A Review on Multimodal Medical Image Fusion: Compendious Analysis of Medical Modalities, Multimodal Databases, Fusion Techniques and Quality Metrics

Muhammad Adeel Azam[1,2], Khan Bahadar Khan[3], Sana Salahuddin[4], Eid Rehman[5], Sajid Ali Khan[5], Muhammad Attique Khan[6], Seifedine Kadry[7], Amir H. Gandomi[8]

[1]Department of Advanced Robotics, Istituto Italiano di Tecnologia, Genova, Italy;
[2]Department of Informatics, Bioengineering, Robotics, and System Engineering, University of Genoa, Italy
[3]Department of Information and Communication Engineering, Faculty of Engineering, The Islamia University of Bahawalpur, Pakistan
[4]Institute of Physics, The Islamia University of Bahawalpur, Bahawalpur, Pakistan
[5]Department of Software Engineering, Foundation University Islamabad, 44000, Pakistan
[6]Department of Computer Science, HITEC University Taxila, Taxila, Pakistan
[7]Department of Applied Data Science, Noroff University College, Kristiansand, Norway
[8]Faculty of Engineering and IT, University of Technology Sydney, Ultimo, NSW 2007, Australia

Abstract

Background and objectives: Over the past two decades, medical imaging has been extensively apply to diagnose diseases. Medical experts continue to have difficulties for diagnosing diseases with a single modality owing to a lack of information in this domain. Image fusion may be use to merge images of specific organs with diseases from a variety of medical imaging systems. Anatomical and physiological data may be included in multi-modality image fusion, making diagnosis simpler. It is a difficult challenge to find the best multimodal medical database with fusion quality evaluation for assessing recommended image fusion methods. As a result, this article provides a complete overview of multimodal medical image fusion methodologies, databases, and quality measurements.

Methods: In this article, a compendious review of different medical imaging modalities and evaluation of related multimodal databases along with the statistical results is provided. The medical imaging modalities are organized based on radiation, visible-light imaging, microscopy, and multimodal imaging.

Results: The medical imaging acquisition is categorized into invasive or non-invasive techniques. The fusion techniques are classified into six main categories: frequency fusion, spatial fusion, decision-level fusion, deep learning, hybrid fusion, and sparse representation fusion. In addition, the associated diseases for each modality and fusion approach presented. The quality assessments fusion metrics are also encapsulated in this article.

Conclusions: This survey provides a baseline guideline to medical experts in this technical domain that may combine preoperative, intraoperative, and postoperative imaging, Multi-sensor fusion for disease detection, etc. The advantages and drawbacks of the current literature are discussed, and future insights are provided accordingly.

KEYWORDS: Multimodal Medical Image Fusion, Multimodal Databases, Fusion Techniques, Image Fusion Quality Metrics.

Introduction

For more than two decades, different medical imaging modalities have been used in clinical applications for disease diagnostic purposes. Generally, it is difficult to extract all required information from a single imaging modality to guarantee clinical precision and strength of the examination for diagnoses. Therefore, multimodal methods combine medical images from various modalities to make a new fused image with rich information that is reliable for clinical usage. Multimodal Medical Image Fusion (MMIF) utilizes images from different sources like X-Rays, Computed Tomography (CT), Single Photon Emission Computed Tomography (SPECT), Ultrasound (US), Magnetic Resonance Imaging (MRI), Infrared and Ultraviolet, Positron Emission Tomography (PET), etc. Images from MRI, X-ray, CT, and US can all show where the lesion is located, how large it is, and what it looks like, as well as the morphological and structural changes it has induced in nearby tissues. To get insight into a tumor's biological processes, soft tissue as well as functional information, the use of PET, fMRI, and SPECT is becoming more common. Functional and structural data of the medical images can be combined to produce more valuable information. Medical image fusion plays a critical part in the treatment of the same human organ, allowing for more accurate disease monitoring and analysis. [1]. Image fusion techniques are widely applicable in the following domains, including remote sensing, machine learning, satellite surveillance, contrast enhancement of images, boosting geometric adjustment, and medical imaging to significantly improve features that have not been observable using a single image, such as malignancies, lesions, cancer cells, etc. [2]. The growing number of research articles available in magazines, books, and journals demonstrates the high interest and importance of multimodal image fusion. Fig. 1 displays the number of publications in the area of multimodal image fusion per year. The results were obtained from PubMed, an online database for biomedical subjects [3].

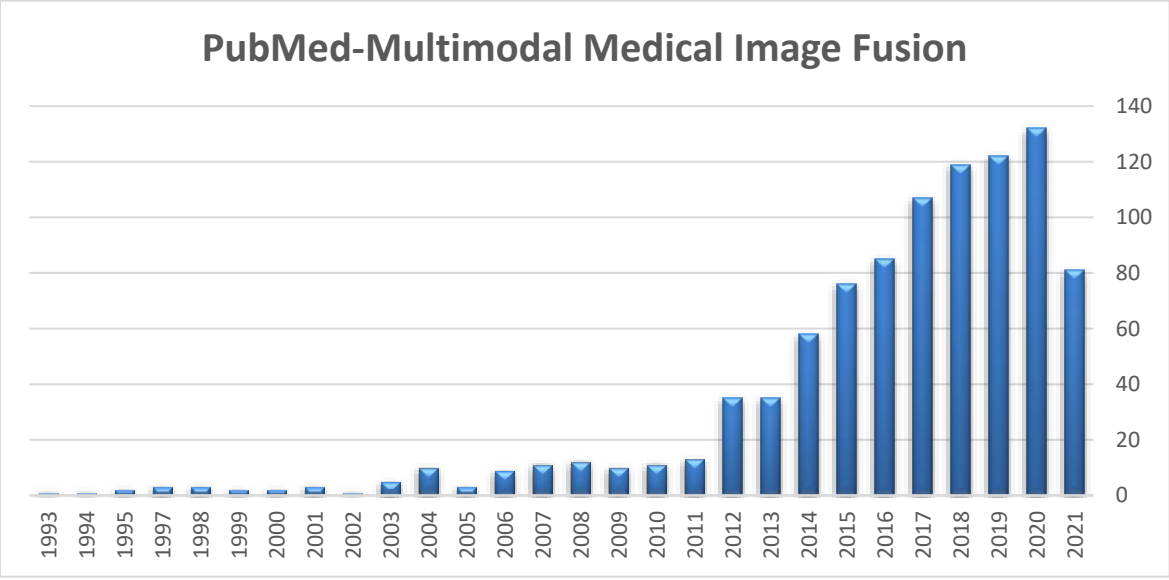

Fig. 1 Multimodal image fusion publications per year obtained from PubMed (1993 to third-quarter year Q3, 2021).

The multimodal medical image fusion techniques and classifications have been summarized in various surveys and review articles [2-6]. For instance, Alex *et al.* [2] cover the various scientific issues addressed in the area of medical image fusion. Define medical image fusion research by the techniques, imaging modalities, and organs examined, but the multimodal databases and fusion quality assessments metrics were still missing in this survey. Jiao *et al.* [4] explained multimodal medical image fusion steps, e.g. image fusion decomposition, and reconstruction, and fusion rules, in detail, comparing six fusion methods and using eight image fusion quality metrics. The work focuses exclusively on Harvard medical (AANLIB), but there is still a need to examine more multimodal databases for a greater variety of datasets. Additionally, the author did not address current diseases involved with MMIF in their research. Fatma *et al.* [5] introduced the classification of medical image registration and highlighted medical image fusion and the current diseases based on fusion work. Bikash *et al.* [6] provided a detailed comparison of region-based image fusion techniques using different fusion quality metrics. The images used in this review were from different sources, such as multimodal medical images, infrared rays, and visible image fusion, multi-focus image fusion, etc. The MMIF quality metrics are also discussed in tabular form.

Table 1 summarizes a recent literature review and compares it to our current study. We highlight many critical points that must be included in each MMIF evaluation. The table below is based on the following points:

IM (Imaging Modalities); It describes the presence of multimodal medical imaging modalities used in fusion.

Quant. A (Quantitative Analysis); whether the survey articles presented quantitative comparison on different fusion metrics for various techniques.

Qual. A (Qualitative Analysis); whether the survey articles presented qualitative comparison on different fusion imaging modalities for various techniques.

D.D (Database Description); It shows whether the review articles encapsulated multimodal databases with comprehensive descriptions that use in medical fusion.

T.D (Techniques Description); It represents whether review articles have any theoretical technique description for MMIF.

D.F (disease-based Fusion); the articles included diseased-based fusion work in their contribution or not.

F. B. S (Fusion Basic steps); whether the articles included general steps and rules to perform medical image fusion before going toward advanced techniques.

M. D. (Metrics Description); is there a proper description of MMIF performance metrics in survey articles.

Table. 1 A recent study of the literature on MMIF, as well as our work.

Ref.	Year	Contributions	I.M	Quant. A	Qual. A	D.D	T.D	D.F	F. B. S	M.D
Dolly JM et al. [7]	2019	Classification: (Feature, Decision, and Pixel level). Pros and cons of each technique described.	✓				✓			
Huang B. et al [8]	2020	Classification: (Spatial, Transform, and DL). Comparatively analysis of MRI/CT and PET/MRI. NSCT and CNN framework presented.		✓	✓		✓			✓
Tirupal T. et al [9]	2020	Classification: (Frequency, Spatial and Pixel level). Imaging modalities brief description. Both Qualitative and statistical evaluations were done.	✓	✓	✓		✓		✓	✓
Hermessi H et al. [10]	2021	Classification: (Feature, Decision, and Pixel level). Well-renowned fusion techniques architecture illustrated. CT/MRI, MRI/PET, MRI/SPECT visual and statistical comparison. Comprehensive fusion metrics description. Techniques summary with limitations.	✓	✓	✓		✓		✓	✓
Sebastian J et al. [11]	2021	Classification: (Spatial, Transform, and DL) Advantages and drawbacks of each technique. All articles included in this review are related to brain imaging.	✓				✓			
Diwakar M et al. [6]	2021	A comprehensive review of SWT and DWT techniques. A fair comparison of visual and quantitati on Harvard medical dataset. The few most common fusion metrics are described.		✓	✓		✓		✓	✓
Zhang H et al. [12]	2021	The entire article focuses on DL techniques (CNN, AE, and GAN). MRI/PET visual results with standard statistical results were presented.		✓	✓		✓		✓	
Our work	2022	Classification: (frequency fusion, spatial fusion, decision-level fusion, deep learning, hybrid fusion, and sparse representation). Five MMIF databases description. Comprehensive fusion metrics. Disease-based fusion summarization. Qualitative and quantitative results on various techniques.	✓	✓	✓	✓	✓	✓	✓	✓

We present a review of multimodal medical imaging modalities by overcoming all the shortcomings of the previous works, thoroughly discussing such modalities, freely accessible multimodal databases,

classification of medical image fusion techniques, and associated diseases. The MMIF approaches are classified into five domains: spatial, frequency, deep learning, hybrid, and sparse representation fusion. These five MMIF methods are compared with each other based on fusion quality performance metric results. In addition, some recent multimodal image fusion articles of different disease detection are also summarized. The arrangement of this review article is highlighted in Fig. 2.

The major contributions of this paper are as follows:

- Classification of medical imaging modalities based on the electromagnetic spectrum and invasive/non-invasive methods are shown in Fig. 3 and Table 2, respectively. The categorization based on the energy source used for image acquisition is presented in Fig. 4.
- The detailed comparisons of all online public accessible multimodal medical databases are presented in Table 3. The frequency distribution of the usage of each database over the last five years is shown in Fig. 6, and the articles compared in this review using each mentioned database are summarized in Fig. 6 (c).
- The overall general and basic procedure of medical image fusion techniques are presented in Figs. 11and 12.
- The classification of medical image fusion techniques based on frequency fusion, spatial fusion, decision-level fusion, deep learning, hybrid fusion, and sparse representation fusion is shown in Fig. 13.
- Comparison of visual image results obtained using various MMIF approaches implemented in this is shown in Fig.14.
- The taxonomy of existing works based on different diseases is summarized in Table 6.
- All the possible MMIF quality assessment metrics are explained in Table 7.
- The medical image fusion statistical results of the five different methods using one large database, i.e., AANLIB, are tabulated in Table 8.

Section 2 contains a full discussion of medical modalities. Section 3 discusses open access multimodal medical databases. Section 4 discusses the method of medical image fusion and its categorization discussed in Section 5. Section 6 illustrates the diseases related to each technique. The fusion quality assessments metrics are explained in Section 7, while the review article is concluded in Section 8.

2. Medical Imaging Modalities

In the medical field, each imaging modality has unique information and characteristics. The different medical imaging modalities used for screening and diagnoses of different diseases of the human body lie in the entire electromagnetic (EM) spectrum range, as shown in Fig. 3. Each imaging modality has a different wavelength and frequency and also shows different characteristics [13]. When EM waves strike an object, they are scattered, reflected, or absorbed by the object. MRIs generate a magnetic field that induces the protons in the body to align. Protons are activated and spin out of phase with the magnetic field when a radiofrequency current is applied to the patient. The modalities of X-ray and CT images are highly radiative and harmful to the human body due to their high-frequency range from 3×10^{16} to 3×10^{19} Hz. Nuclear imaging modalities, such as PET and SPECT, use gamma rays for

diagnosing functional activity in the organs of a human body. Gamma rays have a frequency greater than 10^{19} Hz and wavelength less than 10 picometers. X-rays, CT, SPECT, and PET are based on the principle of ionization, while MRI and US are based on non-ionization [14]. Some modern medical imaging is also used nowadays for endoscopic imaging such a While light and Narrowband imaging [15].

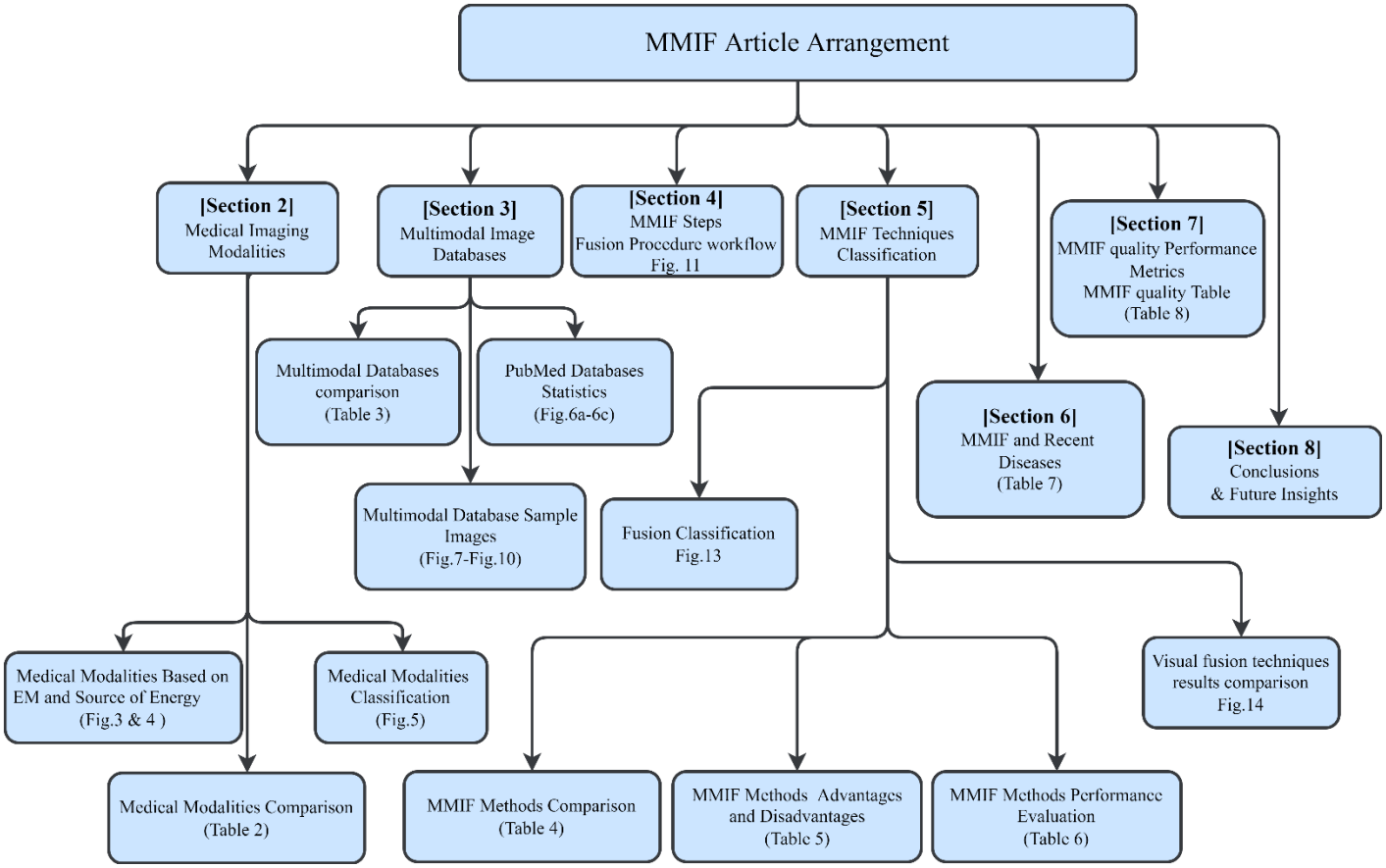

Fig. 2 Arrangement of sections encapsulated in this article.

In addition, the image acquisition process of each modality is different. Some use an internal source of energy for image acquisition, some employ an external source, while others utilize a combined internal and external energy source. The classification of different imaging modalities based on the energy source used for image acquisition is shown in Fig. 4.

Another classification of medical imaging modality is based on invasive and non-invasive methods, as shown in Table 2. Invasive methods refer to the examination of an organ by inserting any object into the body via an incision or needle injection, whereas non-invasive methods utilize some kind of radiation or sound waves. The major classification of medical imaging modalities is shown in Fig. 5.

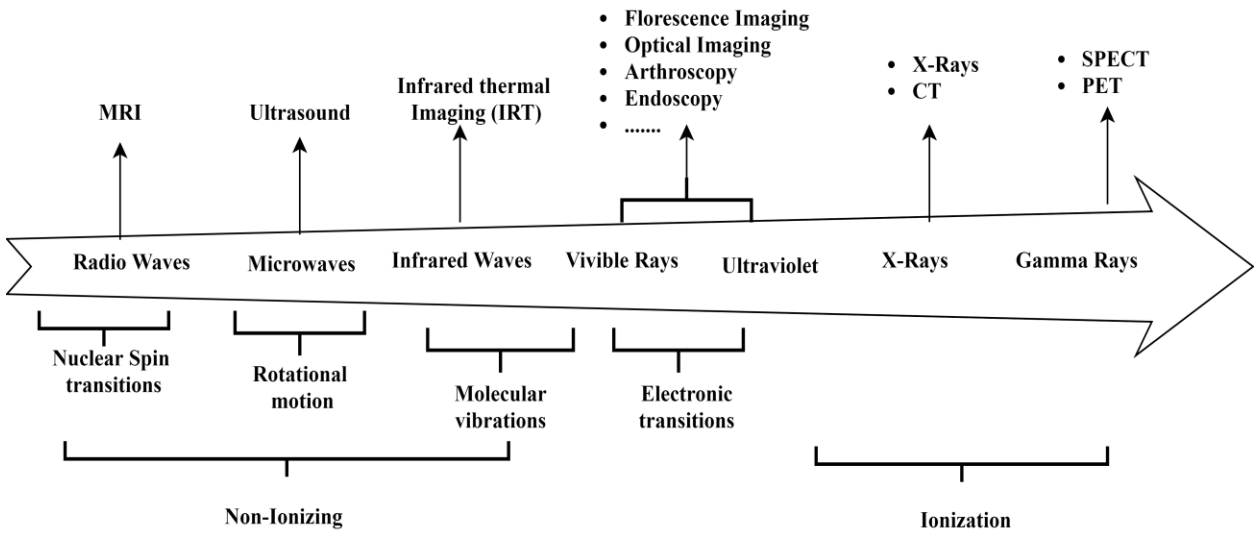

Fig. 3 Biomedical image modalities w.r.t. electromagnetic spectrum.

2.1 Radiology

For diagnosis and treatment purposes, this type of medical imaging uses ionization and radio waves, gamma rays, and sound waves. X-rays, computed tomography (CT), and magnetic resonance imaging (MRI) all involve radiation (ionization and radio waves), whereas the US uses sound waves and PET and SPECT use gamma rays for diagnostic reasons.

X-ray images, first introduced in 1895, are the oldest imaging modality used in the medical field. The X-ray modality is used to diagnose the anatomical structure of broken bones [16]. Although CT images fall under the X-ray imaging domain, X-rays use radiation in one direction, whereas CT rays are emitted into a patient's body from different angles and distances. CT images are presented in the form of cross-section areas of an organ and are comprised of many 2D slices of images, from which 3D CT images can be obtained [17]. CT images are used to diagnose bone fractures,

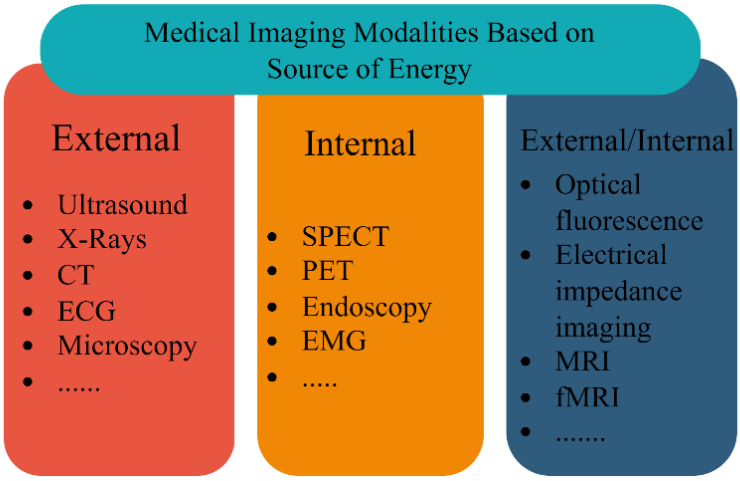

Fig. 4 Biomedical image modalities w.r.t source of energy.

locations of tumors, cancer, cardiac tissues, pulmonary embolism, etc. MRI medical images employ a strong magnetic field inside the human body that induces different orientations of protons. Hence, more detailed images consist of many sequences of slices. Compared to CT images, MRI images contain higher soft-tissue information but require a longer acquisition time. MRI images are also used to examine blood vessels, brain and breast tumors, abnormal tissues, spinal injuries, etc. [17]. MRI is usually non-invasive but sometimes it may be invasive depending on the application and requirement of imaging, with references in many articles MRI is considered as non-invasive so we classified it as a non-invasive group.

Functional MRI (fMRI) is a type of MRI that is used for acquiring functional information of the body, such as brain activity, blood oxygenation level, etc. [18]. The US modality utilizes sound waves and is used to analyze tumors, identify cancer, perform prostate biopsies, and observe the movement of a pre-birth baby in the fetus. [2][12]. Angiography is an invasive technique, in which a thin long tube is inserted into an artery via a small incision to reach the targeted area. After a contrast agent is injected into this tube, a series of X-ray images are obtained to detect the vessels and examine blood flow in arteries of the heart, brain, eye, and hand. SPECT and PET are also invasive techniques, in which radioactive isotopes are injected into a patient's body to obtain functional information for diagnosing cancer in the brain or liver, analyzing blood flow in soft tissues, and detecting brain and pancreatic tumors [2][12][17].

2.2 Visible light Imaging

In visible-light imaging, patients are subjected to visible light source rays instead of invisible rays, from which color or gray-scale images are produced. Dermatology and endoscopy medical images are two important examples of visible light imaging modalities [20]. The dermatology modality is applicable on the skin by using light rays to evaluate skin diseases, such as skin allergies, lesions, infections, and bacterial diseases. The endoscopy modality also uses visible light to examine the internal structure of a human body. These types of images are useful for pre-surgery, during surgery, and post-surgery diagnoses. A small optical camera is usually attached to an endoscopic tube to acquire medical images. Ophthalmology is an application of visible light imaging to study the structure, function, and diseases of the eyes [21]. Another organ modality is otorhinolaryngology, which is the study of the nose, throat, ears, and associated structures of the neck and head.

2.3 Microscopy Imaging

In microscopy modalities, the medical information of small-scale biological objects that cannot be obtained from ordinary other modalities is collected [22]. The two major types of microscopes are electron microscopy and light (optical) microscopy. In an electron microscope, which includes scanning and transmission electron microscopes, the diagnostic object is exposed to a beam of electrons. In a light (optical) microscope, light rays are passed through the microscope's lens to magnify the object and obtain diagnostic medical images. Phase Contrast Microscopy (PCM), Bright Light Microscopy (BLM), fluorescence microscopy, and Dark Field Microscopy (DFM) all employ optical microscopes.

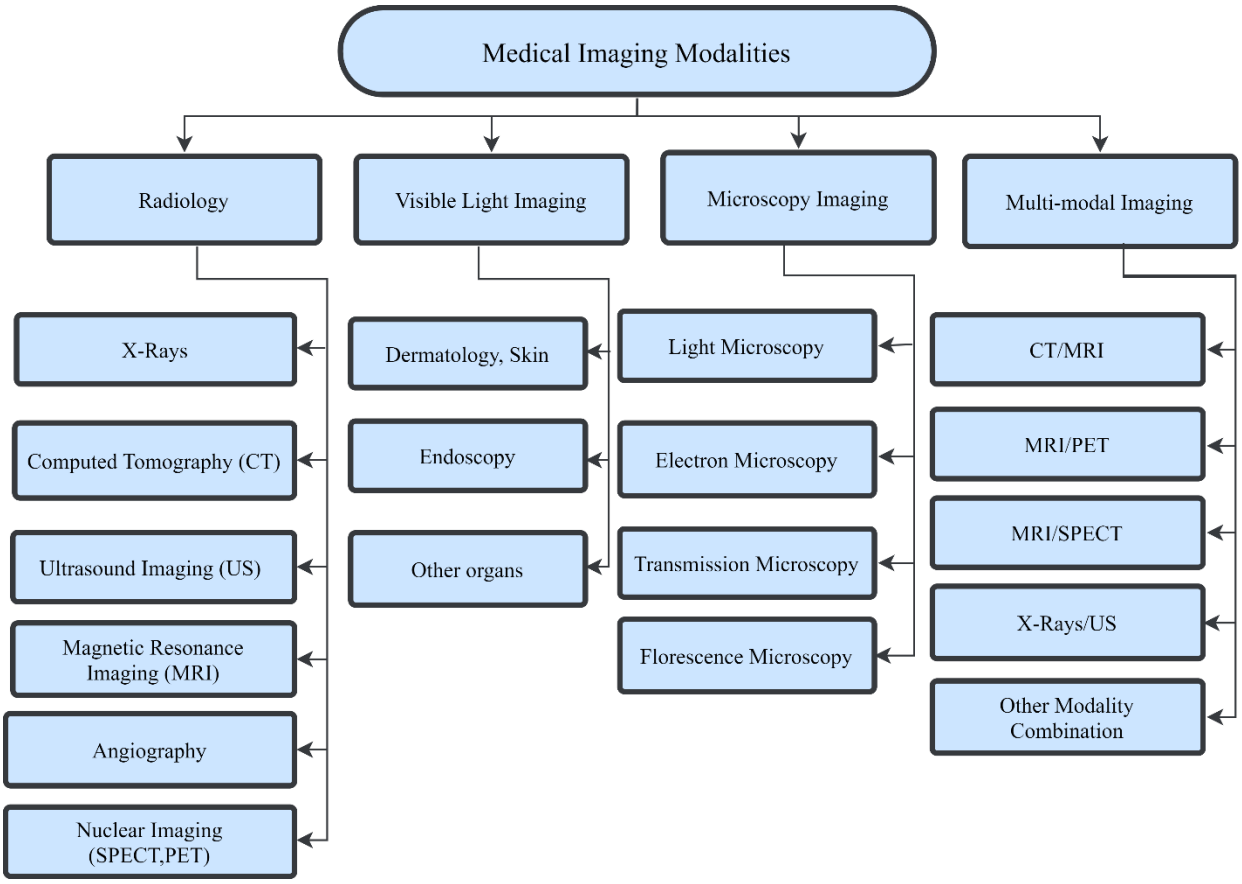

Fig. 5 Classification of medical imaging modalities.

2.4 Multimodal Imaging

Individual imaging modalities can only collect limited medical information. Specifically, anatomical and rigid structure information can only be acquired from CT and X-rays images, while structural with some functional information can be collected from MRI images. Moreover, SPECT and fMRI images reveal only detailed functional information. Medical experts often face difficulties in diagnosing diseases when using a single imaging modality due to the limitation of information. Therefore, the multimodal imaging modalities concept has been introduced, which fuses more than one modality with another via a fusion algorithm to achieve a detailed resultant image. Popular combinations of multimodal modalities include CT-MRI, MRI-SPECT, MRI-PET, X-ray and US, MRI-US, among others [23]. The comparison between the most common imaging modalities is described in Table 2, indicating the characteristics and invasive/non-invasive classification of each modality.

3. Multimodal Image Databases

In this section, we discuss some important multimodal medical datasets that are a crucial initial step of any multimodal medical image fusion technique, particularly for testing and diagnoses. Although many datasets are freely available online for experimental purposes, researchers need medical images from different modalities of the same and different patients for validation of the fusion algorithm. These

images are mostly acquired within the same periods, while other images are taken at different periods for better diagnoses and to evaluate disease progression/recession.

Herein, five free online multimodal databases, namely OASIS, TCIA, BrainWeb Atlas (AANLIB), ADNI, and MIDAS, containing thousands of images were selected for comparison. These datasets are useful for researchers to analyze their image fusion algorithms. The overall comparison of these databases is given in Table 3. These databases were selected because they: 1) are freely available, 2) contain different modalities of images of the same and different patients, and 3) contain many scanned body organs with different diseases and also in different image formats.

Table 2. Comparison of the most prominent MMIF imaging modalities.

Medical Modalities	Abbreviations	Invasive/ Non-Invasive	Characteristics
MRI	Magnetic Resonance Image	Non-Invasive	MRI images show great details about soft and abnormal tissues in the brain and blood flow. MRI can identify a small stroke, lesion, etc., but cannot provide information about activity in the brain or cancer cell information [17], [18].
CT	Computed Tomography	Non-Invasive	CT images show skull, bone structure, large abnormalities, stroke, bleeding, and lesions in the brain, but do not provide great detail about soft tissues [17].
X-Rays	–	Non-Invasive	X-rays are used to diagnose the anatomic structure of a human body. Normal, abnormal, and fractured bones can easily be seen from X-rays images [16].
US	Ultrasound	Non-Invasive	The US works on sound waves of about 20 kHz to several GHz frequency ranges. It is low cost and safer compared to X-Rays and CT images. The US is applicable on both anatomical and functional body organs [2], [19].
PET	Positron Emission Tomography	Invasive	PET scans reveal activity in the brain, as well as sugar activity as energy in the brain. They also give information about cancer cells in the brain using different colors [17][19].
SPECT	Single Photon Emission Computed Tomography	Invasive	SPECT scans indicate the most active area of the brain by analyzing blood flow. SPECT can detect diseases like dementia, seizures, clogged blood vessels, etc. [17], [19].
Endoscopy	–	Invasive	The internal structure of the human body can be seen from endoscopy using a small camera. The main application of endoscopy is to examine the internal structure of the stomach [20].
Microcopy	–	Non-Invasive	Light (optical) microscopy and electron microscopy are used to examine biological information at the micro level [22].

The “Open Access Series of Imaging Studies” (OASIS) database mainly contains medical brain images and three projects, i.e. OASIS-1, OAIS-2, and OASIS-3. The OASIS-1 project includes MRI-T1 and T2 images of a total of 416 subjects, both men and women, ranging 18-96 years of age [24]. The OASIS-1 dataset also contains information about Alzheimer’s disease (AD) and 20 nondemented diseased subjects. The Cancer Imaging Archive (TCIA) mostly includes medical images in Dicom image format [25], different imaging modalities, and multiple body scanned areas. TCIA contains a total of 22 modalities, in which most images are from mammography (MG), followed by CT than MRI images. This database also provides images from a total of 52 anatomical body organs, such as the breast, chest, brain, and colon. Harvard Medical School provides the WholeBrain Atlas (AANLIB) brain image dataset, this dataset is online publically accessible [26]–[28], AANLIB dataset is mainly

categorized into normal and diseased based brain images. Normal brain images are in 2D or 3D, while diseased images are further classified into sub-brain diseases, including brain stroke and tumors, degenerative and infectious diseases, and many other brain-associated diseases. All the images in this database are in GIF file format and easy to use. The AANLIB database focuses on brain images and contains MRI, CT, PET, and SPECT imaging modalities. The "Alzheimer's Disease Neuroimaging Initiative" (ADNI) also provides Alzheimer's brain images from three modalities: CT, MRI, and PET. The MIDAS medical database contains many medical modalities images of various body organs in different image formats [29]. In MIDAS, the RIRE project designed in 2010 focused only on multimodal medical image processing [24].

The PubMed medical database, an online biomedical search engine containing more than 30 million biomedical and life sciences citations, was used for gathering statistical results about these datasets. Herein, results from the past five years were collected from PubMed. The frequency distribution of multimodal fusion articles using these five medical datasets over the previous five years is given in Fig. 6 (a). We collected only publications that were related to these five public open-access databases; moreover, we used a subfilter search in PubMed to identify the most often occurring organ diseases in the human body that would be treated with MMIF. However, because these subfilter searches provide comprehensive information, we were able to eliminate a large number of publications, as mentioned before in Fig.1.

The following keywords were used to search PubMed for the results of the five mentioned databases over the last five years.

- For the OASIS database, keywords (medical image fusion) AND (OASIS OR https://www.oasis-brains.org/) were used, from which 18 research articles were extracted.
- For the AANLIB database, keywords (medical image fusion) AND (AANLIB OR Harvard medical OR http://www.med.harvard.edu/AANLIB/home.html) were used. It can be observed that most of the articles used this database because it contains a variety of modalities and is easy to use. Most articles on medical image fusion were obtained from this database compared to the other four.
- For the ADNI database, keywords (medical image fusion) AND (ADNI OR http://adni.loni.usc.edu/) were used.
- For the TCIA database, keywords (medical image fusion) AND (tcia OR cancer imaging archive OR https //www. cancer imaging archive .net/) were searched.
- For the MIDAS database, keywords (medical image fusion) AND (RIRE OR https://www.insight- journal.org/) were utilized.

The frequency distribution chart of the multimodal databases of the articles used in this review is shown in Fig. 6 (b). Furthermore, the articles in the multimodal image fusion domain examined different body organs for clinical diagnoses of specific human organs. The fusion articles obtained from PubMed based on different body organs, namely Brain, Lungs, Eye/Retina, and Cardiac, are also categorized in Fig. 6 (c). Again, the PubMed medical database was utilized along with keywords (Multimodal medical image fusion AND ("organ name" OR dataset). It can be observed that most articles contain information on the Brain, while only 37 articles are on the Eye/Retina.

Fig. 7 displays some sample images from the Harvard medical database [30], and Fig. 8 shows healthy MRI and MRI images of dementia and Alzheimer's diseases from the OASIS database [24]. The sample brain images from the MIDAS database [24] are displayed in Fig. 9, and those collected from ADNI [31] are provided in Fig. 10.

Table 3. Multimodal database comparisons for the application of MMIF

Dataset	Year	Modalities	Body Organ(s)	Images Categories	Format
OASIS	2010	MRI and PET	Brain	OASIS-3 (2000 MR sessions and PET three different tracers). OASIS-2 (373 MRI sessions with nondemented and Alzheimer's diseases) OASIS-1 (434 MRI sessions with Alzheimer's diseases).	Nifti
TCIA	2014	Mammography, X-rays, US, CT, MRI, PET, SPECT	Brain, Chest, Lungs, Breast, Abdomen, Kidney, Heart, Neck, etc.	Contains a total of 22 modalities and 52 anatomical scanned images. Multiple diseases of human organs.	Dicom
BrainWeb Atlas	1995	CT, MRI, PET, SPECT	Brain	Contains normal and disease-based images (brain strokes, tumors, degenerative, infectious diseases, etc.) and also normal 3D brain images.	GIF
ADNI	2003	MRI, FMRI, PET	Brain	Contains MRI and PET imaging modalities of Alzheimer's disease.	-
MIDAS	2010	CT, MRI, SPECT, PET, US	Heart, Brain, Bones, Head, Liver, etc.	MRI cardiac phantom scan images, CT bones, 3D US phantom, CT scan of the abdominal phantom, Brain MRI, CT, PET, and SPECT.	Dicom Raw, Mhd, hdr

4. Fusion Steps

The multimodal image fusion method is a procedure that integrates many images from one or various imaging modalities to increase accuracy and quality while preserving the complementary information of the images [25]. Medical image fusion mainly concerns MRI, PET, CT, and SPECT [2]. PET and SPECT modalities deliver images with functional information of the body, such as details about metabolism, soft tissue movement, and blood flow, despite having low spatial resolution. MRI, CT, and US provide high spatial resolution images, lending anatomical information about the body. Multimodal images are usually obtained by merging functional images with structural images to yield better information for health specialists to diagnose clinical diseases.

Medical image registration is used to geometrically align two images, then the image fusion technique is applied by overlapping the two input source images to produce a resultant image with excess and complementary information [19]. Two points must be satisfied during the image fusion step: 1) all the appropriate medical information that exists in the input images must be present in the resultant image, and 2) the fused image should not have any extra information that did not exist in the input images. Fusion can be applied on multi-sensor images obtained from a different source of imaging modalities, multi-focus images that are usually taken from the same modality, and multimodal images widely used in medicine. The MMIF procedure, its various steps, and the multimodal fusion classification and methods are described in this section.

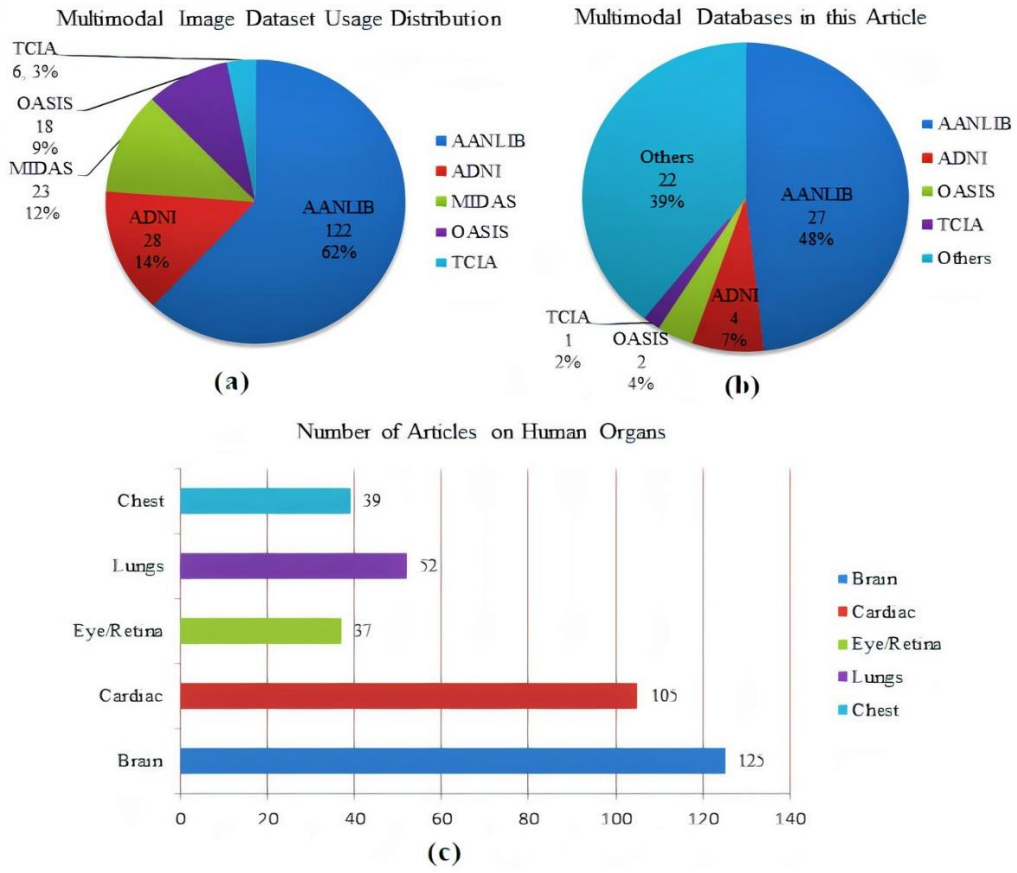

Fig. 6 PubMed database statistical results of MMIF (a) Distribution of usage frequency of multimodal databases w.r.t. medical image fusion of recent five years; (b) Frequency distribution of multimodal databases used in this study and (c) Number of multimodal fusion articles from PubMed w.r.t. human organs in the last five years.

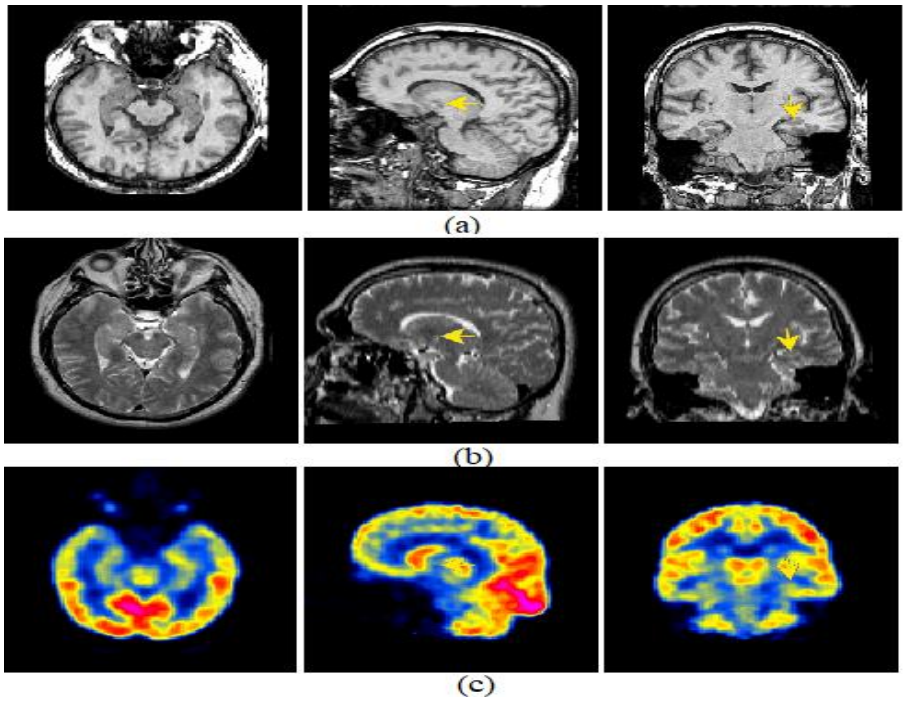

Fig. 7 Harvard medical dataset *(open-access public database)* of normal brain 3D images of slice-60 [30][32]: (a) MR-T1, (b) MR-T2, and (c) PET brain images of Trans-axial, Sagittal, and Coronal directions (left to right).

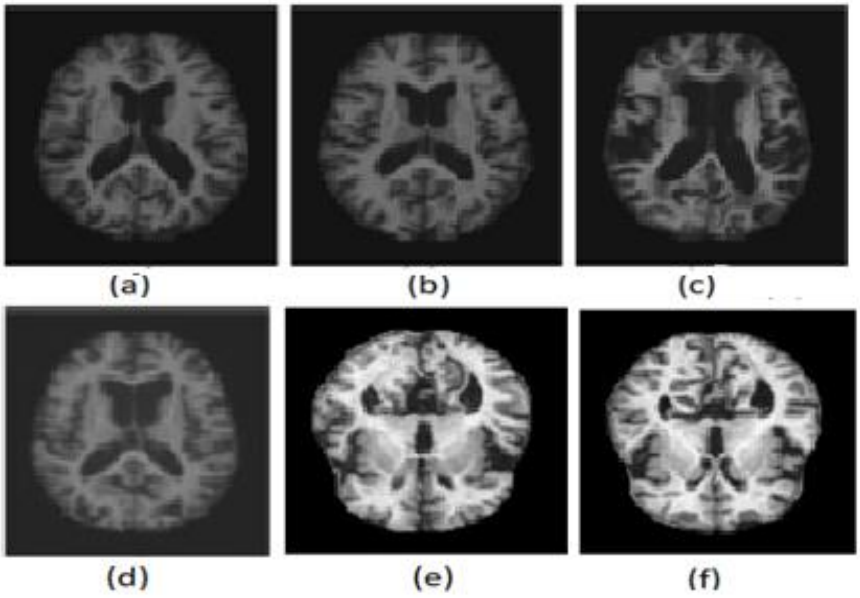

Fig. 8 MRI images from OASIS dataset *(open-access public database)* [24]: (a) non-demented aging disease, (b) high-grade mild dementia, (c) mild dementia disease, (d) less extreme dementia disease, (e) Alzheimer's Disease (AD), and (f) healthy control (HC).

In the multimodal fusion process, initially, the researcher selects the body organ of interest. Then, two or more imaging modalities selected to fuse using the appropriate fusion algorithm. To validate the fusion algorithm, performance metrics are required [34]. In the final step, the resultant fused image contains more information about the scanned area of the body organ than the input images. The overall MMIF procedure shown in Fig. 13.

In the initial stage, the input source image is registered by mapping it with the reference image to relate and match the equivalent images depending on specific features for the image fusion procedure [35]. In the image decomposition stage, the input images are decomposed into sub-images and fusion coefficients using the fusion algorithm. Then, fusion rules are applied to extract important information and multiple features of sub-images that help in the subsequent steps. Finally, the fused image is reconstructed by combining the sub-images using an inverse algorithm known as image reconstruction [36].

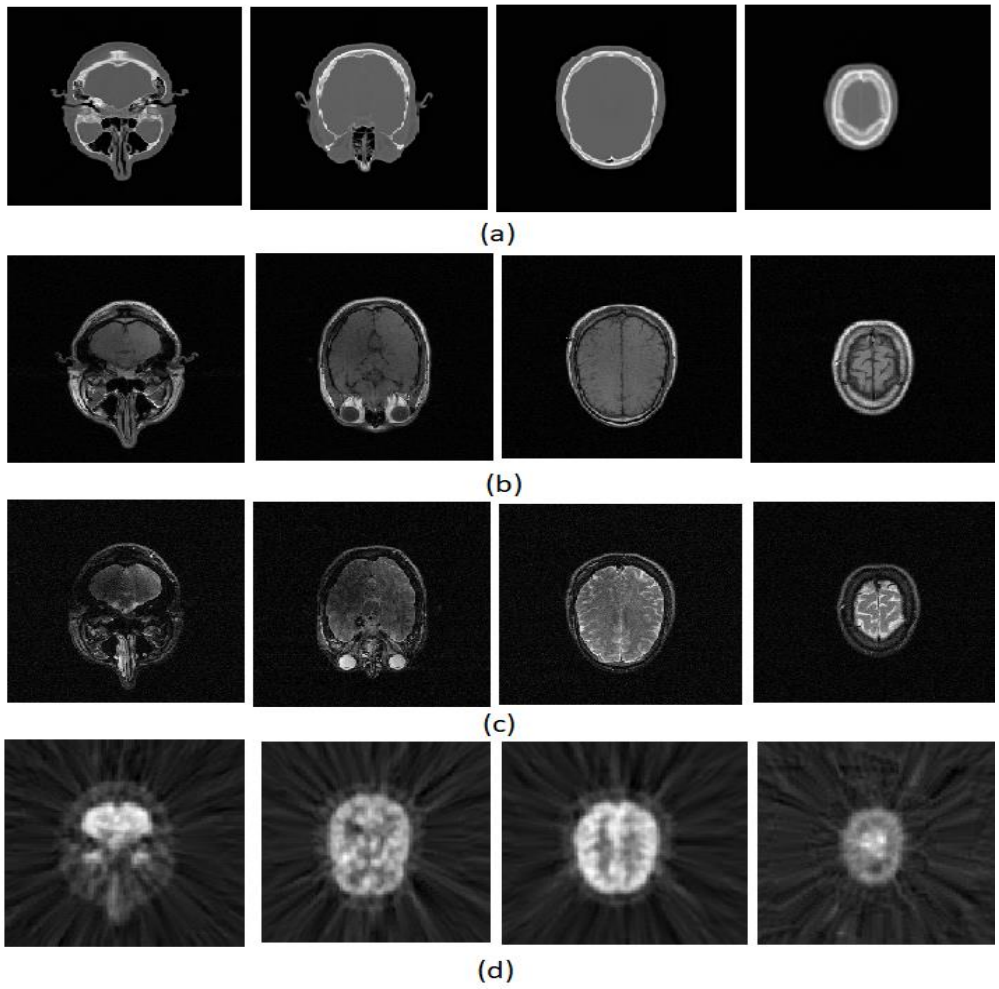

Fig. 9 From MIDAS dataset *(open access public database)*: (a) CT images, (b) T1-weighted MRI images, (c) T2-weighted MRI images, and (d) PET images [24][33].

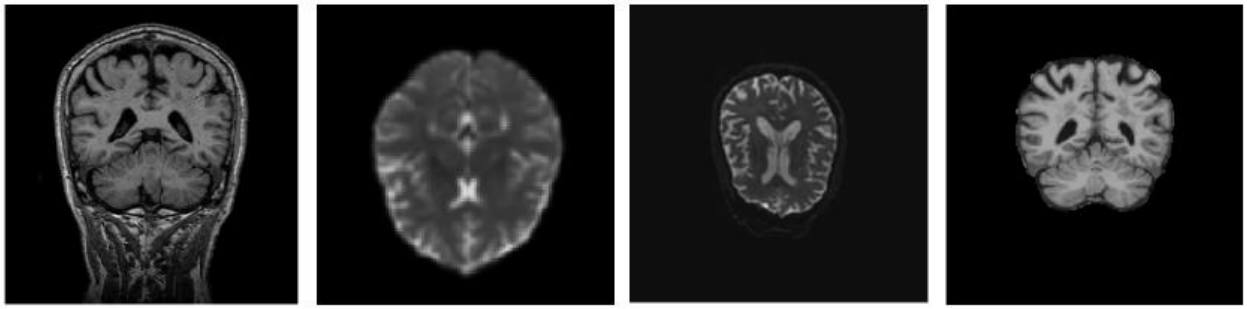

Fig. 10 From ADNI dataset *(open-access public database)*: (a) raw MRI coronal image, (b) axial diffusion tensor image (DTI), (c) raw axial DTI, and (d) post-processed MRI image [31].

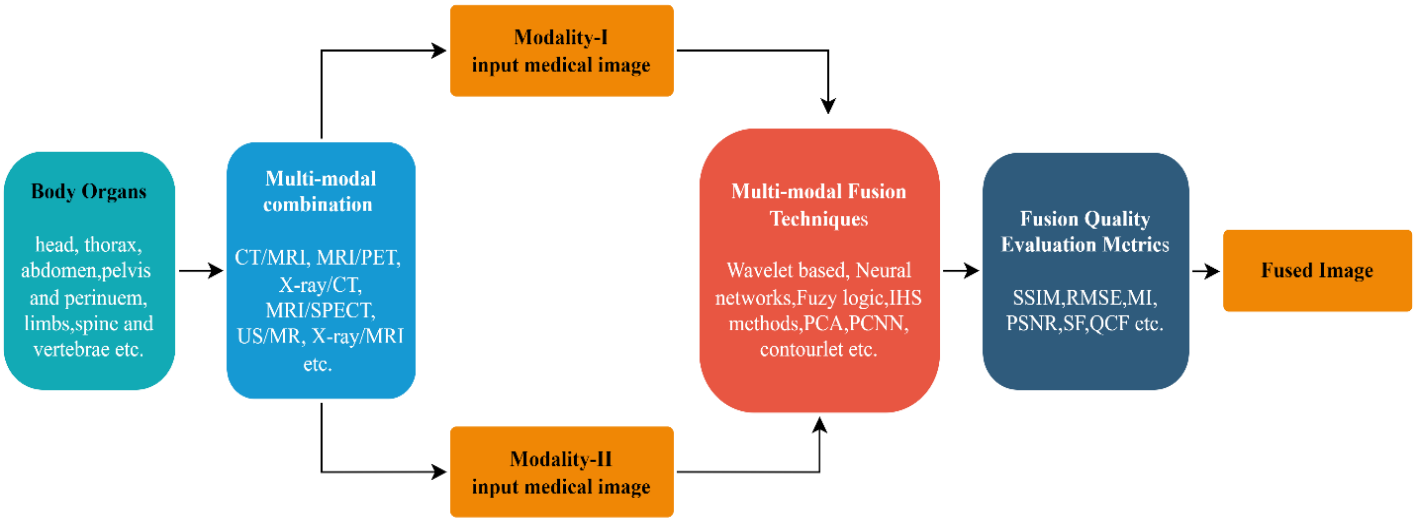

Fig. 11 Overall multimodal medical fusion procedure.

As seen, the image fusion method consists of various simple stages that help to accomplish the objective. Fig. 14 demonstrates the principle steps associated with the image fusion methodology.

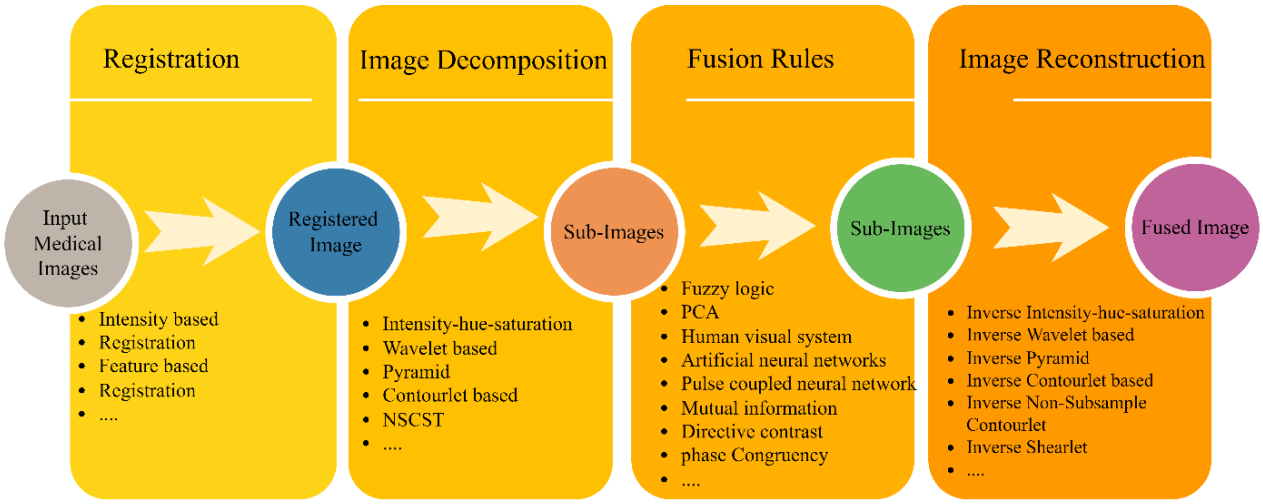

Fig. 12 Medical image fusion step-by-step process.

5. MMIF Techniques Classification

Although there are several image fusion approaches, we concentrated on six dimensions: frequency fusion, spatial fusion, decision-level fusion, deep learning, hybrid fusion, and sparse representation fusion. We begin by providing an overview of pixel-level and feature-level image fusion to aid in the comprehension of broad categorization.

In pixel-level-based fusion methods, images are combined straightforwardly utilizing singular pixels to make the fusion decision [27][37]. It further classified into the frequency domain (section 5.1) and spatial domain (section 5.2). Spatial domain techniques utilize a basic pixel-level strategy, i.e. functioning on the pixel level of images. Since it is implement on the initial images, the resulting

images exhibit less spatial distortion and a lower signal-to-noise ratio (SNR). The methods under this domain are Intensity Hue Saturation (IHS), Simple maximum, Independent Component Analysis (ICA), simple minimum, Principal Component Analysis (PCA), weighted average and simple average method, etc. In the frequency domain method, the images are first convert into the frequency domain via Fourier transform. Then, the fusion process is implement on the frequency quantities, followed by inverse transformation to obtain the final fused image. This method is further characterize into transform fusion techniques and pyramidal fusion methods [32]. Pyramidal techniques are utilized for multi-resolution analysis, while transform fusion methods include wavelet decomposition, Contourlet (CT), and Curvelet Transform fusion, etc. [33][34].

In the feature-based techniques, the complementary features are extracts from input medical images, such as different regions, edges, dimensions, images segments, and shapes [28]. Many researchers have demonstrated that the fusion of medical images based on a combination of regions or objects of images gives more significant fusion results over the fusion of images based on pixel-level [56]. The feature level-based methods are divided into region-based and machine learning-based fusion. Neural Network, PCNN, K-means and Fuzzy clustering are the common machine learning techniques. In region-based techniques, the input images are segmented into sub-images, and then features are determined from these segments or regions.

5.1 Frequency Fusion Methods

In these methods, the input images are first converted into the frequency domain by computing Fourier transform (FT), then the fusion algorithm is applied on the transformed image followed by inverse FT to obtain the resultant fused image [34]. These techniques are further characterized into pyramidal and transform techniques [38]. The pyramidal fusion methods include Gaussian, differentiation, Laplacian and morphological pyramid, Filter Subtract Decimate Pyramid (FSD), and slope pyramid [39]. The transform fusion methods pertain to wavelet decomposition, CT and Curvelet transform. Moreover, wavelet decomposition techniques includes Discrete Wavelet Transform (DWT), Dual-Tree Complex Wavelet (DT-CWT), Lifting Wavelet (LWT), and Redundant Discrete Wavelet (RDWT) methods. Examples of contourlet transform fusion methods are Dual-Tree Complex Contourlet Transform (DT-CCT), and Non-Subsampled Contourlet (NSCT) [40]. Additionally, ridgelet, curvelet, and bandlet are types of curvelet transform fusion methods [41].

Qu *et al.* [42] provided a fusion rule for finding the wavelet modulus maxima of input images with varying bandwidths and levels. To quantify the fusion impact, a metric based on MI measurements is computed. The researchers selected CT and MRI brain input images for fusion. The benefit of this method is the preservation of both components and edge feature information in the fused image. However, this technique failed to meet the requirements of shift invariance. This problem was resolved by RDWT [43], which demonstrated improved performance over DWT due to mutual information dependent on non-linear registration and entropy image information. Moustafa *et al.* [44] introduced four multimodal image fusion methods, namely Laplacian Pyramid, Multi-focus, Wavelet Transform, and Computationally Efficient Pixel-level Image Fusion (CEMIF), which were applied on liver CT and MRI images for detection of hepatic lesions.

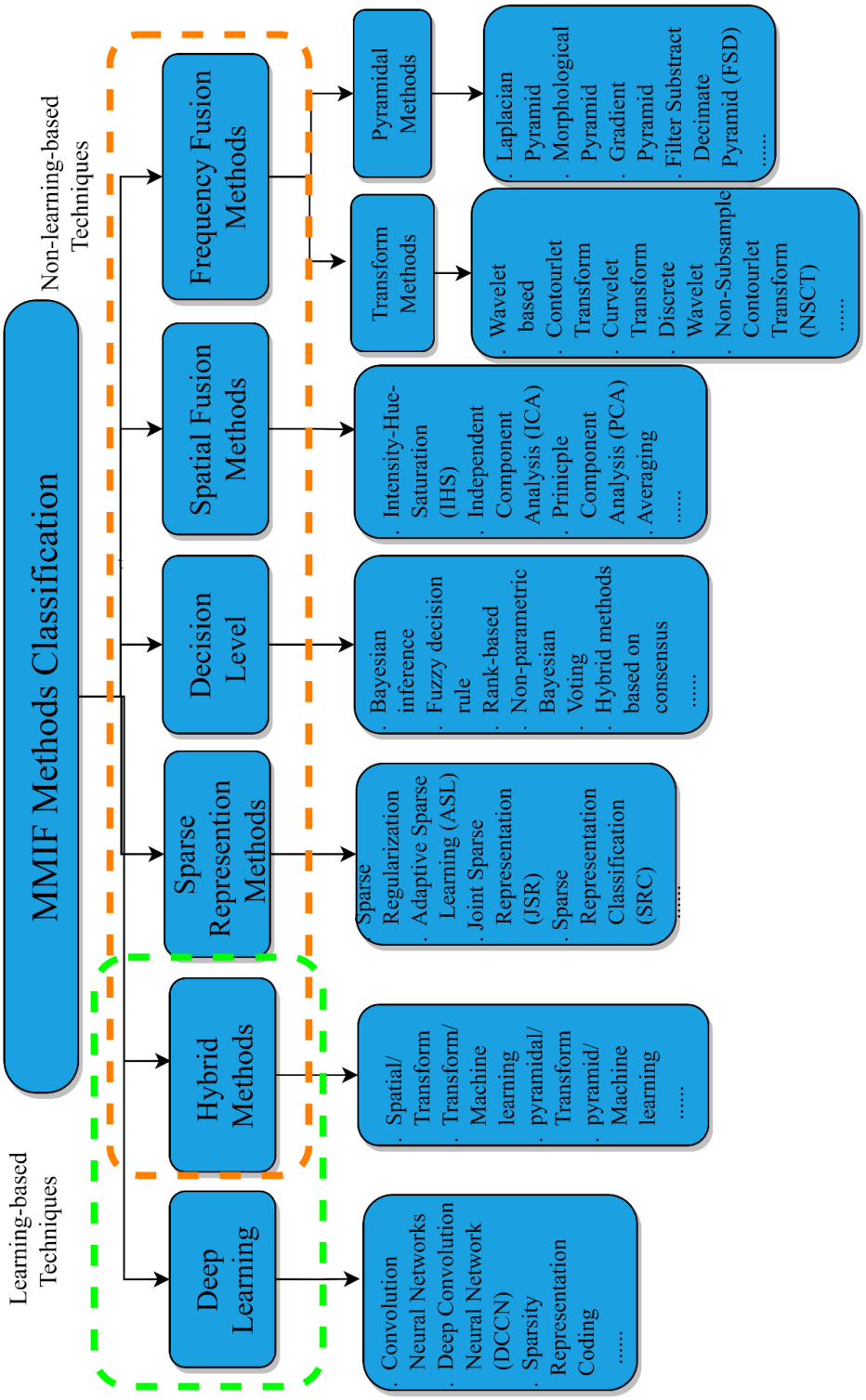

Fig. 13 Classification of medical image fusion techniques.

Liu *et al.* [45] recommended a multimodal image fusion technique based on Multi-wavelet Transform (MWT) and used PET and CT chest images for fusion. Yang *et al.* [46] suggested a new fusion technique based on Contourlet transform (CT), where the resultant fused image was acquired by

taking the inverse CT of the high pass sub-bands. Although this technique demonstrated improved localization, anisotropy, and multiscale characteristic, it could not give significant improvement than other multiscale analysis techniques like wavelet. Three groups of CT and MRI brain images were fused based on Shearlet transform (ST), which yielded the best sparse directional image since the ST method has no restrictions on the number of orientations for sharing. Xue-jun *et al.* [47] suggested another multimodal image fusion method based on LWT to decompose the input image based on high-low frequency weight and by various fusion rules. The input CT and MRI of human brain images were considered for fusion. Bhatnagar *et al.* [48] presented a multimodal image fusion method based on NSCT. The medical images were initially converted by adding low-frequency components with high-frequency components. Two different fusion rules depending on directive contrast and phase congruency were implemented to fuse high and low-frequency coefficients. At last, the resultant fused image was obtained by taking inverse NSCT. Sahu *et al.* [49] proposed a method using Laplacian Pyramid (LP) and DCT, wherein LP decomposes the input image into a distinct low pass image, similar to a pyramidal structure. The quality of fused images increased when the pyramidal level increased, leading to improved edges and information. This method also yielded superior fused results in terms of quantitative and qualitative analysis compared to the Daubechies complex wavelet transform (DCxWT).

Xi *et al.* [50] introduced a Multimodal Medical Volumetric Image Fusion (MVIF) method, which incorporates many features acquired via multiscale geometric analysis of 3D ST using Contextual Hidden Markov Model (CHMM). Firstly, 3D ST is used to decompose the input images into low and high-frequency components. Then, the max fusion rule with local energy is applied on low-frequency sub-bands, and an effective multi-feature fusion rule is implemented for integrating components of high-frequency sub-bands. 3D ST provides better decomposition performance compared to the 2D MST tools for volumetric medical images. Ming *et al.* [51] presented a new method based on NSST, which was used for the decomposition of source medical images into multi-scale, while a “parameter-adaptive pulse-coupled neural network” (PA-PCNN) was used to determine high-frequency band in images. An energy preservation approach applied to determine the low-frequency band of medical images. The resultant fused image was constructed using the inverse NSST technique and four multimodal medical modalities: CT, MRI, PET, and SPECT. Arif *et al.* [52] suggested a multimodal image fusion method based on CT and the Genetic Algorithm (GA). GA is utilized to compute the uncertainties and rambling existing in the image and to optimize the fusion procedure. The methodology verified on brain images using MRI, MRA, PET, and SPECT modalities. All the above-cited techniques are summarized in Table 4.

5.2 Spatial Fusion Methods

Spatial fusion methods based on the pixels of images, where pixel values are manipulated to accomplish the desired results. Spatial domain methods include PCA, IHS, Brovey, High Pass Filtering methods, ICA, Simple maximum, simple average, and weighted average [19]. However, the problem with spatial domain methods is that they generate spatial distortion in the resultant fused image, which is considered a negative factor in the fusion process. Stokking *et al.* [53] proposed an HSV model for a fusion of anatomical and functional information obtained from MRI and SPECT modalities by using a

color encoding pattern. This model performed better than the RGB model and allowed rapid, easy, and intuitive retrospective determination of color encoding of the functional in the fused image. The proposed method was implemented on brain images to evaluate two HSV color manipulation techniques. Daneshvar *et al.* [48] introduced a multi-resolution image fusion technique based on the retinal model, which incorporates a Difference of Gaussian (DoG) operation. The performance results were compared with DWT, HSI, and WT methods. The retina visual-based model gives high spectral features in resultant image with less spatial distortion.

He *et al.* [54] developed an algorithm that integrates the benefits of both IHS and PCA methods to improve the fused image. PET generates medical images with reasonable color and less spatial resolution, while MRI gives suitable spatial resolution with color appearance. The suggested strategy performed better for both human perception and image quality assessments criteria. Bashir *et al.* [55] presented a model based on PCA and SWT, which was tested on a variety of medical images. Results demonstrate that in multimodal fusion, PCA achieved superior results with distinctive contrast and brightness levels. SWT seemed to give superior performance when the images were from different sources like multimodal and multi-sensor images. The above-listed techniques are tabulated in Table 4.

5.3 Decision Level Fusion

Decision Level fusion decides each input image using specific predefine criteria and then merges based on the trustworthiness of each conclusion into global optimum to make the single fused image. These types of techniques produce maximum information using certain rules defined before the fusion process [56]. Dictionary learning and Bayesian techniques are the most prevalent methods used in decision fusion. At the decision level, usually, there are three approaches integrated to get fused images, these approaches are the following (information theory, logical reasoning, and statistics approaches); include joint measures, Bayesian fusion techniques, hybrid consensus methods, voting, and fuzzy decision rules. The Bayesian approach is based on probabilities for combining data from various sensors, these techniques rely on the Bayes hypothesis. Nonparametric Bayesian, HWT Bayesian, and DWT Swarm Optimized are examples of Bayesian techniques [57].

PET and MRI images fused using the nonparametric Bayesian-based on sparse representation (SR-NPB) technique provided by [58]. For both visual and quantitative comparisons, this technique outperformed the other three SRs. In terms of execution time, the new approach is faster. Twenty sets of PET and MRI scans were used for the experiment in their research. The [59] used the BRATS database that includes the MRI brain in their fusion process. The authors proposed a Fractional-BSA-based Bayesian fusion strategy. This approach makes it possible to calculate the appropriate Bayesian parameter for fusion. Fusing images is made easier by using wavelets generated from the original images using the HWT. In [60] study, birds swarm optimization was used to develop a novel Bayesian fusion algorithm. Image fusion is being developed using the BRATS [61] database. Bayesian fusion is performed by optimizing the source images using the BSA algorithm after they have been modified using the Haar-DWT algorithm. According to the findings of the research, the strategy beat the three currently used fusion techniques - NSCT, SWT, and SWT-NSCT.

All images were integrated and matched up initially in [62] research. Multiresolution wavelet decomposition of input images gives more supplementary information than a single resolution approach provides. A comparison of genetic searching strategies using Mutual information (MI) reveals that the fuzzy clustering strategy outperforms the genetic searching technique using MI. A unique medical image fusion technique based on low-rank sparse decomposition and dictionary learning was proposed in [63] study for de-noising and enhancing medical imaging. The dictionary-learning model contains terms of regularization with a low rank and a sparse distribution. The fused image output is constructed by combining the low rank and sparse components of the source images.

5.4 Deep Learning Fusion Methods

These methods consist of multiple layers, where each layer takes input from the previous layer. Deep learning contributes to the layered structure and suitability of the complex framework architecture for large data manipulation [64][65]. The deep learning fusion methods include CCN, Convolution Sparse Representation (CSR) also known as convolution sparse coding techniques, and Deep Convolution Neural Networks (DCCNs). The CNN model of deep learning, which is more popular among all the other techniques [66] [67], is trainable and well-tuned to learn features of input data in the multilayered architecture framework. In CNN, every layer consists of several features maps that hold coefficients known as neurons. In the multiple stages, the features maps connect with every stage using different calculations, including spatial pooling, convolution, and non-linear activation. Another popular deep learning fusion technique is Convolutional Sparse Coding (CSC), which was first proposed by Zeiler *et al.* [68] and originates from the deconvolutional networks. The main target of this technique is to achieve convolutional decomposition of an image under the sparsity constraint. The multistage feature representation of the input image learns from the deconvolutional networks by developing a hierarchical structure of such decomposition. Then, the input image is reconstructed with the help of these multiple decomposition levels in a layered-wise manner. The CSR techniques have yielded promising results in terms of image reconstruction and the feature learning approach.

Wang *et al.* [69] presented a new multimodal image fusion technique based on the fuzzy radial basis function neural network (Fuzzy-RBFNN) to carry out auto-adaptive image fusion. To train the network, GA was implemented, and artificially blurred medical images were included in sample sets. The experimental outcomes showed that this method was more suitable for blurry input images compared with other traditional fusion methods. Wang *et al.* [70] proposed and tested a novel multi-channel m-PCNN for multimodal image fusion on four different medical imaging modalities, which exhibited better performance according to mutual information criteria. Teng *et al.* [71] presented a fusion technique using a neuro-fuzzy logic and hybrid approach, BP, with Least Mean Square (LMS) to train and tune parameters of the membership function. The final fused image from neuro-fuzzy logic was based on a feed-forward neural network reserved texture feature and contained enhanced useful information compared to the BP neural network technique (BPNN). Sivasangumani *et al.* [72] developed a multimodal image fusion method using regional firing characteristic PCNN (RFC-PCNN). The proposed technique was well-suited for increasing medical image fusion quality to determine brain tumors and minimize the effect of artifacts in the resultant fused image.

Liu *et al.* [73] presented a multimodal image fusion technique based on CNN, in which a Siamese convolutional network was implemented to obtain a weighted map that combines the pixel activity information from both input images. The fusion procedure is applied in a multiscale manner using pyramids, and a local similarity-based technique is used to adaptively adjust the fusion mode for the decomposition of coefficients. Hou *et al.* [74] proposed a fusion method based on CNN and a dual-channel spiking cortical model (DCSCM). Initially, NSST was used to generate low and high-frequency coefficients of the images. Then, low-frequency coefficients were fused and applied as input to the CNN framework, where a weighted map was obtained by feature maps of the image and by applying an adaptive selection fusion rule. The high-frequency coefficients were selected as an input to DCSCM. Finally, the fused image was obtained by inverse NSST. This method showed better performance than some current fusion methods. The described techniques based on deep learning MMIF are described in Table 4.

5.5 Hybrid Fusion Methods

Considering that the results of the conventional multimodal image fusion methods are not satisfactory, the basic idea behind hybrid methods is to combine two or more fusion techniques, such as spatial or transform fusion and neural network techniques, to improve the fused image quality and performance. The general advantage of the hybrid methods is to improve the visual quality and decrease the artifacts and noise in the fused images. Daneshvar *et al.* [75] proposed a hybrid fusion method based on Retina-Inspired Model (RIM) and HIS fusion methods, which can maintain high spatial features and additional functional data. The performance and visual results showed that this technique was superior to the Brovey, HIS, and DWT methods. In the proposed method, entropy, discrepancy, mutual information, and averaging gradient were used as fusion quality assessment parameters. Das *et al.* [76] introduced an image fusion technique based on NSCT with PCNN. The input medical image was initially decomposed by NSCT, while the low and high-frequency sub-bands were fused by the fusion rule 'max selection' and PCNN, respectively. The spatial frequency in the NSCT domain is used as the input to PCNN for the subsequent fusion procedure. The resultant fused image was obtained by taking inverse NSCT.

Sharmila *et al.* [77] proposed another multimodal image fusion method based on DWT-Averaging-Entropy-PCA [DWT-Av EN-PCA], which was compared with other current fusion procedures utilizing qualitative and quantitative metrics. Kavitha *et al.* [78] developed a fusion method that incorporates swarm intelligence and a neural network to accomplish a superior fusion output. The image edges were recognized and improved by utilizing Ant Colony Optimization (ACO). The detected edges were then applied as an input to PCNN. The results demonstrated that the proposed hybrid strategy performed much better than the current computational and hybrid intelligent techniques. Ramlal *et al.* [79] proposed an improved hybrid fusion strategy based on NSCT and SWT. Firstly, the input images were decomposed into various sub-bands utilizing NSCT. Then, SWT was employed to decompose the estimation coefficients of NSCT into various sub-bands. Weighted sum modified Laplacian and entropy square of the coefficients were implemented as fusion rules with SWT. The fused output image was acquired by taking inverse NSCT.

The term over-coompetness In recent years, medical image fusion has been progressively used to diagnose different completeness means that the total number of atoms in the dictionary is always greater than the dimension of image signals. Over-completeness gives a sufficient amount of atoms for dictionary learning and allows a correct representation of the signals. As one might expect, SR has attracted noteworthy consideration in the exploration field of image fusion [87][88]. Yang and Li [65] applied the SR technique to the image fusion field for the first time after such various SR-based image fusion strategies were introduced. The works reporting on the sparse representation of MMIF are provided in Table 3. The comparison of each fusion technique is summarized in Table 4.

5.6 Sparse Representation Methods

In the sparse representation (SR) method, an over-complete dictionary is obtained from a sequence of images to achieve a steady and significant representation of the source images [82]. The basic principle of SR representation is based on the treatment of an image signal as a linear combination of less significant atoms from the pre-trained dictionary learning, where the sparse coefficient shows the significant features of the input images. Sparsity refers to the fact that only a minimal number of atoms are necessary to properly reconstruct a signal, resulting in sparse coefficients.

Table 4. MMIF technique classification and major contributions in literature

Research work	Fusion Methods	Year	Multimodal combination	Body Organ	Multimodal Fusion Techniques	Dataset
Stokking *et al.* [53]	Spatial domain	2001	MRI/SPECT	Brain	color-encoding technique with HSV	-
Daneshvar [83]		2007	MRI/PET	Brain	Retina based fusion method	-
He *et al.* [54]		2010	MRI/PET	Brain	IHS and PCA	AANLIB
Bashir *et al.* [55]		2019	X-ray/CT MRI/SPECT	Brain, Leg	Principal Components Analysis (PCA)	-
Zhizhong [84]		2020	CT/MRI	Brain	Gradient similar filtering (GSF)	-
Qu *et al.* [42]	Frequency Domain	2001	CT/MRI	Brain	Discrete Wavelet Transform	-
Moustafa *et al.* [44]		2006	CT/MRI	Liver	Laplacian Pyramid, Multi-focus Technique, Wavelet Transform, (CEMIF)	-
Yang *et al.* [46]		2008	CT/MRI	Brain	Contourlet transform (CT)	-
Singh *et al.* [43]		2009	MRI-T1, MRI-T2	Brain	Redundant Discrete wavelet transform (RDWT)	AANLIB
Liu *et al.* [45]		2010	CT/SPECT	Chest	Multi-wavelet transform (MWT)	-
Xue-jun *et al.* [42]		2010	CT/MRI	Brain	Lifting Wavelet Transform (LWT)	-
Bhatnagar *et al.* [48]		2013	CT and MRI-T1/T2	Brain	Non-subsampled contourlet transform (NSCT)	AANLIB
Sahu *et al.* [49]		2014	CT/MRI	Brain	Laplacian Pyramids	-
Xi XX *et al.* [50]		2017	MRI T1/MRI T2	Brain	3-D Shearlet transform (3D-ST)	AANLIB
Ming *et al.* [51]		2018	MRI/CT/PET	Brain	Parameter-Adaptive PCNN-NSST	ADNI
Arif *et al.* [52]		2019	MRA/MRI PET/SPECT	Brain	Fast curvelet transform through genetic algorithm	AANLIB
Srinivasu *et al.* [85]		2020	MRI/PET/ SPECT	Brain	Empirical wavelet decomposition	AANLIB
Qiu *et al.* [86]		2020	CT/MRI MRT1/ MRT2	Brain	NSCT- separable dictionary learning and Gabor filtering	AANLIB
Wang *et al.* [69]	Deep	2007	CT/MRI	Brain	Fuzzy-RBFNN	-

Wang *et al.*[70]	Learning	2008	CT/MRI	Brain	m-PCNN	-
Teng *et al.* [71]		2010	CT/MRI/ SPECT	Brain	Neuro-Fuzzy Logic	-
Sivasangumani *et al.* [72]		2015	CT/MRI	Brain	Regional firing characteristic PCNN (RFC-PCNN)	-
Liu Y *et al.*[87]		2016	CT/MRI	Brain	Convolutional Sparse Representation (CSR)	-
Liu *et al.* [73]		2017	CT/MRI/ SPECT	Brain	CNN	AANLIB
Hou *et al.* [74]		2018	CT/MRI	Brain	convolutional neural networks and a DCSCM	AANLIB
Liu Y *et al.* [88]		2019	CT/MRI	Brain	Convolutional Sparsity-MCA	AANLIB
Xia KJ *et al.* [89]		2019	CT/MRI/ PET	Brain/ Abdom-en	DCNN	AANLIB/ Public Hospital
Jingming *et al.* [90]		2020	CT/MRI/ SPECT/PET	Brain	parameter-adaptive PCNN	TCIA / AANLIB
Panigrahy *et al.* [91]		2020	MRI/SPECT	Brain	Parameter Adaptive Dual Channel PCNN	AANLIB
Daneshvar *et al.* [75]	Hybrid based methods	2010	PET/MRI	Brain	IHS and Retina-inspired model (RIM)	AANLIB
Das *et al.* [76]		2012	MRI/CT	Brain	NSCT and PCNN	AANLIB
Sharmila [77]		2013	MRI/CT	Brain	DWT-A-EN-PCA	-
Kavitha *et al.*[78]		2014	PET/SPECT MRI	Brain	swarm intelligence and neural network	AANLIB
Ramlal [79]		2019	MRI/CT	Brain	NSCT and SWT	AANLIB
Lina *et al.* [92]		2020	CT/MRI/ SPECT/PET		Homomorphic filter and DWT	AANLIB
Zong *et al.*[93]	Sparse representa-ion (SR) methods	2017	CT/MRI/ SPECT/PET	Brain/Lungs	SR of classified image patches	AANLIB
Jiang *et al.* [94]		2018	CT/MRI	Brain	SR- weighted least squares filter	AANLIB
Zhu *et al.*[95]		2018	MRI/PET/ SPECT	Brain	Image decomposition and SR	AANLIB
Shahdoosti *et al.* [96]		2018	CT/MRI/ PET/SPECT	Brain	Sparse Representation Classification (SRC)	AANLIB
Zhang *et al.*[97]		2019	MRT1/ MRT2/ SPECT	Brain	SR- analysis-synthesis dictionary	-
Maqsood S *et al.* [98]		2019	CT/MRI	Brain	SR-Two-scale Image Decomposition	AANLIB
Liu *et al.* [99]		2020	CT/MRI	Brain	Joint Sparse Representation (JSR)	-
Anantrasirichai *et al.*[100]		2020	CT/MRI/ OCT/fundus	Brain/ Retina	Sparse Regularization	-
Li H *et al.* [101]		2020	CT/MRI	Brain	low-rank sparse decomposition	AANLIB
Das A et al [62]	Decision level	2009	CT/ MRI-T1, MRI-T2	Brain	evolutionary algorithm	AANLIB
Huafeng Li et al [63]		2018	CT/ MRI-T1, MRI-T2	Brain	low-rank sparse dictionaries learning	AANLIB
Shabanzade F et al. [58]		2019	MRI/PET	Brain	nonparametric Bayesian	AANLIB
Bhardwaj J et. Al [59]		2020	MRI-T1, MRI-T2	Brain	Fractional BSA Bayesian	BRATS
Nayak A et al. [60]		2020	MRI-T1, MRI-T2	Brain	BSABayesian	BRATS

To examine the visual performance of each MMIF main approach, we used qualitative findings from the Harvard Atlas [30] database. MRI, CT, and PET images were used as source images for visual

assessment. All tests were carried out on a Windows 10 computer running MATLAB 2020a and Python 3.8, with an Intel ® Core I716010U CPU running at 1.8 GHz and 16 GB of RAM. Figure 10 depicts a visual examination of brain fusion imaging. We compared the visual findings of Li S et al. [102] (Source Link-1), Zhu Z et al. [103] (Source Link-2), Zhang Y et al. [104] (Source Link-3), Yin M et al. [51] (Source Link-4), Liu Y et al. [81] (Source Link-5) and Das S et al. [105] (Source Link-6). The source codes for each approach are all available online through the source links. The advantages and drawback of each fusion techniques summarized in Table 5.

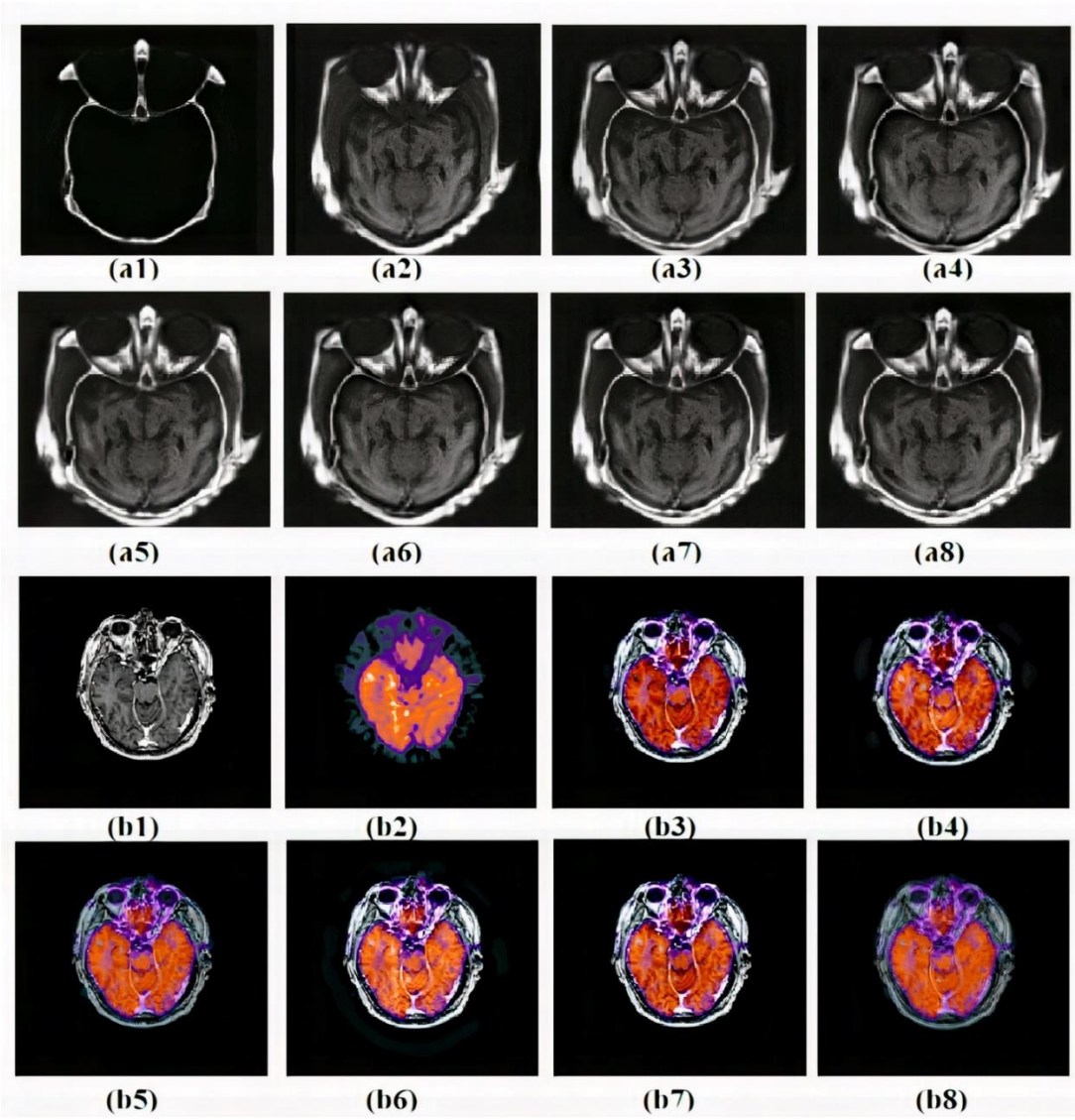

Fig. 14 Comparison of visual image results obtained using various MMIF approaches. (a1) and (a2) are CT and MRI images where (b1) and (b2) are MRI and PET source images. (a3-a8) and (b3-b8) are qualitative fusion results of GFF [102], LE-NSCT [103], CNN [104], NSST PAPCNN [51], MST SR [81] and N-Fuzzy [105] fusion technique respectively.

Table 5. Comparison of MMIF techniques

Techniques	Advantages	Disadvantages
Spatial based	• The easiest image fusion process. • Results in highly concentrated image results produced from the image input. • Highly simplistic, easy to understand, and enforce. • Processing and high sharpening capacity are very quick, computationally efficient, and faster.[106][107]	• The resultant fused image is not promised to be sharp. • Image contrast is diminished by the blurring effect. • Usually, spatial domain fusion causes spectral degradation.[106] • Colors may become distorted and spectrum degradation may occur.[107]
Frequency-based	• Performs better than standard fusion methods in terms of reducing the spectral distortion. • Gives a higher signal-to-noise ratio than an approach based on the pixel level method. • Enhanced results are produced by multilevel fusion where the image is fused In the medical field twice using an appropriate fusion technique. • Both high spatial resolution and decent quality spectral components are included in the output image. • For multi-focus images, strategies provide excellent detailed image quality. [108][109]	• The final fused image may have a less spatial resolution in this process. • The fusion algorithm requires a more complicated procedure than pixel-level techniques. • A good fusion technique is needed for a better result. • Generate output images that are more or less identical.[108][109]
Deep learning	• Neural networks' learning environment makes it much easier to optimize the process of image fusion. • Various input data increase the focus on fusing high-dimensional data to yield a feasible solution. • The methodology can be customized to the needs of the application. • Produce excellent results when there is a lot of input data compared to other fusion techniques. [107]	• Based on dynamic processes with complex parameters. • Many challenges must be addressed, such as local extremum, misidentification, and pace of training convergence. • Requires more time and hardware specification to train the fusion model. • Does not give accurate results for small datasets of images. [107]
Hybrid based	• Complementary detailed information is achieved from the input images. • Improve clarity, contrast, texture, brightness, and edge information in the fused image. • Hybrid techniques prevent difficulties at the pixel level, such as too sensitive noise and blurring. • Minimizes the artifacts in the resultant images.[110]	• Requires detailed knowledge of each technique, otherwise, non-uniform fusion results are produced. • Has a typically complex and time-consuming fusion process. • Cannot be used for large input datasets.[110]
Sparse Representation	• SR coefficients are the most significant factors that improve the final fusion performance. • Retain the visual information better and improve the contrast of the image compared to other techniques. [107] • Preserve information related to the structure of images and maintain the extensive detail of source images.[111]	• Face two main disadvantages: minimal detail preservation capacity and high susceptibility to misregistration. • Often produces visual artifact results in the reconstructed image.[107][111]
Decision level	• The quantity of superfluous and unclear information is reduced. • Increased accuracy by including the information content that is linked to each pixel in the image, which improves than feature level fusion. [112]	• The method gets increasingly difficult to master as time goes on. • In addition, the method becomes more time-consuming and complicated.[112]

6. Multimodal Fusion and Recent Diseases

MMIF has demonstrated exceptional performance for analyzing diseases and improving the precision and performance of the diagnostics area. In the medical field, the information from a single imaging modality does not provide complete information of human body organs. For example, MRI images show only soft-tissue information, while CT images display bone density information. Thus, MMIF becomes a vital zone of research because of its significance for providing high-quality output images for diagnostics and medical treatment [113]. Some recent disease-based multimodal fusion works are summarized in Table 6.

Lei *et al.* [114] presented a fusion technique that implements canonical correlation analysis (CCA) utilizing discriminative feature learning. Multimodal features and their CCA projections were interconnected to show each subject, where both individual and shared information from images of Alzheimer's disease (AD) were obtained. In this work, an accuracy of 86.57% for Mild Cognitive Impairment (MCI vs. NC) disease and 96.93% for AD [vs. normal control (NC)] was achieved. Ahmed *et al.* [115] suggested a multimodal image fusion method based on Multiple Kernel Learning (MKL) to fuse visual features obtained from structural MRI and Diffusion Tensor Imaging (DTI) MD maps of brain images to differentiate between AD and MCI disease. The researchers selected both MRI-T1 and DTI image data of 155 total subjects, where 52 were the Normal Control, 45 had AD, and 58 had MCI.

Table 6. Some recent disease detection-based multimodal fusion methods

Reference	Disease/ disorder/abnormality	Year	Multi-modality	Body organ	Fusion algorithms	Dataset
Lei *et al.* [114]	Alzheimer's disease & MCI	2016	MRI/PET	Brain	Canonical Correlation Analysis (CCA)	ADNI
Ahmed *et al.* [115]	AD, MCI disease	2017	MRI/DTI	Brain	Multiple Kernel Learning	ADNI
Chavan *et al.* [116]	Neurocysticercosis (NCC) is a parasite infection	2017	MRI/CT	Brain	NSRCxWT	Radiopaedia
Piccinelli *et al.* [117]	Coronary Artery Disease	2018	PET/CTA	Heart	Feature extraction using MI	-
Rajalingam [118]	Astrocytoma Disease	2019	MRI/PET/ SPECT	Brain	Hybrid fusion (DFRWTx DTCWT)	AANLIB
Kaur *et al.* [119]	Degenerative and neoplastic brain tumor	2019	CT/MRI/ SPECT	Brain	SWT and PCA	AANLIB
Algarni *et al.* [120]	Tumor detection	2019	CT/MRI	Brain	CNN	-
Xiaoke *et al.* [121]	Alzheimer's Disease	2020	MRI/PET	Brain	Feature selection technique with CMC	ADNI
Jose J et al. [122]	Glioma, Encephalopathy, Mild Alzheimer's	2021	CT/MRI /PET	Brain	AISA-NSST	AANLIB

Chavan *et al.* [116] presented a fusion technique for the diagnosis of lesions caused by neurocysticercosis (NCC) infection in the brain. NCC is a parasitic disease caused by the tapeworm *Taenia solium*, which affects the central nervous structure of the human brain. In this technique, "Nonsubsampled Rotated Complex Wavelet Transform" (NSRCxWT) was used to combine CT and MRI medical images. Piccinelli *et al.* [117] presented a review article on multimodal fusion for the

analysis of coronary artery heart disease. In this article, they covered some fusion multimodalities of PET/SPECT and CTA. In an example case, fusion was implemented by utilizing left ventricle (LV) features extracted from both datasets and then optimized results with mutual information methods. Rajalingam *et al.* [118] developed an MMIF technique by combining the discrete fractional wavelet transform (DFRWT) with DTCWT to diagnose neural astrocytoma, which is a type of cancer in the human brain or human spinal cord system that develops from cells called astrocytes. Kaur *et al.* [119] suggested an MMIF fusion model based on SWT and PCA techniques to analyze degenerative and neoplastic brain tumor disease. The CLAHE technique was implemented for medical image enhancement information in pre-processing before the fusion process.

7. Image Fusion Quality Assessment Metrics

The quality of a fused image is determine by Fusion Quality Assessments Metrics (FQAMs) following subjective/qualitative and objective/quantitative methods. Particularly, subjective methods are based on visual examination to compare the final fused image with original input images. The examination of the fused image considers various parameters like image size, spatial details, color, etc. However, these strategies are inconvenient, costly, time-consuming, and troublesome in many fusion applications because of the absence of ground truth images that are completely fused [6]. The other method is objective quality assessment, which employs some evaluation metrics. The quantitative/objective method further categorized based on whether a reference image is available or not. The ground truth image is the reference image for validation of the fusion algorithm. The ground truth medical image is available in a very rare case or it can be construct manually.

Table 7. MMIF quality performance metrics

FQAMs	Formula	Description	Ref.
MI	$MI_F^{AB} = MI_{FA} + MI_{FB}$	Mutual information determines the similarity between two images. For better fusion, its value should be high. Where A and B are two input images and F is the fused image.	[124]
SD	$SD = \sqrt{\sum_{i=1}^{Q}\sum_{j=1}^{R}(f(i,j) - \bar{u})^2} \Big/ QR$	The intensity variation of the resultant fused medical image is computed. If its value is high, then fusion will be better, and vice versa. Where $\bar{u}$ is the mean value of the resultant image.	[125]
EN	$E = -\sum_{i=0}^{n-1} P(X_i) log P(X_i)$ p(xi) is the normalized histogram of the fusion image's corresponding gray level.	The quantity of information in an image is a measure of entropy, and its value ranges from 0 to 8. The x_i of ith point represents the image gray scale value, and P is probability. Image fusion results are better for a large entropy value	[126]
PSNR	$PSNR = 20\ x\ log10(\frac{Imax}{\sqrt{MSE}})$ $MSE = \frac{1}{QR}\sum_{n=1}^{Q}\sum_{n=1}^{R}\|(I(m,n) - J(m,n))\|$	PSNR computes the ratio of number of intensity levels in the medical image to the correlated pixels in the fused medical image. A higher value of PSNR shows superior. The I denote original image while max indicates the maximum pixel gray level of I. MSE represents mean squre error.I and J are original and fused images respectively. For comprehensive description se ref. [127]	[128]
UQI		UQI is based on the structural information of the final fused resultant images after	

	$UQI = \frac{\sigma_{xy}}{\sigma_x\sigma_y}\frac{2\overline{xy}}{(\overline{X})^2+(\overline{y})^2}\frac{2\sigma_x\sigma_y}{\sigma_x^2\sigma_y^2}$ Where σ denotes variance and μ represents average.	correlation loss, intensity deformation, and brightness deformation have been applied. UQI metrics are inspired by the visual human system. The range of its value metric is from -1 to 1. X and Y represent information transformation from two images.	[129]
EI	$Sx = f * hx, Sy = f * hy$ Where $\sqrt{(S_x{}^2 + S_y{}^2)}$	Higher image edge intensity indicates high image quality and greater image clarity. The edge intensity of image f can be computed utilizing the Sobel operator (S).	[130]
RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{XY}\sum_{i=o}^{X-1}\sum_{j=0}^{Y-1}(R(i,j) - F(i,j))^2}$	RMSE computes the quality of the final fused medical image by relating the ground truth or ideal fused medical image. For good fused image results, its value should be near zero. Where R is input and F denoted fused images, i and j are pixel horizontal and vertical respectively. The x and y show height and width of images.	[131]
SSIM	$\mathrm{SSIM}(A,B\|F) = \frac{(2\overline{F}_A\overline{F}_B + c_1)(2\sigma F_A F_B + c_2)}{(\overline{F}_A{}^2 + \overline{F}_B{}^2 + c_1)(\sigma^2 F_A + \sigma^2 F_B + c_2)}$ Where F_A fused image similarity with image A and F_B shows similarity between fused and image B.	The SSIM assessment metric determines the resemblance between sub-regions of images F_A and F_B with images A and B. where σ^2 denote variance of input images and σ is the average of input images A and B.	[132]
NCC	$\mathrm{NCC} = \frac{\sum_{i,j}[(A_{i,j} - \overline{A_{ij}}).(\hat{F}_{i,j} - \overline{\hat{F}_{ij}})]}{\sqrt{\sum_{i,j}\left[(A_{i,j} - \overline{A_{ij}})^2\right].\sum_{i,j}\left[(\hat{F}_{i,j} - \overline{\hat{F}_{ij}})^2\right]}}$	Normalize Cross Correlation is used to determine feature similarity in both reference images and the resultant fused image. Its value should be high, usually close to 1. $\overline{A_{ij}}$ indicates the average of ground truth images, and $\hat{F}_{i,j}$ is the average of fused images.	[133]
CE	$\mathrm{CE}(A,B,F) = \frac{D(h_A\|\|h_F) + D(h_B\|\|h_F)}{2}$	Cross Entropy measures the resemblance of information content between the source reference images and resultant fused images. The value of CE should be low for better fusion.	[134]
Q0, Qw, QE	$Q_0(x,y) = \frac{1}{\|w\|}\sum_{w\in W}(\lambda(w)\,Q_0(x,f\|w) + (1-\lambda(w))Q_0(y,f\|w))$ $Q_w(x,y,f) = \sum_{w\in W}C(w)(\lambda(w)\,Q_0(x,f\|w) + (1-\lambda(w))Q_0(y,f\|w))$ $Q_E(x,y,f) = Q_W(x,y,f).Q_W(\acute{x}.\acute{y}.\acute{f})^x$	Piella's metric determines prominent information in the resultant image by calculating measurements, i.e. mean luminance, image contrast, and correlation coefficient. It considers important information of image edges. The range of this metric is from 0 to 1, where a value closer to 1 indicates better fusion.	[135], [136]
QXY/F	$Q^{\frac{XY}{F}} = \frac{\sum_{n=1}^{N}\sum_{m=1}^{M}Q^{XF}(n,m)W_A(n,m) + Q^{YF}(n,m)W_B(n,m)}{\sum_{i=1}^{N}\sum_{j=1}^{M}(W_X(i,j) + W_Y(i,j))}$	This evaluation method reveals similarities between edges transmitted during the fusion procedure. The range of this metric is from 0 to 1. X and Y are input images, and F is the resultant image.	[137]

Objective quality evaluation parameters of the reference image include mutual information (MI), root mean square error (RMSE), correlation coefficient (CC), peak signal to noise ratio (PSNR), structural similarity index measure (SSIM), etc. [123]. In the case without ground truth images, the quality metric is calculated using the source medical images and the resultant fused image. Objective quality evaluation parameters without a reference image include entropy (H), standard deviation (SD), spatial frequency (SF), a sum of correlation difference (SCD), cross-entropy (CE), fusion mutual information (FMI), Petrovic metric, Piella metric (Q0, Qw, QE), etc. Some important quality fusion

assessments metrics with mathematical expressions are given in Table 7. To analyses the quantitative result we collected some fusion work from literature with each technique and compared their statistical fusion results using three most common metrics (MI, SD and SSIM). The quantitative quality assessment results are arranged in Table 8.

Table 8. Performance evaluation metric comparison

Methods	**Fusion level**	**Year**	**Harvard Medical School Dataset (AANLIB)**					
			MRI/CT with PET/SPECT			**CT with MRI**		
			MI	**SSIM**	**SD**	**MI**	**SSIM**	**SD**
He *et al.* [54]	Spatial domain	2010	2.9546	-	-	-	-	-
Zhizhong [84]		2020	-	-	-	-	0.8410	-
Singh [43]	Transform domain	2009	-	-	-	0.790	-	-
Bhatnagar[48]		2013	1.5017	0.7972	-	0.9681	0.7795	-
Xi [50]		2017	-	-	-	2.8001	-	66.26
Arif *et al.* [52]		2019	5.4121	0.9751	-	3.3121	**0.8651**	-
Srinivasu [85]		2020	4.4543	**0.9915**	-	-	-	-
Qiu *et al.* [86]		2020	-	-	-	2.2317	-	-
Wang [69]	Deep Learning	2007	-	-	-	**5.0066**	-	68.9896
Liu *et al.* [73]		2017	-	-	-	0.8872	0.6457	-
Hou *et al.* [74]		2018	-	-	-	4.4628	0.7456	74.3093
Liu [66]		2019					0.6397	
Xia [67]		2019	6.344	-	75.422	5.147	-	45.907
Jingming [90]		2020	1.4261	-	**109.515**	0.8599	-	**111.020**
Panigrahy [69]		2020		0.7597			0.7597	
Daneshvar [75]	Hybrid based methods	2010	0.6541	-	-	-	-	-
Das *et al.* [76]		2012	-	-	-	**5.0067**	-	68.9896
Sharmila [77]		2013	-	-	-	-	-	-
Kavitha [78]		2014	**7.2647**	0.6197	64.922	-	-	-
Ramlal. [79]		2019	-	-	-	4.3780	-	58.0671
Rajalingam[60]		2019	4.2991	0.8127	-	-	-	-
Lina [92]		2020	-	-	-	2.8210	0.777	-
Zong [77]	Sparse representa-ion (SR) methods	2017	4.1962	0.7335	68.7610	-	-	-
Jiang [94]		2018	-	-	-	4.673	0.625	-
Zhu *et al.*[95]		2018	2.0520	0.9178	-	-	-	-
Shahdoosti[96]		2018	2.7545	0.8635	-	2.9194	0.7839-	-
Maqsood [98]		2019	-	-	-	4.3580	0.7654	-
Das A et al [62]	Decision level	2009	-	-	-	3.4005	-	-
Huafeng Li et al [63]		2018	4.0708	-	-	4.3575	-	-
Shabanzade F et al. [58]		2019	3.0031	-	-	-	-	-
Bhardwaj J et. Al [59]		2020	-	-	-	1.4860	-	-
Nayak A et al. [60]		2020	-	-	-	1.4764	-	-

*Where MI and SSIM are reference-based metrics while SD is non-reference-based metrics.

8. Conclusions & Future Insights

The subject of medical image fusion briefly discussed in this recent survey. To begin with, the study describes the various medical imaging modalities used in the MMIF in detail. Before moving on to more advanced MMIF approaches, we described several MMIF databases and the main processes with fusion rules workflow. Research on MMIF algorithms is summarized in this publication. Comparative image fusion studies involving frequency fusion, spatial fusion, decision fusion, deep learning, hybrid fusion, and sparse representation fusion are presented in this research. Results and literature evaluation show that most MMIF reviews do not cover all classification strategies fully, as well as multimodal databases and current diseases that are related to MMIF are absent from many of these reviews. In this research, a comparison of several image fusion techniques is presented, along with their advantages and disadvantages. Tables 4, 5, and 6 provide a comparison of several MMIF algorithms based on literature reviews. The quality of the output-fused image is evaluated using a variety of objective criteria in Tables 8 and 9. DL fusion and Hybrid level approaches were shown to enhance the quantitative outcomes of fusion under this review observation. In comparison to other methods of fusing, the results of DL and Hybrid are more reasonable. There are many factors to consider when deciding which imaging approach is best for a certain application, therefore it is impossible to say which one is the best overall. In this review article, the following contribution summarized below:

- Classification of medical modalities used in MMIF according to the EM spectrum, source of energy, and acquisition of imaging.
- Analysis of five medical databases associated with multimodal images to extract statistical results.
- Discussion of the general procedure of MMIF, which further classified into six distinct methods: frequency fusion, spatial fusion, decision fusion, deep learning, hybrid fusion, and sparse representation fusion.
- Comparison of these techniques based on image quality assessment metrics.

In recent years, medical image fusion has been progressively used to diagnose different diseases efficiently, Previously, significant information was obtained from single modalities, but this may easily be increased by using MMIF from different sensors with imaging equipment. Works that utilized such fusion techniques for disease diagnoses are also discussed in this article. Significant fusion quality assessments metrics are highlighted for benchmarking the MMIF methods. In conclusion, researchers have a choice of optimal fusion techniques and datasets for benchmarking their results and further developing suitable models for diagnosing diseases and analyzing different patterns in medical imaging. However, some existing challenges of MMIF pertain to miss-registration and artifacts in the resultant fused image. Such issues can be solved by increasing knowledge of the multimodal fusion domain and achieving improved diagnoses of more recent diseases that cannot be detected by a single modality.

8.1 Advantages and Drawback of previous techniques

Various multimodal medical fusion techniques have covered in this work. It was found that spatial image fusion methods are ineffective in real-world applications. For instance, the PCA, hue intensity

saturation, and Brovey techniques are computationally efficient, fast, and simple, yet they result in illumination variation. The use of principal component analysis to fuse images offers a spatial benefit but suffers from spectral deterioration. While frequency domain methods reduce spectral distortion and can perform better than standard fusion methods, they have a higher signal-to-noise ratio than pixel-level-based methods. Enhanced results can be achieved by multilevel fusion, wherein the image is fused twice using an appropriate fusion technique. Fused images may have a less spatial resolution because of this process. This fusion algorithm is more complex than pixel-level techniques. In literature, the hybrid fusion technique can improve the clarity, contrast, texture, brightness, and edge information of the fused image. This technique can also achieve complementary information from input images with reduced artifacts. However, a thorough understanding of each approach is still need to improve the hybrid fusion algorithm, which yields non-uniform fusion outcomes, and these techniques are typically difficult and time-consuming. The other challenge in the hybrid fusion technique is that large input datasets should not be use. In the sparse representation approach, the coefficients are the most important parameters that increase the fusion performance. By enhancing the image's contrast and retention of visual information, information about image structure and a high level of detail in source images can preserved. Nevertheless, this technique suffers from miss-registration, including minimal detail preservation capacity. Deep learning techniques can make optimizing the image fusion process easier and maybe customized to an application's requirements. When there is a large amount of input data with several dimensions and variety, deep learning approaches outperform other fusion techniques. However, these approaches based on a dynamic process with many parameters, and training a fusion model requires more effort and hardware than other techniques. Another flaw of deep learning methods is that reliable results cannot be produce for smaller image datasets.

8.2 Future direction

Even though researchers have suggested several image fusion algorithms, these approaches even have certain flaws. Specific advanced MMIF techniques and implementation of new image fusion metrics are still in need of the current research. Considering the mentioned shortcomings, some possible future research aspects are as follows:

- Investigate the effectiveness of image of fusion techniques, particularly with advancements of machine learning and deep learning, considering the absence of appropriate image representation methodologies and generally recognized fusion assessment criteria.
- The implementation of evolving methods to enhance image areas of interest before the fusion process is another area of concern.
- Develop innovative multi-scale decomposition algorithms for multimodality image fusion.
- Since most image fusion issues lack accurate ground truth, evaluating the quality of the fused output is highly challenging, we need to develop no-reference metrics with greater descriptive capabilities.
- Improve the image quality of resulting fusion output to manage the limitations of imaging modalities and excessive noise

- Identify solutions to the absence of adequate information in each medical imaging modality, image noise, massive computing cost, and the disparity in dimensions between images in multiple modalities.
- A challenge to collect or build datasets that may deemed trustworthy in various clinical aided-diagnosis such as eye, larynx, and other bodily organs that have not been addressed in the literature to assess fusion methodologies is one of future research in this subject. Only brain imaging data from Harvard dominates the rest of the field, according to this current review.

Supporting Material

(Source Link.1) https://github.com/RRuschel/Image-fusion (MATLAB)

(Source Link.2) https://github.com/zhiqinzhu123/Source-code-of-medical-image-fusion-in-NSCT-domain (MATLAB)

(Source Link.3) https://github.com/uzeful/IFCNN (Python)

(Source Link.4) https://github.com/yuliu316316/NSST-PAPCNN-Fusion (MATLAB)

(Source Link.5) https://github.com/yuliu316316/MST-SR-Fusion-Toolbox (MATLAB)

(Source Link.6) https://github.com/HarrisXia/image-fusion-zoo (MATLAB)

References

[1] B. Rajalingam, R. Priya, and R. Scholar, "Review of Multimodality Medical Image Fusion Using Combined Transform Techniques for Clinical Application," *Int. J. Sci. Res. Comput. Sci. Appl. Manag. Stud. IJSRCSAMS*, vol. 7, no. 3, 2018, [Online]. Available: www.ijsrcsams.com.

[2] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 19, no. 1, pp. 4–19, 2014, doi: 10.1016/j.inffus.2013.12.002.

[3] "PubMed." https://www.ncbi.nlm.nih.gov/pubmed/ (accessed Feb. 10, 2020).

[4] J. Du, W. Li, K. Lu, and B. Xiao, "An overview of multi-modal medical image fusion," *Neurocomputing*, vol. 215, pp. 3–20, 2016, doi: 10.1016/j.neucom.2015.07.160.

[5] F. E. Z. A. El-Gamal, M. Elmogy, and A. Atwan, "Current trends in medical image registration and fusion," *Egypt. Informatics J.*, vol. 17, no. 1, pp. 99–124, 2016, doi: 10.1016/j.eij.2015.09.002.

[6] B. Meher, S. Agrawal, R. Panda, and A. Abraham, "A survey on region based image fusion methods," *Inf. Fusion*, vol. 48, pp. 119–132, 2019, doi: 10.1016/j.inffus.2018.07.010.

[7] J. M. Dolly and A. K. Nisa, "A Survey on Different Multimodal Medical Image Fusion Techniques and Methods," *Proc. 1st Int. Conf. Innov. Inf. Commun. Technol. ICIICT 2019*, 2019, doi: 10.1109/ICIICT1.2019.8741445.

[8] B. Huang, F. Yang, M. Yin, X. Mo, and C. Zhong, "A Review of Multimodal Medical Image Fusion Techniques," *Comput. Math. Methods Med.*, vol. 2020, 2020, doi: 10.1155/2020/8279342.

[9] T. Tirupal, B. C. Mohan, and S. S. Kumar, "Multimodal Medical Image Fusion Techniques – A Review," *Curr. Signal Transduct. Ther.*, vol. 16, no. 2, pp. 142–163, 2020, doi:

10.2174/1574362415666200226103116.

[10] H. Hermessi, O. Mourali, and E. Zagrouba, “Multimodal medical image fusion review: Theoretical background and recent advances,” *Signal Processing*, vol. 183, 2021, doi: 10.1016/j.sigpro.2021.108036.

[11] J. Sebastian and G. R. G. King, “Fusion of Multimodality Medical Images- A Review,” *Proc. - 1st Int. Conf. Smart Technol. Commun. Robot. STCR 2021*, no. October, 2021, doi: 10.1109/STCR51658.2021.9588882.

[12] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, “Image fusion meets deep learning: A survey and perspective,” *Inf. Fusion*, vol. 76, no. May, pp. 323–336, 2021, doi: 10.1016/j.inffus.2021.06.008.

[13] B. Singh *et al.*, “Application of vibrational microspectroscopy to biology and medicine,” *Curr. Sci.*, vol. 102, no. 2, pp. 232–244, 2012.

[14] J. Andreu-Perez, C. C. Y. Poon, R. D. Merrifield, S. T. C. Wong, and G. Z. Yang, “Big Data for Health,” *IEEE J. Biomed. Heal. Informatics*, vol. 19, no. 4, pp. 1193–1208, 2015, doi: 10.1109/JBHI.2015.2450362.

[15] M. A. Azam *et al.*, “Deep Learning Applied to White Light and Narrow Band Imaging Videolaryngoscopy : Toward Real-Time Laryngeal Cancer Detection,” *Laryngoscope*, pp. 1–9, 2021, doi: 10.1002/lary.29960.

[16] M. A. Haidekker, *X-Ray Projection Imaging*. 2013.

[17] “MITA.” https://www.medicalimaging.org/about-mita/medical-imaging-primer/ (accessed Jan. 10, 2020).

[18] V. I. Mikla, *Medical Imaging Technology*. 2013.

[19] F. E. Z. A. El-Gamal, M. Elmogy, and A. Atwan, “Current trends in medical image registration and fusion,” *Egypt. Informatics J.*, vol. 17, no. 1, pp. 99–124, 2016, doi: 10.1016/j.eij.2015.09.002.

[20] 2013 John R.Giudicessi, BA.Michael J.Ackerman., “Imaging Techniques for Kaposi Sarcoma (KS),” *Bone*, vol. 23, no. 1, pp. 1–7, 2008, doi: 10.1038/jid.2014.371.

[21] K. Y. C. Teo *et al.*, “Introduction to Optical Coherence Tomography Angiography,” *Swept-Source Opt. Coherence Tomogr.*, no. 1, pp. 9–14, 2018, doi: 10.1142/9789813239579_0002.

[22] “Microscopy.” https://medical-dictionary.thefreedictionary.com/o (accessed Jan. 10, 2020).

[23] J. Cal-Gonzalez *et al.*, “Hybrid imaging: Instrumentation and data processing,” *Front. Phys.*, vol. 6, no. MAY, 2018, doi: 10.3389/fphy.2018.00047.

[24] M. A. Azam, K. B. Khan, M. Aqeel, A. R. Chishti, and M. N. Abbasi, “Analysis of the MIDAS and OASIS Biomedical Databases for the Application of Multimodal Image Processing,” *Commun. Comput. Inf. Sci.*, vol. 1198, pp. 581–592, 2020, doi: 10.1007/978-981-15-5232-8_50.

[25] “TCIA.” https://www.cancerimagingarchive.net/ (accessed Jan. 05, 2020).

[26] S. D. Ramlal, J. Sachdeva, C. K. Ahuja, and N. Khandelwal, “Multimodal medical image fusion using non-subsampled shearlet transform and pulse coupled neural network incorporated with morphological gradient,” *Signal, Image Video Process.*, vol. 12, no. 8, pp. 1479–1487, 2018, doi: 10.1007/s11760-018-1303-z.

[27] E. Daniel, “Optimum Wavelet-Based Homomorphic Medical Image Fusion Using Hybrid Genetic-Grey Wolf Optimization Algorithm,” *IEEE Sens. J.*, vol. 18, no. 16, pp. 6804–6811, 2018, doi:

10.1109/JSEN.2018.2822712.

[28] Y. Yang, Y. Que, S. Huang, and P. Lin, “Multimodal Sensor Medical Image Fusion Based on Type-2 Fuzzy Logic in NSCT Domain,” *IEEE Sens. J.*, vol. 16, no. 10, pp. 3735–3745, 2016, doi: 10.1109/JSEN.2016.2533864.

[29] A. Torrado-Carvajal *et al.*, “Multi-atlas and label fusion approach for patient-specific MRI based skull estimation,” *Magn. Reson. Med.*, vol. 75, no. 4, pp. 1797–1807, 2016, doi: 10.1002/mrm.25737.

[30] “AANLIB.” http://www.med.harvard.edu/AANLIB/home.html (accessed Jan. 05, 2020).

[31] “ADNI.” http://adni.loni.usc.edu/ (accessed Jan. 05, 2020).

[32] M. A. Azam, K. B. Khan, M. Ahmad, and M. Mazzara, “Multimodal Medical Image Registration and Fusion for Quality Enhancement,” *Comput. Mater. Contin.*, vol. 68, no. 1, pp. 821–840, 2021, doi: 10.32604/cmc.2021.016131.

[33] “MIDAS.” http://www.insight-journal.org/midas/ (accessed Jan. 10, 2020).

[34] M. R. Princess, V. S. Kumar, and M. R. Begum, “Comprehensive and Comparative Study of Different Image Fusion Techniques,” *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, vol. 03, no. 09, pp. 11800–11806, 2014, doi: 10.15662/ijareeie.2014.0309015.

[35] H. A. Mohammed and M. A. Hassan, “The Image Registration Techniques for Medical Imaging (MRI-CT),” *Am. J. Biomed. Eng.*, vol. 6, no. 2, pp. 53–58, 2016, doi: 10.5923/j.ajbe.20160602.02.

[36] H. M. Elhoseny, E. M. El-rabaie, O. S. F. Allah, and F. E. A. El-samie, “Medical Image Fusion: A Literature Review Present Solutions and Future Directions,” *Menoufia J. Electron. Eng. Res.*, vol. 26, no. September, pp. 321–350, 2017, doi: 10.21608/mjeer.2017.63510.

[37] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, “Pixel-level image fusion: A survey of the state of the art,” *Inf. Fusion*, vol. 33, pp. 100–112, 2017, doi: 10.1016/j.inffus.2016.05.004.

[38] K. Parmar and R. Kher, “A comparative analysis of multimodality medical image fusion methods,” *Proc. - 6th Asia Int. Conf. Math. Model. Comput. Simulation, AMS 2012*, pp. 93–97, 2012, doi: 10.1109/AMS.2012.46.

[39] F. Sadjadi, “Comparative Image Fusion Analysais,” pp. 8–8, 2006, doi: 10.1109/cvpr.2005.436.

[40] S. Das and M. K. Kundu, “A neuro-fuzzy approach for medical image fusion,” *IEEE Trans. Biomed. Eng.*, vol. 60, no. 12, pp. 3347–3353, 2013, doi: 10.1109/TBME.2013.2282461.

[41] S. Zarif, I. Faye, and R. Dayang, “A comparative study of different image completion techniques,” *2014 Int. Conf. Comput. Inf. Sci. ICCOINS 2014 - A Conf. World Eng. Sci. Technol. Congr. ESTCON 2014 - Proc.*, vol. 90, no. 19, pp. 12–16, 2014, doi: 10.1109/ICCOINS.2014.6868411.

[42] Q. Guihong, Z. Dali, and Y. Pingfan, “Medical image fusion by wavelet transform modulus maxima,” *Opt. Express*, vol. 9, no. 4, p. 184, 2001, doi: 10.1364/oe.9.000184.

[43] R. Singh, M. Vatsa, and A. Noore, “Multimodal medical image fusion using Redundant Discrete Wavelet Transform,” *Proc. 7th Int. Conf. Adv. Pattern Recognition, ICAPR 2009*, pp. 232–235, 2009, doi: 10.1109/ICAPR.2009.97.

[44] E. Moustafa, “for Detection of Hepatic Lesions and Acute Intra-Cerebral,” *Conf. Inf. Commun. Technol. 2006*, vol. 00.

[45] Y. Liu, “PET / CT Medical Image Fusion Algorithm Based on Multiwavelet Transform.”

[46] L. Yang, B. L. Guo, and W. Ni, “Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform,” *Neurocomputing*, vol. 72, no. 1–3, pp. 203–211, 2008, doi: 10.1016/j.neucom.2008.02.025.

[47] X. J. Wang and Y. Mu, “A medical image fusion algorithm based on lifting wavelet transform,” *Proc. - Int. Conf. Artif. Intell. Comput. Intell. AICI 2010*, vol. 3, no. 8, pp. 474–476, 2010, doi: 10.1109/AICI.2010.337.

[48] S. V and B. R. Kumar, “Directive Contrast Based Multimodal Medical Image Fusion in NSCT with DWT Domain,” *Int. J. Eng. Trends Technol.*, vol. 9, no. 6, pp. 288–294, 2014, doi: 10.14445/22315381/ijett-v9p257.

[49] A. Sahu, V. Bhateja, A. Krishn, and Himanshi, “Medical image fusion with Laplacian Pyramids,” *2014 Int. Conf. Med. Imaging, m-Health Emerg. Commun. Syst. MedCom 2014*, pp. 448–453, 2014, doi: 10.1109/MedCom.2014.7006050.

[50] X. X. Xi, X. Q. Luo, Z. C. Zhang, Q. J. You, and X. Wu, “Multimodal medical volumetric image fusion based on multi-feature in 3-D shearlet transform,” *2017 Int. Smart Cities Conf. ISC2 2017*, no. 1, 2017, doi: 10.1109/ISC2.2017.8090797.

[51] M. Yin, X. Liu, Y. Liu, and X. Chen, “Medical Image Fusion with Parameter-Adaptive Pulse Coupled Neural Network in Nonsubsampled Shearlet Transform Domain,” *IEEE Trans. Instrum. Meas.*, vol. 68, no. 1, pp. 49–64, 2019, doi: 10.1109/TIM.2018.2838778.

[52] M. Arif and G. Wang, “Fast curvelet transform through genetic algorithm for multimodal medical image fusion,” *Soft Comput.*, 2019, doi: 10.1007/s00500-019-04011-5.

[53] R. Stokking, K. J. Zuiderveld, and M. A. Viergever, “Integrated volume visualization of functional image data and anatomical surfaces using normal fusion,” *Hum. Brain Mapp.*, vol. 12, no. 4, pp. 203–218, 2001, doi: 10.1002/1097-0193(200104)12:4<203::AID-HBM1016>3.0.CO;2-X.

[54] C. He, Q. Liu, H. Li, and H. Wang, “Procedia Engineering Multimodal medical image fusion based on IHS and PCA,” *Procedia Eng.*, vol. 7, pp. 280–285, 2010, doi: 10.1016/j.proeng.2010.11.045.

[55] R. Bashir, R. Junejo, N. N. Qadri, M. Fleury, and M. Y. Qadri, “SWT and PCA image fusion methods for multi-modal imagery,” *Multimed. Tools Appl.*, vol. 78, no. 2, pp. 1235–1263, 2019, doi: 10.1007/s11042-018-6229-5.

[56] B. Meher, S. Agrawal, R. Panda, and A. Abraham, “A survey on region based image fusion methods,” *Inf. Fusion*, vol. 48, no. July 2018, pp. 119–132, 2019, doi: 10.1016/j.inffus.2018.07.010.

[57] F. Tabib Mahmoudi, F. Samadzadegan, and P. Reinartz, “Object recognition based on the context aware decision-level fusion in multiviews imagery,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 8, no. 1, pp. 12–22, 2015, doi: 10.1109/JSTARS.2014.2362103.

[58] F. Shabanzade, M. Khateri, and Z. Liu, “MR and PET Image Fusion Using Nonparametric Bayesian Joint Dictionary Learning,” *IEEE Sensors Lett.*, vol. 3, no. 7, pp. 2019–2022, 2019, doi: 10.1109/LSENS.2019.2925072.

[59] J. Bhardwaj and A. Nayak, “Haar wavelet transform–based optimal Bayesian method for medical image fusion,” *Med. Biol. Eng. Comput.*, vol. 58, no. 10, pp. 2397–2411, 2020, doi: 10.1007/s11517-020-02209-6.

[60] J. Bhardwaj and A. Nayak, “Discrete Wavelet Transform and Bird Swarm Optimized Bayesian

Multimodal Medical Image Fusion," *Helix*, vol. 10, no. 1, pp. 07–12, 2020, doi: 10.29042/2020-10-1-07-12.

[61] M. Yousuf, K. B. Khan, M. A. Azam, and M. Aqeel, "Brain Tumor Localization and Segmentation Based on Pixel-Based Thresholding with Morphological Operation," *Commun. Comput. Inf. Sci.*, vol. 1198, no. May, pp. 562–572, 2020, doi: 10.1007/978-981-15-5232-8_48.

[62] A. Das and M. Bhattacharya, "Evolutionary algorithm based automated medical image fusion technique: Comparative study with fuzzy fusion approach," *2009 World Congr. Nat. Biol. Inspired Comput. NABIC 2009 - Proc.*, pp. 269–274, 2009, doi: 10.1109/NABIC.2009.5393715.

[63] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, "Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning," *Pattern Recognit.*, vol. 79, pp. 130–146, 2018, doi: 10.1016/j.patcog.2018.02.005.

[64] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[65] T. Zhou, S. Ruan, and S. Canu, "A review: Deep learning for medical image segmentation using multi-modality fusion," *Array*, vol. 3–4, no. July, p. 100004, 2019, doi: 10.1016/j.array.2019.100004.

[66] Z. Liu *et al.*, "Automatic diagnosis of fungal keratitis using data augmentation and image fusion with deep convolutional neural network," *Comput. Methods Programs Biomed.*, vol. 187, p. 105019, 2020, doi: 10.1016/j.cmpb.2019.105019.

[67] K. S. Choi, J. S. Shin, J. J. Lee, Y. S. Kim, S. B. Kim, and C. W. Kim, "In vitro trans-differentiation of rat mesenchymal cells into insulin-producing cells by rat pancreatic extract," *Biochem. Biophys. Res. Commun.*, vol. 330, no. 4, pp. 1299–1305, 2005, doi: 10.1016/j.bbrc.2005.03.111.

[68] M. D. Zeiler, D. Krishnan, G. W. Taylor, and R. Fergus, "Deconvolutional Networks for Feature Learning," *Cvpr*, pp. 2528–2535, 2010, doi: 10.1109/CVPR.2010.5539957.

[69] Y. P. Wang, J. W. Dang, Q. Li, and S. Li, "Multimodal medical image fusion using fuzzy radial basis function neural networks," *Proc. 2007 Int. Conf. Wavelet Anal. Pattern Recognition, ICWAPR '07*, vol. 2, pp. 778–782, 2007, doi: 10.1109/ICWAPR.2007.4420774.

[70] Z. Wang and Y. Ma, "Medical image fusion using m-PCNN," *Inf. Fusion*, vol. 9, no. 2, pp. 176–185, 2008, doi: 10.1016/j.inffus.2007.04.003.

[71] J. Teng, S. Wang, J. Zhang, and X. Wang, "Neuro-fuzzy logic based fusion algorithm of medical images," *Proc. - 2010 3rd Int. Congr. Image Signal Process. CISP 2010*, vol. 4, pp. 1552–1556, 2010, doi: 10.1109/CISP.2010.5646958.

[72] S. Sivasangumani, P. S. Gomathi, and B. Kalaavathi, "Regional firing characteristic of PCNN-based multimodal medical image fusion in NSCT domain," *Int. J. Biomed. Eng. Technol.*, vol. 18, no. 3, pp. 199–209, 2015, doi: 10.1504/IJBET.2015.070575.

[73] Y. Liu, X. Chen, J. Cheng, and H. Peng, "A medical image fusion method based on convolutional neural networks," *20th Int. Conf. Inf. Fusion, Fusion 2017 - Proc.*, 2017, doi: 10.23919/ICIF.2017.8009769.

[74] R. Hou, D. Zhou, R. Nie, D. Liu, and X. Ruan, "Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model," *Med. Biol. Eng. Comput.*, 2018, doi: 10.1007/s11517-018-1935-8.

[75] S. Daneshvar and H. Ghassemian, "MRI and PET image fusion by combining IHS and retina-inspired models," *Inf. Fusion*, vol. 11, no. 2, pp. 114–123, 2010, doi: 10.1016/j.inffus.2009.05.003.

[76] S. Das and M. K. Kundu, “NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency,” *Med. Biol. Eng. Comput.*, vol. 50, no. 10, pp. 1105–1114, 2012, doi: 10.1007/s11517-012-0943-3.

[77] K. Sharmila, S. Rajkumar, and V. Vijayarajan, “Hybrid method for multimodality medical image fusion using Discrete Wavelet Transform and Entropy concepts with quantitative analysis,” *Int. Conf. Commun. Signal Process. ICCSP 2013 - Proc.*, pp. 489–493, 2013, doi: 10.1109/iccsp.2013.6577102.

[78] C. T. Kavitha and C. Chellamuthu, “Medical image fusion based on hybrid intelligence,” *Appl. Soft Comput. J.*, vol. 20, pp. 83–94, 2014, doi: 10.1016/j.asoc.2013.10.034.

[79] S. D. Ramlal, J. Sachdeva, C. K. Ahuja, and N. Khandelwal, “An improved multimodal medical image fusion scheme based on hybrid combination of nonsubsampled contourlet transform and stationary wavelet transform,” *Int. J. Imaging Syst. Technol.*, vol. 29, no. 2, pp. 146–160, 2019, doi: 10.1002/ima.22310.

[80] Q. Zhang and X. Maldague, “An adaptive fusion approach for infrared and visible images based on NSCT and compressed sensing,” *Infrared Phys. Technol.*, vol. 74, pp. 11–20, 2016, doi: 10.1016/j.infrared.2015.11.003.

[81] Y. Liu, S. Liu, and Z. Wang, “A general framework for image fusion based on multi-scale transform and sparse representation,” *Inf. Fusion*, vol. 24, pp. 147–164, 2015, doi: 10.1016/j.inffus.2014.09.004.

[82] Q. Zhang, Y. Liu, R. S. Blum, J. Han, and D. Tao, “Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: A review,” *Inf. Fusion*, vol. 40, pp. 57–75, 2018, doi: 10.1016/j.inffus.2017.05.006.

[83] S. Daneshvar and H. Ghassemian, “Fusion of MRI and PET images using retina based multi-resolution transforms,” *2007 9th Int. Symp. Signal Process. its Appl. ISSPA 2007, Proc.*, pp. 0–3, 2007, doi: 10.1109/ISSPA.2007.4555524.

[84] Z. Fu, Y. Zhao, Y. Xu, L. Xu, and J. Xu, “Gradient structural similarity based gradient filtering for multi-modal image fusion,” *Inf. Fusion*, vol. 53, no. April 2018, pp. 251–268, 2020, doi: 10.1016/j.inffus.2019.06.025.

[85] S. Polinati and R. Dhuli, “Multimodal medical image fusion using empirical wavelet decomposition and local energy maxima,” *Optik (Stuttg).*, vol. 205, p. 163947, 2020, doi: 10.1016/j.ijleo.2019.163947.

[86] Q. Hu, S. Hu, and F. Zhang, “Multi-modality medical image fusion based on separable dictionary learning and Gabor filtering,” *Signal Process. Image Commun.*, vol. 83, no. December 2019, p. 115758, 2020, doi: 10.1016/j.image.2019.115758.

[87] Y. Liu, X. Chen, R. K. Ward, and J. Wang, “Image Fusion with Convolutional Sparse Representation,” *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, 2016, doi: 10.1109/LSP.2016.2618776.

[88] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, “Medical image fusion via convolutional sparsity based morphological component analysis,” *IEEE Signal Process. Lett.*, vol. 26, no. 3, pp. 485–489, 2019, doi: 10.1109/LSP.2019.2895749.

[89] K. jian Xia, H. sheng Yin, and J. qiang Wang, “A novel improved deep convolutional neural network model for medical image fusion,” *Cluster Comput.*, vol. 22, pp. 1515–1527, 2019, doi: 10.1007/s10586-018-2026-1.

[90] J. Xia, Y. Lu, and L. Tan, “Research of Multimodal Medical Image Fusion Based on Parameter-Adaptive Pulse-Coupled Neural Network and Convolutional Sparse Representation,” *Comput. Math. Methods*

Med., vol. 2020, 2020, doi: 10.1155/2020/3290136.

[91] C. Panigrahy, A. Seal, and N. K. Mahato, "MRI and SPECT Image Fusion Using a Weighted Parameter Adaptive Dual Channel PCNN," *IEEE Signal Process. Lett.*, vol. 27, no. 1070, pp. 690–694, 2020, doi: 10.1109/LSP.2020.2989054.

[92] L. Xu, Y. Si, S. Jiang, Y. Sun, and H. Ebrahimian, "Medical image fusion using a modified shark smell optimization algorithm and hybrid wavelet-homomorphic filter," *Biomed. Signal Process. Control*, vol. 59, p. 101885, 2020, doi: 10.1016/j.bspc.2020.101885.

[93] J. jing Zong and T. shuang Qiu, "Medical image fusion based on sparse representation of classified image patches," *Biomed. Signal Process. Control*, vol. 34, pp. 195–205, 2017, doi: 10.1016/j.bspc.2017.02.005.

[94] W. Jiang *et al.*, "Medical images fusion by using weighted least squares filter and sparse representation," *Comput. Electr. Eng.*, vol. 67, no. March, pp. 252–266, 2018, doi: 10.1016/j.compeleceng.2018.03.037.

[95] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Inf. Sci. (Ny).*, vol. 432, pp. 516–529, 2018, doi: 10.1016/j.ins.2017.09.010.

[96] H. R. Shahdoosti and A. Mehrabi, "Multimodal image fusion using sparse representation classification in tetrolet domain," *Digit. Signal Process. A Rev. J.*, vol. 79, pp. 9–22, 2018, doi: 10.1016/j.dsp.2018.04.002.

[97] Y. Zhang, M. Yang, N. Li, and Z. Yu, "Analysis-synthesis dictionary pair learning and patch saliency measure for image fusion," *Signal Processing*, vol. 167, p. 107327, 2020, doi: 10.1016/j.sigpro.2019.107327.

[98] 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, p. 101810, 2020, doi: 10.1016/j.bspc.2019.101810.

[99] Y. Liu, X. Yang, R. Zhang, M. K. Albertini, T. Celik, and G. Jeon, "Entropy-based image fusion with joint sparse representation and rolling guidance filter," *Entropy*, vol. 22, no. 1, p. 118, 2020, doi: 10.3390/e22010118.

[100] N. Anantrasirichai, R. Zheng, I. Selesnick, and A. Achim, "Image fusion via sparse regularization with non-convex penalties," *Pattern Recognit. Lett.*, vol. 131, pp. 355–360, 2020, doi: 10.1016/j.patrec.2020.01.020.

[101] H. Li, X. He, Z. Yu, and J. Luo, "Noise-robust image fusion with low-rank sparse decomposition guided by external patch prior," *Inf. Sci. (Ny).*, vol. 523, pp. 14–37, 2020, doi: 10.1016/j.ins.2020.03.009.

[102] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, 2013, doi: 10.1109/TIP.2013.2244222.

[103] Z. Zhu, M. Zheng, G. Qi, D. Wang, and Y. Xiang, "A Phase Congruency and Local Laplacian Energy Based Multi-Modality Medical Image Fusion Method in NSCT Domain," *IEEE Access*, vol. 7, pp. 20811–20824, 2019, doi: 10.1109/ACCESS.2019.2898111.

[104] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "IFCNN: A general image fusion framework based on convolutional neural network," *Inf. Fusion*, vol. 54, no. August 2018, pp. 99–118, 2020, doi: 10.1016/j.inffus.2019.07.011.

[105] S. Das and M. K. Kundu, "Corrections to 'A Neuro-Fuzzy Approach for Medical Image Fusion,'" *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, p. 1226, 2015, doi: 10.1109/TBME.2015.2405137.

[106] D. Mishra and B. Palkar, “Image Fusion Techniques: A Review,” *Int. J. Comput. Appl.*, vol. 130, no. 9, pp. 7–13, 2015, doi: 10.5120/ijca2015907084.

[107] S. Bhat and D. Koundal, *Multi-focus image fusion techniques: a survey*, vol. 54, no. 8. Springer Netherlands, 2021.

[108] S. Masood, M. Sharif, M. Yasmin, M. A. Shahid, and A. Rehman, “Image fusion methods: A survey,” *J. Eng. Sci. Technol. Rev.*, vol. 10, no. 6, pp. 186–194, 2017, doi: 10.25103/jestr.106.24.

[109] A. M. Sharma, A. Dogra, B. Goyal, R. Vig, and S. Agrawal, “From pyramids to state-of-the-art: A study and comprehensive comparison of visible-infrared image fusion techniques,” *IET Image Process.*, vol. 14, no. 9, pp. 1671–1689, 2020, doi: 10.1049/iet-ipr.2019.0322.

[110] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, *Multimodal fusion for multimedia analysis: A survey*, vol. 16, no. 6. 2010.

[111] Y. Liu, X. Chen, R. K. Ward, Z. J. Wang, and S. Member, “Sparse Representation,” *Comput. Vis.*, vol. 23, no. 12, pp. 1185–1185, 2021, doi: 10.1007/978-3-030-63416-2_300100.

[112] N. Tawfik, H. A. Elnemr, M. Fakhr, M. I. Dessouky, and F. E. Abd El-Samie, “Survey study of multimodality medical image fusion methods,” *Multimed. Tools Appl.*, vol. 80, no. 4, pp. 6369–6396, 2021, doi: 10.1007/s11042-020-08834-5.

[113] B. Rajalingam, R. Priya, and R. Bhavani, “Hybrid multimodal medical image fusion using combination of transform techniques for disease analysis,” *Procedia Comput. Sci.*, vol. 152, pp. 150–157, 2019, doi: 10.1016/j.procs.2019.05.037.

[114] B. Lei, S. Chen, D. Ni, and T. Wang, “Discriminative learning for Alzheimer’s disease diagnosis via canonical correlation analysis and multimodal fusion,” *Front. Aging Neurosci.*, vol. 8, no. MAY, pp. 1–17, 2016, doi: 10.3389/fnagi.2016.00077.

[115] O. Ben Ahmed, J. Benois-Pineau, M. Allard, G. Catheline, and C. Ben Amar, “Recognition of Alzheimer’s disease and Mild Cognitive Impairment with multimodal image-derived biomarkers and Multiple Kernel Learning,” *Neurocomputing*, vol. 220, pp. 98–110, 2017, doi: 10.1016/j.neucom.2016.08.041.

[116] S. S. Chavan, A. Mahajan, S. N. Talbar, S. Desai, M. Thakur, and A. D’cruz, “Nonsubsampled rotated complex wavelet transform (NSRCxWT) for medical image fusion related to clinical aspects in neurocysticercosis,” *Comput. Biol. Med.*, vol. 81, no. December 2016, pp. 64–78, 2017, doi: 10.1016/j.compbiomed.2016.12.006.

[117] M. Piccinelli, D. C. Cooke, and E. V. Garcia, “Multimodality Image Fusion for Coronary Artery Disease Detection,” *Ann. Nucl. Cardiol.*, vol. 4, no. 1, pp. 74–78, 2018, doi: 10.17996/anc.18-00065.

[118] E. S. K. B and C. S. Bindu, *Emerging Technologies in Computer Engineering: Microservices in Big Data Analytics*, vol. 985. Springer Singapore, 2019.

[119] L. Singh, S. Singh, and N. Aggarwal, *Proceedings of 2nd International Conference on Communication, Computing and Networking*, vol. 46. Springer Singapore, 2019.

[120] A. D. Algarni, “Automated Medical Diagnosis System Based on Multi-modality Image Fusion and Deep Learning,” *Wirel. Pers. Commun.*, vol. 111, no. 2, pp. 1033–1058, 2020, doi: 10.1007/s11277-019-06899-6.

[121] X. Hao *et al.*, “Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer’s disease,” *Med. Image Anal.*, vol. 60, p. 101625, 2020, doi: 10.1016/j.media.2019.101625.

[122] J. Jose *et al.*, “An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion,” *Biomed. Signal Process. Control*, vol. 66, no. February, 2021, doi: 10.1016/j.bspc.2021.102480.

[123] J. Frp, J. Frp, J. D. Frp, E. Sdshu, G. Wkh, and L. Ixvlrq, “comparison of medical image fusion methods using image quaity metrics,” *2018 Int. Conf. Commun. Comput. Internet Things*, vol. 3, pp. 449–454, 2018.

[124] R. Singh and A. Khare, “Multimodal medical image fusion using daubechies complex wavelet transform,” *2013 IEEE Conf. Inf. Commun. Technol. ICT 2013*, no. Ict, pp. 869–873, 2013, doi: 10.1109/CICT.2013.6558217.

[125] W. Wang and F. Chang, “A multi-focus image fusion method based on Laplacian pyramid,” *J. Comput.*, vol. 6, no. 12, pp. 2559–2566, 2011, doi: 10.4304/jcp.6.12.2559-2566.

[126] J. Sen Teh, A. Akhavan, and A. Samsudin, “High quality random source based on multi-core CPUs and chaotic maps,” vol. 00, no. 00, pp. 1–2, 2006, doi: 10.1049/el.

[127] S. Rajkumar and G. Malathi, “A comparative analysis on image quality assessment for real time satellite images,” *Indian J. Sci. Technol.*, vol. 9, no. 34, 2016, doi: 10.17485/ijst/2016/v9i34/96766.

[128] V. P. S. Naidu, “Discrete cosine transform-based image fusion,” *Def. Sci. J.*, vol. 60, no. 1, pp. 48–54, 2010, doi: 10.14429/dsj.60.105.

[129] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002, doi: 10.1109/97.995823.

[130] M. Gad, A. Zaki, and Y. M. Sabry, “2017 , 34 th NATIONAL RADIO SCIENCE CONFERENCE Arab Academy for Science , Technology & Maritime Transport Silicon photonic mid-infrared grating coupler based on silicon-on – insulator technology 2017 , 34 th NATIONAL RADIO SCIENCE CONFERENCE Source Gratin,” *Arab Acad. Sci. Technol. Marit. Transp. II.*, no. Nrsc, pp. 400–406, 2017.

[131] L. F. Zoran, “Quality evaluation of multiresolution remote sensing images fusion,” *UPB Sci. Bull. Ser. C Electr. Eng.*, vol. 71, no. 3, pp. 37–52, 2009.

[132] A. Kaur, L. Kaur, and S. Gupta, “Image Recognition using Coefficient of Correlation and Structural SIMilarity Index in Uncontrolled Environment,” *Int. J. Comput. Appl.*, vol. 59, no. 5, pp. 32–39, 2012, doi: 10.5120/9546-3999.

[133] V. Radhika, K. Veera Swamy, and S. Srininvas Kumar, “Performance evaluation of statistical measures for image fusion in spatial domain,” *1st Int. Conf. Networks Soft Comput. ICNSC 2014 - Proc.*, no. Mi, pp. 348–354, 2014, doi: 10.1109/CNSC.2014.6906716.

[134] W. Pei, G. Wang, and X. Yu, “Performance evaluation of different references based image fusion quality metrics for quality assessment of remote sensing Image fusion,” *Int. Geosci. Remote Sens. Symp.*, pp. 2280–2283, 2012, doi: 10.1109/IGARSS.2012.6351040.

[135] J. Liu, H. Wang, and W. Qin, “A new fusion image quality assessment based on edge and structure similarity,” *2011 IEEE Int. Conf. Cyber Technol. Autom. Control. Intell. Syst. CYBER 2011*, no. 8, pp. 112–115, 2011, doi: 10.1109/CYBER.2011.6011774.

[136] G. Piella and H. Heijmans, “A new quality metric for image fusion,” *IEEE Int. Conf. Image Process.*, vol. 3, pp. 173–176, 2003, doi: 10.1109/icip.2003.1247209.

[137] J. Oliver, “Objective image fusion performance measure,” *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013, doi: 10.1017/CBO9781107415324.004.