Quantitative Evaluation of Multi-Modal Medical Image Fusion Using Diagnostic Performance Metrics

Kiran Kumar B[1], *Hemalatha S*[2]
[1]Postdoc Scholar, Lincoln University College, Malaysia ; Panimalar engineering college, Bangalore trunk road, poonamallee, Chennai Tamil Nadu, India[2]
Email ID: pdf.kiranbeesetti@lincoln.edu.my

Abstract: This multi-modal medical imaging is very important for the precise identification of a certain disorder, because many imaging modalities give the user, not only anatomical, but also functional information at no cost. However, relying on a single modality of images often limits diagnostic accuracy because of the lack of features. In this work, the goal is to quantitatively evaluate multi-modal medical picture fusion techniques with traditional medical diagnostic performance metrics. These images are merged together from a variety of modalities to create a composite image that contains a lot of information from all modalities. The ability to distinguish fused images is carefully measured against single-modality images, with a number of parameters such as accuracy, sensitivity and specificity. Experimental results show that the fused images regularly outperform the individual modality images in all the criteria tested, which means that they provide better feature discrimination and enhanced diagnosis support. The results indicate the effectiveness of multi-modal picture fusion as an effective tool to assist in clinical decision making, and highlight its potential in computer-aided diagnostic systems.

Keywords: Multi-modal medical imaging; Image fusion; Diagnostic performance evaluation; Accuracy; Sensitivity; Specificity; Computer-aided diagnosis

1. Introduction

In modern medicine, imaging plays a crucial role in the visualization of anatomical structures and physiological processes without the need for any invasive procedures, which is crucial for the diagnosis and treatment planning of diseases. Other imaging techniques provide unique and additional information, such as Magnetic Resonance Imaging (MRI), Computed Tomography (CT), Positron Emission Tomography (PET), and Ultrasound. MRI provides excellent contrast resolution of soft tissues, CT is excellent for the bones, and PET provides functional and metabolic information. In complex disease situations, no single imaging modality is sufficient to capture all aspects of the disease in order to allow for proper or accurate diagnosis and interpretation [1]. To address the constraints of single-modality imaging, multi-modal medical picture fusion has arisen as a possible solution. The aim of image fusion is to combine relevant information contained in multiple imaging modalities in a single fused image while preserving the salient features of each modality. Fused pictures are able to enhance the visual perception, improve the representation of features and offer a reliable clinical evaluation using anatomical and functional information. However, the use of image fusion techniques has been widely explored in various fields such as detection of tumors, diagnosis of brain diseases, and monitoring of treatment in these diseases [2]. A major challenge in the objective assessment of the diagnostic use of medical image fusion algorithms has yet to be addressed, despite recent progress in these algorithms. Many existing research projects focus on either qualitative aspects of increasing image quality or are limited to algorithmic aspects, but do not validate these aspects quantitatively from a diagnostic perspective. Moreover, a wide range of image quality metrics are often used to assess images, which do not necessarily reflect their actual diagnostic value [3]. This difference highlights the importance of systematic evaluation frameworks; both based on single modality and fused pictures, based on clinically relevant performance criteria. In this application, the accuracy, sensitivity and specificity are objective diagnostic performance metrics used to test the effectiveness of fused images for clinical diagnosis. These are internationally accepted markers which provide good clinical indicators and are effective in differentiating pathological from non-pathological cases and therefore being representative of a system's ability to differentiate. Such a quantitative assessment, based on these parameters, can aid in the more reliable comparison of single-modality and fused imaging techniques, thus guaranteeing imaging repeatability in different data sets and clinical settings. These motivations are the reason for this work and it is aimed at proposing a quantitative evaluation of multi-modal medical image fusion based on standard diagnostic performance metrics. The proposed examination methodically compares fused pictures with the results obtained from other imaging techniques. The study will show the effectiveness of multi-modal fusion to increase accuracy, sensitivity and specificity of diagnostic aids for improving clinical decision making. This study shows the effectiveness of medical image fusion technique and suggests its use in computer-aided diagnostic (CAD) system [4].

2. Related work

To bridge the intrinsic limitations of using a single modality in medical imaging, multi-modal medical image fusion has been extensively studied. The current fusion methods can be broadly categorized into transform domain based methods, spatial domain based methods, and learning based methods [5]. The advantages and challenges of these classifications differ in each case, not only regarding memory, but also in terms of computational complexity and diagnostic relevance. The early study on medical image fusion focused mainly on the transform domain methods such as wavelet transform, Laplacian pyramid transform, contourlet transform and shearlet transform. The methods are based on decomposing the source image into different frequency sub-bands, and then combine illumination coefficients in a fusion manner. Wavelet fusion techniques have a number of multi-resolution properties and are very efficient in computation, thus they are widely used, particularly for the fusion of CT and MRI data. However, these methods often destroy the complexity of the structure and could also be affected by choice of fusion rules, limiting the ability to diagnose the structure. The spatial domain fusion methods work directly with the intensity values of the pixels and include methods like guided filtering [6], weighted averaging and principal component analysis (PCA). Simple fusion and its ability to reduce modality redundancy have been examined in the context of PCA-based fusion. While it is relatively simple to implement a spatial domain method, it can lead to artifacts, lost contrast or insufficient preservation of the important modality-specific features, especially in complex pathological regions. Learning-based approaches, particularly deep learning and convolutional neural networks (CNNs) have been suggested in recent years in the field of medical image fusion. These methods can learn hierarchical feature representations from multimodal inputs and give multimodal outputs, including structural and functional information [7].

3. Method

The purpose of this research is to propose a quantitative evaluation method for multi-modal medical pictures fusion, focusing on reliable and efficient image pre-processing, fusion and objective diagnostic performance evaluation. The method has been designed to ensure the consistency of position and intensity of the different modalities and to enable the equitable comparison of fused images with single modality images. The overall workflow is divided into six broad steps, namely: data acquisition, pre-processing, picture registration, intensity normalization, image fusion and diagnostic performance evaluation.

3.1 Data Acquisition

Multi-modal medical picture datasets are acquired from publically available and clinically approved datasets to ensure repeatability and generalization. The datasets cover imaging modalities such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET), to provide complementary anatomical and functional information. The datasets consist of images of the same anatomical location, acquired in different imaging conditions and intensities, and thus suitable for fusion evaluation [8].

3.2 Preprocessing Overview

All images are preprocessed prior to fusion, including typical processing operations like noise removal, resizing and intensity scaling. Pretreatment ensures compatibility of multi-modal images for registration and fusion [9] because the images are often different in spatial resolution and orientation. These processes help to reduce artifacts introduced during acquisition and to ensure precise spatial alignment and precise quantitative analyses of the images.

3.3 Image Registration

Precise spatial alignment among modalities is accomplished by image registration. One modality is used as the reference picture, while the other modalities are used as moving pictures. The dynamic images are matched to the reference image by use of a transformation function.

Mathematically, the registered image I_i^{reg} is defined as:

$$I_i^{reg}(x, y) = T_i(I_i(x, y))$$

where $T_i(\cdot)$ represents the spatial transformation applied to the i^{th} modality.

Registration minimizes spatial variations in anatomical components, and therefore reduces misalignment artifacts which negatively affect image fusion quality and the interpretation of the fused images.

3.4 Intensity Normalization

After registration, the intensity normalization is performed to reduce differences between the various imaging methods and scanner characteristics. Normalization ensures that intensity data from different modalities are comparable, and that they retain diagnostic-relevant features.

A normalized image I_i^{norm} is obtained as:

$$I_i^{norm}(x, y) = \frac{I_i^{reg}(x, y) - \mu_i}{\sigma_i}$$

μ_i and σ_i denote the mean and standard deviation of the intensities of the registered pictures of the i^{th} modality. This helps to make the intensity of each modality more uniform, which aids in fusing and extracting features.

3.5 Multi-Modal Image Fusion

The normalized images are then fused to create a single image containing complementary information from all modalities [10]. A weighted fusion approach is used to maintain the salient features of each modality.

The fused image $F(x, y)$is defined as:

$$F(x, y) = \sum_{i=1}^{M} w_i\,(x, y) \cdot I_i^{norm}(x, y)$$

Subject to:

$$\sum_{i=1}^{M} w_i\,(x, y) = 1$$

The weight $w_i(x, y)$ is determined from some local property of the image of the ith modality, e.g. contrast and edge intensity. This fusing process results in a clearer and informer image.

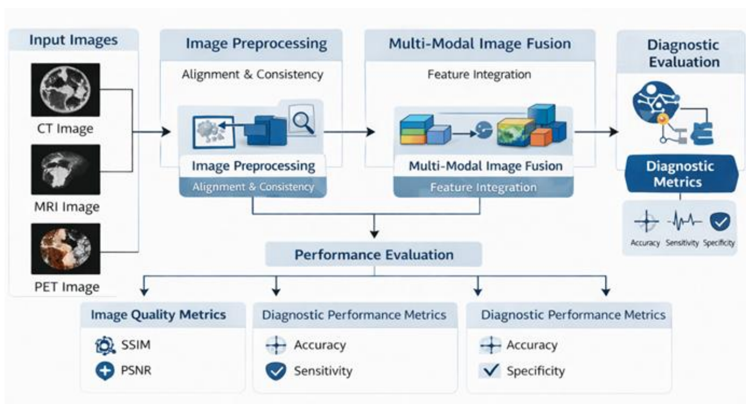

Figure 1: Architecture diagram of the proposed framework

4. Conclusions

This enhancement in image quality and diagnostic parameters can be attributed to the complementary characteristics of multi-modal imaging and the proposed framework's good integration. The method corrects for misalignment in space and inconsistencies in intensity before fusion, so that the features that are relevant for diagnosis in each modality provide something useful to the fused image. The use of clinically accepted metrics in the quantitative evaluation further enhances the validity of the results and can help to prove the applicability of the proposed approach in computer-aided diagnosis systems.

References

1. James, A. P., & Dasarathy, B. V., "Medical image fusion: A survey of the state of the art," *Information Fusion*, vol. 19, pp. 4–19, 2014.

2. Ma, J., Yu, W., Chen, C., Liang, P., Guo, X., & Jiang, J., “Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion,” *Information Fusion*, vol. 41, pp. 214–224, 2018.
3. Sotiras, A., Davatzikos, C., & Paragios, N., “Deformable medical image registration: A survey,” *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.
4. Modat, M., et al., “Global image registration using a symmetric block-matching approach,” *Journal of Medical Imaging*, vol. 1, no. 2, 2014.
5. Ronneberger, O., Fischer, P., & Brox, T., “U-Net: Convolutional networks for biomedical image segmentation,” *MICCAI*, pp. 234–241, 2015.
6. Balakrishnan, G., Zhao, A., Sabuncu, M. R., Guttag, J., & Dalca, A. V., “VoxelMorph: A learning framework for deformable medical image registration,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788–1800, 2019.
7. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P., “Image quality assessment: From error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
8. Pradhan, P., Yadav, R., & Singh, P., “Performance evaluation of medical image fusion techniques using quantitative metrics,” *Biomedical Signal Processing and Control*, vol. 58, 2020.
9. Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X., & Zhang, L., “IFCNN: A general image fusion framework based on convolutional neural network,” *Information Fusion*, vol. 54, pp. 99–118, 2020.
10. Goodfellow, I., Bengio, Y., & Courville, A., *Deep Learning*, MIT Press, 2016.