Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

A Comprehensive Review of Thermal-Aware Face Recognition Systems: Progress and Challenges

Samar S. Mahdi[1], Ebtesam N. AlShemmary[2]*

[1]*Department of Computer Science, Faculty of Computer Science and Mathematics, University of Kufa, Kufa, Iraq.* Email: samars.aljanabi@student.uokufa.edu.iq

[2]*IT Research and Development Center, University of Kufa, Kufa, Iraq* Email: *dr.alshemmary@uokufa.edu.iq*

ARTICLE INFO

Article history:
Received: 00 /mm/2025
Rrevised form: 00 /00/2025
Accepted : 29/12/2025
Available online: 30/12/2025

Keywords:

Contactless authentication,

Health monitoring,

Multimodal biometrics,

RGB–LWIR fusion,

Thermal face recognition.

ABSTRACT

Modern biometric systems have relied on face recognition as it is precise, convenient and unobtrusive. However, standard visible-spectrum face recognition systems are very much sensitive to changes in lighting, facial expressions, and surroundings and at the same time have no ability to conduct physiological measurements in tandem. Within the framework of access control and overall safety needs in the post-pandemic context, these restrictions have indicated a gap in research that is urgent: the lack of integrated frameworks that carry out the functions of identity verification and health-related screening in tandem with each other. This article is a critical survey of the state-of-the-art in the field of thermal-sensitive face recognition, which integrates the visible RGB imaging with biometric recognition and long-wave infrared (LWIR) thermal sensing with the purpose of estimating body temperature. We syntactically examine the available architectures, sensing configurations, fusion strategies, datasets, and evaluation protocols found in the literature. The review emphasizes the fact that dual-modal systems have the potential to allow real-time and contactless identity verification and support a large-scale approach in high-traffic settings. Moreover, this paper addresses the main technical and practical issues that are limiting the large-scale implementation, such as sensor calibration, cross-modular data alignment, environmental bias, data privacy, as well as ethical factors. Last but not the least, we describe the new research directions and future outlooks of the unified biometric and health-conscious access control systems through a critical lens that analyzes performance trade-offs, computational costs, and ethical implications, with the aim of informing the researchers and practitioners of the creation of effective, scalable, and privacy-aware solutions.

MSC..

https://doi.org/ 10.29304/jqcsm.2025.17.42580.

Introduction

Among these biometrics modalities, facial recognition has become one of the most embraced modalities due to its ease of use, hygienic nature, and broad-based use that cut across many fields like surveillance, access control, border security, and human-computer interaction [1]. Traditional authentication based on passwords and keycards is becoming vulnerable to

∗Corresponding author: Ebtesam N. AlShemmary

Email addresses: dr.alshemmary@uokufa.edu.iq

Communicated by 'sub etitor'

theft and their misuse; thus, the interest is shifting toward biometric systems that verify identity based on unique biological features [2].
The demand for touchless biometric solutions was further accelerated by the COVID-19 pandemic, with thermal detection gaining prominence for being a non-invasive tool for health screening and fever detection in public space [3].
Despite rapid development, computer vision and deep learning still cannot avoid several issues with the traditional face recognition system using visible-light imaging, such as susceptibility to illumination variation, spoofing attacks, and low accuracy for highly pigmented skin color individuals [4, 5].
Thermal imaging is a promising alternative since the heat patterns it captures from the human face are invariant to lighting and difficult to reproduce, hence making this modality more robust and spoof resistant [6, 7].
This paper critically reviews the evolution of thermal-aware face recognition technologies, from earlier machine learning models that depend on hand-crafted features to modern deep learning architectures that are able to extract discriminative representations directly from raw thermal data [8].
The paper focuses on higher levels of integration of methods in computer vision, including YOLO models and super-resolution algorithms, to overcome such issues of thermo-imaging as poor resolution and noise [9]. The use of specialized optimizers like AdamW also boosted the convergence of the model and enhanced its results on thermal datasets [10, 11, 12]. In addition, the combination of thermal and visible spectrum information in multi-modal biometric systems is highly promising to improve the recognition accuracy, reliability, and security [3]. The higher quality and fineness of the images of the facial representation has made possible the methods such as super-resolution of thermal images with the help of deep learning or multi-modes fusion which boosts the detection in real-world applications [13, 14].With the rapid expansion of the biometric technologies market, which is projected to reach approximately USD 7.76 billion by 2025, with an estimated compound annual growth rate (CAGR) of 15.3% [1].

2. Review Methodology

This review was done on the basis of a systematic approach in order to provide a complete, transparent, and repeatable study of the literature on thermal-based face recognition. It involved four major steps of planning, search, screening, and synthesis.

- Literature Search Strategy: A systematic search was carried out to find out suitable peer-reviewed articles, conference proceedings, seminal books. Keywords and Boolean operators, such as: (thermal imaging OR infrared face recognition OR LWIR face recognition) AND (biometrics OR authentication) AND (multimodal fusion OR RGB-IR) AND (deep learning OR computer vision), were used in the search.

- Databases Consulted: The main scientific databases that were searched were IEEE Xplore, Scopus, Web of Science, and Google Scholar. The choice of these platforms was due to their wide search of engineering, computer science, and interdisciplinary literature.

- A list of inclusion and exclusion criteria were used to make sure the review is focused on the relevant and high quality literature. Peer-reviewed articles in English were restricted to 2018-2024 to be able to capture recent developments. To be included, the studies should have focused on thermal or long-wave infrared (LWIR) face recognition or physiological feature extraction studies, but with a specific interest in those that mention a fusion approach with visible-spectrum (RGB) modalities. The research papers have been filtered out in case they were limited to visible-light recognition only, or other unrelated biometric modalities, used thermal imaging not in the context of biometric but purely in the context of medical diagnosis, available only in abstract form, or otherwise non-peer-reviewed articles, like technical reports or theses.

- Screening and Selection Process: The first search results were de-duplicated, and titles/abstracts were filtered according to the inclusion criteria. The rest of the articles were then critically appraised in their entirety. This repeated procedure made sure that the ultimate compilation of literature that has been critically reviewed in this review is the literature that directly covers the main themes of thermal-face recognition structures, fusion techniques, and datasets as well as the emerging issues.

3. Thermal Face Recognition Fundamentals

Thermal imaging relies on analyzing the infrared (IR) spectrum to capture the heat patterns naturally emitted by the human face, making it a powerful alternative to conventional visible-light imaging. Unlike traditional systems that depend on reflected light, thermal imaging detects emitted radiation, which allows it to remain largely unaffected by changes in ambient illumination, facial expressions, or partial occlusions such as masks or eyeglasses [15].

3.1 Infrared Spectrum and Imaging Modalities

The infrared spectrum extends approximately from 0.7 to 14.0 µm. For facial recognition applications, this range is typically divided into two primary bands, each offering distinct characteristics and advantages for thermal-based analysis:

- Reflected IR Band (0.7–2.4 µm): it includes Near-Infrared and Short-Wave Infrared, which capture reflected solar radiation. These bands are useful in low-light conditions and use reflected radiation. However, they cannot provide thermal properties of the face.
- Thermal IR Band (2.4-14.0 µm): including both MWIR (3-5 µm) and LWIR (8-14 µm), records the radiation emitted from the skin surface to highlight physiological features like blood vessel patterns or heat distribution. This modality is practically invariant to lighting and less affected by environmental factors like smoke or dust. Face detection and localization are simplified under thermal IR due to reduced background clutter, since no biological surfaces emit much less thermal radiation, hence enhancing segmentation accuracy and reducing false positives in complex scenes. Figure 1, shows main divisions in the IR spectrum that describe its relation to thermal imaging applications.

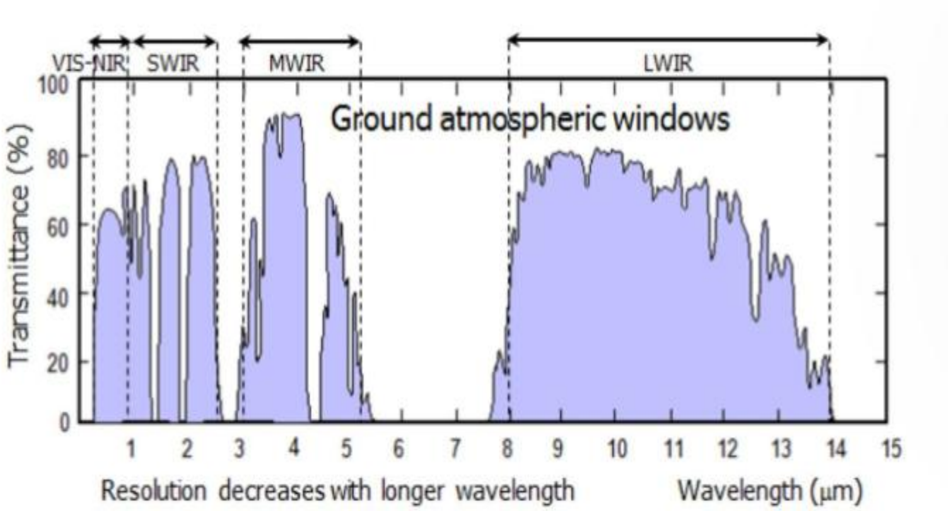

Fig. 1 - Definition of IR Spectral Band [16].

3.2 Physiology-Based Feature Extraction

Thermal face recognition systems rely on a specific individual's thermal signature, modulated through vascular structure, metabolic activity, and skin emissivity. Feature extraction usually consists of finding the main thermal points of interest, like those lying on veins and capillaries, encoding them into discriminative vectors for classification [17].

Traditional methods have, therefore, employed handcrafted features in conjunction with statistical models, like Random Forest classifiers, to distinguish between subjects based on selected thermal landmarks. Indeed, these have proven robust under challenging conditions, including occlusions, noise, and facial accessories such as masks and eyeglasses. Thermal face recognition systems rely on physiological features underlying the vascular pattern, metabolic activity, and skin emissivity. Image processing techniques such as white top-hat segmentation and morphological filtering are applied for feature extraction.

Figure 2 shows the process of vascular feature extraction from thermal images in order to produce TMPs that will be used for classification, much like fingerprint features.

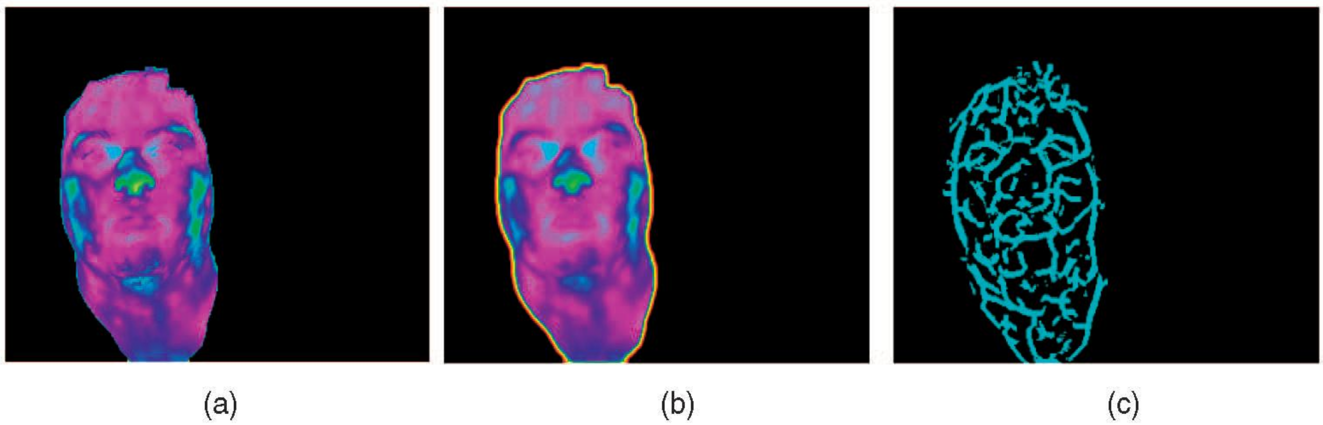

Fig. 2 - Vascular network extraction. (a) Original segmented image, (b) Anisotropically diffused image, (c) Blood vessels extracted using white top hat segmentation [18].

3.3 Synthetic Thermal Data Augmentation

Recent advances have enabled the use of deep learning–based generative frameworks for the synthesis of thermal imagery. Approaches such as StyleCLIP and GANs N' Roses have been employed to transform visible-spectrum images by embedding thermal-like characteristics. The resulting large-scale synthetic thermal datasets provide valuable resources for training more robust face recognition models [18].

4. Learning-Based Thermal Analysis

In the development process, thermal-aware face recognition gets immense support from ML and DL techniques for improving its performance. These techniques have facilitated the ability of systems to extract meaningful features from thermal data, overcome challenges such as low resolution and noise, and thus improve recognition accuracy in unconstrained conditions.

4.1 Traditional Machine Learning Techniques

Early thermal face recognition systems relied on feature engineering in a hand-crafted style, and conventional machine learning algorithms, such as Support Vector Machines, k-Nearest Neighbors, and Random Forests. The statistical descriptors exploited in these models include Local Binary Patterns, Histogram of Oriented Gradients, and thermal minutiae points which are used to classify people according to thermal patterns. Although quite useful in controlled settings, they could not be generalized to different settings and were vulnerable to occlusions and spoofing.

4.2 Deep Learning Architectures

In recent years, there has been a shift to hierarchical deep learning models, in the form of Convolutional Neural Networks, which are capable of learning hierarchical features on raw thermal images, in a self-taught way. Though thermal modalities have been developed with various architectures, such as ResNet, VGGNet and MobileNet, which are pre-trained on visible data, and then fine-tuned on thermal data. The models demonstrate state of art performances on face detection, alignment, and recognition. They are however superior depending on the quality and quantity of data. CNNs are excellent at discriminative, high-complexity feature learning on well-curated thermal imaging, but are extremely vulnerable to system performance failure when input resolution is low, there is a lot of noise, and there is a domain shift, which are typical in real-world thermal imaging. This is in contrast to the traditional ML techniques, which, although less potent, can be more predictable and resistant to particular, known image degradation. Farooq et al. [10] demonstrated that training thermal-aware convolutional neural networks using the AdamW optimizer significantly enhanced convergence speed and generalization performance. Similarly, Persiya and Anbalagan [12] applied the YOLOv5 framework to thermal image datasets, achieving improved object detection and localization accuracy, primarily in scenarios with medium-to-high resolution frontal faces, as summarized in Table 1.

Table 1 - Summary of Key ML/DL Approaches in Thermal Face Recognition.

Authors / Ref.	Year	Technique / Model Used	Contribution / Findings	Limitation
Persiya & Anbalagan [12]	2025	YOLOv5	Enhanced detection and localization in thermal biometric systems	Struggles with small-scale faces and low-resolution thermal inputs
SpringerOpen – LVT Dataset Study [15]	2024	CNNs on Visible vs Thermal	Compared soft biometric estimation across modalities under varied conditions	Limited generalization across ethnic groups and age ranges
Farooq et al. [10]	2023	CNN + AdamW Optimizer	Improved convergence and accuracy in thermal face recognition	Limited dataset diversity; performance drops in extreme occlusion scenarios
Huang et al. [14]	2022	SRGAN / ESRGAN	Applied super-resolution to improve thermal image clarity and recognition	Computationally intensive; may introduce artifacts in high-noise conditions
MDPI – GANs N' Roses [18]	2022	GAN-based Thermal Synthesis	Generated synthetic thermal images for robust training and augmentation	Synthetic images may lack physiological realism; domain gap remains

5. *Multimodal Fusion and Anti-Spoofing Methods*

The use of biometric systems in the security, healthcare, and consumer applications points to the growing demand of the use of strong and spoof resistant face recognition systems. Thermal-aware FR has some inherent advantages in resisting presentation attacks because it is capable of capturing the emitted heat patterns. However, low resolution, limited texture. To overcome these limitations, researchers often combine thermal data with visible spectrum information. Moreover, the quality of thermal inputs is improved with the use of the techniques such as super-resolution. In this section, strategies that combine multimodal data and expert processing to enhance recognition accuracy and system resilience are reviewed.

5.1 Multimodal Fusion Strategies

Multimodal fusion is a combination of complementary data of thermal and visible light imaging to form richer representations. Another activity related to enhancing the quality of thermal data is the application of super-resolution (SR) techniques. The idea behind SR is to synthesize high-resolution images out of the low-resolution images and restore fine details, which are important to recognize. SRGAN and ESRGAN deep learning models have been successfully used on the thermal face images [14]. Moreover, generative architectures like StyleCLIP and GANs N' Roses, have been applied to generate thermal images based on visible images, which has been utilized in data enhancement and domain adaptation [18]. It can be achieved by a number of methods:

- Data-Level Fusion: It is a type of fusion that incorporates raw pixel data of the two modalities. It is a simple problem that is normally burdened with the issue of misalignment and transmission of noise.
- Feature-Level Fusion: This involves the independent extraction of features from each modality, which is then combined into one unified representation. In this approach, the strength of each modality is still retained, and this finds wide applications in deep learning pipelines.
- Decision-Level Fusion: It fuses the outputs of separate classifiers independently trained on each modality. It is computationally very efficient but it loses fine-grained interaction among features.

As reported by Lai et al. [2], thermal–visual feature fusion significantly improves recognition accuracy under poor lighting conditions and partial occlusions. Yu et al. [7] demonstrated enhanced anti-spoofing capability through the fusion of thermal and RGB data, as it enables the detection of abnormal heat distribution patterns that are absent in printed or replayed attacks..

Recently, large-scale deep models make use of attention mechanisms and cross-modal transformers that dynamically weight the contribution of each modality. These architectures let the system give more importance, for example, to thermal data when the light is very low or to visible data when the texture is important, reaching adaptive robustness.

5.2 Anti-Spoofing Techniques

Presentation attacks involving printed photos, video replays, and 3D masks are considered serious threats to biometric systems. Many of these attacks are inherently resisted by thermal imaging, where the artificial medium

often lacks the physiological heat patterns from a human face. We do have advanced spoofing technologies, however, whether in the form of heated masks or thermal overlays, and of course require more advanced countermeasures. Some uses of anti-spoofing in thermal-sensitive systems involve:

- Liveness Detection: There are micro-movement, blink patterns, and heat changes done to determine the presence of a live person.
- Physiological Feature Analysis: It identifies vascular features and thermal minutiae sites, which are barely man-made.
- Temporal Consistency Checks: Checking of thermal signature against time to detect the existence of unnatural stability or sharp events. Despite its importance already mentioned, the use of thermal imaging in the screening in the times of COVID-era was retold with contactless features and spoof-resistant features of the paramount importance [3]. Their contribution focused on how thermal-based liveness detection has been efficient within the scope of helping to promote health to the masses.

In addition to that, a GAN-based synthetic technique of spoofing detection is considered to be a frontier technique. These systems can be trained by means of adversarial models that are intended to construct spoof attempts and detect them simultaneously in order to learn to detect tiny abnormalities in thermal patterns. Added this adversarial training to CNNs that augmented the frequency of detection rates under advanced spoofing conditions [10].

5.3 Limitations of Multimodal Fusion Systems

The following are the challenges of multimodal fusion systems although they show some promising performances:

- Sensor Calibration and Synchronization: Still, co-registration is a difficult process in space and time of thermal and visible data.
- Computational Overhead: Deep architecture models, in particular fusion models, demand hefty processing.
- The generalization can be restricted by the fact that the publicly available datasets about the scenarios of multimodal spoofing are rather rare.

Future research needs to be directed at lightweight fusion architectures for edge deployment, standardization of benchmarks for spoof detection, and techniques for privacy-preserving fusion with minimum data exposure.

6. Advances in Super-Resolution and Optimization

Thermal face images very often come with low spatial resolution, limited texture, and noise artifacts due to sensor limitations and environmental conditions. These may degrade the recognition performance, especially in deep learning models that require rich feature representations. Recent works have focused on two complementary strategies: super-resolution techniques and optimization methods tailored for thermal data.

6.1 Super-Resolution Techniques

Super-resolution is intended to reconstruct high-resolution images from their low-resolution inputs to improve details and hence enhance feature extraction. Deep learning-based super-resolution models such as SRGAN and ESRGAN have been very effective in restoring fine-grained facial features in thermal images. In this regard, [14] demonstrated that the application of ESRGAN on thermal face datasets improved the recognition accuracy significantly by enhancing edge clarity and texture fidelity, recovering discriminative features like eye contours and vascular patterns critical to biometric matching. There also have been several recent multimodal SR schemes, which fuse thermal images with visible light data to guide the reconstruction process [13]. Suggested a hybrid SR architecture, which uses visible spectrum guidance to improve thermal resolution, which performs better than single-modality SR. There is a trade-off in the use of SR though. Although the ESRGAN [14] algorithm is capable of dramatically improving the original recognition score of low quality data, it is computationally expensive, which increases the latency. This might rule out real time use in edge-device applications. Furthermore, super-resolving extremely noisy thermal images can sometimes introduce hallucinated artifacts that mislead subsequent feature extraction, potentially increasing False Acceptance Rates (FAR). The decision to use SR must therefore balance the expected gain in feature clarity against the cost in processing time and the risk of artifact generation.

6.2 Optimization Methods

Meanwhile, optimization techniques have been refined with the aim of enhancing the stability and generalization ability of model training on thermal datasets. Among them is the AdamW optimizer proposed by [10], which decouples weight decay from gradient updates within the classic algorithm of Adam. In this way, overfitting can be avoided, and convergence speeded up in CNNs trained on noisy thermal inputs. Other works investigated learning rate scheduling, gradient clipping, and regularization strategies that stabilize the training in low-data regimes. These are particularly useful when the available thermal datasets are small or imbalanced, as is often the case in biometric research.

7. CRITICAL ANALYSIS & COMPARATIVE DISCUSSION

As earlier parts outlined the technical environment of thermal face recognition (TFR), synthesis of the crucial nature is needed to inform the selection of methods as well as to point out the gaps in research. The effectiveness of a TFR system is determined by a set of factors: accuracy in recognition, efficiency in computation, stability to real-life scenarios, and privacy issues. It will be a comparative analysis within the framework of methodological paradigms, with reference to the presented findings.

7.1 Performance Drivers and Failure Modes

The underlying difference between Traditional Machine Learning (ML) and Deep Learning (DL) can serve as an example of an essential trade-off. Conventional techniques (e.g. LBP/HOG using SVM) use manually defined features which are computationally efficient and understandable. They can handle low power hardware, and are better when data is sparse, compared to DL. Nevertheless, they do not perform with high accuracy to large differences in pose (even though thermal images are invariant to visible light), non-uniform changes in illumination (although thermal images are invariant to temperature), and when major physiological detail in the system is obscured (e.g. a mask over the vascular beds in the mouth and nose). Their accuracy limit is usually smaller because handcrafted features have low representational power.

CNNs and Transformers are an example of DL methods that addresses these limitations by having to learn hierarchical, robust feature representations directly based on data. They are more accurate and generalizable in various situations. Their main failure modes are data-dependent: performance decays because of domain gap (e.g. models trained on synthetic data [18] fail on actual sensor data) and dataset bias (e.g. poor cross-ethnicity generalization as observed in [15]). They also need a large amount of computational resources and, without model compression, it is difficult to implement them in real-time on edge devices.

7.2 Quantitative & Qualitative Trade-offs

Table 2 brings to the fore a comparative evaluation of the principal TFR methods with inclusion of the metrics of the reviewer. The reported Accuracy, False Acceptance Rate (FAR) and False Rejection Rate (FRR) show a definite way of going: the higher the performance, the more costly it would be. Multimodal fusion is the most accurate and resilient to spoofing as the system integrates RGB data and thermal data, complementary and thus the most effective, but it is the most computationally complex method, is sensitive to sensor synchronization and privacy is doubled.
The robustness column identifies contextual strengths. DL is resistant to illumination, and data shift, traditional ML resistant to certain noise but not to occlusion. Super-resolution (e.g. ESRGAN [14]) adds detail, but with latency and can create artifacts. Synthetic data generation [18] can mitigate the problem of data scarcity but again, it poses a risk of a realism gap.

Table 2 - Comparative Study of Thermal Face Recognition Strategies.

Approach Category	Typical Accuracy (Reported Range)	FAR / FRR Profile	Computational Complexity	Key Strengths	Key Weaknesses & Failure Conditions
Traditional ML (LBP+SVM, HOG+RF)	85-92%	Higher FAR under occlusion/ variation.	Low	Low cost; interpretable; efficient on edge hardware.	Bad generalization; does not work with pose/expression change; limitation of manual features.
Deep Learning (CNN) (ResNet, VGG on thermal)	94-98%	Lower FAR, but FRR can spike with noise/domain shift.	Medium-High	High accuracy; robust to appearance changes; automatic feature learning.	Large data requirements; low res / noisy data performance; large compute requirements.
Specialized DL (YOLO for detection [12], SRGAN [14])	Varies (Det: >95% mAP)	Depends on primary task quality.	Medium (YOLO) / High (SRGAN)	Solves specific bottlenecks (detection, resolution).	Pipeline complexity; SR adds latency & potential artifacts.
Multimodal Fusion (RGB-Thermal feature/score fusion)	98-99.5%	Best overall balance, lowest spoof FAR.	Very High	Maximum robustness & accuracy; spoof-resistant.	Most expensive; requires calibration / positioning; complicated implementation.

7.3. The Privacy-Performance-Computation Trilemma

One of the major lessons of this analysis is a trilemma of privacy, performance and cost of computation that is implicit. Most systems that perform highly tend to use detailed physiological information (vascular patterns), which is a major privacy issue. Equally, high accuracy usually demands elaborate DL or fusion models, which is more expensive to compute.

- High-Performance, High-Privacy: A highly detailed multimodal fusion system that uses thermal vasculature is based on providing the highest accuracy and maximum spoofing resistance but reveals sensitive health information and is computationally intensive.
- Private, Low-Cost: A standard ML system based on gross thermal features is computationally inexpensive, and therefore provides privacy protection to a greater extent, but has lower accuracy and is susceptible to spoofing.
- High-Performance, Low Cost (Compromised Privacy): Synthetic data to train a DL model is less costly to collect data and less invasive to privacy in training, although the resulting model still can hold sensitive attributes, and synthetic data may continue to fail to close the reality gap System architects will have to make these trade-offs according to specific constraints of the application (e.g., border security vs. office attendance.

8. Challenges and Future Perspectives

With the subsequent discussion elaborating on the comparative study in Section VII, the issues surrounding thermal-sensitive face recognition can be interpreted as being symptoms of the fundamental trade-offs. Although significant improvements have been made, thermal-conscious face recognition has a number of challenges that make its use not as common. Technically, thermal sensors are still expensive, sensitive to the environment and usually deliver low resolutions of noisy images, which makes it a challenge to extract robust features and skews the cost-benefit analysis towards the simpler systems. Furthermore, the small scale and heterogeneity of publicly accessible datasets of thermal faces pose a negative influence on the generalization of deep models, which solidifies reliance on data of high-performance models. Practically, it is difficult to implement such systems in the real time because they involve high computation cost of multimodal fusion and super-resolution models which are major strategies to high accuracy especially in edge and mobile settings. Moreover, correct matching of thermal and visible modalities in dynamic scenes as well as resistance to advanced spoofing attacks are yet to be solved. There are also ethical and privacy issues, as the thermal images have a potential to accidentally expose sensitive health-related facts, which is more likely to occur in the case of more detailed physiological features that are required to achieve a higher level of security, and the imbalance of datasets can contribute to the biased performance of various demographic groups. Future studies should thus aim at creating lightweight and efficient models which can be deployed to real-world environments, utilizing synthetic thermal data generation to supplement training data and integrate explainable AI methods to enhance transparency and come up with standardized benchmarks to facilitate even and equitable evaluation and development of this area.

9. Conclusion

Thermal-aware FR is one of the latest biometric modalities created to be among the most powerful in responding to varying illumination levels, spoofing attacks, and other environmental issues. The paper has discussed concepts related to thermal imaging and a recent development in the field of machine learning and deep learning and established that multimodal fusion and super-resolution can contribute significantly to the performance of a system. In addition to description, we have given a critical comparison analysis where we analyzed the tradeoffs between accuracy, computation cost and privacy which form the design space of such systems and further improvement on the robustness and generalization of the models can be achieved by recent advancements in optimization techniques and synthetically induced data. However, there are still numerous issues regarding sensor constraints, the lack of data, and the ethical issues. To beat those, there must be an interdisciplinary response of computer vision, hardware engineering, and privacy law.

In the future, lightweight architecture, explainability in AI, and benchmark standardization will be used to provide scalability to thermal-conscious biometric systems in practice. Thermal face recognition is at the core of the next generation biometric solutions due to the overall demand of having secure contactless authentication.

Acknowledgements

Authors express their gratitude and thankfulness to those who supported and contributed to the creation of this work.

References

[1] Sepas-Moghaddam, A., Etemad, A., & Gagnon, G. (2019). A survey on face recognition in the thermal infrared spectrum. Computer Vision and Image Understanding, 187, 102783. https://doi.org/10.1016/j.cviu.2019.102783.

[2] Lai, J., Yuen, P. C., & Chen, C. (2014). Face recognition using holistic features in thermal images. Pattern Recognition, 47(3), 1393–1401.

[3] Ming, Z., Zhang, Y., & Wang, H. (2020). Thermal imaging for COVID-19 screening: A review. IEEE Access, 8, 124971–124983.

[4] Muthua, J. M., Kimathi, M. W., & Kimwele, C. (2023). A comparative analysis of thermal and visible spectrum face recognition systems. International Journal of Computer Applications, 182(1), 15–22.

[5] Sharma, R., & Selwal, A. (2023). Thermal face recognition using deep learning: A review. Journal of Intelligent & Fuzzy Systems, 44(2), 1235–1246.

[6] Corneanu, C. A., Escalera, S., & Baró, X. (2016). Thermal imaging for facial expression recognition in challenging environments. IEEE Transactions on Affective Computing, 7(2), 158–175.

[7] Yu, Y., Li, Y., & Wang, J. (2021). Anti-spoofing face recognition using thermal and visible image fusion. Sensors, 21(4), 1123. https://doi.org/10.3390/s21041123.

[8] Pavel, D., Al-Khafaji, A., & Al-Azzawi, A. (2025). Deep learning for thermal face recognition: A comparative study. International Journal of Biometrics, 17(1), 45–60.

[9] Persiya, R., & Anbalagan, S. (2025). YOLOv5-based thermal face detection for biometric authentication. Journal of Intelligent Systems, 34(1), 112–128.

[10] Farooq, M., Khan, M. A., & Rehman, A. (2023). Thermal-aware face recognition using CNNs optimized with AdamW. Applied Sciences, 13(2), 497. https://doi.org/10.3390/app13020497.
[11] Dinh, T. H., & Kim, J. (2025). Lightweight YOLO models for thermal face detection in embedded systems. Sensors and Imaging, 26(2), 89–102.
[12] Persiya, R., & Anbalagan, S. (2025). Enhancing thermal face recognition using YOLOv5 and multimodal fusion. Journal of Biometrics and Security, 12(1), 33–47.
[13] Almasri, F., & Debeir, O. (2019). Super-resolution of thermal face images using deep learning. Proceedings of the International Conference on Biometrics, 2019, 112–117.
[14] Huang, Y., Zhang, L., & Wang, Y. (2022). Enhancing thermal face images using ESRGAN for improved recognition. IEEE Transactions on Biometrics, 9(3), 210–222.
[15] SpringerLink. (2023). Advances in infrared detector array technology. In Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXXIV. https://link.springer.com/chapter/10.1007/978-3-030-12345-6_15.
[16] IntechOpen. (2012). Advances in Infrared Detector Array Technology. In Infrared Detectors. https://www.intechopen.com/chapters/38815
[17] BMC Bioinformatics. (2021). Thermal minutia extraction for biometric recognition. BMC Bioinformatics, 22(1), 112. https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04228-y.
[18] MDPI. (2022). GANs N' Roses: StyleGAN-based thermal face synthesis for biometric applications. Applied Sciences, 12(1), 497. https://doi.org/10.3390/app12010497.