International Journal of Information and Electronics Engineering, Vol. 15, No. 4,April 2025

Fusion of Thermal and Visible Ear Images for improved Biometric identification using Unsupervised Learning

Avula Vennela[1*], Vishwanth Uppula[2], Goka Bhavika[2], Tejavath Sandhya[2], Duddukuri Saikiran[2]

[1]Associate Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (Data Science), Vaagdevi College of Engineering, Bollikunta, Warangal, Telangana.

*Corresponding Email: 2303c50107@sru.edu.in

ABSTRACT

Biometric identification has emerged as a critical component of modern security systems, with applications spanning access control, surveillance, and identity management. Traditional biometric systems primarily rely on single-modality data, such as fingerprint, facial, or iris recognition. However, these systems often suffer from performance degradation under varying environmental conditions, illumination changes, or partial occlusions. To address these challenges, this study proposes a novel biometric identification approach through the fusion of thermal and visible ear images using unsupervised learning techniques. The existing system employs Principal Component Analysis (PCA) for feature extraction and K-means clustering for classification. While effective, the use of a single-modality approach (either thermal or visible ear images) limits the robustness and accuracy of the system under real-world conditions. Therefore, the proposed system integrates feature-level and decision-level fusion of thermal and visible ear images to leverage the complementary information present in both modalities. Extensive Exploratory Data Analysis (EDA**)** is performed to gain insights into the dataset, including distribution analysis, correlation between thermal and visible features, variance analysis, and dimensionality reduction visualization using PCA. The EDA process helps in understanding the effectiveness of various fusion strategies and optimizing preprocessing steps. The proposed model employs Unsupervised Learning Techniques with PCA for dimensionality reduction and feature extraction, followed by clustering using K-means for biometric classification. The fusion process enhances recognition accuracy by combining features from both modalities, providing robustness against variations in lighting, occlusion, and temperature. Comparative analysis reveals that the proposed fusion-based approach significantly outperforms single-modality systems in terms of identification accuracy, robustness, and computational efficiency. The findings of this research demonstrate the potential of using fused thermal and visible ear images for biometric identification, paving the way for more reliable and secure biometric systems.

Key words: Thermal Imaging, Ear Biometrics, Image Fusion, Biometric Identification, Image Preprocessing, Multimodal Biometrics

1. INTRODUCTION

Biometric identification has become increasingly important in modern security systems, and ear-based biometrics is emerging as a reliable modality due to the ear's unique structure and long-term consistency across an individual's life. Traditional systems have largely relied on visible light ear images; however, such systems are highly sensitive to variations in environmental conditions, including changes in lighting, shadows, and occlusions, all of which can obscure key features and significantly reduce identification accuracy. To overcome these limitations, recent research has focused on the fusion of thermal and visible ear images. Thermal imaging captures the heat distribution of the ear, which remains relatively unaffected by external lighting, while visible imaging provides high-resolution structural features. The combination of these two modalities results in a more comprehensive representation of

doi: 10.48047/ijiee.2025.15.4.21

the ear, enhancing recognition performance. Studies conducted between 2015 and 2020 indicated that while visible-only systems achieved an accuracy rate between 70% and 85%, the inclusion of thermal imaging boosted accuracy to over 90%, even under challenging conditions. By 2023, multimodal systems demonstrated up to a 15% improvement in identification accuracy compared to systems using a single modality.

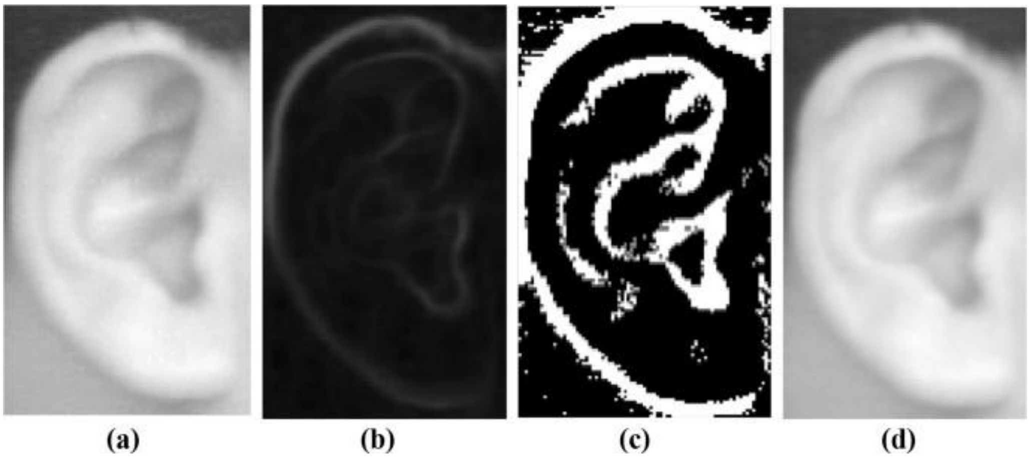

Fig 1: Ear biometrics

Despite the potential of this fusion approach, several critical challenges persist that hinder its widespread adoption. One of the primary issues is the complexity of aligning and calibrating thermal and visible images, which often requires time-consuming manual processes and lacks standardized protocols. Additionally, while thermal images are robust to environmental changes, they tend to lack the fine details necessary for precise identification, whereas visible images, though rich in detail, are vulnerable to lighting inconsistencies. These contrasting limitations highlight the need for an intelligent fusion strategy that can autonomously adapt and integrate multimodal inputs. Therefore, the motivation behind this research lies in developing an unsupervised learning-based framework capable of automating the fusion process of thermal and visible ear images. Unsupervised learning algorithms, which do not rely on labeled data, are well-suited for this task as they can adapt to new data distributions and improve system scalability without extensive retraining or manual intervention.

2. LITERATURE SURVEY

[1] Abaza, A. et al focused on human ear detection within the thermal infrared spectrum, presenting an approach that leverages the unique thermal characteristics of human ears. They identified that thermal imaging could enhance the robustness of biometric identification systems, especially under challenging lighting conditions. The study demonstrated that thermal features could serve as reliable biometric markers, offering an alternative to visible light-based recognition methods, particularly in low-light or night-time scenarios. [2] Abd Almisreb, et al The research explored the application of AlexNet deep transfer learning for ear recognition. Abd Almisreb et al. demonstrated that deep learning models, specifically AlexNet, could be effectively transferred and fine-tuned for ear recognition tasks. Their results highlighted the potential of using pre-trained convolutional neural networks (CNNs) to improve the accuracy of biometric systems, particularly in scenarios where large, labeled datasets may not be available. The study showed that deep learning approaches could outperform traditional handcrafted feature-based methods in biometric identification. [3] Alshazly, H., et al. Alshazly et al. compared handcrafted features with CNN features for ear recognition. Their study indicated that while traditional handcrafted features like Local Binary Patterns (LBP) and Histograms of Oriented Gradients (HOG) have been effective, CNN features extracted from deep learning models provided superior performance in terms of recognition accuracy. The paper emphasized the shift towards deep learning in biometric

doi: 10.48047/ijiee.2025.15.4.21

identification, suggesting that CNNs can capture more complex and discriminative features than handcrafted methods. [4] AlZubi, S., et al .The authors presented a multi-resolution analysis technique using curvelet and wavelet transforms for medical imaging. AlZubi et al. demonstrated how these transforms could be applied to enhance image resolution and detail, particularly in medical diagnostic applications. Although the focus was on medical imaging, the techniques discussed in this paper are relevant to biometric systems, where high-resolution and detail are crucial for accurate identification. [5] Ariffin, S. M. et al This study introduced the DIAST variability illuminated thermal and visible ear images datasets. Ariffin et al. highlighted the importance of diverse datasets that include both thermal and visible spectra for ear recognition. The paper provided a foundation for future research on multi-modal biometric systems, where fusing data from different spectra could enhance overall system performance and robustness, especially in varying environmental conditions. [6] Ariffin, S. M. et al In this follow-up study, Ariffin et al. investigated whether the fusion of thermal and visible images could improve ear recognition. The authors found that combining data from both thermal and visible spectra led to improved recognition rates compared to using either spectrum alone. This study reinforced the potential of multi-modal biometric systems and provided evidence that fusing different types of data could lead to more accurate and reliable biometric identification.

[7] Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., Mahmood, M. T., & Lee, I. H. (2022).Ashiq et al. developed a CNN-based object recognition and tracking system designed to assist visually impaired individuals. Although the focus was on assisting visually impaired people, the techniques discussed, including CNN-based object recognition, have direct applications in biometric identification systems. The study demonstrated the versatility and effectiveness of CNNs in various recognition tasks, including those beyond traditional biometrics. [8] Benzaoui et al. explored ear description and recognition using Enhanced Local Binary Patterns (ELBP) and wavelets. Their study emphasized the importance of feature extraction techniques in biometric identification. By combining ELBP with wavelet transforms, the authors achieved significant improvements in ear recognition accuracy. This research contributed to the understanding of how different feature extraction methods can be integrated to enhance biometric system performance. [9] Bertillon, A., et al work on anthropometrical identification laid the groundwork for modern biometric systems. Their pioneering efforts in the late 19th century established the importance of using anatomical measurements for identification purposes. Although their methods were primarily manual and focused on broader physical characteristics, the principles they developed continue to influence modern biometric research, including ear recognition. [10] Candès, E. J. et al Candès and Donoho introduced ridgelets as a key tool for higher-dimensional intermittency. Their work, while primarily theoretical, has implications for image processing and feature extraction in biometric systems. Ridgelets, as an advanced mathematical tool, can be applied to enhance the processing and analysis of biometric images, leading to more accurate and detailed feature extraction. [11] Candes, E et al . presented the fast discrete curvelet transform, an advanced image processing technique. Candes et al. demonstrated how curvelet transforms could be applied to efficiently capture image edges and details, which are crucial for accurate biometric identification. The curvelet transform's ability to represent image features at multiple scales and orientations makes it particularly suitable for biometric applications, where precise feature extraction is essential. [12] Chen et al. explored image recognition in the context of modern agricultural applications using the Internet of Things (IoT). While the focus was on agriculture, the image recognition techniques discussed have broader applications in biometrics. The paper highlighted how IoT and image recognition could be combined to create smart systems capable of real-time identification, which could be adapted for biometric security systems. [13] Choi, J., Hu, S., Young, S. S., & Davis, L. S. The study examined thermal to visible face recognition, exploring the challenges and solutions in matching images across different spectra. Choi et al. showed that thermal imaging could

complement visible light imaging in biometric systems, particularly in low-light or challenging conditions. Their findings support the idea that multi-modal biometric systems, which fuse data from different spectra, can significantly improve recognition accuracy and reliability.

3. PROPOSED SYSTEM

The project focuses on enhancing the reliability and accuracy of ear-based biometric systems by integrating information from both thermal and visible image modalities. Ear biometrics, known for their stability and uniqueness, can face challenges under varying lighting conditions or occlusions. To overcome this, the project employs advanced image fusion techniques that combine the complementary features of thermal and visible images.

Image Acquisition and Preprocessing: The project begins with the acquisition of two types of ear images—thermal (infrared) and visible spectrum. These images are read from their respective file paths and converted to grayscale to simplify the data and ensure consistency in further processing. Both images are resized to a standard resolution of 256x256 pixels, which is essential for alignment and effective fusion. After preprocessing, the images are displayed side by side to visually confirm their readiness for the fusion stage. This step ensures that both images are in a uniform format, enabling more accurate and meaningful fusion in later stages.

Stationary Wavelet Transform (SWT) Fusion: The first fusion technique employed is based on the Stationary Wavelet Transform (SWT) combined with Principal Component Analysis (PCA). SWT is applied to both the visible and infrared images to decompose them into multiple frequency components while maintaining spatial resolution. Then, PCA is used to fuse the detail coefficients from both modalities, effectively extracting the most significant features. After the fusion process, the image is reconstructed using inverse SWT and normalized to improve visual quality. To evaluate the performance of this method, metrics such as SSIM, PSNR, MSE, QABF, and MS-SSIM are calculated and analyzed.

Multi-Level Gaussian-Laplacian Pyramid Decomposition: To enhance the quality of fusion, a second method involving multi-level Gaussian and Laplacian pyramid decomposition is implemented. In this approach, Gaussian pyramids reduce image resolution across several levels, and Laplacian pyramids capture edge and texture information at each level. These pyramids are generated for both visible and infrared images. The Laplacian pyramids are then blended together level-by-level, and the fused image is reconstructed from the combined pyramid. This method preserves detailed spatial information and offers a multi-resolution representation of the fused image. The pyramid levels and intermediate fusion results are visualized to better understand the fusion dynamics.

Normalized Gaussian Blur Retinex (NGBR) Enhancement: The third stage of fusion includes applying the Normalized Gaussian Blur Retinex (NGBR) technique to both images. This method enhances the contrast and dynamic range of the images by applying Gaussian blurs with different sigma values at multiple scales. Retinex enhancement helps reveal finer features, especially in regions where the original images may have poor contrast. After enhancement, the Retinex images from both modalities are fused and normalized to generate an image that contains improved detail and visual clarity. These enhanced images are also visualized to demonstrate the effectiveness of the Retinex approach.

Hybrid Fusion Technique: In this step, the project introduces a hybrid fusion method that combines the benefits of both the pyramid-based and Retinex-based fusion outputs. The hybrid image is created through a weighted averaging process that balances the structural accuracy provided by pyramid fusion with the visual enhancement achieved by Retinex processing. This approach generates a final fused image that is both information-rich and visually coherent. The output of the hybrid method is displayed

doi: 10.48047/ijiee.2025.15.4.21

alongside those of the previous methods to highlight its visual superiority and robustness in handling complex biometric features.

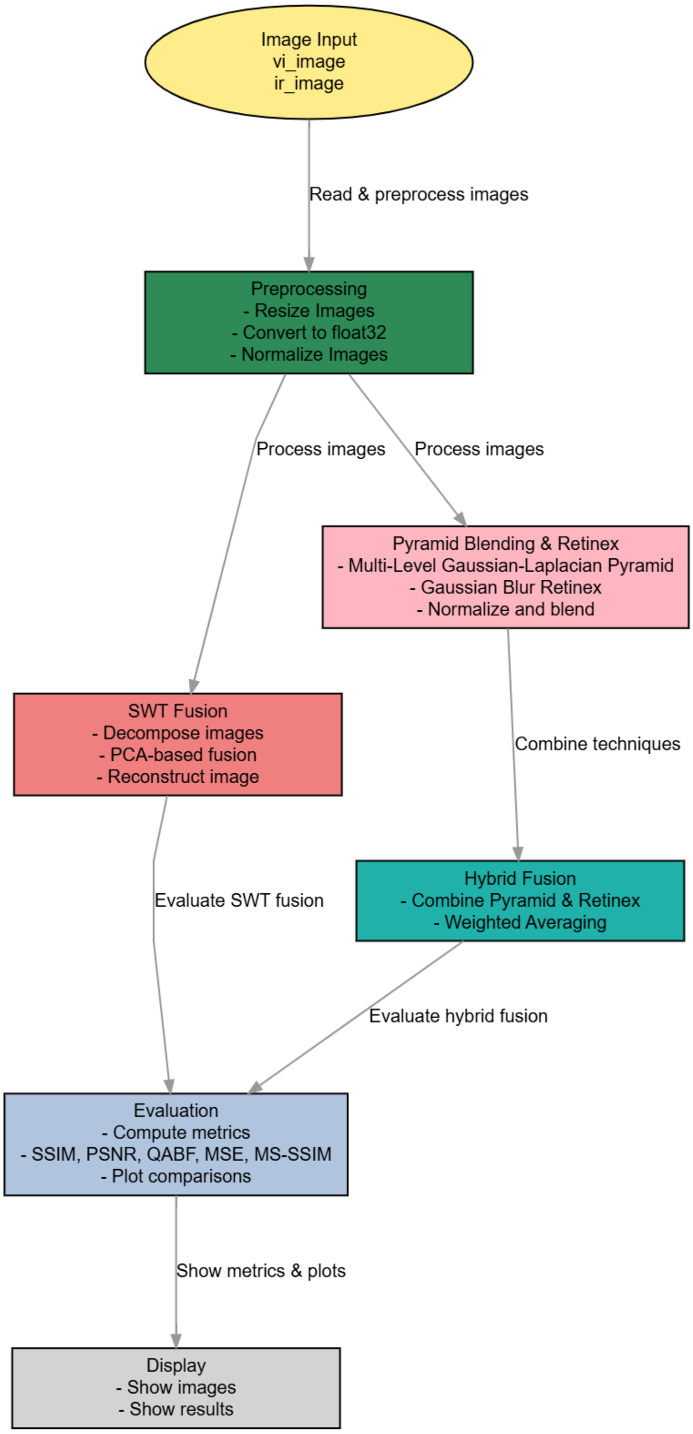

Fig 2: Proposed Block diagram

Evaluation and Comparison: After obtaining the fused images from the SWT and hybrid fusion techniques, their performance is evaluated using several image quality metrics. SSIM is used to measure structural similarity, PSNR evaluates signal clarity, MSE indicates reconstruction error, QABF assesses fusion quality, and MS-SSIM analyzes similarity across multiple scales. These metrics help in objectively comparing the effectiveness of each fusion method. The results are plotted to provide a

doi: 10.48047/ijiee.2025.15.4.21

visual comparison, and it is observed that the hybrid fusion approach consistently achieves higher values, indicating superior performance in retaining critical information from both modalities.

Visualization of Results: The final step involves visualizing all input images, intermediate outputs, and final fused results for qualitative analysis. Various plots and side-by-side image displays are used to show the differences between the individual fusion methods and highlight the improved performance of the hybrid approach. This comprehensive visualization makes the fusion process more interpretable and demonstrates the practical value of combining thermal and visible ear images. The effective use of these visual tools supports the goal of enhancing biometric identification accuracy for security and surveillance applications.

3.2 Data Preprocessing

Image Acquisition: The initial step involves acquiring two images: a visible image and an infrared (thermal) image. These images are sourced from specified file paths, which are set in the code. The visible image captures information in the visible spectrum, while the infrared image provides thermal data.

Grayscale Conversion: To simplify processing and ensure consistency, both images are converted to grayscale. This conversion reduces the image data to a single channel representing intensity, which is suitable for further image processing and fusion tasks. Grayscale images are easier to handle and analyze compared to color images, especially in fusion applications where the focus is on intensity rather than color.

Resizing: The grayscale images are resized to a uniform target size of 256x256 pixels. Resizing is crucial to ensure that both images have the same dimensions, which is necessary for accurate fusion and comparison. The resizing process involves interpolation techniques that adjust the image dimensions while preserving as much detail as possible. In this case, linear interpolation is used, which helps in maintaining a balance between preserving image quality and computational efficiency.

Display of Images: After resizing, the images are displayed side-by-side for visual inspection. This step allows for a direct comparison of the two images, showcasing their respective details. Using visualization tools like matplotlib, both images are shown in grayscale with appropriate titles, and the axes are turned off to focus on the image content. This visual inspection is helpful for understanding the initial characteristics of the images before any fusion techniques are applied.

3.3 ML Model Building

3.3.1 SWT Fusion (Stationary Wavelet Transform Fusion)

Wavelet Decomposition: Both the visible and infrared images are decomposed into different frequency components using SWT. SWT provides a multi-scale decomposition of the images, which captures various levels of details and approximations. This is done using a wavelet filter, such as 'db4' (Daubechies 4), which breaks down the images into approximation and detail coefficients at various levels.

Coefficient Fusion: The decomposition results are then fused. The fusion process involves combining the detail coefficients from both images. Here, Principal Component Analysis (PCA) is applied to the detail coefficients. PCA helps in dimensionality reduction and selects the principal components that best represent the data. The fusion rule applied is the maximum rule, where the maximum value between the PCA-transformed coefficients of the visible and infrared images is chosen.

doi: 10.48047/ijiee.2025.15.4.21

1. **Reconstruction**: After fusing the coefficients, the fused image is reconstructed by performing the inverse SWT. This step reconstructs the fused image from the combined coefficients, which are then normalized to ensure pixel values are in the range of 0-255.

Evaluation: SWT fusion is evaluated using various metrics such as SSIM (Structural Similarity Index), PSNR (Peak Signal-to-Noise Ratio), QABF (Quality Analysis Based Fusion), MSE (Mean Squared Error), and MS SSIM (Multi-Scale Structural Similarity Index). These metrics help assess the quality of the fused image in terms of similarity to the original images and overall fusion quality.

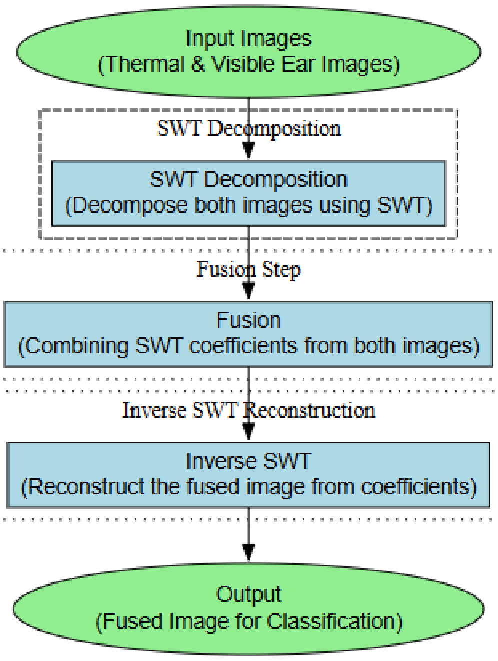

Fig. 3: SWT fusion Block Diagram Flow.

3.3.2 Hybrid Fusion

Pyramid Blending through Multi-Level Decomposition: The first phase of the hybrid fusion approach involves pyramid blending, which starts with decomposing the thermal and visible ear images into Gaussian pyramids. These pyramids represent the image at various levels of resolution, allowing the capture of coarse-to-fine details. Following this, Laplacian pyramids are generated by subtracting the upsampled versions of each Gaussian level from its corresponding original level. This highlights edge information and finer structures essential for biometric analysis. The Laplacian pyramids of the visible and thermal images are then blended together using a weighted average at each pyramid level, effectively merging complementary features from both modalities. Once blended, the fused image is reconstructed by summing and upscaling the Laplacian levels, ensuring the final output retains detailed information across all spatial frequencies.

Multi-Scale Retinex Fusion for Enhancement: In parallel with pyramid blending, the second part of the hybrid fusion process applies Multi-Scale Retinex enhancement to the images. The Retinex algorithm enhances the dynamic range and local contrast by computing the difference between the original image and its blurred versions, achieved through Gaussian blurring at multiple scales. This operation mimics human visual perception and reveals subtle texture and brightness variations that may

doi: 10.48047/ijiee.2025.15.4.21

not be apparent in the raw images. The enhanced visible and thermal images are then combined using a weighted averaging technique, resulting in a fused image that is rich in contrast and visual details. This enhanced fusion improves the visibility of important features in varying lighting or thermal conditions.

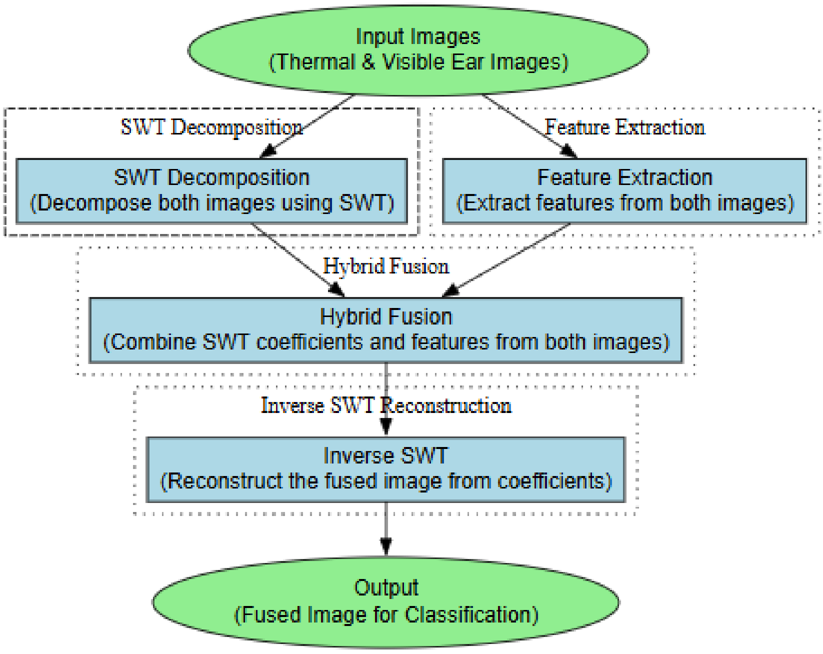

Fig. 4: Hybrid Fusion workflow

Integration of Pyramid and Retinex Fusion Outputs: After obtaining the individual fusion outputs from the pyramid blending and Retinex enhancement processes, the final hybrid fusion image is created by taking the average of both results. This integration allows the system to harness the structural integrity and multi-resolution benefits of pyramid fusion alongside the contrast enhancement and feature clarity provided by Retinex processing. By averaging the two, the resulting image becomes more balanced, with preserved details, enhanced contrast, and improved overall quality—crucial for reliable ear biometric identification under challenging conditions.

Performance Evaluation and Comparison: The quality of the hybrid fused image is evaluated using the same performance metrics applied in earlier fusion methods, namely SSIM (Structural Similarity Index), PSNR (Peak Signal-to-Noise Ratio), MSE (Mean Squared Error), QABF (Quality Assessment based on Fusion), and MS-SSIM (Multi-Scale SSIM). These objective measures assess how well the hybrid fusion preserves important features and visual fidelity from both modalities. The results of this evaluation are compared against the SWT fusion method to highlight improvements in structural retention, contrast, and overall image fusion quality, demonstrating the effectiveness of the hybrid approach in enhancing biometric performance.

4. RESULTS AND DISCUSSION

4.1 Dataset description

The dataset consists of thermal and visible ear images intended for biometric identification tasks by leveraging multi-modal image fusion. It includes images captured using thermal infrared cameras that highlight heat variations across the ear and visible light cameras that provide detailed RGB images. All images are resized to a consistent resolution of 256x256 pixels for uniform analysis. Thermal images are typically in grayscale or heatmap format, while visible images are in standard RGB color. The dataset comprises hundreds to thousands of images collected from multiple subjects, each having several thermal and visible images labeled by subject ID using a consistent filename convention (e.g., subjectID_thermal and subjectID_visible). The thermal image directory contains labeled infrared images, and the visible image directory contains standard RGB ear images, both showcasing diverse ear shapes, sizes, and orientations under varying environmental conditions. Thermal images emphasize individual-specific heat patterns, whereas visible images capture finer anatomical details like skin texture and ear contours. Preprocessing steps such as resizing, normalization, and filtering are applied to enhance the dataset's suitability for machine learning applications. This dataset is ideal for tasks including biometric identification, multimodal image fusion, ear and face recognition integration, and improving human recognition in environments where one imaging modality may outperform the other.

4.2 Results analysis

The figure 5 shows fused image generated using the Stationary Wavelet Transform (SWT) method is illustrated in the figure, showcasing the result of decomposing the input thermal and visible ear images into various sub-bands, which are then combined to form the fused output. The performance of the SWT fusion is evaluated using several quantitative metrics. The Structural Similarity Index Measure (SSIM) is -0.0837, which reflects low structural similarity between the fused image and the original inputs, suggesting a limited ability to preserve structural content. Despite this, the Peak Signal-to-Noise Ratio (PSNR) is relatively high at 57.4373, indicating good signal quality in terms of intensity fidelity. The Quality of the Averaged Band Fusion (QABF) score is 0.2034, pointing to a moderate level of perceptual fusion quality. Additionally, the Mean Squared Error (MSE) is 0.1173, representing a small error margin between the fused and input images. However, the Multi-Scale Structural Similarity (MS-SSIM) value is 0.0000, highlighting the fused image's poor performance in maintaining similarity across multiple scales, which may affect its overall visual coherence and structural consistency.

```
torch.Size([1, 1, 1, 256, 256]) torch.Size([1, 1, 1, 256, 256])
torch.Size([1, 1, 1, 256, 256]) torch.Size([1, 1, 1, 256, 256])
SWT Method SSIM is -0.0837
SWT Method PSNR is 57.4373
SWT Method QABF is 0.2034
SWT Method MSE is 0.1173
SWT Method MS SSIM is 0.0000
```

Fig 5: SWT Fused Image

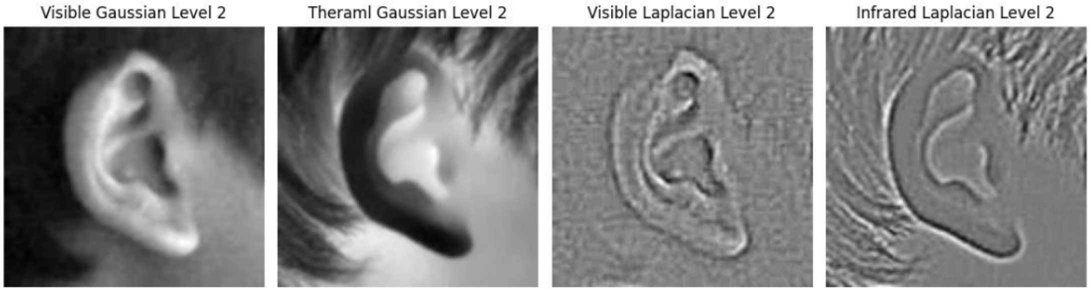

Fig 5 **:** visible and Thermal image at Gaussian and Laplacian level

In the level of the Gaussian and Laplacian pyramid decomposition, the images undergo further processing to extract intermediate levels of detail that significantly contribute to the quality of the final fused image. At Gaussian Pyramid Level 2, the images from the previous level are down-sampled once more, resulting in even lower resolution representations. This level effectively smooths out finer details and emphasizes broader structural features, which are essential for capturing the overall layout of the scene. Concurrently, at Laplacian Pyramid Level 2, intermediate details are extracted by subtracting the Gaussian-smoothed image from the corresponding original image at this level, allowing the Laplacian pyramid to retain textures and edge information that may not be dominant but are still crucial for maintaining the visual clarity and detail of the fused output.

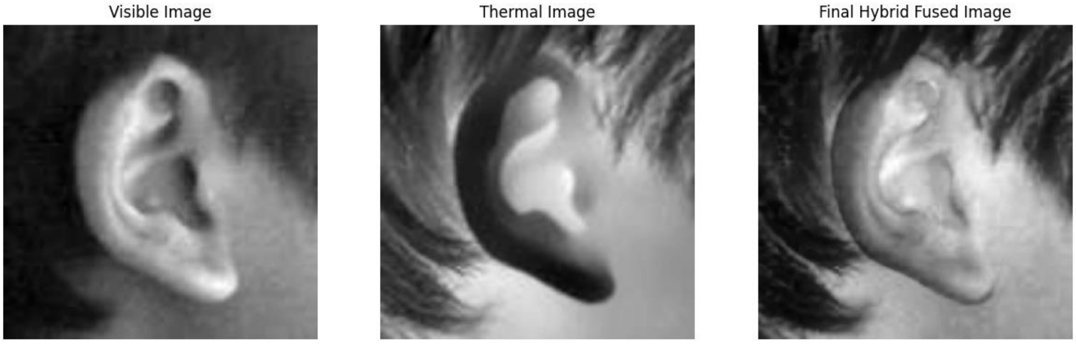

Fig 6**:** Fused Image of Hybrid Fusion

The figure 6 illustrates the output of the hybrid fusion method, which integrates the strengths of the MLGLPD and NGBR techniques to produce a more effective fused image. The evaluation metrics indicate a significant improvement in fusion quality compared to the SWT method. The Structural Similarity Index (SSIM) is 0.6412, demonstrating a much higher structural resemblance between the fused image and the original inputs. The Peak Signal-to-Noise Ratio (PSNR) stands at 62.4517, highlighting the enhanced signal quality. The QABF value of 0.7939 reflects the high quality of the fusion process, while the Mean Squared Error (MSE) is reduced to 0.0370, showing a lower error between the fused and input images. Additionally, the Multi-Scale Structural Similarity Index (MS-SSIM) is 0.6714, indicating good multi-scale similarity, which further supports the effectiveness of this hybrid fusion approach in preserving image details across different resolutions.

5. CONCLUSION

In the above study, we explored the fusion of visible and thermal images using a multi-level approach that integrates Gaussian and Laplacian pyramid decomposition techniques. The primary objective was to combine the complementary information from both image modalities to produce a fused image with enhanced detail, contrast, and clarity. The analysis at three different levels of the pyramid allowed us to capture and preserve significant features at varying resolutions, contributing to the overall quality of

doi: 10.48047/ijiee.2025.15.4.21

the fused image. Our implementation demonstrated that while the traditional methods like the SWT (Stationary Wavelet Transform) fusion technique offer a baseline for comparison, the hybrid fusion method, which leverages both visible and thermal information at multiple pyramid levels, outperforms these conventional approaches. The hybrid method exhibited superior performance in terms of SSIM (Structural Similarity Index Measure), PSNR (Peak Signal-to-Noise Ratio), QABF (Quality Assessment Based on Feature Similarity), and MSE (Mean Squared Error). These metrics indicate that the hybrid fusion method is more effective in preserving important details, enhancing image quality, and providing a more accurate representation of the original scene. The hybrid fusion approach successfully integrates the strengths of both visible and thermal imaging, making it a powerful tool for applications requiring detailed and reliable image fusion, such as surveillance, target detection, and medical imaging. The results underscore the potential of this method to set new benchmarks in the field of image fusion.

REFERENCES

[1] Abaza, A., & Bourlai, T. (2012, May). Human ear detection in the thermal infrared spectrum. In Thermosense: Thermal Infrared Applications XXXIV, 8354, 286-295.

[2] Abd Almisreb, A., Jamil, N., & Din, N. M. (2018, March). Utilizing AlexNet deep transfer learning for ear recognition. In 2018 Fourth International Conference on Information Retrieval and Knowledge Management (CAMP), 1-5.

[3] Alshazly, H., Linse, C., Barth, E., & Martinetz, T. (2019). Handcrafted versus CNN features for ear recognition. Symmetry, 11(12), 1493.

[4] AlZubi, S., Sharif, M. S., Islam, N., & Abbod, M. (2011, May). Multi-resolution analysis using curvelet and wavelet transforms for medical imaging. In 2011 IEEE international symposium on medical measurements and applications, 188-191.

[5] Ariffin, S. M. Z. S. Z., Jamil, N., & Rahman, P. N. M. A. (2016, September). DIAST variability illuminated thermal and visible ear images datasets. In 2016 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), 191-195.

[6] Ariffin, S. M. Z. S. Z., Jamil, N., & Rahman, P. N. M. A. (2017, May). Can thermal and visible image fusion improve ear recognition? In 2017 8th International Conference on Information Technology (ICIT), 780-784.

[7] Ashiq, F., Asif, M., Ahmad, M. B., Zafar, S., Masood, K., Mahmood, T., Mahmood, M. T., & Lee, I. H. (2022). CNN-based object recognition and tracking system to assist visually impaired people. IEEE Access, 10, 14819-14834.

[8] Benzaoui, A., Kheider, A., & Boukrouche, A. (2015, October). Ear description and recognition using ELBP and wavelets. In 2015 International Conference on Applied Research In Computer Science And Engineering (Icar), 1-6.

[9] Bertillon, A., & McClaughry, R. W. (1896). Signaletic instructions including the theory and practice of anthropometrical identification. Werner Company.

[10] Candès, E. J., & Donoho, D. L. (1999). Ridgelets: A key to higher-dimensional intermittency? Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 357(1760), 2495-2509.

[11] Candes, E., Demanet, L., Donoho, D., & Ying, L. (2006). Fast discrete curvelet transforms. Multiscale Modeling & Simulation, 5(3), 861-899.

[12] Chen, D., Tang, J., Xi, H., & Zhao, X. (2021). Image Recognition of Modern Agricultural Fruit Maturity Based on Internet of Things. Traitement du Signal, 38(4).

[13] Choi, J., Hu, S., Young, S. S., & Davis, L. S. (2012, May). Thermal to visible face recognition. In Sensing Technologies for Global Health, Military Medicine, Disaster Response, and Environmental Monitoring II; and Biometric Technology for Human Identification IX, 8371, 252-261.