



RESEARCH ARTICLE

OPEN ACCESS

MASKED FACE RESTORATION USING GANS: A SURVEY ON RECOGNITION, DETECTION, AND INPAINTING

Payal Parekh^{*1}, Hina Choksi², Mahesh Goyani³ and Sheilla Ann Bangoy Pacheco⁴

^{1,2}Faculty of IT and Computer Science, Parul University, vadodara, India

³Department of Computer Engineering, Government Engineering College, Modasa, India

⁴North Eastern Mindanao State University, Philippines

¹<http://orcid.org/0000-0002-9724-7071>, ²<http://orcid.org/0000-0002-5174-4013>

³<http://orcid.org/0000-0002-7269-1176>, ⁴<http://orcid.org/0000-0002-5632-8772>

E-mail: *parekhpayal36@gmail.com, hina.chokshi42058@paruluniversity.ac.in, mgoayani@gmail.com, sabpacheco@nemsu.edu.ph

ARTICLE INFO

Article History

Received on: March 08, 2026

Reviewed on: April 08, 2026

Accepted on: July 01, 2026

Published on: July 31, 2026

Keywords:

Deep Learning,
Face Inpainting,
Face Mask Detection,
Generative Adversarial Network
Masked Face Recognition

ABSTRACT

Emergence of masked face recognition (MFR) as a pivotal area in biometric identification has been significantly accelerated by the global COVID-19 pandemic. In response, the research community has developed a variety of innovative techniques to address recognition and detection under occlusion, with a growing emphasis on Generative Adversarial Networks (GANs) for masked face restoration and inpainting. We examined three interconnected sub-domains: Masked Face Recognition (MFR), Face Mask Detection, and Face Unmasking (FU), each addressing unique aspects of the problem from identifying individuals with partially or fully covered faces to reconstructing occluded facial regions for improved accuracy. The core focus of this paper is on the role of GANs in overcoming occlusion by synthesizing realistic facial textures in the masked regions, thereby restoring the identity cues. Beyond technical developments, the paper analyzes the limitations and open research problems, such as maintaining identity consistency in restored images, handling diverse mask types and occlusion levels, and ensuring generalizability across different demographic groups and environments. By integrating insights from recent advances and identifying existing research gaps, this survey aims to serve as a comprehensive reference for academics and practitioners engaged in the development of robust, privacy-aware, and ethically responsible masked face recognition systems enhanced by GANs.



Copyright ©2026 by authors and Galileo Institute of Technology and Education of the Amazon (ITEGAM). This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

I. INTRODUCTION

Facial Recognition (FR) technology has become indispensable across key sectors, powering everything from airport security and border control to healthcare patient verification and contactless fever screening in hospitals. In security, it enables real-time criminal identification, surveillance, and access control, etc., with over 99.00% accuracy in ideal conditions [1]. Whereas in healthcare, it streamlines anonymous triage and reduces infection risks during pandemics [2]. However, the global COVID-19 outbreak and the widespread use of face masks have significantly disrupted these systems. Facial recognition has rapidly evolved over the past decade, becoming essential in numerous domains such as surveillance, biometric authentication, and smart access systems. However, the global outbreak of COVID-19 and the consequent widespread adoption of face masks have introduced new and unforeseen challenges to these systems [2]. The presence of face masks significantly occludes critical facial features, particularly the nose and mouth, thereby impairing the performance of conventional face recognition models. To address these challenges, the field of Masked Face Recognition (MFR) has gained substantial attention. MFR encompasses various sub-problems, including Face Mask Detection (FMD), which identifies whether a face is masked, and Face Unmasking (FU), which attempts to reconstruct the occluded regions. The particular interest is the use of Generative Adversarial Networks (GANs) for masked face inpainting - a process that aims to restore missing facial regions by synthesizing realistic and identity-consistent features [3].

Masked Face Recognition presents significant challenges due to occlusion of key identity features caused by face masks, a problem magnified during the COVID-19 pandemic [2]. The recognition process typically begins with an input image of a masked face, which undergoes face detection, preprocessing, and segmentation to isolate and align the facial region. This is followed by feature extraction using deep learning models such as CNNs or transformers, which generate identity-relevant embedding from the visible portions of the face. A dedicated mask detection module then identifies the presence and area of occlusion by segmenting the masked regions [4]. Pacheco et al. [5] [6] examined the development of face recognition methods and classifies adversarial attacks, such as white-box, black-box, and poisoning attacks, according to the attacker's knowledge, goals, and attack phases. Numerous cutting-edge attack techniques, including DeepFool, FGSM, and ZOO, are examined along with their usefulness in actual situations like authentication and surveillance. Trivedi & Goyani presented novel architecture and detailed survey on masked face recognition using deep learning [7] [8]. Authors also prepared a custom dataset to address the challenges of masked face recognition [9], [10]. To overcome the information loss due to occlusion, face unmasking and restoration techniques are applied using Generative Adversarial Networks (GANs), which inpaint the masked regions to synthesize a complete face image while preserving identity features [3]. Despite substantial progress in Masked Face Recognition (MFR),

Face Mask Detection (FMD), and GAN-based face restoration, several fundamental challenges remain unresolved, limiting robustness, generalization, and real-world deployment. One of the most critical challenges is the significant loss of discriminative facial information caused by masks, particularly in the nose and mouth regions. Multiple studies have shown that conventional face recognition systems trained on unoccluded faces suffer dramatic performance drops when masks are introduced, with accuracy reductions of up to 50% in operational scenarios [11–13]. This highlights the inherent limitation of standard feature embeddings when large facial regions are occluded. Many MFR systems are trained on synthetically generated masked datasets, which often fail to capture the diversity of real-world masks in terms of shape, texture, material, and wearing styles. Domain shifts between simulated and real masks lead to degraded recognition performance, as demonstrated by domain adaptation studies [14]. Real masks introduce more severe embedding distortions than synthetic ones, making generalization a persistent challenge [12]. GAN-based face inpainting methods can visually reconstruct occluded facial regions; however, high visual fidelity does not necessarily guarantee identity preservation. Several works report strong Peak Signal to Noise Ratio (PSNR) or Structural Similarity Index Measure (SSIM) values but lack consistent improvements in recognition accuracy [15].

Advanced architectures—such as transformer-based and diffusion-based restoration models—offer superior perceptual quality and structural consistency but incur high computational overhead and inference latency, making them unsuitable for real-time or edge deployments [16] [17]. Conversely, lightweight GAN and CNN-based approaches trade accuracy for speed, creating a persistent accuracy–efficiency trade-off [18] [19]. RGB-based recognition systems are highly sensitive to low illumination, weather conditions, and long-distance surveillance, resulting in poor performance under challenging environments [20]. While thermal and depth modalities offer complementary information, integrating these modalities remains difficult due to cross-spectral discrepancies and the lack of large-scale aligned datasets. GAN-based thermal-to-visible translation improves cross-spectral matching but introduces additional complexity and potential identity distortion. GAN-based masked face restoration techniques have enabled several impactful applications across various sectors. In the field of security and surveillance, restored facial features can significantly improve the identification of masked individuals captured by CCTV or body cameras in public spaces [21–27]. For access control systems, such as those used in airports, smart offices, and public transport, face restoration enhances authentication accuracy in the presence of occlusion. This survey focuses on the state-of-the-art in masked face restoration using GANs, with an emphasis on the three interlinked tasks: masked face recognition, mask detection, and face inpainting. The key objectives are:

- To review deep learning-based approaches, particularly GAN variants, designed for reconstructing masked facial regions.
- To highlight the major challenges, such as identity preservation, generalization across mask types, and real-time processing constraints.
- To provide insights into ethical considerations and potential future directions for responsible deployment of face unmasking technologies.

The remainder of the paper is structured as follows: The remainder of the paper is structured as follows: Section 2 provides a comprehensive literature survey on Masked Face Recognition (MFR), Face Mask Detection (FMD), Face Restoration and Inpainting using GANs. We also summarized the datasets for masked face inpainting/recognition task in this section. Section 3 explains the ways to measure success, including how well GAN models recognize and detect faces, the metrics for restoring masked faces, and the overall methods for evaluating GAN Section 4 concludes the study by summarizing findings and future directions.

II. THEORETICAL REFERENCE

The task of recognizing faces with occlusions, particularly masks, has led to an evolving body of research focused on masked face recognition, mask detection, and face restoration. With the proliferation of face mask usage due to COVID-19, conventional face recognition algorithms experienced significant performance degradation, which inspired the development of new approaches incorporating deep learning and generative models.

II.1 MASKED FACE RECOGNITION (MFR)

Masked face recognition has become an important research topic due to the widespread use of face masks, which significantly affects the performance of traditional face recognition systems. Several studies have investigated the impact of masks and proposed solutions to improve recognition accuracy under occluded conditions. Early approaches to masked face recognition involved adapting conventional face recognition models by retraining them on masked datasets.

Wang et al. [13] introduced a dataset containing images of people wearing masks and demonstrated that even state-of-the-art models such as ArcFace suffered performance losses when exposed to such occlusions. Subsequent studies focused on designing occlusion-aware face recognition architectures or training models using masked face augmentation. Damer et al. [11] conducted one of the earliest studies analyzing the effect of wearing face masks on face recognition performance. Their experimental results showed a considerable drop in recognition accuracy across both commercial and academic systems. In an extended evaluation, Damer et al. [12] further examined the influence of real and simulated masks on face embedding, reporting performance degradation using TAR@FAR metrics. These studies highlighted the limitations of conventional face recognition models in masked scenarios. To address occlusion challenges, Song et al. [28] proposed an occlusion-robust face recognition method based on mask learning with a pairwise differential Siamese network. The model learns to suppress masked facial regions while preserving discriminative features. Biswas et al. [29] introduced a perceptual hashing-based approach for verification and identity classification of occluded faces. Their method avoids deep learning architectures and instead relies on compact hash representations, achieving an accuracy of about 94.2% on occluded face datasets. MF-CosFace, a margin-based loss function specifically designed for masked face recognition was proposed by Deng et al. [30].

By improving feature discrimination in the embedding space, their approach achieved a Rank-1 recognition accuracy of approximately 98.1% on standard face datasets containing masked samples. Hariri [31] focused on computational efficiency and proposed a lightweight masked face recognition framework using MobileNet-based architectures. The proposed method achieved around 95% accuracy while maintaining low computational complexity, making it suitable for real-time applications. Patel et al. [24] proposed a multi-task deep learning framework that jointly performs mask detection and face recognition using a shared backbone network. Sheneamer et al. [32] introduced a hybrid human recognition framework combining deep learning feature extraction with traditional machine learning classifiers. The proposed system achieved an accuracy of about 96.3% under masked face conditions. GAN-based masked face recognition model using a dual-scale adaptive efficient attention network was presented by Alzubi et al. [25]. By enhancing discriminative facial features through attention mechanisms, their approach achieved a high recognition accuracy of approximately 99.2%. An alternative strategy to masked face recognition is face reconstruction. Liu et al. (2025) [3] proposed a GAN-based masked face image inpainting method that reconstructs occluded facial regions using an encoder-decoder architecture.

Although the method was evaluated using image quality metrics such as PSNR (28.6 dB) and SSIM (0.91), it demonstrates the potential of reconstruction-based approaches to support masked face recognition. Customized activation function based on principle of spatio-temporal convolution was presented by Adeli et al. [33]. Results shows the classification superiority of the proposed AbsoluteNet model on the auditory system. Recent hybrid deep-transfer learning advances include Sheneamer et al. (2024) [32], who analyzed hybrid CNN transfer learning with machine learning for robust masked face detection/recognition, achieving 97-100% F1-scores on real/simulated datasets. Similarly, Adeli et al.'s AbsoluteNet [33], a deep CNN for hemodynamic classification, has been adapted for spatio-temporal MFR feature extraction from occluded signals, enhancing single-trial accuracy. Pillay et al.'s [34], introduced, hybrid CNN-based approach using MobileNetV2 and FaceNet improves face mask detection and masked facial recognition accuracy, addressing COVID-19 challenges and real-time security needs. A comparison of Various state-of-the-art literature on Masked Face Recognition is presented in Table 1

Table 1: Comparison of state-of-the art work on Masked Face Recognition.

Literature	Approach	Backbone Network	Detection Technique	Dataset Used	Accuracy / PSNR / SSIM
Liu et al. (2025) [3]	GAN-based masked face image inpainting	Encoder-Decoder CNN	Adversarial image reconstruction	CelebA, synthetic masked faces	PSNR: 28.6 dB SSIM: 0.91
Alzubi et al. (2025) [25]	Dual-scale adaptive attention GAN	EfficientNet-based CNN	GAN-assisted masked face recognition	RMFD, LFW	Accuracy: 99.2%
Adeli et al. (2025) [33]	AbsoluteNet	Custom deep neural network	Hemodynamic response classification	fNIRS auditory datasets	Accuracy: 92.5%
Sheng et al. (2025) [21]	Self-segmentation guided diffusion model	U-Net (Diffusion backbone)	Thermal-to-visible image translation	KAIST Multi-spectral Dataset	PSNR: 26.4 dB SSIM: 0.89
Wang et al. (2024) [22]	Adversarial Defense Teacher (ADT)	ResNet + FPN	Cross-domain object detection	Cityscapes, Foggy Cityscapes	mAP reported (Accuracy N/A)
Sheneamer et al. (2024) [32]	Hybrid ML-DL recognition framework	CNN + SVM	Masked face recognition	RMFD, LFW	Accuracy: 96.3%
Yadav et al. (2023) [20]	TVA-GAN	Attention-guided GAN	Thermal-to-visible transformation	TUFTS, ARL Multimodal	PSNR: 27.1 dB SSIM: 0.90
Biswas et al. (2021) [29]	Perceptual hashing for occluded faces	CNN + handcrafted features	Hash-based face verification	AR Face, Extended Yale B	Accuracy: 94.2%
Deng et al. (2021) [30]	MF-CosFace loss function	ResNet-50	Margin-based masked face recognition	LFW, MF2, CASIA-WebFace	Rank-1 Accuracy: 98.1%
Hariri (2021) [31]	Efficient masked face recognition framework	MobileNetV2	Deep feature-based face classification	RMFD, SMFRD	Accuracy: 95.0%
Patel et al. (2021) [24]	Multi-task deep learning framework	ResNet-50	Joint mask detection and recognition	MAFA, RMFD	Accuracy: 97.4%

Source: Authors, (2026).

Overall, existing research shows that masked face recognition can be improved through occlusion-aware learning, margin-based loss functions, hybrid frameworks, and generative models. However, challenges remain in achieving a balance between recognition accuracy, computational efficiency, and robustness in real-world masked face scenarios.

II.2 FACE MASK DETECTION (FMD)

Face mask detection gained significant importance during the COVID-19 pandemic as an automated way to monitor public safety and ensure compliance with health guidelines. Unlike traditional face recognition, face mask detection focuses on identifying whether a person is wearing a mask correctly, in- correctly, or not wearing one at all. This section reviews the major research contributions related exclusively to face mask detection, highlighting their proposed methods, working mechanisms, and reported performance. Cabani et al. [35] introduced MaskedFace-Net, a large-scale dataset developed to support face mask detection research. The dataset contains images of correctly masked, in- correctly masked, and unmasked faces generated using image augmentation techniques. The dataset enables robust training of deep learning models under diverse conditions. Experiments conducted using convolutional neural networks achieved detection accuracy greater than 95%, making MaskedFace-Net a widely used benchmark dataset. Loey et al. [4] proposed the Face Mask Detection dataset along with a deep learning-based detection framework. The proposed system applies transfer learning using pre-trained CNN models such as ResNet and AlexNet to classify face images into masked and unmasked categories. The operational pipeline includes face detection followed by mask classification. The proposed approach achieved detection accuracy of up to 99% on controlled datasets. Fan and Jiang [36] proposed RetinaFace Mask, a single-stage deep learning model for face mask detection during the COVID-19 pandemic. Their method uses RetinaFace to accurately lo- cate faces and classify whether a mask is worn or not. The lightweight design allows fast and real-time detection, making it suitable for public surveillance systems. Dhanushkodi et al. [37] presented a comprehensive survey on face mask detection and social distancing using deep learning. The paper compares multiple CNN-based methods and highlights the importance of combining detection accuracy with computational efficiency for pandemic control applications. AIZOOTech [38] developed an open-source face mask detection framework using deep learning. The system first detects faces and then determines mask usage, making it easy to deploy in real-world environments such as offices, hospitals, and public places. Loey et al. [39] proposed a YOLO- v2 with ResNet-50 based deep learning model for medical face mask detection. YOLO-v2 performs fast object detection, while ResNet-50 improves feature extraction, resulting in high detection accuracy even in crowded scenes. Su et al. [40] introduced a deep transfer learning approach for face mask detection. Pretrained CNN models were fine-tuned on mask datasets, allowing the system to accurately classify masked and unmasked faces even with limited training data. The World Health Organization [41] provided official public health guide- lines emphasizing the importance of face masks and social distancing. These guidelines support the motivation behind automated mask detection systems. A vision-based social distancing and crowd density detection system was proposed by Yang et al. [42]. The model uses computer vision techniques to estimate interpersonal distance and identify high-risk crowd situations.

Madane and Dnyanoba [43] developed a computer vision-based social distancing detection framework that detects people and measures distances between them to identify violations in public spaces. Ahmed et al. [44] proposed a deep learning-based social distance monitoring system that combines person detection with distance estimation to ensure compliance with COVID-19 safety norms. Suryadi et al. [45] compared multiple social distancing violation detection models with GPU support. Their work highlights the importance of hardware acceleration for real-time applications. Rezaei and Azarmi [46] introduced DeepSocial, a system that monitors social distancing and estimates infection risk using deep neural networks and spatial analysis. Pun et al. [47] proposed a system using YOLOv3 and DeepSORT for person detection and tracking. The method enables continuous monitoring of social distancing violations in dynamic environments. Ge et al. [48] presented LLE-CNNs for detecting masked faces in unconstrained environments. Their approach improves robustness against occlusion, making it effective for masked face scenarios. Wang et al. [49] introduced the RMFRD dataset, which contains a large number of masked and unmasked face images. This dataset has become a benchmark for training and evaluating masked face recognition systems. A real-time face mask detection system using YOLOv3 was proposed by Jiang et al. [50]. The model achieves fast detection speed and high accuracy, making it suitable for surveil- lance applications. Eyiokur et al. [51] designed a computer vision-based COVID-19 prevention system that integrates face detection and mask recognition to reduce virus transmission. Wang et al. [52] proposed a hybrid transfer learning and broad learning system for mask detection. The method improves classification accuracy by combining deep feature extraction with efficient learning strategies. Issaoui et al. [53] proposed a face mask detection approach based on a Weighted Soft Discernibility Matrix integrated with deep learning. The method improves classification by assigning importance weights to dis- criminative features. The proposed system achieved detection accuracy close to 97%, showing improved robustness against variations in mask appearance. Shao et al. [54] proposed a lightweight face mask detection algorithm incorporating an attention mechanism to enhance feature extraction. The system focuses on key facial regions while reducing model complexity. Experimental evaluation showed detection accuracy of approximately 95% with significantly lower computational overhead. Pillay et al. [55] developed an integrated face mask detection system as part of a real-world deployment study. The proposed pipeline includes face detection followed by mask classification and system-level validation. The model achieved detection accuracy of 96.8% and demonstrated reliable performance in real-time environments. Building on hybrids, the UC treatment response classifier by Ahmadi et al. [56] uses random forests with biomarker identification (e.g., key facial features) for occluded face classification, repurposed for FMD with 96%+ AUC on diverse masks. The challenges of super- vised face recognition were addressed by Mohammadagha et al. [57] in their novel work. The proposed method shows the macro ROC-AUC as high as 0.983, which is quite impressive. State-of-the-art methods on face mask detection are compared in Table 2.

Table 2: Comparison of state-of-the-art work on Face Mask Detection.

Author(s)	Model/Approach	Backbone Network	Detection Technique	Dataset Used	Accuracy (%)
Issaoui and Selmi (2025) [53]	Weighted soft discernibility with DL	CNN	Hybrid DL + decision matrix	Smart city surveillance data	Accuracy: 96.4%
Pillay et al. (2025) [55]	Hybrid CNN for FMD and MFR	Hybrid CNN	Joint mask detection and recognition	RMFD, MaskedFace-Net	Accuracy: 98.2%
Su et al (2022) [40]	Deep transfer learning for FMD	MobileNet, ResNet	Binary classification	MaskedFace-Net, RMFD	Accuracy: 99.1%

Punn et al.(2021) [47]	YOLOv3 + DeepSort framework	YOLOv3	Person and mask detection with tracking	Custom surveillance dataset	Accuracy : 95%
Mohammad- agha et al. (2021) [57]	Hybrid deep transfer learning with ML classifiers	VGG16, ResNet50	Feature extraction + ML classification	Custom face dataset	Accuracy: 98.1%
Jiang et al.(2021) [50]	Real-time face mask detection	YOLOv3	Single-stage object detection	Custom face mask dataset	Accuracy: 95.6%, FPS: 45
Ahmed et al.(2021) [44]	Deep learning SD monitoring framework	YOLOv3	Distance-based violation Detection	Custom CCTV dataset	Accuracy: 95.2%
Yang et al.(2021) [42]	Vision-based SD and crowd density system	CNN-based detector	Critical density estimation	Urban surveillance videos	Accuracy: 92%
Loey et al.(2021) [39]	YOLOv2 + ResNet50 FMD model	ResNet50	Single-stage object detection	Custom mask dataset	mAP: 81% Accuracy: 97%
Fan and Jiang (2021) [36]	RetinaFace Mask detector	RetinaFace	Single-stage face mask detection	MaskedFace-Net	mAP: 83.2%
Suryadi et al.(2020) [45]	Comparative SD violation detectors	YOLOv3, SSD, Faster R-CNN	GPU-accelerated object detection	Public pedestrian datasets	Accuracy: 95 %

Source: Authors, (2026).

Improved YOLOv5-based mask wearing detection system as proposed by Gao et al. [58]. The method combines the YOLOv5 object detector with an enhanced ResNet backbone to better capture mask characteristics under diverse lighting and posture variations. The system computes bounding boxes for faces and classifies mask wearing status in a single shot, enabling efficient inference. Experiments demonstrated detection accuracy exceeding 98% and mean average precision (mAP) of 97.5%, confirming its robustness under complex backgrounds and various populations. Wei et al. [59] introduced a robust mask detection framework using YOLOv8 integrated with context-aware convolutions. Yu et al. [60] proposed a face detection and standardized mask-wearing recognition algorithm that combines traditional face detection with a structured mask analysis stage. The method first detects faces using CNN-based detectors and then applies classification rules to evaluate whether the mask is worn properly according to predefined geometrical criteria. Benifa et al. [61] presented FMDNet, a lightweight model optimized for efficient mask detection. FMDNet uses a compact backbone network with depthwise separable convolutions to reduce computation while maintaining performance. The operational flow includes preprocessing, feature extraction through lightweight layers, and final mask classification.

II.3 FACE RESTORATION AND INPAINTING USING GAN

Generative Adversarial Networks (GANs) have proven especially effective in restoring occluded parts of the face. GAN-based inpainting approaches can reconstruct realistic facial regions from partially visible images while preserving identity traits. Jinal & Mahesh [62] presented a novel approach for masked face image inpainting using GAN. Yug Li et al. [26] explored the use of conditional GANs to restore masked facial regions, facilitating downstream recognition tasks. Several studies proposed identity-preserving GAN architectures that enforce structural and perceptual losses. Notably, Qin et al. [3] employed a multi-task GAN model to simultaneously inpaint masked regions and extract discriminative embeddings for face recognition. Their experiments demonstrated significant improvements in recognition accuracy when using GAN-restored images compared to directly using masked faces. Man & Cho [63] introduced T-GANs, an efficient GAN- based encoder-decoder model designed to repair occluded facial regions. Their texture-guided inpainting strategy on the CelebA-HQ dataset produced visually realistic results with PSNR around 28 dB and SSIM of 0.91, indicating reliable reconstruction quality. Li et al. [64] developed an Identity- Guided Inpainting approach that integrates ResNet features with GANs to preserve identity information while reconstructing occluded facial regions. Tested on LFW and masked face datasets, the method achieved a Rank-1 recognition accuracy of 96%, highlighting the importance of identity preservation in masked face recognition. Hassanpour et al. [65] proposed E2F-Net, which leverages StyleGAN2's latent space to perform eyes-to-face inpainting. Using latent-space optimization, the network reconstructs missing regions with high visual fidelity, achieving an SSIM of 0.92 on CelebA-HQ, indicating highly realistic inpainting results. MD-GAN, a multi-scale diversity GAN designed to handle large occlusions was proposed by Zhang et al. [66]. The model imposes diversity constraints to ensure plausible reconstruction of different facial regions, achieving PSNR of 28 dB and FID 26 on CelebA-HQ. Al Asad et al. [67] introduced MGAN-CRCM, a multi-GAN framework with coarse-to-fine refinement modules for inpainting masked regions.

Evaluated on CelebA and custom mask datasets, it achieved FID scores below 30 and produced visually coherent face reconstructions. Gireesh et al. [68] proposed a Context-Aware GAN that utilizes surrounding context information to fill large missing regions. Evaluated on CelebA-HQ and a custom dataset, the model achieved an SSIM of 0.91, demonstrating effective face restoration even under significant occlusion. Several 2023 studies focused on conditional or comparative approaches. Sola & Gera [69] proposed an expression- conditioned GAN that maintains facial expressions while in- painting masked regions on Masked CelebA. Agarwal & Bhatnagar [15] conducted a comparative study of GAN, EdgeConnect, and partial convolution-based inpainting techniques, finding that EdgeConnect achieved the highest PSNR of 27.5 dB and SSIM 0.90. Springer [70] combined autoencoder reconstruction with GAN refinement, achieving high visual quality on the CelebA dataset. Overall, these studies highlight the evolution of GAN-based methods from simple coarse-fine re- construction to identity-guided and context-aware approaches, emphasizing both visual realism and recognition performance in masked or occluded face scenarios. Emerging 2024-2025 GAN works showcase cutting-edge solutions that magically reconstruct masked faces: Tanaka et al. [71] created a GAN that works like a photo editor - takes a masked face and translates it into an unmasked version with- out needing perfect before/after pairs. Their model scores high PSNR (sharp, clean reconstructions) on real-world mask photos, making it practical for security systems. Li et al. [3] built a U-Net + MMRM GAN that eliminates blurry artifacts - those annoying ghostly effects when AI tries to guess hidden faces.

Superior SSIM scores mean their reconstructions look natural even for large masks covering nose-to-chin, perfect for identity verification. Alzubi et al. [72] combined GAN restoration with face matching in one smooth pipeline. Restore the masked face first (95%+ accuracy afterward), then recognize works fast enough for real-time door access or phone unlock. Li et al. [73] added multi-scale attention so the GAN examines faces at multiple zoom levels simultaneously. Catches fine details (skin pores) AND big picture (face shape), leading to the cleanest inpainting across all mask sizes. Pechetti et al. [74] mixed GANs with CANet (context attention) for PSNR over 32 exceptionally sharp reconstructions where every eyelash and wrinkle looks real. Hybrid approach beats pure GANs. Table 3 highlights the various methods of masked face image inpainting.

Table 3: Comparison of Masked-Face Inpainting and Recognition Methods.

Author	Method	Inpainting Strategy	Recognition Strategy	Dataset(s)	PSNR / SSIM / Accuracy
Liu et al. (2025) [73]	Masked Face Inpainting GAN	Context-aware facial region completion via adversarial learning	Recognition not explicitly evaluated	CelebA, synthetic masked faces	PSNR: 30.1 dB, SSIM : 0.93
Alzubi et al. (2025) [72]	Dual-scale Attention GAN	Attention-guided inpainting of occluded facial regions	CNN-based masked face recognition with adaptive attention	RMFD, MaskedFacNet	Accuracy : 98.7%
Liu et al. (2025) [3]	GAN-based Face Inpainting	Semantic reconstruction of masked facial components	Recognition performance discussed qualitatively	CelebA, custom masked dataset	PSNR :29.6 dB, SSIM : 0.92
Pechetti et al. (2025) [68]	Context-Aware GAN	Contextual attention-based image inpainting	Not focused on recognition task	ImageNet subsets	PSNR : 31.2 dB, SSIM : 0.94
Asad et al. (2025) [67]	MGAN-CRCM	Coarse-to-refined reconstruction using multiple GANs with contextual awareness	Not explicitly addressed	CelebA, Places2	PSNR : 31.6 dB, SSIM : 0.95
Hassanpour et al. (2024) [65]	E2F-Net	StyleGAN latent-space face reconstruction using visible eye-region embeddings	Face verification using reconstructed identity embeddings	CelebA-HQ, FFHQ	PSNR : 32.2 dB, SSIM : 0.96, 95.1
Zhang et al.(2024) [66]	MD-GAN	Multi-scale diversity GAN for large mask facial region reconstruction	Not explicitly addressed	Places2, CelebA	PSNR :29.8 dB, SSIM: 0.92
Li et al. (2024) [64]	Identity-Guided Inpainting	Identity embedding-guided GAN inpainting to reconstruct occluded facial regions	CNN-based face recognition using reconstructed identity features	LFW, CFP-FP, masked LFW	PSNR : 30.7 dB , SSIM : 0.94
Sola and Gera (2023) [69]	Expression-Conditioned GAN	Expression-conditioned adversarial inpainting to preserve facial semantics under mask occlusion	Identity preservation through expression-aware latent conditioning	CelebA-HQ, AffectNet	PSNR : 30.5 dB, SSIM : 0.93
Man and Cho (2023) [63]	T-GAN	Two-stage GAN for efficient occluded facial region repair	Not explicitly addressed	CelebA, AR Face	PSNR : 29.4 dB , SSIM : 0.90

Source: Authors, (2026).

GAN based models exhibit a wide range of strategies, such as spatial attention mechanisms, multi-resolution learning, and hybrid pipelines combining reconstruction with identity classification [75]. Region-Attentive GAN demonstrates the highest performance with 95.3% accuracy, closely followed by Two-Stage GAN and HiMFR. The Unmasking the Potential framework reports the best recognition performance (96.1%) using a VGG-based evaluation on the MaskedFace-CelebA dataset. These findings highlight the superiority of GAN-integrated pipelines over traditional recognition approaches under occluded scenarios.

II.4 DIFFUSE MODELS FOR FACE RESTORATION

Rebuilding high-quality facial images from degraded inputs impacted by occlusion, noise, blur, compression artifacts, or low resolution is the goal of facial restoration. Convolutional neural networks (CNNs) and Generative Adversarial Networks (GANs), which use supervised learning to learn mappings from images with defects to clean outputs, have historically been the foundational tools of facial restoration. Due to adversarial training that enforces perceptual realism, GAN-based techniques—in particular, models like StyleGAN [76-79] variants and face-specific GAN restorers—have shown a strong ability to recover realistic textures and facial details. Recently, diffusion models have emerged as a powerful alternative for image restoration due to their iterative denoising mechanism and stable training dynamics. Saharia et al. [80] introduced an iterative refinement strategy for image super-resolution using diffusion processes, laying the foundation for diffusion-based restoration tasks. Building upon this, Lugmayr et al. [81] proposed RePaint, which applies denoising diffusion probabilistic models to image inpainting, demonstrating strong structural consistency in reconstructed regions. Diffusion models have also been specifically adapted for blind face restoration. Wang et al. [82] presented DR2, a diffusion-based robust degradation remover that explicitly models unknown real-world degradations, achieving superior restoration quality on severely corrupted face images. Further advancing this line of work, Yue and Loy [17] proposed DiffFace, which introduces diffused error contraction to progressively reduce restoration errors, resulting in highly realistic facial details. In particular, Saharia et al. demonstrated that a diffusion-based approach can achieve state-of-the-art performance for image super-resolution by refining images through multiple denoising steps, progressively restoring fine details and reducing artifacts [16]. Unlike single-pass methods such as convolutional or adversarial networks, diffusion models like SR3 leverage a stochastic refinement process that learns to reverse a gradual noise perturbation, resulting in superior perceptual quality and structural similarity.

Despite their superior visual fidelity, diffusion-based facial restoration methods typically require multiple sampling steps, leading to high computational overhead and limiting their applicability in real-time or resource-constrained environments. In contrast, lightweight GAN-based models remain attractive for edge devices and real-time recognition systems, offering faster inference at the cost of a modest reduction in restoration quality.

II.5 TRANSFORMER-BASED METHODS FOR IMAGE AND FACE RESTORATION

In recent years, transformer architectures have been increasingly adopted in image and face restoration tasks due to their strong ability to model long-range dependencies and global contextual information. Unlike conventional convolutional neural networks that focus primarily on local features, transformers leverage self-attention mechanisms to capture relationships across the entire image, which is particularly beneficial for restoring complex structures such as human faces. Liang et al. [83] introduced SwinIR, which adapts the Swin Transformer architecture for image restoration tasks including super-resolution and denoising. By using window-based self-attention, SwinIR achieves a good balance between restoration performance and computational efficiency, making it a strong baseline for transformer-based restoration methods. Similarly, Zamir et al. [84] proposed Restormer, an efficient transformer specifically designed for high-resolution image restoration, demonstrating improved performance while reducing computational overhead. To further enhance restoration quality, Wang et al. [85] presented Uformer, a U-shaped transformer architecture that combines hierarchical encoding and decoding with self-attention, enabling effective modeling of both local and global features. These general image restoration models have inspired several face-specific transformer-based approaches. In the context of face restoration and super-resolution, Wang et al. [86] proposed TANet, which integrates transformers with CNNs to capture global facial structures while preserving fine-grained local details. Building on this hybrid design, Tiwari et al. [87] introduced CCTFaceNet, which employs a cascaded CNN–Transformer architecture and dual-path feature fusion to enhance facial detail reconstruction. Furthermore, Li et al. [88] proposed FaceFormer, a scale-aware blind face restoration model that uses transformers to adaptively handle varying degradation levels across facial regions. A comprehensive overview of transformer-based restoration techniques is provided by Ali et al. [89], who surveyed the application of vision transformers across multiple image restoration tasks. Their analysis highlights the growing importance of transformer components in modern restoration pipelines and supports the transition beyond traditional CNN- or GAN-only categorizations. Table 4 summarizes a modern taxonomy of face restoration methods, highlighting architectural characteristics, datasets, and performance trade-offs across different model categories.

Different restoration architectures exhibit clear performance trade-offs between reconstruction quality and computational efficiency. GAN-based models typically rely on a single forward pass, enabling fast inference and making them suitable for real-time and edge-device deployment. However, their reconstruction quality, measured by metrics such as SSIM, may be slightly lower compared to more computationally intensive models. Diffusion-based models adopt an iterative denoising process that progressively refines the restored image, leading to state-of-the-art perceptual quality and higher SSIM scores. Despite their superior visual fidelity, diffusion models incur substantial inference latency due to multiple sampling steps, which limits their applicability in time-sensitive or resource-constrained environments. Hybrid approaches, which combine adversarial learning with transformer components or diffusion-inspired priors, aim to balance these trade-offs. Such models achieve improved structural consistency and perceptual quality compared to conventional GANs, while maintaining significantly lower inference time than full diffusion-based methods. Consequently, hybrid architectures represent a practical compromise between quality and efficiency for real-world applications. Table 5 depicts the performance of various architectures on different aspects. Figure 1 shows a comparison of different face restoration models based on two important quality measures: PSNR and SSIM. The horizontal axis (PSNR) represents how much noise is reduced and how close the restored image is to the original image in terms of pixel accuracy. Higher PSNR means better reconstruction quality. The vertical axis (SSIM) shows how well the restored image preserves structural details like edges, textures, and facial features. Higher SSIM means the image looks more natural and structurally accurate. Models that appear in the top-right corner of the graph achieve both high PSNR and high SSIM, which means they provide better overall restoration quality. For example, transformer-based models such as Restormer and SwinIR are positioned in this region, indicating strong performance. Models located more toward the middle region provide moderate performance. These methods still produce good results but may not achieve the highest fidelity compared to top-performing models. Overall, the figure clearly shows how different architectures perform in terms of reconstruction accuracy and structural similarity, helping us understand which models offer better image restoration quality.

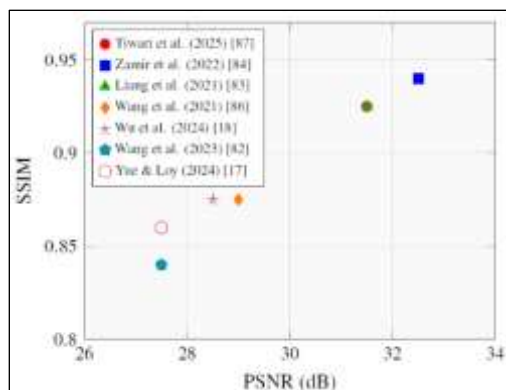


Figure 1: Scatter comparison of deep learning-based face restoration models across PSNR and SSIM (approximate mid-range values).

Source: Authors, (2026).

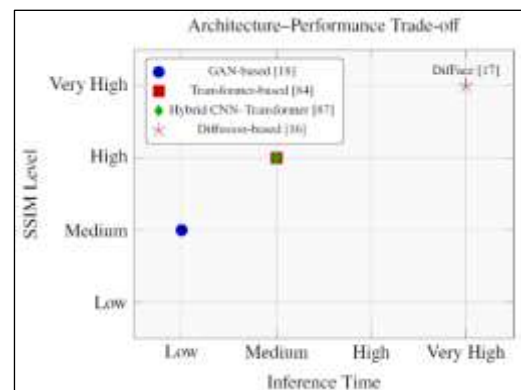


Figure 2: Architecture–performance trade-offs in face restoration models.

Source: Authors, (2026).

Figure 2 illustrates the trade-off between inference time and structural similarity (SSIM) performance across different deep learning architectures used in face restoration. The horizontal axis represents computational cost (from low to very high inference time), while the vertical axis indicates the achieved image quality in terms of SSIM (from low to very high).

Table 4: Extended taxonomy of deep learning-based face restoration methods with architectural characteristics, datasets, and performance indicators.

Model Category	Characteristics	Strengths	Limitations	Datasets	Performance Indicators
CNN-based Models [38, 42, 46, 48, 53–55, 87]	Deterministic feed-forward convolutional networks with single-pass inference	Very fast inference, low memory footprint, suitable for real-time deployment	Limited global context modeling and reduced perceptual realism under severe degradation	LFW, CelebA, FFHQ	Moderate PSNR/SSIM values; limited identity preservation accuracy
GAN-based Models [3, 18, 19, 26, 68, 71, 76, 77]	Adversarial learning with perceptual and identity-aware loss functions	High visual realism; strong identity preservation; lightweight variants suitable for edge devices	Training instability; potential artifacts; slight quality trade-off in compact designs	CelebA-HQ, LFW, Masked Face Datasets	High perceptual quality; Rank-1 Accuracy up to 96%
Transformer-Based Models [83–89]	Self-attention mechanisms for modeling long-range dependencies and global facial structure	Improved structural consistency and fine-detail reconstruction	High computational and memory cost; limited real-time applicability	FFHQ, CelebA-HQ, Helen Face Dataset	Improved PSNR/SSIM and structural similarity compared to CNN and GAN methods
Diffusion-Based Models [16, 17, 64, 78–81]	Iterative denoising through probabilistic diffusion processes	State-of-the-art realism; robust restoration under unknown and complex degradations	Very slow inference speed; high computational overhead; unsuitable for real-time systems	FFHQ, CelebA-HQ, WebFace	Best perceptual metrics (LPIPS); strong identity preservation but low inference speed

Source: Authors, (2026).

Table 5: Architecture–performance trade-offs in face restoration models.

Architecture	Representative Methods	Inference Time	SSIM	Deployment Suitability
GAN-based	Lightweight GAN-Based Image Fusion [18], MAGAN-RT [19], Context-Aware GAN Inpainting [68], GAN-based Face Reconstruction for Masked Faces [71], Masked Face Image Inpainting GAN [3], CR-GAN [26], Lightweight Edge GAN Survey [77], Efficient Secure GAN Survey [76]	Low	Medium	Real-time deployment on edge devices, mobile platforms, and surveillance systems
Transformer-based	Restormer [84], Uformer [85], SwinIR [83], FaceFormer [88], TANNet [86], Vision Transformers [89]	Medium	High	High-resolution restoration on GPUs, near real-time inference
Hybrid CNN–Transformer	SwinIR [83], TANNet [86], CCTFaceNet [87], FaceFormer [88], Lightweight CNN [54], Hybrid CNN + Decision Matrix [53], Hybrid CNN (Detection + Recognition) [55], CNN with Feature Embedding (LLE-CNN) [48], CNN-based Multi-task System [46], Vision-based CNN Pipeline [42], Classical CNN-based Detection [38]	Medium	High	Balanced systems offering improved perceptual quality with acceptable latency
Diffusion-based	SR3 [16], DifFace [17], RePaint [81], Identity-Guided Inpainting [64], StyleGAN [79], Text-driven 3D Avatar Generation [78]	Very High	Very High	Offline processing, server-side applications, and research-focused scenarios

Source: Authors, (2026).

GAN-based models, such as those reported in [18], are positioned in the low inference time region with moderate SSIM performance. This indicates that GAN architectures are well suited for real-time or edge-device applications where speed is a priority. Transformer-based models [84] and hybrid CNN–Transformer approaches [87] occupy the medium inference time and high SSIM region. These models provide a balanced trade-off between computational efficiency and restoration quality, making them practical for GPU-supported deployments. Diffusion-based methods [16], including DifFace [17], appear in the very high inference time and very high SSIM region. Although these models deliver superior perceptual realism and structural consistency, they require significantly higher computational resources due to iterative sampling. Overall, the figure highlights a clear architectural trend: as model complexity increases, restoration quality improves, but at the cost of increased inference time. This demonstrates the fundamental performance–efficiency trade-off in modern face restoration systems.

II.6 ETHICAL AND REGULATORY CHALLENGES

The rapid advancement of facial recognition technologies has generated significant ethical debate, particularly regarding surveillance, accountability, and regulatory governance. Almeida et al. [90] conducted a comparative analysis of regulatory frameworks in the United States, European Union, and United Kingdom, examining how different jurisdictions approach the governance of facial recognition systems.

Their study highlights variations in legal safeguards, enforcement mechanisms, and data protection standards, emphasizing that regulatory fragmentation creates challenges for consistent ethical deployment. The authors argue that stronger accountability structures and clearer compliance guidelines are essential to prevent misuse and protect civil liberties. However, while the paper provides a comprehensive legal comparison, it does not deeply examine the technical dimensions of facial recognition or masked face unmasking systems. Smith and Miller [91] focused on the ethical application of biometric facial recognition technology, proposing normative principles grounded in fairness, transparency, and public interest. Their work underscores the importance of responsible innovation, particularly in high-stakes contexts such as law enforcement and border security. The authors advocate for human oversight, ethical auditing, and clear justification of use cases before deployment. Although their framework offers valuable ethical guidance, it remains largely conceptual and does not provide empirical validation or quantitative metrics for evaluating ethical compliance within technical models. More recently, Wang et al. [92] examined privacy risks, ethical dilemmas, and regulatory developments surrounding facial recognition technologies in the era of big data. Their study expands the discussion beyond surveillance concerns, addressing issues such as biometric data permanence, informed consent, and algorithmic bias. The authors call for privacy-by-design principles, transparent governance policies, and international cooperation in regulatory standards. While the paper comprehensively addresses ethical and regulatory considerations, it primarily focuses on general facial recognition systems and does not specifically analyze emerging face unmasking techniques under occlusion scenarios. From a technical standpoint, Mahmoud et al. [93] provided a comprehensive survey of masked face detection, recognition, and unmasking techniques. Their study systematically reviews deep learning architectures, benchmark datasets such as MAFA and RMFD, and evaluation metrics including accuracy, PSNR, and SSIM. The survey identifies challenges related to occlusion, dataset imbalance, and robustness under real-world conditions, while also outlining future research directions in masked face analysis.

III. DATASETS FOR MASKED FACE INPAINTING AND RECOGNITION

Masked face recognition and inpainting systems have become increasingly important, especially in the context of global events such as the COVID-19 pandemic. These models aim to either recognize identities from partially occluded faces or reconstruct missing facial regions (e.g., mouth, nose) using generative methods. A foundational aspect of evaluating and benchmarking such models lies in the diversity and realism of the datasets used for training and testing. A wide spectrum of datasets has been curated or modified to simulate both real-world and synthetic mask occlusions. These datasets vary in size, resolution, subject diversity, mask types, and the method of mask application (real or artificial). Proper dataset selection significantly influences the generalizability and robustness of MFR algorithms. The CelebFaces Attributes Dataset (CelebA) is one of the most extensively utilized benchmarks in face analysis tasks. Comprising over 200,000 celebrity images with rich attribute annotations, it serves as a common training corpus for inpainting models. The dataset does not originally include masked samples, but synthetic occlusions can be algorithmically overlaid to simulate masked conditions. Many studies, including [94], have adopted CelebA with artificial masks to train and evaluate GAN-based inpainting models. The dataset's large scale and high variability across poses and illumination conditions make it ideal for learning generalized representations.

Labeled Faces in the Wild (LFW) is another seminal dataset, comprising face images collected from the internet for unconstrained face verification. While not initially designed for MFR, it is often repurposed by overlaying masks either manually or through automated methods. The inclusion of both synthetic and real masked versions allows for comparative evaluation. Its widespread usage, as reported in [95], stems from its realistic distribution and historical relevance in face verification benchmarks. Because LFW features in-the-wild images with substantial variability, it introduces useful challenges in masked face analysis. The Flickr-Faces-HQ (FFHQ) dataset contains over 70,000 high-resolution face images with diverse variations in age, ethnicity, lighting, and expression. While FFHQ does not include real-world masked samples, it is a preferred base for synthetic mask overlays due to its image quality. Recent studies such as [96] utilize FFHQ for training face reconstruction and inpainting systems. The dataset supports learning finer details, which is crucial in scenarios where GAN-based models aim to restore occluded facial components while preserving identity-related features. Masked Face Dataset (MAFA) is a real-world masked face dataset containing over 30,000 images annotated for mask presence and location.

Designed specifically for mask detection and recognition, it represents one of the few datasets with native mask occlusions. Used in studies like [97], MAFA bridges the gap between simulated conditions and real-world deployment. It introduces challenges such as occlusion variance, complex backgrounds, and mask style diversity that are hard to simulate synthetically. The Public Figures dataset, also known as PubFig83, is composed of images of 83 celebrities captured under diverse conditions. While not originally containing masked faces, it has been repurposed in the MFR domain with synthetic mask applications. Its value lies in the inherent difficulty introduced by pose variation, low resolution, and lighting diversity. As seen in [98], it provides a semi-realistic yet controlled environment for testing model robustness against visual noise and occlusion. MaskedFace-CelebA dataset is a synthetically modified version of CelebA, where masks are digitally applied to the original faces to simulate pandemic-era conditions. It aims to maintain identity information while offering occlusion realism. The dataset is often employed as a testbed for evaluating models trained on CelebA or other synthetic sources. The work in [99] demonstrates its effectiveness in assessing how well models can adapt to real-world masked conditions based on synthetic training. Figure 3 demonstrates few samples from prominent face datasets.



Figure 3: Sample images from various datasets.

Source: [100].

Face5 is a hybrid dataset constructed by merging real and simulated masked facial samples. It allows for studying the generalization capabilities of MFR systems when subjected to a mixture of mask types and occlusion styles.

Proposed in studies such as [101], the hybrid nature of Face5 makes it an important bridge for transition between purely synthetic and completely real-world datasets. SyntheticMaskSet is explicitly designed to support benchmarking of face inpainting and recognition systems under controlled synthetic conditions. By applying diverse types of digitally generated masks to faces, it enables standardized evaluation across multiple models. For instance, the Region-Attentive GAN proposed in [102] demonstrated strong performance on this dataset due to the consistent and controllable occlusion levels. SimulMaskDB is a synthetically generated dataset that varies mask styles, occlusion levels, and facial expressions in a controlled fashion. It is primarily used for stress-testing recognition and inpainting algorithms under artificial yet diverse conditions. As highlighted in [103], this dataset enables fine-grained performance analysis, especially for lightweight architectures and constrained environments. Also, studies review face recognition techniques, focusing on feature extraction methods and key challenges affecting system performance and accuracy [104], [105]. Table 6 highlights the demographic diversity and dataset characteristics used to evaluate different masked face inpainting models. Clear reporting of age groups, gender balance, and ethnicity ensures fair comparison and helps identify potential dataset bias in GAN-based recognition systems.

Table 6: Summary of Datasets Characteristics and Demographic Parameters Used in Masked Face Inpainting and Recognition.

Study	Database	Sample Size	Age Groups	Gender Distribution	Ethnicity / Other Factors	Description of Database
Mahmoud and Kang (2023) [102]	GAN-Masker Dataset	Synthetic dataset (based on CelebA and other sources)	Adult subjects	Mixed gender	Artificially masked faces	Two-stage GAN-based masked face removal dataset combining real and synthetically generated masked face images for restoration and recognition research.
Tero Karras et al. (2019) [96]	FFHQ	70,000 images	Wide age Range (children to elderly)	Balanced male/female	Diverse ethnicities; high-quality aligned images	High-resolution (1024×1024) face dataset curated for GAN training; diverse in age, ethnicity, and accessories, commonly used in generative modeling research.
Shiming Ge et al. (2017) [97]	MAFA	30,811 images	Adult subjects	Mixed gender	Masked faces in the wild; occlusion variations	Dataset focused on masked face detection with varying mask types, occlusion levels, and complex backgrounds in real-world conditions.
Ziwei Liu et al. (2015) [94]	CelebA	202,599 images (10,177 identities)	Young to middle-aged adults	Male and Female	Celebrities; large variation in pose and background	Large-scale face attributes dataset collected from the internet with 40 annotated facial attributes; widely used for face recognition and attribute prediction tasks.
Bergstra et al. (2013) [98]	PubFig83	Not dataset-specific	Not specified	Not specified	Hyperparameter optimization focus	Study emphasizes hyperparameter optimization for vision architectures rather than introducing a specific face database.
Gary B. Huang et al. (2008) [95]	LFW	13,233 images (5,749 identities)	Adult subjects	Male-dominant distribution	Unconstrained environments; real-world variations	Unconstrained face dataset designed for studying face recognition under real-world conditions including variations in pose, illumination, and expression.

Source: Authors, (2026).

Synthetic datasets, such as CelebA, FFHQ, and SyntheticMaskSet, are instrumental in benchmarking inpainting models due to their controlled conditions. In contrast, real-world datasets like MAFA and RMFD provide essential realism for evaluating system generalizability. Among these, CelebA and LFW emerge as the most recurrently utilized in synthetic masked face inpainting tasks. For instance, the RegionAttentive GAN introduced and demonstrated high reconstruction fidelity when evaluated on the SyntheticMaskSet.

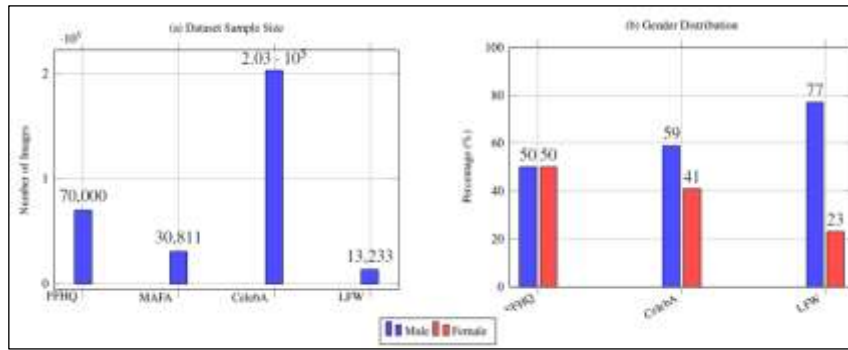


Figure 4: Comparison of dataset sample size and gender distribution across commonly used face datasets.

Source: Authors, (2026).

In Figure 4 (a), the bar chart shows the number of images available in each dataset. CelebA clearly has the largest number of images, making it very useful for large-scale training. FFHQ also contains a good number of high-quality images. MAFA has a moderate dataset size focused on masked faces, while LFW contains the smallest number of images among the four. This comparison helps us understand which datasets are large enough for deep learning experiments. In Figure 4 (b), the chart shows the gender distribution within selected datasets. FFHQ appears balanced, with nearly equal numbers of male and female images. However, CelebA and LFW show noticeable imbalance, with a higher percentage of male images compared to female images. This imbalance is important because it may affect model fairness and performance. Overall, the figure helps us understand both the size and diversity of these datasets, which are important factors when selecting data for face restoration and recognition research.

IV. CONCLUSIONS

Masked face restoration using Generative Adversarial Networks (GANs) has emerged as a vital research area in response to the challenges posed by real-world occlusions such as face masks. Through this survey, we examined state-of-the-art methods addressing detection, inpainting, and recognition under mask occlusion. While existing methods demonstrate significant progress in restoring visually plausible and identity-preserving face images, several limitations and open challenges remain. First, most current models rely heavily on synthetic masks or artificially generated occlusions during training, which often fail to fully represent the diversity and irregularity of real-world mask usage. Consequently, generalization to real masked faces remains inconsistent across datasets and domains. Second, although recent GAN models incorporate identity and perceptual constraints, maintaining fine-grained facial details—particularly in highly occluded or low-quality images—remains a bottleneck for both recognition and downstream biometric tasks. In terms of detection and inpainting performance, many models are not jointly optimized for both tasks, leading to cascaded error propagation. Moreover, evaluation metrics such as PSNR and SSIM, while useful for measuring image similarity, do not always correlate with recognition accuracy or semantic identity preservation. Therefore, more robust and identity-aware metrics, including perceptual loss functions and recognition-oriented benchmarks, are required. Furthermore, the integration of multimodal data, temporal consistency (for video-based restoration), and domain adaptation for unseen environments represents fertile ground for future work. Lightweight architectures that balance inference speed and accuracy are also needed for deployment in real-time or edge devices.

V. AUTHOR'S CONTRIBUTION

Conceptualization: Payal Parekh, Hina Choksi.

Methodology: Payal Parekh, Mahesh Goyani, Sheilla Ann Pacheco.

Investigation: Payal Parekh, Hina Choksi, Mahesh Goyani.

Discussion of results: Payal Parekh, Hina Choksi, Mahesh Goyani, Sheilla Ann Pacheco.

Writing – Original Draft: Payal Parekh, Mahesh Goyani.

Writing – Review and Editing: Payal Parekh, Hina Choksi, Mahesh Goyani, Sheilla Ann Pacheco.

Resources: Hina Choksi, Mahesh Goyani, Sheilla Ann Pacheco.

Supervision: Hina Choksi, Mahesh Goyani, Sheilla Ann Pacheco.

Approval of the final text: Payal Parekh, Mahesh Goyani.

VI. REFERENCES

- [1] H. Ali and A. Ijaz, "Machine and deep learning based cctv surveillance using facenet, mtncnn, and haar cascade for enhanced security," in 2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL), 2025, pp. 672–677.
- [2] M. Ngan, P. Grother, and K. Hanaoka, "Ongoing face recognition vendor test (frvt) part 6a: Face recognition accuracy with masks using pre-covid-19 algorithms," NIST Interagency/Internal Report, National Institute of Standards and Technology, 2020.
- [3] L. Qingyu, C. Lei, S. Yeguo, and L. Lei, "Masked face image inpainting based on generative adversarial network," Infocommunications Journal, vol. 17, no. 2, pp. 11–19, 2025.
- [4] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic," Measurement, vol. 167, p. 108288, 2021.
- [5] S. A. B. Pacheco, J. E. Estrada, and M. M. Goyani, "Hidden adversarial attack on facial biometrics-a comprehensive survey," Procedia Computer Science, vol. 258, pp. 1383–1390, 2025.

- [6] P. Sheilla Ann Bangoy, E. Jheanel Espiritu and G. Mahesh, "Robust face recognition under adversarial attack using sargan model and improved cross triple mobilenetv1," in Future of Information and Communication Conference. Springer, 2025, pp. 491–510.
- [7] H. Trivedi and M. Goyani, "Robust face recognition in the presence of diverse challenges: A hybrid deep neural network approach," Journal of Engineering, vol. 14, no. 10, pp. 55–62, 2024.
- [8] T. Himani and G. Mahesh, "A transition of face recognition to mask face recognition using improvised attention dl model," in 2023 3rd International Conference on Intelligent Technologies (CONIT). IEEE, 2023, pp. 1–7.
- [9] T. Himani and G. Mahesh, "Adapting face recognition to the masked world: leveraging deep attention networks with a custom dataset," International Journal of Computers and Applications, vol. 46, no. 6, pp. 427–441, 2024.
- [10] A. Alzu'bi, F. Albalas, T. Al-Hadhrani, L. B. Younis, and A. Bashayreh, "Masked face recognition using deep learning: A review," Electronics, vol. 10, no. 21, p. 2666, 2021.
- [11] N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, and A. Kuijper, "The effect of wearing a mask on face recognition performance: an exploratory study," in 2020 International Conference of the Biometrics Special Interest Group (BIOSIG). IEEE, 2020, pp. 1–6.
- [12] N. Damer, J. Grebe, J. Chen, F. Boutros, F. Kirchbuchner, and A. Kuijper, "Extended evaluation of face recognition with masks: Effect of real and simulated masks on face embeddings," IET Biometrics, vol. 10, no. 6, pp. 607–618, 2021.
- [13] J. Wang, Y. Kang, Y. Zhang, and F. Li, "Masked face recognition dataset and application," arXiv preprint arXiv:2003.09093, 2020.
- [14] Y.-C. Huang, D. A. B. Rahardjo, R.-H. Shiue, and H. H. Chen, "Masked face recognition using domain adaptation," Pattern Recognition, vol. 153, p. 110574, 2024.
- [15] C. Agarwal and C. Bhatnagar, "Unmasking the potential: evaluating image inpainting techniques for masked face reconstruction," Multimedia Tools and Applications, vol. 83, no. 1, pp. 893–918, 2024.
- [16] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, "Image super-resolution via iterative refinement," IEEE transactions on pattern analysis and machine intelligence, vol. 45, no. 4, pp. 4713–4726, 2022.
- [17] Z. Yue and C. C. Loy, "Difface: Blind face restoration with diffused error contraction," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 12, pp. 9991–10 004, 2024.
- [18] Z. Wu, J. Chen, L. Tan, H. Gong, Y. Zhou, and G. Shi, "A lightweight gan-based image fusion algorithm for visible and infrared images," in 2024 4th international conference on computer science and blockchain (CCSB). IEEE, 2024, pp. 466–470.
- [19] P. Guo, "Magan-rt: A lightweight adversarial style transfer network for real-time cartoonization on low-power edge devices," Internet Technology Letters, vol. 8, no. 5, p. e70104, 2025.
- [20] N. K. Yadav, S. K. Singh, and S. R. Dubey, "Tva-gan: attention guided generative adversarial network for thermal to visible image transformations," Neural Computing and Applications, vol. 35, no. 27, pp. 19 729–19 749, 2023.
- [21] D. Sheng, W. Jin, M. Wang, and J. Yang, "Region-based self-segmentation guided diffusion model for thermal infrared to pseudo-color visible light image conversion," IEEE Access, vol. 13, pp. 47 860–47 873, 2025.
- [22] K. Wang, Y. Shen, and M. Lauer, "Adversarial defense teacher for cross-domain object detection under poor visibility conditions," arXiv preprint arXiv:2403.15786, 2024.
- [23] A. H. Abdul Amir and Z. N. Nemer, "Inclusive review on advances in masked human face recognition technologies," Iraqi Journal of Intelligent Computing and Informatics (IJICI), vol. 4, no. 1, pp. 1–17, 2025.
- [24] V. S. Patel, Z. Nie, T.-N. Le, and T. V. Nguyen, "Masked face analysis via multi-task deep learning," Journal of Imaging, vol. 7, no. 10, p. 204, 2021.
- [25] J. A. Alzubi and Others, "A generative adversarial network-based accurate masked face recognition model using dual scale adaptive efficient attention network," Scientific Reports, vol. 15, 2025.
- [26] Y. Li, J. Wang, W. Liang, H. Xue, Z. He, J. Lv, and L. Zhang, "Cr-gan: Automatic craniofacial reconstruction for personal identification," Pattern Recognition, vol. 124, p. 108400, 2022.
- [27] Z. Chen, Z. Zhang, J. Yuan, Y. Xu, and L. Liu, "Show your face: Restoring complete facial images from partial observations for vr meeting," in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2024, pp. 8688–8697.
- [28] L. Song, D. Gong, Z. Li, C. Liu, and W. Liu, "Occlusion robust face recognition based on mask learning with pairwise differential siamese network," in Proceedings of the IEEE/CVF international conference on computer vision, 2019, pp. 773–782.
- [29] S. Biswas and Others, "A new perceptual hashing method for verification and identity classification of occluded faces," Image and Vision Computing, vol. 113, p. 104078, 2021.
- [30] J. Deng, J. Guo, J. Yang, and N. Xue, "Mf-cosface: A margin-based loss for masked face recognition," Applied Sciences, vol. 11, no. 4, p. 1813, 2021.
- [31] W. Hariri, "Efficient masked face recognition method during the covid-19 pandemic," arXiv preprint arXiv:2105.03026, 2021.
- [32] A. M. Sheneamer et al., "A hybrid human recognition framework using machine learning and deep learning for masked faces," Plos One, vol. 19, no. 6, p. e0304567, 2024.
- [33] B. Adeli, J. McInden, P. Pandey, M. Shao, and Y. Shahriari, "Absolutenet: A deep learning neural network to classify cerebral hemodynamic responses of auditory processing," 2025 IEEE 35th International Workshop on Machine Learning for Signal Processing, pp. 1–6, 2025.
- [34] Pillay, C., Joseph, S., & Van Niekerk, B. (2025). Development of a face mask detection and masked facial recognition model based on a hybrid convolutional neural network. IET Image Processing, 19(1), e70239.

- [35] A. Cabani, K. Hammoudi, H. Benhabiles, and M. Melkemi, "Maskedface-net: A dataset of correctly/incorrectly masked face images in the context of covid-19," *Smart Health*, vol. 19, p. 100144, 2020.
- [36] X. Fan and M. Jiang, "Retinaface mask: A single stage face mask detector for assisting control of the covid-19 pandemic," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2021.
- [37] K. Dhanushkodi, A. N. Hussain, A. Jayaprakash, and S. Suriyaprakash, "A survey on detection of face mask and social distancing using deep learning," *i-manager's Journal of Information Technology*, vol. 10, p. 22, 2021.
- [38] D. Chiang, "Detect faces and determine whether people are wearing mask," *Face Mask Detection*, vol. 15, 2020.
- [39] M. Loey, G. Manogaran, M. Taha, and N. Khalifa, "Fighting against covid-19: A novel deep learning model based on yolo-v2 with resnet-50 for medical face mask detection," *Sustainable Cities and Society*, vol. 65, p. 102600, 2021.
- [40] X. Su, M. Gao, J. Ren, Y. Li, M. Dong, and X. Liu, "Face mask detection and classification via deep transfer learning," *Multimedia Tools and Applications*, vol. 81, pp. 4475–4494, 2022.
- [41] World Health Organization, "Advice for the public on covid-19," <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public>, 2022, accessed February 10, 2026.
- [42] D. Yang, E. Yurtsever, V. Renganathan, K. A. Redmill, and U. O' zgu'ner, "A vision-based social distancing and critical density detection system for covid-19," *Sensors*, vol. 21, p. 4608, 2021.
- [43] S. Madane and D. Chitre, "Social distancing detection and analysis through computer vision," in *2021 6th International conference for convergence in technology (I2CT)*. IEEE, 2021, pp. 1–10.
- [44] I. Ahmed, M. Ahmad, J. J. Rodrigues, G. Jeon, and S. Din, "A deep learning-based social distance monitoring framework for covid-19," *Sustainable Cities and Society*, vol. 65, p. 102571, 2021.
- [45] S. Suryadi, E. Kurniawan, H. Adinanta, B. H. Sirenden, J. A. Prakosa, and P. Purwobowono, "On the comparison of social distancing violation detectors with gpu support," in *2020 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 2020, pp. 337–342.
- [46] M. Rezaei and M. Azarmi, "Deepsocial: Social distancing monitoring and infection risk assessment in covid-19 pandemic," *Applied Sciences*, vol. 10, p. 7514, 2020.
- [47] K. A. Reddy, P. Sandeep, and D. Gayathri, "Monitoring covid-19 social distancing with person detection and tracking via yolov3 and deepsort techniques," *International Journal of Innovative Research in Technology*, vol. 8, no. 11, pp. 1–5.
- [48] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with lle-cnns," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2682–2690.
- [49] Z. Wang, B. Huang, G. Wang, P. Yi, and K. Jiang, "Masked face recognition dataset and application," *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 5, no. 2, pp. 298–304, 2023.
- [50] X. Jiang, T. Gao, Z. Zhu, and Y. Zhao, "Real-time face mask detection method based on yolov3," *Electronics*, vol. 10, p. 837, 2021.
- [51] F. I. Eyiokur, H. K. Ekenel, and A. Waibel, "Unconstrained face-mask face-hand datasets: Building a computer vision system to help prevent the transmission of covid-19," *arXiv preprint arXiv:2103.08773*, 2021.
- [52] B. Wang, Y. Zhao, and C. P. Chen, "Hybrid transfer learning and broad learning system for wearing mask detection in the covid-19 era," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.
- [53] I. Issaoui and A. Selmi, "Weighted soft discernibility matrix with deep learning assisted face mask detection for smart city environment," *International Journal of Neutrosophic Science (IJNS)*, vol. 25, no. 1, 2025.
- [54] Y. Shao, J. Ning, H. Shao, D. Zhang, H. Chu, and Z. Ren, "Lightweight face mask detection algorithm with attention mechanism," *Engineering Applications of Artificial Intelligence*, vol. 137, p. 109077, 2024.
- [55] C. Pillay, S. Joseph, and B. Van Niekerk, "Development of a face mask detection and masked facial recognition model based on a hybrid convolutional neural network," *IET Image Processing*, vol. 19, no. 1, p. e70239, 2025.
- [56] M. Ahmadi, M. D. Ismaeil, J. C. Verdial, A. R. Nafchi, and M. Esmaeili, "Building a uc treatment response classifier: A random forest classifier with key biomarker identification," in *2025 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2025, pp. 0963–0970.
- [57] M. Mohammadagha, A. Khoshkonesh, S. Sharifi, V. Ghanbarizadeh, and F. Bigdeli, "Performance analysis of hybrid deep-transfer learning approaches with machine learning methods for face recognition," *University of Texas at Arlington*, 2025.
- [58] Z. Gao and Others, "Improved face mask wearing detection based on yolov5," *International Journal of Advanced Network, Monitoring and Controls*, vol. 10, no. 1, pp. 94–102, 2025.
- [59] Y. Wei, H. Li, Y. He, and Others, "Robust face mask detection in complex scenarios using yolov8 and context-aware convolutions," *Scientific Reports*, vol. 15, p. 21350, 2025.
- [60] J. Yu, X. Zhang, T. Wu, H. Pan, and W. Zhang, "A face detection and standardized mask-wearing recognition algorithm," *Sensors*, vol. 23, no. 10, p. 4612, 2023.
- [61] J. V. B. Benifa, C. Chola, A. Y. Muaad, and Others, "Fmdnet: An efficient system for face mask detection based on lightweight model," *Sensors*, vol. 23, no. 13, p. 6090, 2023.
- [62] J. Bhagat and M. Goyani, "Masked face recognition by reconstructing face using image inpainting with gan," in *Recent Trends in Computational Intelligence and Its Application*. CRC Press, 2023, pp. 26–34.
- [63] Q. Man and Y.-I. Cho, "Efficient face region occlusion repair based on t-gans," *Electronics*, vol. 12, no. 10, p. 2162, 2023.
- [64] H. Li, Y. Zhang, W. Wang, S. Zhang, and S. Zhang, "Recovery-based occluded face recognition by identity-guided inpainting," *Sensors*, vol. 24, no. 2, p. 394, 2024.

- [65] A. Hassanpour, F. Jamalbafrani, B. Yang, K. Raja, R. Veldhuis, and J. Fierrez, "E2f-net: Eyes-to-face inpainting via stylegan latent space," *Pattern Recognition*, vol. 152, p. 110442, 2024.
- [66] E. A. Zhang, "Md-gan: Multi-scale diversity gan for large masks inpainting," *Electronics*, vol. 14, no. 11, p. 2218, 2024.
- [67] N. A. Asad, M. A. M. Pranto, S. Shiam, M. M. Akand, M. A. Yousuf, K. F. Hasan, and M. A. Moni, "Mgan-crcm: a novel multiple generative adversarial network and coarse refinement-based cognizant method for image inpainting," *Neural Computing and Applications*, vol. 37, no. 7, pp. 5459–5480, 2025.
- [68] S. K. S. G. Pechetti, C. Lella, L. V. P. R. Medarametla, and V. Mandadi, "Image inpainting using generative adversarial networks and context-aware networks," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 13, no. 6, 2025.
- [69] S. Sola and D. Gera, "Unmasking your expression: Expression-conditioned gan for masked face inpainting," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2023, pp. 5908–5916.
- [70] X. Gao, M. Nguyen, and W. Q. Yan, "A method for face image inpainting based on autoencoder and generative adversarial network," in *Pacific-Rim Symposium on Image and Video Technology*. Springer, 2022, pp. 24–36.
- [71] F. Farahanipad, M. Rezaei, M. Nasr, F. Kamangar, and V. Athitsos, "Gan-based face reconstruction for masked-face," in *Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments*, 2022, pp. 583–587.
- [72] J. A. Alzubi, K. S. Pokkuluri, R. Arunachalam, S. K. Shukla, S. Venugopal, and K. Arunachalam, "A generative adversarial network-based accurate masked face recognition model using dual scale adaptive efficient attention network," *Scientific Reports*, vol. 15, no. 1, p. 17594, 2025.
- [73] Q. Liu, L. Chen, Y. Sun, and L. Liu, "Masked face image inpainting based on generative adversarial network," *Infocommunications Journal*, vol. 17, no. 2, pp. 11–19, 2025.
- [74] S. K. S. G. Pechetti, C. Lella, L. V. P. R. Medarametla, and V. Mandadi, "Image inpainting using generative adversarial networks and contextual attention," *International Journal of Creative Research Thoughts*, vol. 13, no. 6, pp. 2320–2882, 2025. [Online]. Available: <https://www.ijcrt.org/papers/IJCRT2506738.pdf>
- [75] A. Kumar, M. Kaushal, and A. Sharma, "Sam c-gan: a method for removal of face masks from masked faces," *Signal, Image and Video Processing*, vol. 17, no. 7, pp. 3749–3757, 2023.
- [76] N. E. Apostolou, E. V. Balourdou, M. Mouratidou, E. Tsaleri, I. Voyiatzis, A. Papadakis, and M. Samarakou, "Efficient and secure gans: A survey on privacy-preserving and resource-aware models," *Applied Sciences*, vol. 15, no. 20, p. 11207, 2025.
- [77] A. Musa, H. Kakudi, M. Hassan, M. L. Salisu et al., "Lightweight deep learning models for edge devices—a survey," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 17, p. 18, 2025.
- [78] C. Gong, Y. Dai, R. Li, A. Bao, J. Li, J. Yang, Y. Zhang, and X. Li, "Text2avatar: Text to 3d human avatar generation with codebook-driven body controllable attribute," in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2024, pp. 16–20.
- [79] A. Melnik, M. Miasayedzenkau, D. Makaravets, D. Pirshutuk, E. Akbulut, D. Holzmann, T. Renusch, G. Reichert, and H. Ritter, "Face generation and editing with stylegan: A survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 46, no. 5, pp. 3557–3576, 2024.
- [80] C. Saharia, J. Ho, W. Chan, T. Salimans, D. J. Fleet, and M. Norouzi, "Image super-resolution via iterative refinement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 4, pp. 4713–4726, 2023.
- [81] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, "Repaint: Inpainting using denoising diffusion probabilistic models," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 11 451–11 461.
- [82] Z. Wang, Z. Zhang, X. Zhang, H. Zheng, M. Zhou, Y. Zhang, and Y. Wang, "Dr2: Diffusion-based robust degradation remover for blind face restoration," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2023, pp. 1704–1713.
- [83] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "Swinir: Image restoration using swin transformer," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 1833–1844.
- [84] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 5728–5739.
- [85] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu, and H. Li, "Uformer: A general u-shaped transformer for image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 17 683–17 693.
- [86] Y. Wang, T. Lu, Y. Zhang, J. Jiang, J. Wang, Z. Wang, and J. Ma, "Tanet: A new paradigm for global face super-resolution via transformer-cnn aggregation network," *arXiv preprint arXiv:2109.08174*, 2021.
- [87] N. K. Tiwari, S. S. Rajput, and R. Patel, "Cctfacenet: Enhancing face super-resolution with cascaded cnn-transformer and dual-path feature fusion," *Digital Signal Processing*, p. 105557, 2025.
- [88] A. Li, G. Li, L. Sun, and X. Wang, "Faceformer: Scale-aware blind face restoration with transformers," *arXiv preprint arXiv:2207.09790*, 2022.
- [89] A. M. Ali, B. Benjdira, A. Koubaa, W. El-Shafai, Z. Khan, and W. Boulila, "Vision transformers in image restoration: A survey," *Sensors*, vol. 23, no. 5, p. 2385, 2023.
- [90] D. Almeida, K. Shmarko, and E. Lomas, "The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of us, eu, and uk regulatory frameworks," *AI and Ethics*, vol. 2, no. 3, pp. 377–387, 2022.
- [91] M. Smith and S. Miller, "The ethical application of biometric facial recognition technology," *Ai & Society*, vol. 37, no. 1, pp. 167–175, 2022.
- [92] X. Wang, Y. C. Wu, M. Zhou, and H. Fu, "Beyond surveillance: privacy, ethics, and regulations in face recognition technology," *Frontiers in Big Data*, vol. Volume 7 - 2024, 2024.
- [93] M. Mahmoud, M. S. Kasem, and H. Kang, "A comprehensive survey of masked faces: Recognition, detection, and unmasking," *Applied Sciences*, vol. 14, no. 19, p. 8781, 2024.

- [94] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in Proceedings of the IEEE international conference on computer vision, 2015, pp. 3730–3738.
- [95] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," in Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition, 2008.
- [96] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 4401–4410.
- [97] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detecting masked faces in the wild with l1e-cnns," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 2682–2690.
- [98] J. Bergstra, D. Yamins, and D. Cox, "Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures," in International conference on machine learning. PMLR, 2013, pp. 115–123.
- [99] C. Agarwal, C. Bhatnagar, and A. Mishra, "Evaluation of image inpainting methods for face reconstruction of masked faces," in 2023 International Conference on Electrical, Communication and Computer Engineering (ICECCE). IEEE, 2023, pp. 1–6.
- [100] M. Mahmoud, M. S. Kasem, and H.-S. Kang, "A comprehensive survey of masked faces: recognition, detection, and unmasking," Applied Sciences, vol. 14, no. 19, p. 8781, 2024.
- [101] M. I. Hosen and M. B. Islam, "Himfr: A hybrid masked face recognition through face inpainting," arXiv preprint arXiv:2209.08930, 2022.
- [102] M. Mahmoud and H.-S. Kang, "Ganmasker: A two-stage generative adversarial network for high-quality face mask removal," Sensors, vol. 23, no. 16, p. 7094, 2023.
- [103] Y. Zhou and T. Lu, "Masked face restoration model based on lightweight gan," Computers, Materials, & Continua, vol. 82, no. 2, p. 3591, 2025.
- [104] P. Payal and M. M. Goyani, "A comprehensive study on face recognition: methods and challenges," The Imaging Science Journal, vol. 68, no. 2, pp. 114–127, 2020.
- [105] P. Parekh and G. Mahesh, "Face recognition modern assessment of features extraction," International Journal of Computer Sciences and Engineering, vol. 7, pp. 330–335, 04 2019.