UNIVERSITY OF SCIENCE AND TECHNOLOGY OF HANOI

ICT DEPARTMENT

Research and Development

BACHELOR THESIS

by

DAO Duy Manh Ha

BI12-141

Information and Communication Technology

Title:

Medical Image Enhancement using Optimization and Deep Learning

Supervisors: Dr. NGHIEM Thi Phuong

Lab name: ICTLab - USTH

Hanoi, September 2024

Attestation

I hereby, DAO Duy Manh Ha, certify that my thesis does not contain plagiarism (copy/-paste) contents from other sources.

In case of plagiarism in my thesis, I know the consequences and I understand that my thesis won't be evaluated. In this case, my internship will be noted as "fail".

October 4, 2024
Signature

DAO Duy Manh Ha

Acknowledgements

"Knowledge is power", and on the path leading to seizing knowledge through scientific research, much of my appreciation is for the University of Science and Technology of Hanoi (USTH), especially the ICTLab, which offered me a great opportunity to professionally work with researchers and other gifted students. The aforementioned chance contributes considerable to my process of grasping knowledge and strengthening my foundation for further scientific research in the future.

Words cannot express my gratitude to my supervisor, Dr. NGHIEM Thi Phuong, for her huge support and perceptive critiques throughout the research. Under this supervision, her profound comprehension and expertise in image processing domain has significantly aided me to embrace the knowledge and write this graduation thesis.

I would like to extend my sincere thanks to other respectable researchers in the ICTLab: Dr. TRAN Giang Son, MSc. HUYNH Vinh Nam, MSc. KIEU Quoc Viet, ... for their immediate assistance for technical issue and detailed answers for many of my complicated questions.

Last but not least, I would be remiss in not mentioning my family, especially my parents, for their great mental support (and also financial). Their belief in me has kept my spirits and motivation high during the research process. I would also like to thank my other friends, who encouraged me a lot to boost my energy whenever I felt demotivated.

(DAO Duy Manh Ha)

Hanoi, September 2024

Contents

List of Acronyms **i**

List of Figures **ii**

List of Tables **iii**

Abstract **iv**

1 Introduction **1**

1.1 Context and Motivation 1

1.2 Objectives 3

1.3 Thesis Structure 3

2 Materials & Methods **4**

2.1 Materials 4

2.2 Image Enhancement Methods 8

2.2.1 CLAHE 10

2.2.2 DnCNN 11

2.2.3 IFCNN 12

2.3 Optimization Methods 13

2.3.1 War Strategy Optimization 15

2.3.2 Marine Predator Algorithm 17

2.3.3 Comprehensive Learning Marine Predator Algorithm 21

2.4 Proposed Method 23

2.4.1 Enhancement Component 23

2.4.2 Proposed Pipeline 24

3 Evaluation & Results **26**

3.1 Evaluation Metrics 26
3.1.1 Entropy 26
3.1.2 Average Gradient 27
3.1.3 Mean Light Intensity (MLI) 27
3.1.4 Contrast Index 27
3.1.5 Structural Similarity Index (SSIM) 28
3.1.6 Edge-based Similarity ($Q^{AB/F}$) 28
3.1.7 Feature Mutual Information (FMI) 29
3.1.8 Visual Information Fidelity for Fusion (VIFF) 29
3.2 Experimental Setup 30
3.3 Results and Discussion 31
3.3.1 Experiment #1 31
3.3.2 Experiment #2 33
3.3.3 Experiment #3 37

4 Conclusion & Future Works **39**
4.1 Conclusion 39
4.2 Future Works 40

References **44**

A DnCNN layers **v**

List of Acronyms

CL Comprehensive Learning.

CLAHE Contrast-limited Adaptive Histogram Equalization.

CLMPA Comprehensive Learning Marine Predator Algorithm.

CT Computed Tomography.

DnCNN Denoising Convolutional Neural Network.

LED Laplacian Edge Detector.

MPA Marine Predator Algorithm.

MRI Magnetic Resonance Image.

PET Positron Emission Tomography.

SPECT Single Photon Emission Computed Tomography.

TWBA The Whole Brain Atlas.

WSO War Strategy Algorithm.

List of Figures

2.1 Classification of medical images 5
2.2 Modalities of Medical Images 6
2.3 Examples of some pairs of MRI-PET images from TWBA dataset 8
2.4 Statistics of the number of literatures investigating different image enhancement techniques 9
2.5 Process of CLAHE 10
2.6 The overall structure of DnCNN 11
2.7 IFCNN Architecture 12
2.8 An Optimization Problem 13
2.9 Classification of Optimization Algorithms 14
2.10 Meta-heuristic Algorithm Classification 15
2.11 The flowchart of WSO 17
2.12 Lévy flight and Brownian flight 18
2.13 Enhancement Component 23
2.14 Pipeline of our studied method 25

3.1 Boxplot of 30 independent runs on 2 different MRI images 31
3.2 Convergence curve of different runs on image MRI-2 with 3 optimization algorithms 32
3.3 Boxplot of 30 independent runs on 2 different PET-gray images 33
3.4 Comparison of different enhancement techniques carried out on an MRI image from dataset M2 34
3.5 Comparison of different enhancement techniques conducted on an PET-gray image from dataset M3 36
3.6 An instance of image from the dataset M4 with different fusion methods and their cropped frame 38

List of Tables

2.1 Types of available modality in TWBA. 7

3.1 Dataset division 30
3.2 Best, worst, mean and standard deviation of fitness function from 30 independent runs on 2 different MRI images. Bold value is the best value relative to others. 31
3.3 Best, worst, mean and standard deviation of objective function from 30 independent runs of different algorithms on image PET-gray-1 and 2 . . . 33
3.4 Experiment #2 on dataset M2. Red value and blue bold value illustrate the best and the second best relative to others. 34
3.5 Experiment #2 on dataset M3, red and blue bold values indicate the best and the second best result relative to others 35
3.6 Results of Experiment#3, bold value is the best relative to other methods 37

A.1 Some layers' details of DnCNN vi

Abstract

In today's world, healthcare is essential to our lives. Accurate early diagnoses through medical imaging significantly improve treatment outcomes. However, various factors like physical limitations and environmental conditions can degrade image quality, including noise, artifacts, or poor contrast. Moreover, medical images consist of multiple modalities, each with its own strengths. As a result, medical professionals often rely on various image types to gather comprehensive information. Nevertheless, interpreting and analyzing multiple images simultaneously can be challenging and potentially confusing for healthcare providers.

Based on these challenges, in this study we aim to introduce a method combining Contrast-limited Adaptive Histogram Equalization (CLAHE), Denoising Convolutional Neural Network (DnCNN), Laplacian Edge Detector (LED) and meta-heuristic optimization algorithms, which can enhance the quality of medical images to a certain extent. Three different optimization algorithms were chosen for this enhancement task. Furthermore, the enhanced images, particularly MRI and PET, are then fused together using the prebuilt image fusion model namely IFCNN and yield acceptable outcomes in comparison with the new method GBF-Robinson.

Keywords: *Medical Imaging, Image Quality, Image Enhancement, Meta-heuristic Optimization, Image Fusion.*

Chapter 1

Introduction

This section outlines the motivation and objective of our works during the internship, and also provides an overview of how this report is organized.

1.1 Context and Motivation

In our modern era, evidence suggests a correlation between contemporary lifestyles and an increased risk of diseases, particularly cancer. The unhealthy habits prevalent in today's society, such as irregular schedules and consumption of food containing harmful chemicals, may contribute to tumor development. Consequently, healthcare has become increasingly crucial in our lives. To address these challenges, researchers and medical professionals have developed various strategies, with medical imaging emerging as one of the most effective solutions. The adage "an image is worth than thousand words" rings especially true in this context, as the visual data from medical images enables doctors to detect diseases earlier and devise appropriate treatment plans for patients. As normal images captured by ordinary camera systems cannot show the details of organs in patients' body, hence physicists has created specialized machines being able to do that using X-ray, sonar, magnetic field, etc. Subsequently, with a specialized imaging machine, a specific type of medical image can be generated.

Alongside these imaging techniques, the widespread adoption of Internet of Things (IoT) technology and Big Data analytics has facilitated easier acquisition of medical images, resulting in a rapidly growing volume of collected data and information. However, the physical devices used for capturing these images have inherent limitations in their sensors, which can lead to imperfections in the generated images under certain conditions. Consequently, the resulting images may be compromised by issues such as noise, artifacts, or poor contrast. These challenges can potentially diminish the overall quality and comprehensibility of the medical images. Consequently, as a feasible response to these challenges, image enhancement methods has been applied to improve the quality and visual comprehensibility of these medical images. Furthermore, it helps researchers and clinicians to seek to extract more meaningful information from improved data, ultimately leading to higher accuracy in diagnosing diseases and improvement in patient outcomes.

Moreover, each medical imaging modality offers unique advantages in diagnostic and therapeutic applications. Hence, to gain a more comprehensive understanding, medical professionals frequently employ multiple imaging modalities simultaneously, capitalizing on the unique strengths of each technique. However, the manual process of comparing and integrating information from various images of the same anatomical region can be challenging and potentially confusing for physicians, possibly leading to diagnostic errors. Addressing this challenge, multimodal image fusion has emerged as a promising solution within the sub-field of image enhancement. This technique involves combining relevant information from multiple image sources into a single, comprehensive image. By doing so, multimodal image fusion effectively tackles the complexities of manual image comparison and integration, thereby supporting higher accuracy in diagnoses.

Considering recent researches in the medical image enhancement domain, P. H. Dinh applied the Grasshopper optimization into the process of improve medical images [6], Salem *et al.* introduced using histogram equalization for enhancement, Yang *et al.* improved the images through the transform domain- wavelet domain [28], Fractional Differential approach was introduced for image enhancement by Li *et al.* [15], Acharya *et al.* used the genetic algorithm based adaptive histogram equalization method [1], while Zhao *et al.* incorporated luminance-level modulation and gradient modulation [33]. Moreover, many medical image fusion method have been developed: Do *et al.* proposed the combination of equilibrium optimization method and the VGG19 network for enhancing the multimodal medical fused image [8], Liu *et al.* introduced a convolutional neural network approach [16], Lahoud *et al.* proposed a real-time fusion method using novel strategy based on convolutional neural network[14], Wang *et al.* used the combination of convolutional neural network and pyramid representation [25], Huang *et al.* applied the multi-generator multi-discriminator conditional generative adversarial network for finding fusion rule.

1.2 Objectives

Based on that premise, we have proposed a method prior to conducting experiments on it to evaluate the overall performance. The main goals of our proposed pipeline are firstly, to improve the quality of medical images and secondly, to enhance the quality of multi-modal medical fused image.

1.3 Thesis Structure

The thesis is organized into 5 distinguished chapters as follow: **Chapter 1 - Introduction**, giving a short introduction about the motivation and goal of this study; **Chapter 2 - Materials & Methods** presenting background knowledge of basic image enhancement method, optimization and describing our proposed methodology; **Chapter 3 - Evaluation & Results** providing qualitative and quantitative results of our experiments; and finally, **Chapter 4 - Conclusion & Future Works** summarizing this study and discussing on future works.

Chapter 2

Materials & Methods

This chapter describes background knowledge applying for our proposed method including image enhancement techniques such as histogram equalization, noise removal using DnCNN, fusion of two different images and mathematical optimization, especially meta-heuristic algorithms like WSO, MPA, CLMPA. Moreover, the dataset and our methodology are also presented in the chapter.

2.1 Materials

Medical imaging plays a crucial role in modern healthcare, providing non-invasive methods to visualize the internal structures and functions of the human body. These techniques have revolutionized diagnosis, treatment planning, and monitoring of various medical conditions. Different imaging modalities offer unique insights into the body's anatomy and physiology, each with its own strengths and applications. Those modalities consists of Magnetic Resonance Image (MRI), Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), Computed Tomography (CT), etc., and its classification is demonstrated in Fig.2.1 .

More specifically, Magnetic Resonance Image (MRI) uses powerful magnets and radio waves to create high-resolution images of soft tissues, excelling in visualizing the brain, spinal cord, and musculoskeletal system. Positron Emission Tomography (PET) employs radioactive tracers to highlight metabolic processes, making it invaluable for detecting cancer and assessing brain function. Single Photon Emission Computed Tomography (SPECT) is similar to PET but uses different tracers and is particularly useful for cardiac

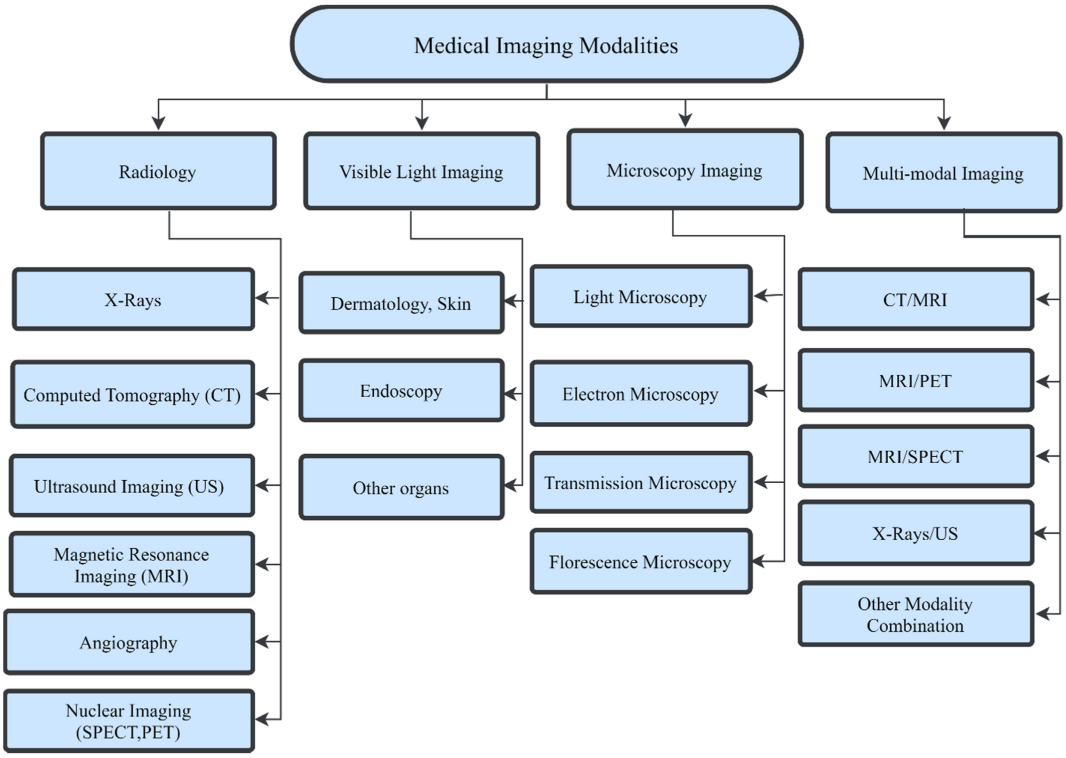


Figure 2.1 – Classification of medical images

and brain imaging. Computed Tomography (CT) combines multiple X-ray images taken from different angles to produce cross-sectional views of bones, blood vessels, and soft tissues, offering quick and detailed scans particularly beneficial in emergency situations. Fig.2.2 shows examples of these modalities.

Regarding the unique strength that each modality offers, MRI excels in providing exceptional soft tissue contrast, making it ideal for examining the brain, spinal cord, and joints without using ionizing radiation. It's particularly effective in detecting subtle changes in tissue composition and identifying small lesions. Meanwhile, PET's strength lies in its ability to visualize metabolic activity and molecular processes, making it invaluable for early cancer detection, staging, and monitoring treatment response. It's also crucial in neurology for diagnosing conditions like Alzheimer's disease. SPECT, while similar to PET, is more widely available and less expensive, offering strong capabilities in cardiac imaging, particularly for assessing blood flow to the heart. It's also useful in brain imaging for conditions like epilepsy and Parkinson's disease. CT's major advantage is its speed and ability to provide detailed images of both soft tissues and bones simultaneously. It's particularly strong in emergency settings, excelling in detecting internal injuries, lung diseases, and complex fractures. CT is also vital in planning radiation therapy and guiding minimally invasive procedures.

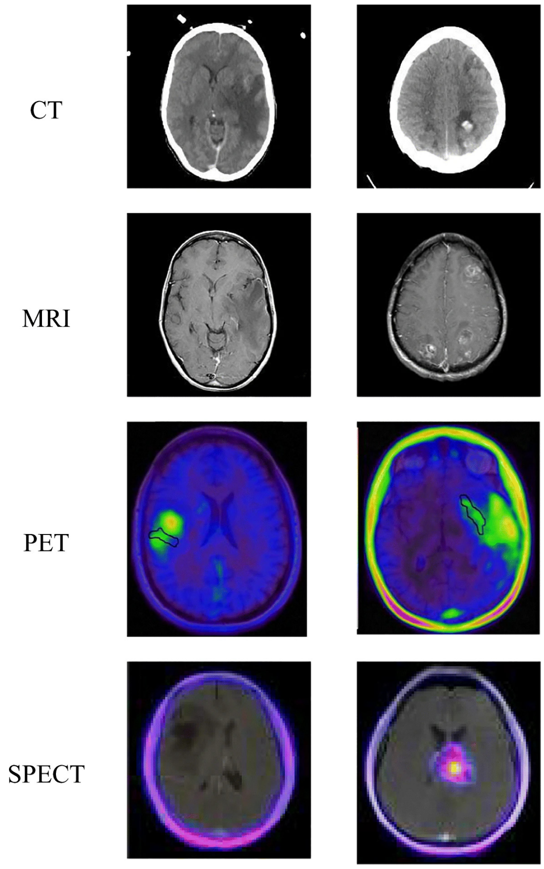


Figure 2.2 – Modalities of Medical Images

In this internship, all the images are extracted from The Whole Brain Atlas (TWBA) dataset, a multimodal medical image resource of the central nervous system. Thanks to being freely accessible, public, and available online at `https://www.med.harvard.edu/aanlib/home.html`, The Whole Brain Atlas (TWBA) has been used in many medical imaging researches. There exists several multimodal medical datasets, such as Alzheimer's Disease Neuroimaging Initiative (ADNI), IXI Dataset, Multimodal Brain Tumour Segmentation Challenge (BraTS) but those need permission or have limited access. Consequently, we decided to use The Whole Brain Atlas dataset thanks to its free and online access.

This dataset was developed by Keith Johnson, MD, and Alex Becker, PhD [24]. It is also supported by the Brigham and Women's Hospital Departments of Radiology and Neurology, Harvard Medical School, the Countway Library of Medicine, and the American Academy of Neurology.

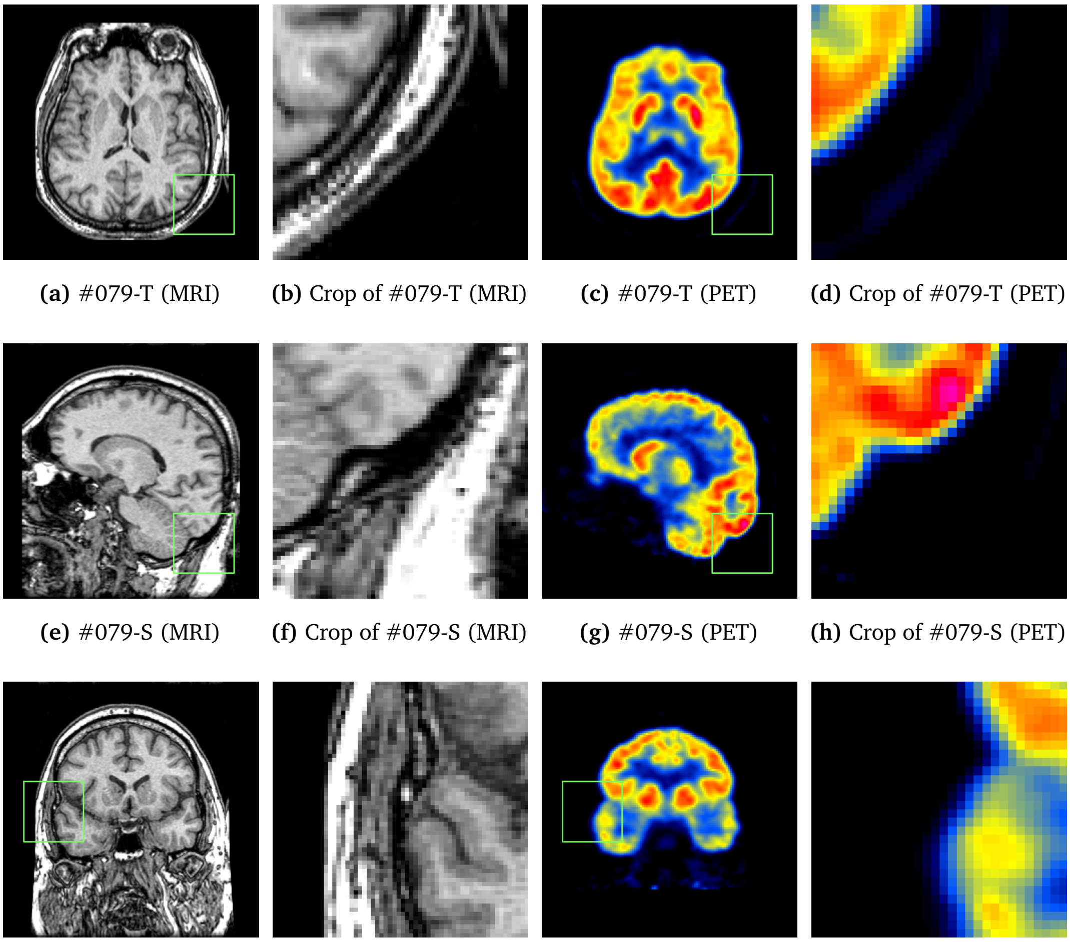
The Whole Brain Atlas (TWBA) has different numbers and types of multimodal images of a brain of one single patient. The site has six main sections based on the diseases that the patients have: Normal Brain, Cerebrovascular, Neoplastic, Degenerative, Inflammatory or Infectious.

Within the scope of this internship, there are 180 images in total extracted from the dataset which will be used, including 90 pairs MRI-PET images. These pairs are divided into 3 sets regarding the capture axis of human brain: transaxial (Set-T), corronal (Set-C) and sagittal (Set-S). Originally each set contains over 120 MRI-PET images, but here we merely extract 30 pairs numbering from 050 to 079 per set.

The following Table 2.1 presents the list of available multimodal images in the TWBA dataset, while Fig.2.3 demonstrates pairs of MRI-PET images and their cropped region extracted from TWBA dataset.

Table 2.1 – Types of available modality in TWBA.

MR: Proton Magnetic Resonance Image
MR-T1: Proton Magnetic Resonance Image, T1-weighted
MR-T2: Proton Magnetic Resonance Image, T2-weighted
MR-PD: Proton Density Weighted
MR-diff: Proton Magnetic Resonance Image diffusion weighted
MR-GAD: T1-Weighted, after Gd-DTPA
MR-CBV: Cerebral blood volume map
CT: Roentgen-ray Computed Tomography, without Iodinated Contrast
CT-i+: Roentgen-ray Computed Tomography, with Iodinated Contrast
SPECT-Tc: Single Photon Emission Computed Tomography with Tc99m-HM-PAO
SPECT-Tl: Single Photon Emission Computed Tomography with Thallium-201
PET-FDG: Positron Emission Tomography with Fluorine-18 Deoxyglucose



(a) #079-T (MRI) **(b)** Crop of #079-T (MRI) **(c)** #079-T (PET) **(d)** Crop of #079-T (PET)

(e) #079-S (MRI) **(f)** Crop of #079-S (MRI) **(g)** #079-S (PET) **(h)** Crop of #079-S (PET)

(i) #079-C (MRI) **(j)** Crop of #079-C (MRI) **(k)** #079-C (PET) **(l)** Crop of #079-C (PET)

Figure 2.3 – Examples of some pairs of MRI-PET images from TWBA dataset

2.2 Image Enhancement Methods

Image enhancement, a crucial sub-field of image processing, has significant practical applications. This area focuses mainly on improving the visual quality and content of images, making them more suitable for subsequent processing tasks such as analysis, detection, segmentation, and recognition [23]. The techniques employed in image enhancement address a wide range of issues, including deblurring, balancing contrast, adjusting brightness and shadows, eliminating noise, and uncovering hidden details. By tackling these challenges, image enhancement methods play a vital role in preparing images for more advanced processing and interpretation.

Following an extensive review of a large amount of literature in the field of image enhancement, Qi *et al.* [19] demonstrated that there is a wide array of techniques have been developed to enhance quality of images as the figure below.

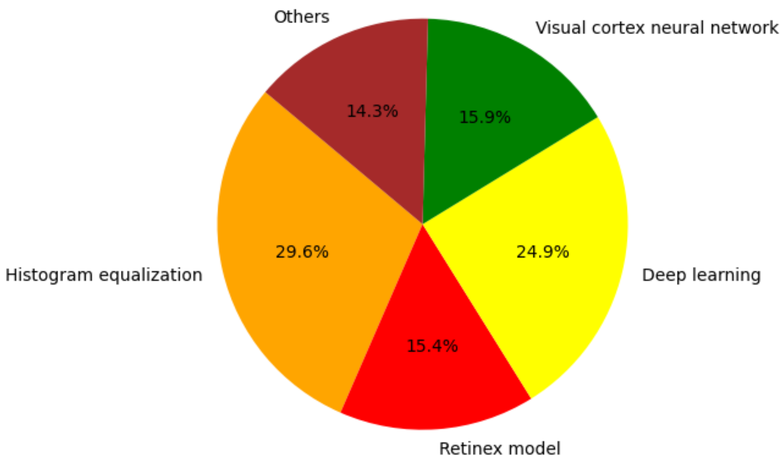


Figure 2.4 – Statistics of the number of literatures investigating different image enhancement techniques

For this internship, our study mainly focused on two specific approaches within the realm of image enhancement: histogram equalization and deep learning techniques. Specifically, we employed Contrast-limited Adaptive Histogram Equalization (CLAHE) to enhance image contrast and brightness. Additionally, we utilized Denoising Convolutional Neural Network (DnCNN) to address noise issues introduced by imaging sensors. The subsequent sections of this chapter will provide in-depth explanations of these two chosen methods and their applications.

Moreover, image enhancement techniques also consists of image fusion, which is defined as a process of combining relevant information from multiple images into a single final image, due to the evidence that there are theoretical and technical limitations of hardware devices that prevent them from capturing effectively and comprehensively the scene [31]. Subsequently, the resultant fused image contains more information and becomes more comprehensive than any of the input images [20].

2.2.1 CLAHE

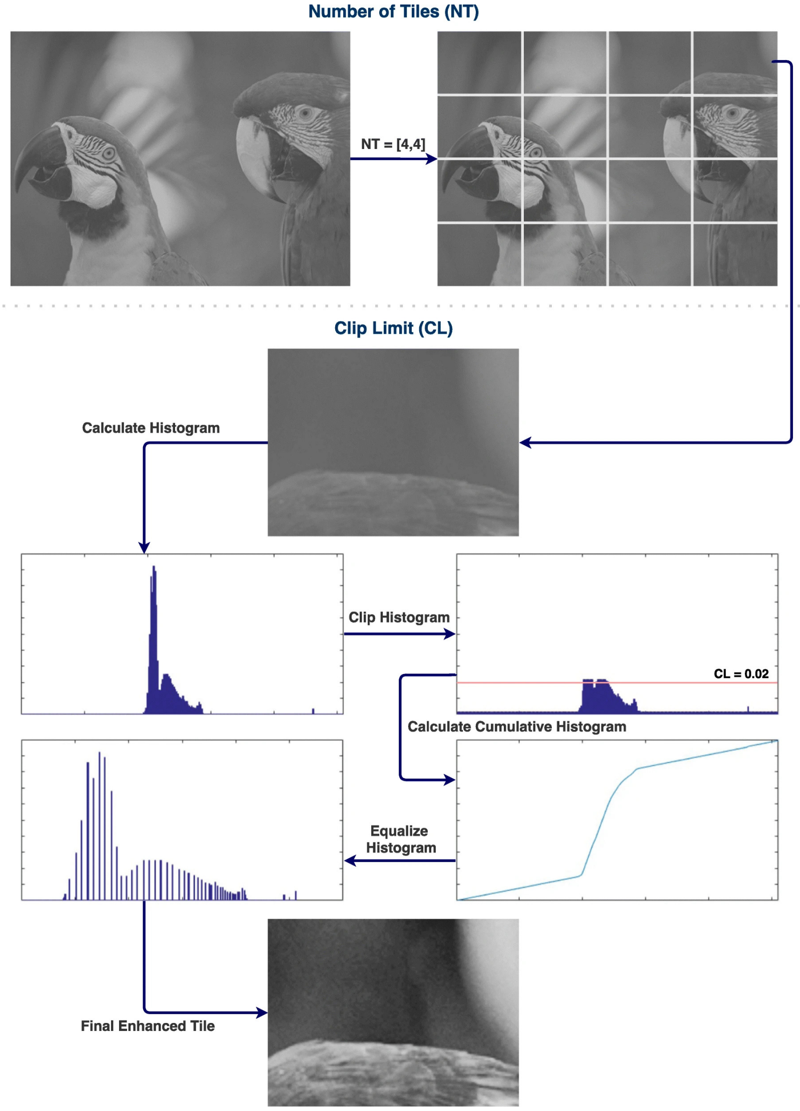


Figure 2.5 – Process of CLAHE

The Contrast-limited Adaptive Histogram Equalization (CLAHE) [34] is a well-known technique for local contrast improvement, was developed from the Adaptive Histogram Equalization (AHE), has been proving its power in several applications. The basic idea of CLAHE is to carry out the histogram equalization technique on non-overlapping sub-regions of the image, then bilinear interpolation is performed to correct inconsistencies between borders. There are two significant parameter in CLAHE comprising the clip limit (CL) and the number of tiles (NT). The former is a numeric value controlling the noise amplification. When the histogram of each sub-region of the image is computed, it is redistributed provided that its height does not exceed the predefined clip limit. Afterwards, the cumulative histogram is calculated and used for equalization. The former is an integer value which determines the number of non-overlapping sub-regions. Based on this value, the image is separated into different (usually squared-shape) non-overlapping areas with the same size [3]. Fig.2.5 illustrates the processing steps of CLAHE.

2.2.2 DnCNN

When it comes to image denoising, its objective is to recover a clean image x from the observation noisy signal y: $y = x + v$ with v is the additive noise [32]. Several former proposed techniques follow the same approach: trying to predict the clean signal x from the noisy one. However, rather than learning a discriminative model with an explicit image prior like others, Zhang *et al.* treat the image denoising problem as a plain discriminative learning problem of seperating the noise from the noisy signal by introducing the Denoising Convolutional Neural Network (DnCNN) model. More specifically, instead of predicting the denoised image x, this novel approach tries to estimate the additive noisy image v. In other words, the proposed model removes the latent clean signal with the operations of hidden layers in an implicit way. The aforementioned model is proved that its performance outstrips earlier models such as BM3D[5] and TNRD[4].

The structure of the network was modified from the VGG network [22] for higher suitability with the task of image denoising. The size of the convolution filters is set to be 3×3 and all pooling layers has been eliminated. As in Fig.2.6, there are 3 different kinds of layer designed for the model. (1) Conv+ReLU: the first hidden layer, comprises 64 kernels size $3 \times 3 \times c$ to create 64 feature maps, prior to passing through the ReLU activation function, with c is the number of image's channels. (2) Conv+BN+ReLU: for the next 18 hidden layers, 64 masks size $3 \times 3 \times 64$ is applied and batch normalization is added between the convolution and ReLU. (3) Conv: the last hidden layer, it uses c kernels with size $3 \times 3 \times 64$ to reconstruct the output. Table.A.1 describes more details of layers.

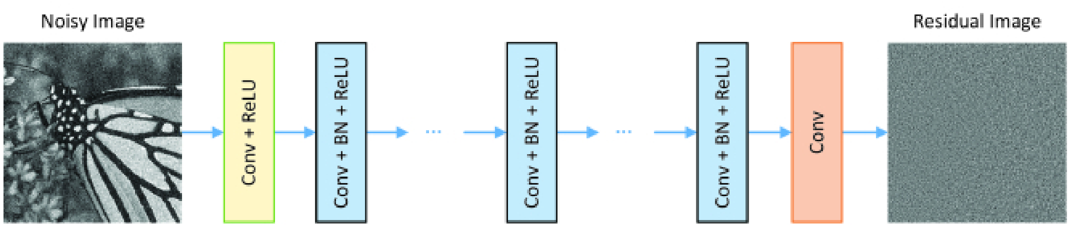


Figure 2.6 – The overall structure of DnCNN

2.2.3 IFCNN

Yu Zhang *et al.* [31] introduced this fusion framework in 2021, called General Image Fusion Framework based on Convolutional Neural Network, which was inspired by the transform-domain image fusion algorithms. Its architecture is based on convolutional neural network as shown in Fig.2.7: it consists of three major modules: feature extraction module, feature fusion module and feature reconstruction module.

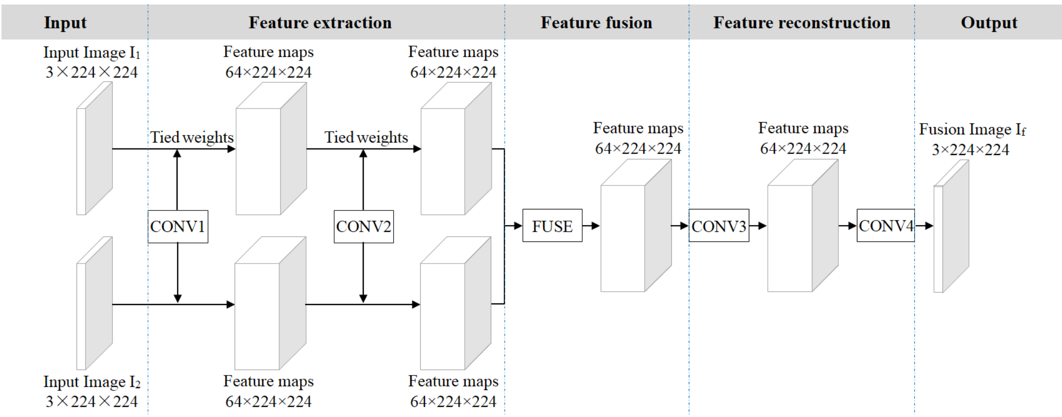


Figure 2.7 – IFCNN Architecture

The first one including 2 convolutional layers is utilized for extracting the extensive low-level features from input images. The layer CONV1 contains 64 kernels size 7×7 with stride 1 and padding 3, which was adopted from the first convolutional layer of the pretrained network ResNet101. Nevertheless, since the CONV1 was trained for classification task, the CONV2 convolutional layer of 64 3 × 3 kernels of stride and padding 1 has been added in order to make it more suitable for feature fusion. Moving onto the second module, the convolutional features of inputs after passing through the CONV2 layer are initially concatenated along the channel dimension. Afterwards, the concatenated features are fused by the element-wise fusion rule (element-wise maximum, element-wise mean, element-wise sum). Within the scope of this research is about medical image, the element-wise maximum fusion rule is utilized. Finally, because of low abstraction level of extracted features after the first 2 convolutional layers, 2 other layers namely CONV3 and CONV4 has been adopted to reconstruct the fusion image from fused convolution features. The CONV3 layers plays a role of tuning the fused convolutional features, its parameters remain the same as CONV2, while CONV4 has 3 1 × 1 kernels with no stride and padding. Furthermore, due to overfitting problem and training performance, layers CONV2 and CONV3 have been equipped with ReLU activation and batch normalization.

2.3 Optimization Methods

Mathematical optimization, also known as mathematical programming, is the name of the process that discovering the best solution within a feasible space. This involves formulating problems where an objective function needs to be minimized or maximized, subject to a set of constraints. The optimization problems arise in various aspects such as engineering, economics, computer science, etc. An example of finding optimal values of a function is illustrated in Fig.2.8.

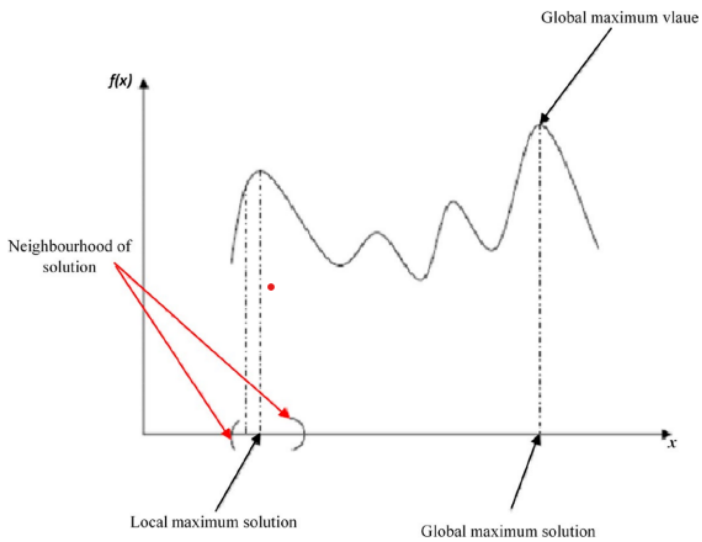


Figure 2.8 – An Optimization Problem

Mathematical optimization techniques can be broadly classified into two major categories: deterministic and stochastic. In the deterministic approach, linear and non-linear programming methods are widely employed. These methods rely on gradient information to find the solution. While deterministic techniques are effective for problems with linear search spaces (unimodal), they often struggle when applied to non-linear search space problems. In such cases, they can easily become trapped in local optima, failing to find the global optimal solution.

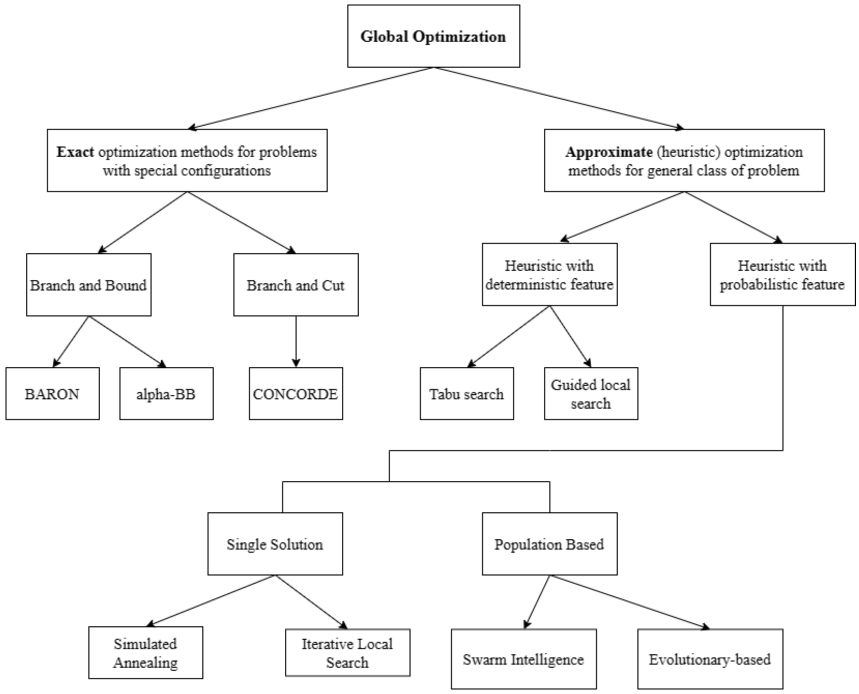


Figure 2.9 – Classification of Optimization Algorithms

Stochastic methods, or estimate methods in other words, is developed as an alternative approach to traditional conventional methods. These algorithms incorporate the generation and utilization of random variables. Their primary purpose is to perform a comprehensive global search across the entire domain, aiming to identify either the globally optimal solution or a near-optimal solution. Meta-heuristic techniques, which fall under the category of stochastic methods, offer several advantages, including simplicity in implementation, independence from the specific problem formulation, flexibility in application, and the ability to operate without requiring gradient information. Meta-heuristics Algorithms would be categorized based on four major sources of inspiration: evolutionary, swarm intelligence, physics-based and human-based methods as demonstrated in Fig. 2.10.

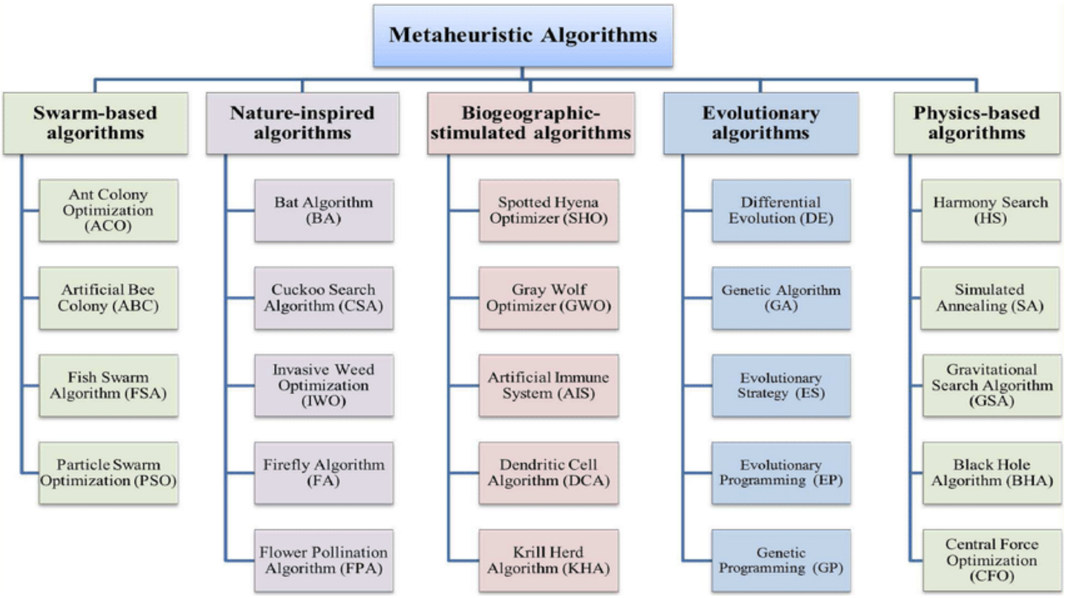


Figure 2.10 – Meta-heuristic Algorithm Classification

Meta-heuristic optimization algorithms possess two crucial features: (1) exploration and (2) exploitation. Exploration refers to the ability to search globally the solution within predefined the space. This capability helps the algorithm escape local optima and prevents it from getting stuck in sub-optimal regions, or it is called local optimal points. On the other hand, exploitation is the ability to search locally around promising solutions in an attempt to refine and improve their quality. Achieving good performance in meta-heuristic algorithms requires a suitable balance between these two features. While exploration allows the algorithm to investigate diverse regions of the search space, exploitation ensures that high-quality solutions are not overlooked by concentrating the search effort in promising areas. The following subsections are reserved for introducing several meta-heuristic algorithms that we applied for our proposed method.

2.3.1 War Strategy Optimization

The War Strategy Algorithm (WSO) was a new meta-heuristic algorithm, which was proposed by T. Ayyarao *et al.* [2]. Its inspiration is based on the strategic movement of army troops during the war. War strategy is modeled as an optimization process wherein each soldier dynamically moves towards the optimum value.The proposed algorithm models two popular war strategies, attack and defense strategies. The positions of soldiers on the battlefield are updated in accordance with the strategy implemented. To improve the algorithms convergence and robustness, a novel weight updating mechanism and a weak soldiers relocation strategy are introduced.

At each iteration, all the soldiers (search agents) in the searching space have the same likeliness to become the King or the Commander depending on their fitness value. The King and the Commander play as leaders in the space, and their movements will guide other soldiers. There exists a probability that the King or Commander might face stiff competition from opponent's soldiers (Local Optima), it means the King/Commander is trapped by local optimum points. As a result, soldiers in the war is guided not only by the King or the Commander, but also by the combination of their movement tactics.

Attack Strategy: In this case, every soldiers updates his position based on the position of the King and the Commander:

$$X_i(t+1) = X_i(t) + 2*r*(C-K) + rand*(W_i*K - X_i(t)) \tag{2.1}$$

where $X_i(t+1)$ is the new position, $X_i(t)$ is the old position, C and K are the Commander and King's position correspondingly, W_i denotes the weight and $rand$ signifies a random vector in range $[0,1]$ with uniform distribution.

Rank and Weight Updating: The rank of each soldier (search agent) is based on his successful history in the war field (search space). This rank reflects how close this soldier is compared to the target (fitness value). In the case that the attack force (fitness value) of the new position (F_n) is smaller than that of the previous position (F_p), the soldier takes the previous position, else, it updates the new position.

$$X_i(t+1) = (X_i(t+1)*(F_n > F_p)) + (X_i(t)*(F_n < F_p)) \tag{2.2}$$

If the soldier updates new position successfully, his rank R_i will be upgraded:

$$R_i = (R_i + 1)*(F_n \leq F_p) + R_i*(F_n < F_p) \tag{2.3}$$

The operation $F_n < F_p$ will return 1 if the value comparison is True and 0 if False. It is the same for other comparison operations in Eq.(2.2) and Eq.(2.3).

The new weight of each soldier will be calculated base on his rank:

$$W_i = W_i \times \left(1 - \frac{R_i}{\text{Max\_Iter}}\right)^2 \tag{2.4}$$

Defense Strategy: The second strategy used for position updating depends on the King's position, the army head and a random soldier, whereas the ranking and weight updating rule remain the same:

$$X_i(t+1) = X_i(t) + 2 * \rho * (K - X_{rand}(t)) + rand * W_i * (C - X_i(t)) \tag{2.5}$$

Replacing/Relocating Weak Soldiers: In each iteration, the weakest soldiers with the worst fitness values will be identified, and they will be relocated randomly in the search space:

$$X_w(t+1) = Lb + rand * (Ub - Lb) \tag{2.6}$$

Another approach may be utilized for relocating weak soldiers:

$$X_w(t+1) = -(1 - rand) * (X_w(t) - median(X)) + K \tag{2.7}$$

This updating rule relocates the weakest soldiers around the median of entire army in the war field. Subsequently, it enhances the convergence ability of the proposed algorithm. The figure below presents the workflow of the WSO algorithm.

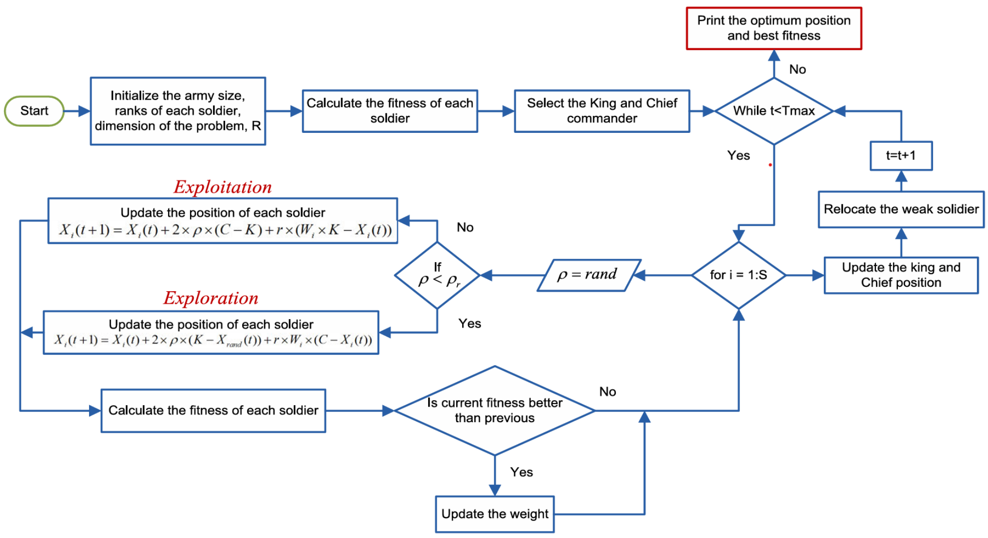


Figure 2.11 – The flowchart of WSO

2.3.2 Marine Predator Algorithm

Marine Predator Algorithm (MPA) is a nature-inspired optimization algorithm, introduced by A. Faramarzi *et al.* [9], which follows the rules that naturally govern in optimal foraging strategy and encounter rate policy between predator and prey in marine ecosystems.

Lévy strategy is a widespread pattern amongst marine predators (e.g. sharks, tunas) when searching for the food in a prey-sparse environment, but when they begin foraging in a prey-abundant area, their moving pattern is commonly switched to Brownian motion [12]. The optimal encounter rate policy in biological interaction between predator and prey is also dependent on the type of the movement that each of the predator/prey is taking and the velocity ratio of prey to predator.

The following points below summarize the governing policies for the optimal foraging, interactions and memories in marine predators:

- Marine predators use Lévy strategy for the environment with a low concentration of prey whilst Brownian movement is employed for the areas with abundant prey.
- Predators show the same proportion of Lévy and Brownian movement during their lifetime of traversing different habitats.

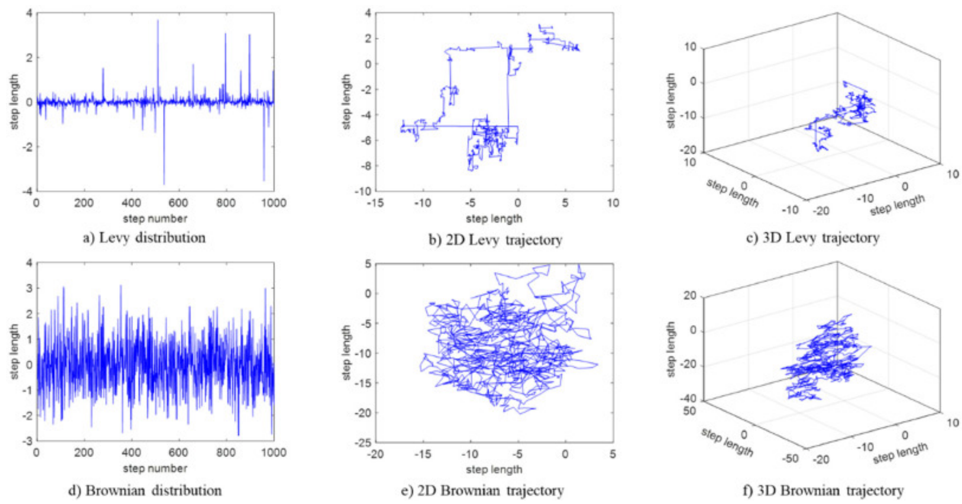


Figure 2.12 – Lévy flight and Brownian flight

- Due to environmental effects such as natural (Eddy Formation) or human-caused (FADs), their behavior is modified in order to hopefully find regions with a different distribution of prey.
- **In low-velocity ratio (v=0.1)**: the best strategy for a predator is Lévy; either prey is moving in Brownian or Lévy model.
- **In the unit velocity ratio (v=1)**: if prey moves in Lévy, the best strategy for predator is Brownian.

- **In high-velocity ratio (v>10)**: the best strategy for a predator is not moving at all. In this case, either prey is moving Brownian or Lévy

- Predators take advantage of good memory in reminding of their associates as well as location of successful foraging.

Based on those highlighted points, MPA optimization process is divided into three main phases of optimization considering different velocity ratio and at the same time simulating the entire life of a predator and prey: (1) in high velocity ratio or when prey is moving faster than predator, (2) in unit velocity ratio or when both predator and prey are moving at almost same pace, and (3) in low velocity ratio when predator is moving faster than prey.

Resembling most of meta-heuristic algorithms, MPA is a population based method, in which the initial solution is uniformly distributed over the search space as the first trial:

$$X_0 = X_{\min} + rand \times (X_{\max} - X_{\min}) \tag{2.8}$$

where X_{max} and X_{min} are respectively the upper and lower bound and *rand* is a uniform random vector in range [0,1].
After randomly initializing the population, the fitness score of each individual will be computed based on the objective function, then the one has the highest fitness level will be considered as a top predator to construct a matrix called *Elite*.

$$Elite = \begin{bmatrix} X^I_{1,1} & X^I_{1,2} & \cdots & X^I_{1,d} \\ X^I_{2,1} & X^I_{2,2} & \cdots & X^I_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ X^I_{n,1} & X^I_{n,2} & \cdots & X^I_{n,d} \end{bmatrix} \tag{2.9}$$

where $\vec{X}^L$ is the top predator vector then is replicated n times to create the *Elite* matrix. n is the number of search agents while d is the number of dimensions. At the end of each iteration, the Elite will be updated if the top predator is substituted by the better predator. The *Prey* matrix is also constructed which has the same dimension with *Elite*, based on this matrix the predators would update their position.

$$Prey = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n,1} & X_{n,2} & \cdots & X_{n,d} \end{bmatrix} \tag{2.10}$$

In equation (2.10), $X_{i,j}$ represents the jth dimension of ith prey.

MPA optimization process comprises three main phases, each phase considers different velocity ratio: (1) in high velocity ratio or when prey is moving faster than predator, (2) in unit velocity ratio or when both predator and prey are moving at almost same pace, and (3) in low velocity ratio when predator is moving faster than prey.

Phase 1: in high velocity ratio or when predator is moving faster than prey.

While $Iter < \frac{1}{3}Max\_Iter$ do:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_B} \otimes \left(\overrightarrow{Elite_i} - \overrightarrow{R_B} \otimes \overrightarrow{Prey_i}\right), \quad i = 1, \ldots, n \tag{2.11}$$

$$\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P \times \overrightarrow{R} \otimes \overrightarrow{stepsize_i} \tag{2.12}$$

where: $\overrightarrow{R_b}$ denotes a vector comprises numbers based on normal distribution representing the Brownian movement and the symbol $\otimes$ presents the element-wise multiplication, $P = 0.5$, $\overrightarrow{R}$ is a random vector of uniform distribution in range $[0, 1]$.

Phase 2: In unit velocity ratio

While $\frac{1}{3}Max\_Iter < Iter < \frac{2}{3}Max\_Iter$ do:

For the first half of the population:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_L} \otimes \left(\overrightarrow{Elite_i} - \overrightarrow{R_L} \otimes \vec{Prey_i}\right), \quad i = 1, 2, \ldots, n/2 \tag{2.13}$$

$$\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P \times \overrightarrow{R} \otimes \overrightarrow{stepsize_i} \tag{2.14}$$

with $\overrightarrow{R_L}$ represents a random vector based on Lévy flight. For the second half of the population:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_B} \otimes \left(\overrightarrow{R_B} \otimes \overrightarrow{Elite_i} - \overrightarrow{Prey_i}\right), \quad i = \frac{n}{2} + 1, \ldots, n \tag{2.15}$$

$$\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P \times CF \otimes \overrightarrow{stepsize_i} \tag{2.16}$$

where $\mathrm{CF} = \left(1 - \frac{Iter}{Max\_Iter}\right)^{2\frac{Iter}{Max\_Iter}}$ is an adaptive parameter used for controlling the step size for predator movement.

Phase 3: In low-velocity ratio or when predator is moving faster than prey.

While $Iter > \frac{2}{3}Max\_Iter$ do:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_L} \otimes \left(\overrightarrow{R_L} \otimes \overrightarrow{Elite_i} - \overrightarrow{Prey_i}\right), \quad i = 1, \ldots, n \tag{2.17}$$

$$\overrightarrow{Prey_i} = \overrightarrow{Elite_i} + P \times CF \otimes \overrightarrow{stepsize_i} \tag{2.18}$$

Another point which causes a behavioral change in marine predators is environmental issues such as the Eddy Formation or Fish Aggregating Devices (FADs) effects. The FADs are considered as local optima and their effect as trapping in these points in search space. Consideration of these longer jumps during simulation avoids stagnation in local optima. Thus, the FADs effect is mathematically presented as:

$$\overrightarrow{Prey_i} = \begin{cases} \overrightarrow{Prey_i} + CF \times [\overrightarrow{X}_{\min} + \overrightarrow{R} \otimes (\overrightarrow{X}_{\max} - \overrightarrow{X}_{\min)}] \otimes \overrightarrow{U} & \text{if } r \leq FADs, \\ \overrightarrow{Prey_i} + [FADs \times (1-r) + r] \times (\overrightarrow{Prey_{r1}} - \overrightarrow{Prey_{r2}}) & \text{if } r > FADs \end{cases} \tag{2.19}$$

where $FADs = 0.2$, $\overrightarrow{U}$ is a binary vector with arrays containing 0 and 1. It is constructed by creating a random vector in range 0 and 1, then modified its entries to 0 if the generated value is less than 0.2, else it changes to 1. r is a random number between 0 and 1. $\overrightarrow{X}_{\min}$ and $\overrightarrow{X}_{\max}$ are the vectors of lower and upper bound of the problem. Finally, $r1$ and r_2 subscripts denotes 2 random indices of the $Prey$ matrix.

The pseudo code for the MPA optimization process is presented as in Algorithm 1

2.3.3 Comprehensive Learning Marine Predator Algorithm

This approach is developed by D. Yousri *et al.* [30], based on the original Marine Predator Algorithm (MPA), which incorporates the Comprehensive Learning (CL) strategy in order to avoiding the premature convergence of the primary algorithm by sharing the best experiences amongst other searching agents [30]. This strategy would be able to improve and control better the search ability during the optimization process, in which searching elements at the current position would learn from the previous best locations of all other agents in the searching space. The mathematical representation of the original MPA remains almost the same, except for three equations of finding the $\overrightarrow{stepsize}$ in Phase 1 and Phase 2 has been reformulate as follow:

Algorithm 1: Pseudo code for the MPA process

Initialize search agents (Prey) population $i = 1, 2, \ldots, N$
while *termination criteria are not responded* **do**
  Calculate the fitness, construct the $Elite$ matrix and accomplish memory saving
  if $Iter < Max\_Iter/3$ **then**
    Update $Prey$ based on Eq.(2.12);
  end
  if $Max\_Iter/3 < Iter < 2Max\_Iter/3$ **then**
    For the first half of the population ($i = 1, 2, \ldots, n/2$);
    Update $Prey$ based on Eq.(2.14);
    For the second half of the population ($i = n/2 + 1, \ldots, n$);
    Update $Prey$ based on Eq.(2.16);
  end
  if *2Max_Iter/3 < Iter* **then**
    Update $Prey$ based on Eq.2.18
  end
  Accomplish memory saving and $Elite$ update;
  Applying FADs effects and update based on Eq.(2.19);
end

- Eq.(2.12) becomes:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_B} \otimes \left(\overrightarrow{Elite_i} - \overrightarrow{R_B} \otimes \overrightarrow{Prey_i} + K \times rand \otimes (Prey_{fi} - Prey_i)\right) \tag{2.20}$$

$$\text{with } K = \left(3 - 1.5 \times \frac{Iter}{Max\_Iter}\right) \tag{2.21}$$

- Eq.(2.14) has changed to:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_L} \otimes \left(\overrightarrow{Elite_i} - \overrightarrow{R_L} \otimes \overrightarrow{Prey_i} + c_1 \times rand \otimes (Prey_{fi} - Prey_i)\right) \tag{2.22}$$

$$\text{with } c_1 = \frac{1}{1 + \exp(-10^{-4} \times Iter/Max\_Iter)} + 2 \times \left(\frac{Iter}{Max\_Iter} - 1\right)^2 \tag{2.23}$$

- Eq.(2.16) has been rewritten as:

$$\overrightarrow{stepsize_i} = \overrightarrow{R_B} \otimes \left(\overrightarrow{R_B} \otimes \overrightarrow{Elite_i} - \overrightarrow{Prey_i} + c_2 \times rand \otimes (Prey_{fi} - Prey_i)\right) \tag{2.24}$$

$$c_2 = \frac{1}{1 + \exp(-10^{-4} \times Iter/Max\_Iter)} + \frac{2 \times Iter^2}{Max\_Iter^2} \tag{2.25}$$

In those defined equations above, the term $Prey_{fi}$ is an exemplar learning from the best locations of other particles, and this term is found using the following algorithm:

Algorithm 2: Pseudo code for the CL strategy

for $i = 1, 2, \ldots, N$ **do**
    Generate random number r
    Calculate the learning probability value $P_{cl,\ i}$
    Choose 2 random particles with indices $f1, f2$
    if $P_{cl,i} > r$ **then**
        if $fobj(U_f 1) < fobj(U_f 2)$ **then**
            exemplar $= Prey_{f1}$
        else
            exemplar $= Prey_{f2}$
    else
        exemplar $= Prey_i$
    End

2.4 Proposed Method

2.4.1 Enhancement Component

This enhancement component, which uses the combination of CLAHE, DnCNN, LED and optimization, is proposed based on a recent study of P. H. Dinh [7]. The schematic visualization in Fig.2.13 illustrates its workflow:

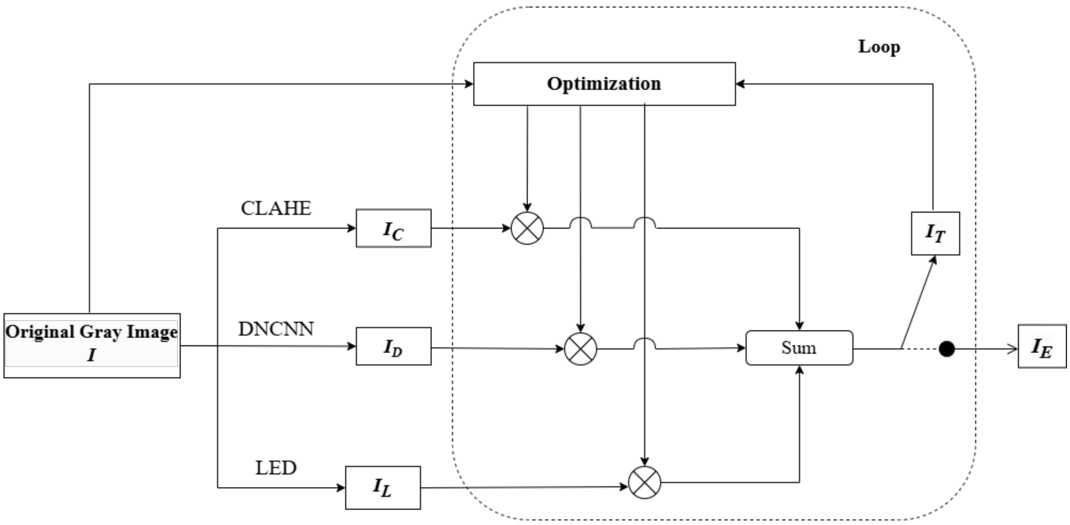


Figure 2.13 – Enhancement Component

The objective function for the optimization problem is defined as:

$$Obj\_Func = \frac{V_T}{M_T} \times \left((E - E_T) + \frac{G - G_T}{PSNR(I_T, I)} \right) \tag{2.26}$$

where:

- I is the original image with size $H \times W$
- I_T is the result image at each iteration of optimization process, having the same size with I, and $I_T = \beta_1 \times CLAHE + \beta_2 \times DnCNN + \beta_3 \times LED$
- V_T and M_T denote the variance and mean of image I_T, in sequence
- E and E_T are entropy of I and I_T, respectively
- G and G_T represent the average gradient of image I and I_T, sequentially
- $PSRN(I_T, I) = 20 \times \log_{10} \frac{MaxPixel}{MSE(I_T,I))}$ with $MaxPixel$ is the maximum gray level of image I, and
  $MSE(I_T, I) = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} (I_T(i,j) - I(i,j))^2$

The goal of the optimization algorithm is to find the optimal value for $\beta_1 \in [0.1, 0.5]$, $\beta_2 \in [0.5, 0.9]$ and $\beta_3 \in [0.2, 0.3]$ provided that the Obj_Func achieves the minimum value. The algorithms that we use for our research is described formerly in Chapter 2.

The below algorithm also demonstrates the steps of the Enhancement Component.

Algorithm 3: The Enhancement Component

Input : A gray image I size $H \times W$
Output : Enhanced Image I_E

1 **Step 1:** Using CLAHE algorithm to generate the first component image
2 **Step 2**: Using DnCNN to generate the second component image
3 **Step 3:** Using Laplacian Edge Detector (LED) to generate the third component image
4 **Step 4:** Enhanced Image I_T is generated using Optimization algorithm to optimize the Eq.(2.26)
5 **Step 5:** With optimal value β_1^*, β_2^* and β_3^*, create the final enhanced image: $I_E = \beta_1^* \times CLAHE + \beta_2^* \times DnCNN + \beta_3^* \times LED$

2.4.2 Proposed Pipeline

The enhancement component mentioned in the former section is integrated into our proposed pipeline as the schematics diagram below:

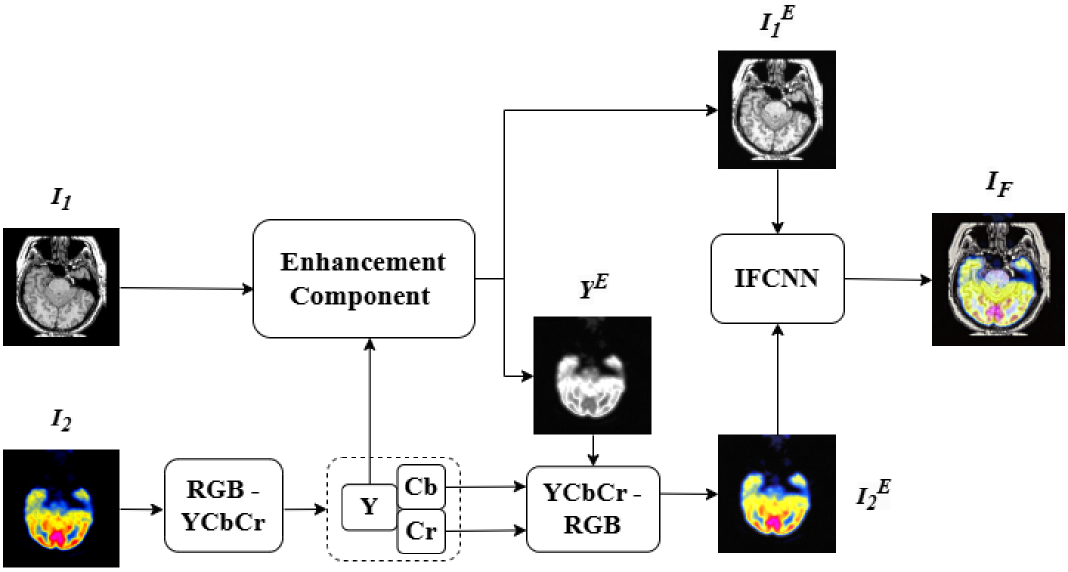


Figure 2.14 – Pipeline of our studied method

Our proposed pipeline processes two types of images: an MRI image and a color PET image. While P.H. Dinh's work only enhanced MRI images [7], our approach extends the enhancement to both MRI and PET images. For the PET image, which is initially in RGB color space, is firstly converted into the YCbCr space. The Y component, representing brightness, undergoes enhancement to improve its quality. After enhancement, this Y component is recombined with the Cb and Cr components and converted back to RGB space.

Both the enhanced MRI image and the processed PET image are then fed into the IFCNN model. This model generates a fused output image that combines information from both enhanced input images.

Algorithm 4: The Proposed Method

Input : Gray MRI image I_1 and color PET image I_2 size 256×256
Output : A color fused image I_F

1 **Step 1:** Convert the color PET image I_2 from RGB to YCbCr color space, subsequently obtain 3 separated component of Y, Cb and Cr
2 **Step 2**: Pass the image I_1 and Y component from Step 1 through the Enhancement Component, obtain 2 enhanced image I_1^E, Y^E: $I_1^E = EnhancementComp(I_1)$, $Y_E = EnhancementComp(Y)$
3 **Step 3:** Merge the gray enhanced component Y_E with Cb and Cr component from step 1, then convert the color image to RGB called I_2^E
4 **Step 4:** Using IFCNN to fused enhanced images I_1^E, I_2^E in order to create fusion image I_F: $I_F = IFCNN(I_1^E, I_2^E)$

Chapter 3

Evaluation & Results

This chapter is preserved for describing details of evaluation metrics that we employ and presenting our results of the research, also with the discussion.

3.1 Evaluation Metrics

The quality index for evaluating the enhanced image and fused image would be conducted in two different approaches [17]: reference and no-reference. The former is to compare the enhance/fusion result image with a known reference image, yet this ground truth is not available all the time. As a result, the latter implementation, is generally preferred to substitute the reference approach when the ground truth image does not exist. In this research, both kinds of image quality assessment is applied for fully evaluating the final image.

3.1.1 Entropy

Entropy of an image is a metric utilizing for evaluating the information brought by an image. Its formula is defined based on the Shannon's formula of information entropy [21]:

$$EN = -\sum_{i=0}^{n-1} p_i \log_2(p_i) \tag{3.1}$$

where n is the number of gray levels, p_i is the probability of gray level i in the image. p_i is calculated by taking ratio of the number of pixels with gray level i and the overall pixels number of image: $p_i = \frac{n_i}{h \times w}$, in which n_i is the number of pixels with gray level i,

h and w are the height and width of the image, respectively. The larger the entropy, the more information obtained from the image.

3.1.2 Average Gradient

The average gradient of the image is computed by taking the mean value of the gradient magnitude of all the pixels in the image:

$$AG = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} G_{i,j} \tag{3.2}$$

with H and W are height and width of the image, $G_{i,j} = \sqrt{(g_x(i,j))^2 + (g_y(i,j))^2}$ is the gradient magnitude of pixel location i, j in the image. In this research, a simple way is used for obtaining the directional gradient g_x and g_y:

$$g_x(i,j) = I(i+1,j) - I(i,j) \tag{3.3}$$

$$g_y(i,j) = I(i,j+1) - I(i,j) \tag{3.4}$$

The higher the average gradient, the clearer image since edges in this image are sharper.

3.1.3 Mean Light Intensity (MLI)

This evaluation metric computes the average gray level of the image:

$$MLI = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} I_{i,j} \tag{3.5}$$

where H and W are height and width of the image, $I(i,j)$ denotes the gray intensity of pixel at location i, j.

3.1.4 Contrast Index

$$CI = \sqrt{\frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} I^2(i,j) - \left(\frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} I(i,j) \right)^2} \tag{3.6}$$

with H and W are height and width of the image, $I(i,j)$ is the gray intensity of pixel at location i, j.

3.1.5 Structural Similarity Index (SSIM)

The image similarity measurement is developed dependent on the assumption that our visual system is highly adapted to structural information. Consequently, a measurement for the loss of structural information would present a good approximation of the perceived image distortion. As a result, Wang *et al.* proposed the Structural Similarity Index in order to compare the structure of the reference image and the distorted image [26], and it was proved being better than the Peak Signal to Noise Ratio (PSNR) or Mean Squared Error (MSE) index. Its formula is denoted as:

$$\begin{aligned} SSIM &= [l(A,B)]^{\alpha} \times [c(A,B)]^{\beta} \times [s(A,B)]^{\gamma} \\ &= \left(\frac{2\mu_A\mu_B + C_1}{\mu_A^2 + \mu_B^2 + C_1}\right)^{\alpha} \left(\frac{2\sigma_A\sigma_B + C_2}{\sigma_A^2 + \sigma_B^2 + C_2}\right)^{\beta} \left(\frac{\sigma_{AB} + C_3}{\sigma_A\sigma_B + C_3}\right)^{\gamma} \end{aligned} \tag{3.7}$$

where $l(A,B)$, $c(A,B)$ and $S(A,B)$ components represent the luminance, contrast and correlation, respectively. To adjust the contribution of each in 3 components, 3 parameter α, β and γ is utilized. Meanwhile, μ_A and μ_B are the average gray value of image A and B, sequentially, σ_A, σ_B denote the variance, σ_{AB} is the covariance. The usage of constant values C_1, C_2 and C_3 is to prevent the instability when the denominator is very close to zero. With $\alpha = \beta = \gamma = 1$ and $C_3 = C_2/2$, Eq.(3.7) becomes:

$$SSIM = \frac{(2\mu_A\mu_B + C_1)(2\sigma_{AB} + C_2)}{(\mu_A^2 + \mu_B^2 + C_1)(\sigma_A^2 + \sigma_B^2 + C_2)} \tag{3.8}$$

3.1.6 Edge-based Similarity ($Q^{AB/F}$)

Xydeas and Petrovic [27] proposed this metric, which measures the relative amount of edge information transferred from the original image A and B into the final fusion image F. Initially, the Sobel kernel is applied to compute the magnitude $g(m,n)$ and direction $\alpha(m,n)$ of gradient at each pixel position of source images and the fused image. Then, the relative magnitude and orientation values of source image A with respect to fused image F is calculated as:

$$G^{AF}_{(m,\ n)} = \begin{cases} \frac{g_F(m,n)}{g_A(m,n)}, \text{if } g_A(m,n) > g_F(m,n) \\ \frac{g_A(m,n)}{g_F(m,n)}, \text{else} \end{cases} \tag{3.9}$$

$$A^{AF}_{(m,\ n)} = 1 - \frac{|\alpha_A(m,n) - \alpha_F(m,n)|}{\pi/2} \tag{3.10}$$

Preserved edge information value is calculated as:

$$Q^{AF}_{(m,n)} = \Gamma_g \Gamma_\alpha \left(1 + e^{K_g(G^{AF}_{(m,n)} - \sigma_g)}\right)^{-1} \left(1 + e^{K_\alpha(A^{AF}_{(m,n)} - \sigma_\alpha)}\right)^{-1} \tag{3.11}$$

where constant values Γ_g, K_g, σ_g and Γ_α, K_α, σ_α determine the exact shape of sigmoid function used for forming the edge strength and orientation. The same computation is applied for calculating $Q^{BF}_{(m,\,n)}$. Finally, the metric is obtained as:

$$Q^{AB/F} = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} \left(Q^{AF}_{(m,n)} w^{AF}_{(m,n)} + Q^{BF}_{(m,n)} w^{BF}_{(m,n)}\right)}{\sum_{m=1}^{M} \sum_{n=1}^{N} (w^{AF}_{(m,n)} + w^{BF}_{(m,n)})} \tag{3.12}$$

3.1.7 Feature Mutual Information (FMI)

It was proposed by [10], which is computed as:

$$FMI^{AB}_F = I_{FA} + I_{FB} \tag{3.13}$$

$$I_{FA} = \sum_{f,a} \left(p_{FA}(x,y,z,w) \times \log_2 \frac{p_{FA}(x,y,z,w)}{p_F(x,y) \times p_A(z,w)} \right) \tag{3.14}$$

$$I_{FB} = \sum_{f,b} \left(p_{FB}(x,y,z,w) \times \log_2 \frac{p_{FB}(x,y,z,w)}{p_F(x,y) \times p_B(z,w)} \right) \tag{3.15}$$

with $p_{FA}(x,y,z,w)$ and $p_{FB}(x,y,z,w)$ are the joint distribution between fused image F with source image A and B, respectively. $p_I(x,y) = \frac{|\nabla I|}{\sum_{x,y} |\nabla I|}$ denotes the marginal distribution with respect to the gradient of the image.

3.1.8 Visual Information Fidelity for Fusion (VIFF)

It was proposed by Han *et al.* [11], which is based on the visual information fidelity to assess objectively fusion performance.

$$VIFF(I_1, I_2, I_F) = \sum_k (p_k \times VIFF_k(I_1, I_2, I_F)) \tag{3.16}$$

$$VIFF_k(I_1, I_2, I_F) = \frac{\sum_b FVID_{k,b}(I_1, I_2, I_F)}{\sum_b FVIND_{k,b}(I_1, I_2, I_F)} \tag{3.17}$$

where:

- p_k is weighting coefficient
- k, b denote the number of sub-bands and blocks in VIF original

- $FVID$ is fusion visual information without distortion
- $FVIND$ is fusion visual information with distortion

3.2 Experimental Setup

180 MRI-PET extracted from TWBA will be divided into datasets as below:

Table 3.1 – Dataset division

Dataset	Description
M1	Two MRI images and two PET images
M2	90 MRI images
M3	90 PET-gray images
M4	90 pairs of MRI-PET images

Experiment #1: Used for comparing the optimization process of different algorithms including WSO, MPA, CLMPA. Dataset M1 is used in this experiment, and on each image, the process depicted in Fig.2.13 is run 30 times with a single optimization algorithm.

Experiment #2: This experiment is set up for evaluating the performance of the Enhancement Component for enhancing images. Dataset M2 and M3 are used, and several former image enhancement approaches are carried out for comparison purpose with our method including CLAHE, EGIF [18], EFF [29]. The evaluation metrics for this experiment include Entropy, Mean Light Intensity (MLI), Contrast Index (CI), Average Gradient (AG) and Structural Similarity index (SSIM).

Experiment #3: Used for evaluating our studied method introduced in Fig.2.14. Dataset M4 is used for the experiment. 90 final fused images then are assessed using qualitative and quantitative evaluation, in comparison with other MRI-PET fusion method namely GBF-Robinson [13].

All the experiments are carried out on personal computer with MATLAB R2023b and Python version 3.10 software, 16GB RAM, CPU Intel Core i5 10400F (2.90 GHz), GPU NVIDIA Geforce GTX 1050Ti 4GB VRAM. The hyperparameters for 3 optimization algorithms are: 50 population (searching agents), 50 iterations for each run.

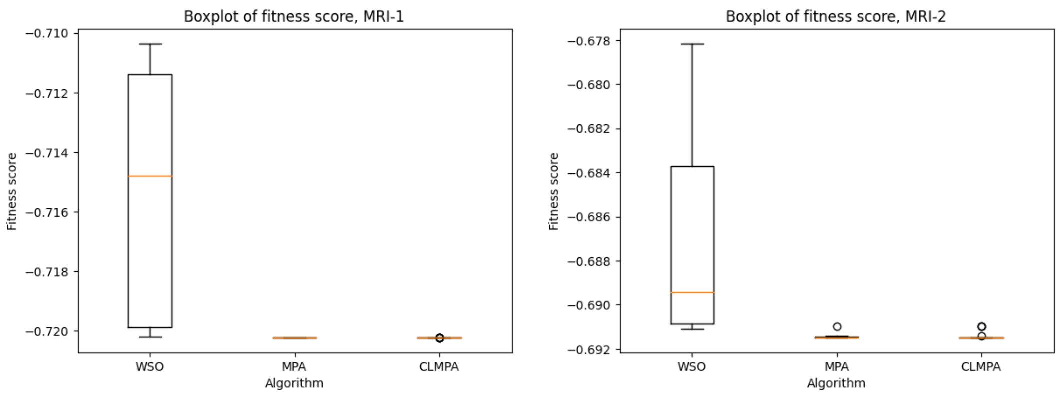
3.3 Results and Discussion

3.3.1 Experiment #1

Table 3.2 and Fig.3.1 show the final result of this experiment on 2 MRI images namely MRI-1 and MRI-2. From the result demonstrated by this table, it can be observed that MPA performed better than two other algorithms on image MRI-2, while on image MRI-1, CLMPA had the best mean value. However, the differences between the best value and others are insignificant, being merely counted in 10^{-3} or 10^{-4} unit. The considerable difference could be seen in the standard deviation column: Marine Predator Algorithm (MPA) had the lowest standard deviation value after 30 independent runs, meanwhile that of WSO was largest. This is the reason why when looking at the boxplot in Fig.3.1, WSO possesses the widest boxplot whilst that of Marine Predator Algorithm (MPA) is the narrowest.

Table 3.2 – Best, worst, mean and standard deviation of fitness function from 30 independent runs on 2 different MRI images. Bold value is the best value relative to others.

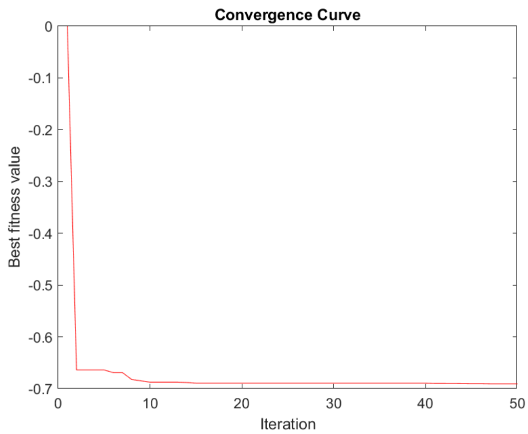
Image	Algorithm	Best	Worst	Mean	Std
MRI-1	WSO	-0.6910536683	-0.6817348899	-0.6874713567	0.0038089172
	MPA	-0.6911578655	**-0.6911361146**	-0.6911507599	**4.816709e-06**
	CLMPA	**-0.6914853412**	-0.690955672	**-0.691427689**	0.00015248
MRI-2	WSO	-0.720200126	-0.7103576069	-0.715274928	0.0039107
	MPA	**-0.720221800**	**-0.720221156**	**-0.720221492**	**1.444739e-07**
	CLMPA	-0.720220937	-0.720030939	-0.720165229	5.49087e-05



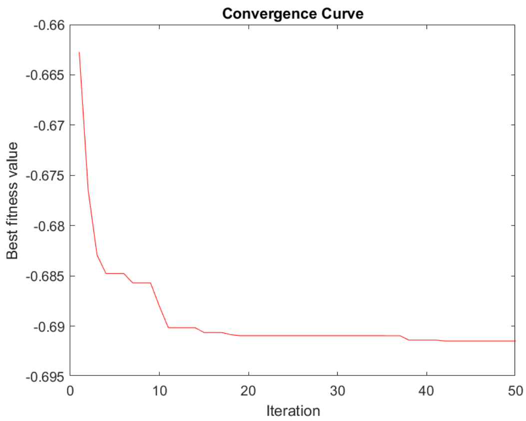
(a) MRI-1 **(b)** MRI-2

Figure 3.1 – Boxplot of 30 independent runs on 2 different MRI images

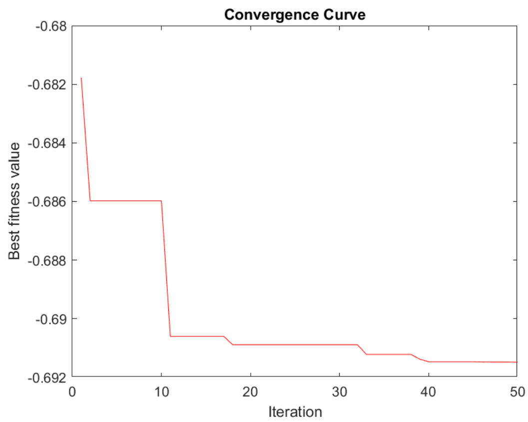
Moreover, Fig.3.2 depicts the convergence curve of 3 algorithms during optimization process of an arbitrary run on image MRI-2: War Strategy Algorithm (WSO) converges very quickly, just after 4 or 5 iterations. Meanwhile, MPA almost converges after over 10 iterations, proving that it searched more in the searching space than WSO. Finally, CLMPA converges very slow, after approximately 40 iterations, illustrating the evidence that this algorithm was designed to have a wider range of searching process and to prevent the pre-mature convergence.



(a) WSO



(b) MPA



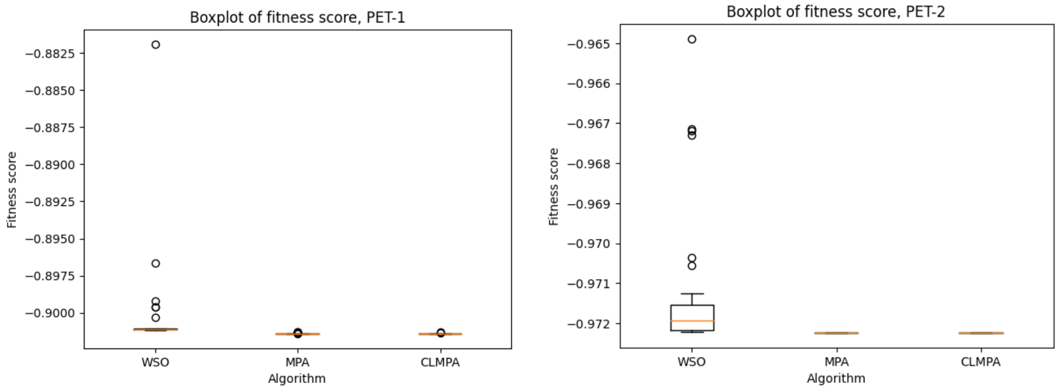
(c) CLMPA

Figure 3.2 – Convergence curve of different runs on image MRI-2 with 3 optimization algorithms

Moving onto Table 3.3 and Fig.3.3 showing the result of experiment #1 on two images PET-gray, the same trend with the results of 2 MRI images above could be witnessed. Once again, MPA had the best performance both on image PET-1 and PET-2 compared with other algorithms. Also, MPA had the smallest standard deviation, while WSO algorithm had the largest standard deviation value after 30 different runs, hence the range of its boxplot is widest. Noticeably, in Fig.**??**, there are much more outliers of WSO than others.

Table 3.3 – Best, worst, mean and standard deviation of objective function from 30 independent runs of different algorithms on image PET-gray-1 and 2

Image	Algorithm	Best	Worst	Mean	Std
PET-gray-1	WSO	-0.901157295	-0.89351999	-0.900776934	0.0013832187
	MPA	**-0.901405541**	**-0.901250634**	**-0.9013918035**	**3.2777395e-05**
	CLMPA	-0.901405126	-0.9011664	-0.901384477	4.717233e-05
PET-gray-2	WSO	-0.97223695	-0.96289473	-0.9707900	0.002403621
	MPA	**-0.97223757**	**-0.9722374**	**-0.9722375**	**4.3680466e-08**
	CLMPA	-0.97223756	-0.9722373	**-0.9722375**	4.4539587e-08



(a) PET-gray-1 **(b)** PET-gray-2

Figure 3.3 – Boxplot of 30 independent runs on 2 different PET-gray images

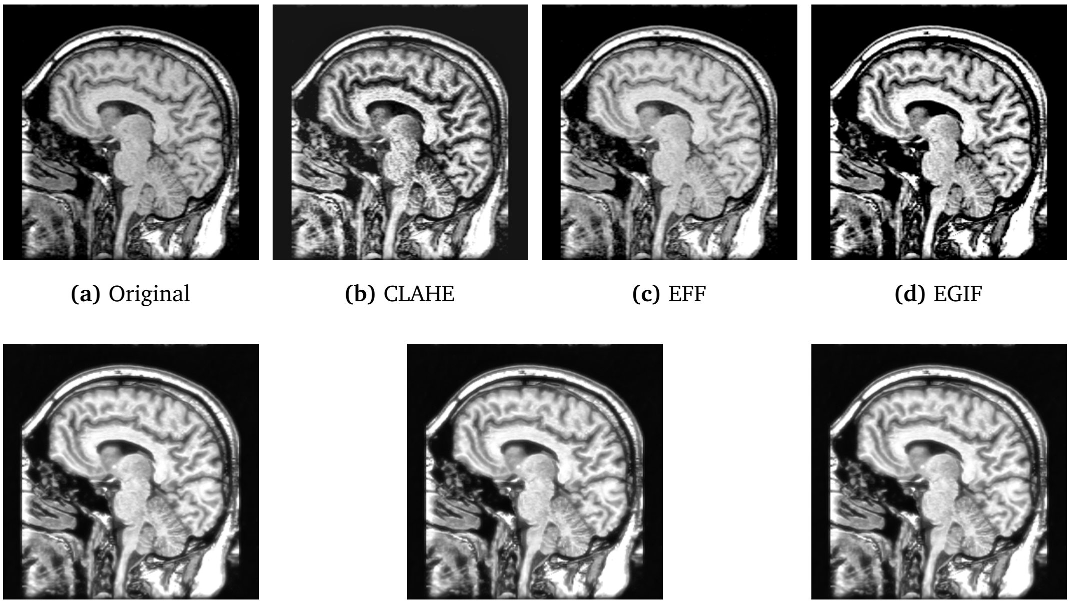
After this experiment, it seems that the MPA's performance is better relative to that of 2 other optimization algorithm.

3.3.2 Experiment #2

The result of Experiment#2 on dataset M2 and M3 are illustrated as Table 3.4 and 3.5, respectively. In the meantime, Fig.3.4 and 3.5 depict the images after applying different image enhancement methods from an original MRI image and a PET-gray from the dataset.

Table 3.4 – Experiment #2 on dataset M2. Red value and blue bold value illustrate the best and the second best relative to others.

	MLI	CI	Entropy	AG	SSIM
Original	0.3045	0.3203	4.6393	0.0803	-
CLAHE	0.3554	0.3078	6.1223	0.1035	0.5447
EFF [29]	0.3606	**0.3629**	5.5651	**0.0894**	**0.9524**
EGIF [18]	0.3250	**0.4665**	4.2310	**0.1581**	0.7153
EC-WSO	**0.3756**	0.3561	**6.6325**	0.0736	0.7242
EC-MPA	**0.3730**	0.3560	**6.6320**	0.0750	**0.7346**
EC-CLMPA	**0.3730**	0.3560	6.6317	0.0750	0.7345



(a) Original **(b)** CLAHE **(c)** EFF **(d)** EGIF

(e) EC-WSO **(f)** EC-MPA **(g)** EC-CLMPA

Figure 3.4 – Comparison of different enhancement techniques carried out on an MRI image from dataset M2

The data presented in Table 3.4 reveals that the Enhancement Component, when implemented with three distinct optimization algorithms, yields comparable outcomes with minor variations between them. This Enhancement Component, utilizing three different algorithms, demonstrates superior Mean Light Intensity (MLI) when compared to alternative enhancement techniques. This superiority is visually evident in Fig.3.4, where the images produced by our approach exhibit noticeably higher overall brightness than those generated by other methods. In terms of Contrast Index (CI) assessment, EGIF and EFF techniques achieve the highest and second-highest scores respectively. While EGIF's top performance significantly overshadows the second, the CI values produced by our proposed technique with 3 algorithms are closely comparable to the second-best result. Next, the entropy measurements resulting from Enhancement Component surpass those of alternative techniques, achieving the highest values. This superiority in entropy indicates that the images enhanced by Enhancement Component contain a greater amount of information compared to those processed by other approaches. Regarding the average gradient (AG), our method underperforms compared to other techniques, ranking lowest among them. This inferior performance is visually corroborated in Fig.3.4, where our resulting images display a slight blur and halo effect, therefore it made the AG value lower. Finally, the Enhancement Component attains the second-highest score, with the EFF technique possessing the top position. Visual inspection of Fig.3.4 confirms these findings: the EFF-enhanced image bears the closest resemblance to the original, while enhancement component's outputs maintains greater similarity to the source image than those produced by CLAHE or EGIF enhancement techniques.

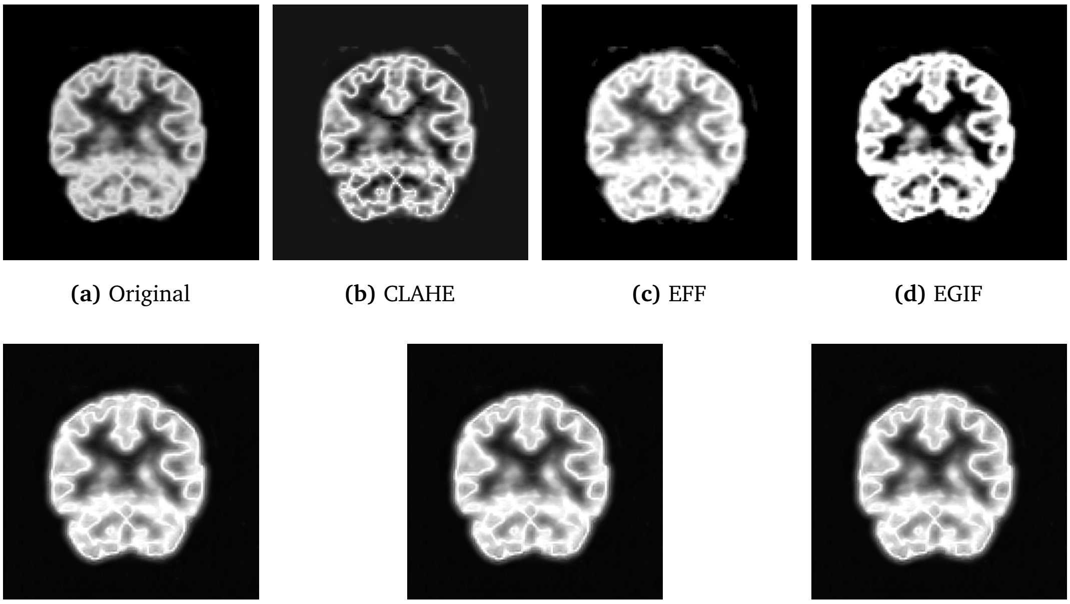
In summary, the Enhancement Component produces enhanced MRI images characterized by increased brightness and enriched information content, while maintaining satisfactory levels of structural similarity and reduced noise. However, these improvements come with the trade-off of minor blur and halo artifacts in the resulting images.

The results presented in Table 3.5, which shows the outcome of experiment 2 on 90 PET-gray images, follow a similar pattern to the previous experiment on the M2 dataset. While CLAHE and EFF methods achieve the highest and second-highest MLI values respectively, enhancement component's results using three algorithms are very close, with differences of less than 0.01. Regarding the CI metric, our approach produces satisfactory results, though they fall short of those obtained by EGIF and EFF techniques. However, in terms of entropy, enhancement component outperforms all other enhancement techniques, achieving the highest values. This indicates that enhancement component's outputs provide more information gain compared to other methods.

Table 3.5 – Experiment #2 on dataset M3, red and blue bold values indicate the best and the second best result relative to others

	MLI	CI	Entropy	AG	SSIM
Original	0.1839	0.3123	3.1988	0.0290	-
CLAHE	**0.2322**	0.2624	4.2399	**0.0294**	0.2899
EFF [29]	**0.2297**	**0.3660**	3.5809	0.0251	**0.8922**
EGIF [18]	0.2109	**0.4024**	3.0927	**0.0442**	**0.8792**
EC-WSO	0.2245	0.3349	**4.9026**	0.0224	0.5100
EC-MPA	0.2241	0.3344	4.8762	0.0223	0.5102
EC-CLMPA	0.2241	0.3344	**4.8800**	0.0224	0.5109

Regarding the AG evaluation, the results for the PET-gray images are comparable to those of the M2 dataset. As visually demonstrated in Fig.3.5, the enhanced PET-gray images exhibit minor blur and halo artifacts. Consequently, enhancement component's AG values are the lowest among the compared techniques. Finally, considering the structural similarity index (SSIM), enhancement component's scores are notably lower than those of EGIF and EFF. This discrepancy can be attributed to two factors: the presence of slight blur and halo effects in our resulting images and the significant noise reduction in our enhanced images compared to the original ones. These factors likely contribute to the lower SSIM scores for our approach, as SSIM measures the structural similarity between the result and original images.



(a) Original **(b)** CLAHE **(c)** EFF **(d)** EGIF

(e) EC-WSO **(f)** EC-MPA **(g)** EC-CLMPA

Figure 3.5 – Comparison of different enhancement techniques conducted on an PET-gray image from dataset M3

To summarize, our proposed enhancement method for PET-gray images demonstrates the highest entropy values, which indicate superior information gain, acceptable levels of brightness and contrast and structural similarity. However, similar to the M2 dataset results, the PET-gray images suffer from minor blur and halo artifacts. Consequently, our method performs poorly in terms of AG metrics. This might stem from the high weight set for the DnCNN component in the Enhancement Component process.

3.3.3 Experiment #3

The result of our proposed method in comparison with GBF-Robinson [13] fusion method are presented as Table 3.6 below:

Table 3.6 – Results of Experiment#3, bold value is the best relative to other methods

	EN	$Q^{AB/F}$	FMI	VIFF
Ours-WSO	**6.8094**	0.4548	0.7647	0.4693
Ours-MPA	6.7944	0.4691	0.7628	0.4695
Ours-CLMPA	6.7946	**0.4692**	0.7628	0.4695
GBF-Robinson [13]	6.0396	0.4275	**0.7908**	**0.9054**

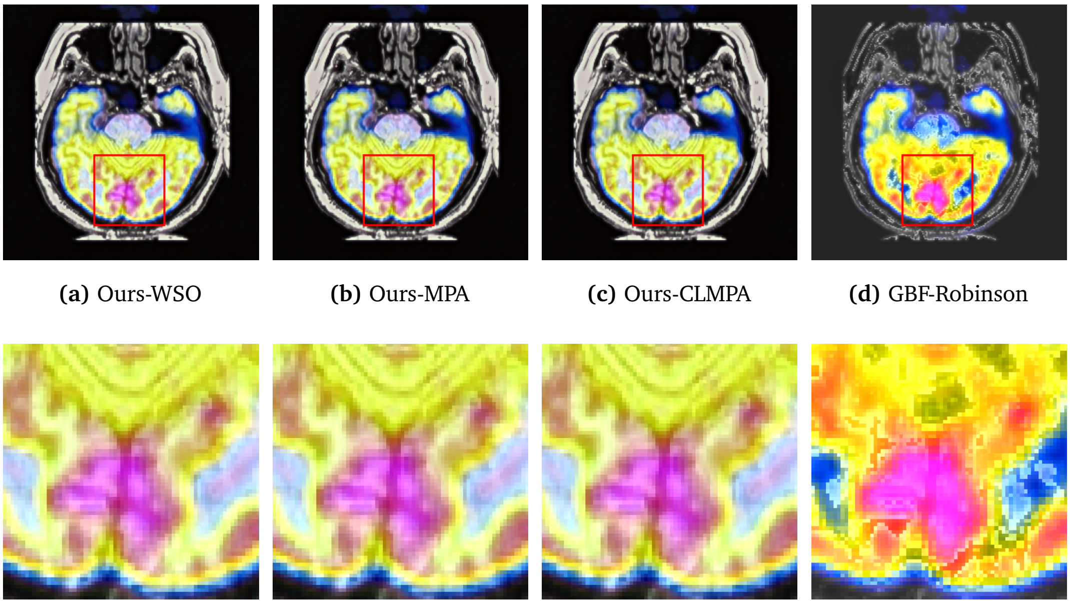
The tabulated results reveal that our method using the WSO algorithm achieves the highest entropy. However, the entropy values obtained using MPA and CLMPA are nearly identical, falling just 0.01 short of WSO's peak value. Notably, our method's entropy values outstrip that of GBF-Robinson by approximately 0.8.

Regarding $Q^{AB/F}$ metrics, CLMPA slightly edges out the others. Yet, the difference between CLMPA and MPA is minimal, with WSO trailing by only about 0.01. In contrast, the GBF-Robinson method's $Q^{AB/F}$ value falls significantly below our three algorithmic approaches.

For FMI metrics, GBF-Robinson possesses the top rank, though our three algorithms yield nearly equivalent results. The gap between GBF-Robinson and our methods is merely 0.03, which can be acceptable.

Lastly, in terms of VIFF, GBF-Robinson demonstrates a sharp superiority, outperforming our proposed method by a difference of around 0.44.

Regarding quantitative evaluation, following Fig.3.6, it reveals minimal variations among our method's results when employing three different algorithms. In contrast to the GBF-Robinson method, the cropped frames demonstrate that our approach generates fused images with reduced noise and more defined edges. Notably, the triangle-shaped area at the upper region of the cropped frame is more visible in our method's fusion images compared to GBF-Robinson's output, in which this feature is less distinguishable.



(a) Ours-WSO (b) Ours-MPA (c) Ours-CLMPA (d) GBF-Robinson

(e) Ours-WSO (f) Ours-MPA (g) Ours-CLMPA (h) GBF-Robinson

Figure 3.6 – An instance of image from the dataset M4 with different fusion methods and their cropped frame

Overall, our proposed pipeline yields final fusion images with more information gain and edge strength, meanwhile performs well with an acceptable level of FMI index. However, the visual information fidelity values of our method are overshadowed sharply by the GBF-Robinson technique.

Chapter 4

Conclusion & Future Works

4.1 Conclusion

Our study proposes an advanced medical image enhancement technique, focusing on MRI and PET images. The method integrates existing enhancement approaches, specifically CLAHE and DnCNN, and then applies an optimization algorithm. This process aims to improve image quality by balancing several factors including increased contrast and brightness levels, noise reduction and more obvious edges. The final step involves fusing the enhanced MRI and PET image pairs using the pre-built IFCNN model, resulting in a higher quality composite image.

Three different experiments was conducted on four divided datasets in order to evaluate the performance of three optimization algorithms, the quality of enhanced image and final fusion MRI-PET image. The results from Experiment#1 indicate that the MPA performed better than two other algorithms with 30 independent runs on four different images. Afterwards, Experiment#2's outputs illustrate that our proposed enhancement technique produced result images with high information gain and brightness, going parallel with acceptable levels of contrast and structural similarity, yet those are affected by blur and halo artifacts. Finally, the qualitative and quantitative evaluation of fused images in comparison with GBF-Robinson technique reveal that our results outperform this technique in entropy and edge based similarity index, while produce a tolerable FMI index, yet those are outstripped by GBF-Robinson regarding VIFF index.

Our study's results demonstrate that our approach can effectively enhance the quality of both individual medical images and fused images from different modalities to a significant degree. These improvements have the potential to advance and refine diagnostic and treatment processes in healthcare. Nevertheless, it's important to note that our research was merely carried out on brain imaging. Given the promising outcomes, we intend to expand the application of this study to medical imaging of other body parts and systems later.

4.2 Future Works

In the future, the 3 distinguished elements of the Enhancement Component would be varied, and also more state-of-the-art optimization algorithms can be studied and applied into our technique. Furthermore, we would like to fine-tune the hyperparameters for optimization algorithms, which might result better in time complexity and performance of those algorithms.

Moreover, as mentioned in the discussion of Experiment#2's results, our enhanced images are degraded slightly by blur and halo effects. This is mainly because the high weight between 0.5 and 0.9 set for the denoised image component. Hence, the range of this factor need to be tuned for less blur artifacts in the improved images. Finally, more fusion rules for combining multi-modal medical images would be considered, in order to yield the higher quality fused images. Considering the fusion rule of multimodal medical images, better fusion models will be applied which might achieve higher quality of final fused images.

In this study, we focused exclusively on the The Whole Brain Atlas (TWBA) dataset, which contains brain images. However, in future research, we plan to extend our enhancement method to other types of medical images like bones imaging or other types of medical imaging.

References

[1] Upendra Kumar Acharya and Sandeep Kumar. "Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement." In: *Optik* 230 (2021), p. 166273.

[2] Tummala. S. L. V. Ayyarao et al. "War Strategy Optimization Algorithm: A New Effective Metaheuristic Algorithm for Global Optimization." In: *IEEE Access* 10 (2022), pp. 25073–25105. DOI: 10.1109/ACCESS.2022.3153493.

[3] Gabriel Fillipe Centini Campos et al. "Machine learning hyperparameter selection for contrast limited adaptive histogram equalization." In: *EURASIP Journal on Image and Video Processing* 2019 (2019), pp. 1–18.

[4] Yunjin Chen and Thomas Pock. "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration." In: *IEEE transactions on pattern analysis and machine intelligence* 39.6 (2016), pp. 1256–1272.

[5] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. "Image denoising by sparse 3-D transform-domain collaborative filtering." In: *IEEE Transactions on image processing* 16.8 (2007), pp. 2080–2095.

[6] Phu-Hung Dinh. "A novel approach based on grasshopper optimization algorithm for medical image fusion." In: *Expert Systems with Applications* 171 (2021), p. 114576.

[7] Phu-Hung Dinh. "A novel approach based on marine predators algorithm for medical image enhancement." In: *Sensing and Imaging* 24.1 (2023), p. 6.

[8] Oanh Cuong Do, Chi Mai Luong, Phu-Hung Dinh, and Giang Son Tran. "An efficient approach to medical image fusion based on optimization and transfer learning with VGG19." In: *Biomedical Signal Processing and Control* 87 (2024), p. 105370.

[9] Afshin Faramarzi, Mohammad Heidarinejad, Seyedali Mirjalili, and Amir H. Gandomi. "Marine Predators Algorithm: A nature-inspired metaheuristic." In: *Expert*

Systems with Applications 152 (2020), p. 113377. ISSN: 0957-4174. DOI: https://doi.org/10.1016/j.eswa.2020.113377.

[10] Mohammad Bagher Akbari Haghighat, Ali Aghagolzadeh, and Hadi Seyedarabi. "A non-reference image fusion metric based on mutual information of image features." In: *Computers & Electrical Engineering* 37.5 (2011), pp. 744–756.

[11] Yu Han, Yunze Cai, Yin Cao, and Xiaoming Xu. "A new image fusion performance metric based on visual information fidelity." In: *Information fusion* 14.2 (2013), pp. 127–135.

[12] Nicolas Humphries et al. "Environmental context explains Lévy and Brownian movement patterns of marine predators." In: *Nature* 465 (June 2010). DOI: https://doi.org/10.1038/nature09116.

[13] Quoc Viet Kieu, Vinh Nam Huynh, Thi Phuong Nghiem, Oanh Cuong Do, and Giang Son Tran. "A NEW METHOD FOR MEDICAL IMAGE FUSION BASED ON GAUSSIAN BLUR FILTER AND ROBINSON COMPASS OPERATOR." In: *Journal of Computer Science and Cybernetics* 40.2 (2024), pp. 135–146.

[14] Fayez Lahoud and Sabine Süsstrunk. "Zero-learning fast medical image fusion." In: *2019 22th international conference on information fusion (FUSION)*. IEEE. 2019, pp. 1–8.

[15] Bo Li and Wei Xie. "Adaptive fractional differential approach and its application to medical image enhancement." In: *Computers & Electrical Engineering* 45 (2015), pp. 324–335.

[16] Yu Liu, Xun Chen, Juan Cheng, and Hu Peng. "A medical image fusion method based on convolutional neural networks." In: *2017 20th international conference on information fusion (Fusion)*. IEEE. 2017, pp. 1–7.

[17] Zheng Liu et al. "Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.1 (2012), pp. 94–109. DOI: 10.1109/TPAMI.2011.109.

[18] Zongwei Lu, Bangyuan Long, Kang Li, and Fajin Lu. "Effective guided image filtering for contrast enhancement." In: *IEEE Signal Processing Letters* 25.10 (2018), pp. 1585–1589. DOI: https://doi.org/10.1109/LSP.2018.2867896.

[19] Yunliang Qi et al. "A comprehensive overview of image enhancement techniques." In: *Archives of Computational Methods in Engineering* (2021), pp. 1–25.

[20] Deepak Kumar Sahu and MP Parsai. “Different image fusion techniques–a critical review.” In: *International Journal of Modern Engineering Research (IJMER)* 2.5 (2012), pp. 4298–4301.

[21] Claude E. Shannon. “A mathematical theory of communication.” In: *Bell Syst. Tech. J.* 27 (1948), pp. 623–656.

[22] Karen Simonyan and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition.” In: *arXiv preprint arXiv:1409.1556* (2014).

[23] Gursharn Singh, Anand Mittal, et al. “Various image enhancement techniques-a critical review.” In: *International Journal of Innovation and Scientific Research* 10.2 (2014), pp. 267–274.

[24] Eric Vidoni. “The Whole Brain Atlas.” In: *Journal of Neurologic Physical Therapy* 36 (June 2012), p. 108. DOI: 10.1097/NPT.0b013e3182563795.

[25] Kunpeng Wang, Mingyao Zheng, Hongyan Wei, Guanqiu Qi, and Yuanyuan Li. “Multi-modality medical image fusion using convolutional neural network and contrast pyramid.” In: *Sensors* 20.8 (2020), p. 2169.

[26] Zhou Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. “Image quality assessment: from error visibility to structural similarity.” In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: 10.1109/TIP.2003.819861.

[27] C. Xydeas and Vladimir Petrovic. “Objective image fusion performance measure.” In: *Electronics Letters* 36 (Mar. 2000), pp. 308 –309. DOI: 10.1049/el:20000267.

[28] Yi Yang, Zhengwei Su, and L Sun. “Medical image enhancement algorithm based on wavelet transform.” In: *Electronics letters* 46.2 (2010), pp. 120–121.

[29] Zhenqiang Ying, Ge Li, Yurui Ren, Ronggang Wang, and Wenmin Wang. “A new image contrast enhancement algorithm using exposure fusion framework.” In: *Computer Analysis of Images and Patterns: 17th International Conference, CAIP 2017, Ystad, Sweden, August 22-24, 2017, Proceedings, Part II 17*. Springer. 2017, pp. 36–46. DOI: https://doi.org/10.1007/978-3-319-64698-5_4.

[30] Dalia Yousri, Ahmed Fathy, and Hegazy Rezk. “A new comprehensive learning marine predator algorithm for extracting the optimal parameters of supercapacitor model.” In: *Journal of Energy Storage* 42 (2021), p. 103035. ISSN: 2352-152X. DOI: https://doi.org/10.1016/j.est.2021.103035. URL: https://www.sciencedirect.com/science/article/pii/S2352152X21007465.

[31] Hao Zhang, Han Xu, Xin Tian, Junjun Jiang, and Jiayi Ma. "Image fusion meets deep learning: A survey and perspective." In: *Information Fusion* 76 (2021), pp. 323–336. DOI: `https://doi.org/10.1016/j.inffus.2021.06.008`.

[32] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising." In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3142–3155. DOI: `10.1109/TIP.2017.2662206`.

[33] Chenyi Zhao et al. "A new approach for medical image enhancement based on luminance-level modulation and gradient modulation." In: *Biomedical Signal Processing and Control* 48 (2019), pp. 189–196.

[34] Karel Zuiderveld. "Contrast limited adaptive histogram equalization." In: *Graphics gems IV*. 1994, pp. 474–485.

Appendix A

DnCNN layers

Table A.1 – Some layers' details of DnCNN

	Name	Type	Activation	Learnable	
0	Input Layer, $50 \times 50 \times 1$ images	Image input	$50 \times 50 \times 1$	-	
1	Conv 1 64 $3 \times 3 \times 1$convolutions with stride [11] and padding [1111] ReLU1	Convolution ReLU	$50 \times 50 \times 64$ $50 \times 50 \times 64$	Weights Bias -	$3 \times 3 \times 1 \times 64$ $1 \times 1 \times 64$
2	Conv2 64 $3 \times 3 \times 1$convolutions with stride [11] and padding [1111] BNorm2 Batch Normalization with 64 channels ReLU2	Convolution Batch Normalization ReLU	$50 \times 50 \times 64$ $50 \times 50 \times 64$ $50 \times 50 \times 64$	Weights Bias Offset Scale -	$3 \times 3 \times 64 \times 64$ $1 \times 1 \times 64$ $1 \times 1 \times 64$ $1 \times 1 \times 64$
...	...	...	...	...	...
19	Conv19 64 $3 \times 3 \times 1$convolutions with stride [11] and padding [1111] BNorm19 Batch Normalization with 64 channels ReLU19	Convolution Batch Normalization ReLU	$50 \times 50 \times 64$ $50 \times 50 \times 64$ $50 \times 50 \times 64$	Weights Bias Offset Scale -	$3 \times 3 \times 64 \times 64$ $1 \times 1 \times 64$ $1 \times 1 \times 64$ $1 \times 1 \times 64$
20	Conv20	Convolution	$50 \times 50 \times 1$	Weights	$3 \times 3 \times 64$
21	Final Regression Layer	Regression output			