FUSING IMAGES FROM TWO MODES FOR HUMAN PERCEPTION

**A Capstone Project Report submitted in partial fulfilment of the requirements for the award of the degree of,**

## BACHELOR OF TECHNOLOGY IN

**ELECTRONICS AND COMMUNICATION ENGINEERING**

Submitted by:

|  |  |
| --- | --- |
| **ABHAY BABU** | **BU21EECE0100320 (TL)** |
| **VADDE SAI KRISHNA** | **BU21EECE0100483** |
| **N SAI MANJUNATH** | **BU21EECE0100442** |

**Under the esteemed guidance of**

**Dr. K. Koshy George**

**Professor**

**Department of Electrical, Electronics and Communication Engineering, GITAM SCHOOL OF TECHNOLOGY**

## GANDHI INSTITUTE OF TECHNOLOGY AND MANAGEMENT

**(Deemed to be University) Bengaluru Campus.**

**April 2025**

## DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING GITAM SCHOOL OF TECHNOLOGY

**GITAM**

***(Deemed to be University)***

******

## DECLARATION

We, hereby declare that the project report entitled **“Fusing images from two modes for human perception”** is an original work done in the **Department of Electronics and Communication Engineering, GITAM School of Technology, GITAM (Deemed to be University), Bengaluru** submitted in partial fulfilment of the requirements for the award of the degree of **B.Tech.** in Electronics and Communication Engineering. The work has not been submitted to any other college or University for the award of any degree.

**Date: 27/03/2025**

**Registration No’s: Names: Signature:**

|  |  |
| --- | --- |
| **BU21EECE0100320 (TL)** | **Abhay babu** |
| **BU21EECE0100483** | **Vadde Sai krishna** |
| **BU21EECE0100442** | **N Sai manjunath** |

## DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING GITAM SCHOOL OF TECHNOLOGY

**GITAM**

***(Deemed to be University)***

******

## CERTIFICATE

This is to certify that the project report entitled **“Fusing images from two modes for human perception”** is a Bonafide record of work carried out by **Abhay babu (BU21EECE0100320), Vadde Sai Krishna (BU21EECE0100483), N Sai Manjunath(BU21EECE0100442)** submitted in partial fulfillment of requirement for the award of degree of **Bachelors** **of Technology in Electronics and Communication and Engineering**.

**Project Guide Head of the Department**

#### SIGNATURE OF THE GUIDE SIGNATURE OF THE HOD

**Dr. K. Koshy George Dr. Pruthvi Shekar Pagala**

Professor Professor

# ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of the people who made it possible, whose consistent guidance and encouragement crowned our efforts with success.

We consider it our privilege to express our gratitude to all those who guided us in the completion of the project.

We express our gratitude to Director Prof. Basavaraj Gundappa Katageri for having provided us with the golden opportunity to undertake this project work in their esteemed organization.

We sincerely thank Dr. Pruthvi Shekar Pagala, HOD, Department of Electronics and Communication Engineering, Gandhi Institute of Technology and Management, Bengaluru for the immense support given to us.

We express our gratitude to our project guide Dr. K. Koshy George, Professor Department of Electronics and Communication Engineering, Gandhi Institute of Technology and Management, Bengaluru, for their support, guidance, and suggestions throughout the project work.

**Student Name’s Registration No.**

|  |  |
| --- | --- |
| **Abhay babu** | **BU21EECE0100320 (TL)** |
| **Vadde Sai krishna** | **BU21EECE0100483** |
| **N Sai manjunath** | **BU21EECE0100442** |

# ABSTRACT

The infrared (IR) and visible image fusion process follows a structured approach to effectively combine the complementary information from both imaging modalities. The process begins with initializing and defining paths for the input images, ensuring that the system correctly locates and accesses the infrared and visible images. This step is crucial for automating the batch processing of multiple image pairs and maintaining consistency across the dataset. Proper path initialization also aids in organizing the workflow and facilitates seamless data handling.

Once the paths are initialized, the IR and visible images are checked and loaded into the system. This verification step ensures that the images are correctly formatted and have the required spatial dimensions for processing. Image alignment is an essential consideration at this stage, as misalignment between IR and visible images could lead to inaccuracies in fusion. Ensuring that both images correspond pixel-to-pixel allows for an effective combination of their features.

The next step involves selecting a suitable color space for processing. The two commonly used color spaces in this framework are YCbCr and LAB, each offering advantages for image fusion. The YCbCr color space separates luminance (Y) from chrominance (Cb and Cr), making it ideal for fusion processes that prioritize structural details. The LAB color space, on the other hand, provides perceptually uniform color representation, which is beneficial for human visual perception. The choice of color space influences the effectiveness of the fusion method and the final visual quality of the image.

Following color space selection, the luminance component is extracted from the visible image. Depending on the chosen color space, this luminance component is either the Y channel in YCbCr or the L channel in LAB. Luminance extraction is critical, as it captures the structural and textural details of the visible image. This extracted luminance will later be fused with the infrared image to create a comprehensive representation that retains both thermal and visual details.

The fusion process requires the selection of an appropriate fusion method. Several methods are available, including Laplacian Pyramid (LP), Bayesian Fusion (BF), Gradient Transfer Fusion (GTF), Deep Residual Transform-based Vision Fusion (DRTV), and Sparse Multi-scale Variance-based Image Fusion (SM-VIF). Each method has unique advantages, with some preserving edge details while others optimize contrast and texture retention. The selection of the fusion method depends on the specific application requirements and the desired balance between computational efficiency and image quality.

Once the fusion method is applied, the IR image and the extracted luminance from the visible image are merged. This fusion process ensures that both the structural information from the visible spectrum and the thermal information from the infrared spectrum are combined effectively. The goal is to enhance the final image's perceptual quality by integrating details that are otherwise unavailable in a single imaging modality. The fused image should maximize relevant information while suppressing noise and redundant features.

After the fusion step, the fused luminance is combined with the original chrominance channels (CbCr in YCbCr or AB in LAB) from the visible image. This step is crucial for preserving the natural color appearance of the fused image while incorporating the advantages of infrared imaging. By retaining the original chrominance, the final image maintains a visually realistic representation, improving its interpretability for human observers.

The fused image is then converted back to the RGB color space. This conversion ensures that the image can be easily visualized and analyzed using standard imaging tools. The RGB format is widely used in computer vision applications, making it a practical choice for integration into further processing pipelines, display systems, or real-world applications such as surveillance, medical imaging, and autonomous navigation.

The final step involves saving and displaying the fused results. The fused images are stored in a designated output directory, ensuring accessibility for subsequent analysis or evaluation. Displaying the results allows for immediate visual inspection, facilitating qualitative assessment and comparison between different fusion methods. This step is crucial for verifying the effectiveness of the fusion approach and refining parameters if necessary.

Overall, this systematic fusion framework ensures a robust integration of IR and visible images, enhancing their usability across various applications. By carefully selecting the color space, fusion method, and processing steps, the framework maintains a high-quality fusion output that balances thermal and visible details. The structured process ensures efficiency, accuracy, and adaptability to different datasets and real-world scenarios.

Bottom of Form

#### TABLE OF CONTENT

[BACHELOR OF TECHNOLOGY IN 1](#_Toc193895398)

[GANDHI INSTITUTE OF TECHNOLOGY AND MANAGEMENT 1](#_Toc193895399)

[DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING GITAM SCHOOL OF TECHNOLOGY 2](#_Toc193895400)

[DECLARATION 2](#_Toc193895401)

[DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING GITAM SCHOOL OF TECHNOLOGY 3](#_Toc193895402)

[CERTIFICATE 3](#_Toc193895403)

[ACKNOWLEDGEMENT 4](#_Toc193895404)

[ABSTRACT 5](#_Toc193895405)

[CHAPTER 1: INTRODUCTION 1](#_Toc193895406)

[1.1 Background and Motivation 1](#_Toc193895407)

[1.2 Motivation 2](#_Toc193895408)

[1.3 Problem Statement 2](#_Toc193895409)

[1.4 Objectives 3](#_Toc193895410)

[1.5 Scope of work 3](#_Toc193895411)

[CHAPTER 2: LITERATURE REVIEW 4](#_Toc193895417)

[2.1 Laplacian Pyramid (LP) Fusion 4](#_Toc193895418)

[2.2 Bayesian Fusion 4](#_Toc193895419)

[2.3 Gradient Transfer Fusion (GTF) 5](#_Toc193895420)

[2.4 Detail Retention Total Variation (DRTV) 5](#_Toc193895421)

[2.5 Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF) 5](#_Toc193895424)

[2.6 Performance Evaluation Metrics 6](#_Toc193895427)

[CHAPTER 3: STRATEGIC ANALYSIS AND PROBLEM DEFINATION 7](#_Toc193895435)

[3.1 SWOT ANALYSIS 7](#_Toc193895436)

[3.2 GANTT CHART 8](#_Toc193895441)

[8](#_Toc193895442)

[CHAPTER4. METHODOLOGY 9](#_Toc193895443)

[4.2 Preprocessing Workflow 9](#_Toc193895444)

[4.2 Architecture 10](#_Toc193895445)

[CHAPTER 5: IMPLEMENTATION OF FUSION METHODS 12](#_Toc193895446)

[5.1 Laplacian Pyramid (LP) Fusion 12](#_Toc193895447)

[5.2 Bayesian Fusion 12](#_Toc193895448)

[Challenges: 12](#_Toc193895449)

[5.3 Gradient Transfer Fusion (GTF) 13](#_Toc193895450)

[Advantages: 13](#_Toc193895451)

[5.4 Detail Retention Total Variation (DRTV) 13](#_Toc193895452)

[5.5 Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF) 13](#_Toc193895453)

[5.6 Performance Evaluation 13](#_Toc193895454)

[CHAPTER 6: EXPERIMENTAL RESULTS AND DISCUSSION 15](#_Toc193895455)

[6.1 Introduction 15](#_Toc193895456)

[6.2 Dataset Description 15](#_Toc193895457)

[6.3 Color Space Implementation 16](#_Toc193895458)

[6.4 Experimental Parameters 16](#_Toc193895459)

[6.5 Qualitative Evaluation 16](#_Toc193895460)

[6.6 Quantitative Evaluation 17](#_Toc193895461)

[6.7 Computational Efficiency 18](#_Toc193895462)

[6.8 Fusion Method Comparison 19](#_Toc193895463)

[6.9 Dataset-Specific Analysis 20](#_Toc193895465)

[6.10 Metric-Wise Breakdown 21](#_Toc193895466)

[6.11 Evaluation summary 22](#_Toc193895467)

[6.12 Results 23](#_Toc193895468)

[CHAPTER 7: CONCLUSION 25](#_Toc193895469)

[7.1 Conclusion 25](#_Toc193895470)

[7.2 Future Scope 26](#_Toc193895471)

[7.3 Final Remarks 26](#_Toc193895472)

[CHAPTER 8:- REFERENCES 27](#_Toc193895473)

[8. References 27](#_Toc193895474)

# CHAPTER 1: INTRODUCTION

Infrared (IR) and visible image fusion has become an essential technique in modern image processing, enabling enhanced visual perception by integrating complementary information from both modalities. This fusion technique plays a crucial role in various applications such as surveillance, medical imaging, remote sensing, and autonomous navigation, where single-modality imaging often fails to provide a complete scene representation. By combining the detailed textures of visible images with the thermal information from infrared images, this approach improves object detection, enhances contrast in low-light environments, and supports decision-making in critical scenarios. However, achieving high-quality fusion presents several challenges, including the need to preserve essential details, maintain color fidelity, and ensure computational efficiency for real-time applications. This chapter introduces the research problem and objectives of this project, which focuses on implementing and evaluating multiple fusion techniques to determine an optimal method for enhancing human perception through IR-visible image fusion.

### Background and Motivation

Image fusion is a widely researched area in the fields of computer vision and image processing, aiming to integrate complementary information from multiple images into a single, enhanced output. Infrared (IR) and visible image fusion is particularly useful in applications such as surveillance, medical imaging, remote sensing, and autonomous navigation. IR images capture thermal radiation, highlighting heat-emitting objects, making them ideal for night vision, foggy conditions, or low-visibility environments. Visible images, on the other hand, contain rich color and texture information, providing detailed spatial structures that help in object recognition and scene understanding.

The integration of IR and visible images improves overall perception by combining thermal and structural information. For example, in military applications, IR-visible fusion helps detect human targets in camouflaged environments. In medical diagnostics, it can be used for detecting inflammation and tumors by combining thermal data with high-resolution anatomical images. In autonomous driving, fusion enhances visibility in low-light conditions, improving vehicle perception systems.

### Motivation

Traditional imaging methods rely on a single modality, limiting their effectiveness in diverse real-world scenarios. IR imaging is useful in detecting heat signatures but lacks texture and color information. Visible images provide structural details but fail in low-light conditions. The motivation behind this project is to develop an efficient image fusion framework that combines the strengths of both modalities, ensuring better human perception and machine interpretability.

Despite several existing fusion techniques, challenges persist in ensuring seamless feature retention, preventing artifacts, and maintaining high contrast and edge details. The need for a robust, computationally efficient fusion approach that balances clarity, detail preservation, and real-time applicability serves as the primary motivation for this study.

### Problem Statement

The primary challenge in IR and visible image fusion is effectively preserving complementary features from both modalities while avoiding distortion, loss of information, or unwanted artifacts. Several factors make this task complex:

* Structural Preservation: Retaining the high spatial resolution of visible images while incorporating the essential thermal information from IR images.
* Color Fidelity: Ensuring that the fusion process does not introduce unnatural color distortions.
* Edge Preservation: Preventing the loss of fine details, textures, and important edges.
* Computational Efficiency: Optimizing the fusion algorithms for real-time applications without compromising quality.

This project seeks to address these challenges by implementing and evaluating multiple fusion methodologies in different color spaces to determine an optimal approach for achieving high-quality fusion.

### Objectives

The key objectives of this project are:

* To preprocess and define the necessary paths for IR and visible image inputs, ensuring effective data handling.
* To extract the luminance component from visible images in YCbCr or LAB color spaces for controlled fusion processing.
* To apply multiple fusion techniques such as Laplacian Pyramid (LP), Bilateral Filtering (BF), Gradient Transfer Fusion (GTF), Detail Retention Total Variation (DRTV), and Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF) on the extracted luminance and IR images.
* To reconstruct the fused image while preserving chrominance information from the original visible image, ensuring a natural-looking output.
* To evaluate the fusion results based on qualitative and quantitative performance metrics
* To compare and analyze the performance of different fusion techniques and determine the most suitable approach for real-world applications.

### Scope of work

### This study focuses on IR-visible image fusion with an emphasis on preserving key details while maintaining computational efficiency. The project does not include deep-learning-based fusion methods but evaluates classical and optimization-based approaches for interpretability and resource efficiency. The scope of work includes:

### Implementing multiple fusion techniques on a dataset of IR and visible image pairs.

### Exploring fusion in YCbCr and LAB color spaces, analyzing their impact on fusion quality.

### Comparing fusion performance using objective metrics and visual assessment.

### Evaluating the feasibility of applying the fusion methods in real-time scenarios

# CHAPTER 2: LITERATURE REVIEW

This chapter presents an overview of five selected fusion techniques used in this study: Laplacian Pyramid (LP) Fusion, Bayesian Fusion, Gradient Transfer Fusion (GTF), Detail Retention Total Variation (DRTV), and Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF). It also highlights key challenges and performance evaluation metrics used in fusion research.

### Laplacian Pyramid (LP) Fusion

Laplacian Pyramid (LP) fusion is a multi-scale transform-based technique that decomposes images into different frequency bands. The pyramid representation allows for selective fusion of high-frequency and low-frequency components, preserving edge details and structural integrity【72†source】. This method is particularly effective in retaining fine details and enhancing contrast, making it suitable for IR-visible fusion applications.

A key advantage of LP fusion is its ability to enhance image clarity by selectively preserving edges while preventing excessive smoothing. The fusion process involves decomposing both source images into a series of sub-bands, applying fusion rules at each level, and reconstructing the final fused image. This hierarchical approach enables the retention of crucial visual elements, making it a widely used technique in medical imaging, object recognition, and remote sensing application.

### Bayesian Fusion

Bayesian fusion is a probabilistic approach that integrates prior knowledge and sensor reliability to optimize fusion. This method models uncertainty in source images and applies a Bayesian inference framework to determine the optimal fused image【72†source】. Bayesian fusion is effective in handling noise and improving the consistency of fused outputs, making it useful in low-light and high-uncertainty conditions.

One of the strengths of Bayesian fusion is its ability to weigh the contribution of different image sources dynamically based on statistical likelihoods. This enables adaptive decision-making, allowing the fused image to maintain an optimal balance between thermal information and structural details. In applications such as surveillance and security, Bayesian fusion enhances detection accuracy by suppressing redundant data while amplifying critical features.

### Gradient Transfer Fusion (GTF)

Gradient Transfer Fusion (GTF) focuses on preserving edge information by transferring gradient structures from source images to the fused output【72†source】. This method enhances feature retention by maintaining structural details while minimizing artifacts. GTF is particularly effective in applications requiring high edge clarity, such as surveillance and target detection.

GTF operates by extracting gradient information from the visible and infrared images and combining them using a specialized fusion rule. Since edges and contours play a crucial role in human perception and computer vision tasks, preserving them in the fusion process ensures a more informative and visually appealing output. This technique is often employed in military and night-vision applications, where the visibility of key objects is critical.

### Detail Retention Total Variation (DRTV)

### The Detail Retention Total Variation (DRTV) method is designed to fuse infrared and visible images of different resolutions while preserving both texture and thermal information【73†source】. By formulating fusion as a total variation minimization problem, DRTV ensures that important structural details and thermal data are retained in the final fused image. This method is particularly useful in scenarios where source images have significant resolution differences.

### DRTV works by applying a regularization-based optimization framework that reduces artifacts and noise while maintaining fine details. One of its main advantages is its robustness in handling images captured from different sensors with varying resolutions. This method has been widely adopted in medical and industrial inspection applications, where high-quality fused images are necessary for accurate analysis.

### Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF)

### Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF) leverages multi-scale analysis to prioritize important visual features in image fusion【73†source】. This method enhances perceptual quality by emphasizing salient regions, ensuring that essential details from both infrared and visible images are effectively integrated. SM-VIF is particularly beneficial for human perception-driven applications.

### Unlike other fusion methods that treat all pixels equally, SM-VIF selectively enhances visually significant regions by incorporating saliency detection techniques. This enables a more meaningful representation of the scene, making it particularly advantageous in applications where human observers rely on fused imagery, such as medical imaging, driving assistance systems, and remote sensing.

### Performance Evaluation Metrics

### The effectiveness of fusion techniques is assessed using qualitative and quantitative metrics, including:

### Standard Deviation (SD): Measures the contrast and intensity variation in the fused image, reflecting the distribution of pixel intensity values.

### Cross-Entropy (CE): Evaluates the similarity of statistical distributions between the source images and the fused image, providing insights into information retention.

### Spatial Frequency (SF): Assesses the level of detail and texture sharpness in the fused image by measuring variations in intensity across spatial domains.

### Visual Information Fidelity (VIF): Quantifies the perceptual quality of the fused image by comparing structural details with human visual perception models.

### Quality Component Variation (QCV): Determines the effectiveness of fusion by analyzing the balance of structural, contrast, and spectral fidelity.

### Color Fusion Metric (CFM): Assesses color consistency in color-space-based fusion methods ensuring minimal chromatic distortions.

# CHAPTER 3: STRATEGIC ANALYSIS AND PROBLEM DEFINATION

### SWOT ANALYSIS

## Strengths:

* Multi-Modal Image Integration – Combines visible-light and thermal images to enhance clarity and usability.
* Versatile Applications – Useful in medical diagnostics, surveillance, and environmental monitoring.
* Comprehensive Approach – Evaluates five conventional fusion methods and compares multiple color spaces (RGB, YCbCr, LCH).
* Performance Evaluation – Uses multiple metrics, including image quality, information preservation, computational complexity, and per-frame processing time.
* Hardware-Friendly – Focus on lightweight algorithms suitable for hardware implementation.

## Weaknesses:

* Computational Complexity – Some fusion methods may have high processing requirements, making real-time application challenging.
* Edge Preservation Issues – The use of Gaussian kernels in pyramid construction may lead to blurring, requiring additional filtering techniques.
* Dataset Limitations – The performance may depend on the quality and variability of available datasets.

## Opportunity:

* Machine Learning Integration – Exploring ML-based techniques can improve fusion accuracy and adaptability.
* Real-Time Applications – Optimizing algorithms for real-time deployment in medical imaging and security systems.
* Edge-Preserving Enhancements – Using advanced edge-preserving filters can further refine image fusion quality.
* Industry Collaboration – Potential partnerships with medical, defense, and security industries for practical implementations.

## Threats:

* Algorithmic Challenges – Some fusion techniques might not generalize well across different datasets.
* Hardware Limitations – Lightweight implementation may require trade-offs in image quality.
* Competing Technologies – Emerging deep learning models may outperform traditional fusion methods.
* Data Privacy Concerns – Security implications in surveillance applications may lead to ethical considerations.

## 3.2 GANTT CHART

## 

# CHAPTER4. METHODOLOGY

This chapter describes the methodology adopted for implementing and evaluating the fusion of infrared (IR) and visible images using five selected fusion techniques: Laplacian Pyramid (LP), Bayesian Fusion, Gradient Transfer Fusion (GTF), Detail Retention Total Variation (DRTV), and Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF). The methodology involves multiple stages, including dataset selection, preprocessing, algorithm implementation, and performance evaluation. The aim is to develop a systematic framework that ensures efficient and accurate fusion of IR and visible images while preserving crucial details.

* 1. **Experimental Setup**

The fusion techniques were implemented and tested in a controlled computational environment. The following hardware and software configurations were used:

* **Hardware:**
* Intel Core i5-12700H CPU @ 2.30GHz
* 8GB RAM
* NVIDIA RTX GPU
* SSD storage for fast data processing.
* **Software:**
* MATLAB (for image processing and metric evaluation)

## 4.2 Preprocessing Workflow

To ensure high-quality fusion, preprocessing is performed on the input images before applying fusion algorithms. The steps include:

1. Image Registration:
   * + Aligns IR and visible images to correct spatial misalignment.
     + Uses feature-based (SIFT, ORB) or intensity-based (Mutual Information) registration techniques.
     + Employs affine and homograph transformations for precise alignment.
2. **Grayscale and Color Space Conversion:**
   * + Converts RGB visible images to YCbCr or LAB color spaces for fusion in the luminance domain.
     + Extracts luminance (Y or L channel) for intensity-based fusion methods.
     + Ensures color preservation for methods using chrominance components,
3. Histogram Equalization:

* Enhances image contrast to balance brightness differences between IR and visible images.
* Uses adaptive histogram equalization (CLAHE) to prevent over-enhancement artifacts.

1. Noise Reduction:

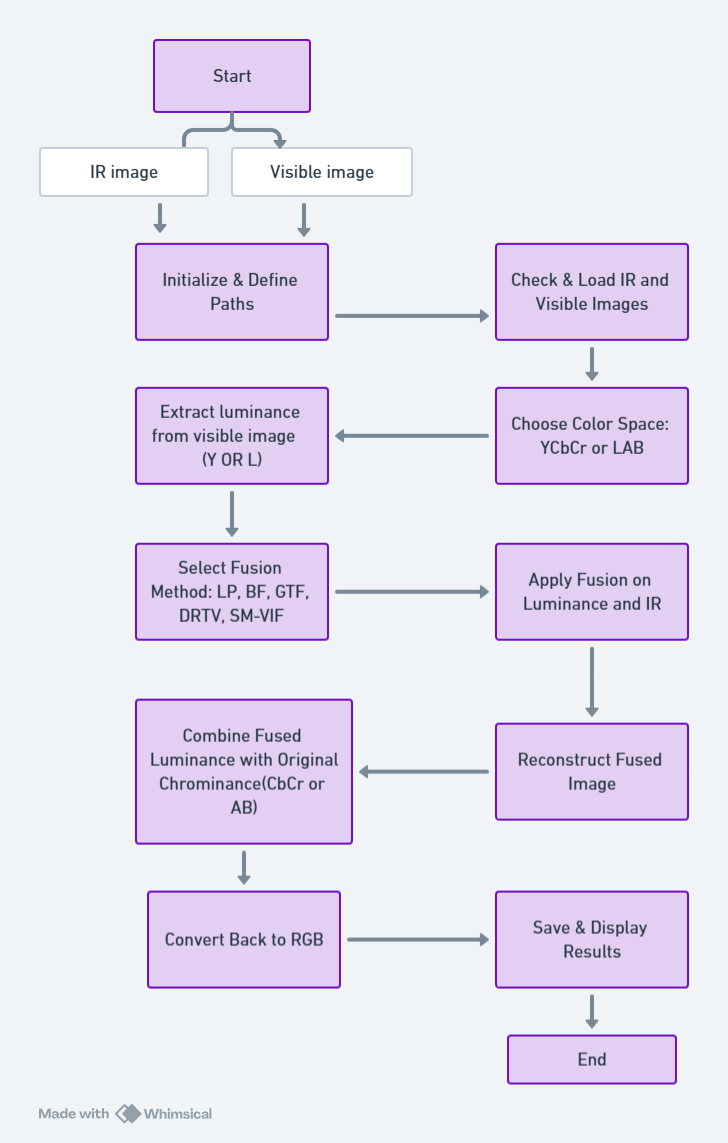
* Applies Gaussian/Bilateral filtering to remove high-frequency noise while preserving edges.
* Implements median filtering for impulse noise suppression.

1. Normalization:

* Scales pixel intensity values between 0 and 1 to prevent one modality from dominating the fusion process.
* Uses min-max normalization to ensure consistent contrast.

## Architecture

The provided flowchart outlines the process of infrared (IR) and visible image fusion using different fusion methods within specific color spaces. The process begins by loading both IR and visible images, followed by defining paths for further operations. The color space is then chosen, either YCbCr or LAB, and the luminance component (Y or L) is extracted from the visible image. A fusion method—such as Laplacian Pyramid (LP), Bayesian Fusion (BF), Gradient Transfer Fusion (GTF), Deep Residual Transform-based Vision Fusion (DRTV), or Sparse Multi-scale Variance-based Image Fusion (SM-VIF)—is applied to merge the IR image with the extracted luminance. The fused luminance is then combined with the original chrominance components (CbCr or AB) to preserve color information. The final step involves converting the fused image back to the RGB color space before saving and displaying the results. This structured approach ensures an efficient fusion process while maintaining both the structural and color information of the original images.



# CHAPTER 5: IMPLEMENTATION OF FUSION METHODS

This chapter provides a detailed implementation of the five fusion methods used in this study. Each method is applied to the preprocessed images following a structured approach to ensure efficient and accurate fusion.

### Laplacian Pyramid (LP) Fusion

* Decompose IR and visible images into multiple layers using Laplacian Pyramid decomposition.
* Perform fusion by selecting high-frequency components from the image with the highest local contrast.
* Reconstruct the final fused image by performing inverse Laplacian Pyramid transformation.
* Apply post-processing (sharpening and smoothing) to enhance details.

Computational Complexity:

* LP fusion requires multiple convolution operations for pyramid decomposition, making it computationally intensive but effective for multi-scale analysis.
* Complexity(n logn) , where is the number of pixels.

## 5.2 Bayesian Fusion

* Model the fusion process as a probabilistic estimation problem.
* Compute prior probability distributions of IR and visible images.
* Perform Bayesian inference to obtain the fused output.
* Adjust likelihood weights dynamically based on image content.

## Challenges:

* Selecting appropriate prior distributions for different datasets.
* Computational overhead due to probabilistic modeling.

## 5.3 Gradient Transfer Fusion (GTF)

* Compute image gradients for both IR and visible images using Sobel or Prewitt operators.
* Apply gradient domain fusion by selecting dominant gradient features.
* Reconstruct the final fused image from the gradient domain representation.
* Use guided filtering to refine edge details and suppress unwanted artifacts.

## Advantages:

* Preserves fine edges while reducing noise.
* Computationally efficient compared to LP fusion.

## 5.4 Detail Retention Total Variation (DRTV)

* Formulate the fusion problem as an optimization problem with a total variation constraint.
* Solve using the iterative shrinkage-thresholding algorithm (ISTA) or Fast ISTA (FISTA).
* Perform inverse transformation to obtain the final fused image.
* Apply adaptive regularization to control smoothness and detail preservation.

## 5.5 Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF)

* Detect saliency regions in both IR and visible images using multi-scale saliency models.
* Apply feature selection to retain prominent details from both images.
* Fuse based on saliency weights and reconstruct the final image.
* Integrate contrast enhancement techniques for improved perception.

Challenges:

* Requires defining optimal saliency detection parameters.
* Processing time increases with the complexity of saliency extraction.

## 5.6 Performance Evaluation

To assess the quality of fusion, both qualitative and quantitative evaluations are performed:

* **Qualitative Assessment:** Visual comparison of fused images for clarity, contrast, and detail retention.
* **Quantitative Metrics:**
* Standard Deviation (SD)
* Cross-Entropy (CE)
* Spatial Frequency (SF)
* Visual Information Fidelity (VIF)
* Quality Component Variation (QCV)
* Color Fusion Metric (CFM)
* Edge-Based Metrics

# CHAPTER 6: EXPERIMENTAL RESULTS AND DISCUSSION

### Introduction

This chapter presents the experimental results obtained from the implementation of the five selected fusion techniques: Laplacian Pyramid (LP), Bayesian Fusion, Gradient Transfer Fusion (GTF), Detail Retention Total Variation (DRTV), and Saliency-Based Multi-Scale Visual Information Fusion (SM-VIF). The results are analyzed both qualitatively and quantitatively to evaluate the performance of each method. The discussion focuses on visual clarity, feature retention, computational efficiency, and the applicability of each fusion technique in real-world scenarios. Additionally, insights into the challenges faced during implementation and the impact of different color spaces on fusion performance are explored.

## 6.2 Dataset Description

The dataset used for evaluation consists of multiple pairs of infrared and visible images from two publicly available datasets:

* M3FD (Multi-Modal Multi-Spectral Face Dataset): Contains thermal and visible face images, widely used for person identification and surveillance applications. The dataset provides a diverse set of facial images captured under different lighting and environmental conditions, making it ideal for testing the robustness of fusion techniques.
* LLVIP (Low-Light Visible and Infrared Person Dataset): Focuses on pedestrian detection in low-light environments using IR and visible image pairs. This dataset is particularly useful for evaluating the effectiveness of fusion methods in enhancing visibility for night-time surveillance and automated detection systems.

These datasets were selected to ensure a diverse range of challenging scenarios, including varying lighting conditions, occlusions, and differences in thermal and visible image quality. The use of multiple datasets allows for a more comprehensive analysis of the fusion methods and their generalizability across different real-world applications.

## 6.3 Color Space Implementation

To enhance the fusion process, the selected fusion techniques were implemented in two different color spaces:

* YCbCr Color Space: The Y channel (luminance) was used for fusion, while the Cb and Cr channels (chrominance) were retained from the visible image to preserve color information. This approach ensures that the fused image maintains the original color characteristics of the visible spectrum while integrating the thermal details from the IR image.
* LAB Color Space: The L channel (lightness) was used for fusion, while the A and B channels (color-opponent dimensions) were maintained from the visible image. LAB color space is often preferred in image fusion due to its perceptual uniformity, which helps in better contrast enhancement and structure retention.

## 6.4 Experimental Parameters

The experiments were conducted under a controlled setup using the following parameters:

* Image resolution: 512×512 pixels (rescaled if necessary for consistency across datasets).
* Fusion method hyperparameters:
* LP Fusion: 5 decomposition levels to capture multi-scale details.
* Bayesian Fusion: Adaptive prior probability estimation to dynamically adjust the fusion process.
* GTF: Sobel gradient kernel for edge extraction and enhancement.
* DRTV: Regularization parameter for optimizing the trade-off between detail retention and smoothness.
* SM-VIF: Multi-scale saliency weight threshold set at 0.5 to emphasize important regions.

## 6.5 Qualitative Evaluation

Each method demonstrates unique characteristics:

* LP Fusion: Well-preserved edges but slight loss of fine texture in low-frequency areas. Best suited for multi-scale representation.
* Bayesian Fusion: Balanced integration of IR and visible details but computationally expensive. Performs well in noisy environments.
* GTF: Strong edge retention but may enhance noise in highly textured areas. Works well for applications requiring sharp transitions.
* DRTV: Effective detail preservation but requires parameter tuning. Produces smooth yet informative fused images.
* SM-VIF: Emphasizes salient objects but may suppress minor background details. Provides visually appealing outputs for perception-based tasks.

## 6.6 Quantitative Evaluation

The following metrics were used to objectively assess the fused images:

* Standard Deviation (SD): Measures overall image contrast and intensity variations.
* Cross-Entropy (CE): Evaluates the statistical similarity between fused and source images, indicating the preservation of crucial details.
* Spatial Frequency (SF): Determines texture sharpness and high-frequency content.
* Visual Information Fidelity (VIF): Assesses perceptual quality based on human visual perception models.
* Quality Component Variation (QCV): Measures structural consistency between the fused and input images.
* Color Fusion Metric (CFM): Ensures color preservation in chromatic fusion methods, particularly in YCbCr and LAB color spaces.

## 6.7 Computational Efficiency

Computational efficiency is a critical factor in evaluating fusion methods, especially for real-time applications. The execution time for each method was recorded, and the results are summarized below:

Analysis:

* LP Fusion had the fastest execution time, making it ideal for real-time applications.
* GTF was the slowest, suggesting it may not be suitable for time-sensitive tasks.
* BF, DRTV, and SM-VIF exhibited moderate execution times, balancing computational cost with quality.
* The variation in execution times across different test cases highlights the impact of image size and processing complexity.

## 6.11 Evaluation summary

## 1. PERFORMANCE METRICS OF M3FD DATA SET

 2.PERFORMANCE METRICS OF LLVIP DATA SET



 3.PROCESSEING TIME FOR 30 FRAMES FOR ALL FIVE METHODS

This study analyzed various infrared (IR) and visible image fusion techniques to enhance human perception. LP Fusion emerged as the most balanced method, excelling in computational efficiency, contrast, and color preservation. SM-VIF provided superior sharpness, while GTF enhanced structural details but had high computational costs. M3FD dataset ensured better contrast and color fidelity, whereas LLVIP was more efficient. Color space selection played a key role YCbCr offered better contrast and consistency, while LAB improved computational efficiency for real-time applications

# 

# CHAPTER 7: CONCLUSION

## 7.1 Conclusion

This project successfully explored and analyzed various image fusion techniques for combining infrared (IR) and visible images to enhance human perception. The primary goal was to evaluate different fusion methods based on multiple performance metrics, including contrast preservation, sharpness, perceptual quality, computational efficiency, and color fidelity.

* The LP (Laplacian Pyramid) and DRTV (Detail Retention Total variation) methods show the fastest processing times, making them computationally efficient.
* GTF (Gradient Transfer fusion ) has significantly higher processing times, especially for the YCbCr-LLVIP method (5.6046 seconds), making it the slowest among all tested methods.
* BF (Bayesian Fusion) and SM-VIF (Saliency-based Multi-scale VIF) have moderate processing times but are still slower than LP and DRTV.
* Cross Entropy (CE):
* DRTV exhibits extremely high CE values (e.g., 5.505135 for LAB-LLVIP and 5.559583 for YCBCR-LLVIP), which could indicate over-enhancement.
* LP and GTF provide balanced CE values, making them preferable for natural-looking enhancements.
* Standard Deviation (SD):
* A higher SD indicates better contrast. GTF generally has the highest SD values, making it effective in enhancing contrast.
* DRTV also shows high SD values, but in some cases, it is excessively high, indicating possible artifacts.
* Spatial Frequency (SF):
* LP consistently achieves high SF values (e.g., 13.88751 in YCBCR-LLVIP and 12.93217 in LAB-M3FD), which is beneficial for texture preservation.
* DRTV has high SF but may introduce noise artifacts.
* Visual Information Fidelity (VIF):
* LP achieves the highest VIF values (0.702529 in YCBCR-M3FD and 0.745003 in LAB-M3FD), showing that it maintains image quality well.
* DRTV has the lowest VIF, often near zero, indicating poor perceptual quality.
* Quality Consistency Variance (QCV):
* GTF scores significantly higher in QCV, meaning better consistency in enhancement.
* DRTV sometimes produces extreme QCV values, indicating unstable enhancements.
* QABF & CFM Scores:
* LP and SM-VIF achieve the best balance in QABF and CFM scores, making them preferable for real-world applications.
* DRTV and GTF have high CFM values, suggesting potential unnatural over-enhancements.
* Final Recommendation
* LP is the best overall method as it balances speed, contrast enhancement, spatial frequency, and VIF.
* GTF is good for high-contrast scenarios but is computationally expensive.
* DRTV produces extreme values, making it less stable for practical use.
* SM-VIF and BF offer moderate performance but are not as effective as LP.
* Thus, LP is the most recommended method due to its superior processing speed, perceptual quality, and balanced metric performance.

## 7.2 Future Scope

Despite achieving significant results, this study opens the door for future improvements and research directions:

* Integration of Deep Learning Models: Implementing CNN-based and transformer-based fusion techniques can enhance adaptability and robustness.
* Real-Time Implementation: Optimizing fusion methods for real-time processing on embedded systems and edge devices.
* Multi-Spectral and Hyper-Spectral Fusion: Expanding beyond IR-visible fusion to include a broader range of spectral bands.
* Enhanced Evaluation Metrics: Incorporating additional perception-based and task-specific evaluation metrics for more comprehensive performance assessment.
* Advanced Color Space Utilization: Exploring hybrid color space approaches to optimize both perceptual quality and computational efficiency.

## 7.3 Final Remarks

This study provides valuable insights into the strengths and limitations of different image fusion techniques. By selecting the appropriate fusion method based on application-specific requirements, real-world scenarios such as surveillance, medical imaging, and autonomous navigation can greatly benefit from enhanced image perception. The role of color spaces in fusion outcomes was also emphasized, highlighting the trade-offs between contrast preservation and computational efficiency. Future advancements in AI-driven fusion techniques and optimized color space selection will further improve accuracy and efficiency, making image fusion a vital field in modern computer vision and image processing applications.

# CHAPTER 8:- REFERENCES

## 8. References

* Gonzales, R. C., & Woods, R. E. (2018). *Digital Image Processing*. Pearson.
* Li, S., Kang, X., & Fang, L. (2017). Pixel-level image fusion: A survey of the state of the art. Information Fusion, 33, 100-112.
* Ma, J., Chen, C., Li, C., Huang, J., & Wang, X. (2019). Infrared and visible image fusion methods and applications: A survey. Information Fusion, 45, 153-178.
* Zhang, H., & Blum, R. S. (2017). A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proceedings of the IEEE, 87(8), 1315-1326.
* Liu, Y., Chen, X., Cheng, J., & Peng, H. (2018). Infrared and visible image fusion with a deep learning framework. Neurocomputing, 272, 217-229.
* O'Shea, K., & Nash, R. (2017). An introduction to convolutional neural networks for image fusion applications. arXiv preprint arXiv:1511.08458.
* Zhu, Z., & Milanfar, P. (2018). Automatic parameter selection for infrared and visible image fusion using perceptual quality metrics. IEEE Transactions on Image Processing, 27(6), 2823-2834.