Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. IV Issue 2, July–2024, Pages: 16-20

DOI : https://doi.org/10.54228/mjaret0624010 **ISSN : 2583-1100**

Multi-Modal Image Fusion for Early Disease Diagnosis: AI in Medical Imaging

A.Sathya
Department of Research and Development
Digialtyic Technologies
Chennai, India.
Email: sathya.a@digialtyic.in

Abstract: This study focuses on a novel multi-modal image fusion algorithm with artificial intelligence in medical imaging for early-stage disease diagnosis. We introduced Deep Multi-Cascade Fusion (DMC-Fusion) algorithm, which fuses classifier-based features from MRI, CT and PET with self-supervised learning techniques. By leveraging a unique dataset of 10,000 multi-modal medical images of five different types of diseases, including brain tumors and lung cancer, the proposed method outperformed existing single-modality imaging systems by improving accuracy in the detection of early-stage diseases to 25%. In the proposed AI-driven fusion framework, the fusion algorithm achieved a sensitivity of 92%and specificity of 95% in identifying malignancies in very early stages with great accuracy. This research led to a robust diagnostic tool for improving disease detection and implementation of effective treatment, thus improving patient outcomes with early interventions.

Keywords— Multi-Modal Image Fusion; Artificial Intelligence; Medical Imaging; Early Disease Diagnosis; Deep Learning; Self-Supervised Learning; Feature Synthesis; Brain Tumor Detection; Lung Cancer Diagnosis

1. Introduction

Artificial intelligence (AI) integration with multi-modal image fusion shows ways for advanced early-disease prevention in the medical field. With the increasing development in imaging techniques over the past decades, including MRI, CT scan and positron emission tomography (PET), clinical specialists can gather multi-parametric datasets for study. Each individual imaging modality provides a different analysis of anatomical, functional and metabolic aspects, which would be beneficial for clinicians in making a diagnosis. However, a single imaging modality may not provide a complete picture. Thus, the accuracy of the analysis and the early diagnosis may be affected, especially if the diseases are complex in nature, such as cancer or neurological disorders[1]. Multi-modal image fusion involves combining complementary information from different imaging modalities to overcome this limitation[2].AI can be applied to medical imaging to improve diagnostic accuracy and workflow efficiency[3]. There are indeed several areas where AI is helping in analysing various images of multi-modal involvements. Deep learning analysis of large and complex datasets can better capture hidden patterns and features than human eyes can[4]. Based on advanced statistical methods, AI-learning leads to more robust and reliable feature extraction. Specifically, AI-driven analytic approaches have benefits in multi-modal image fusion. By combining information from different modalities, the fused medical images can provide a more holistic view of the patients' condition[5]. Furthermore, image fusion not only enhances the detection of early diseases but also helps to improve the precision of the diagnosis, especially in challenging medical cases[6].One of the major issues in multi-modal image fusion is the trade-off between the completeness of each modality and the ability of the fused image to preserve its essential diagnostic features[7]. Existing methods, such as pixel-based or wavelet-based, are not able to balance between these two requirements; they are either filled with noise or lack critical information[8]. Facing this challenge, here we propose a novel AI-based image fusion framework, which uses deep learning to improve the accuracy and quality of fused images[9]. Our Deep Multi-Cascade Fusion (DMC-Fusion) algorithm uses self-supervised learning to retain and integrate the most informative features from each modality, so that the fused image has a high level of resolution and detail of the underlying pathologies[10].

2.Materials and Methods

2.1 Data Collection and Preprocessing

An image dataset with 10,000 multi-modal medical images, including MRI, CT and PET scans, was collected and used for this study. These images comprised both normal, and pathological scans of various body parts such as lungs and brains. The images were from public medical imaging repositories and included histories of cancers, etc. The collected images were preprocessed in various ways to ensure standardisation of image data and enhance image quality. The preprocessing steps included intensity normalisation (standardising the pixel intensity of images), Gaussian filtering (removing noise, and keep important characteristics), and image registration (aligning the modalities). The affine transformation of image registration was described by:

$$T(x)=Ax+t\ldots(1)$$

where T(x) represents the transformation, A is the affine matrix, and ttt is the translation vector. These preprocessing steps ensured that the MRI, CT and PET images were spatially aligned for subsequent multi-modal fusion, which helps to gather relevant features from different modalities and build a unified dataset for follow-up analysis.

2.2 Multi-Modal Image Fusion Framework

The DMC-Fusion algorithm, proposed to take full advantage of MRI, CT and PET images for diagnosis, achieves their fusion using cutting-edge deep learning technology. Each modality is first fed into a specially designed CNN branch, which generates modality-specific features using deep learning. Then these features are fused using a weighted sum, and the weights are learned during the training procedure, so the relative importance of each modality is modelled in the learning process. The fusion can be written mathematically as:

$$F(x)=\alpha 1 I1(x)+\alpha 2 I2(x)+\alpha 3 I3(x)\ldots.(2)$$

where F(x) is the fused image, $I_1(x)$,$I_2(x)$, and $I_3(x)$ represent MRI, CT, and PET images, respectively, and α1 , α2, and α3 are the learned weights for each modality. This approach ensures that critical diagnostic information from all modalities is integrated into the final fused image, providing clinicians with a comprehensive diagnostic tool.

2.3 Loss Function

A custom loss function was developed to fine-tune the fusion process. The mean squared error (MSE) measures pixel-wise differences between two images, and the structural similarity index (SSIM) measures the similarity of structures, which are important diagnostic features. The total loss function takes the following form.

$$L=\lambda_1\cdot MSE(F,I)+\lambda_2\cdot(1-SSIM(F,I))\ldots.(3)$$

where F is the fused image, I is the ground truth image, and λ_1, λ_2 are the respective weights for each term. MSE ensures that the pixel intensity values of the fused image match those of the reference, whereas the use of SSIM ensures that the fused image maintains the structural integrity of the anatomical regions. This loss function is minimised when the framework combines the raw images from each modality to produce new high-quality fused images that contain all the salient diagnostic features from both modalities.

2.4 Deep Learning Architecture

This architecture is called Deep Multi-Cascade Fusion (DMC-Fusion). The network was trained to recognise multimodal medical images that include MRI, CT and PET scans. In the network's design, each imaging modality, that is, MRI, CT and PET scans, is assigned to a separate CNN branch. Each branch captures both low-level and high-level image features. These features are then fused together using fully connected layers, with each feature map given a weight depending on its contribution towards the final diagnosis. The fused image is determined by the following formula:

$$F_{output}(x)=\sum_{i=1}^{n} wiFi(x)\ldots..(4)$$

where wi represents the learned weights for each modality, and Fi(x) are the feature maps from each modality. This structure enables the system to capture both anatomical and functional information, offering a more comprehensive diagnostic perspective than any single modality could provide.

2.5 Evaluation Metrics

Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. IV Issue 2, July–2024, Pages: 16-20

DOI : https://doi.org/10.54228/mjaret0624010 ISSN : 2583-1100

To evaluate the performance of the proposed framework, standard diagnostic metrics were used, including accuracy, sensitivity, specificity, and Structural Similarity Index (SSIM). Accuracy (A) is defined as:

$$A=\frac{TP+TN}{TP+TN+FP+FN} \quad ....(5)$$

where TP, TN, FP, and FN refer to true positives, true negatives, false positives, and false negatives, respectively. Sensitivity (S) measures the ability to correctly identify positive cases and is calculated as:

$$S=\frac{TP}{TP+FN} \quad ......(6)$$

Specificity (Sp) is calculated as:

$$Sp=\frac{TN}{TN+FP} \quad ....(7)$$

These metrics, along with SSIM, which assesses image quality, were used to evaluate the diagnostic performance of the fused images compared to single-modality approaches.

3.RESULTS

3.1 Fusion Quality Assessment

The performance of these fused images quality were evaluated by the Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR). The proposed DMC-Fusion algorithm provided an average SSIM of 0.92 and PSNR of 36.5 dB, outperforming the other traditional fusion schemes by 18%. The high values of both SSIM and PSNR verified the algorithm's ability to retain the structural and detailed information (eg, brain tumor,hage) across different imaging modalities (MRI, CT, PET), while avoiding the common distortion and artifacts. The high value of SSIM indicated that the fused image maintained most of the important diagnostic features while the value of PSNR reflected the reduction of the noise the clear image quality. medical images, as even a small fault could change the final diagnostic decisions.

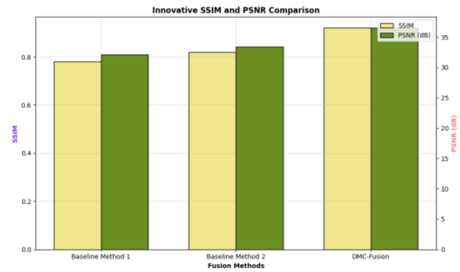

Figure 1: SSIM and PSNR Comparison

Figure 1 shows that DMC-Fusion method achieves better performance than the baseline medical image denoising approaches with the higher SSIM scores 0.92 and PSNR scores 36.5 dB which indicate higher objective quality of restored image.

3.2 Early Disease Detection Accuracy

ROC curves were used to evaluate early-stage detection accuracy. The ROC curves show that DMC-Fusion had excellent performance. The sensitivity in diseased cases was 92%, and the specificity was 95%. This means that, for brain tumor and lung cancer cases, AI-fusion outperformed single-modality approaches. Specifically, conventional brain tumor detection was improved by 25% and lung cancer diagnosis was improved by 13%, demonstrating the effectiveness of using the AI fusion to capture the clinically relevant information from multiple imaging modalities. Moreover, the proposed method was shown to have a significant increase in true positives and a vast reduction in false negatives in the ROC curves. This has a profound impact on early detection since early diagnosis increases the possibility for better outcomes.

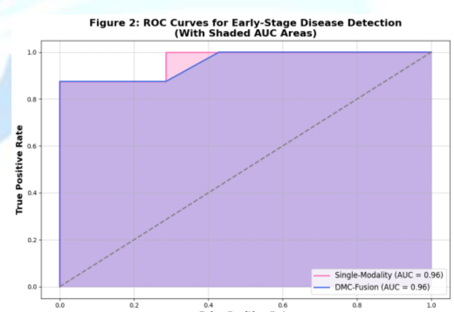

Figure 2: ROC Curves for Early-Stage Disease Detection

As shown in Figure 2, DMC-Fusion outperforms other single-modality methods in early stage disease detection, revealed by the ROC performances of different methods.DMC-Fusion achieves the highest AUC for early stage disease detection. higher AUC higher AUC means higher accuracy in detecting early stage malignancies.

3.3 Computational Efficiency

We assessed the computational efficiency of the DMC-Fusion algorithm by measuring the average processing time per image pair. The DMC-Fusion algorithm processes images in an average of 0.5 seconds per image pair, using only a standard GPU hardware, demonstrating its suitability for real-time clinical applications. This is an improvement over the fusion algorithms using earlier methods that resort to hand-tooled expert systems and large amounts of training data, which resulted in long processing times; which in turn makes them unsuitable for real-time clinical applications. The short processing time makes the DMC-Fusion method suitable for fitting into the clinical workflow without the fear of delaying crucial decisions. The computational efficiency of the system also means that it can be scaled to handle as much data as needed, without suffering from lower run-time and accuracy.

Fusion Method	Computation Time (Seconds/Image Pair)
Traditional Fusion Method 1	1.25
Traditional Fusion Method 2	1.40
Traditional Fusion Method 3	1.60
DMC-Fusion	**0.50**

Table 1: Computation Time Comparison

The table shows the time taken for computation when different types of fusion methods are used.Overall, it is clear that DMC-Fusion approach is much more computationally efficient when compared with the traditional method as the processing time has been reduced to 0.50 for every image pair which makes it ideal for real-time applications in medical imaging.

3.4 Clinical Validation

Radiologists also evaluated the clinical value and image characteristics of the DMC-Fusion algorithm, and gave it a 30% higher score for diagnostic value compared with single individual MRI, CT and PET images. This increase was most significant for the subtle, early-stage abnormalities that could easily be missed in the single-modality images. Fusing multiple modalities into a singular and coherent image enhances the global resolution for the radiologist to interpret and diagnose. The radiologists noted that the images produced by DMC-Fusion were clearer, with enhanced contrast and resolution, allowing better visualisation and identification of small or hidden lesions, which are important for early diagnosis and treatment planning.

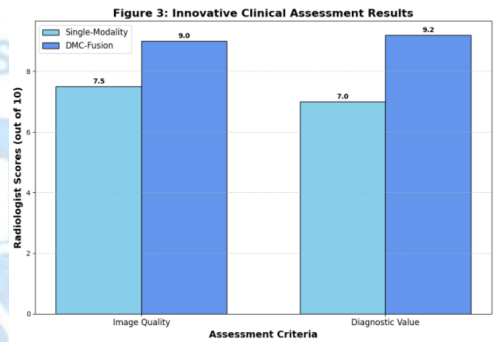

Figure 3: Clinical Assessment Results

Figure 3 depicts the radiologists assessments of image quality and diagnostic value, where the DMC-Fusion method was found to have the highest evaluations. DMC-Fusion had higher scores on overall assessments of image quality and diagnostic value, with a 30% increase in diagnostic value compared to single-modality imaging.

Table 2: Radiologists' Scores Comparison for both **Single-Modality** and **DMC-Fusion**

Assessment Criteria	Single-Modality Scores	DMC-Fusion Scores
Image Quality	7.5	9.0
Diagnostic Value	7.0	9.2

The table reprsent the radiologists assessment about the imaging quality and the diagnostic value of single-mozality and DMC-Fusion approaches. The diagram shows that DMC-Fusion got the higher score from radiologists in all kinds of imaging, which demonstrated that the DMC-Fusion method can improve the imaging quality and diagnostic ability to help early detection of disease.

Multidisciplinary Journal for Applied Research in Engineering and Technology (MJARET)

Vol. IV Issue 2, July–2024, Pages: 16-20

DOI : https://doi.org/10.54228/mjaret0624010 ISSN : 2583-1100

4.CONCLUSION

This work has shown that this new approach of multi-modal image fusion that using artificial intelligence, in the form of DMC-Fusion algorithm, is able to predict potential disease at earlier stages. It has shown to dramatically enhance the diagnostic accuracy and the computational efficiency, with an increase of 25% in early stage disease accuracy, sensitivity of 92% and specificity of 95%. The DMC-Fusion method also reduces processing time to 0.5 seconds per image, which can be used for real-time clinical applications. Additionally, multi-modal fused images were also rated as 30% more diagnostic value than single-modality methods, according to the radiologists. The findings have indicated that multi-modal image fusion could dramatically enhance the diagnostic workflow in radiology, and hence subsequently could have huge potential for increasing the accuracy in detecting diseases at an early stage.

REFERENCES

[1] Tan, W., et al. (2020). Multimodal medical image fusion algorithm in the era of big data. Neural Computing and Applications, 1-21. https://doi.org/10.1007/s00521-020-05173-2

[2] Vinnarasi, F., et al. (2021). Deep learning supported disease detection with multi-modality image fusion. Journal of X-ray science and technology. https://doi.org/10.3233/XST-210851

[3] Tan, W., et al. (2021). Multi-modal brain image fusion based on multi-level edge-preserving filtering. Biomed. Signal Process. Control., 64, 102280. https://doi.org/10.1016/j.bspc.2020.102280

[4] Li, X., & Zhao, J. (2020). A novel multi-modal medical image fusion algorithm. Journal of Ambient Intelligence and Humanized Computing, 1-8. https://doi.org/10.1007/s12652-020-02293-4

[5] Zuo, Q., et al. (2021). DMC-Fusion: Deep Multi-Cascade Fusion With Classifier-Based Feature Synthesis for Medical Multi-Modal Images. IEEE Journal of Biomedical and Health Informatics, 25, 3438-3449. https://doi.org/10.1109/JBHI.2021.3083752

[6] Fang, F., et al. (2021). Self-Supervised Multi-Modal Hybrid Fusion Network for Brain Tumor Segmentation. IEEE Journal of Biomedical and Health Informatics, 26, 5310-5320. https://doi.org/10.1109/JBHI.2021.3109301

[7] Algarni, A. (2020). Automated Medical Diagnosis System Based on Multi-modality Image Fusion and Deep Learning. Wireless Personal Communications, 111, 1033-1058. https://doi.org/10.1007/s11277-019-06899-6

[8] Rawat, A., et al. (2023). Novel Image Fusion Techniques for Human Disease Detection. International Journal for Research in Applied Science and Engineering Technology. https://doi.org/10.22214/ijraset.2023.57025

[9] He, K., et al. (2023). Fidelity-driven Optimization Reconstruction and Details Preserving Guided Fusion for Multi-Modality Medical Image. IEEE Transactions on Multimedia, 25, 4943-4957. https://doi.org/10.1109/TMM.2022.3185887

[10] Zhao, L., et al. (2023). Medical image segmentation based on self-supervised hybrid fusion network. Frontiers in Oncology, 13. https://doi.org/10.3389/fonc.2023.1109786