

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



Đinh Phú Hùng

**NÂNG CAO HIỆU QUẢ TỔNG HỢP HÌNH ẢNH Y HỌC
THEO TIẾP CẬN TỐI ƯU HOÁ**

LUẬN ÁN TIẾN SĨ NGÀNH KHOA HỌC MÁY TÍNH

Hà Nội - Năm 2024

BỘ GIÁO DỤC
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Đinh Phú Hùng

NÂNG CAO HIỆU QUẢ TỔNG HỢP HÌNH ẢNH Y HỌC
THEO TIẾP CẬN TỐI ƯU HOÁ

LUẬN ÁN TIẾN SĨ NGÀNH KHOA HỌC MÁY TÍNH

Mã số: 9 48 01 01

Xác nhận của Học viện
Khoa học và Công nghệ

Người hướng dẫn 1
(Ký, ghi rõ họ tên)

Người hướng dẫn 2
(Ký, ghi rõ họ tên)



Nguyễn Thị Trung

Nguyễn Long Cường

Nguyễn Hùng Đức

Hà Nội - Năm 2024

LỜI CAM ĐOAN

Tôi xin cam đoan luận án: "**Nâng cao hiệu quả tổng hợp hình ảnh y học theo tiếp cận tối ưu hoá**" là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Luận án sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được công bố chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 31 tháng 01 năm 2024

Tác giả luận án



Đinh Phú Hùng

LỜI CẢM ƠN

Luận án tiến sĩ được thực hiện tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam (Viện HLKH&CN Việt Nam), dưới sự hướng dẫn khoa học tận tình của PGS. TS. Nguyễn Long Giang và TS. Nguyễn Huy Đức. Tác giả xin bày tỏ lòng biết ơn chân thành và sự kính trọng sâu sắc đối với các thầy trong Tập thể hướng dẫn khoa học, những người không chỉ truyền đạt nhiều kiến thức quý báu, kinh nghiệm nghiên cứu khoa học mà còn khuyến khích, động viên tác giả vượt qua những khó khăn trong chuyên môn và cuộc sống. Sự chuyên nghiệp, nghiêm túc trong nghiên cứu và những định hướng đúng đắn của các thầy là tiền đề quan trọng giúp tác giả có được những kết quả trình bày trong luận án này.

Tác giả xin trân trọng cảm ơn Ban lãnh đạo Viện Công nghệ Thông tin, Ban Giám đốc Học viện Khoa học và Công nghệ, Phòng Đào tạo, các Phòng Ban chức năng của Học viện và đặc biệt các nhà giáo, nhà khoa học tại Viện HL KH&CN Việt Nam đã quan tâm giúp đỡ, tạo mọi điều kiện thuận lợi về cơ sở vật chất, nguồn học liệu và các thủ tục hành chính cho tác giả trong quá trình học tập, nghiên cứu và hoàn thành luận án này.

Trong suốt thời gian học tập và nghiên cứu tại Học viện Khoa học và Công nghệ, Viện HL KH&CN Việt Nam, tác giả đã nhận được sự hỗ trợ về tài chính từ Quỹ Đổi mới Sáng tạo VinGroup (VinIF), Viện Nghiên cứu Dữ liệu lớn (VinBigData) thông qua Học bổng tiến sĩ trong nước năm học 2021-2022. Qua đây, tác giả xin bày tỏ sự biết ơn sâu sắc với sự hỗ trợ kịp thời của Quỹ Đổi mới Sáng tạo VinGroup để tác giả có thể toàn tâm tập trung cho việc học tập, nghiên cứu và đạt được các kết quả trong luận án.

Tác giả xin chân thành cảm ơn tới Ban Giám hiệu Trường Đại học Thủy lợi, Khoa Công nghệ thông tin (CNTT), Bộ môn Mạng & An toàn thông tin và các đồng nghiệp trong Khoa CNTT đã luôn động viên, giúp đỡ tác giả trong công tác để tác giả chuyên tâm vào việc nghiên cứu và hoàn thành luận án.

Đặc biệt tác giả xin được bày tỏ lòng biết ơn sâu sắc tới Bố, Mẹ, và các thành viên trong gia đình, những người đã luôn động viên và truyền cảm hứng cho tác giả trong suốt quá trình nghiên cứu.

Hà Nội, ngày 31 tháng 01 năm 2024

Tác giả luận án



Đinh Phú Hùng

MỤC LỤC

LỜI CAM ĐOAN	i
LỜI CẢM ƠN.....	ii
MỤC LỤC.....	iii
Danh mục các từ viết tắt.....	v
Danh mục các bảng	vii
Danh mục hình vẽ.....	viii
MỞ ĐẦU	1
CHƯƠNG 1. TỔNG QUAN VỀ BÀI TOÁN TỔNG HỢP HÌNH ẢNH VÀ MỘT SỐ KIẾN THỨC NỀN TẢNG	8
1.1. Giới thiệu	8
1.2. Tổng quan các nghiên cứu tổng hợp hình ảnh.....	12
1.2.1. Các cách tiếp cận truyền thống	12
1.2.2. Các cách tiếp cận dựa trên học sâu.....	18
1.3. Một số hạn chế của các phương pháp tổng hợp hình ảnh và giải pháp	26
1.3.1. Hạn chế liên quan đến chất lượng ảnh	26
1.3.2. Hạn chế liên quan đến phương pháp tổng hợp ảnh	28
1.4. Kiến thức nền tảng.....	32
1.4.1. Giới thiệu về thuật toán Meta-heuristic.....	32
1.4.2. Hàm năng lượng cục bộ (LEF).....	39
1.4.3. Thuật toán FR_PCO	39
1.4.4. Cấu trúc Tensor và toán tử phát hiện đặc trưng	40
1.4.5. Bộ lọc hướng dẫn quay vòng (RGF)	40
1.4.6. Bộ lọc đường cong trung bình có trọng số (WMCF).....	41
1.4.7. Cân bằng biểu đồ xám thích nghi giới hạn độ tương phản (CLAHE)	42
1.4.8. Toán tử phát hiện biên Laplace (LED).....	43
1.4.9. Bộ lọc khuếch tán dị hướng (ADF)	43
1.4.10. Phục hồi hình ảnh dựa trên tổng biến thiên (TVIR).....	44
1.5. Chỉ số đánh giá	46
1.5.1. Chỉ số cường độ sáng	47
1.5.2. Chỉ số tương phản	47
1.5.3. Chỉ số lượng thông tin	47
1.5.4. Chỉ số độ sắc nét.....	47
1.5.5. Chỉ số bảo toàn thông tin.....	47
1.5.6. Chỉ số thông tin tương hỗ.....	49
1.5.7. Chỉ số độ tương đồng dựa trên đặc trưng.....	50

1.6. Kết luận chương 1	51
CHƯƠNG 2. NÂNG CAO CHẤT LƯỢNG HÌNH ẢNH DỰA TRÊN PHÂN RÃ BA THÀNH PHẦN VÀ THUẬT TOÁN MPA	53
2.1. Ý tưởng	53
2.2. Đề xuất phương pháp nâng cao chất lượng ảnh	54
2.2.1. Đề xuất thuật toán phân rã hình ảnh ba thành phần	54
2.2.2. Thiết kế hàm mục tiêu	56
2.2.3. Phương pháp nâng cao chất lượng ảnh IE_TCID_MPA.....	57
2.2.4. Độ phức tạp tính toán của thuật toán IE_TCID_MPA.....	59
2.3. Thực nghiệm và đánh giá.....	60
2.3.1. Dữ liệu thực nghiệm.....	60
2.3.2. Thiết lập thực nghiệm.....	61
2.3.3. Kết quả thực nghiệm	63
2.4. Kết luận chương 2.....	72
CHƯƠNG 3. TỔNG HỢP ẢNH DỰA TRÊN THUẬT TOÁN TỔNG HỢP THÍCH NGHI KẾT HỢP VỚI CÁC BIẾN THỂ CỦA HÀM NĂNG LƯỢNG CỤC BỘ	74
3.1. Ý tưởng	74
3.2. Đề xuất thuật toán tổng hợp cho thành phần chi tiết (FM_CVLEF)	74
3.2.1. Cải thiện đặc trưng cấu trúc Tensor bởi hàm năng lượng cục bộ	74
3.2.2. Thuật toán FM_CVLEF	75
3.3. Đề xuất thuật toán tổng hợp cho thành phần cơ sở (AFM_MPA)	77
3.3.1. Thiết kế hàm mục tiêu	77
3.3.2. Thuật toán AFM_MPA	78
3.4. Đề xuất phương pháp tổng hợp hình ảnh.....	79
3.4.1. Thuật toán phân rã hình ảnh ba thành phần (TCID)	79
3.4.2. Phương pháp tổng hợp ảnh AFM_CVLEF	80
3.5. Độ phức tạp tính toán của thuật toán AFM_CVLEF.....	82
3.6. Thực nghiệm và đánh giá.....	83
3.6.1. Dữ liệu thực nghiệm.....	83
3.6.2. Thiết lập thực nghiệm.....	83
3.6.3. Kết quả thực nghiệm	87
3.7. Kết luận chương 3.....	95
KẾT LUẬN	97
DANH MỤC CÁC CÔNG TRÌNH CỦA TÁC GIẢ.....	99
TÀI LIỆU THAM KHẢO	100

Danh mục các từ viết tắt

Viết tắt	Tiếng anh	Diễn giải
AFM_CVLEF	Image fusion methods based on adaptive fusion methods and combining variations of the local energy function	Thuật toán tổng hợp hình ảnh dựa trên phương pháp tổng hợp thích nghi và các biến thể của hàm năng lượng cục bộ
AFM_MPA	Adaptive fusion methods based on MPA algorithm	Thuật toán tổng hợp thích nghi dựa trên thuật toán tối ưu hoá MPA
BBHE	Brightness preserving bi-histogram equalization	Cân bằng hai biểu đồ mức xám bảo toàn độ sáng
BPFHE	Brightness preserving bi-level fuzzy histogram equalization	Cân bằng biểu đồ mức xám mờ hai mức bảo toàn cường độ sáng
CLAHE	Contrast Limited Adaptive Histogram Equalization	Cân bằng biểu đồ mức xám thích nghi giới hạn độ tương phản
CNPS-NSST	Coupled Neural P Systems in Nonsubsampled Shearlet Transform Domain	Các hệ thống P thần kinh kết hợp trong miền NSST
CRT	Contourlet transform	Biến đổi Contourlet
CSE	Contrast and structure extraction	Trích xuất độ tương phản và cấu trúc
CT	Computed Tomography	Ảnh chụp cắt lớp vi tính
CVT	Curvelet transform	Biến đổi Curvelet
DHE	Dynamic histogram equalization	Cân bằng biểu đồ mức xám động
DT-CWT	Dual-Tree Complex Wavelet Transform	Biến đổi Wavelet phức hợp cây kép
DTNP-NSCT	Dynamic threshold neural P systems in nonsubsampled contourlet transform	Các hệ thống P thần kinh ngưỡng động trong miền NSCT
DTNPS	Dynamic threshold neural P systems	Hệ thống P nơ-ron ngưỡng động.
DWT	Discrete wavelet transform	Biến đổi Wavelet rời rạc
FM_CVLEF	Fusion methods based on combining variations of the local energy function	Thuật toán tổng hợp dựa trên kết hợp các biến thể của hàm năng lượng cục bộ
GAN	Generative Adversarial Networks	Mạng đối kháng sinh

IE_TCID_MPA	Image enhancement method based on three-component image decomposition and the MPA algorithm	Thuật toán nâng cao chất lượng ảnh dựa trên phân dã ảnh ba thành phần và thuật toán MPA
JBF-LGE	Joint bilateral filter and local gradient energy	Bộ lọc hai chiều đồng thời và năng lượng gradient cục bộ
MPA	Marine predators algorithm	Thuật toán động vật biển ăn thịt
MRI	Magnetic Resonance Imaging	Ảnh cộng hưởng từ
MSGA	Multi-scale geometric analysis	Phân tích hình học đa tỉ lệ
MST	Multi-scale transform	Biến đổi đa tỉ lệ
NMHE	Non-Parametric Modified Histogram Equalization	Cân bằng biểu đồ mức xám được sửa đổi phi tham số
NSCT	Nonsubsampled Contourlet Transform	Biến đổi Contourlet không lấy mẫu con
NSST	Nonsubsampled Shearlet Transform	Biến đổi Shearlet không lấy mẫu con
PC-LLE-NSCT	A phase congruency and local laplacian energy in NSCT domain	Đồng dư pha và năng lượng laplace cục bộ trong miền NSCT
PET	Positron Emission Tomography	Ảnh chụp cắt lớp phát xạ Positron
SPECT	Single Photon Emission Computed Tomography	Ảnh chụp cắt lớp đơn Photon
SR	Sparse representation	Biểu diễn thưa
ST	Shearlet transform	Biến đổi Shearlet
SWT	Stationary wavelet decomposition	Biến đổi Wavelet tĩnh
TCID	Three-component image decomposition	Phân rã hình ảnh ba thành phần
TLD-SR	Three-layer decomposition and sparse representation	Phân rã ba lớp và biểu diễn thưa

Danh mục các bảng

Bảng 2.1. Một số thuật toán tối ưu	62
Bảng 2.2. Một số thuật toán nâng cao chất lượng ảnh	62
Bảng 2.3. Một số thuật toán tổng hợp hình ảnh y học	63
Bảng 2.4. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70T).....	64
Bảng 2.5. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70S).....	64
Bảng 2.6. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70C)	64
Bảng 2.7. Giá trị p-values từ kiểm định thống kê Wilcoxon	65
Bảng 2.8. Các chỉ số đánh giá từ các thuật toán nâng cao chất lượng ảnh trên tập D0	66
Bảng 2.9. Các chỉ số đánh giá từ các thuật toán nâng cao chất lượng ảnh trên tập D1	66
Bảng 2.10. Các tham số tối ưu thu được trên hai tập dữ liệu D0 và D1.	68
Bảng 2.11. Thời gian chạy trung bình của các thuật toán trên tập D1.....	68
Bảng 2.12. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (T).	69
Bảng 2.13. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (S).....	70
Bảng 2.14. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (C).	71
Bảng 3.1. Mô tả các tập dữ liệu thực nghiệm	83
Bảng 3.2. Một số thuật toán tối ưu được sử dụng để so sánh	84
Bảng 3.3. Một số thuật toán tổng hợp hình ảnh	84
Bảng 3.4. Trung bình và độ lệch chuẩn thu được từ các thuật toán tối ưu với 30 lần chạy độc lập trên tập dữ liệu K4	88
Bảng 3.5. Bảng các giá trị p-values thu được từ kiểm định Wilcoxon	88
Bảng 3.6. Kết quả thu được bởi các thuật toán tổng hợp khác nhau	89
Bảng 3.7. Các tham số tối ưu thu được	89
Bảng 3.8. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K1	90
Bảng 3.9. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K2	91
Bảng 3.10. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K3	91
Bảng 3.11. Thời gian chạy trung bình của các thuật toán trên (K1&K2&K3).....	95

Danh mục hình vẽ

Hình 1.1. Minh hoạ một số kiểu hình ảnh y học.....	9
Hình 1.2. Sơ đồ chẩn đoán bệnh có sử dụng tổng hợp hình ảnh.....	10
Hình 1.3. Sơ đồ phân đoạn khối u có sử dụng tổng hợp hình ảnh.....	11
Hình 1.4. Số lượng các công bố về tổng hợp hình ảnh y học (2006 đến 09/2023)...	11
Hình 1.5. Các cách tiếp cận tổng hợp hình ảnh	12
Hình 1.6. Sơ đồ chung của quá trình tổng hợp hình ảnh.....	12
Hình 1.7. Sơ đồ chung cho các thuật toán phân rã hình ảnh cơ sở - chi tiết.....	15
Hình 1.8. Sơ đồ chung cho các thuật toán tổng hợp hình ảnh dựa trên bộ lọc	16
Hình 1.9. Sơ đồ cho thuật toán tổng hợp hình ảnh dựa trên biểu diễn thưa	17
Hình 1.10. Sơ đồ mạng học sâu được sử dụng để trích xuất đặc trưng cho các hình ảnh đầu vào trong mô hình tổng hợp hình ảnh.....	19
Hình 1.11. Sơ đồ mạng học sâu được sử dụng để trích xuất đặc trưng cho thành phần chi tiết trong mô hình tổng hợp hình ảnh.	20
Hình 1.12. Sơ đồ mạng CNN phân 2 lớp	21
Hình 1.13. Sơ đồ của mô hình tổng hợp sử dụng mạng CNN	22
Hình 1.14. Sơ đồ quá trình huấn luyện mạng Mã hoá - Giải mã	23
Hình 1.15. Sơ đồ của mô hình tổng hợp hình ảnh sử dụng mạng Mã hoá - Giải mã.....	23
Hình 1.16. Sơ đồ mạng GAN tổng hợp hình ảnh với 1 bộ sinh và 1 bộ phân biệt ...	24
Hình 1.17. Sơ đồ mạng GAN tổng hợp hình ảnh với 1 bộ sinh và 2 bộ phân biệt ...	25
Hình 1.18. Minh hoạ ảnh được tăng cường bởi cân bằng biểu đồ mức xám.....	27
Hình 1.19. Minh hoạ thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình	29
Hình 1.20. Minh hoạ cường độ sáng trung bình của các thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình.....	29
Hình 1.21. Minh hoạ độ tương phản trung bình của các thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình.....	30
Hình 1.22. Sơ đồ của thuật toán MPA	36
Hình 1.23. Các mặt nạ của bộ lọc WMCF	42
Hình 2.1. Sơ đồ phân rã hình ảnh ba thành phần	55
Hình 2.2. Minh hoạ phân rã 3 thành phần cho ảnh <i>I</i>	58
Hình 2.3. Kết quả tăng cường thành phần cấu trúc.....	58
Hình 2.4. Minh hoạ ba thành phần đặc trưng.....	58
Hình 2.5. Sơ đồ thuật toán tăng cường chất lượng ảnh IE_TCID_MPA.....	59
Hình 2.6. Minh hoạ 12 hình ảnh MRI trong tập D0.....	60
Hình 2.7. Minh hoạ 12 hình ảnh MRI có tương phản thấp trong tập D1.....	61
Hình 2.8. Minh hoạ 6 cặp MRI-PET trong tập D3	61

Hình 2.9. Giá trị của hàm tối ưu thu được từ các thuật toán trên tập D2.....	64
Hình 2.10. Hình ảnh đầu ra từ các thuật toán tăng cường ảnh trên tập D1.....	66
Hình 2.11. Bốn chỉ số đánh giá chất lượng ảnh trên tập dữ liệu D1.....	67
Hình 2.12. Hình ảnh kết quả từ các thuật toán nâng cao chất lượng trên D0.	67
Hình 2.13. Bốn chỉ số đánh giá chất lượng ảnh trên tập dữ liệu D0.....	67
Hình 2.14. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-T).....	69
Hình 2.15. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-S).....	70
Hình 2.16. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-C).....	71
Hình 2.17. Hình ảnh tổng hợp trước và sau áp dụng thuật toán đề xuất.....	72
Hình 3.1. Minh hoạ đặc trưng cấu trúc Tensor và kết hợp với hàm năng lượng cục bộ	75
Hình 3.2. Sơ đồ thuật toán FM_CVLEF	76
Hình 3.3. Minh hoạ hình ảnh thu được từ thuật toán FM_CVLEF	77
Hình 3.4. Sơ đồ thuật toán AFM_MPA	78
Hình 3.5. Sơ đồ thuật toán phân rã hình ảnh ba thành phần	80
Hình 3.6. Sơ đồ tổng hợp hình ảnh đề xuất AFM_CVLEF	81
Hình 3.7. Minh hoạ ba cặp hình ảnh (MRI-PET) trong tập thực nghiệm.....	83
Hình 3.8. Biểu đồ box plot minh hoạ giá trị hàm fitness thu được đối với 30 lần chạy độc lập của các thuật toán tối ưu khác nhau trên tập dữ liệu K4	87
Hình 3.9. Minh hoạ cho năm chỉ số đánh giá bởi tám thuật toán trên tập K1	91
Hình 3.10. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068T từ K1	92
Hình 3.11. Minh hoạ các vùng ảnh nhỏ được cắt từ Hình 3.10	92
Hình 3.12. Minh hoạ cho năm chỉ số đánh giá từ bởi tám thuật toán trên tập K2	92
Hình 3.13. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068S từ K2	93
Hình 3.14. Minh hoạ các vùng ảnh nhỏ được cắt từ Hình 3.13	93
Hình 3.15. Minh hoạ cho năm chỉ số đánh giá từ bởi tám thuật toán trên tập K3	93
Hình 3.16. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068T từ K3	94
Hình 3.17. Minh hoạ các vùng ảnh nhỏ được cắt từ Hình 3.16	94

MỞ ĐẦU

1. Tính cấp thiết của luận án

Tổng hợp hình ảnh y học là một trong những chủ đề nghiên cứu quan trọng trong lĩnh vực xử lý ảnh. Bài toán này liên quan đến việc kết hợp hai hoặc nhiều hình ảnh y học được chụp từ các phương thức khác nhau để tạo ra một hình ảnh có chất lượng cao và chứa được các thông tin bổ sung từ từng hình ảnh riêng lẻ. Việc tổng hợp hình ảnh này mang lại rất nhiều ứng dụng trong thực tế. Nó không chỉ có ứng dụng trong việc giúp cho các bác sĩ và chuyên gia chẩn đoán hình ảnh trong việc đưa ra các quyết định chẩn đoán các bệnh được chính xác hơn mà còn hỗ trợ nâng cao hiệu quả trong các ứng dụng chẩn đoán bệnh [1]. Hiện nay, các kiểu hình ảnh y học thường được sử dụng phổ biến có thể kể đến như: ảnh chụp cắt lớp vi tính (CT), ảnh chụp cắt lớp phát xạ Positron (PET), ảnh cộng hưởng từ (MRI), ảnh chụp cắt lớp đơn Photon (SPECT). Mỗi loại hình ảnh này đều chứa các thông tin bổ sung lẫn nhau. Ví dụ, hình ảnh chụp cắt lớp vi tính (CT) có độ phân giải cao và có thể định vị chính xác các tổn thương, nhưng lại cung cấp ít thông tin về các cấu trúc mô mềm. Hình ảnh MRI có thể chứa nhiều thông tin quan trọng về cấu trúc và chức năng của các phần khác nhau trong cơ thể. Thứ nhất là cấu trúc giải phẫu: MRI cho phép hiển thị rất chi tiết về các cấu trúc bên trong cơ thể như não, tim, phổi, gan, thận, xương, cơ và nhiều cơ quan khác. Thứ hai là sự phân bố mô và tế bào: MRI có thể phân biệt các loại mô khác nhau như mô cứng, mô mềm, mô nước, mô mỡ, mô vôi. Điều này có thể giúp xác định sự phân bố của các cấu trúc và bệnh lý. Thứ ba là hiển thị mạch máu: MRI có khả năng hiển thị mạch máu trong cơ thể. Điều này rất hữu ích trong việc phát hiện các vấn đề về tuần hoàn và lưu thông máu. Cuối cùng là cung cấp thông tin về bệnh lý và tình trạng sức khỏe: Hình ảnh MRI có thể phát hiện các bất thường trong cấu trúc và tính chất của các cơ quan, giúp cảnh báo và theo dõi các tình trạng bệnh lý như ung thư, bệnh tim mạch, bệnh thần kinh và nhiều loại bệnh khác. Trong khi đó, hình ảnh PET có thể cung cấp thông tin về sự hoạt động của các cơ quan như não, tim, gan, phổi, thận và nhiều cơ quan khác. Chẳng hạn, hình ảnh PET não có thể cho thấy các khu vực não đang hoạt động trong các nhiệm vụ nhất định. PET có khả năng phát hiện sự tích tụ không bình thường của phân tử trong cơ thể, chẳng hạn như dấu vết của ung thư. Điều này giúp xác định các vùng bất thường và theo dõi sự thay đổi sau khi áp dụng liệu pháp. Các hình ảnh PET và SPECT thường được sử dụng để phát hiện các bệnh về mạch máu và khối u. Trên thực tế, các hình ảnh y học riêng lẻ không thể cung cấp đầy đủ các thông tin cần thiết cho việc chẩn đoán. Vì vậy, việc tổng hợp hình ảnh cho phép tạo ra một hình ảnh mà chứa các thông tin bổ sung từ cả hai hoặc

nhiều hình ảnh đầu vào. Một số kiểu tổng hợp hình ảnh phổ biến hiện nay như: MRI với PET, MRI với SPECT và MRI với CT.

Hiện nay, bài toán này nhận được rất nhiều sự quan tâm của các nhà nghiên cứu bởi những ứng dụng của nó trong thực tế. Số lượng các công bố liên quan đến tổng hợp hình ảnh y học không ngừng gia tăng trong những năm gần đây [1]. Nhìn chung, có thể chia các cách tiếp cận ra làm hai nhóm chính: nhóm các cách tiếp cận truyền thống [2–19] và nhóm các cách tiếp cận dựa trên học sâu [20–24]. Đối với các phương pháp dựa trên cách tiếp cận truyền thống đều phải trải qua ba bước chính: phân rã hình ảnh đầu vào, tổng hợp các thành phần trên miền phân rã và cuối cùng các thành phần đã được tổng hợp trên miền phân rã sẽ được biến đổi ngược so với thuật toán phân rã ảnh ban đầu để thu được hình ảnh tổng hợp. Ưu điểm của các phương pháp dựa trên cách tiếp cận truyền thống là có thể linh hoạt trong việc sử dụng các phương pháp phân rã hình ảnh cũng như thiết kế các thuật toán tổng hợp cho các thành phần cơ sở và chi tiết một cách phù hợp. Đối với cách tiếp cận dựa trên học sâu, các mạng học sâu thường có hai nhiệm vụ chính. Thứ nhất là được sử dụng để biểu diễn các đặc trưng cho hình ảnh để từ đó xây dựng một phương pháp tổng hợp cho các thành phần chi tiết trong ảnh. Thứ hai là được huấn luyện để học cách tổng hợp hình ảnh theo một chu trình khép kín. Các phương pháp dựa trên học sâu thường yêu cầu một lượng dữ liệu lớn để đảm bảo mô hình học sâu hoạt động tốt. Tuy nhiên, đối với bài toán tổng hợp hình ảnh, lượng dữ liệu hình ảnh y học đã được căn chỉnh hình học theo từng cặp như: MRI và PET, MRI và CT, MRI và SPECT còn hạn chế (chỉ vài trăm cặp ảnh theo nguồn dữ liệu “The Whole Brain Atlas”: <http://www.med.harvard.edu/AANLIB/>). Vì vậy, việc xây dựng các phương pháp tổng hợp hình ảnh dựa trên học sâu đối với nguồn dữ liệu còn hạn chế về số lượng đang đặt ra nhiều thách thức cho các nhà nghiên cứu [25].

Thách thức đặt ra đối với bài toán tổng hợp hình ảnh y học là đảm bảo được hình ảnh không bị suy giảm về mặt chất lượng trong quá trình tổng hợp cũng như bảo toàn được các thông tin chi tiết từ hình ảnh gốc. Các hạn chế liên quan đến sự suy giảm chất lượng hình ảnh tổng hợp và mất các thông tin chi tiết từ hình ảnh đầu vào có thể quan sát được thông qua các nghiên cứu về tổng hợp hình ảnh hiện nay. Hạn chế đầu tiên liên quan đến chất lượng của hình ảnh đầu vào. Đây là vấn đề chung cho tất cả các thuật toán tổng hợp hình ảnh. Trong thực tế, hình ảnh y học thường gặp phải một số vấn đề về chất lượng do quá trình thu thập, như hình ảnh bị mờ, nhiễu và có độ tương phản thấp. Những hình ảnh có chất lượng thấp có khả năng cao làm ảnh hưởng đến hiệu quả của quá trình tổng hợp cũng như chất lượng của hình ảnh tổng hợp đầu ra thu được. Vì vậy, một số nghiên cứu gần đây đã thực hiện việc cải thiện

chất lượng hình ảnh gốc nhằm cải thiện chất lượng của hình ảnh được tổng hợp, Ví dụ, Ullah và các cộng sự [26] đã sử dụng bộ lọc Laplace cục bộ nhanh (Fast Local Laplacian Filter - FLLF) để tăng cường chất lượng cho hình ảnh đầu vào. Phương pháp tăng cường này giúp bảo toàn và làm nổi bật thông tin chi tiết của các đường biên. Maqsood và các cộng sự [27] đã áp dụng thuật toán cân bằng biểu đồ mức xám được điều chỉnh phi tham số (Non-Parametric Modified Histogram Equalization - NPMHE) để nâng cao độ tương phản của hình ảnh đầu vào. Ngoài ra, Li và các cộng sự [16] đã đề xuất một phương pháp tổng hợp hình ảnh y học cho phép khử nhiễu trước khi thực hiện quá trình tổng hợp ảnh. Tuy nhiên, các phương pháp này là cơ bản và mỗi phương pháp chỉ tập trung vào việc cải thiện một khía cạnh nhất định của chất lượng hình ảnh đầu vào. Các hạn chế của từng phương pháp chưa được xét đến. Ví dụ, việc sử dụng bộ lọc FLLF trong nghiên cứu của Ullah và các cộng sự [26] có thể làm cho hình ảnh đầu vào khếch đại nhiễu vì bộ lọc FLLF nhạy cảm với nhiễu. Việc sử dụng NPMHE trong nghiên cứu của Maqsood và các cộng sự [27] có thể làm cho hình ảnh bị tăng cường quá mức và nhiễu có thể xảy ra. Do đó, những phương pháp nâng cao chất lượng trên chưa thực sự mang lại hiệu quả đáng kể trong việc nâng cao chất lượng của hình ảnh được tổng hợp. Vì vậy, vấn đề đặt ra là cần phải có một phương pháp tăng cường chất lượng ảnh mà đảm bảo hạn chế được nhiễu phát sinh trong quá trình tăng cường ảnh.

Hạn chế thứ hai liên quan đến các phương pháp tổng hợp được thiết kế cho các thành phần cơ sở (thành phần tần số thấp) và thành phần chi tiết (thành phần tần số cao). Đối với việc tổng hợp các thành phần cơ sở, thuật toán trung bình thường được sử dụng vì tính đơn giản và độ phức tạp tính toán thấp. Thuật toán trung bình sẽ sử dụng trọng số 0.5 đối với thành phần cơ sở của cả hai hình ảnh đầu vào và thực hiện lấy tổng sau khi đã nhân trọng số 0.5 để thu được thành phần tổng hợp cơ sở. Thuật toán trung bình giúp cho việc tổng hợp diễn ra nhanh chóng vì độ phức tạp tính toán thấp. Tuy nhiên, việc sử dụng thuật toán tổng hợp trung bình có thể gây ra một số vấn đề nhất định cho hình ảnh tổng hợp, như làm giảm độ tương phản và cường độ sáng của hình ảnh được tổng hợp. Điều này có thể được giải thích như sau: Hình ảnh cộng hưởng từ (MRI) có độ tương phản và cường độ sáng cao hơn hình ảnh chụp cắt lớp phát xạ (PET). Khi thuật toán trung bình được sử dụng mỗi thành phần MRI và PET chỉ đóng góp 50% về cường độ sáng. Do đó sẽ dẫn tới cường độ sáng và độ tương phản của hình ảnh tổng hợp nhỏ cường độ sáng và độ tương phản của hình ảnh MRI. Ví dụ, Zhang và các cộng sự [28] đã sử dụng thuật toán trung bình để tổng hợp cho các thành phần cơ sở dẫn đến hình ảnh được tổng hợp bị suy giảm cường độ sáng và độ tương phản. Để khắc phục hạn chế của thuật toán tổng hợp trung bình, một số

nguyên cứu đã đề xuất sử dụng thuật toán lựa chọn Max cho thành phần cơ sở. Khi đó, thành phần tổng hợp cơ sở sẽ lựa chọn giá trị Max tại mỗi điểm (i, j) của cả hai thành phần cơ sở của hai hình ảnh đầu vào. Mặc dù cường độ sáng và độ tương phản của hình ảnh được cải thiện, nhưng thuật toán lựa chọn Max có thể dẫn đến mất thông tin trong hình ảnh tổng hợp. Lý do là các giá trị trong ma trận của thành phần cơ sở MRI thường cao hơn các giá trị trong ma trận của thành phần cơ sở PET. Do đó khi thuật toán lựa chọn Max được sử dụng, thông tin ở lớp cơ sở của hình ảnh được tổng hợp sẽ được lựa chọn chủ yếu từ hình ảnh MRI mà không sử dụng thông tin từ hình ảnh PET. Ví dụ, Li và các cộng sự [29] đã sử dụng thuật toán lựa chọn Max để tổng hợp cho các thành phần cơ sở dẫn đến hình ảnh được tổng hợp bị mất các thông tin từ hình ảnh đầu vào. Vì vậy, việc tổng hợp cho các thành phần cơ sở vẫn còn có những hạn chế. Hình ảnh được tổng hợp bị suy giảm về mặt chất lượng cũng như bị mất thông tin. Do đó, thách thức đặt ra là cần có một phương pháp tổng hợp cho thành phần cơ sở mà không làm suy giảm cường độ sáng, độ tương phản cũng như mất thông tin của hình ảnh tổng hợp. Đối với việc tổng hợp cho các thành phần chi tiết, các thuật toán lựa chọn Max cũng thường được sử dụng. Tuy nhiên, thuật toán lựa chọn Max sẽ gây ra việc mất thông tin trong hình ảnh tổng hợp. Lý do là bởi vì thành phần chi tiết của hình ảnh MRI thường có cường độ sáng cao hơn so với thành phần chi tiết của hình ảnh PET. Vì vậy, khi áp dụng thuật toán lựa chọn Max, hình ảnh tổng hợp sẽ chứa các thông tin chi tiết từ hình ảnh MRI mà bỏ qua thông tin chi tiết của hình ảnh PET. Để khắc phục hạn chế của thuật toán lựa chọn Max trong việc tổng hợp thành phần chi tiết, một số nghiên cứu gần đây đã đề xuất thuật toán tổng Laplacian được sửa đổi (A sum-modified-Laplacian - SML). Ví dụ, Liu và các cộng sự [30] đã sử dụng thuật toán SML để tổng hợp cho các thành phần chi tiết cho ảnh. Tuy nhiên, những thuật toán này cho thấy kém hiệu quả trong trường hợp ảnh có nhiễu dù ít bởi vì toán tử phát hiện biên Laplace trong thuật toán SML sẽ có xu hướng khuếch tán nhiễu. Do đó, thách thức đặt ra là cần có một phương pháp tổng hợp cho thành phần chi tiết mà giữ được các thông tin chi tiết từ hình ảnh đầu vào.

Vì vậy, việc thiết kế các phương pháp tổng hợp nhằm giải quyết các hạn chế nêu trên chính là nội dung chính được trình bày trong luận án này.

2. Mục tiêu của luận án

Mục tiêu chung của luận án: Đề xuất phương pháp để nâng cao hiệu quả tổng hợp hình ảnh y học.

Mục tiêu cụ thể: Đề xuất một số phương pháp nhằm cải thiện hiệu quả tổng hợp hình ảnh y học như sau:

- Thứ nhất, để giải quyết vấn đề về hình ảnh đầu vào có chất lượng thấp. Luận án đề xuất một phương pháp nâng cao chất lượng hình ảnh y học. Phương pháp tăng cường ảnh đề xuất sử dụng thuật toán tối ưu hoá MPA (Marine Predators Algorithm) [31] nhằm hướng tới cải thiện về mặt cường độ sáng, độ tương phản và chi tiết của hình ảnh cộng hưởng từ não đồng thời hạn chế nhiễu phát sinh trong quá trình tăng cường ảnh.
- Thứ hai, để giải quyết vấn đề chưa hiệu quả trong các thuật toán tổng hợp cho các thành phần cơ sở và các thành phần chi tiết. Luận án đề xuất một phương pháp tổng hợp hình ảnh mới bao gồm 2 thuật toán được thiết kế để tổng hợp cho các thành phần cơ sở và thành phần chi tiết trong ảnh. Đối với việc tổng hợp các thành phần cơ sở, một thuật toán tổng hợp thích nghi dựa trên thuật toán tối ưu hoá MPA được đề xuất. Phương pháp tổng hợp thích nghi nhằm đảm bảo hình ảnh tổng hợp không bị suy giảm về mặt chất lượng. Đối với việc tổng hợp các thành phần chi tiết, đề xuất một thuật toán tổng hợp dựa trên kết hợp các biến thể của hàm năng lượng cục bộ. Phương pháp tổng hợp cho các thành phần chi tiết này nhằm mục đích đảm bảo cho hình ảnh tổng hợp giữ được các thông tin chi tiết từ hình ảnh đầu vào.

3. Đối tượng nghiên cứu của luận án:

Luận án tập trung nghiên cứu và tìm hiểu một số đối tượng liên quan đến việc tổng hợp hình ảnh y học như:

- Một số thuật toán nâng cao chất lượng ảnh, tổng hợp hình ảnh y học.
- Một số thuật toán tối ưu hoá.
- Một số chỉ số đánh giá chất lượng hình ảnh tổng hợp, dữ liệu ảnh y học (cặp ảnh MRI và PET đã được căn chỉnh hình học) dùng cho thực nghiệm.

4. Phạm vi nghiên cứu của luận án

- Nghiên cứu phương pháp nâng cao chất lượng hình ảnh y học dựa trên thuật toán tối ưu hoá để cải thiện chất lượng của hình ảnh tổng hợp.
- Nghiên cứu thuật toán tổng hợp cho các thành phần cơ sở dựa trên thuật toán tối ưu hoá để đảm bảo chất lượng của hình ảnh tổng hợp không bị suy giảm.
- Nghiên cứu thuật toán tổng hợp cho các thành phần chi tiết dựa trên kết hợp các hàm năng lượng cục bộ và các biến thể của nó nhằm đảm bảo cho hình ảnh tổng hợp không bị mất các thông tin chi tiết.
- Các phương pháp tổng hợp được nghiên cứu và đề xuất trong luận án được thiết kế để tổng hợp cho các cặp hình ảnh y học (MRI và PET) đã được căn chỉnh hình học.

- Việc ứng dụng của các hình ảnh y học đã được tổng hợp đối với từng ứng dụng cụ thể trong thực tế không nằm trong nghiên cứu của luận án này.

5. Phương pháp nghiên cứu

Nghiên cứu lý thuyết: Tổng hợp các nghiên cứu liên quan đến bài toán tổng hợp hình ảnh được công bố trong thời gian gần đây. Trên cơ sở phân tích các vấn đề còn tồn tại của các cách tiếp cận tổng hợp hình ảnh hiện nay, đề xuất các phương pháp tổng hợp hình ảnh hiệu quả nhằm mục đích khắc phục các hạn chế còn tồn tại.

Nghiên cứu thực nghiệm: Các phương pháp đề xuất được cài đặt, so sánh và đánh giá với một số thuật toán tổng hợp hình ảnh y học mới đề xuất gần đây để kiểm tra sự hiệu quả của phương pháp tổng hợp ảnh đề xuất. Dữ liệu thực nghiệm là các hình ảnh y học được thu thập từ nguồn dữ liệu “The Whole Brain Atlas” tại địa chỉ: <http://www.med.harvard.edu/AANLIB/>. Nguồn dữ liệu này đã chứa sẵn các cặp hình ảnh y học đã được căn chỉnh hình học theo từng lát cắt ứng với các trục.

6. Các đóng góp của luận án

Với mục tiêu nâng cao sự hiệu quả tổng hợp hình ảnh y học, luận án có hai nhóm đóng góp chính như sau:

- Nhóm thứ nhất, đề xuất một phương pháp nâng cao chất lượng ảnh y học. Phương pháp đề xuất giải quyết một số vấn đề về chất lượng hình ảnh như cường độ sáng và độ tương phản bị thấp. Phương pháp này đã được công bố tại công trình [CT1]. Phương pháp nâng cao chất lượng ảnh đề xuất đã được ứng dụng để tiền xử lý cho các hình ảnh y học đầu vào nhằm cải thiện hiệu quả của việc tổng hợp hình ảnh. Hướng tiếp cận này đã được công bố tại công trình [CT2].
- Nhóm thứ hai, đề xuất một phương pháp tổng hợp hình ảnh mới bao gồm hai thuật toán được thiết kế để tổng hợp cho các thành phần cơ sở và thành phần chi tiết trong ảnh. Đối với việc tổng hợp thành phần cơ sở, thuật toán tổng hợp thích nghi được đề xuất nhằm đảm bảo cho hình ảnh tổng hợp không bị suy giảm chất lượng trong quá trình tổng hợp. Phương pháp tổng hợp cho các thành phần cơ sở này đã được công bố tại các công trình [CT4, CT6]. Đối với việc tổng hợp cho các thành phần chi tiết, thuật toán đề xuất kết hợp các biến thể của hàm năng lượng cục bộ để đảm bảo cho hình ảnh tổng hợp không bị mất các thông tin chi tiết từ hình ảnh đầu vào. Phương pháp tổng hợp cho các thành phần chi tiết này đã được công bố tại các công trình [CT3, CT4]. Bên cạnh đó, phần kết luận của luận án cũng thảo luận về các công việc nghiên cứu tiếp trong tương lai. Trong đó đề cập đến một hướng tiếp cận sử dụng kỹ thuật học chuyển đổi với các mạng học sâu đã được huấn luyện sẵn để tổng hợp cho các

thành phần chi tiết trong ảnh. Hướng tiếp cận này cũng đã được công bố tại công trình [CT5].

7. Bố cục của luận án

Luận án được thiết kế bao gồm phần mở đầu và ba chương nội dung. Thêm vào đó là phần kết luận, danh mục các công trình của tác giả và danh mục các tài liệu tham khảo. Cụ thể, nội dung của các phần được tóm tắt như sau:

Chương 1 sẽ trình bày tổng quan về bài toán tổng hợp hình ảnh và một số các kiến thức nền tảng cần thiết để thuận tiện cho việc hiểu các thuật toán đề xuất ở các chương sau. Các nghiên cứu liên quan đến bài toán tổng hợp hình ảnh y học được trình bày cụ thể và những nghiên cứu đó được phân theo từng nhóm tiếp cận. Từ nhóm các cách tiếp cận truyền thống cho đến các nhóm các cách tiếp cận dựa trên học sâu. Trên cơ sở đó, luận án phân tích các hạn chế của một số phương pháp tổng hợp hình ảnh hiện tại và nêu rõ mục tiêu để cải thiện các hạn chế đó.

Chương 2 và Chương 3 trình bày các đóng góp chính của luận án. Mỗi chương được thiết kế là một nhóm phương pháp đề xuất để cải thiện hiệu quả của việc tổng hợp hình ảnh. Cụ thể, Chương 2 đề xuất một phương pháp tăng cường chất lượng ảnh dựa trên thuật toán tối ưu hoá MPA và ứng dụng của phương pháp tăng cường chất lượng ảnh đề xuất để cải hiệu quả cho một số thuật toán tổng hợp hình ảnh mới nhất hiện nay.

Chương 3 trình bày chi tiết phương pháp tổng hợp ảnh đề xuất để cải thiện hiệu quả cho việc tổng hợp hình ảnh. Phương pháp đề xuất bao gồm hai thuật toán chính được thiết kế để tổng hợp cho các thành phần cơ sở và chi tiết của ảnh. Cụ thể, thuật toán tổng hợp thích nghi cho thành phần cơ sở và thuật toán tổng hợp cho thành phần chi tiết dựa trên việc kết hợp hàm năng lượng cục bộ và các biến thể của nó.

Cuối cùng, phần kết luận nêu các đóng góp chính của tác giả trong luận án và những công việc của tương lai.

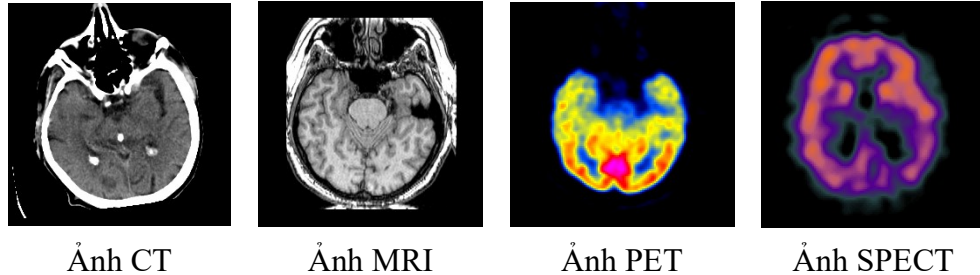
CHƯƠNG 1. TỔNG QUAN VỀ BÀI TOÁN TỔNG HỢP HÌNH ẢNH VÀ MỘT SỐ KIẾN THỨC NỀN TẢNG

Nội dung trong Chương 1 này giới thiệu về bài toán tổng hợp hình ảnh và các nghiên cứu liên quan để giải quyết cho bài toán tổng hợp hình ảnh. Các nghiên cứu liên quan sẽ được phân vào nhóm những phương pháp tiếp cận. Nhìn chung, có hai nhóm tiếp cận chính để giải quyết cho bài toán tổng hợp hình ảnh y học là: nhóm tiếp cận dựa trên phương pháp truyền thống và nhóm tiếp cận dựa trên học sâu. Hạn chế của các phương pháp tổng hợp hình ảnh hiện tại cũng được phân tích chi tiết làm tiền đề cho việc đề xuất các thuật toán ở các chương tiếp theo. Bên cạnh đó, một số kiến thức nền tảng cần thiết để thuận tiện cho việc hiểu thuật toán đề xuất và các chỉ số đánh giá phổ biến cho hình ảnh tổng hợp cũng được trình bày trong chương này.

1.1. Giới thiệu

Ngày nay, việc sử dụng các hình ảnh y học trong chẩn đoán ngày càng trở lên phổ biến. Các loại hình ảnh y học hiện nay cũng rất đa dạng. Một số hình ảnh y học thường dùng trong thực tế có thể được đề cập như: ảnh chụp cắt lớp vi tính (CT), ảnh chụp cắt lớp phát xạ Positron (PET), ảnh chụp cộng hưởng từ (MRI), ảnh chụp cắt lớp đơn Photon (SPECT). Mỗi một loại hình ảnh y học được đề cập ở trên đều chứa những thông tin riêng mà những loại hình ảnh khác không có. Ví dụ, ảnh CT sử dụng tia X để phát hiện thông tin của các cấu trúc mật độ cao như xương. Hình ảnh từ MRI không chỉ mang lại nhiều thông tin quan trọng về cấu trúc và chức năng của các phần khác nhau trong cơ thể, mà còn cung cấp những chi tiết quan trọng về cấu trúc giải phẫu bên trong như não, tim, phổi, xương, cơ, và nhiều cơ quan khác. Ngoài ra, MRI còn khả năng phân biệt các loại mô khác nhau như mô cứng, mô mềm, và mô vôi, từ đó giúp xác định sự phân bố của cấu trúc và nâng cao khả năng chẩn đoán về bệnh lý. Hơn nữa, MRI không chỉ làm rõ về mô cấu trúc mà còn cung cấp hình ảnh trực quan về mạch máu bên trong cơ thể. Khả năng này đặc biệt hữu ích trong việc phát hiện vấn đề về tuần hoàn và lưu thông chất lỏng trong cơ thể. Cuối cùng, ảnh MRI không chỉ là công cụ chẩn đoán mạnh mẽ mà còn có thể phát hiện các bất thường trong cấu trúc và tính chất của các cơ quan, giúp theo dõi và cảnh báo về các tình trạng bệnh lý như ung thư, bệnh tim mạch, bệnh thần kinh, và nhiều loại bệnh khác. Trong khi đó, hình ảnh PET không chỉ cung cấp thông tin về hoạt động của các cơ quan như não, tim, gan, phổi, thận và nhiều bộ phận khác, mà còn giúp hiển thị rõ các khu vực não hoạt động trong các nhiệm vụ cụ thể. PET có khả năng phát hiện sự tích tụ không bình thường của phân tử trong cơ thể, như dấu vết của ung thư, từ đó giúp xác định các vùng bất thường và theo dõi sự thay đổi sau khi áp dụng liệu pháp. Các

hình ảnh PET và SPECT thường được sử dụng trong các nhiệm vụ phát hiện các bệnh về mạch máu và khối u. Hình 1.1 minh họa một số kiểu hình ảnh y học phổ biến thường được dùng trong thực tế. Những hình ảnh này được lấy từ nguồn: “The Whole Brain Atlas”: <https://www.med.harvard.edu/AANLIB>.

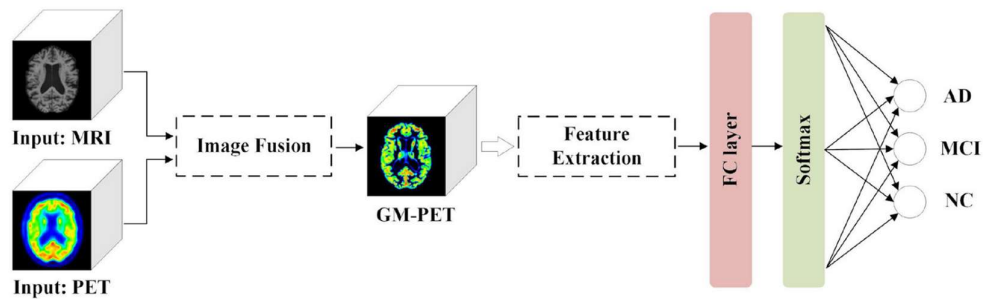


Hình 1.1. Minh họa một số kiểu hình ảnh y học

Trong quá trình chẩn đoán bệnh, mỗi loại hình ảnh y học riêng lẻ không cung cấp đủ thông tin cần thiết để giúp các bác sĩ đưa ra chẩn đoán chính xác về bệnh liên quan. Do đó, việc tạo ra một hình ảnh y học mang thông tin bổ sung từ từng hình ảnh riêng lẻ sẽ giúp cho các bác sĩ có đầy đủ thông tin cần thiết để đưa ra các chẩn đoán chính xác hơn về các loại bệnh liên quan. Tổng hợp hình ảnh y học mang đến rất nhiều ứng dụng. Ứng dụng đầu tiên mà chúng ta có thể thấy là trợ giúp cho các bác sĩ trong việc chẩn đoán. Giả sử chúng ta có một bệnh nhân có triệu chứng mất trí nhớ và được đang được theo dõi về bệnh Alzheimer, một loại bệnh liên quan đến sự suy thoái và mất mát tế bào não. Bác sĩ cho bệnh nhân chụp MRI để quan sát các cấu trúc bên trong của bộ não. Bởi vì hình ảnh MRI có khả năng biểu thị chi tiết các cấu trúc não. Trong trường hợp người bình thường, hình ảnh MRI chụp não không cho thấy bất cứ sự biến đổi khác thường nào trong cấu trúc của não. Tuy nhiên, trong trường hợp người bệnh bị Alzheimer, MRI có thể cung cấp thông tin về việc có sự suy thoái, mất mát mô và sự biến đổi trong cấu trúc não của bệnh nhân. Nếu bác sĩ quan sát được có sự biến đổi trong cấu trúc não của bệnh nhân dựa vào ảnh MRI thì cũng chưa thể biết được chính xác là người bệnh bị chứng mất trí nhớ nhẹ hay là đã bị bệnh Alzheimer. Đồng thời cũng chưa thể đưa ra phác đồ điều trị cho bệnh nhân trong trường hợp này. Khi đó hình ảnh PET được bác sĩ chỉ định chụp cho bệnh nhân. PET sử dụng chất phát xạ positron để theo dõi sự hoạt động của các cơ quan và mô trong cơ thể. Trong trường hợp người bệnh bị Alzheimer, PET có thể sử dụng một loại chất phát xạ mà các vùng não bị tổn thương trong bệnh Alzheimer sẽ thu thập nhiều hơn. Điều này sẽ cho phép nhìn thấy các khu vực của não bị ảnh hưởng nhiều hơn và nhận biết sự thay đổi hoạt động não. Khi tổng hợp cả hai loại hình ảnh này, sự kết hợp cả thông tin về cấu trúc và hoạt động của não sẽ cho bác sĩ có cái nhìn toàn diện hơn về bệnh lý. Điều này giúp hiểu rõ hơn về sự tương quan giữa các biến đổi cấu trúc và

hoạt động của bệnh Alzheimer. Bác sĩ có thể thấy rằng một khu vực của não có cấu trúc bị suy thoái dẫn đến sự suy giảm hoạt động của nó trong hình ảnh PET. Từ đó, ta có thể xác định rằng sự suy thoái cấu trúc này ảnh hưởng đến khả năng hoạt động của vùng đó. Thông tin này có thể hỗ trợ việc đánh giá tình trạng bệnh của bệnh nhân và quản lý điều trị. Tuy nhiên, việc quan sát các vùng quan tâm trong ảnh một cách chính xác từ cả hai hình ảnh MRI và PET đối với bác sĩ là khó khăn và mất thời gian. Hơn nữa, bác sĩ phải quan sát trên một loạt các lát cắt của bộ não tương ứng trên cả hai hình ảnh MRI và PET. Điều này gây khó khăn đối với các bác sĩ trong việc chẩn đoán được chính xác. Vì vậy, việc tổng hợp hình ảnh sẽ tạo ra một hình ảnh chứa được các thông tin bổ sung từ từng hình ảnh riêng lẻ. Nó cung cấp cái nhìn toàn diện hơn về cả cấu trúc và hoạt động của cơ thể, đặc biệt là trong việc theo dõi và chẩn đoán các bệnh liên quan đến não như Alzheimer.

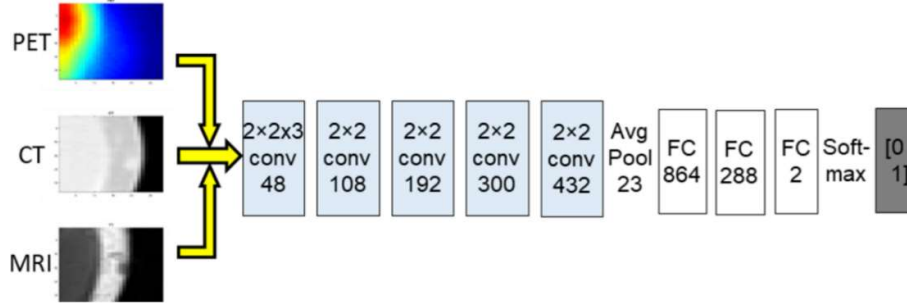
Ứng dụng thứ hai của việc tổng hợp hình ảnh là giúp nâng cao khả năng chẩn đoán trong các hệ thống phân loại bệnh tự động. Thông thường, việc phân lớp giữa người bình thường, người mắc bệnh suy giảm trí nhớ nhẹ và người bệnh Alzheimer sẽ được thực hiện thông qua sử dụng hình ảnh MRI. Tuy nhiên, các nghiên cứu gần đây đã đề xuất sử dụng hình ảnh tổng hợp để nâng cao hiệu quả trong các mô hình chẩn đoán bệnh. Ví dụ, Song và các cộng sự [32] đã tiến hành tổng hợp hai hình ảnh y học MRI và PET sau đó mới xây dựng mạng phân lớp bệnh Alzheimer sử dụng thông tin từ hình ảnh tổng hợp. Cách tiếp cận này đã cho thấy sự cải thiện đáng kể về hiệu quả của mô hình phân lớp bệnh Alzheimer. Hình 1.2 được lấy từ nghiên cứu của Song và các cộng sự [32] để minh họa cho việc sử dụng hình ảnh tổng hợp MRI và PET trong mô hình phân lớp bệnh Alzheimer.



Hình 1.2. Sơ đồ chẩn đoán bệnh có sử dụng tổng hợp hình ảnh

Một ứng dụng khác của tổng hợp hình ảnh y học là ứng dụng phân vùng khối u. Các nghiên cứu trước đây chỉ sử dụng thông tin từ hình ảnh CT hoặc MRI để thực hiện phân vùng khối u. Tuy nhiên, để tăng sự chính xác, các nghiên cứu gần đây đã sử dụng hình ảnh tổng hợp. Ví dụ, Guo và các cộng sự [33] đã sử dụng cả ba hình ảnh MRI, CT và PET tổng hợp với nhau để tạo ra hình ảnh tổng hợp có nhiều thông tin hơn. Sau đó, hình ảnh tổng hợp thu được tiếp tục sử dụng để xác định ảnh có chứa

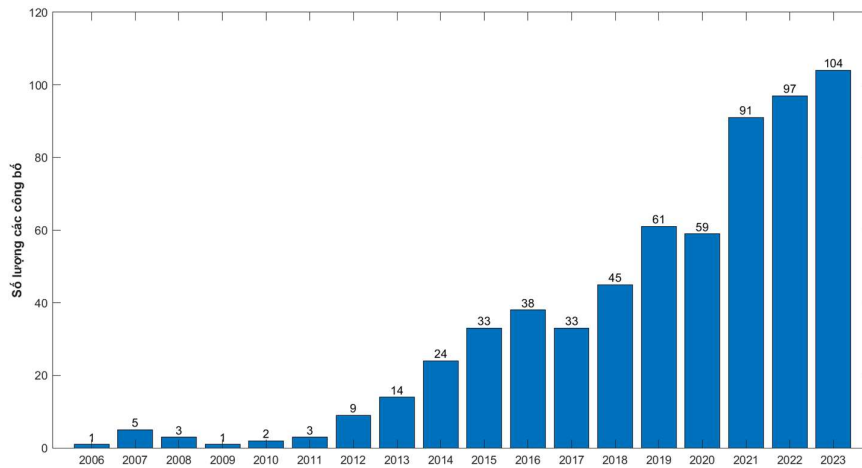
khối u hay không. Các tiếp cận này đã nâng cao hiệu quả của mô hình phân đoạn hình ảnh y học của tác giả. Hình 1.3 được lấy từ nghiên cứu của Guo và các cộng sự [33] để minh họa cho việc sử dụng cả 3 hình ảnh MRI, CT và PET trong mô hình của tác giả Guo và các cộng sự đề xuất.



Hình 1.3. Sơ đồ phân đoạn khối u có sử dụng tổng hợp hình ảnh

Như vậy, mục đích của việc tổng hợp hình ảnh là tạo ra một hình ảnh có chứa nhiều thông tin chi tiết hơn về tình trạng cấu trúc và hoạt động của cơ thể, đặc biệt là trong trường hợp não. Việc tổng hợp hình học có nhiều ứng dụng trong thực tế. Nó không chỉ trợ giúp các bác sĩ nâng cao hiệu quả của việc chẩn đoán mà còn được ứng dụng trong các hệ thống chẩn đoán tự động. Các cách kết hợp thường gặp trong thực tế như: MRI với PET hoặc SPECT, CT với PET, và CT với SPECT.

Trong những năm gần đây, bài toán tổng hợp hình ảnh nhận được sự quan tâm của rất nhiều nhà nghiên cứu trên thế giới. Để thấy được tổng quan các nghiên cứu về tổng hợp hình ảnh y học, số lượng các công bố về tổng hợp hình ảnh y học trong 18 năm trở lại đây (từ năm 2006 đến 09/2023) đã được thu thập từ nguồn dữ liệu Scopus (<https://www.scopus.com/>) với cụm từ truy vấn: “medical image fusion” được tìm trong “Title, Abstract và Keyword” của các bài báo như minh họa trên Hình 1.4.



Hình 1.4. Số lượng các công bố về tổng hợp hình ảnh y học (2006 đến 09/2023)

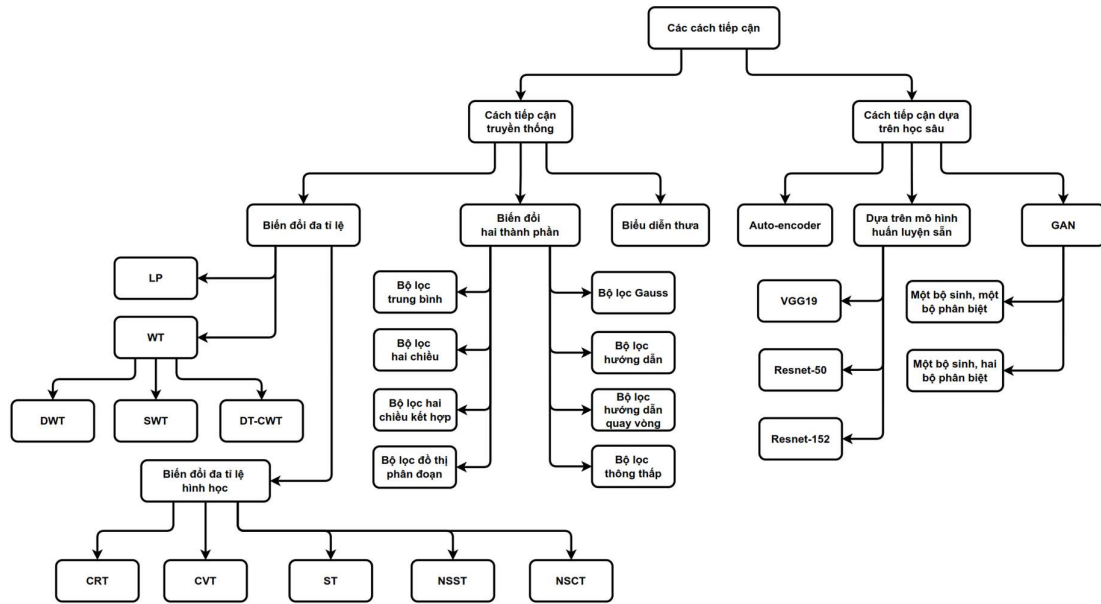
Nhận thấy rằng số lượng các nghiên cứu về tổng hợp hình ảnh y học có thể dao động qua các năm tuy nhiên xu hướng tổng thể là ngày càng tăng. Điều này cho

thấy sự quan tâm và tầm quan trọng của việc nghiên cứu các phương pháp hiệu quả cho bài toán tổng hợp hình ảnh y học.

1.2. Tổng quan các nghiên cứu tổng hợp hình ảnh

Cho đến nay, có rất nhiều các cách tiếp cận khác nhau đã được đề xuất để giải quyết cho bài toán tổng hợp hình ảnh y học. Dựa trên các nghiên cứu [25,34] và [34], Hình 1.5 trình bày sơ đồ các cách tiếp cận chính cho bài toán tổng hợp hình ảnh. Nhìn chung, các phương pháp tổng hợp hình ảnh có thể được phân làm 2 nhóm chính:

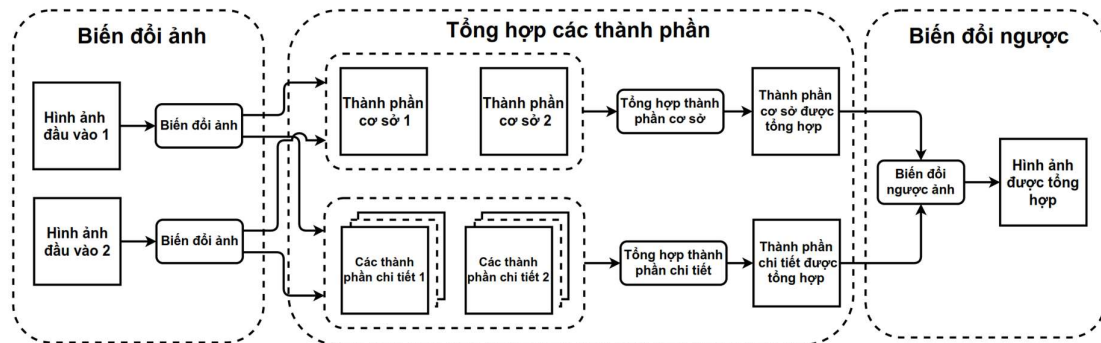
- Các cách tiếp cận truyền thống.
- Các cách tiếp cận dựa trên học sâu.



Hình 1.5. Các cách tiếp cận tổng hợp hình ảnh

1.2.1. Các cách tiếp cận truyền thống

Các cách tiếp cận truyền thống thường phải trải qua 3 bước xử lý chính trong quá trình tổng hợp hình ảnh. Cả ba bước trên được mô tả chi tiết trong Hình 1.6 [34].



Hình 1.6. Sơ đồ chung của quá trình tổng hợp hình ảnh

Bước thứ nhất là biến đổi ảnh. Các hình ảnh đầu được phân rã thành các thành phần cơ sở và các thành phần chi tiết trên miền biến đổi bằng cách sử dụng một phương pháp biến đổi ảnh cụ thể. Bước thứ hai sẽ tiến hành tổng hợp riêng cho các thành phần cơ sở và các thành phần chi tiết bởi các phương pháp tổng hợp chuyên biệt được thiết kế để tổng hợp cho các thành phần đó. Bước thứ ba là biến đổi ngược các thành phần tổng hợp thu được ở bước hai để thu được hình ảnh tổng hợp.

Cho đến nay, có rất nhiều các phương pháp biến đổi ảnh khác nhau. Những phương pháp biến đổi ảnh có thể được phân vào các nhóm như sau:

- Dựa trên biến đổi đa tỉ lệ (MST).
- Dựa trên biến đổi cơ sở - chi tiết (dựa trên các bộ lọc).
- Dựa trên biểu diễn thưa (SR).

1.2.1.1. Các phương pháp dựa trên biến đổi đa tỉ lệ

Các phương pháp biến đổi đa tỉ lệ (MST) sẽ thực hiện việc biến đổi hình ảnh đầu vào thành một số các hình ảnh đầu ra với các tỉ lệ khác nhau. Mục đích của phương pháp biến đổi này là thu được các đặc trưng của ảnh đầu vào dưới các tỉ lệ khác nhau. Các phương pháp thuộc nhóm MST cũng rất đa dạng. Thứ nhất là các cách tiếp cận dựa trên biến đổi kim tự tháp Laplace (LP). Phương pháp LP đã được áp dụng trong một số nghiên cứu tổng hợp hình ảnh. Ví dụ, Li và các cộng sự [3] đã đề xuất một mô hình tổng hợp hình ảnh y học dựa trên kết hợp biến đổi kim tự tháp với phương pháp biểu diễn thưa vào năm 2019. Trong nghiên cứu này, phương pháp biến đổi kim tự tháp được thiết lập với 4 cấp độ tỉ lệ để phân rã cho các hình ảnh đầu vào thành các thành phần chi tiết và thành phần cơ sở. Sau đó, phương pháp biểu diễn thưa (SR) được áp dụng để tổng hợp cho các thành phần cơ sở với thuật toán tổng hợp là lựa chọn Max-L1. Các thành phần chi tiết được tổng hợp bởi thuật toán tổng Laplace được sửa đổi (SML). Wang và các cộng sự [35] đã đề xuất sử dụng biến đổi kim tự tháp kết hợp với phương pháp biểu diễn thưa thích nghi để xây dựng một mô hình tổng hợp hình ảnh vào năm 2020. Các hình ảnh đầu vào được biến đổi kim tự tháp với 4 cấp để thu được 4 cấp tỉ lệ khác nhau. Hình ảnh có tỉ lệ nhỏ (độ phân giải thấp) được đặt phía trên đỉnh của kim tự tháp, hình ảnh có tỉ lệ lớn nhất (độ phân giải lớn nhất) được đặt ở đáy kim tự tháp. Sau đó, các hình ảnh sau khi được biến đổi kim tự tháp mà có cùng tỉ lệ sẽ được tổng hợp lại với nhau bằng cách sử dụng phương pháp biểu diễn thưa thích nghi. Cuối cùng, các thành phần sau khi tổng hợp được biến đổi ngược kim tự tháp để thu được hình ảnh tổng hợp đầu ra.

Thứ hai là các cách tiếp cận dựa trên biến đổi Wavelet rời rạc (DWT). Việc ứng dụng phương pháp biến đổi Wavelet rời rạc cho bài toán tổng hợp hình ảnh đã xuất hiện trong một số nghiên cứu tổng hợp hình ảnh. Ví dụ, Xu và các cộng sự [36]

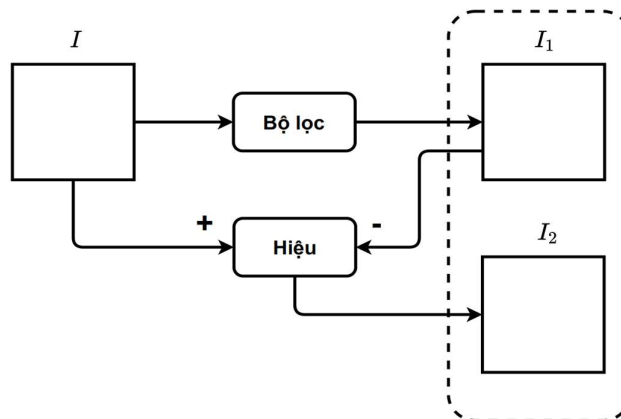
đã kết hợp biến đổi DWT với phương pháp tối ưu hoá SSO (Shark smell optimization) và WCO (World cup optimization) để tổng hợp hình ảnh y học. Tuy nhiên, phép biến đổi DWT không phải là bất biến dịch chuyển, vì các hệ số Wavelet của DWT thay đổi khi tín hiệu được dịch chuyển. Vì vậy, nó gây ra các đặc trưng ngoài ý muốn trong hình ảnh tổng hợp. Biến đổi SWT (Stationary wavelet transform) rời rạc là một phép biến đổi bất biến dịch chuyển. Vì vậy, nó đã khắc phục được hạn chế của phép biến đổi DWT và được áp dụng trong một số mô hình tổng hợp hình ảnh. Chao và các cộng sự [37] đã đề xuất một phương pháp tổng hợp hình ảnh dựa trên biến đổi SWT và mạng hàm cơ sở bán kính (Radial Basis Functions networks – RBF). Trong nghiên cứu này, phép biến đổi SWT được sử dụng ở mức thứ hai cho hình ảnh đầu vào để thu được thành phần cơ sở và 6 thành phần chi tiết. Các thành phần cơ sở và chi tiết đều được đưa qua mạng RBF để tổng hợp. Các thành phần sau khi được tổng hợp bởi mạng RBF sẽ được biến đổi ngược SWT để thu được hình ảnh tổng hợp. Một số cách tiếp cận tổng hợp hình ảnh khác thuộc nhóm Wavelet có thể tìm thấy trong các nghiên cứu như DWT [38], DT-CWT [39]. Nhìn chung, các cách tiếp cận dựa trên LP hoặc biến đổi Wavelet vẫn còn có những hạn chế nhất định trong việc biểu diễn các đặc trưng của hình ảnh đầu vào. Việc thiếu các thông tin về hướng dẫn đến mất các thông tin chi tiết trong hình ảnh tổng hợp.

Thứ ba là các cách tiếp cận dựa trên phân tích hình học đa tỉ lệ (MSGa) được áp dụng rộng rãi hơn trong các nghiên cứu về tổng hợp hình ảnh. Trong nhóm này, một số phương pháp biến đổi phổ biến có thể kể đến như CVT, CRT, ST, NSCT và NSST. Tuy nhiên, các phép biến đổi CVT, CRT và ST vẫn có nhược điểm là không có tính chất bất biến dịch chuyển bởi vì quá trình lấy mẫu xuống. Vì vậy, hai phương pháp biến đổi NSCT và NSST thường được ứng dụng phổ biến trong bài toán tổng hợp hình ảnh y học. Ví dụ, Li và các cộng sự [40] đã giới thiệu một mô hình tổng hợp hình ảnh dựa trên phương pháp biến đổi NSCT và hệ thống P-nơ-ron ngưỡng động (DTNPS) vào năm 2021. Trong nghiên cứu này, các hình ảnh đầu vào được biến đổi bởi phương pháp NSCT để thu được các thành phần tần số thấp và các thành phần tần số cao. Các thành phần tần số thấp được tổng hợp bởi một thuật toán cải tiến từ thuật toán tổng Laplace được sửa đổi (gọi là INSML). Các thành phần tần số cao được tổng hợp bởi hệ thống DTNPS. Các thành phần tần số thấp và cao đã tổng hợp tiếp tục được biến đổi ngược NSCT để thu được hình ảnh tổng hợp. Jose và các cộng sự [13] đã ứng dụng phương pháp biến đổi NSST và kết hợp nó với thuật toán tối ưu AIS (Adolescent identity search algorithm) để tạo ra mô hình tổng hợp hình ảnh vào năm 2021. Phương pháp NSST được sử dụng để biến đổi các hình ảnh đầu vào, thu được các thành phần tần số thấp và các thành phần tần số cao. Các thành phần tần số

thấp được tổng hợp dựa trên một hàm năng lượng cục bộ. Các thành phần tần số cao được xử lý bởi toán tử hình thái sau đó sử dụng thuật toán tối ưu hoá AIS để tổng hợp. Các thành phần tần số thấp và cao sau khi được tổng hợp sẽ được áp dụng biến đổi ngược NSST để thu được hình ảnh tổng hợp. Một số các nghiên cứu tổng hợp hình ảnh khác dựa trên MSGA có thể được đề cập như: NSCT [41], [22], [42], [43], [44] và NSST [8], [19], [45], [46], [47], [10], [12], [48], [9]. Nhìn chung, các cách tiếp cận dựa trên MSGA có một số nhược điểm. Nhược điểm thứ nhất là nó có thể mất các thông tin chi tiết. Lý do là bởi các tần số cao có thể bị giảm đáng kể, dẫn đến việc mất các đặc trưng nhỏ hoặc chi tiết quan trọng trong dữ liệu. Nhược điểm thứ hai là chi phí tính toán cao. Đặc biệt là khi biến đổi ảnh thành nhiều tỉ lệ khác nhau. Do đó, các thuật toán tổng hợp hình ảnh dựa trên biến đổi đa tỉ lệ thường tốn thời gian chạy do phải thực hiện quá trình phân rã hình ảnh đầu vào [49].

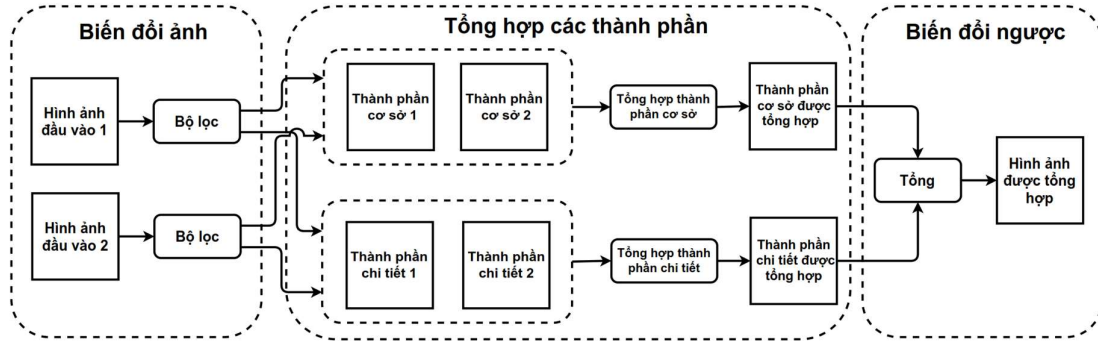
1.2.1.2. Các phương pháp dựa trên biến đổi cơ sở - chi tiết

Phân rã ảnh cơ sở-chi tiết là kỹ thuật phân rã một hình ảnh đầu vào thành hai thành phần (cơ sở và chi tiết) hay còn gọi là phương pháp dựa trên bộ lọc. Trong đó, thành phần cơ sở chứa thông tin tổng thể về hình dạng và cấu trúc chính của ảnh. Nó cho thấy các thay đổi chậm về cường độ sáng trong ảnh. Thành phần chi tiết chứa sự biến đổi nhanh về cường độ sáng trong ảnh. Cách tiếp cận này đã cho thấy sự hiệu quả cả về hiệu năng tổng hợp hình ảnh và hiệu quả tính toán khi so với các cách tiếp cận dựa trên các phương pháp biến đổi đa tỉ lệ. Để thực hiện phân rã ảnh theo kỹ thuật này, các bộ lọc thường được sử dụng như bộ lọc trung bình, bộ lọc Gauss, bộ lọc thông thấp (Low-pass filter - LPF), bộ lọc hai chiều (Bilateral filter - BF), bộ lọc hai chiều kết hợp (Cross Bilateral filter - CBF), bộ lọc hướng dẫn (Guided filter - GF), bộ lọc hướng dẫn quay vòng (Rolling guidance filter - RGF). Hình 1.7 [34] minh họa cho sơ đồ phân rã một hình ảnh đầu vào (I) thành hai thành phần, cơ sở (I_1) và chi tiết (I_2) theo thuật toán phân rã cơ sở - chi tiết.



Hình 1.7. Sơ đồ chung cho các thuật toán phân rã hình ảnh cơ sở - chi tiết

Các phương pháp tổng hợp hình ảnh dựa trên bộ lọc được thực hiện như minh hoạ trong Hình 1.8 [34]. Các hình ảnh đầu vào được tách thành hai thành phần cơ sở và chi tiết dựa trên một bộ lọc cho trước. Các thành phần cơ sở được tổng hợp bởi thuật toán tổng hợp được thiết kế riêng cho các thành phần cơ sở và các thành phần chi tiết được tổng hợp bởi thuật toán được thiết kế riêng cho các thành phần chi tiết. Các thành phần cơ sở và chi tiết sau khi được tổng hợp sẽ được lấy tổng lại với nhau để thu được hình ảnh tổng hợp.

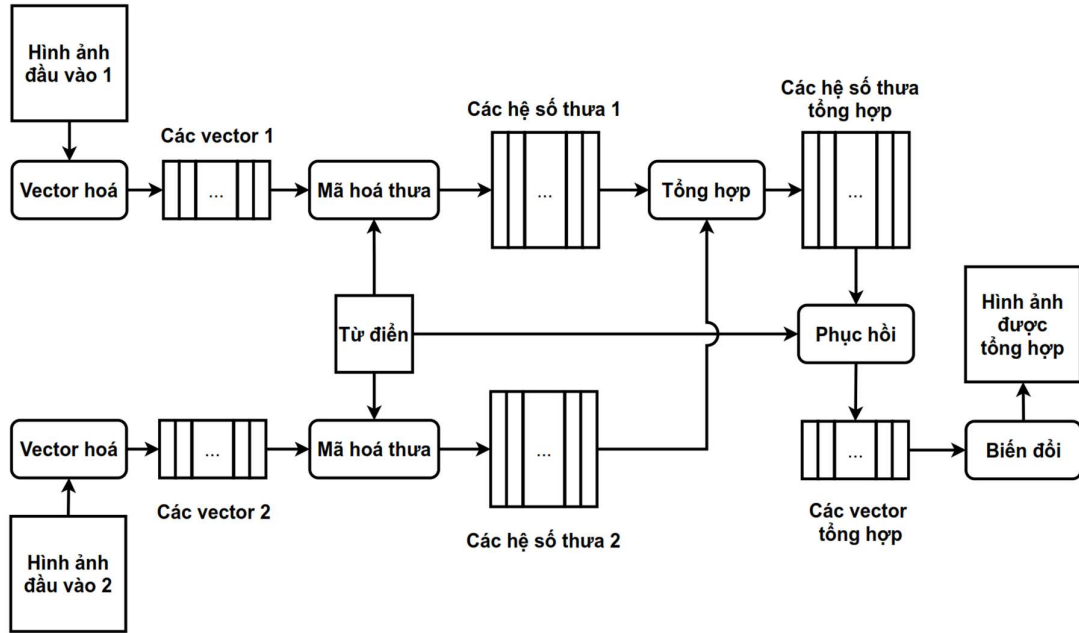


Hình 1.8. Sơ đồ chung cho các thuật toán tổng hợp hình ảnh dựa trên bộ lọc

Các cách tiếp cận dựa trên biến đổi cơ sở - chi tiết đã được ứng dụng trong rất nhiều các nghiên cứu về tổng hợp hình ảnh. Ví dụ, Li và các cộng sự [50] đã sử dụng bộ lọc thông thấp để phân rã cách hình ảnh y học đầu vào thành hai thành phần. Các thành phần cơ sở được tổng hợp dựa trên phương pháp biểu diễn thưa. Các thành phần chi tiết được tổng hợp với nhau bởi một hàm năng lượng cục bộ. Li và các cộng sự [29] sử dụng bộ lọc hướng dẫn quay vòng RGF để tách các hình ảnh đầu vào ra thành các thành phần cơ sở và chi tiết. Các thành phần cơ sở được tổng hợp lại với nhau bằng thuật toán lựa chọn Max. Các thành phần chi tiết được tổng hợp lại với nhau bởi một thuật toán dựa trên hàm năng lượng Gradient cục bộ. Một số các nghiên cứu tổng hợp hình ảnh khác dựa trên cách tiếp cận phân rã hình ảnh thành hai thành phần cơ sở và chi tiết mà sử dụng các bộ lọc có thể được đề cập như bộ lọc LPF [51], [52], bộ lọc RGF [53], [54], bộ lọc SGF (Segment graph filter) [55], bộ lọc Gauss [56], bộ lọc CBF [57], [58], và bộ lọc GF [44]. Việc phân rã hình ảnh dựa trên các bộ lọc cũng có một số nhược điểm nhất định. Nhược điểm thứ nhất là khó khăn trong việc lựa chọn các tham số của bộ lọc cho phù hợp. Trên thực tế, để lựa chọn các tham số của bộ lọc cho phù hợp ứng với từng bài toán cụ thể thông qua việc chạy thực nghiệm là tốn thời gian. Hạn chế thứ hai là các bộ lọc có thể gây mất thông tin khi phân rã nếu kích thước của cửa sổ lọc lớn và số lượng thành phần cần phân rã nhiều [49].

1.2.1.3. Các phương pháp dựa trên biểu diễn thưa

Biểu diễn thưa (SR) là một kỹ thuật cho phép biểu diễn một tín hiệu có thể xấp xỉ dưới dạng tổ hợp tuyến tính của một số lượng nhỏ các thành phần chính bằng cách sử dụng một từ điển dư thừa. Hình 1.9 [34] minh họa sơ đồ tổng hợp hình ảnh dựa trên phương pháp biểu diễn thưa.



Hình 1.9. Sơ đồ cho thuật toán tổng hợp hình ảnh dựa trên biểu diễn thưa

Việc tổng hợp hình ảnh dựa trên biểu diễn thưa được thực hiện qua một số bước. Bước đầu tiên là trích xuất các khối ảnh để vector hoá. Mỗi hình ảnh đầu vào được tách thành các khối có kích thước 8×8 được lấy từ trên xuống dưới, từ trái sang phải. Số lượng điểm ảnh chồng nhau được thiết lập là 7. Mỗi khối được chuẩn hoá và chuyển về một vector dạng cột. Bước tiếp theo là mã hoá thưa cho các vector ở bước đầu tiên, thu được các hệ số thưa bằng cách sử dụng một từ điển. Sau đó, các hệ số thưa thu được sẽ được tổng hợp lại với nhau để thu được các hệ số thưa tổng hợp. Các hệ số thưa tổng hợp này sẽ được phục hồi thành các vector tổng hợp bằng cách sử dụng từ điển. Cuối cùng các vector đã tổng hợp sẽ được biến đổi để thu được hình ảnh tổng hợp. Phương pháp biểu diễn thưa cũng có thể được áp dụng để tổng hợp cho thành phần cơ sở hoặc chi tiết trong ảnh. Phương pháp biểu diễn thưa đã được ứng dụng trong nhiều nghiên cứu về tổng hợp hình ảnh y học. Ví dụ, Li và các cộng sự [50] đã đề xuất một phương pháp tổng hợp hình ảnh dựa trên biểu diễn thưa và hàm năng lượng cục bộ vào năm 2023. Trong nghiên cứu này, các hình ảnh đầu vào được phân rã bởi sử dụng một bộ lọc thông thấp để thu được các thành phần cơ sở và các thành phần chi tiết. Các thành phần cơ sở được tổng hợp lại với nhau bằng

cách sử dụng phương pháp biểu diễn thưa. Các thành phần chi tiết được tổng hợp lại với nhau bởi một phương pháp tổng hợp dựa trên hàm năng lượng cục bộ. Sau đó, các thành phần cơ sở và chi tiết đã được tổng hợp từ bước trước được lấy tổng lại với nhau để thu được hình ảnh tổng hợp. Barba-J và các cộng sự [18] đã đề xuất một phương pháp tổng hợp hình ảnh dựa trên biểu diễn thưa. Các hình ảnh đầu vào được biến đổi thành các thành phần tần số thấp và tần số cao bởi biến đổi Hermite rời rạc. Sau đó thuật toán lựa chọn Max được áp dụng để tổng hợp thành phần tần số cao. Phương pháp SR được áp dụng để tổng hợp cho các thành phần tần số thấp. Việc sử dụng phương pháp SR để tổng hợp hình ảnh y học có thể tìm thấy trong một số các nghiên cứu khác như: [53], [59], [60], [8], [15], [55], [19], [16], [35], [54], [27]. Tuy nhiên, các cách tiếp cận dựa trên biểu diễn thưa cũng đã cho thấy hạn chế về thời gian thực hiện thuật toán [1].

1.2.2. Các cách tiếp cận dựa trên học sâu

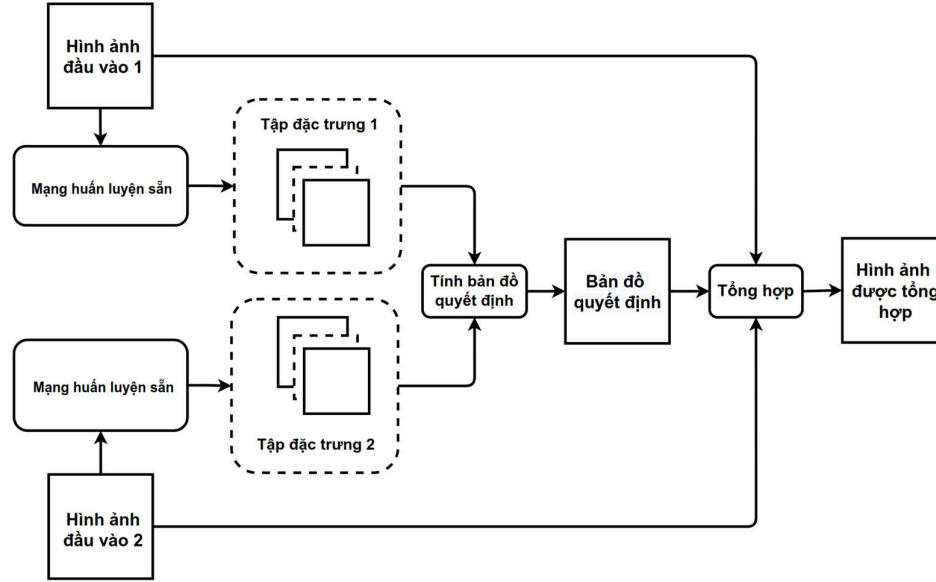
Các cách tiếp cận dựa trên mô hình học sâu đã được ứng dụng trong rất nhiều các bài toán của xử lý ảnh. Đặc biệt trong bài toán tổng hợp hình ảnh, cách tiếp cận dựa trên học sâu được áp dụng để giải quyết cho bài toán tổng hợp hình ảnh có thể phân làm hai nhóm theo Zhou và các cộng sự [25].

- Nhóm 1: Các mạng học sâu chỉ tham gia vào một giai đoạn cụ thể trong quá trình tổng hợp.
- Nhóm 2: Các mạng học sâu được thiết kế để tổng hợp hình ảnh theo một mô hình khép kín.

1.2.2.1. Các cách tiếp cận dựa Nhóm 1

Trong cách tiếp cận này, các mạng học sâu được ứng dụng trong việc trích xuất và biểu diễn các đặc trưng các cho hình ảnh đầu vào. Việc trích xuất đặc trưng có thể tiến hành cho các hình ảnh đầu vào hoặc chỉ trích xuất cho một trong các thành phần (cơ sở hoặc chi tiết) của hình ảnh đầu vào.

Đối với trường hợp mạng học sâu được sử dụng để trích xuất đặc trưng cho toàn bộ hình ảnh đầu vào. Các bước trong mô hình tổng hợp hình ảnh được mô tả như sau: Đầu tiên, các hình ảnh đầu vào được đẩy qua các mạng học sâu đã được huấn luyện sẵn để thu được tập các đặc trưng với các tỉ lệ khác nhau ở các tầng tích chập. Sau đó, các đặc trưng thu được trải qua một số bước xử lý để thu được bản đồ quyết định. Cuối cùng, bản đồ quyết định được sử dụng để tổng hợp cho 2 hình ảnh đầu vào. Hình 1.10 [25] minh họa cho một mô hình tổng hợp hình ảnh trong đó các mạng học sâu được sử dụng để trích xuất đặc trưng cho hai hình ảnh đầu vào.

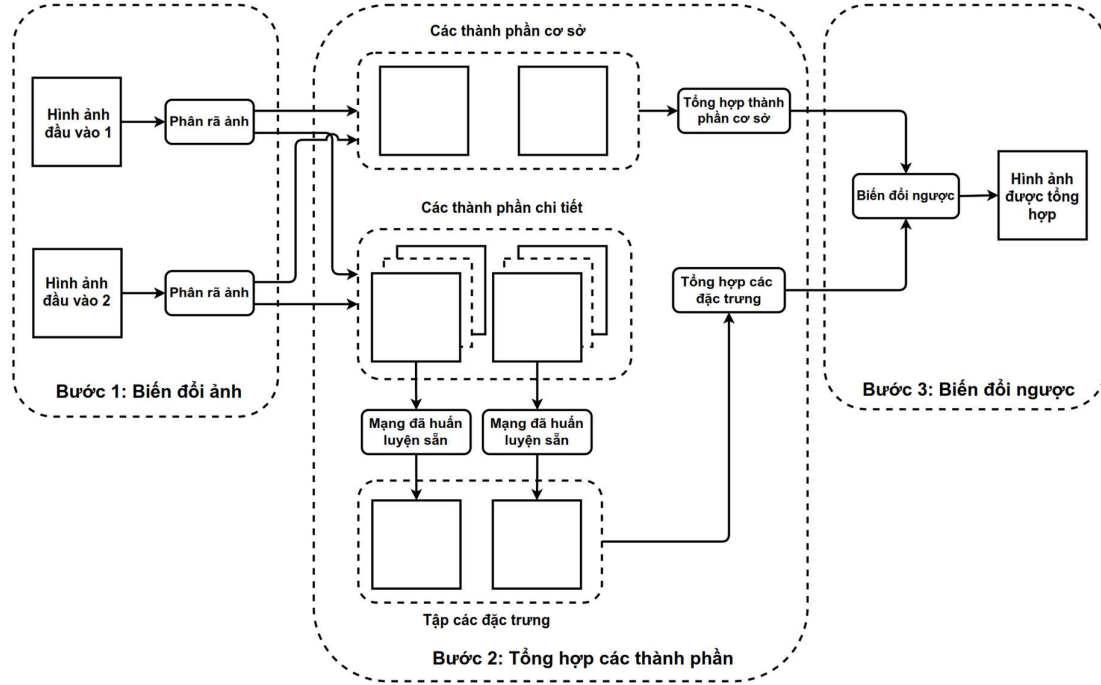


Hình 1.10. Sơ đồ mạng học sâu được sử dụng để trích xuất đặc trưng cho các hình ảnh đầu vào trong mô hình tổng hợp hình ảnh.

Các tiếp cận trên có thể tìm thấy trong nghiên cứu của Zhou và các cộng sự [61]. Trong đó, một mạng VGG19 đã được huấn luyện sẵn dùng để trích xuất các đặc trưng trong phương pháp tổng hợp hình ảnh được đề xuất bởi tác giả. Các hình ảnh đầu vào được đẩy qua mạng VGG19, trong đó 5 lớp tích chập (Conv1_1, Conv1_2, Conv2_1, Conv3_1 và Conv4_1) được sử dụng để thu được tập các đặc trưng với các tỉ lệ khác nhau. Sau đó, toán tử trung bình và chuẩn L_1 được sử dụng để thao tác trên tập các đặc trưng nhằm thu được từ mỗi tầng tích chập một hình ảnh đặc trưng duy nhất. Tiếp theo các đặc trưng từ mỗi tầng tích chập sẽ được chuyển về cùng tỉ lệ với hình ảnh đầu vào (thao tác up-sampling). Phương pháp Softmax được sử dụng đối với các đặc trưng để thu được bản đồ quyết định cho các thành phần đặc trưng. Bản đồ quyết định này sau đó được áp dụng để tổng hợp cho các đặc trưng của từng tầng tích chập. Cuối cùng, hình ảnh tổng hợp thu được bằng cách sử dụng toán tử Max đối với các hình ảnh tổng hợp của từng tầng tích chập.

Đối với trường hợp mạng học sâu được sử dụng để trích xuất đặc trưng cho một thành phần của hình ảnh đầu vào. Hình 1.11 [25] minh họa cho sơ đồ tổng hợp hình ảnh trong trường hợp này. Trong đó mạng học sâu được sử dụng để trích xuất đặc trưng cho các thành phần tần số cao. Các bước trong mô hình tổng hợp hình ảnh được mô tả như sau: Đầu tiên, các hình ảnh đầu vào sẽ được phân rã thành các thành phần cơ sở và thành phần chi tiết bởi một thuật toán phân rã ảnh cụ thể. Sau đó, các thành phần cơ sở được tổng hợp lại với nhau bởi một thuật toán tổng hợp được thiết kế trước. Các thành phần chi tiết sẽ được đẩy qua một mạng học sâu được huấn luyện sẵn. Sau đó tập các đặc trưng thu được sẽ trải qua một số bước xử lý để thu được bản

đồ quyết định và nó được sử dụng cho việc tổng hợp các thành phần chi tiết. Cuối cùng, các thành phần tổng hợp cơ sở và chi tiết sẽ được biến đổi ngược của phương pháp phân rã ảnh ban đầu để thu được hình ảnh tổng hợp.



Hình 1.11. Sơ đồ mạng học sâu được sử dụng để trích xuất đặc trưng cho thành phần chi tiết trong mô hình tổng hợp hình ảnh.

Một số các nghiên cứu tổng hợp hiện nay đã áp dụng theo cách này. Ví dụ, mạng học sâu Resnet-152 được sử dụng trong việc trích xuất đặc trưng cho thành phần chi tiết có thể tìm thấy trong nghiên cứu của Zhang [28]. Đầu tiên, một bộ lọc thông thấp được áp dụng cho các hình ảnh đầu vào. Thuật toán tổng hợp trung bình được áp dụng để tổng hợp cho các thành phần cơ sở. Các thành phần chi tiết được đẩy vào mạng Resnet-152 để trích xuất các đặc trưng bởi sử dụng 5 tầng tích chập. Chuẩn L_1 và toán tử trung bình được áp dụng cho các đặc trưng lấy ra ở các tầng tích chập, thu được một hình ảnh đặc trưng duy nhất ứng với từng tầng tích chập mà nó có các kích thước khác nhau. Sau đó, kỹ thuật nội suy song tuyến tính được sử dụng để chuyển các đặc trưng với các kích thước khác nhau về cùng với kích thước của hình ảnh đầu vào. Toán tử Max được sử dụng đối với các đặc trưng này để thu được bản đồ quyết định cho việc tổng hợp. Sau đó, nó được sử dụng để tổng hợp cho hai thành phần chi tiết. Từ các thành phần tổng hợp cơ sở và chi tiết sẽ được kết hợp lại với nhau để thu được hình ảnh tổng hợp.

Nhìn chung, các cách tiếp cận dựa trên một mạng đã được huấn luyện sẵn (như VGG-16, VGG-19, Resnet-50, Resnet-101,...) để trích xuất đặc trưng thì sẽ phụ thuộc vào chất lượng của mạng học sâu đó trong việc biểu diễn đặc trưng. Trên thực

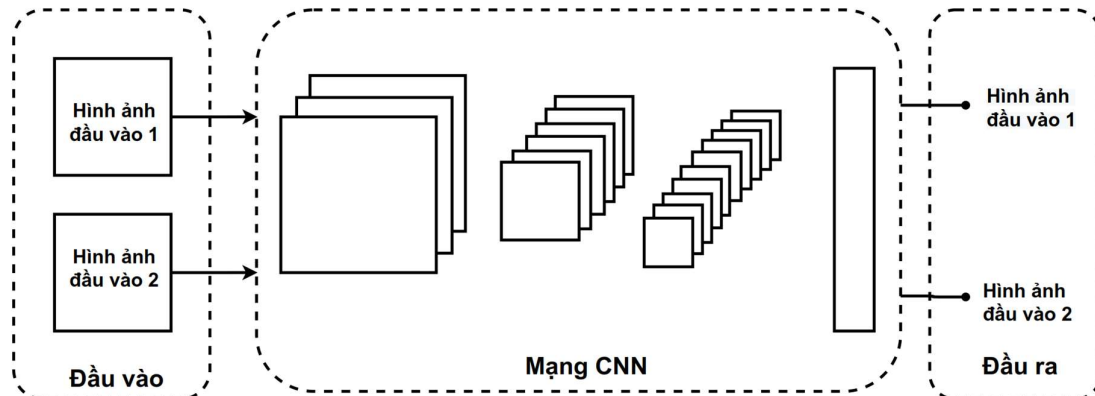
thế, các mạng học sâu này được huấn luyện dựa trên một tập dữ liệu lớn mà không phải từ chính tập các hình ảnh y học (như MRI, CT, PET và SPECT). Do đó, các mạng học sâu được huấn luyện sẵn có thể chưa biểu diễn tốt các đặc trưng cho các hình ảnh y học. Để khắc phục hạn chế về ít dữ liệu, một số giải pháp có thể được kể đến như: Thứ nhất, sử dụng kỹ thuật học chuyển đổi dựa trên các mạng đã được huấn luyện sẵn như: VGG-16, VGG-19, Resnet-18, Resnet-50, Resnet-101, Resnet-152, DenseNet-201... Thứ hai, chia mỗi hình ảnh thành các vùng nhỏ hơn, điều này sẽ làm tăng thêm dữ liệu để huấn luyện.

1.2.2.2. Các cách tiếp cận dựa Nhóm 2

Các cách tiếp cận thuộc nhóm này sẽ học cách để tổng hợp các hình ảnh đầu vào thành một hình ảnh đầu ra theo một mô hình khép kín. Các mạng học sâu được sử dụng trong cách tiếp cận này sẽ được huấn luyện từ tập hình ảnh dữ liệu đầu vào để học các trọng số. Những mạng học sâu này có thể kể đến như: Mạng nơ ron tích chập (CNN), mạng mã hoá-giải mã, mạng sinh đối kháng (GAN).

Các cách tiếp cận dựa trên mạng nơ ron tích chập (CNN)

Phương pháp tổng hợp hình ảnh dựa trên mạng CNN thực hiện trích xuất các đặc trưng ẩn, tổng hợp các đặc trưng và tái tạo hình ảnh bằng cách thiết kế cấu trúc mạng và hàm mất mát. Các tiếp cận dựa trên mạng CNN cho phép tránh các hạn chế của các thuật toán tổng hợp được thiết kế thủ công. Trong giai đoạn huấn luyện, các trọng số của mạng CNN được học dựa trên dữ liệu ảnh đầu vào để có thể phân được hai lớp ứng với hai kiểu hình ảnh đầu vào. Hình 1.12 [25] minh họa cho sơ đồ mạng CNN được huấn luyện để có thể phân hai lớp ứng với hai kiểu hình ảnh đầu vào cần tổng hợp.

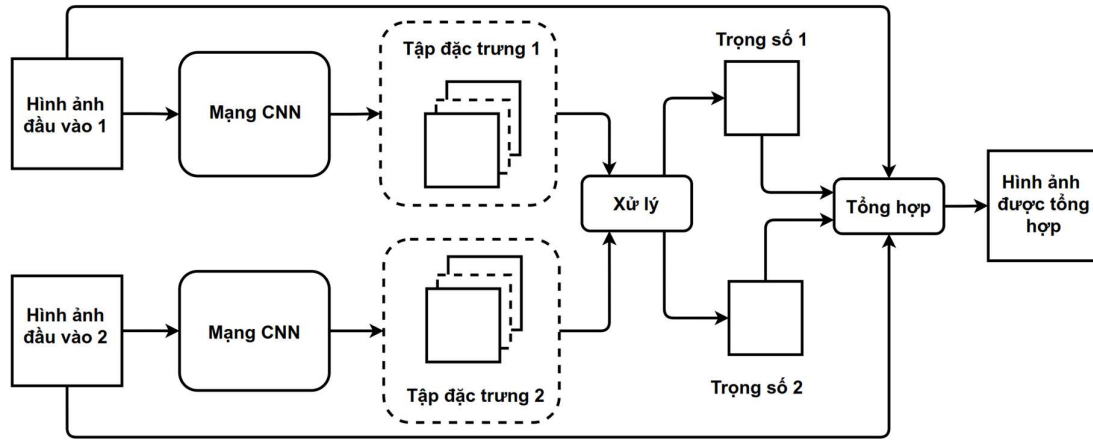


Hình 1.12. Sơ đồ mạng CNN phân 2 lớp

Quá trình tổng hợp hình ảnh sau khi mạng CNN học được các trọng số được mô tả trong Hình 1.13 [25]. Các bước tổng hợp hình ảnh dựa trên mạng CNN được mô tả như sau: Đầu tiên, các hình ảnh đầu vào được đưa vào mạng CNN để trích xuất

các đặc trưng. Sau đó, các đặc trưng thu được từ các lớp của mạng sẽ được giảm chiều bởi thuật toán trung bình để thu được các đặc trưng hai chiều. Tiếp theo, các đặc trưng này được nội suy để điều chỉnh kích thước bằng với kích thước của ảnh đầu vào. Sau đó, thuật toán lựa chọn Max thường được sử dụng để thu được ma trận trọng số phục vụ cho việc tổng hợp. Cuối cùng, ma trận trọng số và các hình ảnh đầu vào sẽ được tính toán để thu được hình ảnh tổng hợp đầu ra.

Trong quá trình tổng hợp, đầu ra của các kết quả trung gian là không cần thiết và CNN học cách ánh xạ trực tiếp từ đầu vào đến đầu ra. So với các thuật toán kết hợp hình ảnh truyền thống, CNN có thể thích ứng với các nhiệm vụ tổng hợp hình ảnh bằng cách học các tham số thích hợp của các bộ lọc tích chập và các tham số của mô hình CNN có thể được tối ưu hóa thông qua việc huấn luyện.

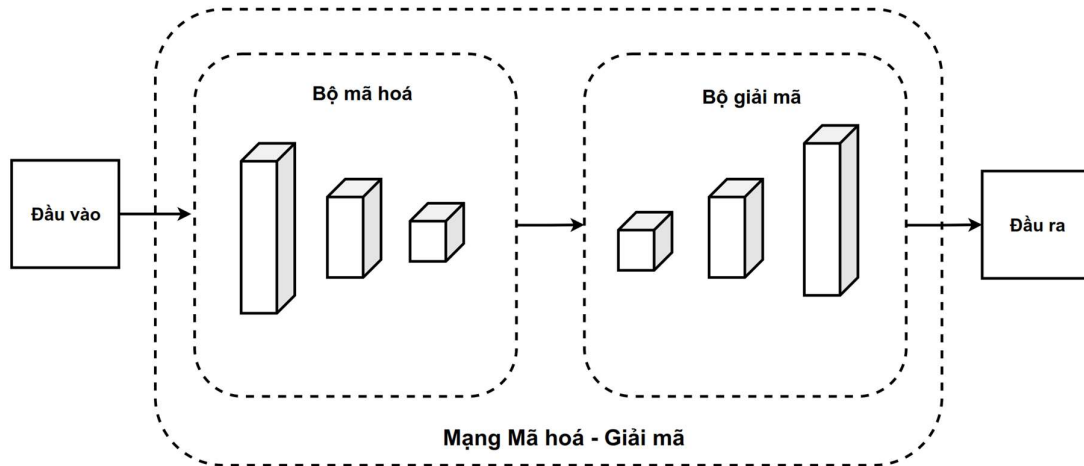


Hình 1.13. Sơ đồ của mô hình tổng hợp sử dụng mạng CNN

Hơn nữa, mạng CNN có thể được sử dụng để trích xuất đặc trưng cho một thành phần của ảnh đầu vào trong mô hình tổng hợp hình ảnh. Ví dụ, Wang và các cộng sự [41] đã đề xuất một mạng CNN để tổng hợp cho các thành phần chi tiết trong ảnh. Trong nghiên cứu của Wang, các hình ảnh đầu vào được biến đổi NSCT để thu được các thành phần cơ sở và chi tiết. Để tổng hợp cho các thành phần cơ sở, Wang đã áp dụng mạng tích chập đầy đủ dùng cho phân đoạn hình ảnh mà đã được đề xuất bởi Long và các cộng sự [62] để tạo ra bản đồ quyết định và sử dụng nó để tổng hợp cho các thành phần cơ sở. Mạng CNN do Wang đề xuất được áp dụng đối với các thành phần chi tiết để thu được bản đồ quyết định cho các thành phần chi tiết. Sau đó, áp dụng nó để tổng hợp cho các thành phần chi tiết. Cuối cùng, các thành phần cơ sở và chi tiết đã được tổng hợp sẽ được biến đổi ngược NSCT để thu được hình ảnh tổng hợp. Một số nghiên cứu khác sử dụng mạng CNN để tổng hợp hình ảnh như: Ding và các cộng sự [63] và Yousif và các cộng sự [59].

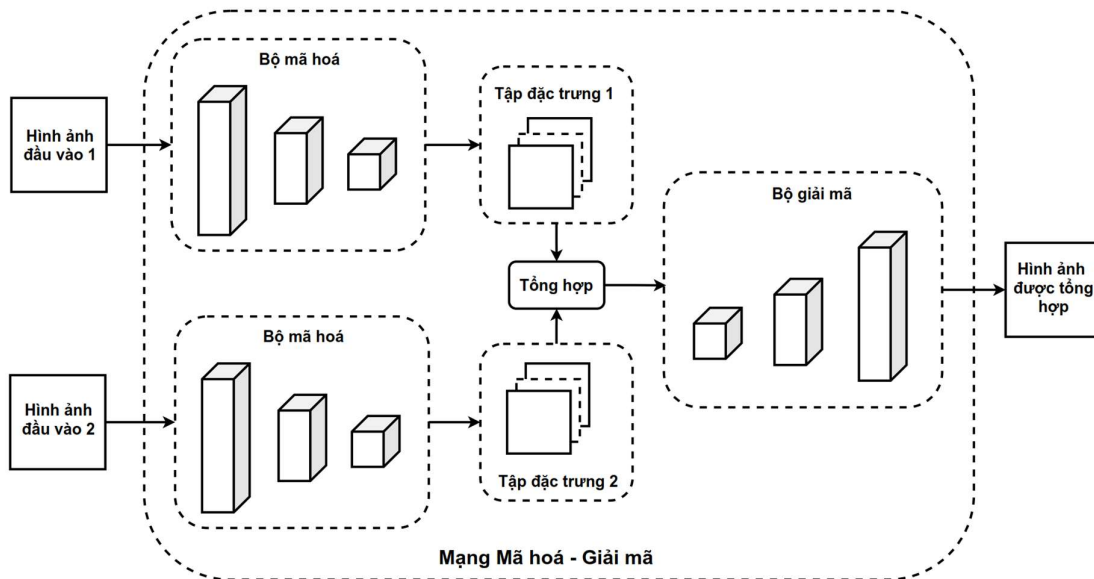
Các cách tiếp cận dựa trên mạng Mã hoá - Giải mã

Mô hình chung của các cách tiếp cận dựa trên mạng Mã hoá - Giải mã đều phải trải qua hai quá trình: Huấn luyện mạng và sử dụng mạng để tổng hợp hình ảnh. Đầu tiên là quá trình huấn luyện, mạng mã hoá và giải mã sẽ học các trọng số cần thiết để có thể tái tạo được ảnh đầu vào. Sau khi hoàn tất quá trình huấn luyện, mạng mã hoá và giải mã được sử dụng cho quá trình tiếp theo là tổng hợp hình ảnh. Hình 1.14 [25] minh hoạ cho mạng Mã hoá-Giải mã dùng trong quá trình huấn luyện.



Hình 1.14. Sơ đồ quá trình huấn luyện mạng Mã hoá - Giải mã

Quá trình tiếp theo là tổng hợp hình ảnh sử dụng mạng Mã hoá – Giải mã đã được huấn luyện trước đó. Hình 1.15 [25] mô tả mô hình tổng hợp hình ảnh sử dụng mạng Mã hoá - Giải mã.



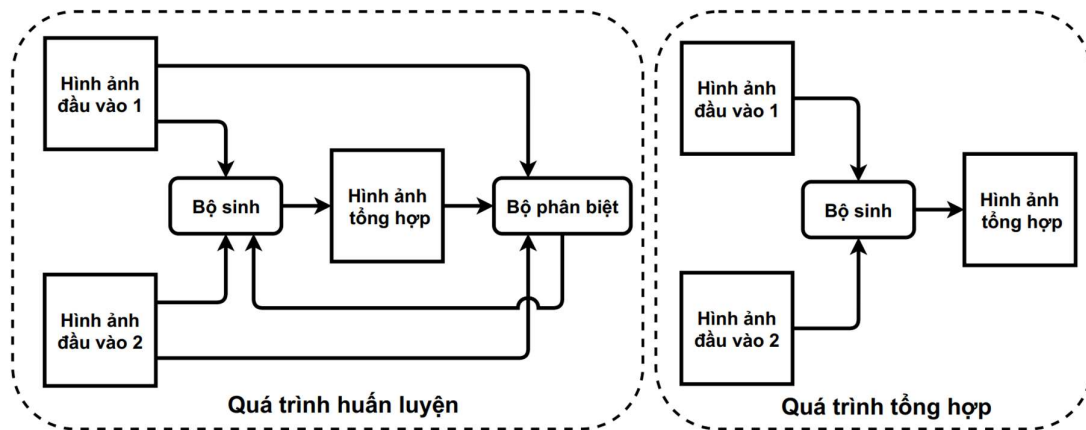
Hình 1.15. Sơ đồ của mô hình tổng hợp hình ảnh sử dụng mạng Mã hoá - Giải mã

Quá trình tổng hợp được tiến hành như sau: Đầu tiên, bộ mã hoá sẽ được sử dụng để trích xuất đặc trưng cho các hình ảnh đầu vào. Sau đó, các đặc trưng sẽ được tổng hợp với nhau theo một thuật toán tổng hợp cụ thể. Những đặc trưng sau khi tổng hợp sẽ được đưa qua bộ giải mã để thu được hình ảnh tổng hợp. Ví dụ, Li và các cộng sự [64] đã đề xuất một mô hình tổng hợp hình ảnh sử dụng một mạng mã hoá và giải mã. Ban đầu, bộ mã hoá và giải mã được huấn luyện để tái tạo cho hình ảnh đầu vào. Sau khi quá trình huấn luyện kết thúc. Các trọng số của mạng mã hoá và giải mã sẽ được sử dụng để mã hoá cho các hình ảnh đầu vào thu được một tập các đặc trưng được mã hoá. Sau đó, tập các đặc trưng của hai hình ảnh này được xử lý bằng cách sử dụng toán tử chuẩn L_1 và toán tử trung bình để thu được đặc trưng cho từng ảnh đầu vào. Sau đó hàm Softmax được sử dụng cho các đặc trưng này để thu được đặc trưng tổng hợp. Mạng giải mã được sử dụng để biến đổi đặc trưng tổng hợp về hình ảnh tổng hợp. Một số cách tiếp cận tổng hợp hình ảnh y học khác dựa trên mạng mã hoá-giải mã như trong các nghiên cứu [20], [65], [66] và [22].

Cách tiếp cận dựa trên mạng đối kháng sinh (GAN)

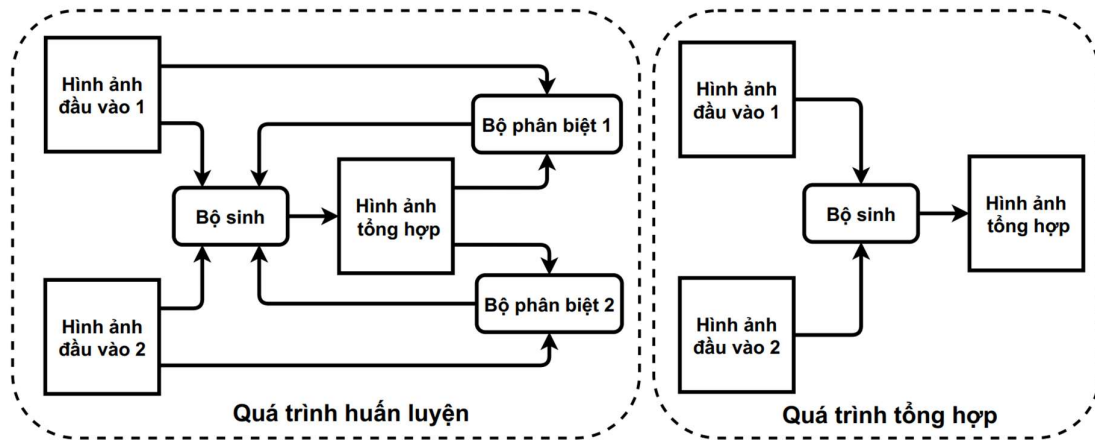
Cho đến nay, có rất nhiều các mô hình tổng hợp hình ảnh y học dựa trên mạng GAN đã được đề xuất. Nhìn chung, có thể nhóm các cách tiếp cận này vào hai nhóm chính như sau:

- **Nhóm thứ nhất:** Có một bộ sinh và một bộ phân biệt. Hình 1.16 [25] minh họa mạng GAN với một bộ sinh và một bộ phân biệt được sử dụng để tổng hợp hình ảnh. Mạng sinh sẽ sử dụng dữ liệu của các hình ảnh đầu vào để học cách tạo ra hình ảnh tổng hợp. Bộ phân biệt sẽ được huấn luyện để có thể phân biệt được 3 hình ảnh (2 hình ảnh đầu vào và 1 hình ảnh tổng hợp được tạo ra). Sau quá trình huấn luyện thì bộ sinh được sử dụng để tổng hợp hình ảnh. Một số nghiên cứu tổng hợp hình ảnh y học thuộc nhóm này như DSAGAN [67], MGM-GAN [24] và TA-cGAN [68].



Hình 1.16. Sơ đồ mạng GAN tổng hợp hình ảnh với 1 bộ sinh và 1 bộ phân biệt

- **Nhóm thứ hai:** Một bộ sinh và hai bộ phân biệt. Hình 1.17 [25] minh họa mạng GAN với một bộ sinh và hai bộ phân biệt được sử dụng để tổng hợp hình ảnh. Mạng sinh sẽ sử dụng dữ liệu của các hình ảnh đầu vào để học cách tạo ra hình ảnh tổng hợp. Bộ phân biệt 1 sẽ được huấn luyện để có thể phân biệt được cặp hình ảnh (hình ảnh đầu vào 1 và hình ảnh tổng hợp được tạo ra). Bộ phân biệt 2 sẽ được huấn luyện để có thể phân biệt được cặp hình ảnh (hình ảnh đầu vào 2 và hình ảnh tổng hợp được tạo ra). Sau quá trình huấn luyện thì bộ sinh được sử dụng để tổng hợp hình ảnh. Một số nghiên cứu tổng hợp hình ảnh y học thuộc nhóm này như DDcGAN [69] và MMFGAN [70].



Hình 1.17. Sơ đồ mạng GAN tổng hợp hình ảnh với 1 bộ sinh và 2 bộ phân biệt

Nhìn chung, việc xây dựng một mạng học sâu để tổng hợp hình ảnh vẫn còn có hạn chế. Hạn chế thứ nhất là việc tổng hợp hình ảnh có thể chưa được hiệu quả khi có ít dữ liệu để huấn luyện mạng. Lý do là bởi vì các tham số học được trong mạng sẽ phụ thuộc nhiều vào số lượng dữ liệu ảnh đầu vào dùng để huấn luyện. Đối với tập dữ liệu không lớn, việc học ra các tham số của mạng có thể không thực sự tối ưu. Hay nói cách khác, mạng với những tham số như vậy chưa thể biểu diễn tốt các đặc trưng của hình ảnh. Đối với bài toán tổng hợp hình ảnh, tập dữ liệu ảnh y học được sử dụng cho việc tổng hợp hình ảnh còn khá hạn chế. Theo Li và các cộng sự [40] đã thu thập các hình ảnh y học từ nguồn “The Whole Brain Atlas” được cung cấp tại trang Web: <http://www.med.harvard.edu/AANLIB/> và lưu trữ toàn bộ những hình ảnh y học thu thập được tại địa chỉ: https://github.com/MorvanLi/image-fusion-zoom/tree/main/Medical_Image_Fusion_Methods/Havard-Medical-Image-Fusion-Datasets. Số lượng hình ảnh y học ở nguồn này hiện có gồm: MRI-PET (269 cặp), MRI-SPECT (357 cặp) và MRI-CT (184 cặp). Số lượng các hình ảnh y học trên không thực sự là nhiều. Do đó, các mạng học sâu chưa thể phát huy sự hiệu quả của nó trong các mô hình tổng hợp hình ảnh [25]. Hạn chế thứ hai là thiếu đi sự linh động so với các phương pháp tổng hợp truyền thống. Ví dụ, trong trường hợp hình ảnh tổng hợp

thu được có chất lượng thấp như cường độ sáng, độ tương phản thấp, mất chi tiết hoặc bị biến dạng về màu sắc thì việc điều chỉnh được mạng để giải quyết các vấn đề đó cũng đặt ra các thách thức. Ngược lại, đối với các phương pháp truyền thống, việc thiết kế các phương pháp tổng hợp để khắc phục các vấn đề của hình ảnh tổng hợp trở lên đơn giản hơn. Cụ thể, nếu hình ảnh có cường độ sáng và độ tương phản thấp, có thể thiết kế các quy tắc tổng hợp cho các thành phần cơ sở để giải quyết vấn đề này. Nếu ảnh bị mất chi tiết, có thể thiết kế các quy tắc tổng hợp cho các thành phần chi tiết để giải quyết hạn chế đó.

1.3. Một số hạn chế của các phương pháp tổng hợp hình ảnh và giải pháp

Thông qua việc giới thiệu các nghiên cứu liên quan đến bài toán tổng hợp hình ảnh trong mục 1.2. Có thể thấy được rằng có nhiều cách tiếp cận khác nhau để giải quyết cho bài toán tổng hợp hình ảnh y học. Tuy nhiên, những phương pháp đó vẫn tồn tại những hạn chế nhất định trong việc tổng hợp hình ảnh. Trong mục này, hai hạn chế chính ảnh hưởng tới hiệu quả của việc tổng hợp hình ảnh sẽ được trình bày chi tiết.

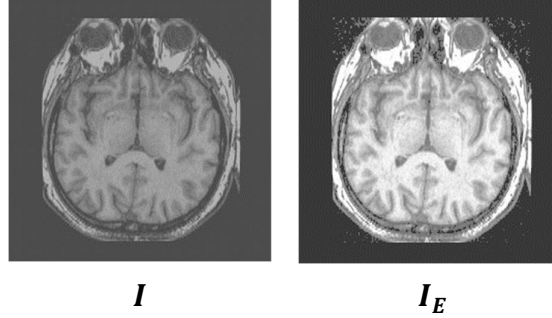
1.3.1. Hạn chế liên quan đến chất lượng ảnh

Hình ảnh y học thường gặp các vấn đề về chất lượng như mờ, cường độ sáng và độ tương phản thấp. Đây là vấn đề khách quan và phổ biến cho tất cả các thuật toán tổng hợp hình ảnh. Chất lượng hình ảnh đầu vào thấp sẽ có khả năng cao làm giảm hiệu quả của quá trình tổng hợp hình ảnh. Nhìn chung, một số phương pháp chính thường được sử dụng để cải thiện chất lượng của ảnh có cường độ sáng và độ tương phản thấp như:

- Phương pháp cân bằng biểu đồ mức xám và các biến thể của nó.
- Phương pháp sử dụng thuật toán tối ưu hoá.

Phương pháp cân bằng biểu đồ mức xám có hiệu quả đối với các ảnh có độ tương phản thấp. Phương pháp này là một trong những phương pháp được sử dụng rộng rãi để tăng cường độ tương phản của hình ảnh. Nó sử dụng một hàm để ánh xạ các mức xám từ hình ảnh đầu vào có độ tương phản thấp sang hình ảnh tăng cường có độ tương phản tốt hơn. Cách tiếp cận này đã trở nên phổ biến do tính đơn giản của nó trong việc tính toán. Tuy nhiên, đối với các hình ảnh cộng hưởng từ có vùng xám chiếm đa số trong ảnh thì phương pháp này tỏ ra không hiệu quả. Hình ảnh thu được sau tăng cường có thể bị tăng cường quá mức. Hạn chế này dẫn đến việc loại bỏ các chi tiết cấu trúc trong hình ảnh cộng hưởng từ não. Mặt khác, cách tiếp cận này cũng làm tăng nhiễu trong hình ảnh, dẫn đến độ sáng không đồng đều xuất hiện trong hình ảnh được tăng cường. Hình 1.18 minh hoạ cho việc cân bằng biểu đồ mức xám đối

với một hình ảnh cộng hưởng từ đầu vào I có độ tương phản và cường độ sáng thấp được lấy từ nguồn “The whole brain atlas”. I_E là hình ảnh thu được sau khi cân bằng biểu đồ mức xám. Dễ dàng nhận thấy rằng, mặc dù hình ảnh sau cân bằng I_E đã được cải thiện về mặt cường độ sáng. Tuy nhiên, nó đã bị cân bằng quá mức và xuất hiện các nhiễu. Do đó, phương pháp cân bằng biểu đồ mức xám là không hiệu quả trong việc cải thiện chất lượng đối với các hình ảnh cộng hưởng từ não.



Hình 1.18. Minh họa ảnh được tăng cường bởi cân bằng biểu đồ mức xám

Vì vậy, các biến thể của phương pháp cân bằng biểu đồ đã được đề xuất. Ví dụ, Subramani và cộng sự [71] đã giới thiệu phương pháp cân bằng biểu đồ chênh lệch mức xám mờ để nâng cao hình ảnh y học. Magudeeswaran và các cộng sự [72] đã đề xuất một phương pháp cân bằng biểu đồ mức xám mờ hai mức bảo toàn cường độ sáng để cải thiện độ tương phản cho các hình ảnh cộng hưởng từ não. Phương pháp này phân vùng biểu đồ mức xám của ảnh đầu vào thành các biểu đồ mức xám con và sau đó cân bằng độc lập từng biểu đồ con. Subramani và các cộng sự [73] đã giới thiệu một phương pháp tăng cường độ tương phản của ảnh cộng hưởng từ não dựa trên cân bằng biểu đồ mức xám thích nghi (AHE - Adaptive histogram equalization) dựa trên logic mờ. Phương pháp này bao gồm hai giai đoạn. Đầu tiên, logic mờ được áp dụng cho một hình ảnh đầu vào và sau đó đầu ra của nó được xử lý bởi phương pháp AHE. Phương pháp này không chỉ duy trì độ sáng trung bình mà còn cải thiện độ tương phản của hình ảnh.

Bên cạnh các phương pháp dựa trên cân bằng biểu đồ mức xám. Trong những năm gần đây, cách tiếp cận dựa trên thuật toán tối ưu hoá cũng đã được áp dụng để nâng cao chất lượng cho hình ảnh nói chung và hình ảnh y học nói riêng đặc biệt về khía cạnh cải thiện cường độ sáng và độ tương phản. Ví dụ, Chen và cộng sự [74] đã sử dụng thuật toán ABC (Artificial bee colony) để nâng cao độ tương phản của hình ảnh. Hình ảnh đầu vào được biến đổi dựa trên hàm Beta không đầy đủ với 2 tham số cần tối ưu là α và β . Một hàm tối ưu được tác giả đề xuất dựa trên các tham số như: độ sắc nét tính theo toán tử phát hiện biên Canny, lượng thông tin trong ảnh tính theo Entropy và cuối cùng là độ tương phản tính theo phương sai của ma trận ảnh. Thuật

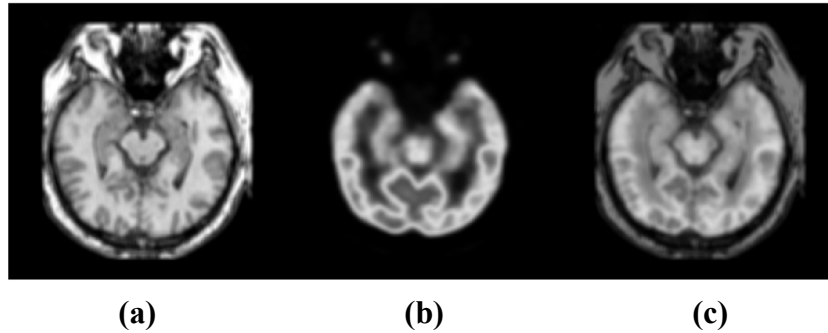
toán tối ưu hoá ABC được sử dụng để tìm hai tham số tối ưu α và β . Kamoona và các cộng sự [75] đã cải tiến thuật toán tối ưu Cuckoo và áp dụng nó vào bài toán nâng cao chất lượng ảnh xám. Hình ảnh đầu vào được nâng cao chất lượng bằng một phương pháp biến đổi cục bộ-toàn cục. Hình ảnh tăng cường sau đó được sử dụng trong một hàm tối ưu được xây dựng dựa trên các thành phần như biên ảnh được tính theo toán tử Sobel, lượng thông tin trong ảnh tính theo Entropy và cường độ sáng. Trong việc cải thiện chất lượng hình ảnh cộng hưởng từ não, các thuật toán tối ưu cũng đã được ứng dụng trong một số nghiên cứu. Ví dụ, Rao [76] đã đề xuất một mô hình nâng cao chất lượng ảnh (gọi là DHE-PSO) bằng cách kết hợp phương pháp cân bằng biểu đồ mức xám động (DHE) với thuật toán tối ưu hóa bầy đàn PSO (Particle swarm optimization) để xác định các tham số tối ưu. Phương pháp này nhằm duy trì độ sáng trong hình ảnh chụp cộng hưởng từ não trong khi tăng cường các vùng mô của nó. Tuy nhiên, các chi tiết cấu trúc trong các vùng mô bị loại bỏ. Acharya và cộng sự [77] đã giới thiệu một phương pháp nâng cao độ tương phản của hình ảnh cộng hưởng từ não dựa trên thuật toán tối ưu bầy đàn PSO kết hợp với cân bằng biểu đồ mức xám. Pashaei và cộng sự [78] đã cải tiến thuật toán tối ưu hoá AOA (Arithmetic optimization algorithm) và kết hợp nó với phương pháp cân bằng biểu đồ mức xám thích nghi có giới hạn độ tương phản để nâng cao chất lượng cho hình ảnh y học.

Để giải quyết vấn đề chất lượng của hình ảnh đầu vào thấp, một số nghiên cứu tổng hợp hình ảnh gần đây đã tiến hành cải thiện chất lượng của hình ảnh đầu vào trước khi tiến hành quá trình tổng hợp. Ví dụ, Ullah và các cộng sự [26] đã sử dụng bộ lọc Laplace cục bộ nhanh (FLLF) để bảo toàn cạnh và tăng cường chi tiết thông tin đường viền. Maqsood và các cộng sự [27] đã sử dụng phương pháp cân bằng biểu đồ mức xám được sửa đổi phi tham số (NMHE) để tăng độ tương phản cho hình ảnh đầu vào. Trên thực tế, các cách tiếp cận này chỉ cải thiện một phần trong chất lượng ảnh đầu vào. Tuy nhiên, nó cũng có những hạn chế nhất định là nhạy cảm với nhiễu và hình ảnh thu được bị tăng cường quá mức. Do đó, những phương pháp trên chưa thực sự cải thiện được nhiều về chất lượng của hình ảnh tổng hợp.

1.3.2. Hạn chế liên quan đến phương pháp tổng hợp ảnh

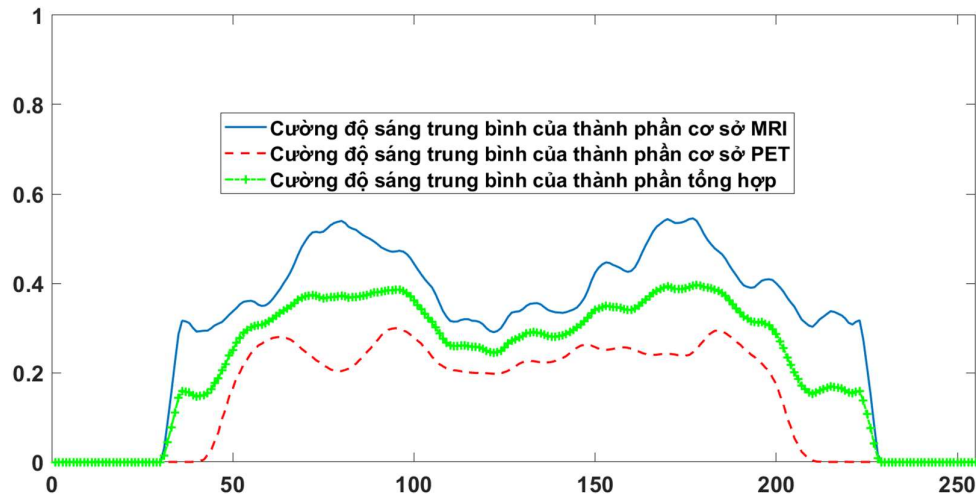
Thông qua quan sát các phương pháp tổng hợp hình ảnh, các phương pháp tổng hợp được thiết kế cho các thành phần cơ sở và chi tiết trong ảnh vẫn còn có những hạn chế nhất định. Cụ thể, đối với việc tổng hợp các thành phần cơ sở, các thuật toán thuật toán trung bình hoặc thuật toán lựa chọn Max thường được sử dụng vì tính đơn giản và độ phức tạp tính toán thấp. Sự lựa chọn này giúp cho việc tổng hợp diễn ra nhanh chóng. Tuy nhiên, việc sử dụng những thuật toán này có thể gây ra một số vấn đề cho hình ảnh tổng hợp, như làm mất thông tin, giảm độ tương phản

và cường độ sáng. Hình 1.19 minh họa cho ba thành phần bao gồm: thành phần cơ sở của hai hình ảnh MRI (Hình 1.19 a), PET (Hình 1.19 b) và thành phần cơ sở tổng hợp theo thuật toán trung bình (Hình 1.19 c). Bằng trực quan, có thể thấy rằng thành phần tổng hợp theo thuật toán trung bình có cường độ sáng và độ tương phản là kém hơn so với thành phần cơ sở của hình ảnh MRI.



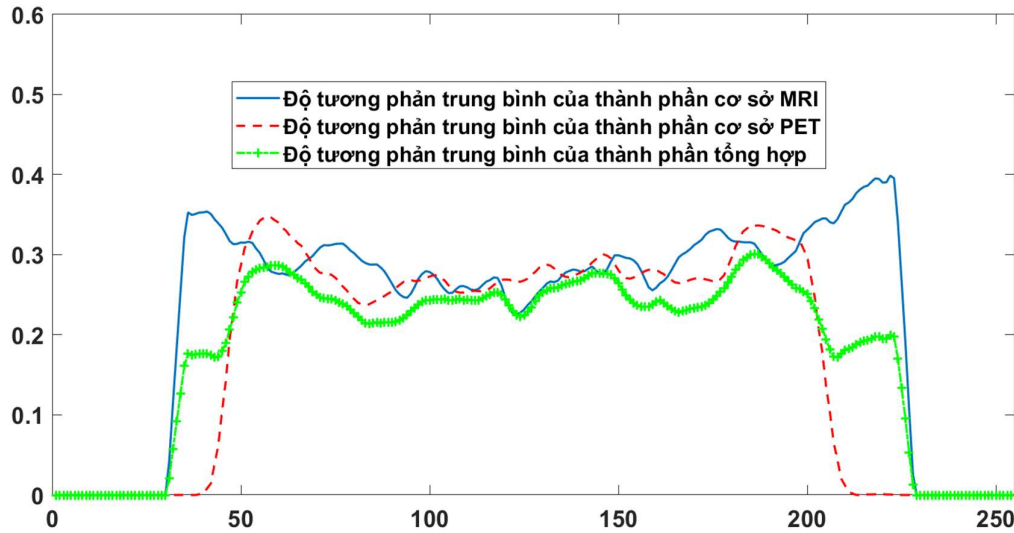
Hình 1.19. Minh họa thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình

Để thấy rõ hơn nữa về sự suy giảm cường độ sáng và độ tương phản trong hình ảnh tổng hợp được thực hiện theo thuật toán trung bình cho các thành phần cơ sở, Hình 1.20 và Hình 1.21 biểu diễn cường độ sáng và độ tương phản trung bình theo cột của ma trận của các thành phần cơ sở của hai hình ảnh MRI, PET và thành phần cơ sở tổng hợp theo thuật toán trung bình. Từ Hình 1.20, nhận thấy rằng đường màu xanh lá cây biểu diễn cho cường độ sáng của hình ảnh tổng hợp theo thuật toán trung bình là thấp hơn so với đường màu xanh nước biển biểu diễn cho cường độ sáng của hình ảnh MRI.



Hình 1.20. Minh họa cường độ sáng trung bình của các thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình

Từ Hình 1.21, nhận thấy rằng đường màu xanh nước biển biểu diễn cho độ tương phản trung bình theo cột của thành phần tổng hợp là thấp hơn so với độ tương phản trung bình theo cột của hai thành phần cơ sở của MRI và PET.



Hình 1.21. Minh họa độ tương phản trung bình của các thành phần cơ sở của ảnh MRI, PET và thành phần tổng hợp theo thuật toán trung bình

Do đó, thuật toán trung bình áp dụng cho các thành phần cơ sở sẽ làm cho thành phần tổng hợp bị suy giảm cường độ sáng và độ tương phản. Hạn chế trình bày ở trên có thể tìm thấy trong một số nghiên cứu về tổng hợp hình đã được công bố gần đây. Ví dụ, Maqsood và các cộng sự [27] đã đề xuất một thuật toán tổng hợp hình ảnh y học dựa trên phân rã hình ảnh hai tỉ lệ và biểu diễn thưa. Thuật toán phân rã hai tỉ lệ được áp dụng để phân rã các hình ảnh đầu vào thành các thành phần cơ sở và các thành phần chi tiết. Thuật toán tổng hợp trung bình đã được sử dụng để tổng hợp cho các thành phần cơ sở. Liu và các cộng sự [79] đã giới thiệu một phương pháp tổng hợp hình ảnh dựa trên biểu diễn thưa tích chập. Trong nghiên cứu này một bộ lọc thông thấp đã được sử dụng để phân rã hình ảnh đầu vào thành các thành phần cơ sở và chi tiết. Các thành phần cơ sở đã được tổng hợp lại với nhau theo thuật toán trung bình. Thuật toán tổng hợp trung bình có thể tìm thấy trong một số nghiên cứu tổng hợp hình ảnh y học khác như Zhang và các cộng sự [28]. Kết quả hình ảnh tổng hợp đầu ra thu được từ các thuật toán đề cập ở trên có cường độ sáng và độ tương phản thấp.

Một số nghiên cứu khác đã tìm cách khắc phục hạn chế của thuật toán trung bình trong việc tổng hợp cho các thành phần cơ sở bằng cách sử dụng thuật toán lựa chọn Max. Mặc dù thuật toán lựa chọn Max đã khắc phục được vấn đề suy giảm về cường độ sáng của hình ảnh tổng hợp. Tuy nhiên, nó lại có hạn chế là làm mất thông tin của hình ảnh tổng hợp. Điều này có thể được giải thích là vì cường độ sáng của

thành phần cơ sở của hình ảnh MRI thường lớn hơn cường độ sáng của thành phần cơ sở của hình ảnh PET. Khi thuật toán lựa chọn Max được sử dụng, hình ảnh tổng hợp sẽ lấy toàn bộ thông tin từ hình ảnh MRI mà bỏ qua thông tin của hình ảnh PET. Thuật toán lựa chọn Max cũng có thể tìm thấy trong nghiên cứu của Liu và các cộng sự [29]. Trong nghiên cứu này, các hình ảnh đầu vào đã được phân rã thành các thành phần cơ sở và thành phần chi tiết bởi sử dụng bộ lọc hai chiều đồng thời (JBF). Sau đó, thuật toán lựa chọn Max đã được áp dụng để tổng hợp cho các thành phần cơ sở. Babu và các cộng sự [80], đã áp dụng đồng thời cả thuật toán lựa chọn Max và thuật toán trung bình để tổng hợp cho các thành phần tần số thấp trong ảnh. Đầu tiên, hình các ảnh đầu vào đã được biến đổi DWT để thu được các thành phần tần số thấp và thành phần tần số cao. Khi đó, thuật toán lựa chọn Max được sử dụng để tổng hợp các thành phần tần số thấp. Tiếp theo thành phần tổng hợp cùng với các thành phần tần số cao của nó được biến đổi ngược DWT để thu được hình ảnh tổng hợp. Hình ảnh này lại tiếp tục được biến đổi NSST để thu được các thành phần tần số thấp và tần số cao. Sau đó thuật toán trung bình được áp dụng để tổng hợp các thành phần tần số thấp.

Đối với việc tổng hợp cho thành phần chi tiết, một số thuật toán tổng hợp thường được sử dụng như thuật toán lựa chọn Max, thuật toán tổng Laplace được sửa đổi (SML). Những thuật toán tổng hợp cho các thành phần chi tiết vẫn có hạn chế trong việc bảo tồn các thông tin chi tiết từ hình ảnh đầu vào. Điều này có thể được giải thích như sau: Độ sắc nét của hình ảnh cộng hưởng từ (MRI) thường tốt hơn so với hình ảnh chụp cắt lớp phát xạ (PET). Trong trường hợp áp dụng thuật toán lựa chọn Max để tổng hợp, các thông tin của hình ảnh tổng hợp thường chính là thông tin của hình ảnh MRI bởi vì cường độ sáng của hình ảnh PET là nhỏ hơn so với MRI. Ví dụ, Barba-J và các cộng sự [18] đã áp dụng thuật toán lựa chọn Max cho các thành phần chi tiết. Thuật toán tổng hợp thành phần chi tiết SML cũng chưa thực sự hiệu quả. Việc chọn các điểm ảnh có độ lớn Laplace cao có thể gây mất thông tin chi tiết vì các chi tiết nhỏ hoặc mịn trong hình ảnh có thể có giá trị Laplace thấp. Các vùng hình ảnh có chi tiết như các đường viền, các cấu trúc nhỏ, hay các khu vực trơn, thường có độ lớn Laplace thấp. Những điểm ảnh này có thể bị bỏ qua trong quá trình tổng hợp, dẫn đến mất thông tin chi tiết trong hình ảnh tổng hợp. Ví dụ, Li và các cộng sự [3] đã áp dụng thuật toán SML cho các thành phần chi tiết. Kết quả cho thấy hình ảnh tổng hợp bị mất các thông tin chi tiết.

Vì vậy, nội dung của luận án này sẽ tập trung giải quyết hai hạn chế đã nêu ở trên bằng cách đề xuất hai nhóm phương pháp như sau:

Nhóm 1:

- Đề xuất phương pháp tăng cường chất lượng ảnh cộng hưởng từ não để giải quyết vấn đề ảnh có độ tương phản và cường độ sáng thấp.

Nhóm 2:

- Đề xuất một thuật toán tổng hợp thích nghi cho các thành phần cơ sở để giải quyết vấn đề về sự suy giảm về mặt chất lượng ảnh trong quá trình tổng hợp.
- Đề xuất một thuật toán tổng hợp hiệu quả cho thành phần chi tiết để giải quyết vấn đề hình ảnh tổng hợp bị mất các thông tin đặc trưng từ hình ảnh đầu vào.

1.4. Kiến thức nền tảng

Một số kiến thức nền tảng cần thiết cho việc hiểu các thuật toán đề xuất sẽ được trình bày một cách tóm tắt trong mục này. Những kiến thức này bao gồm: Giới thiệu về thuật toán Meta-heuristic và thuật toán tối ưu hoá MPA, hàm năng lượng cục bộ (LEF), thuật toán tổng hợp dựa trên hàm năng lượng cục bộ kết hợp với toán tử la bàn Prewitt, cấu trúc Tensor và đặc trưng của nó, bộ lọc hướng dẫn quay vòng (RGF), bộ lọc đường cong trung bình có trọng số (WMCF), cân bằng biểu đồ mức xám thích nghi giới hạn độ tương phản (CLAHE), toán tử phát hiện biên Laplace (LED), bộ lọc khuếch tán dị hướng (Anisotropic diffusion filter - ADF) và phục hồi ảnh dựa trên tổng biến thiên (Total variation image reconstruction - TVIR).

1.4.1. Giới thiệu về thuật toán Meta-heuristic

Thuật toán meta-heuristic là thuật toán tối ưu hóa được sử dụng để tìm ra giải pháp tối ưu cho các vấn đề phức tạp không thể giải được bằng các phương pháp truyền thống [81]. Ngoài ra, các thuật toán meta-heuristic có thể xử lý các không gian tìm kiếm lớn và phức tạp, điều mà các phương pháp tối ưu hóa truyền thống không thể làm được. Thuật toán meta-heuristic hoạt động bằng cách cải tiến lặp đi lặp lại một giải pháp cho đến khi nó đủ tốt để được coi là giải pháp cuối cùng. Chúng bắt đầu với giải pháp ban đầu, sau đó sử dụng một bộ quy tắc hoặc phương pháp phỏng đoán để sửa đổi giải pháp. Mục tiêu là tìm ra giải pháp tốt hơn giải pháp hiện tại. Quá trình được lặp lại cho đến khi tìm được giải pháp đủ tốt. Thuật toán meta-heuristic thường được sử dụng cho các vấn đề NP-khó. Tuy nhiên, các thuật toán meta-heuristic thường có thể tìm ra giải pháp tốt cho những vấn đề này trong một khoảng thời gian hợp lý.

Một trong những hạn chế chính của thuật toán meta-heuristic là không phải lúc nào chúng cũng có thể tìm ra giải pháp tối ưu toàn cục. Do tính chất ngẫu nhiên của thuật toán nên không có gì đảm bảo rằng thuật toán sẽ tìm ra giải pháp tốt nhất. Ngoài ra, các thuật toán có thể tốn kém về mặt tính toán, đặc biệt khi xử lý các không

gian tìm kiếm lớn. Nhìn chung, thuật toán meta-heuristic là công cụ tối ưu hóa mạnh mẽ có thể được sử dụng để giải quyết các vấn đề phức tạp.

1.4.1.1. Phân loại thuật toán Meta-heuristic

Các thuật toán meta-heuristic có thể được phân loại theo nhiều cách dựa trên các tiêu chí khác nhau. Theo Rajwar và cộng sự [81] có thể phân loại thuật toán meta-heuristic như sau:

- **Lấy cảm hứng từ tự nhiên và không lấy cảm hứng từ tự nhiên:** Thuật toán meta-heuristic có thể được phân loại dựa trên việc chúng có được lấy cảm hứng từ quá trình tự nhiên hay không. Ví dụ, thuật toán meta-heuristic lấy cảm hứng từ tự nhiên bao gồm các thuật toán như thuật toán di truyền (Genetic algorithm - GA) và thuật toán tối ưu hóa bầy đàn (Particle swarm optimization - PSO). Thuật toán meta-heuristic không lấy cảm hứng từ tự nhiên bao gồm các thuật toán Simulated annealing (SA) và thuật toán Tabu search (TS).
- **Dựa trên quy trình xác định và ngẫu nhiên:** Thuật toán meta-heuristic có thể được phân loại dựa trên việc chúng sử dụng các quy trình xác định hay ngẫu nhiên. Thuật toán meta-heuristic xác định sử dụng các quy trình xác định để tạo ra các giải pháp mới. Ví dụ, thuật toán Hill climbing (HC). Thuật toán meta-heuristic ngẫu nhiên sử dụng các quy trình ngẫu nhiên để tạo ra các giải pháp mới. Ví dụ, thuật toán di truyền (GA) và thuật toán SA.
- **Dựa trên quỹ đạo và dựa trên quần thể:** Thuật toán meta-heuristic cũng có thể được phân loại dựa trên việc chúng tập trung vào việc tìm kiếm một giải pháp duy nhất hay nhiều giải pháp. Thuật toán meta-heuristic dựa trên quỹ đạo tập trung vào việc tìm kiếm một giải pháp duy nhất và sử dụng cải tiến lặp lại để tinh chỉnh giải pháp đó. Ví dụ, thuật toán Hill climbing (HC), thuật toán Simulated annealing (SA) và thuật toán Tabu search (TS). Thuật toán meta-heuristic dựa trên quần thể tập trung vào việc tìm kiếm nhiều giải pháp và sử dụng tập hợp các giải pháp để khám phá không gian tìm kiếm. Ví dụ, các thuật toán di truyền (GA) và tối ưu hóa bầy đàn (PSO).
- **Dựa trên tìm kiếm cục bộ và dựa trên tìm kiếm toàn cục:** Thuật toán meta-heuristic có thể được phân loại dựa trên việc chúng tập trung vào việc khám phá không gian tìm kiếm cục bộ hay không gian tìm kiếm toàn cục. Thuật toán meta-heuristic dựa trên tìm kiếm cục bộ tập trung vào việc tìm kiếm các giải pháp gần với giải pháp hiện tại. Ví dụ, thuật toán Hill climbing (HC) và thuật toán Tabu search (TS). Thuật toán meta-heuristic dựa trên tìm kiếm toàn cục tập trung vào việc khám phá toàn bộ không gian tìm kiếm. Ví dụ, thuật toán di truyền (GA).

1.4.1.2. Một số thuật toán tối ưu hoá dùng trong luận án

Trong mục này, một số thuật toán được sử dụng trong các thực nghiệm sẽ được giới thiệu tóm tắt. Thêm vào đó, thuật toán tối ưu hoá MPA được sử dụng trong các phương pháp đề xuất sẽ được trình bày chi tiết.

Một số thuật toán tối ưu hoá sử dụng trong các thực nghiệm so sánh bao gồm:

- Thuật toán MVO (Multi-Verse Optimizer) được giới thiệu bởi Mirjalili và các cộng sự [82] vào năm 2016. Thuật toán này được lấy cảm hứng từ lý thuyết đa vũ trụ trong vật lý. Ba khái niệm chính của lý thuyết đa vũ trụ (hố trắng, hố đen và hố sâu) được mô hình hóa toán học để xây dựng MVO.
- Thuật toán WOA (Whale Optimization Algorithm) được giới thiệu bởi Mirjalili và cộng sự [83] vào năm 2016. Thuật toán này được lấy cảm hứng từ chiến lược kiếm ăn bằng lưới bong bóng của cá voi lưng gù.
- Thuật toán SSA (Salp Swarm Algorithm) được giới thiệu bởi Mirjalili và các cộng sự [84] vào năm 2017. Thuật toán này lấy cảm hứng chính từ hành vi bầy đàn của Salps khi di chuyển và tìm kiếm thức ăn trong đại dương.
- Thuật toán SCA (Sine cosine algorithm) được giới thiệu bởi Mirjalili [85] vào năm 2016. Thuật toán này tạo ra nhiều giải pháp ứng cử viên ngẫu nhiên ban đầu và yêu cầu chúng dao động ra ngoài hoặc hướng tới giải pháp tốt nhất bằng cách sử dụng mô hình toán học dựa trên hàm Sin và Cosin.
- Thuật toán GWO (Grey wolf optimizer) được giới thiệu bởi Mirjalili và các cộng sự [86] vào năm 2014. Thuật toán này lấy cảm hứng từ loài sói xám (Canis lupus). Nó mô phỏng hệ thống phân cấp lãnh đạo và cơ chế săn mồi của loài sói xám trong tự nhiên.
- Thuật toán DA (Dragonfly algorithm) được giới thiệu bởi Mirjalili [87] vào năm 2016. Thuật toán này được lấy cảm hứng từ các hành vi bầy đàn động và tĩnh của chuồn chuồn nhân tạo trong tự nhiên.
- Thuật toán ALO (Ant Lion Optimizer) được giới thiệu bởi Mirjalili [88] vào năm 2016. Thuật toán này lấy cảm hứng từ thiên nhiên và nó mô phỏng cơ chế săn mồi của kiến sư tử trong tự nhiên.

Thuật toán tối ưu hoá MPA

Thuật toán MPA là một thuật toán meta-heuristic được truyền cảm hứng từ tự nhiên. Thuật toán này được đề xuất bởi Faramarzi và các cộng sự [31] vào năm 2020 và nó đã được sử dụng rộng rãi trong rất nhiều các bài toán của xử lý ảnh như: tổng hợp ảnh y học [89] và tăng cường chất lượng ảnh [90]. Ý tưởng của thuật toán MPA là bắt chước hành vi săn mồi của các loài ở biển lớn như: cá mập, cá voi... Thuật toán MPA tuân theo các quy tắc chi phối tự nhiên trong chiến lược tìm kiếm thức ăn tối

ưu. Hai loại chuyển động chính mà con mồi và kẻ săn mồi thường tuân theo là Brownian và Lévy. Thuật toán 1.1 trình bày chi tiết thuật toán MPA và Hình 1.22 minh họa cho sơ đồ của thuật toán MPA.

Chuyển động Brownian:

Chuyển động Brown là một quá trình ngẫu nhiên trong đó độ dài bước của chúng được xác định bởi Phân phối chuẩn (Gauss) với giá trị trung bình bằng 0 ($\mu = 0$) và đơn vị phương sai ($\sigma^2 = 1$). Hàm mật độ xác suất tại điểm x cho chuyển động này được tính theo công thức (1.1).

$$f_B(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \quad (1.1)$$

Chuyển động Lévy:

Chuyển động Lévy là một kiểu đi ngẫu nhiên có kích thước bước được xác định bởi phân phối Lévy theo công thức (1.2).

$$f_L(x; \alpha, \gamma) = \frac{1}{\pi} \int_0^\infty \exp(-\gamma q^\alpha) \cos(qx) dx \quad (1.2)$$

Trong đó, tham số α xác định chỉ số phân phối và kiểm soát thang đo thuộc tính của quá trình trong khi tham số γ chọn đơn vị tỷ lệ. Tích phân trong công thức (1.2) có nghiệm chỉ trong một số trường hợp. Khi α bằng 2, nó đại diện cho phân bố Gauss và khi α bằng 1, nó thể hiện phân phối Cauchy. Giải pháp cho tích phân trong biểu thức (1.2) có thể tính xấp xỉ theo công thức (1.3) khi x có giá trị rất lớn.

$$f_L(x; \alpha, \gamma) \approx \frac{\gamma \Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right)}{\pi x^{(1+\alpha)}} , x \rightarrow \infty \quad (1.3)$$

Trong đó, Γ là kí hiệu cho hàm Gamma. Với α là một số nguyên dương thì hàm Γ được tính theo công thức (1.4).

$$\Gamma(\alpha) = (\alpha - 1)! \quad (1.4)$$

Một thuật toán chính xác và nhanh cho chuyển động Lévy với giá trị tham số α tùy ý nằm trong khoảng 0.3 và 1.99 được tính theo công thức (1.5).

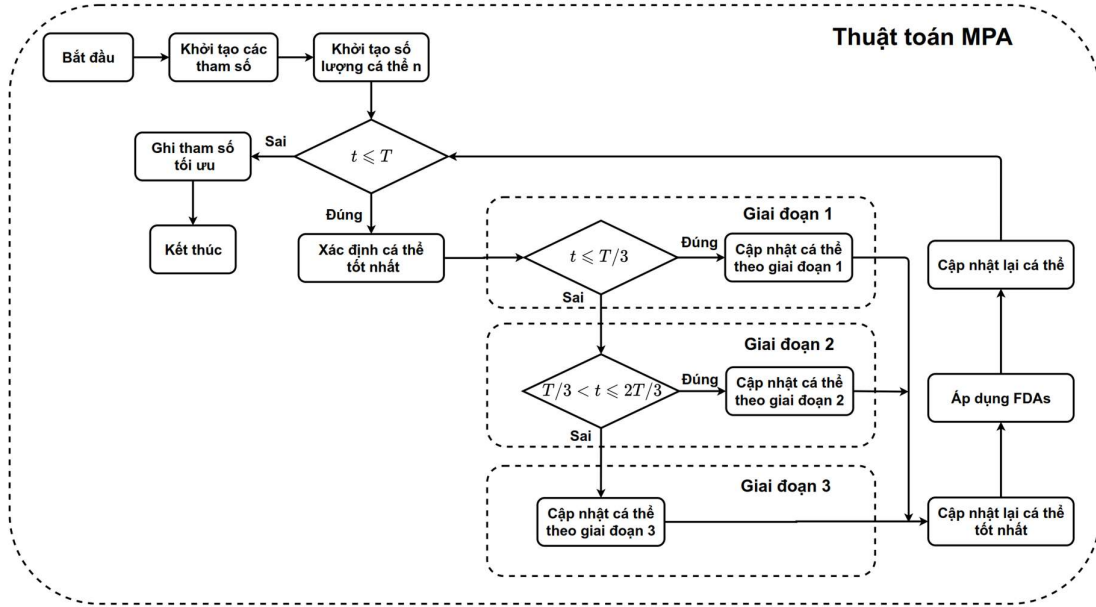
$$\text{Levy}(\alpha) = 0.05 \times \frac{x}{|y|^{1/\alpha}} \quad (1.5)$$

Trong đó, x và y là hai biến phân phối chuẩn được tính theo các công thức (1.6) và (1.7) với độ lệch chuẩn σ_x tính theo công thức (1.8) và $\sigma_y = 1$ và $\alpha = 1.5$.

$$x = \text{Normal}(0, \sigma_x^2) \quad (1.6)$$

$$y = \text{Normal}(0, y_x^2) \quad (1.7)$$

$$\sigma_x = \left[\frac{\Gamma(1 + \alpha) \sin\left(\frac{\pi\alpha}{2}\right)}{\Gamma\left(\frac{1 + \alpha}{2}\right) \alpha 2^{\frac{(\alpha-1)}{2}}} \right] \quad (1.8)$$



Hình 1.22. Sơ đồ của thuật toán MPA

Những điểm nổi bật trong chính sách tìm kiếm thức ăn có thể được tóm tắt như sau:

- Những kẻ săn mồi ở biển sử dụng chiến lược di chuyển Lévy cho môi trường có mật độ con mồi thấp trong khi sử dụng chuyển động Brownian cho những khu vực có nhiều con mồi.
- Do các tác động của môi trường như: sự hình thành dòng xoáy trong tự nhiên hoặc do các thiết bị thu hút cá của con người gây ra, chúng thay đổi hành vi của mình để hy vọng tìm thấy các khu vực có sự phân bố con mồi khác nhau.
- Ở tỷ lệ vận tốc giữa kẻ săn mồi và con mồi là thấp ($v = 0.1$), chiến lược tốt nhất cho kẻ săn mồi là chuyển động Lévy. Con mồi di chuyển theo Brownian hoặc Lévy.
- Ở tỷ lệ vận tốc giữa kẻ săn mồi và con mồi là như nhau ($v = 1$), nếu con mồi di chuyển theo chuyển động Lévy, chiến lược tốt nhất cho kẻ săn mồi là chuyển động Brownian.
- Ở tỷ lệ vận tốc giữa kẻ săn mồi và con mồi là cao ($v \geq 10$), chiến lược tốt nhất cho kẻ săn mồi là không di chuyển chút nào. Trong trường hợp này, con mồi di chuyển theo Brownian hoặc Lévy.

Giai đoạn 1: Khi động vật ăn thịt di chuyển nhanh hơn con mồi. Tình huống này xảy ra trong các lần lặp lại tối ưu hóa ban đầu. Ở tỷ lệ vận tốc cao ($v \geq 10$), chiến lược tốt nhất cho động vật ăn thịt là không di chuyển. Mô hình toán học của quy tắc này được minh họa như các công thức (1.9) và (1.10) tương ứng.

$$\overrightarrow{SS}_i(x, y) = \overrightarrow{R}_B(x, y) \otimes (\overrightarrow{E}_i(x, y) - \overrightarrow{R}_B(x, y) \otimes \overrightarrow{Pr}_i(x, y)) \quad (1.9)$$

$$\overrightarrow{Pr_i}(x, y) = \overrightarrow{Pr_i}(x, y) + K. \vec{R} \otimes \overrightarrow{SS_i}(x, y) \quad (1.10)$$

Trong đó:

- $\overrightarrow{SS_i}(x, y)$ là một vector chứa kích thước bước di chuyển của con mồi.
- $E(x, y)$ là một ma trận được xây dựng trên các giải pháp phù hợp nhất.
- $\vec{R} \in [0,1]$ tuân theo một phân phối đồng nhất.
- $K = 0.5$.
- $\otimes$ là toán tử nhân entry-wise.
- $\vec{R_B}$ là một vector các số ngẫu nhiên được sinh ra theo phân phối chuẩn.

Giai đoạn 2: Khi cả kẻ săn mồi và con mồi đều di chuyển với cùng một tốc độ. Nó bắt chước cả hai đang tìm kiếm con mồi của mình. Do đó, một nửa số cá thể trong bầy được chỉ định để thăm dò và nửa còn lại để khai thác. Trong giai đoạn này, con mồi chịu trách nhiệm khai thác và động vật ăn thịt chịu trách nhiệm thăm dò. Dựa trên quy tắc tỷ lệ vận tốc đơn vị ($v \approx 1$), nếu con mồi di chuyển theo chuyển động Lévy, chiến lược tốt nhất cho kẻ săn mồi là chuyển động Brownian. Do đó, con mồi di chuyển theo chuyển động Lévy trong khi kẻ săn mồi di chuyển theo chuyển động Brownian và được mô tả như các công thức (1.11), (1.12), (1.13) và (1.14).

Một nửa đầu số cá thể trong bầy được tính theo các công thức (1.11) và (1.12):

$$\overrightarrow{SS_i}(x, y) = \overrightarrow{R_L}(x, y) \otimes (\overrightarrow{E_i}(x, y) - \overrightarrow{R_L}(x, y) \otimes \overrightarrow{Pr_i}(x, y)) \quad (1.11)$$

$$\overrightarrow{Pr_i}(x, y) = \overrightarrow{Pr_i}(x, y) + K. \vec{R} \otimes \overrightarrow{SS_i}(x, y) \quad (1.12)$$

Một nửa còn lại số cá thể trong bầy được tính theo các công thức (1.13) và (1.14):

$$\overrightarrow{SS_i}(x, y) = \overrightarrow{R_B}(x, y) \otimes (\overrightarrow{R_B}(x, y) \otimes \overrightarrow{E_i}(x, y) - \overrightarrow{Pr_i}(x, y)) \quad (1.13)$$

$$\overrightarrow{Pr_i}(x, y) = \overrightarrow{E_i}(x, y) + K. CF \otimes \overrightarrow{SS_i}(x, y) \quad (1.14)$$

Trong đó: CF là một tham số thích nghi để kiểm soát kích thước bước cho chuyển động của động vật ăn thịt. Tham số t là vòng lặp hiện tại và T là số vòng lặp tối đa.

$$- CF = \left(1 - \frac{t}{T}\right)^{\frac{2 \cdot t}{T}}$$

Giai đoạn 3: Khi động vật ăn thịt di chuyển nhanh hơn con mồi. Kịch bản này xảy ra trong giai đoạn cuối của quá trình tối ưu hóa. Ở tỷ lệ vận tốc thấp ($v = 0.1$), chiến lược tốt nhất cho động vật ăn thịt là di chuyển theo chuyển động Lévy. Giai đoạn này được trình bày dưới dạng các công thức (1.15) và (1.16).

$$\overrightarrow{SS_i}(x, y) = \overrightarrow{R_L}(x, y) \otimes (\overrightarrow{R_L}(x, y) \otimes \overrightarrow{E_i}(x, y) - \overrightarrow{Pr_i}(x, y)) \quad (1.15)$$

$$\overrightarrow{Pr_i}(x, y) = \overrightarrow{E_i}(x, y) + K. CF \otimes \overrightarrow{SS_i}(x, y) \quad (1.16)$$

Trong đó:

- $\vec{R_L}$ là một vector được sinh ra bởi phân phối Lévy.

Một điểm khác gây ra sự thay đổi hành vi ở các loài săn mồi biển là các vấn đề môi trường như sự hình thành dòng xoáy hoặc các hiệu ứng của thiết bị thu hút cá (FADs - Fish Aggregating Devices). Thông thường cá mập dành hơn 80% thời gian của chúng ở vùng lân cận của các FADs và 20% còn lại chúng sẽ nhảy xa hơn ở các không gian khác nhau để tìm môi trường có sự phân bố con mồi khác [31]. Do đó, hiệu ứng FADs được trình bày dưới dạng toán học như công thức (1.17) và (1.18) sẽ giúp cập nhật lại vị trí của giải pháp. Tránh việc bị kẹt lại các cực trị địa phương.

Nếu $r \leq FADs$

$$\overrightarrow{Pr}_i(x, y) = \overrightarrow{Pr}_i(x, y) + CF[\vec{X}_{min} + R \otimes (\vec{X}_{max} - \vec{X}_{min})] \otimes \vec{U} \quad (1.17)$$

Nếu $r > FADs$

$$\overrightarrow{Pr}_i(x, y) = \overrightarrow{Pr}_i(x, y) + [FADs(1 - r) + r](\overrightarrow{Pr}_{r_1} - \overrightarrow{Pr}_{r_2}) \quad (1.18)$$

Trong đó:

- $FADs = 0.2$.
- $\vec{U}$ là một vector nhị phân, tham số $r \in (0, 1)$.
- $\vec{X}_{max}$ và $\vec{X}_{min}$ là các vector chứa cận trên và cận dưới.
- r_1 và r_2 biểu thị các chỉ số ngẫu nhiên của ma trận con mồi.

Thuật toán 1.1: MPA

Đầu vào: Số vòng lặp, số cá thể, cận dưới, cận trên, số chiều, hàm mục tiêu

Đầu ra: Các tham số tối ưu cần tìm của hàm mục tiêu

Bước 1: Khởi tạo ngẫu nhiên các tham số và vị trí của các cá thể trong bầy

Bước 2: Thực hiện vòng lặp của thuật toán

Bước 2.1: Tính giá trị hàm mục tiêu để xác định cá thể tốt nhất.

Bước 2.2: Trong 1/3 vòng lặp đầu tiên.

Thực hiện cập nhật lại vị trí cá thể theo các công thức (1.9) và (1.10).

Bước 2.3: Trong 1/3 vòng lặp tiếp theo.

Đối với 1 nửa số cá thể đầu tiên:

Thực hiện cập nhật lại vị trí cá thể theo các công thức (1.11) và (1.12).

Đối với 1 nửa số cá thể cuối:

Thực hiện cập nhật lại vị trí cá thể theo các công thức (1.13) và (1.14).

Bước 2.4: Trong 1/3 vòng lặp cuối cùng.

Thực hiện cập nhật lại vị trí cá thể theo các công thức (1.15) và (1.16).

Bước 2.5: Cập nhật lại vị trí cá thể tốt nhất

Bước 2.6: Áp dụng ảnh hưởng của các thiết bị thu hút cá (FADs)

Thực hiện cập nhật lại vị trí cá thể theo các công thức (1.17) và (1.18).

Bước 3: Sau khi kết thúc vòng lặp của MPA, lấy ra các tham số tối ưu tìm được.

1.4.2. Hàm năng lượng cục bộ (LEF)

Hàm năng lượng cục bộ (LEF) của một hình ảnh đề cập đến năng lượng của một vùng nhỏ hoặc cửa sổ của hình ảnh. Nó được tính bằng tổng bình phương của các giá trị pixel trong vùng đó. Năng lượng cục bộ của ảnh có thể được sử dụng để mô tả kết cấu, độ sắc nét hoặc các thuộc tính khác của ảnh trong vùng đó. Ứng dụng của hàm năng lượng cục bộ đã xuất hiện trong nhiều nghiên cứu khác nhau như tổng hợp hình ảnh [91], [92].

Cho một hình ảnh đầu vào I có kích thước $M \times N$. Gọi W là một cửa sổ cục bộ có kích thước $k \times k$. Hàm năng lượng cục bộ tại điểm (i, j) kí hiệu là $LEF(i, j)$ được tính toán theo công thức (1.19).

$$LEF(i, j) = \sum_{k_1=0}^{k-1} \sum_{k_2=0}^{k-1} W(k_1, k_2) I^2(i + k_1, j + k_2) \quad (1.19)$$

1.4.3. Thuật toán FR_PCO

Thuật toán tổng hợp FR_PCO được đề xuất bởi Dinh [93] vào năm 2021. Thuật toán này được xây dựng dựa trên hàm năng lượng cục bộ kết hợp với toán tử la bàn Prewitt và dùng để tổng hợp cho các thành phần chi tiết trong ảnh. Thuật toán 1.2 mô tả chi tiết các bước của thuật toán FR_PCO.

Thuật toán 1.2: FR_PCO

Đầu vào: H_1 và H_2 là hai thành phần chi tiết

Đầu ra: Thành phần H được tổng hợp theo thuật toán FR_PCO

Bước 1: Tính hàm năng lượng cục bộ của các thành phần H_p ($p = \overline{1,2}$) theo các mặt nạ của toán tử la bàn Prewitt W_{PCO}^k ($k = \overline{1,8}$) và theo ma trận toàn giá trị 1 (W), kích thước $m \times m$ như công thức (1.20) và (1.21).

$$LEF_PCO_p^k(i, j) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} W_{PCO}^k(u, v) H_p^2(i + u, j + v) \quad (1.20)$$

$$LEF_p(i, j) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} W(u, v) H_p^2(i + u, j + v) \quad (1.21)$$

Bước 2: Tính tổng giá trị của trị tuyệt đối của các hàm năng lượng $LEF_PCO_p^k$ và LEF_p ($k = \overline{1,8}, p = \overline{1,2}$) theo các công thức (1.22) và (1.23).

$$M_1 = \sum_{k=1}^8 |LEF_PCO_1^k| + |LEF_1| \quad (1.22)$$

$$M_2 = \sum_{k=1}^8 |LEF_PCO_2^k| + |LEF_2| \quad (1.23)$$

Bước 3: Tính thành phần tổng hợp H theo công thức (1.24).

$$H(i, j) = \begin{cases} H_1(i, j) & \text{nếu } M_1(i, j) \geq M_2(i, j) \\ H_2(i, j) & \text{nếu } M_1(i, j) < M_2(i, j) \end{cases} \quad (1.24)$$

1.4.4. Cấu trúc Tensor và toán tử phát hiện đặc trưng

Cấu trúc Tensor là một công cụ hiệu quả để thu được các đặc trưng của hình ảnh. Các ứng dụng của nó trong xử lý hình ảnh có thể kể đến như: tổng hợp hình ảnh Y học, lấy nét mở rộng trong kính hiển vi [94] và phát hiện góc [95].

Cho ma trận ảnh I , cấu trúc Tensor được định nghĩa theo công thức (1.25).

$$ST = \begin{bmatrix} \sum_w E_i^2 & \sum_w E_i E_j \\ \sum_w E_i E_j & \sum_w E_j^2 \end{bmatrix} \quad (1.25)$$

Trong đó:

- w là một cửa sổ cục bộ.
- E_i và E_j là các đạo hàm riêng theo hướng i và j tương ứng.

Hơn nữa, các giá trị riêng của cấu trúc Tensor cung cấp thêm thông tin về hình học cục bộ của hình ảnh đầu vào. Các giá trị riêng (v_1 và v_2) có thể được tính toán theo các công thức (1.26) và (1.27) tương ứng.

$$v_1 = \frac{1}{2} \left(\sum_w E_i^2 + \sum_w E_j^2 + \sqrt{\left(\sum_w E_i^2 - \sum_w E_j^2 \right)^2 + 4 \left(\sum_w E_i E_j \right)^2} \right) \quad (1.26)$$

$$v_2 = \frac{1}{2} \left(\sum_w E_i^2 + \sum_w E_j^2 - \sqrt{\left(\sum_w E_i^2 - \sum_w E_j^2 \right)^2 + 4 \left(\sum_w E_i E_j \right)^2} \right) \quad (1.27)$$

Toán tử phát hiện đặc trưng nổi bật của cấu trúc Tensor (STSDO - Structure tensor salient detection operator) [96] có thể được tính toán theo công thức (1.28).

$$STS = \sqrt{(v_1 + v_2)^2 + 0.5(v_1 - v_2)^2} \quad (1.28)$$

1.4.5. Bộ lọc hướng dẫn quay vòng (RGF)

Bộ lọc RGF được Zhang và các cộng sự [97] giới thiệu vào năm 2010. Đây là một trong những bộ lọc khử nhiễu tiên tiến nhất vẫn bảo toàn được đường biên. Bộ lọc này được ứng dụng rộng rãi trong rất nhiều các bài toán của xử lý ảnh. Bộ lọc RGF được thực hiện theo hai bước chính:

- Xoá bỏ các cấu trúc nhỏ bởi sử dụng bộ lọc Gauss.
- Phục hồi đường biên.

Chi tiết về bộ lọc RGF được mô tả như trong Thuật toán 1.3.

Thuật toán 1.3: RGF

Đầu vào:

- Ảnh I , σ_s , σ_r .
- Số lượng vòng lặp: t

Đầu ra: Ảnh sau khi lọc I_{RGF}

Bước 1: Thực hiện lọc Gauss đối với hình ảnh đầu vào (I), thu được hình ảnh I_G theo công thức (1.29).

$$I_G(p) = \frac{1}{K_p} \sum_{q \in N(p)} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) I(q) \quad (1.29)$$

Trong đó:

- p và q là tọa độ trong ảnh.
- σ_s là độ lệch chuẩn của bộ lọc Gauss.
- $N(p)$ là tập các điểm ảnh lân cận có tâm là p .
- K_p dùng để chuẩn hoá và nó tính theo công thức (1.30).

$$K_p = \sum_{q \in N(p)} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2}\right) \quad (1.30)$$

Bước 2: Áp dụng bộ lọc hướng dẫn để phục hồi đường biên

Gọi J^1 là kết quả thu được từ Bước 1 bởi bộ lọc Gauss. Kí hiệu J^{t+1} là đầu ra thu được từ vòng lặp thứ t . J^{t+1} được tính theo công thức (1.31).

$$J^{t+1}(p) = \frac{1}{K_p} \sum_{q \in N(p)} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2} - \frac{\|J^t(p) - J^t(q)\|^2}{2\sigma_r^2}\right) I(q) \quad (1.31)$$

Trong đó:

- K_p dùng để chuẩn hoá và nó được tính theo công thức (1.32).

$$K_p = \sum_{q \in N(p)} \exp\left(-\frac{\|p - q\|^2}{2\sigma_s^2} - \frac{\|J^t(p) - J^t(q)\|^2}{2\sigma_r^2}\right) I(q) \quad (1.32)$$

- σ_r là độ lệch chuẩn của bộ lọc hướng dẫn.
-

1.4.6. Bộ lọc đường cong trung bình có trọng số (WMCF)

Gong và Goksel [98] đã đề xuất bộ lọc đường cong trung bình có trọng số (WMCF) vào năm 2019 và nó đã được sử dụng hiệu quả trong việc tổng hợp hình

ảnh [99]. Bộ lọc WMCF được trình bày chi tiết trong Thuật toán 1.4 và 8 ma trận mặt nạ của nó được minh họa trong Hình 1.23.

$$\begin{array}{cccc}
 \begin{bmatrix} 1/6 & 1/6 & 0 \\ 1/3 & -1 & 0 \\ 1/6 & 1/6 & 0 \end{bmatrix} & \begin{bmatrix} 1/6 & 1/3 & 1/6 \\ 1/6 & -1 & 1/6 \\ 0 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1/6 & 1/6 \\ 0 & -1 & 1/3 \\ 0 & 1/6 & 1/6 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 0 \\ 1/6 & -1 & 1/6 \\ 1/6 & 1/3 & 1/6 \end{bmatrix} \\
 m_1 & m_2 & m_3 & m_4 \\
 \begin{bmatrix} 1/6 & 1/3 & 1/12 \\ 1/3 & -1 & 0 \\ 1/12 & 0 & 0 \end{bmatrix} & \begin{bmatrix} 1/12 & 1/3 & 1/6 \\ 0 & -1 & 1/3 \\ 0 & 0 & 1/12 \end{bmatrix} & \begin{bmatrix} 0 & 0 & 1/12 \\ 0 & -1 & 1/3 \\ 1/12 & 1/3 & 1/6 \end{bmatrix} & \begin{bmatrix} 1/12 & 0 & 0 \\ 1/3 & -1 & 0 \\ 1/6 & 1/3 & 1/12 \end{bmatrix} \\
 m_5 & m_6 & m_7 & m_8
 \end{array}$$

Hình 1.23. Các mặt nạ của bộ lọc WMCF

Thuật toán 1.4: WMCF

Đầu vào: Ảnh I

Đầu ra: Ảnh sau khi được lọc I_{WMCF}

Bước 1: Tính ma trận d_i bằng cách nhân chập ảnh I với 8 mặt nạ m_i ($i = \overline{1,8}$) theo công thức (1.33).

$$d_i = m_i \otimes I, i = \overline{1,8} \quad (1.33)$$

Trong đó: $\otimes$ là phép toán nhân chập.

Bước 2: Hình ảnh sau khi lọc I_{WMCF} thu được bằng cách lựa chọn phần tử nhỏ nhất trong 8 ma trận d_i và được tính theo công thức (1.34).

$$I_{WMCF} = d_m \quad (1.34)$$

Trong đó: $m = \arg \min(|d_i|, i = \overline{1,8})$

1.4.7. Cân bằng biểu đồ xám thích nghi giới hạn độ tương phản (CLAHE)

Thuật toán cân bằng CLAHE được Zuiderveld [100] đề xuất vào năm 1994. Thuật toán này được mô tả chi tiết trong Thuật toán 1.5. Thuật toán này được sử dụng phổ biến để cải thiện độ tương phản của hình ảnh.

Thuật toán 1.5: CLAHE

Đầu vào: Ảnh I , vùng nhỏ có kích thước $w \times w$, ngưỡng giới hạn tương phản L

Đầu ra: Ảnh được tăng cường I_{CLAHE}

Các bước của thuật toán CLAHE như sau:

Bước 1: Chia hình ảnh đầu vào I thành các vùng có kích thước $w \times w$.

Bước 2: Thực hiện cân bằng biểu đồ mức xám đối với từng vùng $w \times w$ đã chia ở bước 1.

Bước 3: Áp dụng ngưỡng giới hạn tương phản (L) để kiểm soát độ tương phản đối với từng vùng. Những giá trị mức xám quá lớn sẽ được cắt bỏ hoặc giảm bớt.

Bước 4: Kết hợp các vùng nhỏ lại để tạo ra ảnh kết quả I_{CLAHE} .

Thuật toán này cho phép tăng cường độ tương phản trên từng vùng nhỏ của ảnh. Điều này giúp làm nổi bật các đặc trưng quan trọng và tạo ra hình ảnh rõ ràng hơn. Hơn nữa, giới hạn tương phản trong thuật toán CLAHE còn giúp tránh việc cường nhiễu bị tăng cường trong ảnh.

1.4.8. Toán tử phát hiện biên Laplace (LED)

Toán tử phát hiện biên LED (Laplace Edge Detection) là một phương pháp có khả năng phát hiện các biên mạnh và các chi tiết nhỏ. Điều này giúp nó rất hữu ích trong việc tìm ra các vùng biên có độ lớn thay đổi nhanh trong mức xám. Việc phát hiện biên theo phương pháp Laplace có thể được thực hiện một cách đơn giản là nhân chập hình ảnh đầu vào với mặt nạ Laplace như công thức (1.35).

$$K = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (1.35)$$

1.4.9. Bộ lọc khuếch tán dị hướng (ADF)

Bộ lọc ADF được đề xuất bởi Perona và Malik [101]. Ý tưởng cơ bản của bộ lọc ADF là dựa vào khái niệm về tốc độ khuếch tán tại mỗi điểm ảnh dựa trên Gradient (đạo hàm) của hình ảnh. Bộ lọc ADF được mô tả chi tiết trong thuật Toán 1.6.

Thuật toán 1.6: ADF

Đầu vào: Ảnh I , số lần lặp L , khoảng thời gian Δt , hệ số K

Đầu ra: Ảnh được lọc I_{ADF}

Thực hiện vòng lặp L

Bước 1: Tính đạo hàm tại mỗi điểm ảnh theo công thức (1.36).

$$G_x = \frac{\partial I}{\partial x} \quad \text{và} \quad G_y = \frac{\partial I}{\partial y} \quad (1.36)$$

Bước 2: Tính cường độ biến đổi đạo hàm tại mỗi điểm theo công thức (1.37).

$$G = \sqrt{G_x^2 + G_y^2} \quad (1.37)$$

Bước 3: Tính hàm khuếch tán C dựa trên cường độ biến đổi đạo hàm và hệ số K theo công thức (1.38).

$$C = \exp\left(-\frac{G^2}{K^2}\right) \quad (1.38)$$

Bước 4: Tính tốc độ khuếch tán tại mỗi điểm theo công thức (1.39)

$$S_D = C \cdot G \quad (1.39)$$

Bước 5: Cập nhật cường độ sáng tại mỗi điểm theo công thức (1.40)

$$I = I + \Delta t \cdot S_D \quad (1.40)$$

Bước 6: Thu được hình ảnh sau khi lọc I_{ADF} khi kết thúc vòng lặp.

Cường độ biến đổi tại mỗi điểm ảnh xác định tốc độ khuếch tán tại điểm đó. Các vùng có cường độ biến đổi cao hơn sẽ có tốc độ khuếch tán thấp hơn, và ngược lại. Bộ lọc ADF thường được sử dụng trong xử lý ảnh y học do ít làm hỏng các thông tin chi tiết. Thuật toán này có hiệu suất làm mịn tốt, đảm bảo hiệu quả giảm nhiễu, bảo toàn thông tin đường biên.

1.4.10. Phục hồi hình ảnh dựa trên tổng biến thiên (TVIR)

Phục hồi hình ảnh là một vấn đề nghịch đảo trong đó mục tiêu là khôi phục hình ảnh rõ nét từ hình ảnh mờ hoặc nhiễu. Về mặt toán học, có thể mô tả việc phục hồi hình ảnh như công thức (1.41).

$$g = Hf + n \quad (1.41)$$

Trong đó:

- $f \in R^{MN \times 1}$ là một vector biểu thị hình ảnh chưa xác định (có tiềm năng rõ nét) với kích thước $M \times N$.
- $g \in R^{MN \times 1}$ là một vector biểu thị hình ảnh quan sát được.
- $n \in R^{MN \times 1}$ là một vector biểu thị nhiễu.
- $H \in R^{MN \times MN}$ là một biến đổi tuyến tính biểu diễn toán tử nhân chập.

Mục tiêu của việc phục hồi ảnh là khôi phục được f từ g . Hiện nay, có một số các nghiên cứu đã được đề xuất để giải quyết bài toán trên. Một số nghiên cứu có thể đề cập như: khử tích chập Richardson [102], cực tiểu hoá bình phương nhỏ nhất chính quy [103]. Trong phần này giới thiệu một phương pháp phục hồi ảnh do Chan và các cộng sự [104] đề xuất.

Bài toán khôi phục ảnh có thể chuyển về bài toán tìm f để cực tiểu hoá hàm cho trong công thức (1.42).

$$\min_f \frac{\mu}{2} \|Hf - g\|^2 + \|f\|_{TV} \quad (1.42)$$

Công thức (1.42) còn được biết đến là cực tiểu hoá TV/L2.

Trong đó:

- $\|\cdot\|^2$ là chuẩn bình phương.
- $\|f\|_{TV}$ là chuẩn TV đẳng hướng được tính theo công thức (1.43) .

$$\|f\|_{TV} = \sum_i \sqrt{\beta_x^2 [D_x f]_i^2 + \beta_y^2 [D_y f]_i^2 + \beta_t^2 [D_t f]_i^2} \quad (1.43)$$

Trong đó:

- D_x, D_y và D_t là các toán tử sai phân hữu hạn dọc theo các hướng ngang, dọc và thời gian.
- β_x, β_y và β_t là các hằng số.
- $[f]_i$ là thành phần thứ i của vector f .

Vấn đề cốt lõi là giải quyết vấn đề tối ưu hoá bằng cách đi tìm cực tiểu của biểu thức (1.44).

$$\min_f \frac{\mu}{2} \|Hf - g\|^2 + \|u\|_{TV} \text{ thỏa mãn } u = Df \quad (1.44)$$

Trong đó:

- μ là tham số chuẩn hoá.

Để giải quyết vấn đề tối ưu (1.44), phương pháp Augmented Lagrangian được sử dụng để chuyển tối ưu có ràng buộc thành tối ưu không ràng buộc như công thức (1.45).

$$L(f, u, y) = \frac{\mu}{2} \|Hf - g\|^2 + \|u\|_1 - y^T(u - Df) + \frac{p_r}{2} \|u - Df\|^2 \quad (1.45)$$

Trong đó:

- p_r là tham số chuẩn hoá của thành phần $\|u - Df\|^2$.
- y là hệ số nhân Lagrange liên quan đến ràng buộc $u = Df$.

Sử dụng phương pháp lặp để giải quyết các vấn đề con như mô tả trong các biểu thức (1.46), (1.47) và (1.48).

$$f_{k+1} = \arg \min_f \frac{\mu}{2} \|Hf - g\|^2 - y_k^T(u_k - Df) + \frac{p_r}{2} \|u_k - Df\|^2 \quad (1.46)$$

$$u_{k+1} = \arg \min_u \|u\|_1 - y_k^T(u - Df_{k+1}) + \frac{p_r}{2} \|u - Df_{k+1}\|^2 \quad (1.47)$$

$$y_{k+1} = y_k - p_r(u_{k+1} - Df_{k+1}) \quad (1.48)$$

Theo tác giả Chan và các cộng sự, f tối ưu được tìm ra từ biểu thức (1.46) được trình bày trong biểu thức (1.49).

$$f = F^{-1} \left[\frac{F[\mu H^T g + p_r D^T u - D^T y]}{\mu |F[H]|^2 + p_r (|F[D_x]|^2 + |F[D_y]|^2 + |F[D_t]|^2)} \right] \quad (1.49)$$

Giá trị u tối ưu được tìm ra từ biểu thức (1.47) được trình bày trong biểu thức (1.50).

$$u = \max \left\{ v - \frac{1}{p_r}, 0 \right\} \cdot \frac{v_x}{v} \quad (1.50)$$

Trong đó:

- F là toán tử biến đổi Fourier.
- $v = \max \left\{ \sqrt{|v_x|^2 + |v_y|^2 + |v_t|^2}, \epsilon \right\}$, với ϵ là một hằng số nhỏ ($\epsilon = 10^{-6}$).

Thuật toán 1.7 bình bày chi tiết các bước của thuật toán TVIR.

Thuật toán 1.7: TVIR

Đầu vào:

- g và H
- Các tham số μ , β_x , β_y và β_t .
- Các tham số mặc định: $p_r = 2$ và $\alpha_0 = 0.7$

Đầu ra:

- Hình ảnh khôi phục f
-

Bước 1: Khởi tạo các tham số: $f_0 = g$, $u_0 = Df_0$, $y = 0$, $k = 0$

Bước 2: Tính các ma trận: $F[D_x]$, $F[D_y]$, $F[D_t]$ và $F[H]$

Bước 3: Thực hiện vòng lặp:

Bước 3.1: Tính f theo biểu thức (1.49).

Bước 3.2: Tính u theo biểu thức (1.50).

Bước 3.3: Cập nhật lại y theo biểu thức (1.48).

Bước 3.4: Cập nhật lại tham số p_r theo công thức (1.51).

$$p_r = \begin{cases} \gamma p_r & \text{nếu } \|u_{k+1} - Df_{k+1}\|_2 \geq \alpha \|u_k - Df_k\|_2 \\ p_r & \text{ngược lại} \end{cases} \quad (1.51)$$

Bước 3.5: Kiểm tra điều kiện hội tụ.

Nếu $\|f_{k+1} - f_k\|_2 / \|f_k\|_2 \leq tol$ thì kết thúc.

1.5. Chỉ số đánh giá

Một chỉ số đánh giá riêng lẻ có thể không phản ánh đầy đủ tất cả các khía cạnh của hình ảnh tổng hợp, chẳng hạn như độ sắc nét, độ tương phản hay sự bảo toàn thông tin. Vì vậy, việc sử dụng nhiều chỉ số đánh giá sẽ cung cấp một cái nhìn toàn diện hơn về chất lượng hình ảnh tổng hợp. Hiện nay, có nhiều các chỉ số đánh giá khác nhau thường được dùng để đánh giá chất lượng của hình ảnh tổng hợp. Trong phạm vi luận án, bảy chỉ số phổ biến được lựa chọn ra để đánh giá chất lượng cho hình ảnh tổng hợp. Những chỉ số đánh giá này được phân vào ba nhóm:

Nhóm chỉ số đánh giá chất lượng:

- Chỉ số cường độ sáng (MLI - Mean of light intensity).
- Chỉ số độ tương phản (CI - Contrast index).
- Chỉ số lượng thông tin (Entropy).
- Chỉ số độ sắc nét (AG - Average gradient).

Nhóm chỉ số đánh giá bảo toàn thông tin:

- Chỉ số $Q^{AB/F}$ [105]: tính toán lượng thông tin đường biên được chuyển từ hình ảnh gốc sang hình ảnh tổng hợp.

Nhóm chỉ số đánh giá dựa trên thông tin tương hỗ:

- Chỉ số lượng thông tin tương hỗ [106] (MI - Mutual information)

- Chỉ số lượng thông tin tương hỗ dựa trên đặc trưng [107] (FMI - Feature mutual information).

Cho A và B là hai hình ảnh đầu vào kích thước $M \times N$. Gọi F là hình ảnh tổng hợp từ hai hình ảnh A và B . Các chỉ số đánh giá cho hình ảnh tổng hợp được mô tả chi tiết trong các mục sau.

1.5.1. Chỉ số cường độ sáng

Chỉ số cường độ sáng cho biết mức độ sáng tối của một hình ảnh. Chỉ số cường độ sáng được tính theo công thức (1.52).

$$MLI = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N F(i, j) \quad (1.52)$$

1.5.2. Chỉ số tương phản

Chỉ số tương phản của ảnh được tính thông qua phương sai của ma trận ảnh theo công thức (1.53).

$$CI = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N F^2(i, j) - \left(\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N F(i, j) \right)^2 \quad (1.53)$$

1.5.3. Chỉ số lượng thông tin

Lượng thông tin trong ảnh được tính thông qua Entropy. Chỉ số này được tính theo công thức (1.54).

$$E = - \sum_{i=0}^{255} p_i * \log_2(p_i) \quad (1.54)$$

Trong đó, p_i là xác suất xuất hiện mức xám thứ i ở trong ảnh F .

1.5.4. Chỉ số độ sắc nét

Chỉ số độ sắc nét của ảnh được tính bằng trung bình cộng của độ lớn biên của tất cả các điểm ảnh. Chỉ số này được tính theo công thức (1.55).

$$AG = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (1.55)$$

Trong đó, G_x và G_y được tính theo các công thức (1.56) và (1.57)

$$G_x(x, y) = F(x + 1, y) - F(x, y) \quad (1.56)$$

$$G_y(x, y) = F(x, y + 1) - F(x, y) \quad (1.57)$$

1.5.5. Chỉ số bảo toàn thông tin

Xydeas và Petrovic [105] đã đề xuất một phép đo để đo lường lượng thông tin cạnh (đường biên) được truyền từ các hình ảnh đầu vào A và B sang hình ảnh tổng

hợp F . Phương pháp này sử dụng toán tử phát hiện biên Sobel để tính độ lớn của biên $g(m, n)$ và thông tin hướng $\alpha(m, n)$ cùng nhau tại mỗi điểm ảnh trong cả hai hình ảnh đầu vào (A và B) và hình ảnh tổng hợp (F). Thuật toán 1.8 trình bày chi tiết các bước để tính chỉ số $Q^{AB/F}$.

Thuật toán 1.8: $Q^{AB/F}$

Đầu vào:

- Hai hình ảnh đầu vào A và B .
- Hình ảnh tổng hợp F .
- Các tham số cần thiết: $\Gamma_g = 0.9994$, $K_g = -15$, $\sigma_g = 0.5$ và $\Gamma_\alpha = 0.9879$, $K_\alpha = -22$, $\sigma_\alpha = 0.8$.

Đầu ra: Chỉ số $Q^{AB/F}$

Bước 1: Tính độ lớn các đường biên ($g_A(m, n)$, $g_B(m, n)$, $g_F(m, n)$) và các giá trị góc ($\alpha_A(m, n)$, $\alpha_B(m, n)$, $\alpha_F(m, n)$) tại mỗi điểm ảnh (m, n) cho các hình ảnh A , B , F theo phương pháp phát hiện biên Sobel.

Bước 2: Tính giá trị tương đối của độ lớn đường biên tại mỗi điểm ($G_{(m,n)}^{AF}$) và giá trị góc ($A_{(m,n)}^{AF}$) của hình ảnh A so với hình ảnh F theo các công thức (1.58) và (1.59).

$$G_{(m,n)}^{AF} = \begin{cases} \frac{g_F(m, n)}{g_A(m, n)} & \text{nếu } g_A(m, n) > g_F(m, n) \\ \frac{g_A(m, n)}{g_F(m, n)} & \text{ngược lại} \end{cases} \quad (1.58)$$

$$A_{(m,n)}^{AF} = 1 - \frac{|\alpha_A(m, n) - \alpha_F(m, n)|}{\frac{\pi}{2}} \quad (1.59)$$

Bước 3: Tính giá trị tương đối của độ lớn đường biên tại mỗi điểm ($G_{(m,n)}^{BF}$) và giá trị góc ($A_{(m,n)}^{BF}$) của hình ảnh B so với hình ảnh F theo các công thức (1.60) và (1.61).

$$G_{(m,n)}^{BF} = \begin{cases} \frac{g_F(m, n)}{g_B(m, n)} & \text{nếu } g_B(m, n) > g_F(m, n) \\ \frac{g_B(m, n)}{g_F(m, n)} & \text{ngược lại} \end{cases} \quad (1.60)$$

$$A_{(m,n)}^{BF} = 1 - \frac{|\alpha_B(m, n) - \alpha_F(m, n)|}{\frac{\pi}{2}} \quad (1.61)$$

Bước 4: Tính giá trị bảo toàn thông tin cạnh tại mỗi điểm giữa hình ảnh A và hình ảnh F theo công thức (1.62).

$$Q_{(m,n)}^{AF} = \Gamma_g \Gamma_\alpha \left(1 + e^{K_g(G_{(m,n)}^{AF} - \sigma_g)}\right)^{-1} \left(1 + e^{K_\alpha(A_{(m,n)}^{AF} - \sigma_\alpha)}\right)^{-1} \quad (1.62)$$

Bước 5: Tính giá trị bảo toàn thông tin cạnh tại mỗi điểm giữa hình ảnh B và hình ảnh F theo công thức (1.63).

$$Q_{(m,n)}^{BF} = \Gamma_g \Gamma_\alpha \left(1 + e^{K_g(G_{(m,n)}^{BF} - \sigma_g)}\right)^{-1} \left(1 + e^{K_\alpha(A_{(m,n)}^{BF} - \sigma_\alpha)}\right)^{-1} \quad (1.63)$$

Bước 6: Tính chỉ số bảo toàn đường biên $Q_{(m,n)}^{AB/F}$ theo công thức (1.64).

$$Q_{(m,n)}^{AB/F} = \frac{\sum_{m=1}^M \sum_{n=1}^N Q_{(m,n)}^{AF} w_{(m,n)}^{AF} + Q_{(m,n)}^{BF} w_{(m,n)}^{BF}}{\sum_{m=1}^M \sum_{n=1}^N (w_{(m,n)}^{AF} + w_{(m,n)}^{BF})} \quad (1.64)$$

Trong đó, các trọng số $w_{(m,n)}^{AF}$ và $w_{(m,n)}^{BF}$ được định nghĩa như sau:

- $w_{(m,n)}^{AF} = [g_A(m, n)]^L$
 - $w_{(m,n)}^{BF} = [g_B(m, n)]^L$
 - L là một hằng số được chọn bằng 1.
-

1.5.6. Chỉ số thông tin tương hỗ

MI (Mutual Information) là một khái niệm cơ bản của lý thuyết thông tin. Chỉ số MI đo lường sự phụ thuộc thống kê giữa hai biến ngẫu nhiên và lượng thông tin mà một biến chứa về biến kia. Trong việc đánh giá hiệu quả của việc tổng hợp hình ảnh, chỉ số thông tin tương hỗ MI [106] được sử dụng để đo lường độ tương đồng giữa hình ảnh tổng hợp và các hình ảnh gốc. Điều này giúp xác định xem liệu quá trình tổng hợp đã làm giảm hoặc tăng cường thông tin của các hình ảnh gốc hay không. Cụ thể, khi các hình ảnh gốc có nhiều thông tin chồng lấn nhau, việc sử dụng chỉ số MI trong đánh giá hình ảnh tổng hợp cho phép xác định độ tương đồng giữa các thông tin này trước và sau khi tổng hợp. Nếu chỉ số MI giữ nguyên hoặc tăng lên sau khi tổng hợp, điều đó cho thấy quá trình tổng hợp đã giữ được hoặc tăng cường thông tin. Ngược lại, nếu chỉ số MI giảm đi, điều đó cho thấy quá trình tổng hợp đã gây mất mát thông tin. Chỉ số MI được trình bày chi tiết trong Thuật toán 1.9.

Thuật toán 1.9: MI

Đầu vào:

- Hai hình ảnh đầu vào A và B .
- Hình ảnh tổng hợp F .

Đầu ra:

- Chỉ số MI.
-

Bước 1: Tính biểu đồ mức xám của các hình ảnh đầu vào A , B và hình ảnh tổng hợp F , thu được $P_A(a)$, $P_B(b)$ và $P_F(f)$.

Bước 2: Tính biểu đồ mức xám kết hợp $P_{A,F}(a, f)$ và $P_{B,F}(b, f)$ của các hình ảnh đầu vào A, B và hình ảnh tổng hợp F theo các công thức (1.65) và (1.66).

$$P_{A,F}(a, f) = \sum_{m=1}^M \sum_{n=1}^N P_{A,F}(A(m, n), F(m, n)) \quad (1.65)$$

$$P_{B,F}(b, f) = \sum_{m=1}^M \sum_{n=1}^N P_{B,F}(B(m, n), F(m, n)) \quad (1.66)$$

Bước 3: Tính thông tin $MI_{A,F}$ và $MI_{B,F}$ mà nó đại diện cho thông tin được truyền từ hình ảnh A và hình ảnh B sang hình ảnh tổng hợp F theo các công thức (1.67) và (1.68).

$$MI_{A,F} = \sum_{i=1}^{256} \sum_{j=1}^{256} P_{A,F}(i, j) \log_2 \frac{P_{A,F}(i, j)}{P_A(i)P_F(j)} \quad (1.67)$$

$$MI_{B,F} = \sum_{i=1}^{256} \sum_{j=1}^{256} P_{B,F}(i, j) \log_2 \frac{P_{B,F}(i, j)}{P_B(i)P_F(j)} \quad (1.68)$$

Bước 4: Tính thông tin tương hỗ giữa hai hình ảnh đầu vào A và B với hình ảnh tổng hợp F theo công thức (1.69).

$$MI = MI_{A,F} + MI_{B,F} \quad (1.69)$$

1.5.7. Chỉ số độ tương đồng dựa trên đặc trưng

Chỉ số FMI (Feature mutual information) [108] là một trong những chỉ số thường được sử dụng để đánh giá hiệu quả của hình ảnh tổng hợp. Chỉ số này đo lường độ tương đồng giữa các đặc trưng của các hình ảnh gốc với hình ảnh tổng hợp. Công thức tính FMI tương tự như công thức tính MI. Tuy nhiên, nó được tính bằng cách sử dụng các đặc trưng của hai ảnh đầu vào thay vì các giá trị cường độ sáng của chúng. Nói chung, giá trị FMI càng cao thì độ tương đồng giữa hai ảnh đầu vào càng cao và phương pháp tổng hợp ảnh càng tốt. Chỉ số FMI được trình bày chi tiết trong Thuật toán 1.10.

Thuật toán 1.10: FMI

Đầu vào:

- Hai hình ảnh đầu vào A và B .
- Hình ảnh tổng hợp F .

Đầu ra:

- Chỉ số FMI
-

Bước 1: Tính lượng thông tin đặc trưng mà hình ảnh tổng hợp F chứa thông tin của hai hình ảnh đầu vào A và B (I_{FA}, I_{FB}) theo các công thức (1.70) và (1.71).

$$I_{FA} = \sum_{f,a} P_{FA}(x, y, z, w) \log_2 \frac{P_{FA}(x, y, z, w)}{P_F(x, y)P_A(z, w)} \quad (1.70)$$

$$I_{FB} = \sum_{f,b} P_{FB}(x, y, z, w) \log_2 \frac{P_{FB}(x, y, z, w)}{P_F(x, y)P_B(z, w)} \quad (1.71)$$

Bước 2: Tính chỉ số FMI theo công thức (1.72)

$$FMI = I_{FA} + I_{FB} \quad (1.72)$$

Bước 3: Tính chỉ số FMI được chuẩn hoá theo công thức (1.73)

$$FMI = \frac{I_{FA}}{H_F + H_A} + \frac{I_{FB}}{H_F + H_B} \quad (1.73)$$

Trong đó, H_A , H_B và H_F là các Entropy dựa trên biểu đồ mức xám của hình ảnh A, B và F .

1.6. Kết luận chương 1

Trong Chương 1, luận án đã giới thiệu bài toán tổng hợp hình ảnh và các nghiên cứu liên quan để giải quyết cho bài toán này. Hai nhóm tiếp cận chính để giải quyết cho bài toán tổng hợp hình ảnh y học là nhóm các cách tiếp cận truyền thống và nhóm các cách tiếp cận dựa trên học sâu. Trên cơ sở phân tích một số phương pháp tổng hợp hiện tại, sự kém hiệu quả trong việc tổng hợp hình ảnh là do một số nguyên nhân chính sau:

- Thứ nhất là các hình ảnh đầu vào thường có chất lượng thấp như cường độ sáng và độ tương phản thấp, thiếu sắc nét. Hạn chế này ảnh hưởng đáng kể đến các thuật toán tổng hợp hình ảnh.
- Thứ hai là các thuật toán tổng hợp cho các thành phần cơ sở và chi tiết chưa thực sự hiệu quả. Cụ thể, thuật toán trung bình tổng hợp cho thành phần cơ sở có thể dẫn tới sự suy giảm cường độ sáng và độ tương phản của hình ảnh tổng hợp. Thuật toán lựa chọn Max để tổng hợp cho thành phần chi tiết có thể làm cho các hình ảnh tổng hợp bị mất các thông tin chi tiết từ hình ảnh gốc.

Vì vậy, định hướng của luận án sẽ đề xuất các thuật toán mới hiệu quả để giải quyết cho những hạn chế đã đề cập ở trên. Thêm vào đó, chương này cũng giới thiệu một số các kiến thức nền tảng cần thiết để phục vụ cho việc xây dựng thuật toán đề xuất trong các chương tiếp theo. Các kiến thức nền tảng bao gồm: Thuật toán meta-heuristic và thuật toán tối ưu hoá MPA, hàm năng lượng cục bộ (LEF), thuật toán tổng hợp dựa trên hàm năng lượng cục bộ kết hợp với toán tử la bàn Prewitt (FR_PCO), cấu trúc Tensor và toán tử phát hiện đặc trưng của cấu trúc Tensor, bộ lọc hướng dẫn quay vòng (RGF) và bộ lọc đường cong trung bình có trọng số (WMCF), cân bằng biểu đồ xám thích nghi giới hạn độ tương phản (CLAHE), toán

tử phát hiện biên (LED), bộ lọc khuếch tán dị hướng ADF và thuật toán phục hồi hình ảnh dựa trên tổng biến thiên (TVIR).

Cuối Chương 1, luận án cũng giới thiệu một số chỉ số đánh giá thường được sử dụng để đánh giá chất lượng của hình ảnh tổng hợp.

CHƯƠNG 2. NÂNG CAO CHẤT LƯỢNG HÌNH ẢNH DỰA TRÊN PHÂN RÃ BA THÀNH PHẦN VÀ THUẬT TOÁN MPA

Trong Chương 1, luận án đã đi phân tích một số hạn chế của các phương pháp tổng hợp hiện tại. Trong đó, chất lượng hình ảnh đầu vào thấp là một trong những nguyên nhân ảnh hưởng đáng kể đến hiệu quả của các thuật toán tổng hợp. Vì vậy, trong chương 2 này, luận án sẽ đi theo hướng tiếp cận cải thiện chất lượng của hình ảnh đầu vào nhằm hướng tới cải thiện hiệu quả của các thuật toán tổng hợp hình ảnh. Các hình ảnh y học chụp cộng hưởng từ não thường có hạn chế về chất lượng như độ tương phản và cường độ sáng thấp. Vì vậy, thuật toán nâng cao chất lượng ảnh đề xuất nhằm giải quyết hạn chế về chất lượng của các hình ảnh đầu vào ở khía cạnh cải thiện độ tương phản và cường độ sáng. Phần kiểm tra sự hiệu quả của thuật toán nâng cao chất lượng ảnh đề xuất sẽ được đánh giá trong cả hai trường hợp. Đánh giá riêng bởi các chỉ số đánh giá chất lượng ảnh và đánh giá khi kết hợp với các thuật toán tổng hợp hình ảnh mà nó đóng vai trò như bước tiền xử lý cho các ảnh đầu. Các kết quả thực nghiệm đã cho thấy thuật toán nâng cao chất lượng ảnh đề xuất có hiệu quả trong việc nâng cao hiệu quả của việc tổng hợp hình ảnh. Thuật toán đề xuất đã được công bố tại công trình [CT1]. Ứng dụng của phương pháp nâng cao chất lượng ảnh đề xuất để tiền xử lý ảnh đầu vào với mục đích tăng hiệu quả của việc tổng hợp hình ảnh đã được công bố tại công trình [CT2] trong “Danh mục các công trình của tác giả”.

2.1. Ý tưởng

Các hình ảnh chụp cộng hưởng từ não thường có độ tương phản và cường độ sáng thấp. Một số các kỹ thuật tăng cường ảnh hiện nay đều có những hạn chế nhất định. Thứ nhất, các phương pháp nâng cao chất lượng ảnh thường thao tác trực tiếp trên các điểm ảnh. Trong trường hợp hình ảnh có một chút nhiễu, nó có thể dẫn đến khuếch tán nhiễu trong ảnh. Thứ hai, trong quá trình biến đổi ảnh, các chi tiết của ảnh cũng có thể bị suy giảm hoặc tăng lên và nhiễu cũng có thể xuất hiện do ảnh bị tăng cường quá mức. Do đó, để cải thiện những hạn chế đề cập ở trên. Ý tưởng của thuật toán nâng cao chất lượng ảnh đề xuất trong Chương 2 này sẽ không trực tiếp thao tác trên một kênh cường độ sáng của hình ảnh đầu vào. Thay vào đó, phương pháp đề xuất sẽ tách hình ảnh đầu vào thành các thành phần chính có trong ảnh như: thành phần cấu trúc, thành phần kết cấu và thành phần nhiễu. Từ đó, việc tăng cường chất lượng trên từng thành phần sẽ trở nên hiệu quả hơn và kiểm soát được nhiễu phát sinh trong quá trình tăng cường ảnh.

Hiện nay, việc tăng cường chất lượng ảnh dựa trên phân rã ảnh ba thành phần cũng đã xuất hiện trong một số nghiên cứu. Ví dụ, Lim và các cộng sự [109] đã giới

thiệt một phương pháp tăng cường chất lượng ảnh nhiễu và thiếu sáng. Sau khi phân rã thành ba thành phần cấu trúc, kết cấu và nhiễu, hai thành phần cấu trúc và kết cấu sẽ được tăng cường chất lượng và thành phần nhiễu sẽ bị loại bỏ. Hình ảnh tăng cường thu được bằng tổng của hai thành phần cấu trúc và kết cấu đã được tăng cường. Các tiếp cận của Lim và các cộng sự vẫn còn có những hạn chế. Thứ nhất, việc loại bỏ trực tiếp thành phần nhiễu ra khỏi ảnh sẽ ít nhiều dẫn đến mất thông tin chi tiết của hình ảnh. Hơn nữa, việc tăng cường từng thành phần cấu trúc và kết cấu riêng lẻ rồi sau đó tạo ra hình ảnh tăng cường không có cơ chế đảm bảo cho hình ảnh tăng cường được tạo ra có chất lượng tốt. Ảnh có thể bị tăng cường quá mức dẫn đến mất chi tiết và nhiễu có thể phát sinh. Vì vậy, để giải quyết những hạn chế nêu ở trên, cách tiếp cận của phương pháp đề xuất như sau:

- Thứ nhất, đề xuất một phương pháp phân rã ba thành phần dựa trên bộ lọc Gauss, phương pháp phục hồi hình ảnh TVIR và bộ lọc ADF.
- Thứ hai, sử dụng cấu trúc Tensor để bổ sung thêm các đặc trưng cho ảnh với mục đích bù đắp thông tin chi tiết có thể mất mát trong quá trình tổng hợp cũng như tăng thêm chi tiết cho ảnh.
- Thứ ba, sử dụng một thuật toán tối ưu hoá MPA để tìm ra các hệ số thích hợp ứng từng thành phần được tăng cường với một hàm tối ưu đề xuất để kiểm soát chất lượng của hình ảnh đầu ra. Điều này giúp đảm bảo cho hình ảnh sau khi tăng cường không bị nhiễu phát sinh cũng như bị tăng cường quá mức.

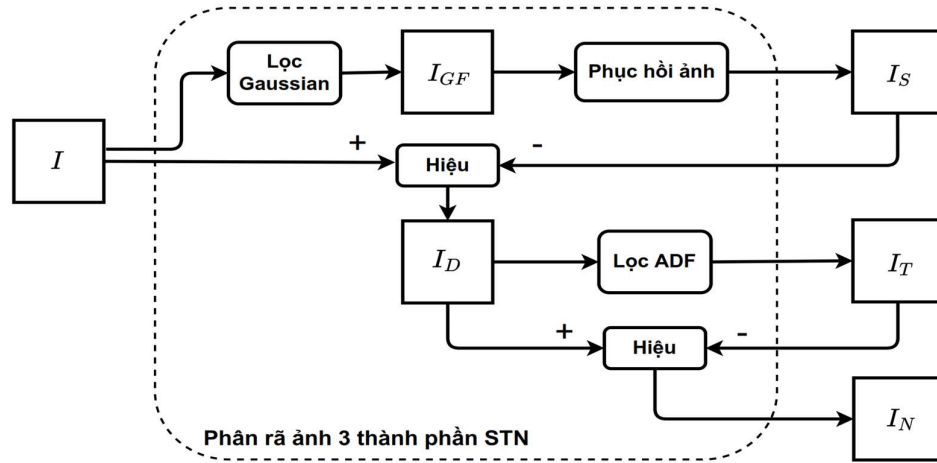
2.2. Đề xuất phương pháp nâng cao chất lượng ảnh

Phần này mô tả chi tiết các bước của phương pháp nâng cao chất lượng ảnh. Phương pháp nâng cao chất lượng ảnh đề xuất dựa trên phân rã ba thành phần và thuật toán tối ưu hoá MPA (được gọi là IE_TCID_MPA). Thuật toán đề xuất có thể mô tả ngắn gọn qua ba giai đoạn chính. Giai đoạn thứ nhất là phân rã hình ảnh đầu vào thành ba lớp: lớp cấu trúc, lớp kết cấu và lớp nhiễu. Giai đoạn thứ hai là bổ sung thêm lớp thành phần chi tiết và tăng cường các thành phần đã có. Giai đoạn thứ ba là tìm các hệ số tối ưu cho từng thành phần thông qua thuật toán tối ưu hoá MPA. Hình ảnh tăng cường thu được thông qua tổng các thành phần nhân với hệ số tối ưu tìm được tương ứng.

2.2.1. Đề xuất thuật toán phân rã hình ảnh ba thành phần

Một hình ảnh có thể coi là sự tổng hợp của nhiều lớp chứa các thông tin khác nhau. Ba lớp thông tin chính thường biểu diễn cho một hình ảnh như: lớp cấu trúc, lớp kết cấu và lớp nhiễu. Lớp cấu trúc thường được biết đến là lớp cơ sở của ảnh. Lớp kết cấu thường được biết đến như lớp chứa chi tiết của ảnh và không chứa nhiễu. Lớp

nhiều biểu diễn cho nhiều xuất hiện trong ảnh. Việc tách một hình ảnh ra thành 3 lớp thông tin như vậy có ý nghĩa rất quan trọng trong rất nhiều các ứng dụng của xử lý ảnh như phân đoạn ảnh, phát hiện đối tượng trong ảnh và nâng cao chất lượng ảnh. Việc phân rã ảnh ba thành phần như vậy đã được giới thiệu trong một số nghiên cứu. Ví dụ, Lim và các cộng sự [109] đã giới thiệu một phương pháp phân rã 3 thành phần: cơ sở, chi tiết và nhiễu và ứng dụng nó để nâng cao chất lượng cho ảnh nhiễu và tối. Trong nghiên cứu của Lim và các cộng sự, để có thể tách được thành phần cấu trúc và thành phần kết cấu trong ảnh, thuật toán FTVd (Fast total variation de-convolution) [110] đã được sử dụng. Thành phần kết cấu có chứa nhiễu và vì vậy nó đã được khử nhiễu bởi bộ lọc BM3D [111]. Sau đó, thành phần nhiễu được xác định bằng hiệu giữa thành phần kết cấu và thành phần đã khử nhiễu.



Hình 2.1. Sơ đồ phân rã hình ảnh ba thành phần

Thuật toán 2.1: STN

Đầu vào: Ảnh xám I và các tham số cần thiết.

Đầu ra: Ba hình ảnh I_S , I_T , I_N .

Bước 1: Hình ảnh đầu vào I được làm mờ bởi bộ lọc Gauss, thu được một hình ảnh I_{GF} .

Bước 2: Thuật toán phục hồi ảnh tổng biến thiên (TVIR) [104] được áp dụng cho hình ảnh I_{GF} để thu được thành phần cấu trúc I_S .

Bước 3: Thành phần chi tiết I_D được xác định theo công thức (2.1).

$$I_D = I - I_S \quad (2.1)$$

Bước 4: Thành phần kết cấu I_T thu được bằng cách áp dụng bộ lọc khuếch tán dị hướng ADF [101] cho thành phần I_D .

Bước 5: Thành phần nhiễu I_N được xác định theo công thức (2.2).

$$I_N = I_D - I_T \quad (2.2)$$

Tuy nhiên, việc thực hiện phân tách giữa hai lớp cấu trúc và kết cấu trực tiếp từ ảnh đầu vào có thể sẽ không hiệu quả trong trường hợp nếu ảnh có nhiều dù ít hay nhiều. Vì vậy, trong mục này, một thuật toán phân rã hình ảnh thành 3 lớp được đề xuất. Việc phân rã một hình ảnh đầu vào thành 3 thành phần được minh họa chi tiết trong Thuật toán 2.1 và Hình 2.1.

2.2.2. Thiết kế hàm mục tiêu

Hàm mục tiêu (FN - Fitness function) đóng một vai trò quan trọng trong thuật toán IE_TCID_MPA vì nó được sử dụng để kiểm soát chất lượng của hình ảnh đầu ra. Vì vậy, phần này trình bày ý tưởng của việc đề xuất hàm mục tiêu. Dựa trên hàm chỉ số tương phản Michelson như được tính trong công thức (2.3). Hàm mục tiêu này đã được ứng dụng trong một số nghiên cứu về nâng cao chất lượng hình ảnh thiếu sáng [112], [113].

$$FN = \frac{V}{M}(E_2 - E_1) \quad (2.3)$$

Trong đó:

- V , M và E_2 lần lượt là phương sai, trung bình cường độ sáng và Entropy của hình ảnh thu được sau khi tăng cường.
- E_1 là Entropy của hình ảnh trước khi tăng cường.

Dựa trên hàm mục tiêu FN , một hàm mục tiêu mới (O_F) được đề xuất như mô tả trong công thức (2.4). Hàm mục tiêu đề xuất được bổ sung thêm thành phần tỉ số cực đại trên nhiễu PSNR (Peak signal-to-noise ratio), Gradient của hai hình ảnh trước và sau khi tăng cường. Mục đích của việc bổ sung thêm các thành phần đó nhằm kiểm soát độ sắc nét của hình ảnh, sự tương đồng của hình ảnh tăng cường với hình ảnh đầu vào. Điều này nhằm tạo ra sự hài hoà giữa các yếu tố đánh giá chất lượng ảnh như cường độ sáng, độ tương phản và độ sắc nét. Một hình ảnh bị mờ hoặc quá sắc nét thì cũng sẽ không phải là tốt đối với một bức ảnh bởi vì các chỉ số đánh giá khác như cường độ sáng, độ tương phản cũng có thể bị suy giảm.

$$O_F = \frac{V}{M} \left((E_2 - E_1) + \frac{G_2 - G_1}{PSNR(I_{Temp}, I)} \right) \quad (2.4)$$

- I là hình ảnh trước khi tăng cường.
- I_{Temp} là hình ảnh tăng cường tạm thời tính theo từng vòng lặp của thuật toán tối ưu.
- V là phương sai của hình ảnh I_{Temp} .
- M là trung bình cường độ sáng của hình ảnh I_{Temp} .
- E_1 và E_2 lần lượt là Entropy của hình ảnh I và I_{Temp} tương ứng.

- G_1 và G_2 là Gradient của I và I_{Temp} tương ứng.
- PSNR là tỉ số cực đại trên nhiễu được tính toán theo công thức (2.5).

$$PSNR(I_{Temp}, I) = 10 \log_{10} \frac{Max^2}{MSE} \quad (2.5)$$

Với MSE được tính toán theo công thức (2.6) và Max nhận giá trị 255.

$$MSE = \frac{1}{H \cdot W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} (I_{Temp}(i, j) - I(i, j))^2 \quad (2.6)$$

2.2.3. Phương pháp nâng cao chất lượng ảnh IE_TCID_MPA

Trong phần này, chi tiết từng bước của phương pháp tăng cường IE_TCID_MPA được mô tả chi tiết trong Thuật toán 2.2. Hình 2.5 là sơ đồ hoạt động của phương pháp IE_TCID_MPA.

Thuật toán 2.2: IE_TCID_MPA

Đầu vào:

- Ảnh xám I ($H \times W$).
- Các tham số cần thiết của thuật toán MPA và phương pháp phân rã ảnh.

Đầu ra: Ảnh được tăng cường I_E

Bước 1: Phân rã ảnh

Phân rã ảnh I thành: I_S , I_T và I_N bởi áp dụng thuật toán phân rã 3 thành phần (Thuật toán 2.1) được giới thiệu trong trong mục 2.2.1. Các thành phần thu được như minh hoạ trong Hình 2.2.

Bước 2: Tăng cường các thành phần thu được và bổ sung thành phần chi tiết

Bước 2.1: Thuật toán CLAHE [100] được sử dụng để nâng cao chất lượng cho thành phần cấu trúc I_S , thu được thành phần I_{CLAHE} như minh hoạ trong Hình 2.3.

Bước 2.2: Toán tử xác định thành phần nổi bật của cấu trúc Tensor [96] được áp dụng cho thành phần I_{CLAHE} , thu được thành phần I_{STS} .

Bước 2.3 Toán tử phát hiện biên Laplace (LED) là được áp dụng cho thành phần I_T , thu được thành phần biên I_{LED} .

Bước 2.4: Tạo ra một thành phần đặc trưng I_{T_E} bằng cách áp dụng toán tử Max đối với hai thành phần I_T và I_{LED} như công thức (2.7). Hình 2.4 minh hoạ cho thành phần I_{T_E} .

$$I_{T_E} = \text{Max}(I_T, I_{LED}) \quad (2.7)$$

Bước 3: Tính xác định tham số tối ưu dựa trên MPA và tính ảnh tăng cường

Bước 3.1: Khởi tạo ngẫu nhiên các tham số tối ưu cần tìm ($\beta_1 \in [0.9, 1.5]$, $\beta_2 \in [0.1, 1.5]$, $\beta_3 \in [0.9, 1.5]$ và $\beta_4 \in [0.1, 1.5]$) và tính hình ảnh tăng cường tạm thời I_{Temp} theo công thức (2.8).

$$I_{Temp} = \beta_1 \cdot I_{CLAHE} + \beta_2 \cdot I_{STS} + \beta_3 \cdot I_{T_E} + \beta_4 \cdot I_N \quad (2.8)$$

Bước 3.2: Thực hiện vòng lặp của thuật toán MPA để cập nhật tham số

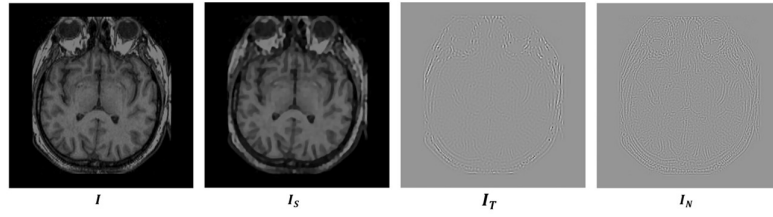
Nếu điều kiện kết thúc chưa thỏa mãn:

Bước 3.2.1: Tính giá trị của hàm tối ưu O_F theo công thức (2.4).

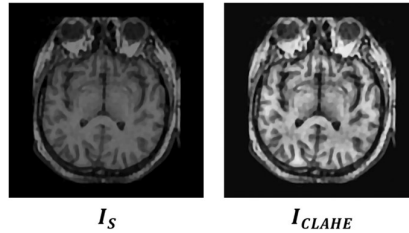
Bước 3.2.2: Cập nhật lại các tham số tối ưu β_1 , β_2 , β_3 và β_4 theo thuật toán MPA.

Nếu điều kiện kết thúc thỏa mãn:

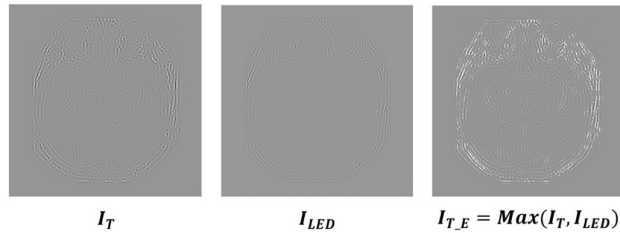
Bước 3.2.3: Thu được các tham số tối ưu và tính hình ảnh tăng cường I_E tương ứng với các tham số tối ưu thu được giống như công thức (2.8).



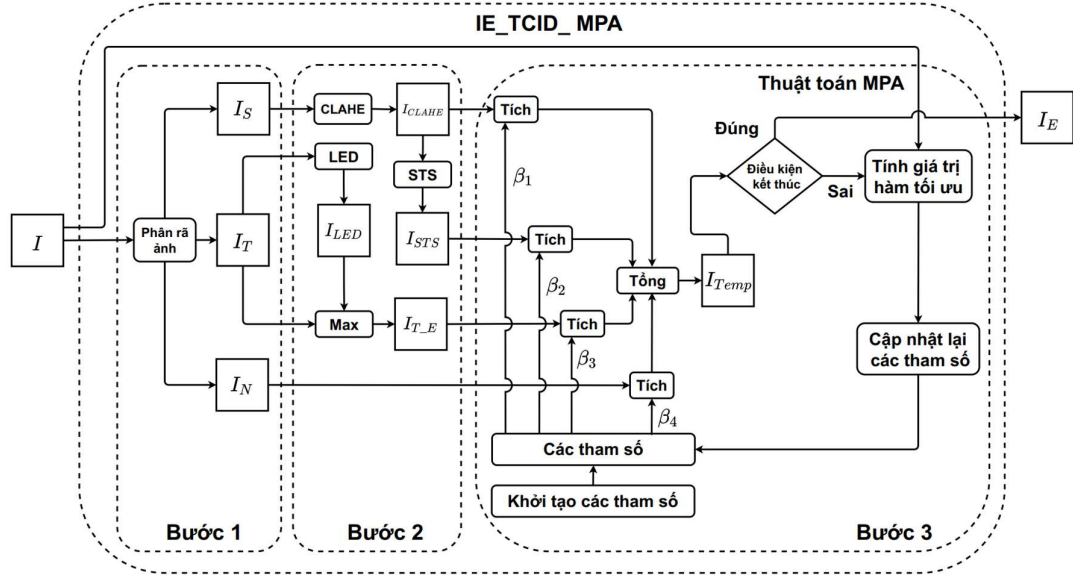
Hình 2.2. Minh họa phân rã 3 thành phần cho ảnh I



Hình 2.3. Kết quả tăng cường thành phần cấu trúc



Hình 2.4. Minh họa ba thành phần đặc trưng



Hình 2.5. Sơ đồ thuật toán tăng cường chất lượng ảnh IE_TCID_MPA

2.2.4. Độ phức tạp tính toán của thuật toán IE_TCID_MPA

Thuật toán IE_TCID_MPA bao gồm 3 bước:

- Bước 1: Phân rã ảnh 3 thành phần.
- Bước 2: Tăng cường thành phần phân rã và bổ sung thành phần chi tiết theo toán tử phát hiện thành phần đặc trưng của cấu trúc Tensor STS.
- Bước 3: Sử dụng thuật toán MPA để tìm các tham số tối ưu cho hàm mục tiêu O_F và tính hình ảnh đầu ra.

Tại bước 1: Thuật toán phân rã ảnh ba thành phần TCID bao gồm bộ lọc Gauss, bộ lọc ADF và phương pháp phục hồi hình ảnh TVIR [104]. Đối với một hình ảnh đầu vào kích thước $N \times N$, gọi $w \times w$ là kích thước của bộ lọc Gauss, khi đó độ phức tạp tính toán của bộ lọc Gauss là $O(N^2 \cdot w^2)$. Gọi l là số vòng lặp của bộ lọc ADF, khi đó độ phức tạp tính toán của bộ lọc ADF là $O(N^2 \cdot l)$. Gọi k là số vòng lặp của phương pháp phục hồi hình ảnh TVIR, khi đó độ phức tạp tính toán của phương pháp TVIR là $O(N^2 \cdot k)$. Như vậy, độ phức tạp tính toán ở Bước 1 sẽ là $O(N^2 \cdot w^2 + N^2 \cdot l + N^2 \cdot k)$.

Tại bước 2: Độ phức tạp tính toán cho một số phương pháp tăng cường ảnh được xác định như sau. Phương pháp CLAHE có độ phức tạp tính toán là xấp xỉ $O(N^2)$ trong trường hợp số lượng các ô chia từ ảnh là nhỏ. Phương pháp LED thực hiện phép nhân chập giữa ảnh đầu vào với ma trận cửa sổ kích thước 3×3 . Do đó nó có độ phức tạp tính toán là $O(N^2 \cdot 3^2)$. Phép toán lấy Max giữa hai ma trận sẽ có độ phức tạp tính toán là $O(N^2)$. Như vậy, độ phức tạp tính toán tại Bước 2 sẽ là $O(N^2 + N^2 \cdot 9 + N^2) \approx O(N^2)$.

Tại bước 3: Thuật toán MPA được sử dụng để tìm các tham số tối ưu cho hàm mục tiêu O_F . Theo tác giả Faramarzi và các cộng tác sự [31], thuật toán MPA có độ phức tạp tính toán là $O(t.(n.d + Cof.n))$. Trong đó t là số vòng lặp của thuật toán MPA, n là số cá thể trong bầy, d là chiều của giải pháp và Cof là chi phí đánh giá hàm mục tiêu. Hàm mục tiêu của thuật toán IE_TCID_MPA (công thức 2.4) bao gồm các hàm tính phương sai, trung bình cường độ sáng, hiệu Entropy giữa hình ảnh tổng hợp với hình ảnh MRI, hiệu Gradient giữa hình ảnh tổng hợp với hình ảnh MRI và cuối cùng là hàm tỉ số cực đại trên nhiễu (PSNR) giữa hình ảnh tổng hợp với hình ảnh MRI. Những hàm này đều có độ phức tạp tính toán là $O(N^2)$ nên hàm mục tiêu có độ phức tạp $O(N^2)$. Như vậy, phần tìm các tham số tối ưu cho việc tổng hợp các thành phần sẽ có độ phức tạp tính toán là $O(t.(n.d + N^2.n))$. Cuối cùng, độ phức tạp tính toán của hàm để tính ma trận ảnh đầu ra I_E khi đã tìm được các tham số tối ưu là $O(N^2)$. Do đó, độ phức tạp tính toán tại bước 3 sẽ là $O(t.(n.d + N^2.n) + N^2)$.

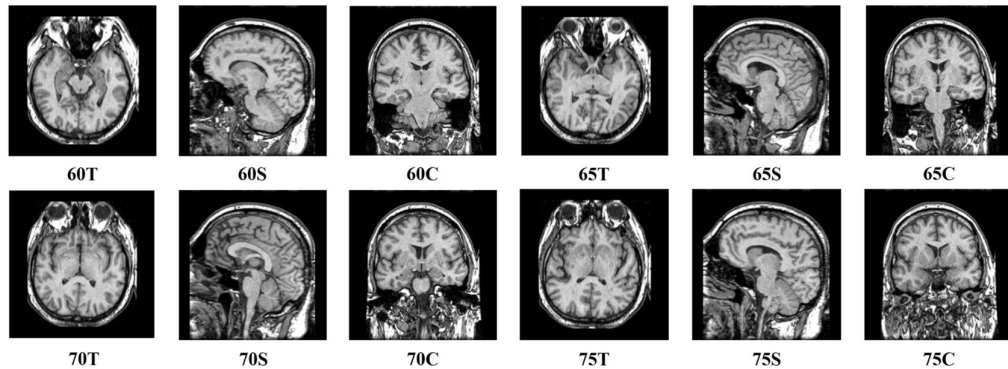
Như vậy, độ phức tạp tính toán của thuật toán IE_TCID_MPA sẽ là $O(N^2.w^2 + N^2.l + N^2.k + N^2 + t.(n.d + N^2.n) + N^2)$.

2.3. Thực nghiệm và đánh giá

2.3.1. Dữ liệu thực nghiệm

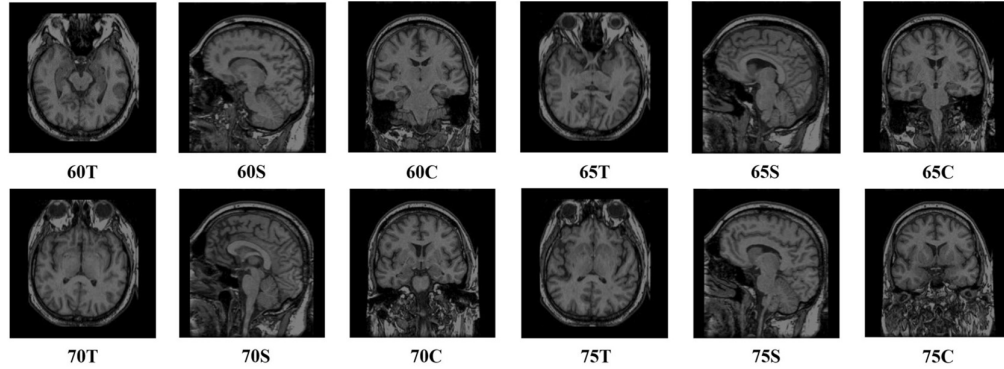
Chín mươi cặp hình ảnh y học MRI và PET (180 hình ảnh) có kích thước 256×256 đã được sử dụng cho các thực nghiệm. Những hình ảnh này được thu thập từ lát cắt thứ 50 đến 79 theo các trục Transaxial (T), Sagittal (S) và Coronal (C) từ nguồn <http://www.med.harvard.edu/AANLIB/>. 180 hình ảnh y học được chia thành các tập dữ liệu để phục vụ cho các thực nghiệm như sau:

- Tập dữ liệu (D0) bao gồm 90 hình ảnh MRI (30 hình ảnh theo mỗi trục T, S và C) được sử dụng để đánh giá thuật toán nâng cao chất lượng hình ảnh đề xuất. Một số hình ảnh MRI trong D0 được minh họa trong Hình 2.6.



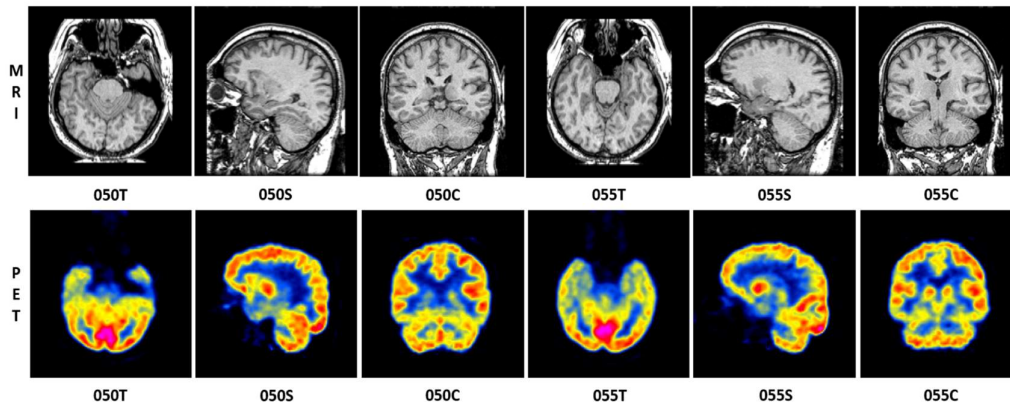
Hình 2.6. Minh họa 12 hình ảnh MRI trong tập D0

- Tập dữ liệu (D1) được tạo từ tập dữ liệu D0 bằng cách giảm độ tương phản và cường độ sáng được sử dụng để đánh giá thuật toán nâng cao chất lượng ảnh IE_TCID_MPA. Một số hình ảnh MRI trong tập D1 được minh họa trong Hình 2.7.



Hình 2.7. Minh họa 12 hình ảnh MRI có tương phản thấp trong tập D1

- Tập dữ liệu (D2) bao gồm ba hình ảnh (70T, 70S, 70C) được sử dụng để kiểm tra tính hiệu quả của thuật toán MPA mà nó được sử dụng trong thuật toán nâng cao chất lượng ảnh IE_TCID_MPA.
- Tập dữ liệu (D3) bao gồm 90 cặp hình ảnh MRI và PET được sử dụng để đánh giá thuật toán IE_TCID_MPA trong việc cải thiện hiệu quả của các thuật toán tổng hợp hình ảnh y học hiện tại. Những hình ảnh này bao gồm Set-T (60 hình ảnh MRI và PET), Set-S (60 hình ảnh MRI và PET) và Set-C (60 hình ảnh MRI và PET) theo trục T, S và C. Một số cặp hình ảnh trong tập dữ liệu (D3) được minh họa trong Hình 2.8.



Hình 2.8. Minh họa 6 cặp MRI-PET trong tập D3

2.3.2. Thiết lập thực nghiệm

Cấu hình máy tính và phần mềm đã sử dụng để tiến hành thực nghiệm:

- Intel Core i9 10900K, bộ xử lý 3.7 GHz và RAM 64 GB.

- Hệ điều hành Windows 10. Phần mềm Matlab R2020b.

Một số thực nghiệm được thiết kế để đánh giá hiệu quả của thuật toán nâng cao chất lượng ảnh đề xuất IE_TCID_MPA như sau:

Thực nghiệm 1:

Để kiểm tra sự hiệu quả của thuật toán tối ưu hoá MPA trong thuật toán tăng cường chất lượng ảnh IE_TCID_MPA, bảy thuật toán tối ưu hóa dựa trên meta-heuristic khác đã được sử dụng để so sánh. Các thuật toán tối ưu hóa này bao gồm DA [87], ALO [88], GWO [86], MVO [82], SSA [84], SCA [85] và WOA [83] (xem chi tiết trong Bảng 2.1). Mỗi thuật toán được sử dụng chạy 30 lần trên tập dữ liệu D2 (070T, 070S, 070C). Bốn chỉ số: Tốt nhất, Kém nhất, Trung bình và Độ lệch chuẩn được sử dụng để đánh giá hiệu quả của các thuật toán tối ưu. Hơn nữa, một kiểm định phi tham số (Wilcoxon) [114] được sử dụng để kiểm tra kết quả thực nghiệm.

Bảng 2.1. Một số thuật toán tối ưu

STT	Các thuật toán	Năm
1	DA (Dragonfly algorithm)	2015
2	ALO (Ant Lion Optimizer)	2015
3	GWO (Grey Wolf Optimizer)	2014
4	MVO (Multi-Verse Optimizer)	2015
5	SSA (Salp Swarm Algorithm)	2017
6	SCA (Sine Cosine Algorithm)	2016
7	WOA (Whale Optimization Algorithm)	2016

Thực nghiệm 2:

Để kiểm tra hiệu quả của thuật toán nâng cao chất lượng ảnh đề xuất, một số thuật toán nâng cao chất lượng hình ảnh khác đã được sử dụng để so sánh. Các thuật toán này gồm: FCCE [115], EFF [116], EGIF [117], RRM [118], FFM [119], SDD [120] được mô tả trong Bảng 2.2. Bốn chỉ số đánh giá chất lượng hình ảnh (MLI, CI, E và AG) và cả hai tập dữ liệu D0 và D1 được sử dụng trong thực nghiệm này.

Bảng 2.2. Một số thuật toán nâng cao chất lượng ảnh

STT	Các thuật toán	Năm
1	NE (No Enhancement)	
2	FCCE (Fuzzy-Contextual Contrast Enhancement)	2017
3	EFF (Exposure Fusion Framework)	2017
4	EGIF (Effective Guided Image Filtering)	2018
5	RRM (Robust Retinex Model)	2018
6	FFM (Fractional-Order Fusion Model)	2019
7	SDD (Semi-Decoupled Decomposition)	2020

Thực nghiệm 3:

Để kiểm tra hiệu quả của thuật toán nâng cao chất lượng ảnh IE_TCID_MPA trong việc cải thiện hiệu quả của các thuật toán tổng hợp hình ảnh, nó đã được sử dụng để tiền xử lý ảnh đầu vào. Một số thuật toán tổng hợp hình ảnh y học được đề xuất gần đây đã được lựa chọn để so sánh sự hiệu quả trong trường hợp trước và sau khi sử dụng thuật toán IE_TCID_MPA. Các thuật toán này được mô tả trong Bảng 2.3. Các chỉ số đánh giá (MLI, CI, E, AG, $Q^{AB/F}$ và MI) và tập dữ liệu D3 được sử dụng trong thực nghiệm này.

Bảng 2.3. Một số thuật toán tổng hợp hình ảnh y học

STT	Các thuật toán	Kí hiệu	Năm
1	PC-LLC-NSCT [121]	Alg1	2019
2	NSST-MSMG-PCNN [122]	Alg2	2020
3	MLCF-MLMG-PCNN [99]	Alg3	2021
4	JBF-LGE [29]	Alg4	2021
5	CSE [123]	Alg5	2021
6	CNPS-NSST [124]	Alg6	2021
7	DTNP-NSCT [40]	Alg7	2021

2.3.3. Kết quả thực nghiệm

Các tham số cần thiết cấu hình cho các thuật toán tối ưu như sau:

- Số lượng con mồi: $n = 50$.
- Số lượng vòng lặp của thuật toán MPA: $l_{max} = 50$.
- Cận dưới: [0.9; 0.1; 0.9; 0.1]
- Cận trên: [1.5; 1.5; 1.5; 1.5]
- Số chiều của tham số tối ưu: $dim = 4$.

Bên cạnh các tham số cần thiết cấu hình ở trên, các tham số riêng cho từng thuật toán tối ưu (nếu có) được thiết lập mặc định. Riêng thuật toán MPA sử dụng một số tham số riêng như sau:

- Tham số $K = 0.5$.
- Tham số $FADs = 0.2$.
- Tham số r là ngẫu nhiên nằm trong $[0,1]$.

Kết quả thực nghiệm 1: Kết quả của thực nghiệm 1 được minh họa trong các Bảng 2.4, 2.5 và 2.6. Kết quả của 30 lần chạy khác của mỗi thuật toán tối ưu trên tập dữ liệu D2 cũng được hiển thị trong biểu đồ Hình 2.9.

Bảng 2.4. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70T)

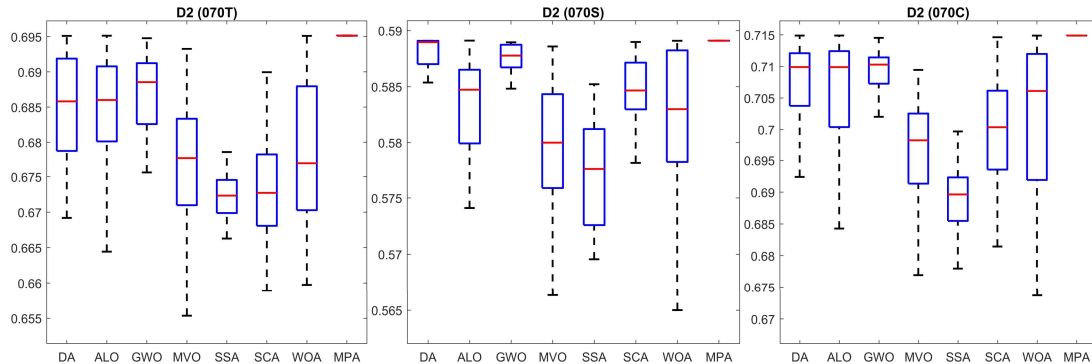
Thuật toán	Tốt nhất	Kém nhất	Trung bình	Độ lệch chuẩn
DA	0.6951114140	0.6691981549	0.6843984697	0.0083739791
ALO	0.6951376300	0.6560385934	0.6837011663	0.0094079738
GWO	0.6947800377	0.6756086978	0.6871735337	0.0053541210
MVO	0.6932673088	0.6553480335	0.6760255938	0.0105207954
SSA	0.6784831232	0.6535586079	0.6718182663	0.0045815877
SCA	0.6899561101	0.6588596616	0.6734006875	0.0077866788
WOA	0.6951131201	0.6597476869	0.6782555567	0.0104211391
MPA	0.6951698787	0.6951116338	0.6951370386	0.0000177660

Bảng 2.5. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70S)

Thuật toán	Tốt nhất	Kém nhất	Trung bình	Độ lệch chuẩn
DA	0.5891014865	0.5830187029	0.5880647400	0.0014460701
ALO	0.5891124202	0.5633896257	0.5823641279	0.0057202657
GWO	0.5889588040	0.5848183257	0.5874600779	0.0013551200
MVO	0.5885907411	0.5663212448	0.5791229963	0.0059970759
SSA	0.5852157777	0.5695596321	0.5770870362	0.0046830438
SCA	0.5889922005	0.5743679148	0.5843664631	0.0033185882
WOA	0.5890997537	0.5650032866	0.5820477472	0.0067166869
MPA	0.5891135503	0.5890983433	0.5891057043	0.0000041417

Bảng 2.6. Bốn chỉ số đánh giá từ 30 lần chạy độc lập trên tập dữ liệu D2 (70C)

Thuật toán	Tốt nhất	Kém nhất	Trung bình	Độ lệch chuẩn
DA	0.7148948786	0.6923901588	0.7077459356	0.0066049362
ALO	0.7047934850	0.7148991946	0.7047934850	0.0116574628
GWO	0.7145312501	0.6902332930	0.7084368980	0.0053736986
MVO	0.7094556379	0.6685034219	0.6955996646	0.0104442968
SSA	0.7072471110	0.6779709192	0.6895636698	0.0062503111
SCA	0.7146313502	0.6814454094	0.6998585259	0.0084873370
WOA	0.7148912743	0.6736590274	0.7011148428	0.0129814423
MPA	0.7149109070	0.7148841067	0.7148959516	0.0000053200



Hình 2.9. Giá trị của hàm tối ưu thu được từ các thuật toán trên tập D2.

Rõ ràng độ lệch chuẩn thu được của hàm tối ưu sau 30 lần chạy độc lập là thấp nhất khi sử dụng thuật toán MPA, trong khi giá trị trung bình đạt giá trị lớn nhất khi so sánh với các thuật toán tối ưu khác. Các kết quả thực nghiệm này giải thích tại sao thuật toán MPA đã được lựa chọn trong thuật toán tăng cường ảnh đề xuất. Hơn nữa, từ Bảng 2.7, kết quả thu được từ kiểm định Wilcoxon có ý nghĩa thống kê vì giá trị p-values nhỏ hơn 0.05.

Bảng 2.7. Giá trị p-values từ kiểm định thống kê Wilcoxon

Tập dữ liệu	Thuật toán	P-values
D2 (070T)	MPA với DA	3.019859359162151e-11
	MPA với ALO	2.609852787955851e-10
	MPA với GWO	3.019859359162151e-11
	MPA với MVO	3.019859359162151e-11
	MPA với SSA	3.019859359162151e-11
	MPA với SCA	3.019859359162151e-11
	MPA với WOA	4.077164846825348e-11
D2 (070S)	MPA với DA	3.474196606831804e-10
	MPA với ALO	5.072313498155892e-10
	MPA với GWO	3.019859359162151e-11
	MPA với MVO	3.019859359162151e-11
	MPA với SSA	3.019859359162151e-11
	MPA với SCA	3.019859359162151e-11
	MPA với WOA	4.504322112705322e-11
D2 (070C)	MPA với DA	8.101362326146523e-10
	MPA với ALO	1.596405099603837e-07
	MPA với GWO	3.019859359162151e-11
	MPA với MVO	3.019859359162151e-11
	MPA với SSA	3.019859359162151e-11
	MPA với SCA	3.019859359162151e-11
	MPA với WOA	3.689725853981014e-11

Kết quả thực nghiệm 2:

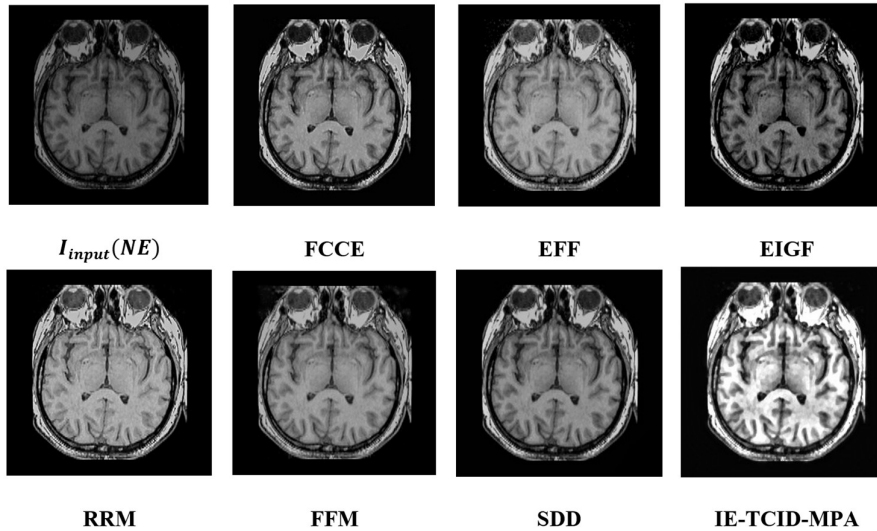
Kết quả của thực nghiệm 2 với bốn chỉ số đánh giá (MLI, CI, Entropy và AG) được trình bày trong các Bảng 2.8, Bảng 2.9 và các Hình 2.11 và Hình 2.13. Đầu ra của các thuật toán tăng cường ảnh được minh họa trong các Hình 2.10, và Hình 2.12. Dễ dàng nhận thấy rằng các chỉ số MLI, CI, Entropy và AG thu được từ thuật toán nâng cao chất lượng ảnh được đề xuất IE_TCID_MPA là cao nhất so với các chỉ số của các thuật toán còn lại trong cả hai tập dữ liệu D0 và D1.

Bảng 2.8. Các chỉ số đánh giá từ các thuật toán nâng cao chất lượng ảnh trên tập $D0$

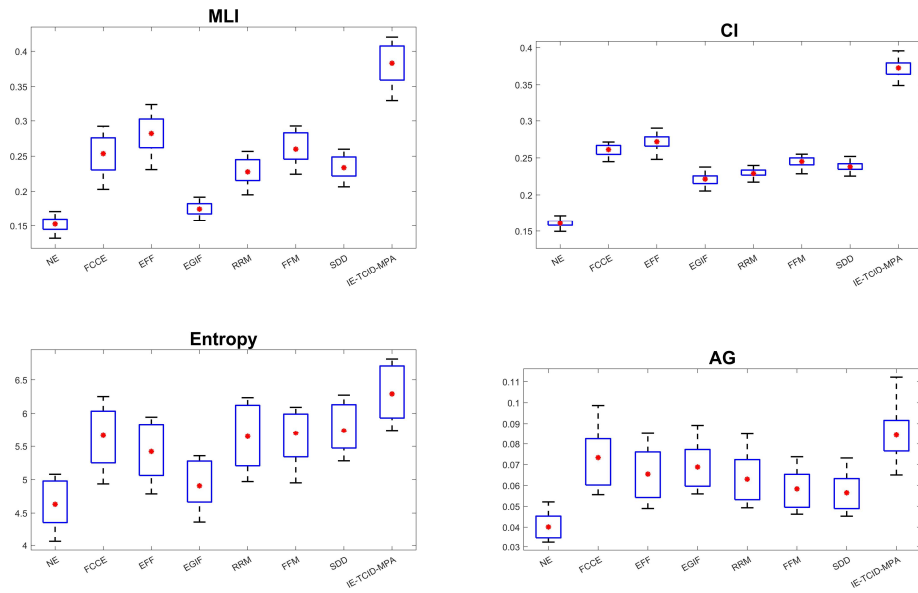
Thuật toán	MLI	CI	Entropy	AG
NE	0.3050	0.3216	4.6314	0.0802
FCCE	0.3250	0.3379	5.5607	0.0853
EFF	0.3606	0.3629	5.5651	0.0927
EGIF	0.3199	0.3761	4.8231	0.1151
RRM	0.3591	0.3594	5.6699	0.1036
FFM	0.3450	0.3401	5.9318	0.0883
SDD	0.3669	0.3666	5.8616	0.0925
IE_TCID_MPA	0.3975	0.4036	6.3156	0.1474

Bảng 2.9. Các chỉ số đánh giá từ các thuật toán nâng cao chất lượng ảnh trên tập $D1$

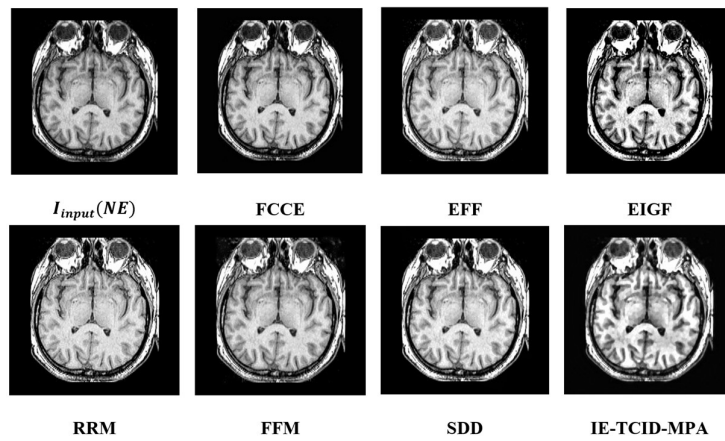
Thuật toán	MLI	CI	Entropy	AG
NE	0.1525	0.1608	4.6314	0.0401
FCCE	0.2536	0.2612	5.6641	0.0734
EFF	0.2821	0.2718	5.4223	0.0655
EGIF	0.1750	0.2215	4.9067	0.0688
RRM	0.2278	0.2290	5.6485	0.0630
FFM	0.2599	0.2451	5.7019	0.0584
SDD	0.2336	0.2379	5.7394	0.0565
IE_TCID_MPA	0.3835	0.3731	6.2892	0.0846



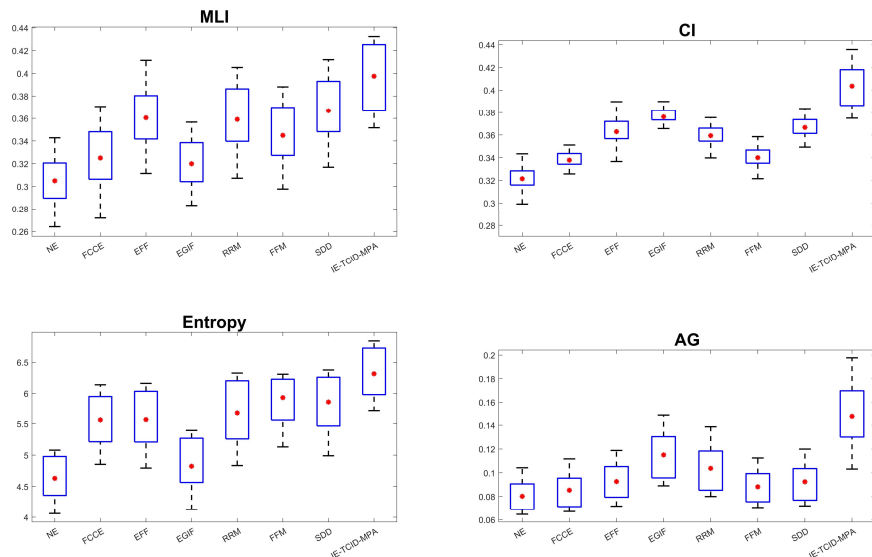
Hình 2.10. Hình ảnh đầu ra từ các thuật toán tăng cường ảnh trên tập $D1$.



Hình 2.11. Bốn chỉ số đánh giá chất lượng ảnh trên tập dữ liệu D1.



Hình 2.12. Hình ảnh kết quả từ các thuật toán nâng cao chất lượng trên tập dữ liệu D0.



Hình 2.13. Bốn chỉ số đánh giá chất lượng ảnh trên tập dữ liệu D0.

Các tham số tối ưu (β_1 , β_2 , β_3 và β_4) thu được từ thuật toán nâng cao chất lượng ảnh đề xuất được trình bày trong Bảng 2.10.

Bảng 2.10. Các tham số tối ưu thu được trên hai tập dữ liệu D0 và D1.

Tập dữ liệu	β_1	β_2	β_3	β_4
D0	1.0011	1.1998	1.1999	1.1285
D1	1.2907	0.7967	0.9545	0.9920

Đối với các hình ảnh có chất lượng thấp trong tập dữ liệu (D1), có thể dễ dàng nhận thấy rằng thuật toán đề xuất đã cải thiện chất lượng hình ảnh một cách đáng kể. Tham số tối ưu β_1 là 1.209 cho phép thành phần I_{CLAHE} được cải thiện đáng kể về cường độ sáng bởi vì sau khi cân bằng bởi phương pháp CLAHE vẫn chưa đủ cường độ sáng do ảnh bị tối và tương phản thấp. Các tham số tối ưu β_2 , β_3 và β_4 thu được đều thấp hơn 1, nên chỉ cho phép bổ sung thêm một lượng nhỏ các đặc điểm cấu trúc và kết cấu vào hình ảnh được tăng cường. Đối với ảnh trong tập dữ liệu (D0) mà ảnh có cường độ sáng không bị yếu. Tham số tối ưu β_1 là 1.0011 cho thấy rằng thành phần I_{CLAHE} đã đảm bảo đủ cường độ sáng cho hình ảnh mà không cần bổ sung thêm. Các tham số tối ưu β_2 , β_3 và β_4 , đều lớn hơn 1, điều này cho thấy rằng các thành phần I_{STS} , I_{T_E} và I_N chỉ cần thêm một lượng lớn thông tin vào hình ảnh. Do đó, các tham số thích nghi cho phép điều chỉnh lượng thông tin cần thiết cho bốn thành phần I_{CLAHE} , I_{STS} , I_{T_E} và I_N để tạo ra hình ảnh có chất lượng cao. Những kết quả này cho thấy rằng thuật toán đề xuất có hiệu quả trong việc cải thiện chất lượng hình ảnh.

Vì thuật toán nâng cao hình ảnh đề xuất được dựa trên thuật toán tối ưu hóa MPA nên nó tiêu tốn một lượng lớn thời gian chạy. Từ Bảng 2.11, có thể dễ dàng thấy rằng thời gian chạy trung bình của thuật toán đề xuất là lớn nhất so với thời gian chạy của các thuật toán nâng cao chất lượng ảnh khác.

Bảng 2.11. Thời gian chạy trung bình của các thuật toán trên tập D1.

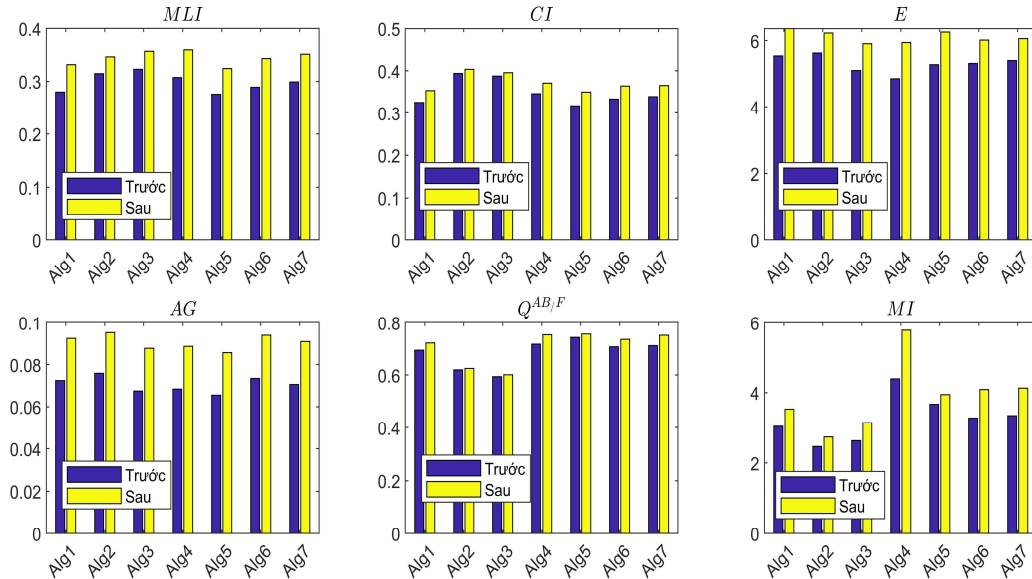
Thuật toán	Thời gian chạy trung bình (s)
FCCE	0.047856
EFF	0.298958
EGIF	0.031338
RRM	2.807581
FFM	0.797329
SDD	0.448706
IE_TCID_MPA	24.747323

Với số cá thể trong đàn $n = 50$ và với số lượng vòng lặp là 50, thuật toán đề xuất mất khoảng 25 giây để hoàn thành. Đây là điểm hạn chế của thuật toán đề xuất.

Kết quả thực nghiệm 3: Kết quả của thực nghiệm thứ ba bao gồm 6 chỉ số đánh giá và bảy thuật toán tổng hợp hình ảnh được trình bày trong các Bảng 2.12, Bảng 2.13 và Bảng 2.14. Các số liệu trong các bảng Bảng 2.12, Bảng 2.13 và Bảng 2.14 được biểu diễn bởi biểu đồ trong các Hình 2.14, Hình 2.15 và Hình 2.16 để thuận tiện cho việc theo dõi trực quan.

Bảng 2.12. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (T).

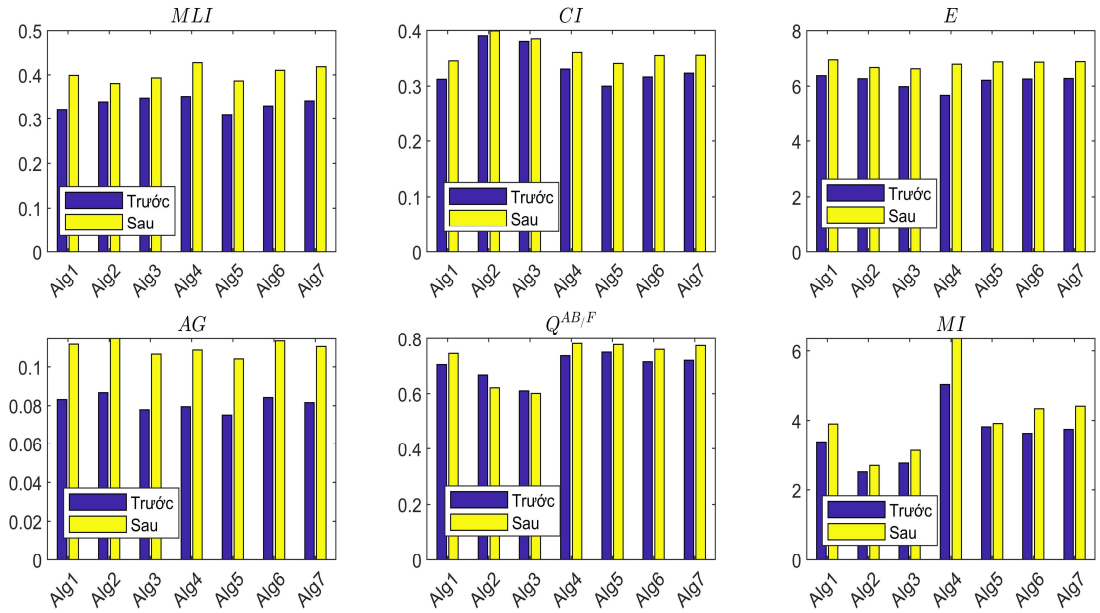
Thuật toán	Kiểu	MLI	CI	E	AG	$Q^{AB/F}$	MI
Alg1	Trước	0.2795	0.3249	5.5404	0.0724	0.6942	3.0354
	Sau	0.3318	0.3529	6.3630	0.0925	0.7222	3.5302
Alg2	Trước	0.3148	0.3937	5.6302	0.0759	0.6199	2.4631
	Sau	0.3462	0.4035	6.2270	0.0952	0.6255	2.7315
Alg3	Trước	0.3231	0.3873	5.1023	0.0674	0.5942	2.6319
	Sau	0.3571	0.3957	5.9033	0.0877	0.6015	3.1434
Alg4	Trước	0.3074	0.3455	4.8495	0.0684	0.7178	4.3910
	Sau	0.3599	0.3706	5.9439	0.0887	0.7537	5.7807
Alg5	Trước	0.2756	0.3165	5.2769	0.0655	0.7434	3.6663
	Sau	0.3243	0.3496	6.2598	0.0857	0.7564	3.9412
Alg6	Trước	0.2888	0.3329	5.3130	0.0735	0.7070	3.2731
	Sau	0.3432	0.3637	6.0178	0.0939	0.7357	4.0908
Alg7	Trước	0.2991	0.3388	5.3990	0.0706	0.7120	3.3447
	Sau	0.3515	0.3653	6.0600	0.0910	0.7512	4.1313



Hình 2.14. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-T).

Bảng 2.13. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (S).

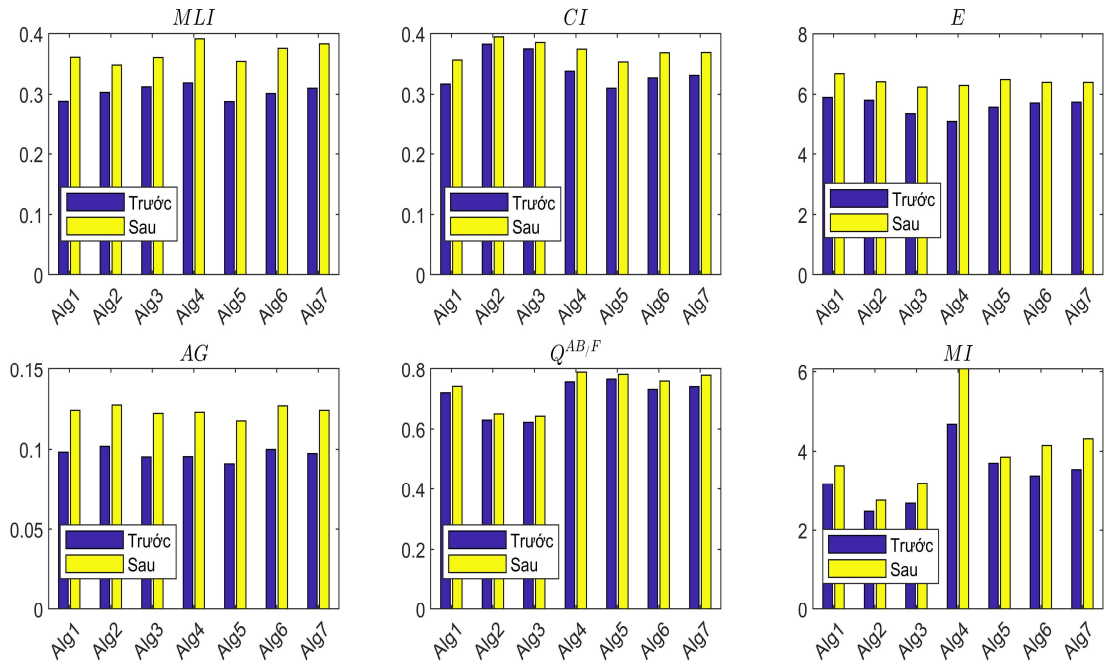
Thuật toán	Kiểu	MLI	CI	E	AG	$Q^{AB/F}$	MI
Alg1	Trước	0.3219	0.3124	6.3709	0.0831	0.7059	3.3813
	Sau	0.3990	0.3453	6.9472	0.1118	0.7465	3.9006
Alg2	Trước	0.3392	0.3905	6.2654	0.0867	0.6678	2.5171
	Sau	0.3808	0.3992	6.6739	0.1148	0.6216	2.7008
Alg3	Trước	0.3477	0.3806	5.9830	0.0778	0.6108	2.7661
	Sau	0.3930	0.3849	6.6219	0.1067	0.6013	3.1306
Alg4	Trước	0.3512	0.3307	5.6692	0.0794	0.7381	5.0347
	Sau	0.4280	0.3606	6.7940	0.1088	0.7818	6.3580
Alg5	Trước	0.3107	0.3001	6.2098	0.0751	0.7507	3.8228
	Sau	0.3864	0.3407	6.8762	0.1041	0.7782	3.9165
Alg6	Trước	0.3297	0.3164	6.2581	0.0842	0.7157	3.6292
	Sau	0.4103	0.3552	6.8673	0.1136	0.7611	4.3368
Alg7	Trước	0.3413	0.3235	6.2772	0.0815	0.7211	3.7479
	Sau	0.4185	0.3554	6.8787	0.1107	0.7742	4.4097



Hình 2.15. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-S).

Bảng 2.14. Các chỉ số đánh giá trước và sau tăng cường ảnh trên tập D3 (C).

Thuật toán	Kiểu	MLI	CI	E	AG	$Q^{AB/F}$	MI
Alg1	Trước	0.2886	0.3171	5.8940	0.0982	0.7203	3.1741
	Sau	0.3611	0.3566	6.6787	0.1242	0.7417	3.6318
Alg2	Trước	0.3033	0.3826	5.8035	0.1018	0.6299	2.4732
	Sau	0.3484	0.3947	6.4161	0.1274	0.6505	2.7522
Alg3	Trước	0.3123	0.3747	5.3588	0.0953	0.6230	2.6773
	Sau	0.3608	0.3856	6.2380	0.1222	0.6432	3.1838
Alg4	Trước	0.3186	0.3381	5.1012	0.0955	0.7562	4.6741
	Sau	0.3917	0.3743	6.2949	0.1230	0.7886	6.0716
Alg5	Trước	0.2878	0.3102	5.5751	0.0910	0.7655	3.6959
	Sau	0.3541	0.3533	6.4829	0.1176	0.7813	3.8505
Alg6	Trước	0.3013	0.3269	5.7136	0.1000	0.7309	3.3756
	Sau	0.3758	0.3686	6.3904	0.1270	0.7594	4.1467
Alg7	Trước	0.3101	0.3312	5.7388	0.0973	0.7405	3.5328
	Sau	0.3833	0.3691	6.3973	0.1242	0.7781	4.3110

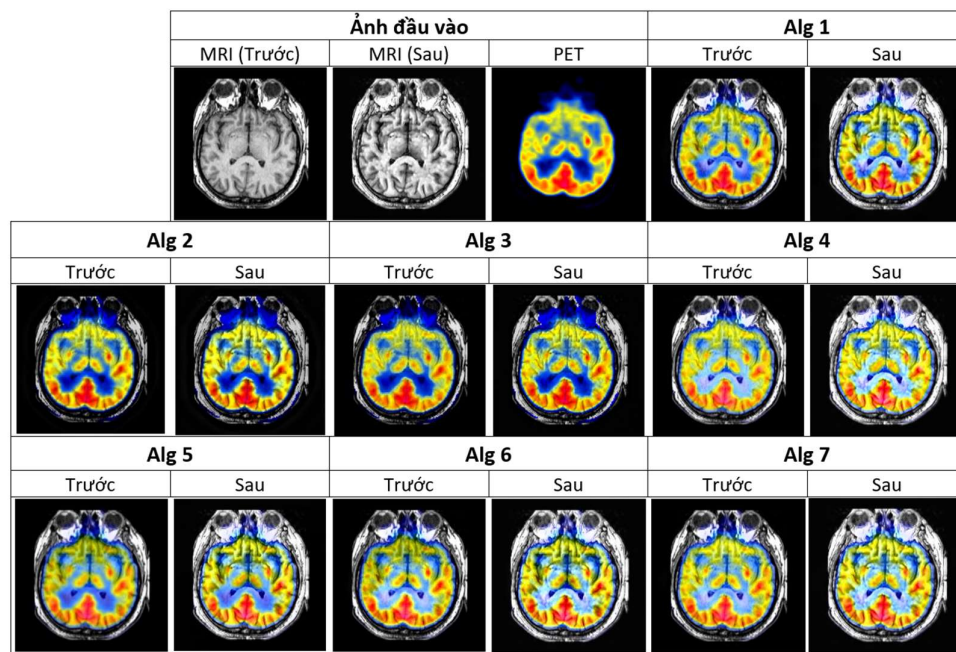


Hình 2.16. Sáu chỉ số đánh giá trên tập dữ liệu D3 (Set-C).

Nhìn chung, chất lượng của các hình ảnh tổng hợp thu được từ các thuật toán đã được cải thiện đáng kể khi áp dụng thuật toán nâng cao hình ảnh đề xuất. Các chỉ số đánh giá đều tăng đáng kể sau khi áp dụng thuật toán tăng cường chất lượng cho hình ảnh đầu vào. Thay đổi đáng kể nhất phải kể đến là độ chi tiết của hình ảnh. Bằng trực quan, có thể dễ dàng nhận thấy rằng hình ảnh tổng hợp đầu ra có độ sắc nét cao

khi sử dụng thuật toán nâng cao chất lượng ảnh đề xuất. Ví dụ, chỉ số AG của thuật toán Alg1 tăng từ 0.0724 lên 0.0925 trên tập dữ liệu (T). Kết quả thực nghiệm đã chỉ ra rằng thuật toán nâng cao hình ảnh đề xuất có hiệu quả đáng kể trong việc cải thiện hiệu quả của các thuật toán tổng hợp hình ảnh.

Hình ảnh đầu ra thu được bằng thuật toán tổng hợp hình ảnh trước và sau khi áp dụng thuật toán tăng cường ảnh đề xuất được hiển thị trong Hình 2.17. Bằng trực quan, có thể nhận thấy rằng, các hình ảnh cộng hưởng từ được tăng cường đã cho thấy sự cải thiện đáng kể trong hình ảnh tổng hợp. Các hình ảnh tổng hợp không những có cường độ sáng, độ tương phản tốt hơn mà cả các chi tiết của ảnh cũng được cải thiện đáng kể.



Hình 2.17. Hình ảnh tổng hợp trước và sau áp dụng thuật toán đề xuất

2.4. Kết luận chương 2

Trong chương này, một phương pháp nâng cao chất lượng ảnh IE_TCID_MPA đã được đề xuất và phương pháp này đã được ứng dụng để cải thiện hiệu quả cho một số thuật toán tổng hợp hình ảnh được đề xuất gần đây. Phương pháp nâng cao chất lượng ảnh đề xuất bao gồm 2 thuật toán chính.

- Thứ nhất là một thuật toán phân rã hình ảnh thành 3 thành phần. Thuật toán phân rã hình ảnh này được xây dựng dựa trên bộ lọc Gauss và bộ lọc ADF.
- Thứ hai là một thuật toán nâng cao chất lượng ảnh dựa trên thuật toán tối ưu hoá MPA.

Phương pháp nâng cao chất lượng ảnh đề xuất đã được ứng dụng để kiểm tra sự hiệu quả so với các thuật toán nâng cao chất lượng ảnh khác và nó được ứng dụng

trong việc cải thiện hiệu quả của một số thuật toán tổng hợp hình ảnh đã có. Hai tập dữ liệu D0 và D1 gồm 90 hình ảnh mỗi tập được sử dụng trong phần thực nghiệm. Bốn chỉ số đánh giá MLI, CI, Entropy và AG cùng với 6 thuật toán nâng cao chất lượng ảnh khác được lựa chọn để đánh giá. Cụ thể, trên tập dữ liệu D0 có cường độ sáng và độ tương phản không thấp, bốn chỉ số đánh giá MLI, CI, Entropy và AG đạt lần lượt 0.3975, 0.4036, 6.3156 và 0.1474. Những giá trị này đạt được là cao nhất so với các phương pháp tăng cường ảnh đã được lựa chọn để so sánh. So với 4 chỉ số của hình ảnh đầu vào, 4 chỉ số đánh giá MLI, CI, Entropy và AG của phương pháp IE_TCID_MPA tăng tương ứng khoảng 30%, 25%, 36% và 84%. Trên tập dữ liệu D1 có cường độ sáng và độ tương phản thấp, bốn chỉ số đánh giá MLI, CI, Entropy và AG đạt lần lượt 0.3835, 0.3731, 6.2892 và 0.0846. Những giá trị này là cao nhất so với các chỉ số của các phương pháp tăng cường ảnh còn lại và cao hơn đáng kể so với các chỉ số của ảnh đầu vào. So với 4 chỉ số của hình ảnh đầu vào, 4 chỉ số đánh giá MLI, CI, Entropy và AG của phương pháp IE_TCID_MPA tăng tương ứng khoảng 151%, 132%, 36% và 111%. Các thực nghiệm đã cho thấy thuật toán nâng cao chất lượng ảnh đề xuất cho phép cải thiện đáng kể chất lượng của hình ảnh tối và có độ tương phản thấp.

Trong thực nghiệm kiểm tra sự hiệu quả của phương pháp tăng cường chất lượng ảnh IE_TCID_MPA, tập dữ liệu D3 gồm 90 cặp hình ảnh y học (30 cặp theo mỗi trục T, S và C) được sử dụng. Sáu chỉ số đánh giá gồm MLI, CI, E, AG, $Q^{AB/F}$ và MI được sử dụng cùng với 7 thuật toán tổng hợp hình ảnh khác được sử dụng để đánh giá. Ví dụ, trên tập dữ liệu D3(T), sáu chỉ số đánh giá thu được trên thuật toán tổng hợp Alg1 sau khi áp dụng phương pháp tăng cường ảnh đề xuất đạt lần lượt là 0.3318, 0.3529, 6.3630, 0.0925, 0.7222 và 3.5302. So với lúc trước khi tăng cường ảnh, các chỉ số này đã tăng lần lượt là khoảng 19%, 9%, 15%, 28%, 4% và 16%. Các kết quả tương tự cũng thu được trên các thuật toán tổng hợp khác cũng như trên các tập dữ liệu còn lại D3(S) và D3(C). Như vậy, khi hình ảnh đầu vào được tiền xử lý bằng phương pháp nâng cao chất lượng ảnh IE_TCID_MPA, các kết quả thực nghiệm cũng cho thấy có sự cải thiện đáng kể về mặt chất lượng ở hình ảnh tổng hợp so với trước khi ảnh được tăng cường chất lượng. Vì vậy, thuật toán nâng cao chất lượng ảnh đề xuất có vai trò quan trọng trong việc nâng cao hiệu quả của các thuật toán tổng hợp hình ảnh.

Kết quả của thuật toán nâng cao chất lượng hình ảnh đề xuất đã được công bố tại công trình [CT1]. Ứng dụng của thuật toán này trong việc nâng cao hiệu quả của việc tổng hợp hình ảnh được công bố tại công trình [CT2] trong phần “Danh mục các công trình của tác giả”.

CHƯƠNG 3. TỔNG HỢP ẢNH DỰA TRÊN THUẬT TOÁN TỔNG HỢP THÍCH NGHI KẾT HỢP VỚI CÁC BIẾN THỂ CỦA HÀM NĂNG LƯỢNG CỤC BỘ

Trong chương 2, chúng ta đã thấy rằng chất lượng của hình ảnh đầu vào thấp sẽ ảnh hưởng tới hiệu quả của việc tổng hợp hình ảnh. Bằng cách tiếp cận nâng cao chất lượng hình ảnh đầu vào đã cho thấy sự cải thiện trong hiệu quả trong việc tổng hợp hình ảnh mà không cần phải cải thiện trong chính các thuật toán tổng hợp hình ảnh. Tuy nhiên, một số thuật toán tổng hợp cho các thành phần cơ sở và chi tiết trong ảnh vẫn còn có những hạn chế nhất định như làm suy giảm cường độ sáng, độ tương phản và làm mất các thông tin chi tiết. Do đó, để cải thiện hơn nữa hiệu quả của việc tổng hợp hình ảnh thì cách tiếp cận cải thiện các thuật toán tổng hợp cho các thành phần cơ sở và các thành phần chi tiết vẫn là vấn đề cấp thiết đặt ra. Vì vậy, nội dung trong Chương 3 này sẽ hướng đến đề xuất thuật toán tổng hợp hiệu quả cho các thành phần cơ sở và chi tiết trong ảnh. Kết quả của Chương 3 đã được công bố tại công trình [CT3, CT4, CT5, CT6] trong phần “Danh mục các công trình của tác giả”.

3.1. Ý tưởng

Như đã đề cập trong Chương 1 (mục 1.3), các thuật toán tổng hợp cho các thành phần cơ sở và chi tiết đều vẫn còn có những hạn chế nhất định.

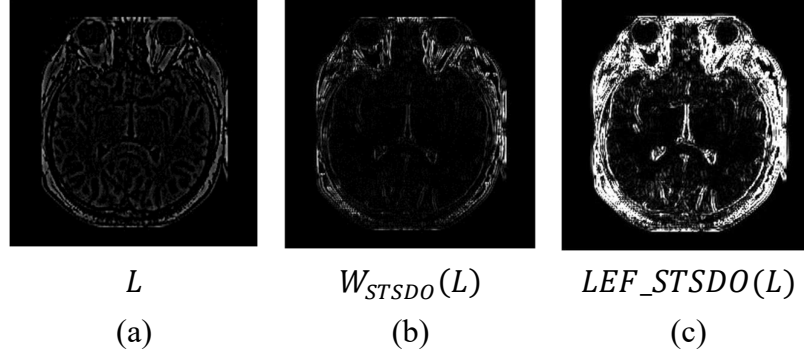
- Đối với hạn chế trong tổng hợp các thành phần cơ sở là sự suy giảm cường độ sáng và độ tương phản của hình ảnh được tổng hợp. Để giải quyết hạn chế này, ý tưởng của thuật toán tổng hợp đề xuất cho các thành phần cơ sở là sử dụng các tham số thích nghi cho các thành phần cơ sở của hình ảnh MRI và PET. Để tìm ra các tham số thích nghi phù hợp, thuật toán tối ưu hoá MPA sẽ được sử dụng. Trong đó, một hàm mục tiêu được đề xuất có vai trò kiểm soát chất lượng cho hình ảnh được tổng hợp đầu ra.
- Đối với hạn chế trong tổng hợp thành phần chi tiết là hình ảnh tổng hợp thường bị mất các thông tin chi tiết có trong hình ảnh đầu vào. Để giải quyết cho hạn chế này, ý tưởng đề xuất cho thuật toán tổng hợp cho thành phần chi tiết là kết hợp hàm năng lượng cục bộ cùng với các biến thể của nó.

3.2. Đề xuất thuật toán tổng hợp cho thành phần chi tiết (FM_CVLEF)

3.2.1. Cải thiện đặc trưng cấu trúc Tensor bởi hàm năng lượng cục bộ

Toán tử phát hiện đặc trưng cấu trúc Tensor (STSDO) là một công cụ hiệu quả để thu được đặc trưng của hình ảnh. Hình 3.1 (b) minh họa cho đặc trưng cấu trúc Tensor thu được từ một thành phần chi tiết (Hình 3.1 (a)). Từ Hình 3.1 (b) nhận thấy

rằng một số cấu trúc được toán tử STSDO phát hiện thành công. Tuy nhiên, một số các đặc trưng yếu và nhỏ thì toán tử STSDO không thể phát hiện được.



Hình 3.1. Minh hoạ đặc trưng cấu trúc Tensor và kết hợp với hàm năng lượng cục bộ

Vì vậy, để cải thiện các đặc trưng thu được từ toán tử STSDO, một công thức kết hợp đặc trưng thu được từ toán tử STSDO với hàm năng lượng cục bộ (gọi là LEF_STSDO) được đề xuất và được tính theo công thức (3.1).

$$LEF_STSDO(L) = W_{STSDO}(L) \odot LEF(L) \quad (3.1)$$

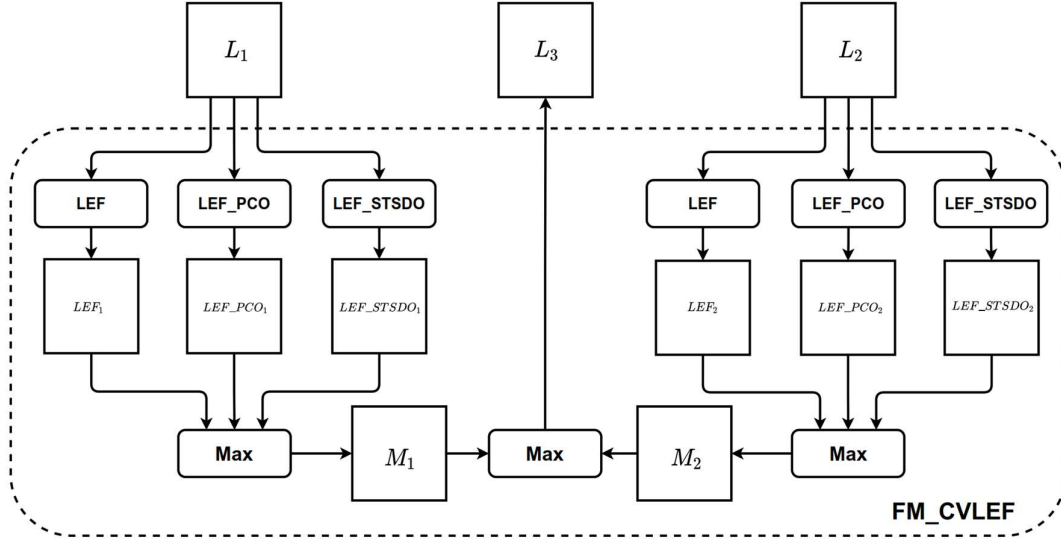
Trong đó

- L là một thành phần chi tiết của một hình ảnh đầu vào.
- $W_{STSDO}(L)$ là ma trận đặc trưng thu được sử dụng toán tử phát hiện đặc trưng của cấu trúc Tensor cho L .
- $LEF(L)$ là hàm năng lượng cục bộ của L .
- $\odot$ là toán tử nhân Hadamard.

Hình 3.1 (c) minh hoạ cho hình ảnh đặc trưng thu được sau khi được kết hợp với cấu trúc Tensor. Dễ dàng nhận thấy rằng, các đặc trưng yếu và nhỏ đã được phát hiện trong ảnh. Cách cải thiện đặc trưng của cấu trúc Tensor bằng cách kết hợp nó với hàm năng lượng cục bộ đã được công bố tại công trình [CT3] trong phần “Danh mục các công trình của tác giả”.

3.2.2. Thuật toán FM_CVLEF

Trong mục này, một thuật toán tổng hợp hiệu quả cho thành phần chi tiết được giới thiệu. Thuật toán này được xây dựng dựa trên kết hợp các biến thể của hàm năng lượng cục bộ (được gọi là FM_CVLEF). Biến thể đầu tiên của hàm năng lượng cục bộ là kết hợp hàm năng lượng cục bộ với toán tử la bàn Prewitt mà nó được đề xuất bởi Dinh [93] vào năm 2021. Biến thể thứ hai là kết hợp hàm năng lượng cục bộ với đặc trưng của cấu trúc Tensor (như đã trình bày trong mục 3.2.1). Hình 3.2 minh hoạ chi tiết các bước của thuật toán FM_CVLEF.



Hình 3.2. Sơ đồ thuật toán FM_CVLEF

Thuật toán 3.1: FM_CVLEF

Đầu vào: Hai lớp chi tiết L_1 và L_2

Đầu ra: Lớp tổng hợp chi tiết L_3

Bước 1: Tính hàm năng lượng cục bộ cho các lớp L_1 và L_2 theo các công thức (3.2) và (3.3).

$$LEF_1(i, j) = \sum_{u=0}^{k-1} \sum_{v=0}^{k-1} W_{LE}(u, v) L_1^2(i + u, j + v) \quad (3.2)$$

$$LEF_2(i, j) = \sum_{u=0}^{k-1} \sum_{v=0}^{k-1} W_{LE}(u, v) L_2^2(i + u, j + v) \quad (3.3)$$

Trong đó, W_{LE} là một ma trận có toàn số 1 và có kích thước $k \times k$.

Bước 2: Tính hàm năng lượng cục bộ sử dụng toán tử la bàn Prewitt [93] cho các lớp L_1 và L_2 theo các công thức (3.4) và (3.5).

$$LEF_PCO_1^k(i, j) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} W_{PCO}^k(u, v) L_1^2(i + u, j + v) \quad (3.4)$$

$$LEF_PCO_2^k(i, j) = \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} W_{PCO}^k(u, v) L_2^2(i + u, j + v) \quad (3.5)$$

Trong đó, W_{PCO}^k ($k = \overline{1,8}$) là một ma trận mặt nạ thứ k của toán tử la bàn Prewitt có kích thước $m \times m$ với $m = 3$.

Bước 3: Tính hàm năng lượng cục bộ kết hợp với đặc trưng của cấu trúc Tensor cho các lớp L_1 và L_2 bằng cách sử dụng công thức (3.1) như được giới thiệu trong mục 3.3.1. Thu được LEF_STSDO_1 và LEF_STSDO_2 theo các công thức (3.6) và (3.7).

$$LEF_STSDO_1 = W_{STSDO}(L_1) \odot LEF(L_1) \quad (3.6)$$

$$LEF_STSDO_2 = W_{STSDO}(L_2) \odot LEF(L_2) \quad (3.7)$$

Trong đó, $W_{STSDO}(L_1)$ và $W_{STSDO}(L_2)$ lần lượt là đặc trưng của cấu trúc Tensor thu được từ hai lớp L_1 và L_2 . $LEF(L_1)$ và $LEF(L_2)$ lần lượt là các hàm năng lượng cục bộ thu được từ hai lớp L_1 và L_2 . Kí hiệu $\odot$ là toán tử nhân Hadamard.

Bước 4: Tính giá trị lớn nhất của hàm năng lượng cục bộ và các biến thể của nó cho hai lớp L_1 và L_2 theo các công thức (3.8) và (3.9).

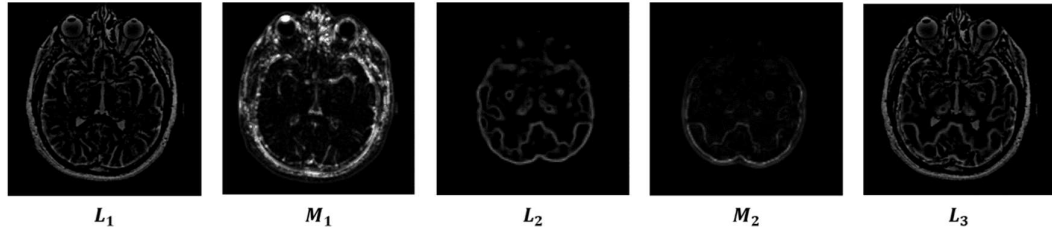
$$M_1 = \text{Max}(LEF_1, LEF_PCO_1^k, LEF_STSDO_1) \quad (3.8)$$

$$M_2 = \text{Max}(LEF_2, LEF_PCO_2^k, LEF_STSDO_2) \quad (3.9)$$

Bước 5: Thuật toán tổng hợp FM_CVLEF được xác định theo công thức (3.10). Trong đó: (i, j) là tọa độ của các điểm ảnh trong ma trận ảnh.

$$L_3(i, j) = \begin{cases} L_1(i, j) & \text{nếu } M_1(i, j) \geq M_2(i, j) \\ L_2(i, j) & \text{nếu } M_1(i, j) \leq M_2(i, j) \end{cases} \quad (3.10)$$

Hình 3.3 minh họa cho việc tổng hợp hai thành phần L_1 và L_2 bởi thuật toán FM_CVLEF.



Hình 3.3. Minh họa hình ảnh thu được từ thuật toán FM_CVLEF

3.3. Đề xuất thuật toán tổng hợp cho thành phần cơ sở (AFM_MPA)

3.3.1. Thiết kế hàm mục tiêu

Hàm mục tiêu đóng một vai trò quan trọng trong việc kiểm soát chất lượng của hình ảnh tổng hợp đầu ra. Dựa trên hàm chỉ số tương phản Michelson theo công thức (2.3), một hàm mục tiêu F được đề xuất như công thức (3.11).

$$F = \frac{\mu_T}{R \times \sigma_T^2} (H_T - H_M)(H_T - H_Y) \quad (3.11)$$

Trong đó:

- μ_T và σ_T^2 là cường độ sáng trung bình và phương sai của hình ảnh xám tổng hợp tạm thời (L_T^F) trong mỗi vòng lặp của thuật toán MPA.

- H_M , H_Y và H_T lần lượt là giá trị Entropy của các hình ảnh I_M (hình ảnh I_{MRI} được tăng cường), hình ảnh Y (hình ảnh xám của I_{PET}) và hình ảnh tổng hợp tạm thời L_T^F .
- Giá trị R là được tính toán thông qua công thức (3.12).

$$R = \rho_3 R(I_M, L_T^F) + (1 - \rho_3) R(Y, L_T^F) \quad (3.12)$$

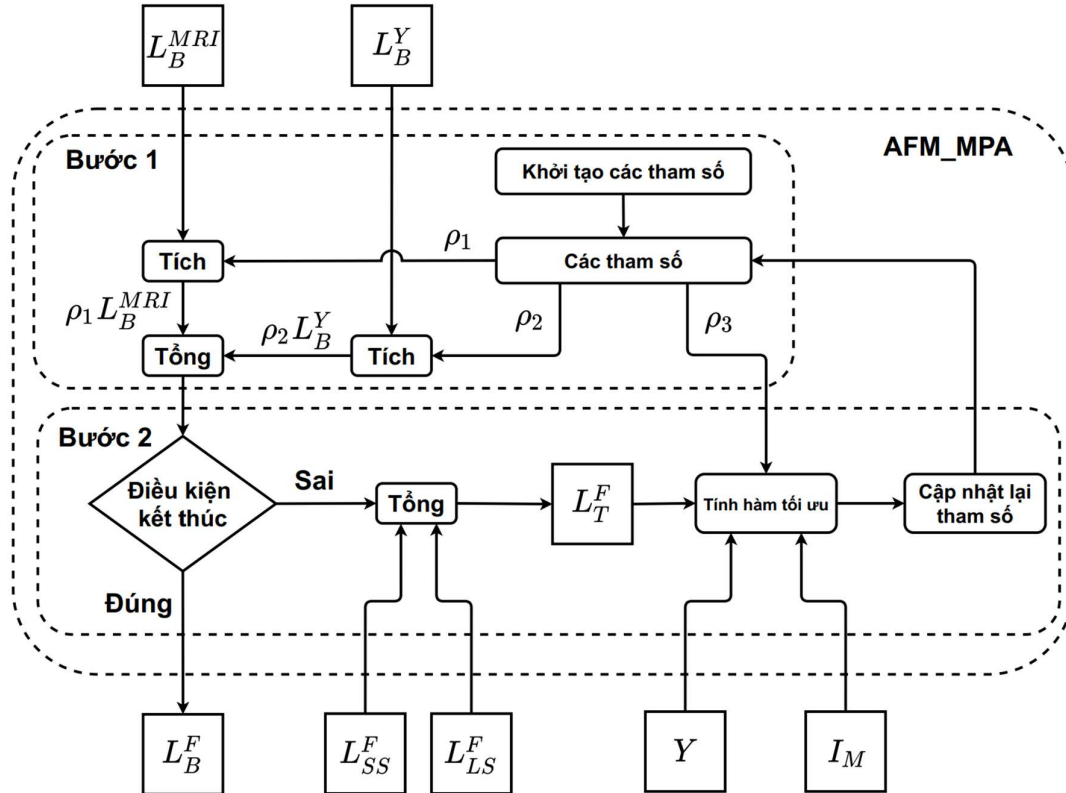
Với $R(I_M, L_T^F)$ và $R(Y, L_T^F)$ được tính theo các công thức (3.13) và (3.14).

$$R(I_M, L_T^F) = \frac{1}{S \times T} \sum_{s=1}^S \sum_{t=1}^T (L_T^F(s, t) - I_M(s, t))^2 \quad (3.13)$$

$$R(Y, L_T^F) = \frac{1}{S \times T} \sum_{s=1}^S \sum_{t=1}^T (L_T^F(s, t) - Y(s, t))^2 \quad (3.14)$$

3.3.2. Thuật toán AFM_MPA

Các thành phần cơ sở thường được tổng hợp với nhau theo thuật toán trung bình vì độ phức tạp tính toán thấp. Tuy nhiên, thuật toán tổng hợp trung bình gây ra sự suy giảm về cường độ sáng cho hình ảnh tổng hợp. Vì vậy, trong mục này, một thuật toán tổng hợp thích nghi dựa trên thuật toán tối ưu hoá MPA cho các thành phần cơ sở (được gọi là AFM_MPA) được đề xuất. Các bước của thuật toán AFM_MPA được minh hoạ trong Hình 3.4.



Hình 3.4. Sơ đồ thuật toán AFM_MPA

Các tham số thích nghi trong thuật toán tổng hợp được tìm được ra dựa trên việc sử dụng thuật toán tối ưu hoá MPA để tìm cực đại cho hàm mục tiêu F như công thức (3.11). Thuật toán AFM_MPA đảm bảo cho hình ảnh tổng hợp không bị suy giảm chất lượng về khía cạnh cường độ sáng và độ tương phản. Chi tiết các bước của thuật toán AFM_MPA được trình bày chi tiết trong Thuật toán 3.2.

Thuật toán 3.2: AFM_MPA

Đầu vào:

- Hai thành phần cơ sở: (L_B^{MRI}, L_B^Y) .
- Hai thành phần tổng hợp của các lớp chi tiết có cấu trúc nhỏ và lớn: (L_{SS}^F, L_{SS}^F) .
- Hai hình ảnh xám của hình ảnh MRI và PET: (I_M, Y) .
- Các tham số cần thiết của thuật toán MPA.

Đầu ra: Thành phần cơ sở được tổng hợp L_B^F .

Bước 1: Tính thành phần tổng hợp cơ sở dựa vào các tham số khởi tạo

Bước 1.1: Khởi tạo ngẫu nhiên các tham số cần tìm $\rho_1 \in [0.8, 1], \rho_2 \in [0, 0.2]$ và $\rho_3 \in [0.9, 1]$.

Bước 1.2: Tổng hợp hai thành phần cơ sở (L_B^{MRI}, L_B^Y) với các tham số (ρ_1, ρ_2) theo công thức (3.15).

$$L_B^F = \rho_1 L_B^{MRI} + \rho_2 L_B^Y \quad (3.15)$$

Bước 2: Thực hiện vòng lặp của thuật toán MPA để cập nhật lại các tham số
Nếu điều kiện kết thúc chưa thỏa mãn:

Bước 2.1: Tính hình ảnh tổng hợp tạm thời L_T^F theo công thức (3.16).

$$L_T^F = L_B^F + L_{LS}^F + L_{SS}^F \quad (3.16)$$

Bước 2.3: Tính giá trị hàm tối ưu F theo công thức (3.11).

Bước 2.4: Cập nhật lại các tham số ρ_1, ρ_2 và ρ_3 theo thuật toán MPA.

Nếu điều kiện kết thúc thỏa mãn:

Bước 2.5: Trả về thành phần tổng hợp cơ sở L_B^F ứng với các tham số tối ưu tìm được.

3.4. Đề xuất phương pháp tổng hợp hình ảnh
3.4.1. Thuật toán phân rã hình ảnh ba thành phần (TCID)

Quá trình tổng hợp ảnh bắt đầu bằng việc phân rã các hình ảnh đầu vào. Mục đích của việc phân rã hình ảnh là để tách hình ảnh đầu vào thành các lớp có thông tin bổ sung lẫn nhau. Việc phân tách này tạo điều kiện thuận lợi cho việc tổng hợp ở giai đoạn sau đó được hiệu quả hơn. Thông thường, các thuật toán phân rã hình ảnh tạo ra một lớp cơ sở và một hoặc nhiều lớp chi tiết. Trong các nghiên cứu trước đây,

phương pháp phân tách hình ảnh thành hai lớp thường được sử dụng. Trong đó, lớp cơ sở thu được bằng cách sử dụng bộ lọc trung bình [44], bộ lọc Gauss [56] hoặc bộ lọc thông thấp [79]. Tuy nhiên, các bộ lọc này có thể làm mất thông tin chi tiết trong ảnh, dẫn đến các lớp chi tiết không đầy đủ. Để giải quyết những hạn chế này, thuật toán phân tách hình ảnh ba lớp (gọi là TCID - Three component image decomposition) dựa trên các bộ lọc RGF và WMCF được đề xuất. Hình 3.5 là sơ đồ của thuật toán phân rã ảnh. Các bước thuật toán của thuật toán phân rã ảnh TCID được trình bày trong Thuật toán 3.3.

Thuật toán 3.3: TCID

Đầu vào: Một hình ảnh xám I

Đầu ra: Ba lớp được phân rã, L_B , L_{SS} , và L_{LS}

Bước 1: Áp dụng bộ lọc RGF cho hình ảnh đầu vào I , thu được một lớp cơ sở L_B .

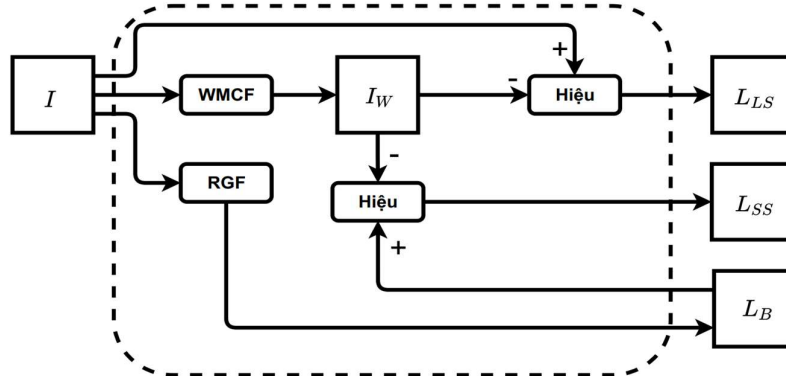
Bước 2: Áp dụng bộ lọc WMCF cho hình ảnh đầu vào I , thu được một hình ảnh I_w .

Bước 3: Lớp có tỉ lệ cấu trúc nhỏ L_{SS} thu được theo công thức (3.17).

$$L_{SS} = I - I_w \quad (3.17)$$

Bước 4: Lớp có tỉ lệ cấu trúc lớn L_{LS} thu được theo công thức (3.18).

$$L_{LS} = L_B - I_w \quad (3.18)$$

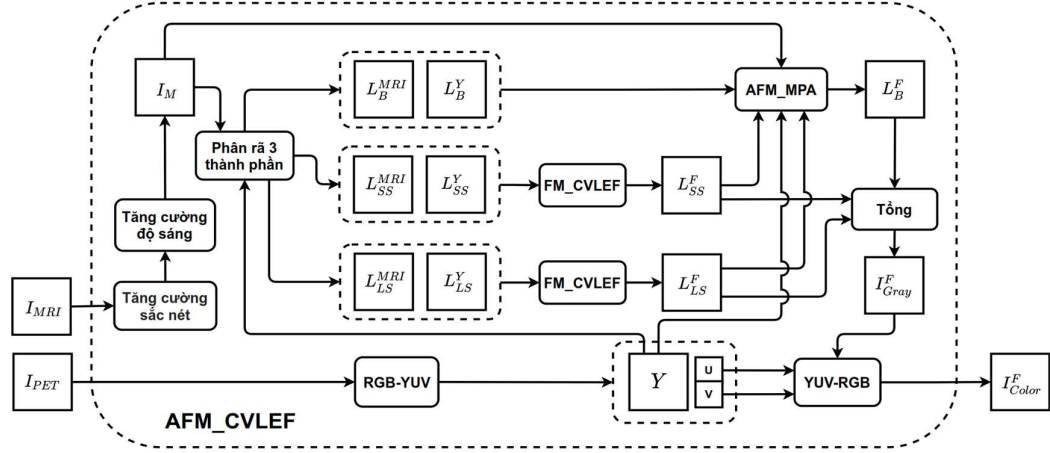


Hình 3.5. Sơ đồ thuật toán phân rã hình ảnh ba thành phần

3.4.2. Phương pháp tổng hợp ảnh AFM_CVLEF

Trong phần này, một phương pháp tổng hợp hình ảnh mới được giới thiệu (gọi là AFM_CVLEF). Phương pháp này dựa trên kết hợp thuật toán tổng hợp thích nghi (AFM_MPA) cho các thành phần cơ sở và thuật toán tổng hợp cho các thành phần chi tiết (FM_CVLEF) dựa trên hàm năng lượng cục bộ và các biến thể của nó. Chi tiết các bước của thuật toán tổng hợp đề xuất AFM_CVLEF được trình bày trong Hình 3.6. và Thuật toán 3.4. Nhìn chung, phương pháp tổng đề xuất AFM_CVLEF bao gồm 4 giải đoạn chính: tiền xử lý hình ảnh đầu vào, phân rã các hình ảnh đầu vào

bởi thuật toán phân rã ba thành phần (TCID), tổng hợp thành phần cơ sở bởi sử dụng thuật toán AFM_MPA và tổng hợp thành phần chi tiết bởi thuật toán FM_CVLEF.



Hình 3.6. Sơ đồ tổng hợp hình ảnh đề xuất AFM_CVLEF

Thuật toán 3.4: AFM_CVLEF

Đầu vào: Hai hình ảnh I_{MRI} và I_{PET} kích thước $S \times T$, các tham số cần thiết.

Đầu ra: Hình ảnh được tổng hợp I_{Color}^F .

Bước 1: Nâng cao chất lượng của hình ảnh đầu vào I_{MRI} bởi hai thuật toán cơ bản. Làm sắc nét ảnh bởi sử dụng toán tử Laplace và sau đó nâng cao cường độ sáng và độ tương phản bởi thuật toán BLLI [125]. Thu được hình ảnh tăng cường I_M .

Bước 2: Biến đổi hình ảnh I_{PET} từ không gian màu RGB sang không gian màu YUV. Thu được ba thành phần Y , U và V .

Bước 3: Phân rã hai hình ảnh xám I_M và Y bởi thuật toán TCID (như được trình bày trong mục 3.4.1). Thu được các lớp tương ứng cho hai hình ảnh xám I_M và Y : (L_B^{MRI}, L_B^Y) , (L_{SS}^{MRI}, L_{SS}^Y) và (L_{LS}^{MRI}, L_{LS}^Y) .

Bước 4: Các thành phần chi tiết (L_{SS}^{MRI}, L_{SS}^Y) và (L_{LS}^{MRI}, L_{LS}^Y) được tổng hợp với nhau bởi thuật toán FM_CVLEF theo các công thức (3.19) và (3.20).

$$L_{SS}^F = FR_CVLEF(L_{SS}^{MRI}, L_{SS}^Y) \quad (3.19)$$

$$L_{LS}^F = FR_CVLEF(L_{LS}^{MRI}, L_{LS}^Y) \quad (3.20)$$

Bước 5: Các lớp cơ sở (L_B^{MRI}, L_B^Y) được tổng hợp với nhau bởi thuật toán AFM_MPA theo công thức (3.21).

$$L_B^F = AFM_MPA(L_B^{MRI}, L_B^Y, L_{SS}^F, L_{LS}^F, I_M, Y) \quad (3.21)$$

Bước 6: Ảnh xám tổng hợp I_{Gray}^F được tính theo công thức (3.22).

$$I_{Gray}^F = L_B^F + L_{SS}^F + L_{LS}^F \quad (3.22)$$

Bước 7: Hình ảnh màu tổng hợp I_{Color}^F thu được bằng cách chuyển các thành phần I_{Gray}^F , U và V về không gian màu RGB.

3.5. Độ phức tạp tính toán của thuật toán AFM_CVLEF

Phương pháp tổng hợp AFM_CVLEF bao gồm bốn phần: Tiền xử lý ảnh, phân rã hình ảnh đầu vào, tổng hợp thành phần chi tiết và tổng hợp thành phần cơ sở.

Phần tiền xử lý ảnh: Với hình ảnh đầu vào có kích thước $N \times N$, thuật toán làm sắc nét ảnh bởi toán tử Laplace có độ phức tạp tính toán là $O(N^2 \cdot 3^2)$. Thuật toán tăng cường sáng BLLI có độ phức tạp tính toán là $O(N^2)$. Như vậy, phần tiền xử lý ảnh có độ phức tạp tính toán là xấp xỉ $O(N^2)$.

Phương pháp phân rã hình ảnh TCID bao gồm hai bộ lọc RGF và WMCF. Với kích thước của bộ lọc Gauss bên trong bộ lọc RGF là $k \times k$ ($k \geq 3$) và l ($l \geq 1$) là số vòng lặp của bộ lọc RGF. Khi đó bộ lọc RGF sẽ có độ phức tạp tính toán là $O(N^2 \cdot k^2 \cdot l)$. Bộ lọc WMCF sử dụng 8 mặt nạ có kích thước 3×3 và một phép toán lấy min của 8 ma trận kết quả sau khi nhân chập ảnh với 8 mặt nạ. Do đó, nó có độ phức tạp tính toán là $O(8 \cdot N^2 \cdot 3^2 + N^2) \approx O(N^2)$. Vì vậy, thuật toán phân rã hình ảnh TCID sẽ có độ phức tạp tính toán là $O(N^2 \cdot k^2 \cdot l)$.

Thuật toán tổng hợp thành phần chi tiết FM_CVLEF bao gồm 3 hàm: hàm năng lượng cục bộ LEF, hàm năng lượng cục bộ sử dụng toán tử la bàn Prewitt (LEF_PCO), và hàm năng lượng cục bộ sử dụng đặc trưng của cấu trúc Tensor (LEF_STSDO). Với $w \times w$ ($w \geq 3$) là kích thước của cửa sổ cục bộ, hàm năng lượng cục bộ LEF có độ phức tạp tính toán là $O(N^2 \cdot w^2)$. Hàm LEF_PCO sử dụng 8 mặt nạ la bàn kích thước 3×3 . Vì vậy, nó có độ phức tạp tính toán là $O(8 \cdot N^2 \cdot 3^2)$. Hàm LEF_STSDO bao gồm 1 hàm năng lượng cục bộ LEF có độ phức tạp tính toán là $O(N^2 \cdot w^2)$ và một hàm phát hiện đặc trưng của cấu trúc Tensor (STSDO) có độ phức tạp $O(N^2)$. Việc tính LEF_STSDO được thực hiện bằng phép nhân Hadamard giữa hàm LEF và hàm STSDO nên phép nhân này có độ phức tạp là $O(N^2)$. Vì vậy, hàm LEF_STSDO có độ phức tạp tính toán là $O(N^2 \cdot w^2)$.

Thuật toán tổng hợp thành phần cơ sở AFM_MPA sẽ có độ phức tạp tính toán chính là độ phức tạp tính toán của thuật toán MPA. Theo tác giả Faramarzi và các cộng tác sự [31], thuật toán MPA có độ phức tạp tính toán là $O(t \cdot (n \cdot d + Cof \cdot n))$. Trong đó t là số vòng lặp của thuật toán, n là số cá thể trong bầy, d là số chiều của giải pháp và Cof là chi phí đánh giá hàm mục tiêu. Hàm mục tiêu của thuật toán AFM_MPA (công thức 3.11) bao gồm các hàm tính phương sai, trung bình cường độ sáng, tổng sai số bình phương giữa hình ảnh tổng hợp với hai hình ảnh xám MRI và PET, hiệu Entropy giữa hình ảnh tổng hợp với hình ảnh MRI và hiệu Entropy giữa hình ảnh tổng hợp với hình ảnh xám PET. Những hàm này đều có độ phức tạp tính toán là $O(N^2)$ nên hàm mục tiêu có độ phức tạp $O(N^2)$. Do đó, thuật toán tổng hợp thành phần cơ sở AFM_MPA có độ phức tạp tính toán là $O(t \cdot (n \cdot d + N^2 \cdot n))$. Vì

vậy, thuật toán tổng hợp AFM_CVLEF sẽ có độ phức tạp tính toán là $O(N^2 + N^2.k^2.l + N^2.w^2 + t.(n.d + N^2.n))$.

Trong trường hợp bộ dữ liệu là lớn và có p cặp ảnh. Khi đó, độ phức tạp tính toán của phương pháp AFM_CVLEF sẽ là $O(p.(N^2 + N^2.k^2.l + N^2.w^2 + t.(n.d + N^2.n)))$.

3.6. Thực nghiệm và đánh giá

Phần này trình bày chi tiết về dữ liệu được sử dụng trong các phần thực nghiệm, cách thiết lập các thực nghiệm và kết quả thực nghiệm thu được.

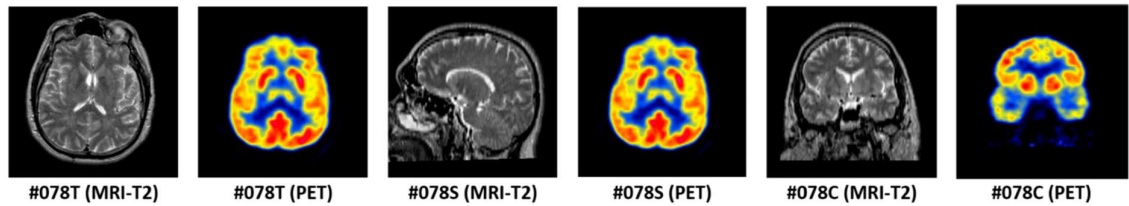
3.6.1. Dữ liệu thực nghiệm

Tổng số 156 hình ảnh bao gồm 78 cặp hình ảnh MRI và PET đã được sử dụng trong nghiên cứu này. Những hình ảnh này được lấy từ nguồn “The Whole Brain Atlas” (<http://www.med.harvard.edu/AANLIB/>) và sự phân chia các hình ảnh này thành các tập dữ liệu con trong các thực nghiệm được mô tả chi tiết như trong Bảng 3.1.

Bảng 3.1. Mô tả các tập dữ liệu thực nghiệm

Nhóm	Số lượng ảnh	Mô tả
K1	26 cặp MRI (T2) –PET	Các lát cắt từ 61 tới 86 theo trục T
K2	26 cặp MRI (T2) –PET	Các lát cắt từ 61 tới 86 theo trục S
K3	26 cặp MRI (T2) –PET	Các lát cắt từ 61 tới 86 theo trục C
K4	3 cặp MRI (T2) –PET	Lát cắt thứ 78 dọc theo các trục T, S, C.

Hình 3.7 minh họa một số hình ảnh được sử dụng trong phần thực nghiệm.



Hình 3.7. Minh họa ba cặp hình ảnh (MRI-PET) trong tập thực nghiệm

3.6.2. Thiết lập thực nghiệm

Môi trường và phần mềm cho các thực nghiệm được thiết kế như sau:

- Máy tính PC với Vi xử lý Intel core i9-10900K, 3.7 Hz. Bộ nhớ Ram: 64G
- Hệ điều hành Windows 10.
- Phần mềm Matlab R2020b.

Các thực nghiệm được thiết kế để kiểm tra hiệu quả của thuật toán tổng hợp AFM_CVLEF như sau:

Thực nghiệm số 1: Để giải thích tại sao thuật toán MPA được lựa chọn trong thuật toán đề xuất AFM_CVLEF, một số thuật toán tối ưu khác đã được lựa chọn để so sánh. Những thuật toán tối ưu này được mô tả trong Bảng 3.2. Dữ liệu thực nghiệm được sử dụng trong phần thực nghiệm này là K4. Mỗi thuật toán được chạy 30 lần khác nhau. Hai chỉ số đánh giá được sử dụng trong phần thực nghiệm này là trung bình và độ lệch chuẩn.

Bảng 3.2. Một số thuật toán tối ưu được sử dụng để so sánh

STT	Các thuật toán	Năm
1	Multi-Verse Optimizer (MVO) [82]	2016
2	Whale Optimization Algorithm (WOA) [83]	2016
3	Salp swarm algorithm (SSA) [84]	2017
4	Sine cosine algorithm (SCA) [85]	2016
5	Grey wolf optimizer (GWO) [86]	2014

Thực nghiệm số 2: So sánh thuật toán FM_CVLEF với một số thuật toán tổng hợp thành phần chi tiết khác như sau:

- Thuật toán lựa chọn Max (R_{Max}).
- Thuật toán cực đại hàm năng lượng cục bộ (R_{MLE}) [126].
- Thuật toán dựa trên mạng thần kinh kết hợp xung thích ứng tham số ($R_{PA-PCNN}$) [127].
- Thuật toán dựa trên tổng Laplace được sửa đổi (R_{SML}) [3].

Hai chỉ số đánh giá được lựa chọn trong thực nghiệm này là $Q^{AB/F}$ và FMI.

Thực nghiệm số 3: So sánh thuật toán đề xuất AFM_CVLEF với một số thuật toán tổng hợp ảnh khác được đề xuất gần đây. Các thuật toán được mô tả trong Bảng 3.3.

Bảng 3.3. Một số thuật toán tổng hợp hình ảnh

STT	Các thuật toán tổng hợp	Mô tả	Năm
1	PC-LLE-NSCT (G1) [121]	Thuật toán PC-LLE-NSCT sử dụng phương pháp biến đổi NSCT để phân rã mỗi hình ảnh đầu vào thành một thành phần tần số thấp và nhiều thành phần tần số cao. Thành phần tần số thấp được tổng hợp theo phương pháp đồng nhất pha (Phase Congruency). Các thành phần tần số	2019

		cao được tổng hợp theo phương pháp hàm năng lượng Laplacian cục bộ.	
2	TLD-SR (G2) [16]	Thuật toán TLD-SR sử dụng bộ lọc thông thấp và bộ lọc IGF (Interval Gradient Filter) để phân rã hình ảnh đầu vào thành 3 thành phần (cấu trúc, kết cấu và thành phần tần số cao). Thành phần cấu trúc được tổng hợp lại với nhau theo phương pháp lựa chọn trị tuyệt đối Max. Thành phần kết cấu được tổng hợp lại với nhau theo phương pháp tần số không gian lân cận (Neighbor spatial frequency). Thành phần tần số cao được tổng hợp với nhau dựa trên phương pháp biểu diễn thưa.	2021
3	JBF-LGE (G3) [29]	Thuật toán JBF-LGE thực hiện phân rã ảnh bởi bộ lọc song phương kết hợp (JBF). Các thành phần cơ sở sẽ được tổng hợp lại với nhau theo phương pháp lựa chọn max. các thành phần chi tiết được tổng hợp với nhau dựa trên hàm năng lượng Gradient lân cận.	2021
4	CSE (G4) [123]	Thuật toán CSE thực hiện như sau: Bước đầu tiên, sẽ loại bỏ các biến dạng khỏi hình ảnh nguồn và sau đó, trích xuất hai phần thông tin quan trọng: độ tương phản cục bộ và cấu trúc nổi bật. Cả độ tương phản cục bộ và cấu trúc nổi bật sau đó được kết hợp để có được bản đồ trọng số. Khi đó các trọng số thu được cho đi qua bộ lọc được dẫn hướng nhanh để loại bỏ sự gián đoạn và nhiễu. Cuối cùng, bản đồ trọng số được kết hợp với các	2021

		hình ảnh nguồn bằng cách sử dụng phương pháp phân rã kim tự tháp để có được hình ảnh tổng hợp cuối cùng.	
5	CNPS-NSST (G5) [124]	Thuật toán CNPS-NSST sử dụng phương pháp biến đổi NSST để phân rã hình ảnh đầu vào. Các thành phần tần số thấp được tổng hợp theo dựa trên hệ thống P nơ ron kết hợp. Các thành phần tần số cao được tổng hợp dựa trên phương pháp cải tiến từ tổng Laplacian được sửa đổi.	2021
6	DTNP-NSCT (G6) [40]	Thuật toán DTNP-NSCT sử dụng phương pháp biến đổi NSCT để phân rã hình ảnh đầu vào. Các thành phần tần số thấp được tổng hợp theo phương pháp cải tiến từ tổng Laplacian được sửa đổi. Các thành phần tần số cao được tổng hợp dựa trên hệ thống P nơ ron ngưỡng động.	2021
7	ACO (G7) [128]	Thuật toán ACO sử dụng bộ lọc co-occurrence filter để phân rã hình ảnh đầu vào. Các thành phần cơ sở được tổng hợp theo phương pháp tối đa hoá hàm năng lượng lân cận. Các thành phần chi tiết được tổng hợp theo phương pháp dựa trên bộ lọc DTNP.	2022

Những thuật toán này lựa chọn là những thuật toán mới được công bố trong 3 năm gần nhất (tính từ thời điểm nghiên cứu được gửi cho tạp chí vào năm 2022) và được phân vào 3 nhóm con của cách tiếp cận truyền thống. Nhóm thứ nhất là cách tiếp cận dựa trên biến đổi đa tỉ lệ. Thuật toán tổng hợp hình ảnh được lựa chọn trong nhóm này là CSE. Thuật toán này sử dụng biến đổi kim tự tháp để phân rã hình ảnh đầu vào. Nhóm thứ hai là cách tiếp cận dựa trên biến đổi đa tỉ lệ hình học. Có ba thuật toán tổng hợp hình ảnh được lựa chọn trong nhóm này bao gồm: PC-LLE-NSCT, DTNP-NSCT và CNPS-NSST. Trong đó, hai thuật toán PC-LLE-NSCT và DTNP-NSCT sử dụng biến đổi NSCT. Thuật toán còn lại CNPS-NSST sử dụng phép biến

đổi NSST để phân rã hình ảnh đầu vào. Nhóm thứ ba là các tiếp cận dựa trên bộ lọc. Có ba thuật toán được lựa chọn trong nhóm này bao gồm: TLD-SR, JBF-LGE và ACO. Trong đó, thuật toán TLD-SR sử dụng bộ lọc IGF (Interval Gradient Filter), thuật toán JBF-LGE sử dụng bộ lọc thông thấp và thuật toán ACO sử dụng bộ lọc ACOF (Adaptive co-occurrence filter) để phân rã hình ảnh đầu vào. Năm chỉ số đánh giá được lựa chọn trong thực nghiệm này bao gồm: MLI, CI, AG, $Q^{AB/F}$ và FMI.

3.6.3. Kết quả thực nghiệm

Các tham số cần thiết để chạy các thuật toán tối ưu như sau:

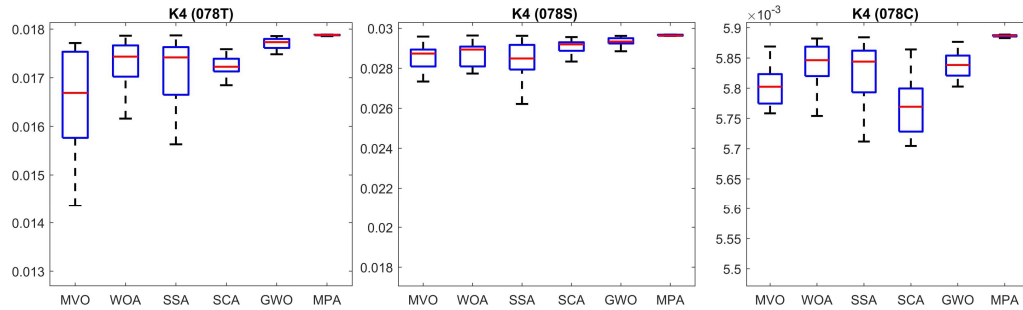
- Số lượng cá thể $n = 50$.
- Số lượng vòng lặp tối đa: $l_{max} = 50$.
- Cận dưới: $[0.8; 0; 0.9]$.
- Cận trên: $[1; 0.2; 1]$.
- Số chiều của giải pháp: $dim = 3$.

Bên cạnh các tham số cần thiết, các tham số riêng cho từng thuật toán tối ưu thuật (nếu có) được thiết lập mặc định. Riêng thuật toán MPA sử dụng một số tham số riêng như sau:

- Tham số $h = 0.5$,
- $FADs = 0.2$, k là số ngẫu nhiên thuộc $[0,1]$.

Kết quả thực nghiệm số 1:

Mỗi thuật toán tối ưu được chạy 30 lần độc lập trên tập dữ liệu K4. Hai chỉ số đánh giá là trung bình và độ lệch chuẩn được sử dụng để đánh giá hiệu quả cho mỗi thuật toán tối ưu. Bảng 3.4 và Hình 3.8 trình bày chi tiết các kết quả thu được.



Hình 3.8. Biểu đồ box plot minh họa giá trị hàm fitness thu được đối với 30 lần chạy độc lập của các thuật toán tối ưu khác nhau trên tập dữ liệu K4

Từ bảng Bảng 3.4, nhận thấy rằng các giá trị trung bình thu được từ thuật toán MPA là lớn nhất và các giá trị độ lệch chuẩn thu được từ thuật toán MPA là nhỏ nhất khi so sánh với các thuật toán tối ưu khác. Điều này giải thích tại sao thuật toán MPA được lựa chọn trong thuật toán tổng hợp hình ảnh đề xuất.

Bảng 3.4. Trung bình và độ lệch chuẩn thu được từ các thuật toán tối ưu với 30 lần chạy độc lập trên tập dữ liệu K4

Tập dữ liệu	Thuật toán	Trung bình	Độ lệch chuẩn
K4 (T)	MPA	0.017876671935888	0.000009033932259
	MVO	0.016518230357696	0.001090932265423
	WOA	0.017277469311324	0.000534054285376
	SSA	0.016899097232778	0.001118836345425
	SCA	0.017217245270695	0.000289792732718
	GWO	0.017669608631168	0.000189772085703
K4 (S)	MPA	0.029647134938738	0.000026405604409
	MVO	0.028356602594419	0.001161220657539
	WOA	0.028698747803551	0.000690713025417
	SSA	0.027886822958534	0.002416877831579
	SCA	0.029093335806103	0.000290569097601
	GWO	0.029300807571747	0.000241217843724
K4 (C)	MPA	0.005885862817517	0.000003159852779
	MVO	0.005790921812435	0.000067984350292
	WOA	0.005828160907535	0.000074232890256
	SSA	0.005816524467147	0.000069518552257
	SCA	0.005774386053179	0.000043737425495
	GWO	0.005836384455938	0.000020781565237

Hơn nữa, một kiểm định phi tham số Wilcoxon được sử dụng. Kết quả các giá trị p-values thu được từ kiểm thử được trình bày trong Bảng 3.5. Nhận thấy các giá trị p-values đều nhỏ hơn 0.05. Kết quả này cho thấy kết quả có ý nghĩa về mặt thống kê.

Bảng 3.5. Bảng các giá trị p-values thu được từ kiểm định Wilcoxon

Tập dữ liệu	Thuật toán	P-values
K4 (T)	MPA vs MVO	3.019859359162151e-11
	MPA vs WOA	3.689725853981014e-11
	MPA vs SSA	6.065757009046759e-11
	MPA vs SCA	3.019859359162151e-11
	MPA vs GWO	3.689725853981014e-11
K4 (S)	MPA vs MVO	3.019859359162151e-11
	MPA vs WOA	1.464306887715034e-10
	MPA vs SSA	6.695518965500180e-11
	MPA vs SCA	3.019859359162151e-11
	MPA vs GWO	4.077164846825348e-11
K4 (C)	MPA vs MVO	3.019859359162151e-11
	MPA vs WOA	8.993406027014880e-11
	MPA vs SSA	3.019859359162151e-11
	MPA vs SCA	3.019859359162151e-11
	MPA vs GWO	4.504322112705322e-11

Kết quả thực nghiệm số 2:

Kết quả của việc sử dụng các thuật toán khác nhau được trình bày trong Bảng 3.6. Hai chỉ số đánh giá cho việc bảo toàn các thông tin chi tiết và sự tương đồng giữa các hình ảnh đầu vào và hình ảnh tổng hợp dựa trên thuyết thông tin được sử dụng. Cả hai chỉ số $Q^{AB/F}$ và FMI thu được từ thuật toán FM_CVLEF là cao nhất khi so với các chỉ số này thu được từ các thuật toán tổng hợp còn lại. Kết quả này cho thấy hình ảnh tổng hợp được tạo ra từ thuật toán đề xuất FM_CVLEF chứa nhiều thông tin từ hình ảnh đầu vào và ít bị biến dạng hoặc mất thông tin khi so với hình ảnh được tạo ra từ các thuật toán tổng hợp còn lại.

Bảng 3.6. Kết quả thu được bởi các thuật toán tổng hợp khác nhau

Tập dữ liệu	Thuật toán	$Q^{AB/F}$	FMI
K1	R_{Max}	0.4829	0.8144
	R_{MLE}	0.7124	0.8720
	$R_{PA-PCNN}$	0.6669	0.8658
	R_{SML}	0.6984	0.8632
	FM_CVLEF	0.7440	0.8737
K2	R_{Max}	0.4868	0.8092
	R_{MLE}	0.7271	0.8749
	$R_{PA-PCNN}$	0.6893	0.8689
	R_{SML}	0.6802	0.8619
	FM_CVLEF	0.7539	0.8782
K3	R_{Max}	0.4957	0.8073
	R_{MLE}	0.7454	0.8728
	$R_{PA-PCNN}$	0.7152	0.8664
	R_{SML}	0.7146	0.8600
	FM_CVLEF	0.7716	0.8743

Kết quả thực nghiệm số 3:

Trước tiên, các tham số tối ưu thu được từ thuật toán đề xuất AFM_CVLEF trên ba tập dữ liệu K1, K2 và K3 được trình bày trong Bảng 3.7.

Bảng 3.7. Các tham số tối ưu thu được

Tập dữ liệu	ρ_1	ρ_2	ρ_3
K1	0.9927	0.0453	0.999
K2	0.9829	0.0524	0.999
K3	0.9942	0.0575	0.999

Tham số ρ_3 gần xấp xỉ bằng 1 báo hiệu rằng hình ảnh tổng hợp có sự tương đồng cao nhất đối với hình ảnh cộng hưởng từ MRI. Tham số ρ_1 khá lớn (gần bằng 1). Kết quả này cho thấy rằng hình ảnh cộng hưởng từ đóng góp một lượng đáng kể thông tin quan trọng cho hình ảnh tổng hợp. Tham số ρ_2 thu được khá nhỏ cho thấy nó chỉ đóng góp một lượng ít thông tin cho hình ảnh tổng hợp. Kết quả này cho thấy rằng hình ảnh PET chỉ đóng góp một lượng rất nhỏ thông tin cho hình ảnh tổng hợp đầu ra.

Tiếp theo, kết quả của thuật toán tổng hợp hình ảnh đề xuất AFM_CVLEF với các thuật toán tổng hợp hình ảnh khác trên ba tập dữ liệu K1, K2 và K3 được mô tả trong các bảng Bảng 3.8, Bảng 3.9 và Bảng 3.10. Những kết quả của các chỉ số đánh giá này được minh họa trực quan trên các biểu đồ Hình 3.9, Hình 3.12 và Hình 3.15. Các hình ảnh tổng hợp bởi các thuật toán tổng hợp khác nhau và thuật toán đề xuất được trình bày trong các Hình 3.10, Hình 3.13 và Hình 3.16. Bằng trực quan, dễ dàng nhận thấy rằng hình ảnh tổng hợp được tạo từ thuật toán đề xuất AFM_CVLEF có cường độ sáng, độ tương phản và độ sắc nét đều tốt hơn so với hình ảnh được tạo ra từ các thuật toán tổng hợp hình ảnh còn lại.

Để quan sát rõ hơn các chi tiết trong các hình ảnh tổng hợp, một vùng ảnh nhỏ được cắt ra từ các hình ảnh tổng hợp trong các hình Hình 3.10, Hình 3.13 và Hình 3.16. Những vùng ảnh nhỏ này được minh họa trong các Hình 3.11, Hình 3.14 và Hình 3.17 tương ứng.

Bảng 3.8. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K1

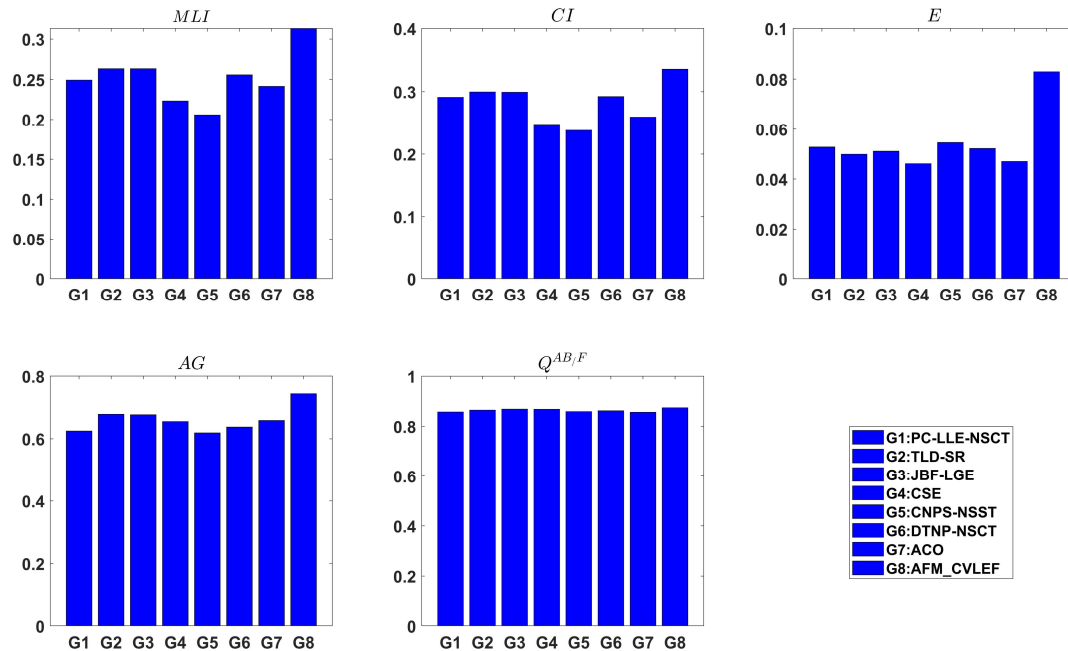
DL	Thuật toán	MLI	CI	AG	$Q^{AB/F}$	FMI
K1	G1	0.2492	0.2910	0.0528	0.6248	0.8569
	G2	0.2634	0.2997	0.0499	0.6786	0.8646
	G3	0.2634	0.2992	0.0511	0.6771	0.8682
	G4	0.2233	0.2479	0.0461	0.6552	0.8681
	G5	0.2060	0.2382	0.0546	0.6192	0.8583
	G6	0.2558	0.2922	0.0522	0.6376	0.8619
	G7	0.2415	0.2596	0.0470	0.6587	0.8562
	AFM_CVLEF	0.3131	0.3356	0.0829	0.7440	0.8737

Bảng 3.9. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K2

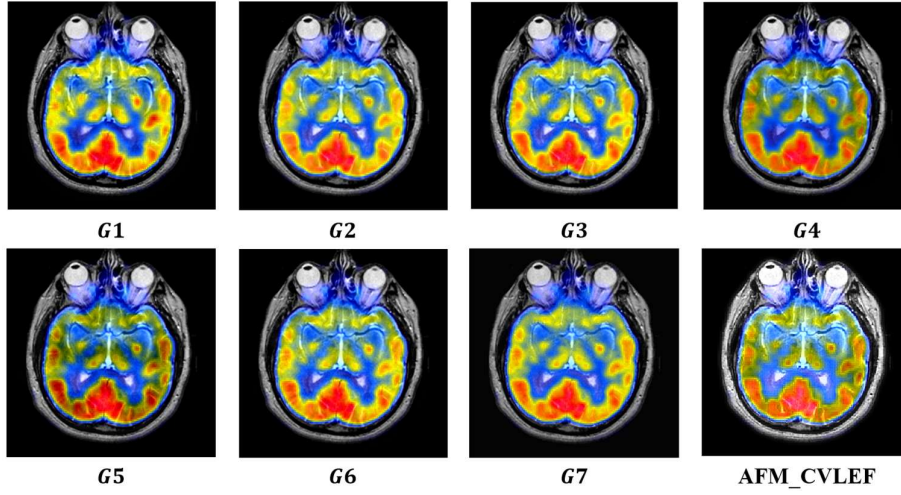
DL	Thuật toán	MLI	CI	AG	$Q^{AB/F}$	FMI
K2	G1	0.2555	0.2816	0.0506	0.6432	0.8675
	G2	0.2700	0.2905	0.0474	0.6818	0.8740
	G3	0.2713	0.2920	0.0487	0.7099	0.8765
	G4	0.2317	0.2410	0.0435	0.6703	0.8762
	G5	0.2162	0.2305	0.0517	0.6344	0.8694
	G6	0.2626	0.2835	0.0497	0.6537	0.8714
	G7	0.2754	0.2727	0.0463	0.6996	0.8695
	AFM_CVLEF	0.3313	0.3217	0.0740	0.7539	0.8782

Bảng 3.10. Các chỉ số đánh giá thu được từ thực nghiệm trên tập dữ liệu K3

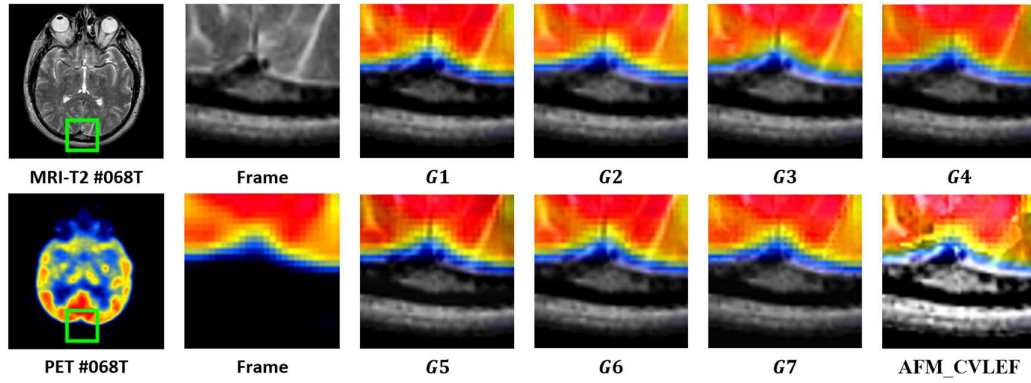
DL	Thuật toán	MLI	CI	AG	$Q^{AB/F}$	FMI
K3	G1	0.2060	0.2648	0.0452	0.6628	0.8674
	G2	0.2175	0.2757	0.0428	0.6987	0.8726
	G3	0.2181	0.2761	0.0435	0.7133	0.8766
	G4	0.1925	0.2380	0.0398	0.6914	0.8716
	G5	0.1789	0.2251	0.0463	0.6524	0.8665
	G6	0.2120	0.2684	0.0446	0.6732	0.8716
	G7	0.2103	0.2421	0.0400	0.6980	0.8644
	AFM_CVLEF	0.2778	0.3230	0.0693	0.7716	0.8743



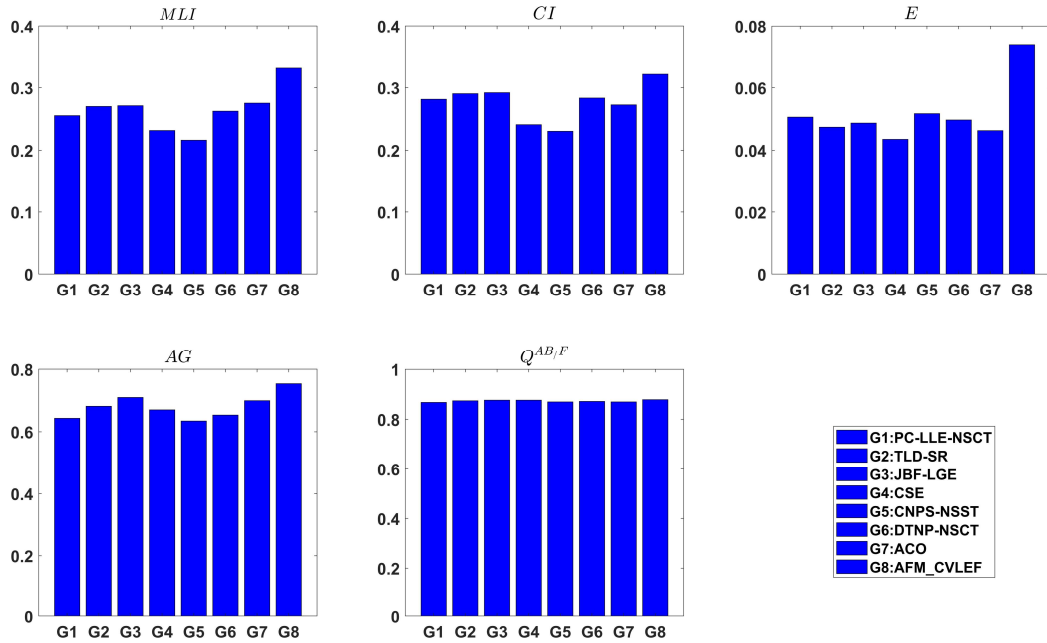
Hình 3.9. Minh họa cho năm chỉ số đánh giá bởi tám thuật toán trên tập K1



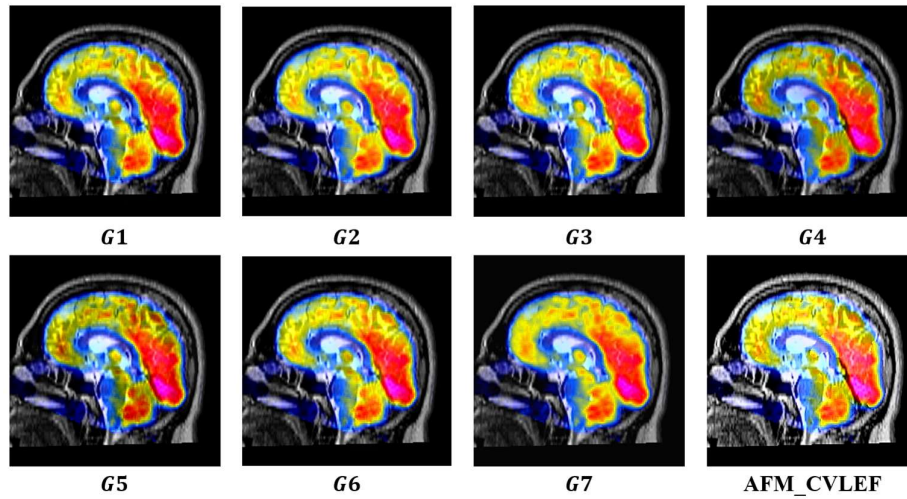
Hình 3.10. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068T từ K1



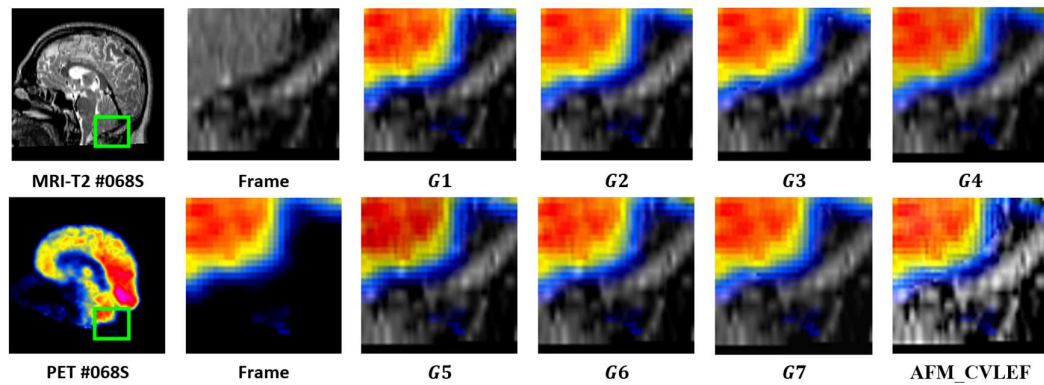
Hình 3.11. Minh họa các vùng ảnh nhỏ được cắt từ Hình 3.10



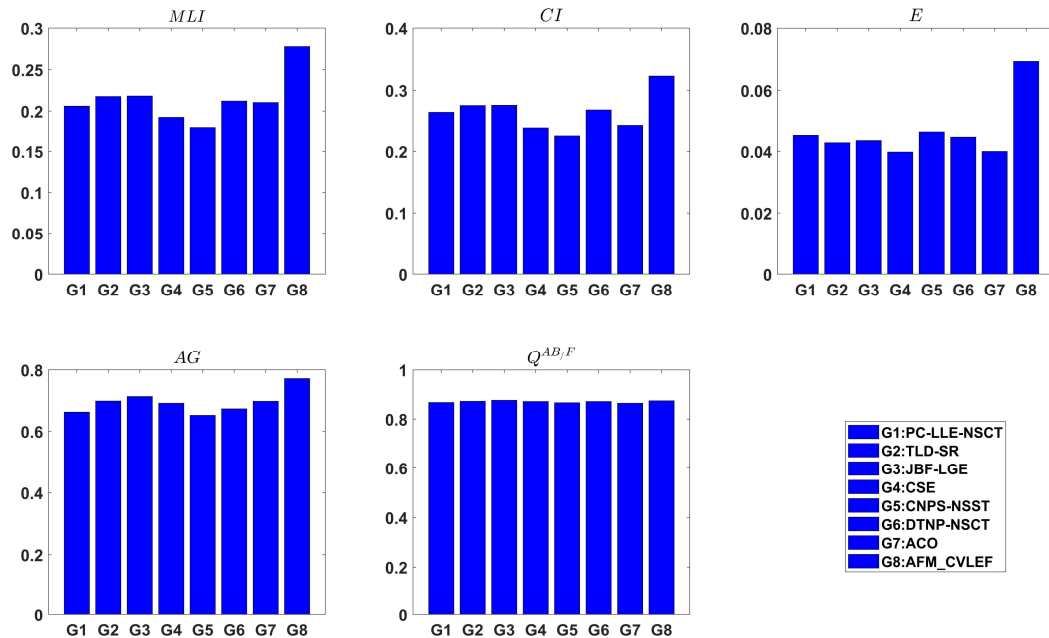
Hình 3.12. Minh họa cho năm chỉ số đánh giá từ bởi tám thuật toán trên tập K2



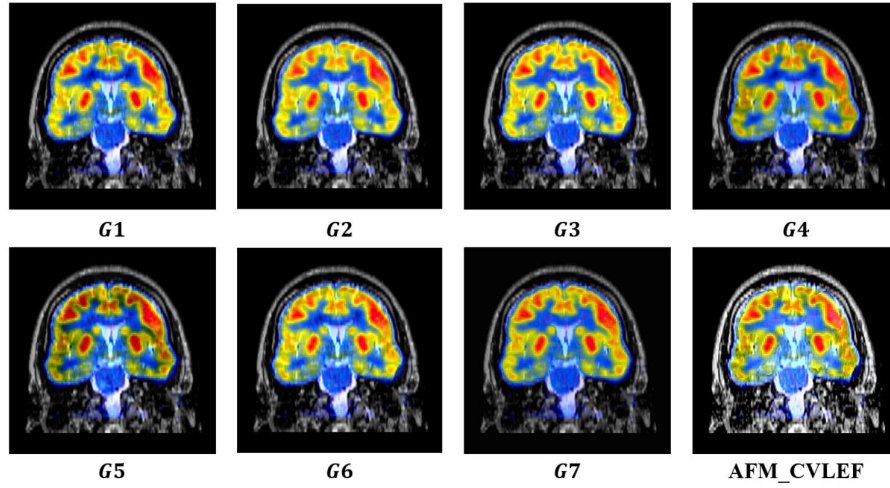
Hình 3.13. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068S từ K2



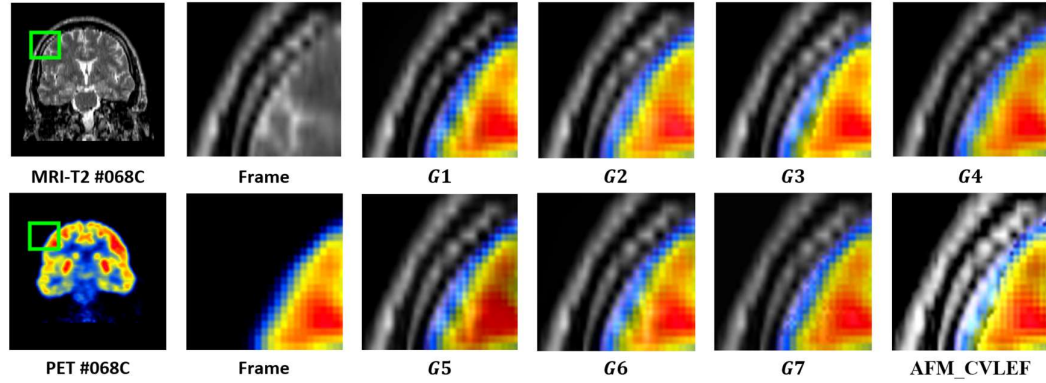
Hình 3.14. Minh họa các vùng ảnh nhỏ được cắt từ Hình 3.13



Hình 3.15. Minh họa cho năm chỉ số đánh giá từ bởi tám thuật toán trên tập K3



Hình 3.16. Các hình ảnh tổng hợp thu được từ tám thuật toán khác nhau cho cặp ảnh MRI-PET #068T từ K3



Hình 3.17. Minh họa các vùng ảnh nhỏ được cắt từ Hình 3.16

Về mặt định lượng, từ các bảng kết quả thực nghiệm (Bảng 3.8, Bảng 3.9 và Bảng 3.10), các chỉ số đánh giá thu được từ thuật toán tổng hợp đề xuất AFM_CVLEF là tốt hơn so với các thuật toán tổng hợp hình ảnh khác. Thứ nhất về nhóm chỉ số đánh giá chất lượng ảnh. Ba chỉ số MLI, CI và AG thu được từ thuật toán đề xuất AFM_CVLEF là cao nhất khi so sánh với các thuật toán tổng hợp còn lại. Ví dụ, chỉ số cường độ sáng trung bình MLI thu được từ thuật toán đề xuất AFM_CVLEF trên tập dữ liệu K1 là 0.3131, trong khi chỉ số MLI của các thuật toán còn lại là thấp hơn, nằm trong khoảng [0.2060, 0.2634]. Kết quả tương tự đối với hai chỉ số CI và AG. Các kết quả này cho thấy hình ảnh được tạo ra từ thuật toán đề xuất AFM_CVLEF có chất lượng tốt hơn các hình ảnh được tạo ra từ các thuật toán tổng hợp hình ảnh còn lại ở các khía cạnh cường độ sáng trung bình, độ tương phản và độ sắc nét.

Thứ hai là xét chỉ số bảo toàn đường biên từ hình ảnh đầu vào. Chỉ số $Q^{AB/F}$ thu được từ thuật toán đề xuất cũng cao nhất trên cả ba tập dữ liệu K1, K2 và K3. Ví dụ, Chỉ số $Q^{AB/F}$ thu được từ thuật toán đề xuất thu được trên tập dữ liệu K1 là

0.7440. Giá trị này cao nhất so với các giá trị $Q^{AB/F}$ thu được từ các thuật toán tổng hợp hình ảnh còn lại, mà chúng chỉ nằm trong khoảng [0.6192, 0.6786]. Kết quả này cho thấy thuật toán đề xuất bảo toàn tốt các đặc trưng từ hình ảnh đầu vào so với các thuật toán tổng hợp khác.

Thứ ba là xét chỉ số FMI được sử dụng để đánh giá sự tương đồng giữa các hình ảnh đầu vào và hình ảnh tổng hợp, mà nó dựa trên thuyết thông tin. Dễ dàng nhận thấy rằng chỉ số FMI thu được từ thuật toán đề xuất cũng đạt giá trị cao nhất so với các chỉ số FMI của các thuật toán tổng hợp khác. Ví dụ, chỉ số FMI của thuật toán đề xuất đạt giá trị 0.8737 trên tập dữ liệu K1, trong khi chỉ số này của các thuật toán còn lại là thấp hơn, chỉ nằm trong khoảng [0.8562, 0.8682]. Kết quả thu được tương tự trên các tập dữ liệu K2 và K3. Các kết quả này cho thấy rằng, hình ảnh tổng hợp được tạo ra từ thuật toán đề xuất chứa nhiều thông tin từ hình ảnh đầu vào và ít bị biến dạng hoặc mất thông tin khi so với hình ảnh được tạo ra từ các thuật toán tổng hợp còn lại.

Thời gian chạy trung bình của các thuật toán tổng hợp hình ảnh trên cả ba tập dữ liệu K1, K2 và K3 được mô tả trong Bảng 3.11. Thuật toán tổng hợp đề xuất AFM_CVLEF mất trung bình 4.58 giây để hoàn tất quá trình tổng hợp. Đúng vị trí thứ 6 trên 8 thuật toán về thời gian chạy. Thuật toán G4 cho thời gian chạy nhanh nhất là khoảng 0.1 giây trong khi thuật toán G6 chạy lâu nhất với khoảng 17.9 giây. Tuy nhiên, đối với thuật toán tổng hợp đề xuất AFM_CVLEF thời gian này tăng lên khi số lần lặp và số cá thể trong bầy của thuật toán MPA tăng lên. Điều này có thể là hạn chế của thuật toán tổng hợp đề xuất AFM_CVLEF.

Bảng 3.11. Thời gian chạy trung bình của các thuật toán trên (K1&K2&K3)

Thuật toán	Thời gian chạy trung bình (s)
G1	3.02972
G2	14.92673
G3	0.26316
G4	0.10035
G5	3.36244
G6	17.90511
G7	1.05729
AFM_CVLEF	4.57611

3.7. Kết luận chương 3

Trong chương này, một phương pháp tổng hợp hình ảnh đã được đề xuất để cải thiện hiệu quả của việc tổng hợp hình ảnh. Phương pháp tổng hợp hình ảnh đề xuất (AFM_CVLEF) bao gồm 3 thuật toán chính.

Thứ nhất là thuật toán phân rã hình ảnh thành ba thành phần TCID. Thuật toán TCID được xây dựng dựa trên bộ lọc RGF và bộ lọc đường cong trung bình có trong số WMCF.

Thứ hai là thuật toán tổng hợp thích nghi cho các thành phần cơ sở AFM_MPA. Thuật toán AFM_MPA được xây dựng dựa trên thuật toán tối ưu hoá MPA để tìm ra các tham số tối ưu cho các thành phần cơ sở. Thuật toán tổng hợp AFM_MPA cho phép hình ảnh tổng hợp thu được không bị suy giảm về mặt chất lượng trong quá trình tổng hợp.

Thứ ba là thuật toán tổng hợp cho thành phần chi tiết FM_CVLEF. Thuật toán FM_CVLEF được xây dựng dựa trên kết hợp hàm năng lượng cục bộ với các biến thể của nó như hàm năng lượng cục bộ sử dụng toán tử la bàn Prewitt và hàm năng lượng cục bộ sử dụng toán tử phát hiện đặc trưng cấu trúc Tensor. Thuật toán tổng hợp FM_CVLEF này giúp bảo tồn các thông tin chi tiết được mang từ hình ảnh đầu vào đến hình ảnh tổng hợp.

Phương pháp tổng hợp hình ảnh đề xuất (AFM_CVLEF) được xây dựng dựa trên sự kết hợp của các thuật toán: TCID, AFM_MPA và FM_CVLEF. Thuật toán tổng hợp AFM_CVLEF đã được so sánh với bảy thuật toán tổng hợp ảnh khác được đề xuất gần đây để kiểm tra sự hiệu quả. Kết quả của thuật toán tổng hợp hình ảnh đề xuất AFM_CVLEF đã được công bố tại các công trình [CT3, CT4, CT5, CT6] trong phần “Danh mục các công trình của tác giả”.

KẾT LUẬN

Tổng hợp hình ảnh y học là một trong những chủ đề nghiên cứu quan trọng trong lĩnh vực xử lý ảnh. Bài toán này liên quan đến việc kết hợp các hình ảnh y học được chụp từ các cách khác nhau để tạo ra một hình ảnh duy nhất có chất lượng cao và chứa được các thông tin quan trọng từ từng hình ảnh riêng lẻ. Tuy nhiên, việc tổng hợp hình ảnh vẫn còn có những hạn chế nhất định bởi hai nguyên nhân chính. Thứ nhất là các hình ảnh đầu vào vẫn còn có chất lượng thấp như có cường độ sáng, độ tương phản thấp và có thể bị nhiễu. Thứ hai là các phương pháp tổng hợp hiện tại chưa thực sự hiệu quả dẫn đến hình ảnh tổng hợp bị suy giảm về mặt chất lượng và bị mất các thông tin chi tiết từ hình ảnh đầu vào. Các đóng góp chính của luận văn nhằm hướng tới giải quyết hai hạn chế đề cập ở trên bằng cách đề xuất hai nhóm phương pháp chính: Nhóm thứ nhất là một phương pháp cải thiện chất lượng cho hình ảnh đầu. Nhóm thứ hai là thuật toán tổng hợp ảnh hiệu quả.

Luận án đã có những đóng góp như sau:

- (1) Đề xuất một phương pháp nâng cao chất lượng cho hình ảnh cộng hưởng từ não IE_TCID_MPA mà đã được công bố tại công trình [CT1]. Phương pháp nâng cao chất lượng ảnh đề xuất thực hiện phân rã hình ảnh đầu vào thành các lớp chứa các thông tin khác nhau như: lớp cấu trúc, lớp kết cấu và lớp nhiễu. Sau đó, phương pháp này đã thực hiện tăng cường ảnh bằng cách xử lý trên từng lớp thông tin và bổ sung thêm lớp thông tin chi tiết sử dụng thành phần đặc trưng cấu trúc Tensor. Thuật toán tối ưu hoá MPA được sử dụng để tìm các tham số tối ưu cho từng lớp thành phần. Cách tiếp cận này cho phép cải thiện được đáng kể cường độ sáng và độ tương phản của hình ảnh đồng thời hạn chế được nhiễu phát sinh trong quá trình tăng cường ảnh. Phương pháp đề xuất áp dụng cho các hình ảnh đầu vào trước khi tổng hợp đã cho thấy cải thiện đáng kể hiệu quả của một số thuật toán tổng hợp hiện có. Phương pháp IE_TCID_MPA cũng đã được ứng dụng trong việc tiền xử lý ảnh đầu vào nhằm nâng cao hiệu quả của việc tổng hợp hình ảnh y học. Kết quả đã này đã được công bố tại công trình [CT2].
- (2) Đề xuất một phương pháp tổng hợp hình ảnh mới (AFM_CVLEF) để nâng cao hiệu quả của việc tổng hợp hình ảnh. Phương pháp đề xuất nhằm giải quyết hai vấn đề còn tồn tại: (a) suy giảm cường độ sáng, độ tương phản của hình ảnh tổng hợp; (b) sự mất thông tin trong hình ảnh tổng hợp. Để giải quyết cho vấn đề thứ nhất, một thuật toán tổng hợp thích nghi (AFM_MPA) cho các thành phần cơ sở được đề xuất. Trong đó, các tham số thích nghi được xác định dựa trên thuật toán tối ưu hoá MPA. Để giải

quyết cho vấn đề thứ hai, một thuật toán tổng hợp hiệu quả cho các thành phần chi tiết được đề xuất (FM_CVLEF). Thuật toán FM_CVLEF kết hợp hàm năng lượng cục bộ với các biến thể của nó như hàm năng lượng cục bộ sử dụng toán tử la bàn Prewitt và hàm năng lượng cục bộ sử dụng thành phần đặc trưng của cấu trúc Tensor. Kết quả của thuật toán tổng hợp hình ảnh đề xuất đã được công bố tại các công trình [CT3, CT4, CT6] trong mục “Danh mục các công trình của tác giả”.

Tổng hợp hình ảnh y học vẫn còn một số vấn đề cần được tiếp tục nghiên cứu để cải thiện hiệu quả. Luận án chỉ hướng tới giải quyết hai vấn đề: (a) cải thiện chất lượng của hình ảnh đầu vào; (b) cải thiện hiệu quả của việc tổng hợp các thành phần cơ sở và chi tiết. Trong tương lai, một số cách tiếp cận tiềm năng có thể được nghiên cứu để cải thiện hiệu quả của việc tổng hợp như sau:

- Để cải thiện hiệu quả của thuật toán tổng hợp cho các thành phần chi tiết, các mạng học sâu có thể được sử dụng để trích xuất đặc trưng từ hình ảnh. Với một số lượng lớn các đặc trưng thu được từ các mạng học sâu sẽ có khả năng cao giúp cho thuật toán tổng hợp cho thành phần chi tiết bảo toàn được các thông tin chi tiết từ hình ảnh đầu vào. Cách tiếp cận dựa trên mạng học sâu để tổng hợp cho các thành phần chi tiết cũng được công bố tại công trình [CT5]. Nghiên cứu này đã sử dụng mạng VGG-19 sử dụng kỹ thuật học chuyển đổi với việc phân 4 lớp hình ảnh: MRI, CT, PET và SPECT để có thể biểu diễn tốt các đặc trưng cho 4 loại hình ảnh trên. Trên cơ sở đó mạng VGG-19 đã được sửa đổi, một phương pháp tổng hợp cho các thành phần chi tiết đã được giới thiệu [CT5]. Tuy nhiên, việc tận dụng các mạng học sâu khác như Resnet-50, Resnet-101 để xây dựng các phương pháp tổng hợp hiệu quả cho các thành phần chi tiết vẫn là vấn đề đặt ra trong các nghiên cứu tiếp theo trong tương lai.
- Để cải thiện hiệu quả về thời gian chạy cho thuật toán tổng hợp hình ảnh, một số thuật toán tối ưu hoá mới được đề xuất gần đây như CSA (Chameleon swarm algorithm) [129], WSO (White Shark Optimizer) [130], COA (Coati Optimization Algorithm) [131] có thể được áp dụng để thay thế cho thuật toán tối ưu hoá MPA.
- Để cải thiện hơn nữa kết quả của việc tổng hợp thì một hướng tiếp cận tiềm năng là kết hợp cả giai đoạn tăng cường ảnh và giai đoạn tổng hợp hình ảnh vào chung một mô hình tổng hợp. Đồng thời sử dụng duy nhất một hàm tối ưu để kiểm soát chất lượng của ảnh đầu ra.

DANH MỤC CÁC CÔNG TRÌNH CỦA TÁC GIẢ

TẠP CHÍ KHOA HỌC QUỐC TẾ

[CT1]	P. H. Dinh and L. G. Nguyen, 2022, A new medical image enhancement algorithm using adaptive parameters, <i>International Journal of Imaging Systems and Technology</i> , vol. 32, no. 6, pp. 2198–2218. (SCIE, Q2) .
[CT2]	P. H. Dinh , 2023, Combining Spectral Total Variation with Dynamic Threshold Neural P Systems for Medical Image Fusion, <i>Biomedical Signal Processing and Control</i> , vol. 80, pp. 104343. (SCIE, Q1) .
[CT3]	P. H. Dinh , 2022, A novel approach using structure tensor for medical image fusion, <i>Multidimensional Systems and Signal Processing</i> , vol. 33, pp. 1001–1021. (SCIE, Q2) .
[CT4]	P. H. Dinh , 2023, A novel approach using the local energy function and its variations for medical image fusion, <i>The Imaging Science Journal</i> , pp. 660-676. (SCIE, Q2) .
[CT5]	P. H. Dinh and L. G. Nguyen, 2023, Medical image fusion based on Transfer learning techniques and Coupled Neural P Systems, <i>Neural Computing and Applications</i> , (SCIE, Q1) .

HỘI NGHỊ KHOA HỌC

[CT6]	Đinh Phú Hùng , Nguyễn Huy Đức, Nguyễn Long Giang, 2022, Tổng hợp hình ảnh y học dựa trên giải thuật tối ưu MPA, <i>Hội nghị Khoa học công nghệ Quốc gia lần thứ XIV về Nghiên cứu Cơ bản và Ứng dụng Công nghệ thông tin (FAIR2022)</i> , trang 501-509 .
-------	---

TÀI LIỆU THAM KHẢO

- [1] Azam M.A., Khan K.B., Salahuddin S., Rehman E., Khan S.A., Khan M.A., et al., 2022, A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics, *Computers in Biology and Medicine*, 144, pp. 102053.
- [2] Mi S., Zhang L., Peng H., Wang J., 2021, Medical image fusion based on DTNP systems and Laplacian pyramid, *Journal of Membrane Computing*, 3(4), pp. 284–295.
- [3] Li X., Zhang X., Ding M., 2019, A sum-modified-Laplacian and sparse representation based multimodal medical image fusion in Laplacian pyramid domain, *Medical and Biological Engineering and Computing*, 57, pp. 2265–2275.
- [4] Hamidu N.M., Liu Q., Zhang C., 2021, Medical Image Fusion Based on Undecimated Dual-Tree Complex Wavelet Transform and NSCT, *Communications in Computer and Information Science*, 1454, pp. 472–484.
- [5] Wang S., Shen Y., 2020, Multi-modal image fusion based on saliency guided in NSCT domain, *IET Image Processing*, 14(13), pp. 3188–3201.
- [6] Li H., Qiu H., Yu Z., Zhang Y., 2016, Infrared and visible image fusion scheme based on NSCT and low-level visual features, *Infrared Physics and Technology*, 76, pp. 174–184.
- [7] Li Y., Sun Y., Huang X., Qi G., Zheng M., Zhu Z., 2018, An image fusion method based on sparse representation and Sum Modified-Laplacian in NSCT Domain, *Entropy*, 20(7), pp. 522.
- [8] Guo P., Xie G., Li R., Hu H., 2022, Multimodal medical image fusion with convolution sparse representation and mutual information correlation in NSST domain, *Complex and Intelligent Systems*, 9, pp. 317–328.
- [9] Kong W., Miao Q., Lei Y., Ren C., 2022, Guided filter random walk and improved spiking cortical model based image fusion method in NSST domain, *Neurocomputing*, 488, pp. 509–527.
- [10] Gao Y., Ma S., Liu J., Liu Y., Zhang X., 2021, Fusion of medical images based on salient features extraction by PSO optimized fuzzy logic in NSST domain, *Biomedical Signal Processing and Control*, 69, pp. 102852.
- [11] Gai D., Shen X., Cheng H., Chen H., 2019, Medical Image Fusion via PCNN Based on Edge Preservation and Improved Sparse Representation in NSST Domain, *IEEE Access*, 7, pp. 85413–85429.

- [12] Diwakar M., Singh P., Shankar A., 2021, Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain, *Biomedical Signal Processing and Control*, 68, pp. 102788.
- [13] Jose J., Gautam N., Tiwari M., Tiwari T., Suresh A., Sundararaj V., et al., 2021, An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion, *Biomedical Signal Processing and Control*, 66, pp. 102480.
- [14] Asha C.S., Lal S., Gurupur V.P., Saxena P.U.P., 2019, Multi-Modal Medical Image Fusion with Adaptive Weighted Combination of NSST Bands Using Chaotic Grey Wolf Optimization, *IEEE Access*, 7, pp. 40782–40796.
- [15] Hu Q., Hu S., Zhang F., 2021, Multi-modality image fusion combining sparse representation with guidance filtering, *Soft Computing*, 25, pp. 4393–4407.
- [16] Li X., Zhou F., Tan H., 2021, Joint image fusion and denoising via three-layer decomposition and sparse representation, *Knowledge-Based Systems*, 224, pp. 107087.
- [17] Liu F., Chen L., Lu L., Jeon G., Yang X., 2021, Infrared and visible image fusion via rolling guidance filter and convolutional sparse representation, *Journal of Intelligent and Fuzzy Systems*, 40(6), pp. 10603–10616.
- [18] Barba-J L., Vargas-Quintero L., Calderón-Agudelo J.A., 2022, Bone SPECT/CT image fusion based on the discrete Hermite transform and sparse representation, *Biomedical Signal Processing and Control*, 71, pp. 103096.
- [19] Wang L., Dou J., Qin P., Lin S., Gao Y., Wang R., et al., 2021, Multimodal medical image fusion based on nonsubsampling shearlet transform and convolutional sparse representation, *Multimedia Tools and Applications*, 80(30), pp. 36401–36421.
- [20] Li W., Zhang Y., Wang G., Huang Y., Li R., 2023, DFENet: A dual-branch feature enhanced network integrating transformers and convolutional feature learning for multimodal medical image fusion, *Biomedical Signal Processing and Control*, 80, pp. 104402.
- [21] Fu J., Li W., Du J., Xiao B., 2020, Multimodal medical image fusion via laplacian pyramid and convolutional neural network reconstruction with local gradient energy strategy, *Computers in Biology and Medicine*, 126, pp. 104048.
- [22] Tawfik N., Elnemr H.A., Fakhr M., Dessouky M.I., El-Samie F.E.A., 2022, Multimodal Medical Image Fusion Using Stacked Auto-encoder in NSCT Domain, *Journal of Digital Imaging*, 35(5), pp. 1308–1325.

- [23] Ma J., Yu W., Chen C., Liang P., Guo X., Jiang J., 2020, Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion, *Information Fusion*, 62, pp. 110–120.
- [24] Zhan B., Li D., Wu X., Zhou J., Wang Y., 2022, Multi-Modal MRI Image Synthesis via GAN with Multi-Scale Gate Mergence, *IEEE Journal of Biomedical and Health Informatics*, 26(1), pp. 17–26.
- [25] Zhou T., Cheng Q.R., Lu H.L., Li Q., Zhang X.X., Qiu S., 2023, Deep learning methods for medical image fusion: A review, *Computers in Biology and Medicine*, 160, pp. 106959.
- [26] Ullah H., Zhao Y., Abdalla F.Y.O., Wu L., 2022, Fast local Laplacian filtering based enhanced medical image fusion using parameter-adaptive PCNN and local features-based fuzzy weighted matrices, *Applied Intelligence*, 52(7), pp. 7965–7984.
- [27] Maqsood S., Javed U., 2020, Multi-modal Medical Image Fusion based on Two-scale Image Decomposition and Sparse Representation, *Biomedical Signal Processing and Control*, 57, pp. 101810.
- [28] Zhang L., Li H., Zhu R., Du P., 2022, An infrared and visible image fusion algorithm based on ResNet-152, *Multimedia Tools and Applications*, 81(7), pp. 9277–9287.
- [29] Li X., Zhou F., Tan H., Zhang W., Zhao C., 2021, Multimodal medical image fusion based on joint bilateral filter and local gradient energy, *Information Sciences*, 569, pp. 302–325.
- [30] Liu X., Mei W., Du H., 2018, Multi-modality medical image fusion based on image decomposition framework and nonsubsampling shearlet transform, *Biomedical Signal Processing and Control*, 40, pp. 343–350.
- [31] Faramarzi A., Heidarinejad M., Mirjalili S., Gandomi A.H., 2020, Marine Predators Algorithm: A nature-inspired metaheuristic, *Expert Systems with Applications*, 152, pp. 113377.
- [32] Song J., Zheng J., Li P., Lu X., Zhu G., Shen P., 2021, An Effective Multimodal Image Fusion Method Using MRI and PET for Alzheimer's Disease Diagnosis, *Frontiers in Digital Health*, 3, pp. 637386.
- [33] Guo Z., Li X., Huang H., Guo N., Li Q., 2019, Deep learning-based image segmentation on multimodal medical imaging, *IEEE Transactions on Radiation and Plasma Medical Sciences*, 3(2), pp. 162–169.
- [34] Du J., Li W., Lu K., Xiao B., 2016, An overview of multi-modal medical image fusion, *Neurocomputing*, 215, pp. 3–20.

- [35] Wang Z., Cui Z., Zhu Y., 2020, Multi-modal medical image fusion by Laplacian pyramid and adaptive sparse representation, *Computers in Biology and Medicine*, 123, pp. 103823.
- [36] Xu L., Si Y., Jiang S., Sun Y., Ebrahimian H., 2020, Medical image fusion using a modified shark smell optimization algorithm and hybrid wavelet-homomorphic filter, *Biomedical Signal Processing and Control*, 59, pp. 101885.
- [37] Chao Z., Duan X., Jia S., Guo X., Liu H., Jia F., 2022, Medical image fusion via discrete stationary wavelet transform and an enhanced radial basis function neural network, *Applied Soft Computing*, 118, pp. 108542.
- [38] Parvathy V.S., Pothiraj S., 2020, Multi-modality medical image fusion using hybridization of binary crow search optimization, *Health Care Management Science*, 23(4), pp. 661–669.
- [39] El-Hoseny H.M., Abd El-Rahman W., El-Rabaie E.S.M., Abd El-Samie F.E., Faragallah O.S., 2018, An efficient DT-CWT medical image fusion system based on modified central force optimization and histogram matching, *Infrared Physics and Technology*, 94, pp. 223–231.
- [40] Li B., Peng H., Wang J., 2021, A novel fusion method based on dynamic threshold neural P systems and nonsubsampling contourlet transform for multi-modality medical images, *Signal Processing*, 178, pp. 107793.
- [41] Wang Z., Li X., Duan H., Su Y., Zhang X., Guan X., 2021, Medical image fusion based on convolutional neural networks and non-subsampling contourlet transform, *Expert Systems with Applications*, 171, pp. 114574.
- [42] Ibrahim S.I., Makhoul M.A., El-Tawel G.S., 2022, Multimodal medical image fusion algorithm based on pulse coupled neural networks and nonsubsampling contourlet transform, *Medical and Biological Engineering and Computing*, 1, pp. 1–23.
- [43] Li X., Zhao J., 2021, A novel multi-modal medical image fusion algorithm, *Journal of Ambient Intelligence and Humanized Computing*, 12(2), pp. 1995–2002.
- [44] Rajalingam B., Al-Turjman F., Santhoshkumar R., Rajesh M., 2022, Intelligent multimodal medical image fusion with deep guided filtering, *Multimedia Systems*, 28, pp. 1449–1463.
- [45] Shilpa S., Ragesh Rajan M., Asha C.S., Shyam L., 2022, Enhanced JAYA optimization based medical image fusion in adaptive non subsampling shearlet

transform domain, *Engineering Science and Technology, an International Journal*, 35, pp. 101245.

- [46] Zhou J., Xing X., Yan M., Yuan D., Zhu C., Zhang C., et al., 2022, A fusion algorithm based on composite decomposition for PET and MRI medical images, *Biomedical Signal Processing and Control*, 76, pp. 103717.
- [47] Zhang Y., Jin M., Huang G., 2022, Medical image fusion based on improved multi-scale morphology gradient-weighted local energy and visual saliency map, *Biomedical Signal Processing and Control*, 74, pp. 103535.
- [48] Ullah H., Ullah B., Wu L., Abdalla F.Y.O., Ren G., Zhao Y., 2020, Multi-modality medical images fusion based on local-features fuzzy sets and novel sum-modified-Laplacian in non-subsampled shearlet transform domain, *Biomedical Signal Processing and Control*, 57, pp. 101724.
- [49] Liu Y., Wang L., Cheng J., Li C., Chen X., 2020, Multi-focus image fusion: A Survey of the state of the art, *Information Fusion*, 64, pp. 71–91.
- [50] Li X., Wan W., Zhou F., Cheng X., Jie Y., Tan H., 2023, Medical image fusion based on sparse representation and neighbor energy activity, *Biomedical Signal Processing and Control*, 80, pp. 104353.
- [51] Yang Y., Cao S., Wan W., Huang S., 2023, Multi-modal medical image super-resolution fusion based on detail enhancement and weighted local energy deviation, *Biomedical Signal Processing and Control*, 80, pp. 104387.
- [52] Hu Q., Hu S., Ma X., Zhang F., Fang J., 2022, MRI Image Fusion Based on Optimized Dictionary Learning and Binary Map Refining in Gradient Domain, *Multimedia Tools and Applications*, 82, pp. 2539–2561.
- [53] Jie Y., Zhou F., Tan H., Wang G., Cheng X., Li X., 2022, Tri-modal medical image fusion based on adaptive energy choosing scheme and sparse representation, *Measurement*, 204, pp. 112038.
- [54] Chen J., Zhang L., Lu L., Li Q., Hu M., Yang X., 2021, A novel medical image fusion method based on Rolling Guidance Filtering, *Internet of Things*, 14, pp. 100172.
- [55] Li Q., Wang W., Chen G., Zhao D., 2021, Medical image fusion using segment graph filter and sparse representation, *Computers in Biology and Medicine*, 131, pp. 104239.
- [56] Palanisami D., Mohan N., Ganeshkumar L., 2022, A new approach of multi-modal medical image fusion using intuitionistic fuzzy set, *Biomedical Signal Processing and Control*, 77, pp. 103762.

- [57] Zhang Y., Wang M., Xia X., Sun D., Zhou X., Wang Y., et al., 2023, Medical image fusion based on quasi-cross bilateral filtering, *Biomedical Signal Processing and Control*, 80, pp. 104259.
- [58] Goyal B., Dogra A., Lepcha D.C., Koundal D., Alhudhaif A., Alenezi F., et al., 2022, Multi-modality image fusion for medical assistive technology management based on hybrid domain filtering, *Expert Systems with Applications*, 209, pp. 118283.
- [59] Sabeeh Yousif A., Omar Z., Ullah Sheikh U., 2022, An improved approach for medical image fusion using sparse representation and Siamese convolutional neural network, *Biomedical Signal Processing and Control*, 72, pp. 103357.
- [60] Shibu D.S., Priyadharsini S.S., 2021, Multi scale decomposition based medical image fusion using convolutional neural network and sparse representation, *Biomedical Signal Processing and Control*, 69, pp. 102789.
- [61] Zhou J., Ren K., Wan M., Cheng B., Gu G., Chen Q., 2021, An infrared and visible image fusion method based on VGG-19 network, *Optik*, 248, pp. 168084.
- [62] Long J., Shelhamer E., Darrell T., 2015, Fully convolutional networks for semantic segmentation, *Computer Vision and Pattern Recognition*, 1, pp. 3431–3440.
- [63] Ding Z., Zhou D., Li H., Hou R., Liu Y., 2021, Siamese networks and multi-scale local extrema scheme for multimodal brain medical image fusion, *Biomedical Signal Processing and Control*, 68, pp. 102697.
- [64] Li H., Wu X.J., 2019, DenseFuse: A fusion approach to infrared and visible images, *IEEE Transactions on Image Processing*, 28(5), pp. 2614–2623.
- [65] Li W., Li R., Fu J., Peng X., 2022, MSNet: A multi-scale enhanced network based on unique features guidance for medical image fusion, *Biomedical Signal Processing and Control*, 74, pp. 103534.
- [66] Xu H., Ma J., 2021, EMFusion: An unsupervised enhanced medical image fusion network, *Information Fusion*, 76, pp. 177–186.
- [67] Fu J., Li W., Du J., Xu L., 2021, DSAGAN: A generative adversarial network based on dual-stream attention mechanism for anatomical and functional image fusion, *Information Sciences*, 576, pp. 484–506.
- [68] Kang J., Lu W., Zhang W., 2020, Fusion of Brain PET and MRI Images Using Tissue-Aware Conditional Generative Adversarial Network with Joint Loss, *IEEE Access*, 8, pp. 6368–6378.

- [69] Ma J., Xu H., Jiang J., Mei X., Zhang X.P., 2020, DDcGAN: A Dual-Discriminator Conditional Generative Adversarial Network for Multi-Resolution Image Fusion, *IEEE Transactions on Image Processing*, 29, pp. 4980–4995.
- [70] Guo K., Hu X., Li X., 2022, MMFGAN: A novel multimodal brain medical image fusion based on the improvement of generative adversarial network, *Multimedia Tools and Applications*, 81(4), pp. 5889–5927.
- [71] Subramani B., Veluchamy M., 2020, Fuzzy Gray Level Difference Histogram Equalization for Medical Image Enhancement, *Journal of Medical Systems*, 44(6).
- [72] Magudeeswaran V., Ravichandran C.G., Thirumurugan P., 2017, Brightness preserving bi-level fuzzy histogram equalization for MRI brain image contrast enhancement, *International Journal of Imaging Systems and Technology*, 27(2), pp. 153–161.
- [73] Subramani B., Veluchamy M., 2018, MRI brain image enhancement using brightness preserving adaptive fuzzy histogram equalization, *International Journal of Imaging Systems and Technology*, 28(3), pp. 217–222.
- [74] Chen J., Yu W., Tian J., Chen L., Zhou Z., 2018, Image contrast enhancement using an artificial bee colony algorithm, *Swarm and Evolutionary Computation*, 38, pp. 287–294.
- [75] Kamoona A.M., Patra J.C., 2019, A novel enhanced cuckoo search algorithm for contrast enhancement of gray scale images, *Applied Soft Computing Journal*, 85, pp. 105749.
- [76] Rao B.S., 2020, Dynamic Histogram Equalization for contrast enhancement for digital images, *Applied Soft Computing Journal*, 89, pp. 106114.
- [77] Acharya U.K., Kumar S., 2020, Particle swarm optimized texture based histogram equalization (PSOTHE) for MRI brain image enhancement, *Optik*, 224, pp. 165760.
- [78] Pashaei E., Pashaei E., 2023, Gaussian quantum arithmetic optimization-based histogram equalization for medical image enhancement, *Multimedia Tools and Applications*, 82(22).
- [79] Liu Y., Chen X., Ward R.K., Wang J., 2016, Image Fusion with Convolutional Sparse Representation, *IEEE Signal Processing Letters*, 23(12), pp. 1882–1886.
- [80] Babu B.S., Narayana M.V., 2023, Two stage multi-modal medical image fusion with marine predator algorithm-based cascaded optimal DTCWT and

- NSST with deep learning, *Biomedical Signal Processing and Control*, 85, pp. 104921.
- [81] Rajwar K., Deep K., Das S., 2023, An exhaustive review of the metaheuristic algorithms for search and optimization: taxonomy, applications, and open challenges, *Artificial Intelligence Review*, 56(11), pp. 13187–13257.
 - [82] Mirjalili S., Mirjalili S.M., Hatamlou A., 2016, Multi-Verse Optimizer: a nature-inspired algorithm for global optimization, *Neural Computing and Applications*, 27(2), pp. 495–513.
 - [83] Mirjalili S., Lewis A., 2016, The Whale Optimization Algorithm, *Advances in Engineering Software*, 95, pp. 51–67.
 - [84] Mirjalili S., Gandomi A.H., Mirjalili S.Z., Saremi S., Faris H., Mirjalili S.M., 2017, Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems, *Advances in Engineering Software*, 114, pp. 163–191.
 - [85] Mirjalili S., 2016, SCA: A Sine Cosine Algorithm for solving optimization problems, *Knowledge-Based Systems*, 96, pp. 120–133.
 - [86] Mirjalili S., Mirjalili S.M., Lewis A., 2014, Grey Wolf Optimizer, *Advances in Engineering Software*, 69, pp. 46–61.
 - [87] Mirjalili S., 2016, Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems, *Neural Computing and Applications*, 27, pp. 1053–1073.
 - [88] Mirjalili S., 2015, The Ant Lion Optimizer, *Advances in Engineering Software*, 83, pp. 80–98.
 - [89] Dinh P.H., 2022, An improved medical image synthesis approach based on marine predators algorithm and maximum Gabor energy, *Neural Computing and Applications*, 34(6), pp. 4367–4385.
 - [90] Dinh P.H., 2023, A Novel Approach Based on Marine Predators Algorithm for Medical Image Enhancement, *Sensing and Imaging*, 24(1), pp. 1–23.
 - [91] Srivastava R., Prakash O., Khare A., 2016, Local energy-based multimodal medical image fusion in curvelet domain, *IET Computer Vision*, 10(6), pp. 513–527.
 - [92] Polinati S., Dhuli R., 2020, Multimodal medical image fusion using empirical wavelet decomposition and local energy maxima, *Optik*, 205, pp. 163947.
 - [93] Dinh P.H., 2021, Multi-modal medical image fusion based on equilibrium optimizer algorithm and local energy functions, *Applied Intelligence*, 51, pp. 8416–8431.

- [94] Ren Z., Guan P., Lam E.Y., Zhao J., 2021, Extended focused imaging in microscopy using structure tensor and guided filtering, *Optics and Lasers in Engineering*, 140, pp. 106549.
- [95] Wang M., Zhang W., Sun C., Sowmya A., 2020, Corner detection based on shearlet transform and multi-directional structure tensor, *Pattern Recognition*, 103, pp. 107299.
- [96] Zhou Z., Li S., Wang B., 2014, Multi-scale weighted gradient-based fusion for multi-focus images, *Information Fusion*, 20(1), pp. 60–72.
- [97] Zhang Q., Shen X., Xu L., Jia J., 2014, Rolling guidance filter, *European Conference on Computer Vision (ECCV)*, 8691, pp. 815–830.
- [98] Gong Y., Goksel O., 2019, Weighted mean curvature, *Signal Processing*, 164, pp. 329–339.
- [99] Tan W., Thitøn W., Xiang P., Zhou H., 2021, Multi-modal brain image fusion based on multi-level edge-preserving filtering, *Biomedical Signal Processing and Control*, 64, pp. 102280.
- [100] Zuiderveld K., 1994, Contrast Limited Adaptive Histogram Equalization, *Graphics Gems*, 4, pp. 474–485.
- [101] Perona P., Malik J., 1990, Scale-Space and Edge Detection Using Anisotropic Diffusion, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), pp. 629–639.
- [102] Richardson W.H., 1972, Bayesian-Based Iterative Method of Image Restoration*, *Journal of the Optical Society of America*, 62(1), pp. 55–59.
- [103] Mesarović V.Z., Katsaggelos A.K., 1995, Regularized Constrained Total Least Squares Image Restoration, *IEEE Transactions on Image Processing*, 4(8), pp. 1096–1108.
- [104] Chan S.H., Khoshabeh R., Gibson K.B., Gill P.E., Nguyen T.Q., 2011, An augmented Lagrangian method for total variation video restoration, *IEEE Transactions on Image Processing*, 20(11), pp. 3097–3111.
- [105] Xydeas C.S., Petrović V., 2000, Objective image fusion performance measure, *Electronics Letters*, 36(4), pp. 308–309.
- [106] Qu G., Zhang D., Yan P., 2002, Information measure for performance of image fusion, *Electronics Letters*, 38(7), pp. 313–315.
- [107] Haghighat M.B.A., Aghagolzadeh A., Seyedarabi H., 2011, A non-reference image fusion metric based on mutual information of image features, *Computers & Electrical Engineering*, 37(5), pp. 744–756.

- [108] Haghghat M.B.A., Aghagolzadeh A., Seyedarabi H., 2011, A non-reference image fusion metric based on mutual information of image features, *Computers and Electrical Engineering*, 37(5), pp. 744–756.
- [109] Lim J., Heo M., Lee C., Kim C.S., 2017, Contrast enhancement of noisy low-light images based on structure-texture-noise decomposition, *Journal of Visual Communication and Image Representation*, 45, pp. 107–121.
- [110] Wang Y., Yang J., Yin W., Zhang Y., 2008, A new alternating minimization algorithm for total variation image reconstruction, *SIAM Journal on Imaging Sciences*, 1(3), pp. 248–272.
- [111] Dabov K., Foi A., Katkovnik V., Egiazarian K., 2007, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Transactions on Image Processing*, 16(8), pp. 2080–2095.
- [112] Singh H., Kumar A., Balyan L.K., Singh G.K., 2018, Swarm intelligence optimized piecewise gamma corrected histogram equalization for dark image enhancement, *Computers & Electrical Engineering*, 70, pp. 462–475.
- [113] Singh H., Kumar A., Balyan L.K., Singh G.K., 2019, A novel optimally weighted framework of piecewise gamma corrected fractional order masking for satellite image enhancement, *Computers & Electrical Engineering*, 75, pp. 245–261.
- [114] Frank Wilcoxon, 2010, Probability Tables for Individual Comparisons by Ranking Methods, *International Biometric Society Stable*, 3(3), pp. 119–122.
- [115] Parihar A.S., Verma O.P., Khanna C., 2017, Fuzzy-Contextual Contrast Enhancement, *IEEE Transactions on Image Processing*, 26(4), pp. 1810–1819.
- [116] Ying Z., Li G., Ren Y., Wang R., Wang W., 2017, A new image contrast enhancement algorithm using exposure fusion framework, *Computer Analysis of Images and Patterns*, 10425, pp. 36–46.
- [117] Lu Z., Long B., Li K., Lu F., 2018, Effective Guided Image Filtering for Contrast Enhancement, *IEEE Signal Processing Letters*, 25(10), pp. 1585–1589.
- [118] Li M., Liu J., Yang W., Sun X., Guo Z., 2018, Structure-Revealing Low-Light Image Enhancement Via Robust Retinex Model, *IEEE Transactions on Image Processing*, 27(6), pp. 2828–2841.
- [119] Dai Q., Pu Y.F., Rahman Z., Aamir M., 2019, Fractional-order fusion model for low-light image enhancement, *Symmetry*, 11(4), pp. 574.

- [120] Hao S., Han X., Guo Y., Xu X., Wang M., 2020, Low-Light Image Enhancement with Semi-Decoupled Decomposition, *IEEE Transactions on Multimedia*, 22(12), pp. 3025–3038.
- [121] Zhu Z., Zheng M., Qi G., Wang D., Xiang Y., 2019, A Phase Congruency and Local Laplacian Energy Based Multi-Modality Medical Image Fusion Method in NSCT Domain, *IEEE Access*, 7, pp. 20811–20824.
- [122] Tan W., Zhang J., Xiang P., Zhou H., Thitøn W., 2020, Infrared and visible image fusion via NSST and PCNN in multiscale morphological gradient domain, *Optics, Photonics and Digital Technologies for Imaging Applications VI*, 11353, pp. 1–7.
- [123] Sufyan A., Imran M., Shah S.A., Shahwani H., Wadood A.A., 2022, A novel multimodality anatomical image fusion method based on contrast and structure extraction, *International Journal of Imaging Systems and Technology*, 32(1), pp. 324–342.
- [124] Li B., Peng H., Luo X., Wang J., Song X., Pérez-Jiménez M.J., et al., 2021, Medical Image Fusion Method Based on Coupled Neural P Systems in Nonsubsampled Shearlet Transform Domain, *International Journal of Neural Systems*, 31(1), pp. 2050050.
- [125] Dong X., Wang G., Pang Y., Li W., Wen J., Meng W., et al., 2011, Fast efficient algorithm for enhancement of low lighting video, *IEEE International Conference on Multimedia and Expo*, 1, pp. 1–6.
- [126] Lu H., Zhang L., Serikawa S., 2012, Maximum local energy: An effective approach for multisensor image fusion in beyond wavelet transform domain, *Computers & Mathematics with Applications*, 64(5), pp. 996–1003.
- [127] Yin M., Liu X., Liu Y., Chen X., 2019, Medical Image Fusion with Parameter-Adaptive Pulse Coupled Neural Network in Nonsubsampled Shearlet Transform Domain, *IEEE Transactions on Instrumentation and Measurement*, 68(1), pp. 49–64.
- [128] Zhu R., Li X., Huang S., Zhang X., 2022, Multimodal medical image fusion using adaptive co-occurrence filter-based decomposition optimization model, *Bioinformatics*, 38(3), pp. 818–826.
- [129] Braik M.S., 2021, Chameleon Swarm Algorithm: A bio-inspired optimizer for solving engineering design problems, *Expert Systems with Applications*, 174, pp. 114685.

- [130] Braik M., Hammouri A., Atwan J., Al-Betar M.A., Awadallah M.A., 2022, White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems, *Knowledge-Based Systems*, 243, pp. 108457.
- [131] Dehghani M., Montazeri Z., Trojovská E., Trojovský P., 2023, Coati Optimization Algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems, *Knowledge-Based Systems*, 259, pp. 110011.