

MDMFUSION: MÔ HÌNH TỔNG HỢP HÌNH ẢNH MRI VÀ CT DỰA TRÊN PHÂN RÃ ĐA MỨC VÀ DATFUSE

MDMFUSION: AN MRI-CT IMAGE FUSION MODEL BASED ON MULTI-LEVEL DECOMPOSITION AND DATFUSE

Lê Thị Hồng Hà¹,
Nguyễn Thế Cường¹,
Trần Doãn Minh¹,
Nguyễn Văn Cường¹,
Nguyễn Hoàng Long¹

¹ Khoa Kỹ thuật Công nghệ, Trường Đại học Hồng Đức;
*Tác giả liên hệ; Email: lethihongha@hdu.edu.vn

DOI: <https://doi.org/10.70117/hdujs.3030-4628.90.09.2026.1383>

Ngày nộp bài: 01/05/2026; Ngày gửi phản biện: 06/05/2026; Ngày duyệt đăng: 25/09/2026

Từ khóa:

CT, DATFuse, MRI, phân rã đa mức, tổng hợp hình ảnh y học.

TÓM TẮT

Tổng hợp ảnh y học nhằm kết hợp thông tin bổ sung từ các phương thức chụp khác nhau vào một ảnh duy nhất để phục vụ chẩn đoán. Bài báo đề xuất MDMFusion, một mô hình tổng hợp ảnh MRI-CT dựa trên phân rã đa mức MIDTE (Multi-level Image Decomposition with THF and EGIF) kết hợp hai bộ lọc bảo toàn biên là THF (Truncated Huber Filtering) và EGIF (Effective Guided Image Filtering). Thành phần cơ sở được tổng hợp bằng mô hình học sâu DATFuse, còn các thành phần chi tiết được tổng hợp bằng mạng nơ-ron xung cải tiến mPCNN.

Thực nghiệm trên bộ dữ liệu MRI-CT với bốn chỉ số Q_{MLI} , Q_E , Q_{MI} và $Q^{AB/F}$, so sánh với chín phương pháp tiên tiến được công bố trong giai đoạn 2020–2025, cho thấy MDMFusion đạt Q_{MLI} và Q_E cao nhất (lần lượt là 0,27243 và 6,48848), đồng thời đạt kết quả cạnh tranh ở Q_{MI} (2,54383) và $Q^{AB/F}$ (0,48593).

Kết quả cho thấy việc kết hợp phân rã đa mức bảo toàn biên với DATFuse cho thành phần cơ sở và mPCNN cho thành phần chi tiết giúp cải thiện độ sáng và lượng thông tin của ảnh tổng hợp; khả năng bảo toàn thông tin biên ($Q^{AB/F}$) vẫn còn thấp hơn một số phương pháp khác và là hướng cần cải thiện.

Keywords:

MRI-CT, DATFuse, PCNN, THF, EGIF.

ABSTRACT

Medical image fusion combines complementary information from different imaging modalities into a single image to support diagnosis. This paper proposes MDMFusion, an MRI-CT image fusion model based on the multi-level decomposition MIDTE (Multi-level Image

Decomposition with THF and EGIF), which combines two edge-preserving filters: Truncated Huber Filtering (THF) and Effective Guided Image Filtering (EGIF). The base components are fused by the deep learning model DATFuse, and the detail components are fused by a modified pulse-coupled neural network (mPCNN). Experiments on an MRI-CT dataset with four metrics (Q_{MLI} , Q_E , Q_{MI} and $Q^{AB/F}$), compared with nine state-of-the-art methods published in 2020–2025, show that MDMFusion achieves the highest Q_{MLI} (0.27243) and Q_E (6.48848) while remaining competitive in Q_{MI} (2.54383) and $Q^{AB/F}$ (0.48593). These findings suggest that combining edge-preserving multi-level decomposition with DATFuse for the base component and mPCNN for the detail components improves the brightness and information content of fused images; edge-information preservation ($Q^{AB/F}$) remains lower than some existing methods and is a direction for further improvement.

1. ĐẶT VẤN ĐỀ

Tổng hợp hình ảnh y học là một bài toán quan trọng trong lĩnh vực xử lý ảnh. Mục tiêu của bài toán là kết hợp thông tin từ nhiều ảnh đầu vào, thường được thu nhận từ các phương thức chụp khác nhau, nhằm tạo ra một ảnh tổng hợp chứa nhiều thông tin hữu ích hơn so với từng ảnh riêng lẻ. Thông qua quá trình này, các đặc trưng bổ sung giữa các nguồn dữ liệu được tích hợp hiệu quả, giúp biểu diễn đầy đủ hơn cả cấu trúc giải phẫu và hoạt động chức năng của cơ thể người.

Trong các bài toán tổng hợp ảnh y học, tổng hợp ảnh chụp cắt lớp vi tính (Computed Tomography - CT) và cộng hưởng từ (Magnetic Resonance Imaging - MRI) được xem là một hướng nghiên cứu quan trọng nhờ khả năng kết hợp hiệu quả thông tin cấu trúc và thông tin mô mềm. Ảnh CT có ưu thế trong việc biểu diễn cấu trúc xương, mô cứng và các chi tiết giải phẫu với độ phân giải không gian cao, đồng thời hiển thị rõ các tổn thương liên quan đến xương, xuất huyết, vôi hóa và cấu trúc giải phẫu tổng thể. Trong khi đó, ảnh MRI nổi bật ở khả năng mô tả mô mềm, não bộ, tủy sống, dây thần kinh và các tổn thương mô mềm nhờ độ tương phản mô cao. Do mỗi phương thức chỉ phản ánh một phần thông tin của cơ thể, việc tổng hợp ảnh CT và MRI cho phép tạo ra một ảnh duy nhất chứa đồng thời thông tin cấu trúc rõ nét từ CT và thông tin mô mềm chi tiết từ MRI, giúp bác sĩ quan sát toàn diện hơn vị trí, kích thước cũng như mối liên quan của tổn thương với các cấu trúc giải phẫu lân cận, từ đó nâng cao hiệu quả đánh giá và chẩn đoán bệnh lý. Trong thực tiễn lâm sàng, ảnh tổng hợp CT-MRI có nhiều ứng dụng quan trọng như hỗ trợ chẩn đoán chính xác hơn nhờ cung cấp đồng thời thông tin

về cấu trúc giải phẫu và mô mềm, hỗ trợ lập kế hoạch phẫu thuật thông qua việc xác định chính xác vị trí tổn thương và mức độ xâm lấn, đồng thời nâng cao hiệu quả xạ trị và điều trị ung thư khi giúp xác định rõ vùng khối u và mô lành xung quanh. Ngoài ra, việc tích hợp nhiều nguồn thông tin trên cùng một ảnh còn giúp tăng hiệu quả quan sát, giảm sai sót trong quá trình đọc ảnh và cung cấp dữ liệu đầu vào giàu thông tin cho các hệ thống trí tuệ nhân tạo và học sâu trong phân tích ảnh y học. Đặc biệt, trong các ứng dụng liên quan đến chẩn đoán thần kinh, ung thư và các bệnh lý phức tạp, tổng hợp ảnh CT và MRI góp phần cải thiện đáng kể chất lượng thông tin hình ảnh, qua đó nâng cao độ chính xác trong chẩn đoán và hiệu quả điều trị lâm sàng.

Mặc dù đã có nhiều phương pháp tổng hợp hình ảnh y học được công bố và đạt được những kết quả tích cực, các nghiên cứu hiện nay vẫn tồn tại một số hạn chế cần tiếp tục được khắc phục. Trước hết, các phương pháp phân rã đa tỉ lệ như NSCT và NSST [1][2][3] có khả năng biểu diễn đặc trưng tốt nhờ khai thác thông tin đa hướng và đa tỉ lệ, tuy nhiên lại đòi hỏi chi phí tính toán lớn, gây khó khăn cho các ứng dụng thời gian thực hoặc các hệ thống có tài nguyên tính toán hạn chế. Bên cạnh đó, trong chiến lược tổng hợp thành phần cơ sở, một số phương pháp áp dụng các quy tắc đơn giản như quy tắc trung bình [4][5] do có ưu điểm về tốc độ tính toán và dễ triển khai. Tuy nhiên, phương pháp này dễ làm suy giảm cường độ sáng và độ tương phản của ảnh tổng hợp, dẫn đến mất mát thông tin chi tiết, đặc biệt tại các vùng chứa nhiều thông tin quan trọng hoặc có mức năng lượng cao. Ngoài ra, đối với chiến lược tổng hợp thành phần chi tiết, các phương pháp dựa trên mạng nơ-ron xung PCNN (Pulse-Coupled Neural Network), điển hình như AGPCNN (Adaptive Gaussian PCNN) [6] vẫn chưa bảo toàn đầy đủ thông tin biên và các đặc trưng cấu trúc quan trọng, đặc biệt tại các vùng có kết cấu phức tạp hoặc độ tương phản thấp. Điều này có thể dẫn đến hiện tượng mất các chi tiết nhỏ.

Trong bài báo này, chúng tôi đề xuất MDMFusion, một mô hình tổng hợp hình ảnh y học mới, các đóng góp chính của nghiên cứu được tóm tắt như sau:

Đề xuất mô hình tổng hợp ảnh y học MDMFusion (Multi-level Decomposition, DATFuse and modified PCNN Fusion) kết hợp phân rã đa mức, học sâu và mạng nơ-ron xung nhằm nâng cao chất lượng ảnh MRI-CT tổng hợp.

Đề xuất giải thuật phân rã đa mức MIDTE (Multi-level Image Decomposition with THF and EGIF) kết hợp THF (Truncated Huber Filtering) [7] và EGIF (Effective Guided Image Filtering) [8] để tăng khả năng bảo toàn biên và tách hiệu quả thành phần cơ sở - chi tiết.

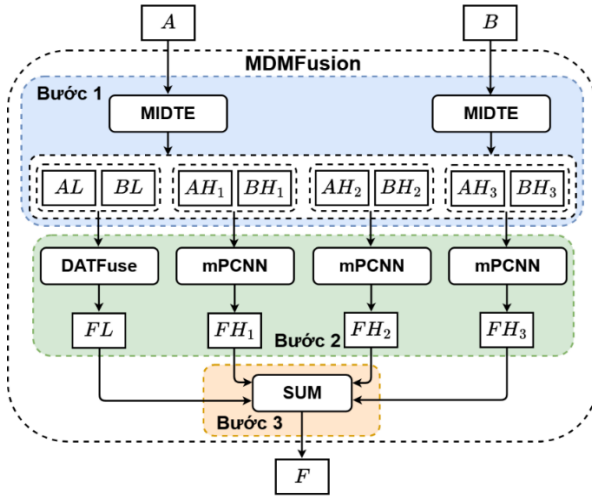
Áp dụng mô hình DATFuse (Dual Attention Transformer) [9] cho tổng hợp thành phần cơ sở nhằm cải thiện độ sáng và độ tương phản của ảnh tổng hợp.

Đề xuất chiến lược tổng hợp thành phần chi tiết dựa trên mPCNN (Modified PCNN) nhằm tăng cường khả năng bảo toàn thông tin biên và cấu trúc giải phẫu quan trọng.

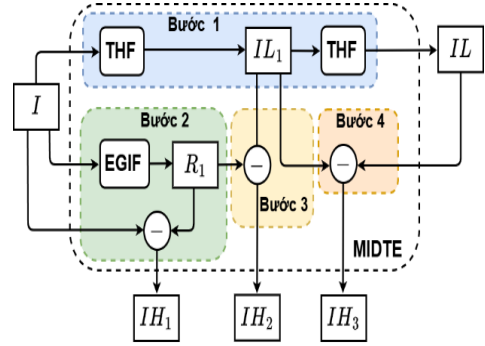
Cấu trúc của bài báo được tổ chức như sau: Phần 2 mô tả chi tiết phương pháp tổng hợp ảnh được đề xuất; Phần 3 phân tích kết quả thực nghiệm và thảo luận; cuối cùng, Phần 4 đưa ra kết luận và định hướng nghiên cứu trong tương lai.

2. MÔ HÌNH ĐỀ XUẤT

2.1. Mô hình tổng hợp hình ảnh MDMFusion



Hình 1. Mô hình tổng hợp hình ảnh MDMFusion



Hình 2. Giải thuật MIDTE

Trong phần này, chúng tôi trình bày MDMFusion, một phương pháp tổng hợp ảnh y học mới được đề xuất cho bài toán tổng hợp ảnh MRI và CT. Khung tổng thể của phương pháp được minh họa trong Hình 1, bao gồm ba giai đoạn chính: phân rã đa mức bằng MIDTE, tổng hợp các thành phần cơ sở và chi tiết thông qua DATFuse và mPCNN, và cuối cùng là tái tạo ảnh. Chi tiết mô hình MDMFusion được mô tả như sau:

Đầu vào: Hai ảnh xám A và B .

Đầu ra: Ảnh tổng hợp xám F .

Bước 1. Phân rã ảnh: Áp dụng phương pháp MIDTE để phân rã hai ảnh đầu vào A và B thành các thành phần cơ sở (AL, BL) và các thành phần chi tiết (AH_1, AH_2, AH_3 và BH_1, BH_2, BH_3), như mô tả trong các công thức (1) và (2).

$$[AL, AH_1, AH_2, AH_3] = \text{MIDTE}(A), \quad (1)$$

$$[BL, BH_1, BH_2, BH_3] = \text{MIDTE}(B). \quad (2)$$

Bước 2. Tổng hợp ảnh:

Tổng hợp hai thành phần cơ sở AL và BL bằng phương pháp DATFuse như trong công thức (3).

$$FL = \text{DATFuse}(AL, BL). \quad (3)$$

Tổng hợp từng cặp thành phần chi tiết bằng phương pháp mPCNN, như được trình bày trong công thức (4) với $i = \overline{1, 2, 3}$.

$$FH_i = \text{mPCNN}(AH_i, BH_i), \quad (4)$$

Bước 3. Tái tạo ảnh: Ảnh tổng hợp cuối cùng F được tái tạo bằng cách kết hợp các thành phần cơ sở và chi tiết đã tổng hợp, như được định nghĩa trong công thức (5).

$$F = \text{SUM}(FL, FH_1, FH_2, FH_3). \quad (5)$$

2.2. Giải thuật phân rã đa mức MIDTE

Thuật toán MIDTE được đề xuất dựa trên chiến lược phân rã ảnh đa mức nhằm tách hiệu quả thông tin cơ sở và thông tin chi tiết từ các ảnh đầu vào. Ý tưởng chính của phương pháp xuất phát từ thực tế rằng trong bài toán tổng hợp ảnh y học, việc bảo toàn đồng thời cấu trúc giải phẫu và các chi tiết biên là yếu tố quan trọng quyết định chất lượng ảnh tổng hợp. Do đó, nghiên cứu kết hợp hai bộ lọc THF và EGIF nhằm khai thác tính bổ sung giữa khả năng bảo toàn biên và tăng cường thông tin cấu trúc. Trước hết, THF được lựa chọn nhờ khả năng làm trơn hiệu quả các vùng đồng nhất trong khi vẫn duy trì tốt các đường biên quan trọng. Bộ lọc này giúp giảm nhiễu và tạo ra lớp cơ sở ổn định cho quá trình phân rã. Tuy nhiên, khi chỉ sử dụng THF, một số thông tin cấu trúc và chi tiết cục bộ có thể chưa được biểu diễn đầy đủ. Vì vậy, EGIF tiếp tục được áp dụng để tinh chỉnh lớp cơ sở trung gian và tăng cường khả năng trích xuất thông tin cấu trúc. Việc kết hợp THF và EGIF giúp tạo ra sự phân tách cân bằng hơn giữa thành phần cơ sở và thành phần chi tiết, đồng thời cải thiện khả năng bảo toàn thông tin biên và kết cấu ảnh. Nhìn chung, MIDTE cung cấp một khung phân rã đa mức có khả năng cân bằng hiệu quả giữa khử nhiễu và bảo toàn chi tiết so với các phương pháp phân rã truyền thống. Thuật toán MIDTE được minh họa trong Hình 2.

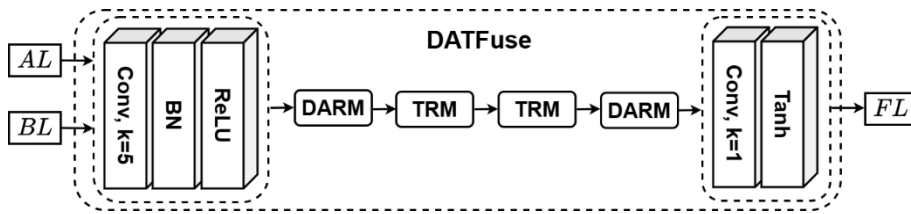
2.3. Giải thuật tổng hợp thành phần cơ sở DATFuse

Tang và cộng sự [9] đã đề xuất DATFuse, một khung tổng hợp ảnh hồng ngoại – khả kiến đầu cuối kết hợp giữa mạng tích chập CNN và mô hình Transformer nhằm khai thác đồng thời thông tin cục bộ và thông tin toàn cục của ảnh. Ý tưởng chính của DATFuse là tận dụng khả năng trích xuất đặc trưng cục bộ hiệu quả của CNN cùng khả năng mô hình hóa các mối quan hệ phụ thuộc dài hạn của Transformer để nâng cao chất lượng ảnh tổng hợp. Cụ thể, mô hình sử dụng các mô-đun dư chú ý kép (Dual Attention Residual Modules - DARMs) nhằm tăng cường khả năng biểu diễn đặc trưng cục bộ, đồng thời áp dụng các mô-đun Transformer (Transformer Modules - TRMs) để học thông tin ngữ cảnh toàn cục và các quan hệ phụ thuộc không gian xa. Nhờ sự kết hợp này, DATFuse có thể tạo ra các ảnh tổng hợp bảo toàn cường độ sáng tốt hơn so với nhiều phương pháp truyền thống.

Trong mô hình này, DATFuse được áp dụng để tổng hợp hai thành phần cơ sở AL và BL . Như minh họa trong Hình 3, khung phương pháp bắt đầu bằng việc ghép hai ảnh thành phần cơ sở AL và BL theo chiều kênh để tạo thành đầu vào hai kênh F_{cat} . Đầu vào này sau đó được xử lý bởi một khối tích chập ban đầu gồm phép tích chập 5×5 , theo sau là chuẩn hóa batch (Batch Normalization - BN) và hàm kích hoạt ReLU nhằm trích xuất các đặc trưng mức thấp, hỗ trợ quá trình xử lý thị giác ban đầu và tối ưu hóa ổn định, như được định nghĩa trong công thức (6).

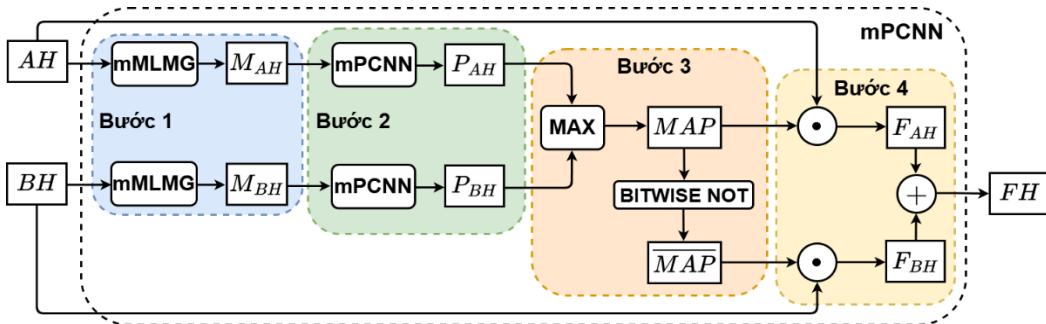
$$F_{in} = \text{ReLU}_{BN}(\text{Conv}_{5 \times 5}(F_{cat})) \quad (6)$$

trong đó: $F_{cat} = \text{Concat}(AL, BL)$. Các đặc trưng được trích xuất sau đó lần lượt được tinh chỉnh thông qua một DARM, hai TRM và một DARM thứ hai, trước khi được ánh xạ thông qua phép tích chập 1×1 kết hợp với hàm kích hoạt $Tanh$ để tái tạo thành phần cơ sở cuối cùng FL .



Hình 3. Giải thuật tổng hợp thành phần cơ sở DATFuse

2.4. Giải thuật tổng hợp thành phần chi tiết mPCNN



Hình 4. Giải thuật tổng hợp thành phần chi tiết mPCNN

Mặc dù sự kết hợp giữa PCNN truyền thống và MLMG (Multi-Level Morphological Gradient) [10][11] đã cho thấy hiệu quả trong phát hiện biên và bảo toàn cấu trúc, phương pháp này vẫn tồn tại một số hạn chế như chi phí tính toán cao do sử dụng nhân liên kết kích thước lớn, hiện tượng dao động tại các vùng đồng nhất và sự phụ thuộc mạnh vào phần tử cấu trúc trong quá trình tính gradient. Những hạn chế này có thể làm giảm độ ổn định cũng như hiệu quả bảo toàn chi tiết trong ảnh tổng hợp. Để khắc phục các vấn đề trên, nghiên cứu đề xuất chiến lược tổng hợp thành phần chi tiết kết hợp giữa MLMG cải tiến (modified MLMG - mMLMG) và PCNN cải tiến (modified PCNN - mPCNN). Ý tưởng chính của chiến lược này là duy trì khả năng phát hiện đặc trưng nổi bật của PCNN đồng thời giảm độ phức tạp tính toán, tăng tính ổn định và nâng cao hiệu quả trích xuất chi tiết. Những cải tiến trên góp phần nâng cao hiệu quả và độ ổn định của PCNN trong khi vẫn duy trì khả năng biểu diễn các chi tiết tinh. Chiến lược này tích hợp mMLMG và mPCNN thành quy trình ba bước cho bài toán tổng hợp thành phần chi tiết, được minh họa trong Hình 4.

Ba cải tiến chính được đề xuất gồm:

mMLMG với phần tử cấu trúc hình vuông: Khác với MLMG truyền thống sử dụng phần tử cấu trúc dạng đĩa nhằm đảm bảo tính đẳng hướng, mMLMG sử dụng phần tử cấu trúc hình vuông. Điều này giúp giảm chi phí tính toán, tăng độ nhạy theo các hướng trục của ảnh và thuận tiện hơn trong việc điều chỉnh tỉ lệ.

mPCNN với nhân liên kết nhị thức: Thay vì sử dụng nhân suy giảm theo khoảng cách có kích thước lớn, mPCNN áp dụng nhân nhị thức kích thước 3x3. Nhân này có khả năng xấp xỉ quá trình làm trơn Gaussian, giúp giảm đáng kể độ phức tạp tính toán và ổn định quá trình kích hoạt trong các vùng đồng nhất mà vẫn bảo toàn tốt thông tin biên.

Cơ chế tính lân cận ổn định: Kích thích lân cận được tính thông qua phép tích chập kết hợp với cơ chế sao chép biên, giúp hạn chế hiện tượng biên giả do đệm giá trị 0, duy trì tính nhất quán không gian tại vùng biên và tạo ra các bản đồ hoạt động mượt hơn cho quá trình tổng hợp.

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Thiết lập thực nghiệm

3.1.1. Chỉ số đánh giá

Để đánh giá khách quan chất lượng ảnh tổng hợp, nghiên cứu sử dụng bốn chỉ số phổ biến gồm Q_{MLI} , Q_E , Q_{MI} và $Q^{AB/F}$.

Q_{MLI} (Mean Luminance Index) là chỉ số đánh giá khả năng bảo toàn cường độ sáng của ảnh tổng hợp so với các ảnh nguồn. Giá trị càng lớn cho thấy ảnh tổng hợp duy trì độ sáng càng tốt.

Q_E (Entropy) phản ánh lượng thông tin chứa trong ảnh tổng hợp. Giá trị entropy càng cao cho thấy ảnh chứa càng nhiều thông tin và chi tiết.

Q_{MI} (Mutual Information) [12] đo lượng thông tin tương hỗ giữa ảnh tổng hợp và các ảnh nguồn, qua đó đánh giá khả năng lưu giữ thông tin từ ảnh đầu vào trong quá trình tổng hợp.

$Q^{AB/F}$ [13] là chỉ số đánh giá khả năng bảo toàn thông tin biên và cạnh của ảnh tổng hợp dựa trên mức độ duy trì gradient và cấu trúc biên từ các ảnh nguồn.

3.1.2. Mô tả thực nghiệm

Thực nghiệm 1: Thực nghiệm được thực hiện nhằm đánh giá hiệu quả của giải thuật tổng hợp thành phần cơ sở DATFuse trong mô hình MDMFusion trên bộ dữ liệu MRI-CT. Hai chỉ số đánh giá được sử dụng gồm: Q_{MLI} và Q_E . Cụ thể, DATFuse được thay thế lần lượt bởi các quy tắc tổng hợp thành phần cơ sở, bao gồm: quy tắc dựa trên toán tử Robert cải tiến [6], CNP [14], INSML+WLE [1], toán tử Prewitt cải tiến [15]. Các biến thể tương ứng được ký hiệu từ T_1 đến T_4 .

Thực nghiệm 2: Thực nghiệm được tiến hành nhằm đánh giá hiệu quả của mô hình MDMFusion trên bộ dữ liệu MRI-CT. Bốn chỉ số đánh giá được sử dụng gồm: Q_{MLI} , Q_E , Q_{MI} và $Q^{AB/F}$. Mô hình đề xuất được so sánh với chín phương pháp tổng hợp ảnh tiên tiến được công bố trong giai đoạn 2020-2025 gồm: CNP_NSST [14], MLMG_PCNN [10], DTNP_NSCT [1], AGPCNN [6], IDCPCNN [15], DCRNN [16], FATFusion [17], MDHU [18] và MACTFusion [19]. Các phương pháp này được ký hiệu lần lượt từ P_1 đến P_9 .

3.1.3. Môi trường thực nghiệm

DATFuse, FATFusion và MACTFusion được triển khai bằng PyTorch và huấn luyện trên GPU NVIDIA A100. Toàn bộ ảnh MRI-CT được lấy từ bộ dữ liệu công khai Harvard Whole Brain Atlas (AANLIB)¹. Các mô hình học sâu được huấn luyện trên bộ

¹ <https://www.med.harvard.edu/aanlib/>

dữ gồm 184 cặp ảnh MRI-CT, không trùng lặp với 24 cặp ảnh sử dụng cho đánh giá. Trong quá trình huấn luyện, kích thước batch được thiết lập bằng 8, số epoch bằng 10 và toàn bộ ảnh được chuẩn hóa về kích thước 256×256 pixel. Bộ tối ưu Adam được sử dụng với tốc độ học khởi tạo 1×10^{-3} . Các thí nghiệm còn lại được thực hiện trên MATLAB R2022b với máy tính sử dụng CPU Intel® Core™ i5-12400 thế hệ thứ 12 (2.50 GHz) và RAM 16 GB. Các tham số của mô hình được lựa chọn theo công bố gốc hoặc được xác định thông qua các thực nghiệm sơ bộ trước khi tiến hành đánh giá chính thức.

3.2. Đánh giá kết quả thực nghiệm

3.2.1. Kết quả thực nghiệm 1

Kết quả thực nghiệm 1 được trình bày trong Bảng 1 với hai chỉ số đánh giá là Q_{MLI} và Q_E . Quan sát cho thấy DATFuse (trong mô hình MDMFusion) đạt kết quả cao nhất ở cả hai chỉ số, với $Q_{MLI} = 0,27243$ và $Q_E = 6,48848$, vượt trội so với các phương pháp T_1 – T_4 . Trong nhóm các phương pháp so sánh, T_1 có Q_{MLI} thấp nhất trong nhóm, với $Q_{MLI} = 0,19546$ và $Q_E = 4,28925$, trong khi các phương pháp T_2 , T_3 và T_4 có giá trị tương đối gần nhau. Cụ thể, T_4 đạt $Q_{MLI} = 0,26608$, cao hơn so với T_3 và T_2 , tuy nhiên giá trị Q_E vẫn thấp hơn đáng kể so với DATFuse. Kết quả này cho thấy DATFuse không chỉ cải thiện khả năng bảo toàn cường độ sáng mà còn nâng cao lượng thông tin trong ảnh tổng hợp.

3.2.2. Kết quả thực nghiệm 2

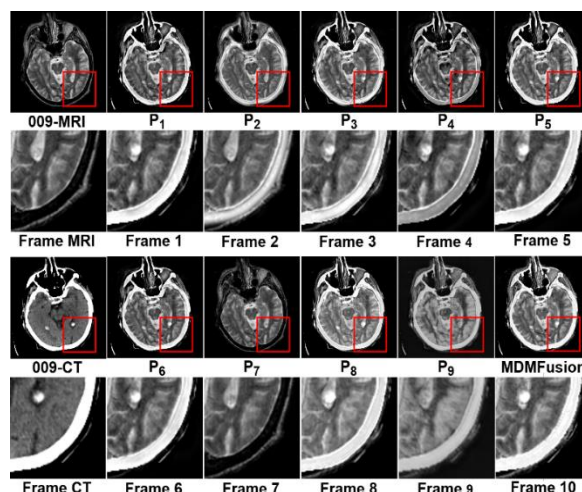
3.2.2.1. Đánh giá trực quan

Đánh giá chất lượng thị giác phản ánh lượng thông tin có giá trị chẩn đoán mà người quan sát có thể khai thác từ các phương thức ảnh nguồn và là một tiêu chí quan trọng để đánh giá các phương pháp tổng hợp ảnh y học phục vụ chẩn đoán lâm sàng dựa trên hình ảnh. Để thực hiện phân tích so sánh, chúng tôi lựa chọn ngẫu nhiên 01 cặp ảnh tổng hợp được tạo ra bởi mười phương pháp tổng hợp tiêu biểu từ bộ dữ liệu MRI-CT; kết quả được trình bày trong Hình 5. Đối với mỗi cặp ảnh, một vùng quan tâm tại cùng vị trí giải phẫu được cắt và phóng đại để quan sát chi tiết.

Hình 5 minh họa kết quả tổng hợp cho cặp ảnh mức xám (trường hợp #009), đại diện cho bộ dữ liệu MRI-CT. Các phương pháp P_2 , P_4 , P_7 và P_9 tạo ra các ảnh tổng hợp có độ sáng và độ tương phản suy giảm rõ rệt; trong khi P_8 vẫn duy trì được độ sáng tương đối tốt nhưng xuất hiện hiện tượng mất tương phản đáng kể. Đáng chú ý, các vùng vôi hóa tăng đậm độ trên ảnh CT bị mất hoàn toàn trong kết quả của P_9 và gần như không còn xuất hiện trong P_2 và P_7 . Bên cạnh đó, nhiều chi tiết giải phẫu quan trọng từ ảnh MRI nguồn bị suy giảm đáng kể trong kết quả của P_8 , P_9 và P_3 . Mặc dù các phương pháp P_1 , P_3 , P_5 và P_6 bảo toàn độ sáng, độ tương phản và các chi tiết quan trọng của ảnh đầu vào tốt hơn so với P_2 , P_4 , P_7 và P_9 , chất lượng thị giác tổng thể cũng như khả năng duy trì các đặc trưng nổi bật của chúng vẫn kém hơn rõ rệt so với phương pháp MDMFusion được đề xuất.

Bảng 1. So sánh DATFuse với các giải thuật thay thế

Ký hiệu	Q_{MLI}	Q_E
T_1	0,19546	4,28925
T_2	0,25606	4,25989
T_3	0,26349	4,27128
T_4	0,26608	4,26663
MDMFusion	0,27243	6,48848



Hình 5. So sánh trực quan và các lát cắt của các ảnh tổng hợp

3.2.2.2. Đánh giá khách quan

Bảng 2. So sánh MDMFusion với các phương pháp tổng hợp ảnh

Ký hiệu	Giải thuật	Q_{MLI}	Q_E	Q_{MI}	$Q^{AB/F}$
P_1	CNP_NSST	0,24530	4,77408	2,48253	0,55075
P_2	MLMG_PCNN	0,23981	4,84915	2,59324	0,37595
P_3	DTNP_NSCT	0,25534	4,97451	2,53026	0,54459
P_4	AGPCNN	0,19741	5,26739	2,00897	0,50340
P_5	IDCPCNN	0,25714	5,01937	2,52228	0,53891
P_6	DCRNN	0,25119	4,70098	2,52993	0,53906
P_7	FATFusion	0,12956	4,69403	2,96910	0,40324
P_8	MDHU	0,26953	4,06632	3,31451	0,57693
P_9	MACTFusion	0,19254	4,57308	2,46191	0,39151
	MDMFusion	0,27243	6,48848	2,54383	0,48593

Bảng 2 trình bày so sánh định lượng giữa mô hình MDMFusion đề xuất và chín phương pháp tổng hợp ảnh tiên tiến từ P_1 đến P_9 trên bộ dữ liệu MRI-CT thông qua bốn chỉ số đánh giá gồm Q_{MLI} , Q_E , Q_{MI} và $Q^{AB/F}$. Kết quả cho thấy MDMFusion đạt hiệu quả cạnh tranh và thể hiện ưu thế rõ rệt ở khả năng bảo toàn độ sáng và lượng thông tin của ảnh tổng hợp.

Cụ thể, với chỉ số Q_{MLI} , MDMFusion đạt giá trị 0,27243, cao nhất trong tất cả các phương pháp được so sánh. Kết quả này vượt qua MDHU (0,26953), IDCPCNN (0,25714) và DTNP_NSCT (0,25534), đồng thời cao hơn so với các phương pháp AGPCNN (0,19741), MACTFusion (0,19254) và FATFusion (0,12956). Điều này cho thấy MDMFusion có khả năng bảo toàn cường độ sáng của ảnh đầu ra hiệu quả.

Đối với chỉ số Q_E , MDMFusion đạt giá trị 6,48848, cao vượt trội so với tất cả các phương pháp còn lại. Trong khi phương pháp AGPCNN đạt giá trị cao nhất trong nhóm so sánh với 5,26739, các phương pháp khác chủ yếu dao động trong khoảng từ 4,0 đến 5,0. Kết quả này cho thấy ảnh tổng hợp của MDMFusion chứa lượng thông tin lớn hơn đáng kể so với các phương pháp so sánh.

Ở chỉ số Q_{MI} , MDMFusion đạt giá trị 2,54383, cao hơn nhiều phương pháp như CNP_NSST (2,48253), AGPCNN (2,00897), DCRNN (2,52993) và MACTFusion (2,46191). Tuy nhiên, kết quả này vẫn thấp hơn một số phương pháp như FATFusion (2,96910), MDHU (3,31451) và MLMG_PCNN (2,59324). Điều này cho thấy mặc dù chưa đạt giá trị cao nhất về thông tin tương hỗ, MDMFusion vẫn duy trì được lượng thông tin đáng kể từ các ảnh nguồn trong quá trình tổng hợp.

Đối với chỉ số $Q^{AB/F}$, MDMFusion đạt giá trị 0,48593, cao hơn FATFusion (0,40324) và MACTFusion (0,39151), tuy nhiên vẫn thấp hơn các phương pháp như MDHU (0,57693), CNP_NSST (0,55075) và DTNP_NSCT (0,54459). Kết quả này cho thấy mô hình đề xuất đạt hiệu quả tương đối tốt trong việc bảo toàn thông tin biên, mặc dù vẫn còn khoảng cách nhất định so với một số phương pháp chuyên sâu về bảo toàn cạnh.

Nhìn chung, MDMFusion cho thấy hiệu quả nổi bật trong việc bảo toàn độ sáng và gia tăng lượng thông tin của ảnh tổng hợp, đồng thời duy trì khả năng lưu giữ thông tin từ ảnh nguồn ở mức cạnh tranh. Mặc dù hiệu quả bảo toàn thông tin tương hỗ và thông tin biên chưa đạt mức tối ưu so với một số phương pháp tiên tiến khác, kết quả thực nghiệm vẫn khẳng định tính hiệu quả và tiềm năng ứng dụng của mô hình đề xuất trong bài toán tổng hợp ảnh y học.

4. KẾT LUẬN

Trong bài báo này, chúng tôi đã đề xuất MDMFusion, một mô hình tổng hợp ảnh y học mới cho bài toán tổng hợp ảnh MRI và CT dựa trên phân rã đa mức MIDTE, mô hình học sâu DATFuse và mạng nơ-ron xung cải tiến mPCNN. Phương pháp đề xuất tận dụng ưu điểm của THF và EGIF trong quá trình phân rã nhằm cân bằng giữa khả năng khử nhiễu và bảo toàn chi tiết, đồng thời sử dụng DATFuse để nâng cao chất lượng tổng hợp thành phần cơ sở và mPCNN để cải thiện khả năng bảo toàn thông tin biên và cấu trúc trong các thành phần chi tiết. Kết quả thực nghiệm trên bộ dữ liệu MRI-CT cho thấy MDMFusion đạt hiệu quả tốt trong việc bảo toàn cường độ sáng, duy trì thông tin từ ảnh nguồn và cải thiện chất lượng thị giác của ảnh tổng hợp. So với nhiều phương pháp tổng hợp ảnh tiên tiến hiện nay, mô hình đề xuất cho thấy khả năng cạnh tranh rõ rệt, đặc biệt ở các chỉ số Q_{MLI} và Q_E . Mặc dù đạt được những kết quả khả quan, phương pháp đề xuất vẫn còn hạn chế, đặc biệt ở khả năng bảo toàn thông tin biên (chỉ số $Q^{AB/F}$ thấp hơn MDHU, CNP_NSST và DTNP_NSCT) và thông tin tương hỗ (Q_{MI}). Trong các nghiên cứu tiếp theo, chúng tôi sẽ tập trung cải thiện chiến lược tổng hợp thành phần chi tiết, tối ưu độ phức tạp tính toán và mở rộng đánh giá trên nhiều bộ dữ liệu y học đa phương thức khác nhằm nâng cao khả năng ứng dụng thực tiễn của mô hình.

TÀI LIỆU THAM KHẢO

- [1] B. Li, H. Peng, and J. Wang, “A novel fusion method based on dynamic threshold neural P systems and nonsubsampling contourlet transform for multi-modality medical images,” *Signal Processing*, vol.178, pp.107793, 2021, DOI: 10.1016/J.SIGPRO.2020.107793.
- [2] C. Panigrahy, A. Seal, and N. K. Mahato, “Parameter adaptive unit-linking dual-channel PCNN based infrared and visible image fusion,” *Neurocomputing*, vol. 514, pp.21-38, 2022, DOI: 10.1016/J.NEUCOM.2022.09.157.
- [3] C. Panigrahy, A. Seal, C. Gonzalo-Martín, P. Pathak, and A. S. Jalal, “Parameter adaptive unit-linking pulse coupled neural network based MRI–PET/SPECT image fusion,” *Biomed. Signal Process. Control*, vol. 83, pp.104659, 2023, DOI: 10.1016/J.BSPC.2023.104659.
- [4] B. Rajalingam, F. Al-Turjman, R. Santhoshkumar, and M. Rajesh, “Intelligent multimodal medical image fusion with deep guided filtering,” *Multimed. Syst.*, 28(4):14491463, 2022, DOI: 10.1007/S00530-020-00706-0/METRICS.
- [5] S. Maqsood and U. Javed, “Multi-modal Medical Image Fusion based on Two-scale Image Decomposition and Sparse Representation,” *Biomed. Signal Process. Control*, vol. 57, pp.101810, 2020, DOI: 10.1016/J.BSPC.2019.101810.
- [6] P. Vajpayee, C. Panigrahy, and A. Kumar, “Medical image fusion by adaptive Gaussian PCNN and improved Roberts operator,” *Signal Image Video Process.*, 17(7): 3565-3573, 2023, DOI: 10.1007/S11760-023-02581-4/METRICS.
- [7] W. Liu, P. Zhang, Y. Lei, X. Huang, J. Yang, and M. Ng, “A Generalized Framework for Edge-Preserving and Structure-Preserving Image Smoothing,” *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(10):6631-6648, 2022, DOI: 10.1109/TPAMI.2021.3097891.
- [8] Z. Lu, B. Long, K. Li, and F. Lu, “Effective Guided Image Filtering for Contrast Enhancement,” *IEEE Signal Process. Lett.*, 25(10):1585-1589, 2018, DOI: 10.1109/LSP.2018.2867896.
- [9] W. Tang, F. He, Y. Liu, Y. Duan, and T. Si, “DATFuse: Infrared and Visible Image Fusion via Dual Attention Transformer,” *IEEE Transactions on Circuits and Systems for Video Technology*, 33(7):3159-3172, 2023, DOI: 10.1109/TCSVT.2023.3234340.
- [10] W. Tan, W. Thitøn, P. Xiang, and H. Zhou, “Multi-modal brain image fusion based on multi-level edge-preserving filtering,” *Biomed. Signal Process. Control*, vol. 64, pp. 102280, 2021, DOI: 10.1016/J.BSPC.2020.102280.
- [11] W. Tan, P. Tiwari, H. M. Pandey, C. Moreira, and A. K. Jaiswal, “Multimodal medical image fusion algorithm in the era of big data,” *Neural Comput. Appl.*, pp. 1-21, 2020, DOI: 10.1007/S00521-020-05173-2/FIGURES/18.
- [12] G. Qu, D. Zhang, and P. Yan, “Information measure for performance of image fusion,” *Electron. Lett.*, 38(7):313315, 2002, DOI: 10.1049/EL:20020212.
- [13] C. S. Xydeas and V. Petrović, “Objective image fusion performance measure,” *Electron. Lett.*, 36(4):308-309, 2000, DOI: 10.1049/EL:20000267.

- [14] B. Li *et al.*, “Medical Image Fusion Method Based on Coupled Neural P Systems in Nonsubsampled Shearlet Transform Domain,” 31(1), 2020, DOI: 10.1142/S0129065720500501.
- [15] A. Sinha *et al.*, “Multi-modal medical image fusion using improved dual-channel PCNN,” *Med. Biol. Eng. Comput.*, 62(9):2629-2651, 2024, DOI: 10.1007/S11517-024-03089-W/METRICS.
- [16] N. Goyal *et al.*, “Dual-channel Rybak neural network based medical image fusion,” *Opt. Laser Technol.*, vol. 181, 2025, DOI: 10.1016/j.optlastec.2024.112018.
- [17] W. Tang and F. He, “FATFusion: A functional-anatomical transformer for medical image fusion,” *Inf. Process. Manag.*, 61(4), 2024, DOI: 10.1016/j.ipm.2024.103687.
- [18] Y. Jie, X. Li, H. Tan, F. Zhou, and G. Wang, “Multi-modal medical image fusion via multi-dictionary and truncated Huber filtering,” *Biomed. Signal Process. Control*, vol. 88, pp. 105671, 2024, DOI: 10.1016/J.BSPC.2023.105671.
- [19] X. Xie *et al.*, “MACTFusion: Lightweight Cross Transformer for Adaptive Multimodal Medical Image Fusion,” *IEEE J. Biomed. Health Inform.*, 29(5): 3317-3328, 2025, DOI: 10.1109/JBHI.2024.3391620.

Tuyên bố về xung đột lợi ích: Các tác giả tuyên bố không có xung đột lợi ích.

Thông tin tài trợ (nếu có): Nghiên cứu này không nhận được tài trợ từ bên ngoài.