



Artificial Intelligence for Face Recognition in Security Systems: A Review of Algorithms and Challenges

El-Sayed M. El-kenawy^{1,*} Anis Ben Ghorbal²

¹ School of ICT, Faculty of Engineering, Design and Information & Communications Technology (EDICT), Bahrain Polytechnic, PO Box 33349, Isa Town, Bahrain

² Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh 11432, Saudi Arabia

Emails: sayed.elkenawy@polytechnic.bh · assghorbal@imamu.edu.sa

Received: November 20, 2024 Revised: December 18, 2024 Accepted: January 15, 2025 ★ Corresponding author

ABSTRACT

Facial recognition technology (FRT) is acknowledged as one of the successful advancements of biometric applications in security, surveillance, health care, and innovative solutions. The past decade has seen major improvements in deep learning, pre-trained convolutional neural networks (CNNs), and ensemble methods, which have substantially improved FRT accuracy and efficiency. Nonetheless, several issues remain, including facial expression variation, illumination changes, demographic bias, adversarial attacks, and backdoor threats. These limitations require new approaches and tools to enhance the reliability and ethical use of FRT. This review discusses state-of-the-art algorithms, application areas, security vulnerabilities, open problems, and the ethical and social consequences of deploying FRT systems.

Keywords: Facial recognition technology ▪ Deep learning ▪ Neural networks ▪ Biases ▪ Adversarial attacks ▪ Ethics

1. INTRODUCTION

Facial recognition technology (FRT) is widely regarded as a revolutionary element in contemporary biometric applications, supporting progress in security, monitoring, healthcare, smart cities, and identity management. In the last decade, deep learning technologies, pre-trained CNN models, and ensemble approaches have contributed to increased accuracy and efficiency. However, the field continues to face expression bias, illumination bias, demographic shift bias, and sensitivity to adversarial and backdoor attacks. These limitations signify the need for new measures and mechanisms that can improve reliability while also addressing ethical and social consequences.

1.1 State-of-the-Art Algorithms and Techniques

FRT has developed rapidly through improved algorithms and methodology. Transfer learning has become a common strategy in contemporary systems, allowing established models such as DenseNet201, ResNet152V2, and MobileNetV2 to reduce feature-extraction and learning time on large databases. These models derive stable facial features and enhance recognition performance under changing illumination and occlusion. Ensemble models, including weighted-average and hybrid models, further improve performance by combining multiple feature predictors. Online learning also allows systems to introduce new personnel and update templates in dynamic environments. Together with careful hyperparameter tuning and data conditioning, these developments have established new standards for FRT accuracy, speed, and responsiveness in real-life situations [1].

1.2 Bias and Security Vulnerabilities

Modern FRT suffers from important bias and security risks that can undermine trust. Expression bias occurs when recognition performance changes with facial expression. Demographic disparities linked to age, race, gender, or ethnicity intensify fairness concerns. Security risks include adversarial patches and digital manipulations that exploit algorithmic vulnerabilities. More serious backdoor attacks, such as Facial Identity Backdoor Attacks (FIBA), involve poisoning samples during enrollment so that attackers can bypass systems using universal triggers. These weaknesses require diverse datasets, fair algorithm design, and strong defense mechanisms [2].

1.3 Areas of Application and Effectiveness

The use of FRT has expanded across core infrastructure and essential applications by enhancing security, comfort, and governmental procedures. In surveillance, it improves public security by assisting in identifying people in public areas and supporting law-enforcement investigation. In authentication, it enables convenient access to devices, workplaces, and financial transactions. In smart cities, border control, and electronic voting, FRT supports identity management and organizational efficiency. Nevertheless, integration into these domains raises concerns about privacy invasion, ethical violation, and misuse. Its practical advantages must therefore be balanced with individual rights and public trust.

1.4 Ethical and Governance Challenges

Facial recognition technology creates major ethical and governance risks because it captures and stores highly private biometric data. Concerns include excessive surveillance, misuse, discrimination, lack of consent, and unclear accountability. Existing regulations often lag behind technological adoption, making it difficult to define standards for security, consent, and lawful deployment. Principled governance should emphasize transparency, fairness, accountability, and respect for civil liberties. Legal standards and ethical norms are essential if society is to benefit from FRT while limiting its potential harms [13].

1.5 Future Directions and Open Problems

Future progress depends on addressing unresolved challenges and pursuing innovative research directions. More diverse datasets that include demographic variety, lighting conditions, and facial expressions can reduce bias and improve algorithmic fairness. Robust models against adversarial attacks and backdoor vulnerabilities remain a security priority. Explainable FRT algorithms can promote transparency and accountability, while privacy-preserving methods such as federated learning can mitigate risks associated with data exposure. These open problems provide opportunities to refine FRT systems and ensure fair, reliable, and ethical applications.

2. LITERATURE REVIEW

FRT has become an important component of security and biometric applications in surveillance, authentication, and smart city technologies. This literature review focuses on the progression of FRT and on difficulties such as illumination, pose variation, and expression bias. It considers pre-trained CNNs, novel recognition approaches, transfer learning, en-

semble learning, and resistance to adversarial and backdoor assaults. Ethical and privacy issues are also examined as a significant social dimension of FRT.

Artificial intelligence has become a pivotal driver of innovation in video surveillance. Machine learning, deep learning, neural networks, natural language processing, and expert systems have transformed traditional surveillance practices by enabling intelligent analytics and improved security-system quality [3]. A related smart-lock system based on YOLO and a Jetson Nano board collects user images, transfers data to a server for training through auto-labeling, adjusts brightness, and supports Android-based authentication. Under favorable conditions, the system achieved 99% accuracy and a recognition speed of 0.6 seconds [4].

Broad surveys describe face recognition systems, their components, operational principles, algorithms, strengths, and limitations [5]. Deep-learning surveys in biometrics cover face, fingerprint, iris, palmprint, ear, voice, signature, and gait recognition, highlighting datasets, benchmark performance, and open challenges [6]. Recent comprehensive reviews classify face-recognition techniques into appearance-based and hybrid methods, emphasizing the influence of illumination, pose variation, expressions, occlusion, and aging [7]. Classical and modern approaches, including sparse models, fuzzy set theory, and deep learning, are also analyzed across datasets and application domains [8].

Deep-learning approaches have been explored through pre-trained CNN models including DenseNet201, ResNet152V2, MobileNetV2, SeResNeXt, and Xception. Weighted-average ensemble models and grid-search optimization can improve feature extraction and classification, with studies reporting 99.48% accuracy on the LFW dataset [9]. Robustness studies show that face-recognition algorithms remain vulnerable to physical presentation attacks, disguise and makeup, digital adversarial attacks, and GAN-based morphing or tampering. These works also reveal that age and gender differences affect algorithm performance [2].

Backdoor threats have expanded the security landscape. FIBA attacks operate at the enrollment stage, using a single poisoned example to create a universal trigger that allows attackers to bypass FRS at scale [10]. Other systems combine face detection with voice and biometrics for automated attendance and access control through cascade classification and template matching [11]. Smart-city fog and cloud systems use transfer learning and deep CNNs to handle occlusion, expression, illumination, and pose variation; when evaluated on SDUMLA-HMT, 113, and CASIA datasets, such systems reported 99.06% accuracy, 99.12% precision, 99.07% recall, and 99.10% specificity [12].

Regulatory studies examine the tension between public security, privacy, civil liberties, and ethics. They argue that legal frameworks must evolve to protect human rights and guide responsible FRT implementation [13]. Expression-bias research shows that facial expressions remain a security threat in recognition scenarios. Experiments involving face detectors, recognition models, and databases reveal significant expression-related biases that affect performance and call for more stable face-recognition systems [14].

Table 1 summarizes the reviewed studies, their focus areas,

Table 1. Summary of Literature Review

Ref.	Focus Area	Key Contributions	Evaluation Metrics
[3]	AI in video surveillance systems	Examines advancements in AI-driven surveillance systems and the integration of deep learning, neural networks, and analytical software.	Emphasizes trends, challenges, and impact on security-system quality.
[4]	Smart lock with face recognition	Develops a YOLO-based smart lock on Jetson Nano with auto-labeling, brightness adjustment, and Android authentication.	Achieved 99% accuracy and 0.6-second recognition under favorable conditions.
[5]	Applications of FRT	Reviews components, principles, algorithms, applications, occlusion, and expression bias.	Discusses broader applications and system efficiency.
[6]	Deep learning in biometrics	Surveys over 150 works on face, iris, fingerprint, voice, gait, and other biometric modalities.	Uses benchmark datasets and identifies open challenges.
[7]	Advances in FR techniques	Reviews appearance-based and hybrid methods and challenges such as pose, illumination, occlusion, and expression.	Classifies image- and video-based FR methods.
[8]	Modern FR approaches	Reviews classical and modern FR techniques including deep learning, sparse models, and fuzzy set theory.	Analyzes age, gender, and background complexity in FR systems.
[9]	Deep learning and ensembles	Introduces weighted-average ensemble models using DenseNet201, ResNet152V2, MobileNetV2, and others.	Achieved 99.48% accuracy on LFW.
[2]	Robustness of FR systems	Examines physical, digital, adversarial, disguise, and morphing attacks plus demographic bias.	Discusses real-world robustness without one fixed metric.
[10]	Backdoor vulnerabilities in FRS	Proposes FIBA, an enrollment-stage attack using poisoned examples and universal triggers.	Evaluates scalability, stealth, and trigger capability.
[11]	Face detection with voice and biometrics	Develops automated attendance and access control using cascade classification and template matching.	Highlights attendance and security effectiveness.
[12]	Fog and cloud FR systems	Uses transfer learning, DCNNs, DT, KNN, and SVM for adaptive recognition under occlusion and pose variation.	Reports 99.06% accuracy, 99.12% precision, 99.07% recall, and 99.10% specificity.
[13]	Ethical and governance challenges	Explores privacy, ethics, legal frameworks, and governance structures for FRT.	Focuses on legal and ethical recommendations.
[14]	Facial expression bias in FR	Analyzes expression bias in popular datasets and its effect on face-recognition accuracy.	Uses two detectors, three recognition models, and three databases.
[10]	Practical threats to FRS	Investigates adversarial and backdoor threats and introduces enrollment-stage FIBA.	Analyzes universal triggers and poisoned examples.

contributions, and evaluation criteria.

The reviewed works show that facial recognition systems continue to change through machine learning and deep learning. Although accuracy and robustness have improved, problems remain, especially bias, security threats, privacy, and governance. These gaps motivate innovations in algorithms, dataset selection, defense strategies, and regulatory frameworks. Addressing them can promote safer, fairer, and more sustainable FRT systems.

3. DISCUSSION

This section discusses the main arguments emerging from the literature review, focusing on FRT progress, hurdles, and broader implications. It formulates research priorities and ethical principles for the development of secure, efficient, and fair FRT systems.

3.1 FRT Technological Developments

Due to deep learning, pre-trained CNNs, and ensemble learning, FRT has made significant progress. These technologies have increased recognition accuracy and speed, although occlusions and variations in pose and illumination remain difficult. Transfer learning and hyperparameter tuning enable systems to learn from different datasets with less computing power. However, real-time adaptation and scalability are still not fully solved, especially in demanding applications such as video surveillance and authentication [7, 15].

3.2 Addressing Bias in FRT

Expression, demographic, and illumination biases remain central problems. Studies indicate performance differences related to age, gender, and ethnicity, which raises ethical and fairness concerns. These biases often arise from imbalanced datasets and algorithmic constraints. To address them, researchers must collect more diverse datasets, design fairer algorithms, and integrate interpretability methods into

machine-learning models [16, 17].

3.3 Security Threats and Protection Techniques

FRT systems face adversarial attacks, backdoor vulnerabilities, and enrollment-stage attacks such as FIBA. These attacks target datasets, feature databases, or algorithms and may compromise system integrity. Defense strategies include adversarial training, dataset sanitization, secure database management, robust feature verification, and continuous monitoring of enrollment pipelines [18, 1].

3.4 Ethical Considerations and Governance

The rapid growth of FRT has raised concerns about privacy, surveillance, citizenship freedoms, discrimination, and consent. Governance models based on openness, responsibility, and human-rights protection are needed. Public, private, and civil-sector actors should collaborate so that innovation is not driven solely by technical interests. Regulations should also remain adaptable as technologies and threats evolve [19, 20].

3.5 Future Research Directions

Future work should develop challenging datasets, robust algorithms for adversarial conditions, and real-time processing methods. Privacy-preserving approaches such as federated learning can reduce data-security concerns, while explainable AI (XAI) can improve trust and accountability in decision-making. These directions can help FRT systems achieve higher reliability and ethical alignment [21, 22].

4. CONCLUSION

Facial recognition technology has transformed biometric systems and enabled new forms of security, surveillance, authentication, and identity management across many sectors. The literature shows strong progress in machine learning and deep learning, particularly CNNs and ensemble models, which improve accuracy and efficiency. However, dataset bias, security

vulnerabilities, and ethical concerns remain major barriers to responsible adoption.

The reviewed studies confirm persistent expression bias, demographic disparities, and environmental variability. These biases affect FRT effectiveness and create ethical questions about equity and inclusivity. Security threats such as adversarial attacks and FIBA-style backdoor attacks require reliable defense mechanisms, including adversarial training, database enhancement, and fairness-aware algorithms.

Ethics and governance frameworks are central to FRT adoption. Insufficient regulation increases concerns about privacy, surveillance, and misuse of biometric data. Future policies must emphasize transparency, accountability, and protection of individual rights. Collaboration among researchers, policymakers, and industry leaders is required to develop ethical standards and legal frameworks for responsible use.

Future research must address open problems such as diverse datasets, adversarial robustness, real-time processing, federated learning, and explainable AI. Achieving these goals will support a more secure, fair, and ethically aligned future for FRT applications.

REFERENCES

- [1] L. Li, X. Mu, S. Li, and H. Peng, "A Review of Face Recognition Technology," *IEEE Access*, vol. 8, pp. 139110–139120, 2020, doi: 10.1109/ACCESS.2020.3011028.
- [2] R. Singh, A. Agarwal, M. Singh, S. Nagpal, and M. Vatsa, "On the Robustness of Face Recognition Algorithms Against Attacks and Bias," *AAAI Conference on Artificial Intelligence*, pp. 13583–13589, 2020, doi: 10.1609/aaai.v34i09.7085.
- [3] Reference from source paper: AI applications and advancements in video surveillance systems.
- [4] Reference from source paper: YOLO-based smart lock system with face recognition on Jetson Nano.
- [5] Reference from source paper: Applications and principles of face recognition technology.
- [6] T. Ahmad et al., "Biometrics recognition using deep learning: a survey," *Artificial Intelligence Review*, vol. 56, no. 8, pp. 8647–8695, 2023, doi: 10.1007/S10462-022-10237-X.
- [7] H. L. Gururaj, B. C. Soundarya, S. Priya, J. Shreyas, and F. Flammini, "A Comprehensive Review of Face Recognition Techniques, Trends, and Challenges," *IEEE Access*, vol. 12, pp. 107903–107926, 2024, doi: 10.1109/ACCESS.2024.3424933.
- [8] W. Ali, W. Tian, S. U. Din, D. Iradukunda, and A. A. Khan, "Classical and modern face recognition approaches: a complete review," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 4825–4880, 2021, doi: 10.1007/S11042-020-09850-1.
- [9] J. Selvagesan et al., "Enhancing face recognition performance: a comprehensive evaluation of deep learning models and a novel ensemble approach with hyperparameter tuning," *Soft Computing*, vol. 28, no. 20, pp. 12399–12424, 2024, doi: 10.1007/S00500-024-09954-Y.
- [10] J. Chen et al., "Rethinking the Vulnerabilities of Face Recognition Systems: From a Practical Perspective," 2024. [Online]. Available: <https://arxiv.org/abs/2405.12786v3>
- [11] M. F. Siddiqui, W. A. Siddique, M. Ahmedh, and A. K. Jumani, "Face Detection and Recognition System for Enhancing Security Measures Using Artificial Intelligence System," *Indian Journal of Science and Technology*, vol. 13, no. 9, pp. 1057–1064, 2020, doi: 10.17485/IJST/2020/V013I09/149298.
- [12] D. S. Abdelminaamid, A. M. Almansori, M. Taha, and E. Badr, "A deep facial recognition system using computational intelligent algorithms," 2020, doi: 10.1371/journal.pone.0242269.
- [13] 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. 7, 2024, doi: 10.3389/FDATA.2024.1337465.
- [14] A. Peña, A. Morales, I. Serna, J. Fierrez, and A. Lapedriza, "Facial Expressions as a Vulnerability in Face Recognition," *International Conference on Image Processing*, pp. 2988–2992, 2020, doi: 10.1109/ICIP42928.2021.9506444.
- [15] M. Jha, A. Tiwari, M. Himansh, and V. M. Manikandan, "Face Recognition: Recent Advancements and Research Challenges," *ICCCNT*, 2022, doi: 10.1109/ICCCNT54827.2022.9984308.
- [16] M. Hassaballah and S. Aly, "Face recognition: Challenges, achievements and future directions," *IET Computer Vision*, vol. 9, no. 4, pp. 614–626, 2015, doi: 10.1049/IET-CVI.2014.0084.
- [17] D. Verma, K. Dhanda, M. Kumar, and M. Gulhane, "Facial Recognition Unlocking Potential Facing Challenges in Face and Gender Identification," *IC3SE*, pp. 125–130, 2024, doi: 10.1109/IC3SE62002.2024.10592951.
- [18] Mujiyanto, A. Setyanto, E. Utami, and K. Kus-rini, "Facial Expression Recognition with Deep Learning and Attention Mechanisms: A Systematic Review," *ICICOS*, pp. 12–17, 2024, doi: 10.1109/ICICOS62600.2024.10636857.
- [19] M. K. Rusia and D. K. Singh, "A comprehensive survey on techniques to handle face identity threats: challenges and opportunities," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 1669–1748, 2022, doi: 10.1007/S11042-022-13248-6.
- [20] H. Ghazouani, "Challenges and Emerging Trends for Machine Reading of the Mind from Facial Expressions," *SN Computer Science*, vol. 5, no. 1, 2024, doi: 10.1007/S42979-023-02447-Z.

- [21] K. Marwa and O. Kais, “Current Challenges of Facial Recognition using Deep Learning,” *IEEE SSD*, pp. 1980–1986, 2022, doi: 10.1109/SSD54932.2022.9955857.
- [22] Y. Wen, B. Liu, L. Song, J. Cao, and R. Xie, “Facial Recognition Technology and the Privacy Risks,” in *Face De-identification: Safeguarding Identities in the Digital Era*, pp. 15–20, 2024, doi: 10.1007/978-3-031-58222-6_2.