

Photo-realistic photo synthesis using improved conditional generative adversarial networks

Raghavendra Shetty Mandara Kirimanjeshwara^{1,2}, Sarappadi Narasimha Prasad³

¹School of ECE, Reva University, Bengaluru, India

²Department of Electronics and Communication Engineering, Canara Engineering College, Mangaluru, India

³Department of Electrical and Electronics Engineering, Manipal Institute of Technology Bengaluru Manipal Academy of Higher Education, Manipal, India

Article Info

Article history:

Received Feb 17, 2023

Revised Mar 28, 2023

Accepted Apr 12, 2023

Keywords:

Computer vision

Conditional generative adversarial networks

Image processing

Photo-sketch synthesis

Pix2pix generative adversarial networks

ABSTRACT

There are a wide range of potential uses for both the forward (generating face drawings from actual images) and backward (generating photos from synthetic face sketches). However, photo/sketch synthesis is still a difficult problem to solve because of the distinct differences between photos and sketches. Existing frameworks often struggle to acquire a strong mapping among the geometry of drawing and its corresponding photo-realistic pictures because of the little amount of paired sketch-photo training data available. In this study, we adopt the perspective that this is an image-to-image translation issue and investigate the usage of the well-known enhanced pix2pix generative adversarial networks (GANs) to generate high-quality photo-realistic pictures from drawings; we make use of three distinct datasets. While recent GAN-based approaches have shown promise in image translation, they still struggle to produce high-resolution, photorealistic pictures. This technique uses supervised learning to train the generator's hidden layers to produce low-resolution pictures initially, then uses the network's implicit refinement to produce high-resolution images. Extensive tests on three sketch-photo datasets (two publicly accessible and one we produced) are used to evaluate. Our solution outperforms existing image translation techniques by producing more photorealistic visuals with a peak signal-to-noise ratio of 59.85% and pixel accuracy of 82.7%.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Narasimha Sarappadi Prasad

Department of Electrical and Electronics Engineering

Manipal Institute of Technology Bengaluru Manipal Academy of Higher Education

Manipal, India-576104

Email: sn.prasad@manipal.edu

1. INTRODUCTION

Converting sketch photographs to photo-realistic visuals in a automated manner finds enormous applications in electronic entertainment, arts, security forces, and several domains. A face photograph with poor identification and one that was drawn by police or a drawing expert from an eyewitness's description are two very different things in the realm of law enforcement. Producing a genuine color picture by hand requires an experienced team with drawing and painting expertise in addition to solid investigative abilities. In several application sectors, face recognition is a well-studied topic. Nonetheless, matching drawings to digital facial photos is a crucial law enforcement application that has garnered comparatively little attention. Based on the memories of an eyewitness and the skill of a sketch artist, forensic drawings are made. Due to the limited or approximate description supplied by the eyewitness, forensic drawings include several inaccuracies. Typically,

forensic drawings are compared manually with a database of digitised face photos of identified people. Existing cutting-edge face recognition algorithms cannot be used directly and need extra processing to account for the non-linear fluctuations found in face sketches and digital face photos. A technology that automatically matches a drawing to a digital facial picture may aid law enforcement organisations and make the identification procedure efficient and reasonably quick. It takes the skill of a sketch artist and the witness's description for a forensic drawing to be created. As seen in Figure 1, forensic drawings involve many flaws due to incomplete or imprecise narrative supplied by the eyewitness. Due to the non-linear fluctuations in both drawings and digital face photos, current state-of-the-art face recognition techniques cannot be employed directly and need further processing. The identification process may be sped up and made more accurate by using an automated sketch to digital facial picture matching system, which can be a great help to law enforcement.



Figure 1. Forensic art that exaggerates face features

Automatic synthesis of realistic face portraits from sketch photos using generative adversarial network (GAN) [1] may increase the likelihood of recognition and boost security agencies case-solving efficiency. Because of this, there has been a lot of work done in the field of artificial intelligence (AI) to figure out how to turn drawings into portraits [2], [3]. One common technique for making images is the use of convolutional neural networks (CNNs) for transferring artistic effects from one image to another. However, in the picture generation using CNNs, both the source image and the destination image must be given. To generate a photorealistic picture, one may use standard graphics theory to recreate the curves, skin tone, and lighting of a realistic one [4]–[6]. In practise, typical graphics algorithms work well, but it's costly and tedious to construct and modify the visual context, and each detail of the region needs to be defined precisely. Pix2pix is capable to produce portraiture in the use of trained models, but the significant variety of test results indicate that it is nevertheless far from perfect; the images' contours are sometimes ambiguous, and some results missing intricate details and image texturing. In this research, we implement fixes to these problems into the pix2pix generating model. Edge information is obtained, the image's contour is modified, and portraiture generations converging is limited by new model. Our experiments show that the modified version of pix2pix effectively fixes the edge blurring problem in face image synthesizing, and also serves as a standard for similar image-generating applications.

Several researchers have used adversarial learning to transform one image into another. Training data consists of pairs of input and output pictures; the input photographs are converted into the desired target images by following the examples provided. Current interesting advances in picture production may be attributed to the fast growth of deep learning, particularly the introduction of generative adversarial networks (GAN) [7]. The purpose of this research is to refine the model and derive edge information from the dataset itself. The model's image-generation capabilities will be constrained by the necessity for high similarity between test and training imagery, and usage of a boundary map will further complicate the already challenging task of generating a dataset. The super-resolution technique is the gold standard in computer vision [8], [9].

For the study of dynamic facial expression change, the author [10] used GANs to create static facial expression photos from a natural (expressionless) image. Experimental testing findings indicate that the method yields the superior picture of facial emotion. For verification purposes, the discriminator receives a composite of the image produced by the generator network and the edge produced by the edge network that comes next. Experiments show [11] that the proposed approach can more successfully generate colour portraiture from drawings than preexisting techniques, and that the photos it generates have a more distinct and convincing edge than those generated by a pix2pix model. The average value of the structural similarity index measure (SSIM) is 82.78% when using the recommended approach, whereas the values are 42.99% and 78.60% when using pix2pix and alternative techniques, respectively. The author investigates [12] picture creation led by hand drawing in this study. Due to the strict requirements imposed by the image-to-image translation procedure, when the input sketch is poorly drawn, the output follows the input edges. Instead, we suggest using sketch as a weak constraint, in which the output edges are not required to match the input edges. We solve this issue with a unique method of cooperative picture completion in which the sketch offers the visual context for image

completion or generation. Using joint pictures, we train a contextually generated adversarial network (GAN) to learn the joint distribution of a drawing and its associated image.

In this paper, the author [13] demonstrated a unique generative adversarial network (GAN) technique that generates realistic pictures from 50 categories, including motorbikes, horses, and sofas. A completely automated data augmentation approach for drawings demonstrates that the supplemented data is beneficial to our purpose. A novel network building block that is suitable for both the generator and the discriminator by injecting the input picture at different scales was proposed. The proposed method creates more realistic visuals and obtains much higher Inception scores when analysed against state-of-the-art image translation techniques. Using convolutional neural network (CNN)-based feature extraction from the Modified National Institute of Standards and Technology (MINST) dataset and algebraic fusion of several classifiers trained on multiple varied feature sets (obtained via feature selection applied to the CNN-extracted feature set), author [14] described a system capable of recognising a wide range of images. The author designed [15] a neural algorithm which can discretise and associate the visual information as well as artistic style of natural photos. By combining the subject matter of any given photograph with the aesthetics of other well-known works of art, the algorithm enables the generation of new images of high thoughtful quality. The results provide new light on the deep picture representations that convolutional neural networks learn and show how these networks may be used for sophisticated image creation and manipulation.

Normalized direction-preserving Adam (ND-Adam) is a method proposed in [16] that improves generalisation performance by enabling finer granularity over both the direction and step size of weight vector updates. Following similar reasoning, we increase the generalisation performance of classification problems further by regularising the softmax logits. Not only do researchers hope to close the gap between stochastic gradient descent (SGD) and Adam, but also to provide light on the topic of why some optimization methods are more broadly applicable than others.

Pix2pix, a specialised GANs model, defines the image translation issue as the mapping connection amongst the input and output pixels. While a convolutional classifier called "PatchGANs" is used in the discriminator, an Unsampled-Network (U-Net) structure is used in the generator. Pix2pix gives a universal solution to the picture conversion issue, unlike CNNs and other GANs. Extensive conditional training is employed to automatically learn the loss function of the image translation problem, which is then utilised to restrict the possible directions of image translation and convergence. Whenever translating a sketch into a photo-realistic image or altering the style of actual images, for example, it is common for fine details and realistic textures to be lost in the action of image transformations because the image structures of the semantic picture and target image are so drastically dissimilar [17], [18]. To combat the problem of edge blur in the converted photos, Wang manually inserted edge information to every label image throughout model training.

Subspace learning, sparse representation, Bayesian inference, and deep-learning-based techniques are main four types of existing face-sketch synthesis approaches; the first three fall under the data-driven category, while the final kind is model-driven [19]. The deep-learning-based technique emphasises a model-driven technique in which the mapping function is learned beforehand and utilised to implement the transformations. The various past works and their approaches are summarised in Table 1.

Table 1. Related works summary

References	Aim	Algorithm	Measures
[2]	Facial Photo-Sketch Synthesis	Multi-Adversarial Networks	Image Quality Assessment (IQA)
[3]	Caricature Sketching	Deep learning, pix2pix-Net	Mean Square Error
[4], [13]	Face sketch-photo synthesis, Matching photo-realistic images	Generative Adversarial Networks, SketchyGAN	Semantic Accuracy, Fooling Rate, Inception Score
[5], [11]	High-Resolution Image Synthesis	Conditional GAN	Pixel accuracy, mean intersection-over-union, peak signal-to-noise ratio (PSNR) and SSIM
[6]	Image-to-Image Translation	Cycle-Consistent Adversarial Networks	FCN score
[12]	Face sketch-photo synthesis	Contextual GAN	SSIM, Verification Accuracy
[17]	Sketch-Photo Synthesis	Sparse Representation	Verification Accuracy
[18]	Face Sketch-to-Photo Retrieval	Fuzzy rule based layered classifier	False Positive Rate, Cumulative Match Curve
[20]	Forensic Face Photo-Sketch Recognition	Convolutional neural network	Average rank value
[21]	Limited-labeled Sketch-to-Photo Retrieval	Instance-level Heterogeneous Domain Adaptation (IHDA) framework	Rank-1 Accuracy
[22]	Face Photo-Sketch Synthesis and Recognition	Identity-Aware CycleGAN	SSIM, FSIM

According to some sketch artists, making a drawing is an unknown psychological phenomenon, but a sketch artist usually focuses on face characteristics and texture, which he/she attempts to integrate in the sketch via a combination of soft and noticeable edges. Thus, local descriptors can effectively express facial patterns in drawings and digital face photographs, which inspired the proposed technique. This study uses the improved pix2pix cGAN model to match drawings with digital facial pictures automatically. This study introduces a pre-processing method to improve forensic sketch-digital picture pairings. Pre-processing forensic sketching boosts performance by 4–5%. In this research we have used 3 different datasets: The Chinese University of Hong Kong (CUHK) and Indraprastha Institute of Information Technology Delhi's (IIIT-D's) students face sketch and corresponding digital images, available online [23], [24] and Self-generated datasets of students at Canara Engineering College, Mangalore.

2. METHOD

While examining the current state of sketch-photo synthesis methods, we discovered a potential drawback that might hinder the face image retrieval procedure. The number of nearest neighbours is fixed, hence the pseudo images produced using these approaches have poor resolution as can be seen [25]. Specifically, we investigate GANs in a conditional context using this method. In the same way that generative adversarial networks (GANs) learn a generative model from data, conditional GANs (cGANs) learn a conditional generative model [26]. The ability to condition on an input images and produce an output image makes cGANs useful for “image-to-image translation” applications.

2.1. Pix2pix cGAN architecture

Pix2pix is a kind of conditional GAN (cGANs), production of the target picture is dependent on the source imagery data. Generator and discriminator comprise the network. The generator generates the picture from the input. The discriminator compares the supplied picture to an unknown image and guesses whether it was generated. The loss that is anticipated by the discriminator for the produced images is minimised by updating the generator. To prevent overfitting, the generator is only indirectly directed by the loss functions during training, and it is never directly presented the training dataset. To counter this, a dropout layer is deployed in both the training and testing phases, and also provides a random sample for the generator [7], [27]. It is possible to modify the discriminator model directly, but the generator model can only indirectly modify. Towards this end, a novel composite model might be crafted in which the discriminator model incorporates the output of the generator model as a required input. Stacking the generator model above the discriminator creates a hybrid architecture. The sharpness loss term was suggested for use in Sharp-GAN to produce nuclei with distinct borders [7]. It established harsh penalties for contour pixels that showed little variation from their surrounding counterparts. GANs' hazy-boundary problem was fixed by the sharpness reduction. Total projected loss is defined as,

$$L(G, D) = E_{x,y}[L_1] + E_x[L_2 + L_{s1} + \beta L_{s2}] \quad (1)$$

Wherein L1 and L2 are,

$$L_1 = -\log D(x, y) \quad (2)$$

$$L_2 = -\log (1 - D(x, G(x))) \quad (3)$$

when matched up with a discriminator D that seeks to maximise loss, as in (1)-(3), the generator G seeks to reduce loss L. The generator is revised to reduce the L1 loss, commonly known as the “Mean absolute error”, amongst the output and input. A weighted sum of the adversarial loss obtained out of discriminator's output and L1 losses are used to update the generator in this manner. Figure 2 illustrates the procedural diagram of pix2pix model for sketch to image synthesis.

2.1.1. Pix2pix cGAN training

All available instances of each class's images were used to train two sets of identical pix2pix cGAN models. Each class's pre-processed images were uploaded as random pairs for the source and target images. Image pairs uploaded are scaled such that value of the pixel range from -1 to +1 rather than 0 to 255. GAN models usually reach an equilibrium between the generator and discriminator. Thus, stopping training is difficult to determine. Hence, we routinely stored the model including its weights throughout training cycles to obtain sample images for quality evaluation. Model weights are generated using a random Gaussian distribution with a mean of 0.01 and a standard deviation of 0.02. Discriminator loss is weighted by 50% for every model update to slow down discriminator training, which is faster than generator training. The suggested approach is pictorially shown in Figure 3.

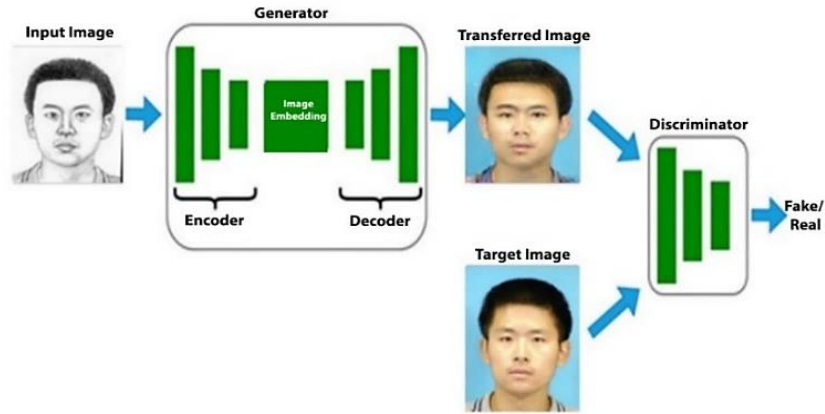


Figure 2. Procedural diagram of pix2pix model for image synthesis

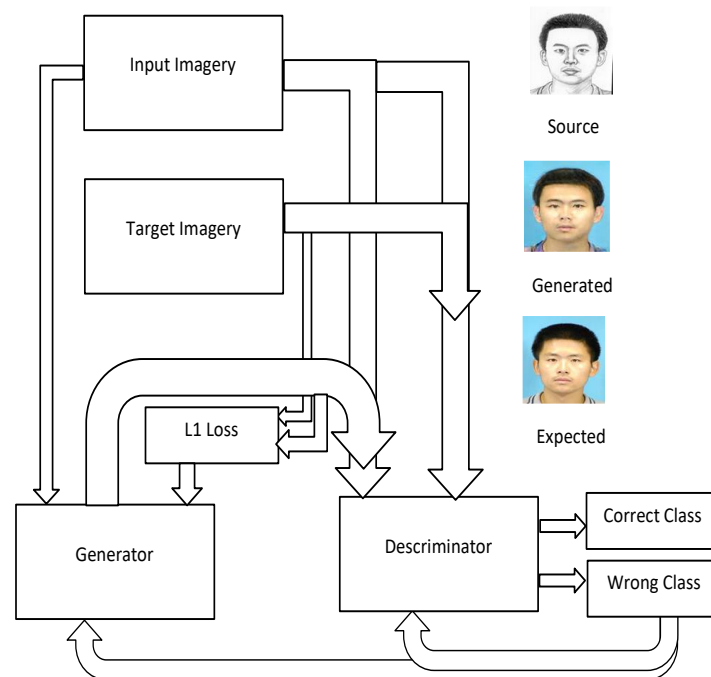


Figure 3. Proposed pix2pix model

2.1.2. Database used

Since collecting face drawings is difficult, there are few datasets of human-drawn sketches and face photographs. To evaluate our method against those already in use, we use 88 subjects from the CUHK student dataset as a training set and the remaining 518 subjects as a testing set, which consists of 123 images from the augmented reality (AR) dataset, 295 images from the XM2VTS dataset, and the remaining 100 images from the CUHK student dataset. We also employ a mixture of the IIITD dataset and data we generate in-house, in addition to the CUHK dataset. There are 238 sketch-digital picture pairings in the IIITD database. Digital photographs from various sources are used to create these drawings. It has 72 sketch-digital picture pairings from the IIIT-D student and staff database, 99 from the labeled faces in the wild (LFW) database, and 67 from the face and gesture recognition network (FG-NET) ageing database. Canara College of Engineering, Mangalore student face sketch photos, together with their equivalent digital photographs captured in a variety of lighting situations, make up the custom dataset.

3. RESULTS AND DISCUSSION

In order to design and implement the proposed model, we utilized the “TensorFlow”, a open-source framework and the Python programming language. The experiment was executed on a Windows 11 OS using

a 5 GHz Intel i7 12700k 12-Core CPU, GPU of 4GB. 200 epochs of experimentation are used. Used batch size is 64 and learning rate is 0.0001. Adam optimizer is used. To characterize the imagery color rendering quality of different models more objectively, we employ two primary indices: the peak signal-to-noise ratio (PSNR) (4) and the structural similarity index measure (SSIM) in (6).

$$PSNR = 10 \log_{10}(2n - 1) 2MSE \quad (4)$$

$$MSE = 1HW_i = 1H_j = 1W(i, j) - y(i, j)2 \quad (5)$$

$$SSIM = (2xx + c1)(2xy + c2)(x'y' + c1)(x'y' + c2) \quad (6)$$

Since the PSNR index, which has its own set of limitations, is not sufficient for describing the quality of the generated photo and its visual characteristics [28], the SSIM index is used for further comparison. Where x and y are the mean values of the true and created false images, respectively. x' and y' show the difference between the authentic and false images, whereas xy shows their covariance. Relative performance comparison is shown in Table 2.

Different generators are compared with all other components fixed. In particular, we evaluate our generator in contrast to the state-of-the-art U-Net and customer reference number (CRN) generator designs as shown in Table 3. Both semantic segmentation scores and findings from human perceptual studies are considered in our performance evaluation. Figure 4(a) shows the training loss curve for D_{fake} , Figure 4(b) shows training loss for D_{real} and Figure 4(c) shows training loss for G_{Gan} model training of 200 epoch. As seen in Figure 4, early in the run, all three losses exhibit considerable randomness before levelling out between epochs 175 and 200. After that point, losses are stable, albeit their variability grows. Figure 5 gives visual output of sketch to image synthesis.

Table 2. PSNR and SSIM comparison with related work

Algorithm	PSNR	SSIM
Cycle GAN [6]	58.56	0.345
Our Approach	59.85	0.356

Table 3. Generator design comparison

Measures	U-Net [6]	CRN [27]	Ours
Pixel Accuracy in Percentage	77.857	78.956	82.7
Mean Intersection-over-union (IoU)	0.391	0.399	0.456

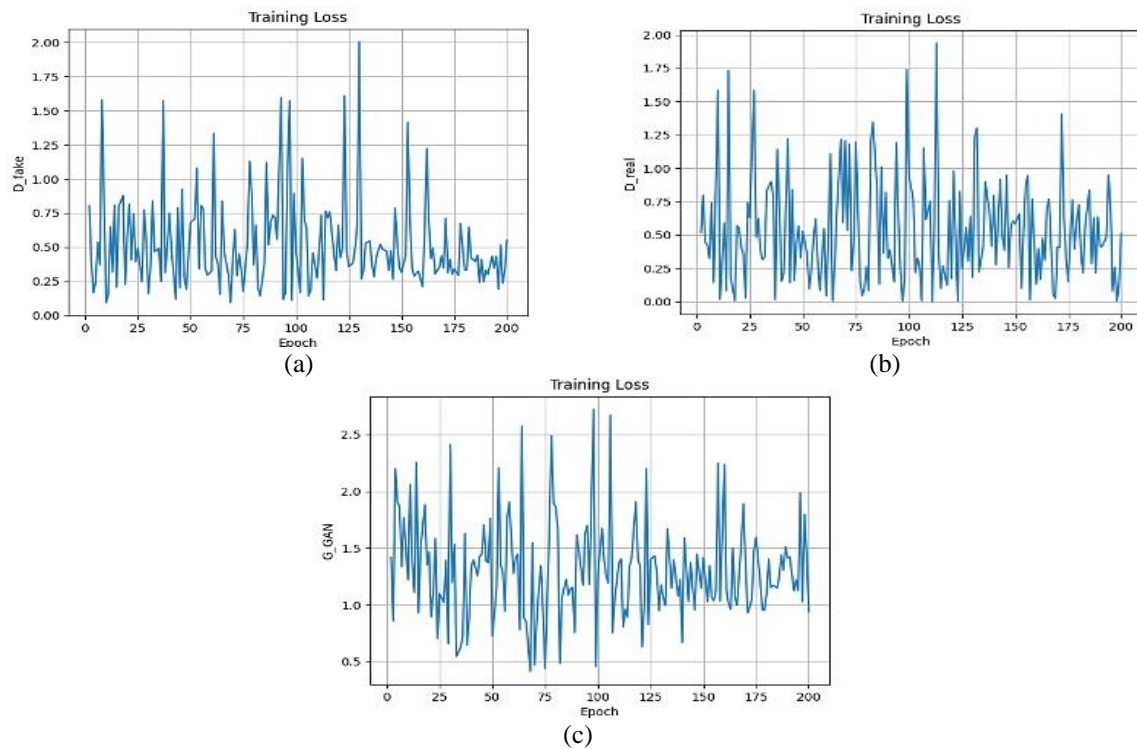


Figure 4. Training loss curve of the proposed model; (a) D_{fake} training loss, (b) D_{real} training loss, and (c) G_{GAN} training loss

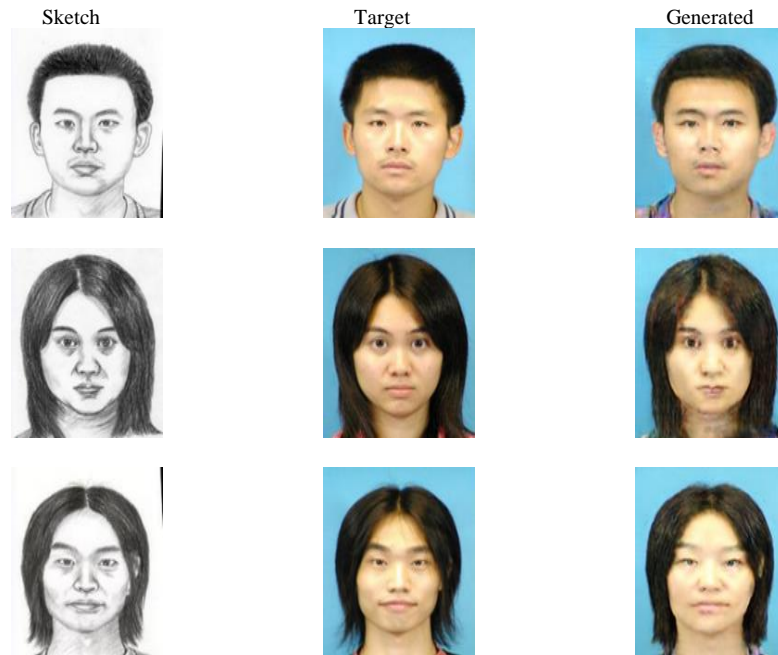


Figure 5. Visual output of sketch-to-image synthesis

4. CONCLUSION

Using the pix2pix generative model, we investigated the issue of photo-sketch synthesis. The suggested approach was designed to aid GANs produce high-resolution photorealistic pictures from drawings. Three datasets are used for analyses, and the outcomes are compared to those generated by the most cutting-edge generative methods available. It has been shown without any reasonable doubt that the suggested strategy significantly enhances visual quality. Extensive paired data experiments prove that our strategy outperforms the alternatives we investigated. Our suggested method has been shown to produce a pixel accuracy of 82.7% in experimental settings. Furthermore, it is suggested to hyper parameter optimization to see whether we can further improve performance by selecting the best possible feature subsets.




REFERENCES

- [1] J. Lin, D. Liu, H. Li, and F. Wu, "Generative adversarial network-based frame extrapolation for video coding," Dec. 2018, doi: 10.1109/VCIP.2018.8698615.
- [2] L. Wang, V. Sindagi, and V. Patel, "High-quality facial photo-sketch synthesis using multi-adversarial networks," in *Proceedings - 13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018*, May 2018, pp. 83–90, doi: 10.1109/FG.2018.00022.
- [3] X. Han *et al.*, "Caricatureshop: Personalized and photorealistic caricature sketching," *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 7, pp. 2349–2361, Jul. 2020, doi: 10.1109/TVCG.2018.2886007.
- [4] H. Kazemi, F. Taherkhani, and N. M. Nasrabadi, "Unsupervised facial geometry learning for sketch to photo synthesis," Sep. 2018, doi: 10.23919/BIOSIG.2018.8552937.
- [5] T. C. Wang, M. Y. Liu, J. Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 8798–8807, doi: 10.1109/CVPR.2018.00917.
- [6] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2017-October, pp. 2242–2251, 2017, doi: 10.1109/ICCV.2017.244.
- [7] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.
- [8] Y. Guo *et al.*, "Closed-loop matters: Dual regression networks for single image super-resolution," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2020, pp. 5406–5415, doi: 10.1109/CVPR42600.2020.00545.
- [9] S. Maeda, "Unpaired image super-resolution using pseudo-supervision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2020, pp. 288–297, doi: 10.1109/CVPR42600.2020.00037.
- [10] Y. Kawai, M. Seo, and Y. W. Chen, "Automatic generation of facial expression using generative adversarial nets," in *2018 IEEE 7th Global Conference on Consumer Electronics, GCCE 2018*, Oct. 2018, pp. 329–330, doi: 10.1109/GCCE.2018.8574866.
- [11] W. Xia, Y. Yang, and J. H. Xue, "Cali-sketch: Stroke calibration and completion for high-quality face image generation from human-like sketches," *Neurocomputing*, vol. 460, pp. 256–265, Oct. 2021, doi: 10.1016/j.neucom.2021.07.029.
- [12] Y. Lu, S. Wu, Y. W. Tai, and C. K. Tang, "Image generation from sketch constraint using contextual GAN," in *Lecture Notes in*




- Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11220 LNCS, Springer International Publishing, 2018, pp. 213–228.
- [13] W. Chen and J. Hays, “SketchyGAN: Towards diverse and realistic sketch to image synthesis,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2018, pp. 9416–9425, doi: 10.1109/CVPR.2018.00981.
- [14] H. Huang Zhao and H. Liu, “Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition,” *Granular Computing*, vol. 5, no. 3, pp. 411–418, Feb. 2020, doi: 10.1007/s41066-019-00158-6.
- [15] L. A. Gatys, A. S. Ecker, and M. Bethge, “Image style transfer using convolutional neural networks,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2016, vol. 2016-Decem, pp. 2414–2423, doi: 10.1109/CVPR.2016.265.
- [16] Z. Zhang, “Improved adam optimizer for deep neural networks,” Jun. 2019, doi: 10.1109/IWQoS.2018.8624183.
- [17] X. Gao, N. Wang, D. Tao, and X. Li, “Face sketch-photo synthesis and retrieval using sparse representation,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 8, pp. 1213–1226, Aug. 2012, doi: 10.1109/TCSVT.2012.2198090.
- [18] M. A. Khan and A. S. Jalal, “A fuzzy rule based multimodal framework for face sketch-to-photo retrieval,” *Expert Systems with Applications*, vol. 134, pp. 138–152, Nov. 2019, doi: 10.1016/j.eswa.2019.05.040.
- [19] C. Galea and R. A. Farrugia, “Forensic face photo-sketch recognition using a deep learning-based architecture,” *IEEE Signal Processing Letters*, vol. 24, no. 11, pp. 1586–1590, Nov. 2017, doi: 10.1109/LSP.2017.2749266.
- [20] F. Yang, Y. Wu, Z. Wang, X. Li, S. Sakti, and S. Nakamura, “Instance-level heterogeneous domain adaptation for limited-labeled sketch-to-photo retrieval,” *IEEE Transactions on Multimedia*, vol. 23, pp. 2347–2360, 2021, doi: 10.1109/TMM.2020.3009476.
- [21] S. Setumin, A. Radman, and S. A. Suandi, “An empirical investigation on the effect of shape exaggeration in face sketch to photo matching,” in *Proceedings of the 2019 IEEE International Conference on Signal and Image Processing Applications, ICSIPA 2019*, Sep. 2019, pp. 302–307, doi: 10.1109/ICSIPA45851.2019.8977748.
- [22] Y. Fang, W. Deng, J. Du, and J. Hu, “Identity-aware CycleGAN for face photo-sketch synthesis and recognition,” *Pattern Recognition*, vol. 102, p. 107249, Jun. 2020, doi: 10.1016/j.patcog.2020.107249.
- [23] “CUHK facesketch database,” 2023, [Online]. Available: <http://mmlab.ie.cuhk.edu.hk/archive/facesketch.html>.
- [24] “IIIT-D sketch database,” 2023, [Online]. Available: <http://iab-rubric.org/index.php/iiit-d-sketch-database>.
- [25] L. Lan *et al.*, “Generative adversarial networks and its applications in biomedical informatics,” *Frontiers in Public Health*, vol. 8, May 2020, doi: 10.3389/fpubh.2020.00164.
- [26] J. Zhao, M. Mathieu, and Y. LeCun, “Energy-based Generative Adversarial Network,” *MM 2017 - Proceedings of the 2017 ACM Multimedia Conference*, pp. 672–680, Sep. 2016, doi: 10.1145/3123266.3123334.
- [27] S. Butte, H. Wang, M. Xian, and A. Vakanski, “Sharp-GAN: Sharpness loss regularized GAN for histopathology image synthesis,” in *Proceedings - International Symposium on Biomedical Imaging*, Mar. 2022, vol. 2022-March, doi: 10.1109/ISBI52829.2022.9761534.
- [28] Q. Chen and V. Koltun, “Photographic image synthesis with cascaded refinement networks,” in *Proceedings of the IEEE International Conference on Computer Vision*, Oct. 2017, vol. 2017-October, pp. 1520–1529, doi: 10.1109/ICCV.2017.168.

BIOGRAPHIES OF AUTHORS



Raghavendra Shetty Mandara Kirimanjeshwara    research Scholar in School of ECE at REVA University, Bengaluru, India. Graduation from Karnataka University Dharwad and Post-graduation from VTU Belagavi. Total 14 years in teaching and 08 years of industry experience in Engineering Field. His areas of Interest: AI, Embedded systems, and renewable Energy. He can be contacted at: r19pec11@gmail.com.



Sarappadi Narasimha Prasad    Professor in Department of Electrical and Electronics Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education (MAHE), Manipal, Karnataka, India-576104. Total experience is 22 Years, completed graduation from Mangalore University, Post-graduation from VTU and Doctorate from Jain University. More than 80 Journals/Conferences in profile and presently guiding 8 research scholars. Area of Interest is AI, Embedded Systems and Signal Processing. He can be contacted at email: sn.prasad@manipal.edu.