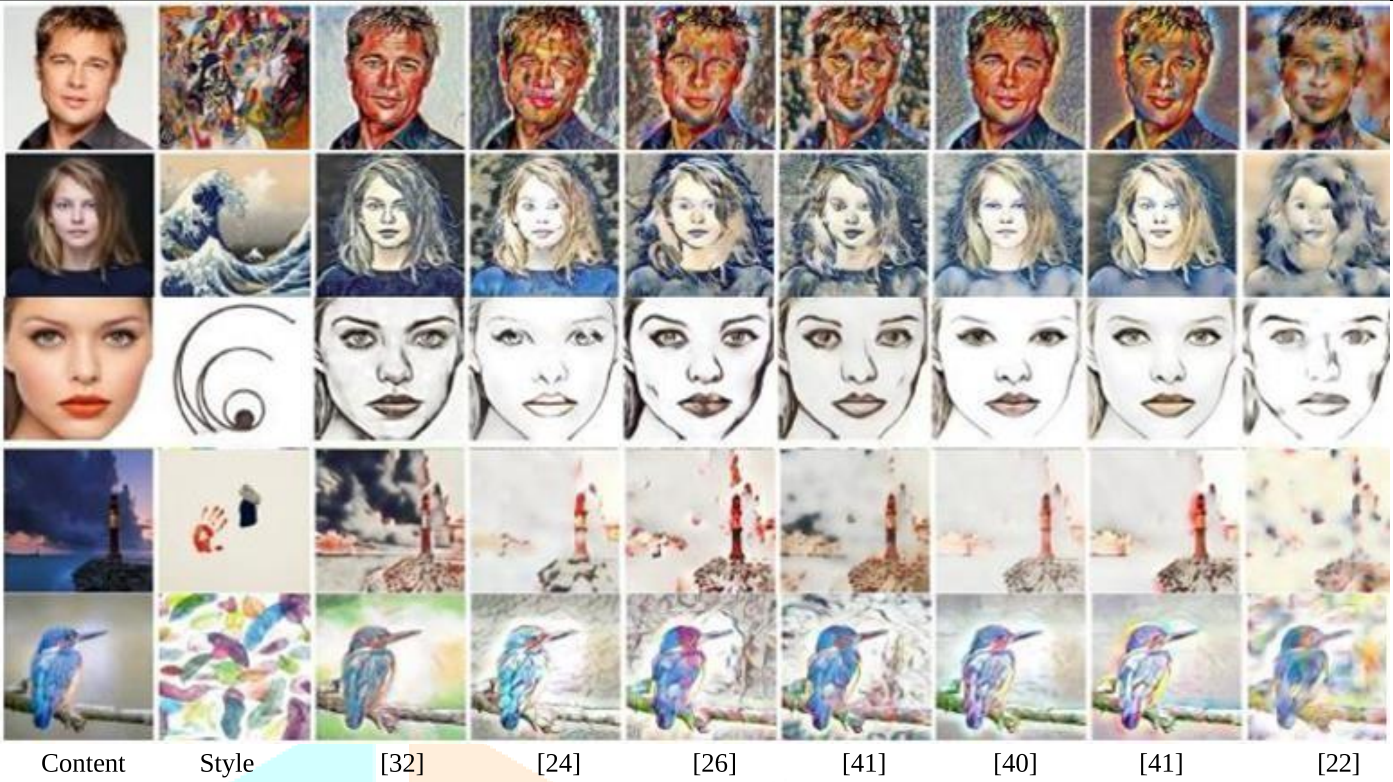


Figure 15: Comparison of [32] with [24],[26],[41],[40],[22] methods.

Methods	Inference Time (sec./image)		
	256x256	512x256	512x512
Avatar-Net [22]	0.124	0.176	0.311
AdaIN [24]	0.038	0.049	0.066
Linear [40]	0.028	0.036	0.049
MCCNet [41]	0.024	0.040	0.057
MAST [41]	0.046	0.073	0.115
SANet [26]	0.043	0.064	0.081
AdaAttN	0.051	0.066	0.112

Table 9: Execution speed comparison at resolutions of 256x256,512x256,512x512.

2.3 Clothing segmentation

Clothing segmentation is a grueling visual task generally enforced within a finely granular semantic segmentation frame. Unlike traditional segmentation, apparel segmentationhas some sphere-specific parcels similar as texture uproariousness, different appearance changes, nonrigid figure distortions, and limited sample literacy. To address these issues, the authors of [33] proposed a semantic position-apprehensive segmentation model, that adapts by hunting primary garment photo with a semantically equivalent (example., disguise) additional example. There is more natural information about the position of various structures in apparel photos by considering the relationships between the apparel image and its exemplar, making the literacy process of the small sample problem more stable and compliant. Likewise, they presented a CNN model grounded on the deformable complications to prize then on-rigid geometry aware features for apparel photos. The model was trained on the CFPD and Fashionista training dataset.

Figure 16: From left to right, the figure depicts raw image, ground truth image, and the network's outcome.

Deep Fashion [34] includes numerous flaws, such as a single apparel-itemper picture, meager milestones, and no per-pixel masks, separating it from real-world scripts. The authors of [35] fill in the blanks by proposing DeepFashion2 as a solution to these problems. Clothing discovery, pose estimation, segmentation, and reclamation are all part of this complex standard. DeepFashion2 has 801K apparel particulars, with rich reflections such as style, standpoint, occlusion, scale, masks, thick milestones, and bounding box on each item. In addition, there are 873K Marketable-Consumer clothing dyads. DeepFashion2 has substantially greater reflections than its predecessors, similar as 8 X in the FashionAI Global Challenge. Match RCNN, a strong birth that constructs on Mask R-CNN to break the four tasks completely, is proposed.

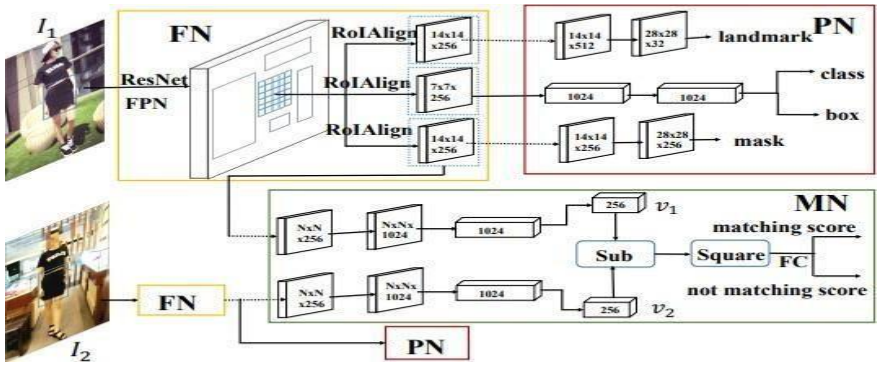


Figure 17: Match R-CNN diagram showing the three key components: match network (MN), feature extraction network (FN) and perception network (PN).

	zoom-in no medium large	viewpoint no wear frontal side or back	scale small moderate large	occlusion slight medium heavy	overall
APmask	**0.703** 0.627 0.526	0.695 **0.697** 0.617	0.634 **0.700** 0.669	**0.720** 0.674 0.389	0.680
$AP^{IOU=0.50}{}_{mask}$	**0.899** 0.815 0.663	0.829 **0.886** 0.843	0.831 **0.900** 0.844	**0.900** 0.878 0.559	0.873
$AP^{IOU=0.75}{}_{mask}$	**0.842** 0.740 0.613	0.792 **0.834** 0.732	0.765 **0.838** 0.786	**0.850** 0.813 0.463	0.812

Table 10: Mask R-CNN detects clothes on several validation subsets, such as zoom-in, viewpoint, scale and occlusion. APbox, APIoU=0.50 box, and APIoU=0.75 box are the evaluation measures.

The authors of [36] proposed a reclamation system grounded on apparel segmentation and apparel image detection. Originally, Mask R-CNN was utilized to descry and member the image to obtain details about the apparel, collar corridor, sleeve order and fund positions, also VGG16 was utilized to prize 512-dimensional features from the apparel and collar corridor, grounded on these details, the comparability between the apparel to be recaptured and the apparel in the repository was computed one after another. The hunt results are conferred to the stoner in decreasing order of resemblance. The findings reveal that the system can concentrate on the complete apparel and their corridor, therefore facilitating reclamation grounded on apparel style. It permits druggies to acclimate the weights of each part and produce seek outcomes that are trendy and match their specific needs. When the image is 180 degrees rotated, the method used in this research paper can only recoup appreciatively arranged apparel photos, and still, it'll affect the retrieval of pocket position information.

Figure 18: Clothing detection and segmentation using Mask R-CNN

The authors of [37] proposed variety of semantic segmentation models capable of linking many fashion details in pre-determined sequences. The authors of [37] trained on 5 different models for image segmentation. Those models include FCN Resnet50, Unet, Atrous Resnet50, DenseNet, SegNet and worked on iMaterialist (Fashion) 2019 dataset (iMFD). It consists of 50000 apparel images. The results are evaluated using metrics such as Intersection over Union (IoU) and Accuracy.

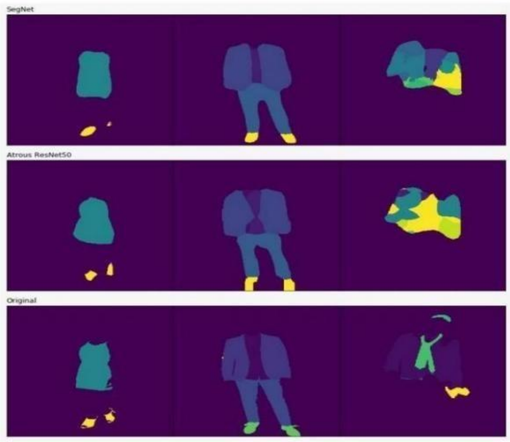


Figure 19: Comparison of SegNet vs Atrous ResNet50

The authors of [38] proposed a method for the problem of segmenting fashion photographs semantically into distinct orders of apparel. For the reason that the significance of both textural information and indications from appearance and the environment, this topic provides unique challenges. At the end of this paper [38], the authors presented a completely convolutional neural network grounded on Feature Pyramid Networks (FPN) and a backbone complying to the ResNeXt armature to achieve this goal. The suggested model achieves advanced outcomes on 2 standard datasets, according to the experimental results. The presented model is advantageous from a computational standpoint because it has a small memory footprint and can be employed without a conditional random field (CRF) without significant quality decline. The technique performs better than all advanced models in all estimated criteria on both datasets by wide margin when compared to all other styles without CRF. Actually, without the CRF, the technique achieves a higher level of delicacy than advanced models that use CRFs.

	Refined Fashionista		CFPD	
	IoU	Accuracy	IoU	Accuracy
FPN	**49.81**	**93.26**	**53.00**	**93.52**
OE	46.40	91.50	51.42	91.52
PSPNet	46.68	92.53	-	-

Table 11: On the revised fashionista dataset and the CFPD dataset without CRF, the mean intersection- over-union and average per-pixel accuracy for FPN and two advanced models.

+CRF	Refined Fashionista		CFPD	
	IoU	Accuracy	IoU	Accuracy
FPN	50.64	**93.62**	54.39	**93.82**
OE	**51.78**	91.74	**54.65**	92.35
PSPNet	47.85	92.93	-	-

Table 12: On the revised fashionista dataset and the CFPD dataset with CRF, the mean intersection-over- union and average per-pixel accuracy for two advanced models and FPN.

III. CONCLUSION

Object identification, image segmentation, object recognition and texture synthesis have all been successfully achieved using convolutional neural networks in recent years. Deep Convolutional Neural Network (DCNN) was trained in 2012 by Krizhevsky et al. ImageNet's object identification efficiency has been improved significantly. Many research investigations and publications have resulted as a result of these changes in the fashion industry, such as clothing parsing, recommendation, clothing classification, retrieval. In this paper, we put forth the key research papers which use Single and Multiple Image Style Transfer for designing apparel like Neural Style Transfer, Generative Networks (containing encoder and decoders) and major image and cloth Segmentation models. Most of the papers on NST are using Gaty's et al style algorithm and pre-trained VGG-19 network for applying style transfer with feature additions to gain appealing outputs. In the field of fashion and generating apparel the dataset plays an important role as demonstrated by [18], [19]. The more diverse and complex the dataset containing style images are, the more aspects in apparel generation can be covered. After generating a resultant image from style transfer, as we are dealing with clothing there's a need to extract foreground (i.e. remove background) for which few papers employed using of Grab-Cut utility [11]. The existing approaches can be further extended. [33], [35], [36], [37], [38] employed Segmentation Models which can be used for separating the clothing part to the rest of the background in content image to allow Style Features of style image to be applied only on the clothing part instead of the entire image by generating mask from the content image (masks generated were either binary or multi-colored). For example, GAN-based models combined with Super-Resolution models can be utilized to produce randomized and high resolution images. The Generative model's generalization capacity could be enhanced. During the testing phase, segmentation can be used to deal with complicated backgrounds andstances can be included. The images generated by the frameworks, however, are low quality, and the generated apparel styles are limited. In the future, the focus should be on producing more high- resolution apparel creations.

REFERENCES

[1] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks,," in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414–2423, Las Vegas, NV, USA, June 2016.

[2] Yongcheng Jing, Yezhou Yang, Zunlei Feng, Jingwen Ye, Yizhou Yu, and Mingli Song. 2019. Neural style transfer: A review. IEEE transactions on visualization and computer graphics (2019)

[3]Abhinav Ravi, Arun Patro, Vikram Garg, Anoop Kolar Rajagopal, Aruna Rajan, and Rajdeep Hazra Banerjee. 2019.Teaching DNNs to design fast fashion. In KDD 2019 Workshop: AI for fashion The fourth international workshop on fashion and KDD, August 2019, Anchorage, Alaska - USA., 8 pages. https://doi.org/10.1145/1122445.1122456

[4] Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.

[5] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. 2016. Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2414–2423

[6] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. 2016. Image super- resolution using deep convolutional networks. IEEE transactions on pattern analysis and machine intelligence 38, 2 (2016), 295–307

[7] Alpana Dubey, Nitish Bhardwaj, Kumar Abhinav, Suma Mani Kuriakose, Sakshi Jain, and Veenu Arora. 2020. AI Assisted Apparel Design. In KDD '20 Workshop on AI for fashion supply chain, 24 August 2020, San Diego, USA. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.

[8] Hanying Wang, Haitao Xiong, Yuanyuan Cai, and Silvia Conforto. 2020. Image Localized Style Transfer to Design Clothes Based on CNN and Interactive Segmentation. Intell. Neuroscience 2020 (2020). DOI:https://doi.org/10.1155/2020/8894309

[9] H. Zhang and K. Dana, "Multi-style generative network for real-time transfer,"in Proceedings of the European Conference on Computer Vision (ECCV), pp. 349– 365, Munich, Germany, September 2018.

[10] T. Q. Chen and M. Schmidt, "Fast patch-based style transfer of arbitrary style," in Proceedings of the NIPS Workshop Constructive Machine Learning, pp. 1–9, Barcelona, Spain, December 2016.

[11] ClothGAN: generation of fashionable Dunhuang clothes using generative adversarial networks, Connection Science, 33:2, 341-358, DOI: 10.1080/09540091.2020.1822780

[12] Cui, Y. R., Liu, Q., Gao, C. Y., & Su, Z. (2018). FashionGAN: Display your fashion design using conditional generative adversarial nets. Computer Graphics Forum, 37(7), 109– 119.https://doi.org/10.1111/cgf.2018.37.issue-7

[13] Kato, N., Osone, H., Oomori, K., Ooi, C. W., & Ochiai, Y. (2019). FashionGAN: GANs-based clothes design:pattern maker is all you need to design clothing. In AH2019: Proceedings of the 10th augmented human international conference (pp. 1– 7). Association for Computing Machinery

[14] S. Jiang, J. Li and Y. Fu, "Deep Learning for Fashion Style Generation," in IEEE Transactions on Neural Networks and Learning Systems, doi: 10.1109/TNNLS.2021.3057892.

[15] E. Simo-Serra and H. Ishikawa, "Fashion style in 128 floats: Joint ranking and classification using weak data for feature extraction," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2016, pp. 298–307.

[16] M. H. Kiapour, X. Han, S. Lazebnik, A. C. Berg, and T. L. Berg, "Where to buy it: Matching street clothing photos in online shops," in Proc. IEEE Int. Conf. Comput. Vis., Dec. 2015, pp. 3343–3351.

[17] S. Jiang, Z. Wang, A. Hertzmann, H. Jin, and Y. Fu, "Visual font pairing," IEEE Trans. Multimedia, vol. 22, no. 8, pp. 2086–2097, Aug. 2020

[18] Bhattacharjee, Rajat Kanti; Nandi, Meghali; Jha, Amrit; Kalita, Gunajit; Barbhuiya, Ferdous Ahmed (2020). [IEEE 2020 IEEE International Conference on Image Processing (ICIP) - Abu Dhabi, United Arab Emirates (2020.10.25-2020.10.28)] 2020 IEEE International Conference on Image Processing (ICIP) - Handloom Design Generation Using Generative Networks. , (), 1606–1610. doi:10.1109/icip40778.2020.9190925

[19] Fayyaz, Raja Asim; Maqbool, Muaz; Hanif, Muhammad (2020). [IEEE 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) - London, ON, Canada (2020.8.30-2020.9.2)] 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) - Textile Design Generation Using GANs. , (), 1– 5. doi:10.1109/ccece47787.2020.9255674

[20] Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[21] IEEE Transactions on Pattern Analysis and Machine Intelligence, Explicit Filterbank Learning for Neural Image Style Transfer and Image Processing July 2021, pp. 2373- 2387, vol. 43, DOI Bookmark: 10.1109/TPAMI.2020.2964205

[22] Avatar-Net: Multi-scale Zero-shot Style Transfer by Feature Decoration, Lu Sheng, Ziyi Lin, Jing Shao, Xiaogang Wang arXiv:1805.03857 [cs.CV]

[23] Arbitrary Style Transfer via Multi-Adaptation Network, Yingying Deng, Fan Tang, Weiming Dong, Wen Sun, Feiyue Huang, Changsheng Xu arXiv:2005.13219 [cs.CV]

[24] Xun Huang and Belongie Serge. 2017. Arbitrary style transfer in real-time with adaptive instance normalization. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).IEEE,1501–1510.

[25] Yijun Li, Chen Fang, Jimei Yang, Zhaowen Wang, XinLu, and Ming-HsuanYang. 2017. Universal style transfer via feature transforms. In Advances in neural information processing systems.386–396.

[26] Dae Young Park and Kwang HeeLee. 2019. Arbitrary Style Transfer with Style- Attentional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.IEEE,5880–5888.

[27] Yuan Yao, Jianqiang Ren, Xuansong Xie, Weidong Liu, Yong-Jin Liu, and Jun Wang.2019. Attention-aware multi-stroke style transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.IEEE,1467–1475

[28] Park, Dae & Lee, Kwang. (2018). Arbitrary Style Transfer with Style-Attentional Networks.

[29] Jing, Yongcheng & Liu, Xiao & Ding, Yukang & Wang, Xinchao & Ding, Errui & Song, Mingli & Wen, Shilei. (2020). Dynamic Instance Normalization for Arbitrary Style Transfer. Proceedings of the AAAI Conference on Artificial Intelligence. 34. 4369-4376. 10.1609/aaai.v34i04.5862.

[30] Chandran, Prashanth & Zoss, Gaspard & Gotardo, Paulo & Gross, Markus & Bradley, Derek. (2021). Adaptive Convolutions for Structure-Aware Style Transfer. 10.1109/CVPR46437.2021.00788.

[31] Chiu, Tai-Yin & Gurari, Danna. (2020). Iterative Feature Transformation for Fast and Versatile Universal Style Transfer. 10.1007/978-3-030-58529-7_11.

[32] Liu, Songhua & Lin, Tianwei & He, Dongliang & Li, Fu & Wang, Meiling & Li, Xin & Sun, Zhengxing & Li, Qian & Ding, Errui. (2021). AdaAttN: Revisit Attention Mechanism in Arbitrary Neural Style Transfer.

[33] Ji, Wei et al. “Semantic Locality-Aware Deformable Network for Clothing Segmentation.” IJCAI (2018).

[34] Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang. Deepfashion: Powering robust clothes recognition and retrieval with rich annotations. In CVPR, 2016

[35] Ge, Yuying & Zhang, Ruimao & Lingyun, Wu & Wang, Xiaogang & Tang, Xiaoou & Luo,
Ping. (2019). DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images.

[36] HUANG, QIUBO & HAN, XIAOPENG & LU, TING & LIU, GUOHUA. (2021). Clothing Image Retrieval Based on Parts Detection and Segmentation. 53-59. 10.1145/3469951.3469961.

[37] Castro, Hassler & Ramirez, Mariana. (2020). Segmentation task for fashion and apparel.

[38] J. Martinsson and O. Mogren, "Semantic Segmentation of Fashion Images Using Feature Pyramid Networks," 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), 2019, pp. 3133-3136, doi: 10.1109/ICCVW.2019.00382.

[39] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In European Conference on Computer Vision, 2016.

[40] Li, X., Liu, S., Kautz, J., Yang, M.H.: Learning linear transformations for fast image and video style transfer. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3809–3817 (2019).

[41] Yingying Deng, Fan Tang, Weiming Dong, Haibin Huang, Chongyang Ma, and Changsheng Xu. Arbitrary video style transfer via multi-channel correlation. arXiv preprint arXiv:2009.08003, 2020.