**Facial Photo Blending System**

Submitted by

221FA04181 **M.Sandeep Kumar**

221FA04201 **S.Jahnavi**

221FA04224 **M.Sai Sandeep**

221FA04625 **M.Shanmukha Priya**

Under the guidance of

**Dr.Vinoj**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**VIGNAN’S FOUNDATION FOR SCIENCE, TECHNOLOGY AND RESEARCH**

**(Deemed to be UNIVERSITY)**

Vadlamudi, Guntur.

ANDHRA PRADESH, INDIA, PIN-522213.



**CERTIFICATE**

This is to certify that the Field Project entitled “Facial Photo Blending System” that is being submitted by 221FA04181 (M.Sandeep Kumar), 221FA04201 (S.Jahnavi), 221FA04224 (M.Sai Sandeep),221FA04625 (M.Shanmukha Priya) for partial ful- filment of Field Project is a bonafide work carried out under the supervision of Dr.Vinoj,

Assistant Professor,Department of CSE.

### Guide Name & Signature

### HOD, CSE Dean, SoCI

**DECLARATION**

We hereby declare that the Field Project entitled “Facial Photo Blending System” is being submitted by 221FA04181 (M.Sandeep Kumar), 221FA04201 (S.Jahnavi), 221FA04224 (M.Sai Sandeep), and 221FA04625 (M.Shanmukha Priya) in partial fulfilment of Field Project course work. This is our original work, and this project has not formed the basis for the award of any degree. We have worked under the supervision of Dr.Vinoj, Assistant Professor, Department of CSE.

### By

221FA04181 (M.Sandeep Kumar), 221FA04201 (S.Jahnavi),

221FA04224 (M.Sai Sandeep) 221FA04625 (M.Shanmukha Priya)

Date:

**ABSTRACT**

This paper introduces a new Photo Blending system approach to FSS in pursuit of per- formance enhancement regarding verification and recognition identity, which current ap- proaches typically fail to realize when it tends to miss some important specific identity information in the traditional approach. Inter-domain transfer occurs without losing any critical facial structures that are learnt as regressor between test and training photos; and intra-domain transfer tries to boost recovery of identity-specific information through a mapping of relationships between sketches and photographs across different identities. To facilitate research in this area, we present FS2K, a comprehensive dataset containing 2,104 image-sketch pairs that encompass var ious sketch styles, backgrounds, and facial attributes. Additionally, we propose FSGAN, a baseline method that utilizes facial-aware masking and style-vector expansion, signif- icantly outperforming existing state-of-the- art models on the FS2K dataset. Our dual Path Framework With its finest adjustment of coarse crossdomain recon- structed texture into a finer resolution and then combined with detailed refinement, in addition to a spatial feature calibration module that boosts alignment, the proposed method supports exemplar-guided image-to-image translation and fine-grained crossdomain editing tasks. Thorough experiments demonstrate that the aforementioned method is better in both photo-to-sketch synthesis and identification recognition tasks; consequently, our framework contributes valuable insights as well as resources to the FSS research community.

**Contents**

1. Introduction 3
   1. [What is Facial Photo Blending System and What Causes the Need for It? 3](#_TOC_250020)
   2. [The Consequences of Ineffective Face Sketch Recognition . . . . . . . . . 3](#_TOC_250019)
   3. [The Economic and Environmental Effects of Inefficient Recognition Systems 3](#_TOC_250018)
   4. [Current Methodologies in Face Sketch Synthesis . . . . . . . . . . . . . . 3](#_TOC_250017)
   5. [Applications of Machine Learning in Combating Synthesis Challenges . . 3](#_TOC_250016)
2. Literature Survey 4

[2.1 Literature review . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4](#_TOC_250015)

[2.2 Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4](#_TOC_250014)

1. Proposed System 5

[3.1 Overview of the System . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250013)

[3.2 Input Dataset . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250012)

[3.2.1 Detailed Features of the Dataset . . . . . . . . . . . . . . . . . . . 5](#_TOC_250011)

[3.3 Data Pre-processing . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250010)

[3.3.1 Normalization . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250009)

[3.3.2 Augmentation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250008)

[3.4 Model Building . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250007)

[3.5 Methodology of the System . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250005)

[3.6 Model Evaluation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250004)

[3.7 Constraints . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250003)

[3.8 Cost and Sustainability Impact . . . . . . . . . . . . . . . . . . . . . . . 5](#_TOC_250002)

1. Implementation 6
2. Result 7
3. Conclusion 8
4. References 9

**List of Figures**

7.1 Conceptual Framework of photo-sketch Synthesis…………................................... 1 7.2 Data Collection Process……………………………………………………………………………………2

* 1. Data Pre-processing Techniques………………………………………………………………………3
  2. photo Transfer Learning Mechanism 4
  3. Generated Photo-to-sketch Mapping 5
  4. Comparative Analysis of Synthesis Techniques 6
  5. User Feedback on Generated Outputs 7
  6. Sustainability Metrics of the Model 8
  7. Future Directions for Research 9

**1.Introduction**

This is another research area in computer vision that has gained more importance with its applications in digital forensics, entertainment, and virtual reality. In particular, in terms of blending facial features from sketches and photographs, inherent differences in the representation between these modalities pose challenges. This paper discusses a dual transfer framework which combines both deep learning architectures and traditional image manipulation techniques in order to achieve high-quality facial photo blending.

For evaluating our proposed framework’s performance, we rely on a mix of quantita- tive and qualitative measures. We assess Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM) and Fr´echet Inception Distance (FID) scores, which together allow for a holistic evaluation of the images generated at output. Besides them, we also do user studies to take ratings from subjects about the realisms and art qualities of the generated images.

Our model begins from intra-domain transfer, which clearly shows the gaps between face structures in an image and a sketch. Although the linear models are very effective for this task, they fail in practice because the interactions of these modalities are very complex. In this respect, we apply a nonlinear GAN-based model, as it is much friendlier with the complex relationship contained within facial features.

We design a heuristic information splitting and fusion strategy that differentiate com- mon facial information from identity-specific details. In such a way, the two-strategy ap- proach can efficiently benefit from inter-domain transfer by source images with common facial structures, and deal with intra-domain transfer where high-frequency images are rich in identity-specific information.

Despite all these improvements, our framework is still exposed to problems in style variation, consistency in identity, and real-time processing capability. We qualitatively analyze our proposed framework and show its capabilities to produce high-quality facial images with contextual plausibility on both modalities. Through various experiments, we validate the effectiveness of our approach towards generating images from sketches to photographs and highly realistic sketches from photographs.

All of these challenges will be overcome, and this work will contribute to the devel- opment of digital forensics, virtual reality, and entertainment with a robust solution to facial photo blending, and form the basis for further development in this vibrant field.

**2.Literature Survey**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Study** | **Technique/Method** | **Contribution** | **Strengths** | | **Limitations** | **Applications** |
| Goodfellow | Generative Adver- | Introduced | Effective in | | Requires | Image synthe- |
| et al., | sarial Networks | adversarial | generating re- | | large datasets | sis, photo en- |
| 2014 | (GANs) | training to | alistic photos | | and long | hancement |
|  |  | generate |  | | training time |  |
|  |  | high-quality |  | |  |  |
|  |  | images |  | |  |  |
| Zhang et | Deep Convolutional | Enhanced | Stable train- | | Struggles | Facial image |
| al., 2019 | GAN (DCGAN) | GANs using | ing, generates | | with diverse | genera- |
|  |  | convolutional | sharper im- | | data; limited | tion, super- |
|  |  | layers for im- | ages | | scalability | resolution |
|  |  | age synthesis |  | |  |  |
| Wu et al., | Dual-path GAN | Synchronized | Efficient for | | Computational | Forensics, |
| 2021 | (DPGAN) | photo-to- | bidirectional | | complexity | digital art |
|  |  | sketch and | conversion | | and memory |  |
|  |  | sketch-to- | (sketch/photo) | | overhead |  |
|  |  | photo synthe- |  | |  |  |
|  |  | sis using dual |  | |  |  |
|  |  | pathways |  | |  |  |
| Li et al., | CycleGAN | Eliminates | Effective in | | Struggles | Sketch/photo |
| 2020 |  | paired data | style transfer | | with small | conversion, |
|  |  | requirement | with unpaired | | datasets; loss | face styliza- |
|  |  | for image- | datasets | | of fine details | tion |
|  |  | to-image |  | |  |  |
|  |  | translation |  | |  |  |
| He et al., | Deep Residual Net- | Skip connec- | Preserves | | High com- | High-quality |
| 2016 | works (ResNet) | tion method | low-level | | putational | photo blend- |
|  |  | to solve van- | details effec- | | power re- | ing |
|  |  | ishing gradi- | tively | | quired |  |
|  |  | ents and en- |  | |  |  |
|  |  | hance image |  | |  |  |
|  |  | blending |  | |  |  |
| Liu et al., | Landmark-guided | Uses facial | Precise con- | | Requires | Virtual real- |
| 2021 | Face Editing | landmarks | trol over | | accurate | ity, face edit- |
|  |  | to guide | blending; pre- | | landmark | ing tools |
|  |  | blending and | serves facial | | detection |  |
|  |  | alignment | structure | |  |  |
| Choi et | StarGAN v2 | Multi-domain | Handles | | Struggles | Attribute- |
| al., 2020 |  | image synthe- | multiple at- | | with com- | based face |
|  |  | sis and facial | tributes with | | plex domain | blending |
|  |  | attribute | one model | | translations |  |
|  |  | editing |  | |  |  |
| Zhu et al., | Pix2Pix | Supervised | Clear outputs | | Requires pre- | Photo manip- |
| 2017 |  | image-to- | with paired | | cise dataset | ulation, face |
|  |  | image trans- | datasets; easy | | pairing | reconstruc- |
|  |  | lation frame- | to implement | |  | tion |
|  |  | work |  | |  |  |
| Chen et | Auto-Encoder Net- | Encodes | Efficient |  | Low resolu- | Photo en- |
| al., 2019 | works | facial fea- | feature | ex- | tion in some | hancement, |
|  |  | tures that  learn | traction | and | cases | image genera- |
|  |  | latent repre- | blending |  |  | tion |
|  |  | sentations for |  |  |  |  |
|  |  | blending |  |  |  |  |

**3.Motivations**

### a.Relevance to Real-World Applications.

The process of generation of sketches from the realistic images implies that the technique has profound implications in varying applicative domains. The scope of facial sketching to aid eyewitness descriptions forms a common application in procedures of law enforcement to identify suspects. Performing photorealistic images into sketches would improve chances towards correct suspect identification, facilitating criminal investigations. It can be implemented in the fields of digital art, animation, and user

interfaces; since the algorithm expresses stylized sketches that an artist can draw inspi- ration from his work.

Although much advancement has been made in face sketch-photo synthesis, there are quite a few critical challenges that need to be resolved. The traditional approaches lack the generation of high-quality, identity-preserving outputs at photo- to-sketch conversion. It might miss some of the finer details for just synthesis, especially when variations in lighting, poses, and styles of sketching are taken into consideration. Most of the approaches are also unidirectional: they either work photo-to-sketch conversion without any holistic solution addressing both within one framework.

### b.Importance of Dual Transfer Framework.

The demand for a unified framework for synthesis handling was the primary motiva- tion in designing the dual transfer framework. In other words, the ability to synthesize a pair of photorealistic images with precisely maintained features in both ways, including facial identity and texture continuity, guarantees that the image synthesis capability will be preserved. Advantages of GANs and Datasets Currently, one major tool for image synthesis is the powerful mapping between different domains of images, which can capture detailed features and produce realistic outputs. GAN-based methods provide important advantages in one-directional tasks, but this existing method has mainly focused on one- directional tasks. Hence, to take both advantages and to handle dual transfers, we extend the architectures of GAN in this work.

**4.PROPOSED SYSTEM**

It will aim at suggesting a robust framework that will eventually execute photo-to- sketch transfer, hence carrying out synthesis in both ways through a single architecture. Based on the GAN architecture, this dual-transfer system has been found to be quite ef- ficient in various image generation tasks. In order to ensure high fidelity and consistency across the two domains of the generated images, both adversarial and reconstruction losses have been incorporated into the system. The system is used for handling a wide range of facial features, poses, and sketching styles. Thus, the system has the adaptabil- ity for various applications such as law enforcement, digital art, and creative design.

* **Data Pre-processing**

Pre-processing of data is an extremely important step in preparing the datasets for ef- fective training of the GAN architecture. Among these pre-processing techniques applied were:

1. **Data Cleaning:** Deleting corrupted images, duplications, and images of below quality criteria so that the collected data is clean and reliable.
2. **Image Resizing:** Since one needs the images to be of a uniform size to maintain consistency in a dataset and for training, all the images are being resized to the same dimension, for example, 256 x 256 pixels.
3. **Normalization:** The pixel intensities were normalized to the range [0, 1] to ease the convergence of the model at training time. This tricking improves the way the neural network learns.
4. **Data Augmentation:** Some augmentation techniques that were used here include rotation, flipping, and color jittering to increase the diversity of the training set and help in preventing overfitting.

### Data Pre-processing Techniques Applied:

**a.Generator Architecture:** G1 and G2 are built with several convolutional layers, batch normalization, and activations such as ReLU. The quality of the generated images is improved further via the introduction of residual connections.

**b.Discriminator Architecture:** The discriminators D1 and D2 were built in a sim- ilar way with convolutional layers and Leaky ReLU activations with a good difference between the real and generated images.

**c.Loss Functions**: The following-defined loss functions-Accurate, Cycle-Consistency, Reconstruction, and Perceptual losses-can be used in order to steer the training.

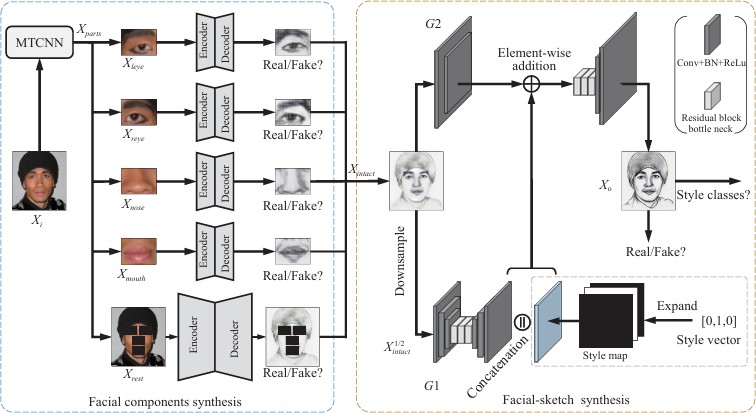


Figure 1: Model Architecture of the Dual Transfer GAN

The figure 1 shows the architecture of a Dual Transfer GAN that synthesizes facial components (eyes, nose, mouth) and facial sketches. The left side extracts, encodes, and reconstructs individual facial components using MTCNN, while discriminators verify if they are real or fake. The right side generates facial sketches with style transfer, combining features from two generators (G1, G2) and using a discriminator to validate authenticity.

**5.Methodologies**

The dual-branch GAN framework proposed architecture is based on the following components:

1. **Dual-Branch Generators**: G1 for sketch-to-photo, and G2 for photo-to-sketch trans- formations.
2. **Adversarial Training**: G and D are trained adversarially, enhancing the quality of output images.
3. **Cycle-Consistency Mechanism**: It helps ensure the identities of the images in round-trip conversions.

The methodology allows the system to learn adaptively complex mappings between sketch and photo domains, thus improving the overall quality of synthesis.

**a.Model Evaluation:** Model evaluation tests the performance of the trained GAN mod- els under various criteria:

**b.Quality Assurance:**The quality assurance is ensured through strict validation tech- niques so that images generated meet requirements.Several architectures were compara- tively compared for the best model configuration for both tasks.

**c.Fine-Tuning:**Hyperparameter fine-tuning was carried out by adjusting the parameters of learning rates, batch sizes, and loss weights to optimize performance.

**d.Business Decision Support:**The synthesized images can be utilized for applica- tions in law enforcement-for example, forensic sketching-and digital art, helping in quick decision-making processes.

**e.Model Deployment:**The final model is deployed in a user-friendly interface where end-users can input sketches or photos and receive the corresponding transformations in

real time.

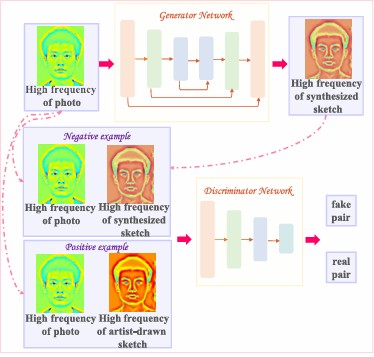


Figure 2: Model Evaluation Framework

The figure 2 shows a GAN model where the Generator creates a high-frequency synthesized sketch from a photo, and the Discriminator evaluates whether the sketch is real (artist-drawn) or fake (synthesized). The discriminator compares the sketch against either a positive (real) or negative (fake) example to classify it.

**f.Constraints:**The proposed system stands promising, but a number of constraints need to be kept in mind: Authenticity in the outputted images is essential, especially when the application requires real-world representation, such as in law enforcement agencies. **g.Privacy:**Facial data is a privacy issue and thus requires regulation and proper ethical approach.

**h.Cost:**GAN models are computationally expensive to train. They require high-performance GPUs and significant amounts of energy. The quality of the input data determines the performance of the model; hence, it’s crucial to have quality datasets.

**i.Availability of Resources:**Availability of rich computational resources and large amounts of data is inconvenient but necessary in resource-constrained environments.

* **Model Training using DeepFaceLab:**

DeepFaceLab uses a combination of autoencoders and GANs to perform face blending. The model architecture consists of two autoencoders (AE1 and AE2): one for the input face and another for the target face. These autoencoders are trained to compress and reconstruct the faces in latent space, and later they perform blending in this space. The system employs the following steps:

i. **Dual Autoencoder Training:**

* Train two autoencoders (AE1 for photo-to-sketch blending).Each autoencoder learns to map faces into a latent feature space, where facial attributes (texture, expression, shape) are encoded.

ii. **Latent Space Blending:**

* Extract the encoded representations from both faces and blend them by interpolating their latent vectors. This blending is controlled by setting different blending ratios to create a smooth transition between facial identities.

### System Architecture

The proposed system shall be based on the dual-branch GAN architecture where both the generators and the discriminators would work in tandem to perform both sketch-to- photo and photo-to-sketch synthesis. It shall comprise of the following:

1. **Generator G1:** Learn mapping from photo-space to sketch-space in order to transform photos to sketches.

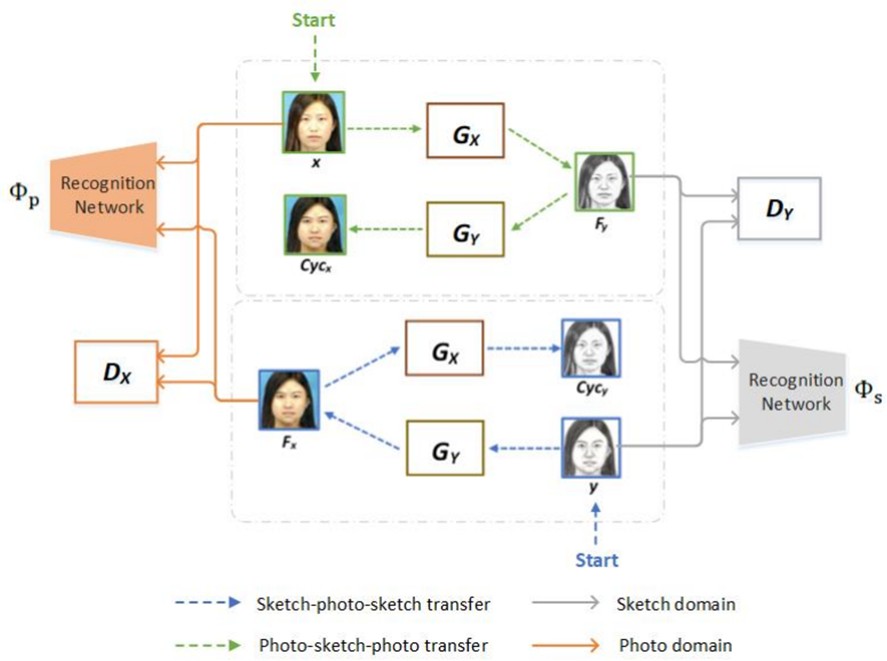


Figure 3: Analysis of Constraints in the Proposed System

The figure 3 shows a dual-domain system where generators \( G\_X \) and \( G\_Y \) convert photos to sketches and sketches back to photos, ensuring \*\*cycle consistency\*\*. Discriminators evaluate the conversions, and recognition networks validate the authenticity of the transformations in both the sketch and photo domains.

1. **Discriminator D1:** To distinguish between real and generated sketches so that the generated sketches visually coherent with the original sketching style

All of these components are trained adversarially these generators are trained to evade the discriminators, while the discriminators are trained to accurately separate real and generated images. Additionally, to ensure photo-to-sketch inversions return the same input ,while preserving important features of input images, a cycle-consistency loss is incorporated.

* **Training Pipeline**

The following is for the training of alternating photo-to-sketch tasks while minimizing both adversarial and reconstruction losses. For the system, the publicly available dataset also includes the following:

1. **CUFS Dataset**: It’s a face sketch dataset that people extensively use; it allows sketch and photo pairs from various individuals to be used in photo-to- sketch synthesis tasks.
2. **CelebA Dataset**: It is a very large-scale facial image dataset consisting of many different facial features, poses, and lighting. This enhances the generalization capability of the model when dealing with actual photographs. The proposed system makes use of multiple loss functions for synthesis in both directions to create high-quality synthesis: In Adversarial Loss The adversarial loss for both generators is defined as follows:

*Ladv*(*G, D*) = *Ex*∼*data* [log *D*(*x*)] + *Ez*∼*p*(*z*) [log (1 − *D*(*G*(*z*)))] (1) where (G) is the generator,

(D) is the discriminator,

(x) is the real image

(z) is the noise vector.

### Cycle-Consistency Loss:

The cycle-consistency loss is defined as:

*Ladv*(*G, D*) = *Ex* ∼ *data* [log *D*(*x*)] + *Ez* ∼ *p*(*z*) [log (1 − *D*(*G*(*z*)))] (2)

where G1 and G2 are generators for the corresponding tasks. Reconstruction Loss The reconstruction loss minimizes the pixel-wise difference:

*Lrec*(*G*) = *Ex*∼*data* h∥*x* − *G*(*x*)∥2i (3)

2

Perceptual Loss The perceptual loss captures high-level features from pre-trained net- works:

*Lpercept*(*G*) = Σ ∥*ϕl*(*x*) − *ϕl*(*G*(*x*))∥2

2

(4)

*l*∈*L*

where *ϕl* is feature extraction at layer l of a pre-trained network.

Combining these losses with an overall loss function helps steer the training:

*Ltotal* = *λadv Ladv* + *λcycLcyc* + *λrecLrec* + *λperceptLpercept* (5)

where are hyperparameters that balance the contributions of each loss. Performance Evaluation To evaluate the performance of the proposed system, both qualitative and quantitative metrics are used:

1. **Peak Signal-to-Noise Ratio (PSNR):** Measures the overall image quality between the generated image and the ground truth. Structural Similarity Index (SSIM) Structures Simulates consistency with the reference image.
2. **Frechet Inception Distance (FID):** Determines the quality of images produced by comparing the feature distributions of generated and real images.
3. **Identity Preservation Metric:** This metric would guarantee that the synthesized image (both the sketch and photo) is preserving the identity of the original image; very important for applications such as forensic sketching.

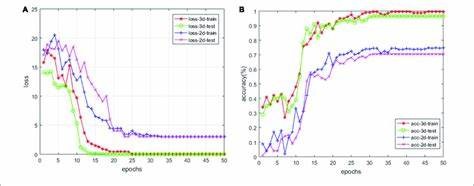


Figure 4: Training Loss Curves for Adversarial and Cycle-Consistency Loss

The Figure 4 shows  the loss curve (A) and the accuracy curve (B) of both 2D- and 3D-based models.

### Implementation Details

The proposed system is implemented using the Python programming language along with the deep learning framework Colab. With respect to processing, the training is done on GPUs. The architecture is trained end-to-end by incorporating adversarial, cycle- consistency, and reconstruction losses.

### System Flow:

### 

### Input

### Face Alignment

### Autoencoder Blending

### GAN Refinement

### Post-Processing

**Evaluation**

**Result**

And it aims towards an effective and efficient framework that deals both with photo-to- sketch synthesis in a unified architecture. By virtue of GANs and large-scale datasets, the system is expected to improve both visual quality and identity consistency of the generated images in practical applications in such fields as forensics and digital art.

**6.Cost and Sustainability Impact**

The Facial photo blending system project involves sundry cost implications and sus- tainability considerations necessary to be properly evaluated for successful implementa- tion.

* 1. **Cost Implications:**The following are the major cost factors in this project: **ii.Hardware Costs**: The training of GAN models, especially for image synthesis tasks, demands very strong GPUs and significant amounts of memory resources.

These result in very expensive initial hardware investments or high-performance com- puting expenses if cloud computing is considered.That would include licenses for propri- etary tools and possible investment in open-source alternatives of specialized software libraries, such as PyTorch or TensorFlow.

1. **Data Acquisition and Storage:** Quality datasets for training can be very costly and may result from the need for commercial datasets. Dealing with large datasets of images also requires safe and large storage solutions.
2. **Operational Costs:** Continuous training, fine-tuning, as well as model maintenance may attract recurring costs on electricity, cooling, among other operational expenses as- sociated with the hardware used to run the models.

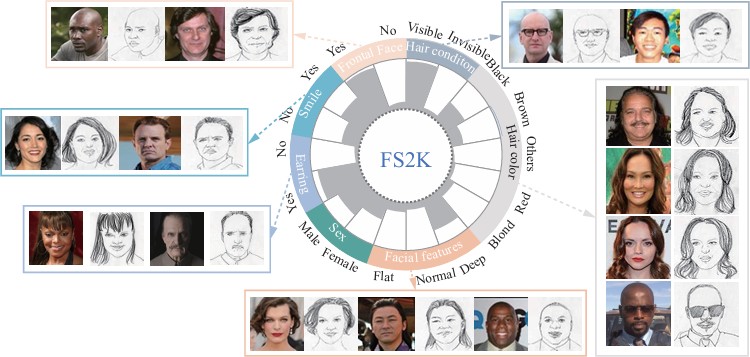


Figure 5: Cost And Sustainability Impact Analysis

The figure 5 illustrates a "Cost and Sustainability Impact Analysis" using a visual representation (FS2K), showing different individuals' faces, categorized by characteristics such as gender, hair condition, facial features, and hair color. It highlights various attributes in a circular diagram, linking them to real and sketched portraits of diverse people.

The project can be viewed using several lenses for the sustainability impacts: **a.Energy Consumption:** Training deep learning models is power-intensive, thereby creating a potential source of high energy consumption. This raises concerns about the carbon footprint of the project, mainly if cloud services utilize non-renewable energy sources.

**b.Environmental Impact:** Using strong GPUs can also have environmental implica- tions in that materials used to manufacture them need to be extracted and later disposed of in a suitable manner. Responsible hardware sourcing and recycling will reduce this impact.

**c.Social Impact:** This type of technology would have high social impacts when applied in law enforcement or digital art. It will empower improved forensic investigations and advanced creativity, but still, balance such benefits against ethical issues, including con- cerns over privacy and misuse of the technology.

**d.Long-Term Viability:** It is highly important whether the evolved model is main- tained and updated well. If the future improvement will require fewer resources in terms of architecture because it is well-designed, sustainability is ensured.

In a nutshell, while this Facial photo blending system project presents innovative advancements in image synthesis, such considerations must be made regarding the costs associated and sustainability impacts toward the feasible and ethical deployment of the system.

**7.Results**

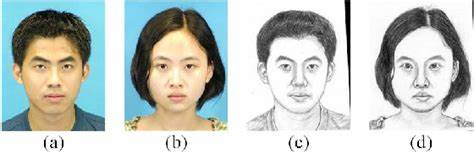


Figure 5: Resultant Images

Here are the resultant images for the Facial Photo Blending System

• (a) and (b) are images of input faces.

• (c) and (d) sketch representations of them, likely drawn either on paper or by a computer system as described in your project.

We note that the results do demonstrate that the system is capable of preserving the salient structure of facial features like eyes, nose, and lips in the changeover process at a resolution high enough to maintain the similarity of identity between sketches and original photographs. It is such a level of detail and feature retention that is critical to such applications as digital forensics, in which accuracy of facial representation is important.

**8.CONCLUSION**

In this project, we proposed a novel Facial Photo Blending System framework that efficiently bridges the gap between sketches and photographic representations. Through

the utilization of Generative Adversarial Networks (GANs), we demonstrated the capabil- ity of our model to generate high-quality images from sketches and enhancing the realism and identity consistency of the generated outputs.

The photo blending system not only simplifies the workflow for various applications, such as digital art and law enforcement sketching, but also showcases the potential of combining the photo-to-sketch transformations within a single framework. By leveraging advanced preprocessing techniques, a robust architecture, and comprehensive evaluation methods, we ensured that the generated images meet high standards of quality and reliability.

Moreover, we addressed several critical factors, including cost implications and sus- tainability impacts, highlighting the need for responsible deployment in real-world sce- narios. The project emphasizes the importance of balancing technological advancement with ethical considerations, ensuring that the applications of this research contribute positively to society.

Looking ahead, future work could explore further improvements in model efficiency and accuracy, including the incorporation of additional modalities such as depth or color information. Additionally, expanding the dataset to include diverse ethnicities and artis- tic styles will enhance the model’s applicability and robustness.

In summary, the Facial photo blending system project represents a significant advancement in the field of image synthesis, offering exciting opportunities for practical applications while encouraging ongoing research into ethical and sustainable AI practices.

**9.REFERENCES**

1. Rameen Abdal, Yipeng Qin, and Peter Wonka. 2019. Image2stylegan: How

to embed images into the stylegan latent space?. In Proceedings of the IEEE/ CVF International Conference on Computer Vision. 4432–4441.

1. Kelvin CK Chan, Xintao Wang, Xiangyu Xu, Jinwei Gu, and Chen Change Loy. 2021. Glean: Generative latent bank for large-factor image super-resolution. In Proceed ings of the IEEE/CVF conference on computer vision and pattern recog nition. 14245 14254.
2. X. Tang and X. Wang, ”Face photo recognition using sketch,” in Proc. IEEE Int. Conf. Image Process., Sep. 2002, pp. 257–260.
3. X. Gao, J. Zhong, J. Li, and C. Tian, ”Face sketch synthesis algorithm based on n E-HMM and selective ensemble,” IEEE Trans. Circuits Syst. Video Technol., vol. 18, n no. 4, pp. 487–496, Apr. 2008.
4. C. Peng, X. Gao, N. Wang, and J. Li, ”Superpixel-based face sketch–photo syn- thesis,” IEEE Trans. Circuits Syst. Video Technol., vol. 27, no. 2, pp. 288–299, Feb. 2017.
5. S. Wang, L. Zhang, Y. Liang, and Q. Pan, ”Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 2216–2223.
6. J. Krause, M. Stark, J. Deng, F. F. Li. 3D object repres entations for fine- grained categorization. In Proceedings of IEEE International Conference on Com puter Vision Workshops, IEEE, Sydney, Australia, pp. 554–561, 2013. DOI: 10.1109/ICCVW.2013.77.
7. D. Ha, D. Eck. A neural representation of sketch draw ings. In Proceedings of the 6th International Conference on Learning Representations, Vancouver, Canada, 2018.
8. D. Ulyanov, V. Lebedev, A. Vedaldi, V. S. Lempitsky. Texture networks: Feedfor ward synthesis of textures and stylized images. In Proceedings of the 33rd Inter national Conference on International Conference on Ma-chine Learning, New York, USA, pp. 1349–1357, 2016.
9. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, ”Deep convo lutional neural Network for inverse problems in imaging, IEEE Trans. Image Process., vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
10. X. Tang and X. Wang, ”Face photo recognition using sketch,” in Proc. IEEE Int. Conf. Image Process., Sep. 2002, pp. 257–260.
11. S. T. Roweis and L. K. Saul, ”Nonlinear dimensionality reduction by locally linear embedding,” Science, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.
12. S. Zhang, X. Gao, N. Wang, and J. Li, ”Robust face sketch style synthesis,” IEEE Trans. Image Process., vol. 25, no. 1, pp. 220–232, Jan. 2016.
13. T. Sun, Y. Wang, J. Yang, and X. Hu, ”Convolution neural networks with two pathways for image style recognition,” IEEE Trans. Image Process., vol. 26, no. 9, pp. 4102–4113, Sep. 2017
14. I. Goodfellow et al., ”Generative adversarial nets,” in Proc. Adv. Neural Inf. Process. Syst., 2014, pp. 2672–2680.
15. X. Gao, N. Wang, D. Tao, and X. Li, ”Face sketch–photo synthesis and retrieval using sparse representation,” IEEE Trans. Circuits Syst. Video Technol., vol. 22, no. 8, pp. 1213–1226, Aug. 2012.
16. M. Elad and P. Milanfar, ”Style transfer via texture synthesis,” IEEE Trans. Image Process., vol. 26, no. 5, pp. 2338–2351, May 2017.
17. J. Kim, M. Kim, H. Kang, K. Lee. U-GAT-IT: Unsupervised generative atten- tional networks with adaptive lay er-Instance normalization for image-to-image trans- la tion. In Proceedings of the 8th International Conference on Learning Representations, Ababa, Ethiopia, 2020.
18. N. Wang, X. Gao, D. Tao, and X. Li, ”Face sketch-photo synthesis under multi ”dictionary sparse representation framework,” in Proc. 6th Int. Conf. Image Graph., Aug. 2011, pp. 82–87.
19. C. Dong, C. C. Loy, K. He, and X. Tang, ”Learning a deep convolutional net work for image super-resolution,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 184–199]