# IMAGE FUSION EMERGINF APPLICATIONS AND TECHNIQUES

Report submitted to GITAM (Deemed to be University) as a partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in (write your respective branch)


**HARIPRIYA KAMSALA**
**Regd. No: BU22EECE0100518**
**Sadiya Samrin K**
**Regd. No: BU22EECE0100317**

DEPARTMENT OF ELECTRICAL, ELECTRONICS AND COMMUNICATION ENGINEERING
GITAM SCHOOL OF TECHNOLOGY
GITAM (DEEMED TO BE UNIVERSITY)
BENGALURU -561203


NOV 2025

# DECLARATION

I/We declare that the project work contained in this report is original and it has been done by me under the guidance of my project guide.

Name: Haripriya Kamsala
Sadiya Samrin K

Date:23/09/2025 Signature of the Student

**Department of Electrical, Electronics and Communication Engineering**
**GITAM School of Technology, Bengaluru-561203**

## CERTIFICATE

This is to certify that Haripriya Kamsala bearing Regd. No.: BU22EECE0100518 and Sadiya Samrin bearing Regd. No.: BU22EECE0100317 has satisfactorily completed Mini Project Entitled in partial fulfillment of the requirements as prescribed by University for VIIth semester, Bachelor of Technology in "Electrical, Electronics and Communication Engineering" and submitted this report during the academic year 2025-2026.

[Signature of the Guide] [Signature of HOD]

# Table of contents

# Chapter 1: Introduction

## 1.1 Overview of the problem statement

In the modern digital era, the ability to capture and process high-quality images plays a vital role across diverse domains such as healthcare, surveillance, autonomous navigation, and remote sensing. However, a single imaging modality often fails to provide complete information about a scene due to limitations such as shallow depth of field, varying illumination, or restricted spectral sensitivity.

Image fusion has emerged as a solution to this challenge by combining information from multiple input images into a single enhanced image. Three important applications of image fusion are **multi-focus fusion**, **multi-exposure fusion**, and **thermal-RGB fusion**. Multi-focus fusion addresses the problem of shallow depth of field, producing an all-in-focus image. Multi-exposure fusion combines differently exposed images to preserve details in both shadows and highlights. Thermal-RGB fusion merges temperature information with visual details, which is particularly useful in low-light environments, medical imaging, and surveillance.

The lack of a unified framework for handling these fusion tasks and the need for systematic comparison of different algorithms form the core problem addressed in this project. The study aims to evaluate and optimize multiple fusion approaches, identifying their suitability for different application domains.

## 1.2 Objectives and Goals

The primary objectives of this project are as follows:

1. To study and analyze existing image fusion techniques across pixel, spatial, and transform domains.
2. To implement traditional approaches for multi-focus, multi-exposure, and thermal-RGB image fusion with suitable preprocessing and postprocessing methods.
3. To compare the performance of different fusion algorithms using qualitative and quantitative metrics.
4. To explore machine learning and deep learning-based methods for improved fusion performance.
5. To identify the most effective approaches for each application scenario and propose an optimized framework for image fusion.
6. To evaluate the potential real-world applications of the developed techniques in areas such as autonomous navigation computer vision, and security.

Through these goals, the project aims to contribute to the development of robust and adaptive image fusion techniques, offering insights into their practical deployment in emerging fields.

# Chapter 2: Literature Review

Image fusion has been an active area of research for several decades, with applications expanding as imaging technologies advance. Early approaches relied on simple pixel-level techniques, while modern methods leverage advanced transform-based, machine learning, and deep learning models.

## Literature Survey

## 1. Image fusion — overview and foundational methods

Image fusion integrates complementary information from multiple imagery sources to form a composite image that is more informative for human interpretation or automated processing than any single input [1]. Foundational work framed fusion at multiple abstraction levels — pixel, feature, and symbol — each with different pre-processing and robustness requirements; pixel-level fusion preserves original information but is highly sensitive to misregistration, while feature- and symbol-level fusion afford more semantic integration at the cost of additional processing such as segmentation or classification [2]. Multiresolution transforms (pyramids, wavelets, contourlets, NSCT) became popular because they separate image content across scales, enabling selective fusion of approximation and detail bands; statistical wavelet-domain modeling also enhanced robustness to noise and texture variations [3], [14]. Sparse representation (SR) approaches marked a shift to *data-driven* basis design: instead of fixed transform bases SR learns overcomplete dictionaries that better capture salient structures (edges, textures) of images and thereby often outperform fixed MSTs in subjective and objective metrics, albeit at higher computational cost and dictionary-dependence [5]. More recently, deep learning has produced end-to-end fusion frameworks capable of learning fusion policies and perceptual objectives directly, but this progress comes with new challenges: large annotated datasets are scarce for many fusion problems, learned models can be hard to interpret, and standardized evaluation remains incomplete [6], [9]. Across paradigms, two persistent practical needs are robust registration (especially for pixel-level fusion) and standardized metrics and benchmarks to allow fair comparison [2], [3], [13].

## 2. Multi-focus image fusion

Multi-focus fusion specifically addresses limited depth-of-field by merging images focused at different depth planes into an all-in-focus image. Piella (2003) offered a general multiresolution (MR) fusion framework from pixels to regions, formalizing how MR decomposition plus coefficient-level fusion yields coherent results and how region-based strategies can reduce pixel-wise artifacts [7]. Traditional transform-domain approaches (DWT, NSCT, curvelet) dominated because they preserve multiscale structure, though they can be sensitive to shift variance and misregistration; redundant transforms such as NSCT or SWT mitigate some of those sensitivities [8], [14]. To better capture local structure, sparse representation (SR) methods emerged: Zhang et al. (2018) reviewed SR-based multi-sensor fusion, detailing representation models, dictionary learning strategies (global/adaptive/coupled) and activity measures; SR often yields better edge and texture preservation than fixed MSTs and is more robust to small misalignments due to overlapping patches, but its effectiveness strongly depends on dictionary quality and it increases computational expense [5]. Feature-based innovations also advanced the field. Liu et al. (2015) demonstrated that **dense SIFT** descriptors provide a powerful local activity measure: using sliding windows, DSIFT produces reliable initial decision maps which are then refined via feature matching and local focus comparison, improving the handling of mis-registered pixels and object edges relative to many earlier methods [9]. Hybrid methods combine strengths: Aymaz & Köse (2019) applied **super-resolution** preprocessing to increase source image detail, then used Stationary Wavelet Transform (SWT) + PCA to fuse subbands; this hybrid reduces edge loss and spatial distortion and yields visually clearer results, at the cost of extra computation and sensitivity to interpolation artifacts if super-resolution is mismatched [10].

# 3. Multi-exposure image fusion (MEF)

Multi-exposure fusion aims to merge images taken at different exposures to produce HDR-like results while preserving both informative content and visual realism. Classical MEF pipelines follow decomposition → activity measurement (exposure, contrast, saturation) → fusion, but struggle to guarantee natural color and consistent appearance across scenes [11]. With deep learning, unsupervised and perceptual strategies emerged to handle the lack of ground truth: early unsupervised networks (e.g., DeepFuse variants) optimized MEF-specific non-reference losses such as MEF-SSIM to preserve structure, but often produced desaturated or pale color outputs because many pipelines fuse only luminance channels [12]. Han et al. (2022) addressed this by introducing DPE-MEF, a deep perceptual enhancement network composed of a **detail enhancement** module (to mine structural and informative content) and a **color enhancement** module (to learn color–brightness relationships), thus balancing informativeness and visual realism. DPE-MEF reported superior perceptual quality and competitive speed on modern GPUs, while highlighting the broader issues: limited real ground truth for HDR fusion, and reliance on synthetic or pseudo-ground truth for supervised training [13]. Other MEF works explored patch-wise structural decomposition, gradient-domain fusion to avoid ghosting, and GAN-based approaches (MEF-GAN) to improve realism; yet color fidelity and artifact suppression remain active challenges [11], [12], [13].

# 4. Thermal (IR) + RGB fusion

Fusing thermal and RGB imagery harnesses complementary strengths: thermal is robust under poor lighting and can reveal heat-based features (useful in inspection and surveillance), while RGB provides high spatial resolution and color detail. Robust cross-modal matching and registration are prerequisites: Jiang et al. (2021) surveyed multimodal image matching and emphasized the importance of descriptors and deep correspondence models for reliable alignment across modalities [1]. Practical sensor studies — e.g., Alexander & Lunderman's

FLIR One Pro reliability study (2021) and application reports on image-based monitoring of infrastructure — underscore that sensor characteristics (resolution, dynamic range, calibration) and environmental factors strongly affect fusion performance [14], [15]. In algorithmic development, RGBT tracking surveys (Feng & Su, 2024) and deep multimodal biomedical surveys (Duan et al., 2024) show that adaptive fusion strategies (attention/gating, three-stream networks, relation

propagation modules) often outperform naive concatenation: gating or attention mechanisms learn to weight modalities by reliability (e.g., thermal at night, RGB by day) and reduce cross-modal contamination [4], [5], [16]. Still, thermal-RGB fusion faces persistent challenges: modality gap in appearance/statistics, sensor misalignment (parallax, differing FOV), resolution mismatch, limited labeled datasets across real-world conditions, and computational constraints for real-time deployment. Domain-specific works (applications in cultural heritage inspection, structural nondestructive testing, and robotics) demonstrate gains but also reveal that calibration and pre-processing (denoising, radiometric correction, geometric alignment) are essential to practical success [18], [19].

# Chapter 3: Strategic Analysis and Problem Definition

## 3.1 SWOT Analysis

A SWOT analysis was conducted to evaluate the strengths, weaknesses, opportunities, and threats associated with this project on image fusion.

**Strengths:**

- Ability to integrate complementary information for improved visualization.
- Wide applicability in healthcare, surveillance, autonomous vehicles, and remote sensing.
- Systematic study of three distinct applications (multi-focus, multi-exposure, thermal-RGB).
- Combination of traditional and modern (ML/DL) approaches allows for comprehensive evaluation.

**Weaknesses:**

- Image registration errors may degrade fusion quality.
- Traditional methods often fail to adapt to non-linear variations.
- Deep learning approaches require large training datasets and significant computational resources.
- Implementation complexity may increase with multi-modal datasets.

**Opportunities:**

- Growing demand for real-time, reliable fusion in AI-driven systems.
- Potential to extend framework into hardware for real-world deployment.
- Contribution to ongoing research in medical imaging, security, and autonomous navigation.
- Possibility of optimizing algorithms for cross-application adaptability.

**Threats:**

- Availability of suitable datasets for all three applications may be limited.
- Risk of overfitting when using deep learning models on small datasets.

- Competing technologies such as hyperspectral imaging may reduce dependency on fusion.
- High computational requirements may pose challenges for real-time applications.

## 3.2 Project Plan - GANTT Chart

The project is planned in phases, each focusing on progressive development. A GANTT chart is used to visualize task scheduling and dependencies.

| Phase | Activity | Duration | Timeline |
|---|---|---|---|
| Phase1 | Problem understanding & Literature survey | 10 weeks | Week 1-10 |
| Phase 2 | Dataset collection & preprocessing methods | 3 weeks | Week 2-5 |
| Phase 3 | Implementation of traditional fusion techniques (multi-focus, multi-exposure, thermal-RGB) | 3 weeks | Week 6-8 |
| Phase 4 | Post-processing & performance evaluation (metrics, comparison with benchmarks) | 5 weeks | Week 8-13 |
| Phase 5 | Implementation of ML/DL approaches | 5 weeks | Week 9-13 |
| Phase 6 | Comparative analysis & unified framework design | 2 weeks | Week 12-13 |
| Phase 7 | Documentation, final report & presentation | 2 weeks | Week 14–15 |

## 3.3 Refinement of Problem Statement

Initial understanding of the problem identified image fusion as a critical tool for combining multiple input images to generate an output with richer information content. However, different application domains (multi-focus, multi-exposure, thermal-RGB) pose distinct challenges such as focus-level detection, illumination balancing, and multi-modal data alignment.

The refined problem statement is as follows:

***"To design, implement, and evaluate a unified framework for image fusion that can adapt to multiple application domains — specifically multi-focus, multi-exposure, and thermal-RGB fusion — by systematically comparing traditional, machine learning, and deep learning approaches, while employing preprocessing and postprocessing techniques to enhance performance and robustness.***

# Chapter 4: Methodology

## 4.1 Description of the approach

The methodology for this project is structured to systematically implement, evaluate, and optimize image fusion techniques across multiple applications. The approach is divided into three main stages:

1. **Preprocessing:**
   - Image registration to align input images for multi-focus, multi-exposure, and thermal-RGB datasets.
   - Noise reduction using filters (e.g., Gaussian, median) to improve fusion quality.
   - Intensity normalization and contrast adjustment to handle illumination variations.
2. **Fusion Techniques:**

   **Traditional Methods:**
   - Multi-scale approaches such as Laplacian pyramid and wavelet transform.
   - Pixel-level and feature-level fusion using gradient, edge, and morphological operations.
   - Weighted averaging and principal component analysis (PCA)-based fusion.

   **Machine Learning and Deep Learning Methods (future phases):**
   - Supervised learning techniques (SVM, Random Forest) for feature selection and fusion decisions.
   - Convolutional Neural Networks (CNN) and advanced deep learning architectures for end-to-end fusion.
3. **Postprocessing and Evaluation:**
   - Contrast enhancement, sharpening, and color correction to improve visual quality.
   - Quantitative evaluation using metrics such as Feature Mutual Information (FMI), Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), and entropy.
   - Comparative analysis to determine the best-performing method for each application domain.

### 4.1.1 Preprocessing Pipeline

- Resizing images to the same dimensions

- Grayscale conversion (if needed)
- Noise removal (Gaussian smoothing, median filtering)
- Histogram equalization or normalization

**4.1.2 Fusion Algorithm**

**Multi-Focus Image Fusion**

**Description of the Algorithm**

The algorithm fuses two or more partially focused images of the same scene into a single all-in-focus image using morphological focus detection, trimap generation, alpha matting, and weighted fusion. This approach ensures that only the sharp, in-focus regions from each source image are retained.

1. **Input Handling**

   - Reads a set of multi-focus grayscale (or converted-to-grayscale) images of the same scene.
   - Ensures all input images are of the same size (resized if necessary).

2. **Focus Measure Calculation (Morphological Filtering)**

   - A focus detection measure is applied using morphological filters (e.g., gradient, Laplacian, or variance operators).
   - Each pixel is assigned a sharpness score, indicating whether it belongs to a focused or defocused region.

3. **Trimap Generation**

   - Based on the focus measure, a trimap is created with three regions:
   - Foreground → definitely in-focus.
   - Background → definitely out-of-focus.
   - Unknown region → uncertain focus.

4. **Alpha Estimation (Matting)**

   - Alpha values (weights between 0 and 1) are estimated for pixels in the unknown region using an image matting algorithm.
   - This produces a smooth transition mask, reducing boundary artifacts between focused and unfocused regions.

5. **Weighted Fusion**

   - Using the alpha matte, the in-focus regions from different images are blended together.
   - Sharp areas are preserved from each input image, while blurred regions are suppressed.

6. Final All-in-Focus Image Generation

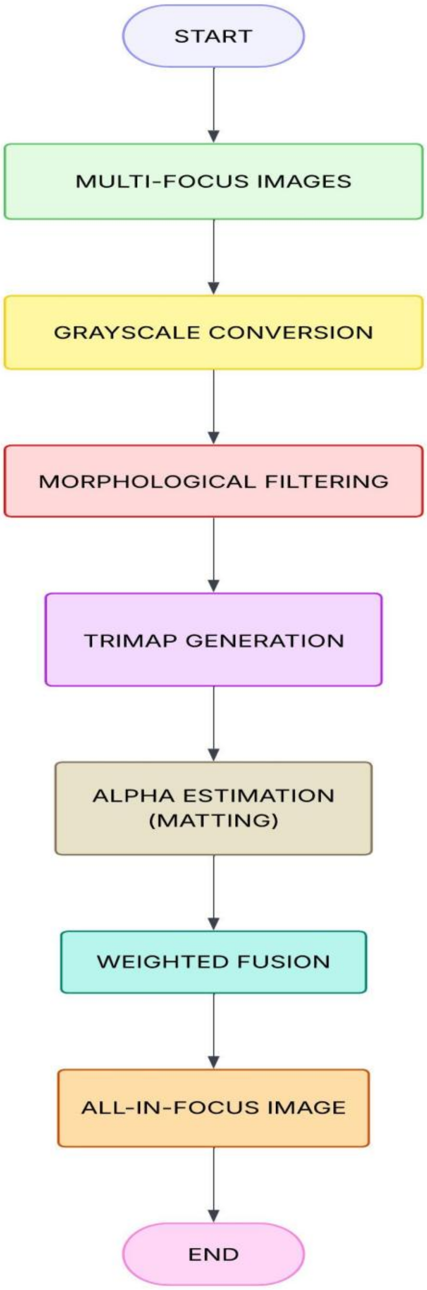
   - The fused image is normalized to 8-bit (uint8).

- The final all-in-focus image is saved with a suffix such as _fused.png.

**Advantages**

- Preserves sharp details from all source images.
- Eliminates blur and depth-of-field limitations of optical lenses.
- Reduces blocking artifacts by using trimap + alpha matting instead of hard decision maps.
- Smooth blending at focus boundaries improves visual quality.
- Effective for digital photography, optical microscopy, and surveillance applications.

**Flowchart :-**



**Multi-Exposure Image Fusion:**

Description of the Algorithm

The algorithm performs multi-exposure image fusion using guided filtering to combine multiple input images with different exposurSe levels into a single, visually balanced fused image.

Step-by-Step Workflow

1. Input Handling

Reads all the preprocessed input images (under-exposed, normally exposed, and over-exposed).
Ensures that all input images are the same size (resized if required).

Converts images into grayscale versions for computing weight maps, while still keeping the RGB channels for the final fusion.

2. Weight Map Generation

Three weight maps are computed for each input image:
Well-exposedness map → gives higher weight to pixels closer to mid-intensity, ensuring proper brightness.
Contrast map → highlights regions with stronger edges and texture.
Saturation map → emphasizes colorful and visually rich regions.
These maps are normalized and combined to generate the final weight map for each input image.

3. Multi-scale Decompositionw

Each input image is decomposed into base layers (low-frequency structure) and detail layers (high-frequency texture).
Similarly, the weight maps are refined using guided filtering to avoid artifacts and halo effects.

4. Fusion Strategy

Base layers are fused by weighted averaging, ensuring a smooth and balanced illumination.
Detail layers are fused by taking maximum contributions from high-frequency details, preserving sharpness and texture.

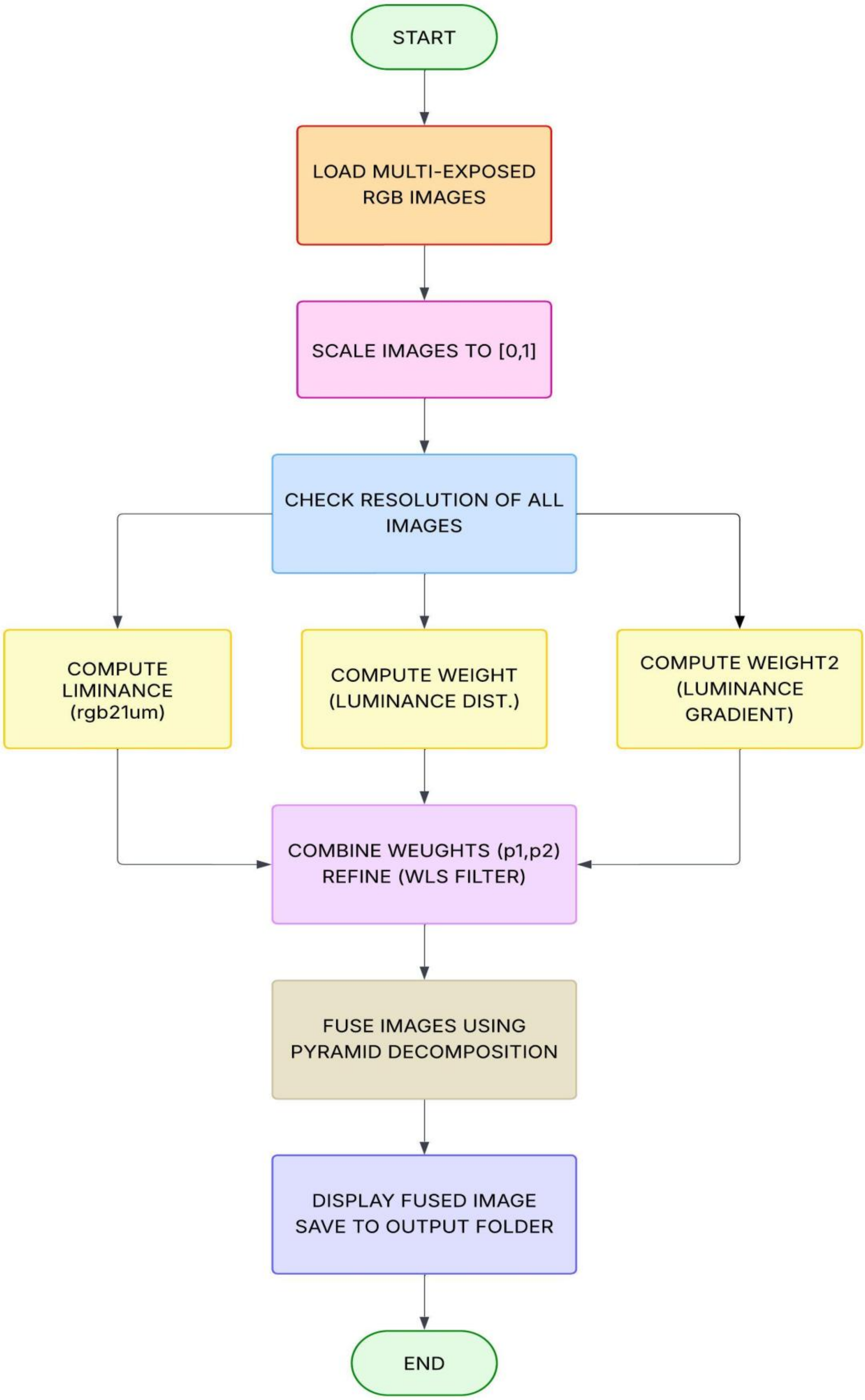
5. Reconstruction

The fused base and fused detail layers are recombined to form the final fused image.
The output image has well-balanced brightness, rich colors, and preserved details from all input exposures.

6. Output Handling

The fused image is normalized and converted into a standard 8-bit format. Finally, it is saved in the specified folder with a unique file name.

Flowchart;-



**Thermal and RGB Fusion**

**Description of the Algorithm**

The algorithm fuses visible RGB images (for detailed textures and colors) and thermal infrared images (for heat signature and night vision) using **2D Discrete Wavelet Transform (DWT)**.

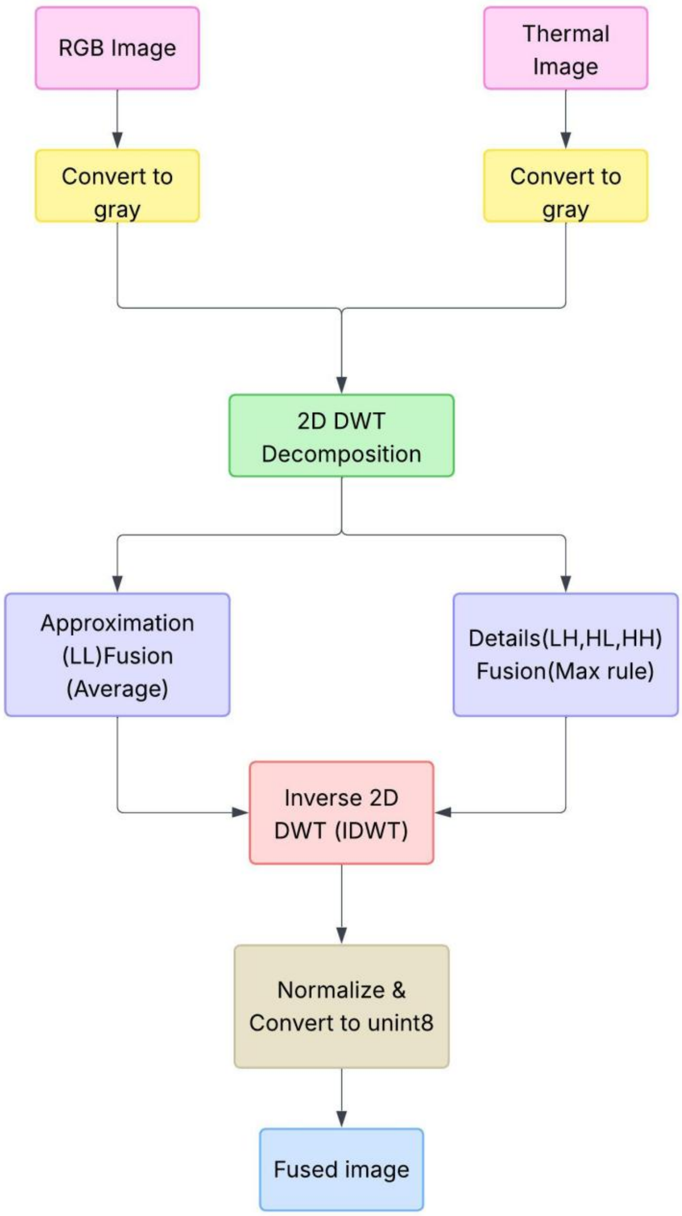
**Step-by-Step Workflow**

1. **Input Handling**
   - Reads all preprocessed RGB and their corresponding thermal images.
   - Ensures both are the same size (resized if needed).
   - Converts thermal images to grayscale if they are RGB.
2. **Wavelet Decomposition**
   - Both RGB (converted to grayscale) and thermal images are decomposed using **DWT with Daubechies-2 (db2)**.
   - This produces **four subbands** for each image:
     - Approximation coefficients (**LL**) → low-frequency (overall image structure).
     - Horizontal detail (**LH**).
     - Vertical detail (**HL**).
     - Diagonal detail (**HH**).
3. **Fusion Strategy**
   - **Approximation (LL)**: averaged between RGB & thermal ((a1+a2)/2) → balances structure.
   - **Detail coefficients (LH, HL, HH)**: maximum selection → keeps the strongest edge/feature.
4. **Inverse DWT**
   - The fused coefficients are recombined using **inverse DWT** to form the final fused image.
5. **Normalization & Saving**
   - The fused image is normalized to 8-bit (uint8).
   - The final fused image is saved with the suffix _fused.png.

## Advantages

- **Preserves structural details** (edges) from both modalities.
- **Retains thermal information** (good for night/low-light).
- **Enhances visual interpretability** by combining RGB detail with thermal contrast.
- Simple, computationally efficient, and effective for autonomous vehicle perception, surveillance, and medical imaging.

Flow Chart:


RGB Image
Thermal Image
Convert to gray
Convert to gray
2D DWT Decomposition
Approximation (LL)Fusion (Average)
Details(LH,HL,HH) Fusion(Max rule)
Inverse 2D DWT (IDWT)
Normalize & Convert to unint8
Fused image

:

### 4.1.3 Postprocessing Pipeline

- Steps applied after fusion to enhance results, e.g.:
- Contrast enhancement
- Sharpening
- Edge enhancement or denoising
- Color adjustment (for color fusion)

## 4.2 Tools and techniques utilized

 **Software Tools:**

- MATLAB for implementing traditional and deep learning fusion algorithms.
- Python (with libraries such as OpenCV, NumPy, TensorFlow/PyTorch) for advanced ML/DL techniques.

 **Techniques:**

- Image preprocessing: filtering, normalization, registration.
- Feature extraction: gradient, edge, DCT, wavelet.
- Fusion methods: pixel-level, multi-scale, ML/DL-based fusion.
- Evaluation: quantitative metrics (FMI, SSIM, PSNR, entropy) and visual inspection.

## 4.3 Design considerations

- Robustness**:** The system should handle variations in image modality (grayscale, color, thermal).
- Scalability**:** Algorithms should scale with image resolution and number of input images.
- Adaptability**:** The fusion framework must support multiple applications without extensive reconfiguration.
- Efficiency**:** Optimization for runtime performance, especially for high-resolution images.
- Quality**:** Preserving important features such as edges, texture, and contrast while minimizing artifacts.
- Extensibility**:** Framework design should allow future integration of advanced ML/DL methods.

# Chapter 5: Implementation

## 5.1 Description of how the project was executed

## 5.2 Challenges faced and solutions implemented

# Chapter 6: Results

## 6.1 outcomes

## 6.2 Interpretation of results

## 6.3 Comparison with existing literature or technologies

# Chapter 7: Conclusion

Here write Suggestions for further research or development and Potential improvements or extensions

# Chapter 8 : Future Work

Here write Suggestions for further research or development
Potential improvements or extensions

# References

1. Jiang, X., Ma, J., Xiao, G., Shao, Z., & Guo, X. (2021). *A review of multimodal image matching: Methods and applications.* Information Fusion, pp. 22–71.
2. Fox, T., Elder, E., & Crocker, I. (2008). *Image Registration and Fusion Techniques.* In PET-CT in Radiotherapy Treatment Planning, pp. 35–51.
3. El-Gamal, F. E.-Z. A., Elmogy, M., & Atwan, A. (2016). *Current trends in medical image registration and fusion.* Egyptian Informatics Journal, 17(2), 99–124.
4. Feng, M., & Su, J. (2024). *RGBT tracking: A comprehensive review.* Information Fusion, 102492.
5. Duan, J., Xiong, J., Li, Y., & Ding, W. (2024). *Deep learning based multimodal biomedical data fusion: An overview and comparative review.* Information Fusion, 102536.
6. Singh, O. P., K. N. Singh, A. K. Singh, A. K. Agrawal, & H. Zhou (2024). *Robust watermarking algorithm based on multimodal medical image fusion.* In Data Fusion Techniques and Applications for Smart Healthcare, pp. 77–107.
7. Piella, G. (2003). *A general framework for multiresolution image fusion: from pixels to regions.* Information Fusion, 4(4), 259–280.
8. Achim, A., Łoza, A., Bull, D., & Canagarajah, N. (2008). *Statistical modelling for wavelet-domain image fusion.* In Image Fusion, pp. 119–138.
9. Zhang, Q., Liu, Y., Blum, R. S., Han, J., & Tao, D. (2018). *Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review.* Information Fusion, 40, 57–75.
10. Liu, Y., Liu, S., & Wang, Z. (2015). *Multi-focus image fusion with dense SIFT.* Information Fusion, 23, 139–155.
11. Aymaz, S., & Köse, C. (2019). *A novel image decomposition-based hybrid technique with super-resolution method for multi-focus image fusion.* Information Fusion, 45, 113–127.
12. Goshtasby, A. A. (2005). *Fusion of multi-exposure images.* Image and Vision Computing. (Classical MEF overview — cited across MEF surveys.)
13. Prabhakar, K. R., et al. (DeepFuse and related unsupervised MEF works — discussed in surveys and DPE-MEF paper).
14. Han, D., Li, L., Guo, X., & Ma, J. (2022). *Multi-exposure image fusion via deep perceptual enhancement (DPE-MEF).* Information Fusion, 79, 248–262.
15. Alexander, Q. G., & Lunderman, C. V. (2021). *Thermal camera reliability study: FLIR One Pro.* US Army Engineer Research and Development Center.
16. Fluke. (2021). *What does infrared mean?* (practical thermal imaging primer).
17. Feng, M., & Su, J. (2024). *RGBT tracking: A comprehensive review.* Information Fusion, 102492. *(duplicate of [4], included for emphasis on IR+RGB tracking literature)*
18. Avdelidis, N. P., & Moropoulou, A. (2004). *Applications of infrared thermography for the investigation of historic structures.* Journal of Cultural Heritage, 5(1), 119–127.
19. Hess, M., Vanoni, D., Petrovic, V., & Kuester, F. (2015). *High-resolution thermal imaging methodology for non-destructive evaluation of historic structures.* Infrared Physics and Technology, 73, 219–225.
20. Kaur, J., & Shekhar, C. (2020). *Multimodal medical image fusion using deep learning.* In Advances in Computational Techniques for Biomedical Image Analysis, pp. 35–56.

Bengaluru City Office No 5/1, First Floor, Prestige Terraces, Union Street, Infantry Road, Bengaluru - 560003 Karnataka, India

GIMSR
GITAM Medical College Rd, Gandhi Nagar, Rushikonda, Visakhapatnam - 530045 Andhra Pradesh, India

GITAM
DEEMED TO BE UNIVERSITY


Bengaluru