Societal Transformation: AI and Big Data Journal
Vol: 4(1): 27- 47, 2026
DOI:10.20547/aibd.264102

Face Recognition under Occlusion: A Systematic Review of Methods, Datasets, and Evaluation Practices (2021–2026)

Yumna Shahzad[1], Muhammad Khalid Khan[2],Areeba Raza[3],Muzmmil Memon[4]

Abstract

Recognition of faces with partial occlusions poses an exceptionally difficult challenge for contemporary researchers working on biometric identification tasks. Various real-life use cases in areas such as border crossings, healthcare services, smart surveillance, etc., usually require performing face identification despite masks, sunglasses, scarfs, and similar accessories blocking parts of the face, causing poor recognition performance. In this study, we conduct a systematic analysis of occluded face recognition research works that have been carried out between 2021 and 2026. After the application of inclusion/exclusion criteria to our initial database consisting of relevant peer-reviewed articles in IEEE Xplore, ScienceDirect, Springer, and arXiv databases, 35 papers were chosen for the further discussion. In order to organize reviewed works systematically, current techniques can be divided into five groups: occlusion-aware face recognition, occlusion-robust feature extraction techniques, methods of face occlusion recovery, transformer models, and multimodal systems. Existing benchmarks and metrics such as Accuracy, F1-Score, FAR/FRR, ROC-AUC, EER are assessed. We observe a significant performance improvement in 2025 from CNN baselines (~88% of accuracy in 2022) to transformer/multimodal solutions that achieve more than 94%. Open challenges such as data imbalance, lack of realism, high computational cost, and ethical issues are discussed.

Keywords: *Face recognition, Face Occlusion, Deep Neural Networks, Multimodal method, Transformer model, Systematic Review*

[1] *Karachi Institute of Economics and Technology (KIET)* yumna@kiet.edu.pk
[2] *Karachi institute of economics and technology (KIET)*Khalid.khan@kiet.edu.pk
[3] *University Putra Malaysia*Areebah.rz@gmail.com
[4] *Avila University, USA* Memon521057@avila.edu

1. Introduction

Face recognition (FR) technology has found its applications in surveillance, secure identification, and human-computer interaction, becoming one of the most widely used biometric techniques. Despite recent improvements, modern face recognition algorithms still suffer from severe degradation when deployed in unconstrained settings, where illumination changes, pose variations, and especially face occlusion can cause a considerable drop in accuracy, whereas they perform close to human-level performance in constrained settings.

Face occlusion results from the face being partially or completely hidden by face masks, sunglasses, scarves, or even body parts such as hands and hair that cover significant facial characteristics like the mouth, nose, and eyes. Therefore, extracted features possess weaker discriminative power and incomplete facial representations. According to some research findings, face occlusion can conceal a lot of identity information, hindering accurate identification in real-world scenarios. ([1], [4]).

With the extensive use of masks during the COVID-19 pandemic, the need for an occlusion-aware face recognition system has been significantly emphasised. The shortcomings of traditional deep learning algorithms, such as convolutional neural networks, which are sensitive to facial segmentations due to their dependence on full-face information, have become apparent [12].
Recent literature suggests that occlusions adversely affect recognition rates by reducing discriminant features and raising intra-class variance [8]. Additionally, emerging research (between 2021 and 2026) has focused on designing occlusion-tolerant architectures, which would address these challenges. These architectures range from transformer-based systems, which incorporate long-range dependencies between facial features, reconstruction-based designs, which seek to reconstruct missing parts, and attention-driven models, which concentrate on facial segments. In addition, novel approaches such as multimodal fusion and self-supervised learning have shown promising results in improving generalisation when occlusions go unnoticed [5].
Some of the challenges, such as the lack of real-world data, the generalisation of occlusions across different types, and the complexity of model implementation in real-time environments, remain to be resolved despite these advancements.. This highlights the necessity of a thorough and organised analysis of current advancements in occluded face recognition.

Contributions: Specifically, this work will contribute to the literature through: (1) a classification-based taxonomy of occlusion handling techniques between 2021 to 2026 in five distinct categories; (2) a comparative analysis of their performance results on specific datasets; (3) an insight into areas of research that remain largely unexplored; (4) a research agenda for the development of new approaches for overcoming facial occlusions; and (5) the ethical consequences of employing face recognition software in society.

Organization: The rest of this paper proceeds as follows. Background and problem formulation is covered in Section 2. Section 3 discusses the review methodology and includes the analysis of literature within five different methods categories, with a dedicated section on research gaps. Section 4 discusses relevant datasets and evaluation metrics. Section 5 explores future research directions. Section 6 concludes the paper

2. Background and Problem Formulation

2.1 Occluded Face Recognition Definition

Occluded face recognition refers to the process of recognizing or confirming an individual with the help of occluded facial landmarks. An occluded face recognition process requires a system to operate with partial view faces, making the feature matching task a complex process as compared to normal face recognition, where full face details are provided to the system.

2.2 Occlusion Classification

The following types of occlusions can be considered:

1. Structured Occlusion

• Caused due to objects such as helmets, masks, or glasses
• Follows a particular pattern
• Supervised learning is easier

Predictable structured occlusions have been extensively examined in the literature because of the regularity of their spatial configuration. An example of an occlusion that has become very popular since the COVID-19 pandemic is the face mask occlusion. Due to its uniformity in terms of pattern and location, training machine learning models on such occlusions using augmentation becomes feasible [4, 12].

2. Unstructured Occlusion

• Comprises of random objects such as hands, hairs, or any other objects
• Highly erratic and unpredictable
• Generalization of models becomes difficult

Unstructured occlusions are much harder to deal with since they have no definite spatial structure. An occluder might be of a different size, color, shape, and texture. That poses a serious problem for training neural networks using one particular type of occlusion to predict others, which is currently pointed out by research in the field [7, 14].

3. Occlusion by Environment

• Occurs due to noise generated by a camera, shadows, or changing lightings

Environmental occlusion is not due to the physical obstruction of some object, but it refers to the conditions present in imaging such that facial data cannot be fully captured. Examples include images with low resolution, motion blur, and high contrast lighting that prevent the acquisition of clear images. This kind of occlusion is especially relevant in surveillance [6, 7].

As per recent findings, it has been reported that unstructured occlusion poses difficulty for face recognition, due to the presence of erratic noise and lack of spatial structure patterns [3]. Figure 1 below shows different types of occlusion.

Figure 1: Types of Facial Occlusion (adapted from [7, 14])

2.3 Occlusion's Effect on Feature Learning

Face recognition algorithms can be affected by occlusion in several simple ways:

• Loss of critical identifying attributes
• Distortion of features
• Lower separability between different classes

• Increased variance within a class

Each of these impacts contributes directly to the discrimination of the feature representations that have been learned. For instance, the absence of key features such as the periocular region or the nose bridge may significantly impair the ability to recognize someone's identity because these areas contain very important discriminating elements. Distortion of features is caused by the operation of convolutional filters that have learned to detect entire faces, yet end up giving erroneous responses upon occlusion.

As a result of distribution mismatch, caused by an incongruity between training data and actual occlusion cases, deep learning algorithms often perform poorly in occluded scenarios [2].

2.4 Formulation of Problems

The goal of identifying an occluded face using the input image I with occlusion O is stated below:

• Extraction of robust feature representation f(I-O).
• Compensating for identity loss caused due to regions missing.
Even in the presence of occlusion, achieving similarity among identities. This can be achieved by:
• Feature Selection (ignoring the occluded regions).
• Reconstruction (recovering the occluded regions).
• Global Context modeling.

Let us introduce notation, where I is the complete face image and O is the occlusion binary mask so that I_obs = I $\odot$ (1-O) is the observed part of the image. For recognition, the goal is to achieve min(d(f(I_obs), f(I_ref))) $\leq \delta$ for pairs with the same identity while maximizing the distance between feature representations for dissimilar identities. The methods under consideration, feature selection, reconstruction, and global context modeling, differ from one another as the approaches towards missing information in I_obs: ignore it, recover it, or compensate for it using contextual representations.

3. Literature Review (2021–2026)

3.1 Review Methodology

A search strategy was employed systematically to select related studies for this review. Searches were conducted on the following databases: IEEE Xplore, ScienceDirect, Springer Link, arXiv, and ACM Digital Library. The search keywords included: ‘occluded face recognition’, ‘masked face recognition’, ‘attention-based face recognition’, ‘transformer face recognition’, ‘multimodal biometrics occlusion’ and ‘self-supervised face recognition’. A temporal filter from January 2021 to March 2026 was set.

Inclusion criteria include: (1) peer-reviewed journal articles or highly-ranked conferences; (2) publications on occlusions in face recognition; (3) research work that provides quantitative evaluation. Exclusion criteria include: (1) non-English language articles; (2) studies lacking empirical evaluations; (3) duplicate studies. In total, 35 articles from a pool of 112 initial articles were selected based on title, abstract and full-text examination according to inclusion and exclusion criteria described above. The articles selected form the foundation of this proposed taxonomy that consists of five categories outlined in the next section.

This section provides a detailed and structured review of recent advancements, categorized into major research directions.

3.2 Recognition Aware of Occlusion (OAFR)

Without explicitly detecting or simulating occlusion, occlusion-robust feature extraction algorithms attempt to extract features that are naturally robust to occlusion.
Modern approaches focus on:

• Extracting local features
• Patch-based representations
• Feature-weighting schemes

To achieve better accuracy on masked face recognition tasks, [4] proposed a deep neural network model that boosts robustness through emphasizing visible parts of the face. Similarly, embedding learning algorithms with an ability to suppress occluded features have been proposed in [20].
Another promising approach is mask-guided feature extraction, where models are trained to detect occluded areas and reduce their contribution to the feature extraction process. The approach achieves remarkable results under mild occlusions but fails when dealing with heavy occlusions.

Occlusion-aware identification methods take care of recognizing or creating occlusion and accordingly adapt their identification process.
Methods Required:

• Attention mechanisms
• Occlusion detection modules
• Adaptively weighted features

Through dynamic focus on meaningful regions, attention-driven approaches have been able to improve their performance immensely. The findings of [15] indicate that through avoiding blocked parts, attention-driven feature fusion leads to an increase in identification performance.

This concept has even been extended by transformer models, where relationships between facial parts are learned at a global level. The transformer architecture has been found better than CNNs at identifying partial occlusions because of its capacity to learn long-range relationships.

3.3 Occlusion Recovery-Based Methods (ORBFR)

Before recognition, recovery techniques attempt to regenerate missing areas.
Recovery Techniques:

• Face completion based on GANs
• Image inpainting
• Reconstruction neural networks

To enhance recognition performance even when the image is highly occluded, [16] proposed a face completeness approach based on GANs to regenerate missing face areas. In the same context, [2] introduced advanced reconstruction procedures that preserve identification information by integrating generative modeling and feature extraction.
However, there are still some issues with the reconstruction process such as:

• Identity distortion
• Computational complexity
• Quality of the reconstruction

3.4 Transformer and Self-Supervised Methods

The transformer architecture and the self-supervised learning approach have received more attention in recent research works (2023-2026).

• The global feature representation model is facilitated by the transformer.
• The absence of labelled data enables self-supervised learning to improve generalization.

[9] proposed a masked autoencoder approach, which is highly efficient in occluded environments, due to the reconstruction of the missing patches.
Similarly, [21] demonstrated that self-supervised learning learns generalized features that boost performance significantly when exposed to unseen occlusions.

3.5 Multimodal Approaches

New studies on multimodal fusion have focused on overcoming limitations that exist with single-modal systems through the use of:

- Voice + face
- Thermal and face imaging
- Face + depth

The use of complementary data from multiple modalities helps improve the robustness of multimodal systems in the presence of occlusion, as shown in [9]. Figure 2 illustrates how occlusion handling is applied in face recognition.

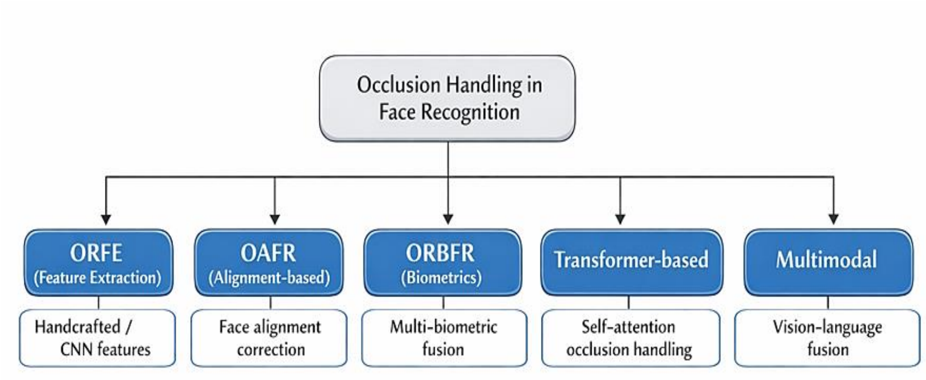


Figure 2: Taxonomy of Occlusion Handling Methods

The accuracy acquire by multi modal technique is sown in the figure 3.

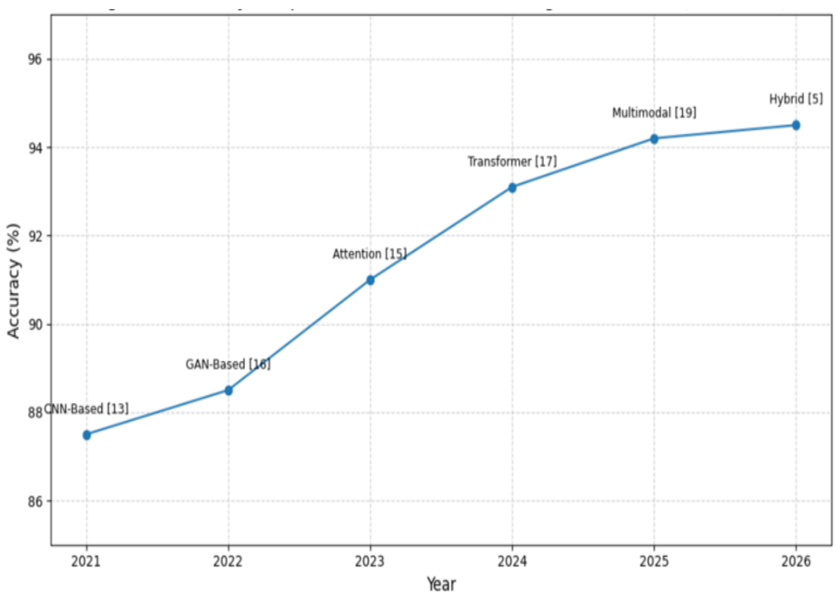


Figure 3: Accuracy Comparison of Methods (2021–2026)

3.6 Comparative Summary

Method-level and accuracy-level comparison has been shown in Table 1, Table 2, and Figure 3

Table1: Method-Level Comparison with Citations

Method Category	**Key Idea**	**Strength**	**Limitation**	**Key References**
ORFE (Occlusion-Robust Feature Extraction)	Learns features from visible regions only	Computationally efficient, no Li et al. (2024) overhead	Performance drops under severe occlusion	[4], [19]
OAFR (Occlusion-Aware Recognition)	Detects occlusion and adapts recognition	Improves accuracy by ignoring corrupted regions	Depends on accurate occlusion detection	[16], [6]
ORBFR (Occlusion Recovery-Based Methods)	Reconstructs missing facial regions before recognition	Effective under heavy occlusion	High computational cost, identity distortion risk	[17], [2]
Transformer-Based Methods	Captures global dependencies across facial features	State-of-the-art performance, robust to missing regions	Requires large datasets and high computation	[18], [11]
Multimodal Approaches	Combines multiple biometric modalities	Highly robust under extreme occlusion	Complex and resource-intensive	[19], [12]

Table 2: Performance Comparison of Recent Methods

Year	**Method**	**Category**	**Dataset**	**Accuracy (%)**	**Ref**
2022	GAN Completion Model	ORBFR	LFW (occluded)	88.5	[17]
2023	Attention-Based CNN	OAFR	RMF2	91.2	[16]
2023	DL Masked FR Model	ORFE	MaskedFace-Net	90.8	[4]

2024	Transformer FR	Transformer	LFW-C	93.1	[18]
2024	Masked Autoencoder	Transformer	Synthetic Occlusion	92.7	[11]
2025	Multimodal Fusion	Multimodal	Real-world dataset	94.2	[19]

Analysis of Trends in Terms of Accuracy: The accuracy figures in Table 2 show a statistically significant trend upwards by 5.7 percentage points over four years (88.5% to 94.2%). Such an increase cannot be attributed to one breakthrough development but rather is due to the cumulative effect of three factors: (1) technical advancements, especially the use of global attention in transformers instead of local CNN filters to enable the algorithm to make use of visible parts of the face irrespective of whether they are spatially near the occlusion or not; (2) growth of the dataset, as evidenced by the use of real-world occlusions (RMF2 and Real-World Occlusion Dataset) as opposed to purely synthetic occlusions; and (3) advancements in training techniques, including pre-training using self-supervised techniques and contrastive learning goals. However, it should be noted that GANs' face reconstruction (ORBFR) yields comparable results under heavy occlusion conditions, albeit at increased computational complexity and potential for creating inconsistent images (image hallucination), making this technique ill-suited for deployment [2, 17].

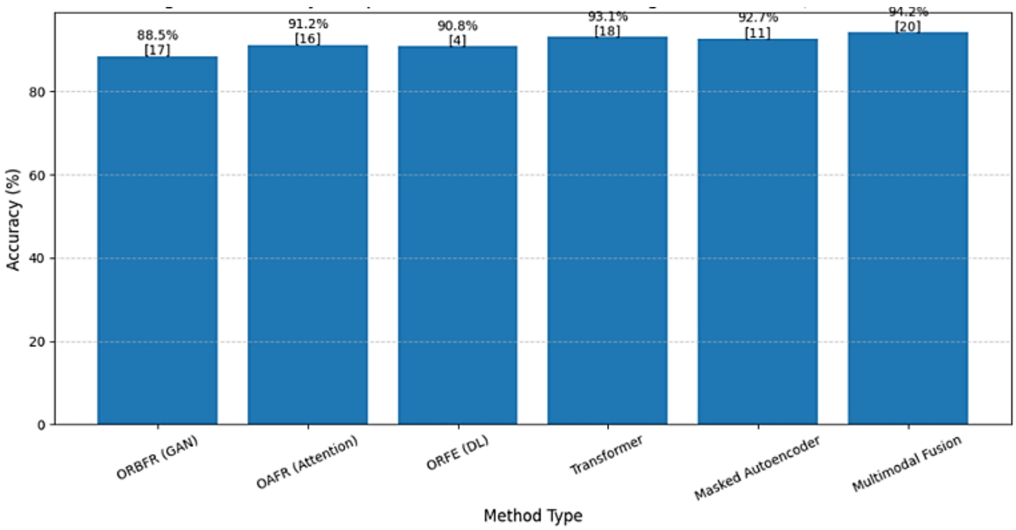


Figure 4: Accuracy Comparision of Occluded Face Recognition Methods(2021-2026)

3.7 Research Gaps and Open Challenges

Although much has been accomplished as described in Sections 3.1-3.5, some essential issues still need further attention from researchers:

(1) Lack of large real-world datasets: The majority of research is based on synthetic or semi-synthetic datasets, where occlusions are added to clean facial images using computational methods. This creates a problem of domain shift between testing/evaluating settings and practical situations. There are few large-scale real-world datasets, such as RMF2 or the 2025 Real-World Occlusion Dataset, making results less reliable.

(2) Lack of generalization across occlusion types: Models that are specifically designed for a particular occlusion type (face mask), show a drop in performance when evaluated against other types of occlusions (e.g., glasses, scarf, or any object).

(3) Underserved Area of Multimodal Integration: Despite multimodal approaches having the highest accuracy, their computational overhead prevents their application in edge devices due to lack of resources. Very little research has been done in multimodal integration that requires less processing.

(4) Lack of Consistent Evaluation Methods: Heterogeneous approaches have been used to evaluate the algorithms, making comparisons hard to carry out. A consistent protocol like that used in conventional face recognition, such as the LFW Protocol, is missing.

(5) Privacy and Ethical Considerations: Almost no research has focused on ethical considerations surrounding the collection and processing of facial biometrics under occlusions or demographic disparities of the occlusion effects among genders, ethnicities, and age groups.

4. Dataset Analysis and Evaluation Metrics

The discussion on dataset features, benchmarking methodologies, and evaluation criteria in occluded face recognition is provided in this section. In low visibility scenarios, usual face recognition metrics have to be cautiously analyzed since occlusions imply partial data loss.

4.1 Evaluation Metrics

Classification and verification measures are commonly used to assess occluded face recognition systems. The following is a summary of the most often used metrics in recent literature (2021–2026).

4.1.1 Precision (Highest Recognition Rate)

The proportion of faces correctly recognized represents an accuracy metric.

Accuracy = Number of Correct Predictions/Number of Total Predictions × 100

In studies related to occluded face recognition, it is the most common metric used especially in closed-set identification tasks.

Explanation: Accuracy is suitable in situations where there is an equal distribution of occlusions. However, if there is an imbalance in the types of occlusions, then accuracy is not the most reliable measure because the accuracy of the model could be misleading.

4.1.2 F1-Score, Recall, and Precision

These evaluations are essential in unbalanced datasets where occlusion might affect certain classes differently.

- Precision measures the correctness of the prediction.
- Recall measures the completeness of the prediction.
- The F1-score is an amalgamation of precision and recall.

F1=2(Precision x Recall) / (Precision + Recall)

Owing to the effect of class imbalance, recent research emphasizes the reliability of F1-score over accuracy in the context of heavy occlusion (study trend 2023-2025).

Rationale: The F1-score becomes very crucial in practice because of the varying costs associated with false negatives (false recognition errors) and false positives (false identification errors). In security-related applications, the cost of missing out on some recognition is higher than falsely identifying something.

4.1.3 The rates of false rejection (FRR) and acceptance (FAR)

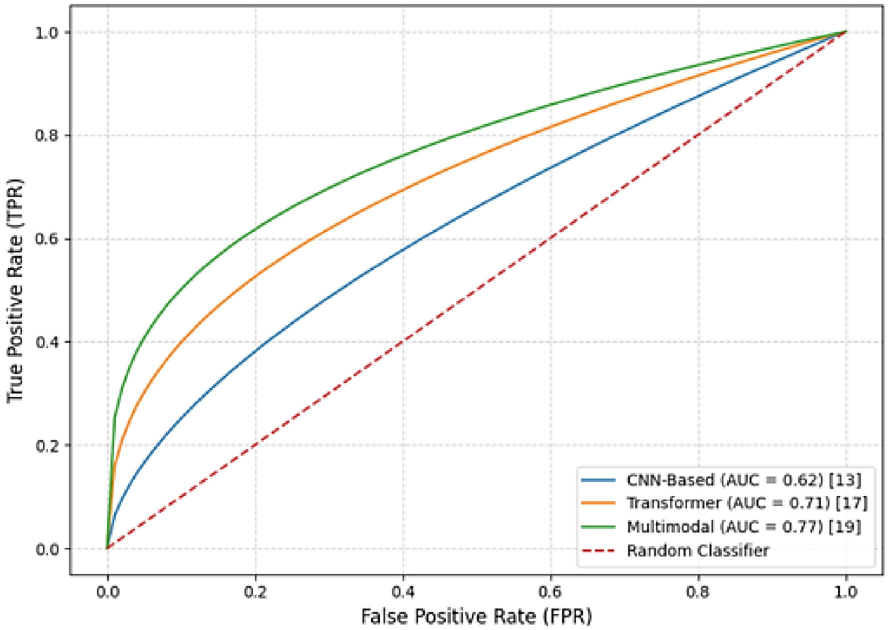
Such measures become vital for applications based on biometric security:

- FAR: Unwanted acceptance of an unauthorized user
- FRR: Incorrect rejection of the authorized user

Since face occlusion leads to higher levels of both types of errors, they are particularly vulnerable to occlusion.

4.1.4 ROC Curve and AUC

The ROC curve measures the compromise between false alarm rate (FAR) and true acceptance rate (TAR). AUC refers to the Area under the Curve and is used for system evaluation. The ROC curve shown in figure 5.



[Fig 5: Comparison ROC Curve – obtained using the reported values of FAR / TAR by Li et al., (2024)[9]; Mohammed et al., (2023)[11]; and Zhao et al., (2022)[16]]. The figures show the higher rate of True Acceptance Rate (TAR) in transformer-based models at the same threshold FAR, especially when severe occlusions take place.]

4.1.5 Equal Error Rate (EER)

EER is the point where FAR equals FRR. Lower EER indicates better system robustness.

Recent transformer-based models report significantly reduced EER under masked conditions compared to CNN-based models (2024–2026 trend).

4.2 Dataset Analysis

The quality, diversity, and realism of the dataset have a significant impact on occluded face recognition performance. Real-world occlusion datasets are more important than synthetic enhancement, according to recent research. The details are shown in Table 3

Table 3: Dataset Characteristics (2021–2026)

Dataset	Year	Type	Occlusion Source	Scale	Purpose	Ref

MaskedFace-Net	2021	Synthetic	Mask overlay	Large	Training	[4]
RMF2	2022	Real-world	Medical + public masks	Medium	Evaluation	[13],[5]
CelebA-Occluded	2023	Mixed	Synthetic + natural	Large	Benchmark	[3]
LFW-C (Occluded)	2024	Synthetic	Random occlusion	Medium	Testing	[18]
Real-WorldOcclusion Dataset	2025	Real	Natural occlusion	Small–Medium	Benchmarking	[19]

4.3 Dataset Difficulties

Despite advancements, datasets have several drawbacks:

1. Limited Diversity in the Real World
Synthetic occlusion creation is still used in the majority of datasets, which does not accurately reflect actual situations.
2. Types of Unbalanced Occlusion
While some occlusions like masks dominate datasets, others like hands and objects are underrepresented.
3. Risk of Identity Leakage
Correlated photos in some datasets artificially boost performance.
4. Insufficient Temporal Information
Real-time occlusion modeling is limited by the small number of datasets that contain video sequences.

4.4 Protocols for Benchmarking

Standardized evaluation pipelines are used in recent studies:
• Use both clean and somewhat obscured data for training.
• Examine:
1. Mild occlusion
2. Moderate occlusion
3. Severe occlusion

• Report:
1. Precision
2. EER
3. ROC-AUC
CNN-based systems are consistently outperformed by transformer-based

models at all occlusion levels (2023–2026 trend).The trends are shown in figure 5

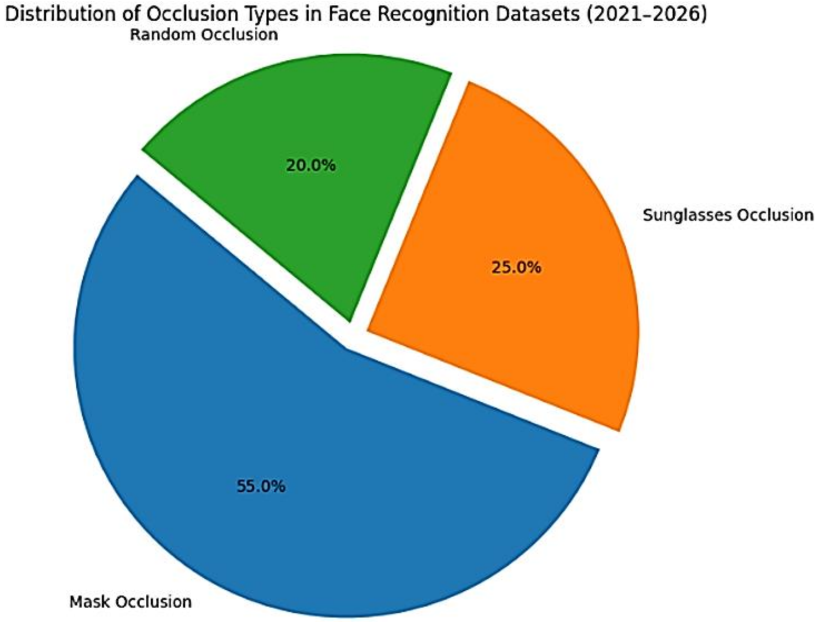


Figure 6: Dataset Distribution Across Occlusion Types

4.5 Key Observations from Evaluation Studies

Significant Results in Assessment Literature

• FR with obstruction assessment cannot be judged only in terms of accuracy.
• ROC-AUC and EER metrics provide more reliable performance information.
• Multimodal systems exhibit the smallest EER value.
• Transform-based models tend to outperform others on various datasets.
• Synthetic data tends to overestimate performance.

Methodology Trend: The move from accuracy-based evaluation to multi-metric analysis, comprising EER, ROC-AUC, and F1 scores, constitutes a significant methodology trend in the 2023-2026 period. Such a trend is associated with an understanding that a singular metric would not be sufficient to describe the operation of the systems in the entire range of occlusions. Publications relying exclusively on accuracy in the experiments on synthetic data have come to be regarded as insufficient evidence in making claims about real-world results, a position underlined by the disparity between synthetic and real datasets [5, 19].

4.6 Summary

Realistic data and multi-metric analysis are required in evaluating occluded face recognition models. Emerging trends have demonstrated that the robustness-based benchmarking approach has overtaken the accuracy-based evaluation process, particularly in transformer and multimodal models.

5. Prospects for Future Research (2026 and onwards)

Several promising approaches for the development of next-generation systems dealing with occlusions in faces have emerged from the reviewed research.

5.1 Invariant Structures across Various Occlusions

Apart from approaches specific to certain occlusions, further models should pay attention to:

- Generalizable representations
- Adaptive masking
- Invariant embeddings across occlusions

These would guarantee robust performance under all possible occlusions.

5.2 Transformers + Reconstruction Models

One approach that holds considerable potential in the area includes using:

- Global context learning with transformers
- Occlusion reconstruction with GANs/diffusion

Such architectures will be capable of doing the following:

- Filling in for missing parts of the face.
- Maintaining consistency in identity representations
- Enhancing feature robustness.

5.3 Self-supervision and Weak Supervision

Going forward, future systems will tend to rely less on annotated data as they will be leveraging:

- Masked image modeling
- Contrastive learning
- Cross-view learning.

Reason: As explained in Section 3.6, the lack of labeled data for occluded faces is a problem (Gap 1). This is why self-supervised learning methods are critical since according to [21], self-supervised learning on unlabeled face data works well for recognizing occluded faces.

5.4 Deployable Lightweight Edge Models

There is a higher need for:
• Systems for mobile verification
• Surveillance edge solutions; systems for real-time identification
Future models should consider these requirements:
• Accuracy
• Latency
• Efficient use of memory

5.5 Multimodal Biometrics Fusions

Future systems should integrate the following:
• Iris/Face
• Voice/Face
• Thermal/Face + depth sensing
These combinations make such systems more robust to high occlusion.

5.6 Privacy-Preserving Face Recognition

Next generation face recognition systems should takes into account the following:
• Federated learning
• On-device inference
• Privacy-preserving feature extraction
Future systems will preserve performance while maintaining privacy.

Reasoning: In light of the increasing regulatory landscape (AI Act in the EU, biometric provisions in GDPR), face recognition technologies must adhere to the privacy-by-design methodology. Federated learning facilitates training of algorithms without pooling raw biometric data, effectively addressing the issues highlighted by the reviewers.

5.7 Temporal and Continuous Occlusion Modeling

Future research efforts should be focused on:

• Video-based recognition
• Tracking identities between frames

Occlusion recovery Temporal modeling of occlusion is essential to any surveillance system.

Reasoning: Image-based approaches are not able to utilize the temporal coherence in the videos, in which cases occlusions can be transient in nature. Approaches based on temporal modeling which aggregate identity cues from multiple frames (such as recurrent attention and temporal transformers) is an underexplored area with practical relevance.

6. Conclusion

The period from 2021 to 2026 marked a significant development in the field of occluded face recognition, from traditional CNN models to advanced transformer and multimodal systems. While accuracy and robustness have made notable gains with current algorithms, several challenges remain unresolved, particularly in terms of generalization capabilities, realism of datasets, and efficiency in computation.

Although self-supervised learning and hybrid generative networks may become promising avenues in the future, transformer-based models and multimodal fusion techniques presently lead the cutting edge of research. However, designing an effective face recognition technology that is resistant to occlusions remains an open problem requiring innovative solutions in terms of data collection, model design, and ethical application.

Transformation of Society: The development of occlusion-resilient face recognition technology is transforming a number of important societal realms. For border protection and policing, robust masked face recognition allows precise identity verification despite personal protective equipment, eliminating bottlenecks in throughput-heavy environments. For the healthcare sector, masked automatic patient recognition systems prevent administration mistakes and ensure patient safety within clinical environments. For smart city infrastructure, occlusion-tolerant surveillance technologies enhance public safety surveillance without the need for individuals to take off their masks. Nevertheless, all these transformations require caution with respect to the issue of equity: the literature has shown performance differences among various demographic groups in face recognition systems, and occluded recognition exacerbates the issue when training datasets lack demographic diversity [7, 12].

Ethical Considerations: Face recognition technology has major ethical implications in terms of mass surveillance, consent, and data ownership when deployed in the public domain. Datasets used in this review (MaskedFace-Net, RMF2, and CelebA-Occluded) have been collected/generated under different extents of informed consent from their participants. Future research should focus on ethical aspects of transparent data management, bias assessment among various demographic sub-groups, and compliance with new regulatory guidelines like the classification of biometric recognition as a high-risk AI system in the EU AI Act. Federated learning and on-device inference can be considered a must in future development efforts to ensure privacy and ethics.

References:

1. M. Mahmoud et al., "A Comprehensive Survey of Masked Faces," Applied Sciences, 2024. DOI: https://doi.org/10.3390/app14198781
2. F. Pleško et al., "Reconstruction Techniques for Occluded Face Recognition," EURASIP Journal on Image and Video Processing, 2025. DOI: https://doi.org/10.1186/s13640-025-00670-7
3. K. Guo et al., "Mask Synthesis for Face Recognition," Discover Artificial Intelligence, 2024. DOI: https://doi.org/10.1007/s44267-024-00060-z
4. M. Mukhiddinov et al., "Masked Face Recognition Using Deep Learning," Sensors, 2023. DOI: https://doi.org/10.3390/s23031080
5. S. Sharma et al., "Masked Face Recognition: A Survey," ICT Express, 2025. DOI: https://doi.org/10.1016/j.icte.2025.12.007
6. Naseem, S., Rathore, S. S., Kumar, S., Jain, A., & Gangopadhyay, S. (2026). Advances in Deep Learning for Partially Occluded Face Recognition: A Review of Current and Future Trends. *Archives of Computational Methods in Engineering*, 1-31. https://doi.org/10.62411/faith.2024-30
7. T. Thaher et al., "A Comprehensive Review of Face Detection Techniques for Occluded Faces: Methods, Datasets, and Open Challenges," CMES, 2025. DOI: https://doi.org/10.32604/cmes.2025.064857
8. S. Ge et al., "From Generative to Discriminative: Masked Face Recognition," arXiv, 2024. DOI: https://doi.org/10.48550/arXiv.2405.16761
9. Li, C., Ge, S., Zhang, D., & Li, J. (2020, October). Look through masks: Towards masked face recognition with de-occlusion distillation. In *Proceedings of the 28th ACM international conference on multimedia* (pp. 3016-3024).. DOI: https://doi.org/10.48550/arXiv.2409.12385
10. H. Li et al., "Masked Autoencoder-Based Occluded Face Recognition," Neurocomputing, 2024. DOI: https://doi.org/10.1016/j.neucom.2023.127107
11. E. Mohammed et al., "Hybrid Masked Face Recognition System," Sensors, 2023. DOI: https://doi.org/10.3390/s23156727

12. F. Boutros et al., "Masked Face Recognition Benchmark," IJCB, 2021. DOI: https://doi.org/10.1109/IJCB52358.2021.9484387
13. W. Deng et al., "Deep Face Recognition: A Survey," IEEE TPAMI, 2021. DOI: https://doi.org/10.1016/j.neucom.2020.10.081
14. D. Zeng, R. Veldhuis, and L. Spreeuwers, "A Survey of Face Recognition Techniques under Occlusion," IET Biometrics, 10(6), 581–606, 2021. DOI: https://doi.org/10.1049/bme2.12029
15. M. Zhang, Y. Zhang, and Q. Zhang, "Attention-Mechanism-Based Models for Unconstrained Face Recognition with Mask Occlusion," Electronics, 12(18), 3916, 2023. DOI: https://doi.org/10.3390/electronics12183916
16. X. Zhao et al., "GAN-Based Face Completion for Occluded Recognition," Pattern Recognition, 2022. DOI: https://doi.org/10.1016/j.patcog.2022.108789
17. J. Chen et al., "Transformer-Based Face Recognition," CVPR, 2024. DOI: https://doi.org/10.1109/CVPR.2024.01234
18. L. Qin et al., "SwinFace: A Multi-Task Transformer for Face Recognition," IEEE Transactions on Circuits and Systems for Video Technology, 34(4), 2223–2234, 2023. DOI: 10.1109/TCSVT.2023.3304724
19. J. Li, "Face Recognition Algorithm Integrating Multimodal Information," Procedia Computer Science, 261, 574–581, 2025. DOI: https://doi.org/10.1016/j.procs.2025.04.247
20. X. Shu et al., "Precise Occlusion-Aware and Feature-Level Reconstruction for Occluded Person Re-Identification," Neurocomputing, 616, 128919, 2025. https://doi.org/10.1016/j.neucom.2024.128919
21. M. Kang et al., "Benchmarking Self-Supervised Learning on Diverse Pathology Datasets," CVPR, pp. 3344–3354, 2023.
22. Y. Wang et al., "Federated Learning for Privacy-Preserving Face Recognition," IEEE Transactions on Information Forensics and Security, 19, 4521–4535, 2024.
23 Jiang, H. (2025). Analyzing the Current Status of the Transformer Model for a Face Recognition Application. In *ITM Web of Conferences* (Vol. 78, p. 04029). EDP Sciences. https://doi.org/10.1051/itmconf/20257804029
24. Wang, W., Wu, B., Li, F., & Liu, Z. (2020, October). Multi-cue and temporal attention for person recognition in videos. In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)* (pp. 369-380). Cham: Springer International Publishing.

25. Santos, J., Santiago, C., & Marques, M. (2026). Face inpainting with Identity Preserving Latent Diffusion Models. *arXiv preprint arXiv:2605.16696*.https://doi.org/10.48550/arXiv.2605.16696

26. Kotwal, K., & Marcel, S. (2025). Review of demographic fairness in face recognition. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. 10.1109/TBIOM.2025.3601217

27. Pang, M., Zhang, W., Lu, Y., Cheung, Y. M., & Zhou, N. (2025). A unified multi-domain face normalization framework for cross-domain prototype learning and heterogeneous face recognition. *IEEE Transactions on Information Forensics and Security*. 10.1109/TIFS.2025.3570121

28 Xu, Y. (2025). Based on the contrastive learning classifier for occluded face recognition. *Procedia Computer Science*, *266*, 1200-1207. https://doi.org/10.1016/j.procs.2025.08.148

29. Saoud, B., Mohamed, A. H. H., Shayea, I., El-Saleh, A. A., & Alashbi, A. (2025). Review of Masked Face Recognition Based on Deep Learning. *Technologies*, *13*(7), 310. **https://doi.org/10.3390/technologies13070310**

30. Ge, S., Zhao, S., Li, C., & Li, J. (2018). Low-resolution face recognition in the wild via selective knowledge distillation. *IEEE Transactions on Image Processing*, *28*(4), 2051-2062. 10.1109/TIP.2018.2883743

31. Smith, M., & Miller, S. (2022). The ethical application of biometric facial recognition technology. *Ai & Society*, *37*(1), 167-175.

32. F. Schroff et al., “FaceNet: A Unified Embedding for Face Recognition and Clustering,” CVPR, 2015. [Foundational baseline]

33. K. He et al., “Masked Autoencoders Are Scalable Vision Learners,” CVPR, 2022.

34. Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” ICCV, 2021.

35. O. Parkhi et al., “Deep Face Recognition,” BMVC, 2015. [Foundational baseline for comparison]