Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 5, October 2024, pp. 3418~3429
ISSN: 2302-9285, DOI: 10.11591/eei.v13i5.7659

Novel entropy-based style transfer of the object in the content image using deep learning

Jyoti Sudhakar Raghatwan[1], Sandhya Arora[2]
[1]Department of Computer Engineering, Smt Kashibai Navale College of Engineering, Savitribai Pule Pune University, Pune, India
[2]Department of Computer Engineering, MKSSS's Cummins College of Engineering, Savitribai Phule Pune University, Pune, India

Article Info

Article history:

Received Oct 3, 2023
Revised Feb 28, 2024
Accepted Mar 20, 2024

Keywords:

Content image
High pass filter
Low pass filter
Style image
VGGNet19
Wavelet transform

ABSTRACT

Recently neural style transfer (NST) has drawn a lot of interest of researchers, with notable advancements in color representation, texture, speed, and image quality. While previous studies focused on transferring artistic style across entire content images, a new approach proposes to transfer style specifically to objects within the content image based on the style image and maintain photorealism. Recent techniques have produced intriguing creative effects, but often only work with artificial effects, leaving real flaws visible in photographs used as references for styles. The suggested approach employs a two-dimensional wavelet transform (WT) to achieve style transfer by adjusting image structure with high-pass and low pass filters (LPF). Preserving the information content and numerical attributes of VGGNet19 through WT-based style transfer using the db5 WT at level 5, we can achieve a peak signal-to-noise ratio (PSNR) value of up to 96.76725. The qualitative result of the proposed methodology is compared with other existing algorithm. Also, the time complexity of the proposed methodology on different hardware platforms has been calculated and presented in the paper. The proposed methodology able to maintains appealing and precise quality of resultant image.

Corresponding Author:

Jyoti Sudhakar Raghatwan
Department of Computer Engineering, Smt Kashibai Navale College of Engineering
Savitribai Phule Pune University
Pune, India
Email: jyotiraghatwan2@gmail.com

1. INTRODUCTION

Style is an abstract description of artistic characteristics used in the development of art. The same things can be shown in a variety of styles, each with its own set of connotations based on context or culture. Imitation or re-representation of many styles is a time-consuming and challenging procedure that necessitates a high level of artistic ability and technical proficiency. With the help of computer technology, the difficulty of this task can be lowered and excellent outcomes produced.

Style transfer is the technique of extracting an image's content components and re-presenting them in a different extracted style to create a redesign impact. Computer-assisted style transfer can be traced back to picture texture synthesis technology that was in use before 2000. To recognize and develop textures, researchers used complicated mathematical models and formulae. Manual modeling, on the other hand, is time-consuming and labor-intensive, and computational power at the time was rather restricted. It can be effectively applied and relatively optimal results are obtained with obvious enhancements in properties like texture, color, and structure by processing high-level abstract features. Picture features are merged in line with human visual habits, resulting in outstanding versatility and convenience of use, as well as the elimination of the need to repeat complex mathematical processes.

Image processing involves the immense utilization of wavelet transform (WT), and to apply on pictures requires the knowledge of its application in two dimensions. WTs are widely used in image processing, and their application on pictures necessitates an understanding of their two-dimensional application.

The noise is successfully eliminated by the conventional filter. However, the picture will get the noise is successfully eliminated by the conventional filter. However, the picture will get blurry. As a result, when reducing picture noise, we should protect the image's edge image. Wavelet analysis is a time-frequency analysis method that modifies frequency band selection based on signal attributes. The spectrum is then matched to the frequency band, which improves time-frequency resolution. Wavelets have the advantage of allowing simultaneous localization in the time and frequency domain. The second major benefit of wavelets is that they are computationally very quick when utilizing the fast WT. The ability to discern small details in a signal is a huge advantage of wavelets.

Style transfer is a computer vision approach that combines two pictures-a content picture and a style reference picture-so that the output picture keeps the basic components of the content picture while appearing to be "painted" in the style reference picture's style.

By using different techniques for the style transfer there was not only an accuracy issue but also there was the issue of the system getting lots of time to execute. The clarity of the picture is not very well by using different techniques. To overcome all the problems, we proposed our system to get good signal-to-noise ratio (SNR). We can prove that our system gives better performance as well as it gives better SNR value by using VGGNet along with WT. Figure 1 shows the resultant style transferred image where Figure 1(a) represent content image, Figure 1(b) represent style image and Figure 1(c) is resultant style transfer image. Proposed approaches can have numerous applications in multiple domains like fashion design, architectural design, entertainment to visualize the objects in many styles in minimal time.

(a) (b) (c)

Figure 1. Style transfer of the content image according to the style image; (a) content image, (b) style image, and (c) style transferred

The organisation of rest of the paper is as follows. Section 2 focuses on literature survey. Proposed method is discussed in section 3. Experimental results and analysis are provided in section 4. Finally, the conclusion and future research directives given in section 5.

2. LITERATURE SURVEY

The important deep learning-based picture style transfer methodologies, such as picture-based iteration and model-based iteration, are covered in this section. The first includes directly optimizing and iterating on a noise image, whereas the second involves iterative noise image optimization. A neural network model can achieve rapid style transfer by using network feed-forward. Following that, the necessary technologies are thoroughly investigated and addressed.

The rapid style transfer framework proposed by Ye *et al*. [1] has looked at the ability of existing ways to preserve single semantic information and proposes a rapid style transfer framework that includes multi-semantic preservation. Experiments reveal that the author's method effectively keeps the original semantic information, such as salience and depth characteristics, resulting in a finished artwork with a visual impression that is improved by emphasizing its geographic focus and depth information. Their method is more capable in terms of semantic preservation and can produce more artwork with distinct sections when compared to other ways. Controllable semantics, a diverse set of data, and a diverse set of emotions.

Wang *et al.* [2] work on the two distinct approaches. The authors divide the degree of style transmission into three categories: content fidelity (CF), global effects (GE), and local patterns (LP). Then, two distinct approaches of using these characteristics to improve stylization quality are presented. The first, dubbed cascade style transfer (CST), uses criteria to direct the use of established neural style transfer (NST) systems in a cascade to absorb their benefits while avoiding their drawbacks. The second, known as the multi-objective network (MO-Net), directly optimises various aspects to balance their performance and produce more harmonic stylistic results.

The authors propose a trying-to-draw interactive style transfer (IST) technique in this paper. This system introduced a consistency loss in new designs. NST [3], [4] can create by adapting reference style to the content picture, exceptional artworks. Current picture-to-picture conversion fine-grained controls are often required for artistic editing, but NST approaches lack these. Zheng *et al.* [5] proposes an IST method to address this problem, which allows users to make a picture that is harmonious while drawing. The style-coherence loss network (SCNet) is introduced, which merges content loss, style loss, and consistency loss in innovative concepts. Their IST approach can be used as a brush to apply a style from any the SCNet, it has introduced a newly designed loss of consistency that includes content loss, style loss, and the newly developed loss of consistency. These losses are estimated using high-level properties extracted from a pre-trained network visual geometry group 16 (VGG-16).

The system proposed modeling a photo’s content as feature reactions from a were before convolutional neural networks (CNN). The style losses modify the style picture by adding new colors and patterns, whereas the content losses keep the high-level abstract data in the input frames. Rather than employing optical flow to directly correct the styled video frames, the author uses coherence loss to make the stylized video inherit the source video’s dynamics and motion patterns, which avoids temporal flickering. Paint on any section of the target content picture that you select. Based on similarity maps, they build a fluid simulation method that uses styles as colors to identify the action scope of the brush interface and spread in style or content pictures. Jing *et al.* [6] goes on to show a range of evaluation methods, including qualitative and quantitative assessments of numerous NST algorithms.

Because the self-supervised semantic network (SSNet) and the style transfer network exchange parameters, the improved alternatives can keep the semantic wholeness of the stylized picture. To address the content missing and blur challenges that are typical in NST, Cui *et al.* [7] presents the bidirectional convolution block (B-block) and feature fusion approach (F-block) for visual smoothness. The method’s efficacy is demonstrated by ablation studies, and the method’s benefits are demonstrated through comparing trials with government baselines.

A de-stylized network is used to recreate the stego picture from the stylized picture. Finally, using an extraction technique, the sent secret picture is recovered from the rebuilt stego picture. The basic goal of the steganography approach is to ensure that the eavesdropper perceives independent and consistent behavior even when the carrier picture’s look and distribution are modified. As indicated in [8] extensive testing is carried out to confirm that the suggested method is practical.

The authors developed a new approach to online video style transfer that avoids the use of video frames in both training and testing. The SCNet, which integrates content, style, and newly created coherence losses, is introduced. These losses are calculated using high-level properties taken from a VGG-16 network that has already been trained. The style loss alters the style picture by adding new colors and patterns, whereas the content loss preserves the input frames high-level abstract data. Rather than using optical flow to directly correct the styled video frames, as Xu *et al.* [9] suggested, they use coherence loss to make the stylized video inherit the source video’s dynamics and motion patterns, which reduces temporal skipping. Any reference style photo can be used using arbitrary style approaches [4], [10]–[13]. Using a pre-trained model, the online optimization methods decouple the content and stylistic information of input photographs. The new artistic picture imitates the style picture in the feature gram matrix while keeping familiarity with the content picture in feature representation thanks to iterative optimization. Online optimization takes time and requires a lot of calculations. Other methods sought to communicate any style with a single forward pass. A common method [9], [13] in the instance feature space matches the statistics between the new artwork and the supplied style picture. Some studies [14] have learned to anticipate dynamic factors for any style. Although such methods provide flexibility, the processing granularity is still a complete picture, which means they cannot suit the needs of all users.

The author introduces class-based styling (CBS). That really can undertake true fashion transmission to specific object categories. There are many techniques in those systems. More emphasis is being placed on passing knowledge to local communities. However, it is usually dealt with using a mask, such as a semantic mask [15] or an instance mask [16]. Many methods [17], including brush area [18], and object mask employing an interactive segmentation method rely on user-created binary masks. Additional procedures include locally retouching for style refinement [19] and altering the semantic mask to generate a new texture picture.

NST technique is used in this system. NST is a common artistic production technique that attempts to transfer the artistic style of a reference picture to a content picture. Other NST approaches have been proposed, including content retention [10], [20] robustness analysis [13], [16] the ceiling of style types geometric modifications, brushstroke simulation [21], and other NST methods. Several recent researches have introduced the user's goal, allowing manual stylization, semantic assistance [15], and disguised style splicing [16] changes to NST.

Style-corpus constrained learning (SCCL) is used in this system that is proposes a new method called SCCL [21]. To restrict the stylized picture with style consistency among multiple samples, a style-corpus with both style-specific and style-agnostic properties is proposed. As a result, the photorealism of the stylization output is improved. A basic fast-to-execute network is trained to replace earlier complicated feature transformation models using an adversarial distillation learning technique, significantly decreasing the computational cost.

Style transfer can be thought of as recreating or summarising texture based on the precise meaning of the target object. Most existing techniques, such as content style interpolation, color-preserving, spatial control, and brush stroke size control, are compatible with the proposed multi-scale guided convolutional network (MSG-Net) as a universal framework for real-time style transfer. MSG-Net is the first style transfer system that provides real-time, comprehensive feed-forward brush size management. Torch, PyTorch, and MXNet framework implementations and pre-trained models by Zhang and Dana [22] will be publicly available.

Photorealistic style transfer must achieve two opposing goals proposed in his system [23]. Yoo *et al.* [23] introduces a whitening and coloring transforms (WCT2) that preserves information about the structure and VGG feature field features are extracted during the visual style and are based on lightening and coloring changes. This approach produces a constant, time-independent video stylization. Moreover, the classification algorithms used in the biomedical field also can be used in the segmentation or classification of objects from the content images [24]–[30]. The majority of relevant research papers on NST were expressly recognized as focusing on real-time NST, narrowing the scope of video style transfer research. To allow the models to be trained, conditional GANs (cGANs) add picture-to-picture translation and a loss function. Custom loss functions or mapping functions are no longer required [31]–[33]. A recent study demonstrated the capacity of CNN to create visually amazing visuals by extracting and recombining image content and style. The process of using CNN to relocate the semantic information of a single image to several styles is known as NST. NST has become a hot topic in academic literature and industry applications since then.

From the literature, it is clear that there is a need to design a style transfer algorithm that equalizes the images such as low and high variations in the style and content image should get transferred with coordination. The proposed approach consists of dividing content and style images into two parts. Then the low variation part of the style image overlapped with the low variation part of the content image. Similarly, the high variation part of the style image overlapped with the high variation part of the content image. The result of the style transferred images is elaborated in the result session followed by the proposed system architecture. The results are compared based on SNR, peak signal-to-noise ratio (PSNR), time complexity, and space complexity in the results and discussion session.

3. METHOD

The issues that have been observed in the conventional style transfer algorithms are as: i) style not only gets transferred to objects in the content image but also to the background of the content image and ii) style gets transferred to the content image without considering the texture of the image. The solution of these two issues has been solved by using the proposed architecture. the novelty of the architecture is the introduction of the WT to preserve structural information of features and masks to get style transfer results.

By breaking down an image into its different frequency components, the WT makes it possible to distinguish between high-frequency (detail) and low-frequency (smooth) information. The technique known as wavelet decomposition is used to do this. Low pass filters (LPF) in the WT context refers to keeping the image's low-frequency elements while removing its high-frequency details. The goal is to retain the image's smoother regions and overall structure, which are frequently linked to low-frequency data. High-pass filtering, on the other hand, aims to keep the high-frequency information in an image while removing the low-frequency elements. Tasks that call for attention to edges, delicate textures, or other high-frequency elements will benefit most from this.

The use of WT in the style transfer along with masking of the object from the content image is the main contribution in this field. In this method, there are two images those are style image and the content image. Firstly, WT of level 5 and db5 is applied on the content image and style image for getting low entropy and high entropy patches of the image as shown in Figure 2.

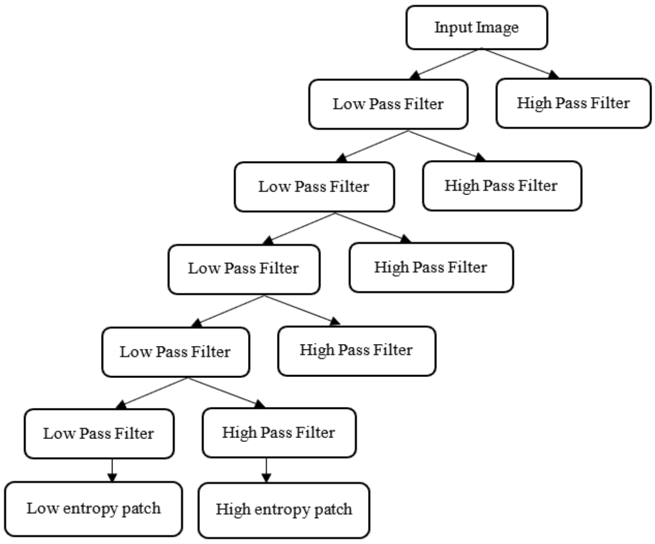


Figure 2. Low entropy patch and high entropy patch of the input image

Then by using VGGNet the low entropy patch of the content image is got style transferred using low entropy patch of the style image. Similarly, the high entropy patch of the content image is style transferred using high entropy patch of the style image. The result of style transferred output of both low entropy patch and high entropy patch is then combined. To get the result of style transferred on only object in the image the mask of the object in the content image is applied to get the final style transferred image. The detail architecture is shown in Figure 3.

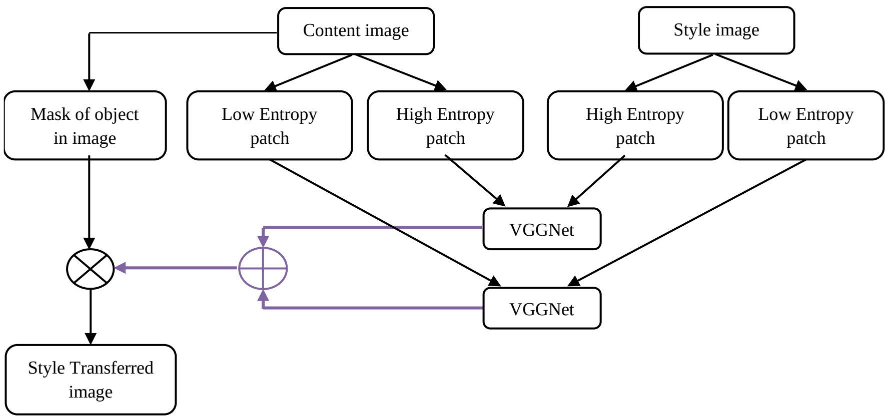


Figure 3. Proposed system architecture

We extract the photos low- and high-entropy content. Level 5 and db5 are used for the decomposition of the image. As shown in (1), the WT is applied to the $image(a,b)$ to produce $WT_{\psi}^{i}(j,m,n)$.

$$WT_{\varphi}(j_0,m,n)=\frac{1}{\sqrt{MN}}\sum_{x=0}^{M-1N-1}.\,image(a,b)\varphi_{j=0}(a,b) \quad (1)$$

$$WT_{\psi}^{i}(j,m,n)=\frac{1}{\sqrt{MN}}\sum_{a=0}^{M=1}.\sum_{b=0}^{N-1}.\,image(a,b)\psi^{i}j_{j,m,n}(a,b)\psi i=\{H,V,D\} \quad (2)$$

Here, j_0 is an arbitrary starting scale $WT_{\varphi}(j_0,m,n)$ coefficients define an approximation of $image(a,b)$ at scale $j_0 WT_{\psi}^{i}(j,m,n)$ parameters for the scales $j \geq j_0$ Given the WT_{φ} and WT_{ψ}^{i}, $image(a,b)$.

$$image(a,b) = \frac{1}{\sqrt{MN}} \sum_{m} \sum_{n} WT_{\varphi}(j_0,m,n)\varphi_{j_0,m,n}(a,b) + \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{j=j_0}^{\infty} \sum_{m} \sum_{n} WT_{\psi}^{i}(j,m,n)\psi_{j,m,n}^{i}(a,b) \quad (3)$$

Texture stochastic gradient descent training of generative adversarial nets.
The number of steps to apply to the discriminator, h, is a hyperparameter.
Low pass filter (LPF)=level 5, db5, less entropy components from texture and content image.
High pass filter (HPF)=level 5, db5, high entropy components from texture and content image.
Here h is used as the depth of the style transfer experiments.
for number of training iterations do
for h steps do

- Sample texture of n LPF samples $\{z^{(1)}, \dots, z^{(n)}\}$ from noise prior $p_g(z)$.
- Sample texture of m HPF $\{x^{(1)}, \dots, x^{(n)}\}$ from data generating distribution $p_{data}(x)$.
- Update the discriminator by ascending its stochastic gradient:

$$\nabla_{\theta_d} \frac{1}{n} \sum_{i=1}^{n} \left[log\ De\left(x^{(i)}\right) + log\ \left(1 - De\left(Gr\left(z^{(i)}\right)\right)\right)\right] \quad (4)$$

end for

- Sample texture of n LPF samples $\{z^{(1)}, \dots, z^{(n)}\}$ from content prior $p_g(z)$.
- Update the generator by descending its stochastic gradient:

$$\nabla_{\theta_g} \frac{1}{n} \sum_{i=1}^{n} log\ \left(1 - De\left(Gr\left(z^{(i)}\right)\right)\right) \quad (5)$$

end for
The gradient-based updates can use any standard gradient-based learning rule.

Once the texture image and content image are passed through this mathematical model the resultant image is the style transferred image. But the complete content image is got style transferred. The main aim of the system is to transfer the style of object only. To do so, before doing style transferring part, mask of the object present in content image is generated. This mask is then super imposed on the style transferred image. By doing this only object in the content image is get style transferred. The result of this is well illustrated in next session.

4. RESULT AND DISCUSSION

WT is used to separate out the smooth variation and drastic variation in the content image and style image. For that the various combinations of the WT have been tried. Here, in the application, the quality of the style-transformed image plays an important role. To get the style transformed image correctly, the LPF part and HPF part of the both content image and style image must be separated properly. As it is very difficult to see if this separation is successfully done or not by using naked eyes it is always better to use the SNR and PSNR values to check it. SNR and PSNR are metrics commonly used in image processing to quantify the quality of a processed or transformed image. However, when it comes to the selection of a perfect WT, these metrics play an important role [34]. The SNR and PSNR of the resultant image are tabulated in Table 1. It shows SNR and PSNR value of the style transferred image for WT of level 2 to level 7 for different values of db1 to db5. From the graph presented in Figure 4, it is clear that WT of db5 and level 5 gives the maximum SNR and PSNR value. This is the reason behind selecting the WT of db5 and level 5 is selected in our research work.

Table 1. SNR and PSNR value of the style transferred image for WT of levels for different values of db

WT (level)	SNR (db1)	PSNR (db1)	SNR (db3)	PSNR (db3)	SNR (db5)	PSNR (db5)	SNR (db7)	PSNR (db7)
2	82.07647	77.97265	84.12774	75.71497	84.53888	80.31194	84.83858	80.59665
3	86.06979	85.20909	86.93953	82.59256	87.80928	80.78454	88.67902	84.24507
4	89.54877	85.07133	90.41852	83.18504	91.28826	78.50791	92.15801	87.55011
5	93.02776	88.37637	93.8975	89.20263	96.76725	91.92889	92.15801	87.55011
6	89.54877	85.07133	88.67902	79.81112	86.93953	82.59256	84.12774	79.92136
7	81.3234	77.25723	80.93483	71.22265	80.54626	76.51895	80.15769	71.34034

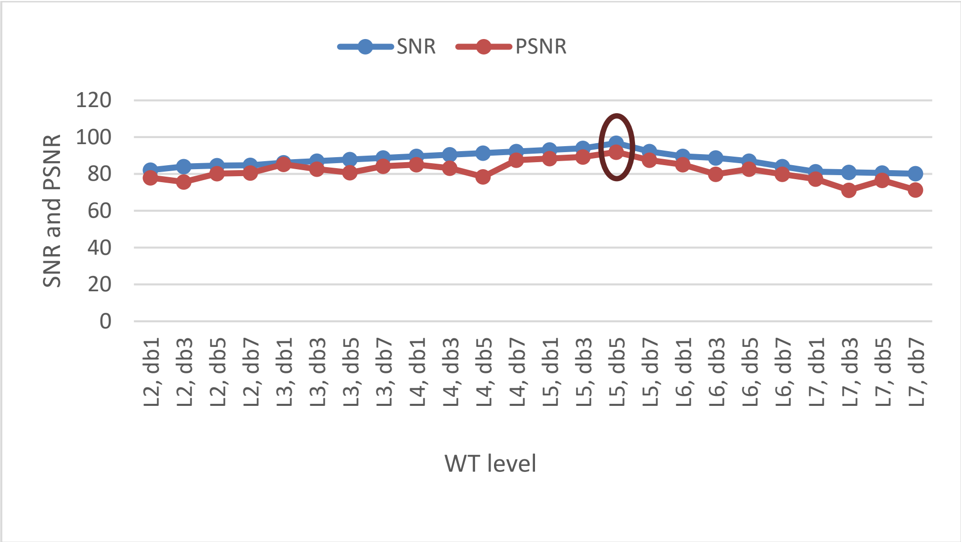


Figure 4. Value of SNR and PSNR of the style transferred image for different for different values of db

The step-by-step transformation of content image and style image is shown Figure 5. Here high entropy patch and low entropy patch of the content image and style image is extracted using (1) to (5). Then VGGNet is applied on low entropy patch of the style image and content image. Parallelly, VGGNet is applied on high entropy patch of the style image and content image. The result of both VGGNet is then combined to get the final style transformed result. Figure 6. Shows the qualitative result of proposed system. Content image and style image is presented in Figures 6(a) and (b) and resultant image is presented in Figure 6(c).

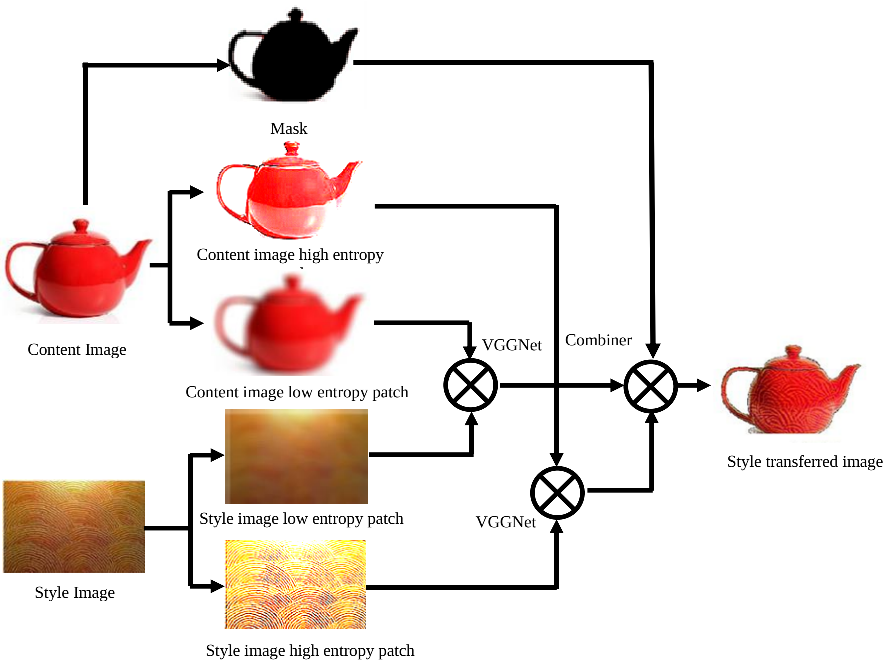
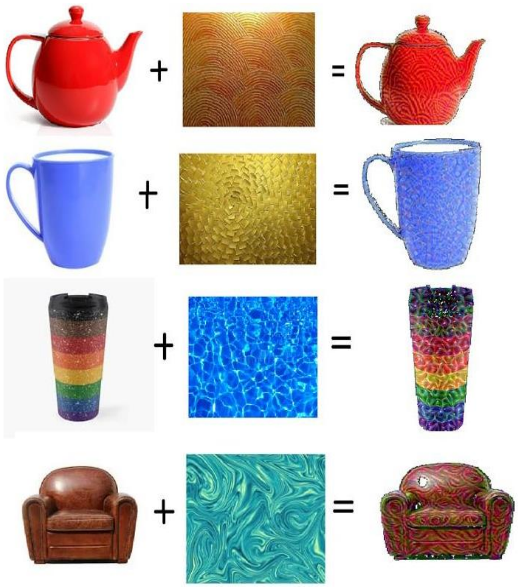


Figure 5. Steps of style transformation of proposed system architecture



(a) (b) (c)

Figure 6. Results of proposed system: (a) content image, (b) style image, and (c) style transferred image

The low entropy patch and high entropy patch get extracted from these two images by using (1), (2), and (3). Parallelly mask of the object in content image get extracted. The high entropy patch and low entropy patch of both style and content image is passed separately through the VGGNet as shown in Figure 3. Then the result of these two VGGNets are get combined pixel wise to get the style transferred image. Extracted mask is now applied on resultant image to get the transfer of style on the object in the content image only, which are presented in Figure 6(c).

Figure 7 illustrates the novelty in the proposed methodology. Figure 7(a) is a content image as first input and Figure 7(b) is style image as second input. The style transferred image using Artsy-GAN [33] transfers the style of the complete image including background Figure 7(c). The proposed methodology transfers the style of the object only Figure 7(d). Moreover, the styled transferred image using the proposed methodology gives results ae per the texture of both style image and content image.

(a) (b) (c) (d)

Figure 7. Style transfer of the content image according to the style image; (a) content image, (b) style image, (c) style transferred image as an output using artsy-GAN [33], and (d) style transferred image as an output using proposed methodology

Structural similarity index measure (SSIM) of style transferred image with content image and style image is calculated for different architectures and tabulated in Table 2. Figure 8 clearly illustrates that the value of SSIM is higher in case of proposed system architecture than any other system architecture.

Table 2. SSIM of style transferred image of different architectures with content image and style image

Architecture	SSIM of style transferred image with content image (%)	SSIM of style transferred image with style image (%)
GAN [31]	43	51
CycleGAN [32]	69	49
Artsy-GAN [33]	72	58
Propose system architecture (PSA)	83	74

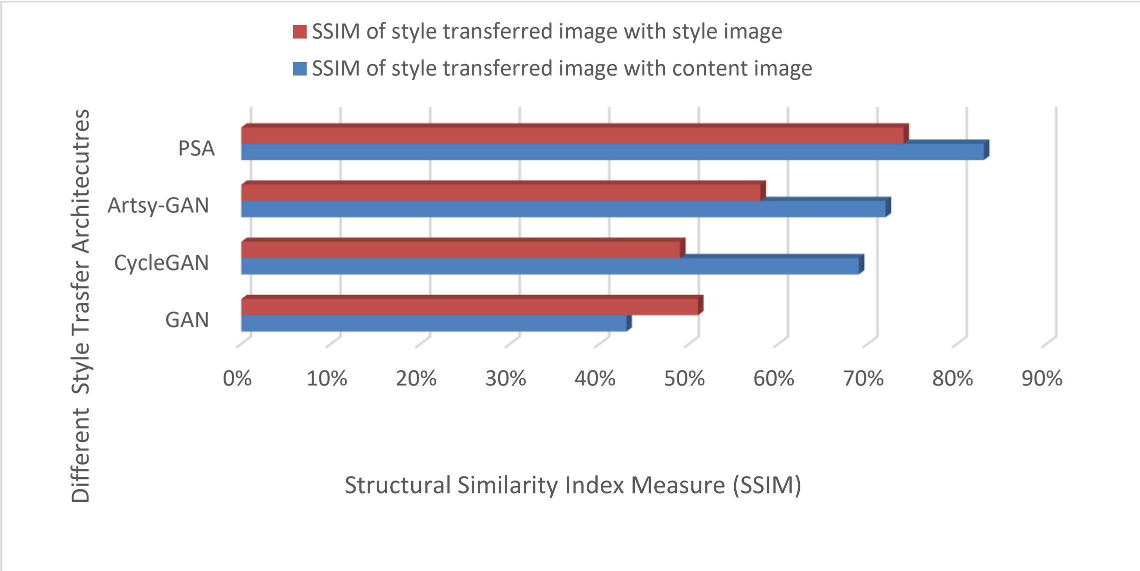


Figure 8. comparison of different architectures on the basis of SSIM

The time and space complexity of the proposed system architecture is calculated are presented in Table 3. The hardware used for the experimentation is central processing unit (CPU) with i5 processor and 8 GB RAM. The relationships between various image sizes and the processing time and memory needed are graphically depicted in Figures 9 and 10.

Table 3. Time and space complexity of the proposed system architecture

Image size	Time (seconds)	Required memory for processing (kb)
50 kb	5.241	115
100 kb	7.2349	230
200 kb	12.234	460
500 kb	17.23043	1150
750 kb	24.03819	1725
1 Mb	29.4829	2300
5 Mb	34.93201	11500
10 Mb	39.91831	23000
15 Mb	54.1239	34500

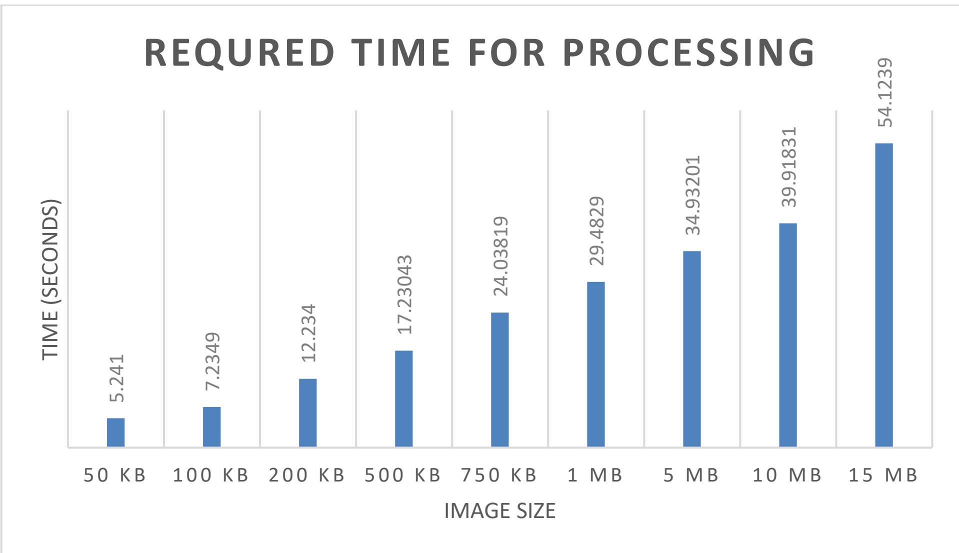


Figure 9. Time complexity of the proposed system architecture with respect to different size of content image

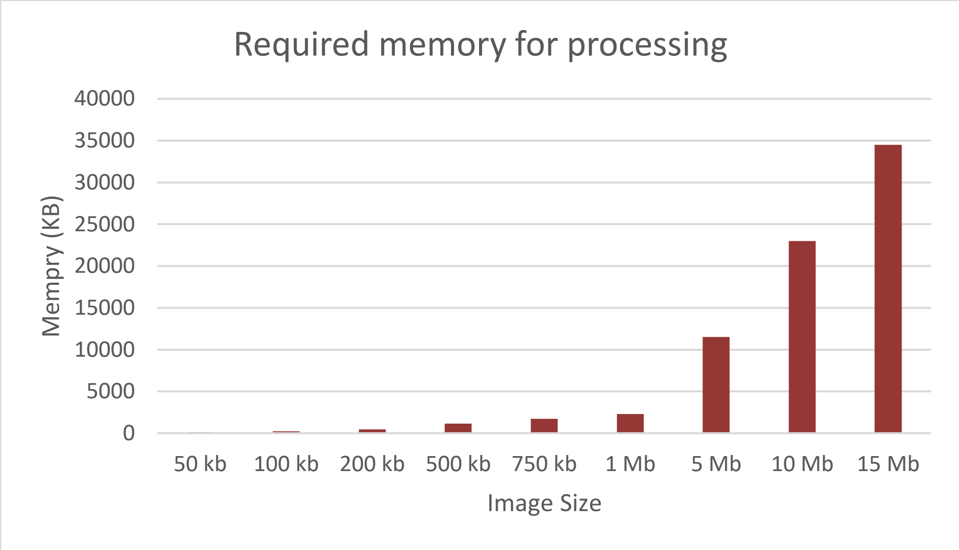


Figure 10. Space complexity of the proposed system architecture with respect to different size of content image

Figures 9 and 10 shows that, both processing and memory usage increases with image size. But it is moderately optimum with large image size. Also, the time complexity is tested on different hardware as tabulated in Table 4. CPU with i3 processor, i5 processor, i7 processor, and Nvidia graphics processing unit (GPU) processor is used to get time complexity.

From Figure 11 it is clear that there is no drastic change in the time when we are using CPU for getting result. So, we can say that the system takes almost same time to get result on any CPU based hardware platform. But if the time is the crucial parameter to get result, then it is always better to use GPU as it takes very less time to get result.

Table 4. Time required to get the style transferred result of size 1 Mb

Hardware platform	Time required to get result (in seconds)
CPU, i3 processor, 8GB RAM	36.2341
CPU, i5 processor, 8GB RAM	29.4829
CPU, i7 processor, 8GB RAM	27.3419
GPU, Nvidia K80	5.234123

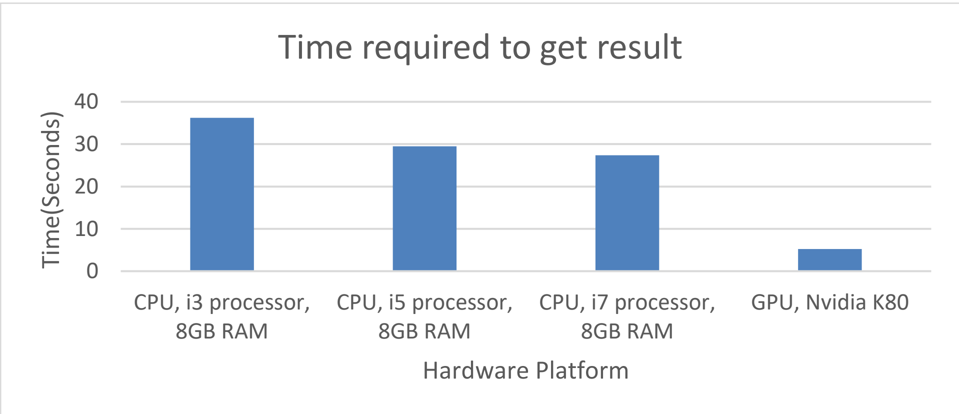


Figure 11. Time required for getting style transferred image on different hardware platform

5. CONCLUSION

In this study, a novel entropy-based technique is suggested. The image's smooth and dramatic variation patch is obtained using the WT, and it is then sent through the VGGNet. The paper presents the outcomes of several WTs as well as the rationale for the sort of WT that was selected. Along with time and space complexity, the SNR and PSNR of the suggested system architecture are also compared. Thanks to the precise recovery of the WT, our method can give stable stylization without any restrictions while maintaining structural information. The stylized output conveys information about the style image while maintaining the overall structure of the content image. To fully extract the information from the style image, this process is performed numerous times by inverting the features provided by the style picture VGG-19 at different depths. In the suggested technique, the style image and the content image are combined using VGG-19. This system is efficient and faster. The further study can focus on giving text information as an input in addition to the content image and style image to the style transfer algorithm unlike traditional image style transfer, where the style is derived from another image, incorporating text introduces the additional complexity of interpreting and applying stylistic elements described in the text. In this case, additional text input with natural language processing (NLP) has to give as an input to the neural network.

REFERENCES

[1] W. Ye, X. Zhu, Z. Xu, Y. Liu, and C. C. Chang, "A comprehensive framework of multiple semantics preservation in neural style transfer," *Journal of Visual Communication and Image Representation*, vol. 82, p. 103378, Jan. 2022, doi: 10.1016/j.jvcir.2021.103378.

[2] Z. Wang *et al.*, "Evaluate and improve the quality of neural style transfer," *Computer Vision and Image Understanding*, vol. 207, p. 103203, Jun. 2021, doi: 10.1016/j.cviu.2021.103203.

[3] M. M. Cheng, X. C. Liu, J. Wang, S. P. Lu, Y. K. Lai, and P. L. Rosin, "Structure-preserving neural style transfer," *IEEE Transactions on Image Processing*, vol. 29, pp. 909–920, 2020, doi: 10.1109/TIP.2019.2936746.

[4] T. Lin *et al.*, "Drafting and revision: laplacian pyramid network for fast high-quality artistic style transfer," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 5137–5146, 2021, doi: 10.1109/CVPR46437.2021.00510.

[5] M. C. L. Zheng, Z. Zhang, K.-R. Zhang, and B. Ren, "Interactive style transfer: all is your palette," arXiv (Cornell University), Mar. 2022, doi: 10.48550/arxiv.2203.13470.

[6] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural style transfer: a review," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 11, pp. 3365–3385, Nov. 2020, doi: 10.1109/TVCG.2019.2921336.

[7] J. Cui *et al.*, "PortraitNET: Photo-realistic portrait cartoon style transfer with self-supervised semantic supervision," *Neurocomputing*, vol. 465, pp. 114–127, Nov. 2021, doi: 10.1016/j.neucom.2021.08.088.

[8] Q. Li *et al.*, "Image steganography based on style transfer and quaternion exponent moments," *Applied Soft Computing*, vol. 110, p. 107618, Oct. 2021, doi: 10.1016/j.asoc.2021.107618.

[9] K. Xu, L. Wen, G. Li, H. Qi, L. Bo, and Q. Huang, "Learning self-supervised space-time CNN for fast video style transfer," *IEEE Transactions on Image Processing*, vol. 30, pp. 2501–2512, 2021, doi: 10.1109/TIP.2021.3052709.

[10] J. Cheng, A. Jaiswal, Y. Wu, P. Natarajan, and P. Natarajan, "Style-aware normalized loss for improving arbitrary style transfer," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 134–143, 2021, doi: 10.1109/CVPR46437.2021.00020.

[11] D. Y. Park and K. H. Lee, "Arbitrary style transfer with style-attentional networks," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 5873–5881, 2019, doi: 10.1109/CVPR.2019.00603.

[12] L. Sheng, Z. Lin, J. Shao, and X. Wang, "Avatar-Net: multi-scale zero-shot style transfer by feature decoration," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 8242–8250, 2018, doi: 10.1109/CVPR.2018.00860.

[13] P. Wang, Y. Li, and N. Vasconcelos, "Rethinking and Improving the Robustness of Image Style Transfer," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 124–133, 2021, doi: 10.1109/CVPR46437.2021.00019.

[14] F. Shen, S. Yan, and G. Zeng, "Neural style transfer via meta networks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 8061–8069, doi: 10.1109/CVPR.2018.00841.

[15] L. Kurzman, D. Vazquez, and I. Laradji, "Class-based styling: Real-time localized style transfer with semantic segmentation," in *Proceedings-2019 International Conference on Computer Vision Workshop*, IEEE, Oct. 2019, pp. 3189–3192, doi: 10.1109/ICCVW.2019.00396.

[16] F. Stahl, M. Meyer, and U. Schwanecke, "IST-style transfer with instance segmentation," in *International Symposium on Image and Signal Processing and Analysis*, IEEE, Sep. 2019, pp. 277–281, doi: 10.1109/ISPA.2019.8868438.

[17] L. N. Alegre and M. M. Oliveira, "SelfieArt: interactive multi-style transfer for selfies and videos with soft transitions," in *Proceedings-2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images*, IEEE, Nov. 2020, pp. 17–22, doi: 10.1109/SIBGRAPI51738.2020.00011.

[18] M. Dubey, J. Otto, and A. G. Forbes, "Data brushes: Interactive style transfer for data art," in *Proceedings of the IEEE VIS Arts Program*, IEEE, Oct. 2019, pp. 1–9, doi: 10.1109/VISAP.2019.8900858.

[19] M. Reimann, B. Buchheim, A. Semmo, J. Dollner, and M. Trapp, "Interactive multi-level stroke control for neural style transfer," *Proceedings-2021 International Conference on Cyberworlds*, pp. 1–8, 2021, doi: 10.1109/CW52790.2021.00009.

[20] D. Kotovenko, A. Sanakoyeu, P. Ma, S. Lang, and B. Ommer, "A content transformation block for image style transfer," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2019-June, pp. 10024–10033, 2019, doi: 10.1109/CVPR.2019.01027.

[21] D. Kotovenko, M. Wright, A. Heimbrecht, and B. Ommer, "Rethinking style transfer: From pixels to parameterized brushstrokes," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2021, pp. 12191–12200, doi: 10.1109/CVPR46437.2021.01202.

[22] H. Zhang and K. Dana, "Multi-style generative network for real-time transfer," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, pp. 349–365, doi: 10.1007/978-3-030-11018-5_32.

[23] J. Yoo, Y. Uh, S. Chun, B. Kang, and J. W. Ha, "Photorealistic style transfer via wavelet transforms," in *Proceedings of the IEEE*

International Conference on Computer Vision, IEEE, Oct. 2019, pp. 9035–9044, doi: 10.1109/ICCV.2019.00913.
[24] A. S. Ladkat, A. A. Date, and S. S. Inamdar, "Development and comparison of serial and parallel image processing algorithms," in *Proceedings of the International Conference on Inventive Computation Technologies*, IEEE, Aug. 2016, pp. 1–4, doi: 10.1109/INVENTIVE.2016.7824894.
[25] A. S. Ladkat *et al.*, "Deep neural network-based novel mathematical model for 3D brain tumor segmentation," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–8, Aug. 2022, doi: 10.1155/2022/4271711.
[26] M. Shobana *et al.*, "Classification and detection of mesothelioma cancer using feature selection-enabled machine learning technique," *BioMed Research International*, vol. 2022, pp. 1–6, Jul. 2022, doi: 10.1155/2022/9900668.
[27] S. G. Munde, A. S. Ladkat, and R. Patil, "Zero blackout avoidance keeping emergency services at priority using machine learning," in *2023 5th International Conference on Electrical, Computer and Communication Technologies*, IEEE, Feb. 2023, pp. 1–5, doi: 10.1109/ICECCT56650.2023.10179676.
[28] G. Kaur and N. Chaudhary, "Brain tumor detection using machine learning hybrid approach," *2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering*, pp. 1076–1081, 2023, doi: 10.1109/UPCON59197.2023.10434384.
[29] S. L. Bangare, G. Pradeepini, and S. T. Patil, "Neuroendoscopy adapter module development for better brain tumor image visualization," *International Journal of Electrical and Computer Engineering*, vol. 7, no. 6, pp. 3643–3654, 2017, doi: 10.11591/ijece.v7i6.pp3643-3654.
[30] J. Surve, D. Umrao, M. Madhavi, T. S. Rajeswari, S. L. Bangare, and M. K. Chakravarthi, "Machine learning applications for protecting the information of health care department using smart internet of things appliances -a review," in *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering*, IEEE, Apr. 2022, pp. 893–898, doi: 10.1109/ICACITE53722.2022.9823642.
[31] S. Azadi, M. Fisher, V. Kim, Z. Wang, E. Shechtman, and T. Darrell, "Multi-content GAN for few-shot font style transfer," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2018, pp. 7564–7573, doi: 10.1109/CVPR.2018.00789.
[32] E. Batziou, K. Ioannidis, I. Patras, S. Vrochidis, and I. Kompatsiaris, "Artistic neural style transfer using CycleGAN and FABEMD by adaptive information selection," *Pattern Recognition Letters*, vol. 165, pp. 55–62, Jan. 2023, doi: 10.1016/j.patrec.2022.11.026.
[33] H. Liu, P. N. Michelini, and D. Zhu, "Artsy-GAN: A style transfer system with improved quality, diversity and performance," in *Proceedings-International Conference on Pattern Recognition*, IEEE, Aug. 2018, pp. 79–84, doi: 10.1109/ICPR.2018.8546172.
[34] H. Du *et al.*, "Disentangling noise patterns from seismic images: noise reduction and style transfer," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022, doi: 10.1109/TGRS.2022.3219117.

BIOGRAPHIES OF AUTHORS

Jyoti Sudhakar Raghatwan has complete M.E. (Information Technology) from Pune University. Currently she is research scholar at SKCOE, Pune. Her research interests include image processing, machine learning, and computer vision. She can be contacted at email: jyotiraghatwan2@gmail.com.

Dr. Sandhya Arora has completed Ph.D. (Computer Science & Engineering) from Jadavpur University, Kolkata. She is working as Professor, in Department of Computer Engineering, MKSSS's Cummins College of Engineering for Women, Pune. She has rich teaching experience of 26 years. Her research interests include soft computing, online learning, intelligent tutoring system, pattern recognition, image processing, data analytics, and machine learning. She can be contacted at email: sandhya.arora@cumminscollege.in.