An Efficient Method for Facial Sketches Synthesization Using Generative Adversarial Networks

Surya Prakasa Rao Reddi*

Department of ECE, GVP College of Engineering (A), Visakhapatnam, A.P. India. E-mail: drspreddi4u@gmail.com

T.V. Madhusudhana Rao

Department of CSE, Vignan's Institute of Information Technology Visakhapatnam, A.P. India. E-mail: madhu11211@gmail.com

P. Srinivasa Rao

Department of CSE, MVGR College of Engineering (A), Vizianagaram, A.P. India. E-mail: psr.sri@gmail.com

Prakash Bethapudi

Department of IT, Vignan's Institute of Engineering for Women, Visakhapatnam, A.P. India. E-mail: prakash.vza@gmail.com

Received September 16, 2021; Accepted December 15, 2021 ISSN: 1735-188X DOI: 10.14704/WEB/V1911/WEB19206

Abstract

The synthesis of facial sketches is an important technique in digital entertainment and law enforcement agencies. Recent advancements in deep learning have shown its possibility in generating images/sketches using attribute guided features. Facial features are important attributes because they determine human faces' detailed description and appearance during sketch generation. Traditionally, the forensic or composite artist has to sketch by interviewing witnesses manually. To automate this process of face sketch generation, a deep learning-based generative adversarial network incorporated with multiple activation functions is proposed for its efficiency improvement. The proposed model is extensively tested using different evaluation metrics such as RMSE, PSNR, SSIM, SRE, SAM, UIQ & BRISQUE.

Keywords

Sketch Synthesis, Face Sketch Generation, Attribute Guided, Generative Adversarial Networks (GANs).

Introduction

The facial sketch synthesis technique has long been studied due to its various applications (C. Hu, et al., 2018); (MHM Krishna Prasad, K Thammi Reddy, 2013); (Jain, A., et al., (2017); (Mingrui Z., et al., 2019); (Jonathan L., Evan S., Trevor D., 2015) in computer vision and pattern recognition (H. Kazemi, et al., 2018); (Ilya S., et al., 2014); (T. V. Madhusudhana Rao, et al., 2020). It has a significant role in law enforcement agencies as well as in digital entertainment. In law enforcement, whenever a crime is committed, the information (P Srinivasa Rao, S Satyanarayana, 2018) available about the suspect is limited because either there is no image/video of the suspect or low-quality images or surveillance videos. In such cases, a sketch drawn by a forensic or composite artist by interviewing the victim or witness is considered an alternative for suspect identification (Sharma, S. K., et al., 2019). Then the sketch is compared with the mugshot database for suspect image identification. This is a time-consuming process, and training a forensic or composite artist is expensive (Bhardwaj, A., et al., 2019).

In this scenario, the generative adversarial networks invented by Ian Good fellow is a class of machine learning techniques used to generate images (Krishna Prasad, et al., 2014). The generative model consists of two different networks, namely the generator and the discriminator (Zhang S., et al., 2018); (Joshi M., et al., 2019); (Jain, A., Kumar, A., & Sharma, S., 2015). The generator tries to generate new samples from the true data (Madhusudhana Rao T.V., et al., 2017), which look like the true data to confuse the discriminator between true samples and the fake samples generated by the generator (S.Vidya sagar Appaji, P.V. Lakshmi, 2020). And the discriminator is trained to distinguish between the fake samples generated by the generator and the true samples of the database. Since generative adversarial networks exist, they showed their capability in accomplishing different tasks related to different applications. In this paper, the modified generative adversarial network incorporated with different activation functions has been studied to improve the efficiency in generating more photo-realistic images (Jain, A., et al., 2018); (Jain, A., & Kumar, A., 2021); (Kumar, S., et al., 2021).

The remainder of the paper is ordered as follows section 2 details the literature work, the methodology is explained in section 3, section 4 describes the experimental setup, performance analysis, and experimental results are discussed in section 5, section 6 gives out the conclusion, and future work and future work are given in section 6.

Literature Work

Great progress over the past few years was made for the development of sketch synthesis methods. Xiang Chen et al. (2019) demonstrated a fully-trained generative adversarial

network by training text and image encoders for text to face generation. The experimental results show that the proposed methods outperform the state-of-the-art methodologies. Sert M. and Boyacı E. (2019) experimented with a methodology by fusing transfer learning CNNs at the feature level. Analyzed different layers of different pretrained ImageNet networks by combining with CNN-SVM pipeline and employed principal component analysis for feature dimension reduction for face sketch recognition and yielded good results than the human recognition. H. Zhang et al. (2019) proposed 2-stage generative adversarial networks called StackGAN++ for realistic photo image synthesis. Stage-I generates shape and colour based on input text, and stage-II generates high-resolution images based on stage-I results. The experimental results show that the proposed model obtained better results than the state-of-art methods for realistic photo image synthesis. Hao T. et al. (2019) explained an Attribute Guided Sketch Generative Adversarial Network, which generates more photo-realistic faces with sharper facial attributes than baseline on three different generative tasks, i.e., face-to-attribute-sketch translation, face colourization and face completion. W. Zhang et al. (2019) illustrated a convolutional neural network method using style transfer to synthesize colour sketches and yielded better results than the state-of-art methods. Mingjin Zhang et al. (2019) interpreted the three-stage drawing process of artists and proposed a bionic face sketch generator using generative adversarial networks in U-Net and achieved better performance than the state-of-art works. S. Bae et al. (2019) elucidated using conditional generative adversarial networks with a target style for generating realistic face sketches using a single network (Jain, A., et al., 2016); (Saleem A., Agarwal A.K., 2016). They obtained better quality sketches than the state-of-art methodologies, which uses multiple different networks (Madhusudhana Rao, T.V., Srinivas, Y, 2017); (T.V. Madhusudhana Rao, P.S. Latha Kalyampudi, 2020). Weiguo Wan, Hyo Jong Lee (2019) elucidated a generative adversarial multi-task learning method for face sketch synthesis and recognition simultaneously using U-Net deep model and achieved better performance than state-of-art methods in face sketch synthesis and recognition. Xinxun X. et al. (2019) demonstrated a method for image retrieval using sketches drawn from free hands by incorporating zero-shot learning by employing a triplet loss and obtained a better performance model than the state-of-art methodologies. Varshaneya V et al. (2019) addressed the problem of sketch generation in vector format by proposing a Variational Auto Encoders (VAEs) and Generative Adversarial Networks based frame for producing qualitative and quantitative images. Nan C. et al. (2017) illustrated a deep generative model for generating high-quality multi-class sketches using CNN based autoencoder and achieved better results than Sketch-RNN and Sketch-pix2seq. Wang Y. et al. (2019) illustrated attribute-label based facial image and video generation incorporating 2D and 3D deep conditional generative adversarial networks and generated

realistic faces from attribute labels. Deepanshu Wadhwa et al. (2019) explained a multilayered approach by sketch to image generation using cGAN based pix2pix model and then face recognition by One-Shot Learning using FaceNet and yielded good results on multiple datasets with the generated images performing with an accuracy close to that with the original images (Vidya Sagar Appaji setti ,P Srinivasa Rao., 2018); (P Srinivasa Rao, Krishna Prasad, P.E.S.N., 2017).

As per the literature, several Generative Adversarial Networks are used to generate sketches from images and their attributes. This paper proposes a feature guided Generative Adversarial Network is proposed to improve the accuracy of the model (Agarwal, T., et al., 2014).

Methodology

In order to sketch a suspect in a crime, a composite or a forensic artist should be assigned, and the artist should interview the victim or witness for the features of the suspect's face, which is a time-consuming process. To automate this process, a deep-learning-based generative adversarial network is proposed in this paper for the generation of realistic photo sketches.

The proposed system architecture is highlighted in Fig. 1. Initially, the dataset is divided into training, testing, and validation sets. A few pre-processing techniques, such as resizing the image into 28×28 (as CPU is used for execution, large-sized images cannot be processed on CPU due to insufficient resources), is considered for pre-processing the data.

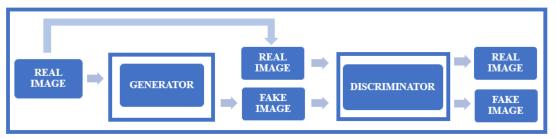


Fig. 1 System architecture

The generator architecture used in the proposed model is highlighted in Fig. 2. The generator model is made up of an input layer, four hidden layers, and one output layer with 4 Conv2DTranspose layers with (256, 128, 128, 64) filters having kernel sizes of 5x5 each with strides (1x1, 2x2, 2x2, 2x2), another Conv2DTranspose layer having output shape of 28x28x1 as an output layer. LeakyReLU activation function is used in all the layers except for the last layer, where the Tanh activation function is used.

http://www.webology.org

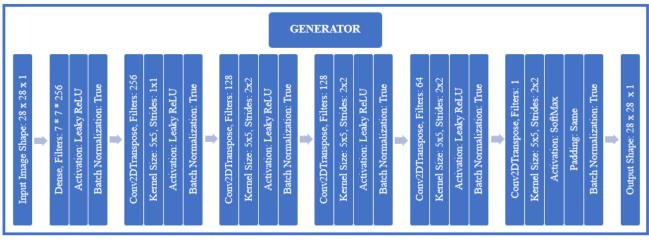


Fig. 2 Generator architecture

LeakyReLU:
$$f(h_{\theta}(x)) = h_{\theta}(x)^{+} = Max(\alpha * h_{\theta}(x), h_{\theta}(x))$$
 (1)
 $f(x) = \tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$ (2)

The architecture of the discriminator is used in the proposed model is highlighted in Fig. 3. The discriminator is made up of an input layer, three hidden layers, and one output layer with 3 Conv2D layers with (64, 128, 256) filters having kernel sizes 5x5 each with strides 2x2, and 1 fully connection layer having 1024 filters and another fully connected layer with 2 filters for output classes. LeakyReLU activation function is used in all the layers except for the last layer, where the SoftMax activation function is used.

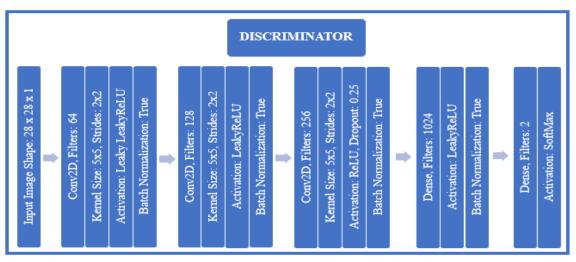


Fig. 3 Discriminator architecture

SoftMax:
$$f(y_i) = \frac{e^{y_i}}{\sum_i e^{y_j}}$$

(3)

The generative adversarial network is trained, bypassing real images from the true data (2015) distribution to generate the fake images. The training continues until the generator generates the fake images that look like the real images. Simultaneously the discriminator is trained to differentiate the real and fake images by extracting feature maps from the images. The trained generator model is used to generate sketches with the feature vector given as an input.

Experimental Setup

The overall experimentation was carried out in a system with the Windows 10 Operating System (64-bit) with Intel® Core[™] i5-8250 CPU @ 1.80 GHz Processor, 8.00 GB Ram, and 2 TB HDD installed with Anaconda, Python platform with supporting Keras packages and Tensorflow as backend (Table 1).

The experiment was performed on a publicly available CUHK Face Sketch Database (CUFS) student dataset, which consists of 188 images with their respective sketches drawn by an artist. The distribution dataset into training, testing, and validation are shown in Table-1. Sample sketches from the dataset are shown in Fig. 4.

DESCRIPTION	STUDENT	
Training Set	88	
Testing Set	90	
Validation Set	10	

Table 1 Description of dataset distribution



Fig. 4 Sample student sketches

Performance Analysis and Experimental Results

The proposed model is tested extensively for its performance based on various evaluation metrics such as root mean square error (RMSE), peak signal to noise ratio (PSNR), structural similarity index measure (SSIM), feature-based similarity index measure (FSIM), information theoretic-based static similarity measure (ISSM), signal to reconstruction error

ration (SRE), spectral angle mapper (SAM) and universal image quality index (UIQ), which are defines as,

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i,j) - k(i,j)]^2$$
(5)

$$PSNR = 20 \cdot log_{10}(MAX_I) 10 \cdot log_{10}(MSE)$$
(6)

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2^{\sigma}_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$
(7)

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_m(x)}{\sum_{x \in \Omega} PC_m(x)}$$
(8)

$$I(x, y) = \frac{C(x, y).EHS(x, y).(a+b)+e}{a.C(x, y).EHS(x, y)+b.EHS(x, y)+c.S(x, y)+e}$$
(9)

$$SRE = 10 \log_{10} \frac{\mu_x^2}{\|\hat{x} - x\|^2/n}$$
(10)

$$SAM = COS^{-1} \frac{\sum_{i=1}^{nb} t_i r_i}{\sqrt{\sum_{i=1}^{nb} t_i^2} \sqrt{\sum_{i=1}^{nb} r_i^2}}$$
(11)

$$UIQ = \frac{4_{xy}^{0}xy}{(\sigma_{x}^{2} + \sigma_{y}^{2})[(\bar{x})^{2} + (\bar{y})^{2}]}$$
(12)

Where (I, K) or (x, y) or (x^{*}, x) or (t, r) are the two comparing images.

All the evaluation metrics considered for the analysis of the proposed generative adversarial network trained with different epochs are shown with their corresponding results in Table 2.

METRIC\EPOCHS	1000	2000	3000	4000	5000
MSE	0.000167453	0.000154679	0.000148554	0.000156705	0.000133433
PSNR	75.52213255	76.21138695	76.56234089	76.09834377	76.68162954
SSIM	0.99982528	0.999894765	0.999920954	0.999890907	0.999990275
FSIM	0.658236601	0.749826309	0.734162682	0.767701874	0.754903748
ISSM	0.605784571	0.611654525	0.612454542	0.658456655	0.685485483
SRE	9.100475174	9.445101715	9.620579091	9.388580058	9.780223268
SAM	48.67347	47.066326	50.05102	46.14796	52.117348
UIQ	0.119020803	0.245594467	0.305478666	0.225691163	0.329932939
BRISQUE	96.24768161	96.76655081	97.8271309	97.17145689	97.98439725

Table 2 Evaluation metrics & their results

From Table 2, it can be inferred that the image quality assessment technique with reference images called structural similarity index measurement has obtained the highest similarity compared with the reference image, which is 0.99%, and without any reference, the similarity obtained using the BRISQUE quality metric is 97.98%. The sample outputs and the reference image are shown in Fig. 5. The samples used for training and testing are

28x28x1 size, having low resolution due to the CPU environment used for training the GAN.

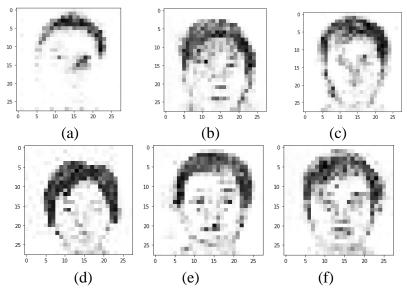


Fig. 5 (a) – (e) represents the output of the generator model after each of 1000, 2000, 3000, 4000 & 5000 epochs and (f) is the reference image

Conclusion and Summary

This paper presents an effective method for generating photo-realistic facial sketches using generative adversarial networks by experimenting with a different activation function. The proposed model generator comprises four hidden layers having Conv2DTranspose layers with their corresponding activation functions. Similarly, the proposed generative adversarial network model discriminator comprises three hidden layers having Conv2D layers with their corresponding activation functions. Among all the activation functions used in the proposed model, the LeakyReLU, Tanh, and SoftMax activation functions and their combinations obtained better results after 5000 epochs with SSIM similarity of 0.99%. Also, the results showcased that the automated process of generating facial sketches using deep learning reduces human effort and requires less time when compared to a sketch drawn by an artist with freehand.

Future Work

Since the experiment is carried out in a CPU environment, low-resolution images are used for computation due to the lack of a GPU environment. Different methods can be used for comparison, and different hybrid architectures can be implemented in generator and discriminator models. These issues can be considered for future work.

References

- Hu, C., Li, D., Song, Y. Z., Xiang, T., & Hospedales, T.M. (2018). Sketch-a-classifier: Sketchbased photo classifier generation. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 9136-9144. http://doi.org/10.1109/CVPR.2018.00952
- Rao, P.S., Reddy, K.T., & Prasad, M.K. (2013). A Novel and Efficient Method for Protecting Internet Usage from Unauthorized Access Using Map Reduce. *International Journal of Information Technology and Computer Science*, 3, 49-55.
- Zhu, M., Wang, N., Gao, X., Li, J., & Li, Z. (2019). Face Photo-Sketch Synthesis via Knowledge Transfer. Proceedings of the Twenty-Eight International Joint Conference on Artificial Intelligence, 1048-1054. https://doi.org/10.24963/ijcai.2019/147.
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, 3431-3440. https://arxiv.org/abs/1411.4038.
- Kazemi, H., Iranmanesh, M., Dabouei, A., Soleymani, S., & Nasrabadi, N.M. (2018). Facial attributes guided deep sketch-to-photo synthesis. In IEEE Winter Applications of Computer Vision Workshops (WACVW), 1-8. http://doi.org/10.1109/WACVW.2018.00006
- Sutskever, I., Vinyals, O., & Le, Q.V. (2014). Sequence to sequence learning with neural networks. *In Advances in neural information processing systems*, 3104-3112.
- Rao, T.M., Rao, P.S., & Kalyampudi, P.L. (2020). Iridology based Vital Organs Malfunctioning identification using Machine learning Techniques. *International Journal of Advanced Science and Technology*, 29(5), 5544-5554.
- Rao, P.S., & Satyanarayana, S. (2018). Privacy preserving data publishing based on sensitivity in context of Big Data using Hive. *Journal of Big Data*, 5(1), 1-20.
- Zhang, S., Ji, R., Hu, J., Gao, Y., & Lin, C.W. (2018). Robust Face Sketch Synthesis via Generative Adversarial Fusion of Priors and Parametric Sigmoid. *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence*, 1163-1169. https://doi.org/10.24963/ijcai.2018/162.
- Madhusudhana Rao T.V., Srinivasa Rao, P., & Srinivas, Y. (2017). A Secure Framework for Cloud Using Map Reduce. *Journal of Advanced Research in Dynamical and Control Systems (IJARDCS)*, 9(Sp-14), 1850-1861.
- Chen, X., Qing, L., He, X., Luo, X., & Xu, Y. (2019). FTGAN: A fully-trained generative adversarial networks for text to face generation. arXiv preprint arXiv:1904.05729.
- Sert, M., & Boyacı, E. (2019). Sketch recognition using transfer learning. *Multimedia Tools and Applications*, 78(12), 17095-17112. https://doi.org/10.1007/s11042-018-7067-1
- Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D.N. (2019). Stackgan++: Realistic image synthesis with stacked generative adversarial networks. *IEEE transactions* on pattern analysis and machine intelligence, 41(8), 1947-1962. http://doi.org/10.1109/TPAMI.2018.2856256
- Tang, H., Chen, X., Wang, W., Xu, D., Corso, J.J., Sebe, N., & Yan, Y. (2019). Attribute-guided sketch generation. In 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019), 1-7.

- Zhang, W., Li, G., Ma, H., & Yu, Y. (2019). Automatic color sketch generation using deep style transfer. *IEEE computer graphics and applications*, 39(2), 26-37. http://doi.org/10.1109/MCG.2019.2899089
- Zhang, M., Wang, N., Li, Y., & Gao, X. (2019). Bionic face sketch generator. *IEEE transactions* on cybernetics, 50(6), 2701-2714.
- Bae, S., Din, N. U., Javed, K., & Yi, J. (2019). Efficient generation of multiple sketch styles using a single network. *IEEE Access*, 7, 100666-100674. http://doi.org/10.1109/ACCESS.2019.2931544
- Wan, W., & Lee, H.J. (2019). Generative adversarial multi-task learning for face sketch synthesis and recognition. *In IEEE International Conference on Image Processing (ICIP)*, 4065-4069. http://doi.org/10.1109/ICIP.2019.8803617
- Xu, X., Wang, H., Li, L., & Deng, C. (2019). Semantic adversarial network for zero-shot sketchbased image retrieval. arXiv preprint arXiv:1905.02327.
- Balasubramanian, S., & Balasubramanian, V.N. (2019). Teaching gans to sketch in vector format. arXiv preprint arXiv:1904.03620.
- Cao, N., Yan, X., Shi, Y., & Chen, C. (2019). AI-sketcher: a deep generative model for producing high-quality sketches. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 33(01), 2564-2571. https://doi.org/10.1609/aaai.v33i01.33012564
- Wang, Y., Dantcheva, A., & Bremond, F. (2018). From attribute-labels to faces: face generation using a conditional generative adversarial network. *In Proceedings of the European Conference on Computer Vision (ECCV) Workshops*. https://doi.org/10.1007/978-3-030-11018-5 59
- D. Wadhwa, U. Maharana, D. Shah, V. Yadav and P. Pandey, "Human Sketch Recognition using Generative Adversarial Networks and One-Shot Learning," 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India, 2019, pp. 1-6
- Wadhwa, D., Maharana, U., Shah, D., Yadav, V., & Pandey, P. (2019). Human Sketch Recognition using Generative Adversarial Networks and One-Shot Learning. In Twelfth International Conference on Contemporary Computing (IC3), 1-6. http://doi.org/10.1109/IC3.2019.8844885
- Rao, P.S., Prasad, M.K., & Reddy, K.T. (2015). An efficient semantic ranked keyword search of big data using map reduce. *International Journal of Database Theory and Application*, 8(6), 47-56.
- Jain, A., Kumar, A., & Sharma, S. (2015). Comparative Design and Analysis of Mesh, Torus and Ring NoC. Procedia Computer Science, 48, 330-337.
- Jain, A., Dwivedi, R., Kumar, A., & Sharma, S. (2017). Scalable design and synthesis of 3D mesh network on chip. In Proceeding of International Conference on Intelligent Communication, Control and Devices, 661-666.
- Jain, A., Gahlot, A. K., Dwivedi, R., Kumar, A., & Sharma, S.K. (2018). Fat Tree NoC Design and Synthesis. In Intelligent Communication, Control and Devices, 1749-1756.
- Srinivasa Rao, P., Krishna Prasad, M.H.M., & Thammi Reddy, K (2014). A Efficient Data Integration Framework in Hadoop Using MapReduce. Computational Intelligence Techniques for Comparative Genomics, 129-137.

- Sharma, S. K., Jain, A., Gupta, K., Prasad, D., & Singh, V. (2019). An internal schematic view and simulation of major diagonal mesh network-on-chip. *Journal of Computational and Theoretical Nanoscience*, 16(10), 4412-4417.
- Madhusudhana Rao, T.V., & Srinivas, Y. (2017). A Secure Framework for Cloud Using Map Reduce. Journal of Advanced Research in Dynamical and Control Systems, 9(Sp-14), 1850-1861.
- Jain, A., & Kumar, A. (2021). Desmogging of still smoggy images using a novel channel prior. Journal of Ambient Intelligence and Humanized Computing, 12(1), 1161-1177.
- Kumar, S., Jain, A., Shukla, A.P., Singh, S., Raja, R., Rani, S. & Masud, M. (2021). A Comparative Analysis of Machine Learning Algorithms for Detection of Organic and Nonorganic Cotton Diseases. *Mathematical Problems in Engineering*.
- Rao, T.M., Rao, P.S., & Kalyampudi, P.L. (2020). Iridology based Vital Organs Malfunctioning identification using Machine learning Techniques. *International Journal of Advanced Science and Technology*, 29(5), 5544-5554.
- Jain, A., Dwivedi, R., Kumar, A., & Sharma, S. (2016). Network on chip router for 2D mesh design. *International Journal of Computer Science and Information Security*, 14(9).
- Appaji, S.V.S., Lakshmi, P.V., & Rao, P.S. (2019). Maximizing Joint Probability in Visual Question Answering Models. *International Journal of Advanced Science and Technology*, 29(3), 3914-3923.
- Bhardwaj, A., Kaur, S., Shukla, A.P., & Shukla, M.K. (2019). Performance Comparison of Despeckling filters on the Basis of Incremental Iteration in Ultrasound Imaging. In International Conference on Power Electronics, Control and Automation (ICPECA), 1-5.
- Setti, V.S.A., & Srinivasa Rao, P. (2018). A Novel Scheme for Red Eye Removal with Image Matching. *Journal of Advanced Research in Dynamical & Control Systems, 10*(13-Special Issue).
- Joshi, M., Agarwal, A.K., & Gupta, B. (2019). Fractal image compression and its techniques: a review. Soft Computing: Theories and Applications. Advances in Intelligent Systems and Computing, 235-243. https://doi.org/10.1007/978-981-13-0589-4_22
- Rao, P.S., & Prasad, P.K. (2017). A Secure and Efficient Temporal Features Based Framework for Cloud Using MapReduce. In International Conference on Intelligent Systems Design and Applications, 114-123.
- Agarwal, T., Agarwal, A.K., & Singh, S.K. (2014). Cloud computing security: issues and challenges. *In Proceedings of SMART*, 10-14.
- Saleem, A., & Agarwal, A.K. (2016). Analysis and design of secure web services. In Proceedings of Fifth International Conference on Soft Computing for Problem Solving, 453-462. https://doi.org/10.1007/978-981-10-0451-3_41

http://www.webology.org