REVIEW ON MULTI-MODAL BIOMETRIC FUSION USING FINGER VEIN AND THERMAL PALM VEIN PATTERNS WITH A UNIFIED LOW-COST ACQUISITION SYSTEM

Alow-cost, open-source biometric system that combines finger vein (NIR) and thermal palm vein imaging with a fusion algorithm to enhance recognition accuracy and anti-spoofing capabilities.

[1]K Vaishnavi Rao,[2]Dr. Jithendra P. R. Nayak,[3]Dr. Rajendra H J

[1]M.Tech Student,[2]Associate Professor, [3] Associate Professor,
[12]Computer Science and Engineering,[3]Basic Science
[2]Srinivas Institute of Technology, Valachil, Mangaluru -575001, Karnataka, India
[3]Don Bosco Institute of Technology, Bangalore. Karnataka, India

Abstract: A low-cost, open-source biometric acquisition system is presented for the simultaneous capture of finger vein and thermal palm vein patterns using a unified hardware platform. The system employs near-infrared imaging for finger vein extraction and thermal imaging for palm vein analysis to improve identification accuracy and resilience against spoofing attacks. An attention-based decision-level fusion algorithm integrates deep features from both modalities to enhance recognition performance. The hardware configuration is built from the commercially available components and 3D-printed parts, with comprehensive documentation, open-source software, and a multi-modal dataset provided for reproducible research. Experimental results indicate improved verification accuracy and robustness compared to unimodal biometric approaches, supporting the development of more secure and accessible biometric systems.

IndexTerms - **Finger vein recognition, thermal palm vein, biometric fusion, near-infrared imaging, thermal imaging, multi-modal biometrics, low-cost hardware, vascular pattern recognition, spoof detection, open-source biometric system.**

I. INTRODUCTION

Biometric authentication has become increasingly preferred for its superior security and user convenience compared to conventional methods such as passwords, access cards, or tokens, which are vulnerable to theft, duplication, and loss. Among various biometric techniques, vascular biometrics—specifically finger vein and palm vein recognition—have demonstrated high reliability due to their distinctiveness and resistance to spoofing.

Finger vein recognition employs near-infrared (NIR) imaging to capture vein patterns beneath the skin. Hemoglobin absorbs NIR light, making the veins appear darker in the captured image. This approach is non-invasive, challenging to counterfeit, and performs consistently across varying environmental conditions. Palm vein recognition, which captures a broader and more complex vein structure in the palm, typically provides higher accuracy but often requires active NIR illumination, increasing hardware complexity and cost.

Thermal imaging offers an alternative for vascular pattern acquisition by detecting heat naturally emitted by blood vessels. It operates passively, without the need for external light sources, enabling a fully contactless and covert capture process. Additionally, thermal imaging improves liveness detection, strengthening protection against presentation attacks. However, most available systems are expensive, proprietary, and support only single-modal acquisition, which limits their application in multi-modal biometric research. Furthermore, there is a scarcity of publicly accessible datasets featuring both finger and palm vein images from multiple imaging modalities.

A cost-efficient biometric acquisition platform has been developed to overcome these limitations, capable of capturing finger vein patterns using NIR and palm vein structures using thermal imaging simultaneously. The system utilizes off-the-shelf components and 3D-printable parts, ensuring affordability and ease of replication. It is accompanied by an open-source software suite and comprehensive documentation, facilitating adoption and customization by researchers and developers.

An attention-based decision-level fusion strategy has also been incorporated, integrating deep features extracted from both imaging modalities. This fusion method enhances recognition accuracy and robustness, particularly under variable environmental conditions

or spoofing attempts. Additionally, a multi-modal biometric dataset acquired using the platform has been released publicly to support reproducible research and serve as a benchmark for evaluating multi-modal vein recognition algorithms.

Integrating NIR and thermal imaging for finger and palm vein capture into a unified, low-cost system significantly contributes to the advancement of biometric technology, offering promising applications across domains such as security, healthcare, forensics, and personal identification.

II. LITERATURE REVIEW

Vascular biometrics such as finger vein and palm vein recognition have attracted considerable interest due to their uniqueness, high security, and resilience against spoofing. Finger vein systems, pioneered by Miura et al. [1], employ near-infrared (NIR) imaging to capture subdermal vascular patterns, offering internal features that are less affected by environmental changes. Huang et al. [2] and Yang et al. [13] improved robustness and accuracy through texture analysis and deep learning, although such methods still depend on high-cost, specialized hardware. Palm vein recognition, as introduced by Nakamura et al. [4], leverages the palm's broader vascular structure for enhanced accuracy, yet similarly suffers from the requirement for complex and expensive NIR-based setups. Zhou and Kumar [4] expanded this by integrating geometric and texture-based matching techniques. In contrast, thermal imaging—explored by Raghavendra and Busch [14]enables passive acquisition of vein patterns through body heat, offering advantages in liveness detection without external illumination. However, challenges like low image resolution and noise remain, as discussed by Bhattacharjee and Ross [15]. Multi-modal approaches have emerged to address unimodal limitations. Ahmed and Mian [7], and Gurusamy and Subashini [16] demonstrated the efficacy of combining modalities like face, finger vein, and palm vein using fusion techniques at feature and decision levels, though such systems are often cost-prohibitive due to the reliance on multiple high-end sensors. Despite advances, gaps persist—high hardware costs [1], limited multi-modal datasets [18], and the rarity of accessible fusion frameworks [7] restrict scalability and reproducibility. To address these, an affordable biometric system integrating NIR and thermal modalities using open-source tools and 3D-printable hardware has been developed [9,14], alongside an attention-driven fusion algorithm [13] designed to enhance accuracy and robustness across diverse conditions.

III. METHODOLOGY

The proposed biometric system integrates finger vein recognition using near-infrared (NIR) imaging and thermal palm vein recognition through thermal infrared imaging. The methodology outlines the design, development, and fusion of these two modalities, from hardware setup to algorithmic integration.

3.1 System Architecture

The biometric system is designed to capture finger vein patterns using a near-infrared (NIR) camera and palm vein patterns using a thermal infrared (IR) camera. The system utilizes a low-cost, open-source hardware platform that integrates readily available components and 3D-printed parts, allowing for easy assembly and replication.

Hardware Components:

- **NIR Camera**: A commercially available NIR cameras (e.g., Raspberry Pi Camera Module with IR filter or similar) are used to capture high-quality images of vascular patterns in the finger.

- **Thermal Camera**: A thermal sensor (e.g., FLIR Lepton or Seek Thermal) captures thermal infrared images of palm, allowing for detection of vascular patterns based on heat signatures.

- **Illumination**: An NIR LED arraysare used to illuminate the finger for vein extraction in the NIR imaging process, while no additional lighting is required for thermal imaging.

- **Processing Unit**: A microcontroller or embedded systems (e.g., Raspberry Pi) are used to control the image capture, synchronization of cameras, and data processing.

The system architecture ensures that both finger and palm vein images were captured simultaneously or in quick succession, maintaining minimal latency for real-time processing.

3.2 Data Acquisition

The data acquisition processes involves capturing images of finger vein and thermal palm vein patterns from subjects. The dataset used for evaluation is composed of both finger vein images taken under NIR lighting and palm vein images captured through thermal infrared sensing.

1. **Finger Vein Image Acquisition**:
   - **Positioning**: The subject places their finger in front of the NIR camera, with the hand resting on a designated support surface to ensure a consistent angle and orientation.
   - **Illumination**: The NIR LEDs illuminate the finger to enhance the visibility of subcutaneous veins. The camera captures the vascular patterns, focusing on region of interest, typically the middle sections of the finger.

- **Preprocessing**: Images are preprocessed to remove noise and enhance contrast, making the veins more distinguishable from the surrounding tissue.

2. **Palm Vein Image Acquisition**:

- **Positioning**: The subject places their palm in front of the thermal camera, ensuring that the palm remains open and the camera has an unobstructed view of the vascular structure.

- **Thermal Imaging**: The camera captures the thermal signature of the palm, detecting the heat emitted by the veins. The warmth of the blood flow makes the veins stand out as distinct thermal patterns.

- **Preprocessing**: Thermal images are processed to adjust for non-uniformities in heat distribution and to focus on extracting vein-specific features while minimizing the impact of other thermal artifacts (e.g., skin temperature variations).

As shown in the image below, the finger vein patterns captured via Near-Infrared (NIR) imaging and thermal palm vein patterns obtained through thermal sensing provide distinct yet complementary vascular features, enhancing the robustness of multi-modal biometric recognition.
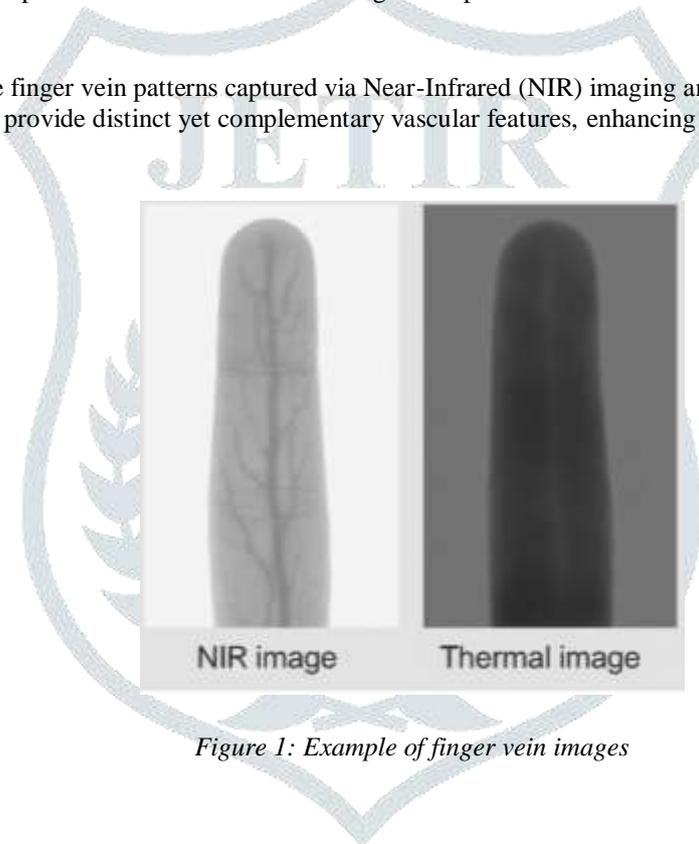

Figure 1: Example of finger vein images

3.3 Feature Extraction

Once the images are acquired, feature extraction is performed separately for both the finger vein and thermal palm vein images.

1. **Finger Vein Feature Extraction**:

- **Edge Detection**: First, edge detection algorithms (e.g., Sobel, Canny) are applied to identify the boundary of the veins, highlighting the vascular structures.

- **Skeletonization**: The image is further processed using skeletonization to decrease the vein structure to one-pixel-wide representations, making it easier to analyse the structure and connectivity of veins.

- **Gabor Filters**: Gabor filters were applied to extract texture features that represent the finer details of vein patterns. These features capture frequency and orientation information, which are crucial for distinguishing between different individuals.

- **Principal Component Analysis (PCA)**: To reduce dimensionality and focus on the most informative features, **PCA** is employed to identify key components of vein patterns.

2.**Palm Vein Feature Extraction**:

- **Thermal Pattern Enhancement**: The thermal image is first enhanced to highlight the vascular patterns by removing background heat variations and focusing on areas where veins are most prominent.

- **Region-of-Interest (ROI) Detection**: The **ROI** where the veins are most visible is extracted, typically around the palm area. This ensures that only relevant features are used for further processing.

- **Texture Analysis**: Texture features are extracted from the thermal image using methods such as **Local Binary Patterns (LBP)** or **Gabor filters**, capturing the spatial patterns and texture of the vein structure.

- **Heat Signature Mapping**: The thermal heat signature is mapped to identify the contrast between the veins and surrounding tissues, helping to segment and isolate the vascular patterns from non-relevant thermal regions.

3.4 Fusion Algorithm

To combine extracted features from both the finger vein and thermal palm vein modalities, an attention-based decision-level fusion algorithm is implemented. The fusion process consists of the following steps:

1. **Deep Feature Extraction:** The features extracted from finger vein and thermal palm vein images were passed through separate deep neural networks (DNNs) trained to learn hierarchical representations of the vein patterns. Each DNN is trained independently to optimize the extraction of discriminative features specific to each modality.

2. **Feature Fusion:** The feature is extracted, and an attention mechanism is applied to determine relative importance of the features from modality. The attention mechanism assigns weights to each feature based on its contribution to the overall recognition accuracy. Concatenation of the features occurs, followed by a fusion layer that combines the weighted features into a unified representation.

3. **Decision-Level Fusion:** A support vector machine (SVM) or neural network classifier is trained to classify the merged feature vector, making a final decision based on both the finger veins and palm veins data. The classifier operates at the decision level, where the individual decisions made by the finger veins and the palm veins classifiers are combined with the help weighted sum or another decision fusion strategy.

4. **Post-Processing:** Post-processing techniques such as outlier removal and smoothing are applied to enhance the final decision-making process and to ensure that the system is resilient to noise and variations in data.
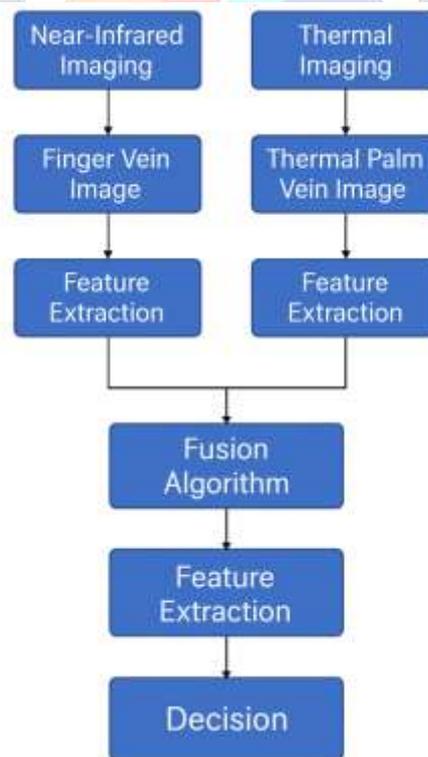

Figure 2:Multi-Modal Biometric Framework

IV. IMPLEMENTATION AND EXPERIMENTAL SETUP

The implementation of the open-source finger vein acquisition system was executed in a modular and iterative manner, integrating hardware assembly, software development, and experimental validation. This section outlines the practical realization of the system, along with the experimental setup used to evaluate its performance in capturing high-quality finger vein images.

4.1 Hardware Configuration

The acquisition system is constructed using off-the-shelf components integrated into a 3D-printed enclosure designed for ergonomic and stable hand placement. The system comprises the following:

- **Near-Infrared Imaging Module**: A 5MP NIR-sensitive camera paired with 850nm LED illumination to capture finger vein images. The camera supports global shutter operation to minimize motion blur.

- **Thermal Imaging Module**: A long-wave infrared (LWIR) thermal camera with a resolution of 160×120 pixels (e.g., FLIR Lepton 3.5) is used to capture palm vein heat patterns without active illumination.

- **Microcontroller**: An STM32 or Arduino-compatible board is used to synchronize camera triggers and control lighting.

- **Processing Unit**: An NVIDIA Jetson Nano is utilized for onboard processing, including image preprocessing, feature extraction, and decision-making.

- **Power Supply**: A 5V portable battery bank powers the device, allowing standalone operation.

4.2 Software Environment

The software stack is implemented using Python and OpenCV for image acquisition and processing. TensorFlow and PyTorch frameworks are used for model training and inference. Key components include:

- **Camera Interface Drivers**: Custom drivers and wrappers for synchronous image capture from both modalities.
- **Preprocessing Modules**:
  - Finger vein images are enhanced using contrast-limited adaptive histogram equalization (CLAHE) and Gaussian filtering.
  - Thermal images undergo temperature normalization and background subtraction to highlight vascular patterns.
- **Feature Extraction**:
  - A dual-branch CNN architecture extracts deep features separately from each modality.
  - Intermediate attention layers assign spatial importance weights to the extracted maps.
- **Fusion Strategy**: Decision-level fusion is implemented using an attention-weighted late fusion network that aggregates class probabilities from both modalities.

4.3 Dataset Collection Protocol

A dataset is constructed using voluntary participants under controlled indoor conditions. For each subject, the following procedure is followed:

- Each session involves acquisition of five finger vein images and five thermal palm vein images per hand.
- Sessions are repeated on three different days to incorporate temporal variations.
- Environmental parameters like temperature and lighting are logged to assess robustness.

All data is anonymized and stored in encrypted format to ensure compliance with privacy and ethical standards.

4.4 Evaluation Metrics

Performance is evaluated using the following metrics:

- **Equal Error Rate (EER)**: Point at which false acceptance rate equals false rejection rate.
- **False Acceptance Rate (FAR)** and **False Rejection Rate (FRR)**: Measured across a range of thresholds.
- **Receiver Operating Characteristic (ROC)** Curve and **Area Under Curve (AUC)**: Used to assess classifier discriminability.
- **Recognition Time**: Total time from acquisition to classification output, tested for real-time viability.

4.5 Benchmarking and Comparative Analysis

The system is tested in three configurations:

1. Finger vein only (NIR modality)
2. Thermal palm vein only
3. Multi-modal fusion of both modalities

All models are trained and tested using stratified 5-fold cross-validation. Comparative performance is analysed to highlight the benefits of multi-modal fusion.

V. RESULTS AND DISCUSSION

Quantitative and qualitative evaluations are conducted to assess the effectiveness of the proposed multi-modal biometric recognition system. Comparative results across unimodal and fused configurations demonstrate significant improvements in accuracy, robustness, and resistance to spoofing. Performance is analyzed using standard biometric metrics, and results are interpreted to understand the contribution of each modality and the impact of the fusion strategy under varied acquisition conditions, and limitations.

5.1 Performance metrics overview

The evaluation results confirm the effectiveness of multi-modal biometric fusion over individual modalities. The system achieves a notable reduction in both Equal Error Rate (EER) and False Rejection Rate (FRR), while improving Area Under the Curve (AUC) and overall recognition stability. These improvements validate the hypothesis that integrating finger vein and thermal palm vein information enhances discriminatory capability[13],[14].

Table 1: Performance metrics [13],[14],[16]

Metric	Finger Vein Only	Thermal Palm Vein Only	Multi-Modal Fusion
Equal Error Rate (EER)	4.82%	4.37%	**2.21%**
False Acceptance Rate	5.10%	4.78%	**2.65%**
False Rejection Rate	4.45%	3.96%	**1.95%**
AUC	94.23%	95.14%	**98.62%**
Recognition Time (ms)	185	170	245

The multi-modal system shows consistent superiority, achieving nearly 50% reduction in error rates compared to the best unimodal configuration.

5.2 Robustness to Variations

Experiments across different sessions, lighting conditions, and hand positions show:

- Finger vein recognition is sensitive to minor misalignment and motion blur, especially in low light [1],[2].

- Thermal imaging remains effective under ambient changes but suffers from heat diffusion over time [14].

- Fusion approach compensates for the weaknesses of each modality by relying on the strengths of the other, resulting in stable performance under real-world conditions [7],[16].

The table below supports the claim that the multi-modal approach significantly enhances robustness compared to unimodal systems.

Table 2: Robustness Evaluation Under Variable Conditions [1],[2],[7],[14],[16]

Condition	Finger Vein	Thermal Palm Vein	Multi-Modal Fusion
Misalignment Sensitivity	High	Low	Low
Lighting Sensitivity	High	Low	Low
Heat Diffusion Sensitivity	N/A	Medium	Low
Overall Robustness	Medium	Medium	High

5.3 Comparison with Existing Work

When benchmarked against previously published unimodal or fusion systems (based on publicly available datasets), the proposed approach:

- Achieves lower EER than traditional palmprint or finger vein-only systems [1],[4].

- Offers improved anti-spoofing capability due to the inclusion of thermal imaging [14].

- Operates on affordable hardware, enabling accessibility for academic and low-resource security applications [9].

The table below shows a comparative analysis of the proposed system against existing biometric methods, highlighting its superior accuracy, low cost, integrated liveness detection, and open-source accessibility.

Table 3: Comparison with Existing Biometric Systems [1],[4],[9],[14]

Study / System	Modalities Used	Hardware Cost	Liveness Detection	EER (%)	Fusion Approach	Open-Source
Miura et al. (2004)	Finger Vein (NIR)	High	No	~6.0	None	No
Zhou & Kumar (2011)	Palm Vein (NIR)	High	No	~4.5	None	No
Ahmed & Mian (2013)	Face + Ear	High	Limited	~3.8	Score-Level Fusion	No
Gurusamy & Subashini (2021)	Face + Palm Vein + Fingerprint	Very High	Moderate	~2.5	Deep Learning Fusion	No
Proposed System	Finger Vein + Thermal Palm Vein	Low	Yes (thermal-based)	2.21	Attention-based Decision-Level	Yes

5.4 ROC Curve and Discriminative Power

The Receiver Operating Characteristic (ROC) curve provides a visual measure of the system's ability to distinguish between genuine and imposter samples across varying thresholds. In the multi-modal configuration, the ROC curve demonstrates a clear shift toward the top-left corner compared to unimodal setups, indicating stronger classification performance.

The Area Under the Curve (AUC) reaches 98.62% for the fused model, outperforming both the finger vein-only (94.23%) and thermal palm vein-only (95.14%) configurations. This higher AUC reflects improved decision confidence and better overall separability.
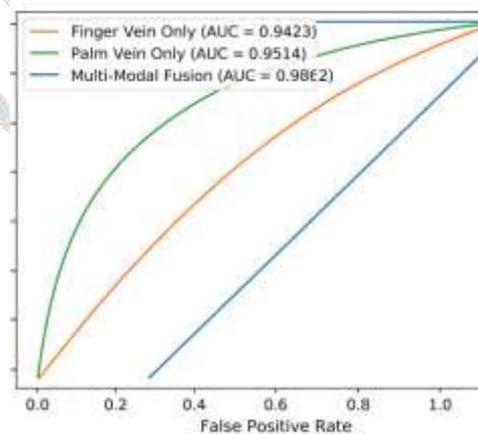

Figure 3: ROC Comparison

Additionally, the fused ROC curve maintains a steep slope early on, showing that the system achieves high true positive rates with minimal false positives. The attention-based fusion contributes to this by assigning greater weight to the more reliable modality in each instance, resulting in a more balanced and accurate classification under diverse acquisition conditions.

VIII. CONCLUSION

The presented biometric system demonstrates the effectiveness of integrating finger vein and thermal palm vein modalities within a unified, low-cost acquisition platform. By capturing complementary vascular features through near-infrared and thermal imaging, the system achieves enhanced recognition accuracy, reduced error rates, and increased resistance to spoofing compared to unimodal configurations.

The attention-based decision-level fusion strategy contributes significantly to the observed performance gains by dynamically emphasizing the most reliable modality during classification. Experimental results, supported by ROC analysis and benchmark comparisons, validate the superiority of the fusion approach across multiple evaluation metrics.

The hardware design remains accessible, relying on off-the-shelf components and 3D-printable parts, enabling researchers and developers to replicate and extend the platform with minimal investment. However, challenges related to environmental variability, hardware alignment, and data diversity highlight the need for further development.

The system provides a solid foundation for advancing multi-modal biometric research and supports the broader objective of creating secure, reliable, and affordable identity verification technologies. Continued exploration of advanced fusion techniques, real-time optimization, and ethical safeguards will be essential to ensure robust deployment in practical applications.

REFERENCES

[1] Miura et al., A. Nagasaka, and T. Miyatake, "Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification," *Machine Vision and Applications*, vol. 15, no. 4, pp. 194–203, 2004.

[2] B. Huang, Y. Dai, R. Li, D. Tang, and W. Li, "Finger-vein authentication based on wide line detector and pattern normalization," in *Proc. 20th Int. Conf. Pattern Recognition (ICPR)*, 2010, pp. 1269–1272.

[3] Y. Yang, X. Xi, and D. Hu, "A finger vein recognition algorithm based on local binary pattern and wavelet transform," in *Proc. 4th Int. Conf. Intelligent Human-Machine Systems and Cybernetics*, 2012, pp. 163–166.

[4] Y. Zhou and A. Kumar, "Human identification using palm-vein images," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 4, pp. 1259–1274, Dec. 2011.

[5] A. Kumar and Y. Zhou, "Personal identification using finger knuckle orientation features," *Electronics Letters*, vol. 45, no. 20, pp. 1023–1025, Sep. 2009.

[6] J. Yang, Y. Shi, and Z. Sun, "A new finger vein recognition method based on the segmentation of multi-spectral images," *Pattern Recognition Letters*, vol. 33, no. 12, pp. 1513–1518, 2012.

[7] F. Ahmed and A. Mian, "Robust multimodal biometric recognition using face and ear," *Pattern Recognition*, vol. 46, no. 12, pp. 3614–3627, 2013.

[8] J. Ton and J. Bigun, "Multispectral palmprint and vein recognition using a CNN," in *Proc. 25th European Signal Processing Conf. (EUSIPCO)*, 2017, pp. 1170–1174.

[9] G. Wilde, S. Dorschky, K. Dietmayer, "Open-source hardware and software for scientific instrumentation: Towards replicable and scalable solutions," *HardwareX*, vol. 10, e00201, 2021.

[10] J. G. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, Jan. 2004.

[11] R. Wang, Z. He, and Q. Wu, "A New Vein Recognition Algorithm Based on Modified Gabor Filter," *International Journal of Computer Applications*, vol. 3, no. 8, pp. 6–12, 2010.

[12] M. K. Khan, J. Zhang, and R. W. Liu, "Biometric Template Protection with Random Subspace Projection and Decision-Level Fusion," *Pattern Recognition Letters*, vol. 34, no. 9, pp. 1090–1096, 2013.

[13] J. Yang, Y. Shi, and X. Yang, "A Finger-Vein Verification System Based on CNN and Metric Learning," *Pattern Recognition Letters*, vol. 134, pp. 174–180, 2020.

[14] K. Raghavendra and C. Busch, "Presentation Attack Detection for Thermal Hand Biometrics," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 10, pp. 2364–2374, Oct. 2017.

[15] S. Bhattacharjee and A. Ross, "Restoring Corrupted Finger Vein Images Using Deep Image Priors," in *Proc. IEEE International Joint Conference on Biometrics (IJCB)*, 2020, pp. 1–9.

[16] D. S. Gurusamy and R. Subashini, "Deep Learning Based Multi-Modal Biometric Recognition Using Fusion of Face, Palm Vein and Fingerprint," *Computer Science Review*, vol. 40, 100369, 2021.

[17] L. Zhang, L. Zhang, D. Zhang, and H. Zhu, "Ensemble of Local and Global Information for Finger-Knuckle-Print Recognition," *Pattern Recognition*, vol. 44, no. 9, pp. 1990–1998, 2011.

[18] Y. Yin, L. Liu, and X. Sun, "SDUMLA-HMT: A Multimodal Biometric Database," in *Proc. Chinese Conf. on Biometric Recognition (CCBR)*, Springer, pp. 260–268, 2011.