UNIVERSITY OF KASDI MERBAH OUARGLA

Faculty of New Technologies of Information and Communication

Department of Electronics and Telecommunications

MASTER

Option: Telecommunication

Specialty: Systems of Telecommunication

Presented by:

BELALEM ABDELKARIM

FAKROUNE SALEM

Topic:

Fusion of visible and thermal Solar Panels images

Was Publicly Debated in: june 2024

The Jury:

Dr. A. Benchabane	MCA	President	UKM Ouargla
Dr. S. Souri	MAA	Examiner	UKM Ouargla
Dr. F. Charif	Pr	Supervisor	UKM Ouargla
Dr. H. Dida	Dr	Co-Supervisorr	UKM Ouargla

Aknowledgment

First and foremost, we express boundless gratitude to Allah, the Most Magnificent and Merciful, for guiding us throughout our academic journey. It is certain that without His assistance, this endeavor would not have come to fruition.

We wish to extend our heartfelt appreciation to those individuals who have collaborated with us from the inception to the completion of this research, especially our supervisor Mrs. CHARIF FELLA and Co-Supervisor Mr. DIDA Houdaifa. We also acknowledge the support of Dr. KHALED BENSID, the Head of the Electronics and Communications Department, for his assistance in the successful completion of this work.

We are appreciative of all the professors at the University of Ouargla, particularly those in the electronics and communication department. Additionally, we are grateful to our friends and colleagues for their unwavering support, patience, and valuable friendship.

Lastly, we would like to express our gratitude to our families for their encouragement and unwavering support throughout this journey.

Dedicate

I dedicate this modest work to those who are the source of my inspiration and my courage.

To my dear mother, who always gives me hope to live and who has never stopped praying for me.

To my dear father Abdelkader, for his encouragement and support, And above all for his sacrifice so that nothing will hinder the course of my studies.

To all the professors and teachers who have followed me throughout my schooling and who have allowed me to succeed in my studies.

To my cherished siblings.

***BELALEM ABDELKARIM**

I dedicate this modest work to those who are the source of my inspiration and my courage.

To my beloved mother, whose unwavering belief in me and continuous prayers have been a guiding light.

To my dear father, Ahmed, for his motivation, guidance, and the sacrifices he has made to ensure my education progresses smoothly.

To all the educators who have guided me through my academic journey and contributed to my success.

To my cherished siblings.

***FAKROUNE SALEM**

summary

A study has been conducted on image fusion methods in this field. After that, it was decided to adopt one type of method and follow it to develop an effective image fusion method.

Image merging, which involves combining two or more images from different or similar sources to create a new image with more comprehensive information.

The research specifically focuses on comparing the effectiveness of the convolutional neural networks (CNN) method with other techniques such as LATLRR and NSST when applied to thermal and color image fusion. By evaluating six factors, it is clear that the CNN method achieves the best results among the three methods.

The study also includes a comprehensive discussion of the types of image fusion, levels of processing, and the importance of image fusion in multiple domains, in addition to visual and quantitative evaluations of the performance of different fusion methods

ملخص:

تم إجراء دراسة حول طرق دمج الصور في هذا المجال. بعد ذلك، تم تقرير اعتماد نوع واحد من الأساليب، واتباعه لتطوير طريقة فعالة لدمج الصور.

دمج الصور، الذي يشمل الجمع بين صورتين أو أكثر من مصادر مختلفة أو مماثلة لإنشاء صورة جديدة تحتوي على معلومات أكثر شمولاً.

يركز البحث بشكل خاص على مقارنة فعالية طريقة الشبكات العصبية التلافيفية CNN مع تقنيات أخرى مثل NSST و LATLRR

عند تطبيقها على دمج الصور الحرارية والملونة. من خلال تقييم ستة عوامل، يتضح أن طريقة الشبكات العصبية التلافيفية تحقق أفضل النتائج بين الطرق الثلاثة.

تتضمن الدراسة أيضًا مناقشة شاملة حول أنواع دمج الصور، مستويات المعالجة، وأهمية دمج الصور في مجالات متعددة، إضافة إلى تقييمات مرئية وكمية لأداء طرق الدمج المختلفة.

Contents

List of Figures

List of Tables

List of abbreviations

CNN	Convolutional Neural Networks
NSST	non-subsampled shearlet transform
LATLRR	latent low-rank representation

General Introduction

Image fusion is a process that involves merging multiple images of the same scene or subject to create a single, comprehensive image. This technique is widely used in various fields, including remote sensing, medical imaging, surveillance, and military applications. In the context of thermal imaging, the technique utilizes infrared radiation to capture heat emitted by an object or scene. [1]

Thermal imaging technology offers numerous benefits such as the ability to detect hidden objects, identify potential hazards, and monitor temperature changes. To enhance the quality and information content of thermal images, different image fusion techniques are employed. Two common techniques used for this purpose are CNN (Convolutional Neural Network) and LatLRR (Latent Low-Rank Representation).[2]

Additionally, the Non-Subsampled Shearlet Transform Convolutional Neural Network (NSST-PCNN) technique is also used for image fusion, particularly in infrared applications. This method leverages advanced algorithms to merge infrared images and enhance their quality, providing a more detailed and accurate representation of the scene. [3]

Overall, the combination of thermal imaging and image fusion techniques like CNN, LatLRR, and NSST-CNN in infrared applications offers a powerful tool for various industries, allowing for improved detection capabilities, hazard identification, and temperature monitoring. [4]

This study presents a novel multi-focus image fusion methodology that leverages guided filter-based focus region detection to overcome the limitations associated with spatial domain-based techniques. Initially, Using a guided filter based on focus region detection to generate accurate focus maps . Subsequently, Applying the pixel-wise maximum rule to match accurate focus maps and create an initial decision map the final Fusing the source images with , the final decision map using pixel-wise weighted-averaging technique . [5]

The rest of this study is organized as follows. In Chapter 1, we give information about image fusion and image thermal on general. Chapter 2 chooses work on three image merging techniques and choose the best one that gives a good and targeted result. Finally, in Chapter 3, the experimental results and performance analysis are given.[6]

Chapter I

An Overview of Thermal Solar Panel and Image Fusion

I.1 Introduction

Image fusion combines infrared and visible images to create a more informative and visually enhanced image. The process involves balancing the thermal information from infrared images with the contextual information from visible images. Various techniques and algorithms are used in image fusion, including weighted averaging, principal component analysis, and multi-scale transform methods. The resulting fused image can have numerous applications, including surveillance, remote sensing, medical imaging, and material inspection, and can improve decision-making and situation awareness. Overall , image fusion is a powerful tool for extracting valuable insights from complex scenes.

I.2 Thermal Imaging

Thermal imaging involves using a thermal camera to capture infrared radiation emitted by an object and creating a visual representation of it. This process falls under the category of infrared imaging science. Thermography enables the observation of temperature differences, with warm objects appearing prominently against cooler backgrounds when viewed through a thermal imaging camera. Algorithms can be used to interpret this data and generate an image. [7].

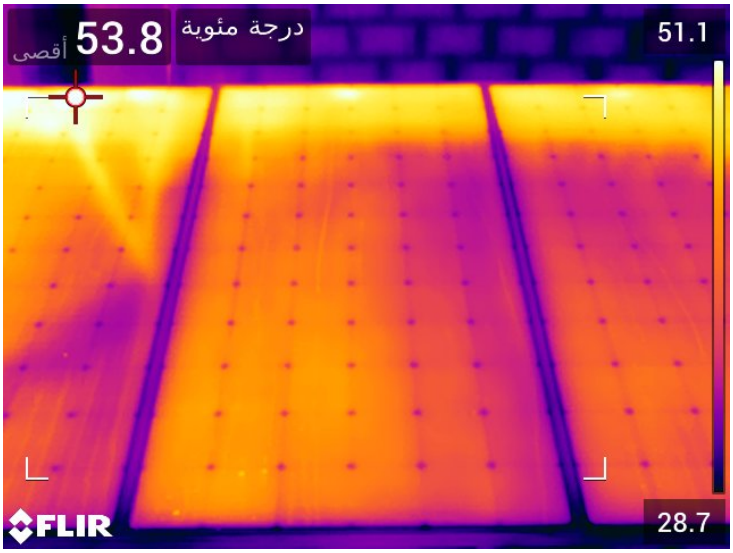


Fig. I.1: Thermal image of a solar panel

I.2.1 Thermal image analysis techniques

Using these features makes it easier to find what you want to highlight in the image. The three most important functions for improving thermal images are thermal framing, isotherms, and plates. [8]

Thermal framing

Thermal framing is the adjustment of the scale of the image to optimize the contrast for the needs of the analysis. An example is shown in Fig. I.2. [9]

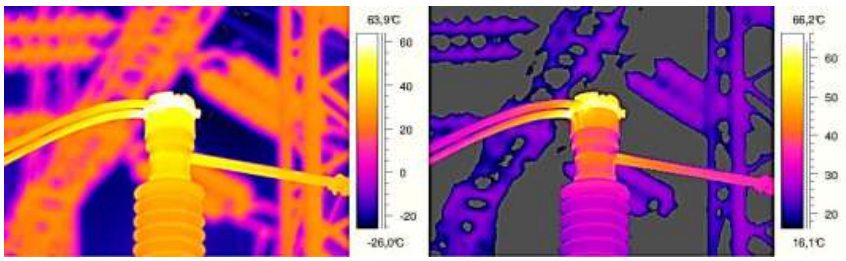


Fig. I.2: image automatically adjusted and thermally framed on the component

Isotherme

The isotherm only replaces one color range with another that contrasts more with the colors used in the image. An example is shown in Fig. I.3. [9]

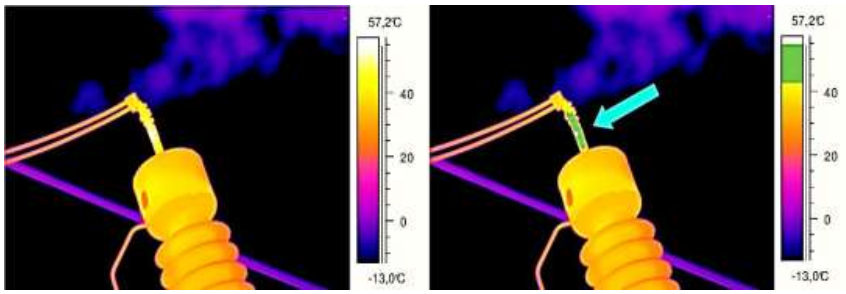


Fig. I.3: High voltage transformer head with an isotherm.

Pallets

A palette assigns different colors to mark defined apparent temperature levels. An example is shown in Fig. I.4 [9]

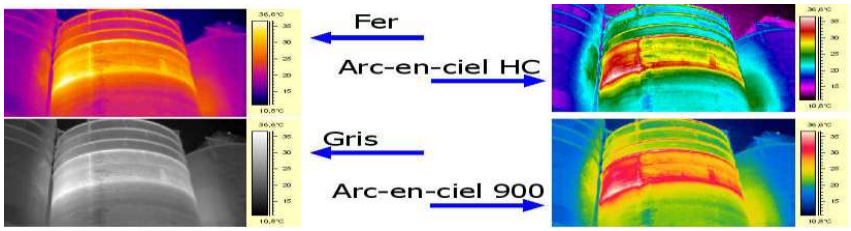


Fig. I.4: Thermal imaging of the storage tank is taken from various platforms

I.2.2 The infrared camera

The infrared camera does not measure temperatures, but radiation, whereas, visualized by the thermograph, the thermal image that the camera provides can be transformed into thermograms, into temperature images. This is exactly what we are looking for to determine the state of health of equipment and above all predict what will happen in the future, in predictive maintenance. [10]

Fig. I.5: One of the used thermal imaging cameras

I.3 Solar Panel

What Is A Solar Panel ? A Solar panels (also known as "PV panels") is a device that converts light from the sun, which is composed of particles of energy called "photons", into electricity that can be used to power electrical loads. [11]

I.3.1 Types Solar Panel

Monocrystalline Solar Panels (Mono-SI)

This type of solar panels (made of monocrystalline silicon) is the purest one. You can easily recognise them from the uniform dark look and the rounded edges.[12]

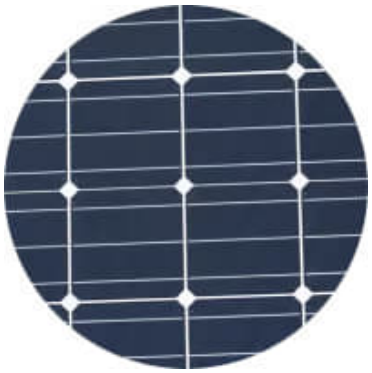


Fig. I.6: Monocrystalline Solar Panels (Mono-SI) .

Polycrystalline Solar Panels (p-SI)

You can quickly distinguish these panels because this type of solar panels has squares, its angles are not cut, and it has a blue, speckled look. which is a faster and cheaper process than that used for monocrystalline panels. [12]

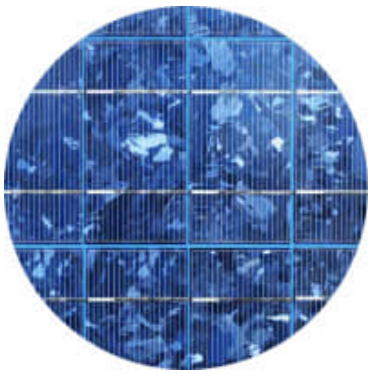


Fig. I.7: Polycrystalline Solar Panels (p-SI) .

Thin-Film Solar Cells (TFSC)

These types of solar panels are the easiest to produce and economies of scale make them cheaper than the alternatives due to less material being needed for its production.[12]

Fig. I.8: Thin-Film Solar Cells (TFSC).

Concentrated PV Cell (CVP and HCVP)

Those multi-junction types of solar panels have an efficiency rate of up to 41 , which among all photovoltaic systems, is the highest sofar. [12]

Fig. I.9: Concentrated PV Cell (CVP and HCVP).

I.3.2 The efficiency of solar panels with its pros and cons

Solar Cell Type	Efficiency	Advantages	Disadvantages
Monocrystalline Solar Panels (Mono-SI)	20%	High efficiency, optimized for commercial use, high lifetime value	Expensive
Polycrystalline Solar Panels (p-Si)	15%	Lower price	Sensitive to high temperatures, lower lifespan, slightly less space efficiency
Thin-Film Amorphous Silicon Solar Panels (A-SI)	7-10%	Relatively low costs, easy to produce, flexible	Shorter lifespan, shorter warranties
Concentrated PV Cell (CVP)	41%	Very high performance, very high efficiency rate	Solar tracker and cooling system needed

Table I.1: A table showing the advantages, disadvantages, and effectiveness of solar panels [13]

I.4 fusion image

Image fusion The process of blending several images to produce an image with superior details required for subjective as well as objective analysis. Normally two or more images captured through various sensors and a composite image is generated having deep knowledge of object and scene.[14] The objects in the resultant image are more clear, more detectable and more recognizable. The goal behind this process is to extract valuable and complementary details from a data set. A well-designed and optimal algorithm is required, which will transfer maximum information from data set to a comoposite image. These images cater to a variety of applications under different domains.[15]

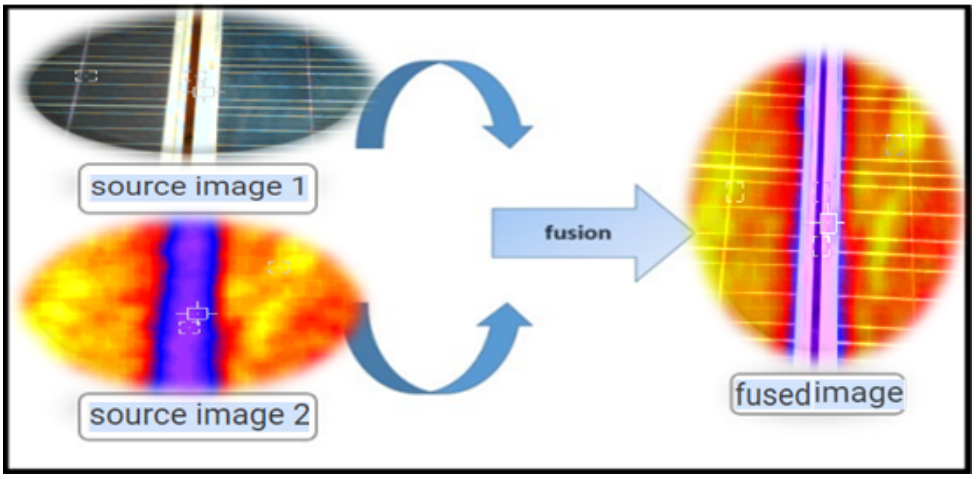


Fig. I.10: General diagram of the image fusion process.

I.4.1 Categories of image fusion

I.4.1.1 Multi-focus image fusion (MFF)

Multi-focus image fusion (MFF) : is a multiple image compression technique using input images with different focus depths to make one output image that preserves all information [16]

Fig. I.11: Multi-focus image fusion (MFF).

I.4.1.2 Multi-exposure image fusion (MEF)

is a integrate images with multiple exposure levels into a full exposure image of high quality. It is an economical and effective way to improve the dynamic range of the imaging system and has broad application prospects. [17]

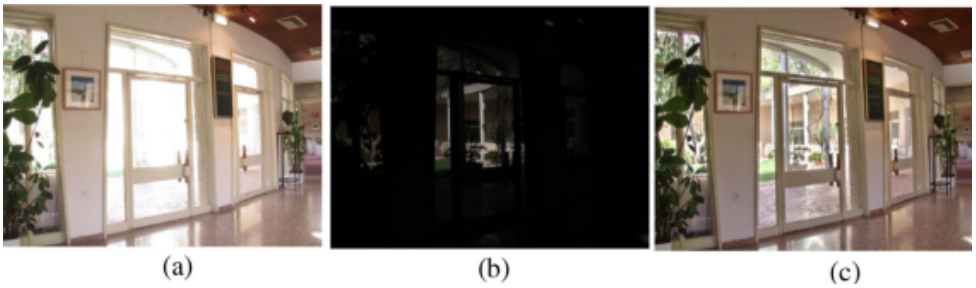


Fig. I.12: Multi-exposure image fusion (MEF).

I.4.1.3 Visible and Infrared (Vis-IR) Image Fusion

Vis-IR image fusion exemplifies multimodal fusion. Various images from visible and IR sensors are fused in this process.[18]

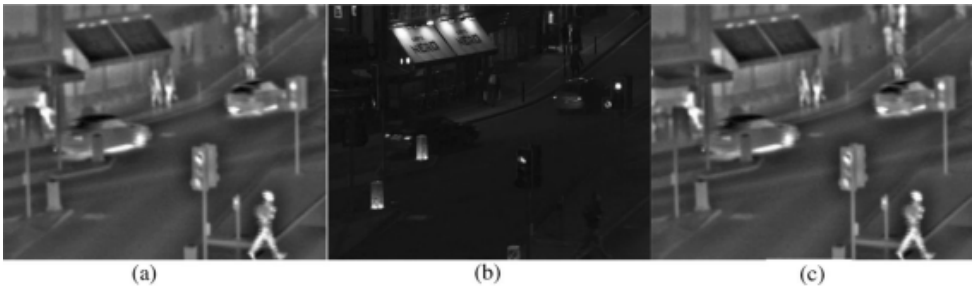


Fig. I.13: Visible and Infrared Image Fusion.png

I.4.1.4 Biomedical Image Fusion

Biomedical Image Fusion : is a combined together like information regarding hard tissues are provided by (CT) and soft tissues specific position is provided by (SPECT) imaging. All these fused thermal images show more depth of field. so in that case, multimodal image fusion is a cheap and more patient-friendly method.[19] [18]

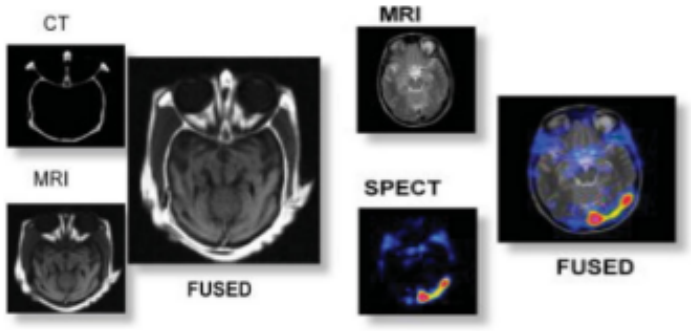


Fig. I.14: Visible and Infrared Image Fusion.png

I.4.2 The level of image fusion (Processing)

The processing of images involves a hierarchical approach, comprising three primary stages: pixel-level processing, feature extraction, and decision-making. A detailed breakdown of each stage is provided below. [20] [21]

I.4.2.1 Pixel-level Fusion

Pixel-level methods combine two images by altering one of the images to maximize its similarity with the other, as measured by a specific criteria. [20]

I.4.2.2 Feature-level Fusion

Feature-level fusion involves identifying and extracting relevant objects or features from different image modalities, often using segmentation techniques. The resulting comparative features or regions from each modality are then combined and analyzed using factual methods to gain a more comprehensive understanding.[20]

I.4.2.3 Decision-level Fusion

Decision-level fusion is a method of information extraction that combines labeled data from independently processed images. By consolidating the data, this approach strengthens basic translation and provides a deeper understanding of the observed objects. The key advantage of this method is that by leveraging higher-level representations, it enables more robust and reliable fusion of multimodal data.[20]

I.4.3 Avantages and inconvenients of the fusion of images

I.4.3.1 Avantages

- it has a high resolution used for multi-scale images.
- Color image fusion is possible.
- Signal capacity is maintained in all domains during image fusion.
- Storage and transmission of data are eliminated with image fusion.

I.4.3.2 Disadvantages

- Not easily visible, primarily due to the camera.
- Visualization may not be effective if source images are not clear, impacting performance.
- Normal risks associated with data manipulationduring image fusion.
- Image fusion may impact the quality of the final donation.[22]

I.5 Evaluation

To verify the effectiveness of a fusion algorithm we need some quantitative and qualitative measurements. Many fusion metrics have been proposed. These metrics work according to the objective of image fusion. [23]

I.5.1 Visual evaluation

Visual analysis is necessary to check the quality of the images obtained by fusion, it allows to identify and identify defects. These defects can affect the quality of the image (high resolution, pixelated appearance, etc.), or the geometry (distortion of line elements, blurring of lines, grouping of objects, etc.). [24]

I.5.2 Quantitative evaluation

In this work, we will use Petrovic statistics to provide more detailed information on the advantages and disadvantages of the fusion method by estimating the informational contribution of each image source such that (Q_G , Q_P , Q_Y , $Q^{XY/F}$, $L^{XY/F}$, $N^{XY/F}$)

I.5.2.1 Edge-Based Metric Q_G

The Q_G gradient-based index evaluates the success of edge information transferred from the source images to the fused image. Q_G is defined as follows: [24]

$$Q_G = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(Q^{AF}(i,j)W^{A}(i,j)+Q^{BF}(i,j)W^{B}(i,j))}{\sum_{i=1}^{m}\sum_{j=1}^{n}(W^{A}(i,j)+W^{B}(i,j))} \tag{I.1}$$

where $Q^{AF}(i,j) = Q_g^{AF}(i,j)$,while $Q_g^{AF}(i,j)$ and $Q_O^{AF}(i,j)$, denote the edge strength and orientation preservation values at pixel (i,j), respectively. $W^A(i,j)$ is the weight coefficient for each edge that shows the importance of that edge.The size of each source image is m x n. $Q^{BF}(i,j)$ and $W^B(i,j)$ are similarly defined. The larger the value of Q_G is, the more edge information will be transferred from the source images to the fused image

I.5.2.2 Metric based on phase congruence Q_P

This measure was proposed by Zhao et al. Based on the image match, this match provides an absolute measure of an image feature to determine an evaluation metric. Q_P consists of three basic correlation coefficients to measure the amount of edge and angle information transferred from the source images to the fused image. Through these parameters, this scale can be defined as follows:[24]

$$Q_P = (P_P)^{\alpha}(P_M)^{\beta}(P_m)^{\gamma} \tag{I.2}$$

Where p, m and n denote the phase congruence, maximum and minimum moments, respectively.

I.5.2.3 Metric based on structural similarity Q_Y

Proposed the Q_Y similarity measure which mainly relies on structural information about the human visual system. This metric is based on SSIM to calculate the similarity ratio between images, as shown in the following equation:[24]

$$Q_Y = \begin{cases} \lambda(\omega)SSIM(A,F\mid\omega)+(1-\lambda(\omega))SSIM(B,F\mid\omega), SSIM(A,B\mid\omega)\geq 0.75 \\ max(SSIM(A,F\mid\omega),SSIM(B,F\mid\omega)), SSIM(A,B\mid\omega)<0.75 \end{cases} \tag{I.3}$$

Or :

$$\lambda = \frac{s(A\mid\omega)}{s(A\mid\omega)+s(B\mid\omega)} \tag{I.4}$$

Here w represents a window size, (w) is a local weight while S(A| (w)) and S(B | (w)) are the variance of source images A and E in window w, respectively. SSIM(A,F) and SSIM(B,F) represent the SSIM map calculated in a sliding manner between A and F. For two source images X and Y, the definition of the structural similarity metric SSIM

measure proposed by Wang et al. is given as follows:

$$SSIM(X,Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_X\sigma_Y + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X\sigma_Y + C_2)} \tag{I.5}$$

I.5.2.4 Total fusion performance $Q^{XY/F}$

Consider two source images X,Y and a merged image F.The total performance of the fusion $Q^{XY/F}$ of the source and merged images of size M × N is calculated by:[24]

$$Q^{XY/F} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M}(Q^{XF}(n,m)W^X(n,m) + Q^{YF}(n,m)W^Y(n,m))}{\sum_{n=1}^{N}\sum_{m=1}^{M}(W^X(n,m) + W^Y(n,m))} \tag{I.6}$$

The total performance of the fusion satisfies:

$$0 \leq\ Q^{XY/F} \leq\ 1 \tag{I.7}$$

* If $Q^{XY/F} = 0$ then this implies a complete loss of source information.
* If $Q^{XY/F} = 1$ then indicates the 'ideal fusion' without loss of source information.

I.5.2.5 Loss of fusion $L^{XY/F}$

It measures the edge information lost in the fusion task.This information is not presented in the merged image but in the source images. The mathematical expression is given by:[24]

$$L^{XY/F} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M} r(n,m)((1 - Q^{XF}(n,m)W^X(n,m) + (1 + Q^{YF}(n,m)W^Y(n,m))}{\sum_{n=1}^{N}\sum_{m=1}^{M}(W^X(n,m) + W^Y(n,m))} \tag{I.8}$$

The fusion loss range is :

$$0 \leq\ L^{XY/F} \leq\ 1 \tag{I.9}$$

* If $L^{XY/F} = 0$ the fusion loss is low.
* If $L^{XY/F} = 1$ the fusion loss is high.

The value of $L^{XY/F}$ should be lower to obtain better performance from any fusion algorithm.

I.5.2.6 Fusion Artifacts $N^{XY/F}$

Unnecessary visual information may be introduced into the combined image that has no relevance to the source images.These artifacts should be avoided. The mathematical expression of $N^{XY/F}$ is given by: [24]

$$N^{XY/F} = \frac{\sum_n\sum_m AM_{n,m}((1 - Q^{XF}(n,m)W^X(n,m) + (1 - Q^{YF}(n,m)W^Y(n,m))}{\sum_n\sum_m(W^X(n,m) + W^Y(n,m))} \tag{I.10}$$

This fusion metric satisfies :

$$0 \leq N^{XY/F} \leq 1 \tag{I.11}$$

* If $N^{XY/F} = 0$ then there is no fusion artifacts.

* If $N^{XY/F} = 1$ then there is a serious degradation in image quality due to noise or artifacts.

It is found that the fusion information score, loss and fusion artifacts are complementary to each other. The sum of these measurements must give the unit:

$$Q^{XY/F} + L^{XY/F} + N^{XY/F} = \mathbf{1}$$

I.6 Conclusion

The significant progress in image fusion research indicates the importance of this research in various real applications. combining multiple image fusion methods is also observed to be successful in this field. there still exist many challenges in image fusion and objective fusion performance evaluation, resulting from image noise, resolution difference between images, and limitations of the imaging hardware. Therefore, it is expected that novel researches on image fusion would continue to grow in the upcoming years. [25] [26]

Chapter II

Some Image Fusion methods

II.1 Introdution

Over the last thirty years, a diverse array of techniques for the fusion of multi-focused images has emerged. These are predominantly categorized into transform domain and spatial domain methods. Within these methodologies, source images are decomposed into various transform coefficients. [27]These are then merged according to defined criteria within the transform domain fusion techniques. Subsequently, the combined image is generated by reconstructing these fused coefficients. This study has reviewed a number of methodologies applied in the domain of image fusion. This particular chapter will outline the techniques utilized in the fusion of multi-focus images. [28]

II.2 Convolutional Neural Network (CNN)

A CNN is composed of multiple building blocks, known as layers, which work together to form the architecture. In this section, we will delve into the details of these building blocks and their roles in the CNN architecture.[29]

CNN is a type of Artificial Neural Network (ANN) that excels at processing spatial data, such as images, with remarkable generalization abilities. Its deep feed-forward architecture enables it to learn highly abstracted features of objects, making it more efficient at identification. A deep CNN model consists of multiple processing layers that learn various features of input data at different levels of abstraction. The initial layers extract high-level features with lower abstraction, while the deeper layers extract low-level features with higher abstraction.[30]

II.2.1 Concepts of Convolutional Neural Network

A Convolutional Neural Network (CNN), also known as ConvNet, is a type of Artificial Neural Network (ANN) with a deep feed-forward architecture. Compared to other networks with FC layers, CNN has exceptional generalization capability and can efficiently learn and identify highly abstracted features of spatial data, particularly objects. A deep CNN model is comprised of a finite number of processing layers that can learn various features of input data, such as images, at multiple levels of abstraction. The initial layers of the model learn and extract high-level features with lower abstraction, while the deeper layers learn and extract low-level features with higher abstraction. The basic conceptual model of CNN is depicted in Figure (II.1), with different types of layers explained in subsequent sections.[30] [31]

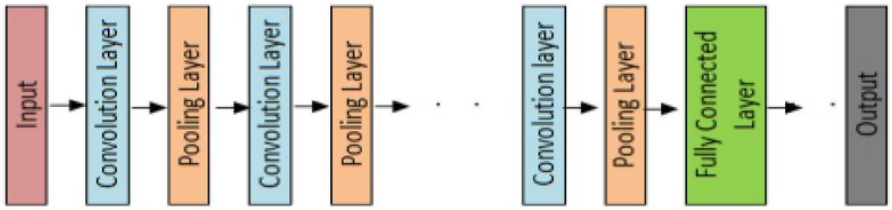


Fig. II.1: Conceptual model of CNN

Why are Convolutional Neural Networks preferred over traditional neural networks in the field of computer vision ?

- One of the main advantages of CNNs is their weight sharing feature, which reduces the number of trainable parameters in the network, helping to prevent overfitting and improve generalization.
- In CNNs, the classification and feature extraction layers work together, resulting in a more organized output that is dependent on the extracted features.
- Implementing a large network is easier with CNNs compared to other types of neural networks. CNNs have become a popular choice for various computer vision applications such as image classification, object detection, face detection, speech recognition, vehicle recognition, facial expression recognition, and text recognition.[32]

CNN has become a popular tool for achieving successful outcomes in a variety of computer vision tasks such as image classification, object detection, face detection, speech recognition, vehicle recognition, facial expression recognition, and text analysis. [32]

Now, a brief overview of the various components or fundamental elements of CNN is provided below.

II.2.2 CNN architecture

CNNs are designed with a specific assumption in mind: that the input data will be images. As a result, the architecture of CNNs is tailored to effectively process and analyze this type of data.[33]

II.2.3 Network Layers

A Convolutional Neural Network (CNN) is composed of multiple building blocks, known as layers, which work together to form the architecture. In this section, we will delve into the details of these building blocks and their roles in the CNN architecture [33]

II.2.3.1 Convolutional Layer

The convolutional layer is a vital element in the architecture of any CNN. It comprises a group of convolutional kernels, or filters, that are used on the input image (depicted as N-dimensional matrices) to create an output feature map.

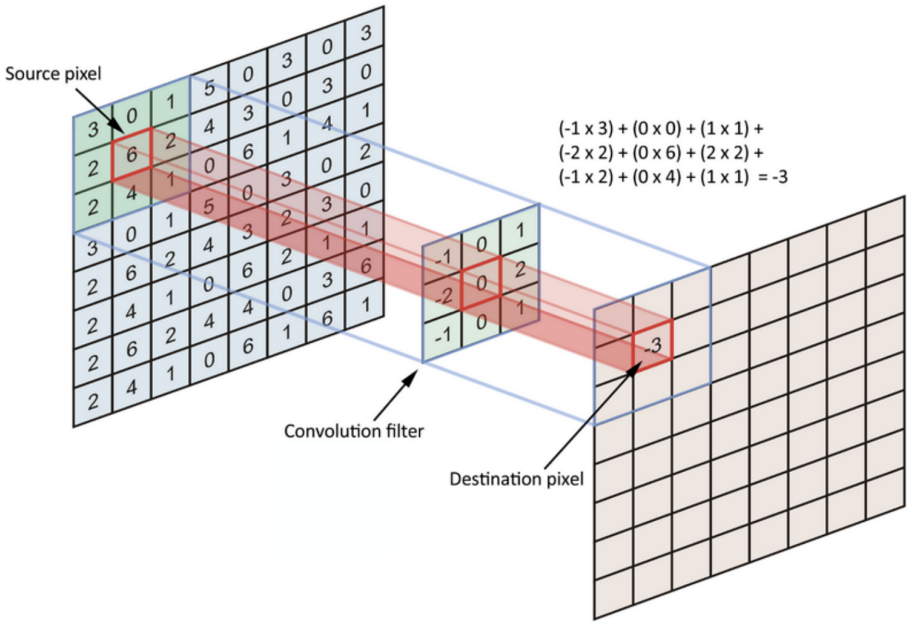


Fig. II.2: convolutional layer in cnn

II.2.3.2 Pooling Layer

Pooling layers in (CNNs) downsample feature maps produced by convolutional operations, reducing their size while preserving the most important features. The pooling process involves specifying the region size and stride, similar to convolution. Various pooling techniques exist, including max pooling, min pooling, average pooling, and more, with max pooling being the most widely used. However, a key limitation of pooling layers is that they can compromise the overall performance of CNNs by focusing on feature presence rather than their exact location in the input image.

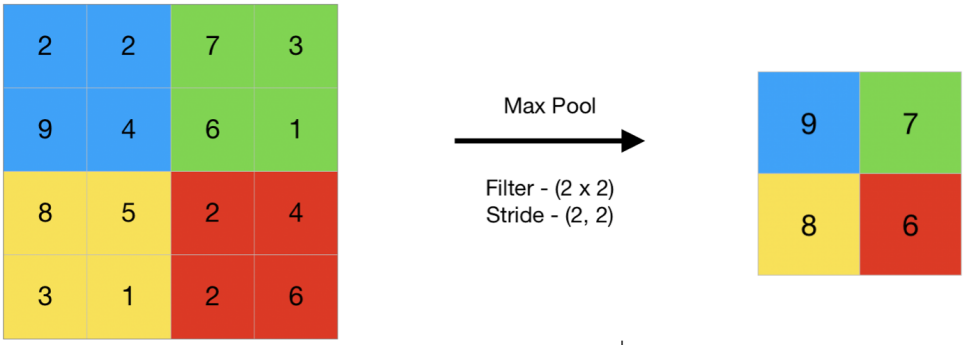


Fig. II.3: pooling layer in cnn

II.2.3.3 Activation Functions (Non-Linearty)

The primary role of an activation function in a neural network is to transform the input into an output, which is calculated by combining the weighted sum of the neuron's inputs with a bias term (if applicable). Essentially, the activation function determines whether a neuron is activated or not for a given input, producing a corresponding output.

In Convolutional Neural Network (CNN) architectures, non-linear activation layers are typically used after each learnable layer (such as convolutional and fully connected layers). These non-linear layers enable the CNN model to learn complex patterns and map inputs to outputs in a non-linear fashion. A crucial property of an activation function is that it must be differentiable, allowing for error backpropagation to train the model effectively. The most popular activation functions used in deep neural networks, including CNNs, are described below.

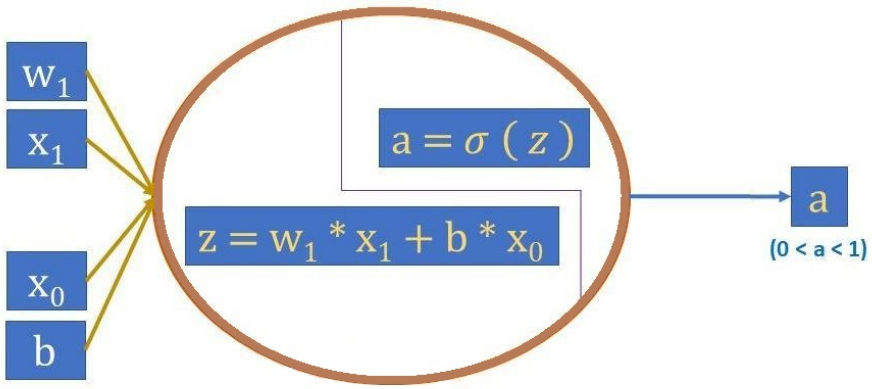


Fig. II.4: activation functions non-linearity in cnn

II.2.3.4 Fully Connected (FC) Layer

In most CNN architectures designed for classification, the final components are fully-connected layers, where each neuron in a layer is linked to every neuron in the previous

layer. The last fully-connected layer serves as the output layer, also known as the classifier, of the CNN architecture.

Fully-connected layers are a type of feed-forward artificial neural network (ANN) that follows the traditional multi-layer perceptron neural network (MLP) principle. They receive input from the last convolutional or pooling layer, which is a set of feature maps. These feature maps are then flattened into a vector, which is fed into the fully-connected layer to produce the final output of the CNN.

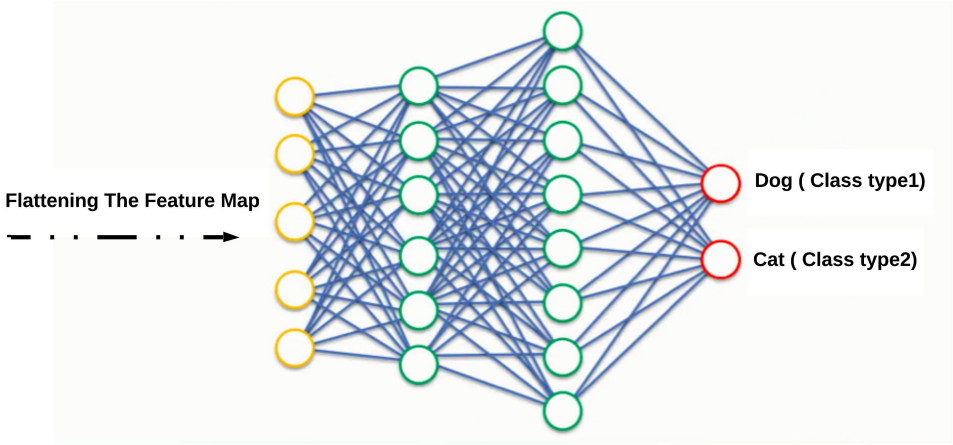


Fig. II.5: The architecture of Fully Connected Layers

II.2.4 CNN model

A CNN is a common type of deep learning model designed to learn a complex hierarchy of features in signal or image data. It is structured as a multi-stage neural network where each stage represents a different level of abstraction. This network consists of feature maps containing neurons that are connected through operations such as convolution, activation, and pooling.In Convolutional Neural Networks (CNNs), three fundamental architectural concepts are used: [34]

- Local Receptive Fields: Neurons in a layer are connected to a small group of neighboring neurons from the previous layer, allowing for local convolutional operations. This concept is inspired by the visual cortex organization in mammals.
- Shared Weights: The weights of a convolutional kernel are consistent across different locations within a feature map at a given layer. This helps in reducing the number of trainable weights significantly.
- Sub-sampling: This concept involves reducing the dimensions of the feature maps by techniques like max-pooling, which helps in capturing the essential information while reducing computational complexity.

These concepts together allow CNNs to efficiently process spatial data such as images.

II.2.5 CNNs for image fusion

II.2.5.1 Feasibility

In summary, fusion in image processing can be seen as a classification problem, where the activity level measurement is a form of feature extraction and the fusion rule acts as a classifier. CNNs can be used for image fusion due to their end-to-end framework, which includes convolutional layers for feature extraction and fully-connected layers for classification. Existing fusion methods often involve designing local filters or using pre-designed bases for activity level measurement, which can be seen as similar to convolving with those bases in CNN architecture.[35]

II.2.5.2 Superiority

Similar to visual object classification applications, the CNN-based fusion method offers two key advantages over existing methods.

Firstly, it eliminates the need for manually designing complex activity level measurement and fusion rules, as the focus shifts to designing the network architecture instead. With the availability of user-friendly CNN platforms like Caffe and MatConvNet, researchers can easily implement the network design.

Secondly, the CNN model allows for the joint generation of activity level measurements and fusion rules through learning. The resulting solution can be considered optimal to some degree, making it likely to be more effective than manually designed rules. As a result, the CNN-based method shows great potential for producing higher quality fusion results compared to traditional methods.[35]

II.2.6 Fusion scheme

II.2.6.1 Focus detection

A and B are the names given to the two original image data sources. If the data sources are colored images, they will be converted to grayscale before merging. Assume A and B are grayscale versions of the original images, stored as A = A and B = B in grayscale format. These grayscale versions are input into a trained CNN architecture to generate a scale map, denoted as parameter S. Each value of S ranges from 0 to 1. The adjusted versions of the images, A' or B', are determined based on the value of S approaching 1 or 0. A 2-pixel footprint is overlaid on relevant sections of each source image in S, representing two nearby parameters. The values of each parameter in S are provided for all pixels within its corresponding patch in M, along with the average of the overlaid pixels, to create a focus map of the same size as the original images. [36]

Fig. II.6: Initial segmentation. (a) Focus map (b) binary segmentation map

In a set of 16x16 patches extracted from the source images, the more a value approaches 1 or 0, the more focused the corresponding patch from source images $\hat{A}$ or $\hat{B}$ appears. Two adjacent coefficients in the set S align their patches in each source image with a step of two pixels. To generate a focus map (referred to as M) with the same dimensions as the source images, we map the value of each coefficient in S to all pixels within its corresponding patch in M and compute the average over the overlapping pixels. In Figure (II.10)(a), it is clear that the focus details are accurately captured. Areas with intricate features tend to have values near 1 (white) or 0 (black), while simpler regions typically have values closer to 0.5 (gray) based on intuition. [36]

II.2.7 Focus detection

The first step in segmentation In order to maintain the relevant information to the highest extent, there is a need for additional processing of the focus map M.

In our approach, similar to many techniques for multi-focus image fusion in spatial domain [20, 19, 18], we also employ the commonly used "choose-max" technique to handle M. Specifically, we apply a constant threshold of 0.5 to divide M into a binary map T, following the classification principle of the trained CNN model. This process essentially involves segmenting the focus map.[36]

$$T(x,y) = \begin{cases} 1 & \text{if } M(x,y) \geq 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{II.1}$$

The binary map obtained is displayed in Fig (II.6)(b). The optical illusion in the focus map shown in Fig (II.6)(a) should be noted, where gray areas appear darker than their actual intensity on a white background and brighter on a black background. The focus map accurately classifies almost all gray pixels, showing that the trained CNN model can achieve accurate performance even in areas without many features in the source images.[36]

II.2.7.1 Consistency verification

Observing Fig (II.6)(b) reveals that the binary segmented map is prone to including certain misclassified pixels, which can be readily rectified through the small region removal technique. Specifically, any region smaller than a specified area threshold is inverted within the binary map. It is worth noting that in some instances, the source images may contain exceedingly tiny voids. In such rare cases, individuals have the option to manually adjust the threshold, even setting it to zero, thereby bypassing the region removal strategy. In this study, the area threshold is consistently established at 0.01 × H × W, where H and W represent the individual height and width of each source image, respectively.[36] .

Fig. II.7: (a) Initial Decision Map (b) Initial Fused Image (c) Final Decision Map (d) Fused Image

In Fig (II.7)(a), you can see the initial decision map generated by following this method. Fig (II.7)(b) shows the combined image created by applying the initial decision map using the weighted-average rule. It is apparent that there are some unwanted flaws near the borders that separate the clear and blurry regions. Just like in prior studies [18], we employ the guided filter to enhance the accuracy of the initial decision map. The guided filter, recognized for its ability to maintain edges, transfers the structural characteristics of a guide image to the filtered result of the input image. The initial fused image acts as the guide image to direct the filtration process of the initial decision map. The guided filtering procedure involves two adjustable parameters: the local window radius noted as 'r' and the regularization parameter. ,[36] denoted as ϵ

II.2.7.2 Fusion

At last, using the obtained decision map D, we compute the fused image F using the pixel-wise weighted-average rule.[36]

$$F(x,y) = D(x,y)A(x,y) + (1 - D(x,y))B(x,y) \tag{II.2}$$

II.3 Latent Low-Rank Representation (LATLRR)

Latent Low-Rank Representation (LatLRR) is a popular method for combining visible and infrared images. It breaks down images into three key components: the base, salient, and sparse parts. While current approaches focus on merging the base and salient parts, this study argues that all three components should be included in the fusion process to achieve accurate image reconstruction. Additionally, incorporating Convolutional Neural Networks (CNNs) with LatLRR is challenging, especially when sparse parts are involved.The results show that incorporating sparse parts significantly improves fusion performance. The proposed strategy involves using deep learning to fuse base and sparse parts, and summation to combine salient parts. This approach enhances the performance of LatLRR-based methods and provides valuable insights for future advancements in image fusion.[37]

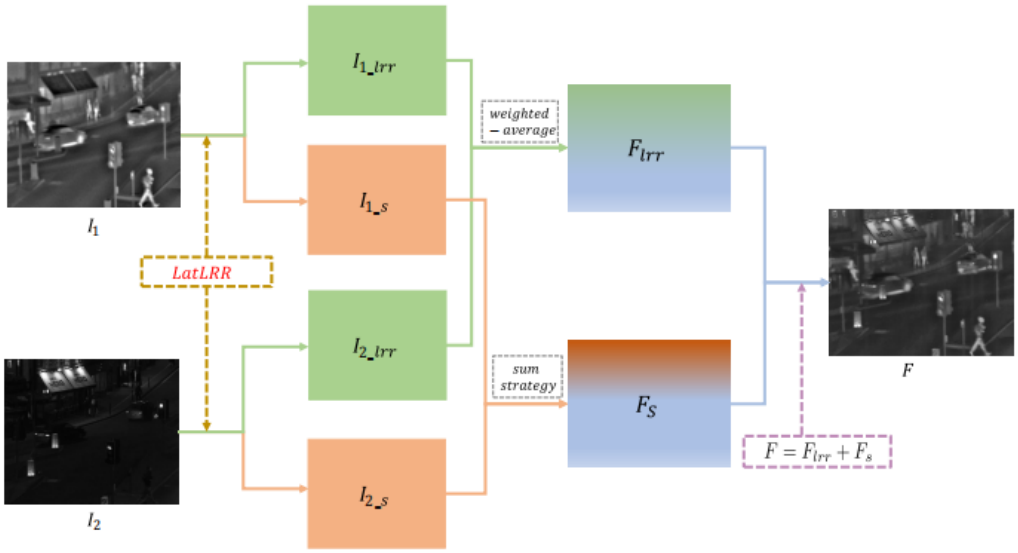


Fig. II.8: The framework of proposed method

II.3.1 Strategy LatLRR method

In this paragraph, we will explain the proposed image merging method in detail. The approach involves decomposing input images into low-rank and salient parts using Latent Low-Rank Representation (LatLRR), as illustrated in Fig.9. The decomposition process yields a low-rank part (Xlrr) and a salient part (Xs) for each input image (X). The proposed fusion method, shown in Fig.7, takes two source images (I1 and I2) as input and applies LatLRR to obtain low-rank and salient parts for each image. The low-rank and salient parts are then fused separately using a weighted-average strategy. Finally, the fused image (F) is reconstructed by combining the fused low-rank part (Flrr) and salient part (Fs).[37]

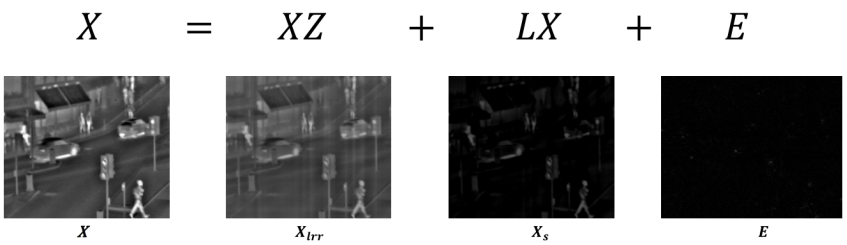


Fig. II.9: The framework of proposed method

II.3.1.1 Fusion of low-rank parts

The low-rank components of the source images capture more general structural details and overall brightness. Therefore, in our image fusion approach, we employ a weighted averaging technique to combine these low-rank components. This fused low-rank component is computed using Equation (II.3).[37]

$$Flrr(i,j) = w1I1lrr(i,j) + w2I2lrr(i,j) \tag{II.3}$$

The low-rank components of the source images contain essential global structure and brightness information. To effectively combine these components, our fusion method employs a weighted average approach. Specifically, the fused low-rank component is calculated using Equation (II.3), where the weights w1 and w2 are assigned to the coefficients of the low-rank components of I1 and I2, respectively, at each position (i,j). This strategy helps preserve the global structure and brightness information while minimizing redundant information, we choose w1 = 0.5 and w2 = 0.5. [37]

II.3.1.2 Fusion of salient parts

The key components of the model incorporate both local structural details and prominent characteristics, as illustrated in Fig.II.10

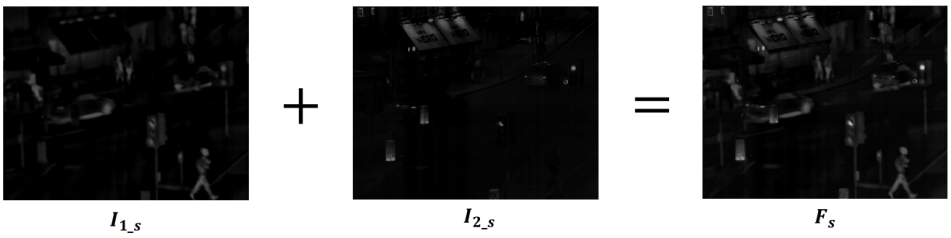


Fig. II.10: The fusion procedure of salient parts

The key point is that the salient parts of each source image, denoted as I1s and I2s, contain complementary information that needs to be preserved in the fused image. using a simple sum strategy to fuse the salient parts, as represented in Equation (II.4).In other words, the idea is to add the salient features from each source image to create a

fused image that retains all the important information from both sources. This approach ensures that no salient features are lost during the fusion process. [37]

$$Fs(i,j) = s1I1s(i,j) + s2I2s(i,j) \tag{II.4}$$

In Eq. (II.4), (i, j) is for coefficients of I1s, I2s, and Fs. Choosing s1 = 1, s2 = 1

II.3.1.3 The reason of choose sum strategy

In this section, we'll justify our approach of using a simple summation strategy to combine the salient parts. Specifically, we select coefficients from two salient parts that are located in the same row, as shown in Fig. II.11(a) and Fig. II.11.4(b).

(a) (b)

Fig. II.11: The saliency part of infrared image(a) and The saliency part of visible image(b)

In Fig. II.11.(a) and Fig. II.11.(b), infrared salient part > visible part in red boxes. Visible part > infrared in third red box. Salient features from source images must be in fused image without loss (sum strategy retains features weight-average reduces them).[37]

II.3.1.4 Reconstruction

After calculating the fused low-rank part Flrr and salient part Fs using Equation (II.3) and Equation (II.4), the fused image F is reconstructed through Equation (II.5)[37]

$$F(i,j) = Flrr(i,j) + Fs(i,j) \tag{II.5}$$

II.3.2 Summary

Here is a summary of the fusion method proposed based on LatLRR:

- The source images are decomposed by LatLRR into the low-rank parts IC lrr and the salient parts IC lrr .
- Two fusion strategies, weighted-average fusion, and sum strategy, are utilized to fuse the low-rank parts and salient parts. This results in the fused low-rank part Flrr and the fused salient part Fs.
- Finally, the fused image F is obtained using Eq (II.5).[37]

II.4 NSST and Improved PCNN

In this study, we present a new algorithm for merging visible and infrared images., taking advantage of the unique characteristics of each type of image. The proposed method uses a non-subsampled shearlet transform (NSST) to decompose the source images into high and low frequency components. The low frequency components are then fused using a modified spatial frequency approach, which incorporates a pulse-coupled neural network (PCNN) and adaptive link strength adjustment. For the high frequency components, a self-adaptive fusion rule is applied based on local area variance and gradient. Finally, the fused low and high frequency components are reconstructed using the inverse NSST transform, resulting in a fused image that effectively combines the important information from both infrared and visible images. Experimental results demonstrate that the proposed method outperforms other image fusion techniques based on non-subsampled contourlet transform (NSCT) and NSST. [38]

II.4.1 Strategy NSST method

II.4.1.1 Low Frequency Components Fusion Rule

After applying the non-subsampled shearlet transform (NSST), the fundamental information of the image is extracted, which involves removing the texture and fine details. Therefore, fusing the low-frequency components is crucial. While the adjacent domain energy extraction method helps preserve the essential edges of the source image, it tends to weaken the edge information in lower-brightness areas. To address this, a modified spatial frequency (MSF) approach is employed to stimulate pulse-coupled neural network (PCNN) neurons. This approach effectively mitigates the Gibbs phenomenon and provides a more comprehensive representation of image details. The specific process is outlined as follows:[38]

Calculate the modified spatial frequency MSF

The spatial frequency (SF) method involves using a sliding window to extract low-frequency components and calculating the gradient energy in both row and column directions. The low-frequency components, denoted as I C (i, j), are obtained through a Non-Subsampled Shearlet Transform (NSST) decomposition. The definition of MSF (Multiscale Spatial Frequency) is as follows:

$$MSF^I = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (RF + CF + MDF + SDF) \tag{II.6}$$

In Equation (II.6). The variables RF, CF, MDF, and SDF represent row frequency, column frequency, main pair angular frequency, and pair angular frequency, respectively. Additionally, M and N signify the dimensions of the rectangular window used for I MSF. [38]

Calculate the adaptive link coefficients

Reflects the average color gamut of the image. It can capture fine details such as edges, borders and textures. Mathematically, the average gradient can be expressed as follows:

$$\overline{G}^I(i;j) = \frac{1}{(M-1)(N-1)} \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} [(C^I(i+1,j) - C^I(i,j)^2 + (C^I(i,j+1) - (C^I(i,j))^2]/2 \tag{II.7}$$

According to equation (II.8), the term B represents the variance.[39]

$$\beta_{ij}^I = \frac{1}{1 + exp(-\overline{G}^I(i,j))} \tag{II.8}$$

Calculate low Frequency Fusion Components

Here is a paraphrased version:

- The adaptive connection coefficients I ij are calculated using the method described earlier.
- The parameters of the PCNN model are initialized, with L(0) = U(0) = Y(0) = T(0), and the maximum number of iterations set to N max.
- The input images are fed into the PCNN model as motivational signals, and the total number of firings for the two source images are recorded as T(n) and T(n) at each iteration.
- The low-frequency fusion result is obtained by applying Equation (II.9) to the PCNN output.[38]

$$C^F(i,j) = \begin{cases} C^A(i,j), T_{ij}^A(n) \geq T_{ij}^B(n) \\ (C^B(i,j), otherwise \end{cases} \tag{II.9}$$

II.4.1.2 High Frequency Components Fusion Rule

Following the decomposition of NSST, the high frequency components primarily represent the edge details and texture information in images, significantly impacting the overall visual quality of the fusion image. The variability of the local area indicates the magnitude of changes within regional components, serving as a measure of image clarity in that specific area. Furthermore, the local average gradient provides insight into finer details such as boundaries and textures in the region. Therefore, in order to enhance the high frequency components, an adaptive weighting method based on local average gradient and variance is employed.[38]

II.5 Conclusion

Image fusion is a critical area of study in image processing, with a primary focus on pixel-level fusion currently. This process involves combining multiple images from various sensors to create a new image with enhanced information suitable for tasks like target recognition and image comprehension. This chapter introduces methods such as LATLRR and NSST for image fusion, and also utilizes the CNN method for hot spots images in our research.[40]

Chapter III

Result and discussion

III.1 Introduction

In this chapter we will present the simulation results of the method CNN which it combines the advantages of several techniques to provide better quality image. Finally, we will compare it with other methods on the same database.

III.2 Experiment

In this section, we present a comparative study of image fusion for solar panels. The performance of 3 representative fusion methods is evaluated on 6 pairs of solar panel images, using 6 objective metrics.

III.3 Proposed working method for CNN

In this chapter we studied a image fusion for solar panels method based on focus region detection using Convolutional Neural Networks(CNN). So we have to determine the fusion of the images to get a more clean and precise images using the CNN method. The important steps of this method are summarized as follows:

- Loading and Preparation: Data is loaded, variables are initialized, and model weights and biases are prepared.
- Layer 1: Data is passed through the first layer of the neural network with the ReLU activation function applied to obtain the layer's output.
- Layer 2: Repeat the previous steps for the second layer while calculating the new data input and output.
- Layer 3: Combine the data from the previous two layers and pass it through this layer to get the final output.
- Feature layer: Apply feature extraction to the data collected from the previous layer.

- Softmax output layer: Calculate the final results by applying the softmax function to the previous outputs.
- Create a concentration map: Calculate the concentration map using the final results.
- Image Merging: Use the output map to combine images using the specified merging technique.
- Calculate performance metrics: Calculate merger performance metrics such as QG, QP, QY, QXY/F, LXY/F, NXY/F.
- Display results: Display the merged image and different metrics to evaluate the quality of the merge.

III.4 Data collection

In our study of this experiment, we need a thermal camera and solar panels to collect data. In our search for thermal camera, we found it at Sonelgaz Foundation. After taking the necessary administrative measures, it was agreed to take the camera outside the institution to photograph solar panels from another institution (Zergoun green energy), which specializes in manufacturing them. After it became clear that the data was insufficient to complete the study, we went to the photovoltaic power station in the Al- hadjira area, where we were received after taking the necessary procedures and photographed using the institution's camera, which took thermal and color images .

III.5 Parameters setting

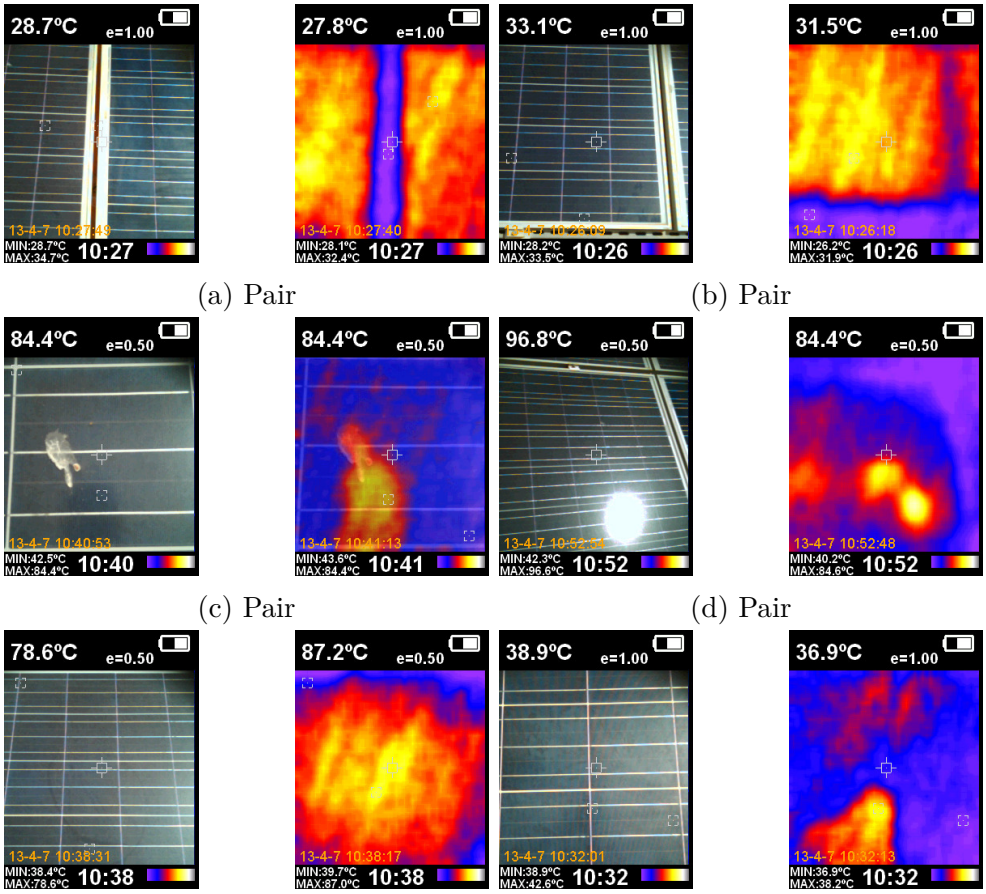
During experiment, we use 6 pairs solar panel images as test images, and calculate the above six objective evaluation metrics (view chapter1)to evaluate the influence of three parameters on fusion performace Besides, considering the parameters of the compared fusion methods have been already optimized by their authors in program codes, so we keep them unchangeable in the experiments.

III.6 Method validation

The assessment of the fused image quality is based on visual analysis and quantitative analysis. In quantitative analysis we used several statistical metrics (Q_G , Q_P , Q_Y , $Q^{XY/F}$, $L^{XY/F}$, $N^{XY/F}$). In qualitative analysis, it depends on visuality to locate and identify faults that may affect the image quality, deformation of linear elements, smearing of contours, aggregation of objects, color saturation or dominant colors. This analysis is necessary to verify the quality of the images obtained by the fusion.

III.7 Experimental settings

In order to evaluate the performance of the studied method, 6 pairs of solar panel images used as test images are shown in fig(III.1). The studied fusion method (CNN) is compared with two resentative methods, which are : NSST stands for Non-Subsampled Shearlet Transform and it analyzes images by analyzing lines and edges at different slant angles. NSST is characterized by its ability to effectively represent irregular shapes and enhance image details. LATLRR stands for Local Activity-tuned Low Rank and Sparse Representation and focuses on representing data using the concept of low-rank and sparse representation. LATLRR relies on improving image quality by combining low-rank and sparse information. To obtain the results of these methods, we use the original codes of these methods, which are provided onine by their authors.



(a) Pair (b) Pair

(c) Pair (d) Pair

(e) Pair (f) Pair

Fig. III.1: 6 pairs of solar panel images used as test images.

III.8 Comparison with other image fusion methods

In this subsection, we compare with the fused images obtained by different fusion methods both in visual effect and objective evaluation metrics.

III.8.1 Quantitative analysis

The objective performance evaluation is included in this subsection, like shown in table (III.2), Table(III.3), Table(III.4) List of values for the six objective evaluation metrics Q_G , Q_P , Q_Y , $Q^{XY/F}$, $L^{XY/F}$, $N^{XY/F}$ with different methods CNN, LATLRR, NSST, respectively on 6 pairs solar panel images. Meanwhile, Table(III.1) Average values of the six metrics above with the three methods.The better results are shown in bold. It can be seen that the metric values of the three fusion methods can be divided into two levels. First level, the value of LATLRR are small, which indicate that the fused images of transform domain methods lose more image spatial information than spatial domain methods. Second level, the values of CNN and NSST methods are larger than the above methods. This indicates the fused images of CNN have better effects, which is consistent with subjective analysis. The fused images of the proposed fusion method CNN perform well in both subjective and objective evaluation, and present satisfactory fusion effects.

Table III.1: Average evaluation metrics of different methods on 6 pairs of color and thermal images

Metrics Methods	Q_G	Q_P	Q_Y	$Q^{XY/F}$	$L^{XY/F}$	$N^{XY/F}$
CNN	**0,6390**	**0,7834**	0,7591	**0,9751**	0,0194	0,0047
LATLRR	0,4807	0,5657	0,7293	0,9401	**0,0588**	0.0016
NSST	0,6032	0,5864	**0,8325**	0,9547	0,0414	**0,0147**

Table III.2: Evaluation metrics for CNN methods on 6 pairs of color and thermal images

Metrics Pairs	Q_G	Q_P	Q_Y	$Q^{XY/F}$	$L^{XY/F}$	$N^{XY/F}$
A	0.5905	0.7346	0.7453	0.9717	0.0230	0.0053
B	0.6355	0.7803	0.7327	0.9784	0.0172	0.0044
C	0.5846	0.8277	0.6970	0.9856	0.0118	0.0016
D	0.7206	0.7706	0.8566	0.9664	0.0260	0.0076
E	0.5949	0.7824	0.7326	0.9698	0.0241	0.0060
F	0.7084	0.8053	0.7905	0.9818	0.0145	0.0037

Table III.3: Evaluation metrics for LATLRR methods on 6 pairs of color and thermal images

Metrics / Pairs	Q_G	Q_P	Q_Y	$Q^{XY/F}$	$L^{XY/F}$	$N^{XY/F}$
A	0.4949	0.5763	0.7327	0.9284	0.0703	0.0012
B	0.5000	0.6228	0.7134	0.9334	0.0657	0.0008
C	0.4467	0.4594	0.7080	0.9512	0.0476	0.0021
D	0.5063	0.6607	0.7605	0.9395	0.0585	0.0020
E	0.4680	0.5607	0.7374	0.9203	0.0782	0.0014
F	0.4688	0.5144	0.7263	0.9638	0.0339	0.0023

Table III.4: Evaluation metrics for NSST methods on 6 pairs of color and thermal images

Metrics / Pairs	Q_G	Q_P	Q_Y	$Q^{XY/F}$	$L^{XY/F}$	$N^{XY/F}$
A	0.5772	0.6128	0.8032	0.9403	0.0527	0.0070
B	0.6109	0.6377	0.7886	0.9518	0.0431	0.0051
C	0.5610	0.4830	0.8268	0.9588	0.0384	0.0028
D	0.6502	0.7027	0.8673	0.9563	0.0410	0.0027
E	0.5819	0.5714	0.8440	0.9454	0.0500	0.0045
F	0.6385	0.5111	0.8661	0.9741	0.0234	0.0005

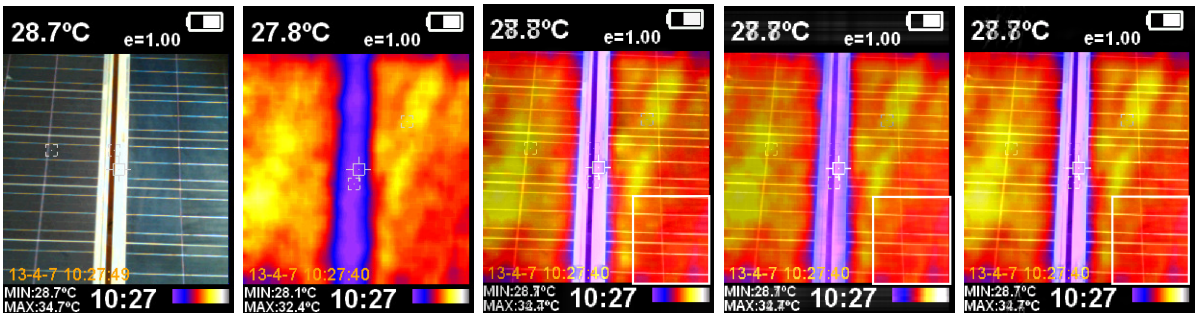
III.8.2 Qualitative analysis

The combined results of the different merging methods for these two examples are shown in Figure III.2-3. In Figures (Figure III.2) and (Figure III.3), subfigures (a) and (b) show a color image and a thermal image, and the combined results obtained by the CNN, LATLRR and NSST fusion methods are shown in subfigures (c), (d), and (e), respectively.

These images and combined results are analyzed and discussed in detail below.

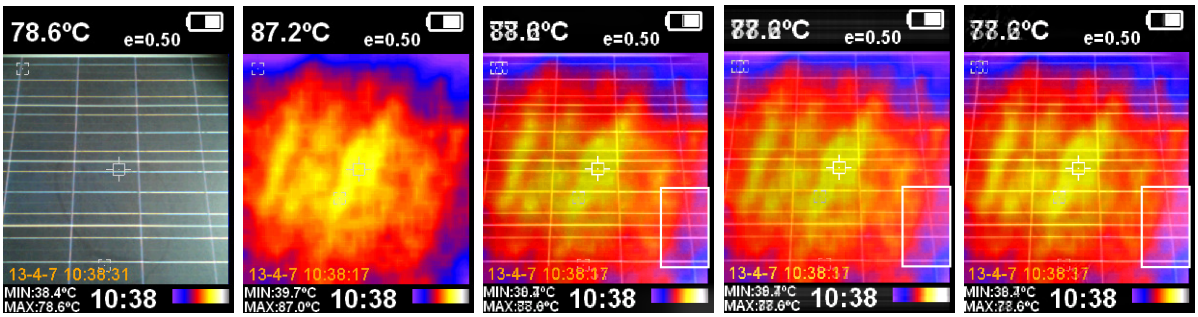
Figure III.2 shows the source images and the combined results of the “solar panel” images. The white frame is also observed in the figure, and it can be found that the clarity and contrast of the temperature gradient in the white frame region are well preserved by LATLRR and NSST, as shown in subfigures (d) and (e), respectively. It can be seen from subfigure (c) that the temperature gradient in the white frame area in the image fused by CNN is more pronounced than the above two methods.

Figure III.3 shows the source images and the merged results of the “other solar panel” images, as noted by the white frame in the figure. NSST can get good visual effect. Through careful observation, it can be seen that the temperature gradient contrast in the white frame area in Figure (III.3) outperforms the clarity and contrast of LATLRR compared to CNN, as shown in Figure (d) and Figure (c).



(a) Source(A) (b) Source(A) (c) CNN (d) LATLRR (e) NSST

Fig. III.2: source images and fused images by different methods



(a) Source(E) (b) Source(E) (c) CNN (d) LATLRR (e) NSST

Fig. III.3: source images and fused images by different methods.

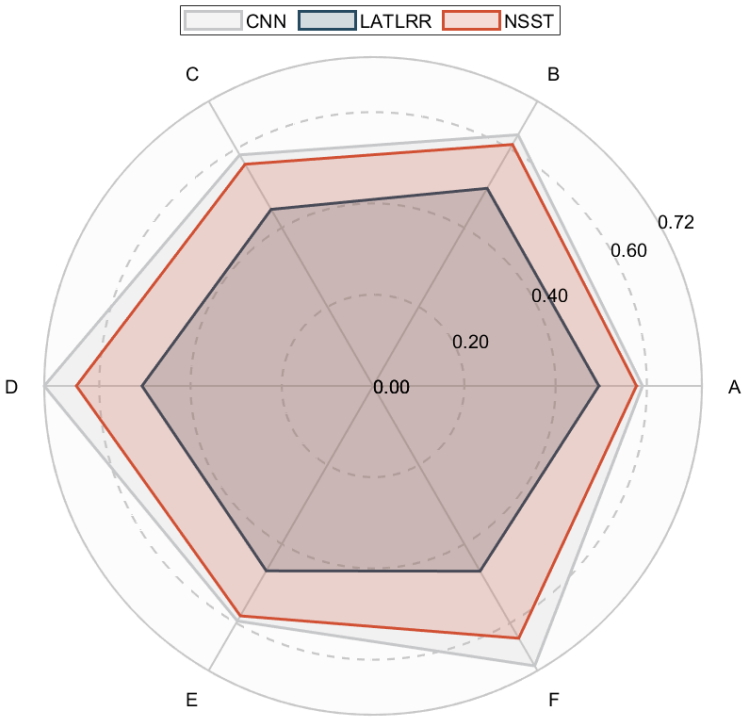


Fig. III.4: Objective performance of different fusion methods on metric Q_G

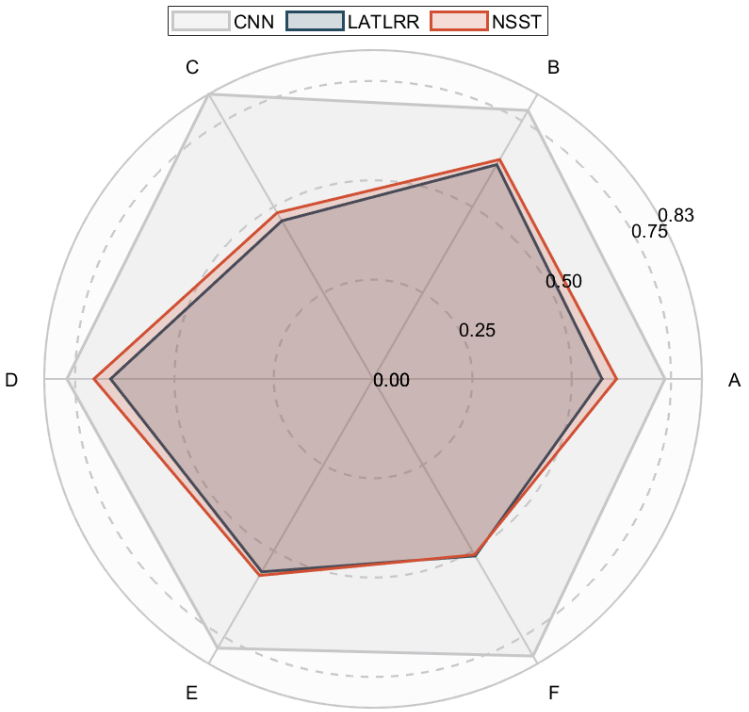


Fig. III.5: Objective performance of different fusion methods on metric Q_P

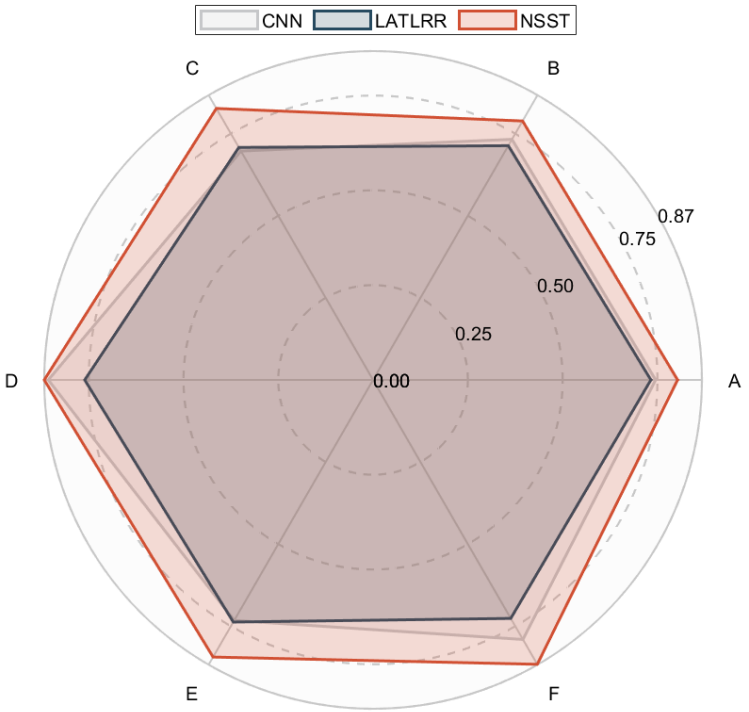


Fig. III.6: Objective performance of different fusion methods on metric Q_Y

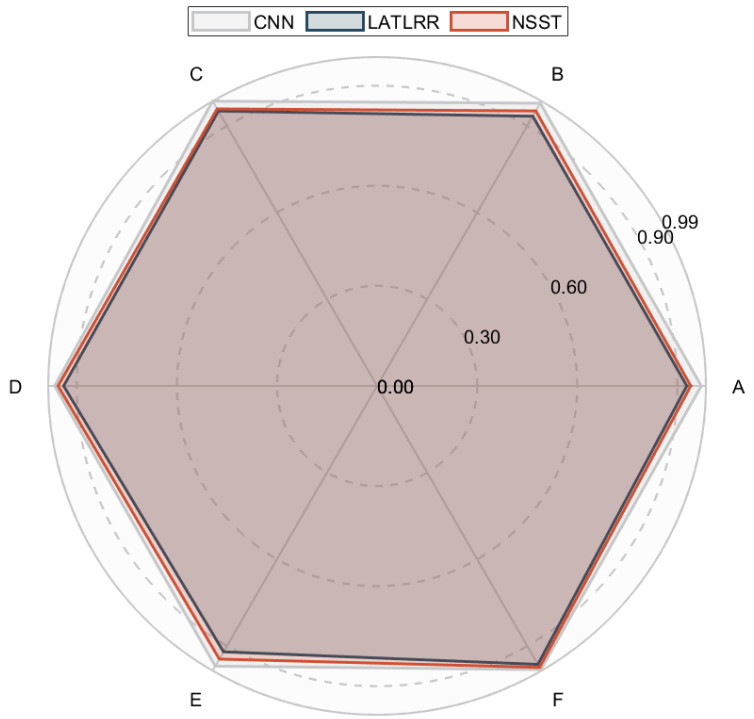


Fig. III.7: Objective performance of different fusion methods on metric $Q^{XY/F}$

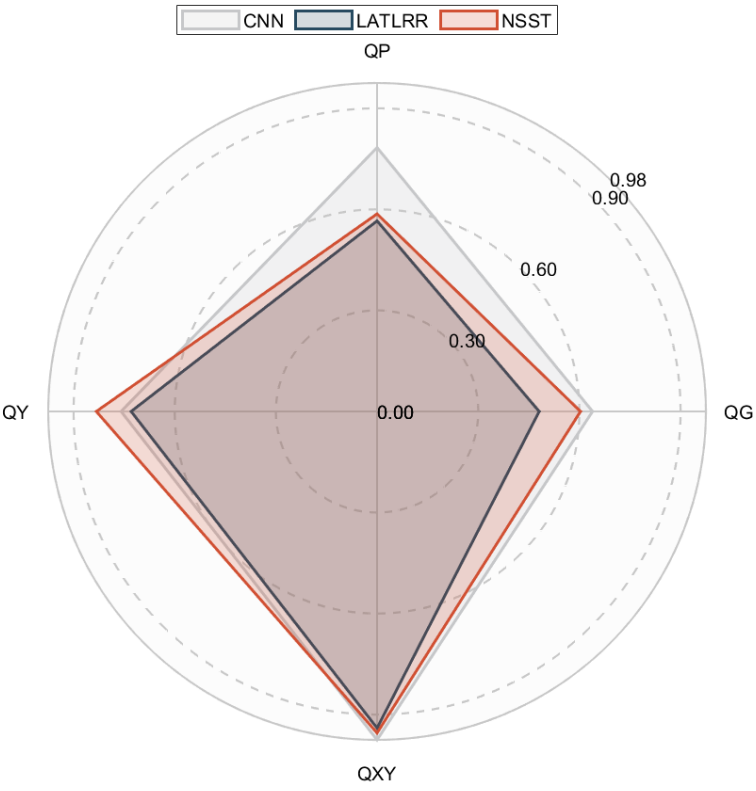


Fig. III.8: Average some evaluation metrics for different methods

Figure (III.4),(III.5),(III.6), (III.7), and (III.8) provide further insights into the objective performance of different merging methods on a single dataset. These figures show a hexagonal grid diagram (radar) used to compare the performance of three different models: CNN, LATLRR, and NSST.

In Figure (III.4), (III.5), (III.6), and Figure (III.7), six axes (A, B, C, D, E, F) represent the studied merged images, where it is noted that the blood CNN method Its area in the radar chart is greater than LATLRR and NSST in the three scales (Q_G, Q_P, $Q^{XY/F}$), with the exception of the scale (Q_Y), which gives a good result for the NSST integration method, as shown in the figure. (III.6), this indicates that the CNN fusion method gives most of the data in the images fusion between color and thermal images.

As for Figure (III.8), which represents the average of some of the four metrics studied (Q_G, Q_P, $Q^{XY/F}$, Q_Y), we notice that the CNN fusion method prevails in the three metrics, and thus the This method gives the best fusion results compared to other methods.

III.9 Conclusion

Image fusion is the process of fusion two or more images from different or same sources to create a new image with more comprehensive information. The final section of the study focused on comparing the effectiveness of the CNN method with other techniques such as LATLRR and NSST when applied to color and thermal images. By evaluating six factors, it was clear that the CNN method achieved the best results among the three methods.

Final conclusion

Image fusion plays a vital role in various fields, including merging thermal images with color images, which is considered an important field of research. Being an effective way to collect information, combining thermal images with color images has attracted the attention of researchers in the field of image processing. This approach is widely used in the field of medicine, but in this case our attention was directed to another field, namely renewable energy, where we fused thermal images with color images, and then discussed the different methods and types of image fusion.

Many methods address the problem of image fusion, one of which is merging thermal and color images. In order to discover this, we took a method (CNN) that can be used to analyze both thermal and color images. For thermal images, artificial neural networks can learn efficient representations of thermal data and identify regions of interest in images. For color images, it can be used to classify objects in images and reveal patterns and relationships between different elements. In short, it can also be used to analyze and process images regardless of their type. The functioning of a (CNN) includes several layers of transformations and filters that are used to extract important visual features from images. The process begins with the convolutional layer. Then followed by the Pooling layer. Followed by Pooling and Activation layers. Next, Fully Connected layers are used for final classification of the image based on the features extracted from it. Among several image fusion methods, we compared our CNN method to two representative methods. Specifically, in transform domain methods: (NSST), (LATLRR). In our experience, we clearly saw that the (CNN) method has the best results according to the results provided qualitatively (merged image results) and quantitatively (target evaluation metrics).

Due to time constraints, we could not conduct further experiments with other parameter setting values. In future work, we may address this issue and compare it with more image fusion methods using more metrics for better evaluation.

Bibliography

[1] Harpreet Kaur, Deepika Koundal, and Virender Kadyan. Image fusion techniques: a survey. *Archives of computational methods in Engineering*, 28(7):4425–4447, 2021.

[2] Ayush Dogra, Bhawna Goyal, and Sunil Agrawal. Medical image fusion: A brief introduction. *Biomedical & Pharmacology Journal*, 11(3):1209, 2018.

[3] Zeyu Wang, Xiongfei Li, Haoran Duan, Yanchi Su, Xiaoli Zhang, and Xinjiang Guan. Medical image fusion based on convolutional neural networks and non-subsampled contourlet transform. *Expert Systems with Applications*, 171:114574, 2021.

[4] Dominic J Mccafferty. The value of infrared thermography for research on mammals: previous applications and future directions. *Mammal Review*, 37(3):207–223, 2007.

[5] Cheng-Shu You and Suh-Yuh Yang. A simple and effective multi-focus image fusion method based on local standard deviations enhanced by the guided filter. *Displays*, 72:102146, 2022.

[6] Jiayi Ma, Yong Ma, and Chang Li. Infrared and visible image fusion methods and applications: A survey. *Information fusion*, 45:153–178, 2019.

[7] Mritunjay Rai, Tanmoy Maity, and RK Yadav. Thermal imaging system and its real time applications: a survey. *Journal of Engineering Technology*, 6(2):290–303, 2017.

[8] Antoni Rogalski and Krzysztof Chrzanowski. Infrared devices and techniques. In *Handbook of optoelectronics*, pages 633–686. CRC Press, 2017.

[9] Kurt Ammer. The influence of colour scale on the accuracy of infrared thermal images. In *Infrared Imaging: A casebook in clinical medicine*, pages 4–1. IOP Publishing Bristol, UK, 2015.

[10] Rubén Usamentiaga, Pablo Venegas, Jon Guerediaga, Laura Vega, Julio Molleda, and Francisco G Bulnes. Infrared thermography for temperature measurement and non-destructive testing. *Sensors*, 14(7):12305–12348, 2014.

[11] Robert N Castellano. *Solar panel processing*. Archives contemporaines, 2010.

[12] Azhar M Ghazali and Abdul Malek Abdul Rahman. The performance of three different solar panels for solar electricity applying solar tracking device under the malaysian climate condition. *Energy and environment Research*, 2(1):235, 2012.

[13] Greenmatch.co.uk. Advantages anddisadvantages of solar energy. urlhttps://www.greenmatch.co.uk/blog/2014/08/5-advantages-and-5-disadvantages-of-solar-energy. April 31, 2024.

[14] A Ardeshir Goshtasby and Stavri G Nikolov. Guest editorial: Image fusion: Advances in the state of the art. *Information Fusion: Special Issue on Image Fusion: Advances in the State of the Art*, 8:114–118, 2007.

[15] Zheng Li, Yongcheng Wang, Ning Zhang, Yuxi Zhang, Zhikang Zhao, Dongdong Xu, Guangli Ben, and Yunxiao Gao. Deep learning-based object detection techniques for remote sensing images: A survey. *Remote Sensing*, 14(10):2385, 2022.

[16] Levent Karacan. Multi-image transformer for multi-focus image fusion. *Signal Processing: Image Communication*, 119:117058, 2023.

[17] Fang Xu, Jinghong Liu, Yueming Song, Hui Sun, and Xuan Wang. Multi-exposure image fusion techniques: A comprehensive review. *Remote Sensing*, 14(3):771, 2022.

[18] Yanling Chen, Lianglun Cheng, Heng Wu, Fei Mo, and Ziyang Chen. Infrared and visible image fusion based on iterative differential thermal information filter. *Optics and Lasers in Engineering*, 148:106776, 2022.

[19] Haithem Hermessi, Olfa Mourali, and Ezzeddine Zagrouba. Multimodal medical image fusion review: Theoretical background and recent advances. *Signal Processing*, 183:108036, 2021.

[20] Mark Vincent Mangrich. Hierarchical feature extraction: A stepwise approach to image classification. 2000.

[21] Luigi Landini, Luca T Mainardi, Vincenzo Positano, Alistair A Young, Maria Santarelli, Leslie Ying, Walid Elias Kyriakos, AJ den Dekker, Martin Styner, Yaroslav O Halchenko, et al. *Advanced image processing in magnetic resonance imaging*. CRC press, 2018.

[22] Fella Charif, Kenza Benzid, and Kheira Zaatout. *Fusion d'images médicales multimodales utilisant les méthodes meta-heuristiques.* PhD thesis, UNIVERSITY OF KASDI MERBAH OUARGLA.

[23] Liping Yan, Yulei Liu, Bo Xiao, Yuanqing Xia, and Mengyin Fu. A quantitative performance evaluation index for image fusion: Normalized perception mutual information. In *Proceedings of the 31st Chinese Control Conference*, pages 3783–3788. IEEE, 2012.

[24] Erik Blasch, Xiaokun Li, Genshe Chen, and Wenhua Li. Image quality assessment for performance evaluation of image fusion. In *2008 11th International Conference on Information Fusion*, pages 1–6. IEEE, 2008.

[25] Shutao Li, Xudong Kang, Leyuan Fang, Jianwen Hu, and Haitao Yin. Pixel-level image fusion: A survey of the state of the art. *information Fusion*, 33:100–112, 2017.

[26] Xiaoyan Jiang, Zuojin Hu, Shuihua Wang, and Yudong Zhang. Deep learning for medical image-based cancer diagnosis. *Cancers*, 15(14):3608, 2023.

[27] Rabia Zafar, Muhammad Shahid Farid, and Muhammad Hassan Khan. Multi-focus image fusion: algorithms, evaluation, and a library. *Journal of Imaging*, 6(7):60, 2020.

[28] Ayodeji Olalekan Salau, Shruti Jain, and Joy Nnenna Eneh. A review of various image fusion types and transform. *Indonesian Journal of Electrical Engineering and Computer Science*, 24(3):1515–1522, 2021.

[29] Shadman Sakib, Nazib Ahmed, Ahmed Jawad Kabir, and Hridon Ahmed. An overview of convolutional neural network: Its architecture and applications. 2019.

[30] Saad Albawi, Oguz Bayat, Saad Al-Azawi, and Osman N Ucan. Social touch gesture recognition using convolutional neural network. *Computational Intelligence and Neuroscience*, 2018(1):6973103, 2018.

[31] Laith Alzubaidi, Jinglan Zhang, Amjad J Humaidi, Ayad Al-Dujaili, Ye Duan, Omran Al-Shamma, José Santamaría, Mohammed A Fadhel, Muthana Al-Amidie, and Laith Farhan. Review of deep learning: concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, 8:1–74, 2021.

[32] Sakshi Indolia, Anil Kumar Goswami, Surya Prakesh Mishra, and Pooja Asopa. Conceptual understanding of convolutional neural network-a deep learning approach. *Procedia computer science*, 132:679–688, 2018.

[33] Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9:611–629, 2018.

[34] Manjunath Jogin, MS Madhulika, GD Divya, RK Meghana, S Apoorva, et al. Feature extraction using convolution neural networks (cnn) and deep learning. In *2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)*, pages 2319–2323. IEEE, 2018.

[35] Yu Liu, Xun Chen, Hu Peng, and Zengfu Wang. Multi-focus image fusion with a deep convolutional neural network. *Information Fusion*, 36:191–207, 2017.

[36] Akansha Sharma, Hitakshi Gupta, and Yashika Sharma. Image fusion with deep leaning using wavelet transformation. *J. Emerg. Technol. Innov. Res*, 8:2826–2834, 2021.

[37] Min Li, Xianjie Yuan, Zhidan Luo, and Xiaohua Qiu. Infrared and visual image fusion based on nsst and improved pcnn. In *Journal of Physics: Conference Series*, volume 1069, page 012151. IOP Publishing, 2018.

[38] Weiwei Kong, Longjun Zhang, and Yang Lei. Novel fusion method for visible light and infrared images based on nsst–sf–pcnn. *Infrared Physics & Technology*, 65:103–112, 2014.

[39] Baohua Zhang, Xiaoqi Lu, Haiquan Pei, and Ying Zhao. A fusion algorithm for infrared and visible images based on saliency analysis and non-subsampled shearlet transform. *Infrared Physics & Technology*, 73:286–297, 2015.

[40] Shrida Kalamkar et al. Multimodal image fusion: A systematic review. *Decision Analytics Journal*, page 100327, 2023.