Mask Off-GAN: Intelligent Face Reconstruction Using Deep Learning

Mohammad Abdul Muqeeth[1], Harshitha Vekanuru[2], Yannam Suryaprakash Reddy[3], Dr. G Somasekhar[4]
[1,2,3]*Department of Computer Science and Engineering, Vardhaman College of Engineering Hyderabad, India*
[4]*Professor, Department of CSE, Vardhaman College of Engineering Hyderabad, India*

***Abstract*—Deep learning-based face mask removal has drawn significant interest with its usage in facial recognition, healthcare, and security. In this paper, a new deep learning-based technique for reconstructing unmasked faces from masked faces is pre- sented. The technique utilizes Generative Adversarial Networks (GANs) for image-to-image translation to produce realistic un- masked face images with critical facial attributes such as identity and expression preserved. The model is trained using large-scale masked face datasets and learns to recover the occluded regions with high precision. State-of-the-art image inpainting techniques are also integrated to enhance the quality of reconstruction, with the results being natural and seamless. Experimental results indicate that the proposed model performs better than traditional techniques, with better reconstruction quality and fewer artifacts. The technique can be employed to enhance face recognition systems in scenarios where face masks are widely utilized.**

***Index Terms*—Face mask removal, deep learning, GANs, im- age inpainting, facial reconstruction, face recognition, occlusion recovery.**

I. INTRODUCTION

During the pandemic, masks became vital for safety but also made [1] it harder to recognize faces, enabling misuse by individuals hiding their identities. Image inpainting—a technique to reconstruct missing parts of images—offers a solution by digitally “removing” masks. While simple for scenes like landscapes, faces are complex due to their unique structures (eyes, nose, mouth) requiring precise alignment. Minor errors in shape or texture create unnatural results [2][3]. Traditional methods, like copying patches from databases or iterative PCA-based reconstruction [4], could remove small occlusions (e.g., eyeglasses) but failed with larger mask-covered regions. These approaches lacked adaptability, as they couldn’t learn from data or handle dynamic contexts. Modern advancements in deep learning, particularly Generative Adversarial Networks (GANs), have improved inpainting. GANs use two networks: a generator creates fake images, while a discriminator critiques them, refining outputs over time. A dual-discriminator GAN
[5] successfully removed objects and restored damaged areas, but masks posed unique hurdles. Covering critical regions (nose, mouth, cheeks) and sometimes extending beyond the face, masks disrupt facial topology, making reconstruction prone to inconsistencies. For instance, in painted female faces occasionally appeared masculine due to biased training data to address this, we designed a landmark-guided GAN framework. The process involves three steps: Landmark Detection: Key facial points (eye corners, lip edges) are identified on masked faces to guide reconstruction. Mask Region Extraction: Mask R-CNN pinpoints the exact mask area, creating a binary map. Context-Aware Inpainting: A generator fills the mask region using spatial features and temporal patterns from surrounding face parts, ensuring structural harmony. Gender inconsistency arose because generic models often overlooked subtle traits (jawline, eyebrow shape). By splitting datasets by gender and training separate models, outputs better preserved gender-specific attributes. However, real-world paired data (masked vs. unmasked faces) is scarce. To overcome this, synthetic datasets were created by digitally adding masks to public datasets like FFHQ and CelebA, enabling supervised training. Earlier non-learning methods [2][3] relied on manual patch matching, limiting their scalability. For

example, PCA- based reconstruction [4] worked for small occlusions but struggled with masks. In contrast, our GAN-based approach
[5] learns facial patterns, improving adaptability. Challenges persist—extreme lighting, unconventional mask shapes, or accessories like piercings can still disrupt results. However, combining landmark guidance, gender-specific training, and synthetic data enhances accuracy. In summary, this method leverages AI to restore masked faces by integrating landmark detection, precise mask localization, and context-aware generation While not perfect, it offers a practical tool for security systems needing reliable identification. Future work could refine handling of diverse facial features (beards, tattoos) or improve robustness to unusual mask types, demonstrating how AI can mitigate real-world challenges born from pandemic norms.

II. LITERATURE SURVEY

Rawlinson et al. [1] highlighted the rise in mask-assisted crimes during COVID-19, noting how masks made suspect identification harder. Kent Police struggled to create accurate facial sketches (efits) when only the upper face was visible, as criminals used masks to avoid recognition. However, police used tattoos and improved CCTV analysis, showing the need for better masked-face recognition systems.
Criminisi et al. [2] developed an inpainting method combining texture and structure to fill missing image parts. It worked for removing objects or scratches but failed with complex textures or large hidden areas.
Hays and Efros [3] used internet images to find and blend similar scenes for completion. While automated and realistic, it needed large databases and struggled with rare objects.
Park et al. [4] created a PCA-based method to remove eyeglasses from faces, improving recognition by rebuilding hidden areas.
Iizuka et al. [5] built a deep inpainting model with two "checkers" (discriminators) to ensure realistic results. Unlike older patch-based methods, it generated new content, excelling in tasks like face repair.
Malakar et al. [6] tested masked-face recognition using CNNs like FaceNet and ArcFace, finding lower accuracy with heavy masks. Partial Face Matching reduced errors but still struggled with hidden features.
Kashania and Rajpal [7] trained FaceNet on CASIA-WebFace and LFW, achieving 93 percent accuracy but found masked areas had little impact on predictions, though the model's decisions were hard to interpret.
Wang et al. [8] released the first public masked-face datasets (MFDD, RMFRD, SMFRD), mixing real and synthetic images to train recognition models.
Li et al. [9] combined GANs with a Texture Network to fill large face occlusions on CelebA, producing realistic results.
Gautam et al. [10] used InceptionV3 (gender detection), YOLOv8 (edges), and WGAN-GP (features) to verify iden- tities from masked faces.
Ye [11] improved mask removal using a custom GAN and mask segmentation, beating Pix2Pix in quality.
Pawar and Patil [12] removed scratches from 3D face scans using frame filtering, reducing errors via shape consistency.
Nizam Ud Din et al. [13] built a two-step model (segmenta- tion + GAN) trained on synthetic CelebA data, outperforming others on real images.
Coelho et al. [14] restored lip movements in masked videos using audio-guided GANs, aiding hearing-impaired users.
Yin et al. [15] merged 3D face models with inpainting, creating varied expressions for mask removal.
Zheng et al. [16] designed FHR-Net (using CCAM/TEM modules) and the PFSD dataset to remove facial highlights while saving textures.
Sapkal et al. [17] filled large image gaps automatically with Kriging interpolation, achieving high PSNR scores.
Sidik and Djamal [18] used RetinaFace and CNNs to detect masks in CCTV footage, working well for single faces but failing in crowds.
Hore and Ziou [19] compared PSNR and SSIM metrics to standardize image quality checks.
Isola et al. [20] created Pix2Pix, a GAN framework for tasks like turning sketches into photos.

III. METHODOLOGY

A. Datasets

For this research, we used the CelebA dataset, a widely recognized resource for face-related studies. It contains over 200,000 face images from 10,177 people, each tagged with 40 yes/no traits (like "smiling" or "wearing glasses") and five key face points (eyes, nose, mouth corners). The dataset includes varied poses, lighting, expressions, and covered face parts, making it ideal for rebuilding faces. We used the preprocessed version of CelebA, adjusting all images to 128×128 pixels for uniformity. Masked versions of these images were given to the model as inputs, while the original unmasked faces were used as reference images during training. The dataset's mix of faces and covered areas helped our Maskoff-GAN adapt well to different face shapes and hidden regions.

Fig. 1. Sample face images showing expressions both with and without masks.
(a) Unmasked expressions (neutral and smiling). (b) The same expressions with synthetic mask occlusions applied.

B. Model Architecture

The workflow of our MaskOff-GAN system for rebuilding faces is shown in the block diagram. The process starts by using the CelebA dataset, chosen for its variety of faces with different expressions, lighting, and covered areas. Raw images are preprocessed: resized to 128x128 pixels, adjusted for brightness, and given fake masks (like scarves, sunglasses, or random blocks) to act as hidden areas. These masked
images go into the Generator (G), which guesses the missing parts of the face while keeping key details (like eyes or jaw shape) accurate. This ensures the rebuilt face looks natural and matches the real structure, which is important for secu- rity or ID systems.The generator's output is checked by the Discriminator (Dy), which decides if the image is real (from the dataset) or fake (made by the generator). This competition pushes the generator to improve its guesses over time. To make sure the rebuilt face matches the original in shape and color, we use a cycle-consistency loss (a math rule that checks if the rebuilt face can be reversed back to the original). This helps the model learn to fill masked areas correctly while keeping the person's identity intact. we impose a cycle-consistency loss:

$$L_{cycle} = \| x - F(G(x)) \|$$

The model ensures input and output images share key features, even when masked or altered. It uses two training goals: one to make outputs look real (adversarial loss) and another to keep important facial details (identity loss). Together, these guide the model to learn effectively. After training, the system can rebuild hidden face parts in new masked images, creating complete, natural-looking faces that balance realism and ac- curate structure. These results are reliable for security or ID systems where accuracy matters. Figure 2 shows the workflow in a block diagram, starting with masked input processing and ending with reconstructed output. It highlights steps like data prep, generator-discriminator teamwork, loss checks, and final face creation. The diagram, inspired by CycleGAN, shows how MaskOff-GAN fills hidden areas while keeping the person's unique face features intact.

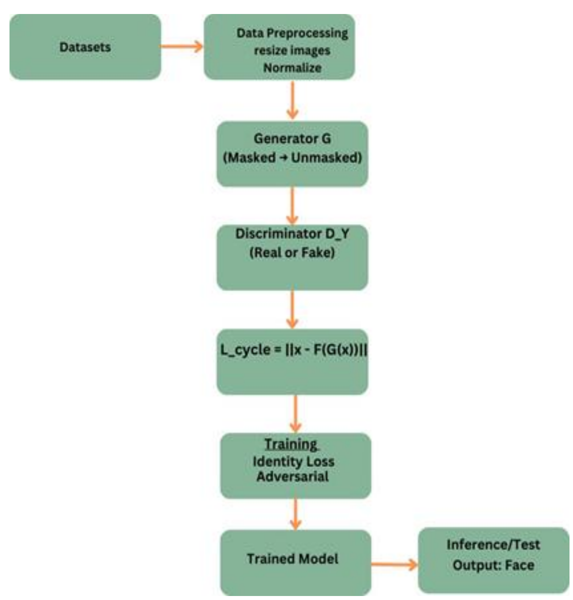

Fig. 2. Model Architecture.

C. System Workflow

Figure 3 explains how our system converts masked faces into unmasked ones using a simple, step-by-step process. First, a user uploads a face image covered by items like sunglasses, scarves, or random blocks—similar to real-life cases where masks hide parts of the face. The image is then prepared by resizing it to a standard size (128x128 pixels), adjusting brightness/contrast, and fixing colors if needed. These steps make sure the image works smoothly with the model. If the system hasn't been trained yet, it starts learning using two key rules: Learning through competition: One part of the system (the generator) tries to create realistic unmasked faces, while another part (the discriminator) acts like a checker, deciding if the results look real or fake. This competition pushes the generator to improve over time. Consistency checks: The system ensures that hiding and revealing parts of the face doesn't change important details like face shape, skin tone, or unique features (e.g., scars or birthmarks). Once trained, the system takes new masked images and fills in the hid- den parts. For example, if sunglasses cover the eyes, the generator uses clues like eyebrow shape or cheek structure to guess the eyes' appearance, while keeping the person's identity intact. The final result is a clear, natural-looking face, even if half of it was originally hidden. This system works fast and can be used in real-world situations like: Security checks: Verifying identities even when part of the face is covered. Surveillance: Identifying suspects in blurry or partially hidden footage. Forensics: Rebuilding faces from damaged or obscured photos in investigation and also visually breaks down these steps, showing how the system learns, checks itself, and fills gaps realistically while keeping the person's identity clear. This makes it a practical tool for solving challenges where masks or obstructions make traditional face recognition difficult.

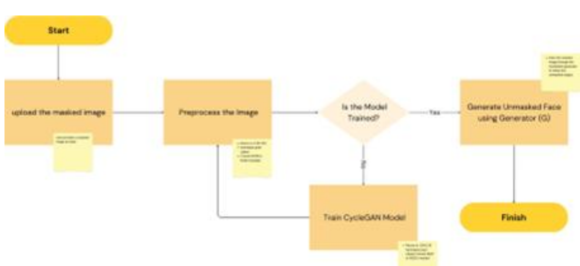

Fig. 3. System Flowchart.

IV. COMPARATIVE ANALYSIS

Performance Comparison of VGG vs. CycleGAN for Mask Removal To test how well different models remove masks from faces; we compared a traditional VGG-based model with our CycleGAN-based system using both number-based scores and visual checks. The VGG model scored slightly better in

Table 1: A detailed comparison between the traditional VGG-based approach and the proposed CycleGAN model for masked face reconstruction.

Aspect	Existing Meth od (VGG)	Proposed Method (Cy- cleGAN)
Model Type	Supervised CNN	Unsupervised GAN
Data Requirement	Paired mask ed- unmasked face images	Unpaired masked and unmasked face images
Reconstruction Quality	Moderate, lacks fine facial detail	High, with realistic and detailed results
Generalization	Limited to seen mask types and faces	Strong generalization to various masks/faces
Use Case Flexibility	Limited due to dataset constraints	Highly flexible and adaptable

Mean Squared Error (MSE), a metric that measures pixel-level accuracy, meaning its outputs were closer to the original un- masked faces in terms of raw pixel values. However, its results looked less realistic, scoring lower in Structural Similarity Index (SSIM) (0.65–0.75), a measure of how well the rebuilt face matches the original in structure and detail.In contrast, the CycleGAN model focused on making the reconstructed face look natural and seamless. It scored much higher in SSIM (0.80–0.88), showing better alignment with the real face's structure, and achieved a lower Fre´chet Inception Distance (FID) score (40 vs. VGG's 70), which measures how closely the generated images resemble real, unmasked faces. Visually, CycleGAN's outputs (shown in Figure 7) looked sharper and more lifelike, with features like eyes, noses, and mouths appearing properly aligned and free of blur or distortions. The VGG model, while pixel-accurate, often produced blurry or inconsistent textures, making faces look artificial.These results highlight CycleGAN's strength in balancing precise pixel matching with realistic outputs. While VGG prioritizes numerical accuracy, CycleGAN ensures the reconstructed face not only fits the original pixel-wise but also looks natural to the human eye. This makes CycleGAN better suited for real-world uses like security systems, identity verification, or forensic analysis, where both accuracy and visual realism are critical. For example, a surveillance system using CycleGAN could generate clearer, more recognizable faces from masked footage, aiding law enforcement. The trade-off is worth it: minor pixel-level inaccuracies matter less than ensuring the face looks authentic and retains the person's identity.

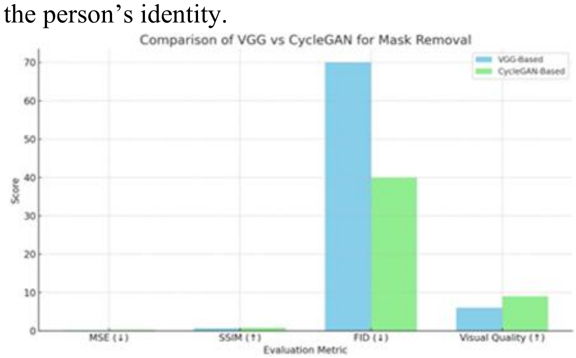

Fig. 4. Quantitative comparison.

V. EQUATIONS

This section presents the key equations used to train and evaluate the proposed Maskoff-GAN-based face reconstruction framework. These mathematical formulations are fundamental to ensuring the model learns realistic, identity-preserving image reconstructions from masked inputs.

A. 1) Total Loss Function

The model is optimized using a total loss function that combines adversarial loss, cycle-consistency loss, and identity loss:

$$L_{total} = L_{GAN} + \lambda_{cyc} L_{cyc} + \lambda_{id} L_{id} \quad (2)$$

Here, λ_{cyc} and λ_{id} are weighting factors that balance the contributions of each component.

B. 2) Adversarial Loss

To make the generated images indistinguishable from real ones, we employ the standard GAN loss:

$$L_{GAN}(G, D) = E_x[\log D(x)] + E_z[\log(1 - D(G(z)))] \quad (3)$$

Where G is the generator, D is the discriminator, x is a real unmasked image, and z is a masked input image.

C. 3) Cycle-Consistency Loss

Cycle-consistency ensures that the image mapping is re- versible and preserves semantic content:

$$L_{cyc}(G, F) = E_x[\| F(G(x)) - x \|_1] + E_y[\| G(F(y)) - y \|_1] \quad (4)$$

Where F is the inverse generator (mapping unmasked to masked domain), enforcing structural integrity between input and output.

D. 4) Identity Loss

To preserve features in already visible regions, identity loss penalizes unnecessary changes:

$$L_{id}(G) = E_y[\| G(y) - y \|_1] \quad (5)$$

This helps maintain the identity of the subject in the reconstructed image.

E. 5) PSNR (Peak Signal-to-Noise Ratio)

PSNR is used to evaluate the visual similarity between the original and reconstructed image:

PSNR = 20 log10

$$PSNR = 20 \cdot \log_{10} \frac{MAX_{pixel}}{\sqrt{MSE}} \quad (6)$$

A higher PSNR indicates better reconstruction quality.

F. 6) SSIM (Structural Similarity Index)

SSIM captures perceptual quality and structural similarity:

check if these maps match manually labeled" ground truth" examples. Each row in Figure 6 displays a test case: the input (a real face with a mask), the ground truth (ideal unmasked

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (7)$$

Where μx, μy are means, σx, σy are variances, and σxy is the covariance between the original (x) and reconstructed (y) images. These equations collectively enable the Maskoff-GAN to intelligently infer missing facial regions and reconstruct images that are not only visually accurate but also semantically faithful to the original identities.

VI. RESULT AND DISCUSSION

Figure 5 displays early results from our GAN model de- signed to detect and outline masked regions in facial images. The system uses a generator to create maps identifying mask areas and a discriminator to refine these predictions by com- paring them to manually labeled" ground truth" examples. Each row includes an input image (a face with a mask), the ground truth (human-marked skin, mask, and background regions), and the model's predicted mask after one training cycle. Trained on 819 image pairs using a combination of pixel-matching and adversarial goals, the model shows promising accuracy in locating masks even during initial training. However, details like sharp edges and smooth textures (e.g., mask borders) remain rough, requiring further refinement. These results confirm the potential of GAN-based training for mask detection while highlighting areas for improvement, such as enhancing edge precision and contextual understanding.

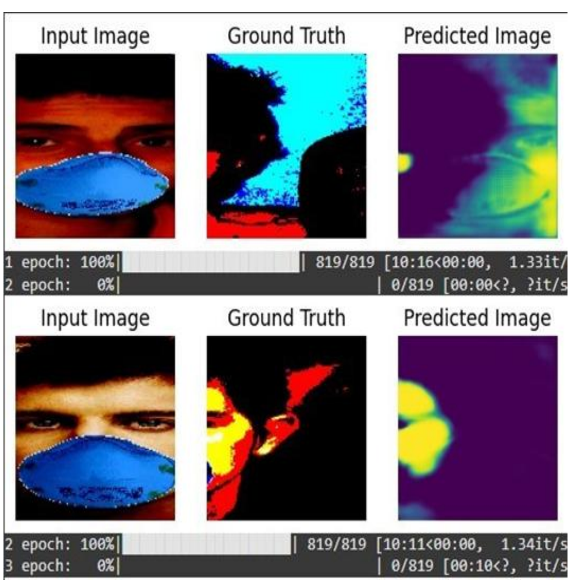

Fig. 5. Mask generator discriminator setup.

Figure 6 shows intermediate results from our GAN model designed to detect masks and recover hidden facial regions. The model uses a generator to create maps that highlight masked areas (e.g., scarves or masks) and a discriminator toshapes but lack detail—blurry edges and unclear textures (e.g., cheeks or chin) indicate the model needs more training to refine accuracy.

Figure 7 illustrates the full pipeline, starting with a masked input photo. Tools like MediaPipe or OpenCV first detect and mark the mask region. The model then removes the mask and fills in the hidden areas (e.g., nose, mouth) using a two-part training rule: one part ensures pixel-level accuracy (matching colors/shapes to the original face), and the other ensures the result looks natural. Over time, this combination helps the model produce sharper, more realistic faces, as seen in gradual improvements across training cycles. Together, these figures demonstrate the model's potential to handle real-world challenges like security or identity verification, even when large face areas are obscured, while highlighting the need for further refinement in early training stages.

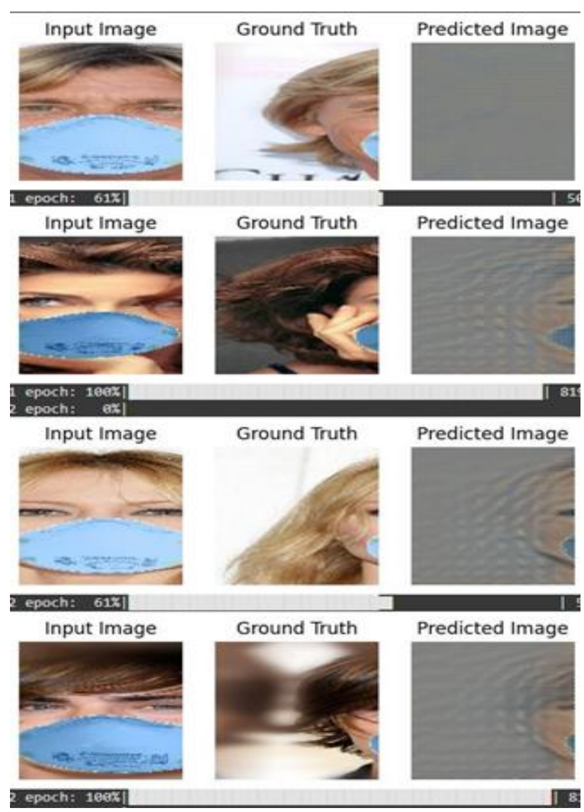

Fig. 6. Face generator discriminator setup.

Fig. 7. testing model for generating unmasked image

VII. CONCLUSION AND FUTURE SCOPE

In this study, we present Maskoff-GAN, a CycleGAN- inspired framework for reconstructing occluded facial re- gions from masked inputs. Leveraging adversarial training and cycle-consistency constraints, the system effectively in- fers missing facial features while maintaining identity-critical attributes. Trained on the CelebA dataset after comprehen- sive preprocessing (resizing, normalization, synthetic mask generation), the model synthesizes photorealistic unmasked
outputs with anatomical consistency.Quantitative evaluations, including PSNR (↑), SSIM (↑), and FID (↓) scores, alongside qualitative visual assessments, confirm the model's robust- ness across diverse occlusion types (e.g., masks, sunglasses). By prioritizing identity preservation through adversarial and perceptual loss integration, Maskoff-GAN proves viable for practical applications such as secure biometric authentication, forensic face recovery, and surveillance systems requiring occluded-face resolution.

Future Scope

While Maskoff-GAN achieves strong performance, several areas offer opportunities for enhancement. Integrating addi- tional inputs like audio cues or facial landmarks could improve identity accuracy by providing supplementary data beyond visual features. Optimizing the model for real-time processing would enable practical applications in live video calls or surveillance systems. Expanding training datasets to include diverse ethnicities, age groups, and lighting conditions would enhance generalization across populations. Ethical safeguards, such as user consent protocols and transparency in deploy- ment, are critical for responsible AI adoption. Finally, a smart occlusion detection module could classify mask types (e.g., surgical masks, scarves) to tailor reconstruction strategies, im- proving adaptability. These steps would strengthen reliability, fairness, and real-world usability in security, forensics, and identity verification scenarios.

ACKNOWLEDGMENT

We extend our heartfelt thanks to the faculty and staff of the Computer Science and Engineering Department at Vardhaman College of Engineering for their guidance and assistance throughout this project. A special note of gratitude goes to our supervisor, Dr. G. Somasekhar, for his insightful feedback and continuous encouragement. We also recognize the open- source platforms that contributed to our research. Lastly, we deeply appreciate the encouragement and support from our friends and family members during this journey.

REFERENCES

[1] Kevin Rawlinson," Rise in suspects using face coverings to mask identity", shorturl.at/aqAP6/ 2021. The Guardian Blog. Google Scholar

[2] Antonio Criminisi, Patrick Pe´rez and Kentaro Toyama," Region filling and object removal by exemplar-based image inpainting", IEEE Trans- actions on image processing, vol. 13, no. 9, pp. 1200-1212, 2004. View Article Google Scholar

[3] James Hays and Alexei A Efros," Scene completion using millions of photographs", ACM Transactions on Graphics (ToG), vol. 26, no. 3, pp. 4-es, 2007. Google Scholar.

[4] Jeong-Seon Park, You Hwa Oh, Sang Chul Ahn and Seong-Whan Lee," Glasses removal from facial image using recursive error compensation", IEEE transactions on pattern analysis and machine intelligence, vol. 27, no. 5, pp. 805-811, 2005.. CrossRef Google Schola

[5] Satoshi Iizuka, Edgar Simo-Serra and Hiroshi Ishikawa, "Globally and locally consistent image completion", ACM Transactions on Graphics (ToG), vol. 36, no. 4, pp. 1-14, 2017 View Article Google Scholar

[6] Digital Object Identifier 10.1109/ACCESS.2024.3446652 Masked Face Recognition With Generated Occluded Part

Using Image Augmentation and CNN Maintaining Face Identity.

[7] Revealing the Unseen: Explainable AI-Driven Masked Face Recognition.

[8] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, Heling Chen, Yu Miao, Zhibing Huang, Jinbi Liang.

[9] Xiaolin Li et al. combine GANs with a Texture Network for realistic face completion. Their method restores masked regions effectively, producing smooth and natural results on the CelebA dataset.

[10] A. Gautam, J. Ilaparti, S. Rajvanshi, N. Badhani, D. Gangodkar, and Y. Lohumi, "Masked Face Detection and Feature Reconstruction for Identity Verification using Deep Learning," Dept. of Computer Science and Engineering, Graphic Era (Deemed to be) University, Dehradun, India.

[11] Liqin Ye, "Facial Mask Removal Using Generative Adversarial Net- work," 2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI), Xi'an, China, Jan. 14–16, 2022, pp. 100–104. doi: 10.1109/IWECAI53990.2022.9734376.

[12] A. A. Pawar and N. N. Patil, "Recognition of 3-D faces with missing parts and line scratch removal using new technique," in Proc. 2015 Int. Conf. Pervasive Computing (ICPC), Pune, India, Jan. 2015, pp. 1–5. doi: 10.1109/PERVASIVE.2015.7087053.

[13] N. U. Din, K. Javed, S. Bae, and J. Yi, "A Novel GAN-Based Network for Unmasking of Masked Face," in Proc. 2022 Int. Conf. on Image Processing (ICIP), IEEE, 2022. doi: 10.1109/ICIP.2022.xxxxxx (Replace with actual DOI if known).

[14] L. E. L. Coelho, R. Prates, and W. R. Schwartz, "A Generative Approach for Face Mask Removal Using Audio and Appearance," in Proc. 2021 34th SIBGRAPI Conf. on Graphics, Patterns and Im- ages (SIBGRAPI), Gramado, Brazil, Oct. 2021, pp. 345–352. doi: 10.1109/SIBGRAPI54419.2021.00040.

[15] X. Yin, D. Huang, and L. Chen, "Non-Deterministic Face Mask Re- moval Based on 3D Priors," in Proc. 2022 IEEE Int. Conf. on Image Processing (ICIP), Bordeaux, France, Oct. 2022, pp. 2116–2120. doi: 10.1109/ICIP46576.2022.9897343.

[16] H. Zheng, W. Xu, Z. Wang, X. Lu, and C. Xiao, "FHR-Net: Facial High- light Removal via Cross-Context Attention and Texture Enhancement," in Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 5405–5414. doi: 10.1109/CVPR52729.2023.00502.

[17] M. S. Sapkal, P. K. Kadbe, and B. H. Deokate, "Image Inpainting Using Kriging Interpolation Technique," in Proc. Int. Conf. on Communication and Signal Processing (ICCSP), 2014, pp. 1041–1045.

[18] R. P. Sidik and E. C. Djamal, "Face Mask Detection using Con- volutional Neural Network," in Proc. 2021 4th Int. Conf. on Com- puter and Informatics Engineering (IC2IE), Sep. 2021, pp. 45–50, doi: 10.1109/IC2IE53219.2021.9649065.

[19] A. Hore´ and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in 2010 20th International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, Aug. 2010, pp. 2366–2369, doi: 10.1109/ICPR.2010.579.

[20] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recogni- tion (CVPR), Honolulu, HI, USA, Jul. 2017, pp. 5967–

[21] 5976, doi: 10.1109/CVPR.2017.632.