FUSION OF MULTIMODAL NEUROLOGICAL IMAGES

Ph. D. Thesis

by

SNEHA SINGH

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
AUGUST, 2019

FUSION OF MULTIMODAL NEUROLOGICAL IMAGES

A THESIS

Submitted in partial fulfilment of the
requirements for the award of the degree

of

DOCTOR OF PHILOSOPHY

in

ELECTRICAL ENGINEERING

by

SNEHA SINGH

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE – 247667 (INDIA)
AUGUST, 2019

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE

STUDENT'S DECLARATION

I hereby certify that the work presented in the thesis entitled **"FUSION OF MULTIMODAL NEUROLOGICAL IMAGES"** is my own work carried out during a period from December, 2016 to August, 2019 under the supervision of Dr. R. S. Anand, Professor, Department of Electrical Engineering, Indian Institute of Technology Roorkee, Roorkee.

The matter presented in the thesis has not been submitted for the award of any other degree of this or any other Institute.

Dated: **November 26, 2019** (**SNEHA SINGH)**

SUPERVISOR'S DECLARATION

This is to certify that the above mentioned work is carried out under my supervision.

Date: **November 26, 2019**

(R.S. Anand)
Supervisor

The Ph.D. Viva-Voce Examination of **SNEHA SINGH**, Research Scholar, has been held on **November 26, 2019**.

Chairperson, SRC **Signature of External Examiner**

This is to certify that the student has made all the corrections in the thesis.

Signature of Supervisor
Dated: **November 26, 2019**

Head of the Department

ABSTRACT

Medical imaging plays a vital role in modern medicine. With the advent of modern technology, there is a tremendous improvement in the capabilities of several medical imaging modalities such as X-ray, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound (US) and functional imaging modalities such as single photon emission computed tomography (SPECT), positron emission tomography (PET) etc. which are extensively prescribed by the clinicians/radiologists for diagnosis purposes. In usual, the diagnostic procedures based on the perception of medical images are performed in a subconscious way which is based on the conclusion drawn upon how the clinicians understand and interpret them. However, due to several sources of medical images used by the clinicians and radiologists, a big problem of information overloading occurs. Moreover, none of the medical imaging modality is able to provide comprehensive and accurate information, especially in critical diseases such as brain hemorrhage, tumor, cancer, other nervous system disorders, any accidental injuries, etc. For example, anatomical imaging (CT/MR) provides morphological information about the human body, but do not reflect the functional status, whereas functional imaging (SPECT/PET) provides the physiological information, but do not reveal anatomical information. Therefore, it is necessary to correlate one modality of medical images to other to provide the significant diagnostic information that requires lots of years of experience and this process is very rigorous, costly and time consuming and has the chance of lots of human errors. Moreover, the advanced imaging modalities prescribed by the doctors multiple times, are too much costlier that also puts an extra financial and mental burden on an individual. Therefore, there is a need to develop some effective multimodal medical image fusion (MIF) approaches to merge all of the features taken from each of individual modality into a single composite fused image that has a significant clinical information and is suitable for the effective diagnostic analysis.

Thus, in the above perspective, the medical image fusion algorithms should fulfil the following three principal criteria:

1. The MIF algorithm must be capable to preserve maximum diagnostic information from the input images with perceptible visual quality and without introducing any spatial and spectral distortion.
2. The true tissue information either anatomical or functional, including the edges and other diagnostic details should also be reflected properly.
3. The fusion algorithm must be computationally efficient, stable and robust.

With the above background, the main objective of the present research work has been chosen as to design and develop the effective fusion approaches by integrating all the complementary and contrasting information from the different image datasets of same organ and tissues so that the fused image is more useful and acceptable to the human visual

system and machine perception. Accordingly, the entire research work has been planned and carried out under the following two major objectives.

1. A comparative evaluation of several existing fusion approaches has been carried out and new efficient CT and MR image fusion approaches have been designed, developed and implemented to improve the fusion performance by preserving the clinically relevant information present in the source CT and MR images with higher contrast level and without introducing any artifacts.
2. Based on the post-analysis of the existing and developed fusion approaches, new suitable anatomical and functional (MR, SPECT, PET) image fusion approaches have been designed to reflect the anatomical details produced by the MR/CT images without disturbing any functional status of the tissue reflected in the SPECT/PET images.

In order to achieve the first objective of the initial phase of work by developing and implementing the MIF approach, three different fusion techniques are proposed in the present work that will also lead to fullfil the different sub-objectives.

Based on the prominent features and advantages of multiscale transformation techniques presented in the literature, the first fusion approach based on the nonsubsampled shearlet transform (NSST) is proposed in the present work. In the proposed NSST domain medical image fusion (NSST-MIF), anatomical (CT-MR) image fusion is performed in the NSST domain using a modified pulse coupled neural network. The proposed fusion approach incorporates the regional energy (RE) based activity level measure to fuse the low frequency (lf) NSST coefficients and novel sum modified Laplacian (NSML) motivated PCNN to fuse the high frequency (hf) NSST components which help to reflect more amount of informative contents present in the source CT and MR images. The performance of the proposed approach is compared with eight different fusion approach in which WT, NSCT and NSST with different fusion rules such as averaging of lf coefficient fusion, maximum and spatial frequency motivated PCNN based hf subband fusion are utilized. Their performance is not only analyzed and evaluated in terms of visual perception, but also in terms of different performance measures such as entropy (En), standard deviation (STD), mutual information (MI), spatial frequency (SF), image quality index (IQI) and Xydeas edge index (XEI). Based on the experimental results, it is observed that the proposed NSST-MIF approach is able to fuse the CT and MR images in a better way without distorting the information and showing a significant improvement in detectability of the source images.

Based on the findings obtained from the results presented in the literature, it observed that the curvelet transform also produce better results, however, it uses a parabolic scaling law to resolve the two-dimensional singularities along C^2 curves. To represent the diagnostic edge detail more efficiently, discrete ripplet transform with two new additional parameters is utilized in second proposed fusion approach which provides a new tight frame with a sparse representation for the source images with discontinuities along C^d curves, where $d = 2$ refers to parabolic scaling the same as curvelets and $d = 3$, refers that ripplet has the cubic scaling and so forth. In the second proposed image fusion approach named as DRT-MIF, firstly, DRT has been applied to decompose the source images, individually in one lf and several hf ripplet components which are fused by computing NSML and novel sum modified spatial frequency (NMSF) motivated PCNN model that is able to capture the fine details present in the reference images. This model helps to preserve redundant information also. The PCNN model is utilized for lf and hf DRT coefficients based on the firing times and improved feeding inputs as NMSF for hf components and NSML for lf ripplet subband. Fusion rules help to capture the suitable differences and provide the resultant images with high contrast and clarity. Finally, fused images are reconstructed by applying the inverse DRT. The results of the proposed DRT-MIF method is compared with wavelet transform (WT), nonsubsampled contourlet transform (NSCT) and NSST with PCNN based fusion approaches including the proposed NSST-MIF. It is observed from their comparative results that the proposed DRT-MIF approach provides a better quality of fused images by preserving the edge and important morphological information. Moreover, the proposed approach provides higher values of En, MI, SF and XEI than NSST-MIF and other existing fusion methods.

Based on the analysis of experimental results obtained earlier in two proposed fusion methods, it is observed that both, the NSST and DRT both the decomposition methods provide better fusion results. DRT helps to reflect the higher-order singularities whereas, NSST overcomes the problem of shift-invariance and helps to lose the important information. Considering their motivation, a cascaded medical image fusion (C-MIF) framework has been proposed for CT and MR images in DRT and NSST domain. At the first stage decomposition, a PCNN model motivated by different feeding inputs such as NSML and NMSF is utilized to fuse the lf and hf DRT coefficients, respectively. The NSST decomposition is used in the second stage where regional energy is computed to fuse the lf NSST approximation coefficients and for hf NSST detail coefficients, the sum of absolute difference (SAD) and absolute maximum (AM) based fusion rules are applied to provide the richer representation of the edge detail information with improved contrast, respectively. The fusion performance of the proposed C-MIF approach is validated by extensive simulations performed on a different CT-MR image dataset and a detailed comparison is made with WT, dual-tree complex wavelet transform (DTCWT), NSCT, NSST, stationary wavelet transform (SWT) decomposition based and other fusion approaches. It is observed from their comparative analysis that the C-MIF approach gets more informative content in the fused image by computing higher values of En and MI as compared to the others. Moreover, the C-MIF approach ensures to retain the contrast and edge detail information by producing higher STD, SF and XEI values than others, thus providing the fused images with better visual quality.

In order to achieve the next objective of the present work, four different multimodal image fusion methods are proposed to fuse MR-SPECT, MR–PET along with CT-MR Images. All these four MIF approaches are developed and implemented in such a way that all these methods help to provide more robust and clean structural detail information without introducing any artifacts and without altering the functional information of the tissue reflected in the source images.

To achieve the fusion of anatomical and functional images, an improved fusion approach has been proposed that uses the entire features extracted by NSST and adaptive PCNN model (ADPCNN) to retain the desirable contrast and detail information in the fused results. In the proposed approach (ADP-MIF), the ADPCNN model is applied to fuse the lf NSST decomposed coefficients with adaptive linking strength parameter based on local visibility and NSML motivated feeding input which helps to provide higher sensitivity and clarity in the visual perception. For fusing the hf NSST coefficients, a local log Gabor energy (LLGE)-based fusion is used to extract optimal texture feature with broad spectral information. The fusion performance is compared with existing fusion methods. It is observed from the fusion results that the proposed ADP-MIF approach is able to reproduce the significant visual information with the preservation of structural and spectral content and provides a clear picture of edge details available in the source images. Furthermore, the fusion performance of the proposed ADP-MIF approach is also compared with twenty seven existing image fusion methods for CT-MR images. From the results, it is concluded that the proposed ADP-MIF approach helps to provide a significant improvement in terms of the visual quality of fused images by providing additional diagnostic information especially for the fusion of anatomical with functional images.

In the next proposed fusion approach, a hybrid multimodal medical image fusion (H-MIF) approach based on the NSST and SWT decomposition has been proposed for the anatomical with functional and anatomical with anatomical images. In the proposed approach, SWT is applied only on the lf NSST subband and letting the hf NSST components to remain the same. After SWT decomposition, another lf and hf SWT subbands are produced. To fuse the lf SWT coefficients, an ADPCNN model motivated by NSML based fusion rule is utilized while LLGE based fusion rule is applied for hf SWT coefficients to extract the salient features available in the source image and to retain the color and edge details without introducing any artifacts. Finally, AM and SAD based fusion rules are applied to remaining hf NSST subbands to retain more information related to edge details. The proposed H-MIF maintains the spatial and spectral details well with sharp minute details shown by higher scores of En and IQI. The higher values of FMI obtained for the proposed method indicate that the brighter minute features in source images are preserved properly with the appropriate consistency and localization. Moreover, similar performance has also been reflected for CT-MR image fusion by the proposed H-MIF approach that achieves good complementary information with more structural details but it suffers from contrast reduction of fused images having lower STD value that may be acceptable with its ability to retain more diagnostic information.

Based on the results and limitations (chromatic imbalance, overbrightness, sensitive to random noise, etc.) of the state of the art methodologies, a unified multimodal fusion framework named as SDL-MIF has been proposed here using multiscale geometric analysis with sparse representation (SR) and guided filtering. The proposed SDL-MIF approach is based on the sparse K-SVD dictionary learning and guided filtering in the NSST domain in which an overcomplete dictionary is learned (training of medical image dataset) to capture complex details of medical images and sparsely represented lf NSST subband for better visual feature (luminance, contrast) projection without any spectral distortion. Fusion rule using a dictionary learning (DL) based SR is utilized to improve the comprehensive information in lf NSST subband, while guided filtering based rule is adopted to fuse hf NSST subbands, which is able to extract the salient features from the source images and reflect color and edge detail properly in the fused outcomes without incorporating any artifacts. Several experimental results are performed on MR-CT, MR-PET and MR-SPECT dataset to validate the proposed SDL-MIF method and showed a detailed comparative analysis with the other available fusion methods. Based on the comparative experimental results, it is observed that the proposed SDL-MIF method is able to preserve the significant information of multimodal input images by producing better visual quality of fused images with improved contrast.

In the next fusion approach, a feature level multimodal image fusion framework (CNN-MIF) has been proposed using two-scale $\ell_1 - \ell_0$ hybrid layer decomposition with convolutional neural network (CNN) based feature mapping and structural patch clustering. In the proposed CNN-MIF approach, ℓ_n-norm based two scale hybrid layer decomposition is utilized to preserve the desired edges and intensity variations at each scale. A pre-trained CNN model followed by consistency verification is used to extract the prominent features from each of the decomposed base layer components and to generate the pixel activity and fusion weight map. For each output feature map, RE based activity measure is computed and refined in the consistency verification step to optimize the activity weight map for merging the decomposed base layers. The two-scale detail layers are merged by utilizing clustering based pre-learned multichannel dictionary with saliency matching rule to efficiently map the structural details of the layers. Moreover, the color components associated with both the source images are also combined using pixel saliency measure and finally all three components, i.e. fused base layer, fused detail layer and color component get merged to reconstruct the fused image. From the experimental results, it is observed that the proposed CNN-MIF approach highlights the ability to preserve the layer information (structural, fine details, brightness, and color) and increases fusion accuracy. A ℓ_n-norm based two-scale hybrid layer decomposition method is used to separate out the main information from the textural details in the spatial domain. The experimental results also show that the proposed CNN-MIF approach can efficiently extract the complex structure and maintains the spectral information as well without introducing any processing artifacts. It is further observed that the proposed CNN-MIF approach achieve higher performance measures than the other fusion approaches which itself signifies an improvement in the results of the proposed CNN-MIF approach.

For the purpose of implementing and evaluating the performance of the above discussed proposed methods, the multimodal CT, MR, SPECT and PET neurological images were acquired from the multimodal image database available at Harvard whole brain atlas (http://www.med.harvard.edu/AANLIB/home.html).

ACKNOWLEDGEMENTS

I express my gratitude to my supervisor Dr. R.S. Anand, Professor, Department of Electrical Engineering, IIT Roorkee, for his guidance and suggestions for the completion of this study.

I express my sincere thanks to the members of my research committee Prof. G.N. Pillai, Prof. P. Sumathi, Department of Electrical Engineering, IIT Roorkee and Prof. Dharmendra Singh, Department of Electronics and Communication Engineering, IIT Roorkee for their constructive discussions and valuable suggestions during the meetings of my research committee. I am thankful to Prof. N.P. Padhy, Dean, Academic Affairs, IIT Roorkee and Prof. B. Das, H.O.D. Department of Electrical Engineering, IIT Roorkee for providing necessary support and facilities throughout this study and in the submission of this thesis. I am equally indebted to Prof. S.P. Srivastava, Prof. Pramod Agarwal, Prof. Manoj Tripathi, Prof. Barjeev Tyagi, Department of Electrical Engineering, IIT Roorkee, who has always been my source of motivation and pursuance of my work. I am also thankful to all other faculty members of the Department of Electrical Engineering, IIT Roorkee for their support.

I am thankful to Mr. Dinesh Sharma, Mr. Aamir Ahmed, Mr. Rishab, Mr. Mohan Singh and the staff of Biomedical Lab, Library and office of the Department of Electrical Engineering for their co-operation and assistance. I extend thanks to all non-teaching staff, Department of Electrical Engineering, IIT Roorkee for their day to day co-operation.

I am extremely thankful to Dr. Deep Gupta, Department of ECE, VNIT Nagpur for his every help and invaluable support rendered during my research work. I am also grateful to the reviewers for their constructive suggestion and invaluable advice for improving the quality of my publications derived from this work.

I am highly thankful to my friends and fellow researchers Dr. Jayendra Kumar, Dr. Arun Balodi, Dr. Yogesh Kumar Sariya. Dr. Nishant Jain and Dr. Soumi Ray for their necessary help during the work. I am also thankful to my fellow researchers Ms. Prachi Sharma, Ms. Shanti Chandra, Mr. Gaurav and Mr. Ashish Rohila for their kind support during the work.

I cannot forget my brother Anurag and sister Meha who never hesitated to render their help at the time of some critical difficulties and spared their time whenever required during my Ph.D. work. I express my thanks for their help in need. I also extend my thanks to all those whom I have missed to mention.

I also sincerely acknowledge the financial support and assistantship provided by the MHRD, Government of India, for this study.

I am indebted to my beloved son ARAV, parents and in-laws for their pains, encouragements and consistently supports during my Ph.D. work. They are my pillars and without their support, this work was not possible to conclude. I express my deepest appreciation to my well-wishers whose blessings helped me to reach this end. I owe greatly to my Father for his unwavering patience, understanding and encouragement, as the present work is finally concluded, because of his personal sacrifices and supports for which I shall ever remain indebted.

Last but not least, I am thankful to the Almighty, who gave me the strength and health for completing the work.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ADPCNN	Adaptive PCNN
AM	Absolute Maximum
ANFIS	Adaptive Neural-Fuzzy Inference System
ANN	Artificial Neural Network
BIFNN	Biologically Inspired Feedback Neural Network
CBF	Cerebral Blood Flow
CNN	Convolutional Neural Network
CNT	Contourlet Transform
CSF	Cerebrospinal Fluid
CSR	Convolutional Sparse Representation
CT	Computed Tomography
CVT	Curvelet Transform
D	Dimensional
DL	Dictionary Learning
DRT	Discrete Ripplet Transform
EP	Edge Preserving
En	Entropy
FCM	Fuzzy C-Means
FMI	Feature Mutual Information
GDGF	Gradient-Domain Guided Image Filtering
GF	Guided Filtering
GSR	Group Sparse Representation
hf	High Frequency
HSV	Hue-Saturation-Value
HVS	Human Visual System
ICA	Independent Component Analysis
IQI	Image Quality Index
JSR	Joint Sparse Representation
lf	Low frequency
LE	Local Energy
LGE	Log Gabor Energy
LLGE	Local Log Gabor Energy
LP	Laplacian Pyramid
MCA	Morphological Component Analysis

MI	Mutual Information
MIF	Medical Image Fusion
MOD	Method of Optimal Directions
m-PCNN	Multichannel PCNN
MR	Magnetic Resonance
MST	Multiscale Transform
NFIS	Neural-Fuzzy Inference System
NMSF	Novel Sum Modified Spatial Frequency
NSCT	Nonsubsampled Contourlet Transform
NSLP	Nonsubsampled Laplacian Pyramid
NSML	Novel Sum Modified Laplacian
NSST	Nonsubsampled Shearlet Transform
OMP	Orthogonal Matching Pursuit
PCA	Principal Component Analysis
PCNN	Pulse Coupled Neural Network
PET	Positron Emission Tomography
RE	Regional Energy
RT	Ripplet Transform
SAD	Sum of Absolute Difference
SF	Spatial Frequency
ShF	Shearing Filter
SML	Sum-Modified-Laplacian
SOFM	Self-Organizing Feature Mapping
S-PCNN	Simplified PCNN
SPD	Structural Patch Decomposition
SPECT	Single Photon Emission Computed Tomography
SR	Sparse Representation
STD	Standard Deviation
SVD	Singular Value Decomposition
SWT	Stationary Wavelet Transform
Tc	Technetium
TI	Thallium
WGF	Weighted Guided Image Filtering
WT	Wavelet Transform
XEI	Xydeas Edge Index

CHAPTER 1: INTRODUCTION

This chapter presents an introduction to the research work carried out in this thesis. It starts with some background on medical imaging and different imaging modalities. This chapter also presents an exhaustive review of the work done on different fusion algorithms for multimodal medical images. The objectives for present study have been decided on the basis of these reported works. The organization of the present thesis is also given at the end of this chapter.

1.1 Motivation

A large number of critical neurological disorders and diseases have become an important public health issue in the world or especially in developing countries. The diagnosis of neurological diseases is delayed or missed due to absence of generally no symptoms or signs until the diseases have reached an advanced stage. Also, the symptoms may vary from patient to patient. In addition to this, some symptoms may not be very specific to a particular disease and may resemble with the symptoms of others. This needs some hidden information from the human body where medical imaging plays an important role. Nowadays, multimodal medical imaging sensor technology is considered by the clinicians as a prominent solution for the detection and prognosis of many severe neurological disorders such as brain hemorrhage, tumor, cancer, nervous system disorders, accidental injuries, etc.. This is the main reason for attracting towards the fusion of multimodal images which is capable to provide different insights of the human anatomy to identify and monitor the severity of injuries and diseases.

There are several sources of medical imaging modalities such as X-ray, ultrasound, computed tomography (CT), magnetic resonance (MR) imaging, single photon emission computed tomography (SPECT), and positron emission tomography (PET) etc. giving different insights of the human body and opted by the radiologist time to time-based on their perception. However, due to several sources of medical images used by the clinicians lead to a big problem of information overloading occurs because of none of the medical imaging modality is able to provide comprehensive and accurate information especially in case of critical diseases as mentioned above. For example, anatomical modality such as MR imaging reveals the soft tissue structure with higher resolution, while the CT imaging produces hard tissue information of different organs, however, it is limited in soft tissue contrast which is required to differentiate normal and diseased tissues. On the other side, the functional imaging modality like SPECT and PET provides the physiological information by which the actual functional activity of lesion can be known [213]. However, SPECT and PET images are not able to provide anatomical information that lead to a problem for the localization of lesion in functional images. Therefore, it is necessary to correlate one modality of medical images to another one for providing the significant diagnostic information which requires domain expertise. Additionally, this process is very rigorous, costly, time consuming and sensitive to human errors. Moreover, the advanced imaging modalities are too much costlier that also put an extra burden on the individuals.

Another possible reason to fuse the multimodal images is to combine all anatomical and functional information without introducing any structural distortion and without altering the functional content that would be helpful to the clinicians by providing a complete pictorial view of the tissue in a single image. With the fusion perspective, the characteristic details of the reference images to be combined can be classified in three different classes such as 1) common features, that are available in all the reference images, 2) complementary features, that are uniquely available in one of the reference images, and 3) redundant features that do not carry any useful information but sometimes important to enhance the overall image quality.

Therefore, there is a need to integrate all the common, redundant and complementary diagnostic information taken from each individual outcome of different imaging modalities to produce a single composite fused image. It may help to provide a meaningful quantifiable interpretation which is more suitable for clinical diagnosis by providing the exact location and orientation of the defected tissues. These motivations motivated to search for the better image fusion approaches for anatomical with anatomical and functional images to provide an efficient and accurate diagnosis.

1.2 Medical Imaging Modalities: A Brief Overview

Medical imaging refers to many different processes and technologies that aims to provide different insights of the human anatomy in order to monitor, diagnose, treatments or many therapeutic processes based on the image perception. The basic principle of medical imaging follow some fundamental steps; 1) a sensor system generates energy or radiations which can penetrate the human body/anatomical structure, 2) radiations passed through the different tissues or organs that are absorbed or attenuated according to atomic behavior and density level of different anatomical structures and generates a reflected waves, 3) these reflected waves are detected by the detectors associated with the energy source and 4) detected waves are mathematically processed for image formation, computer control and display. Normally, medical imaging modalities are classified based on the energy source either internal and external energy sources or combination of both [39] such as X-ray imaging, X-ray CT imaging, Ultrasound imaging, MR imaging, Thermography, SPECT, PET , and Optical tomography etc.. Additionally, based on tissue details reflected by the different modalities are broadly categorized into anatomical and functional imaging modalities. The anatomical modalities provide the soft and dense tissue or organ details such as size, shape and localization of region of interest, while the functional imaging modalities localize the metabolic activity information such as the brain and cardiac functions, etc. with respect to the tissue or organ of interest. The modalities which are used as primary inputs to image fusion in the present work are discussed in the following sections.

1.2.1 X-ray computed tomography

Computed tomography (CT) is one of the most commonly used tests in neurology and neurosurgery. It plays an important role in the acute evaluation of stroke, head, traumatic injuries. CT images are formed by passing X-rays in circular motion through the object of interest (body parts) at different angles as the patient moves through a gantry containing a rotating X-ray source mounted opposite to the detector array that is being used to collect the X-ray that have passed through the body. Numerous data points collected in this manner are synthesized to form tomographic images and displayed on a gray scale matrix. The idea behind to form the CT images is that the different types of tissue in the human body attenuate the X-ray by different amounts which provides the contrast image because the amount of attenuation depends on the density of tissues. Lower density materials like air and water have less attenuation and are displayed as low densities (dark), whereas the higher density materials like bones, calcification and acute blood have high attenuation and are displayed as high density (bright) on CT. Other soft tissues are represented by multiple shades of gray. However, the contrast of CT image for soft tissues is very poor. A sample of CT images is shown in Figure 1.1. The different tissues and their appearance based on their tissue characteristics on CT image are presented in Table 1.1.

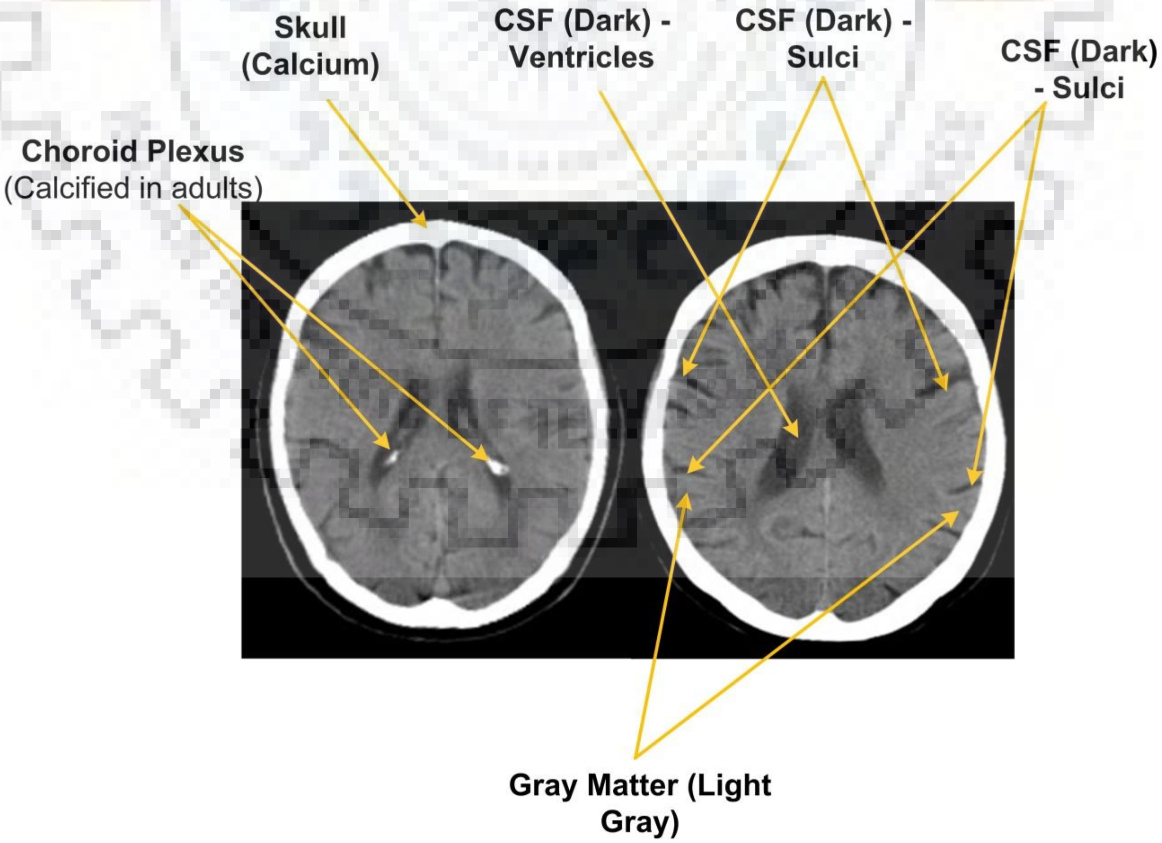


Figure 1.1 High density (bright) and low density (dark) on CT image in a normal patient

Table 1.1 Density based tissue appearance in CT images

Tissue	CT Intensity
Bone/Calcium	Bright
Air	Dark
Water	Dark
Fat	Dark
Infarct	Dark
Cerebrospinal fluid (CSF)	Dark
Bleed	Bright
Tumor	Dark (unless calcified)
Multiple sclerosis (MS) plaque	Dark (often isodense)

1.2.2 Magnetic resonance imaging

Magnetic resonance (MR) imaging is a non-invasive diagnostic technique in radiology which is based on the magnetization properties of atomic nuclei. MR imaging provides exquisite detail of brain, spinal cord and vascular anatomy with an ability to visualize the anatomy in three different axial, sagittal and coronal planes. The construction of the MR scanner is composed of four basic components such as radio frequency (RF) transmission system (40 to 130MHz), a giant circular-shaped permanent magnet to produce strong uniform magnetic field (0.5T to 3T), an RF coil, and three gradient coils to linearly vary the field across the imaging axis. This imaging technology follows the fundamental property of magnetization of certain tissue nuclei present in the body and hydrogen is the only substance that provides adequate small magnetic nuclei (single proton) and found in abundance in different body tissues. A human body contains overall 63% of hydrogen nuclei present in fat and water. Any unpaired protons present in hydrogen atom have the property of nuclear-spinning with the spin of half and opposite about its central axis and produce a signal in the presence of a magnetic field.

In the MR imaging process, subject is placed in a powerful, uniform and external magnetic field which is used to align the hydrogen nuclei (protons) spins in a direction parallel to the field. These protons oscillate at a frequency named Larmor frequency proportional to the external magnetic field. When an RF pulse having a frequency equal to the Larmor frequency is applied, then these protons gain energy and tilt. After removal of RF pulse, protons release energy and try to return to their equilibrium state (relaxation state). The emitted RF pulse signal is picked up by RF coil and is used to form the grayscale image. Different types of images are generated by applying the varying RF pulse sequences. The arrival of protons or hydrogen nuclei to the equilibrium state does not take place immediately, but it takes some time and involves two processes called longitudinal relaxation (T1) and transverse relaxation (T2). One is the factor which measures the time taken to realign the spinning protons in phase with the external magnetic field with time constant T1 and other is the signal strength decrease in time with loss of phase coherence of hydrogen nuclei at a time constant T2. Both T1 and T2 are tissue-dependent. The magnetic gradient field is used to localize spins in space which provides the spatial information enabling an image to be formed. The received signal strength depends on proton density, T1 and T2 relaxation time. Contrast among different tissues is dependent upon how these three parameters differ between the tissues. For most of the soft tissue, proton density is homogeneous and therefore does not contribute in a major way to signal differences seen in an image. So T1 and T2 are responsible for the major contrast between soft tissues. T1 and T2 are mainly influenced by the viscosity or rigidity of a tissue. In general, higher value of viscosity gives smaller T1 and T2. Furthermore, the tissue having shorter T1 is displayed brightly in T1-weighted image (such as fatty tissue) and dark for the tissue having larger T1 (CSF), while the opposite is true for T2-weighted images. MR imaging provides good contrast between the different soft tissues of the body, which make it useful in imaging the brain, muscles, heart and cancers. The physical characteristics of certain tissue based on relaxation time for different types of anatomical materials present inside the human body are shown in Table 1.2 with two different contrast MR images. A sample of MR-T1 weighted and MR-T2 weighted images [68] is shown in Figure 1.2 (a) and (b), respectively.

Table 1.2 Density based tissue appearance in MR-T1 and MR-T2 weighted images

Tissue	MR-T1	MR-T2
Air	Dark	Dark
Fat	Bright	Bright
Dense Bone	Dark	Dark
Water	Dark	Bright
Infarct	Dark	Bright
Bleed	Bright	Bright
Tumor	Dark	Bright
MS plaque	Dark	Bright
Calcification	Dark	Dark
CSF	Dark	Bright
High protein fluids	Bright	Bright
Subacute Hemorrhage	Bright	Bright
Cerebral Infarction	Dark	Bright
Flow void	Dark	Dark


(a) (b)

Figure 1.2 (a) MR-T1 weighted image (b) MR-T2 weighted image

1.2.3 Single photon emission computed tomography

A single photon emission computed tomography (SPECT) scan is a nuclear imaging technique that displays the functional activity of tissues or organs which may help to diagnose and localize seizures, tumors, head injuries, blockage and stroke. In SPECT imaging, a radioactive tracer (radioactive pharmaceutical substance) is injected intravenously to the patients that emits single or multiple gamma ray photons with very high energy. The tracer molecules get metabolized according to the physiological condition of the patient. The SPECT images are acquired by using a scintillating camera known as gamma camera rotated 360º around (normally for optimal reconstruction) the patient’s body to scan the object of interest at multiple angles. These specialized gamma cameras consist of a lead collimator, sodium iodide crystal, and several photomultiplier tubes. The collimator projects the detected gamma photons to the given directions and passes through the scintillation crystal which limits higher photon energy to visible light range followed by a photomultiplier tube. This absorbs the light energy and converts into the electronic signal for image reconstruction. The resultant reconstruction of tomographic images can be performed in two ways such as iterative and filtered backpropagation methods [39].

SPECT image is a map of the distribution of radioactive tracer throughout the body which observes the level of physiological activity based on the absorption of radionuclide. A radionuclide is an unstable radioactive isotope that releases energy by emitting ionizing rays (alpha, beta or gamma). The most widely used isotope in SPECT imaging is Technetium-99m (^{99m}Tc) which is added with different carrier molecules used in brain, kidney, heart and skeletal imaging. Other substance as Iodine-123 (^{123}I) is a pure gamma emitter and used especially in thyroid investigations. Moreover, Thallium-201 is used in brain scan whereas Xenon-133 is used to measure cerebral blood flow (CBF) in various brain regions. The SPECT images made with Technetium and Thallium radioactive isotope of a patient (51year women) suffered from anaplastic astrocytoma [68] are shown in Figure 1.3 (a) and (b), respectively.

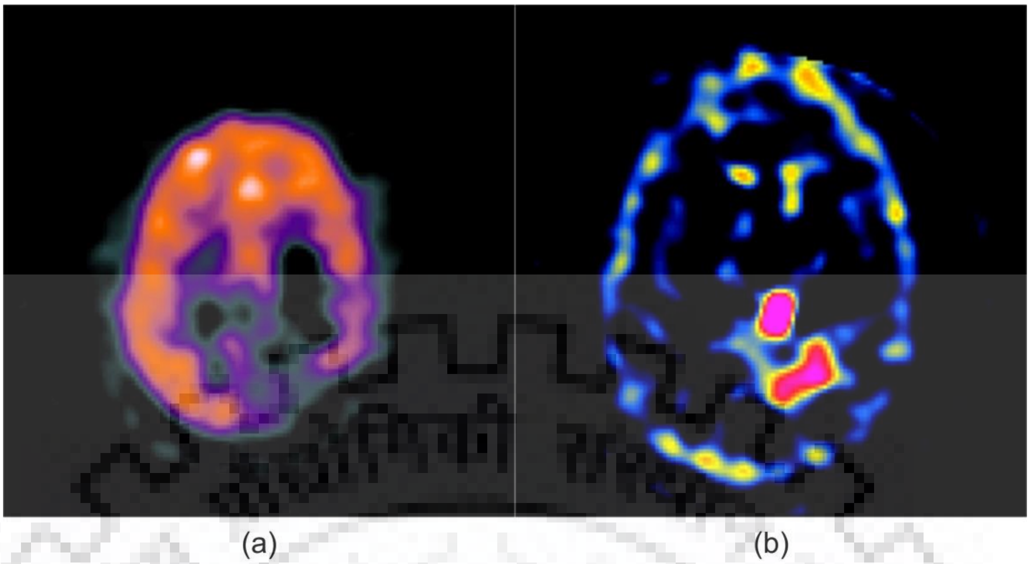


Figure 1.3 SPECT image made with (a) Technetium (Tc) (b) Thallium (TI)

1.2.4 Positron emission tomography

Positron emission tomography (PET) and SPECT scans mainly differ by the type of radioisotope used to form radiopharmaceutical or tracer. PET scan uses positron emitting radioisotopes e.g. Fluorine-18 (F-18). Unlike the SPECT imaging, PET imaging monitors the amount of glucose uptake in the areas with abnormalities or unusual metabolic activity. These abnormalities may result due to several neural disorders such as dementia, Alzheimer’s, tumor, head injury, Parkinson's disease and epilepsy. Similar to the SPECT imaging, in PET imaging modality, a short life radioisotope is injected which is circulated by the blood and amalgamated into the object of interest to measure the biologically active molecules such as oxygen, water and glucose. The physical and chemical properties of radionuclides used to trace the functional activity of subject in PET imaging are different from that used in SPECT.

PET imaging system consists of a detector, scintillating crystal and photomultiplier tubes [39]. The radioactive tracer is injected into the patient that emits the positron. The emitted positron collides with the free electron present in the tissues. Then complete annihilation of positron and electron pair takes place that emits a pair of two gamma rays exactly at 180° angle. These coincident gamma ray pair is detected by the ring detector placed around the patient and image is reconstructed. The most commonly tracer used in PET scan is F-18 combined with glucose to form fluorodeoxyglucose (FDG) which indicates the metabolic activity of tissue with respect to the consumed glucose in that regions (normal, hyperactive or dead cells) and an image is traced based on the concentration of uptake. The cancerous regions which have higher glucose metabolism will take more FDG and results as the bright spots in PET images. The concentration of glucose is higher (hot-spots) in hyperactive region and thus it appears brighter than the normal region, while for dead cells region, this concentration is very low (cold-spots) and appears darker than the normal visualization. The primary applications of PET scans are oncology (cancer detection and treatment), neurology (treatment and surgery of the brain and nervous system) and cardiology (treatment of the heart). Figure 1.4 shows the PET scans of normal brain and a 70 year old male patient having the mild Alzheimer’s disease [68].

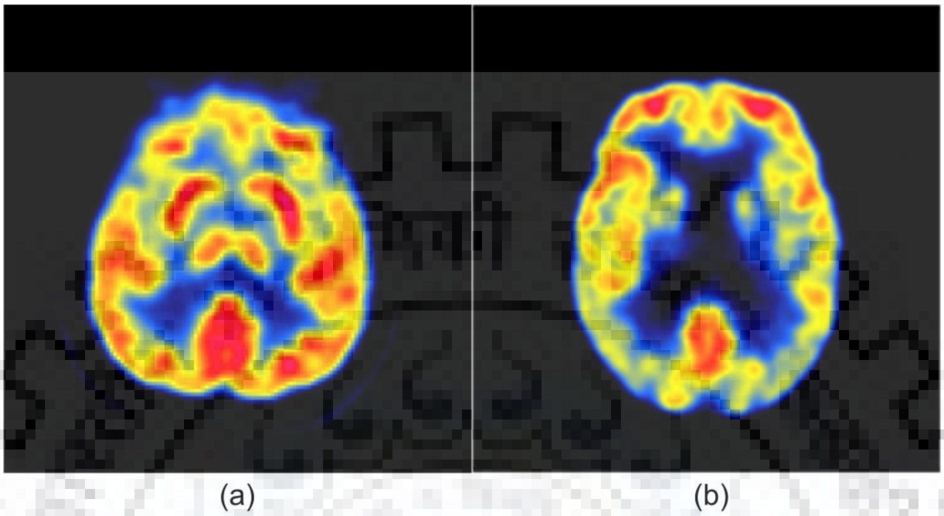


Figure 1.4 Sample of PET images (a) Normal (b) Mild Alzheimer’s

1.3 Literature Review

In recent years, many medical imaging techniques are being extensively used for visualizing the complementary information of the different imaging modalities in order to identify several diseases or disorders present in the human body. The aim of medical image fusion is to combine the information captured from these multimodal images in a single composite image which is more able to boost the diagnostic accuracy and can precisely spot the severity of the deformities.

Over the last few years, many research works have been reported on the image fusion techniques [7, 52, 62, 66, 71, 96]. which are broadly classified into three categories as pixel [96, 113], features [29, 78] and decision level fusion [141]. Pixel level fusion is considered as a low-level fusion that performs directly on the image pixels and able to preserve the actual state measures with less computational complexity. In feature level fusion, different image features such as color, edges, textures, and fine detail, etc. extracted from multiple sources are integrated and finally generate a consolidated weight map for maximum feature matching based fusion with respect to each source input. Decision level fusion corresponds to combine the feature map score generated using different classifiers and take the fusion decisions using maximum voting (or class similarity), fuzzy logic or forecasting basis. All the image fusion techniques can be developed in both the spatial and transform domain as per the application and fusion quality needed. The following sections discuss the brief details about the several image fusion methods such as spatial domain, transform domain, sparse representation based, deep learning based, and hybrid image fusion as stated in the literature.

1.3.1 Spatial domain image fusion methods

Spatial domain image fusion methods are implemented directly on the input images using local pixel statistics. This provides more detailed information, fast execution and low computational complexity. However, these methods suffer from color distortion, poor contrast and pixel consistency [115]. In the spatial domain category, the first experimental demonstration of image fusion was reported in the 1980s that is based on gray level average, weighted average and maximum rule based fusion. Though, this is easy to implement and time-efficient, but also limited by poor contrast and spatial localization. There would also be a loss of the few structures in the fused images. Since then, the studies on image fusion have been investigated a range of spatial domain techniques. However, they suffer from spectral degradation. Partitioning of the image-based fusion method was presented [61] that use the selection of block based on its saliency or activity. In this method, the selection of the block, their size and saliency criteria decide the quality of fused images. It also results in a complete loss of information at each location that would affect the diagnosis based on the analysis of radiologist/clinicians. In the same line, other image fusion methods are also categorized based on numerical and statistical information of the pixels such as hue-saturation-value (HSV) [31], principal component analysis (PCA) [138, 171], independent component analysis (ICA) [29, 124, 125]. Though, color space transform and dimensionality reduction techniques have been successfully used to merge the high resolution grayscale images (MR/CT) with low resolution color image (PET/SPECT) using global fusion strategy for retaining the spatial details of inputs. But, such type of global approaches is not often considered the local variance and consistency between the inputs hence produces the spectral distortion.

To improve the performance of spatial domain image fusion methods, block or region based (overlapped patches) approaches were employed in many research works. The key principle of such approaches is to select image blocks or partitioned regions from the source images with some pre-defined activity measures, such as image spatial frequency (SF), regional energy (RE), sum-modified-Laplacian (SML), variance etc. However, the fused images using a block based approach usually suffer from blocking effect as an image block may capture both clear and flat areas, simultaneously [96]. For region-based methods, it is difficult to achieve the fused result with superior quality since the patchwise image partitioning needs a proper selection of patch size and overlapping steps. In the recent years, several new spatial domain image fusion approaches such as image matting [95], higher-order singular value decomposition (SVD) [102, 129], guided filtering (GFF) [94, 202], weighted guided filtering (WGF) [101] and gradient-domain guided image filtering (GDGF)

[218] based methods have been introduced. These methods can extract sufficient structural information from the source images with higher computational efficiency and spatial consistency required in the fused image but suffer from spatial smoothing or staircase effects at the image boundary. Some authors have successfully employed a neural network and fuzzy logic based fusion approaches for selecting the local pixels or region segment in the source image inspired from the basic principle of the human visual system (HVS) model [35, 88, 89, 168].

1.3.2 Biologically inspired neural network based fusion methods

The artificial neural network (ANN) model basically requires a training set to compute the appropriate weight (set of network model parameters) to analyze and predict the behavior of the given test data in a self-adaptive way. Besides the ability to adapt the variability of the input details (due to different modalities), the fusion performance is affected by the amount and quality of the training inputs. In order to offer the robustness to the ANN based fusion approaches, many other hybrid versions of neural network model have been documented such as self-organizing feature mapping (SOFM) neural network [215] and pulse coupled neural network (PCNN) [73, 88, 180, 182]. The feature mapping type model offers a multilevel fusion scheme that exhibits a self-organizing two-layer neural network used to connect the statistical distribution of the image pixels and clustering. Thus, this method provides efficient performance for complex detail fusion in terms of both the statistical and computational measures,. However, this model leads to bias the network only towards the learned feature patterns of the training examples and may fail to recognize new patterns, hence limits the fusion performance.

On the other hand, PCNN model is a special type of neural network inspired by the cat’s visual cortex system and has been effectively applied in several image processing applications as edge detection, denoising, segmentation, feature extraction, fusion and classification [36, 67, 181]. From many works based on the PCNN model reported in literature, it has been observed that this approach outperforms to several other methods. However, the conventional PCNN consist a complex structure with nine different parameters (three bias potentials, three attenuation time constants, two Gaussian synaptic weights and one linking strength parameter), which comes as a difficult task to select and optimize their values based on the different applications. Thus, to overcome the limitations of parameter tuning complexity of the conventional PCNN, many authors have been proposed the different modified versions of PCNN model such as simplified PCNN (S-PCNN) [48, 105, 176, 177], multi-channel PCNN (m-PCNN) [10, 180], and adaptive PCNN [48, 209], etc.. Furthermore, the performance of these methods is greatly dependent on the input stimuli used to motivate the PCNN and the choice of neighbor pixels linking factor for the appropriate firing map. In the last decade, PCNN model and its variants used in the transform domain are promising for multimodal medical image fusion (MIF). The basic PCNN model is utilized in the wavelet domain [99] and later on, it is applied in multimodal medical image fusion [98]. Qu et al. [145] proposed an image fusion algorithm based on spatial frequency motivate PCNN model in NSCT domain and later on, modified spatial frequency is adapted in same transform domain reported by Das and Kundu [34]. In such methods, PCNN based fusion rules for both the types of subband coefficients have been modelled in a similar manner, while the informative contents carrying by both the low and high frequency subband coefficients are different. In the last few years, the utilization of PCNN model for multisensor medical image fusion has been reported by several researchers [35, 48, 108, 147, 176, 184, 185, 188, 209] in state-of-the-art. In the present work, an advantage point of the PCNN is taken into consideration and a suitable feature measure is used to motivate the simplified and adaptive PCNN for subband processing.

1.3.3 Fuzzy logic based image fusion methods

In the image fusion process, pixel contrast and the visual saliency can be effectively modelled the nonlinear human visual system (HVS) to decide the visually important or unimportant pixel with respect to its neighboring pixels, hence affect the object perception and fusion quality. Although, deciding the image pixel's significance is a subjective mechanism, but to deal with such uncertainty, a fuzzy logic based approaches have been successfully employed by several researchers [12, 63, 120, 146, 150, 163]. The major limitation of such techniques is the selection of appropriate membership function and fuzzy set that can maximize the fusion accuracy and thus improve the quality of fused images. To consider this point, several other fusion techniques based on neural-fuzzy inference system (NFIS) have been reported in state of the art such as neuro-fuzzy logic based fusion [35, 162], fuzzy clustering [118], type-2 fuzzy logic based fusion [201] and intuitionistic fuzzy set based method [8] etc. In NFIS model, the neural network automatically exploits the tuning of membership function that is used as a decision making operator for a significant pixel selection and thus provides an improved fusion performance with less execution time. To further improve the quality of image fusion, neural networks are extended to adaptively select the membership function according to the image pixel's visual parameters using adaptive neural-fuzzy inference system (ANFIS) [183]. Another approach based on fuzzy clustering is used to decide the number of cluster based on the inherent fuzziness or uncertainty and correlation of the input dataset [118].

1.3.4 Transform domain image fusion methods

Transform domain image fusion methods are more efficient compared to spatial domain technique and showed a good dominance over the others as they are able to preserve the overall information very well, but limited by luminance difference and spatial inconsistency [96]. Because of this, multiscale decomposition techniques have been widely used for analyzing the images at different scales and directions to get the desired image fusion results. The key concept in multiscale decomposition based image fusion approaches are the decomposition of both the source images into a low frequency subband and several high frequency subbands at different scales and directions. The second step is to combine the different decomposed coefficients of source images in a single fused representation based on some fusion rules. Finally, a fused image is reconstructed by taking the inverse transformation of fused coefficients.

To boost the performance of image fusion approaches, Toet et al. [165, 166] introduced different pyramid schemes for data fusion. However, this approach suffers from blocking artifacts because fusion scheme failed to provide any spatial orientation selectivity in decomposition processes. To overcome such limitation, Li et al. [86] proposed a wavelet transform (WT) based image fusion approach using maximum selection rule. Pajares et al. [135] also presented a fusion approach with a similar or different resolution level of multiple images. In [142], Pradhan et al. reported another WT based fusion approach for multispectral and panchromatic images. Palsson et al. [136] discussed a model-based fusion approach in the WT domain and PCA. Yang et al. [198] also presented a WT based image fusion approach in which visibility and variance based fusion rules are selected to fuse the low and high frequency subband coefficients, respectively. Though, WT based image fusion methods are capable to capture the one dimensional singularity. It means that WT is able to reflect only limited directional information, thus it causes artifacts along the edges and may lose the important diagnostic information. Besides this, Chavan et al. [24] proposed a nonsubsampled rotated complex WT based fusion approach for the diagnostic purpose and post-treatment review of Neurocysticercosis by combining CT and MR image data of the same patients and also shown the extracted complementary and edge related features. In another study, the dual-tree complex WT based fusion approach is presented in which after decomposing the reference images into low and high subimage coefficients, its relevant features are selected using principal component analysis (PCA) [15]. Furthermore, several other multiscale decompositions as complex WT [154], stationary wavelet transform (SWT) [6, 63], wavelet packets [9] have been introduced to design an effective platform which provides better localization of image contour and consistent texture details with superior visual image quality.

To overcome the limitation of WT, ridgelet transform is introduced to extract the edges [40], but it did not do well to capture the curve edge details. So, Starck et al. have introduced curvelet transform (CVT) [156] which is capable to capture the 2-D singularities of any arbitrary curve. Considering the advantages of CVT, several researchers presented CVT based image fusion approaches [27, 90, 133]. CVT based fusion methods show better fusion results with more edge details, however, its fusion performance suffers from the problem of shift-invariance. In addition to this, CVT can also not built directly in discrete domain [121]. Furthermore, some researchers have been turned towards the contourlet transform (CNT) and proposed CNT based image fusion methods for medical images [4, 5, 16, 33, 195] and others [106, 122, 143, 197]. Moreover, Srivastava et al. [155] proposed the local energy based fusion rule that is more effective than the single-pixel fusion rule. However, the results still suffer from the lack of shift-invariance that leads to ringing effect in the resultant images. All these limitations of the WT, CVT and CNT decomposition are resolved by evolving a concept of the nonsubsampled version of contourlet transform (NSCT).which is widely used in image fusion schemes [30, 105, 159, 176, 199]. Yang et al. [196] also proposed a fusion rule as a combination of the neural model and stationary wavelet based NSCT. In 2012, Das and Kundu further presented a modification in the computation of spatial frequency based on the NSCT [34]. Chai et al. [23] utilized NSCT decomposition alongwith the novel sum modified Laplacian (NSML) to fuse the image components and preserved the local features available in the source images. Bhatnagar et al. [18] used phase congruency and directive contrast based fusion rule for multimodal medical images in NSCT domain. Padma et al. [47] also discussed the NSCT based medical image fusion approach based on energy computation as a fusion rule that showed the suppression of artifacts compared with the WT but leads to computationally inefficient. Yang et al. [200] also used NSCT to fuse low frequency and high frequency decomposed components using sum modified Laplacian (SML) and log Gabor energy, respectively. Based on the observations of all aforementioned methods, it is presented that they provide better visual quality results, however, their decomposition suffers from the limited number of the directions [74]. In another fusion method, Yang et al. proposed an NSCT based fusion approach using fuzzy PCNN [203]. They have further [201] presented an approach using a local type-2 fuzzy entropy based high band fusion and local energy based low band fusion in NSCT domain which enhances the detail information in the fused outcome and significantly improve the visual quality compared to other popular fusion schemes. Such methods improve the overall image contrast and quality but limited by less directional sensitivity for complex contours present in the image. Furthermore, Jun et al. [187] introduced discrete ripplet transform (DRT) type I which generalizes the CVT by adding two new parameters that assured to present the singularities along an arbitrarily shaped curve. The DRT is also able to overcome the limitations of other transformation approaches by providing the sparse representation of an image object [32, 70]. In another DRT domain fusion method [51], the authors presented a fusion approach for remote sensing data and presented better results than the WT and other decomposition techniques. In [70], authors proposed a hybrid approach using the WT and DRT in which the approximation component obtained after the WT decomposition, is further decomposed using the DRT. However, the results are still suffered from shift-invariance effect.

To overcome the limitations of directional components and provide the fused image with better visual appearance, nonsubsampled shearlet transform (NSST) having multiscale and multidirectional property is utilized in several image fusion approaches [46, 58, 65, 74, 116, 147, 170, 172]. The NSST also overcomes the problem of CNT and NSCT by providing a sufficient number of the directional coefficients of the source images [48, 75]. Ganasala and Kumar [48] utilized different regional features to compute the linking strength of PCNN model and to motivate the PCNN model for fusing low and high frequency image coefficients in NSST and achieved good fusion results. Guorong et al. [54] reported another NSST domain fusion approach for multi-focus images and showed the superiority of NSST over the WT and NSCT based fusion techniques. In [175], the author presented multimodal medical image fusion approach based on shift-invariant shearlet transform in which averaging and maximum fusion rules were utilized to fuse the decomposed coefficients. Yin et al. [209] introduced parameter adaptive PCNN using a weighted score of local energy (LE) and NSML in NSST domain to improve the fusion accuracy by preserving the local information with edge details. Though, the NSST based fusion results look promising but lack in pixel contrast and preservation of tiny edges. The redundant and shift-invariance property promises superior performance for several image fusion applications. In addition to this, the PCNN and fuzzy logic based pixel selection have been widely adopted in many fusion approaches in transform domain [35, 76, 147, 177, 185, 188, 203, 217]. The present work has also been carried out by using NSST and others.

1.3.5 Sparse representation based image fusion methods

The main idea behind all aforementioned conventional transform domain approaches is that the significant information available in the source images can be extracted robustly from the decomposed coefficients. In addition to the transform domain approaches, sparse representation (SR) model has also been widely applied in several image fusion applications [64, 96, 110, 132, 185, 191, 207, 208, 212, 219]. The basic principle of SR based model is that an image can be compactly represented as a linear combination form of non-zero pixel elements or transformed coefficients using an overcomplete dictionary learned from the training dataset (size of a dictionary must be higher or equal to the dimension of the image, i.e. an over-completeness). Yang and Li [191] proposed an SR based image fusion approach for multi-focus images in which firstly, sparse coefficients (as pixel activity measure) are merged for two source images. Sparse coefficients are calculated on local patches of the image and fused using choose-max rule which significantly improve the contrast, invariance, and stability. Since SR based rules work on a patchwise image partitioning in the spatial domain, they are very much sensitive to random noise that can cause spatial variability and computational inefficiency. Because of the overcomplete dictionary learning phase, this method is computationally more complex as compared to the transform domain fusion techniques. However, an offline overcomplete dictionary learning approach outperforms the other fusion methods in terms of execution time, visual clarity, noise suppression and information preservation, but still lack of spatial consistency and smoothing.

In order to improve the performance of the SR based theory in image fusion, many researchers have explored such method and introduced different variants of SR such as joint sparse representation (JSR) [212] and group sparse representation (GSR) [93]. Intuitively, the multimodal data inputs do not necessarily have feature similarities between them, hence could be uncorrelated to each other. Thus, the defined features with similar sparsity coefficient may not signify the multiple prominent features sensitive to HVS model. The JSR and GSR based fusion approaches are able to define those unreliable features extracted from the different source inputs using multiple sub-dictionaries in feature space for a given kernel matrix. Liu et al. [109] employed SR based fusion in the multiscale domain to overcome most of the limitations of classical transform-based and SR based fusion approaches which can be easily observed from their experimental results. However, the multiscale transform (MST)-SR-based rule leads to the chromatic imbalance and causes over brightness, as pixel saliency is not well captured during the fusion process. A feature-level fusion method is developed for visual tracking based on joint SR [80], but the unreliable features are not always correlated with the similar sparsity coefficients in visual tracking application. Thus, the single sparse dictionary can not signify the multiple features. In another approach, the joint SR-based fusion framework is utilized for the visual tracking that is able to perform in feature space for a given kernel matrix [81]. Furthermore, multiple SR-based framework has been presented by Lan et al. [82] to explain the different features by decomposing the multiple sparsity models, which ensure more representative learned features. To overcome the drawback of aforementioned SR based approaches, Liu et al. [110] introduced another efficient model named convolutional SR (CSR) that focuses to get the sparse representation of the entire image instead of local patch based approximation.

1.3.6 Deep learning based image fusion methods

Currently, deep learning has been efficiently introduced in several applications of image processing and computer vision such as image segmentation, visual recognition and classification, feature extraction and mapping, fusion, and decision making [1, 14, 57, 69, 84, 104, 111-113, 126]. The main issues with the most of image fusion techniques are the manual selection of pixel activity measures to correlate the pixels with their corresponding neighboring pixels and to find the optimal weight to integrate the activity which can improve fusion quality. However, it is a complex procedure to identify the best activity measure and corresponding weight scores based on the different source images and applications. Based on the above considerations, Liu et al. [112] reported a convolutional neural network (CNN) based fusion method for multifocus images which is further extended for multimodal medical images with the modifications to improve its computational efficiency [111]. A multimodal medical image fusion scheme based on CNN having three main steps namely, feature extraction, selection and forecasting [57] has taken an important role in the processing steps for better performance of pixel to decision level fusion. The basic idea is to train a CNN model by optimally learn the network parameters for automatic feature selection and mapping from the source inputs and activity weight assignment. Hermessi et al. [57] presented CNN based image fusion framework using similarity learning in shearlet domain. Although deep similarity learning provides promising fusion results because of direct mapping of visually important features, however, the main drawback of CNN based fusion methods is computational time (for model training), as the trained model predict focused or defocused features. Thus, for medical images, the slight modifications in the model are required to capture the complex textures and an extensive data is also required for feature modelling which is difficult in the field of multimodal medical domain [53].

1.3.7 Hybrid image fusion methods

In addition to aforementioned fusion scheme categorization, there are three key steps to design an efficient fusion technique such as 1) to capture the pixel characteristics (using transformed coefficients or spatial distribution) for activity mapping, 2) to obtain the optimal weight score for activity measures and integrate using a suitable fusion rule, and 3) to reconstruct, post-processing (if required) and evaluate the quality of fused images and fusion performance. Over the last few years, considering all key parameters, some hybrid fusion frameworks have been proposed for multimodal medical images [16, 72, 91, 117, 205].

An efficient fusion approach based on structural patch decomposition (SPD) was proposed to preserve the salient features of the image and combined patch-wise basis using different fusion rules [117]. The SPD based method is able to retain most of the information, time-efficient and robust to several artifacts (ghosting effect and blocking artifacts). Yang et al. [205] introduced a unified methodology for multimodal image fusion using SPD and fuzzy-discrimination. This scheme gained good results for gray (anatomical) as well as color (functional) medical images and also outperformed the other techniques in terms of visual quality of the fused images. In addition to this, the edge-preserving spatial domain filters and morphological filters have also gained great attention towards the development of image fusion algorithms. The edge-preserving (EP) filters such as bilateral filter [153, 164], guided filter [101, 107, 128, 193, 204, 218], alternating sequential filter [216], weighted least square filter [64], morphological filters and pyramids [123, 148, 169] have been successfully employed to achieve a significant improvement in fusion results with lower computational time. However, morphological filters such as Top hat and Bottom hat class filters are basically mathematical operators which are able to preserve and enhance the important diagnostic contents available in source images such as edge, contrast and brightness with a significant noise suppression with the help of object morphology (shape) based function. Still, there is a scope to improve spatial and spectral quality of the fused images in anatomical and functional image fusion using either an activity level measurement parameter or using a suitable transform, SR or deep learning approach in decorrelated color model.

1.4 Performance Evaluation Measures

Quantitative evaluations are often used to facilitate the conclusions by computing some numerical values and quantify the quality of fused images in comparison to original one. Therefore, our main emphasis is to improve the following performance measures.

(a) *Entropy* (En): The entropy (En) is used to measure the information present in the reference as well as fused images. It is defined as the average number of bits required to quantize the image intensities. Based on the principle of Shannon's information theory, the entropy is given by

$$En = -\sum_{i=0}^{L-1} p(i) log_2 p(i) \tag{1.1}$$

where $p(i)$ is the probability of gray level with the range $[0, \ldots, L-1]$. Higher values of the entropy reflect the more information content present in the fused images.

(b) *Mutual information* (MI): Mutual information is a measure of the quality of fused image by evaluating the amount of information preserved from multiple source images to a fused image. The mathematical expression of MI is given by,

$$MI = I(x_R; x_F) + I(x_S; x_F) \tag{1.2}$$

$$I(x_R; x_F) = \sum_{u=1}^{L} \sum_{v=1}^{L} h_{R,F}(u,v) log_2 \frac{h_{R,F}(u,v)}{h_R(u) h_F(v)} \tag{1.3}$$

$$I(x_S; x_F) = \sum_{u=1}^{L} \sum_{v=1}^{L} h_{S,F}(u,v) log_2 \frac{h_{S,F}(u,v)}{h_S(u) h_F(v)} \tag{1.4}$$

where R and S refer to te source images and F represents a fused one.

(c) *Standard deviation* (STD): It is a measure of the contrast shown by the fused image. Higher STD values present the higher contrast of fused images. It is represented by,

$$STD = \sqrt{\sum_{i=1}^{M} \sum_{j=1}^{N} \frac{\left[I(i,j) - \left(\frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} I(i,j)\right)\right]^2}{M \times N}} \tag{1.5}$$

where $I(i,j)$ denotes pixel values of fused image at position (i,j). With the increment of standard deviation of the fused image, there is no doubt that grayscale will distribute discretely and the contrast between them is becoming more noticeable by obtaining a fused image with more details.

(d) *Spatial frequency* (SF): The overall activity and clarity level of an image can be measured by evaluating the spatial frequency which is represented by [47],

$$SF = \sqrt{RF^2 + CF^2} \tag{1.6}$$

where the RF and CF refer to the row and column frequency, respectively,

$$RF = \sqrt{\frac{1}{M(N-1)} \sum_{i=1}^{M} \sum_{j=2}^{N} (I(i,j-1) - I(i,j))^2} \tag{1.7}$$

$$CF = \sqrt{\frac{1}{(M-1)N} \sum_{i=2}^{M} \sum_{j=1}^{N} (I(i,j) - I(i-1,j))^2} \tag{1.8}$$

where M and N denote the size of an image and $I(i,j)$ denotes a gray level value of fused image. Higher spatial frequency indicates the better fusion quality.

(e) *Image quality index* (IQI): Wang et al. [179] proposed an index which is designed by modelling any image distortion that refers to the combination of three factors such as contrast distortion, luminance distortion and loss of correlation. The Q_0 values may vary within the range of [-1 1]. It is defined by,

$$Q_0(F,R) = \left(\frac{\sigma_{FR}}{\sigma_F \sigma_R}\right) \cdot \left(\frac{2\mu_F \mu_R}{\mu_F^2 + \mu_R^2}\right) \cdot \left(\frac{2\sigma_F \sigma_R}{\sigma_F^2 + \sigma_R^2}\right) \tag{1.9}$$

where $\mu_F,\ \mu_R$ is the mean and σ_F^2, σ_R^2 refer to the variance of the fused and reference image, respectively. In the fusion process, there are two reference image R and S and one fused image F. So the quality metric (Q_0) is evaluated as,

$$Q_0 = \frac{Q_0(F,R) + Q_0(F,S)}{2} \tag{1.10}$$

If the Q_0 achieves values closer to unity that signifies better quality of fused image is obtained using a particular fusion approach.

(f) *Feature mutual information* (FMI): Feature mutual information (FMI) reflects the degree of dependence of two images [56]. Higher values of FMI indicate the higher regional mutual information corresponding window size 3×3 in between the fused and source images. The FMI is given by,

$$FMI = \frac{I_{FR}}{H_F + H_R} + \frac{I_{FS}}{H_F + H_S} \tag{1.11}$$

where $I_{FX} = \sum_{f,a} p_{FX}(i,j,p,q) \log_2 \frac{p_{FA}(i,j,p,q)}{p_F(i,j) \cdot p_X(p,q)}$ and H_X refer to the histogram based entropies of the corresponding image.

(g) *Xydeas edge index* (XEI): The edge strength [190] is defined as a measure of the edge information that is preserved in the fused images from the source images, individually. For two input images R and S, and fused image F, sobel edge detection approach is used to obtain the edge strength $g(i,j)$ and orientation $\alpha(i,j)$ information for each pixel $p(i,j)$, $1 \le j \le N$ and $1 \le i \le M$.

$$g_R(i,j) = \sqrt{s_R^x(i,j)^2 + s_R^y(i,j)^2} \tag{1.12}$$

$$\alpha_R(i,j) = \tan^{-1}\left(\frac{s_R^y(i,j)}{s_R^x(i,j)}\right) \tag{1.13}$$

where $s_R^x(i,j)$and $s_R^y(i,j)$ refer to the output of the horizontal and vertical Sobel templates centred on the pixel $p_A(i,j)$ and convolved with the corresponding image pixels R. The mathematical formulas of the relative strength and orientation values of $G^{FR}(i,j)$and $\alpha^{RF}(i,j)$ of an input image R with respect to the fused image is written as,

$$G^{FR}(i,j) = \begin{cases} \frac{g_F(i,j)}{g_R(i,j)}; & g_R(i,j) > g_F(i,j) \\ \frac{g_R(i,j)}{g_F(i,j)}; & otherwise \end{cases} \tag{1.14}$$

$$\alpha^{FR}(i,j) = 1 - \frac{|\alpha_R(i,j) - \alpha_F(i,j)|}{\pi/2} \tag{1.15}$$

The edge strength (Q_g^{FR}) and orientation preservation (Q_α^{FR}) values are given as,

$$Q_g^{FR} = \frac{\Gamma_g}{1 + e^{\varkappa_g(G^{FR}(i,j) - \sigma_g)}} \tag{1.16}$$

$$Q_\alpha^{FR} = \frac{\Gamma_\alpha}{1 + e^{\varkappa_\alpha(\alpha^{FR}(i,j) - \sigma_\alpha)}} \tag{1.17}$$

The parameters $\Gamma_g, \varkappa_g, \sigma_g$ and $\Gamma_\alpha, \varkappa_\alpha, \sigma_\alpha$ determine the shape of the sigmoid functions utilized to build the edge strength and orientation preservation values.

$$Q^{FR}(i,j) = Q_g^{FR}(i,j) + Q_\alpha^{FR}(i,j) \tag{1.18}$$

With $0 \le Q^{FR}(i,j) \le 1$, a zero value indicates that no edge information is preserved in the fused image $Q^{FR}(i,j) = 1$ that indicates the fused image is obtained with no loss of edge information w.r.t the source images. A similar process is performed to calculate the value of $Q^{FS}(i,j)$ for S image of size $M \times N$. A normalized weighted performance metric $Q^{RS/F}(i,j)$ of a given fusion process that operates on images R and S, and produces F is obtained by the given expression,

$$Q^{RS/F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} Q^{FR}(i,j)w^R(i,j) + Q^{FS}(i,j)w^S(i,j)}{\sum_{i=1}^{M}\sum_{j=1}^{N} w^R(i,j) + w^S(i,j)} \tag{1.19}$$

The edge preservation values, $Q^{FR}(i,j)$ and $Q^{FS}(i,j)$ are weighted by $w^R(i,j)$ and $w^S(i,j)$, respectively, and also note that $0 \le Q^{RS/F}(i,j) \le 1$. In general, edge preservation values which correspond to pixels with high edge strength should influence $Q^{RS/F}(i,j)$ value higher and closer to unity, i.e. less information loss.

1.5 Objectives of the Present Study

The main objective of the present study is to develop the fusion approaches for multimodal neurological medical images. It has been recognized and emphasized in the previous sections that the different medical imaging modalities are being extensively utilized to visualize the complementary diagnostic information. Therefore, the development of an effective fusion approach for multimodal imaging modalities has been targeted as a problem here under consideration. Another main problem is to develop the fusion approach for maximizing the relevant diagnostic information in the fused image without introducing any spatial and spectral distortion. To achieve such objectives, different approaches have been proposed by several researchers as summarized in the previous section. With these objectives, initially it has been necessary to analyze and identify a better approach among the existing fusion approaches. After identification of the suitable approach, the task has been to improve the performance of the identified approach, either by modifying the earlier algorithm or by suggesting a new approach. Thus, based on the above observation, the following objectives have been identified for this study:

1. Design, development and implementation of three different anatomical (CT and MR) image fusion approaches.

   In order to achieve this objective, different fusion methods are proposed in the present study that also leads to further different sub-objectives. For such purposes, efforts have been made to improve the visual quality of their fused images and finally targeted to make the superior performance of the image fusion approach in terms of quantitative measures so that the individual fusion approach will be able to present all clinically relevant diagnostic information present in the source CT and MR images with higher contrast and without introducing any structural distortion.
2. Design, development and implementation of the anatomical and functional (SPECT /PET) image fusion approaches.

   To fulfill this objective, four different multimodal image fusion approaches are proposed in the present study, which are able to reflect the anatomical details produced by the MR/CT images without disturbing any functional status of the tissue reflected by the SPECT/PET images.

1.6 Organization of the Thesis

The research work carried out here in this thesis is reported for the fusion of multimodal neurological images in which the different imaging modalities such as CT, MR, SPECT and PET images are considered to analyze the performance of the fusion approaches. This thesis comprises of seven chapters. The present Chapter 1 introduces the topic and states the objectives. The rest of the thesis is organized as follows:

Chapter 2 provides the details of developed three image fusion approaches in ripplet and nonsubsampled shearlet transform (NSST) domain using different feature motivated pulse coupled neural network (PCNN) model. The performance of fusion approaches is evaluated for two different anatomical images qualitatively and quantitatively. A comparative analysis of the fusion results obtained by the proposed and other existing approaches is also presented at the end of the chapter, individually.

Chapter 3 presents an improved fusion framework developed for two different anatomical images and fusion of functional image with anatomical images using both the features of NSST and PCNN with an adaptive linking strength parameter based on the local visibility of the image. The fusion results and behavior of the proposed approach are explained at the end of this chapter.

Chapter 4 presents a hybrid multimodal medical image fusion developed in NSST and stationary wavelet transform (SWT) domain using an adaptive PCNN model and local log Gabor energy based fusion rules. The quantitative and qualitative analysis of the proposed approach is carried out for the fusion of functional and anatomical images and two different anatomical images, individually. Based on quantitative analysis, advantages and limitations of the proposed approach are mentioned at the end.

Chapter 5 presents an effective multimodal image fusion approach developed using sparse K-SVD dictionary learning and guided filtering in NSST domain. The proposed approach start with the NSST decomposition and then the different fusion rules based on sparse representation and guided filtering are applied on the NSST decomposed subbands. To provide in depth insight, a comparison is also made between the fusion results

Chapter 6 introduces another multimodal medical image fusion approach developed using two-scale hybrid layer decomposition with a convolutional neural network with consistency verification based feature mapping and structural patch clustering. At the end of this chapter, a comparative analysis of the fusion performance obtained by the proposed and other existing methods is presented to fuse both the anatomical with anatomical and anatomical with functional images.

Chapter 7 concludes the work stating the important features and advantages of the proposed methods. The outlook for future work is also given at the end of this chapter.

भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 2: NONSUBSAMPLED SHEARLET AND RIPPLET DOMAIN FUSION FOR CT-MR MEDICAL IMAGES

This chapter presents three different approaches for fusing two different anatomical images as CT and MR imaging in nonsubsampled shearlet and ripplet domain which are used to decompose the source CT and MR images. The methodologies used to present the proposed medical image fusion methods are discussed in this chapter. Different fusion rules are employed on low and high frequency decomposed coefficients. The performance of the proposed fusion approaches is also analyzed for CT and MR neurological images, qualitatively and quantitatively. A detailed comparative analysis has also been done and presented in the last section of this chapter to evaluate the performance of image fusion methods.

2.1 Introduction

In the recent years, clinical diagnosis is mainly focused on analyzing the digital images which are the major assets in the struggle against the critical diseases like Cancer, Hemorrhage and Alzheimer’s and so many others. This is the main reason for attracting towards the fusion of multimodal images because there are several medical imaging modalities, giving a different insight of the human body. So, the main objective of a fusion process is to capture the most relevant diagnostic information from the source images into a single image that has a significant clinical interpretation and suitable for the effective diagnostic analysis. Moreover, It has been analyzed and emphasized in the state-of-the-art approaches presented in Chapter 1 that the fused CT and MR images must preserve hard tissue information present in the CT and soft tissue details present in MR image alongwith the same contrast as that of the reference CT and MR medical images. As discussed in Chapter 1, several researchers have been approached towards the multiresolution transforms which present the image information in a similar way as human perceives the image information [4, 5, 25, 135, 144, 195, 198, 211]. There is also a possibility to analyze the different features of an image correspond to the different resolutions and directions through multiresolution transformation techniques. As discussed in Chapter 1, there are several transformation techniques such as wavelet, ridgelet, curvelet, bandlet, contourlet transform, etc. which can be used for medical image fusion applications. Wavelet transform (WT) is capable to capture only one dimensional singularity which means that it has only limited directional information, thus it causes artifacts along the edges that may contain the important diagnostic information. To overcome the restriction of the WT, ridgelet transform is introduced to extract the edges [40], but it did not do well for capturing the curve and higher order edge information. Donoho et al. have introduced curvelet transform (CVT) to capture the two dimensional singularities of any arbitrary curve [156]. However, the CVT cannot be built directly in discrete domain [121] and it generally leads to shift variance and pseudo-Gibbs oscillations. Thus, some visual distortions are reflected in the resultant images. Jun et al. [187] introduced ripplet transform (RT) type I which generalizes the CVT by adding two new parameters for providing the anisotropy capability that assured to present the singularities along an arbitrary shaped curves. The RT having different features as anisotropy, localization, directionality, multiscale and multiresolution is also able to overcome the limitations of other transformation approaches by providing the sparse representation of an image object. Another decomposition technique called nonsubsampled shearlet transform (NSST) is also introduced to overcome the limitation of the downsampling used in the forward and inverse transformation. The NSST also overcomes the problem of the contourlet transform (CNT) and nonsubsampled contourlet transform (NSCT) by providing a sufficient number of the directional coefficients of the reference images. This is the key motivation to consider the discrete ripplet transform (DRT) and NSST decomposition techniques in the proposed medical image fusion approach that can capture the complex image contours with unrestricted directional details at each scale in order to optimize the reconstruction efficiency of the method.

The subsequent part of this chapter is structured as follows. Next three sections provide the brief idea of NSST, DRT, its decomposition structures and pulse coupled neural network (PCNN) used to implement all three proposed approaches. After that, the proposed NSST based, DRT based and cascaded framework of the NSST and DRT based medical fusion approaches and their implementation are discussed. Next section illustrates the evaluation and analysis of the fusion performance of all three proposed methods. Many efforts have been made to evaluate and compare their fusion results with other existing techniques.

2.2 Non-Subsampled Shearlet Transform

The NSST is an extension of the WT in multidimensional and multidirectional case that combines the multiscale and direction analysis, separately. Firstly, the nonsubsampled laplacian pyramid filter (NSLP) is used to decompose an image into low and high-frequency components, and then direction filtering is employed to get the different subbands and different direction shearlet coefficients. The direction filtering is achieved using the shear matrix, which provides various directions. The brief discussion of the NSST is introduced as follows:

Consider a 2-D affine system with composite dilations as [42, 54],

$$A_{DS} = \left\{\psi_{x,y,z}(x) = |det\,D|^{x/2}\psi(S^y D^x r - z) \colon x, y \in \mathbb{Z}, z \in \mathbb{Z}^2\right\} \qquad (2.1)$$

where D and S represent the anisotropic dilation and shear matrix with scale (x), direction (y) and shift (z) parameter, respectively. The anisotropic dilation matrix $\begin{bmatrix} \mathrm{d}\ (\mathrm{d}^{\frac{1}{2}}) & 0\ (0) \\ 0(0) & \mathrm{d}^{\frac{1}{2}}(d) \end{bmatrix}$, where $\mathrm{d} > 0$ controls the scaling factor. The shear matrix $S = \begin{bmatrix} 1\ (1) & \mathrm{s}(0) \\ 0(\mathrm{s}) & 1(1) \end{bmatrix}$

controls the number of directions. In the presented work, $d = 4$ and $s = 1$ are considered as anisotropic and shear matrix parameters, respectively. Hence, the NSST transform function is introduced as follows,

$$\psi_{x,y,z}^{(0)}(x) = 2^{x\frac{3}{2}}\psi^{(0)}\left(S_0^y D_0^x x - z\right) \text{ and } \psi_{x,y,z}^{(1)}(x) = 2^{x\frac{3}{2}}\psi^{(0)}\left(S_1^y D_1^x r - z\right), \qquad (2.2)$$

where $x \geq 0, -2^x \leq y \leq 2^x - 1, z \in \mathbb{Z}^2$, $\hat{\psi}^{(0)}(\xi) = \hat{\psi}^{(0)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\hat{\psi}_2(\xi_2/\xi_1)$ and $\hat{\psi}^{(1)}(\xi) = \hat{\psi}^{(1)}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_2)\hat{\psi}_2(\xi_1/\xi_2)$.

As mentioned above, the NSLP and shearing filters (ShF) are utilized to provide the multiscale and multidirectional decomposition. At each NSLP decomposition level, one high frequency and one low frequency sub images are produced and further the low frequency subband is decomposed, iteratively. At the decomposition level $m = 3$, an image is decomposed into $m + 1 = 4$ subbands with the same size of the source image in which one subband image is the low frequency component and other m images are the high frequency subband images. Shearing filter is used in higher frequency subimages decomposition without sub-sampling which satisfies the shift invariance property. Using the ShF at k levels, the high frequency subband images obtained from the NSLP at each decomposition level, 2^k directional subband image coefficients are produced with the same size as of the source images. Three level decomposition of the NSST is shown in Figure 2.1 which illustrates the NLSP and its corresponding directional decompositions. In the presented work, the number of shearing directions is taken to be 8, 8 and 4 from finer to coarser scale using three levels NSST decomposition of an image. The tiling and the frequency supports of the NSST are presented in the Figure 2.2 (a) and (b), respectively.

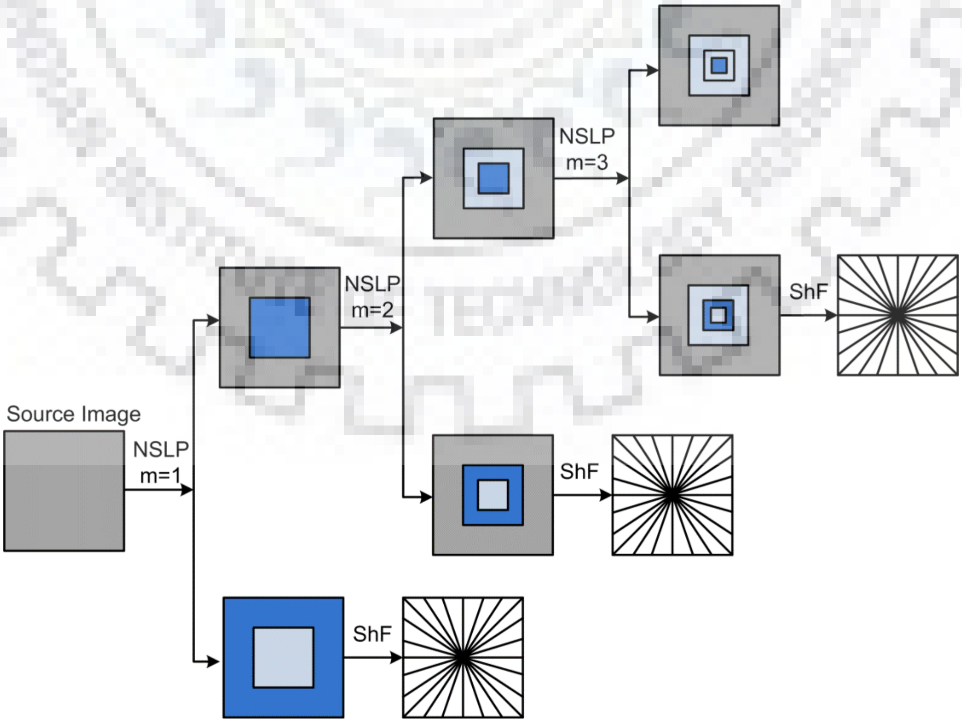


Figure 2.1 Three level multiscale and multidirectional decomposition of the NSST

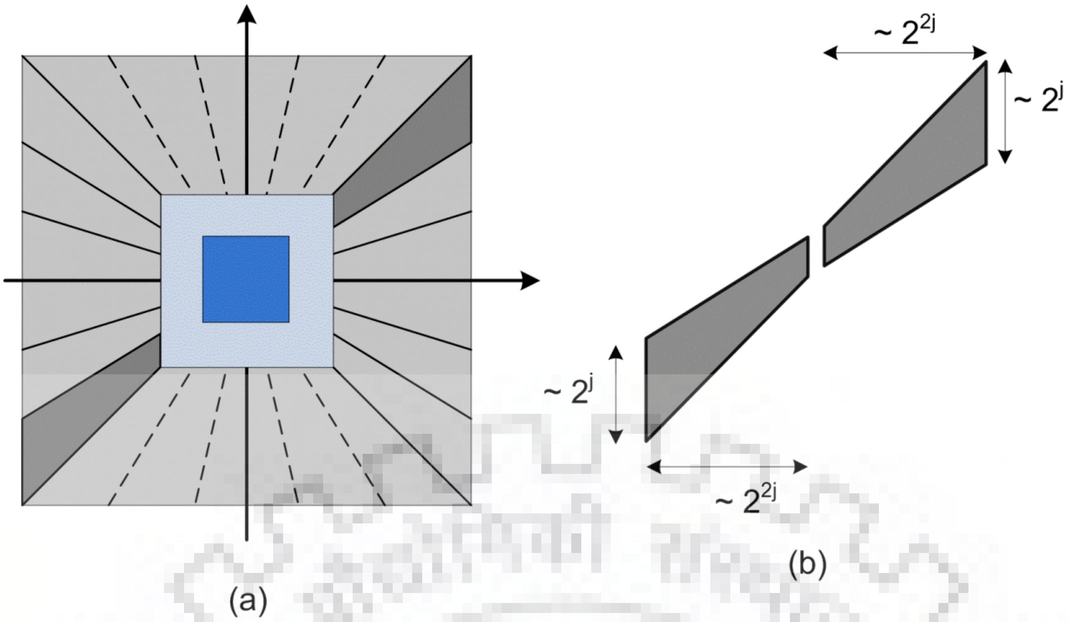


Figure 2.2 (a) The tiling of the frequency plane induced by the NSST (b) The size of the frequency support of the NSST

An example of three level NSST decomposition of a zone plate image is also shown in Figure 2.3.

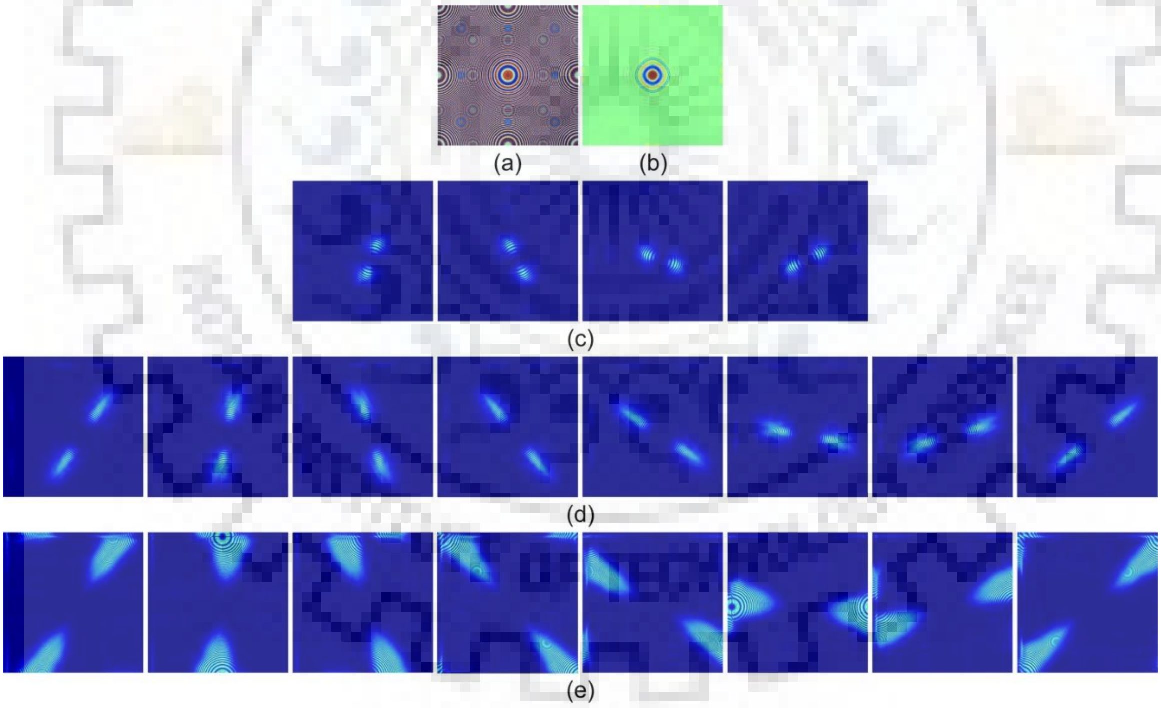


Figure 2.3 NSST decomposition of (a) Original zoneplate image (b) Approximate NSST component. The detail NSST components at (c) scale 3 (d) scale 2 (e) scale 1

2.3 Ripplet Transform

A higher dimensional framework called ripplet transform (RT) is able to characterize an image at various scales and directions.To realize anisotropic directionality, the CVT uses a parabolic scaling law as mentioned in [156]. From this perspective, the anisotropic properties of the CVT guarantees to resolve the two dimensional singularities along C^2 curves [156] . On the other side, the RT also provides a new tight frame with sparse representation for the source images with discontinuities along C^d curves [187]. If $d = 1$, then ripplet does not show the anisotropy behavior. For $d = 2$, it has parabolic scaling same as the curvelets and for $d = 3$, ripplet has the cubic scaling and so forth. So, it is assumed that the RT is a generalized form of curvelet transform by the inclusion of two new parameters i.e. support (c) and degree (d). The anisotropic capabilities of the RT are capable to efficiently represent the singularities along the random curve shape due to the addition of these two new parameters c and d.

2.3.1 Continuous ripplet transform

The continuous RT is defined as inner product of two dimensional integrable function $s(\vec{x})$ and ripplets $p_{a,\vec{b},\theta}(\vec{x})$ as follows [187]:

$$R(a,\vec{b},\theta) = \langle s, p_{a,\vec{b},\theta}\rangle = \int s(\vec{x})\,\overline{p_{a,\vec{b},\theta}(\vec{x})}\,d\vec{x} \tag{2.3}$$

where $R(a,\vec{b},\theta)$ is the ripplet coefficients and $\overline{(\cdot)}$ shows the conjugate operation. The ripplet function is defined as

$$p_{a,\vec{b},\theta}(\vec{x}) = p_{a,\vec{0},0}\left(R_\theta(\vec{x}-\vec{b})\right) \tag{2.4}$$

where $p_{a,\vec{0},0}(\vec{x})$ is a ripplet element function, $R_\theta = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix}$ is a rotation matrix, which rotates θ radians. $\vec{x}$ and $\vec{b}$ are two dimensional vectors.

The $p_{a,\vec{0},0}(\vec{x})$ can also be represented in the frequency domain as follows:

$$\hat{p}_a(r,\omega) = \frac{1}{\sqrt{c}}\,a^{\frac{1+d}{2d}}\,W(a\cdot r)\,V\left(\frac{a^{1/d}}{c\cdot a}\cdot\omega\right) \tag{2.5}$$

where $\hat{p}_a(r,\omega)$ is the Fourier transform of the ripplet element function $p_{a,\vec{0},0}(\vec{x})$ in polar co-ordinate. a, $\vec{b}$ and θ are the scale, position and rotation parameter, respectively. $W(r)$ and $V(\omega)$ represent the radial and angular window, respectively, having compact supports on [1/2, 2] and [-1, 1], respectively, that satisfy the two admissibility conditions as follows:

$$\int_{1/2}^{2} W^2(r)\,\frac{dr}{r} = 1 \text{ and } \int_{-1}^{1} V^2(t)\,dt = 1 \tag{2.6}$$

These two windows divide the polar frequency domain into wedges shown in Figure 2.4 (a). The main issue with the continuous ripplet transform is that it can only capture the behavior of the high frequency components of the original signal. So the full continuous ripplet transform is established by the combination of finer scale RT and coarse scale isotropic WT which represent the characteristic of high and low frequency components, respectively. The approximated image can be reproduced by the inverse of ripplet transform and expressed as given below [187]:

$$s(\vec{x}) = \int R(a,\vec{b},\theta)\,p_{a,\vec{b},\theta}(\vec{x})\,da\,d\vec{b}\,d\theta/a^3 \tag{2.7}$$

where R is a ripplet coefficient of the input signal, a, $\vec{b}$ and θ refer to scale, position and rotation parameter, respectively.

2.3.2 Discrete ripplet transform

In the field of digital image processing, discrete transforms are needed for their computerized algorithm implementation. So, discrete RT (DRT) is evaluated by discretizing the parameters of ripplets. The parameter a is sampled at dyadic intervals whereas $\vec{b}$ and θ are sampled at equally spaced intervals. The frequency response of ripplet function is given by,

$$\hat{p}_x(r,\omega) = \frac{1}{\sqrt{c}} a^{\frac{1+d}{2d}} W(2^{-x} \cdot r) V\left(\frac{2^{-x(1/d-1)}}{c} \cdot \omega - z\right) \tag{2.8}$$

where W and V satisfy the following conditions and $a_x = 2^{-x}$, $\vec{b}_y = [c \cdot 2^{-x} \cdot y_{1,} 2^{-x/d} \cdot y_{2,}]^T$ *and* $\theta_z = (2\pi/c) \cdot (2^{-[x(1-1/d)]}) \cdot z$, where $\vec{y} = [y_1\, y_2\,]^T$. $(\cdot)^{\mathrm{T}}$ denotes the transpose of a vector and $x, y_{1,} y_{2,} z \in \mathbb{Z}$.

$$\textstyle\sum_{x=0}^{\infty} |W(2^{-x} \cdot r)|^2 = 1 \text{ and } \sum_{z=-\infty}^{\infty} \left| V\left(\frac{2^{-x(1/d-1)}}{c} \cdot \omega - z\right)\right|^2 = 1 \tag{2.9}$$

For a fixed value of c, degree d is used to control the resolution in the directions at each high pass band. For given a fixed value of d, parameter c controls the number of directions at all high pass bands. For a particular combination of the parameters c and d is used to find the total number of the directions at each subband together. The decomposition of a 2-D image $r(i,j)$ with size $n \times m$ done by the DRT is expressed in terms of the DRT coefficients $R_{x,\vec{y},z}$

$$R_{x,\vec{y},z} = \textstyle\sum_{i=0}^{n-1} \sum_{j=0}^{m-1} r(i,j) \overline{p_{x,\vec{y},z}(i,j)} \tag{2.10}$$

After applying the inverse DRT, an image $\hat{r}(x,y)$ can be approximated as,

$$\hat{r}(i,j) = \textstyle\sum_x \sum_{\vec{y}} \sum_z R_{x,\vec{y},z} p_{x,\vec{y},z}(i,j) \tag{2.11}$$

Figure 2.4 (b) shows a real MR image and decomposition of the image processed using the DRT is illustrated in Figure 2.4 (c).

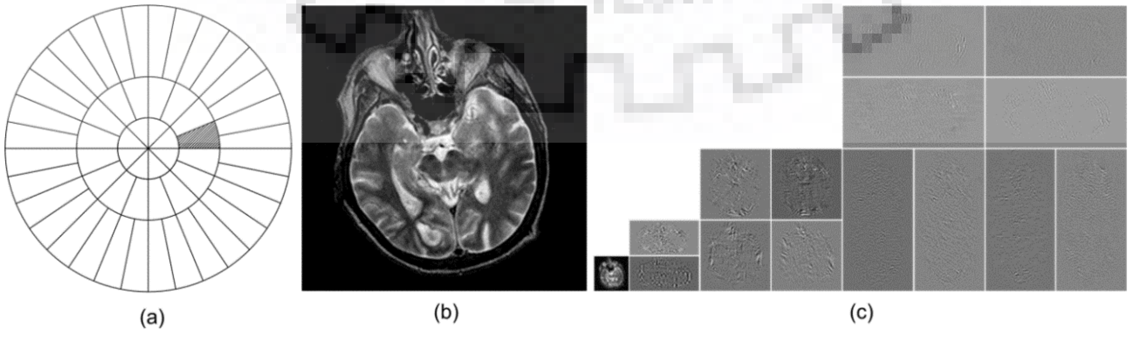


Figure 2.4 (a) The tiling of the polar frequency domain. The dashed wedge corresponds to the frequency transform of the element function. (b) Source MR image (c) Subbands after DRT decomposition of source image with support ($c = 1$) and degree ($d = 4$)

2.4 Pulse Coupled Neural Network

Pulse coupled neural network is a biologically inspired neural network [2, 44, 108, 188] which is efficiently utilized in several image processing applications [127]. The PCNN is a single layer 2-D array of laterally linked pulse coupled neurons with one to one correspondence between the pixels and neurons. Each pixel is connected to a particular neuron and this neuron is further connected to its neighbouring neurons within the linking range. Figure 2.5 shows a PCNN model that consists of three parts such as receptive field, linking modulation field and pulse generator. The receptive field receives signal from other neurons and source input. Furthermore, this signal is divided into linking $L_{i,j}$ and feeding $F_{i,j}$ inputs. The mathematical formulation of the PCNN model is written as,

$$F_{i,j}[n] = S_{i,j} + e^{-\alpha_F}F_{i,j}[n-1] + V_F\sum_{k,l} M_{i,j,k,l}Y_{i,j}[n-1] \tag{2.12}$$

$$L_{i,j}[n] = e^{-\alpha_L}L_{i,j}[n-1] + V_L\sum_{k,l} W_{i,j,k,l}Y_{i,j}[n-1] \tag{2.13}$$

$$U_{i,j}[n] = F_{i,j}[n]\left(1 + \beta L_{i,j}[n]\right) \tag{2.14}$$

$$T_{i,j}[n] = e^{-\alpha_T}T_{i,j}[n-1] + V_T Y_{i,j}[n-1] \tag{2.15}$$

$$Y_{i,j}[n] = \begin{cases} 1, & U_{i,j} > T_{i,j} \\ 0, & otherwise \end{cases} \tag{2.16}$$

where i and j represents the pixel location, k and l refer to the dislocation in a symmetric neighbouring pixel, $n = \{1,2,3,\ldots,n_{max}\}$ refers to the current iteration and n_{max} is the total number of iterations. The PCNN model consist of a feeding and linking field acquired six parameters such as three decay constants $(\alpha_F, \alpha_L, \alpha_T)$ and three normalizing constants (V_F, V_L, V_T) for feeding $(F_{i,j})$, linking $(L_{i,j})$ and threshold $(T_{i,j})$ inputs, respectively. In Eqs. (2.14), (2.15) and (2.16), $U_{i,j}$ is the internal activity of neuron and $T_{i,j}$ is the dynamic threshold and $Y_{i,j}$ refers to the pulse output of neurons, respectively.

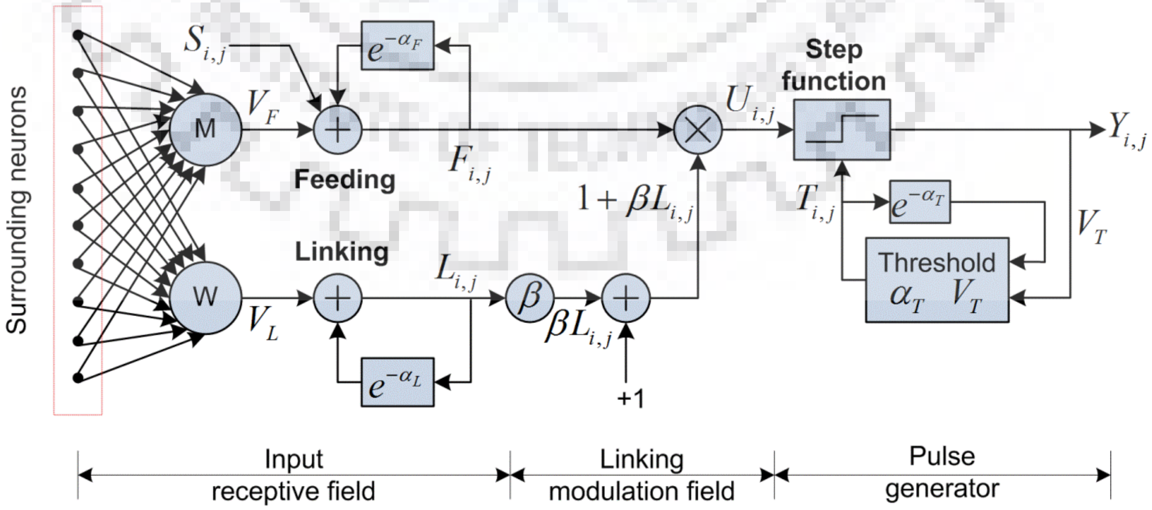


Figure 2.5 Pulse coupled neural network

In the presented work, an improved model [34, 147, 203] is used and its mathematical formulation is given as,

$$\left.\begin{aligned}
&F_{i,j}[n] = S_{i,j} \\
&L_{i,j}[n] = e^{-\alpha_L} L_{i,j}[n-1] + V_L \textstyle\sum_{k,l} W_{i,j,m,n} Y_{k,l}[n-1] \\
&U_{i,j}[n] = F_{i,j}[n]\left(1 + \beta L_{i,j}[n]\right) \\
&T_{i,j}[n] = e^{-\alpha_T} T_{i,j}[n-1] + V_T Y_{i,j}[n-1] \\
&Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n] \\ 0, & otherwise \end{cases}
\end{aligned}\right\} \tag{2.17}$$

2.5 Proposed NSST Domain CT-MR Image Fusion Approach

For fusing the CT and MR neurological images and implementing the above aspects, the proposed NSST domain fusion approach using regional energy and novel sum modified Laplacian (NSML) are formulated in the following steps,

1. Decomposition of Source CT and MR images using the NSST.
2. Fusion of low frequency NSST subband coefficients.
3. Fusion of High Frequency NSST subbands coefficients.
4. Estimate the reconstructed fused image by taking the inverse NSST of these NSST subband coefficients.

The process flow of the proposed NSST domain multimodal image fusion approach (NSST-MIF) is shown in Figure 2.6. The NSST transform and the PCNN model are employed to present the proposed fusion method that incorporates the regional energy to fuse the low frequency coefficients. High frequency coefficients are also fused using a pulse coupled neural network model that is used as a biologically inspired feedback neural model. It also retains the edges and detail information from the source images. The novel sum modified Laplacian (NSML) in the NSST domain is given as an input to motivate the neuron of PCNN model for high frequency coefficients fusion. The main motivation to consider the NSST over the other transformation technqiues is because of its important properties as discussed in [43], a) inherently sparse approximation model of the input, b) well localized (trapezoid shape tilling in frequency space with increased directional coefficients by 2^{j+1} factor on each scale), c) less Pseudo-Gibbs type residual error in approximation, d) computationally more efficient (inverse transform needs only a summation of shearing filter coefficients instead the inversion of the directional filter bank as required in NSCT [30]).

Implementation steps

Step 1: Start with the given CT and MR images $X = X(i,j)$ and $Y(i,j)$ respectively.

Step 2: Decompose the original images using NSST into one low frequency (lf) subband and a series of high frequency (hf) coefficients at level $m = 3$ and direction $k = [2,3,3]$ based on repeated experiments from coarser to finer scale of decomposition.

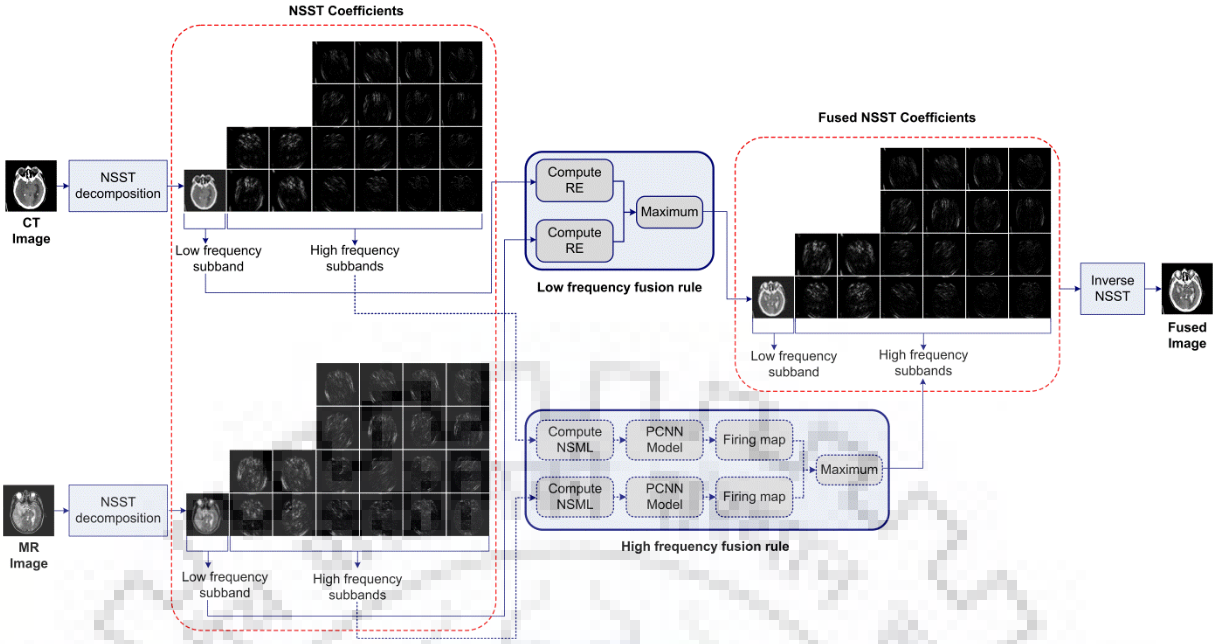


Figure 2.6 Process flow of the proposed NSST domain fusion (NSST-MIF) approach

$$[lf_X^{NSST}, hf_X^{NSST}] = NSST(X) \text{ and } [lf_Y^{NSST}, hf_Y^{NSST}] = NSST(Y) \tag{2.18}$$

Step 3: For low frequency subband fusion, an activity level measurement such as the regional energy is used because of the lf subband contains most of the signal energy and detail information of the original image. The regional energy is evaluated as,

$$G_Z(i,j) = \sum_{i=1}^{a} \sum_{j=1}^{b} lf_Z^{NSST}(a+i, b+j)^2\, w(i,j) \tag{2.19}$$

where Z denotes the original images either X or Y and $w(i,j) = \frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$ denotes a 3×3 template. $a \times b$ is the size of the template.

Step 4: After evaluating the regional energy of both the lf subband of the CT and MR images, a maximum selection rule is applied in choosing the coefficients that have the highest activity measure. It is given as follows:

$$lf_F^{NSST} = \begin{cases} lf_X^{NSST}; & if\ |G_X(i,j)| \geq |G_Y(i,j)| \\ lf_Y^{NSST}; & if\ |G_X(i,j)| < |G_Y(i,j)| \end{cases} \tag{2.20}$$

where lf_F^{NSST} denotes the final fused low frequency subband and F refers to the fused image.

Step 5: For the high frequency subbands, compute the NSML as follows:

$$NSML(i,j) = \sum_a \sum_b w(i,j).F(i+a, j+b) \tag{2.21}$$

and

$$F(i,j) = \begin{bmatrix} \left|2hf_Z^{NSST}(i,j) - hf_Z^{NSST}(i-1,j) - hf_Z^{NSST}(i+1,j)\right| \\ + \left|2hf_Z^{NSST}(i,j) - hf_Z^{NSST}(i,j-1) - hf_Z^{NSST}(i,j+1)\right| \end{bmatrix} \qquad (2.22)$$

where w is the normalized window and it is defined as,

$$w(i,j) = \begin{bmatrix} 1/15 & 2/15 & 1/15 \\ 2/15 & 3/15 & 2/15 \\ 1/15 & 2/15 & 1/15 \end{bmatrix} \qquad (2.23)$$

Step 6: Apply the NSML of each high frequency components as an input to activate the PCNN and generate the pulse of neurons using the following mathematical equations,

$$\left.\begin{aligned} &F_{i,j}^{Z}[n] = NSML_{i,j}^{Z} \\ &L_{i,j}^{Z}[n] = e^{-\alpha_L} L_{i,j}^{Z}[n-1] + V_L \textstyle\sum_{k,l} W_{i,j,k,l}^{Z} Y_{k,l}^{Z}[n-1] \\ &U_{i,j}^{Z}[n] = F_{i,j}^{Z}[n]\left(1 + \beta_{i,j} L_{i,j}^{Z}[n]\right) \\ &T_{i,j}^{Z}[n] = e^{-\alpha_T} T_{i,j}^{Z}[n-1] + V_T Y_{i,j}^{Z}[n-1] \\ &Y_{i,j}^{Z}[n] = \begin{cases} 1, & U_{i,j}^{Z} > T_{i,j}^{Z} \\ 0, & otherwise \end{cases} \end{aligned}\right\} \qquad (2.24)$$

Step 7: Evaluate the firing times (sum of the $Y_{i,j}^{Z} = 1$ for $U_{i,j}^{Z} > T_{i,j}^{Z}$) in n iteration, as follows,

$$t_{i,j}^{Z}[n] = t_{i,j}^{Z}[n-1] + Y_{i,j}^{Z}[n] \qquad (2.25)$$

Step 8: if $n = n_{max}$, then iteration stops and the high frequency subimage coefficients are fused as the following rule based on their firing time as evaluated in step 7.

$$hf_F^{NSST} = \begin{cases} hf_X^{NSST}; & if\ t_{i,j}^{X}[n_{max}] \geq t_{i,j}^{Y}[n_{max}] \\ hf_Y^{NSST}; & if\ t_{i,j}^{X}[n_{max}] < t_{i,j}^{Y}[n_{max}] \end{cases} \qquad (2.26)$$

Step 9: Perform the inverse NSST ($NSST^{-1}$) on the fused LF and HF subband coefficients to reconstruct the final fused image (F).

$$F = NSST^{-1}\left(lf_F^{NSST}, hf_F^{NSST}\right) \qquad (2.27)$$

2.6 Proposed DRT Domain CT-MR Image Fusion Approach

This section presents DRT domain medical image fusion approach based on PCNN model named as DRT-MIF approach. As mentioned above, after DRT decomposition of the source CT/MR images, one low frequency and a series of high-frequency subband coefficients are obtained. A low-frequency component conveys the useful diagnostic information of the source CT/MR image, whereas high-frequency subbands depict the details by varying its directions and scales based on the DRT parameters. In addition to the fusion of low and high-frequency coefficients, NSML and novel sum modified spatial frequency (NMSF) are computed and applied to motivate the neural model as an external input. The novel sum modified spatial frequency is computed to include the directional contents that reflect both the clarity and activity level. The NSML is utilized as an activity level measurement parameter for subband coefficients which reflect most of the informative content present in both the reference images. It is also able to represent the information related to the contours and boundaries of multiple objects present in the reference images. Moreover, the salient implementation steps involved in the process flow of proposed fusion approach are shown in Figure 2.7.

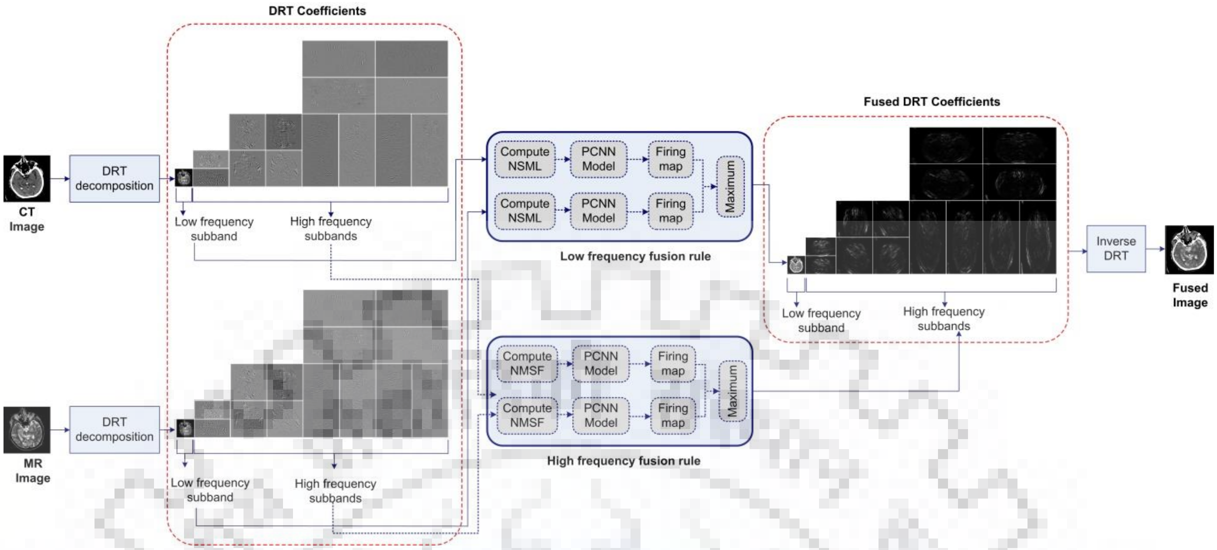


Figure 2.7 Process flow of the proposed DRT domain fusion approach (DRT-MIF)

Let $\boldsymbol{X} = X(i,j)$ and $\boldsymbol{Y} = Y(i,j)$ be two source images acquired from different medical imaging modality such as the CT and MR, respectively, and the implementation of the above aspects, the proposed fusion algorithm is formulated as follows,

Implementation steps

Step 1: Start with the decomposition of reference (CT and MR) images using the DRT (with the parameter $c = 1$ and $d = 4$) into one lf and several hf coefficients based on the successive experiments from coarser to the finer scale of decomposition.

$$\left[lf_X^{DRT}, hf_X^{DRT}\right] = DRT\left(X_{i,j}\right) \text{ and } \left[lf_Y^{DRT}, hf_Y^{DRT}\right] = DRT\left(Y_{i,j}\right) \qquad (2.28)$$

Step 2: For the lf coefficients fusion, firstly compute the NSML as mentioned in Eqs. (2.21) to (2.23).

Step 3: For hf coefficients fusion, firstly compute the NMSF that is utilized to express the clarity and activity level of the input source images in a template of size 3×3. This measure is also able to detect the edges present in the input reference images. The computed NMSF motivates the PCNN and generates the pulse of neurons. The mathematical expression for the NMSF is expressed as,

$$NMSF = \sqrt{\begin{array}{l}\frac{1}{M(N-1)}\left[\sum_{i=1}^{M}\sum_{j=2}^{N}\left(hf_Z^{DRT}(i,j-1) - hf_Z^{DRT}(i,j)\right)^2\right] \\ +\frac{1}{(M-1)N}\left[\sum_{i=2}^{M}\sum_{j=1}^{N}(hf_Z^{DRT}(i,j) - hf_Z^{DRT}(i-1,j))^2\right] + D_f\end{array}} \qquad (2.29)$$

where Z denotes the original image either X or Y and $w(m,n)$ denotes a 3×3 template considered in the present work. M and N refer to the image size and the third term D_f is an additional diagonal frequency in the neighborhood pixels which is added in the expression of the spatial frequency and it is expressed as,

$$D_f = \left[\begin{array}{c} \sqrt{\frac{1}{(M-1)(N-1)} \left[\sum_{i=2}^{M} \sum_{j=2}^{N} \left(hf_Z^{DRT}(i,j) - hf_Z^{DRT}(i-1,j-1) \right)^2 \right]} + \\ \sqrt{\frac{1}{(M-1)(N-1)} \left[\sum_{i=2}^{M} \sum_{j=2}^{N} (hf_Z^{DRT}(i-1,j) - hf_Z^{DRT}(i,j-1))^2 \right]} \end{array} \right]^2 \tag{2.30}$$

Step 4: Apply the NSML and NMSF computed for lf components and each individual hf coefficients, respectively, to activate the PCNN model and build up the pulse of neurons using the following equations,

For the lf coefficients fusion

$$F_{i,j}^{Z,P}[n] = NSML_{i,j}^{Z} \tag{2.31}$$

For the hf coefficients fusion

$$F_{i,j}^{Z,P}[n] = NMSF_{i,j}^{Z} \tag{2.32}$$

and

$$\left. \begin{aligned} & L_{i,j}^{Z,P}[n] = e^{-\alpha_L} L_{i,j}^{Z,P}[n-1] + V_L \sum_{k,l} W_{i,j,k,l}^{Z,P} Y_{i,j}^{Z,P}[n-1] \\ & U_{i,j}^{Z,P}[n] = F_{i,j}^{Z,P}[n] \left(1 + \beta L_{i,j}^{Z,P}[n] \right) \\ & T_{i,j}^{Z,P}[n] = e^{-\alpha_T} T_{i,j}^{Z,P}[n-1] + V_T Y_{i,j}^{Z,P}[n-1] \\ & Y_{i,j}^{Z,P}[n] = \begin{cases} 1, & U_{i,j}^{Z,P} > T_{i,j}^{Z,P} \\ 0, & otherwise \end{cases} \end{aligned} \right\} \tag{2.33}$$

where P refers to the term lf or hf and Z denotes the original images either X or Y.

Step 5: Evaluate the firing times (sum of the $Y_{i,j}^{Z,P} = 1$ for $U_{i,j}^{Z,P} > T_{i,j}^{Z,P}$) in n iteration for both the subbands fusion, as follows,

$$t_{i,j}^{Z,P}[n] = t_{i,j}^{Z,P}[n-1] + Y_{i,j}^{Z,P}[n] \tag{2.34}$$

Step 6: If $n = n_{max}$, process stops and the lf and hf coefficients are fused, based on the fusion rule given below

$$lf_F^{DRT} = \begin{cases} lf_X^{DRT}; & if\, t_{i,j}^{X,lf}[n_{max}] \geq t_{i,j}^{Y,lf}[n_{max}] \\ lf_Y^{DRT}; & if\, t_{i,j}^{X,lf}[n_{max}] < t_{i,j}^{Y,lf}[n_{max}] \end{cases} \tag{2.35}$$

$$hf_F^{DRT} = \begin{cases} hf_X^{DRT}; & if\, t_{i,j}^{X,lf}[n_{max}] \geq t_{i,j}^{Y,lf}[n_{max}] \\ hf_Y^{DRT}; & if\, t_{i,j}^{X,lf}[n_{max}] < t_{i,j}^{Y,lf}[n_{max}] \end{cases} \tag{2.36}$$

Step 7: Finally, reconstruct the fused image by applying the inverse DRT.

$$F = DRT^{-1}(lf_F^{DRT}, hf_F^{DRT}) \tag{2.37}$$

2.7 Proposed CT-MR Image Fusion Approach based on a Cascaded Framework in DRT and NSST Domain

In this proposed fusion approach, a cascaded framework is presented for multimodal medical information in DRT and NSST domain. The DRT and NSST having different features are utilized in a cascade manner that provides several directional decomposition coefficients and increases shift invariance information in the fused images. Such cascade fusion approach is framed based on the DRT and the PCNN model at stage 1 in which NSML and NMSF are applied as inputs to the PCNN model for low and high frequency subimage coefficients in the DRT domain, respectively. The NSST decomposition is used at second stage where regional energy is computed to fuse the low frequency NSST approximation coefficients and for high frequency NSST detail coefficients, the sum of absolute difference (SAD) and absolute maximum (AM) based fused model is proposed and used to provide the richer representation of the edge detail information with the improved contrast, respectively. The cascaded fusion framework can preserve more details in the reference images and further enhance the visualization of the fused images. Moreover, the salient implementation steps involved in the process flow of the proposed cascaded image fusion approach (C-MIF) is shown in Figure 2.8.

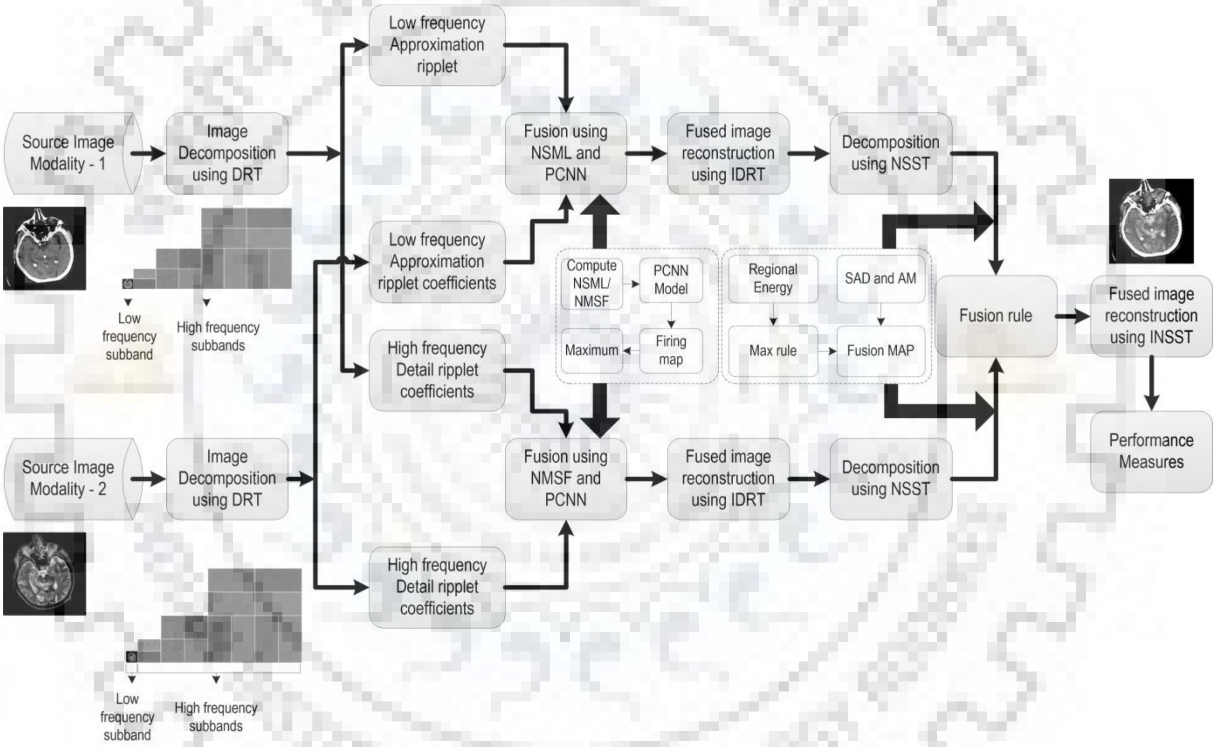


Figure 2.8 Process flow of the proposed cascaded framework in DRT and NSST domain for CT-MR medical images

Let $\boldsymbol{r} = r(i,j)$ and $\boldsymbol{s} = s(i,j)$ be two source images acquired from different medical imaging modality such as CT and MR, respectively, and the implementation of the above aspects, the proposed fusion algorithm is framed as given below,

Implementation steps

Step 1: Start with the decomposition of the source (CT and MR) images using DRT (with the parameters $c = 1$ and $d = 4$) into one low frequency (lf) and a series of high frequency (hf) coefficients.

$$\left[lf_{(r)}^{DRT}, hf_{(r)}^{DRT}\right] = DRT(r) \text{ and } \left[lf_{(s)}^{DRT}, hf_{(s)}^{DRT}\right] = DRT(s) \tag{2.38}$$

Step 2: For fusing the lf coefficients, compute the NSML which reflects the edge details of low frequency subband image coefficients, as mentioned in Eqs. (2.21) to (2.23).

Step 3: For fusing the hf coefficients, compute the NMSF that is utilized to express the clarity and activity level of the input source images in a template of size 3×3 and also able to detect the gradient of the input source images, as mentioned in Eqs. (2.29) and (2.30).

Step 4: Apply the NSML and NMSF computed for lf components and each individual hf coefficients, respectively, to motivate the PCNN and build up the pulse of neurons using the mathematical relations as follows,

For lf coefficients fusion

$$F_{i,j}^{(\boldsymbol{z}),p}[n] = NSML_{i,j}^{(\boldsymbol{z})} \tag{2.39}$$

For hf coefficients fusion

$$F_{i,j}^{(\boldsymbol{z}),\boldsymbol{p}}[n] = NMSF_{i,j}^{(z)} \tag{2.40}$$

and

$$\left.\begin{aligned} L_{i,j}^{(\boldsymbol{z}),p}[n] &= e^{-\alpha_L} L_{i,j}^{(\boldsymbol{z}),p}[n-1] + V_L \textstyle\sum_{m,n} W_{i,j,m,n}^{(\boldsymbol{z}),p} Y_{i,j}^{(\boldsymbol{z}),p}[n-1] \\ U_{i,j}^{(\boldsymbol{z}),p}[n] &= F_{i,j}^{(\boldsymbol{z}),p}[n]\left(1 + \beta L_{i,j}^{(\boldsymbol{z}),p}[n]\right) \\ T_{i,j}^{(\boldsymbol{z}),p}[n] &= e^{-\alpha_T} T_{i,j}^{(\boldsymbol{z}),p}[n-1] + V_T Y_{i,j}^{(\boldsymbol{z}),p}[n-1] \\ Y_{i,j}^{(\boldsymbol{z}),p}[n] &= \begin{cases} 1, \ U_{i,j}^{(\boldsymbol{z}),p} > T_{i,j}^{(\boldsymbol{z}),p} \\ 0, \ otherwise \end{cases} \end{aligned}\right\} \tag{2.41}$$

where p refers to the term lf or hf and $\boldsymbol{z}$ denotes the original images either $r(i,j)$ or $s(i,j)$.

Step 5: Evaluate the firing times (sum of the $\mathrm{Y}_{\mathrm{i,j}}^{(\boldsymbol{z}),\mathrm{p}} = 1$ for $U_{i,j}^{(\boldsymbol{z}),p} > T_{i,j}^{(\boldsymbol{z}),p}$) in n iteration for both the subbands fusion, as follows,

$$t_{i,j}^{(\boldsymbol{z}),p}[n] = t_{i,j}^{(\boldsymbol{z}),p}[n-1] + Y_{i,j}^{(\boldsymbol{z}),p}[n] \tag{2.42}$$

Step 6: If $n = n_{max}$, then iteration stops and based on the firing times, the lf and hf subband coefficients (p) are fused using the fusion rules given below,

$$p_{(\boldsymbol{F})}^{DRT} = \begin{cases} p_{(\mathrm{r})}^{DRT}; & if t_{i,j}^{(\boldsymbol{r}),p}[n_{max}] \geq t_{i,j}^{(\boldsymbol{s}),p}[n_{max}] \\ p_{(\mathrm{s})}^{DRT}; & if t_{i,j}^{(\boldsymbol{r}),p}[n_{max}] < t_{i,j}^{(\boldsymbol{s}),p}[n_{max}] \end{cases} \tag{2.43}$$

Step 7: Apply the inverse DRT on subband coefficients calculated in step 6 and obtain the images $z_1 = r_1(i,j)$ and $s_1(i,j)$.

Step 8: Apply the NSST for decomposing the images obtained in step 7 and compute the two different NSST coefficients $lf^{NSST}(i,j)$ and $hf^{NSST}(i,j)$.

Step 9: Compute the regional energy for the corresponding $lf^{NSST}(i,j)$ using the mathematical formulation given as per Eq. (2.19),

$$C_{(\boldsymbol{z_1})}^{NSST} = RE\left(lf_{(\boldsymbol{z_1})}^{NSST}(i,j)\right) = \sum_{a=-1}^{1} \sum_{b=-1}^{1} w(a,b)\left(lf_{(\boldsymbol{z_1})}^{NSST}(i+x, j+y)\right)^2 \tag{2.44}$$

where $w(a,b) = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$ denotes a 3×3 template.

Step 10: After computing the values for regional energy, maximum fusion rule is applied for fusing the $lf^{NSST}(i,j)$ coefficients as follows

$$lf_{(F_1)}^{NSST} = \begin{cases} lf_{(\mathbf{r_1})}^{NSST} & \left|C_{(\mathbf{r_1})}^{NSST}\right| \geq \left|C_{(\mathbf{s_1})}^{NSST}\right| \\ lf_{(\mathbf{s_1})}^{NSST} & \left|C_{(\mathbf{r_1})}^{NSST}\right| < \left|C_{(\mathbf{s_1})}^{NSST}\right| \end{cases} \tag{2.45}$$

Step 11: For fusing the directional $hf^{NSST}(i,j)$ coefficients, sum of the absolute difference (SAD) and absolute maximum (AM) based fused rule is utilized with in a 3×3 template as follows

$$SAD_{(z_1)}^{NSST} = \begin{bmatrix} \sum_{a=-1}^{1} abs\left(hf_{(z_1)}^{NSST}(i+a,j) - hf_{(z_1)}^{NSST}(i+a,j+1)\right) \\ + \sum_{a=-1}^{1} abs\left(hf_{(z_1)}^{NSST}(i+a,j) - hf_{(z_1)}^{NSST}(i+a,j-1)\right) \\ + \sum_{a=-1}^{1} abs\left(hf_{(z_1)}^{NSST}(i,j+a) - hf_{(z_1)}^{NSST}(i+1,j+a)\right) \\ + \sum_{a=-1}^{1} abs\left(hf_{(z_1)}^{NSST}(i,j+a) - hf_{(z_1)}^{NSST}(i-1,j+a)\right) \end{bmatrix} \tag{2.46}$$

and

$$AM_{(z_1)}^{NSST} = max\left(\left|hf_{(z_1)}^{NSST}(i+x,j+y)\right|\right),\ -1 \leq x,y \leq 1 \tag{2.47}$$

Step 12: Based on the computed value of AM and SAD that represent the contrast between the edges and the dominant edges presented in the source images, a fusion rule is framed for fusing the $hf^{NSST}(i,j)$.

$$hf_{(F_1)}^{NSST} = \begin{cases} hf_{(r_1)}^{NSST}; & AM_{(r_1)}^{NSST} > AM_{(s_1)}^{NSST} \\ hf_{(s_1)}^{NSST}; & AM_{(r_1)}^{NSST} < AM_{(s_1)}^{NSST} \\ hf_{(r_1)}^{NSST}; & SAD_{(r_1)}^{NSST} > SAD_{(s_1)}^{NSST} \\ hf_{(s_1)}^{NSST}; & otherwise \end{cases} \tag{2.48}$$

Step 13: Finally, reconstruct the fused image by applying the inverse NSST.

2.8 Experimentation

Three different experiments are conducted on the source CT and MR images acquired from the open image source database the whole brain atlas (http://www.med.harvard.edu/AANLIB/home.html) to analyze the fusion performance of all three proposed approaches, individually. Moreover, a detailed comparative evaluation has been done for each one of the proposed fusion technique with the state-of-the–art methods. The performance measures mentioned in Chapter 1 are considered to assess the performance of the proposed fusion methods quantitatively. All the CT and MR image pairs of the dataset utilized in the presented work belong to the same patient. Several experiments are conducted to analyze the performance of the fusion methods on these pre-registered CT and MR images. To conduct the experiments, the NSST decomposition level = [2, 3, 3] and PCNN parameters were set as $\alpha_L = 0.3,\ \alpha_T = 0.1,\ \beta = 0.2,\ \ V_L = 1,\ V_T = 10,\ W = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ \ 1\right]$

and maximum number of iterations $n_{max} = 200$. All these parameters are decided based on the repeated experiments. Four different experiments are discussed as given below:

Experiment 1: Analysis and evaluation of the proposed NSST-MIF approach

Experiment 2: Analysis and evaluation of the proposed DRT-MIF approach

Experiment 3: Analysis and evaluation of the proposed C-MIF approach

Experiment 4: Comparative evaluation of the proposed approaches and others

2.9 Results and Discussions

2.9.1 Experiment 1: Analysis and evaluation of the proposed NSST-MIF approach

To assess the performance of the proposed NSST-MIF approach, different experiments were conducted on the given CT and MR neurological image dataset shown in Figure 2.9 (a) and (b) respectively. Fused images obtained by the proposed approach are also shown in Figure 2.9 (c_1) to (c_9). The proposed approach NSST-MIF formulated by the regional energy applied for fusing the lf coefficients and NSML motivated PCNN mapping for hf subband coefficients gives better quality fused images. It shows more information contents and edge details than the input CT and MR images. Visually, the fused images shown in Figure 2.9 have more information and edge details available in the reference CT and MR images. The quantitative analysis is also done for all the fused images The bar diagram shown in Figure 2.10 represents an illustrative idea about the parameter variation such as the En, STD and SF, evaluated by the NSST-MIF approach as compared to their original values of all nine data sets of CT and MR images. It is clearly observed from the results presented in Table 2.1 and Figure 2.10 that the fused images preserve more information by achieving higher En and MI values. The proposed NSST-MIF approach also gains higher values of the STD compared to the reference CT and MR images. Figure 2.10 (c) shows the higher SF values for the fused image obtained by the proposed NSST-MIF than all the input CT and MR images.

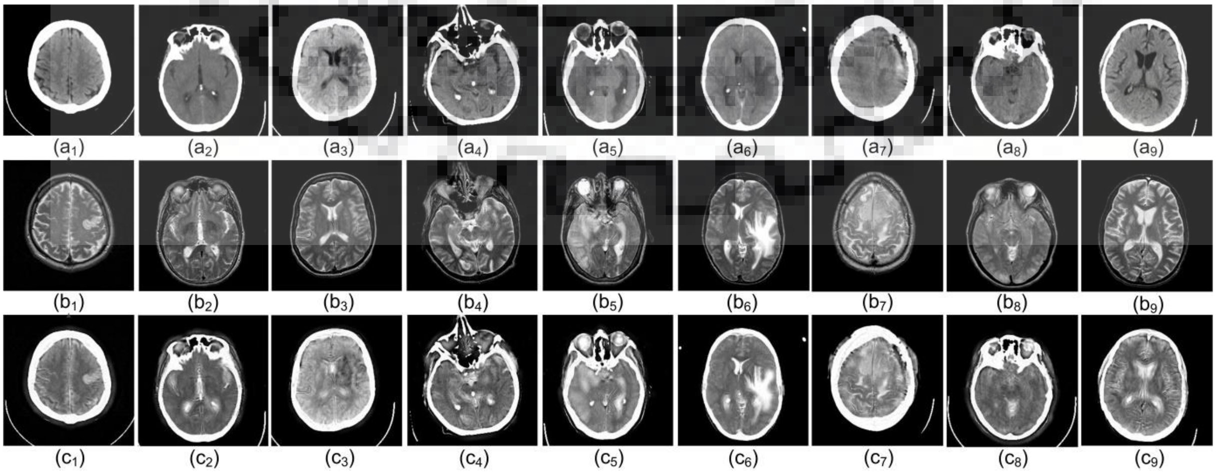
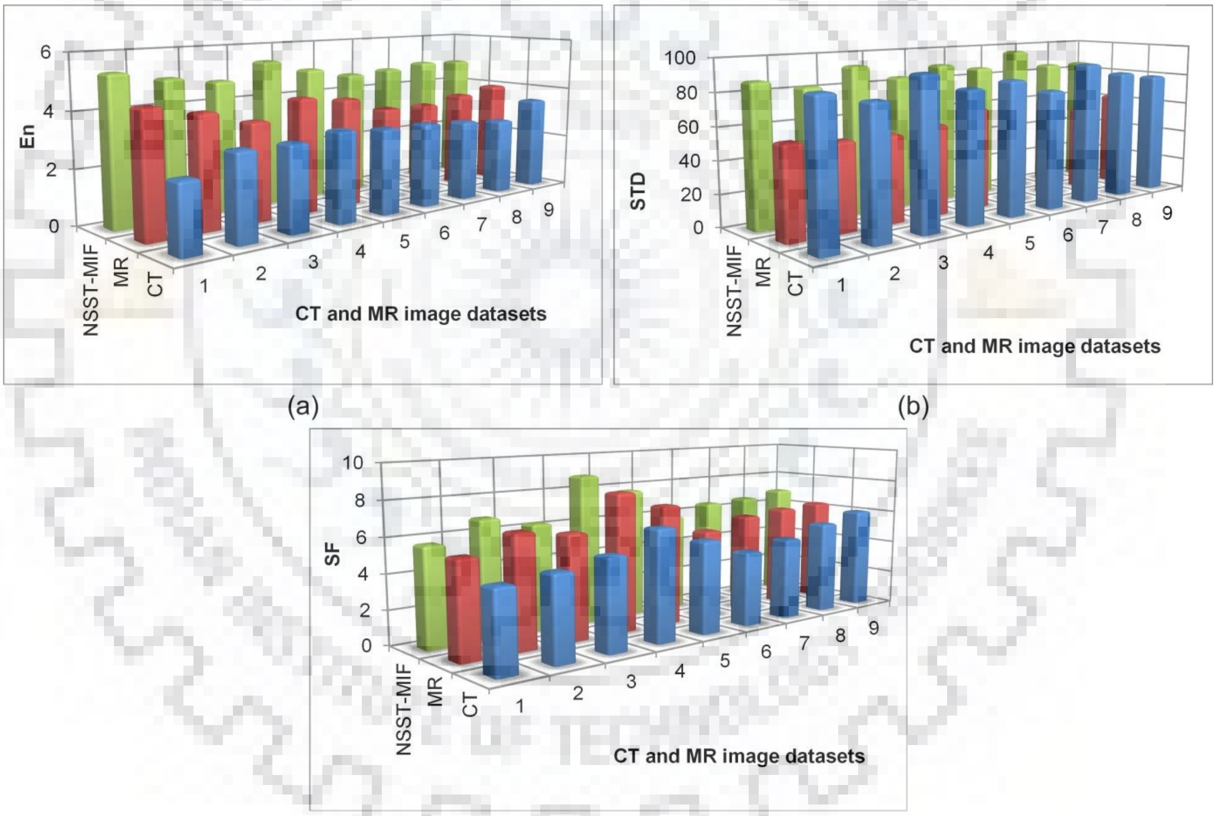


Figure 2.9 (a) Source CT images, (b) MR images (c) Fused images obtained by the proposed NSST-MIF approach

Table 2.1 Performance measures obtained by the proposed NSST-MIF approach for the fused images illustrated in Figure 2.9

Dataset	Performance Measures					
	En	MI	STD	SF	IQI	XEI
# 1	5.3430	3.2956	87.1860	5.7153	0.4839	0.4685
# 2	5.0335	3.0853	82.0347	6.8577	0.4627	0.5496
# 3	4.8078	3.0724	92.6860	6.2712	0.4833	0.5002
# 4	5.4243	2.8509	83.8141	8.8396	0.4649	0.5255
# 5	4.9927	2.9886	89.9084	7.7340	0.4602	0.5188
# 6	4.6794	3.2026	86.1525	5.8769	0.5592	0.5514
# 7	4.7696	3.1514	96.1604	6.4666	0.5857	0.5970
# 8	4.9426	2.8933	85.0326	6.5816	0.5153	0.5930
# 9	4.9018	3.0973	84.8521	6.9213	0.4557	0.4969



(c)

Figure 2.10 Performance of the proposed NSST-MIF approach compared with the source CT and MR image dataset

Furthermore, the superiority of the proposed NSST-MIF approach is investigated by comparing its results with the following existing fusion methods:

Method 1: Wavelet based image fusion (WT_AVG_MAX) in which the averaging is used for the low frequency approximation fusion and maximum selection rule for the fusion of high frequency detail subimages [86, 92].

Method 2: Fusion method based on the NSCT decomposition with averaging and maximum (NSCT_AVG_MAX) for low and high frequency coefficients, as described in [92] with decomposition level = [4, 8, 8, 16].

Method 3: NSST based image fusion method with similar fusion rules as method 2 (NSST_AVG_MAX) as shown in [121] with decomposition level = [2, 3, 3].

Method 4: Image fusion method in the NSCT domain using maximum selection and the spatial frequency motivated PCNN (NSCT_MAX_SF_PCNN) as described in [145] with decomposition level = [1 2 4] and PCNN parameter $\alpha_L = 0.06931$, $\alpha_T = 0.2$, $\beta = 0.2$, $V_L = 1$, $V_T = 20$, $n = 200$ and $W = [\,0.707, 1, 0.707;\ 1, 0, 1;\ 0.707, 1, 0.707]$. Maximum selection and SF_PCNN are used for the fusion of low and high frequency subbands, respectively.

Method 5: NSCT based CT and MR fusion method using maximum selection rule and the MSF motivated PCNN model (NSCT_MAX_MSF_PCNN) as discussed in [34] with similar parameters as mentioned in method 4.

Method 6: The PCNN based fusion method in NSST domain with maximum selection rule as described in [50] and named as NSST_MAX_SF_PCNN. In this method, each PCNN neuron is motivated by the SF of the high frequency subimages of the NSST.

Method 7: NSST based image fusion using max selection and MSF motivated PCNN model (NSST_MAX_MSF_PCNN) with the decomposition level = [2, 3, 3] and the PCNN parameters as $\alpha_L = 0.3$, $\alpha_T = 0.1$, $\beta = 0.2$, $V_L = 1$, $V_T = 10$, $W = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & \sqrt{2} & 1 \\ \sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{bmatrix}$ and the maximum number of iterations $= 200$.

Method 8: Image fusion using the PCNN model in NSCT domain (NSCT_RE_NSML_PCNN) in which regional energy is used for fusing the low frequency subband and the NSML motivated PCNN model is utilized for the fusion of high frequency NSCT coefficients as described in [23] with the similar parameters and decomposition level as discussed in method 4.

Method 9: The proposed CT and MR fusion method where NSML and PCNN is used for hf subband coefficients and a fusion map is generated by computing the RE for lf NSST coefficients. To implement the proposed fusion approach, similar parameters are considered as used in method 7.

For the comparative analysis of the fusion results produced by the aforementioned methods, nine different pairs of the CT and MR brain images as shown in Figure 2.11 (a) and (b), respectively, are used. Their corresponding fused images are also shown in Figure 2.11 (c)-(k). From the visual analysis of all these results, it is observed that the proposed method is able to successfully preserve both the feature information such as bony structure of the CT images and soft tissue information of the MR image with better resolution as compared to other methods. For better presentation that also helps for visual analysis, the CT and MR image pair 4 and 8 are shown in Figure 2.12 and Figure 2.13, respectively. From these results mentioned in Figure 2.12 and Figure 2.13, it can be easily observed that the fused images provided by the proposed method are better than others.

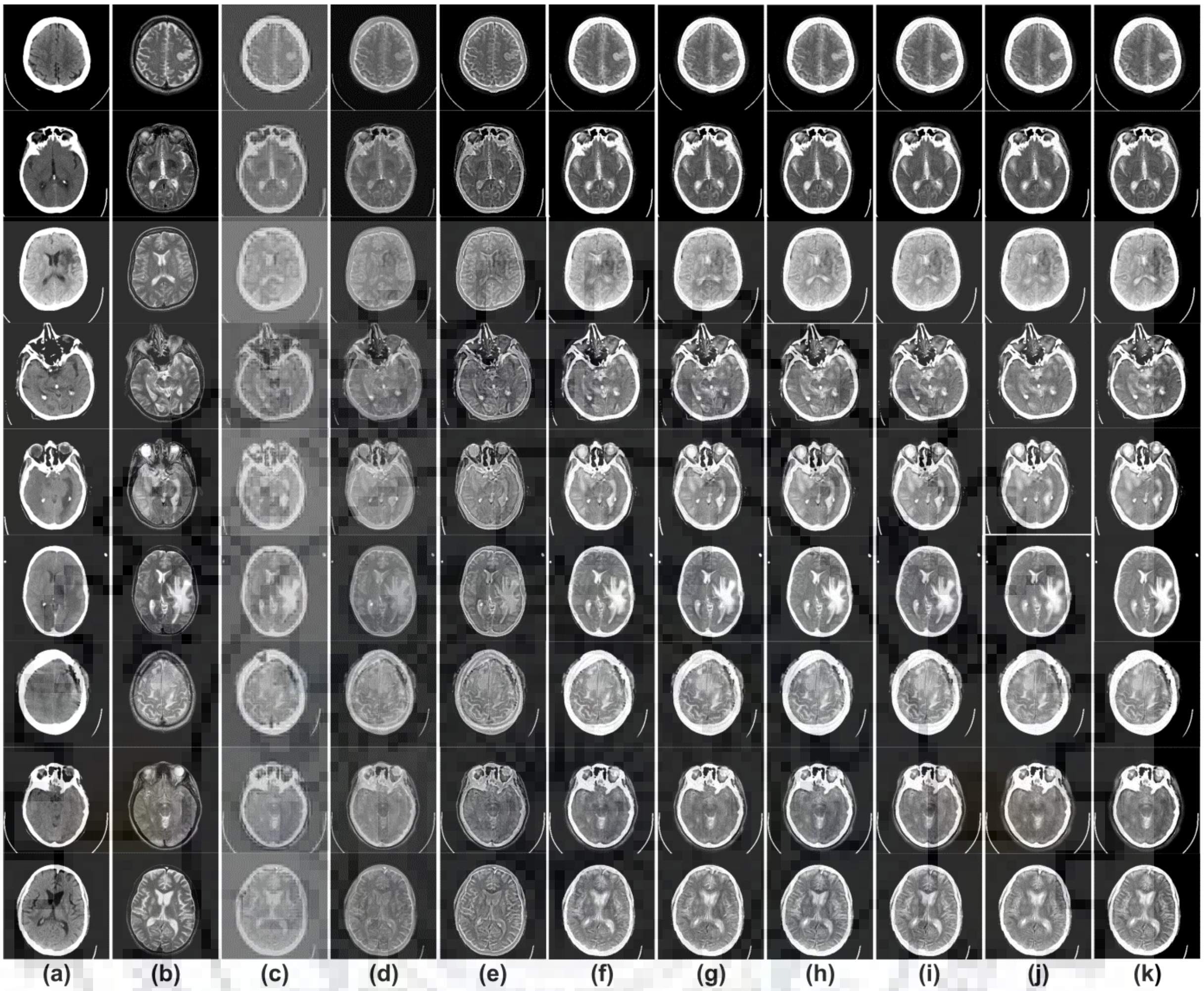


Figure 2.11 Comparative results of the different fusion methods applied to the source (a) CT images and (b) MR images. Rest of the columns show the fused image produced by the (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Method 6 (i) Method 7 (j) Method 8 (k) Proposed NSST-MIF method

Furthermore, the subjective evaluations of the proposed NSST-MIF and other fusion methods are supported by the quantitative results using six different performance measures as mentioned in Chapter 1. These objective outcomes obtained for different datasets using different fusion schemes are listed in Table 2.2 to Table 2.7. From the results mentioned in Table 2.2, it is seen that the proposed fusion method gains higher entropy than the input CT and MR images and other methods. For some cases, the WT based method 1 also achieves higher entropy values than others, but still it is less than the proposed method. Table 2.3 shows the standard deviation values of the source CT, MR and fused images produced by the different fusion methods. The higher value of the standard deviation of fused images reveals better contrast than the input images and other fused images produced by the other methods. Table 2.4 also shows the highest value of the MI that indicates the large amount of information from the input CT and MR images is retained in the fused image generated by the proposed fusion method. Table 2.5 shows the spatial frequency as another performance metric value for the input and fused images obtained by the different fusion techniques. The higher SF values for the proposed method shows that the fused images obtained by the proposed method have more information present in the source CT and MR images and also presents more activity and clarity level than the given input images. Besides this, the maximum value of IQI shown in Table 2.6 ensures better quality of the fused images produced by the proposed fusion method as compared to others. Moreover, another parameter Xydeas edge index (XEI) is also evaluated for the entire CT and MR image dataset and compared with all fusion methods as mentioned above and listed in Table 2.7. From the results presented in Table 2.7, higher value of XEI of the proposed method than others signifies that the proposed method is able to preserve the edge information of the original CT and MR images in the fused images.

Table 2.2 Comparative evaluation of entropy (En) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	5.2127	5.0814	5.2057	5.1121	5.2727	5.2989	5.3401	5.2839	5.3430
# 2	4.8940	4.8612	4.8638	4.9348	4.9361	4.9914	4.9941	4.9362	5.0335
# 3	4.4542	4.5273	4.6208	4.6876	4.6935	4.7247	4.7301	4.7231	4.8078
# 4	5.2507	5.2802	5.2286	5.2871	5.3184	5.3486	5.3937	5.3255	5.4243
# 5	4.8650	4.7992	4.8337	4.8924	4.9108	4.9528	4.9768	4.9413	4.9927
# 6	4.1145	4.2413	4.2791	4.2841	4.3675	4.4379	4.5661	4.4298	4.6794
# 7	4.0108	4.0834	4.1262	4.1826	4.1897	4.5431	4.6669	4.2629	4.7696
# 8	4.7939	4.8264	4.8554	4.8586	4.8604	4.8970	4.9093	4.8878	4.9426
# 9	4.5105	4.6513	4.6864	4.6937	4.724	4.7826	4.799	4.7392	4.9018

Table 2.3 Comparative evaluation of standard deviation (STD) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	65.380	63.982	67.512	86.128	86.275	86.695	86.841	86.546	87.186
# 2	59.750	57.758	61.896	81.305	81.587	81.737	81.954	81.544	82.035
# 3	66.890	65.842	68.378	91.637	91.667	92.174	92.394	92.162	92.686
# 4	57.008	60.982	63.388	83.266	83.368	83.492	83.721	83.393	83.814
# 5	65.749	63.765	67.744	89.121	89.489	89.653	89.898	89.643	89.908
# 6	62.414	60.548	63.579	85.309	85.631	85.948	86.106	85.642	86.153
# 7	69.762	69.192	71.432	95.027	95.402	95.789	95.793	95.421	96.16
# 8	62.674	60.946	65.385	84.304	84.363	84.659	84.675	84.631	85.033
# 9	61.429	59.674	62.947	83.952	84.361	84.433	84.78	84.394	84.852

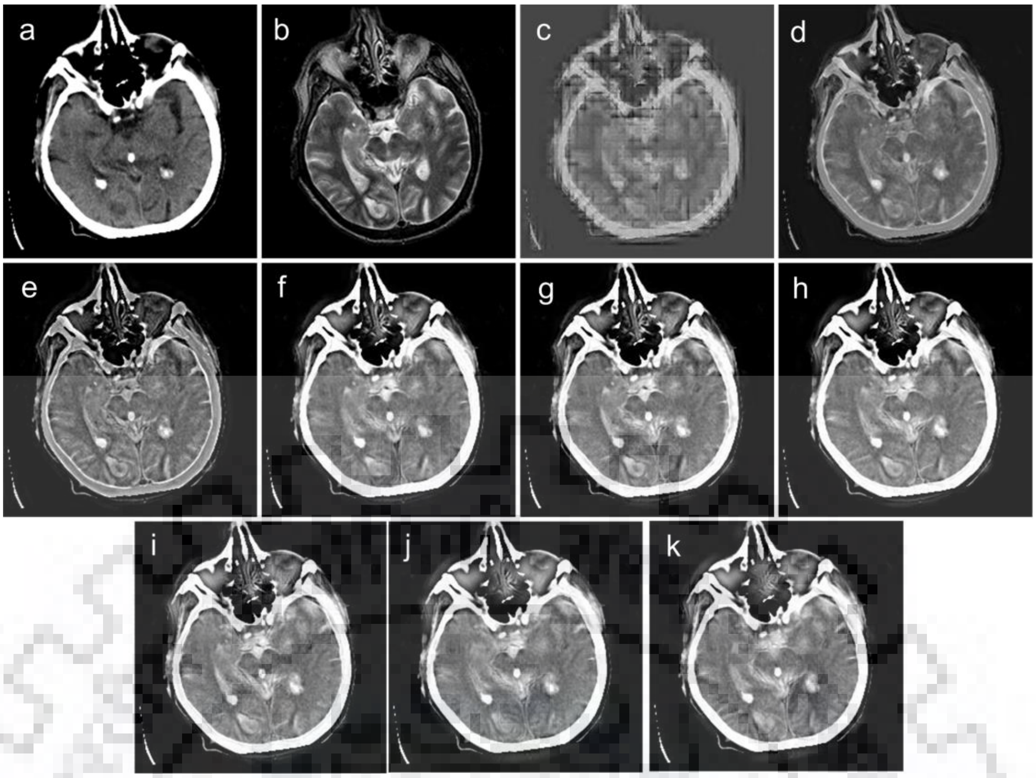


Figure 2.12 Comparative analysis of visual results obtained by different fusion methods applied to the image pair-4 (a) CT image (b) MR image. The fused image provided by the (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Method 6 (i) Method 7 (j) Method 8 (k) Proposed NSST-MIF method

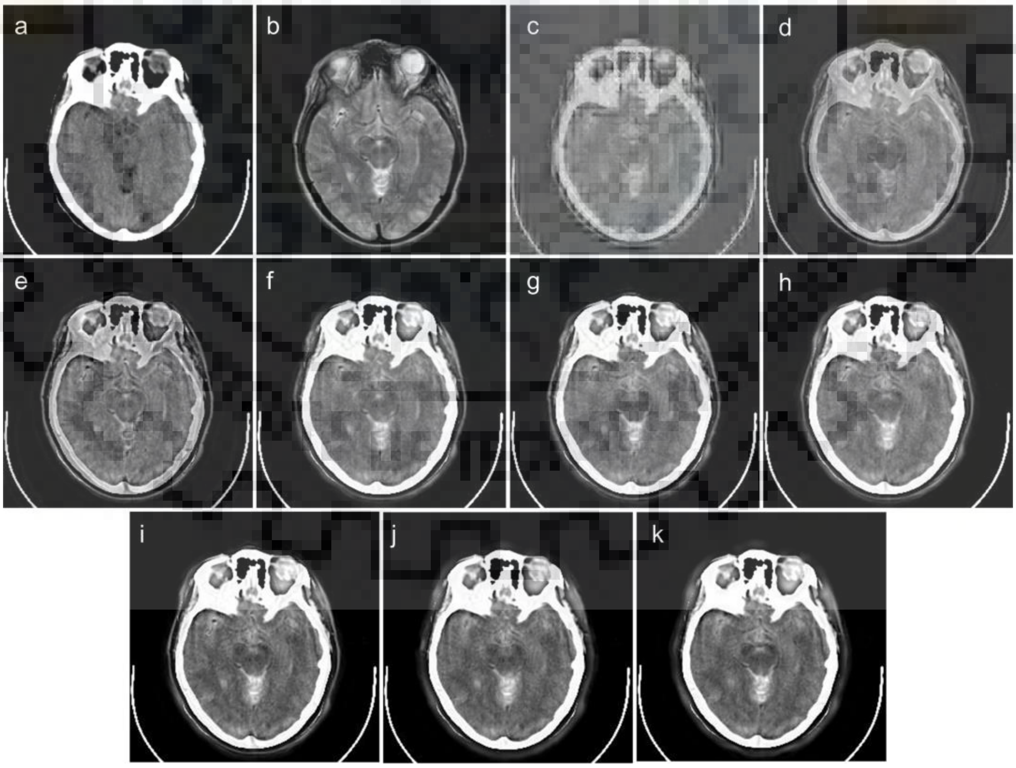


Figure 2.13 Comparative analysis of visual results obtained by different fusion methods applied to the image pair-8 (a) CT image (b) MR image. The fused image provided by the (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Method 6 (i) Method 7 (j) Method 8 (k) Proposed NSST-MIF method

Table 2.4 Comparative evaluation of mutual information (MI) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	2.5634	3.0854	3.2190	3.1058	3.1384	3.2226	3.2462	3.1822	3.2956
# 2	2.524	2.7632	2.9961	3.0139	3.0489	3.0644	3.0792	3.0604	3.0853
# 3	2.3354	2.4483	2.6345	2.9971	3.0389	3.031	3.0594	3.0454	3.0724
# 4	2.1773	2.4156	2.7054	2.7727	2.8223	2.8432	2.8469	2.8377	2.8509
# 5	2.3399	2.5264	2.7793	2.9413	2.9532	2.9722	2.9865	2.9641	2.9886
# 6	2.6323	2.7924	3.1108	3.1173	3.1358	3.1570	3.1730	3.1544	3.2026
# 7	2.4357	2.5998	2.8338	3.0357	3.0390	3.0913	3.0978	3.0902	3.1514
# 8	2.3288	2.5726	2.8409	2.8499	2.8607	2.8813	2.8877	2.8624	2.8933
# 9	2.4312	2.6263	2.8091	3.0268	3.0406	3.0502	3.0569	3.0475	3.0973

Table 2.5 Comparative evaluation of spatial frequency (SF) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	5.1216	4.8845	5.2468	5.3696	5.421	5.5828	5.6992	5.5582	5.7153
# 2	6.6219	6.2814	6.5585	6.3475	6.6242	6.6959	6.8099	6.3185	6.8577
# 3	5.5161	5.6126	6.0983	5.6933	5.8879	6.1315	6.148	5.8938	6.2712
# 4	7.1815	7.1322	7.5095	7.2480	7.3292	7.6077	7.6801	7.4228	8.8396
# 5	6.5576	6.2480	6.7447	6.3264	6.6292	6.8515	6.9478	6.6472	7.7340
# 6	5.5684	5.3986	5.5623	5.4073	5.5829	5.7063	5.7910	5.6298	5.8769
# 7	6.3112	5.8942	5.9743	6.0055	6.2845	6.3453	6.3712	6.2919	6.4666
# 8	5.8714	6.0644	6.1661	6.1553	6.2186	6.4386	6.5236	6.3458	6.5816
# 9	6.0779	6.2712	6.5367	6.5796	6.7005	6.8568	6.8796	6.7025	6.9213

Table 2.6 Comparative evaluation of image quality index (IQI) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	0.2822	0.3642	0.4204	0.4415	0.4542	0.4709	0.4764	0.4616	0.4839
# 2	0.3052	0.3311	0.4097	0.4128	0.4346	0.4482	0.4526	0.4401	0.4627
# 3	0.3290	0.3347	0.3997	0.4173	0.4354	0.4396	0.4490	0.4387	0.4833
# 4	0.3244	0.3945	0.4502	0.4278	0.4365	0.4553	0.4597	0.4423	0.4649
# 5	0.2964	0.3683	0.4336	0.4225	0.4303	0.4470	0.4583	0.4388	0.4602
# 6	0.2883	0.3404	0.4872	0.5131	0.5304	0.5482	0.5495	0.5422	0.5592
# 7	0.3655	0.3498	0.5218	0.5681	0.5702	0.5756	0.5825	0.5746	0.5857
# 8	0.3781	0.3702	0.4452	0.4662	0.4722	0.5023	0.5058	0.4855	0.5153
# 9	0.2856	0.3008	0.4295	0.4333	0.4477	0.4518	0.4527	0.4511	0.4557

Table 2.7 Comparative evaluation of edge index (XEI) values obtained by the proposed NSST-MIF with other fusion schemes

Image Pairs	Fusion Methods								
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	Method 7	Method 8	Proposed
# 1	0.2769	0.3847	0.438	0.4299	0.4375	0.4494	0.4631	0.4480	0.4685
# 2	0.2586	0.3624	0.4592	0.4636	0.5088	0.5134	0.5164	0.5098	0.5496
# 3	0.2592	0.3516	0.4059	0.4172	0.4234	0.4261	0.4347	0.4236	0.5002
# 4	0.2502	0.3592	0.4963	0.4692	0.4866	0.5122	0.5143	0.4962	0.5255
# 5	0.2405	0.3478	0.4847	0.4723	0.4729	0.4919	0.4964	0.4843	0.5188
# 6	0.2669	0.3514	0.4736	0.4957	0.4975	0.5112	0.5283	0.5008	0.5514
# 7	0.2758	0.4208	0.5163	0.4805	0.4825	0.5231	0.5399	0.5211	0.5970
# 8	0.2929	0.4116	0.5448	0.5490	0.5772	0.5797	0.5856	0.5784	0.5931
# 9	0.2395	0.3358	0.4102	0.4164	0.4203	0.4244	0.4295	0.4211	0.4969

Moreover, Table 2.8 provides a comparison of the averaged performance measure (mean±standard deviation) obtained by the different fusion methods. From the analysis of the results presented in Table 2.8, some remarkable points can be summarized as,

a) The NSST based fusion approach obviously outperforms the other fusion methods. The proposed NSST-MIF method achieves 62.77%, 28.56% higher entropy than source the CT and MR images, respectively. Moreover, the proposed method gains approx 6.62%, 6.08%, 5.14%, 4.57%, 3.75%, 2.09%, 1.17% and 3.14% higher entropy than all fusion methods 1 to 8 respectively. These results ensure that the more information lies in the fused images obtained by the produced method than others.

b) The proposed fusion approach also reveals an improvement in the contrast resolution of the fused images as compared to others by achieving the 37.96 %, 41.10%, 33.02%, 0.57-1.0% and 0.21-0.41% larger standard deviation of the fused images than WT_AVG_MAX, NSCT_AVG_MAX, NSST_AVG_MAX, NSCT_PCNN and NSST_PCNN based fusion methods, respectively. The proposed method also gains approx 4.32% and 47.49% higher STD values than both the CT and MR images.

c) The higher values of MI and SF metrics of the fused images are produced by the NSST-MIF. It gains approx 11.73%, 7.84-11.74% and 4.10-5.23% higher SF values than WT, NSCT and NSST based fusion approach, respectively. Moreover, it has 26.97% and 1.44-16.21% larger MI values than the WT and NSCT based fusion method, respectively. It achieves approx 29.22% and 7.52% higher SF values than the input reference CT and MR image, respectively.

d) Finally, the proposed NSST-MIF method also gets higher values of both the quality and edge index that provide better quality of the fused images with better edge transformation from the input reference images. The averaged IQI values obtained by the proposed NSST-MIF method increases by approx. 4.59-9% and 1.93-3.05% from NSCT and NSST based PCNN models. Moreover, the proposed approach achieves approx 9.53-14.46% larger XEI values than NSCT based PCNN models.

Table 2.8 Comparative analysis of the averaged performance measures (mean ± standard deviation) obtained by the proposed method and others

Methods	Performance Measures					
	EN	MI	STD	SF	IQI	XEI
CT	3.0647 ± 0.3035	-	83.9079 ± 5.4953	5.2680 ± 0.6784	-	-
MR	3.8803 ± 0.3366	-	59.3509 ± 4.9843	6.3309 ± 0.8517	-	-
Method 1	4.6785 ± 0.4401	2.4186 ± 0.1404	63.4506 ± 3.8991	6.0920 ± 0.6451	0.3172 ± 0.0351	0.2623 ± 0.0177
Method 2	4.7024 ± 0.3941	2.6424 ± 0.2044	62.0401 ± 3.5699	5.9626 ± 0.6356	0.3411 ± 0.0268	0.3556 ± 0.0311
Method 3	4.7444 ± 0.3707	2.8810 ± 0.1908	65.8068 ± 3.1501	6.2664 ± 0.6702	0.4441 ± 0.0387	0.4699 ± 0.0468
Method 4	4.7703 ± 0.3585	2.9845 ± 0.1133	86.6721 ± 4.4163	6.1258 ± 0.5964	0.4558 ± 0.0524	0.4660 ± 0.0421
Method 5	4.8081 ± 0.3713	3.0086 ± 0.1102	86.9047 ± 4.4341	6.2976 ± 0.6017	0.4679 ± 0.0494	0.4785 ± 0.0490
Method 6	4.8863 ± 0.3073	3.0348 ± 0.1218	87.1756 ± 4.5285	6.4685 ± 0.6267	0.4821 ± 0.0493	0.4924 ± 0.0508
Method 7	4.9307 ± 0.2848	3.0482 ± 0.1267	87.3512 ± 4.4871	6.5389 ± 0.6230	0.4874 ± 0.0486	0.5009 ± 0.0510
Method 8	4.8366 ± 0.3499	3.0271 ± 0.1189	87.0419 ± 4.4968	6.3123 ± 0.5819	0.4750 ± 0.0503	0.4870 ± 0.0503
Proposed	**4.9883** ± 0.2508	**3.0708** ± 0.1423	**87.5363** ± 4.5573	**6.8071** ± 0.9693	**0.4968** ± 0.0470	**0.5334** ± 0.0435

2.9.2 Experiment 2: Analysis and evaluation of the proposed DRT-MIF approach

This experiment presents the subjective analysis of fusion performance obtained by the proposed DRT-MIF approach. To perform all the experiments, the following parameters are selected based on the successive experiments with DRT parameters as $c = 1$ and $d = 4$ for the decomposition levels of the input reference images and the PCNN parameters $\alpha_L = 0.3$, $\alpha_T = 0.1$, $\beta = 0.2$, $V_L = 1$, $V_T = 10$, $W = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2};\ 1\ \sqrt{2}\ \ 1\right]$ and iteration=150. To conduct this experiment, nine pairs of different source images have been considered as similar used in experiment 1 and shown in Figure 2.14 (a) and (b) and the fused images are shown in Figure 2.14 (c), correspondingly. From the resultant images presented in Figure 2.14 (c), it is visualized that the output fused images provide more diagnostic information extracted by the proposed DRT-MIF approach. It is also verified by the quantitative measures computed for the proposed DRT-MIF approach and presented in Table 2.9. From the results listed in Table 2.9, it is observed that all the fused images have better quantitative measures by achieving the significant values of the En, MI, STD, SF, and XEI. To analyze the superiority of the proposed approach in a better way, a bar graph is also shown in Figure 2.15 that is able to present the improved values of En, STD and SF computed for the fused images compared to the reference images. The higher En values indicate more informative content preserved in the fused images and higher contrast is shown by higher STD value.

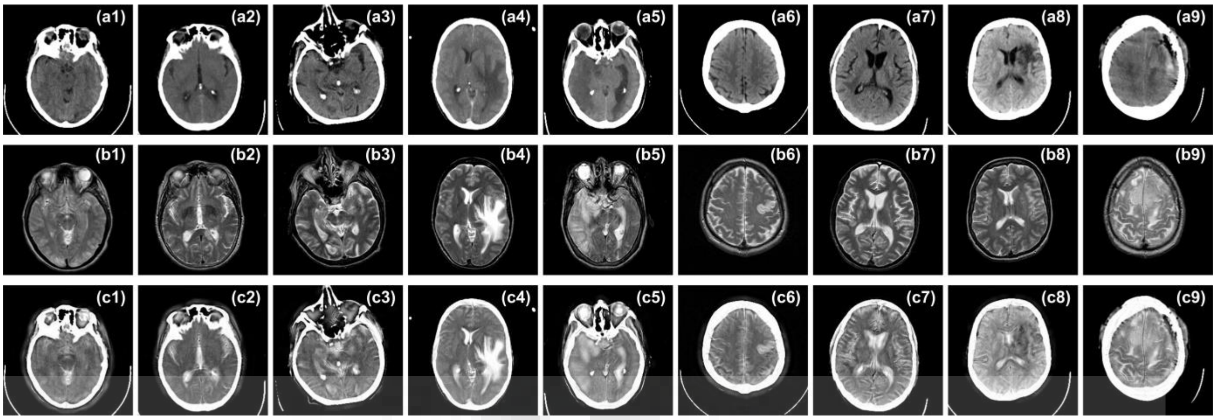


Figure 2.14 (a) Source CT images (b) Source MR images (c) Fused images obtained by the DRT-MIF approach

Table 2.9 Performance indices obtained by the DRT-MIF approach for fused images

Dataset	Performance Measures				
	EN	MI	STD	SF	XEI
# 1	4.9681	2.9125	85.9565	6.5978	0.6142
# 2	5.1254	3.2145	82.7658	6.9045	0.5541
# 3	5.4298	2.8754	84.8644	8.8431	0.5286
# 4	4.8954	3.2214	87.2412	8.9176	0.5784
# 5	5.2125	3.0452	91.0145	7.7455	0.5544
# 6	5.5232	3.3745	89.5877	5.6754	0.4916
# 7	5.0147	3.1712	85.2565	6.9454	0.5014
# 8	4.9876	3.1420	93.0147	6.2787	0.5115
# 9	4.9512	3.2419	96.6752	6.4785	0.6210

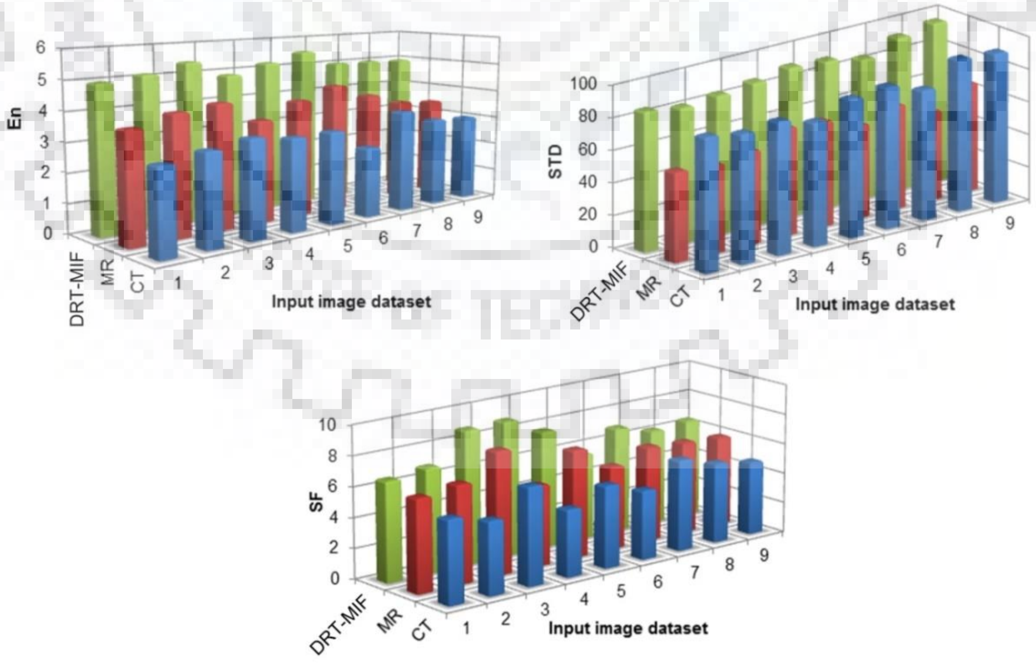


Figure 2.15 Performance of the proposed DRT-MIF approach compared with the source CT and MR image dataset

Furthermore, a comparative study is presented among the fused images obtained by the proposed DRT-MIF approach and other existing methods given as follows.

Method 1: Wavelet based image fusion (WT_AVG_MAX) in which the averaging is used for the low frequency approximation fusion and maximum selection rule for the fusion of high frequency detail subimages [86, 92]. Three level decomposition is used for decomposing the source CT and MR images.

Method 2: Fusion method based on the NSCT decomposition with averaging and maximum (NSCT_AVG_MAX) for lf and hf subband coefficients, as described in [92] with decomposition level = [4, 8, 8, 16].

Method 3: NSCT based CT and MR fusion method using maximum selection rule and the MSF motivated PCNN model (NSCT_MAX_MSF_PCNN) as discussed in [34] with the NSCT decomposition level = [1 2 4] and PCNN parameter $\alpha_L = 0.06931,\ \alpha_T = 0.2,\ \beta = 0.2,\ V_L = 1,$ $V_T = 20,\ n = 200$ and $W = [\,0.707, 1, 0.707;\ 1, 0, 1;\ 0.707, 1, 0.707]$. Maximum selection and SF_PCNN based fusion rules are used for the fusion of low and high frequency subbands, respectively.

Method 4: Image fusion using PCNN model in NSCT domain named as the NSCT_RE_NSML_PCNN in which regional energy is used for fusing the low frequency subband and the NSML motivated PCNN model is utilized for the fusion of high frequency NSCT coefficients as described in [23] with the similar parameters and decomposition level as considered in method 3.

Method 5: Image fusion using the NSST domain with max selection rule and the PCNN model (NSST_MAX_SF_PCNN) as described in [50]. In this approach, a neuron is stimulated by the SF computed for high frequency NSST coefficients with the decomposition level = [2, 3, 3] and the PCNN parameters as $\alpha_L = 0.3,\ \alpha_T = 0.1,\ \beta = 0.2,\ V_L = 1,\ V_T = 10,$ $W = \begin{bmatrix} \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \\ 1 & 0 & 1 \\ \frac{1}{\sqrt{2}} & 1 & \frac{1}{\sqrt{2}} \end{bmatrix}$ and $n = 200$.

Method 6: The NSST based fusion approach (NSST-MIF) as discussed in section 2.5.

Method 7: The proposed DRT-MIF approach with similar parameters to perform the above experiment.

To analyze the comparison between the fusion results obtained by the different approaches as mentioned above, the fused images are shown in Figure 2.16 (c)-(i). From the results as mentioned in Figure 2.16, it is visualized that the proposed DRT-MIF approach succeeded to retain both the soft tissue content and bony information in comparison to other fusion approaches. Moreover, to validate the subjective results obtained by the other fusion approaches, a comparative analysis is also done for the different computed quantitative measures listed in Table 2.10 to Table 2.14. The results mentioned in Table 2.10 and Table 2.12 explain the superiority of the proposed DRT-MIF approach over the other existing fusion methods by achieving higher En and STD values. Table 2.11 and Table 2.13 present larger values of MI and SF values that also signify the preservation of more information in the fused images obtained by the DRT-MIF approach compared to the other methods. Moreover, Table 2.14 shows higher XEI values to indicate better preservation of the edge details by the proposed DRT-MIF approach. To support all the quantitative results, the averaged performance indices evaluated by the all the fusion approaches mentioned above are provided in Table 2.15 from which some of the following points are summarized.

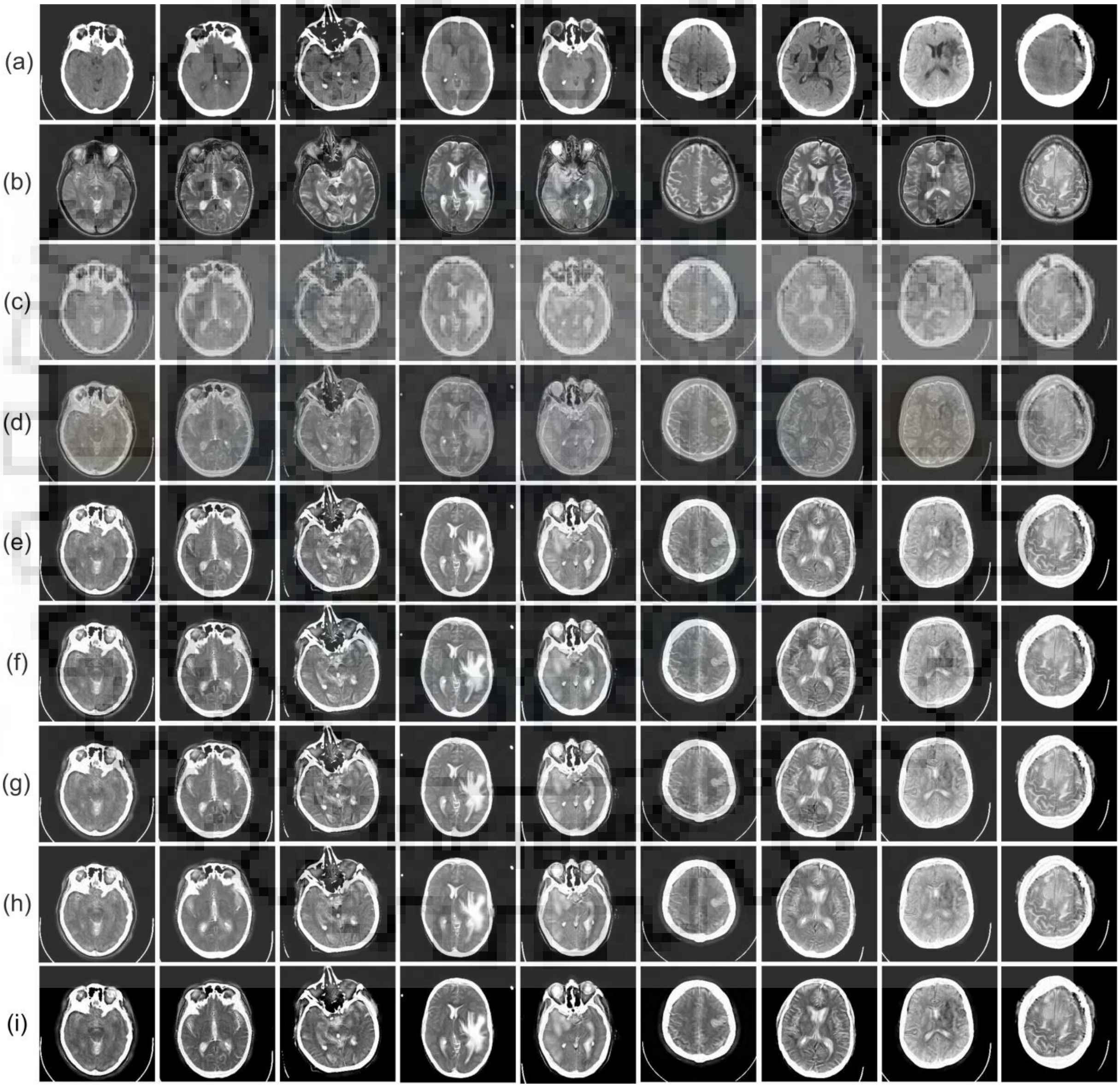


Figure 2.16 Comparative visual results obtained by the fusion methods applied to the reference (a) CT and (b) MR images. The fused image obtained by the (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Method 6 (i) Proposed DRT-MIF method

Table 2.10 Comparative evaluation of entropy (En) values computed for the proposed DRT-MIF and other existing fusion approaches

Image Pairs	Fusion Methods						
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	DRT-MIF
#1	4.7939	4.8013	4.8604	4.8878	4.8970	4.9426	4.9681
#2	4.8940	4.8601	4.9361	4.9362	4.9914	5.0335	5.1254
#3	5.2507	5.296	5.3184	5.3255	5.3486	5.4243	5.4298
#4	4.1145	4.2301	4.3675	4.4298	4.4379	4.6794	4.8954
#5	4.8650	4.7970	4.9108	4.9413	4.9528	4.9927	5.2125
#6	5.2127	5.1201	5.2727	5.2839	5.2989	5.3430	5.5232
#7	4.5105	4.6494	4.724	4.7392	4.7826	4.9018	5.0147
#8	4.4542	4.5035	4.6935	4.7231	4.7247	4.8078	4.9876
#9	4.0108	4.0645	4.1897	4.2629	4.5431	4.7696	4.9512

Table 2.11 Comparative evaluation of mutual information (MI) values computed for the proposed DRT-MIF and other existing fusion approaches

Image Pairs	Fusion methods						
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	DRT-MIF
#1	2.3288	2.5628	2.8607	2.8624	2.8813	2.8933	2.9125
#2	2.5240	2.7564	3.0489	3.0604	3.0644	3.0853	3.2145
#3	2.1770	2.4050	2.8223	2.8377	2.8432	2.8509	2.8754
#4	2.6323	2.7863	3.1358	3.1544	3.1570	3.2026	3.2214
#5	2.3399	2.5112	2.9532	2.9641	2.9722	2.9886	3.0452
#6	2.5634	3.0783	3.1384	3.1822	3.2226	3.2956	3.3745
#7	2.4310	2.6086	3.0406	3.0475	3.0502	3.0973	3.1712
#8	2.3354	2.4772	3.0389	3.0454	3.0310	3.0724	3.1420
#9	2.4357	2.5961	3.0390	3.0902	3.0913	3.1514	3.2419

Table 2.12 Comparative evaluation of standard deviation (STD) values computed for the proposed DRT-MIF and other existing fusion approaches

Image Pairs	Fusion Methods						
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	DRT-MIF
#1	62.674	60.324	84.363	84.631	84.659	85.033	85.957
#2	59.750	57.183	81.587	81.544	81.737	82.035	82.766
#3	57.008	60.234	83.368	83.393	83.492	83.814	84.864
#4	62.414	60.492	85.631	85.642	85.948	86.153	87.241
#5	65.749	63.178	89.489	89.643	89.653	89.908	91.015
#6	65.38	63.842	86.275	86.546	86.695	87.186	89.588
#7	61.429	59.097	84.361	84.394	84.433	84.852	85.257
#8	66.89	65.102	91.667	92.162	92.174	92.686	93.015
#9	69.762	68.909	95.402	95.421	95.789	96.160	96.675

Table 2.13 Comparative evaluation of spatial frequency (SF) values computed for the proposed DRT-MIF and other existing fusion approaches

Image Pairs	Fusion Methods						
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	DRT-MIF
#1	5.8714	6.0536	6.2186	6.3458	6.4386	6.5816	6.5978
#2	6.6219	6.2705	6.6242	6.3185	6.6959	6.8577	6.9045
#3	7.1815	7.1192	7.3292	7.4228	7.6077	8.8396	8.8431
#4	5.5684	5.3852	5.5829	5.6298	5.7063	5.8769	8.9176
#5	6.5576	6.2101	6.6292	6.6472	6.8515	7.7340	7.7455
#6	5.1216	4.8754	5.421	5.5582	5.5828	5.7153	5.6754
#7	6.0779	6.2537	6.7005	6.7025	6.8568	6.9213	6.9454
#8	5.5161	5.6107	5.8879	5.8938	6.1315	6.2712	6.2787
#9	6.3112	5.8846	6.2845	6.2919	6.3453	6.4666	6.4785

Table 2.14 Comparative evaluation of edge index (XEI) values computed for the proposed DRT-MIF and other existing fusion approaches

Image Pairs	Fusion Methods						
	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6	DRT-MIF
#1	0.2929	0.3930	0.5772	0.5784	0.5797	0.5930	0.6142
#2	0.2586	0.3460	0.5088	0.5098	0.5134	0.5496	0.5541
#3	0.2502	0.3401	0.4866	0.4962	0.5122	0.5255	0.5286
#4	0.2669	0.3456	0.4975	0.5008	0.5112	0.5514	0.5784
#5	0.2405	0.3328	0.4729	0.4843	0.4919	0.5188	0.5544
#6	0.2769	0.3784	0.4375	0.448	0.4494	0.4685	0.4916
#7	0.2395	0.3118	0.4203	0.4211	0.4244	0.4969	0.5014
#8	0.2592	0.3443	0.4234	0.4236	0.4261	0.5002	0.5115
#9	0.2758	0.4086	0.4825	0.5211	0.5231	0.597	0.621

Table 2.15 Averaged performance evaluation parameters computed by the proposed DRT-MIF and other fusion approaches

Methods	Performance Measures				
	En	MI	STD	SF	XEI
Source CT	3.0647 ± 0.3035	-	83.9079 ± 5.4953	5.2680 ± 0.6784	-
Source MR	3.8803 ± 0.3366	-	59.3509 ± 4.9843	6.3309 ± 0.8517	-
Method 1	4.6785 ± 0.4401	2.4186 ± 0.1404	63.4506 ± 3.8991	6.0920 ± 0.6451	0.2623 ± 0.0177
Method 2	4.7057 ± 0.3802	2.6478 ± 0.2069	62.5210 ± 3.5133	5.9763 ± 0.6386	0.3694 ± 0.0297
Method 3	4.8081 ± 0.3713	3.0086 ± 0.1102	86.9047 ± 4.4341	6.2976 ± 0.6017	0.4785 ± 0.0490
Method 4	4.8366 ± 0.3499	3.0271 ± 0.1189	87.0419 ± 4.4968	6.3123 ± 0.5819	0.4870 ± 0.0503
Method 5	4.8863 ± 0.3073	3.0348 ± 0.1218	87.1756 ± 4.5285	6.4685 ± 0.6267	0.4924 ± 0.0508
Method 6	4.9883 ± 0.2508	3.0708 ± 0.1423	87.5363 ± 4.5573	6.8071 ± 0.9693	0.5334 ± 0.0435
DRT-MIF	**5.1231** ± 0.2230	**3.1332** ± 0.1616	**88.4863** ± 4.4637	**7.1541** ± 1.1257	**0.5506** ± 0.0471

2.9.3 Experiment 3: Analysis and evaluation of the proposed C-MIF approach

In order to investigate the proposed cascaded medical image fusion (C-MIF) approach, all the CT-MR image pairs used as similar for previous experiments are considered. The following parameters for DRT decomposition ($c = 1$ and $d = 4$) at stage-1, NSST decomposition level $= [2, 3, 3]$ at stage-2 are considered. For the fusion at stage-1, a PCNN model is utilized with the following selected parameters, $\alpha_L = 0.3$, $\alpha_T = 0.1$, $\beta = 0.2$, $V_L = 1$, $V_T = 10$, $W = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ \ 1\right]$ and iterations= 200.

For analyzing the qualitative performance of the proposed C-MIF approach, four image pairs out of CT-MR image dataset is shown in Figure 2.17(a) and (b), respectively. Their corresponding fused images are presented in Figure 2.17(c). From the subjective visual analysis of the fused images, it is experienced that the fused images have more informative content and edge details than the source images. These results were discussed with the radiologist also to meet all the requirements from a clinical point of view. The objective performance measures are also evaluated for all the fused images and presented in Table 2.16. The bar graph shown in Figure 2.18 represents an illustrative idea about the parameter variations such as En, STD and SF calculated by the proposed fusion approach against their original values of all nine source image pairs. From the quantitative results presented in Table 2.16 and Figure 2.18, it is clearly observed that the fused images having higher En values than the source CT and MR image depict more diagnostic information in the fused images. Moreover, the proposed approach provides higher values of MI that indicate more detailed and textural information present in the fused images. The proposed approach also gains higher STD values, presenting a better contrast compared to the reference CT and MR images. Table 2.16 indicates good SF values for the fused images which are verified by achieving better quality of fused images compared to all the input CT and MR images.

Table 2.16 Performance measures obtained by the proposed C-MIF approach

Image Pairs	Performance Measures				
	En	MI	STD	SF	XEI
# 1	4.9987	2.9347	85.9578	6.6024	0.6218
# 2	5.1623	3.2654	82.7842	6.9098	0.5682
# 3	5.5147	2.8824	84.8868	8.8542	0.5346
# 4	4.8542	3.2456	87.2521	8.9248	0.5878
# 5	5.2241	3.0874	91.1241	7.7845	0.5584
# 6	5.5274	3.3942	89.6024	5.6842	0.4968
# 7	5.1422	3.1847	85.2625	6.9684	0.5086
# 8	4.9894	3.2126	93.0189	6.2811	0.5215
# 9	4.9627	3.2275	96.6842	6.4818	0.6224

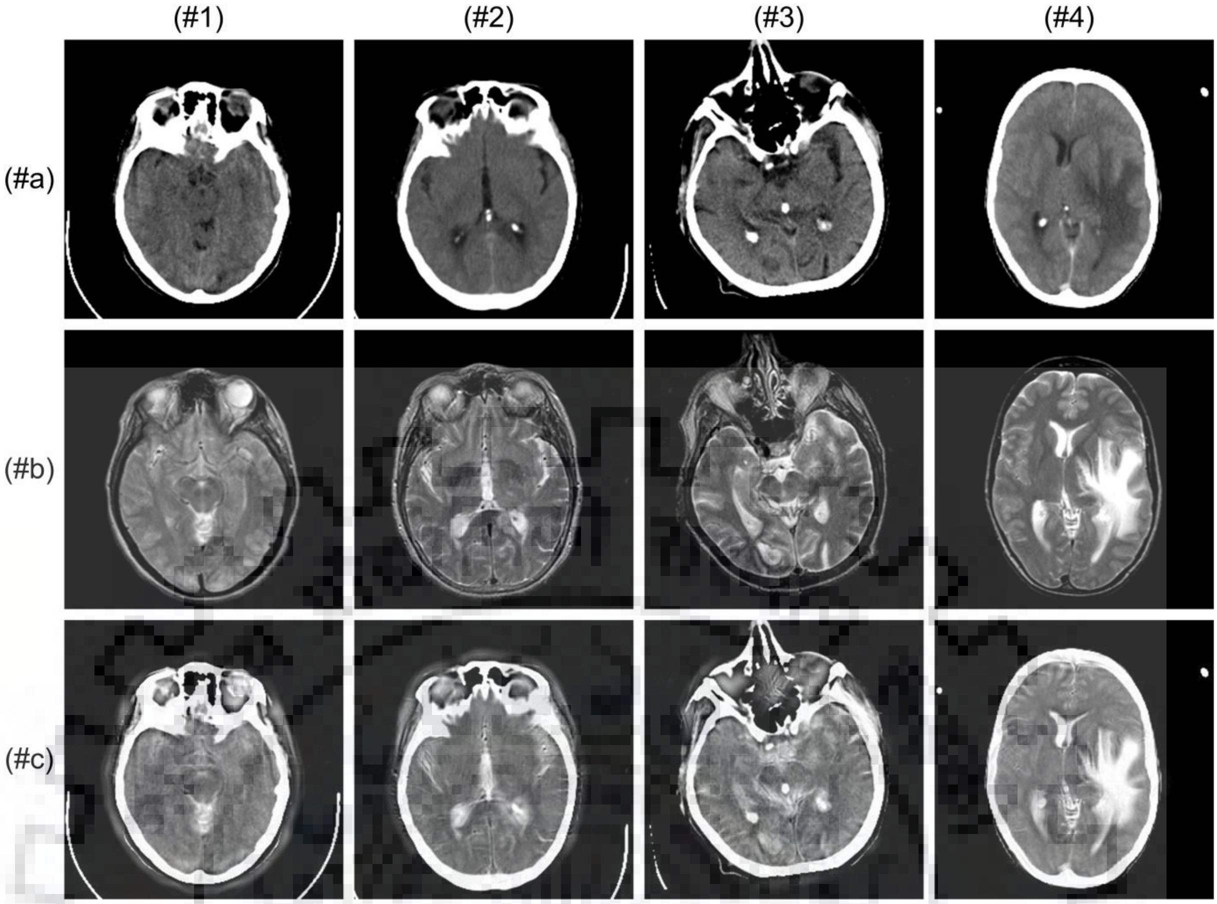


Figure 2.17 Source (a) CT images, (b) MR images (c) visual performance of fused images evaluated by the proposed C-MIF approach

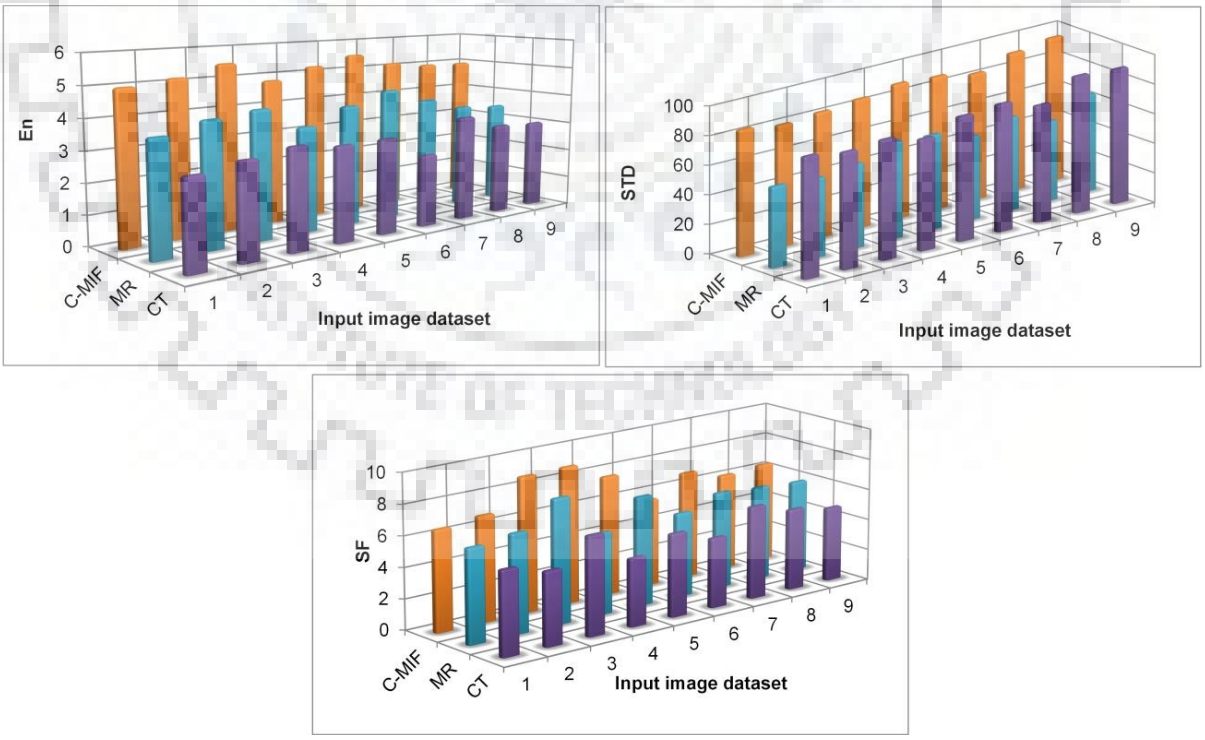


Figure 2.18 Bar graph to illustrate the fusion performance of the proposed C-MIF approach compared to source CT and MR images

Furthermore, a comparison of subjective and objective analysis is made for the fused images obtained by the proposed C-MIF approach and other existing fusion methods. For such comparative evaluation, different fusion approaches along with the proposed C-MIF are applied to complete CT-MR image datasets considered here, out of which some of them are shown in Figure 2.19, Figure 2.20, and Figure 2.21. Moreover, the superiority of the proposed C-MIF approach is assessed by comparing both the quantitative and qualitative results with the other existing fusion methods given as method 1, 2, 3, 5 and 6 described in subsection 2.9.2 of Chapter 2.

The fused image corresponding to the three image pairs as shown in Figure 2.19 to Figure 2.21 (a and b) are obtained from all the fusion methods considered alongwith the proposed C-MIF approach which are shown in Figure 2.19 to Figure 2.21 (c) to (h), respectively. For showing the proper visualization and improvement in the fused images, a rectangle is marked by red color on the source CT image of every image pair and the corresponding region is zoomed and shown inside the image.

Furthermore, two different information spots (marked by red and yellow arrow) are also marked to show the specific improvements in that particular region. Based on the insight of each fused image corresponds to the marked region, it is observed that the fused region obtained by the proposed C-MIF approach is more visible that having more soft and hard tissues information taken from the source CT and MR images as compared to the other existing fusion methods.

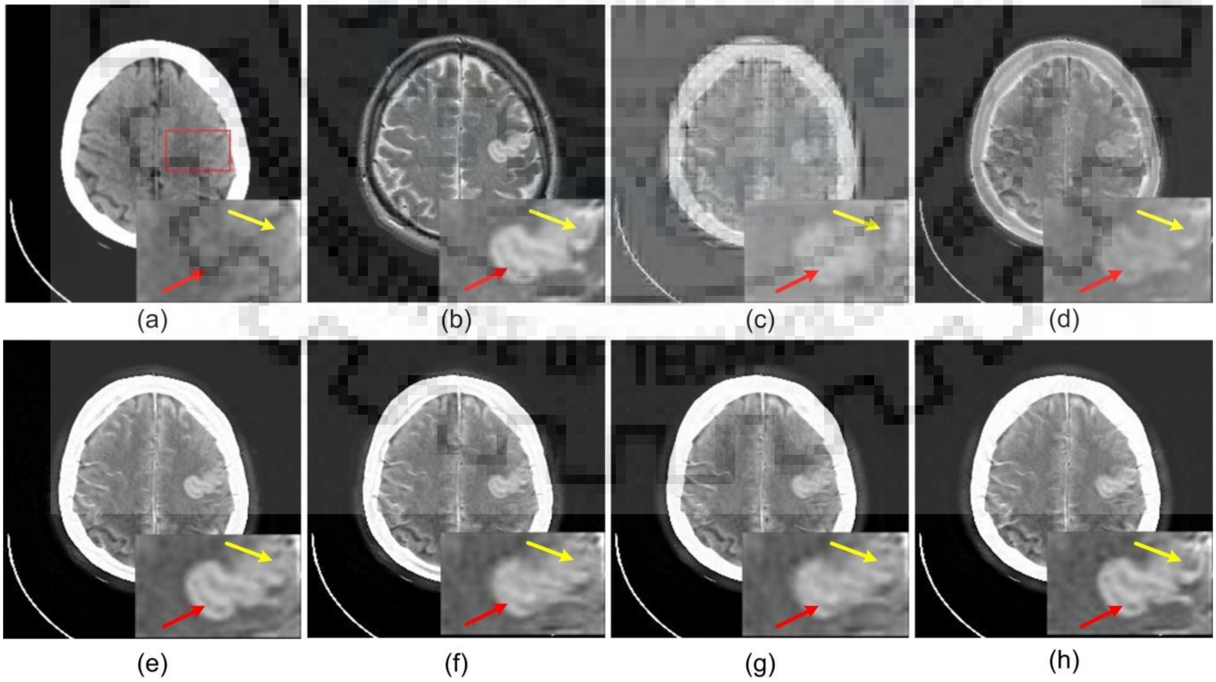


Figure 2.19 Comparison between visual performance evaluated by the fusion approaches for set-1 of (a) CT image and (b) MR image, using (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Proposed C-MIF approach

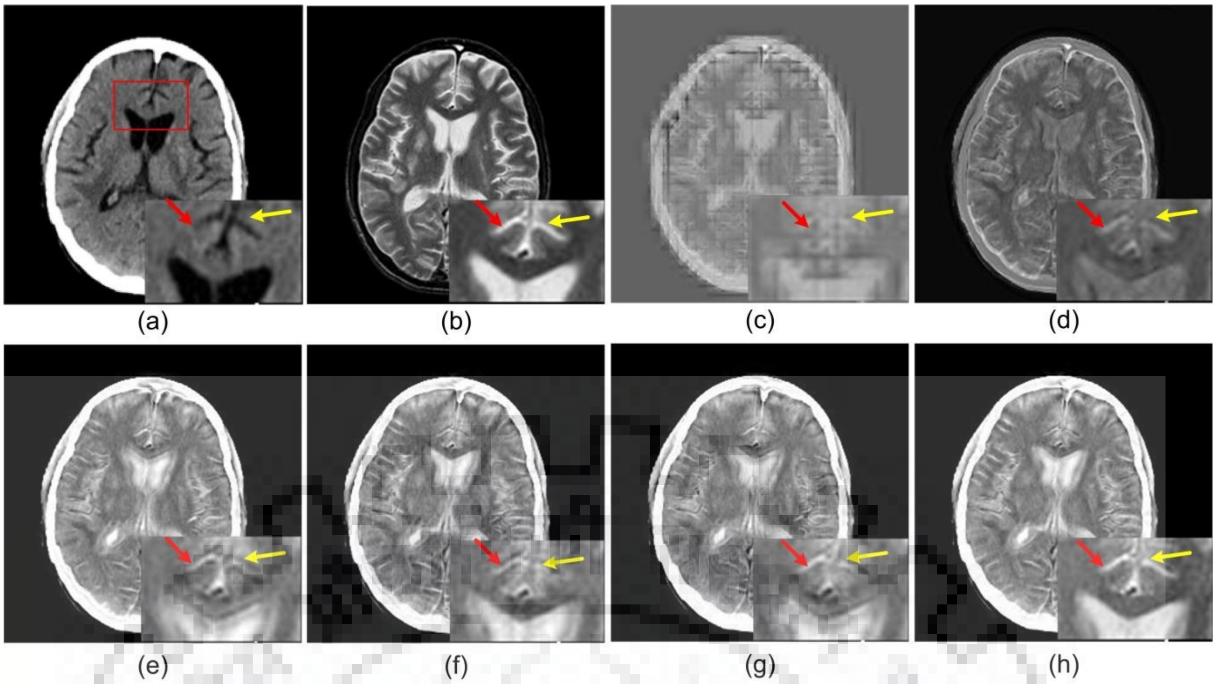


Figure 2.20 Comparison between visual performance evaluated by the fusion approaches for set-2 of (a) CT image and (b) MR image, using (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Proposed C-MIF approach

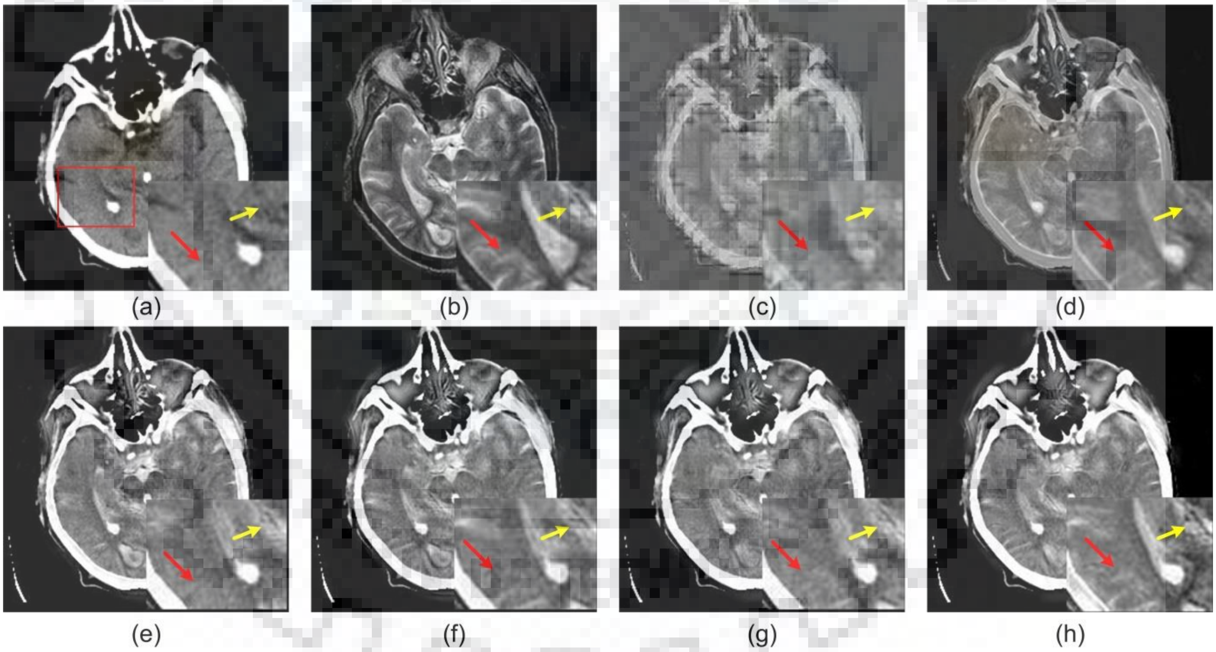


Figure 2.21 Comparison between visual performance evaluated by the fusion approaches for set-3 of (a) CT image and (b) MR image, using (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Proposed C-MIF approach

For validating the subjective observations, some quantitative analysis has also been done and for such analysis, Table 2.17 to Table 2.21 present all aforesaid performance measures that are evaluated for the image pairs processed by all fusion methods considered here. From the measures indicated in the Table 2.17 and Table 2.18, it is easily understood that the proposed C-MIF approach carried more informative content in the fused image by computing higher values of the En and MI as compared to the other method 1-5, respectively. This is also shown in Table 2.19 to Table 2.21 that the proposed C-MIF approach is able to retain the contrast and edge detail information by producing higher STD, SF and XEI values, respectively. Furthermore, Table 2.22 provides a support to all the qualitative and quantitative results by indicating the averaged fusion performance measures in terms of mean and standard deviation. From the averaged results mentioned in Table 2.22, the following points are given below,

(a) The proposed C-MIF approach gets higher En, MI and SF values by approx. 68.14%, 5.48 and 36% from the reference CT image and 32.8%, 49.13% and 13.19% of the MR images, respectively. It means that the fused images produced by the proposed C-MIF approach provides much more information as compared to the reference CT MR images.

(b) The proposed C-MIF approach also achieved higher En values by approx. 10.14%, 9.5%, 7.17%, 5.46% and 3.3% in comparison to the method 1 to 5, respectively. In the same fashion, the proposed C-MIF approach has produced larger MI values by 30.6%, 19.32%, 5.01%, 4.11% and 2.89% as compared to the method 1 to 5, respectively which is also larger than the reference CT-MR images. It indicates more relative information than reference images and other fused images obtained by the other approaches.

(c) Besides this, higher STD and SF values provided using the proposed C-MIF approach largely by 39.5%, 41.57%, 1.85%, 1.53% and 1.2 % and 17.62%, 19.9%, 13.78%, 10.78% and 5.27%, respectively in comparison to the method 1 to 5, respectively which refers to better contrast and clarity level found in the fused images.

(d) In terms of the preservation of the edge detail present in the source CT and MR images, the proposed C-MIF approach shows its superiority again by outperforming the other approaches and provides approx. 4.65%-13.3% and 16.6%-51% larger edge index values than the NSST and NSCT based approaches as discussed above.

Table 2.17 Comparative investigations of entropy (En) values computed by the fusion approaches applied on the reference image pairs.

Fusion Methods	Image dataset pair								
	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Method 1	4.7939	4.8940	5.2507	4.1145	4.8650	5.2127	4.5105	4.4542	4.0108
Method 2	4.8013	4.8601	5.2960	4.2301	4.7970	5.1201	4.6494	4.5035	4.0645
Method 3	4.8604	4.9361	5.3184	4.3675	4.9108	5.2727	4.7240	4.6935	4.1897
Method 4	4.8970	4.9914	5.3486	4.4379	4.9528	5.2989	4.7826	4.7247	4.5431
Method 5	4.9426	5.0335	5.4243	4.6794	4.9927	5.3430	4.9018	4.8078	4.7696
C-MIF	4.9987	5.1623	5.5147	4.8542	5.2241	5.5274	5.1422	4.9894	4.9627

Table 2.18 Comparative investigations of mutual information (MI) values computed by the fusion approaches applied on the reference image pairs.

Fusion Methods	Image dataset pair								
	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Method 1	2.3288	2.524	2.1771	2.6323	2.3399	2.5634	2.4310	2.3354	2.4357
Method 2	2.5628	2.7564	2.4050	2.7863	2.5112	3.0783	2.6086	2.4772	2.5961
Method 3	2.8607	3.0489	2.8223	3.1358	2.9532	3.1384	3.0406	3.0389	3.0390
Method 4	2.8813	3.0644	2.8432	3.1570	2.9722	3.2226	3.0502	3.0310	3.0913
Method 5	2.8933	3.0853	2.8509	3.2026	2.9886	3.2956	3.0973	3.0724	3.1514
C-MIF	2.9347	3.2654	2.8824	3.2456	3.0874	3.3942	3.1847	3.2126	3.2275

Table 2.19 Comparative investigations of standard deviation (STD) values computed by the fusion approaches applied on the reference image pairs.

Fusion Methods	Image dataset pair								
	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Method 1	62.674	59.75	57.008	62.414	65.749	65.380	61.429	66.890	69.762
Method 2	60.324	57.183	60.234	60.492	63.178	63.842	59.097	65.102	68.909
Method 3	84.363	81.587	83.368	85.631	89.489	86.275	84.361	91.667	95.402
Method 4	84.659	81.737	83.492	85.948	89.653	86.695	84.433	92.174	95.789
Method 5	85.033	82.035	83.814	86.153	89.908	87.186	84.852	92.686	96.160
C-MIF	85.958	82.784	84.887	87.252	91.124	89.602	85.262	93.019	96.684

Table 2.20 Comparative investigations of spatial frequency (SF) values computed by the fusion approaches applied on the reference image pairs.

Fusion Methods	Image dataset pair								
	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Method 1	5.8714	6.6219	7.1815	5.5684	6.5576	5.1216	6.0779	5.5161	6.3112
Method 2	6.0536	6.2705	7.1192	5.3852	6.2101	4.8754	6.2537	5.6107	5.8846
Method 3	6.2186	6.6242	7.3292	5.5829	6.6292	5.4210	6.7005	5.8879	6.2845
Method 4	6.4386	6.6959	7.6077	5.7063	6.8515	5.5828	6.8568	6.1315	6.3453
Method 5	6.5816	6.8577	8.8396	5.8769	7.7340	5.7153	6.9213	6.2712	6.4666
C-MIF	6.6024	6.9098	8.8542	8.9248	7.7845	5.6842	6.9684	6.2811	6.4818

Table 2.21 Comparative investigations of edge index (XEI) values computed by the fusion approaches applied on the reference image pairs.

Fusion Methods	Image dataset pair								
	# 1	# 2	# 3	# 4	# 5	# 6	# 7	# 8	# 9
Method 1	0.2929	0.2586	0.2502	0.2669	0.2405	0.2769	0.2395	0.2592	0.2758
Method 2	0.3930	0.3460	0.3401	0.3456	0.3328	0.3784	0.3118	0.3443	0.4086
Method 3	0.5772	0.5088	0.4866	0.4975	0.4729	0.4375	0.4203	0.4234	0.4825
Method 4	0.5797	0.5134	0.5122	0.5112	0.4919	0.4494	0.4244	0.4261	0.5231
Method 5	0.5930	0.5496	0.5255	0.5514	0.5188	0.4685	0.4969	0.5002	0.5970
C-MIF	0.6218	0.5682	0.5346	0.5878	0.5584	0.4968	0.5086	0.5215	0.6224

Table 2.22 Averaged performance analysis using the proposed C-MIF approach and others fusion methods

Methods	Performance Measures				
	En	MI	STD	SF	XEI
Source CT	3.0647 ± 0.3035	-	83.9079 ± 5.4953	5.2680 ± 0.6784	-
Source MR	3.8803 ± 0.3366	-	59.3509 ± 4.9843	6.3309 ± 0.8517	-
Method 1	4.6785 ± 0.4401	2.4186 ± 0.1404	63.4506 ± 3.8991	6.0920 ± 0.6451	0.2623 ± 0.0177
Method 2	4.7057 ± 0.3802	2.6478 ± 0.2069	62.5210 ± 3.5133	5.9763 ± 0.6386	0.3694 ± 0.0297
Method 3	4.8081 ± 0.3713	3.0086 ± 0.1102	86.9047 ± 4.4341	6.2976 ± 0.6017	0.4785 ± 0.0490
Method 4	4.8863 ± 0.3073	3.0348 ± 0.1218	87.1756 ± 4.5285	6.4685 ± 0.6267	0.4924 ± 0.0508
Method 5	4.9883 ± 0.2508	3.0708 ± 0.1423	87.5363 ± 4.5573	6.8071 ± 0.9693	0.5334 ± 0.0435
Proposed C-MIF	**5.1529** ± 0.2378	**3.1594** ± 0.1637	**88.5082** ± 4.4684	**7.1657** ± 1.1289	**0.5578** ± 0.0465

2.9.4 Experiment 4: Comparative evaluation of the proposed approaches and others

This section presents a further investigation of the proposed and several state-of-the-art methods applied to another CT-MR image pair shown in Figure 2.22. In addition to all the discussion mentioned above, another experiment has been performed on the CT-MR dataset shown in Figure 2.22 that is commonly used to compare the fusion performance. This experiment also adds an impact on the superiority of the proposed approaches over the other fusion methods. The performance measures are mentioned in Table 2.23 that indicates a detailed comparison between several fusion approaches and proposed approaches in terms of En, and STD values. Based on the results obtained for the second dataset, it is clearly visualized that the proposed approaches prove its superiority and outperform the others by providing higher En and STD values. Furthermore, the proposed approaches show better subjective and objective fusion results as compared to all considered existing fusion approaches.

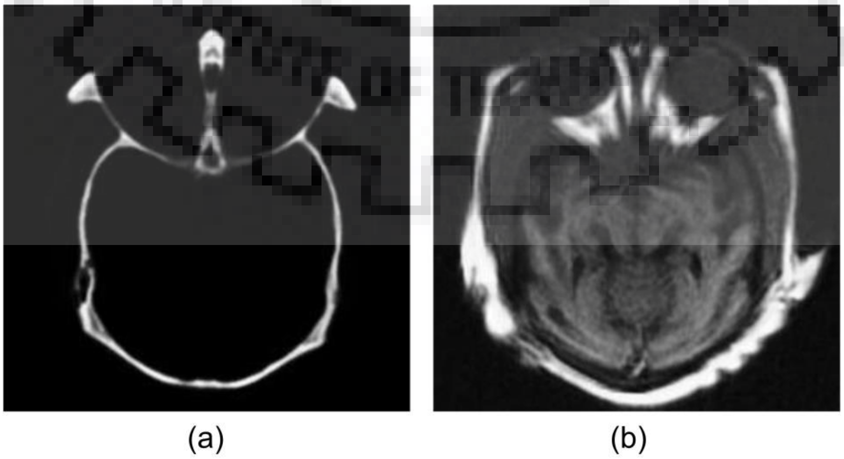

(a) (b)

Figure 2.22 Input reference (a) CT and (b) MR image to evaluate the fusion performance

Table 2.23 A detailed comparison of the performance measures obtained by several fusion approaches applied on the source images shown in Figure 2.22

Fusion Methods	En	STD	Fusion Methods	En	STD
Method [34, 60]	4.982	33.65	Method (PCA) [171]	4.613	-
Method [34, 198]	6.729	57.97	Method (SHARP) [164]	4.851	-
Method [115, 134]	-	54.15	Method (SWT)	5.031	-
Method [115]	-	64.70	Method (DWT) [37]	5.280	-
Method [34, 195]	6.387	53.82	Method (DTCWT) [85]	5.479	-
Method [154]	5.990	32.90	Method (LWT) [186]	5.439	-
Method [92, 154]	5.960	32.55	Method (MWT) [174]	5.2622	-
Method [173]	-	34.85	Method [70]	6.730	60.32
Method [189]	-	20.89	Method [34, 145]	6.771	59.83
Method [176]	6.192	60.02	Method [34]	6.767	59.85
Method [19]	4.227	60.77	Method [195]	6.387	53.82
Method [8]	5.005	65.66	Method [50]	6.780	60.02
Method [86, 92]	6.096	41.56	Proposed NSST-MIF	6.835	62.17
Method [92]	6.199	40.56	Proposed DRT-MIF	6.848	65.29
Method [121]	6.065	40.22	Proposed C-MIF	6.949	65.65

2.10 Summary

In this chapter, three different medical image fusion approaches are proposed using the different decomposition methods (NSST, DRT and a cascaded farmework of DRT and NSST) and PCNN model named as NSST-MIF, DRT-MIF and C-MIF. In the proposed methods, the features of the NSST and DRT are used that provide multiscale and direction analysis of the reference images and help to extract more edge detail information from the source images. Moreover, the NSST and DRT also improve the visualization of the fuse images by reflecting the higher order singularities. The PCNN model whose feeding input is motivated by the NSML and NMSF is able to improve the multimodal medical image analysis by increasing the precision and the performance of diagnosis based on computational methods. Several experiments have been performed for the analysis of the performance of all three proposed fusion methods and also compared to other existing fusion approaches. From the experimental results, it is observed that the fusion rules in the proposed DRT-MIF and C-MIF approaches provide more activities, clarity levels and diagnostic details present in the source images. However, the proposed approaches take more time as compared to wavelet based fusion techniques, but it is acceptable at the cost of improved quality of fused images. The proposed C-MIF approach provides improved performance as compared to the others by preserving more diagnostic information in reference to the source CT and MR images, alongwith the better visual quality of the fused images.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 3: MULTIMODAL MEDICAL IMAGE FUSION BASED ON ADAPTIVE PCNN IN NSST DOMAIN

This chapter presents a multimodal medical image fusion for anatomical-functional (MR-SPECT) images and anatomical-anatomical (CT-MR) images. In this chapter, an improved fusion framework is proposed that utilizes the features of both the NSST and PCNN with adaptive linking strength parameter based on the local visibility of the image. In the proposed fusion method, two different fusion rules are applied to low and high frequency NSST coefficients, separately. The methodologies used to present the fusion method are also discussed in this chapter. The performance of the proposed approach is investigated for both the anatomical-anatomical and anatomical-functional images based on the quantitative and qualitative manner and detailed comparative analysis has also been done and presented in the results section of this chapter.

3.1 Introduction

Nowadays, multimodal medical imaging technology is considered by the radiologist as a prominent solution for the detection and prognosis of many severe neurological disorders and accidental injuries. Based on the types of imaging modality to be fused, multimodality image fusion can be categorized into two different fusion modalities. The first category is the fusion of anatomical details with functional information and the second one is the anatomical with anatomical information fusion. The main difference between anatomical and functional imaging modality is that the anatomical imaging reveals hard and soft tissue structural information in grayscale with higher resolution, while functional imaging modality like SPECT provides the functional information with lower resolution in a pseudo color format. Moreover, it has been analyzed and emphasized in previous chapters that the image fusion is an important process in image processing, but it becomes more important in multimodal medical image fusion such that the fusion of anatomical and functional information can be done without introducing any structural distortion and altering the functional contents which provides a complete pictorial view of the tissue in a single image. It provides a meaningful quantifiable interpretation which is most suitable for clinical diagnosis by providing the exact location and orientation of the defected tissues [29, 49]. From the comparative analysis presented in Chapter 2, it has been analyzed that the NSST decomposition has several advantages over the other multiscale decomposition methods. Moreover, the NSST has multiscale and multidirectional property that also helps to retain more detail information presented in the source images [58]. Besides this, the NSST has some other important properties [43] such as 1) inherently sparse approximation model of the input, 2) well localized (trapezoid shape tilling in frequency space with increased directional coefficients by 2^{j+1} factor on each scale), 3) less Pseudo-Gibbs type residual error in approximation, and 4) computationally more efficient (inverse transform needs only a summation of shearing filter coefficients instead of the inversion of the directional filter bank as required in NSCT).

In the presented work, the adaptive pulse coupled neural network (ADPCNN) model is applied to fuse the low frequency (lf) NSST decomposed coefficients rather than the conventional PCNN model with fixed linking strength parameter as discussed in Chapter 2. In ADPCNN model, linking strength parameter is estimated based on the local visibility of the image by inspiring the human visual system (HVS) model and adaptively varies to achieve the high internal activity from those particular regions of source images that have significant features than the low activity regions. Moreover, to fuse the lf NSST subband, the novel sum modified Laplacian (NSML) based focus feature of the source images is utilized to motivate the adaptive PCNN model. This local windowing based fusion rule can adaptively select the brighter region with higher weights and less important regions with minimal weights corresponding to its neighboring pixels. In order to extract optimal texture feature with broad spectral information, a local log Gabor energy (LLGE) based fusion is employed to fuse the high frequency (hf) NSST coefficients that helps to provide more robust fusion and clean structural details.

The remaining part of this chapter is structured as follows. The next two sections describe the methodology used to implement the proposed adaptive PCNN based medical image fusion (ADP-MIF) approach. The proposed approach and its implementation steps are presented in the subsequent section. Furthermore, the different experimentation and their results are discussed in the next section of the chapter

3.2 Adaptive Pulse Coupled Neural Network

The PCNN is the best model of the biologically inspired feedback neural network (BIFNN) [2, 44, 108] which is utilized in several image processing applications such as segmentation, noise removal, fusion, etc. [34, 36, 38, 152, 214]. In PCNN, each pixel available in the input data is attached with an individual neuron which is also associated with its surrounding neurons. This framework does not need to train, but has an ability to get significant knowledge from a typical background. The PCNN neuron has three different segments as receptive, linking and a pulse generator as illustrated in Figure 2.5 of Chapter 2. In the proposed fusion framework, an improved adaptive PCNN (ADPCNN) model is used and given as [34, 203],

$$\left.\begin{aligned}
&F_{i,j}[n] = S_{i,j} \\
&L_{i,j}[n] = e^{-\alpha_L} L_{i,j}[n-1] + V_L \textstyle\sum_{k,l} W_{i,j,m,n} Y_{k,l}[n-1] \\
&U_{i,j}[n] = F_{i,j}[n]\big(1 + \beta L_{i,j}[n]\big) \\
&T_{i,j}[n] = e^{-\alpha_T} T_{i,j}[n-1] + V_T Y_{i,j}[n-1] \\
&Y_{i,j}[n] = \begin{cases} 1, & U_{i,j}[n] > T_{i,j}[n] \\ 0, & otherwise \end{cases}
\end{aligned}\right\} \qquad (3.1)$$

where (i,j) refer as pixel location, $S_{i,j}$ is input applied externally, $U_{i,j}$, $Y_{i,j}$ are presented as activities related to internal and external state of the neuron. $T_{i,j}$ is a threshold value. The

ADPCNN has feeding and linking input as $F_{i,j}$ and $L_{i,j}$, respectively with parameters α_{L} and α_{T} as delay constants and n as iteration. The parameter (β) refers to a linking strength parameter. To achieve the high internal activity from those particular regions of source images that have significant features than low activity regions, there is a need to get an adaptive linking strength parameter for all neurons in the PCNN instead of fixed one. The parameter (β) regulates the linking input and generates the total internal pulse activity when summed up with the feeding input of the PCNN as given in Eq. (3.1). From the human visual system (HVS) model, the visual threshold and contrast can be imitated as nonlinear system characteristics which decide visually significant or minor feature pixels with respect to its neighbors. Hence, instead of taking fixed linking strength parameter, it is intuitive to use variable parameter to modulate the total internal activity pulse for significant feature regions and suppress for insignificant feature regions. Thus, the linking strength parameter (β) is estimated using local visibility of the image inspired by the human visual system as given below,

$$\mu\left(V_{i,j}^{I}\right) = {}^{1}\!\Big/\!\left[1 - exp\left(\overline{V_{i,j}^{I}} - V_{i,j}^{I}\right)\right] \tag{3.2}$$

where the mathematical formulation of local visibility of an image (I) is given by [17],

$$V_{i,j}^{I} = \begin{cases} \frac{1}{r \times c} \sum_{i=1}^{r} \sum_{j=1}^{c} (m_I)^{-(1+v)} \left|I_{i,j} - m_I\right| & m_I \neq 0 \\ I_{i,j} & m_I = 0 \end{cases} \tag{3.3}$$

where r and c refer to the number of pixels in row and column, m_I is the average value of the pixels in the image $I_{i,j}$ at $(i,j)^{th}$ pixel location. The visual constant parameter (v) is chosen within the range of 0.6 and 0.7. Image component would be more important with larger $V_{i,j}^{I}$ and $\overline{V_{i,j}^{I}}$ represented as the mean of the local visibility. Moreover, the feeding input is not provided as a conventional neural model. The novel sum modified modified Laplacian (NSML) feature of the decomposed NSST lf coefficient is computed and NSML motivated feeding input is applied to the ADPCNN model for implementing the proposed fusion framework. The NSML is used as the feature input in the receptive field and decides the total internal activity with respect to the neighboring pixels for fusion at each location.

3.3 Log Gabor Filtering in NSST Domain

Gabor filters are mainly utilized for extracting the image features [26, 199, 200, 206]. The traditional Gabor filtering approach has the limitation of spatial localization and evenly, it is also not able to suppress the dc components of the filtering response. Thus, log Gabor filters are applied to integrate all the detailed components present in hf NSST subimage coefficients, simultaneously as presented in [200, 206] to detect the focused region. Two dimensional Log Gabor filter is formulated in polar coordinate systems and given by the mathematical formulation as,

$$g\left(\omega^{c},\theta^{d}\right)=\frac{exp\left(-0.5\left(ln\left(\frac{\omega^{c}}{\omega_{0}^{c}}\right)\right)^{2}/\left(ln\left(\frac{\varkappa}{\omega_{0}^{c}}\right)\right)^{2}\right)}{exp\left(0.5\left(\theta^{d}-\theta_{0}^{d}\right)^{2}/\left(\varkappa_{\theta^{d}}\right)^{2}\right)} \tag{3.4}$$

where ω_0^c and θ_0^d refer to the center frequency and orientation direction, respectively. $\varkappa$ and $\varkappa_{\theta^d}$ are the parameters used to compute radial and orientation bandwidth given as B_r and B_o [200].

$$\left.\begin{aligned} B_r &= 2\sqrt{2/ln2}\times\left|\ln\left(\frac{\varkappa}{\omega_0^c}\right)\right| \\ B_o &= 2\varkappa_{\theta^d}\sqrt{2ln2} \end{aligned}\right\} \tag{3.5}$$

3.4 Proposed NSST Domain Image Fusion Approach Based on Adaptive PCNN

This section provides the complete implementation steps with detailed mathematical formulation involved in the proposed medical image fusion using the adaptive PCNN in NSST domain (ADP-MIF). In the proposed fusion framework, NSST decomposition is utilized to preserve the most important diagnostic information. After NSST decomposition of source CT/MR/SPECT images, one lf and several hf subband components are obtained. In the next stage, the NSML and local log Gabor energy (LLGE) are computed to fuse the lf and hf NSST decomposed subbands, respectively. Additionally, for the lf NSST subband fusion, NSML is applied to motivate the ADPCNN model as an external input which reflects the more amount of informative content related to the contours and boundaries of multiple objects exist in both the CT-MR/MR-SPECT images. The salient implementation process included in the development of the proposed ADP-MIF approach is presented in Figure 3.1.

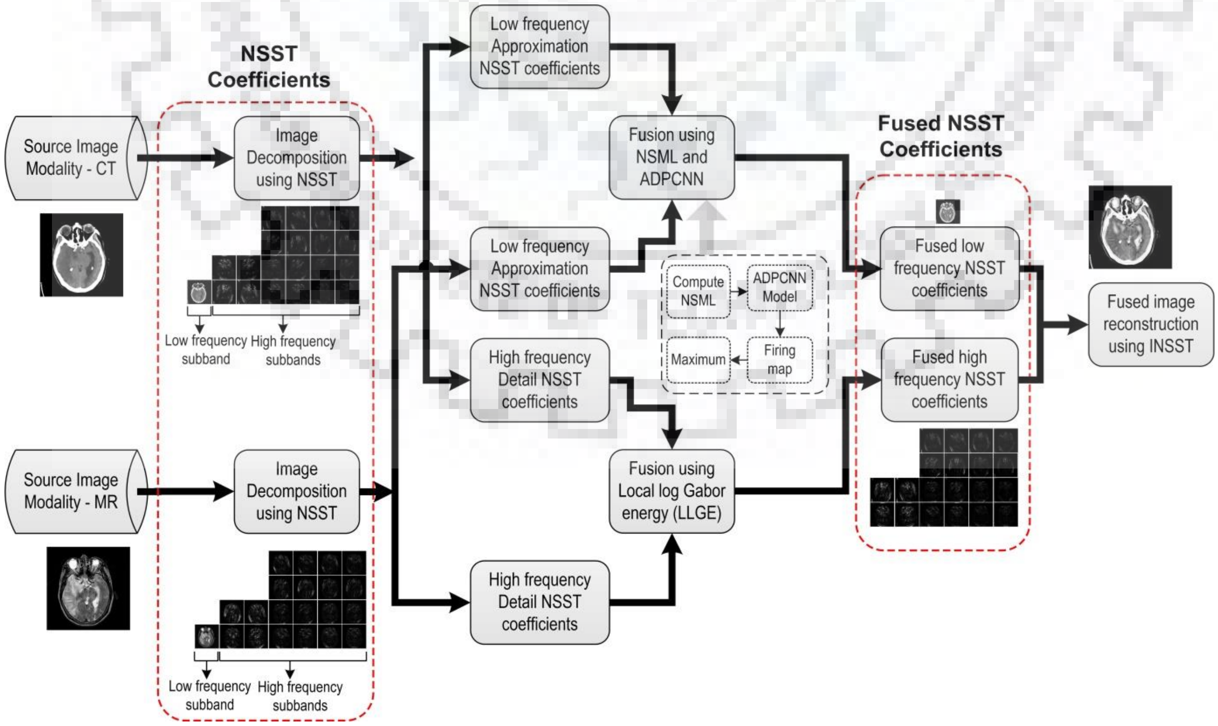


Figure 3.1 Process flow of the proposed ADP-MIF framework in NSST domain

Let $\boldsymbol{R} = R(i,j)$ and $\boldsymbol{S} = S(i,j)$ be two reference images as CT-MR/MR-SPECT, and the implementation of all aforementioned methods, the ADP-MIF model is implemented as given below,

A. Conversion of RGB color space to $\boldsymbol{l\alpha\beta}$ color space

Covert the RGB model of the functional SPECT image into $l\alpha\beta$ color space using the formulation given below as discussed in [18]. Here, $l\alpha\beta$ color space is considered in the presented work, because if the application is centered on human visual perception as in case of medical imaging which may vary from observer to observer, the lαβ color space would work better and this assumes that the image processing is ideally done by the HVS. The RGB color space is firstly converted to LMS cone space and then convert to $l\alpha\beta$ color space. The LMS shows primary color channels (additive combination of narrowband waves) using their wavelengths (long, medium and short same as the receptors wavelength of the human eye) which are correlated, but not the same, hence does not affect the natural color perception of the scene. While, the HSI model separates the brightness from the chromatic components of the RGB images and these components are fundamentally uncorrelated hence relatively less sensitive to the brightness change which is good for the application like color detection and object tracking.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3.6}$$

The additional skew is suppressed by converting the LMS to logarithmic color space.

$$X = lg\, L\,, Y = lg\, M, Z = lg\, S \tag{3.7}$$

Now, convert the logarithmic color space into $l\alpha\beta$ color space by the given mathematical expression,

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{3.8}$$

where l, α and β refer to the achromatic channel, a chromatic yellow-blue and a red-green channels, respectively.

B. Decomposition of source images using the NSST

Decompose of the reference images (CT and MR) or l part of the SPECT images using NSST having shearing directions [4, 8, 8] into lf and hf sub-bands based on the successive trails from coarser to finer decomposition level.

$$\left[lf_{\boldsymbol{R}}^{NSST}, hf_{\boldsymbol{R}}^{NSST}\right] = NSST\big(R(i,j)\big) \text{ and } \left[lf_{\boldsymbol{S}}^{NSST}, hf_{\boldsymbol{S}}^{NSST}\right] = NSST\big(S(i,j)\big) \tag{3.9}$$

C. Fusion of low frequency NSST coefficients

For fusing the lf components, the following steps are involved as given below.

(a) Firstly, estimate the NSML as follows,

$$NSML_{i,j}^{Z} = \sum_a \sum_b w(a,b).X(i+a, j+b) \tag{3.10}$$

$$X(i,j) = \begin{bmatrix} \left|2lf_Z^{NSST}(i,j) - lf_Z^{NSST}(i-1,j) - lf_Z^{NSST}(i+1,j)\right| + \\ \left|2lf_Z^{NSST}(i,j) - lf_Z^{NSST}(i,j-1) - lf_Z^{NSST}(i,j+1)\right| \end{bmatrix} \tag{3.11}$$

$$w(a,b) = \begin{bmatrix} 1/15 & 2/15 & 1/15 \\ 2/15 & 3/15 & 2/15 \\ 1/15 & 2/15 & 1/15 \end{bmatrix} \tag{3.12}$$

where Z depicts the original CT, MR or SPECT image either R or S and $w(a,b)$ is a 3×3 template used in the presented work.

(b) Compute the adaptive linking strength parameter $\beta_{i,j}$ of lf subband coefficients as mentioned above.

$$\beta_{i,j} = \mu\left(V_Z^{NSST}\right) = {}^{1}\Big/_{\left[1 - exp\left(\overline{V_Z^{NSST}} - V_Z^{NSST}\right)\right]} \tag{3.13}$$

$$V_Z^{NSST} = \begin{cases} \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{1}{r \times c} (m_I)^{-(1+v)} \left|lf_Z^{NSST} - m_I\right|; & m_I \neq 0 \\ lf_Z^{NSST}; & m_I = 0 \end{cases} \tag{3.14}$$

(c) Apply the NSML of the lf components to activate the ADPCNN neural model with the adaptively computed the linking strength and boost the neuron pulse by the given mathematical formulations,

$$\left.\begin{aligned} & F_{i,j}^{Z}[n] = NSML_{i,j}^{Z} \\ & L_{i,j}^{Z}[n] = e^{-\alpha_L} L_{i,j}^{Z}[n-1] + V_L \sum_{k,l} W_{i,j,k,l}^{Z} Y_{k,l}^{Z}[n-1] \\ & U_{i,j}^{Z}[n] = F_{i,j}^{Z}[n]\left(1 + \beta_{i,j} L_{i,j}^{Z}[n]\right) \\ & T_{i,j}^{Z}[n] = e^{-\alpha_T} T_{i,j}^{Z}[n-1] + V_T Y_{i,j}^{Z}[n-1] \\ & Y_{i,j}^{Z}[n] = \begin{cases} 1, \ U_{i,j}^{Z} > T_{i,j}^{Z} \\ 0, \ otherwise \end{cases} \end{aligned}\right\} \tag{3.15}$$

(d) Compute the firing times (sum of the $Y_{i,j}^{Z} = 1$for $U_{i,j}^{Z} > T_{i,j}^{Z}$) in n iterations,

$$t_{i,j}^{Z}[n] = t_{i,j}^{Z}[n-1] + Y_{i,j}^{Z}[n] \tag{3.16}$$

(e) If $n = n_{max}$, process stops, fuse the lf components based on the fusion rule by comparing the firing time evaluated in the previous step as follows,

$$lf_F^{NSST} = \begin{cases} lf_{\boldsymbol{R}}^{NSST}; & if\, t_{i,j}^{\boldsymbol{R}}[n_{max}] \geq t_{i,j}^{\boldsymbol{S}}[n_{max}] \\ lf_{\boldsymbol{S}}^{NSST}; & if\, t_{i,j}^{\boldsymbol{R}}[n_{max}] < t_{i,j}^{\boldsymbol{S}}[n_{max}] \end{cases} \tag{3.17}$$

D. Fusion of high frequency NSST coefficients

(a) For fusing the hf subband, compute the convolution between all the NSST hf components and log Gabor filter as discussed above,

$$G_{i,j}^{k,l} = hf_Z^{NSST} * g_{i,j}^{k,l} \tag{3.18}$$

where $g_{i,j}^{k,l}$ refer to the log Gabor filter coefficients in k and l number of scale and directions, respectively.

(b) Compute the local log Gabor energy (LLLG) in the NSST domain at every pixel of hf coefficients using the formulation,

$$LLGE_{i,j} = \frac{1}{(1+2a)\times(1+2b)} \sum_{i=-a}^{a} \sum_{j=-b}^{b} LGE_{i+a,j+b} \tag{3.19}$$

where $(1+2a)\times(1+2b)$ refers to the template size with $a = 3$ and $b = 3$ and LGE refers to the log Gabor energy in the NSST domain for all decomposed hf coefficients and estimated as,

$$LGE = \sum_{k=1}^{K} \sum_{l=1}^{L} \left(Re\left(G_{i,j}^{k,l}\right)^2 + Im\left(G_{i,j}^{k,l}\right)^2 \right)^{0.5} \tag{3.20}$$

(c) Now, the hf NSST subband components are fused by the fusion rule formulated based on the computation of the LLGE aforementioned.

$$hf_F^{NSST} = \begin{cases} hf_{\boldsymbol{R}}^{NSST}; & if\ LLGE_{i,j}^{\boldsymbol{R}} \geq\ LLGE_{i,j}^{\boldsymbol{S}} \\ hf_{\boldsymbol{S}}^{NSST}; & if\ LLGE_{i,j}^{\boldsymbol{R}} <\ LLGE_{i,j}^{\boldsymbol{S}} \end{cases} \tag{3.21}$$

E. Reconstruct the image using inverse NSST

Apply the inverse NSST on the fused lf and hf components to reconstruct the resultant image, as follows

$$F = NSST^{-1}\left(lf_F^{NSST}, hf_F^{NSST}\right) \tag{3.22}$$

F. Reconstruction of $l\alpha\beta$ color space to RGB color space

For the MR-SPECT images, include α and β computed in step A with the fused image obtained from step E by the given mathematical formulation and finally convert it back to the RGB color format by using equation given below. The schematic of the proposed fusion approach is shown in Figure 3.2 for the MR-SPECT image in which the shaded block of proposed NSST based ADP-MIF approach is considered as shown in Figure 3.1.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} \tag{3.23}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} 10^X \\ 10^Y \\ 10^Z \end{bmatrix} \tag{3.24}$$

3.5 Experimentation

To assess the fusion performance of the proposed ADP-MIF method, a large number of the experiments have been performed and a comparison is made with several existing fusion methods. The fused outcomes of the proposed ADP-MIF and other fusion approaches are evaluated and analyzed for the MR-SPECT neurological images alongwith the CT-MR images in order to get both the anatomical and functional information taken from the source images, respectively. After getting the resultant fused images, both the subjective and objective analysis has been done. The former analysis reflects a detailed comparison of source and fused images in terms of visual appearance. The quantitative analysis of the proposed and existing fusion methods is also done by a set of predefined performance measures opted for verifying the acceptability of visual observation. To assess the fusion performance, several experiments are conducted using different pairs of MR-SPECT and CT-MR images out of which ten MR-SPECT image pairs are presented here to reflect the comparative performance of the proposed ADP-MIF and other approaches. Moreover, four CT-MR image pairs are also presented here to provide strong evidence of the superiority of the proposed ADP-MIF approach. To conduct the experiments, NSST decomposition level $= [2, 3, 3]$ is considered and ADPCNN parameters are set as $\alpha_T = 0.2$, $\alpha_L = 1, V_L = 1$, $V_T = 20$, $W = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ \ 1\right]$ and iterations$= 200$. Two different experiments are discussed as given below,

Experiment 1: To analyze and evaluate the fusion performance of the proposed ADP-MIF approach applied to the anatomical-functional images (MR-SPECT images) and presents a comparative analysis of the proposed and other existing fusion methods, qualitatively and quantitatively.

Experiment 2: To analyze and evaluate the fusion performance of the proposed ADP-MIF approach applied to the anatomical-anatomical images (CT-MR images) and presents a comparative analysis of the proposed and other existing fusion methods, qualitatively and quantitatively.

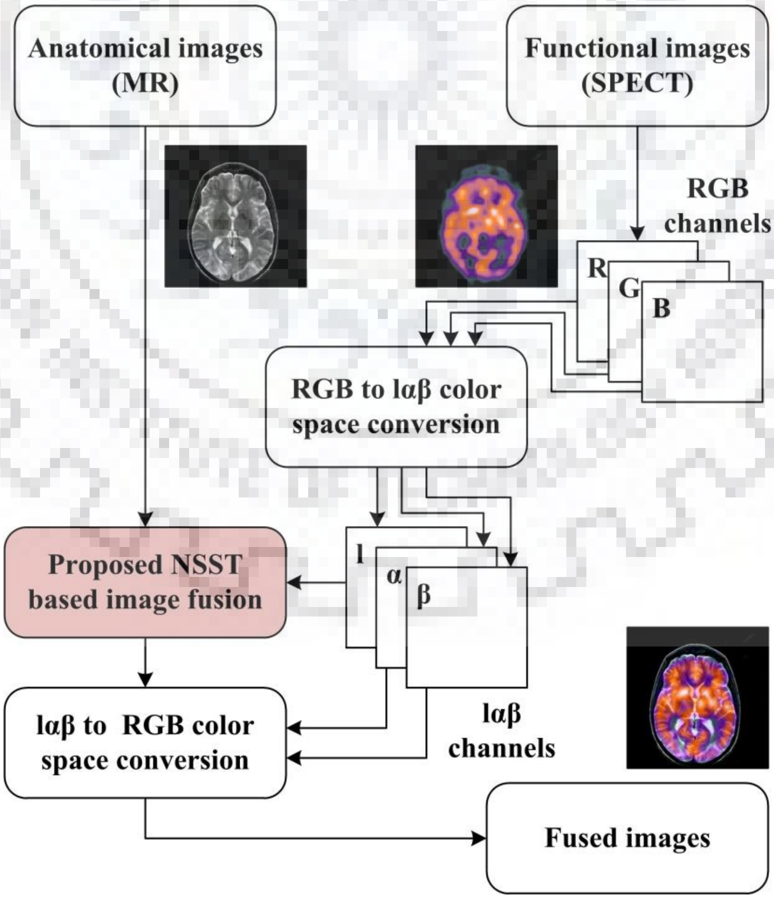


Figure 3.2 Methodology used to fuse MR and SPECT medical images

3.6 Results and Discussions

3.6.1 Experiment 1: Comparative analysis and evaluation of anatomical-functional image fusion results

To conduct such an experiment and analyze the objective fusion performance of the proposed ADP-MIF approach, a large dataset of the MR-SPECT image pairs are utilized, out of which ten pairs of MR and SPECT images are shown in Figure 3.3 (a) and (b), respectively, and the corresponding fused resultant images are shown in Figure 3.3 (c). The proposed approach ADP-MIF formulated by the NSML motivated adaptive PCNN mapping for fusing the lf coefficients and local log Gabor energy based fusion rule for hf subband coefficients provide a better quality of fused images. This shows more information contents and edge information than the source images. Visually, the fused images shown in Figure 3.3 have more information alongwith the edge and spectral information as compared to the source MR and SPECT images. The quantitative analysis is also done for all the fused images listed in Table 3.1.

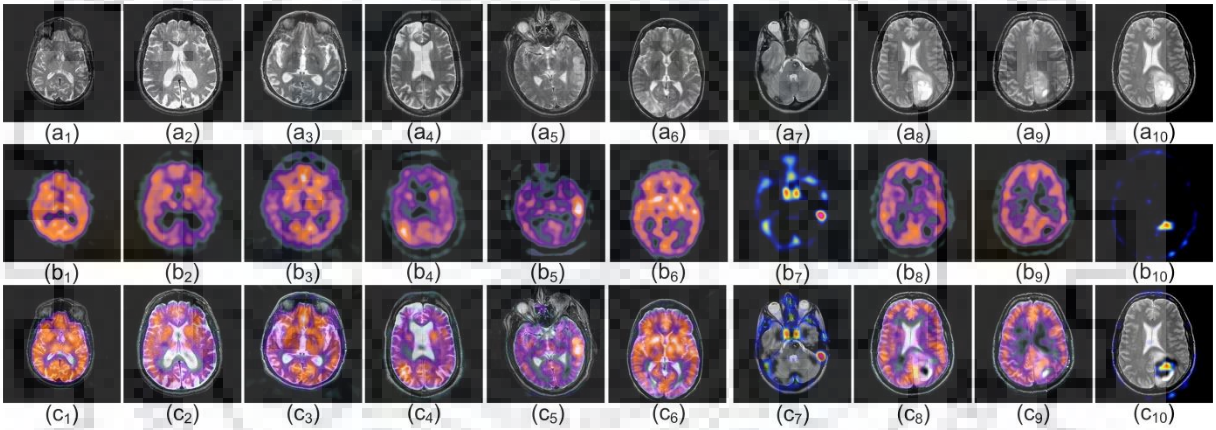


Figure 3.3 (a) Source MR images, (b) SPECT images (c) Fused images obtained by the proposed ADP-MIF approach

Table 3.1 Performance measures obtained by the proposed ADP-MIF approach for the fused images illustrated in Figure 3.3

Image Dataset	Performance Measures					
	En	STD	SF	FMI	IQI	XEI
# 1	4.0628	65.1155	6.8849	0.52065	0.78228	0.66438
# 2	5.7197	87.0779	8.4712	0.52865	0.68537	0.75416
# 3	5.6552	75.8878	7.9455	0.48696	0.64195	0.68312
# 4	5.3051	72.2671	7.5789	0.51231	0.69089	0.71564
# 5	6.0809	62.3435	8.4666	0.50561	0.57534	0.74294
# 6	5.0830	74.2916	7.4773	0.49531	0.72044	0.62447
# 7	4.1757	60.8131	6.770	0.47871	0.74534	0.64242
# 8	5.4344	77.2511	7.5854	0.50871	0.71224	0.66273
# 9	5.1914	65.4189	7.4523	0.49737	0.72694	0.64581
# 10	5.3646	69.1356	7.1312	0.48585	0.71989	0.88506

Furthermore, the superiority of the proposed ADP-MIF approach is investigated by comparing the fusion results shown in Figure 3.4 obtained from the following fusion results.

Method 1: Image fusion approach using discrete cosine harmonic wavelet transform with 4 level decomposition discussed as similar in [79].

Method 2: Fusion method in NSCT domain with decomposition level [2 2 3] with two different fusion rules as discussed in [18].

Method 3: Fusion approach based on the multiresolution singular value decomposition technique as described in [129] by applying the same division of the image with 2×2 blocks and arranging the corresponding block in a 4×1 vector with the similar parameters.

Method 4: Fusion scheme based on the standard deviation in WT domain [137] by partitioning each one of the reference images into 8×8 block.

Method 5: Pixel-level fusion scheme based on the region based segmentation and spectral variance as discussed in [157].

Method 6: Fusion approach based on the hybrid multiscale decomposition with guided image filtering concept as described in [218].

Method 7: Another fusion approach based on the NSST decomposition and PCNN model using sum of direction gradient with the similar parameters as mentioned in [48].

Method 8: Fusion approach developed in NSCT domain as described in [200] by getting the similarity between the previously fused and reference image.

Method 9: Image fusion based on NSCT decomposition where the SML based fusion rule is applied on both hf and lf subband as discussed in [203].

Method 10: The proposed ADP-MIF approach based on the NSST decomposition with NSST decomposition level [2 3 3] and ADPCNN parameters $\alpha_{\mathrm{T}} = 0.2$, $\alpha_L = 1, \mathrm{V_L} = 1$, $\mathrm{V_T} = 20$, $\mathrm{W} = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ \ 1\right]$ and iterations$= 200$. The alpha variable of the ADPCNN model controls the firing time of the neuron and defines the rate of decay of the threshold in iterative series for the pulse generator unit. The larger value of alpha can make the running time faster, whereas the smaller value can make the firing process very slow but performs the pixel’s selection in a precise way.

To evaluate the superiority of the proposed approach, a detailed assessment is done by comparing its performance with the other methods in terms of both the subjective and quantitative analysis. A visual comparison of all ten MR-SPECT brain images processed with all above-said ten approaches is shown in Figure 3.4 (c) to (l), respectively. In addition, the zoom regions of other fused images are also shown in Figure 3.5 and Figure 3.6, to visualize the differences more clearly between the proposed ADP-MIF approach and state of the art approaches. From these visual results, it is clearly visualized that the proposed approach has an ability to incorporate the structural and detailed information acquired from the MR image and functional information from the SPECT brain images. The fused images obtained by the proposed ADP-MIF approach are more bright and clear. This visual comparison is also supported by computing the quantitative measures such as En, STD, SF, FMI, IQI and XEI as discussed in Chapter 1.

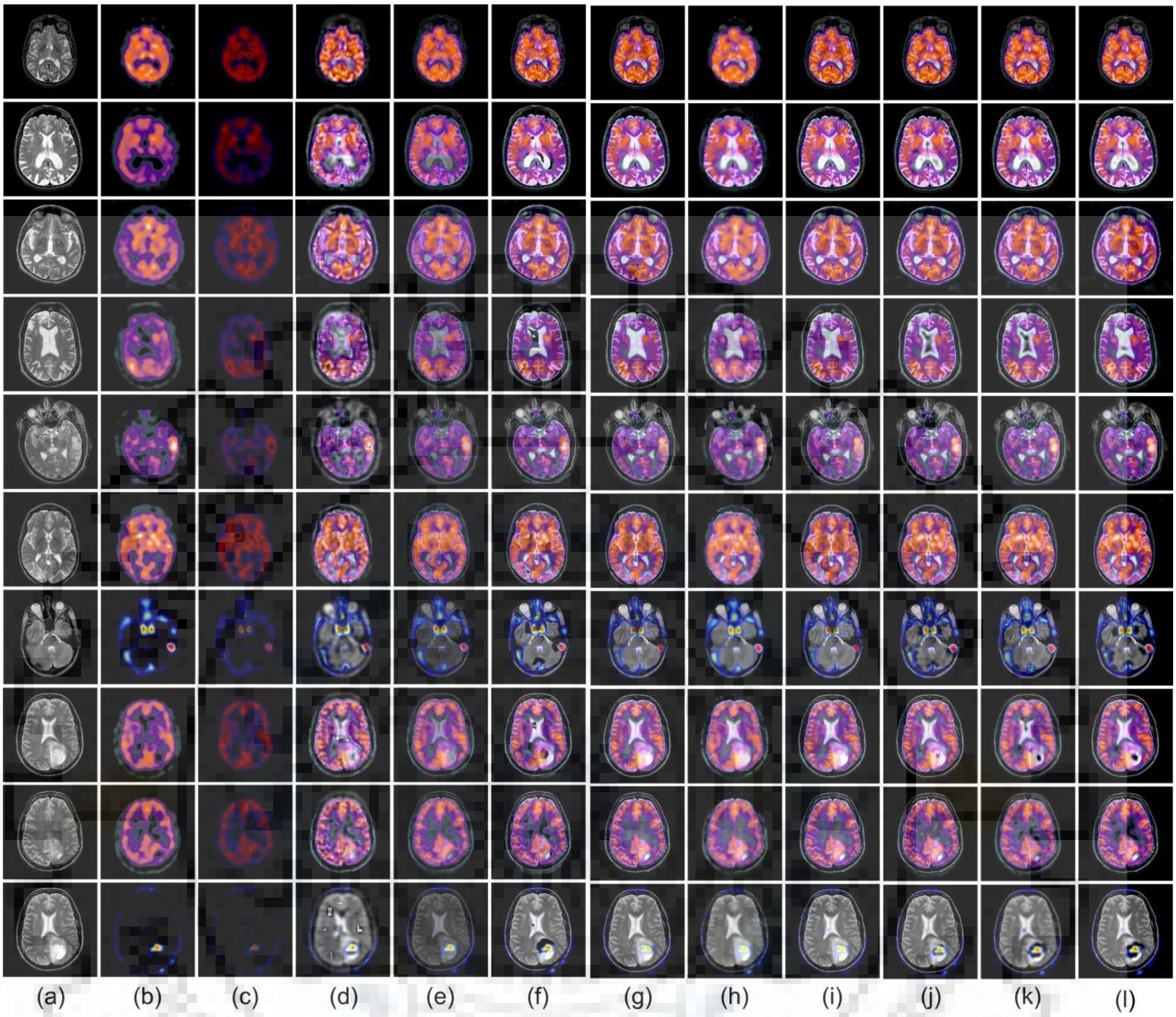


Figure 3.4 Multimodal fusion results of real neurological (a) MR images (b) SPECT image, fused images obtained by the (c) Method 1 (d) Method 2 (e) Method 3 (f) Method 4 (g) Method 5 (h) Method 6 (i) Method 7 (j) Method 8 (k) Method 9 (l) Method 10

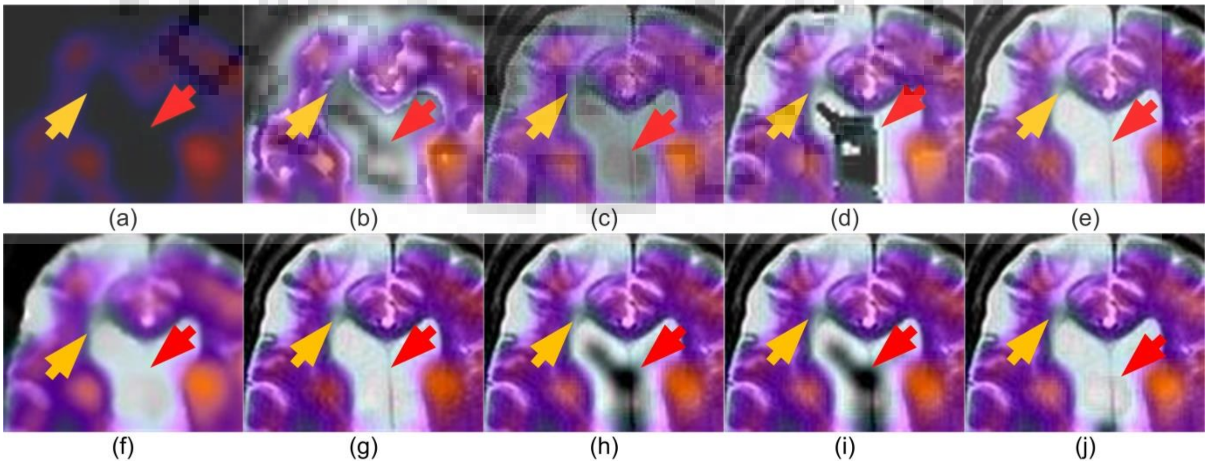


Figure 3.5 Zoomed region of the fusion results of another MR-SPECT image pair-4, fused images obtained by the (a) Method 1 (b) Method 2 (c) Method 3 (d) Method 4 (e) Method 5 (f) Method 6 (g) Method 7 (h) Method 8 (i) Method 9 (j) Method 10

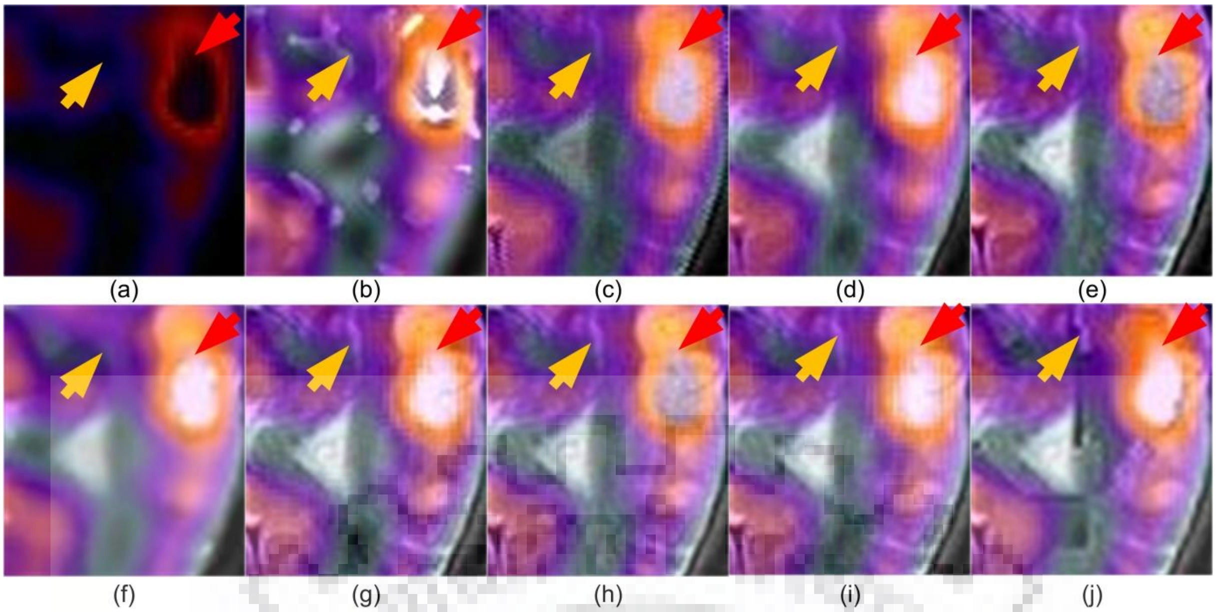


Figure 3.6 Zoomed region of the fusion results of another MR-SPECT image pair-5, fused images obtained by the (a) Method 1 (b) Method 2 (c) Method 3 (d) Method 4 (e) Method 5 (f) Method 6 (g) Method 7 (h) Method 8 (i) Method 9 (j) Method 10

The quantitative results of En, STD, FMI, SF, IQI, and XEI are shown in Table 3.2 to Table 3.7 respectively. From the results mentioned in Table 3.2, it is noticed that the higher En value is obtained by the proposed ADP-MIF approach as compared to the state of the art approaches considered here that indicates the large extent of information available in the resultant fused images. The results of Table 3.3 and Table 3.4 also show the higher STD and SF values computed by the proposed approach. However, for some of the images, method 5 achieves a higher STD value and method 9 provide higher SF value for some cases, but there is not so much difference between the STD and SF value computed by the method 5 and 9 as compared to the proposed approach that also refers to the better contrast and clarity level found in the fused images. Moreover, IQI and XEI values mentioned in Table 3.6 and Table 3.7, respectively, are also higher for all the fused MR-SPECT images obtained by the proposed ADP-MIF approach in comparison to all the other methods which mean that the higher extent of edge information is available in the fused images produced by the proposed ADP-MIF approach. Alongwith all these edge information, the proposed ADP-MIF approach provides higher values of FMI shown in Table 3.5 among the state-of-the-art fusion approaches as mentioned above. From the experimental results, it is observed that the proposed ADP-MIF approach yields superior results in terms of all En, FMI, STD, SF, IQI and XEI measures for the different multimodal image data. Finally, it is concluded that the proposed ADP-MIF approach is able to reproduce the significant visual appearance with the preservation of structural and spectral information content and provides a clear picture of edge details available in any one of the source images.

Table 3.2 Comparative evaluation of entropy (En) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	2.0254	3.5339	3.6953	3.7282	3.4439	3.5606	3.7333	3.4439	3.6138	**4.0628**
# 2	2.6881	4.8561	5.2561	5.078	4.7909	5.4938	4.9631	4.7918	4.8987	**5.7197**
# 3	3.0406	5.1138	5.3453	5.3329	4.7014	5.0917	5.4337	4.7129	5.1718	**5.6552**
# 4	2.2653	4.6860	4.8964	4.8299	4.5878	5.0327	5.0555	4.5983	4.7683	**5.3051**
# 5	1.9941	5.6554	5.6653	5.9643	5.5244	5.9569	5.8069	5.5484	5.6811	**6.0809**
# 6	2.6931	4.3396	4.6467	4.4713	4.2514	4.7047	4.5779	4.2602	4.4072	**5.0830**
# 7	0.6853	3.4875	3.2359	3.4306	3.2591	4.1252	3.7976	3.2608	3.5353	**4.1757**
# 8	2.4429	4.6192	4.8980	4.8357	4.4904	5.1386	4.9056	4.5141	4.6971	**5.4344**
# 9	2.213	4.5028	4.7606	4.6944	4.3316	4.8066	4.8387	4.3769	4.5904	**5.1914**
# 10	0.9791	4.1890	4.0250	4.5860	3.7862	5.0633	4.2289	3.8408	4.1946	**5.3646**

Table 3.3 Comparative evaluation of standard deviation (STD) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	26.5782	63.1479	60.6575	63.3754	63.2687	64.8247	64.8167	63.2686	64.3592	**65.1155**
# 2	22.4354	78.7286	67.7183	85.3506	86.5982	83.1075	86.3611	87.0768	85.8915	**87.0779**
# 3	26.8139	73.0270	66.1248	75.1290	75.6913	74.6549	75.7106	75.6752	75.6264	**75.8878**
# 4	19.5312	65.4292	57.3648	69.1780	71.9444	70.2975	71.9812	71.9412	68.9441	**72.2671**
# 5	17.4453	58.1466	51.1160	61.8237	**62.3481**	60.0039	62.1543	62.0524	62.1164	62.3435
# 6	27.3973	70.7510	65.3198	71.8408	70.7808	73.1026	73.4840	70.7769	72.7192	**74.2916**
# 7	20.2637	51.9536	42.7765	57.1234	56.1113	58.4903	59.4757	56.1109	56.0538	**60.8131**
# 8	22.8211	74.4049	63.384	74.0926	**77.2567**	74.4726	76.1346	76.2006	74.5230	77.2511
# 9	21.1350	64.4275	57.6457	63.2451	64.2506	63.2522	64.2506	65.4121	64.0905	**65.4189**
# 10	10.5608	61.8478	39.6933	65.6362	67.0165	65.8348	67.1324	69.1279	65.1317	**69.1356**

Table 3.4 Comparative evaluation of spatial frequency (SF) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	2.5165	6.5840	5.2886	6.1233	4.4161	6.8171	6.8321	6.8171	6.8344	**6.8849**
# 2	2.6132	8.1903	6.6201	7.6456	5.6249	8.4227	8.450	8.4226	8.462	**8.4712**
# 3	2.8899	7.6582	7.0562	7.2111	5.7014	7.7908	7.7804	7.7718	7.7785	**7.9455**
# 4	2.2964	7.0169	6.5096	6.8198	5.2404	7.3449	7.5024	7.345	7.5129	**7.5789**
# 5	2.4499	7.7021	6.1793	7.2659	5.4627	8.3680	8.4354	8.3679	8.3879	**8.4666**
# 6	3.0283	7.1581	6.6514	6.8999	5.336	7.3409	7.4159	7.3411	7.4079	**7.4773**
# 7	2.4130	6.1584	5.2439	6.1333	5.3138	6.5044	6.6749	6.5044	6.6610	**6.770**
# 8	2.8446	7.2492	6.7894	7.0403	5.3334	7.3861	7.5593	7.3862	**7.5865**	**7.5854**
# 9	2.7619	7.0137	6.4038	6.7468	4.9509	7.2331	7.419	7.2334	7.3659	**7.4523**
# 10	1.4265	6.3370	4.8658	6.3159	4.9113	6.9711	7.0735	6.9703	7.0223	**7.1312**

Table 3.5 Comparative evaluation of feature mutual information (FMI) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	0.27806	0.18787	0.37188	0.4131	0.37002	0.20209	0.37616	0.43551	0.46028	**0.52065**
# 2	0.25626	0.18053	0.34978	0.42226	0.50922	0.19535	0.51047	0.51605	0.51894	**0.52865**
# 3	0.31679	0.21907	0.36864	0.40354	0.42592	0.23986	0.41993	0.40388	0.45493	**0.48696**
# 4	0.28427	0.20079	0.34893	0.4002	0.39373	0.2241	0.3985	0.44404	0.4365	**0.51231**
# 5	0.24118	0.18702	0.31269	0.40616	0.47322	0.18638	0.4771	0.49248	0.48399	**0.50561**
# 6	0.29621	0.20146	0.35451	0.39459	0.36829	0.23992	0.36809	0.39856	0.43704	**0.49531**
# 7	0.28001	0.19795	0.30804	0.36646	0.34962	0.20516	0.34961	0.42024	0.41504	**0.47871**
# 8	0.3004	0.20832	0.33438	0.39268	0.36863	0.23832	0.37299	0.43986	0.41824	**0.50871**
# 9	0.29741	0.20719	0.32852	0.39032	0.35116	0.23893	0.3555	0.42116	0.40054	**0.49737**
# 10	0.25214	0.17615	0.29989	0.38459	0.39114	0.19459	0.39268	0.43824	0.43357	**0.48585**

Table 3.6 Comparative evaluation of image quality index (IQI) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	0.66306	0.26527	0.76704	0.29501	0.29305	0.43285	0.49278	0.26594	0.42713	**0.78228**
# 2	0.46944	0.21601	0.62852	0.35744	0.36082	0.17496	0.56218	0.34031	0.26979	**0.68537**
# 3	0.4646	0.23308	0.62826	0.33708	0.33013	0.47159	0.32448	0.34533	0.29871	**0.64195**
# 4	0.5009	0.25517	0.65983	0.3268	0.34137	0.3984	0.3475	0.33209	0.39856	**0.69089**
# 5	0.36914	0.16686	0.47953	0.43004	0.43713	0.19678	0.43198	0.45504	0.38255	**0.57534**
# 6	0.55534	0.26531	0.71294	0.33584	0.31536	0.49379	0.31437	0.30267	0.44466	**0.72044**
# 7	0.66682	0.29615	0.69477	0.29187	0.30123	0.46715	0.31284	0.26608	0.49199	**0.74534**
# 8	0.51701	0.31071	0.69091	0.33316	0.34388	0.48633	0.35196	0.3858	0.49132	**0.71224**
# 9	0.53865	0.3186	0.70828	0.3356	0.33994	0.54317	0.3534	0.37795	0.51628	**0.72694**
# 10	0.52372	0.23239	0.62463	0.35289	0.3497	0.43319	0.34662	0.41267	0.49534	**0.71989**

Table 3.7 Comparative evaluation of Xydeas edge index (XEI) values obtained by the proposed ADP-MIF with the other fusion schemes for MR-SPECT image pairs shown in Figure 3.4

Image Pairs	Fusion Methods									
	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	Proposed
# 1	0.05215	0.14255	0.38679	0.5417	0.63271	0.20378	0.63797	0.66398	0.64742	**0.66438**
# 2	0.01945	0.10527	0.40506	0.64269	0.73098	0.20857	0.73038	0.75344	0.73748	**0.75416**
# 3	0.0336	0.16384	0.39094	0.61883	0.67603	0.25083	0.68211	0.68305	0.67452	**0.68312**
# 4	0.01923	0.14427	0.37523	0.62359	0.67623	0.25658	0.68119	0.71533	0.69059	**0.71564**
# 5	0.02124	0.13189	0.40263	0.64278	0.71089	0.22041	0.71552	0.74173	0.71922	**0.74294**
# 6	0.05325	0.15714	0.36602	0.53061	0.59914	0.23527	0.60397	0.62438	0.61047	**0.62447**
# 7	0.05522	0.16277	0.37763	0.50097	0.62181	0.32542	0.63383	0.62092	0.61271	**0.64242**
# 8	0.02932	0.17151	0.40636	0.57464	0.63144	0.2397	0.6369	0.66252	0.63988	**0.66273**
# 9	0.02866	0.17087	0.38218	0.57859	0.61903	0.2133	0.6214	0.64548	0.63143	**0.64581**
# 10	0.02415	0.10022	0.4799	0.78219	0.87486	0.2409	0.8825	0.87246	0.8791	**0.88506**

3.6.2 Experiment 2: Comparative analysis and evaluation of anatomical-anatomical image fusion results

To consider the performance of the proposed ADP-MIF approach, several other experiments are performed on the anatomical CT-MR brain images. Four image pairs out of complete CT-MR dataset are shown in Figure 3.7. For presenting the comparative visual performance of the fused images, some of the existing methods developed previously are considered as fusion method (FM-1) [137], FM-2 [157], FM-3 [48], FM-4 [203] (mentioned above as Method 4, Method 5, Method 7 and Method 9), respectively, FM-5 (the C- MIF approach presented in Chapter 2) and the proposed ADP-MIF approach. Their fusion results are shown in Figure 3.8 (a) to (f) respectively. After observing the fused images, it is visualized that the resulting fused images obtained by the proposed ADP-MIF approach have better visual ability in terms of better contrast and edge information which is also verified by the quantitative results correspond to the En, STD, SF and XEI measures. All the evaluated parameters are mentioned in Table 3.8 from which, it is observed that the proposed ADP-MIF method gets higher values compared to all state-of-the art methods considered here.

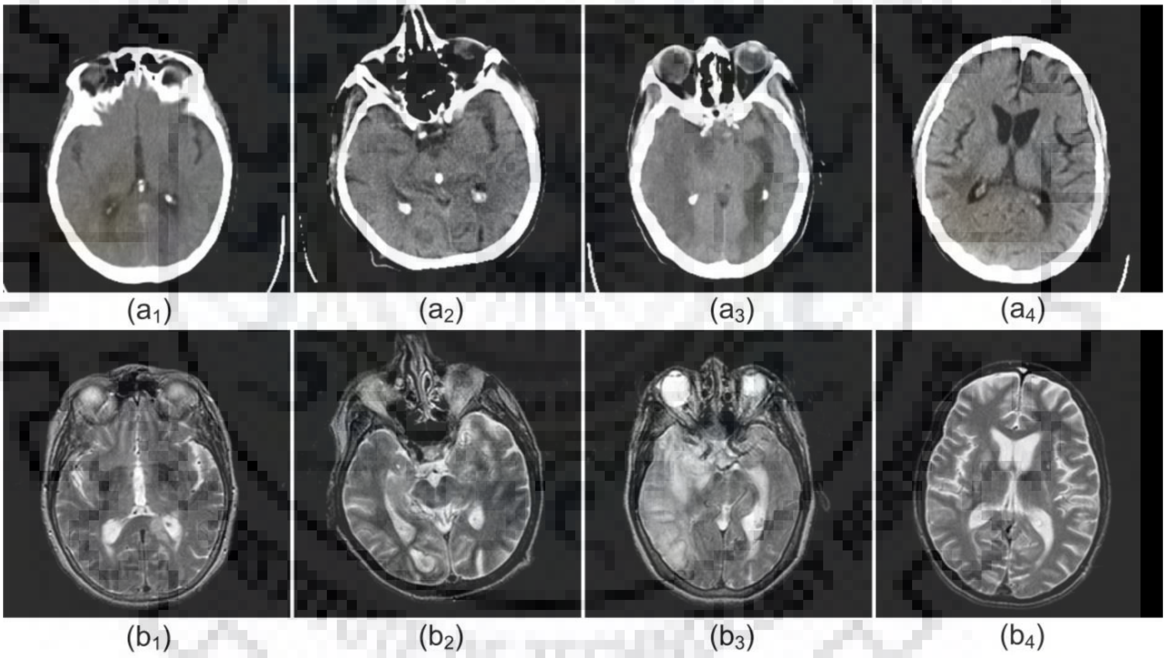


Figure 3.7 (a) Source MR images, (b) SPECT images for evaluating the performance of the proposed ADP-MIF approach

Furthermore, another dataset of CT-MR as shown in Figure 3.9 is considered to compare the fusion performance evaluated by the other several fusion approaches developed recently. After getting the fused images presented in Figure 3.9 (c) to (n), a comparison is made in terms of the En, STD, and MI values computed for each individual fused image obtained by all the fusion methods as shown in Table 3.9. Furthermore, the proposed ADP-MIF approach produces higher En, STD and MI values as compared to the other approaches considered here. Finally, based on all experimental results, it is concluded that the presented ADP-MIF approach has an ability to produce a fused image with better visual quality. In addition, the proposed approach also shows its superiority in terms of quantitative results among all the considered approaches.

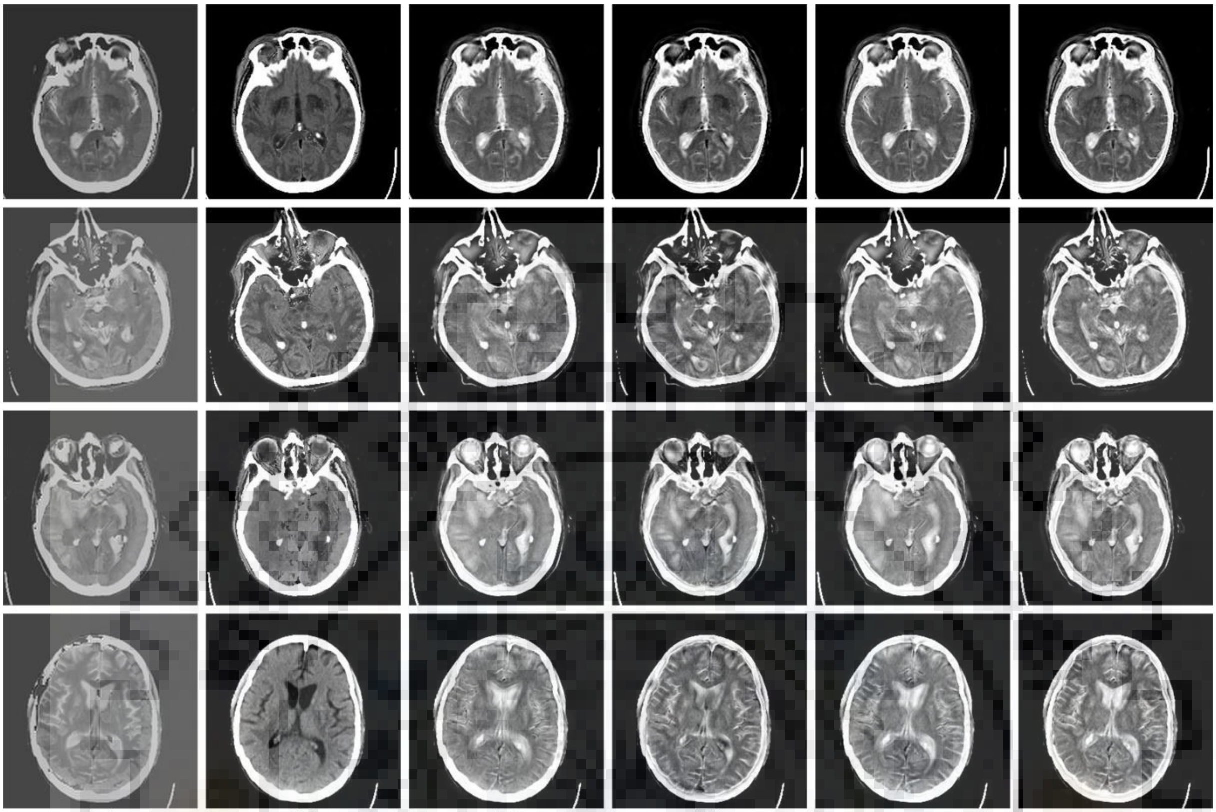

(a) (b) (c) (d) (e) (f)

Figure 3.8 Comparative visual results of fused images obtained by the (a) FM-1 (b) FM-2 (c) FM-3 (d) FM-4 (e) FM-5 (f) proposed ADP-MIF approach

Table 3.8 Comparative analysis of quantitative parameters evaluated for CT-MR image fusion

Image Pairs	Performance Measures	Fusion Methods							
		CT	MR	FM-1	FM-2	FM-3	FM-4	FM-5	Proposed
# 1	En	3.0911	4.0592	4.7054	4.4939	4.7692	4.8087	5.1623	**5.1824**
	STD	79.854	53.881	60.823	82.486	79.859	81.464	82.784	**82.891**
	SF	4.8211	6.4724	6.7467	2.2071	5.8048	6.5477	6.9098	**6.9415**
	XEI	-	-	0.4104	0.5296	0.5328	0.5461	0.5682	**0.5702**
# 2	En	3.3004	4.2197	5.076	5.0684	5.2205	5.2361	5.5147	**5.5412**
	STD	81.822	56.84	60.803	82.383	82.372	83.448	84.887	**85.012**
	SF	6.4645	8.1102	7.8695	3.3249	5.8048	7.5436	**8.8542**	8.651
	XEI	-	-	0.4369	0.5084	0.5114	**0.5375**	0.5346	0.5372
# 3	En	3.1549	4.0256	4.901	4.6585	4.8321	4.8843	5.2241	**5.3124**
	STD	84.708	63.885	60.823	81.486	85.214	87.183	91.124	**91.148**
	SF	5.428	7.1188	7.1971	2.9732	6.2896	6.5436	7.7845	**7.7881**
	XEI	-	-	0.4043	0.5257	0.5308	0.5547	0.5584	**0.5588**
# 4	En	3.5252	3.9525	4.6585	4.5346	4.7512	4.7173	5.1422	**5.1425**
	STD	79.285	62.851	60.926	82.053	80.108	81.997	**85.263**	85.257
	SF	6.0419	6.2869	6.3218	2.2343	5.6939	5.8631	6.9684	**6.9746**
	XEI	-	-	0.4318	0.5077	0.5049	0.5046	0.5086	**0.5099**

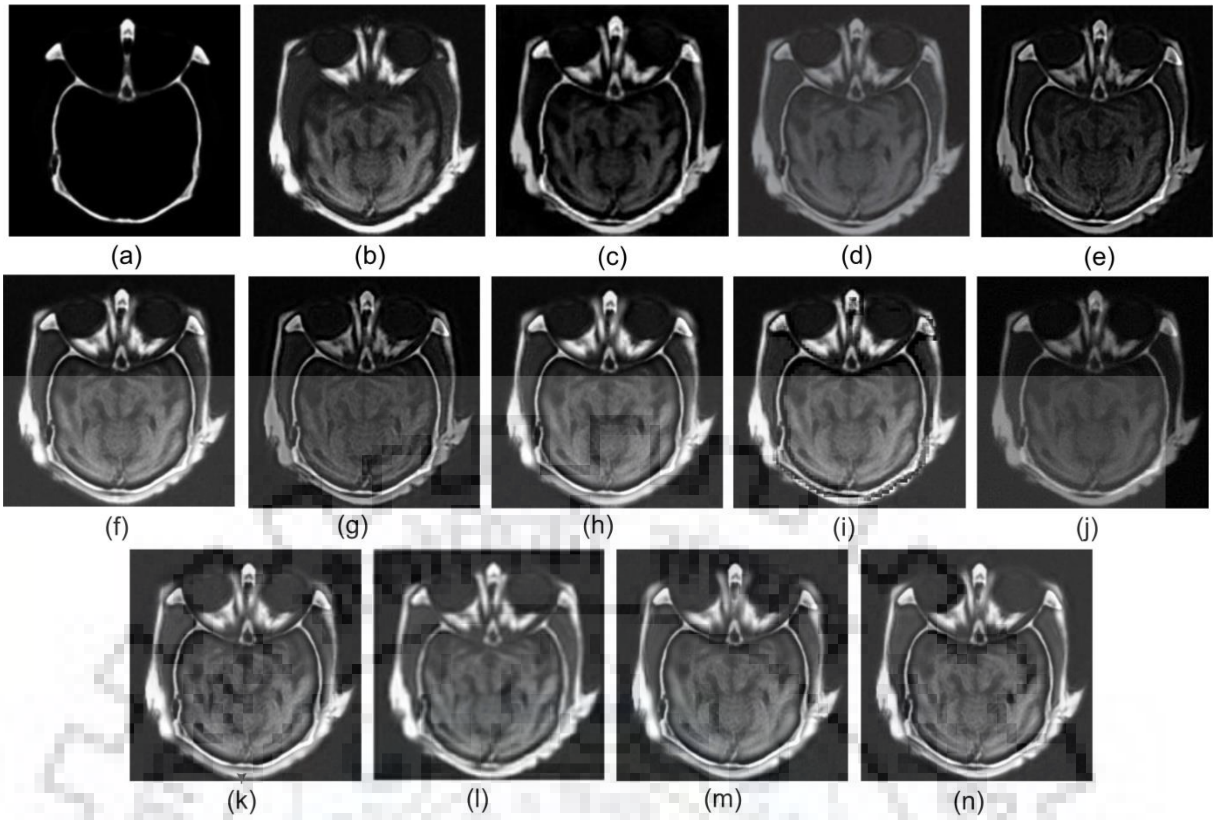


Figure 3.9 Multimodal fusion results of (a) CT image (b) MR image, and fused images obtained by the fusion (c) method [195] (d) method [154] (e) method [92] (f) method [34] (g) method [121] (h) method [55] (i) method [50] (j) method [173] (k) method [189] (l) method [23, 55] (m) NSST-MIF (n) proposed ADP-MIF approach

Table 3.9 Comparative analysis of quantitative parameters evaluated for CT-MR image pair shown in Figure 3.9

Fusion Methods	Performance Measures		
	En	STD	MI
Method [195]	6.387	53.82	-
Method [154]	5.990	32.90	-
Method [92]	6.199	40.56	2.748
Method [34]	6.767	59.85	3.452
Method [121]	6.065	40.22	3.318
Method [55]	6.777	62.03	3.774
Method [50]	6.780	60.02	3.793
Method [173]	-	34.85	3.586
Method [189]	-	20.89	5.889
Method [23, 55]	6.801	60.11	4.10
NSST-MIF	<u>6.835</u>	<u>62.17</u>	<u>4.155</u>
Proposed ADP-MIF	**6.898**	**65.42**	**6.314**

Apart from all the discussion mentioned above, another experiment has also been performed on the standard CT-MR image dataset shown in Figure 3.9 (a) and (b) respectively, which is commonly used to compare the fusion performance. This experiment adds an impact on the superiority of the presented ADP-MIF approach over the others. A detailed comparison of the performance measures evaluated for several dominated medical image fusion approaches developed earlier is presented in Table 3.10. Based on the fusion results obtained for the dataset shown in Figure 3.9, it is clearly observed that the fusion approach presented in this chapter proves its superiority and outperforms the other methods by getting the higher performance measures. Moreover, the presented approach also shows better objective results alongwith the better visual quality of the fused images having more detail information.

Table 3.10 A detailed comparisons of the performance measures obtained by several medical image fusion approaches applied to the source images shown in Figure 3.9

Fusion Methods	Performance Measures		
	En	STD	MI
Wang and Ma, 2008 [180]	-	-	2.41
Yang et al. 2010 [34, 198]	6.729	57.97	2.71
Li et al., 2011 [92, 154]	5.960	32.55	-
Kavitha et al. 2012 [70]	6.730	60.32	-
Wang et al., 2013 [176]	6.192	60.02	-
Liu et al., 2014 [115]	-	64.70	6.27
Bhatnagar et al., 2015 [19]	4.227	60.77	-
Srivastava et al., 2016 [155]	5.694	-	-
Lui et al., 2016 [108]	-	-	5.43
Xia et al, 2018 [185]		55.41	2.24
Tang et al, 2018 [160]	-	60.51	6.11
Proposed ADP-MIF	**6.898**	65.42	**6.31**

3.7 Summary

This chapter presents an improved neurological image fusion approach using the NSML motivated adaptive PCNN model in NSST domain. The NSST used in the presented ADP-MIF approach provides multiscale and directional decomposition. In the proposed approach, local visibility is considered to estimate the adaptive linking strength parameter that varies adaptively to get the high internal activity from those particular regions of source images that have significant features than the low activity regions. The NSML motivated adaptive PCNN model also provides higher sensitivity and clarity in the visual perception. Moreover, this model also captures more detail information present in the input images and increases the visual accuracy of the fused images. The LLGE based fusion rule helps to extract optimal texture feature and provides broad spectral information so that this approach is able to produce more robust and clean structural details corresponding to the maximum energy contained feature regions

For analyzing the fusion results computed by the ADP-MIF approach, several experimentations have been done on the multimodal MR-SPECT and CT-MR image dataset and their performance has been evaluated for both the subjective and quantitative point of view. Moreover, the fusion results achieved by the proposed approach have been compared with other existing approaches available at present. From the objective analysis presented above, it is evident that the NSST decomposition helps to take out edge/detail information in a better way from the reference images and the fusion rules also provide more contrast, clarity level and diagnostic details in the fused resultant images. Based on all the comparative experimental approach and their results, it is emphasized that the proposed approach is able to produce better visual quality fused images as compared to the other existing methods and also retains the diagnostic details appropriately.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 4: ADAPTIVE PCNN BASED HYBRID MULTIMODAL IMAGE FUSION IN NSST AND SWT DOMAIN

This chapter presents another multimodal medical image fusion approach for fusing the anatomical to functional and anatomical to anatomical images. In this chapter, a hybrid medical image fusion (H-MIF) approach is proposed that is based on the NSST and stationary wavelet transform (SWT). The proposed approach also utilizes the features of an adaptive PCNN with varying linking strength parameter based on the local visibility of an image component. Different fusion rules are applied to the low and high frequency of SWT decomposed image subbands and high frequency NSST image coefficients. The methodologies used to present the fusion method are also discussed in this chapter. The performance of the proposed H-MIF approach is explored with the CT, MR, SPECT and PET images in terms of the quantitative and qualitative manner. A detailed comparative analysis has also been done and presented in the results section of this chapter.

4.1 Introduction

From the experimental results discussed in the previous chapters, it is observed that the NSST decomposition has several advantages over to the other multiscale decomposition methods. The NSST has a multiscale and multidirectional property that helps to retain more detail information presented in the source images [58, 175]. Moreover, as discussed in Chapter 1, the PCNN is efficiently utilized in several image processing applications [50, 65, 83, 147, 152, 173, 177]. The modified and improved models of PCNN have also been reported by several researchers in the past years [34, 160, 188, 203]. On the basis of the reviewed model reported in [181], it is analyzed that PCNN based fusion models yield the resultant images with better visual appearance, but still suffer from the proper selection of parameters that may result in loss of some detail information. Therefore, an adaptive PCNN (ADPCNN) model is presented in Chapter 3 in which linking strength parameter is estimated based on the local visibility of the image inspired from the human visual system (HVS) model and adaptively varies to achieve the high internal activity from those particular regions of source images which have significant features than the low activity regions. Therefore, in the proposed approach, the ADPCNN model is employed to fuse the lf SWT subband coefficients followed by the NSST decomposition of source images. Moreover, a novel sum modified Laplacian (NSML) is considered as a feeding input to the ADPCNN in SWT domain and the linking strength parameters (β) is computed adaptively based on the visual clarity of the lf SWT components. The hf SWT subband components are fused using a texture based fusion rule in order to improve the fine details and structural information, hence Local log Gabor energy (LLGE) is computed for each hf SWT component to capture the region which maximizes the respective energy. Furthermore, the hf NSST coefficients are fused using the absolute maximum (AM) and sum of absolute difference (SAD) based fusion rules which help to provide more information related to the pixel contrast and edge details.

The remaining parts of this chapter are structured as follows. The next contains a brief discussion about the methodology used to develop the proposed algorithm. The proposed approach is presented in the subsequent section. Furthermore, the different experimentation and their results are discussed in the last section of the chapter.

4.2 Stationary Wavelet Transform

Most of the traditional multiscale transform techniques are downsampling based decomposition methods that have some limitations such as blocking effects, artifacts and also not able to reflect image detailed information accurately [66, 135, 150]. Evenly in conventional non-redundant wavelet transform, signal is convolved with lowpass and highpass filters and then downsampled to get the next level decomposition, thus the resultant size of the signal is reduced at the next higher level [11]. Nason and Silverman [131] proposed stationary wavelet transform (SWT) to overcome the problem of shift invariance in conventional discrete wavelet transform (DWT) based on multiscale and multidirection decomposition. Instead of down-sampling after applying the lowpass and highpass filters, SWT does not decimate the source images and modifies the filters at each decomposition level by padding them out with zeros. So, the size of the signal is obtained similarly as the approximation signal at the next higher level. SWT can preserve more detailed information available in the reference image by its redundant properties at each scale as compared to other traditional multiscale transformation methods [161, 194]. Therefore, the decompose coefficients can have most of the information of the source images. Thus, SWT is known as átrous algorithm [13, 45]. SWT also shows low computational cost [28, 63].

For a given source image $R(x,y)$ of size $m \times n$, j^{th} level discrete SWT decomposition is given as follows [194]:

$$A^{j+1}(a_x,a_y) = \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} h^j_{k_x} h^j_{k_y}\, C^j(a_x+k_x, a_y+k_y) \tag{4.1}$$

$$D_h^{j+1}(a_x,a_y) = \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} g^j_{k_x} h^j_{k_y}\, C^j(a_x+k_x, a_y+k_y) \tag{4.2}$$

$$D_v^{j+1}(a_x,a_y) = \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} h^j_{k_x} g^j_{k_y}\, C^j(a_x+k_x, a_y+k_y) \tag{4.3}$$

$$D_d^{j+1}(a_x,a_y) = \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} g^j_{k_x} g^j_{k_y}\, C^j(a_x+k_x, a_y+k_y) \tag{4.4}$$

where $A^{j+1}(a_x,a_y)$ refer to the approximation coefficients at j^{th} level and $D_h^{j+1}(a_x,a_y)$, $D_v^{j+1}(a_x,a_y)$ and $D_d^{j+1}(a_x,a_y)$ are the detail coefficients of horizontal, vertical and diagonal direction at j^{th} level, respectively. $a_x = 1,2,3,\dots,m$ and $a_y = 1,2,3,\dots,n$. $h^j_{k_x}$ and $h^j_{k_y}$ refer to low pass filters, while $g^j_{k_x}$ and $g^j_{k_y}$ are high pass filters. C^j represents the low frequency coefficients at j^{th} level. k_x and k_y are the displacement at x-and y-axis, respectively. The inverse discrete SWT is represented as,

$$\tilde{C}^j = \begin{bmatrix} \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} \tilde{h}_{k_x}^j \tilde{h}_{k_y}^j \tilde{A}^{j+1}(a_x + k_x, a_y + k_y) \\ + \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} \tilde{g}_{k_x}^j \tilde{h}_{k_y}^j \widetilde{D}_h^{j+1}(a_x + k_x, a_y + k_y) \\ + \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} \tilde{h}_{k_x}^j \tilde{g}_{k_y}^j \widetilde{D}_v^{j+1}(a_x + k_x, a_y + k_y) \\ + \sum_{k_x=-\infty}^{k_x=\infty} \sum_{k_y=-\infty}^{k_y=\infty} \tilde{g}_{k_x}^j \tilde{g}_{k_y}^j \widetilde{D}_d^{j+1}(a_x + k_x, a_y + k_y) \end{bmatrix} \tag{4.5}$$

where $\tilde{h}_{k_x}^j$ and $\tilde{h}_{k_y}^j$ refer to the reconstruction low pass filters and $\tilde{g}_{k_x}^j$ and $\tilde{g}_{k_y}^j$ refer to as high pass filters. $\tilde{C}^j$ is the reconstructed low frequency coefficients at j^{th} level. $\tilde{A}^{j+1}$ is $(j+1)^{th}$ level modified approximation coefficients and $\widetilde{D}_h^{j+1}$, $\widetilde{D}_v^{j+1}$ and $\widetilde{D}_d^{j+1}$ refer to the modified detail coefficients of horizontal, vertical and diagonal at $(j+1)^{th}$ level.

4.3 Proposed NSST and SWT Domain Hybrid Multimodal Image Fusion Approach

This section presents a hybrid medical image fusion model (H-MIF) in which nonsubsampled shearlet transform (NSST) as a decomposition technique is used to collect the important diagnostic information by providing one low-frequency (lf) and several high frequency (hf) subbands. After getting the decomposed lf NSST subband, SWT is applied only on lf NSST components and left hf NSST components remain same. After the SWT decomposition, another lf and hf SWT subbands are produced. Furthermore, lf SWT coefficients get fused based on the NSML as discussed in the previous chapter, which is applied to motivate the ADPCNN discussed in Chapter 3 as an external input. It reflects more information about the contours and boundaries of multiple objects present in multimodal medical images. In the proposed H-MIF approach, NSML is used as the feature input in the receptive field and decides the total internal activity with respect to the neighboring pixels for fusion at each location. For fusing the remaining hf SWT subbands and integrating all the detailed components, local log Gabor energy based fusion rule is applied that helps to capture the region, which maximizes the respective energy and to provide more robust and clear structural details. Finally, the remaining hf NSST components are also fused by the SAD and AM based fusion rule which reflects the contrast information among the fine and dominant edge available in the reference images. The detail implementation steps are shown in Figure 4.1.

Let $\boldsymbol{R} = R(i,j)$ and $\boldsymbol{S} = S(i,j)$ be two reference anatomical and functional images, respectively, and the proposed H-MIF framework is implemented as given below.

Step 1: Firstly convert the RGB color model of the functional image into $l\alpha\beta$ color space using the given formulation as discussed in [18]. In the proposed approach, the $l\alpha\beta$ color space is considered because if the application is centered on human visual perception as on the medical imaging, which may vary from observer to observer, the lαβ color space would work better with the assumption that the image processing is ideally done by the HVS.

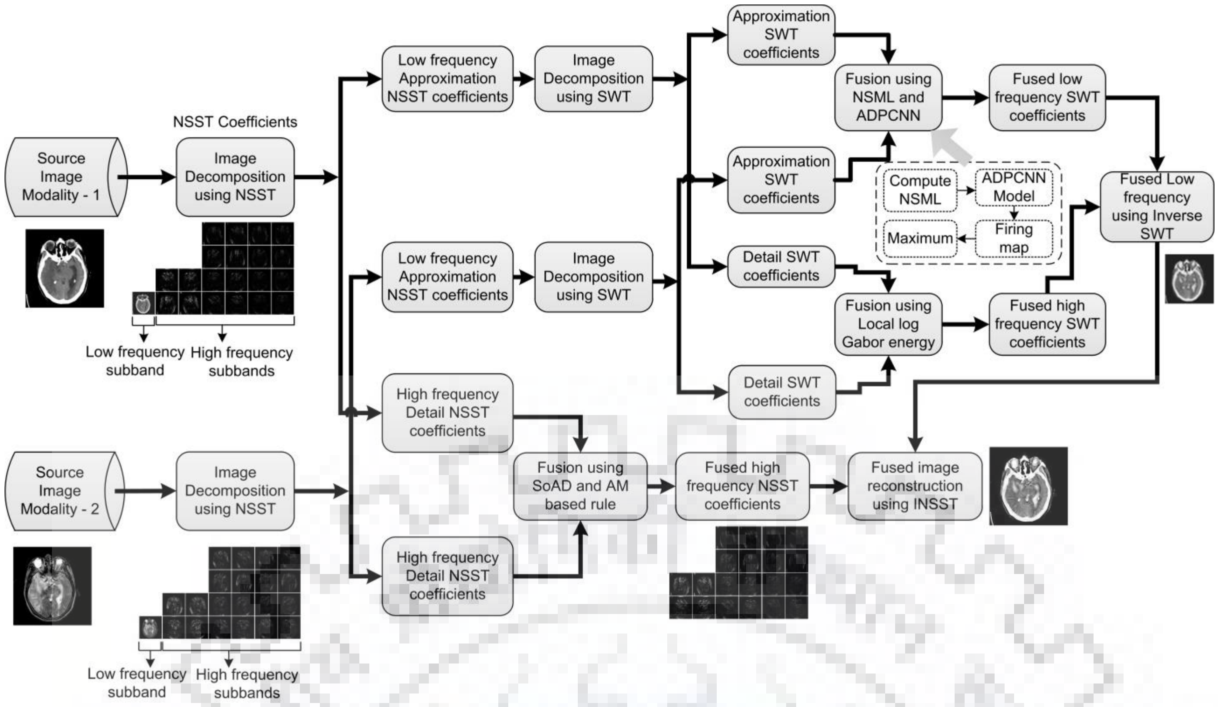


Figure 4.1 Process flow of the proposed H-MIF method

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{4.6}$$

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} \lg L \\ \lg M \\ \lg S \end{bmatrix} \tag{4.7}$$

where l, α and β refer to the achromatic channel, a chromatic yellow-blue and a red-green channels, respectively.

Step 2: Apply the NSST to decompose the source (CT and MR) images or l part of the functional (SPECT and PET) images with shearing directions [4, 8, 8] into lf and hf sub-bands based on the successive trails from coarser to finer decomposition level.

$$\left.\begin{aligned} \left[lf_{\boldsymbol{R}}^{NSST}, hf_{\boldsymbol{R}}^{NSST}\right] &= NSST\big(R(i,j)\big) \\ \left[lf_{\boldsymbol{S}}^{NSST}, hf_{\boldsymbol{S}}^{NSST}\right] &= NSST\big(S(i,j)\big) \end{aligned}\right\} \tag{4.8}$$

Step 3: Decompose the $X_Z = lf_Z^{NSST}$, where Z refers to source images $\boldsymbol{R}$ and $\boldsymbol{S}$, using SWT at two levels with 'sym2' and produces approximation and detail coefficients refer as low and high-frequency components as $L_{X_Z}^{SWT}$ and $H_{X_Z}^{SWT}$.

$$\left.\begin{aligned} \left[L_{X_{\boldsymbol{R}}}^{SWT}, H_{X_{\boldsymbol{R}}}^{SWT}\right] &= SWT(X_{\boldsymbol{R}}) \\ \left[L_{X_{\boldsymbol{S}}}^{SWT}, H_{X_{\boldsymbol{S}}}^{SWT}\right] &= SWT(X_{\boldsymbol{S}}) \end{aligned}\right\} \tag{4.9}$$

Step 4: To fuse lf SWT subband $\left(L_{X_Z}^{SWT}\right)$, estimate the NSML which is able to extract the information in lf coefficients as follows,

$$NSML_{i,j}^{L_{X_Z}} = \sum_a \sum_b w(a,b).G(i+a, j+b) \tag{4.10}$$

where

$$G(i,j)=\begin{bmatrix}\left|2L_{X_Z}^{SWT}(i,j)-L_{X_Z}^{SWT}(i-1,j)-L_{X_Z}^{SWT}(i+1,j)\right|+\\ \left|2L_{X_Z}^{SWT}(i,j)-L_{X_Z}^{SWT}(i,j-1)-L_{X_Z}^{SWT}(i,j+1)\right|\end{bmatrix} \tag{4.11}$$

and

$$w(a,b)=\begin{bmatrix}0.0667 & 0.1333 & 0.0667\\ 0.1333 & 0.2 & 0.1333\\ 0.0667 & 0.1333 & 0.0667\end{bmatrix} \tag{4.12}$$

where $w(a,b)$ is a 3×3 template considered to implement the proposed work.

Step 5: Compute the adaptive linking strength parameter $\beta_{i,j}$ of lf subband coefficients as mentioned above.

$$\beta_{i,j}=\mu\left(V_{X_Z}^{SWT}\right)={}^{1}\Big/\left[1-exp\left(\bar{V}_{X_Z}^{SWT}-V_{X_Z}^{SWT}\right)\right] \tag{4.13}$$

$$V_{X_Z}^{SWT}=\begin{cases}\sum_{i=1}^{r}\sum_{j=1}^{c}\frac{1}{r\times c}(m_I)^{-(1+v)}\left|L_{X_Z}^{SWT}-m_I\right|; & m_I\neq 0\\ L_{X_Z}^{SWT}; & m_I=0\end{cases} \tag{4.14}$$

where r and c refer to the number of pixels in row and column, m_I is the average value of the pixels in the lf SWT subband at $(i,j)^{th}$ pixel location and $V_{X_Z}^{SWT}$ local visibility of the lf subband. The visual constant parameter (v) is chosen within the range of 0.6 and 0.7. Image component would be more important with larger $V_{X_Z}^{SWT}$ and $\bar{V}_{X_Z}^{SWT}$ represent as the mean of the local visibility.

Step 6: Apply the NSML to activate ADPCNN model alongwith the computed adaptive linking strength parameter $\left(\beta_{i,j}\right)$ to boost the neuron pulse by the given mathematical formulation,

$$\left.\begin{aligned}&F_{i,j}^{X_Z}[n]=NSML_{i,j}^{L_{X_Z}}\\ &L_{i,j}^{X_Z}[n]=e^{-\alpha_L}L_{i,j}^{X_Z}[n-1]+V_L\sum_{k,l}W_{i,j,k,l}^{X_Z}Y_{k,l}^{X_Z}[n-1]\\ &U_{i,j}^{X_Z}[n]=F_{i,j}^{X_Z}[n]\left(1+\beta_{i,j}L_{i,j}^{X_Z}[n]\right)\\ &T_{i,j}^{X_Z}[n]=e^{-\alpha_T}T_{i,j}^{X_Z}[n-1]+V_T Y_{i,j}^{X_Z}[n-1]\\ &Y_{i,j}^{X_Z}[n]=\begin{cases}1, & U_{i,j}^{X_Z}>T_{i,j}^{X_Z}\\ 0, & otherwise\end{cases}\end{aligned}\right\} \tag{4.15}$$

Step 7: Compute the firing times (sum of the $Y_{i,j}^{Z}=1$for $U_{i,j}^{Z}>T_{i,j}^{Z}$) in n iterations,

$$t_{i,j}^{X_Z}[n]=t_{i,j}^{X_Z}[n-1]+Y_{i,j}^{X_Z}[n] \tag{4.16}$$

Step 8: If $n=n_{max}$, the process stops and the coefficients are combined based on the fusion rule by firing time evaluated in the previous step as presented below.

$$L_F^{SWT}=\begin{cases}L_{X_R}^{SWT}; & t_{i,j}^{X_R}[n_{max}]\geq t_{i,j}^{X_S}[n_{max}]\\ L_{X_S}^{SWT}; & t_{i,j}^{X_R}[n_{max}]<t_{i,j}^{X_S}[n_{max}]\end{cases} \tag{4.17}$$

Step 9: For fusing the hf SWT subband $\left(H_{X_Z}^{SWT}\right)$, compute the local log Gabor energy (LLGE) in SWT domain.

$$LLGE_{i,j} = \frac{1}{(1+2a)\times(1+2b)} \sum_{i=-a}^{a} \sum_{j=-b}^{b} LGE_{i+a,j+b} \tag{4.18}$$

where $(1+2a)\times(1+2b)$ refers to the template size with $a=3$ and $b=3$ and LGE refers to the log Gabor energy in the SWT domain for all hf coefficients estimated as,

$$LGE = \sum_{k=1}^{K} \sum_{l=1}^{L} \left(Re\left(H_{X_Z}^{SWT} * g_{i,j}^{k,l}\right)^2 + Im\left(H_{X_Z}^{SWT} * g_{i,j}^{k,l}\right)^2 \right)^{1/2} \tag{4.19}$$

where $g_{i,j}^{k,l}$ refers to the log Gabor filter coefficients in k and l number of scale and directions, respectively.

Step 10: Now, fuse the hf SWT subband coefficients using the formulated fusion rule based on the computation of the LLGE aforementioned.

$$H_F^{SWT} = \begin{cases} H_{X_R}^{SWT}; & LLGE_{i,j}^{X_R} \geq LLGE_{i,j}^{X_S} \\ H_{X_S}^{SWT}; & LLGE_{i,j}^{X_R} < LLGE_{i,j}^{X_S} \end{cases} \tag{4.20}$$

Step 11: Apply the inverse SWT on fused L_F^{SWT} and H_F^{SWT} to compute the fused lf subband.

$$X_F = SWT^{-1}\left(L_F^{SWT}, H_F^{SWT}\right) \tag{4.21}$$

Step 12: Compute the AM and the SAD in a 3×3 template for hf_Z^{NSST} subband.

$$AM_Z^{NSST} = max\left(\left|hf_Z^{NSST}(i+x, j+y)\right|\right), -1 \leq x, y \leq 1 \tag{4.22}$$

$$SAD_Z^{NSST} = \begin{bmatrix} \sum_{a=-1}^{1} abs\left(hf_Z^{NSST}(i+a,j) - hf_Z^{NSST}(i+a,j+1)\right) \\ +\sum_{a=-1}^{1} abs\left(hf_Z^{NSST}(i+a,j) - hf_Z^{NSST}(i+a,j-1)\right) \\ +\sum_{a=-1}^{1} abs\left(hf_Z^{NSST}(i,j+a) - hf_Z^{NSST}(i+1,j+a)\right) \\ +\sum_{a=-1}^{1} abs\left(hf_Z^{NSST}(i,j+a) - hf_Z^{NSST}(i-1,j+a)\right) \end{bmatrix} \tag{4.23}$$

Step 13: Apply the fusion rule based on AM and SAD to get fused hf_F^{NSST} subband.

$$hf_F^{NSST} = \begin{cases} hf_{\boldsymbol{R}}^{NSST}; & AM_{\boldsymbol{R}}^{NSST} > AM_{\boldsymbol{S}}^{NSST} \\ hf_{\boldsymbol{S}}^{NSST}; & AM_{\boldsymbol{R}}^{NSST} < AM_{\boldsymbol{S}}^{NSST} \\ hf_{\boldsymbol{R}}^{NSST}; & SAD_{\boldsymbol{R}}^{NSST} > SAD_{\boldsymbol{S}}^{NSST} \\ hf_{\boldsymbol{S}}^{NSST}; & otherwise \end{cases} \tag{4.24}$$

Step 14: Apply the inverse NSST on the fused lf and hf components to reconstruct the resultant image, as follows

$$F = NSST^{-1}\left(X_F, hf_F^{NSST}\right) \tag{4.25}$$

Step 15: To fuse the anatomical and functional images, include α and β computed in step 1 with the fused image obtained from step 14 by the given mathematical formulation and finally convert it back to the RGB color format.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} F \\ \alpha \\ \beta \end{bmatrix} \tag{4.26}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} 10^X \\ 10^Y \\ 10^Z \end{bmatrix} \tag{4.27}$$

The schematic of the proposed fusion approach for anatomical and functional images is shown in Figure 4.2 in which the shaded block of proposed H-MIF approach is considered as shown in Figure 4.1.

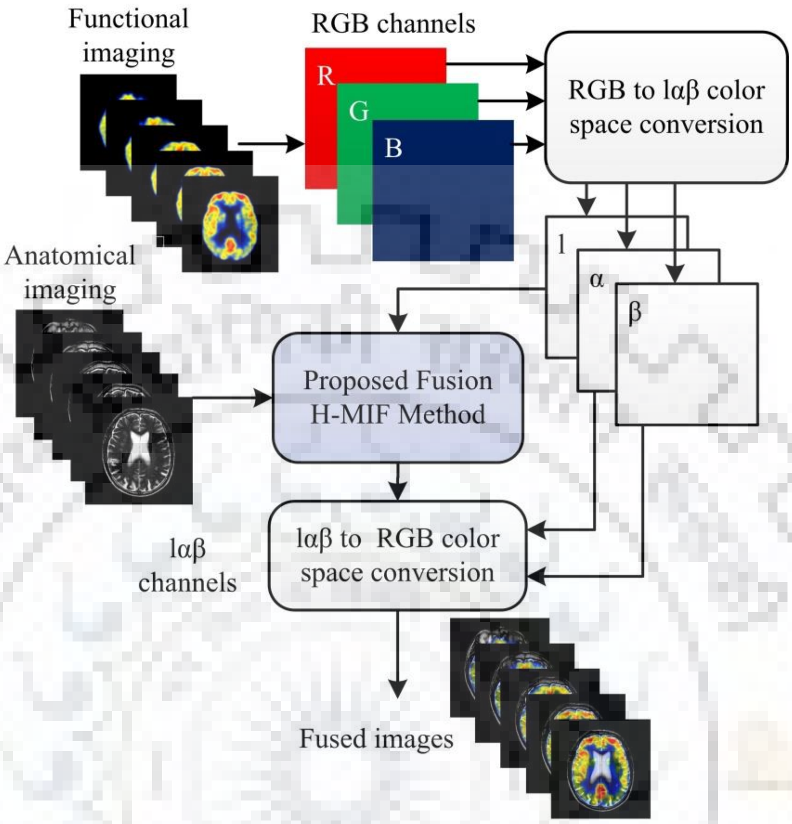


Figure 4.2 Process flow to fuse MR to SPECT/ PET images

4.4 Experimentation

To investigate the efficacy of the proposed H-MIF method, a large number of the experiments have been performed and a comparison is made with several existing fusion methods in both the term of qualitatively and quantitatively. The qualitative results speculate the comparative assessment of visual perception among the source and fused images while the quantitative investigation demonstrates the usefulness of the perceived results through a set of predefined performance measures. Fused outcomes of the proposed H-MIF and other approaches are evaluated and analyzed for the MR-SPECT and MR-PET neurological images alongwith the CT-MR images in order to get more diagnostic information with edge details available in both the source images, respectively. To assess the fusion performance, several experiments are conducted using different pairs of MR-SPECT and MR-PET images and the results are presented to reflect the comparative performance of the proposed H-MIF and other approaches. Moreover, the fusion of CT-MR image pairs is also presented in this chapter to provide a strong evidence of the superiority of the proposed H-MIF approach in terms of performance measures. These performance measures do not only show the comparative analysis, but also verify the acceptability of visual observation by showing the preservation of the information and structural contents available in the source images. To conduct the experiments, NSST decomposition level $= [2, 3, 3]$ and ADPCNN parameters were set as $\alpha_{\mathrm{T}} = 0.2$, $\alpha_L = 1, \mathrm{V_L} = 1$, $\mathrm{V_T} = 20$, $\mathrm{W} = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2};\ 1\ \sqrt{2}\ \ 1\right]$ and iterations$= 200$. Two different experiments are discussed as given below,

Experiment 1: To analyze and evaluate the fusion performance of the proposed H-MIF approach applied on the anatomical-functional images (MR-SPECT and MR-PET images) and present a comparative analysis of the proposed and other existing fusion methods, qualitatively and quantitatively.

Experiment 2: To analyze and evaluate the fusion performance of the proposed H-MIF approach applied to the anatomical-anatomical images (CT-MR images) and present a comparative analysis of the proposed and other existing fusion methods, qualitatively and quantitatively.

4.5 Results and Discussions

4.5.1 Experiment 1: Comparisons of anatomical-functional image fusion results

This section presents the results of different existing image fusion methods and the proposed H-MIF method with different experimentations. To conduct an experiment and analyze the objective fusion performance of the proposed H-MIF approach, a large dataset of MR-SPECT image pairs are utilized, out of which four pairs of MR and SPECT images are shown in the first and second row of Figure 4.3, respectively, with two regions marked by red and blue rectangles. MR images show the structural information of soft tissues, however, the metabolic changes are not clearly perceived through it due to low accuracy. The SPECT image provides the functional information required to detect the growth of any abnormality. Their corresponding fused images for visual inspection of all four MR-SPECT image pairs are shown in Figure 4.4 (f) to Figure 4.7(f). The proposed approach H-MIF formulated by the NSML motivated ADPCNN mapping for fusing the lf SWT subband and the LLGE based rule to fuse hf SWT subband coefficients provides a better quality of fused images. This shows more information contents and edge information than the source images. Visually, the fused images have more information alongwith the edge and spectral information as compared to the source MR and SPECT images. Furthermore, the superiority of the proposed H-MIF approach is investigated by comparing the fusion results shown in Figure 4.4 to Figure 4.7 obtained from the following fusion results.

Method 1: Fusion approach based on the multiresolution singular value decomposition technique by applying the image division with 2×2 blocks and arranging the corresponding block in a 4×1 vector with the similar parameters as discussed similar in [129].

Method 2: Fusion scheme based on the standard deviation in WT domain [137] by decomposing each one of the reference images into 8×8 block.

Method 3: Fusion approach based on the NSST decomposition and PCNN model using sum of directional gradient with the similar parameters as mentioned in [48].

Method 4: Image fusion based on NSCT decomposition where the SML based fusion rule is applied on both hf and lf subband as discussed in [203].

Method 5: Fusion approach based on the hybrid multiscale decomposition with guided image filtering concept as described in [218].

Method 6: The proposed H-MIF approach based on the NSST and SWT decomposition ADPCNN parameters $\alpha_{\mathrm{T}} = 0.2, \alpha_L = 1, \mathrm{V_L} = 1, \mathrm{V_T} = 20,\ \mathrm{W} = \frac{1}{\sqrt{2}}\left[1\ \sqrt{2}\ \ 1;\ \sqrt{2}\ 0\ \sqrt{2}; 1\ \sqrt{2}\ \ 1\right]$ and iterations= 200. The alpha variable of the ADPCNN model controls the firing time of the neuron and defines the rate of decay of the threshold in the iterative series for the pulse generator. The larger value of alpha can make the running time faster, whereas the smaller value can make the firing process very slow but performs the pixel's selection very precisely.

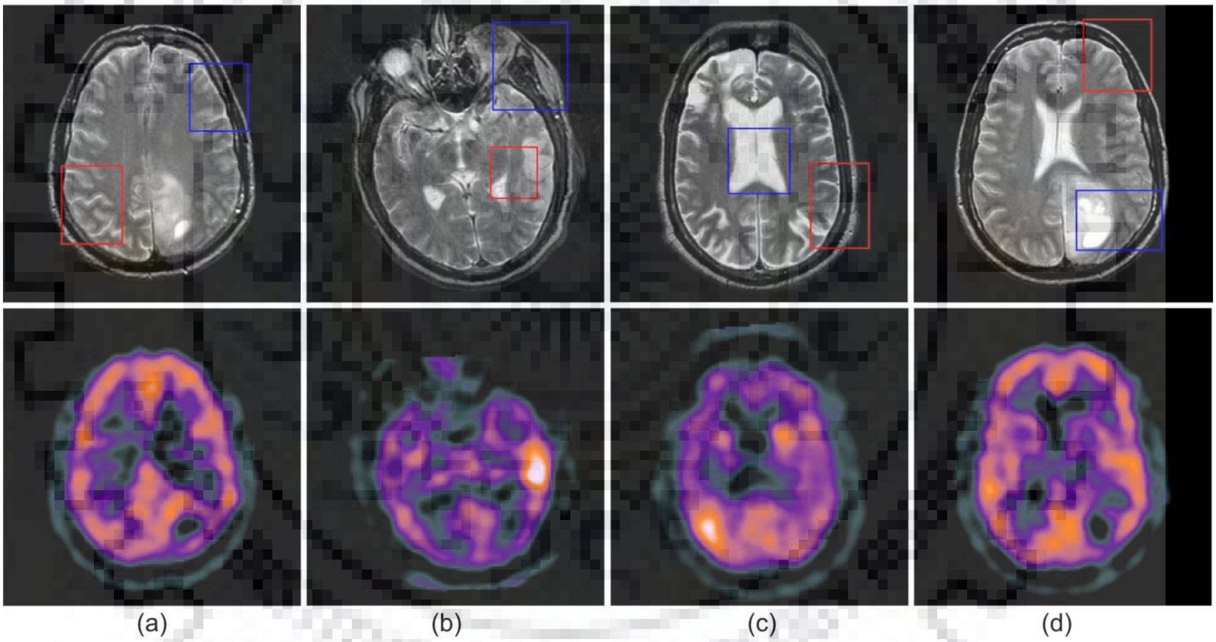


Figure 4.3 Source MR-SPECT image data (a) Pair 1 (b) Pair 2 (c) Pair (3) (d) Pair 4 with two marked regions by red and blue for showing the comparative analysis of fusion performance

To evaluate the superiority of the proposed H-MIF approach, a detailed assessment is done by comparing the performance of other fusion methods in terms of both the subjective and quantitative analysis. A visual comparison of all four pairs of MR-SPECT brain images is made with all abovesaid six approaches as shown in Figure 4.4 to Figure 4.7. Figure 4.4 to Figure 4.7 (a)-(e) present the fused images of all five existing fusion methods 1 to 5, while Figure 4.4 (f) to Figure 4.7 (f) present the fusion results obtained by the proposed H-MIF method. The method presented by Nirmala et al. [137] does not show good visual results as it suffers from some spectral distortion. The method proposed by Ganasala and Kumar [48] shows good information and edge preservation from the source MR images, however, it suffers from poor spectral contrast and consistency hence shows over-brightness in some regions. Yang et al. [203] approach blurs most of the important edge details. The method presented by Zhu et al. [218] preserves the spatial consistency very well but smoothed out the structural details. From Figure 4.4 (f) to Figure 4.7 (f) as the fusion results of the proposed H-MIF approach, it is visualized that the proposed method outperforms the other fusion methods by reflecting more amount of diagnostic information and edge details with sharp localization of abnormal tissues (conserve the proper spectral variance near brighter area) that are also pointed by the arrows of the different colors. For more and clear visualization of the fused images, the zoomed regions of two rectangles marked by the red and blue color shown in Figure 4.3, are also presented in Figure 4.8 and Figure 4.9 respectively. From the zoomed results presented in Figure 4.8 and Figure 4.9, it is observed that the pixel contrast becomes little vague near boundary regions, while for blue box false brightness near abnormality is introduced due to spatial inconsistency. As shown in Figure 4.8 (d) and Figure 4.9 (d), the zoomed area for red box indicates a good contrast with smooth edge details while blue box indicates inappropriate pixel saliency. In Figure 4.8 (e) and Figure 4.9 (e), it is clearly observed that most of the structural details are smoothed out.

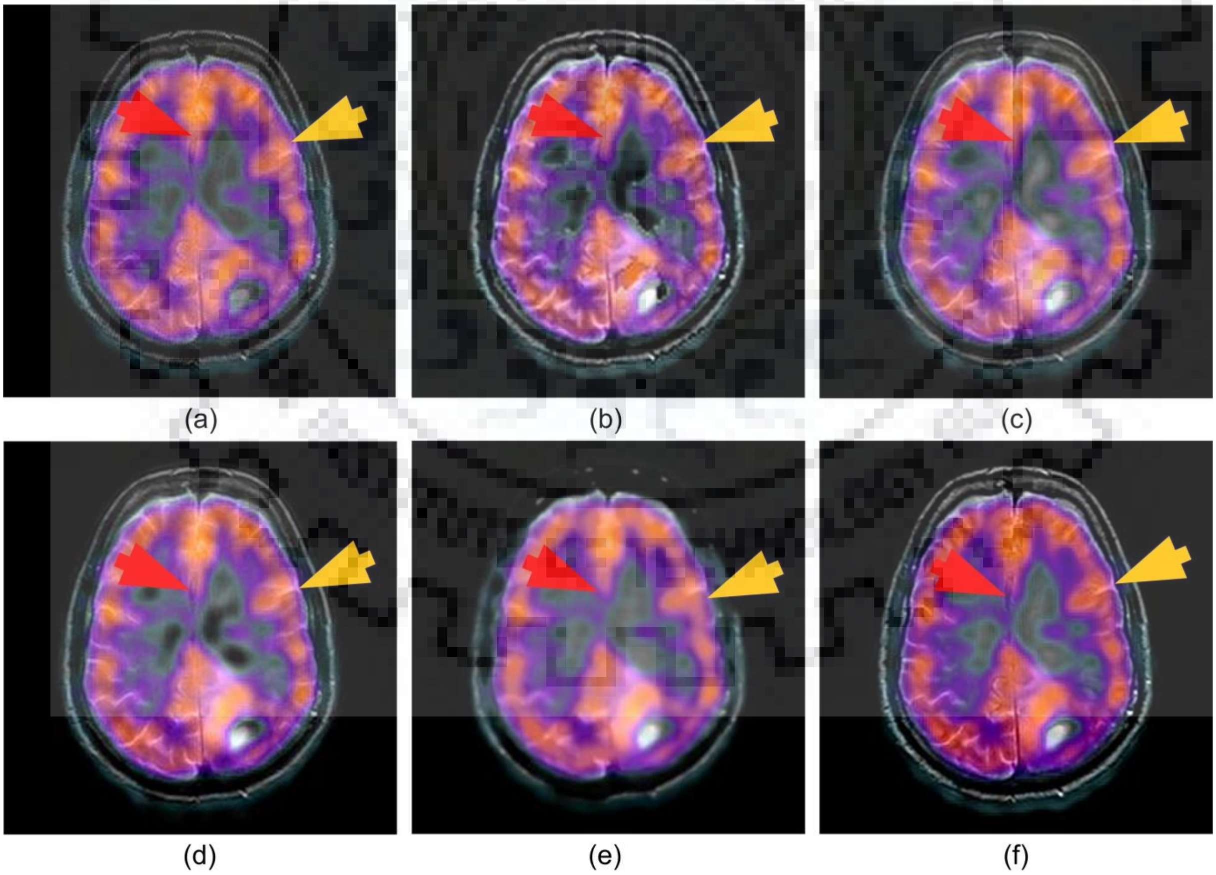
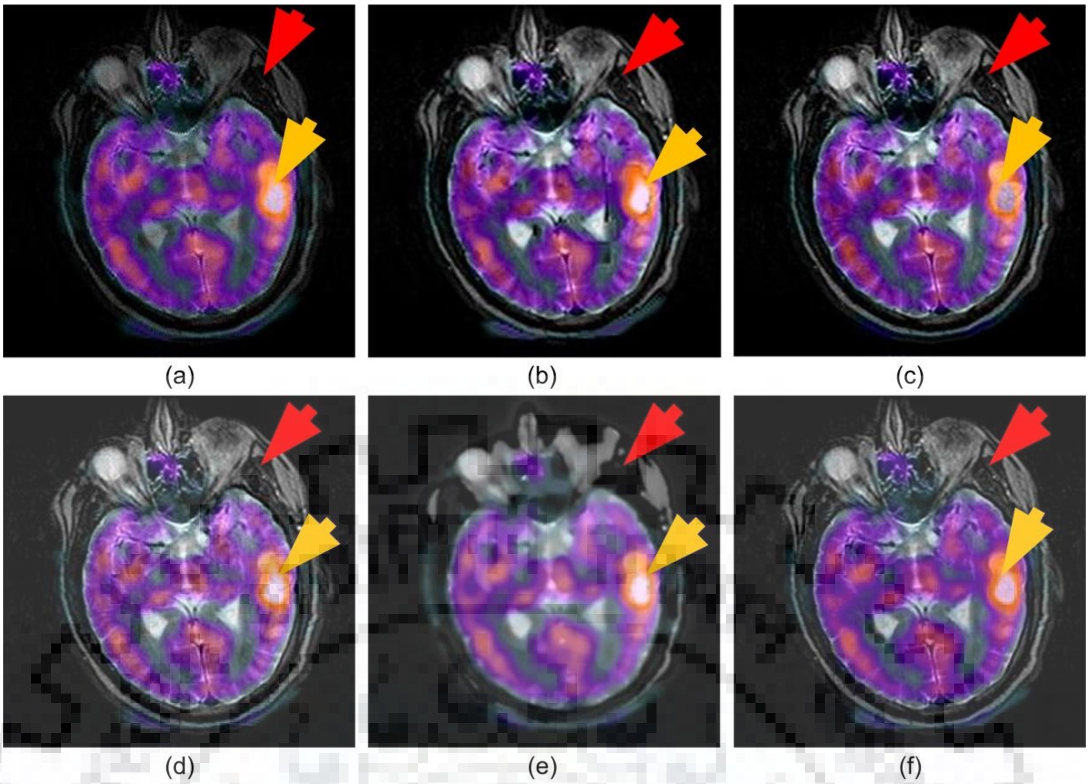
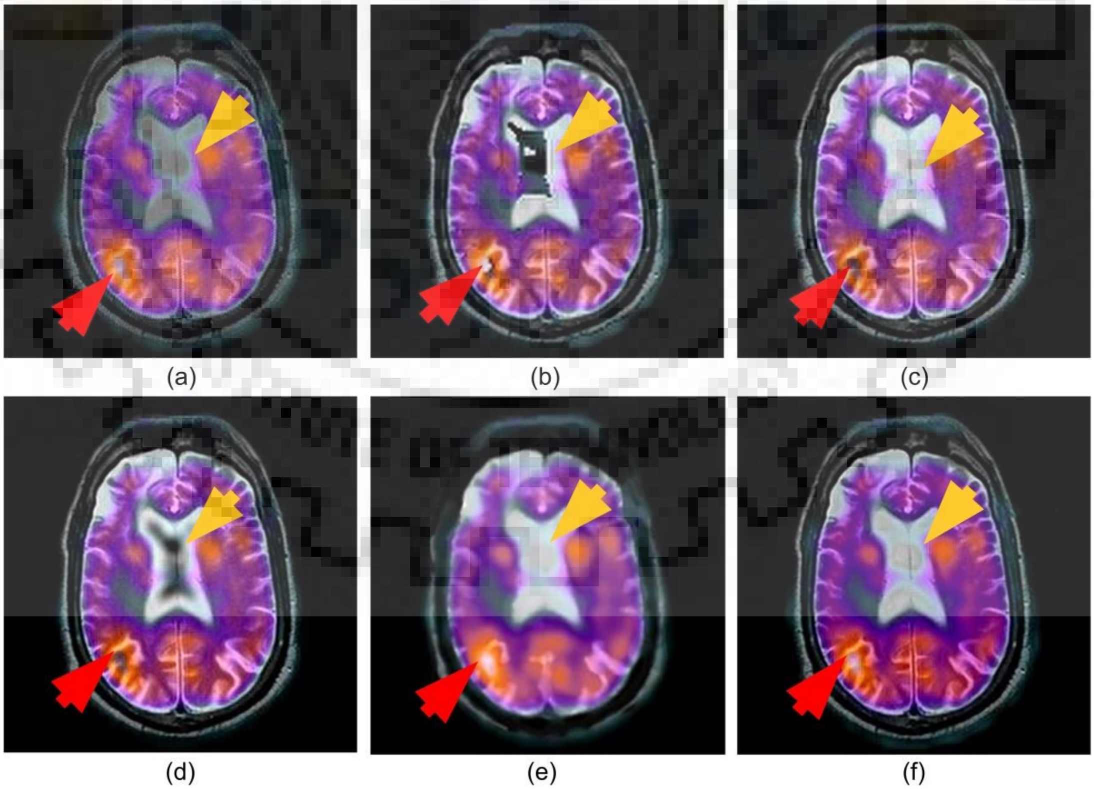


Figure 4.4 Comparative analysis of the fusion results for pair-1 obtained by (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF


(a) (b) (c)

(d) (e) (f)

Figure 4.5 Comparative analysis of the fusion results for pair-2 obtained by (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF


(a) (b) (c)

(d) (e) (f)

Figure 4.6 Comparative analysis of the fusion results for pair-3 obtained by (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF

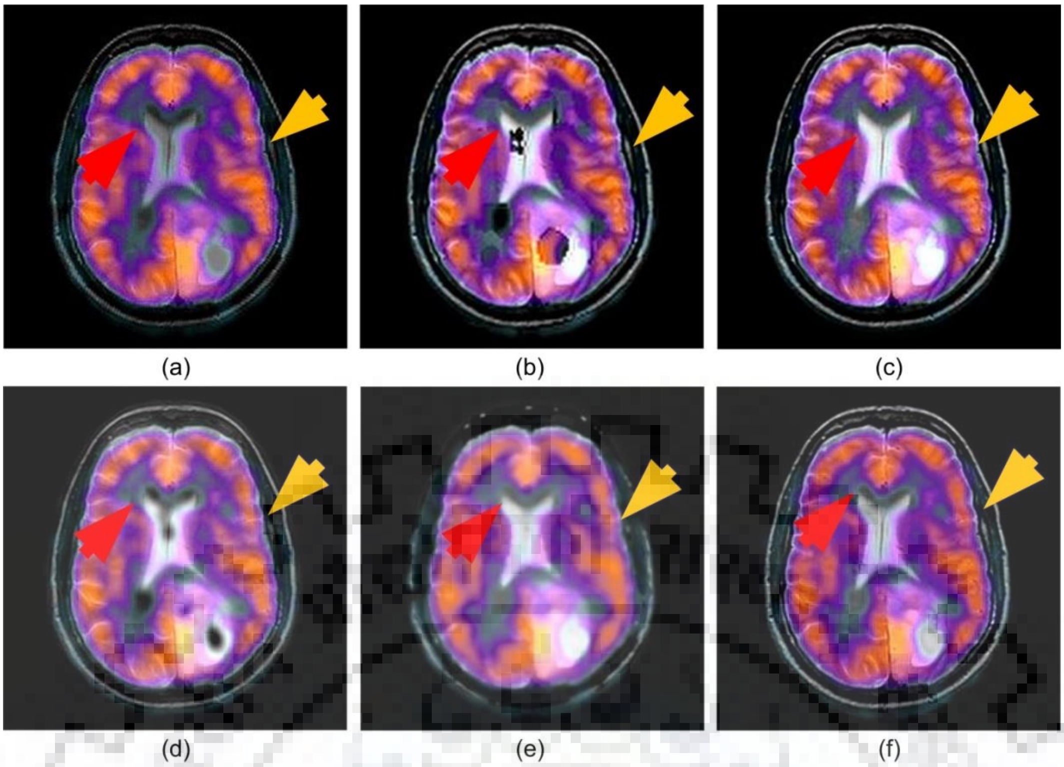

(a) (b) (c)
(d) (e) (f)

Figure 4.7 Comparative analysis of the fusion results for pair-4 obtained by (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF

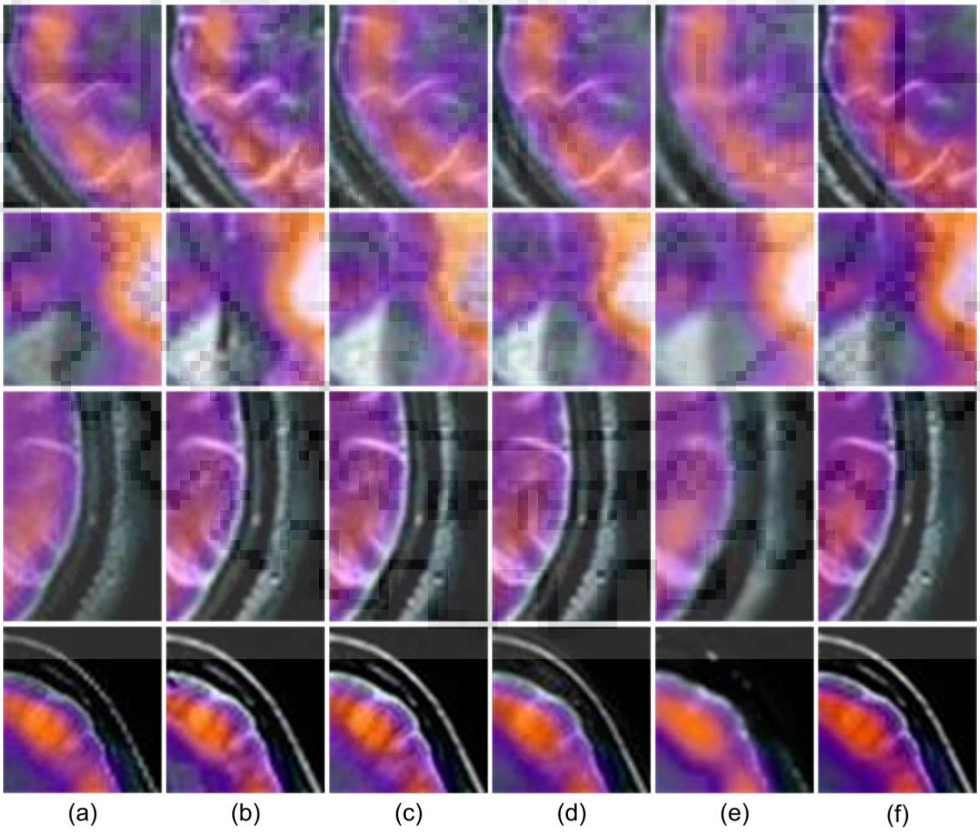

(a) (b) (c) (d) (e) (f)

Figure 4.8 Comparisons of the zoomed regions marked on Figure 4.3 by the red color (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF

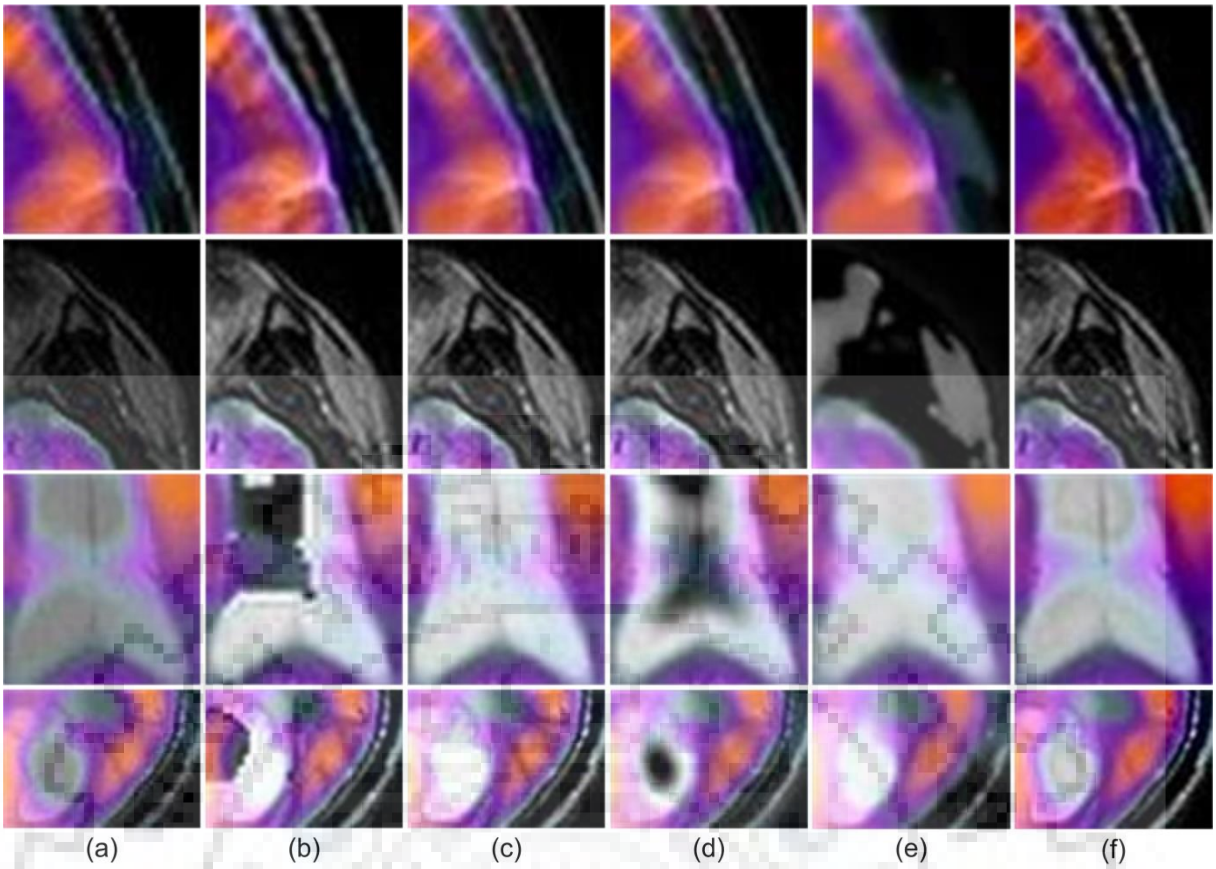


Figure 4.9 Comparisons of the zoomed regions marked on Figure 4.3 by the blue color (a) Method 1 (Naidu, 2011 [129]) (b) Method 2 (Nirmala et al. 2016 [137]) (c) Method 3 (Ganasala and Kumar, 2016 [48]) (d) Method 4 (Yang et al., 2017 [203]) (e) Method 5 (Zhu et al. (HMSD-GDGF), 2018 [218]) (f) Proposed H-MIF

In order to verify the performance of the visual observation, the quantitative evaluation of all above mentioned fusion methods with the proposed technique is shown in Table 4.1 to Table 4.4 by considering the performance measures i.e. EN, STD, FMI, IQI and XEI. From the results, it is observed that the method 1 provides the higher IQI values to show better fused quality images but not able to reflect significant edge preservation (XEI) as well as FMI values. Ganasala and Kumar method [48] shows higher values of STD for most of the MR-SPECT image pair with the significant values of En to represent the information in fused outcome, but lacks to preserve overall IQI due to poor spectral contrast and color tone consistency. Yang et al. approach [203] provides the satisfactory performance in terms of the image quality and edge preservation presented by IQI and XEI values, but somewhere shows less variance in the neighboring pixels near sharp edges, hence gains lower STD values compared to method 3. The method proposed by Zhu et al. [218] maintains the image information very well but underperforms for fine details and pixel variations, thus shows the lowest value of FMI and XEI. The performance of the proposed H-MIF method shows higher values of En, FMI and XEI which verify the visual analysis of the fused result by preserving the more edge details.

Table 4.1 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-1

Measures	Source Images		Fusion Methods					
	MR	SPECT	Method 1	Method 2	Method 3	Method 4	Method 5	H-MIF
En	3.4229	3.9282	4.7606	4.6944	4.8387	4.5904	4.8066	**5.2294**
IQI	-	-	**0.70828**	0.3356	0.35343	0.51628	0.54317	0.60784
FMI	-	-	0.32852	0.39032	0.3555	0.40054	0.23893	**0.49175**
STD	55.9916	53.213	57.6457	63.2451	**64.2506**	64.0905	63.2522	59.3715
XEI	-	-	0.38218	0.57859	0.6214	0.63143	0.21332	**0.64421**

Table 4.2 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-2

Measures	Source Images		Fusion Methods					
	MR	SPECT	Method 1	Method 2	Method 3	Method 4	Method 5	H-MIF
En	4.7658	3.6669	5.6653	5.9643	5.8069	5.6811	5.9569	**6.0724**
IQI	-	-	**0.47953**	0.43004	0.43198	0.38255	0.19678	0.43513
FMI	-	-	0.31269	0.40616	0.4771	0.48399	0.18638	**0.48433**
STD	55.7785	46.38127	51.116	61.8237	62.1543	62.1164	60.0039	**62.6116**
XEI	-	-	0.40263	0.64278	0.71552	0.71922	0.22041	**0.71972**

Table 4.3 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-3

Measures	Source Images		Fusion Methods					
	MR	SPECT	Method 1	Method 2	Method 3	Method 4	Method 5	H-MIF
En	3.8952	4.1106	4.8964	4.8299	5.0555	4.7683	5.0327	**5.3412**
IQI	-	-	**0.65983**	0.3268	0.34752	0.39856	0.3984	0.58112
FMI	-	-	0.34893	0.39373	0.39851	0.4365	0.2241	**0.46611**
STD	64.95625	50.76391	57.3648	69.178	**71.9812**	68.9441	70.2975	69.9306
XEI	-	-	0.37523	0.62359	0.68119	0.69059	0.25658	**0.69861**

Table 4.4 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-4

Measures	Source Images		Fusion Methods					
	MR	SPECT	Method 1	Method 2	Method 3	Method 4	Method 5	H-MIF
En	3.6817	4.0593	4.898	4.8357	4.9056	4.6971	5.1386	**5.4175**
IQI	-	-	**0.69091**	0.33316	0.35196	0.49132	0.48633	0.58708
FMI	-	-	0.33438	0.39268	0.37299	0.41824	0.23832	**0.42082**
STD	68.1471	55.40022	63.384	74.0926	**76.1346**	74.523	74.4726	73.4073
XEI	-	-	0.40636	0.57464	0.6369	0.63988	0.23972	**0.65976**

Furthermore, another experiment is conducted on two different image pairs as shown in Figure 4.10 and Figure 4.11 correspond to mild Alzheimer's disease of a patient with different slices of brain MR-T2 and PET-FDG to detect the changes in anterior temporal and posterior parietal regions. Figure 4.10 to Figure 4.11 (c), (d) and (e) show their corresponding fused image pairs obtained by Ganasala and Kumar, 2016 [48], Yang et al., 2017 [203], and the proposed H-MIF approach, respectively. Figure 4.10 and Figure 4.11 show clearly the superiority of the proposed method over the others by providing better visual quality that is pointed out by the arrows in all the fused images. Moreover, a detailed comparison of the averaged objective parameters is presented in Table 4.5. The combined analysis of visual and quantitative results of the method presented in [48] for the given MR-PET images indicate that the color variations and edge information are well preserved. This method also provides a significant values of the STD and XEI with a fair amount of overall information contents taken from both the source images with higher FMI values. However, the fine details near the boundary are smoothed out. Another method presented by Yang et al. [203] shows better tiny edge detail preservation, but loses the actual spectral details and variations hence shows lower values of IQI and MI. The proposed method maintains the spatial and spectral details well with sharp tiny details shown by a higher score of En and IQI. The higher value of FMI for the proposed method indicates that the brighter tiny features in source images are preserved properly with the appropriate consistency and localization.

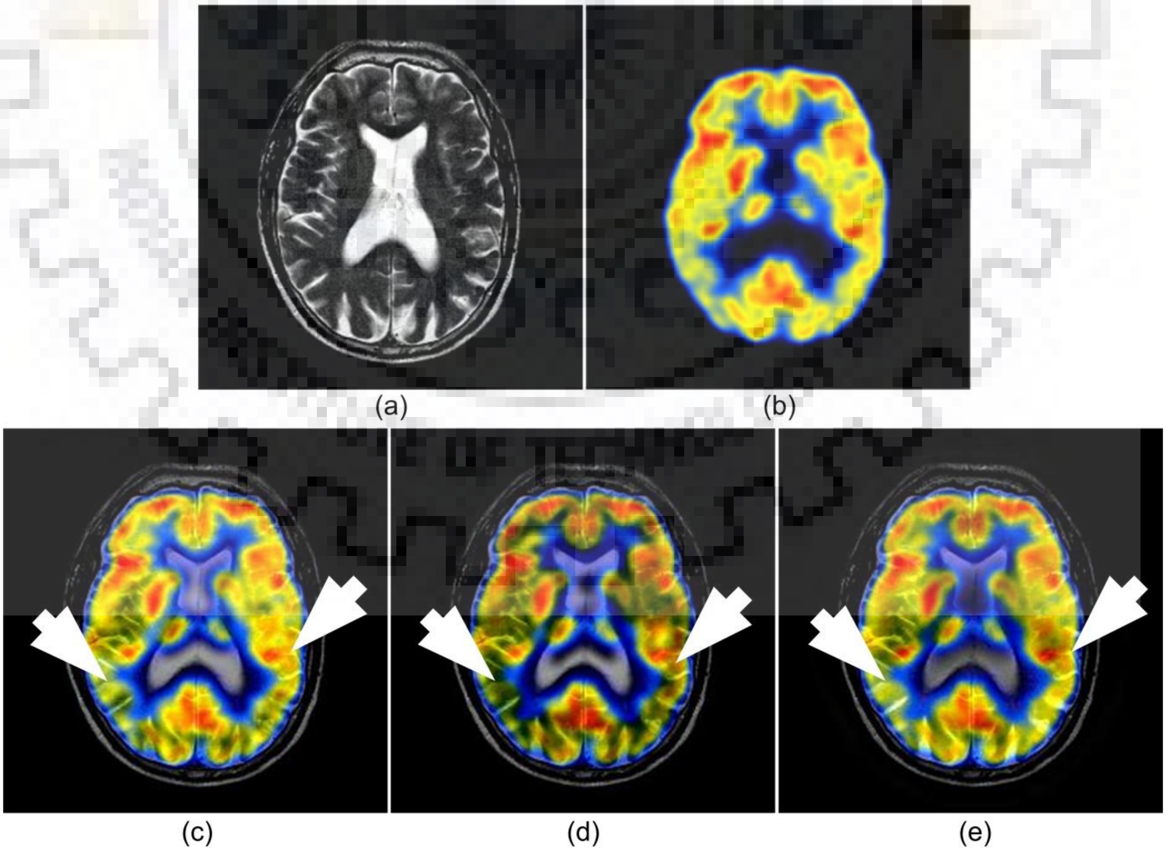
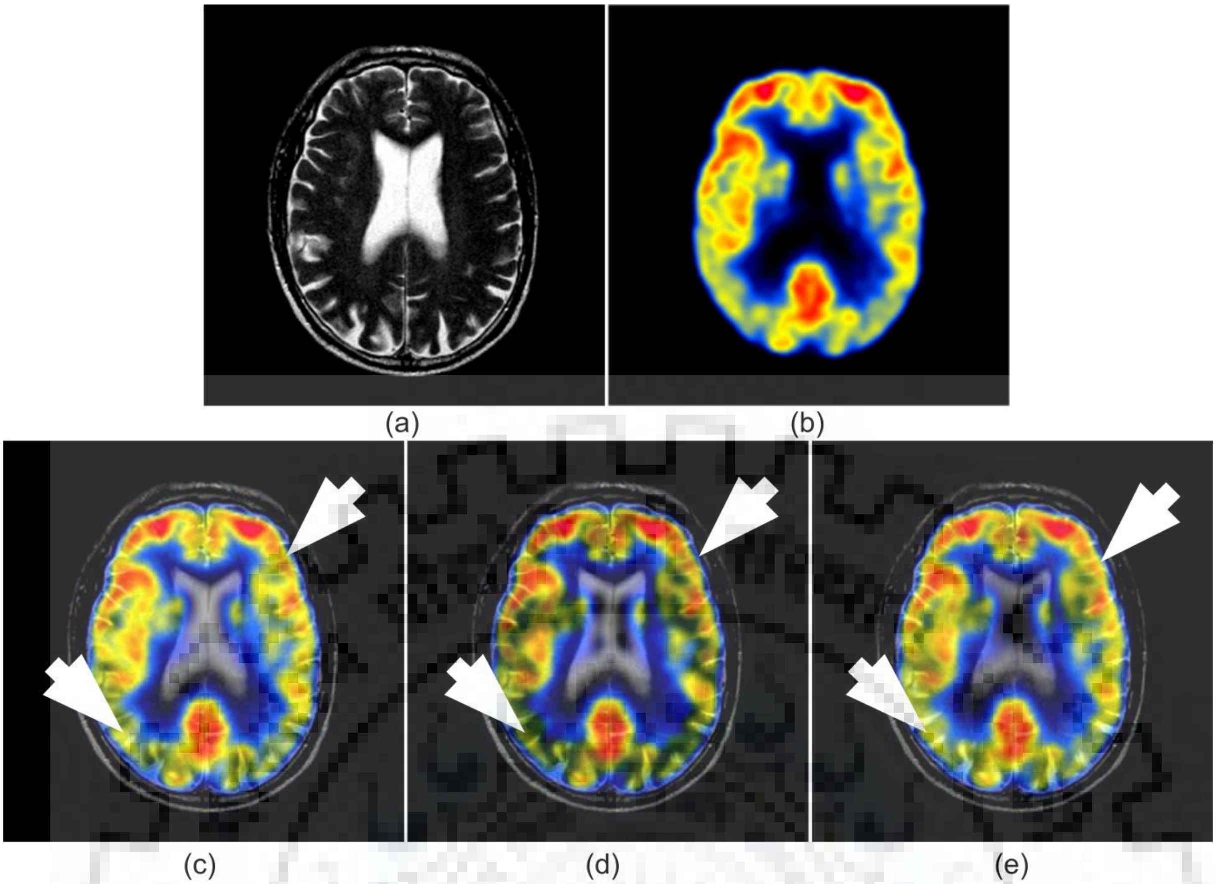

(a) (b)
(c) (d) (e)

Figure 4.10 Comparative analysis of (a) MR image (b) PET image and fusion results of (c) Ganasala and Kumar, 2016 [48] (d) Yang et al., 2017 [203] (e) Proposed H-MIF



(a) (b)

(c) (d) (e)

Figure 4.11 Comparative analysis of (a) MR image (b) PET image and fusion results of (c) Ganasala and Kumar, 2016 [48] (d) Yang et al., 2017 [203] (e) Proposed H-MIF

Table 4.5 Comparative analysis of averaged performance measures for MR-PET images

Performance Measures	Ganasala and Kumar, 2016 [48]	Yang et al., 2017 [203]	Proposed H-MIF
En	4.7493 ± 0.1667	4.5992 ± 0.1723	4.9683 ± 0.1621
MI	2.3231 ± 0.0175	1.9162 ± 0.0095	2.3854 ± 0.0395
IQI	0.5897 ± 0.0143	0.5621 ± 0.0111	0.5916 ± 0.0013
STD	71.674 ± 1.7520	56.607 ± 0.2557	67.700 ± 2.180
FMI	0.4551 ± 0.0830	0.4483 ± 0.0580	0.4846 ± 0.067
XEI	0.5904 ± 0.0021	0.5775 ± 0.0080	0.5952 ± 0.002

4.5.2 Experiment 2: Comparisons of anatomical-anatomical image fusion results

This section presents the comparative analysis of the fusion of anatomical-anatomical (CT-MR) images done by another four state of the art fusion approaches alongwith the proposed H-MIF method. Two image pairs out of complete CT-MR dataset are shown in Figure 4.12 (a)-(b) and Figure 4.13 (a)-(b). To present a comparative visual performance of the fused images, some of the existing methods developed previously are considered as fusion method 3, method 4 and method 5 presented by Ganasala and Kumar, 2016 [48], Yang et al., 2017 [203] and Zhu et al. (HMSD-GDGF), 2018 [218], respectively as mentioned above, C-MIF approach discussed earlier in Chapter 2 and the proposed H-MIF method.

Their fusion results are Figure 4.12 (c)-(g) and Figure 4.13 (c)-(g), respectively. In addition to visual assessment, the objective performance measures are also computed and listed in Table 4.6. In Figure 4.12 (c) and Figure 4.13 (c) Ganasala and Kumar approach [48] indicates clear hard tissues with correct brightness and significant En values for both of the image pairs but fails to show small structural details with lower values of MI and IQI as displayed by zoomed areas. For the second set of CT-MR image pair, the method blurs the resultant image, hence does not show the soft tissue details clearly. The visual observation of Yang et al. [203] method indicates that the small edges are clearly perceived in the fused results shown in Figure 4.12 (d) and Figure 4.13 (d), thus achieves a decent score of XEI and STD, though the zoomed area indicates that few details suffer from pixel contrast and inappropriate luminance. Zhu et al. method [218] preserves better saliency and brightness for very small details, hence, results the higher values of En and MI, but still have degraded variance and little smoothed edge details as shown in Figure 4.12 (e) and Figure 4.13 (e). In Figure 4.12 (f) and Figure 4.13 (f), the C-MIF approach discussed in Chapter 2 shows an improvement in the preservation of the informative content and the pixel variations but gets lower IQI due to slight blurring in brighter areas as represented by zoomed regions. The proposed H-MIF method is designed in such a way that the fused outcome gets more structural details with proper pixel variance in order to get better FMI and IQI values. The improved En values also indicate that the proposed H-MIF method achieves good complementary information, hence clearly outperforms the other considered fusion methods in terms of both the visual and objective assessments.

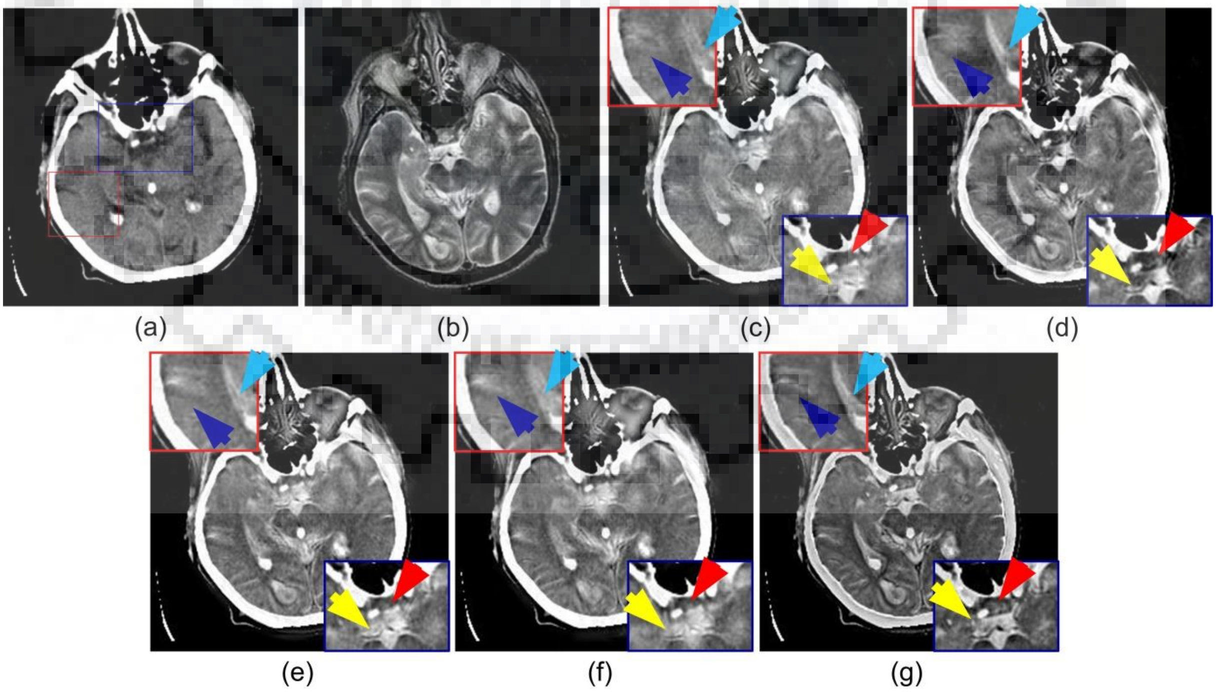


Figure 4.12 Comparative analysis of fusion results obtained for source (a) CT image (b) MR image, by (c) Ganasala and Kumar, 2016 [48] (d) Yang et al., 2017 [203] (e) Zhu et al. (HMSD-GDGF), 2018 [218] (f) C-MIF method (g) Proposed H-MIF method

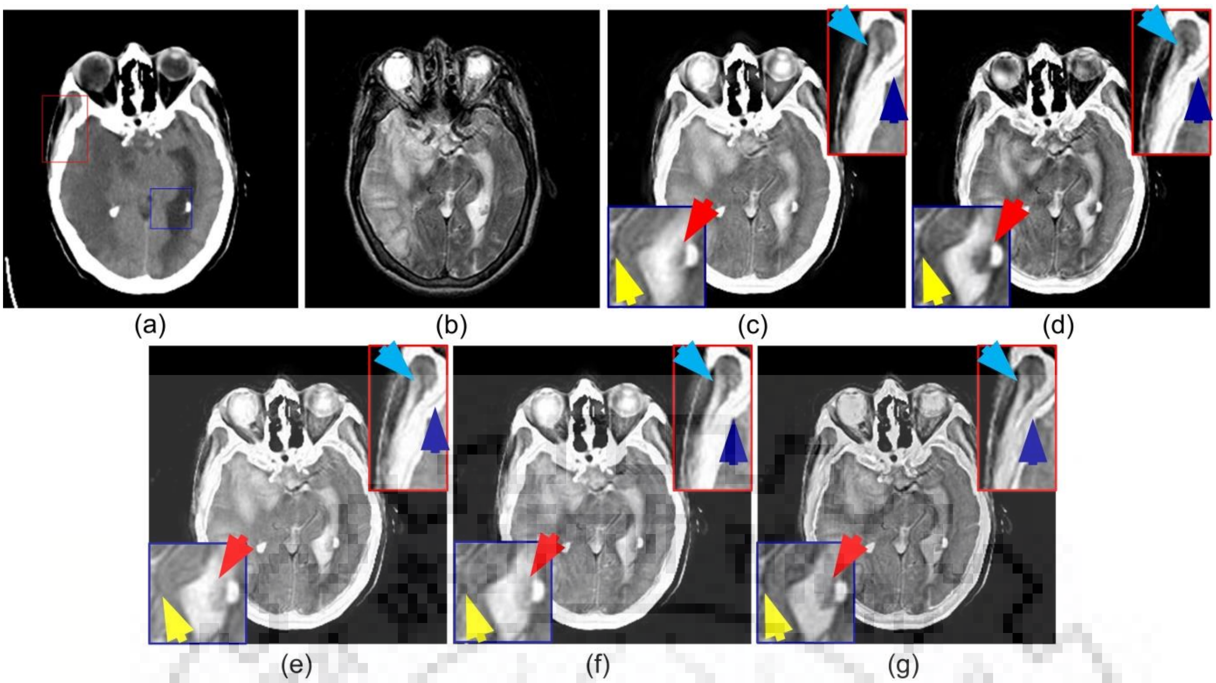


Figure 4.13 Comparative analysis of fusion results obtained for source (a) CT image (b) MR image, by (c) Ganasala and Kumar, 2016 [48] (d) Yang et al., 2017 [203] (e) Zhu et al. (HMSD-GDGF), 2018 [218] (f) C-MIF method (g) Proposed H-MIF method

Table 4.6 Comparative analysis of the performance measure evaluated for CT-MR image pairs

Image Pairs	Methods	Performance Measures				
		En	MI	IQI	STD	XEI
Figure 4.12	CT	3.3004	-	-	81.822	-
	MR	4.2197	-	-	56.840	-
	Ganasala and Kumar, 2016 [48]	5.2205	2.5842	0.4554	82.372	0.5114
	Yang et al., 2017 [203]	5.2361	2.5875	0.4758	83.448	0.5375
	Zhu et al. (HMSD-GDGF), 2018 [218]	5.2728	2.8077	0.466	82.218	0.5013
	C-MIF	5.5147	2.8824	0.4544	**84.887**	0.5346
	Proposed H-MIF	**5.5516**	**2.8877**	**0.5003**	80.105	**0.5592**
Figure 4.13	CT	3.1549	-	-	84.708	-
	MR	4.0256	-	-	63.885	-
	Ganasala and Kumar, 2016 [48]	4.8321	2.8754	0.4989	85.214	0.5308
	Yang et al., 2017 [203]	4.8843	2.8417	0.5301	87.183	0.5547
	Zhu et al. (HMSD-GDGF), 2018 [218]	5.0989	2.8441	0.5267	88.006	0.5681
	C-MIF	5.2241	3.0874	0.5228	**91.124**	0.5584
	Proposed H-MIF	**5.2465**	**3.1142**	**0.5646**	84.433	**0.5835**

4.6 Summary

This chapter presents an effective fusion approach for multimodal medical images with an adaptive PCNN model in the NSST and SWT domain. The proposed fusion approach employs the NSST as multiscale and multidirectional decomposition to maintain both the spatial as well as spectral information very well using shearing filters. Moreover, the SWT provides better textural information which is required for mapping of the several neurological diseases as well, whereas the NSST is able to capture edge/detail information from the source images. In the SWT domain, adaptive PCNN model is utilized to improve the comprehensive information in lf component, while local log Gabor energy is utilized to extract the salient features available in the source image and to retain the color and edge details without introducing any artifacts. To analyze the fusion results of the H-MIF approach, several experimentations have been done on the multimodal MR-SPECT, MR-PET and CT-MR image dataset and their performance have been evaluated for both the subjective and quantitative point of view. Moreover, a detailed comparative analysis is presented to validate the superiority of the proposed H-MIF approach. However, it still suffers from contrast level of fused images having lower STD value that may be acceptable alongwith the ability of the proposed method to retain the more diagnostic detail information. Based on the comparative experimental results, it is ensured that the proposed H-MIF method is able to preserve the significant information of multimodal input images by producing the better visual quality of fused images with a significant improvement in the performance measures.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 5: MULTIMODAL MEDICAL IMAGE FUSION USING SPARSE K-SVD DICTIONARY LEARNING IN NSST DOMAIN

This chapter presents a novel fusion framework for multimodal neurological images which is able to capture small-scale details of input images with original structural details. In its procedural steps, at first, source images get decomposed by the NSST into a low frequency and several high frequency components to separate out the two basic characteristics of source image i.e. principal information and edge details. The low frequency layer gets fused with a sparse representation based model and high frequency components are merged by the guided filtering based approach. Finally, fused images are reconstructed by employing the inverse NSST. The methodologies used to present the fusion method are also discussed in this chapter. The superiority of the proposed approach is confirmed by a large extent of analytical experimentations performed on different real MR-SPECT, MR-PET, and CT-MR neurological image datasets. The detailed comparative analysis has also been done and presented in the results section of this chapter.

5.1 Introduction

As discussed in the previous chapters, multimodal medical imaging sensor technology is considered by the clinicians as a prominent solution to diagnose several neurological diseases in which fusion of anatomical and functional images would be helpful to provide a meaningful quantifiable interpretation which is suitable for clinical diagnosis by providing the exact location and orientation of the defected tissues. For such purposes, several authors moved towards the development of the multimodal image sensor fusion methods [19, 62, 71, 171]. As shown in Chapters 2, 3 and 4, multiscale transform (MST) based fusion approaches showed a good dominance over the other methods as they are able to preserve the overall information very well compared to the spatial domain approaches which often suffer from spectral distortion and poor contrast [115]. In this context, different multiscale decompositions as complex wavelet [24, 154], SWT [6], wavelet packets [9], CVT [155], CNT [16, 33, 195, 197], NSCT [18, 47, 91, 158, 176] and NSST [46, 48, 77, 147, 175, 210] have been introduced to design an effective platform which provide better localization of image contour and consistent texture details with superior visual image quality.

The main idea behind all these conventional transform domain approaches is that the significant information of the reference images can be extracted robustly from the decomposed coefficients. In addition to transform domain fusion technqiues, the sparse representation model has also been applied in several image fusion applications [64, 132, 191, 212]. Yang and Li [191] applied sparse representation (SR) based fusion approach for multifocus images in which sparse coefficients are merged for two reference images as an activity level measurement. Also, these sparse coefficients are calculated on local patches of the training images and then fused by the max rule which significantly improves the contrast, invariance and stability. Since the SR-based rules work on the local patches in the spatial domain, they are very much sensitive to random noise that can cause spatial variability and computational inefficiency. Liu et al. [109] employed SR based fusion in the multiscale domain to overcome most of the limitations of classical transform based and SR based fusion approaches which can be easily observed from their experimental results. However, the MST-SR based rule leads to the chromatic imbalance and causes over brightness, as pixel saliency is not well captured during the fusion process. A feature level fusion method is developed for visual tracking based on joint sparse representation [80], but the unreliable features are not always correlated with the similar sparsity coefficients in visual tracking application; hence the single sparse dictionary cannot signify the multiple features. In another approach, a joint sparse representation based fusion framework is utilized for the visual tracking that is able to perform in feature space for a given kernel matrix [81]. Furthermore, multiple SR based framework has been presented by Lan et al. to explain the different features by decomposing the multiple sparsity models which ensure more representative learned features [82]. Besides this, Li et al. [94] implemented another edge retaining nonlinear filtering as a guided filter to capture sharp features of detail layers and large intensity variations of the base layer of source images. Recently, Zhu et al. [218] adopted another gradient domain guided filtering to fuse the images. Though, it shows the favorable fused results, however, it suffers from spatial smoothing or staircase effects in image boundary.

Considering the results and limitations of the state of the art fusion methodologies aforementioned, an effective multimodal medical image fusion framework is presented in this chapter using multiscale geometric analysis with sparse representation and guided filtering after decomposing the source images into a base layer lf and a series of detail layer hf components. The main motivation to work with the NSST is the several advantages over the other transformation techniques as mentioned in Chapter 3. The main purpose to perform the NSST on source images is that shearlets are optimally sparse for approximation with fine time-frequency localization and directional sensitivity. Furthermore, To fuse the decomposed lf and hf NSST coefficients, the proposed method reports about the utilization of both the features of sparse representation and guided filtering. The combination of these two models shows better preservation of both the spatial and spectral contents even with the presence of noise, artifacts and other spatial or geometrical inconsistencies. The proposed medical image fusion based on sparse dictionary learning and guided filtering (SDL-MIF) approach utilizes the ability to extract the spatial information and can preserve the spatial consistency alongwith robustness to acquisition noise and mis-registration artifacts. The salient contributions of the proposed approach are summarized as follows,

- This chapter proposes a unified fusion framework for multimodal medical images in the NSST domain that can capture the complex image contours with unrestricted directional details at each scale in order to optimize the reconstruction efficiency of the method.

- An overcomplete dictionary is learned (training of medical image dataset) to capture the complex details of medical images and sparsely represented for low frequency subband for better visual feature (luminance and contrast) projection without any spectral distortion.
- Intuitively, most of the edge information lies at high frequency, which is also sensitive to random white Gaussian noise and some obvious artifacts. Hence, the spatial saliency based decision map is constructed and optimized for the high frequency subband fusion to preserve the pixel saliency and spatial consistency.
- Moreover, the performance of the proposed framework is also verified to assess the chromatic balance and spectral features in the fused color image.

The remaining part of this chapter is structured as follows. The next two sections present the different preliminary related works which are adapted to formulate the proposed SDL-MIF framework. The proposed method and its implementation steps are presented in the subsequent section. Furthermore, the different experimentation and their results are discussed in next section of the chapter.

5.2 Dictionary Learning based Sparse Representation

The concept of the sparse representation is that the image data is approximately expressed in terms of a linear combination of the few atoms by an over-complete dictionary. The mathematical expression of the sparse representation is given by $\boldsymbol{X} = \boldsymbol{D}\alpha$ where $\boldsymbol{D} \in R^{k\times m}$ $(k < m)$ is an over-complete dictionary and sparse coefficients refer to α. The terms k and m refer to the size of image data and dictionary, respectively. The main aim of the sparse representation is to find the optimized solution of the sparse coefficient vector which has the most zero value in all vectors. This process is called sparse coding. In recent developments, multi-scale and multi-directional techniques are also an important motivation on sparse coding [64, 132]. Different algorithms available in literature are used for over-complete dictionary learning that is able to present the efficacy of sparse codings such as the method of optimal directions (MOD) [109], orthogonal matching pursuit (OMP) [20], K-singular value decomposition (K-SVD) [3] and and simultaneous OMP [167]. In the learning stage, a mock dictionary is trained from a large number of input data patches inferred from the specific applications. This creates a learned dictionary model where the atoms come from some underlying observations rather than some restricted environments. For n number of training patches of a fixed size $\sqrt{k} \times \sqrt{k}$, the sample training data $[\boldsymbol{Y}]_{i=1}^{n}$ is obtained by random sampling and convert it to column vector in R^k space. The model for dictionary learning is expressed as,

$$\min_{d_k,\{\alpha_i^{\boldsymbol{X}}\}_{i=1}^{n}} \sum_{i=1}^{n}\left\|\alpha_i^{\boldsymbol{X}}\right\|_0, \quad s.t.\ \left\|\boldsymbol{Y}-\boldsymbol{D}\alpha_i^{\boldsymbol{X}}\right\| < \epsilon,\ i \in [1,2,\ldots,n] \tag{5.1}$$

where $\boldsymbol{D}$ is the prototype overcomplete dictionary to be learned and $\{\alpha_i^{\boldsymbol{X}}\}_{i=1}^{n}$ represents an unknown vector of sparse coefficients corresponding to sample data patches $\{\boldsymbol{y}\}_{i=1}^{n}$ and $\epsilon > 0$ is the possible tolerable value of error. In the presented fusion model, a sparse dictionary learned by the training dataset of 120 multimodal medical images. For learning the overcomplete dictionary K-SVD is employed that is most popular among all adaptive learning approaches available in the state-of-the-art. The K-SVD approach comprises mainly the two-step process (a) coding by using the OMP to compute sparse coefficients $\{\alpha_i^{\boldsymbol{X}}\}_{i=1}^{n}$ of each individual patch; and (b) to search a better approximation model by updating the dictionary column by column singular value decomposition (SVD) computation using the optimization equation as mentioned in [3],

$$\min_{d_k,g_k}\left\|E_k - d_k g_k^T\right\|_F^2; \quad s.t.\ \|d_k\|_2 = 1 \tag{5.2}$$

where $E_k = \boldsymbol{Y} - \sum_{j\neq k} d_j \alpha_{j,x}$ is the residual matrix except k^{th} dictionary atom and g_k^T is the k^{th} row in the $\alpha_i^{\boldsymbol{X}}$. This rank-one problem is solved by approximation method and then update all atoms of $\boldsymbol{D}$ in a similar fashion to get a new dictionary and then go for the sparse coding stage to compute the best possible sparse vectors until it terminates at a stopping criterion. This process is executed for only those instances whose present model belongs to the atom d_k to avoid by the introduction of any new nonzero coefficient in $\alpha_i^{\boldsymbol{X}}$. The dictionary computed by using K-SVD outperforms both for real and synthetic images by filling missing pixels and also provide a better image representation. Such model is too fast and efficient as it simultaneously updates the current atom and its correlated sparse coefficients unlike the MOD approach that suffers from relatively high computational complexity because of matrix inversion instead of updating the dictionary [149]. The dictionary learning process is shown in Algorithm 1.

5.3 Guided Filtering

In this section, a brief discussion on a guided fIlter (GFF) is presented that is used to implement the proposed approach. The GFF is employed to filter the input image in accordance with the contents of the guidance image over the size of local window (r) neighboring pixels with the regularization parameter (ε). This regularization parameter decides the degree of blurring while the radius (r) controls the filter size. Intuitively, for the pixel region (within a window) with lower variance than ε (Flat area), the kernel behaves as a low pass filter and the value of current pixel is the average value of the neighboring pixels, whereas in the higher variance pixel regions (edge patches), the current pixel value is preserved and the pixels within the kernel are not smoothed in order to maintain the spatial consistency. The theoretical model of guided filter expressed as,

Algorithm 1 Dictionary Learning Using K-SVD Approach

```
Input: Multimodal training dataset ${\boldsymbol{Y}}_{i=1}^{n}$
Output: Learned dictionary $D^{k^2\times K}$ with $k \ll K$
Parameters: Dictionary size $(K)$, number of iterations $(iter)$, reconstruction error $(\epsilon)$, patch size $(k)$ and the total number of atoms to be trained $(n)$
Initialization: $D_{updated} = D^0, \in R^{k^2\times K}$ (initial dictionary)
Repeat: $iter$ times
    1. Sparse coding: Compute sparse vector using orthogonal matching pursuit using,
       $\alpha_{i,j} = \min_{D^0,\alpha_i} \sum_{i=1}^{n} \|\alpha_i\|_0, \quad s.t. \|\boldsymbol{Y} - D^0\alpha_i\| < \epsilon$
    2. Dictionary update: For each dictionary atom
        • Find the patches belongs to dictionary atom
        • Compute approximation error matrix.
          $E_k = \boldsymbol{Y} - \sum_{j\neq k} d_j \alpha_{j,x}$
        • Apply column by column SVD $(U, \Delta, V)$ on approximation matrix $(E_k)$.
        • Update $D^0$ column as the first column of $U$.
        • Update $\alpha_i$ as the multiplication of $V$ and $\Delta$.
    3. Find the updated sparse coefficients
       $[D_{updated}, \alpha_{updated}] = argmin \|E_k - d_k g_k^T\|_F^2; \quad s.t. \|d_k\|_2 = 1$
end
```

$$F_i = A_n \boldsymbol{S}_x + B_n, \quad \forall x \in W_n \tag{5.3}$$

where F is the filtering output of the guidance image $\boldsymbol{S}$ in a local patch of W_n of size $(2r+1)\times(2r+1)$ centered at pixel x. The parameters A_n and B_n are fixed linear parameters within the given window and estimated by minimizing the given error equation for source image I and filtered output F.

$$E(A_n, B_n) = \sum_{x\in W_n}((A_n \boldsymbol{S}_x + B_n - \boldsymbol{I}_x)^2 + \varepsilon A_n^2) \tag{5.4}$$

The parameters A_n and B_n are estimated by the linear regression analysis presented in [41, 94] as,

$$A_n = \frac{\frac{1}{|W|}\sum_{x\in W_n} \boldsymbol{S}_x \boldsymbol{I}_x - \mu_n^{\boldsymbol{S}} \mu_n^{\boldsymbol{I}}}{\sigma_n^{\boldsymbol{S}} + \varepsilon} \tag{5.5}$$

$$B_n = \mu_n^{\boldsymbol{I}} - A_n \mu_n^{\boldsymbol{S}} \tag{5.6}$$

where $\mu_n^{\boldsymbol{S}}$ and $\sigma_n^{\boldsymbol{S}}$ refer to an average and variance parameter of the guidance image within a local template (W_n), respectively, $\mu_n^{\boldsymbol{I}}$ also refers to the average parameter of the source image within W_n and $|W|$ is computed as the complete pixel values within the same local template (W_n). The resultant output is computed by taking an average of all possible values of the coefficients A_n and B_n using the following formulation,

$$F_x = \frac{1}{|w|}\sum_{n\in W_x} A_n \boldsymbol{S}_x + \frac{1}{|w|}\sum_{n\in W_x} B_n \tag{5.7}$$

5.4 Proposed NSST Domain Image Fusion Approach Based on Sparse Dictionary Learning

This section summarizes the complete implementation steps with the mathematical formulations involved in the proposed fusion method based on sparse K-SVD learning and guided filtering approach. In the proposed fusion framework, the first step is to get the multiscale decomposition of the source images in a lf and multiple hf subbands, then the

SR and GFF based fusion rules are applied to fuse the energy and detailed information corresponds to pixel characteristics. In the last step, fused image reconstruction and performance evaluation are done. The detailed implementation scheme of the proposed approach is presented in Figure 5.1.

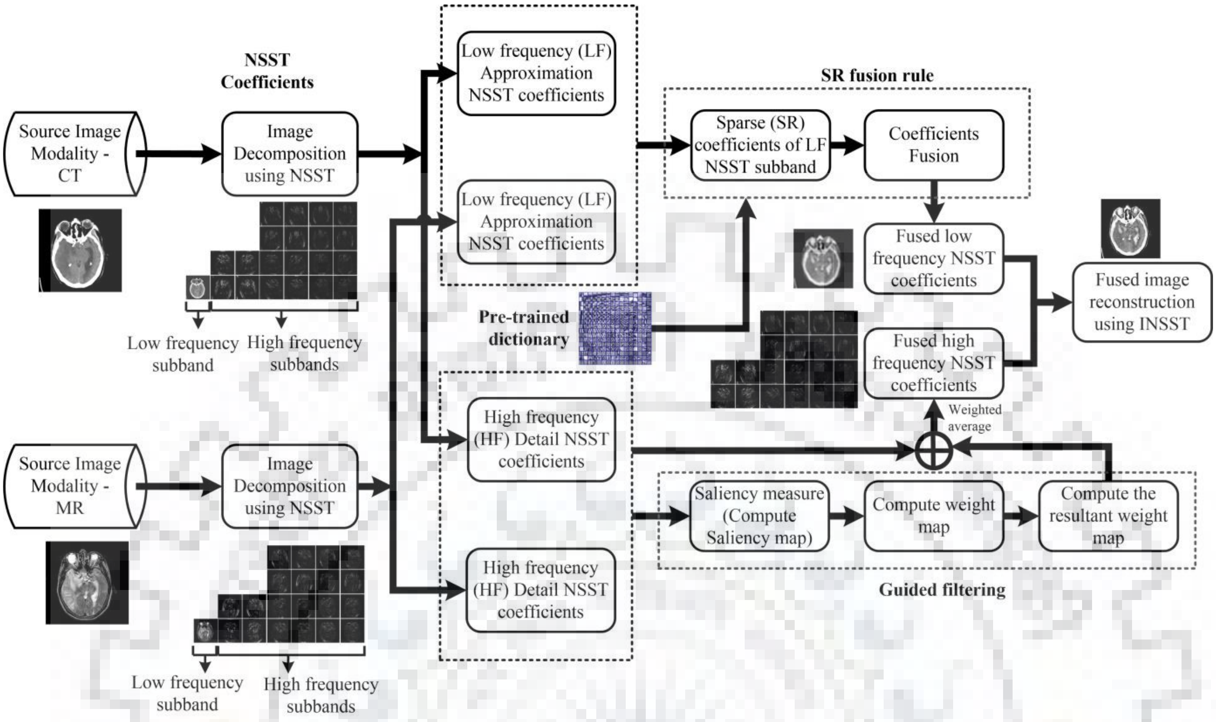


Figure 5.1 Process flow of the proposed SDL-MIF method

Implementation steps:

Let $\boldsymbol{R} = R(x, y)$ and $\boldsymbol{S} = S(x, y)$ be two reference images as CT-MR/MR-SPECT/MR-PET, and the implementation of all aforementioned approaches, the proposed SDL-MIF method is implemented as,

Step 1: Covert the RGB model of the functional SPECT image into LMS color space using the formulation given as [18],

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.3811 & 0.5783 & 0.0402 \\ 0.1967 & 0.7244 & 0.0782 \\ 0.0241 & 0.1288 & 0.8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{5.8}$$

The additional skew is suppressed by converting the LMS to logarithmic color space.

$$X = \lg L, = \lg M\ , Z = \lg S \tag{5.9}$$

Step 2: Convert the logarithmic color space into $l\alpha\beta$ color space using the mathematical expression given as,

$$\begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{5.10}$$

where l, α and β refer to the achromatic channel, chromatic yellow-blue and red-green opponent channel, respectively.

Step 3: Decompose l^{th} part the reference images (SPECT/MR) using NSST having a level of decomposition [2, 3, 3, 4] into lf and series of hf sub-band layers based on the successive trails from finer to coarser decomposition level.

$$\left[lf_{\boldsymbol{R}}^{NSST}, hf_{\boldsymbol{R}_l}^{NSST}\right] = NSST(\boldsymbol{R}) \tag{5.11}$$

and

$$\left[lf_{\boldsymbol{S}}^{NSST}, hf_{\boldsymbol{S}_l}^{NSST}\right] = NSST(\boldsymbol{S}) \text{ , for } l = 1, 2, 3, \ldots, 2^k \tag{5.12}$$

where l refers to the number of the subband at k^{th} direction at each decomposition level.

Step 4: ***Low frequency subband fusion*****:** Sparse representation based dictionary learning approach is used to fuse the lf NSST coefficients that is shown in Figure 5.2. To implement such an approach, the size of the overcomplete dictionary is considered 256 and image patch size is 8×8. The following steps are considered to fuse lf components,

(1) A sliding window technique is used to divide $lf_{\boldsymbol{R}}^{NSST}$ and $lf_{\boldsymbol{S}}^{NSST}$ into all likely image patches having size $\sqrt{k} \times \sqrt{k}$ with the step size of 1 and represented as $\left(p_{\boldsymbol{X}}^{lf}(i)\right)_{i=1}^{n}$, where $\boldsymbol{X}$ refers to the $\boldsymbol{R} = R(x,y)$ or $\boldsymbol{S} = S(x,y)$ and parameter $n = (M - \sqrt{k} + 1) \times (N - \sqrt{k} + 1)$ is the total number of computed patches within each lf NSST component and $M \times N$ is the size of source images.

(2) Rearrange $\left(p_{\boldsymbol{X}}^{lf}(i)\right)_{\boldsymbol{i}=1}^{n} = \left[p_{\boldsymbol{X}}^{lf}(1), p_{\boldsymbol{X}}^{lf}(2), \ldots, p_{\boldsymbol{X}}^{lf}(n)\right]$ into column vector $C_i^{\boldsymbol{R}}$ and $C_{\boldsymbol{i}}^{\boldsymbol{S}}$.

(3) Normalize the average value of each column vector $C_{\boldsymbol{i}}^{\boldsymbol{X}}$ to zero to obtain $\tilde{C}_{\boldsymbol{i}}^{\boldsymbol{X}}$ so that all patches contain only structural details. Normalized $k \times 1$ vectors are represented as $\tilde{C}_{\boldsymbol{i}}^{\boldsymbol{R}}$ and $\tilde{C}_{\boldsymbol{i}}^{\boldsymbol{S}}$.

$$\tilde{C}_i^{\boldsymbol{X}} = C_i^{\boldsymbol{X}} - \mu_i^{\boldsymbol{X}} \cdot \mathbf{1}_{k \times 1} \tag{5.13}$$

where the term $\boldsymbol{X}$ is used for each input $\boldsymbol{R}$ and $\boldsymbol{S}$ separately.

(4) Compute the sparse coefficients $\tilde{\alpha}_i^{\boldsymbol{R}}$ and $\tilde{\alpha}_i^{\boldsymbol{S}}$ of column vectors $C_i^{\boldsymbol{R}}$ and $C_{\boldsymbol{i}}^{\boldsymbol{S}}$, respectively using the orthogonal matching pursuit (OMP) approach [119, 132] as per given formulation.

$$\tilde{\alpha}_i^{\boldsymbol{R}} = \min_{D,\{\alpha_i^{\boldsymbol{R}}\}_{i=1}^{n}} \sum_{i=1}^{P} \left\|\alpha_i^{\boldsymbol{R}}\right\|_0 ; \qquad s.t. \ \left\|\tilde{C}_i^{\boldsymbol{R}} - \boldsymbol{D}\alpha_i^{\boldsymbol{R}}\right\|_2 < \epsilon \tag{5.14}$$

$$\tilde{\alpha}_i^{\boldsymbol{S}} = \min_{D,\{\alpha_i^{\boldsymbol{S}}\}_{i=1}^{n}} \sum_{i=1}^{P} \left\|\alpha_i^{\boldsymbol{S}}\right\|_0 ; \qquad s.t. \ \left\|\tilde{C}_i^{\boldsymbol{S}} - \boldsymbol{D}\alpha_i^{\boldsymbol{S}}\right\|_2 < \epsilon \tag{5.15}$$

where $\boldsymbol{D}$ is the learned dictionary using the K-SVD algorithm and the error tolerance > 0.

(5) Get fused sparse vector by applying “max-L1” rule to fuse $\tilde{\alpha}_{\boldsymbol{i}}^{\boldsymbol{R}}$ and $\tilde{\alpha}_{\boldsymbol{i}}^{\boldsymbol{S}}$ using the mathematical expression given below

$$\alpha_i^{\boldsymbol{F}} = \begin{cases} \tilde{\alpha}_i^{\boldsymbol{R}} ; & if \ \left\|\tilde{\alpha}_i^{\boldsymbol{R}}\right\|_1 > \left\|\tilde{\alpha}_i^{\boldsymbol{S}}\right\|_1 \\ \tilde{\alpha}_i^{\boldsymbol{S}} ; & otherwise \end{cases} \tag{5.16}$$

where $\alpha_i^{\boldsymbol{F}}$ refers to the fused sparse coefficient vector of the lf NSST subband.

(6) Finally, the lf fused image patch is computed iteratively.

$$\tilde{C}_i^{\boldsymbol{F}} = \begin{cases} \boldsymbol{D}\alpha_i^{\boldsymbol{F}} + \mu_i^{\boldsymbol{R}} \cdot 1; & if\ \alpha_i^{\boldsymbol{F}} = \alpha_i^{\boldsymbol{R}}, \forall i \in (1 \ldots\ldots P) \\ \boldsymbol{D}\alpha_i^{\boldsymbol{F}} + \mu_i^{\boldsymbol{S}} \cdot 1; & if\ \alpha_i^{\boldsymbol{F}} = \alpha_i^{\boldsymbol{S}}, \forall i \in (1 \ldots\ldots P) \end{cases} \tag{5.17}$$

(7) Repeat all above four steps as mentioned for all the computed image patches. After that, reshape the fused sparse vector back to image patches $\left(p_{\boldsymbol{F}}^{lf}(i)\right)_{i=1}^{n}$ and then place the fused patches to its original position in fused lf subband denoted by $lf_{\boldsymbol{F}}^{NSST}$. All the overlapped pixel regions in $lf_{\boldsymbol{F}}^{NSST}$ are averaged out as the image patches are overlapped.

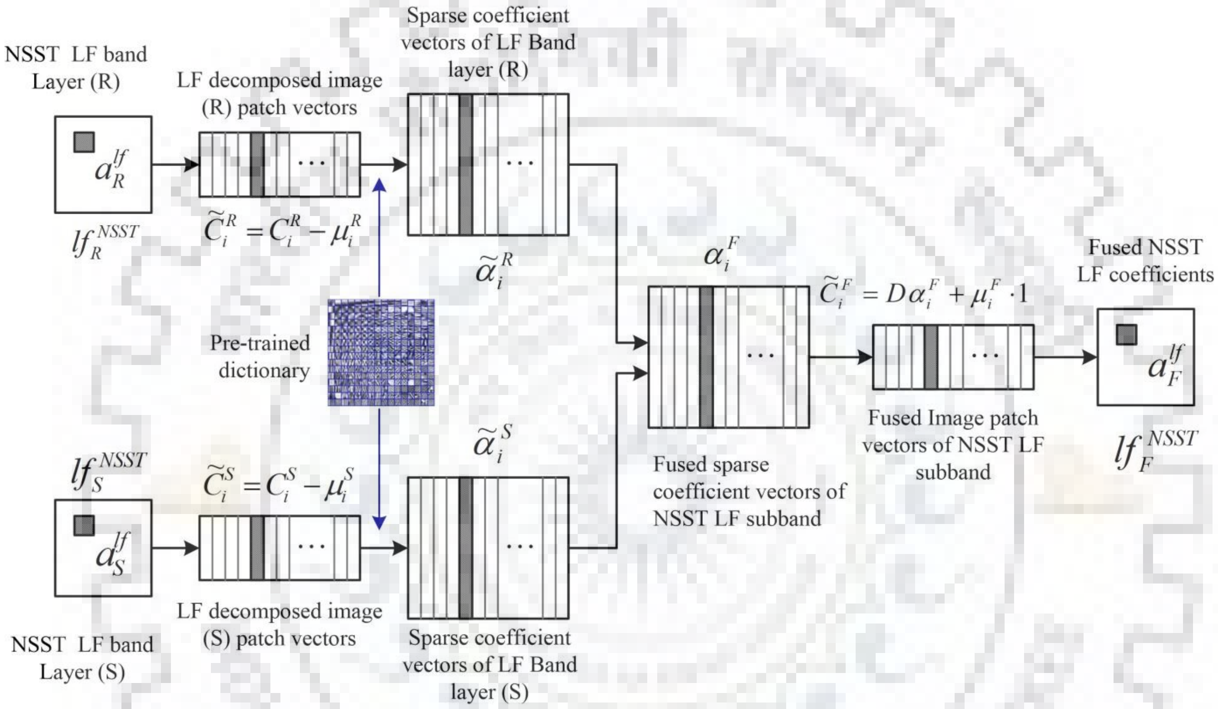


Figure 5.2 Schematic diagram for low-frequency NSST coefficient fusion

Step 5: *High frequency subband fusion*: For hf NSST subband fusion, the concept of guided filtering [94] is employed that improves the contrast level with noise filtering as much as possible. In the proposed method, the saliency map is constructed for hf subbands which contain most of the structural details, intuitively as compared to the lf subband. In order to highlight the detailed information, a Laplacian kernel is used and followed by Gaussian smoothing of the zero-mean input image to preserve the edges, reduce the noise sensitivity (controlled by the degree of smoothing) and also to provide more robust fusion performance. Since the Fourier transform of Gaussian is again Gaussian with no sharp cutoff frequency, hence it allows the high-frequency component in accordance with the decay rate of its point-spread distribution function. In addition, computational complexity and detection accuracy are the two main factors of any saliency map technique. Most of the methods such as contrast based [139], graph based [178], and deep learning based [100] is driven using multiple parameters to achieve the desired results, while Gaussian function estimates more robust saliency map using minimum parameters (only filter size and blurring degree) with less complexity. The complete high-frequency subband fusion process is shown in Algorithm 2. The following implementation points are in consideration given as,

(1) Firstly, Laplacian filtering is employed to each high pass subband image $hf_{\boldsymbol{X}_l}^{NSST}$, individually to get the edge information at each level of detail coefficients.

$$H_{\boldsymbol{X}_l} = hf_{\boldsymbol{X}_l}^{NSST} * \mathcal{L} \qquad (5.18)$$

where $\mathcal{L}_{3\times 3}$ Laplacian mask and $\boldsymbol{X} \in [\boldsymbol{R}, \boldsymbol{S}]$.

(2) Construct the saliency map for better characterization of the spatial saliency level of the detail information. The local smoothing is used for this purpose by applying a Gaussian filter,

$$\wp_{\boldsymbol{X}_l} = \left|H_{\boldsymbol{X}_l}\right| * g_{a,\sigma} \qquad (5.19)$$

where g is a Gaussian filter of size $(2a+1) \times (2a+1)$ and σ is the level of blurring. The values of a and σ are considered 4 and 2, respectively for further experiments.

(3) Now, compare the saliency level to get the corresponding weight map that is formulated as,

$$W_{\boldsymbol{X}_l}^{x} = \begin{cases} 1; & if\ \wp_{\boldsymbol{X}_l}^{x} = \max(\wp_{\boldsymbol{R}_l}^{x}, \wp_{\boldsymbol{S}_l}^{x}) \\ 0; & otherwise \end{cases} \qquad (5.20)$$

where $\wp_{\boldsymbol{X}}^{x}$ is the saliency value for the pixel x in the reference detail subimages.

(4) For optimizing the weights, guided filtering is employed instead of the optimization based rule to meet the efficiency of searching performance. Hence, guided filtering is applied to each $\wp_{\boldsymbol{X}_l}^{x}$ with its corresponding input detail subband $hf_{\boldsymbol{X}_l}^{NSST}$ considering as a guidance image.

$$D_{\boldsymbol{X}_l}^{W} = G_{r_h,\varepsilon_h}\left(W_{\boldsymbol{X}_l}, hf_{\boldsymbol{X}_l}^{NSST}\right) \qquad (5.21)$$

where the parameters r_h and ε_h refer to the guided filter parameters that are used to control the mask size and the degree of smoothness, respectively. The value of r_h and ε_h are chosen as 7 and 10^{-6}, respectively, for the proposed fusion method. $D_{\boldsymbol{X}_l}^{W}$ refers to the desired weight map corresponds to the detail information available in both the reference images. Now, normalize the weight map for both the images such that for each pixel, it sums up to 1. For restoring the edge information, small mask size and less blur level are chosen so that weight map corresponds to sharp and edge details are obtained to present the spatial consistency.

(5) Reconstruct the fused detail information given as follows,

$$hf_{\boldsymbol{F}_l}^{NSST} = D_{\boldsymbol{R}_l}^{W} \times hf_{\boldsymbol{R}_l}^{NSST} + D_{\boldsymbol{S}_l}^{W} \times hf_{\boldsymbol{S}_l}^{NSST} \qquad (5.22)$$

```
Algorithm 2: High-frequency subband fusion
Input: NSST decomposed high frequency subband components $hf_{\boldsymbol{X}_l}^{NSST}$, $\boldsymbol{X} \in [\boldsymbol{R}, \boldsymbol{S}]$.
Output: High frequency subband fused component $hf_{\boldsymbol{F}_l}^{NSST}$.
Parameters: decomposition level $(l)$, Gaussian kernel size $(a)$, degree of smoothing $(\sigma)$, guided filter size $(r_h)$ and blurring level $(\varepsilon_h)$.
Repeat: For all values of $l = 1,2,3,\ldots,2^k$
    1. Apply Laplacian filter kernel to $hf_{\boldsymbol{X}_l}^{NSST}$ to get the edge information at each level $H_{\boldsymbol{X}_l}$.
    2. Compute the local average of the absolute value $H_{\boldsymbol{X}_l}$ using Gaussian kernel to obtain a saliency map $\wp_{\boldsymbol{X}_l}$.
    3. Weight map construction: From the saliency map $\wp_{\boldsymbol{X}_l}$, compute the weight value $W_{\boldsymbol{X}_l}^{x}$ using of pixel $x$ in image $\boldsymbol{X} \in [\boldsymbol{R}, \boldsymbol{S}]$.
    4. Weight map refinement: Compute the optimized weight map using the Eq. $D_{\boldsymbol{X}_l}^{W} = G_{r_h,\varepsilon_h}\left(W_{\boldsymbol{X}_l}, hf_{\boldsymbol{X}_l}^{NSST}\right)$, where $G_{r_h,\varepsilon_h}$ refer edge preserving guide filter.
    5. Subband layer fusion: Compute the fused detail layer by Eq. (5.22).
end
```

Step 6: Apply the inverse NSST on the fused lf and hf components to reconstruct the resultant image as follows,

$$F = NSST^{-1}\left(lf_{\boldsymbol{F}}^{NSST}, hf_{\boldsymbol{F}_l}^{NSST}\right) \tag{5.23}$$

Step 7: In case of fusing the MR-SPECT images, include α and β part computed in step 2 with the fused image obtained from step 6 by the given mathematical formulation and convert it finally in RGB color format. The schematic of the proposed SDL-MIF fusion model is shown in Figure 5.3 for MR-SPECT and MR-PET images in which the shaded block of the proposed SDL-MIF approach is considered similar to as shown in Figure 5.1.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & -2 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} l \\ \alpha \\ \beta \end{bmatrix} \tag{5.24}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 4.4679 & -3.5873 & 0.1193 \\ -1.2186 & 2.3809 & -0.1624 \\ 0.0497 & -0.2439 & 1.2045 \end{bmatrix} \begin{bmatrix} 10^{X} \\ 10^{Y} \\ 10^{Z} \end{bmatrix} \tag{5.25}$$

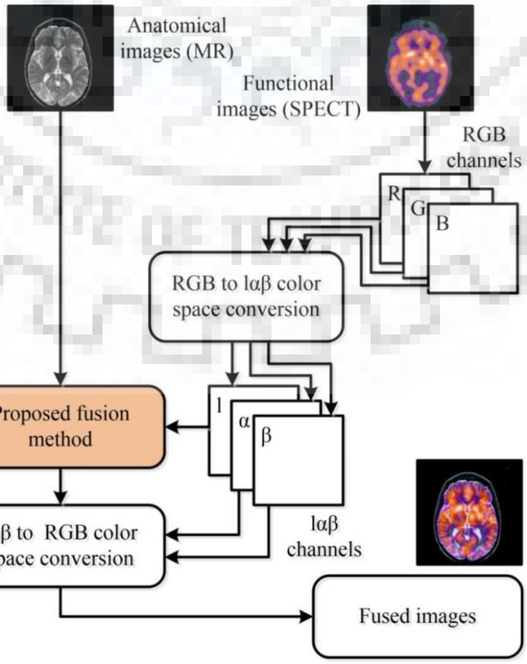


Figure 5.3 Process flow for the presented approach for anatomical and functional image fusion

The complete summarized process flow of the proposed medical image fusion based on sparse K-SVD dictionary learning approach (SDL-MIF) is shown in Algorithm 3.

```
Algorithm 3  Proposed Method
Input: Pre-registered image pair $R^{M\times N}$ and $S^{M\times N}$, pre-learned dictionary $(D^{k^2\times K})$
Output: Fused Images
Parameters: MST decomposition level $(l)$, patch size $(k)$, reconstruction error $(\epsilon)$, guided filter size $(r_h)$ and blurring level $(\varepsilon_h)$
  1. Convert the images R and S from RGB space to $l\alpha\beta$ color space by Eq. (5.8) to (5.10).
  2. Apply the NSST on $l^{th}$ component of $l\alpha\beta$ color space and get the decomposed $lf_{\boldsymbol{X}}^{NSST}$ and $hf_{\boldsymbol{X}}^{NSST}$ coefficients, where $\boldsymbol{X} = (R, S)$
  3. Low frequency subband fusion
     Loop: Repeat $\forall i \in (1,2,3,\dots,n)\ where, n = (M - \sqrt{k} + 1) \times (N - \sqrt{k} + 1)$
        • Divide the $lf_X^{NSST}$ components, $\boldsymbol{X} = (R, S)$ into patches $\left(p_{\boldsymbol{X}}^{lf}(i)\right)_{i=1}^{n}$.
        • Compute the column vectors $C_i^{\boldsymbol{X}}$ from the patches.
        • Compute $\tilde{C}_i^{\boldsymbol{X}}$ and its sparse coefficients $\tilde{\alpha}_i^{\boldsymbol{X}}$, individually.
        • Get the fused sparse vector $\alpha_i^{\boldsymbol{F}}$ and fused patch mean $\mu_i^{\boldsymbol{F}}$ using 'max-L1' rule.
        • Compute the fused column vector $\tilde{C}_i^{\boldsymbol{F}}$ using Eq.(5.17).
     Packing: Reshape each $\tilde{C}_i^{\boldsymbol{F}}$ and $p_{\boldsymbol{F}}^{lf}$ and plug the patches into its native co-ordinates.
        • Aggregate overlapped fused patches and get sparsely fused $lf_{\boldsymbol{F}}^{NSST}$ coefficients.
  4. High frequency subband fusion
     Weight map construction: For each decomposition level $hf_{\boldsymbol{X}_l}^{NSST}$, compute $H_{\boldsymbol{X}_l}$, $\wp_{\boldsymbol{X}_l}$ and $W_{\boldsymbol{X}}^{i}$ using Eq. (5.18), (5.19) and (5.20), respectively.
     Guided weight optimization: Apply guided filter $G_{r_h,\varepsilon_h}$ using Eq. (5.21).
     Layer Fusion: Get the fused detail layer using Eq. (5.22).
  5. Inverse NSST of the $lf_{\boldsymbol{F}}^{NSST}$ and $hf_{\boldsymbol{F}}^{NSST}$ components to reconstruct the fused image.
  6. Reconvert the fused output from $l\alpha\beta$ color space to RGB color space.
```

5.5 Experimentation

To evaluate the performance of the proposed method, different experiments are conducted on 65 pairs of multimodal medical images which contain 14 pairs of CT-MR, 29 pairs of MR-SPECT and 22 pairs of MR-PET images. To implement the proposed approach, NSST decomposition is employed with different decomposition levels from 1 to 4. Based on the successive experiments by evaluating the visual and quantitative analysis, the decomposition level is considered as [2,3,3,4] from coarser to a finer scale. After that, the SR-based rule is applied to merge the lf approximation band. For the same, the over-complete dictionary of size 64×256 is trained with the 120 multimodal medical image dataset and 10 outdoor images taken of different contrast, focus and camera settings. The main reason to consider large medical images and few natural image sets is to define the area overlapping (co-adaptation of the image atomic structures) and to extend the work for further visualizing the performance generalization. The medical dataset contains complex structures with high pixel contrast, while the natural dataset exhibits generic regularity in pixel saliency. An overcomplete dictionary trained with even large amount of medical images cannot represent the suitable structures of the natural image data, but the dictionary with few appropriate samples of natural data combined with larger samples of medical data may be able enough to represent appropriate natural image features. The dictionary trained in the presented work is particularly designed for generalization of two areas, keeping more medical image samples as the testing data belonging to the medical image class. The patch block size and error tolerance are considered 8 and 0.1, respectively, in a similar way as described in [191]. The iteration number for dictionary training using K-SVD is set to 200. The guided filter parameters (r_h, ε_h) are chosen as default values $r_h = 7$ and $\varepsilon_h = 10^{-6}$ as mentioned in [94]. From the experimental results, it is observed that too large filter size will smooth out the edges and too small will introduce artificial edges at the boundary. Similarly, the degree of blurring should also be low during the fusion of high band information.

For CT and MR image dataset, the proposed method is compared with the other existing fusion algorithms such as NSCT domain PCNN fusion proposed by Das and Kundu (2012) [34], guided filtering (GFF) developed by Li et al. (2013) [94], multiscale transform with a sparse representation (MST-SR) by Liu et al. (2015) [109], shearlet domain PCNN fusion developed by Ganasala and Kumar (2016) [48], fuzzy PCNN and NSCT based fusion given by Yang et al. (2017) [203], cascaded fusion model in ripplet and nonsubsampled shearlet domain C-MIF method as discussed in Chapter 2 and other fusion model using type-2 fuzzy techniques with NSCT given by Yang et al. (2017) [201]. Moreover, for fused MR-SPECT and PET image analysis, other exiting fusion approaches developed by Allanki and Bindu, 2016 [157], Bhatnagar et al., 2013 [18], Yang and Li (OMP), 2012 [192],Yang et al., 2017 [203],Yang et al., 2015 [200] and Zhu et al. (HMSD-GDGF), 2018 [218], Yang et al. (2016) [201] are also considered here. The parameters for these algorithms are set as follows:

For MST-SR approach presented by Liu et al. [109] overcomplete dictionary of size 64×256 is used to train with 40 images and learned with the K-SVD rule. The number of iterations for dictionary learning is 180 and other parameters are considered as default values mentioned in the corresponding individual method. Moreover, to investigate the efficacy of the proposed SDL-MIF approach objectively, several performance measures (En, STD, MI, IQI, FMI, XEI as discussed in previous chapters) are adapted to present the comparative evaluation. The subjective results speculate the comparative assessment of visual perception among the source and fused images while the quantitative investigation demonstrates the usefulness of the perceived results through various quality evaluation measures. Higher index values show a better fusion performance for all considered objective measures. The performance measures do not only show the comparative result analysis, but also verify the acceptability of visual observation by preserving the information and structural contents of the source images.

5.6 Results and Discussions

5.6.1 Experiment 1: Comparisons of anatomical-anatomical image fusion results

This experiment discusses the results of fusion obtained by the proposed SDL-MIF method. The preprocessing step ensures decent geometrical matching and confined resolution for further processing of input data. The experimental results will reveal the fusion of these multimodal neurological images, providing a single fused image having soft and hard both the tissues and structural information from the different MR and CT images. The set-1 (MR-T2 and CT) given in Figure 5.4 (a) and (b) includes the pair-1, pair-2, and pair-3 shows an acute stroke of Alexia, multiple embolic infarcts (refractory focal seizures) and cerebral toxoplasmosis (formation of cystic structure around focal lesions), respectively with their fused results using the proposed method as shown in Figure 5.4 (c).

In the first pair of source images, the infarct is seen in the CT and MR image on the left side of the splenium (thick posterior part) of the corpus callosum of the brain and left posterior cerebral artery territory. In the third pair of the image, MR image reveals focal lesions involving basal ganglia and formation of cystic structure, while CT reveals the calcification. In all three pairs of CT and MR images shown in Figure 5.4, it is observed that the significant information content is increased in the fused images compared to individual source images. This is also evident by providing higher values of the evaluation metrics En, STD, MI and XEI for the proposed SDL-MIF method. Furthermore, to evaluate the superiority and validate the proposed fusion method with the addition of several experiments done so far, the quantitative evaluation and visual comparison is given with selected seven decently performed state-of-the-art approaches based on single as well as integrated transformation techniques in the cascaded model.

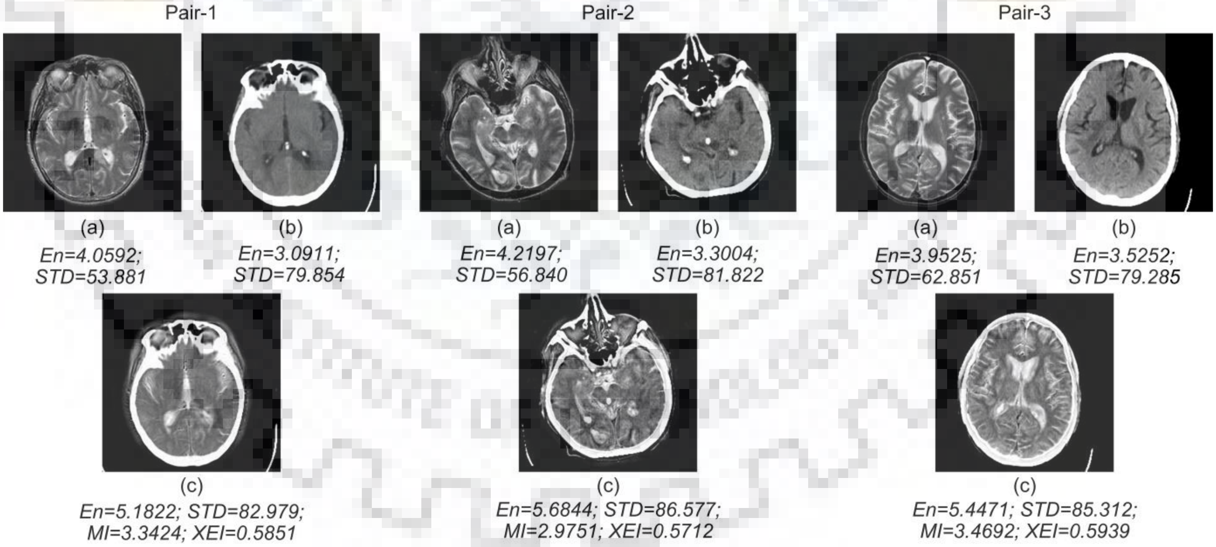
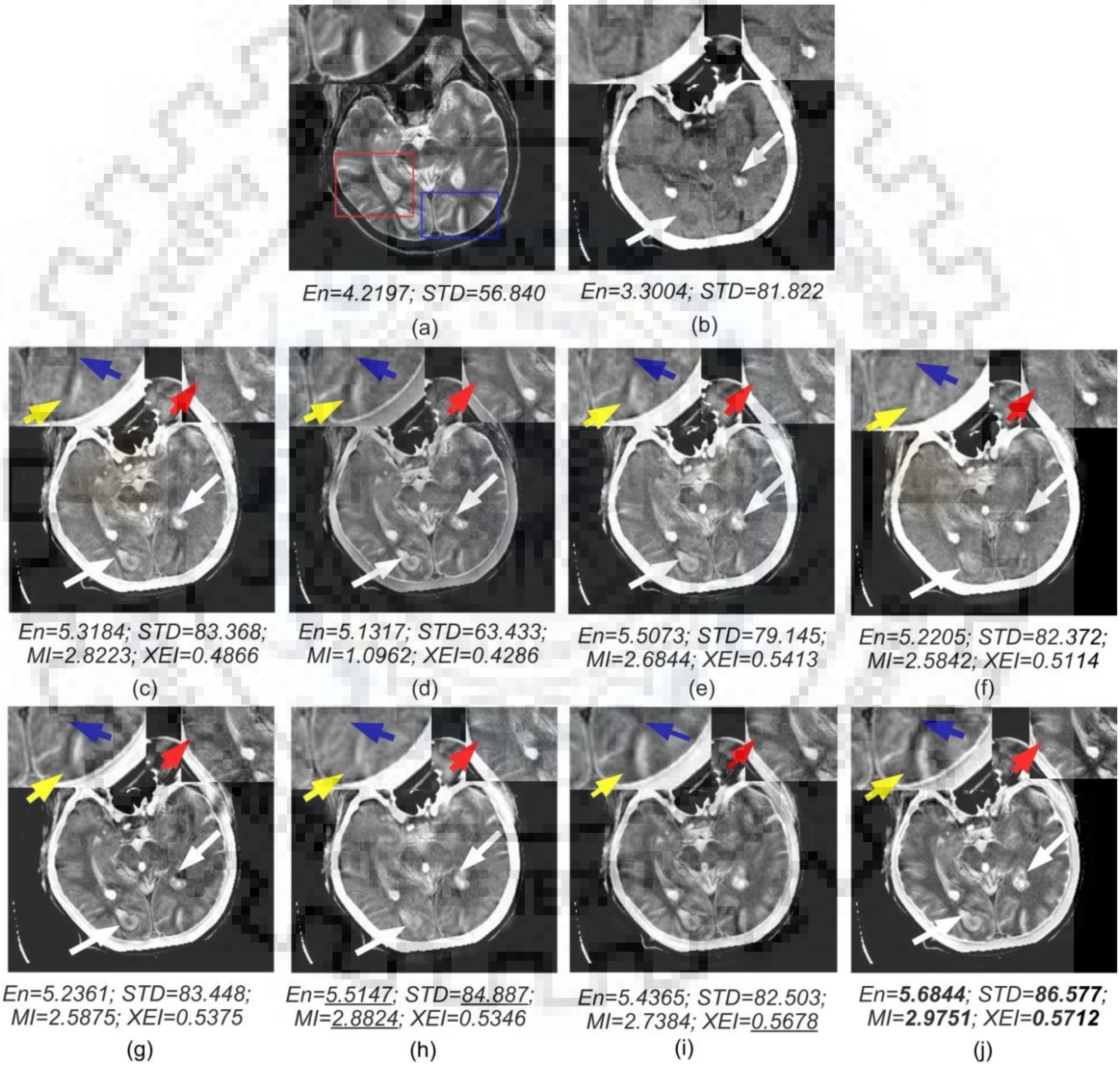


Figure 5.4 (a) Source MR image (b) Source CT Image (c) Image fused by the proposed SDL-MIF method

Figure 5.5 and Figure 5.6 show the fusion results of two different sets of CT and MR images. Figure 5.5 (a) and (b) shows the MR and CT scans, respectively, of a patient having multi-infarcts. Those are pointed with different arrows in the CT image of this case. Similarly, Figure 5.6 shows MR and CT scans of another patient suffered from Fatal stroke and developed an enlarged right pupil that can be seen in the CT and MR images for this case. Figure 5.5 (c) and Figure 5.6 (c) present the results of Das and Kundu model that is based on external features motivated PCNN in NSCT domain. This method is able to restore the average information with a correct level of contrast, but fails to capture the small edge details as can be seen in a zoomed region. Figure 5.5 (d) and Figure 5.6 (d) show the fused images produced by the GFF model which depict the robustness against artifacts with true chromatic balance and genuine brightness without altering the edge details. However, this method suffers from low contrast and poor information capture (tiny details of the first image are smoothed out). From Figure 5.5 (e) and Figure 5.6 (e), MST-SR based fusion model captures the information and structural details very well that is also validated by significantly improved entropy values. However, the results in over brightness as spatial consistency is not well preserved as indicated by the lower value of STD and MI.

Figure 5.5 (f) and Figure 5.6 (f) show the fusion results produced by another fusion model developed by Ganasala and Kumar [48] based on NSST with feature motivated PCNN. This method is unable to capture the pixel variations around the focal part and occipital axial regions shown by zoomed regions of the source images. The fusion results of Yang et. al. [203] presented a fuzzy adaptive PCNN based fusion method is shown in Figure 5.5 (g) and Figure 5.6 (g). This result outperforms other previously mentioned approaches in terms of better visual quality with a decent level of edge index and average information but little blurred at fluidic regions. In Figure 5.5 (h) and Figure 5.6 (h), fusion results of the C-MIF approach preserve the significant information indicated by the higher En and MI values as well as image brightness (validated by the higher STD values) closely similar to the original one but edge details are little smoothed at the area has tiny variations. Figure 5.5 (i) and Figure 5.6 (i) show the fused images obtained by another NSCT domain method using type-2 fuzzy logic [201] that are able to capture the available detail from the source images in the fused result. However, some representative regions (zoomed area) show lack of image sharpness and poor edge boundary definition which are also validated with quantitative results. The STD values show low sharpness and the moderated scores of XEI shows that the edges are preserved well but fine lines near the edge boundary are not defined clearly. The proposed SDL-MIF method produces the fused images shown in Figure 5.5 (j) and Figure 5.6 (j) having all principal and tiny details with appropriate brightness and contrast level, as the resultant image is neither falsely brighter nor has blurred of tiny edges. Besides this, the proposed SDL-MIF method is able to preserve pixel spatial consistency, precisely and edge information is also restored without adding any artifacts.

In addition to this, the visual assessment of fused images developed by all the methods has also been validated objectively. Li et al. [94] and Liu et al. [109] methods are producing comparatively lower STD values for the resultant fused images which show that these methods suffer from contrast reduction. Though, these methods yield higher En values indicating higher information content in the fused images compared to source images. The proposed method yields the highest values of all the evaluation measures En, MI, STD, and XEI in both the cases. Moreover, a detailed quantitative comparison of complete CT-MR dataset has also been made and the objective evaluation measures are mentioned in Table 5.1 in terms of average and standard deviation of all performance measures which reflect the higher averaged value of En, MI, STD and XEI measures computed by the proposed SDL-MIF method compared to all other considered fusion methods.



En=4.2197; STD=56.840
(a)

En=3.3004; STD=81.822
(b)

En=5.3184; STD=83.368; MI=2.8223; XEI=0.4866
(c)

En=5.1317; STD=63.433; MI=1.0962; XEI=0.4286
(d)

En=5.5073; STD=79.145; MI=2.6844; XEI=0.5413
(e)

En=5.2205; STD=82.372; MI=2.5842; XEI=0.5114
(f)

En=5.2361; STD=83.448; MI=2.5875; XEI=0.5375
(g)

En=5.5147; STD=84.887; MI=2.8824; XEI=0.5346
(h)

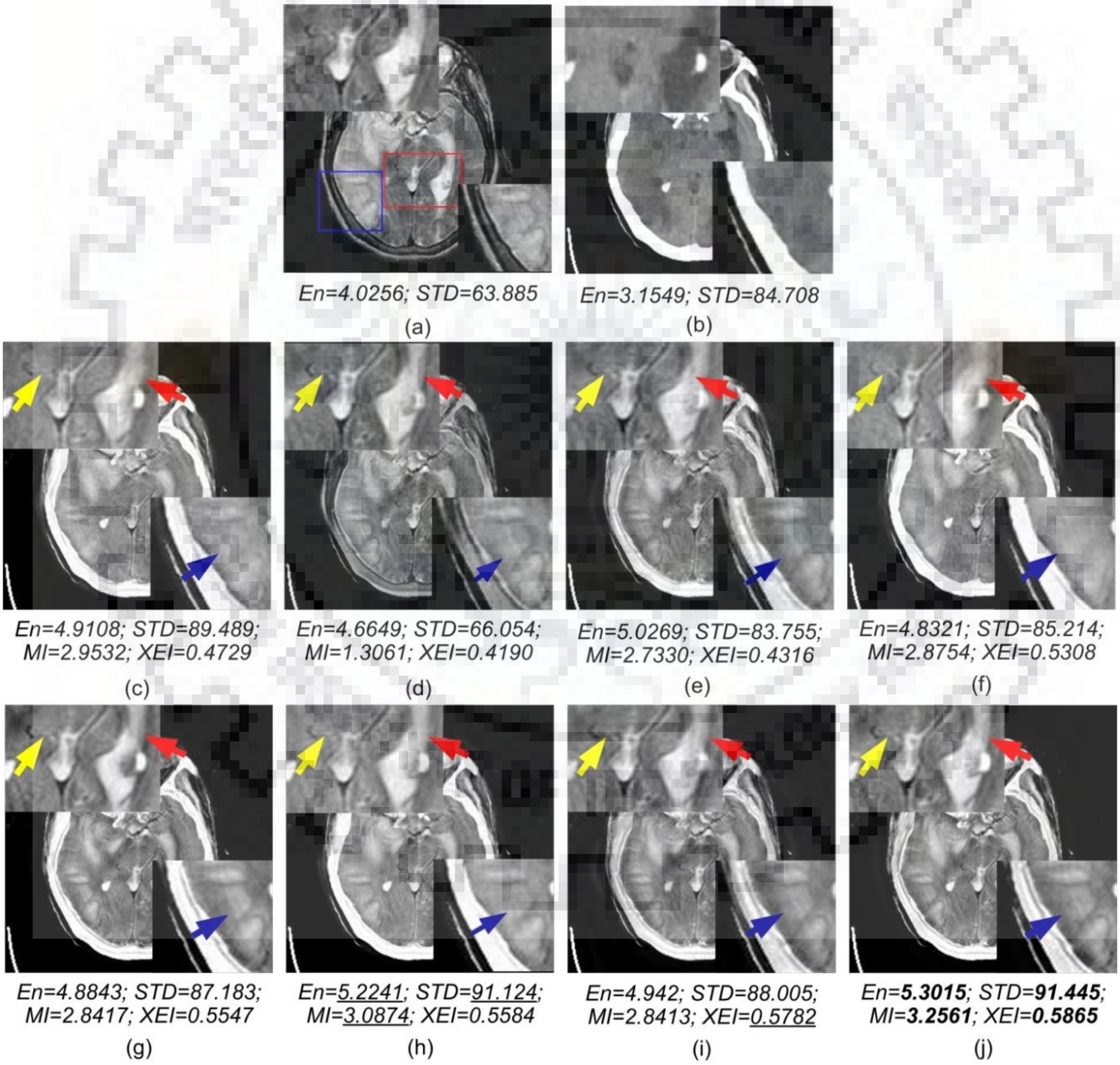
En=5.4365; STD=82.503; MI=2.7384; XEI=0.5678
(i)

*En=**5.6844**; STD=**86.577**; MI=**2.9751**; XEI=**0.5712***
(j)

Figure 5.5 (a) MR images with two zoom regions marked by red and blue color (b) CT image, fused images along with the corresponding zoomed regions obtained by (c) Das and Kundu, 2012 [34] (d) Li et al. (GFF), 2013 [94] (e) Liu et al. (MST-SR), 2015 [109] (f) Ganasala and Kumar, 2016 [48] (g) Yang et al., 2017 [203] (h) C-MIF method (i) Yang et al. (NSCT-Type 2 Fuzzy), 2016 [201] (j) Proposed SDL-MIF method

Table 5.1 Averaged evaluation measures (mean ± standard deviation) for CT-MR image dataset

Fusion Methods	Performance Measures			
	En	MI	STD	XEI
Source MR	3.8803 ± 0.3366	-	59.3509 ± 4.9843	-
Source CT	3.0647 ± 0.3035	-	83.9079 ± 5.4953	-
Das and kundu, 2012 [34]	4.8081 ± 0.3713	3.0086 ± 0.1102	86.9047 ± 4.4341	0.4785 ± 0.0490
Li et al. (GFF), 2013 [94]	4.6501 ± 0.2949	1.3093 ± 0.2232	64.0659 ± 3.7542	0.4458 ± 0.0530
Liu et al. (MST-SR), 2015 [109]	5.1369 ± 0.2105	2.8894 ± 0.1578	82.8205 ± 3.9480	0.4824 ± 0.0416
Ganasala and Kumar, 2016 [48]	4.9135 ± 0.2426	2.8656 ± 0.1751	84.0917 ± 4.3443	0.5199 ± 0.0513
Yong et al., 2017 [203]	5.0273 ± 0.2126	2.9026 ± 0.2145	84.9058 ± 4.3424	0.5477 ± 0.0436
C-MIF	5.1529 ± 0.2378	3.1594 ± 0.1637	88.5082 ± 4.4684	0.5578 ± 0.0465
Yang et al. (2016) [201]	4.8686 ± 0.4170	2.9393 ± 0.1469	85.8544 ± 4.2374	0.5659 ± 0.1328
Proposed SDL-MIF	5.5652 ± 0.2011	3.4027 ± 0.2128	89.4922 ± 3.8922	0.5881 ± 0.0438



En=4.0256; STD=63.885
(a)

En=3.1549; STD=84.708
(b)

En=4.9108; STD=89.489; MI=2.9532; XEI=0.4729
(c)

En=4.6649; STD=66.054; MI=1.3061; XEI=0.4190
(d)

En=5.0269; STD=83.755; MI=2.7330; XEI=0.4316
(e)

En=4.8321; STD=85.214; MI=2.8754; XEI=0.5308
(f)

En=4.8843; STD=87.183; MI=2.8417; XEI=0.5547
(g)

En=5.2241; STD=91.124; MI=3.0874; XEI=0.5584
(h)

En=4.942; STD=88.005; MI=2.8413; XEI=0.5782
(i)

*En=**5.3015**; STD=**91.445**; MI=**3.2561**; XEI=**0.5865***
(j)

Figure 5.6 (a) MR images with two zoom regions marked by red and blue color (b) CT image, fused images along with the corresponding zoomed regions obtained by (c) Das and Kundu, 2012 [34] (d) Li et al. (GFF), 2013 [94] (e) Liu et al. (MST-SR), 2015 [109] (f) Ganasala and Kumar, 2016 [48] (g) Yang et al., 2017 [203] (h) C-MIF method (i) Yang et al. (NSCT-Type 2 Fuzzy), 2016 [201] (j) Proposed SDL-MIF method

5.6.2 Experiment 2: Comparisons of anatomical-functional image fusion results

The MR-T2 represents structural details of soft tissues to highlight the fat and water (unusual brightness of fatty and water-based tissue indicates the growth of abnormality). The metabolic alterations (shape and size) of such abnormalities are not discriminated in MR imaging due to low accuracy. Hence, functional imaging modality as SPECT and PET is referred to reflect the significant information. The pair-1 and pair-2 of MR-T2 and 99mTc-SPECT are shown in Figure 5.7 (a) and (b), respectively. The pair-1and pair-2 correspond to cerebrovascular disease (brain attack) and neoplastic disease (Astrocytoma-type of brain tumor), respectively. The pair-3 and pair-4 shown in Figure 5.7 (a) and (b) correspond to mild Alzheimer disease of a patient with different slices of brain MR-T2 and PET-FDG to detect the changes in anterior temporal and posterior parietal regions. The corresponding fused images are presented in Figure 5.7 (c) obtained by the proposed SDL-MIF method that produces a better quality of fused images having bright regions of the MR image and color content taken from the SPECT/PET images.

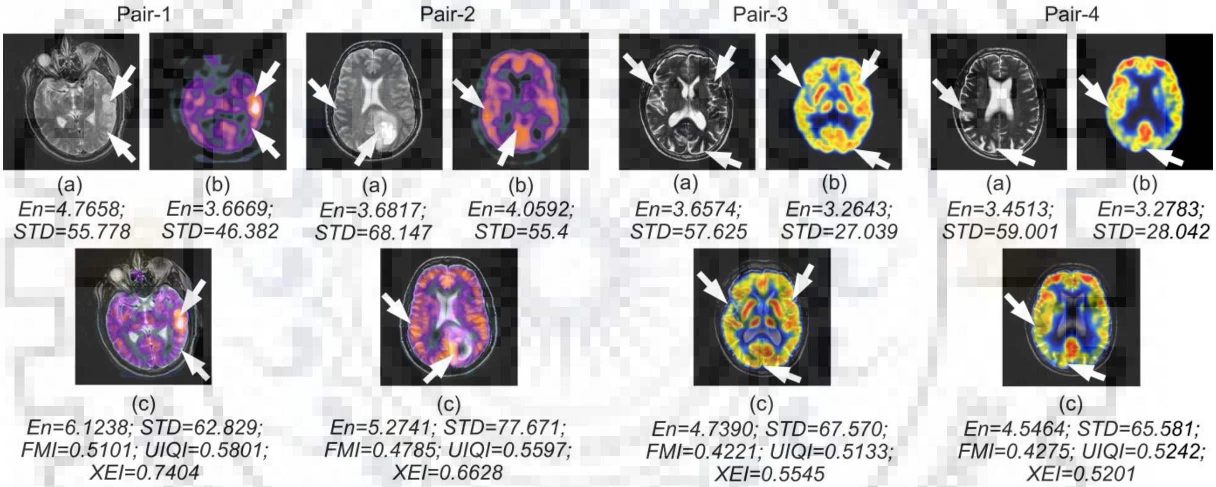


Figure 5.7 Fused images using the proposed SDL-MIF method on various test neurological images along with quantitative values. Pair 1 and 2 input (a) MR image (b) SPECT image (c) Fused image, Pair-3 and 4 input image (a) MR image (b) PET image (c) Fused image

In all four sets of source images (Pair-1, 2, 3 and 4, respectively), the details along with the spectral information have also been preserved in the fused images. To validate the results of the proposed method, higher values of quantitative measures En and STD compared to source images are also mentioned. Moreover, other measures mentioned in Table 5.2 and Table 5.3 as FMI, IQI, and XEI are also computed for fused images. To illustrate the ability of the proposed SDL-MIF method, another multimodal image pair comprising MR-T2 and 99mTc-SPECT is shown in Figure 5.8 (a) and (b), respectively. This case belongs to a patient that has sub-acute stroke failed to explore the left half space. In this case, the high signal intensity in the superior frontal Gyrus region and hyperperfusion in the right posterior parietal lobe is seen in MR-T2 and 99mTc-SPECT image. Similarly, another pair of MR-T2 and the 99mTc-SPECT image is shown in Figure 5.9 (a) and (b), respectively, where MR image reveals an area of mixed intensity in the left occipital area, while a border of the cystic lesion is seen in 99mTc-SPECT. Therefore, two pairs out of 29 MR-SPECT image dataset are presented in Figure 5.8 (a)-(b) and Figure 5.9 (a)-(b) for comparative analysis. The objective analysis of these image pairs is mentioned in Table 5.2 and Table 5.3 with a statistical comparison between proposed and other fusion methods. In Figure 5.8 (c) and Figure 5.9 (c), Allanki and Bindu [157] method based on spectral parameter variance shows good spectral balance and restores the detail information as well, but the abnormal zone present in the source images become little hazy.

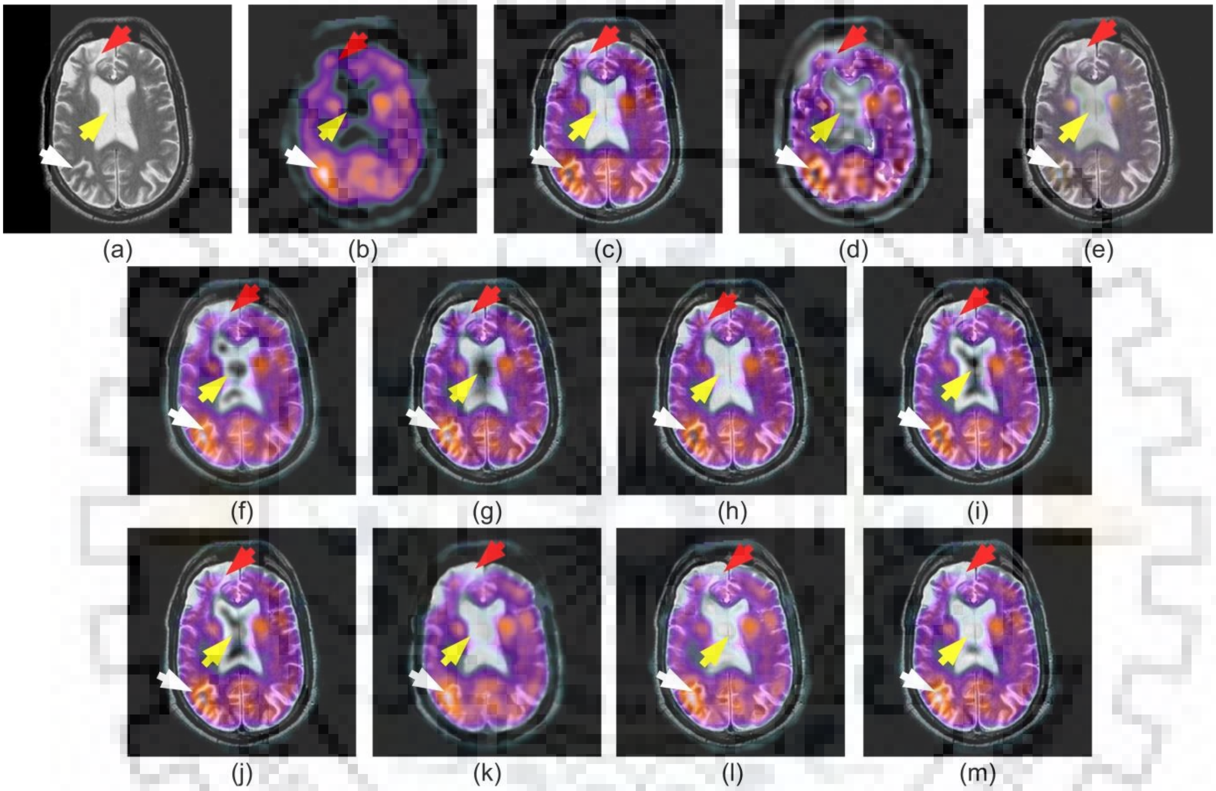


Figure 5.8 Input image (a) MR image (b) SPECT image, qualitative comparisons of fused images obtained by (c) Allanki and Bindu, 2016 [157] (d) Bhatnagar et al., 2013 [18] (e) Li et al. (GFF), 2013 [94] (f) Yang and Li (OMP), 2012 [192] (g) Liu et al., (MST-SR) 2015 [109] (h) Ganasala and Kumar, 2016 [48] (i) Yang et al., 2017 [203] (j) Yang et al., 2015 [200] (k) Zhu et al. (HMSD-GDGF), 2018 [218] (l)Yang et al. (NSCT-Type2 Fuzzy), 2016 [201] (m) Proposed SDL-MIF method

The quantitative results also indicate high contrast images, but comparatively low information content. Figure 5.8 (d) and Figure 5.9 (d) show the fusion results of recent work of Bhatnagar et al. [23] that incorporates directive contrast in NSCT domain. This provides less significant edge information with a blurred area of tonal difference for the complete dataset. However, in Figure 5.8 (e) and Figure 5.9 (e) the GFF based method gives a moderate result for complementary and edge details with noise suppression, while limited by low variation of color intensity and proportions. In Figure 5.8 (f) and Figure 5.9 (f), Yang and

Li [192] approach show better preservation of the abnormal tissue area, but edge details become little vague in visual presentation.

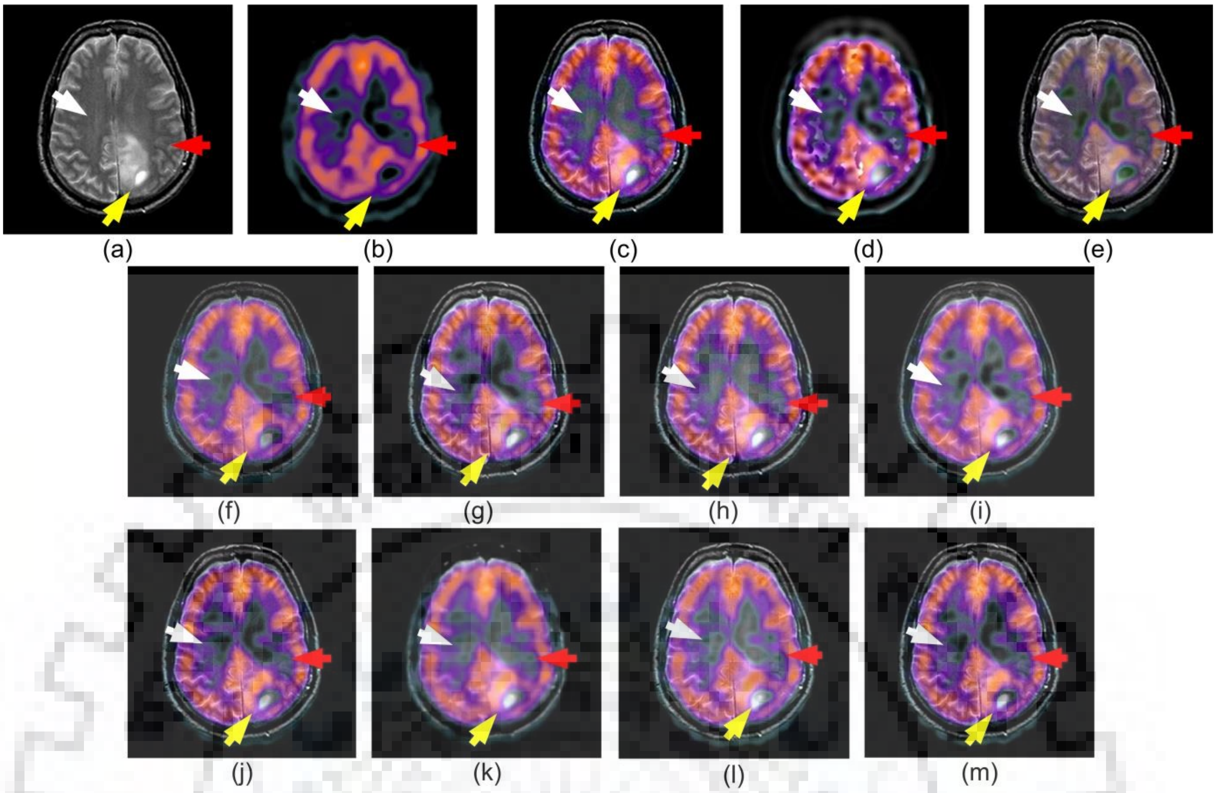


Figure 5.9 Input image (a) MR image (b) SPECT image, qualitative comparisons of fused images obtained by (c) Allanki and Bindu, 2016 [157] (d) Bhatnagar et al., 2013 [18] (e) Li et al. (GFF), 2013 [94] (f) Yang and Li (OMP), 2012 [192] (g) Liu et al., (MST-SR) 2015 [109] (h) Ganasala and Kumar, 2016 [48] (i) Yang et al., 2017 [203] (j) Yang et al., 2015 [200] (k) Zhu et al. (HMSD-GDGF), 2018 [218] (l)Yang et al. (NSCT-Type2 Fuzzy), 2016 [201] (m) Proposed SDL-MIF method

Similarly, the SR-based fusion framework proposed by Liu et. al. [109] (Figure 5.8 (g) and Figure 5.9 (g)) presents competitive objective measures (in terms of En, STD, FMI, IQI, and XEI values) compared to all above four methods, but some focal lesions having different information and morphology are combined with others, causing poor visual perception for some areas. Figure 5.8 (h) and Figure 5.9 (h) show an appropriate score of average informative content and spatial consistency, but the fine details are blurred in some salient regions indicated out by arrows in subjective results which are also validated by the evaluated measure mentioned in Table 5.2 and Table 5.3. Figure 5.8 (i) and (j) and Figure 5.9 (i) and (j) obtained by the recent works [200, 203] show that the fused results have clear visual quality with good edge preservation, but arrows labeled regions are little fuzzy, whereas the recent work of Zhu et al. [218] yielded lower value of the En, STD, FMI, and XEI compared to others as mentioned above and shows the fused images in Figure 5.8 (k) and Figure 5.9 (k). Also, from the visual results, it is clearly seen that the principal information is well preserved to some extent, but results in poor edge details and a blurred vision of the fused image. Figure 5.8 (l) and Figure 5.9 (l), show the fusion results computed by Yang et al. [201] that are good to preserve the brightness properly in color functional images, but some detail information is lost in the fused outcomes with lower contrast level. Moreover, the quantitative result also defines a lower value of XEI and FMI, while the En shows a significant score. The corresponding fused images obtained by the proposed SDL-MIF method is shown in Figure 5.8 (m) and Figure 5.9 (m). These images show that the method can extract most of the informative content and fine edge details without incorporating any artifact and smoothing effects. The regions marked by arrows in Figure 5.8 and Figure 5.9, suffered from chromatic imbalance and blurring in most of the cases used for comparative analysis for MR-SPECT image dataset, while the proposed framework is able to reproduce the significant visual features with the preservation of structural and spectral information content. Thus, it provides a clear picture of any tiny details and abnormal tissues available in any one of the source images.

Table 5.2 Performance assessment of existing fusion approaches with the proposed SDL-MIF method for MR-SPECT image pair shown in Figure 5.8

Image Seq.	Performance Measures				
	En	STD	FMI	IQI	XEI
(c)	4.5878	71.9444	0.39373	0.34137	0.67623
(d)	4.6861	65.4292	0.20079	0.25517	0.14427
(e)	4.5575	59.5072	0.38052	0.34543	0.60297
(f)	4.7034	66.6461	0.38596	0.37870	0.69423
(g)	4.8008	67.6616	0.43867	0.39833	0.70040
(h)	5.0555	71.9812	0.39851	0.34752	0.68119
(i)	4.7683	68.9441	0.43650	0.39856	0.69059
(j)	4.5983	71.9412	0.44404	0.33209	0.71533
(k)	5.0327	70.2975	0.22410	0.39840	0.25658
(l)	5.2762	72.6751	0.27783	0.38912	0.49117
(m)	**5.4035**	**72.6461**	**0.48596**	**0.41357**	**0.75781**

Table 5.3 Performance assessment of existing fusion approaches with the proposed SDL-MIF method for MR-SPECT image pair shown in Figure 5.9

Image Seq.	Performance Measures				
	En	STD	FMI	IQI	XEI
(c)	4.3316	64.2506	0.35116	0.33994	0.61903
(d)	4.5028	64.4275	0.20719	0.31860	0.17087
(e)	4.5323	53.3025	0.34184	0.40834	0.60361
(f)	4.6602	60.2365	0.35542	0.40834	0.60480
(g)	4.7599	62.8014	0.41782	0.41274	0.61487
(h)	4.8387	64.2506	0.35550	0.35343	0.6214
(i)	4.5904	64.0905	0.40054	0.51628	0.63143
(j)	4.3769	65.4121	0.42116	0.37795	0.64548
(k)	4.8066	63.2522	0.23893	0.54317	0.21332
(l)	5.0614	64.7547	0.31565	0.42817	0.43183
(m)	**5.2619**	**65.9203**	**0.45952**	**0.70961**	**0.64693**

Furthermore, Table 5.4 is presented to provide a snapshot of a detailed comparative analysis of all quantitative measures computed by all 10 fusion methods for all the MR-SPECT image dataset. This comparative analysis is tabulated to assess the image quality in term of information and edge detail content. The averaged measures indicate that the fusion outcomes obtained by the proposed method have most of the information with appropriate spatial consistency and spectral balance. The fused results also retain the high degree of structural similarity and a mutual connection to the source images.

Table 5.4 Averaged evaluation measures (mean ± standard deviation) for MR-SPECT images

Fusion Methods	Performance Measures				
	En	STD	FMI	IQI	XEI
Method [157]	4.1225 ± 0.6795	68.112 ± 8.8127	0.4114 ± 0.0222	0.3152 ± 0.0635	0.6669 ± 0.1316
Method [18]	4.3566 ± 0.6876	65.146 ± 7.5963	0.1948 ± 0.0135	0.2455 ± 0.0495	0.1394 ± 0.0253
Method [94]	4.3049 ± 0.6636	57.343 ± 9.1360	0.3855 ± 0.0515	0.3777 ± 0.0607	0.6648 ± 0.0621
Method [192]	4.4827 ± 0.7329	65.131 ± 8.3147	0.4130 ± 0.0472	0.3852 ± 0.0632	0.6708 ± 0.0619
Method [109]	4.6985 ± 0.6626	65.822 ± 7.4958	0.4440 ± 0.0302	0.3973 ± 0.0646	0.6773 ± 0.0650
Method [48]	4.7641 ± 0.6204	68.739 ± 8.2240	0.4031 ± 0.0516	0.3783 ± 0.0406	0.6753 ± 0.0673
Method [203]	4.5966 ± 0.6623	67.262 ± 8.5285	0.4447 ± 0.0349	0.4020 ± 0.0404	0.6842 ± 0.0645
Method [200]	4.3564 ± 0.6686	68.209 ± 9.0919	0.4493 ± 0.0348	0.3376 ± 0.0598	0.6879 ± 0.0621
Method [218]	4.8645 ± 0.6922	67.468 ± 7.3426	0.2132 ± 0.0211	0.3988 ± 0.1305	0.2362 ± 0.0366
Method [201]	4.8688 ± 0.6712	69.113 ± 7.2812	0.2593 ± 0.0379	0.4165 ± 0.1227	0.6169 ± 0.1221
Proposed SDL-MIF	**5.1978** ± 0.5431	69.511 ± 8.6855	**0.4928** ± 0.0150	**0.4748** ± 0.0458	**0.7258** ± 0.0606

In addition to evaluate the performance of the proposed SDL-MIF method, another comparison is also done with the sparse dictionaries trained by the different images such as either only on natural images or on medical and outdoor images as discussed above. Such comparison is done on all the source MR-CT and MR-SPECT image dataset. Figure 5.10 shows a detailed visual and quantitative comparison between the fused images obtained by the proposed SDL-MIF method based on different dictionaries built with only natural or medical and outdoor images as shown in Figure 5.10 (c1), (c2) and (d1), (d2) respectively. From Figure 5.10, it is clearly shown that fused results obtained by the proposed method in which dictionary trained by only natural images is able to capture the main information well, but limited due to poor structural detail extraction. The small portions at the boundary of the CT image (shown by the red arrow) for pair-1 and 2 are poorly captured while for pair-3 the MR image focal area is not well defined. Similarly, for the color images, the result shown in Figure 5.10 (c2) shows comparatively lower contrast and smoothed edge details than the fused image as shown in Figure 5.10 (d2) for all image pairs shown in Figure 5.10. The quantitative scores computed for all the image pairs show the lower values of En, MI, XEI and FMI metrics. Moreover, the higher values of the STD computed for Figure 5.10 (d1) and (d2) validates the superiority (higher contrast level) of the proposed SDL-MIF method in which dictionary is trained with a combination of medical and outdoor images. From all the experimental results, the overall comparative analysis concludes that the overcomplete dictionary learned of the proposed SDL-MIF method works better than the dictionary trained using natural images only.

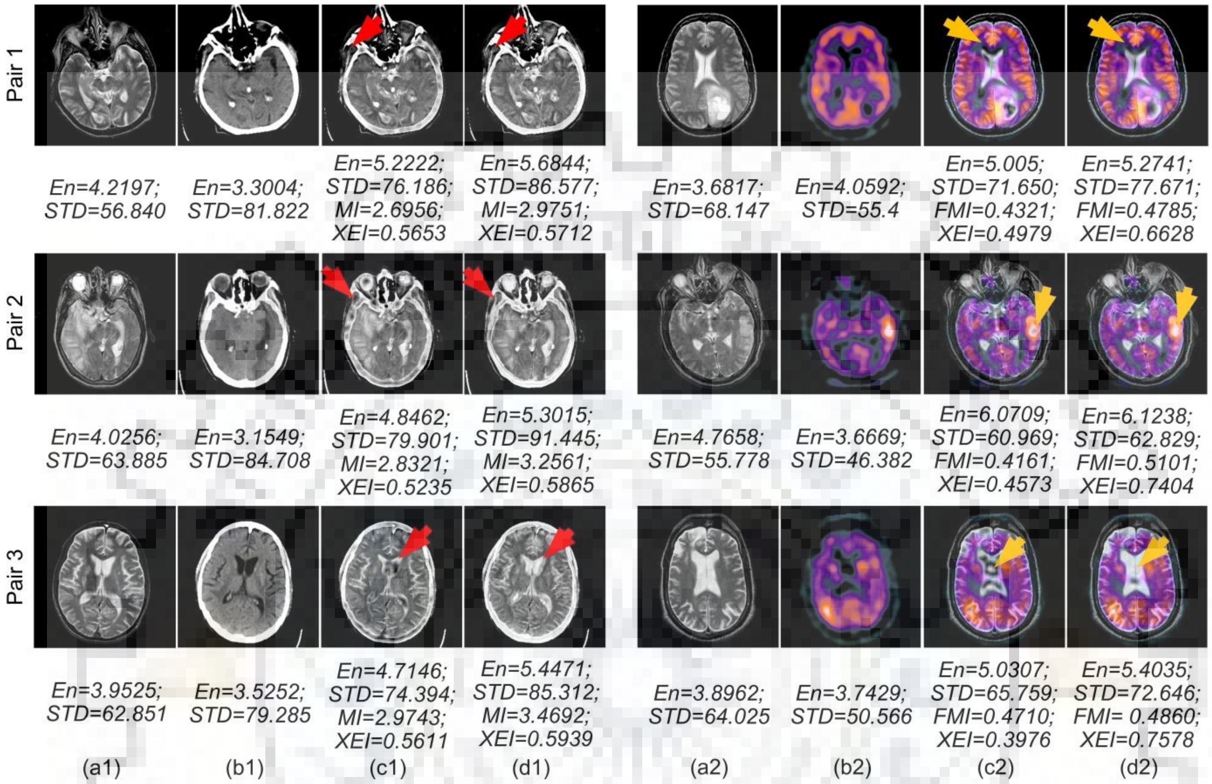


Figure 5.10 (a1) and (a2) Source MR images (b1) Source CT images (b2) Source SPECT images, fused images based on the proposed method in which dictionary is trained by (c1) and (c2) only natural images (d1) and (d2) medical and outdoor images

5.6.3 Computational complexity analysis

In this section, the computational efficiency of the proposed approach is discussed and compared with the other state-of-the-art methods. The comprehensive algorithm complexity and computational time of the proposed method are compared with the different multimodal image fusion methods and given in Table 5.5. All the experiments have been performed using MATLAB 2014a and a machine having Intel™ Core2Duo 2.93 GHz processor with 8 GB RAM on Windows 10. The computational complexity of the proposed approach is explained as follows.

For the number of image pixels (M), dictionary size (K), patch size (N), the number of computations required for the NSST decomposition with n^{th} level and j^{th} shear directional filter coefficient is $O(2^{n_j} + 1)$, the sparse representation needs $O(KT_oLN)$ where, L is total patch number and T_o is the total non-zero entries in the sparse vector matrix and the guided filter requires $O(M)$. Hence, the total complexity of the proposed method is $O\left(M\left(\left(2(2^{n_j}+1)\right)+KT_oLN+1\right)\right)$.

In the proposed method, most of the time has been taken by the dictionary construction (≈115 sec for one image of size 256×256) that can be significantly reduced by offline learning method (pre-learned dictionary) for the same multimodal medical image dataset. Moreover, the averaged running time required for the processing of different fusion methods is also computed and listed in Table 5.5; from which it is evident that the proposed algorithm takes ≈ 4-5 second to complete the process of the proposed SDL-MIF method with the pre-trained dictionary. In some of the cases, shown in Table 5.5, some state of the art methods take less processing time compared to the proposed method, but still, the proposed approach produces better fusion quality along with detailed information available in the source images and definitely outperforms the other methods.

Table 5.5 Comparison of averaged running time taken by different fusion methods

Fusion Methods	Running Time (second)
Yang et al., 2010 [198]	0.5820 ± 0.0305
Das and Kundu., 2012 [34]	62.9089 ± 4.6848
Bhateja et al., 2015 [16]	9.7980 ± 0.4876
Li et al. (GFF), 2013 [94]	0.4552 ± 0.0507
Yang et al., 2015 [200]	7.8319 ± 0.5935
Liu et al., (MST-SR), 2015 [109]	22.5098 ± 0.9849
Ganasala and Kumar, 2016 [48]	6.1528 ± 0.4892
Yang et al., 2017 [203]	89.2823 ± 6.7748
Zhu et al. (HMSD-GDGF), 2018 [218]	28.3125 ± 0.6421
Yang et al. (NSCT-Type2 Fuzzy), 2016 [201]	10.1611 ± 0.5842
Proposed SDL-MIF	5.3122 ± 0.0416

5.7 Summary

This chapter presents an effective fusion framework based sparse K-SVD dictionary learning and guided filtering. The proposed method employed the NSST as multiscale and multidirectional decomposition to maintain both the spatial as well as spectral information well using shearing filters. The dictionary learning based on sparse representation based fusion rule is adopted to improve the comprehensive information in the low-frequency NSST subband, while guided filtering is employed on high frequency NSST components to extract the salient features available in the source image and to retain the color and edge details very well without incorporating any artifacts. For anatomical-functional image fusion, the color scale conversion from RGB space to $l\alpha\beta$ space is also used that shows the descent correlation between color channels. Hence, significant results are found over conventional grayscale K-SVD modeling to learn each color channel, separately. To avoid overfitting, assuming that the coarser scale of the image has relatively less noise, leads to better results at the reconstruction. Several experimental results have been presented on MR-CT, MR-PET and MR-SPECT dataset to validate the proposed method and showed a detailed comparative analysis with the other available state of the art methods. Based on all the comparative experimental results, it is assured that the proposed SDL-MIF method is able to preserve the significant information of multimodal input images by producing the better visual quality of fused images with improved contrast.

CHAPTER 6: MULTIMODAL MEDICAL IMAGE FUSION USING HYBRID LAYER DECOMPOSITION WITH CNN BASED FEATURE MAPPING AND STRUCTURAL CLUSTERING

This chapter presents a feature level multimodal medical image fusion with the use of two-scale $\ell_1 - \ell_0$ *hybrid layer decomposition scheme to maximize the structural details with significant noise and artifact suppression. The proposed fusion approach utilizes a convolutional neural network (CNN) with consistency verification and structural patch clustering (fuzzy c-means based) for the decomposed base and detail layer fusion, respectively. At first, a color space transform is used to separate the luminance and chrominance components from both of the source images. In the second step, a pre-trained CNN model is used to extract the prominent features of each decomposed base layer components. Then for each output feature map, a regional energy based activity measure is computed. This activity measure is further refined in the consistency verification step to optimize the activity weight map for merging the decomposed base layers. The two-scale detail layers are merged by utilizing clustering based pre-learned multichannel dictionary with saliency matching rule to efficiently map the structural details of the layers. The color components associated with both of the images are also combined using the pixel saliency measure. Finally, the fused base layer, detail layers, and color components are merged to get the resultant fused image. The methodologies used to present the fusion method are also discussed in this chapter. The superiority of the proposed approach is confirmed by a large extent of analytical experimentations on the different real MR-SPECT, MR-PET, and CT-MR neurological image datasets. A detailed comparative analysis has also been presented in the results section of this chapter.*

6.1 Introduction

The motivation for the presented work is to preserve the maximum diagnostic details from the input multimodal medical images with perceptible visual quality. Another possible reason is to integrate morphologically different anatomical and functional details without introducing any spatial and phase distortion. Over the last few years, many research works have been reported on the multimodal medical image fusion [62]. As discussed earlier, fusion techniques are broadly categorized in pixel, feature, and decision level. Usually, these levels of abstraction for image fusion are designed in spatial and transform domain as per the application and fusion quality needed. Spatial domain algorithms are executed directly on the input image using local pixel features, however; this suffers from poor contrast and spatial localization. So, researchers moved towards the transform domain using WT, SWT, CVT, NSCT, and NSST etc. [6, 22, 24, 47, 48, 54, 85, 97, 116, 147, 151, 158, 197, 209] for different types of imaging.

In recent year neural network and fuzzy logic based pixel selection have been popularly adopted in the state of the art methods. Yang et al. have proposed a fuzzy adaptive pulse-coupled neural network (PCNN) in NSCT domain [203]. The fusion quality of such work is quite better over other classic fusion approaches, producing appropriate feature modelling and pixel selection due to the improved uncertainty management of fuzzy logic. Liu et al. [109] utilized NSCT with sparse representation (SR) for fusing the decomposed component to improve the visual quality. The pre-learned overcomplete dictionary approach outperformed the other fusion works in terms of image clarity and information preservation, but lacks in spatial consistency and produces some blurring effect. In addition, NSCT have limited directional feature localization. The NSST has an unrestricted multiscale and multidirectional nonlinear approximation. Though the NSST based fusion results look promising but lack in pixel contrast and preservation of tiny edges as shown in Chapter 2. An edge-preserving decomposition technique named as hybrid $\ell_1 - \ell_0$ prior based layer decomposition [103] considers two separate regularization parameters for base and detail layer in order to retain the week structures and sharp edges. This algorithm efficiently avoids the halo artifacts and enhances the visual quality of image. The guided filtering based image fusion has also become popular [94] as it retains the decent spatial consistency but suffers from spatial smoothing or staircase effects at image boundary. Liu et al. [107] reported another work based on guided filtering in complex shearlet transform (CST) domain. This method shows an improvement by restoring the edges and suppressing the false variance in the image but due to filtering the fused image, it becomes smoother near boundary regions and some fine details get overpowered.

Currently, deep learning has been efficiently introduced in several applications of image processing such as image segmentation, classification, feature mapping, and decision making. A multimodal medical image fusion scheme based on CNN [57] has taken an important role in the processing steps for better performance of pixel to decision level fusion [111]. The major drawbacks of such methods are execution complexity and an extensive data requirement for model training [53]. A convolutional sparsity based morphological component analysis (CS-MCA) based fusion approach is proposed [114] to fuse the medical images by utilizing both the features of sparse representation and morphological component analysis (MCA) to produce a globally sparse overcomplete dictionary and multi-component representation (i.e. cartoon and detail) of the source images, respectively. This shows better results, but still, there is a scope to improve the performance for functional imaging modality.

In this chapter, a novel fusion model based on edge preserving layer decomposition technique using CNN and PCA clustering is proposed for multimodal medical neurological images. The proposed framework efficiently captures the spatial information, maintains the spatial consistency and suppresses the noise and artifacts (mis-registration, halo, and blocking). The major contributions of the proposed fusion method are featured below:

- A novel two-scale composite $\ell_1 - \ell_0$ layer decomposition model is employed to preserve the desired edges and intensity variations at each scale.

- To generate the pixel activity and weight map, a CNN based method followed by consistency verification (binary thresholding, iterative morphological filtering and weighted guided filtering) is employed due to the low accuracy of medical images.
- Fuzzy c-means clustering based local content mapping method is applied to capture the complex detail layer structures followed by PCA based subspace learning for the compact projection of image contents.
- The effectiveness of the proposed framework is demonstrated using extensive experimental results and comparisons made using some published state-of-the-art fusion techniques for gray and color medical images.

The remaining part of this chapter is structured as follows. The next two sections review the brief discussion on some related works adapted to formulate the proposed fusion approach. Next section contains the proposed fusion approach with salient steps for implementation. Furthermore, the different experimentation and their results with comparative analysis are discussed in next section of the chapter.

6.2 Hybrid $\boldsymbol{\ell_1 - \ell_0}$ Decomposition

The composite employment of an $\ell_1 - \ell_0$ sparse regularization parameters the presented layer decomposition is structured as discussed in [103]. The ℓ_1 gradient prior term retains the image edges very well, but due to over-smoothing nature, it shows poor preservation of the structural information. However, the ℓ_0 gradient prior term holds clear piecewise constant property, hence greatly preserves the structural details. To model the two-scale hybrid layer decomposition, firstly apply the $\ell_1 - \ell_0$ priors on the input image to get the first scale decomposed components formulated as follows,

$$\left.\begin{matrix} b_1 = model_{s=1}^{\ell_1,\ell_0}\,(I_X) \\ d_1 = I - b_1 \end{matrix}\right\} \qquad (6.1)$$

where s indicates the scale while I, b_1, and d_1 show the input image, first scale base and detail layer, respectively. The equation for an $\ell_1 - \ell_0$ optimization model $model_{s=1}^{\ell_1,\ell_0}$ is given as follows,

$$min_b \sum_{i=1}^{N} \left\{ \left(I^i - b^i\right)^2 + \lambda_1 \sum_{j=x,y} \left|\partial_j b^i\right| + \lambda_2 \sum_{j=x,y} F\left(\partial_j d^i\right) \right\} \qquad (6.2)$$

where I^i is the i^{th} pixel index of input image (I), b and d refer to base and detail layer obtained by $(I - b)$. $\partial_j \in \mathcal{R}^{2N\times N}$ is the partial derivative operator for base and detail layers along $j = x$ or y directon, and N is the total pixel count. Variables λ_1 and λ_2 are the two regularization parameters with respect to the two gradient sparsity terms ℓ_1 and ℓ_0, respectively and their values are selected based on the spatial properties of the respective layers. The first term in this equation specifies the minimization of squared error between the original image and decomposed base layer in order to correctly replicate the main information, second term indicates ℓ_1 gradient sparsity prior (edge-preserving) imposed on base layer to suppress the outliers and third term indicates ℓ_0 gradient sparsity prior (structure-preserving) imposed on detail layer to suppress the artificial and weak textures with a function $F(\cdot)$ which indicates,

$$F(x) = \begin{cases} 1; & x \neq 0 \\ 0; & else \end{cases} \tag{6.3}$$

The regularization parameter λ_1 and λ_2 control the degree of smoothness for the base and the detail layer, respectively. These regularization parameters are largely dependent on the pixel distribution present in the source image at first scale decomposition. Higher values of these parameters represent the texture rich regions while the lower values represent a region having weaker structural content. Thus, the parameters should be low for the detail layer as it carries more texture information while for the base layer it must be higher in order to suppress the noises and artifacts and provide piecewise smoothness. The variation in statistical behaviour lies more at the first stage of decomposition and reduces as decomposition level increases. For the second scale decomposition, the model in Eq. (6.2) is now applied in layer b_1 to preserve the most of the textural details transferred onto it during first scale decomposition and the remaining part will contain only local brightness. The $model^{\ell_1}_{s=2}$ can be expressed as,

$$min_b \sum_{i=1}^{N} \left\{ \left({b_1}^i - b^i\right)^2 + \lambda_3 \sum_{j=x,y} \left|\partial_j b^i\right| \right\} \tag{6.4}$$

$$\left.\begin{matrix} b_2 = model^{\ell_1}_{s=2}(b_1) \\ d_2 = b_1 - b_2 \end{matrix}\right\} \tag{6.5}$$

where λ_3 is the controlling parameter for final base layer b_2. λ_3 works as weight for ℓ_1 gradient sparsity prior to preserve stronger gradients of the base layer while providing piecewise smoothing effect. To guide the final base layer components weighted guided filtering is performed on b_2 with b_1 to restore the spatial consistency sharp boundaries. The reconstruction is done by summing all three layers d_1, d_2 and b_2, represented as,

$$I_{Rec} = d_1 + d_2 + b_2 \tag{6.6}$$

The hybrid composition of these two priors (ℓ_1 and ℓ_0 based sparsity parameters) brought into the base and detail layer, respectively shows the effective and efficient decomposition model as shown in Figure 6.1. The structural information is enhanced and intact in the two-scale detail layer by imposing ℓ_0 term, while the appropriate luminance is preserved in the base layer with suppressed halo artifacts by applying ℓ_1 sparsity prior based edge-preserving filter. With the use of proper regularization parameters, the fused outcome outperforms clearly to another state-of-the-art approaches. In order to achieve edge-aware filtering for a base layer, a weighted guided filter (WGF) [101] is used which can preserve the local as well as global structural details well alongwith the merits of guided filtering. The WGF is also enhancing the image features by removing the halo artifacts which is the common problem with most of the traditional fusion methods.

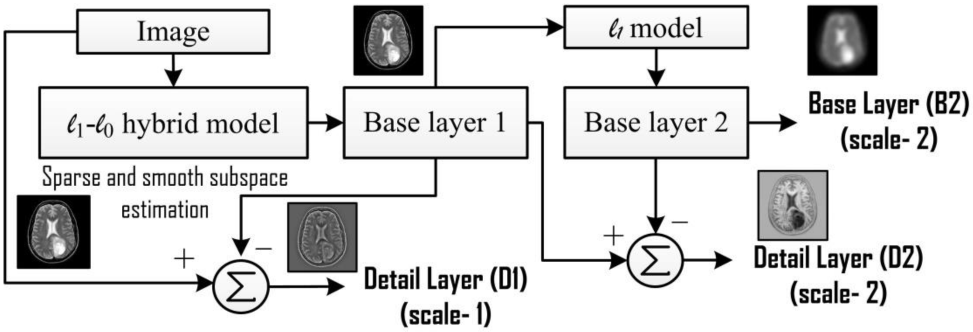


Figure 6.1 A two-scale hybrid layer decomposition

6.3 Convolutional Neural Network

Nowadays, a CNN based method has become the majorly accepted model for the multimodal medical input by which the important visual features and corresponding score map can be estimated. A CNN is a trainable feed-forward neural network having multiple stages corresponds to a certain number of feature maps. These feature maps are connected to its next stage using three spatial operations such as convolution, activation ($ReLU$) and sub-sampling (max-pooling). Each stage automatically extracts some features from its previous stage which carries most prominent details in accord to the sensitivity of the mammal visual cortex. The constructed features maps are spatially invariant and more generalized (e.g. edges, blobs and textures) at shallow stages while gets more domain-specific at deeper stages. A Siamese based CNN architecture [111] used in the proposed work for feature mapping and base layer fusion has two channels which share similar weights and bias. This is also computationally easy to be trained compared to the other networks. The pre-trained CNN model has three convolutional layers followed by non-linear activation function i.e. $ReLU$ for each layer and one max-pooling layer. In the proposed fusion framework, a CNN model taken from [111] is trained with 2 million patches of Gaussian filtered and real images taken from ImageNet dataset [112] using a deep learning platform 'Caffe' The main advantage to consider the Gaussian blurred patches is that the trained classifier has an ability to be a generalized model instead of the application specific.

The training images are partitioned into several patches of 16×16 size which convolve with the layer kernel of size 3×3, single stride and having 64 filter channels at the first stage. Each convolutional layer extracts the image features and provides the map with depth size of $64 \times 2^{i-1}$, where $i = (1, 2, 3)$ is a layer index. Features extracted by the convolutional layers are contracted using max-pooling layer to get the most important uncommon features (based on maximum votes of patch members) available in the predefined patch window. However, max-pooling shrinks the spatial resolution too, hence in the next phase; the subsampled feature maps are integrated with the third convolution kernel to acquire shift invariance. The max-pooling layer kernel size is set to 2×2 with a stride of 2. After the third stage, concatenation of feature maps is done and a 2-D feature vector is generated. This feature vector feds to the two-channel soft-max layer to compute the probability distribution in order to generate a final score map. For an input source image of size $N \times M$, the size of the output feature map is $(N/2 - 2^i + 1) \times (M/2 - 2^i + 1)$, where $i = 3$ and the learning rate is initially set to 10^{-4}. In soft max layer, the logistic loss function of the output layer is minimized using a gradient descent approach to obtain the optimized weights and bias for each layer. The network contains two subsequent convolutional layers followed by max-pooling layer and the third convolutional layer as shown in Figure 6.2.

At each convolutional layer, the feature mapping is performed by computing the nonlinear relationship between the layer input and the shared weights activated with some bias. Let r_k and $\hat{r}_l$ are representing the k^{th} input map and l^{th} output map of convolutional layer i, respectively. The nonlinear formulation of the input and convolutional operator (weight matrix) can be expressed as,

$$\hat{r}_l = ReLU_f\left(\frac{bias_l + \sum_k convw_{k,l} * r_k}{\varphi}\right) \tag{6.7}$$

where the $ReLU$ nonlinear activation function [130] can be expressed as $ReLU_f(\varphi) = max(\varphi, 0)$. The operator $(*)$ indicates the convolution operation, $convw_{k,l}$ indicates the convolutional kernel imposed on k^{th} input map to gain l^{th} output map, and $bias_l$ is the bias for each i^{th} layer. The convolutional operation performed within the local mask of size $n \times n$ which is kept same for all three layers while the pooling layer reduces the resolution of the layer output by two.

In the proposed framework, CNN based base layer fusion is obtained in three key steps; 1) initial feature map generation using pre-trained CNN model, 2) activity measure and fusion using weighted average based rule, and 3) consistency verification of the fused image for adequate contrast and brightness in the resultant image. The detailed steps to fuse the base layer are explained in the next section. The key problem of using a CNN model with MST-based approach is that the CNN models are usually trained in spatial domain; whereas the fusion approaches performed using CNN for transformed coefficients may fail to maintain a fair amount of spatial consistency (artifacts present after decomposition and reconstruction phase of the MST) needed for the perceptually better image [57].

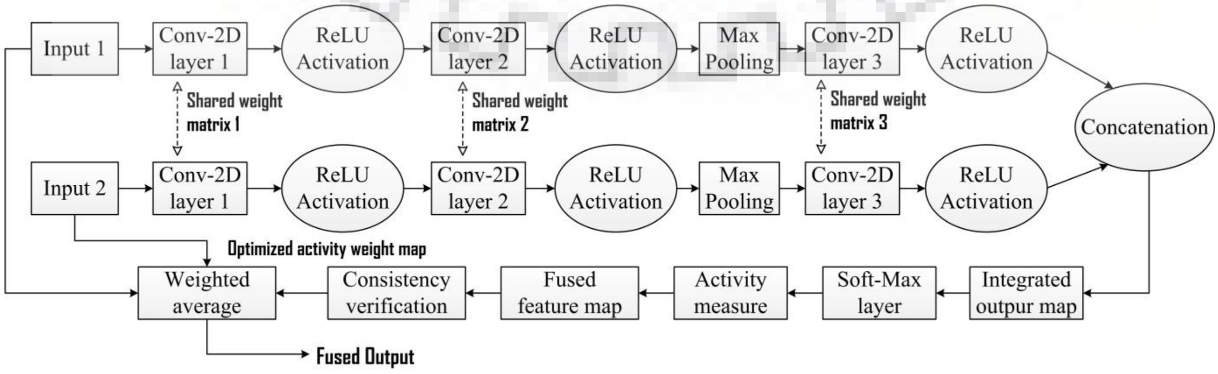


Figure 6.2 Flowchart of CNN based fusion framework

6.4 Clustering PCA based Statistical Feature Mapping

This section focuses on the preservation and enhancement of structural features of an image based on patchwise statistical PCA clustering. This approach is composed of three basic steps as 1) patch partitioning based on geometrical details of the layers (i.e. flat area and higher variance area), 2) clustering the maximum variance regions such as edges and textures in different classes based on minimum Euclidean distance and 3) learning the PCA basis for each class and combine each sub-basis to form a composite dictionary. For a given input data, the structural patch decomposition approach is applied based on statistical measures of an image such as fine details, local content and the color variations. The decomposed patch groups are clustered using fuzzy c-means (FCM) based unsupervised learning to generate the different cluster groups for each measure. The FCM based clustering approach is selected because the fuzzy function can handle the visual uncertainty better than other nonlinear system and hence successfully models the human visual system [35]. Moreover, it can correctly discriminate the significant and insignificant pixels for contrast and fine detail enhancement to boost overall image clarity. A PCA basis is generated for each cluster to capture the most prominent feature of the image patches. The fusion framework based on statistical clustering is shown in Figure 6.3.

For a given input image, firstly overlapped patch-based image partitioning takes place with a fixed window and step size. After the image partitioning, each patch is decomposed into two major components, namely local mean and detailed structure of an image. The local mean μ is calculated by averaging the pixels presented in patch p and representing the flat areas of p. The detail structure denoted by p_d is calculated by subtracting the patch mean from the extracted image patch p as follows,

$$p_d = p_{N\times M} - \mu_{N\times M} \tag{6.8}$$

The detail structure component well preserves the local structural information for better visibility and interpretation of the main object while suppressing the noise and artifacts. With the use of the above defined components, each patch p is reconstructed using following mathematical expression given as,

$$p = \overline{p_d} + \bar{\mu} \tag{6.9}$$

If the input image has three channels, then both of the decomposed components are to be calculated for each color channel R, G, and B separately and then each patch is reconstructed using,

$$p = \begin{bmatrix} \bar{p}_d^r \\ \bar{p}_d^g \\ \bar{p}_d^b \end{bmatrix} + \frac{1}{3}\left(\bar{\mu}^r + \bar{\mu}^g + \bar{\mu}^b\right) \tag{6.10}$$

where the first component represents the color structure extracted from the three channels while the second component represents the average of local mean of the three channel color patches.

In the next phase, patch clustering is evolved to capture the pixels in one cluster group based on structural similarity. For clustering the pixels, FCM based learning approach is adopted. Each cluster is then defined through subspace projection using PCA to get the most prominent detail structures of the input data. Finally, a PCA basis is generated representing the most common and significant features in the image detail structure which is further applied to transform the testing image patches into PCA subspace [87].

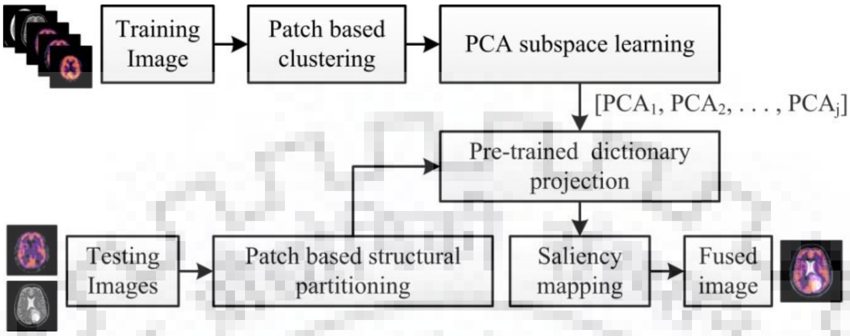


Figure 6.3 Flowchart of the PCA clustering based fusion

6.5 Proposed Fusion Approach based on CNN Feature Mapping and Structural Clustering

The proposed fusion method is implemented in six main stages to accomplish the desired fusion result. These stages involve color space transform (RGB to uncorrelated luminance and color components), layer decomposition, CNN based base layer fusion, PCA clustering based detail layer fusion, layer reconstruction and color component fusion followed by the RGB space re-conversion. In this section, the complete implementation steps involved in the proposed method are summarized and the process flow of the proposed approach is shown in Figure 6.4.

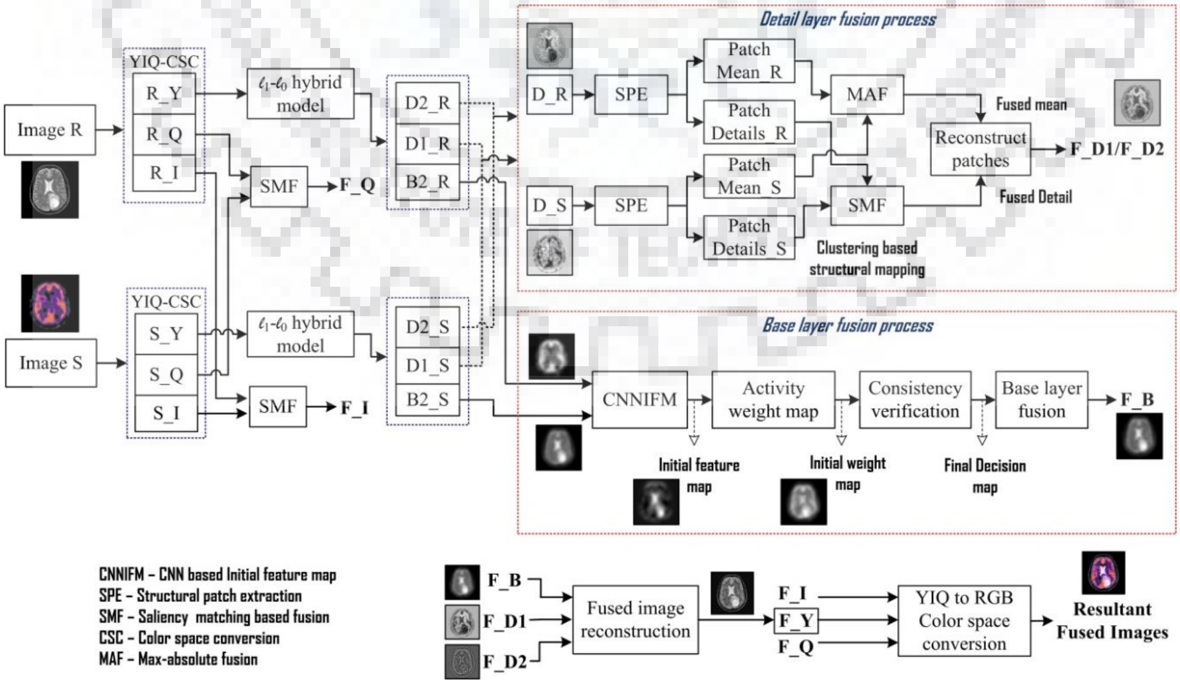


Figure 6.4 Process flow of the proposed method

For the given input source image pairs $\boldsymbol{R} = R(x,y)$ and $\boldsymbol{S} = S(x,y)$, the proposed multimodal medical image fusion model based on the CNN (CNN-MIF) is implemented as described below.

Implementation Step:

Step 1: Covert the RGB model of the input image into YIQ color space as per given mathematical formulation [140].

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5959 & -0.2746 & -0.3213 \\ 0.2115 & -0.5227 & 0.3112 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{6.11}$$

Step 2: Decompose the luminance component (Y_X) of reference image $X \in (R,S)$ using hybrid two-scale layer decomposition method into one base layer B_2^X and two detail layers D_1^X and D_2^X getting from both the scales.

Step 3: Base layer fusion- A convolutional neural network based fusion approach is used to merge the base layer components from each decomposed data. To implement such an approach, a pre-trained CNN model is used and the output layer map for each input is then optimized using feature-based fine tuning followed by consistency verification. The detailed process of the fusion is structured as given in steps below:

1. Consider the base layer content $B_2^{\boldsymbol{R}}$ and $B_2^{\boldsymbol{S}}$ that uses a three layer CNN model to extract the activity measure for content mapping. For the i^{th} convolutional layer and c channels of the layer, the output map can be expressed as,

   $$O_X^{i,c} = ReLU_{f^i}\left(\varphi(B_2^X)\right),\ X \in (R,S) \tag{6.12}$$

   where at each i^{th} layer $\varphi(\cdot)$ denotes the feature map acquired from Eq. (6.7). The first layer uses B_2^X as input image patch while the later stages use the $(i-1)^{th}$ layer output as the input patch for the i^{th} layer feature extraction denoted by $O_X^{i,c}$.
2. Calculate the initial weight map $W_X(y,z)$ at 2-D soft-max layer using the concatenated output feature maps O_X of the third convolution layer and expressed as,

   $$W_X(y,z) = \frac{O_X(y,z)}{O_{\boldsymbol{R}}(y,z)+O_{\boldsymbol{S}}(y,z)} \tag{6.13}$$
3. Calculate the fused weight map using block-based regional energy (RE) as an activity measure for individual output maps of the two channels and taking a weighted average with upsampling to get a final activity map same in size with the reference input and expressed by mathematical formulation given below.

   $$map_X(y,z) = RE_b\big(W_X(y,z)\big) \tag{6.14}$$

   where RE refers to a weighted regional energy of the image $W(y,z)$, where the region is referred to a size of the weighted window $w_b(y,x)$ for the neighbourhood processing as given below.

   $$RE_b(y,z) = \sum_{(j,i)\in w} w_b(j,i)W^2(j,i) \tag{6.15}$$

where width $b = 2l + 1$, $y - l \leq j \leq y + l$, $z - l \leq i \leq z + l$ and the normalized weight value is defined in 5×5 neighboring pixels for efficient feature masking.

$$w_b = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix} \tag{6.16}$$

$$\overline{map}_X(y,z) = \frac{1}{9}\left[\sum_{t=-1}^{1}\sum_{u=-1}^{1} map_X(y+t, z+u)\right] \tag{6.17}$$

$$W_{A_X}(y,z) = \frac{\overline{map}_X(y,z)}{\overline{map}_R(y,z) + \overline{map}_S(y,z)} \tag{6.18}$$

$$\widetilde{W}_X(y+a, z+a) = W_{A_X}(y,z) \tag{6.19}$$

where b and a refers to block size and upsample unit, respectively. The block size of a weighted window can be 3×3, 5×5, 7×7 or higher based on the required masking. Though smaller window preserves more tiny details, but not able to filter the noise well. The larger size of a window makes the image smoother and robust to misregistration, while attenuates tiny edges and produces a visually less perceptible image. Block size $b = 5$ is considered here to preserve the useful high frequency features (i.e. edges) mapped from the CNN with required artifact suppression. The upsampling step size is set as $a = 8$ in order to reconstruct the initial feature mapped image resized to the source image using the relation $a = (0, 1, \ldots, 2^{i-1} - 1)$, where i refers to the layers in a convolutional network. Though, lower or higher upsample units can be applied to resize the image but smaller size does not consider the remarkable features to restore in the fused outcome while the higher size could make the prominent features less clear and reduced the contrast. Finally, combine the resultant activity maps to preserve most of the prominent contents (brightness and structural information) relative to the two base layer coefficients as follows.

$$W_F(y,z) = \widetilde{W}_{\boldsymbol{R}}(y,z) \times B_2^{\boldsymbol{R}} + \widetilde{W}_{\boldsymbol{S}}(y,z) \times B_2^{\boldsymbol{S}} \tag{6.20}$$

4. <u>*Post-processing of decision map:*</u> To get the optimized decision map for more better fusion results, post-processing is done in three steps, namely (1) initial segmentation map using binary thresholding with threshold $median\left(W_F(y,z)\right)$ to acquire a binary segmented map (2) morphological filtering to reduce mis-classified pixels (smoothen tiny areas) using ‘bwareopen’ and close operation and (3) weighted guided filtering based weight map optimization given below. The final decision map improves the spatial consistency and suppresses noise and artifacts. Let for the target segmented image $t(y,z)$ getting after step 3, the fused guidance image $g(y,z)$, optimized final decision map $D(y,z)$, and fused base layer image B_2^F are obtained according to the expressions given below.

$$g(y,z) = B_2^{\boldsymbol{R}} \times t(y,z) + B_2^{\boldsymbol{S}} \times (1 - t(y,z)) \tag{6.21}$$

$$D(y,z) = G_{r,\epsilon}\big(g(y,z), t(y,z)\big) \tag{6.22}$$

$$B_2^F = D(y,z) \times B_2^{\boldsymbol{R}} + \big(1 - D(y,z)\big) \times B_2^{\boldsymbol{S}} \tag{6.23}$$

where $G_{r,\epsilon}$ is the weighted guidance filter (WGF) [101] with r and ϵ as filter radius and degree of smoothing, respectively.

Step 4: Detail layer fusion- For the detail layer fusion, the concept of statistical clustering based PCA transform [87] is employed, which improves the contrast level significantly and enhances the overall image clarity. The following procedures are considered to implement the fusion process given below.

1. *PCA basis learning:* Generate a compact PCA basis using an offline learning process to boost the computational efficiency. The training stage is done with 130 medical images of size 256×256, (CT, MR, PET, and SPECT). A sliding-window based patch extraction (with patch size 8 and stride 1) is applied to the training dataset and two structural components (patch mean and patch details) are computed for each patch. An FCM based structural clustering is used to index the patch in N groups having almost similar features but different intensities. For each cluster, cluster centroid and PCA based approximation model are computed by choosing the uppermost principal components. The resultant principal subspaces are combined to generate a PCA dictionary (P), consisting of major elements of the clustered patch.
2. Partition the detail layer components D_i^X into overlapped patches (size 8 and step 1) p_i^X, where $X \in (R,S)$ and $i \in (1,2)$ is the decomposition scale.
3. From each patch, compute the statistical information as patch mean and patch structure similarly expressed as above and denoted by μ_i^X and p_i^X, respectively.
4. Compute the class of each patch based on the minimum Euclidean distance between the centroids of each cluster and patch $p_i^{dx}(y,z)$. Apply the learned PCA dictionary on patch component p_i^{dx} to extract the most important local structural information s_i^{dx} as given below.

   $$s_i^{dx} = P \cdot p_i^{dx} \tag{6.24}$$

   The larger coefficients value of s_i^{dx} defines the principal details, while the smaller value defines the fine texture details.
5. Reconstruct the approximated patch detail model $\tilde{p}_i^{dx}$ using,

   $$\tilde{p}_i^{dx} = P^T \cdot s_i^{dx} \tag{6.25}$$

6. Fuse the patch mean of detail layer from both the inputs using ‘max(abs)’ as,

   $$\mu_i^F = \begin{cases} \mu_i^{\boldsymbol{R}}; & |\mu_i^{\boldsymbol{R}}| > |\mu_i^{\boldsymbol{S}}| \\ \mu_i^{\boldsymbol{S}}; & otherwise \end{cases} \tag{6.26}$$

7. Fuse the patch structural information using the saliency measure based on local contrast variation to preserve the edge details well followed by matching the saliency measures. For two inputs $\tilde{p}_i^{d_{\boldsymbol{R}}}$ and $\tilde{p}_i^{d_S}$, the required computation are expressed as,

$$\tilde{p}_i^{d_F} = d_i(y,z) \cdot \tilde{p}_i^{d_{\boldsymbol{R}}} + \left(1 - d_i(y,z)\right) \cdot \tilde{p}_i^{d_S} \tag{6.27}$$

$$d_i(y,z) = \begin{cases} \tau; & M_i(y,z) \le \tau, LC_i^{\boldsymbol{R}}(y,z) > LC_i^{\boldsymbol{S}}(y,z) \\ 1-\tau; & M_i(y,z) \le \tau, LC_i^{\boldsymbol{R}}(y,z) \le LC_i^{\boldsymbol{S}}(y,z) \\ 0.5 + 0.5\left(\frac{1-M_i(y,z)}{1-\tau}\right); & M_i(y,z) > \tau, LC_i^{\boldsymbol{R}}(y,z) > LC_i^{\boldsymbol{S}}(y,z) \\ 0.5 - 0.5\left(\frac{1-M_i(y,z)}{1-\tau}\right); & M_i(y,z) > \tau, LC_i^{\boldsymbol{R}}(y,z) \le LC_i^{\boldsymbol{S}}(y,z) \end{cases} \tag{6.28}$$

$$LC_i^X(y,z) = \sum_{w\in 3} \left|\tilde{p}_i^{d_X}(y+w, z+w)\right|^2 \tag{6.29}$$

$$M_i(y,z) = \frac{2\sum \tilde{p}_i^{d_{\boldsymbol{R}}}(y+w,z+w) * \tilde{p}_i^{d_{\boldsymbol{S}}}(y+w,z+w)}{\sum_{X\in(\boldsymbol{R},\boldsymbol{S})} LC_i^X(y,z)} \tag{6.30}$$

where $d_i(y,z)$ and $LC_i^X(y,z)$ represent a decision weight map for fusion and local contrast based patch saliency measure, respectively. $M_i(y,z)$ represents a match measure within $(-1,1)$. $\tilde{p}_i^{d_F}$ denotes the fused structural patch detail layer $i \in (1,2)$. τ is a visual threshold value to control the saliency-based match measure that is mapped between two different source images. The aforementioned formulations present an interdependency between $LC(y\,(z)$, and $M(y,z)$ followed by a decision weight map $d(y,z) \in (0,1)$. If the local contrast based pattern similarity is higher than τ, then the decision weight map will vary nearly about 0.5 based on the dominant patterns between two images. However, if the similarity is lower than τ at a given spatial position or the salient patterns are less correlated, the image having more significant saliency will be weighted more typically by $d_{max}(y,z) = \tau$ in the decision map.

The effect of τ is analyzed on the fused detail layer based on several successive experiments. A higher value indicates that some irrelevant features dominate in the high frequency information while lower value rejects useful features and provides the results as a loss of pixel saliency. Additionally, no such remarkable differences are found by changing its value slightly except either keep it below 0.5 or above 0.9. Thus, in the present work, the parameter $\tau = 0.7$ is considered to generate the best possible decision weight map according to Burt's method [21].

8. Combine the fused patches computed in step 6 and 7 to reconstruct the fused detail layers as given below.

$$D_i^F = \tilde{p}_i^{d_F} + \mu_i^F \tag{6.31}$$

Step 5: Reconstruct the resultant fused luminance image (F^Y) using three fused layers and can be expressed as,

$$F^Y = B_2^F + D_2^F + D_1^F \tag{6.32}$$

Step 6: Reconvert the fused image to RGB color space using YIQ to RGB transform to get the final fused result.

$$\begin{bmatrix} \mathrm{R} \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0.956 & 0.621 \\ 1 & -0.272 & -0.647 \\ 1 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} F^Y \\ I \\ Q \end{bmatrix} \tag{6.33}$$

6.6 Experimental Details

This section represents the experimental details and performance analysis of the proposed CNN-MIF approach. To evaluate the performance of the proposed CNN-MIF method, seventy three multimodal medical image pairs are used (14 CT-MR, 29 MR-SPECT, and 30 MR-PET images). To implement the proposed CNN-MIF approach, hybrid norm based two-scale layer decomposition is employed with CNN based base layer fusion and statistical clustering based detail layer fusion. The two-scale layer decomposition technique requires the appropriate setting of the three parameters namely λ_1, λ_2, and λ_3. The degree of smoothness of the first layer is controlled by λ_1 and λ_2 for the base and detail layer obtained at scale 1, respectively. However, the second scale has only λ_3 parameter that controls the smoothing level of the base layer obtained at scale 2.

The range of the regularization smoothing parameters of gradient priors ℓ_1 and ℓ_0 for two scale decomposition, i.e. λ_1, λ_2 for scale 1 and λ_3 for scale 2 lies in [0 1]. Several successive experiments are performed to select the appropriate values of these parameters and analyze its effects on multimodal medical images. In the present work, standard deviation of the Gaussian filtered image (GSTD) is computed to measure the smoothness of corresponding layers at each scale and to estimate the pixel distribution. Higher GSTD values indicate more structural information available in the images, hence the lower value of the smoothing parameter is considered to retain more structural contents in detail layers. Higher value of the smoothing parameter is selected to remove noise and artifacts in Gaussian filtered images having lower GSTD values. The gradient prior ℓ_1 facilitates an edge-preserving smoothing and outlier-rejection favorable for base layer components, whereas the prior term ℓ_0 provides structure-preserving smoothing to enhance the important textural contents with fine structural details required for the detail layer contents. Hence, The regularization parameters λ_1 and λ_2 are set as 0.4, $0.01\lambda_1$, respectively. λ_3 is set to 0.05 as it controls the second scale base layer and due to low accuracy of medical images, it kept to a smaller value than the layer one. In the base layer fusion, CNN based initial decision map is optimized using the WGF controlled by radius $r = 1$ and blurring degree $\epsilon = 0.01$.

A pre-trained PCA dictionary and CNN model are used to reduce the computational cost and the proposed CNN-MIF method is tested offline testing basis. For the objective an subjective performance analysis of fusion results, the proposed method is compared with several existing fusion methods such as wavelet transform (WT) based image fusion proposed by Nirmla et al., 2016 [137], PCNN based fusion model on NSCT domain proposed by Das and Kundu, 2012 [34], guided filtering (GFF) based fusion approach developed by Li et al., 2013 [94], NSST domain PCNN based fusion model proposed by Ganasala and Kumar, 2016 [48], fuzzy adaptive PCNN based fusion in NSCT domain by Yang et al., 2017 [203], gradient domain guided filtering (GDGF) based fusion proposed by Zhu et al., 2018 [218], CNN based medical image fusion by Liu et al., 2017 [111], CNN based fusion by Hermessi et al., 2018 [57] and NSST domain CNN model by Hou et al., 2019 [59].and the C-MIF approach as discussed in Chapter 2. Moreover, the other approaches developed by following authors such as, Bhatnagar et al., 2013 [18], Yang et al., 2015 [200], and Yang et al. (NSCT-Type2 Fuzzy), 2016 [201] are also considered here to analyze the peformance of color medical image fusion.

To investigate the efficacy of the proposed CNN-MIF approach objectively, different performance measures (En, MI, STD, IQI, FMI, SF and XEI) as discussed earlier are adapted to present the comparative evaluation. The subjective results speculate the comparative assessment of visual perception among the source and fused images while the quantitative investigation demonstrates the usefulness of the perceived results through various image quality evaluation measures. Higher index values show a better fusion performance for all the considered objective measures. The performance measures do not only show the comparative analysis of fusion results, but also verify the acceptability of visual observation by preserving the information and structural contents of the source images.

6.7 Results and Discussions

6.7.1 Comparison of anatomical-anatomical image fusion

This section discusses the different simulation methods and their results for the proposed CNN-MIF method and other existing fusion approaches. The pre-processing step ensures decent geometrical matching and confined resolution for further processing of input data. The experimental results will reveal the fusion of these multimodal neurological images, providing a single fused image having both the soft and hard tissue and structural information taken from the different MR and CT images. To present the performance of the proposed CNN-MIF, four sets of the multimodal medical dataset are taken for performance evaluation.

The set-1 shown in Figure 6.5 includes four different pairs of CT and MR images shown in Figure 6.5 (a) and Figure 6.5 (b), respectively with their fused images obtained using the proposed CNN-MIF method as shown in Figure 6.5 (c). For all four pairs of CT and MR images, significant information content is preserved nicely and enhanced in the resultant fused image and compare to the source images. The objective fusion scores are also demonstrated with the higher values of En, STD, MI, SF, and XEI to present the fusion performance of the proposed CNN-MIF framework.

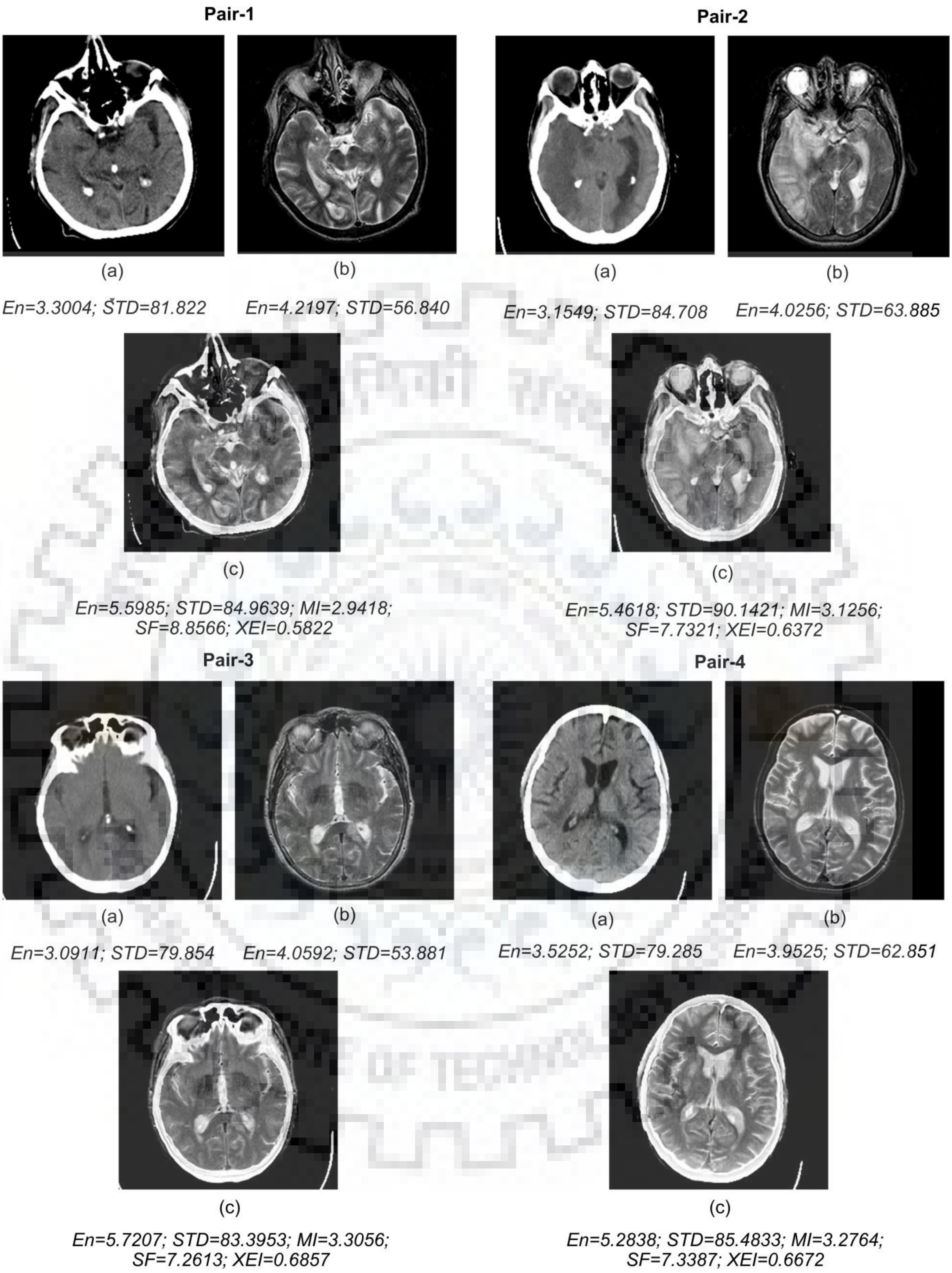


Figure 6.5 Fusion performance evaluated by the proposed approach on (a) Source CT image (b) Source MR image (c) Fused image

Furthermore, to validate the effectiveness of the proposed fusion scheme, several experiments have been done so far and their quantitative and qualitative results are compared with ten exsiting fusion methods. The fusion results of different CT and MR image pairs are shown in Figure 6.6 to Figure 6.9. Moreover, a cropped view is also shown on each image to visualize the subjective performance of each method more clearly. From the visual results, it can be clearly observed that the fused outcome characterizes both the soft and hard tissue structures present in the source MR and CT images, respectively. Figure 6.6 (a) to Figure 6.9 (a) shows WT based image fusion as discussed in [137]. Though the image is fused correctly, but limited by severe blocking effect and poor contrast. In Figure 6.6 (b) to Figure 6.9 (b), the fusion method based on PCNN in NSCT domain, can effectively restore the average information with decent contrast, but fails to capture the fine line textures as shown by red and yellow arrows in the zoomed box. The GFF based method shown by Figure 6.6 (c) to Figure 6.9 (c) stores spatial consistency well compared to the previous two methods, but does not solve the problem of poor contrast and fails to capture the adequate information of the CT images for all image pairs.

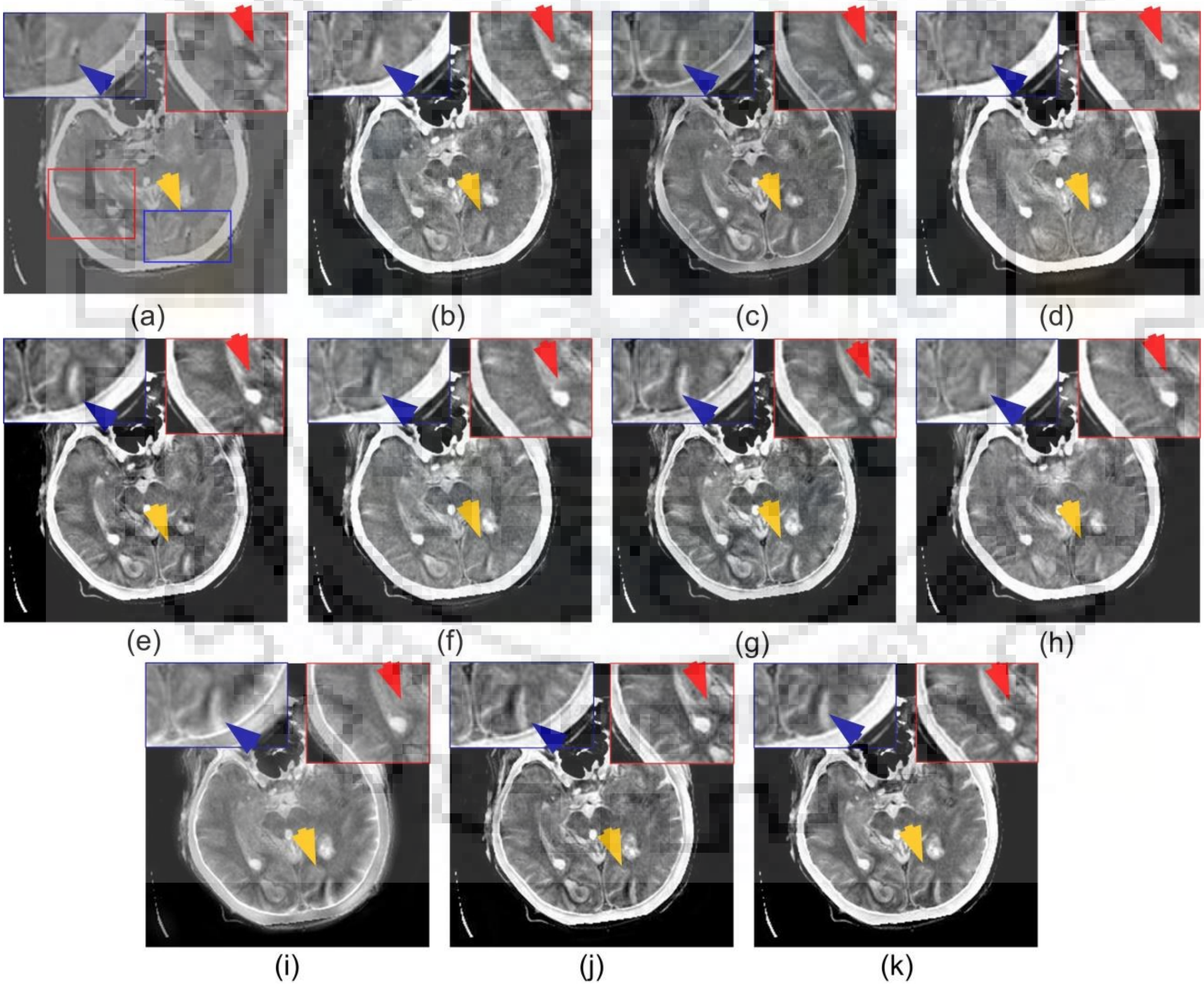


Figure 6.6 CT-MR image pair-1 shown in Figure 6.5. Comparison of fused images obtained by (a) Nirmla et al., 2016 [137] (b) Das and Kundu, 2012 [34] (c) Li et al. (GFF), 2013 [94] (d) Ganasala and Kumar, 2016 [48] (e) Yang et al., 2017 [203] (f) Zhu et al. (HMSD-GDGF), 2018 [218] (g) Liu et al., 2017 [111] (h) C-MIF method (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

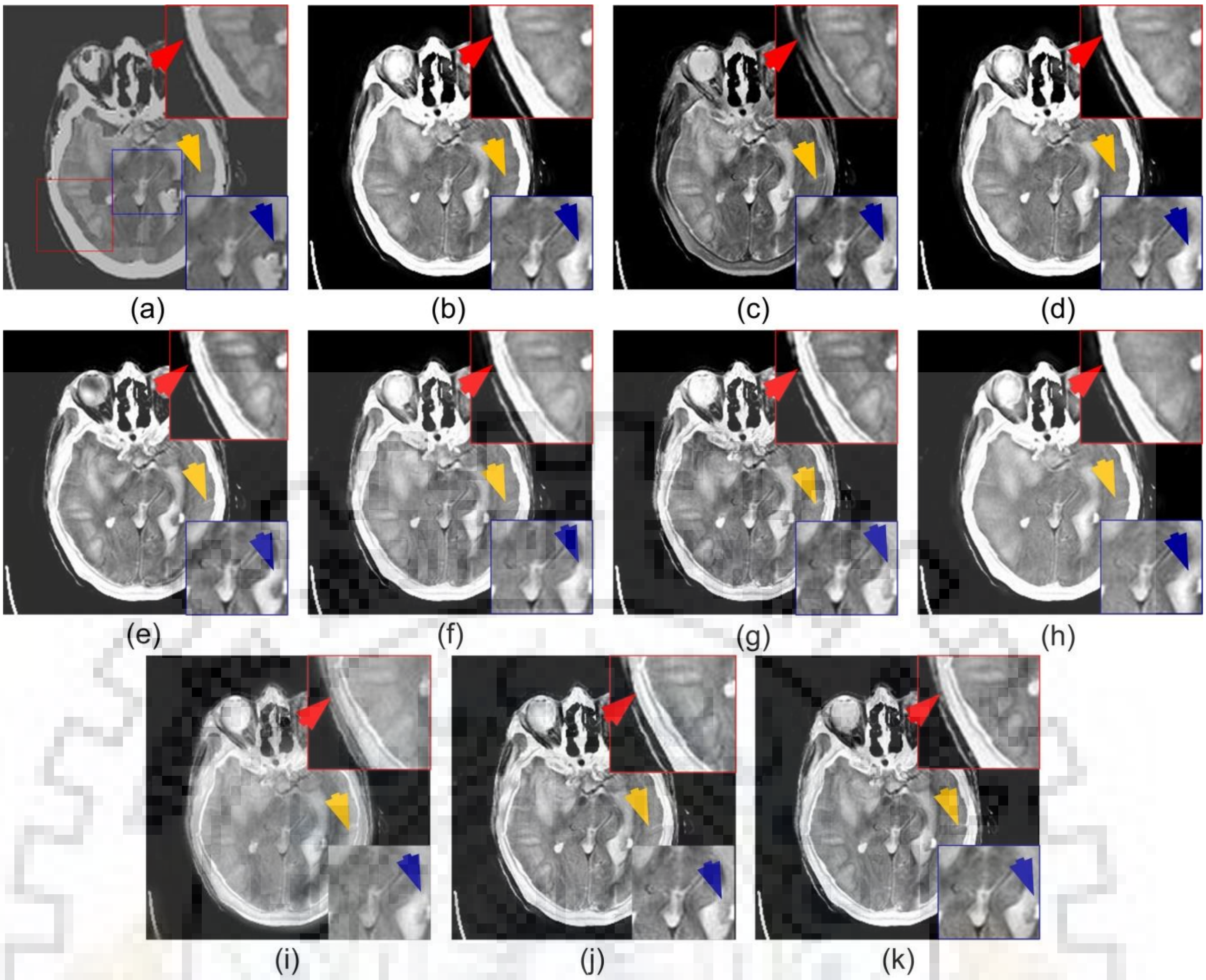

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure 6.7 CT-MR image pair-2 shown in Figure 6.5. Comparison of fused images obtained by (a) Nirmla et al., 2016 [137] (b) Das and Kundu, 2012 [34] (c) Li et al. (GFF), 2013 [94] (d) Ganasala and Kumar, 2016 [48] (e) Yang et al., 2017 [203] (f) Zhu et al. (HMSD-GDGF), 2018 [218] (g) Liu et al., 2017 [111] (h) C-MIF method (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

On the other hand, the NSST-PCNN based method successfully captures the pixel activity that provides a good contrast in the image as clearly seen in Figure 6.6 (d) to Figure 6.9 (d), however, still unable to capture the edges around centre focal and occipital axial regions (small edges are smoothed out). Figure 6.6 (e) to Figure 6.9 (e) depicts a fuzzy adaptive PCNN based fusion. Though the visual results clearly deliver better visual quality than the previous ones in terms of luminance and structural detail preservation, but the fluidic regions become a little hazy (red and yellow arrows). In Figure 6.6 (f) to Figure 6.9 (f), the GDGF based fusion model can preserve structural details well, but suffers from inappropriate pixel consistency, hence results over brightness in the resultant image (shown by the yellow arrow). In Figure 6.6 (g) to Figure 6.9 (g), CNN based fusion outperforms the other compared techniques by providing satisfactory fusion results, but fails to depict a CT image properly and the edge outlines are also not clear in the resultant image. From the visual results of previously proposed work based on cascaded fusion model (C-MIF) as shown in Figure 6.6 (h) to Figure 6.9 (h), it can be clearly seen that the method preserves the significant information from both the source images with natural contrast level, however, still suffers from smoothed edge details in the area has small but frequent variations (as indicated by red and blue arrows). CNN based similarity learning approach [57] retains most of the feature information, but had blurred contours near fine edges with over brightness as shown in Figure 6.6 (i) to Figure 6.9 (i). This method shows a lower value of edge index values compared to the other methods. Figure 6.6 (j) to Figure 6.9 (j) present the fuse images obtained by Hou et al. [59] which is able to capture edge details and retain the overall information very well, however, limits to maintain the pixel contrast. The fusion result of the proposed CNN-MIF method is shown in Figure 6.6 (k) to Figure 6.9 (k) for all four reference image pairs which characterize all principals and tiny details with required brightness and pixel contrast. In addition to this, the proposed method is also able to preserve sharp edges and visual consistency quite well and provides the halo-free image.

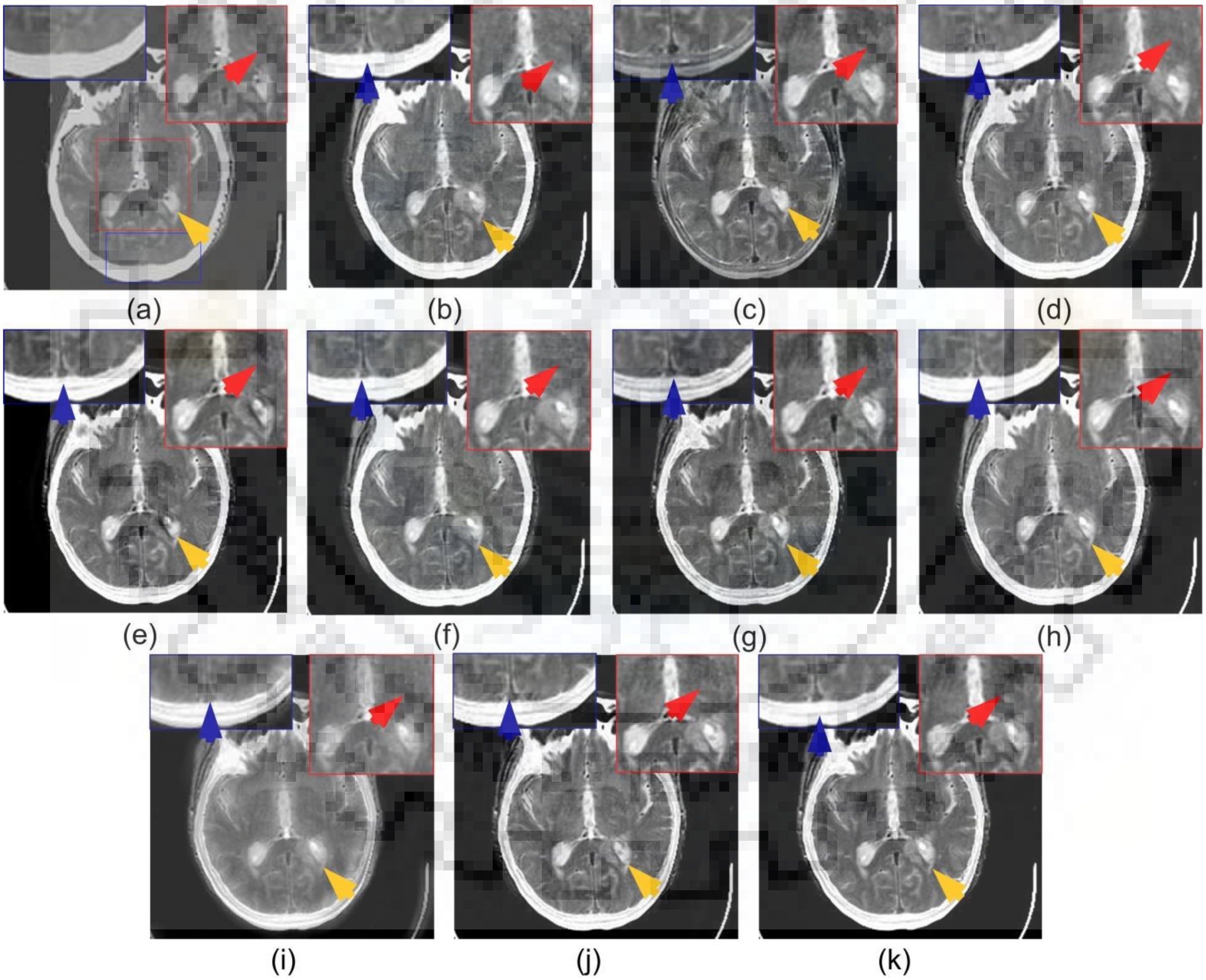


Figure 6.8 CT-MR image pair-3 shown in Figure 6.5. Comparison of fused images obtained by (a) Nirmla et al., 2016 [137] (b) Das and Kundu, 2012 [34] (c) Li et al. (GFF), 2013 [94] (d) Ganasala and Kumar, 2016 [48] (e) Yang et al., 2017 [203] (f) Zhu et al. (HMSD-GDGF), 2018 [218] (g) Liu et al., 2017 [111] (h) C-MIF method (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

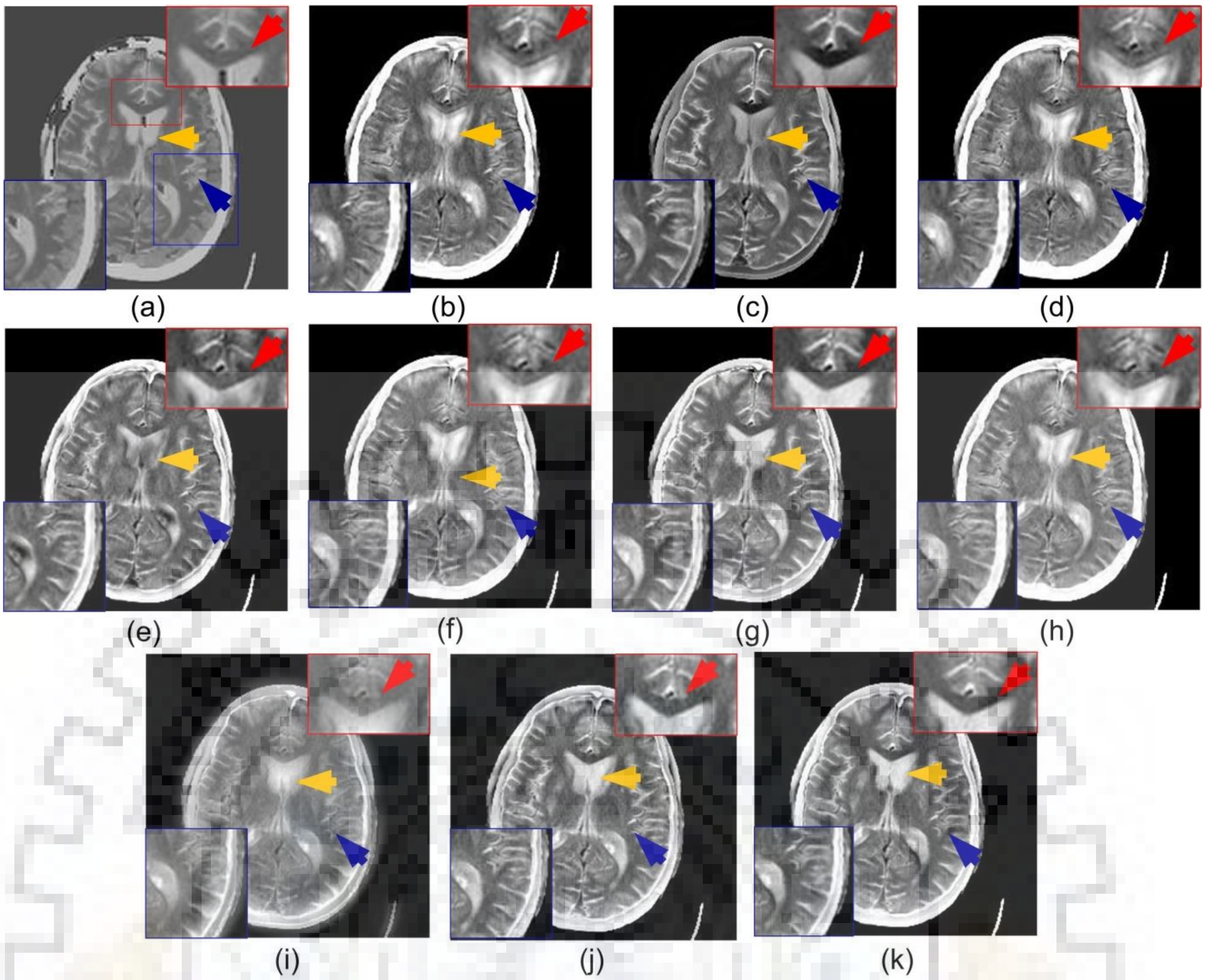


Figure 6.9 CT-MR image pair-4 shown in Figure 6.5. Comparison of fused images obtained by (a) Nirmla et al., 2016 [137] (b) Das and Kundu, 2012 [34] (c) Li et al. (GFF), 2013 [94] (d) Ganasala and Kumar, 2016 [48] (e) Yang et al., 2017 [203] (f) Zhu et al. (HMSD-GDGF), 2018 [218] (g) Liu et al., 2017 [111] (h) C-MIF method (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

Furthermore, the fusion results of all the methods are also analyzed objectively to validate the corresponding subjective performances. Table 6.1 to Table 6.4 show the objective fusion results for all four pairs of the testing dataset and their averaged scores are shown in Table 6.5. From the qualitative results shown in Table 6.1 to Table 6.4, it is clearly seen that the proposed method achieves highest edge index XEI and subsequent information from the source. The cascaded fusion framework shows the second highest performance among all other fusion methods. The CNN based fusion achieves good SF and XEI index for pair 3 and 4 but lacks in sharp edge restoration and achieves a lower value of STD for all the data pairs. The NSCT-PCNN based fusion yields higher information content in the fused images with a high En metric score compared to WT, NSST-PCNN, GFF, CNN, and GDGF based methods but lacks due to low SF values. Table 6.5 reflects that the proposed method outperforms the other compared methods and scoring the higher averaged values of En, MI, STD, SF and XEI performance hence, justifies the visual performance of the proposed CNN-MIF method.

Table 6.1 Comparative analysis of the performance measures evaluated for CT-MR image pair-1

Methods	Performance Measures				
	En	STD	MI	SF	XEI
Nirmla et al., 2016 [137]	5.0760	60.803	2.6754	7.8695	0.4369
Das and Kundu, 2012 [34]	5.3184	83.368	2.8223	7.3292	0.4866
Li et al. (GFF), 2013 [94]	5.1317	63.433	1.0962	7.8214	0.4286
Ganasala and Kumar, 2016 [48]	5.2205	82.372	2.5842	5.8048	0.5114
Yang et al., 2017 [203]	5.2361	83.448	2.5875	7.5436	0.5375
Zhu et al. (HMSD-GDGF), 2018 [218]	5.2728	82.2183	2.8077	7.7862	0.5013
Liu et al (CNN), 2017 [111]	5.1036	79.4688	2.6576	8.3656	0.5675
C-MIF	5.5147	84.887	2.8824	8.8542	0.5346
Hermessi et al., 2018 [57]	5.1254	71.79	2.5340	7.5836	0.4336
Hou et al., 2019 [59]	5.5137	80.064	2.6633	8.3548	0.4958
Proposed CNN-MIF	**5.5985**	**84.964**	**2.9418**	**8.8566**	**0.5822**

Table 6.2 Comparative analysis of the performance measures evaluated for CT-MR image pair-2

Methods	Performance Measures				
	En	STD	MI	SF	XEI
Nirmla et al., 2016 [137]	4.901	60.823	2.4851	7.1971	0.4043
Das and Kundu, 2012 [34]	4.9108	89.489	2.9532	6.6292	0.4729
Li et al. (GFF), 2013 [94]	4.6649	66.054	1.3061	6.2101	0.419
Ganasala and Kumar, 2016 [48]	4.8321	85.214	2.8754	6.2896	0.5308
Yang et al., 2017 [203]	4.8843	87.183	2.8417	6.5436	0.5547
Zhu et al. (HMSD-GDGF), 2018 [218]	5.0989	88.006	2.8441	7.0792	0.5681
Liu et al (CNN), 2017 [111]	4.7416	84.192	2.7226	7.7298	0.6298
C-MIF	5.2241	**91.124**	3.0874	**7.7845**	0.5584
Hermessi et al., 2018 [57]	5.2938	81.122	2.6920	6.9385	0.5021
Hou et al., 2019 [59]	5.3007	84.757	2.7757	7.7423	0.5439
Proposed CNN-MIF	**5.4618**	90.142	**3.1256**	7.7321	**0.6372**

Table 6.3 Comparative analysis of the performance measures evaluated for CT-MR image pair-3

Methods	Performance Measures				
	En	STD	MI	SF	XEI
Nirmla et al., 2016 [137]	4.7054	60.823	2.7542	6.7467	0.4104
Das and Kundu, 2012 [34]	4.9361	81.587	3.0489	6.6242	0.5088
Li et al. (GFF), 2013 [94]	4.6573	58.294	1.4823	6.4681	0.474
Ganasala and Kumar, 2016 [48]	4.7692	79.859	3.0513	5.8048	0.5328
Yang et al., 2017 [203]	4.8087	81.464	2.9881	6.5477	0.5461
Zhu et al. (HMSD-GDGF), 2018 [218]	4.8141	80.935	3.0815	6.8501	0.5418
Liu et al (CNN), 2017 [111]	4.8860	76.169	2.8743	7.2448	0.6533
C-MIF	5.1623	82.784	3.2654	6.9098	0.5682
Hermessi et al., 2018 [57]	5.2991	72.629	2.8224	6.6074	0.5108
Hou et al., 2019 [59]	5.208	78.576	2.9878	7.4378	0.5729
Proposed CNN-MIF	**5.7207**	**83.395**	**3.3056**	**7.2613**	**0.6857**

Table 6.4 Comparative analysis of the performance measures evaluated for CT-MR image pair-4

Methods	Performance Measures				
	En	STD	MI	SF	XEI
Nirmla et al., 2016 [137]	4.6585	60.926	2.7941	6.3218	0.4318
Das and Kundu, 2012 [34]	4.724	84.361	3.0406	6.7005	0.4203
Li et al. (GFF), 2013 [94]	4.5797	62.779	1.4205	6.8758	0.4616
Ganasala and Kumar, 2016 [48]	4.7512	80.108	3.0113	5.6939	0.505
Yang et al., 2017 [203]	4.7173	81.997	2.9531	5.8631	0.5046
Zhu et al. (HMSD-GDGF), 2018 [218]	4.5953	83.695	3.0517	6.8237	0.5388
Liu et al (CNN), 2017 [111]	4.871	81.716	3.0315	7.2616	0.6635
C-MIF	5.1422	85.263	3.1847	6.9684	0.5086
Hermessi et al., 2018 [57]	5.3659	75.363	2.8634	6.490	0.4524
Hou et al., 2019 [59]	5.3873	79.650	2.9672	7.6164	0.5511
Proposed CNN-MIF	**5.2838**	**85.488**	**3.2764**	**7.3387**	**0.6672**

Table 6.5 Averaged evaluation measures (mean ± standard deviation) for CT-MR images

Fusion Methods	Performance Measures				
	En	STD	MI	SF	XEI
Source MR	3.8803 ± 0.3366	59.3509 ± 4.9843	-	6.3309 ± 0.8517	-
Source CT	3.0647 ± 0.3035	83.9079 ± 5.4953	-	5.2680 ± 0.6784	-
Nirmla et al. [137]	4.8122 ± 0.1812	60.8751 ± 0.0445	2.9276 ± 0.2097	7.0909 ± 0.4656	0.4505 ± 0.1289
Das and Kundu [34]	4.8081 ± 0.3713	86.9047 ± 4.4341	3.0086 ± 0.1102	6.2976 ± 0.6017	0.4785 ± 0.0490
Li et al. (GFF) [94]	4.6501 ± 0.2949	64.0659 ± 3.7542	1.3093 ± 0.2232	6.3345 ± 0.6796	0.4458 ± 0.0530
Ganasala and Kumar [48]	4.9135 ± 0.2426	84.0917 ± 4.3443	2.8656 ± 0.1751	5.8675 ± 0.4616	0.5199 ± 0.0513
Yang et al. [203]	5.0273 ± 0.2126	84.9058 ± 4.3424	2.9026 ± 0.2145	6.4654 ± 0.6297	0.5477 ± 0.0436
Zhu et al. [218]	4.9127 ± 0.3124	83.1357 ± 3.157	2.9225 ± 0.1497	7.1182 ± 0.4261	0.5384 ± 0.0448
Liu et al (CNN) [111]	4.7219 ± 0.2021	80.6907 ± 3.2741	2.7915 ± 0.1742	7.2672 ± 0.6489	0.6118 ± 0.1451
C-MIF	5.1529 ± 0.2378	88.5082 ± 4.4684	3.1594 ± 0.1637	7.1657 ± 1.1289	0.5578 ± 0.0465
Hermessi et al., 2018 [57]	5.3857 ± 0.3374	75.7403 ± 3.3521	2.7978 ± 0.1726	7.1092 ± 0.4521	0.4954 ± 0.0403
Hou et al., 2019 [59]	5.467 ± 0.3760	81.2753 ± 2.055	2.9184 ± 0.1784	7.4922 ± 0.3891	0.5616 ± 0.0324
Proposed CNN-MIF	**5.6749** ± 0.2554	**89.9548** ± 4.2868	**3.4304** ± 0.1244	**7.5627** ± 0.6894	**0.6559** ± 0.0419

6.7.2 Comparison of anatomical-functional image fusion

To assess the experimental results of anatomical and functional image fusion, five different pairs of color medical images (MR-SPECT/PET) are considered and a comparative analysis is presented to demonstrate the superiority of the work. In the testing dataset, MR-T2 images show structural details of the fluidic and fatty tissues (unusual shape or size change of the tissues indicates the presence of an abnormality). Due to the poor accuracy of the physiological details present in the MR images, functional activity based another imaging modality (SPECT/PET) is used to clarify the significant information about the metabolic functioning of tissues. From Figure 6.10 (a) and (b), pair-1, 2, and 3 depicts MR-T2 and 99mTc-SPECT images, while the pair-4 and 5 show MR-T2 and PET-FDG images, respectively. The fused images obtained by the proposed method for these inputs are presented in Figure 6.10 (c). The resultant images present a better quality image with the significant structural and color content captured from the input images. To validate the performance of the proposed method, the quantitative measures are also compared with few objective metrics scored by the source images.

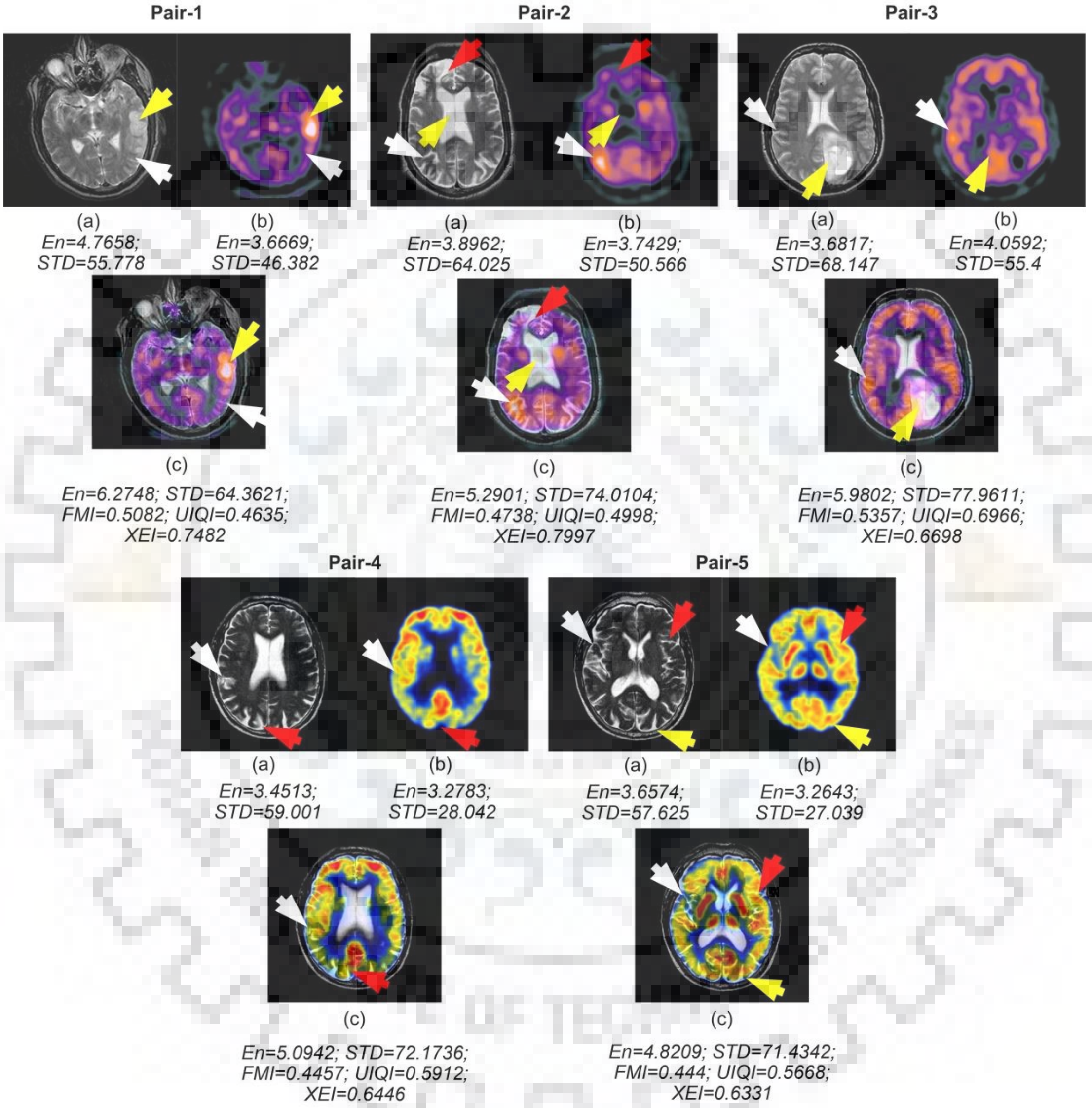


Figure 6.10 Fused images using the proposed method on various test neurological images along with quantitative metrics. Pair-1, 2 and 3 Input image (a) MR image (b) SPECT image (c) Fused image, Pair-4 and 5 Input image (a) MR image (b) PET image (c) Fused image

The comprehensive objective analysis of testing image pairs is given in Table 6.6 to Table 6.8 with a statistical comparison between the proposed and some previously proposed multimodal fusion methods. In Figure 6.11 to Figure 6.13, the comparisons of visual outcome are given with eight state-of-art approaches and the proposed CNN-MIF approach. Figure 6.11 (c) to Figure 6.13 (c) demonstrate the work proposed by Bhatnagar et al. [18] based on directive contrast in the NSCT domain, shows a good amount of luminance restoration, but the edge details present in the source images are completely blurred in the fused image. The objective parameters also indicate a lower value of FMI and SF to define the poor structural details such as color, and edge preservation while the En achieves a moderated metric value makes the overall fused image quality index lower than other listed methods. Furthermore, this method lacks to show the edges near the boundary region for all testing image pairs. Figure 6.11 (d) to Figure 6.13 (d) shows the GFF based [94] fusion results that clearly preserves the spatial information well, but limited by poor spectral variation and proportions. From Table 6.6 to Table 6.8, it can be seen that this approach obtains better IQI and XEI compared to the previous approaches, however, still achieves lowest STD due to poorly defined fine details. In Figure 6.11 (e), another method based on prominent region detection in NSCT domain presented by Yang et al. [200], has nicely captured the abnormal tissue details with good pixel contrast, but the structural details shown by the red arrow have become little vague and visually less perceptible in the resultant. Similarly, the averaged objective performance of both the testing image pairs shows a comparatively lower value of En and IQI to represent the lower information content, hence degrades the image quality.

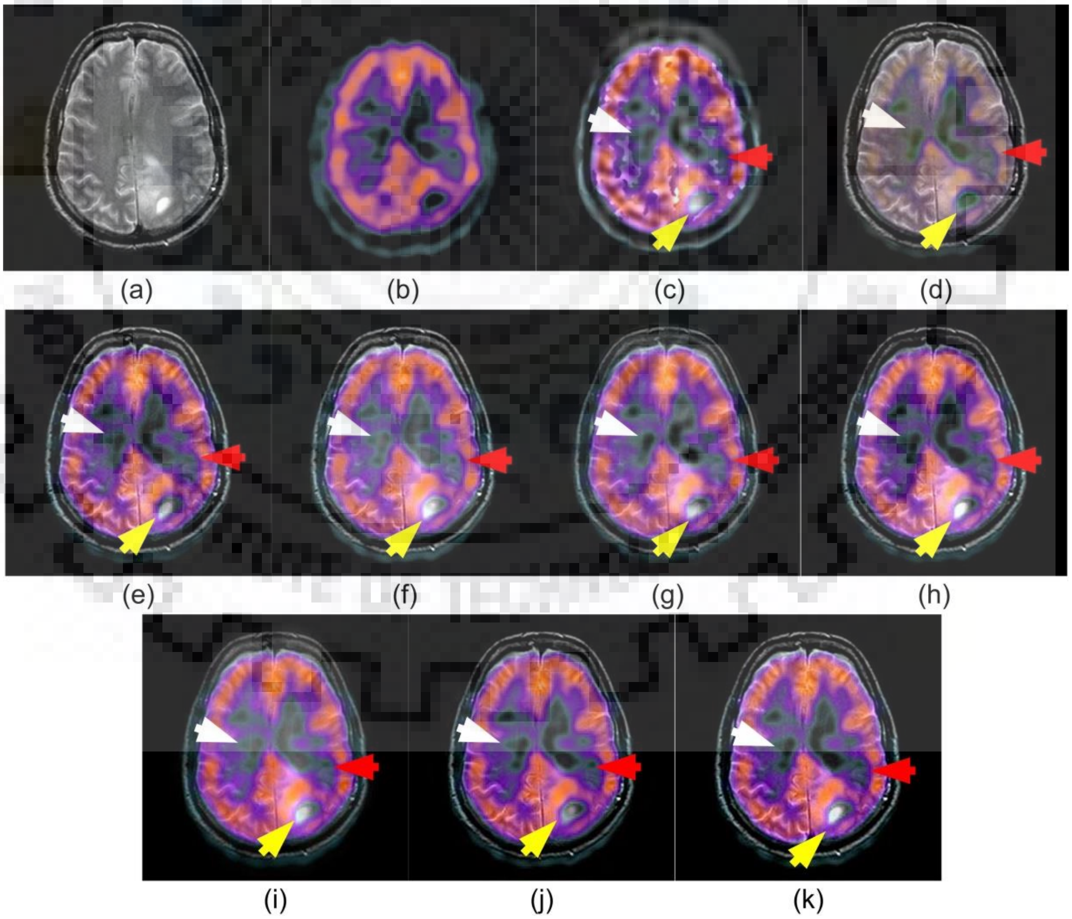


Figure 6.11 Input image pair-1 (a) MR image (b) SPECT image, qualitative comparisons of fused images obtained by (c) Bhatnagar et al., 2013 [18] (d) Li et al. 2013 [94] (e) Yang et al., 2015 [200] (f) Yang et al. 2016 [201] (g) Yang et al., 2017 [203] (h) Liu et al., 2017 [111] (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

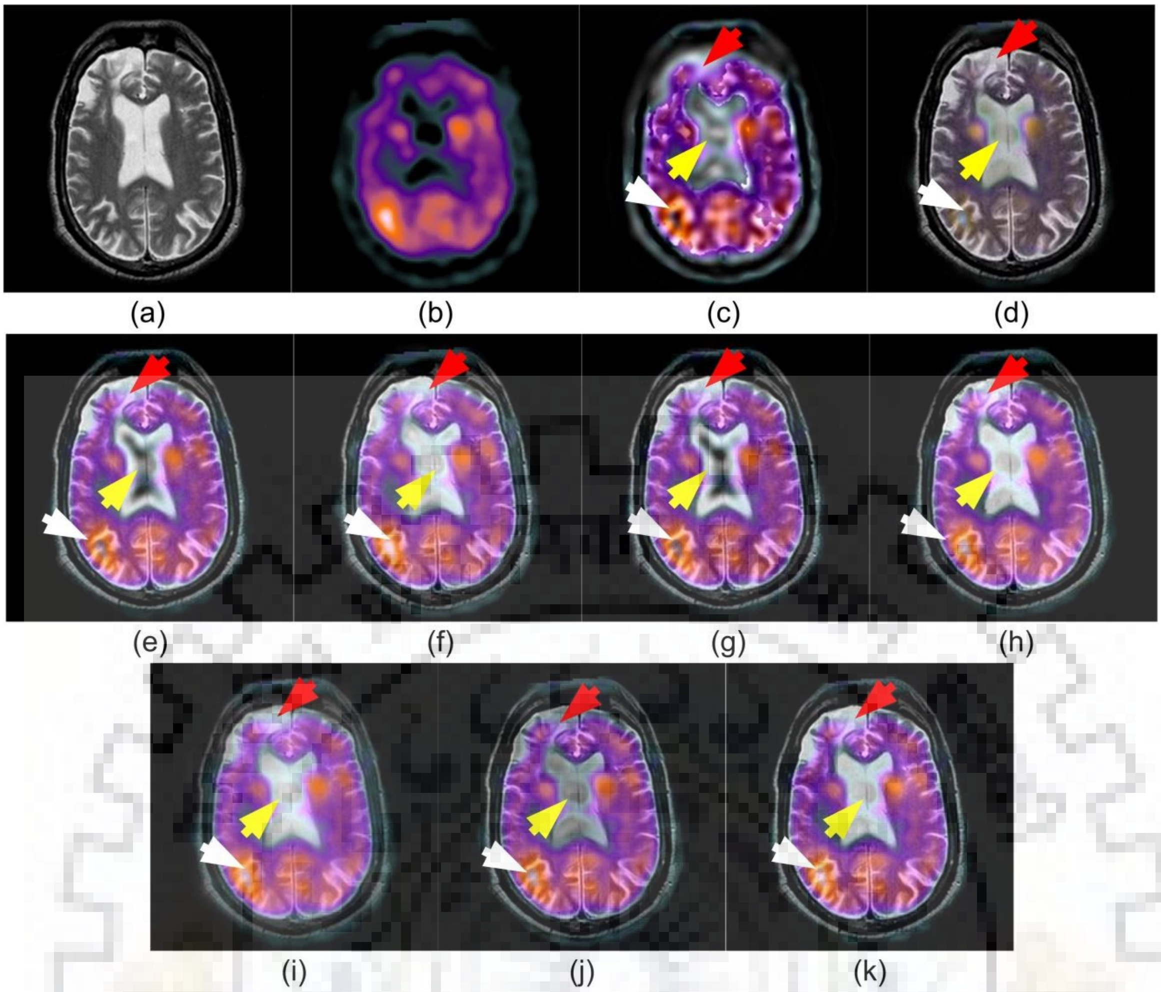


Figure 6.12 Input image pair-2 (a) MR image (b) SPECT image, qualitative comparisons of fused images obtained by (c) Bhatnagar et al., 2013 [18] (d) Li et al. 2013 [94] (e) Yang et al., 2015 [200] (f) Yang et al. 2016 [201] (g) Yang et al., 2017 [203] (h) Liu et al., 2017 [111] (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

Another fusion approach proposed by the same authors using type-II fuzzy logic for multimodal medical image data [201] has gained significant information content about the source inputs, but the tissue area shown by the yellow arrows for all datasets suffers from poor spatial contrast as shown in Figure 6.11 (f) to Figure 6.13 (f). The average performance as listed in Table 6.9 also signifies the lower values of FMI, SF and XEI indices achieved by this method compared to the other three methods discussed previously. In Figure 6.11 (g) to Figure 6.13 (g), fuzzy adaptive PCNN based fusion approach has been employed for performance assessment and comparison. The subjective results show better spatial consistency and spectral preservation, but lacks of few basic anatomical details such as some focal lesions having different information and morphology (shown by the white arrow) are roughly defined that causing a poor visual perception of some areas. The lower value of En obtained by the respective method verifies the visual performance assessment. Figure 6.11 (h) to Figure 6.13 (h) show the fused results obtained using a CNN based approach proposed by Liu et al. [111] that achieves a good amount of average informative content and spatial consistency, hence depicts an overall good quality fused image. The objective results have also scored a good value of En and SF. However, some tiny edge details appeared less sharp in the salient regions as indicated by the red and yellow arrows in the results for both the tested image pairs and also validated by poor value of STD and XEI metrics as given in Table 6.6 to Table 6.8. Figure 6.11 (i) to Figure 6.13 (i) shows clearly the preservation of structural content while the fine edges are blurred and inconsistent in the fused image. Figure 6.11 (j) to Figure 6.13 (j) shows the overall detail information in the fused images with moderated SF values, but limited by poor contrast and lower STD. The fused images obtained by the proposed method are shown in Figure 6.11 (k) to Figure 6.13 (k) which shows the ability to extract the object and edge detail without incorporating any artifact (transform technique is sensitive to false edges). The fine lines near the boundary area are clearly visible with appropriate sharpness. The regions marked by yellow arrows in the testing image pairs are representing proper localization of spatial information and consistency with respect to the nearby neighboring pixels. The method achieves the correct level of pixel contrast and comparable to the source image dataset.

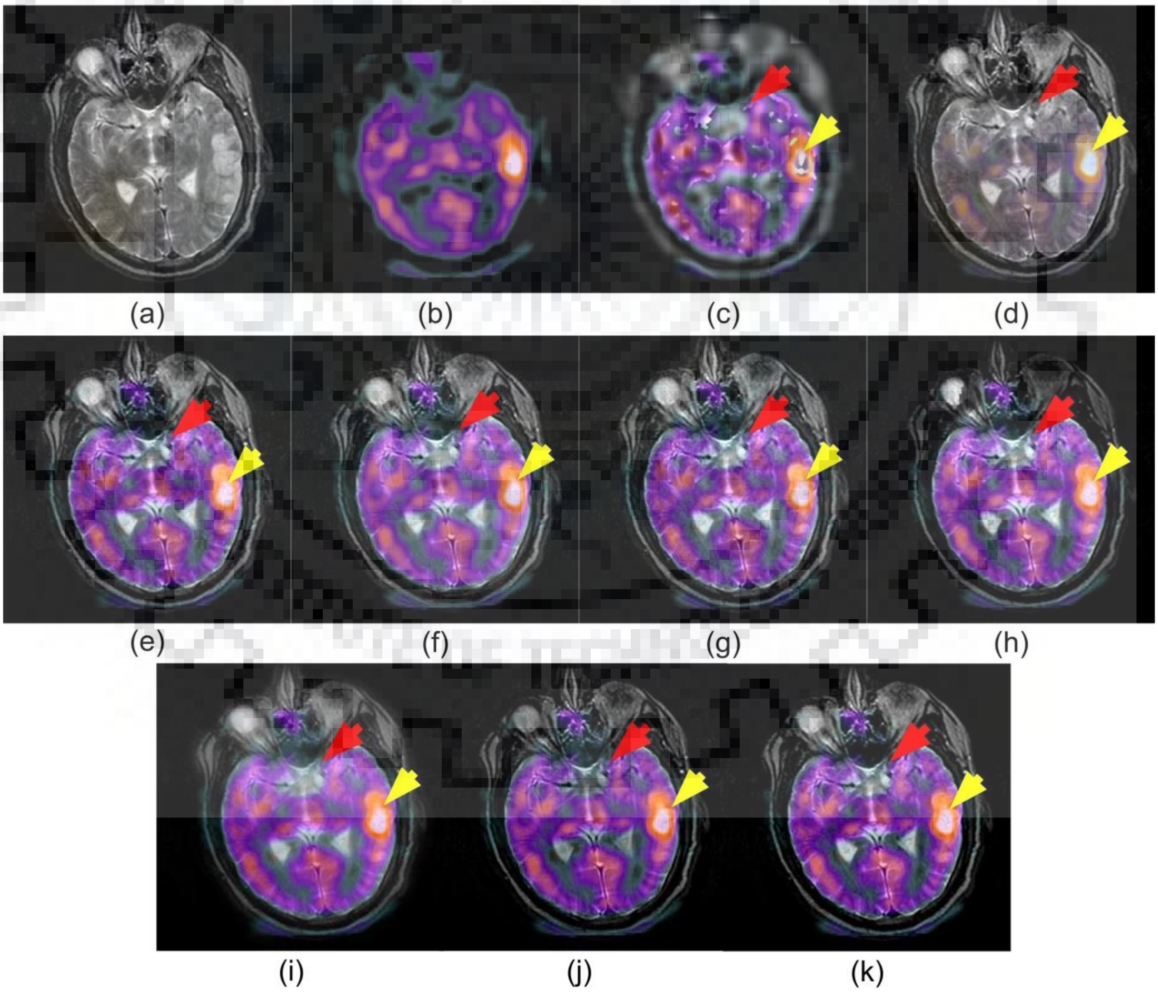

(a) (b) (c) (d)
(e) (f) (g) (h)
(i) (j) (k)

Figure 6.13 Input image pair-3 (a) MR image (b) SPECT image, qualitative comparisons of fused images obtained by (c) Bhatnagar et al., 2013 [18] (d) Li et al. 2013 [94] (e) Yang et al., 2015 [200] (f) Yang et al. 2016 [201] (g) Yang et al., 2017 [203] (h) Liu et al., 2017 [111] (i) Hermessi et al., 2018 [57] (j) Hou et al., 2019 [59] (k) Proposed CNN-MIF

Furthermore, the objective performance analysis has also shown that the proposed framework is able to capture the significant visual features with higher FMI score and can preserve the significant luminance and structural information (represented by the higher value of En and XEI metrics). Hence, by analyzing the computed performance measures for the proposed work listed in Table 6.6 to Table 6.8, it is clear that the method provides visually perceptible information with fine details localization available in any source images individually.

Table 6.6 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-1

Fusion Methods	Performance Measures					
	En	STD	FMI	SF	IQI	XEI
Bhatnagar et al., 2013 [18]	4.5028	64.4275	0.2072	6.2531	0.3186	0.1709
Li et al. (GFF), 2013 [94]	4.5323	53.3025	0.3418	6.4931	0.4083	0.6036
Yang et al., 2015 [200]	4.3769	65.4121	0.4212	7.2324	0.3780	0.6455
Yang et al. (NSCT-Fuzzy Type 2), 2016 [201]	5.0614	64.7547	0.3157	6.7829	0.4282	0.4318
Yang et al., (NSCT-Fuzzy-PCNN) 2017 [203]	4.5904	64.0905	0.4005	7.3659	0.5163	0.6314
Liu et al. (CNN), 2017 [111]	4.8457	62.0423	0.4230	6.9828	0.5314	0.6347
Hermessi et al., 2018 [57]	5.1564	63.0452	0.2567	5.9659	0.3869	0.3936
Hou et al., 2019 [59]	5.0719	58.4451	0.2543	6.6938	0.3836	0.4047
Proposed CNN-MIF	**5.1298**	**65.8217**	**0.4986**	**7.6333**	**0.6337**	**0.6564**

Table 6.7 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-2

Fusion Methods	Performance Measures					
	En	STD	FMI	SF	IQI	XEI
Bhatnagar et al., 2013 [18]	4.6860	65.4292	0.2008	6.2755	0.2558	0.1443
Li et al. (GFF), 2013 [94]	4.5575	59.5072	0.3805	6.8999	0.3455	0.6031
Yang et al., 2015 [200]	4.5983	71.9412	0.4440	7.3450	0.3321	0.7153
Yang et al. (NSCT-Fuzzy Type 2), 2016 [201]	5.2762	72.6751	0.2778	6.9753	0.3891	0.4912
Yang et al., (NSCT-Fuzzy-PCNN) 2017 [203]	4.7683	68.9441	0.4365	7.5129	0.3986	0.6960
Liu et al. (CNN), 2017 [111]	4.8582	69.6076	**0.4762**	7.0191	0.4047	0.7036
Hermessi et al., 2018 [57]	5.2195	70.5628	0.2396	5.7947	0.3560	0.4633
Hou et al., 2019 [59]	5.0865	59.9375	0.2362	6.7957	0.3607	0.3725
Proposed CNN-MIF	**5.2901**	**74.0104**	0.4738	**7.6118**	**0.4998**	**0.7997**

Table 6.8 Comparative analysis of the performance measures evaluated for MR-SPECT image pair-3

Fusion Methods	Performance Measures					
	En	STD	FMI	SF	IQI	XEI
Bhatnagar et al., 2013 [18]	5.1412	57.1206	0.2087	6.0231	0.2468	0.1415
Li et al. (GFF), 2013 [94]	5.4785	54.6073	0.2312	7.8485	0.358	0.5559
Yang et al., 2015 [200]	5.5484	62.0524	0.4925	8.3679	0.455	0.7417
Yang et al. (NSCT-Fuzzy Type 2), 2016 [201]	6.2067	64.2226	0.4301	7.8535	0.453	0.5511
Yang et al., (NSCT-Fuzzy-PCNN) 2017 [203]	5.6811	62.1164	0.484	8.3879	0.3826	0.7192
Liu et al. (CNN), 2017 [111]	5.4754	59.3141	0.4959	7.8703	0.4166	0.6694
Hermessi et al., 2018 [57]	5.6154	61.2140	0.2154	6.4328	0.3461	0.5153
Hou et al., 2019 [59]	5.5461	56.7211	0.2135	7.6451	0.3589	0.5217
Proposed CNN-MIF	**6.2748**	**64.3621**	**0.5082**	**8.5375**	**0.4635**	**0.7482**

The averaged comparative analysis of the color medical images is presented in Table 6.9 to assess the overall fusion performance using different state of the art approaches and the proposed in terms of information, color and structural content. The averaged measures indicate that the fusion outcomes obtained by the proposed method contain most of the information with appropriate spatial consistency and spectral balance. The fused results also retain the higher degree of structural similarity and able to mutually correspond to the source images.

Table 6.9 Averaged evaluation measures (mean ± standard deviation) for MR-SPECT images

Fusion Methods	Performance Measures					
	En	STD	FMI	SF	IQI	XEI
Bhatnagar et al., 2013 [18]	4.3566 ± 0.6876	65.146 ± 7.5963	0.1948 ± 0.0135	5.7691 ± 0.7218	0.2455 ± 0.0495	0.1394 ± 0.0253
Li et al., 2013 [94]	4.3049 ± 0.6636	57.343 ± 9.1360	0.3855 ± 0.0515	7.0841 ± 0.7412	0.3777 ± 0.0607	0.6648 ± 0.0621
Yang et al., 2015 [200]	4.3564 ± 0.6686	68.209 ± 9.0919	0.4493 ± 0.0348	7.0941 ± 0.8285	0.3376 ± 0.0598	0.6879 ± 0.0621
Yang et al., 2016 [201]	4.8688 ± 0.6712	69.114 ± 7.2812	0.3593 ± 0.0379	6.9678 ± 0.5698	0.4165 ± 0.1227	0.6169 ± 0.1221
Yang et al., 2017 [203]	4.5966 ± 0.6623	67.262 ± 8.5285	0.4447 ± 0.0349	7.2612 ± 0.6741	0.4020 ± 0.0404	0.6842 ± 0.0645
Liu et al. (CNN), 2017 [111]	4.6984 ± 0.6372	65.509 ± 6.4233	0.4619 ± 0.0377	7.2421 ± 0.6006	0.4454 ± 0.1363	0.4166 ± 0.1206
Hermessi et al., 2018 [57]	5.7734 ± 0.3598	68.326 ± 5.8216	0.2374 ± 0.0326	6.1468 ± 0.2660	0.3175 ± 0.0952	0.3234 ± 0.1010
Hou et al., 2019 [59]	5.0227 ± 0.5112	56.033 ± 7.5614	0.2253 ± 0.0259	6.8656 ± 0.5813	0.3282 ± 0.0896	0.3498 ± 0.0954
Proposed CNN-MIF	5.2184 ± 0.5908	69.385 ± 7.7885	0.4953± 0.0377	7.5582 ± 0.5778	0.4928 ± 0.1175	0.7306 ± 0.1235

6.7.3 Computational complexity analysis

In this section, the computational efficiency of the proposed CNN-MIF is discussed and compared with the state-of-the-art methods. The comprehensive algorithm complexity and computational time of the proposed CNN-MIF method are compared with eight different fusion methods and given in Table 6.10. All the experiments have been performed using MATLAB 2014a and a machine having Intel™ Core2Duo 2.93 GHz processor with 8 GB RAM on Windows 10. The computational complexity of the proposed approach is explained as follows.

The number of computations required for the layer decomposition method is $O(Nlog(N))$. The base layer and detail layer fusion process are taking approximately 15 seconds and 4 seconds, respectively. Since in the test stage, a pre-trained generalized CNN model [111] is adopted for base layer feature mapping that reduces the corresponding computation significantly. From the rest of the process, the most time-consuming operation is a PCA basis generation that cost approximately 172 seconds per image for structural clustering (FCM based) and dictionary construction for training dataset. The computational cost is significantly reduced by using a pre-computed PCA dictionary. The average execution time of the various state of the art approaches is given with the proposed one. From the results, it can be clearly visualized that the proposed method moderately performed in time complexity and takes about 19~20 seconds as running time. The parameter settings for the implementation of other methods listed in Table 6.10 are considered similarly as mentioned in their respective methods.

Table 6.10 Comparison of averaged running time (mean ± standard deviation) taken by different fusion methods

Fusion Methods	Running time (second)
Nirmla et al., 2010 [137]	0.5852 ± 0.0265
Das and Kundu., 2012 [34]	62.909 ± 4.6848
Li et al. (GFF), 2013 [94]	0.4552 ± 0.0507
Yang et al., 2015 [200]	7.8319 ± 0.5935
Ganasala and Kumar, 2016 [48]	6.1528 ± 0.4892
Yang et al. (NSCT-Type2 Fuuzy), 2016 [201]	10.161 ± 0.5842
Yang et al., 2017 [203]	89.282 ± 6.7748
Zhu et al. (HMSD-GDGF), 2018 [218]	28.313 ± 0.6421
Liu et al. (CNN), 2017 [111]	20.511 ± 0.0341
Hermessi et al., 2018 [57]	466.57 ± 15.521
Hou et al., 2019 [59]	42.096 ± 6.9140
Proposed CNN-MIF	19.312 ± 0.0219

6.8 Summary

In this chapter, a unified fusion framework is presented based on two-layer norm decomposition, feature learning and decision mapping (for base and detail layer with a different approach). The proposed CNN-MIF method has an ability to preserve the layer information (structural, fine details, brightness, and color) with relatively low computational complexity, thus improves overall fusion accuracy. In the proposed CNN-MIF method, novel ℓ_n-norm based two-scale hybrid layer decomposition is utilized to separate out the main information from the texture details in the spatial domain. Different from many state of the art methods based on sparse representation or CNN in MST domain, the proposed CNN-MIF method works on the spatial domain to maintain the pixel consistency and suppress the noise and artifacts, which is usually considered as one of the basic problems of MST based fusion approaches. Based on the visual sensitivity of the HVS model most prominent features are extracted from both the base and detail layer and combined using the feature and fine texture preserving fusion rules. Several experimental results have been presented on MR-CT, MR-PET and MR-SPECT dataset to validate the proposed method and showed a detailed comparative analysis with the other existing methods. The experimental results show that the proposed CNN-MIF method can efficiently extract the complex structure and maintain the spectral information as well without producing any processing artifacts. The comparative results demonstrated that the proposed CNN-MIF method is superior to many recent works in terms of the preservation of significant information with improved contrast and fusion quality of multimodal images.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CHAPTER 7: CONCLUSIONS AND SCOPE FOR FUTURE WORK

This chapter presents the conclusions based on the performance of all the proposed multimodal medical image fusion methods discussed in the previous chapters. Different fusion approaches for anatomical-anatomical and functional-anatomical medical images have been proposed to gratify all the requirements for precise diagnosis and treatment planning of several neurological diseases. The chapter also suggests several possible directions of the present work to be investigated in the future.

7.1 Conclusions

This thesis has focused on the multimodal image fusion for neurological images, especially CT, MR, SPECT and PET. The major objectives of this study are to design and develop the effective fusion methods and to improve the visual perception of multimodal medical images by fusing all the anatomical and functional information without introducing any structural distortion as well as without altering the functional content which are helpful to the radiologists by providing a complete pictorial view of the soft and hard tissues in a single fused image. Besides the fusion of anatomical and functional neurological images, attention has also been given to develop more generalized algorithms (proposed ADP-MIF, H-MIF, SDL-MIF and CNN-MIF) so that they can be applied not only to fuse MR-SPECT and MR-PET images but also for fusing the CT and MR images.

The accurate fusion of the multimodal images provides a meaningful quantifiable interpretation which is most suitable for clinical diagnosis by providing the exact location and orientation of the defected tissues. Thus, three fusion approaches such as NSST-MIF, DRT-MIF and C-MIF have been developed for specially anatomical to the anatomical (CT-MR) images to capture both the relevant features taken from CT and MR images, while four other fusion methods such as ADP-MIF, H-MIF, SDL-MIF and CNN-MIF have also been developed for particularly anatomical to functional image fusion to boost the diagnostic accuracy and precisely spot the severity of the deformities. Moreover, these fusion approaches work well on the CT and MR images by enhancing the visualization of the fused images.

Based on the several experimental results, different conclusions have been drawn at various stages of the presented work, which are summarized below.

7.1.1 Proposed nonsubsampled shearlet domain CT-MR image fusion approach

In the first approach presented in Chapter 2, anatomical image (CT-MR) fusion has been performed in the NSST domain by using NSML motivated PCNN as a bio-inspired neural network model to fuse the hf NSST subband coefficients and regional energy based an activity level measure is used to fuse lf subband coefficients. The performance of the proposed NSST-MIF method is compared with eight transform domain methods WT, NSCT and NSST with different fusion rules such as averaging of lf coefficient fusion, maximum and spatial frequency based rule for hf subband fusion. Quantitative analysis of the experimental results using various performance measures such as entropy (En), standard deviation (STD), mutual information (MI), spatial frequency (SF), image quality index (IQI) and Xydeas edge index (XEI) shows that the proposed NSST-MIF method is able to fuse the CT and MR images in a better way than the other existing methods considered in Chapter 2 without altering the information.

From the comparative evaluation and analysis of the experimental results, it is observed that the proposed NSST-MIF approach gains approx 62.77% and 28.56% higher En value, 4.32% and 47.49% higher STD value, 29.22% and 7.53% higher SF value than the source CT and MR image dataset, respectively. Moreover, the proposed NSST-MIF approach gets higher STD values of approx. 37.96%, 41.1%, 33.02%, 1.01% and 0.41% higher STD and 6.62%, 6.08%, 5.14%, 4.57%, 2.09% higher En values than the WT_AVG, NSCT_AVG and NSST_AVG (averaging based rule), NSCT_PCNN and NSST_PCNN, respectively. It indicates better fusion performance provided by the proposed NSST-MIF approach by providing more information with a better contrast level. Furthermore, the proposed NSST-MIF method gains an improvement by approx 1.93-9% and 9.5%-14.5% higher IQI and XEI values than the NSCT and NSST based conventional PCNN models. Thus, from the observation, it is clear that the proposed method shows a significant improvement in terms of fusion performance and detectability of the source images.

7.1.2 Proposed ripplet domain CT-MR image fusion approach

In order to get fused image having more edge detail and good contrast for both the soft and hard tissues, the second anatomical image fusion approach (DRT-MIF) is presented in Chapter 2 which is developed based on the combination of the DRT and PCNN by improving the feeding inputs. It provides more details present in the source CT and MR images and further enhances the visualization of the fused images. In the proposed DRT-MIF approach, the PCNN model is utilized for both the lf and hf DRT subimage coefficients based on the firing times and improved feeding inputs by estimating the NSML and NMSF which is able to capture the suitable differences and provide the resultant images with higher contrast and clarity. From the quantitative analysis of the experimental results shown in Chapter 2, it is observed that the proposed DRT-MIF shows an improved fusion performance than the NSST-MIF and other WT and NSCT based fusion methods by producing higher En and STD value by approx (2.72%, 9.51%, 5.9%-8.88% and 4.85%) and (1.1%, 39.46%, 1.7%-41.54% and 1.51%), respectively. Moreover, the proposed approach gains approx (67.17% and 32.03%), (5.46% and 49.1%) and (35.81% and 13.01%) higher En, STD and SF values than the source CT and MR images. From the comparative results, it is also visualized that the proposed DRT-MIF approach takes more time as compared to WT and NSST-MIF approach, but it is acceptable at the cost of improved quality of fused images having more detailed information which is also evident by producing the higher performance measures such as En, MI, SF and XEI by approx 4.85%, 3.26%, 10.87% and 11.96%, respectively, than the NSST-MIF approach. Thus, it is concluded that the proposed DRT-MIF approach helps to preserve diagnostic information as much as possible and provides a better quality of fused images.

7.1.3 Proposed cascaded model for CT-MR image fusion in NSST and DRT domain

The efforts in the earlier approaches were focused on the improvement in the fusion performance of CT-MR images with a better visualization. These efforts have been successful to a considerable extent resulting in more effective among the fusion approaches while improving or maintaining the same qualitative performance. Thus, a cascaded fusion framework (C-MIF) is developed based on the concept of the DRT and PCNN model in the first stage and NSST decomposition at stage 2 in which NSML and NMSF are also utilized for lf and hf subimage coefficients in the DRT domain. At the second stage, the SAD and AM based fusion model is used to provide the richer representation of the edge detail information and contrast, respectively.

From the comparative evaluation and analysis shown in Chapter 2, it is observed that the proposed C-MIF approach gains approx (68.14% and 32.8%), (5.49% and 49.13%) and (36.03% and 13.19%) higher averaged En, STD and SF values compared to the input source CT and MR images. It indicates a better quality of fused images having more informative content with better contrast. Moreover, the proposed C-MIF approach outperforms both the NSST-MIF and DRT-MIF methods. The proposed C-MIF approach also achieves approx 3.3%, 2.9%, 1.12%, 5.27%, 4.58% higher averaged En, MI, STD, SF and XEI values, respectively, than the NSST-MIF approach. Also, approximately 0.6%, 0.84% and 1.31% higher En, MI and XEI values are achieved by the proposed C-MIF method as compared to the DRT-MIF approach. Furthermore, the proposed C-MIF approach provides the fused images with better perceptual quality than the other decomposition based fusion methods. It is validated by achieving approx (6.54%-9.51%, 4.11%-5.1%, 1.53%-1.89%, 10.78%-19.91% and 14.54%-51.01%) and (3.3-5.46%, 2.89%-4.11%, 1.12%-1.53%, 5.27%-10.78% and 4.58%-13.29%) higher En, MI, STD, SF and XEI values than the NSCT and NSST based fusion approaches, respectively. In view of the above observations, it is concluded that the CT-MR image fusion method presented in Chapter 2 contributes significantly compared to the state-of-the-art methods.

7.1.4 Proposed multimodal image fusion approach based on adaptive PCNN and log Gabor filtering

In the approach presented in Chapter 3, an improved fusion framework is proposed to fuse the anatomical as well as functional images with the anatomical neurological images. In this approach, the PCNN model with adaptive linking strength parameter based fusion rule is used to fuse the anatomical and functional images in the NSST domain. For anatomical and functional image fusion, the $l\alpha\beta$ color space model which has better decorrelation of the achromatic and chromatic channel is utilized. The proposed ADP-MIF approach is able to retain the required information without losing the resolution of the disease morphology. Moreover, the adaptive neural model based on local visibility and log Gabor energy based rules are applied to lf and hf components, respectively. Quantitative analysis of the experimental results using various performance measures shows that the proposed ADP-MIF approach is able to capture more detail information present in the input images and increase the visibility of the fused images with more contrast and clarity level without introducing any spectral distortion.

From the quantitative analysis of the experimental results shown in Chapter 3, it is observed that the proposed ADP-MIF approach works well for fusing both the anatomical and functional image with anatomical images. For the anatomical and functional image fusion, the proposed ADP-MIF approach shows an improved fusion performance by achieving En, STD, FMI, IQI and XEI values by approx. 7.19%-13.1%, 1.39%-2.82%, 0.6-1.64%, 17.44%-39.84% and 2.06%-2.62% higher than the NSCT based fusion methods. Moreover, the proposed fusion approach outperforms the HMSD-GDGF fusion approach by providing performance measures approx. 1.29%, 2.51% and 18.38% higher En, STD and IQI values, respectively. Furthermore, the proposed ADP-MIF approach also gains 3.43%, 0.61%, 12.13%, 24.8% and 3.97% higher En, STD, FMI, IQI and XEI values, respectively than the existing NSST based fusion approaches considered in Chapter 3. In addition to the fusion of MR-SPECT images, the performance of CT-MR image fusion is also evaluated for the proposed ADP-MIF approach. Based on the quantitative evaluation, it is observed that the proposed ADP-MIF approach gains approx. (6.14%, 5.7%, 11.48% and 5.74%), (3.35%, 3.59%, 6.08% and 2.43%) and (2.75%, 2.73%, 5.9% and 1.11%) higher averaged En, MI, SF and XEI values than the NSST-MIF, DRT-MIF and C-MIF approaches. Thus, it is concluded that the proposed ADP-MIF approach helps to provide a significant improvement in terms of the visual quality of fused images by providing additional diagnostic information especially for the fusion of anatomical with functional images.

7.1.5 Proposed adaptive PCNN based hybrid multimodal image fusion approach in NSST and SWT domain

In the next approach presented in Chapter 4, a hybrid medical image fusion framework is proposed which is based on both the NSST and SWT decomposition. In this approach, an adaptive PCNN model motivated by NSML based fusion rule was used to fuse lf SWT coefficients while LLGE based fusion rule has been applied for hf SWT coefficients to extract the salient features available in the source image and to retain the color and edge details without introducing any artifacts. The basic idea here was to improve the fusion results by using the different feeding input to motivate the PCNN processing lf component instead of the absolute values of lf coefficients. Absolute maximum and sum of absolute difference based fusion rule is also applied to hf NSST subbands to retain more information related to edge details. From the quantitative evaluation and analysis of the experimental results shown in Chapter 4, it is observed that the proposed H-MIF approach outperforms the other existing methods mentioned in Chapter 4 whereas in some of the cases, other approaches get higher IQI and STD values as compared to the proposed H-MIF approach. Based on comparative averaged results evaluated for the MR-SPECT image fusion, It is observed that the proposed H-MIF approach gets approx. (2.68%-8.67% and 4.85%), (2.99%-114% and 13.62%), (17.61%-18.56% and 24.98%) and (4.08%-201.4%, 5.45%) higher En, FMI, IQI and XEI values than the NSCT and NSST based fusion methods, respectively considered in the Chapter4. The proposed H-MIF approach also outperforms the ADP-MIF approach by achieving approx. 1.38%, 1.33%. 0.2% and 1.42% higher En, FMI, IQI and XEI values, respectively. However, the ADP-MIF approach gets STD value higher by a value of 2.7% than the proposed H-MIF approach. Similar performance has also been reflected for fusion of CT-MR images by the proposed approach that achieves good complementary information with more structural details, but it suffers from contrast level of fused images having lower STD value that may be acceptable with the ability to retain more diagnostic information.

7.1.6 Proposed multimodal image fusion approach using sparse K-SVD dictionary learning

The efforts in the previous sections have been made to evaluate and analyze the performance of the fusion approaches proposed in Chapters 2 to 4 which are based on conventional PCNN or adaptive PCNN model with different fusion rules in the NSST, DRT and hybrid domain. Considering the results and limitations of the previously proposed fusion approaches, a unified multimodal fusion framework (SDL-MIF) based on sparse K-SVD dictionary learning and guided filtering in the NSST domain is proposed in which an overcomplete dictionary is learned to capture the complex detailed information available in the source images. Fusion rule using a dictionary learning based sparse representation approach is utilized to improve the comprehensive information in lf NSST subband while guided filtering based fusion rule is adopted to fuse hf NSST subbands which is able to extract the salient features from the source images and reflect the color and edge detail properly in the fused outcome.

From a comparison of subjective evaluation and analysis shown in Chapter 5, it is observed that the proposed SDL-MIF approach is able to provide more informative content in the fused images with appropriate spatial consistency and spectral balance. The fusion performance of the proposed SDL-MIF approach is also analyzed in terms of all quantitative measures for all MR-CT, MR-SPECT and MR-PET images. From the quantitative results presented for CT-MR image fusion, it is observed that the proposed SDL-MIF approach gets higher En (11.57%, 8.63%, 8%, 5.11% and 3.08%), MI (10.81%, 8.6%, 7.7%, 4.83% and 4.14%), STD (2.23%, 1.14%, 1.11%, 1.61% and 8.68%) and XEI (10.25%, 6.81%, 5.43%, 4.27% and 2.58%) values than the other proposed NSST-MIF, DRT-MIF, C-MIF, ADP-MIF and H-MIF approaches, respectively. Moreover, from the experimental results shown in Chapter 5 for the fusion of anatomical with functional images, it is observed that the proposed SDL-MIF approach yield approx. 5.49%, 0.6%, 9.1%, 0.67% and 3.38% higher En, STD, FMI, IQI and XEI values, respectively than the proposed ADP-MIF approach. The proposed SDL-MIF approach outperforms the proposed H-MIF method by achieving 4.1%, 7.6%, 0.42% and 1.92% higher En, FMI, IQI and XEI values, respectively. The proposed SDL-MIF approach is also able to overcome the limitation of the proposed H-MIF approach by getting 3.24% higher STD values. Based on all the comparative results analysis, it is visualized that the proposed SDL-MIF method is able to preserve the significant information of multimodal input images by producing the better visual quality of fused images with improved contrast.

7.1.7 Proposed multimodal image fusion approach based on CNN feature mapping and structural clustering

In the last part of the work, a feature level multimodal image fusion approach (CNN-MIF) is proposed using two-scale $\ell_1 - \ell_0$ hybrid layer decomposition with CNN based feature mapping and structural patch clustering. In the proposed CNN-MIF approach, ℓ_n-norm based two scale hybrid layer decomposition is utilized to preserve the desired edges and intensity variations at each scale. A pre-trained CNN model followed by consistency verification is also used to extract the prominent features of each of the decomposed base layer components and to generate the pixel activity and fusion weight map. For merging the two-scale detail layer, a clustering based pre-learned multichannel dictionary with the saliency matching rule is utilized to efficiently map the structural details of the layers. Moreover, the color components associated with both the source images are also combined using pixel saliency measure and finally all three components, i.e. fused base layer, fused detail layer and color component get merged to reconstruct the fused image. The visual results and quantitative measures presented in Chapter 6 verify the effectiveness of the proposed CNN-MIF approach.

From a comparison of subjective evaluation and analysis shown in Chapter 6, it is observed that the proposed CNN-MIF approach is able to preserve sharp edges and visual consistency quite well. Furthermore, the objective analysis shows that the proposed CNN-MIF approach captures the significant visual features and also preserves the significant luminance and structural information which were verified by achieving higher FMI, En and XEI values. From a comparative result analysis presented for MR-SPECT image fusion, it is observed that the proposed CNN-MIF approach achieves approx. higher En (5.91%, 4.47% and 0.4%), FMI (9.58%, 8.14% and 0.51%), IQI (4.22, 4.1% and 3.62%) and XEI (4.06%, 2.6% and 0.7%) values than the other proposed ADP-MIF, H-MIF and SDL-MIF approaches, respectively. Moreover, it is also observed that the proposed CNN-MIF approach also outperforms the other existing methods mentioned in Chapter 6. Further, another comparative analysis is presented for CT-MR image fusion from which it is observed that the proposed CNN-MIF approach gains averaged En, MI, STD and XEI values higher by (13.76%, 11.7%, 2.76% and 22.97%), (10.77%, 9.49%, 1.66% and 19.12%), (10.13%, 8.58%, 1.63% and 17.59%), (7.18%, 5.69%, 2.13% and 16.29%), (5.11%, 4.99%, 9.24% and 14.41%) and (1.97%, 0.82%, 0.52% and 11.53%) than the other proposed NSST-MIF, DRT-MIF, C-MIF, ADP-MIF, H-MIF, SDL-MIF and CNN-MIF approaches.

Finally, it is stated that the developed methods such as NSST-MIF, DRT-MIF, C-MIF and ADP-MIF, H-MIF, SDL-MIF, CNN-MIF used for the multimodal image fusion (CT-MR, MR-SPECT, MR-PET) contributed significantly toward the state-of-art fusion methods.

7.2 Scope for the Future Work

Although the present work could have been able to contribute in the area of multimodal medical image fusion, the following suggestions are made for future work in this area:

(1) In this study, the proposed fusion methods were tested on brain CT/MR/SPECT/PET images. The performance of these proposed methods may be analyzed for other organ’s images for different diseases with the proper registration of source images.

(2) The performance and efficiency of the proposed methods may be analyzed based on the fusion of other imaging modalities such as CT/MR with ultrasound image, MR with optical images etc.

(3) Some more advanced fusion rules based on feature extraction, classification and prediction may be designed to localize the application-specific task used for different clinical purposes.

(4) In this study, a pre-trained CNN model is used for the automatic feature mapping of input images. In addition to such an approach, a fine tuning or transfer learning-based model training on large medical dataset may be designed to get better fusion results.

(5) The research work was carried out on 2-D slices of CT/MR/SPECT/PET images for each patient as 3-D datasets were not available. In future, the proposed fusion approaches may be extended for 3-D images.

(6) The fused images obtained using the proposed methods may be validated by the expert radiologists for a large database.

(7) It would still be a wonderful idea to develop an expert system for fusing the multimodal medical images that may assist the radiologists in taking decisions.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

LIST OF PUBLICATIONS FROM PRESENT WORK

International Journals

[1] Sneha Singh and R.S. Anand, "Multimodal medical image sensor fusion model using sparse K-SVD dictionary learning in nonsubsampled shearlet domain," *IEEE Transaction on Instrumentation and Measurement*, pp. 1-15, 2019. (DOI: 10.1109/TIM.2019.2902808)

[2] Sneha Singh and R.S. Anand, "Multimodal medical image fusion hybrid layer decomposition with CNN-based feature mapping and structural clustering," *IEEE Transaction on Instrumentation and Measurement*, pp. 1-13, 2019 (DOI: 10.1109/TIM.2019.2933341).

[3] Sneha Singh, R.S. Anand and D. Gupta, "CT and MR image information fusion scheme using a cascaded framework in ripplet and NSST domain," *IET Image Processing*, vol. 12, no. 5, pp. 696-707, 2018

[4] Sneha Singh and R.S. Anand, "Multimodal neurological image fusion based on adaptive biological inspired neural model in nonsubsampled shearlet domain," *International Journal of Imaging Systems and Technology,* vol. 29, pp. 50-64, 2019.

[5] Sneha Singh and R.S. Anand, "Ripplet domain fusion approach for CT and MR medical image information," *Biomedical Signal Processing and Control*, vol. 46, pp. 281-292, 2018.


भारतीय प्रौद्योगिकी संस्थान रुड़की
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

REFERENCES

[1] Afifi A. J. and Hellwich O., "Object depth estimation from a single image using fully convolutional neural network," in *International Conference on Digital Image Computing: Techniques and Applications* 2016, pp. 1-7.

[2] Agrawal D. and Singhai J., "Multifocus image fusion using modified pulse coupled neural network for improved image quality," *IET Image Processing,* vol. 4, pp. 443-451, 2010.

[3] Aharon M., Elad M., and Bruckstein A., "rmK-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing,* vol. 54, pp. 4311-4322, 2006.

[4] Al-Azzawi N. and Abdullah W. A. K. W., "Medical image fusion schemes using contourlet transform and PCA," in *Image Fusion and Its Applications*: InTech, 2011, pp. 98-110.

[5] Al-Azzawi N. A., Sakim H. A. M., and Abdullah A. K. W., "An efficient medical image fusion method using contourlet transform based on PCM," in *IEEE Symposium on Industrial Electronics & Applications*, 2009, pp. 11-14.

[6] Aymaz S. and Köse C., "Multi-focus image fusion using stationary wavelet transform (SWT) with principal component analysis (PCA)," in *10th International Conference on Electrical and Electronics Engineering*, 2017, pp. 1176-1180.

[7] Bai L., Xu C., and Wang C., "A review of fusion methods of multi-spectral image," *Optik,* vol. 126, pp. 4804-4807, 2015.

[8] Balasubramaniam P. and Ananthi V. P., "Image fusion using intuitionistic fuzzy sets," *Information Fusion,* vol. 20, pp. 21-30, 2014.

[9] Balasubramanian R. and Bhatnagar G., "A new contrast based image fusion using wavelet packets," in *IEEE Conference on Applications of Intelligent Systems* 2008, pp. 141-145.

[10] Bao L., Zhao D., and Zhou D., "Image fusion algorithm Based on m-PCNN," in *IEEE Second International Workshop on Education Technology and Computer Science*, 2010, pp. 235-238.

[11] Barj E. M., Afifi M., Idrissi A. A., Nassim K., and Rachafi S., "Speckle correlation fringes denoising using stationary wavelet transform. Application in the wavelet phase evaluation technique," *Optics & Laser Technology,* vol. 38, pp. 506-511, 2006.

[12] Barra V. and Boire J. Y., "A general framework for the fusion of anatomical and functional medical images," *NeuroImage,* vol. 13, pp. 410-424, 2001.

[13] Bayram D. and Åžeker S., "Redundancy-based predictive fault detection on electric motors by stationary wavelet transform," *IEEE Transactions on Industry Applications,* vol. 53, pp. 2997-3004, 2017.

[14] Bernal J., Kushibar K., Asfaw D. S., Valverde S., Oliver A., Martí R., and Lladó X., "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review," *Artificial intelligence in medicine,* vol. 95, pp. 64-81, 2018.

[15] Bhat M. and Karki M. V., "Feature selection based on PCA and PSO for multimodal medical image fusion using DTCWT," *ArXiv preprint arXiv:1701.08918,* 2017.

[16] Bhateja V., Patel H., Krishn A., Sahu A., and Lay-Ekuakille A., "Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains," *IEEE Sensors Journal,* vol. 15, pp. 6783-6790, 2015.

[17] Bhatnagar G., Jonathan Wu Q. M., and Liu Z., "Human visual system inspired multi-modal medical image fusion framework," *Expert Systems with Applications,* vol. 40, pp. 1708-1720, 2013.

[18] Bhatnagar G., Wu Q. M. J., and Liu Z., "Directive contrast based multimodal medical image fusion in NSCT domain," *IEEE Transactions on Multimedia,* vol. 15, pp. 1014-1024, 2013.

[19] Bhatnagar G., Wu Q. M. J., and Liu Z., "A new contrast based multimodal medical image fusion framework," *Neurocomputing,* vol. 157, pp. 143-152, 2015.

[20] Bruckstein A. M., Donoho D. L., and Elad M., "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM review,* vol. 51, pp. 34-81, 2009.

[21] Burt P. J. and Kolczynski R. J., "Enhanced image capture through fusion," in *4th IEEE International Conference on Computer Vision*, 1993.

[22] Cai J., Cheng Q., Peng M., and Song Y., "Fusion of infrared and visible images based on nonsubsampled contourlet transform and sparse K-SVD dictionary learning," *Infrared Physics & Technology,* vol. 82, pp. 85-95, 2017.

[23] Chai Y., Li H., and Zhang X., "Multifocus image fusion based on features contrast of multiscale products in nonsubsampled contourlet transform domain," *Optik,* vol. 123, pp. 569-581, 2012.

[24] Chavan S. S., Mahajan A., Talbar S. N., Desai S., Thakur M., and D'Cruz A., "Nonsubsampled rotated complex wavelet transform (NSRCxWT) for medical image fusion related to clinical aspects in neurocysticercosis," *Computers in Biology and Medicine,* vol. 81, pp. 64-78, 2017.

[25] Cheng S. L., He J. M., and Lv Z. W., "Medical image of PET/CT weighted fusion based on wavelet transform," in *2nd International Conference on Bioinformatics and Biomedical Engineering*, 2008, pp. 2523-2525.

[26] Chengjun L. and Wechsler H., "A Gabor feature classifier for face recognition," in *Eighth IEEE International Conference on Computer Vision. ICCV 2001*. vol. 2, 2001, pp. 270-275 vol.2.

[27] Choi M., Kim R. Y., and Kim M.-G., "The curvelet transform for image fusion," *International Society for Photogrammetry and Remote Sensing,* vol. 35, pp. 59-64, 2004.

[28] Coifman R. R. and Donoho D. L., "Translation-invariant de-noising," in *Wavelets and statistics*: Springer, 1995, pp. 125-150.

[29] Cvejic N., Bull D., and Canagarajah N., "Region-based multimodal image fusion using ICA bases," *IEEE Sensors Journal,* vol. 7, pp. 743-751, 2007.

[30] da Cunha A. L., Jianping Z., and Do M. N., "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Transactions on Image Processing,* vol. 15, pp. 3089-3101, 2006.

[31] Daneshvar S. and Ghassemian H., "MRI and PET image fusion by combining IHS and retina-inspired models," *Information Fusion,* vol. 11, pp. 114-123, 2010.

[32] Das S., Chowdhury M., and Kundu M. K., "Medical image fusion based on ripplet transform type-I," *Progress In Electromagnetics Research B,* vol. 30, pp. 355-370, 2011 2011.

[33] Das S. and Kundu M. K., "Fusion of multimodality medical images using combined activity level measurement and contourlet transform," in *IEEE International Conference on Image Information Processing*, 2011, pp. 1-6.

[34] Das S. and Kundu M. K., "NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency," *Medical & Biological Engineering & Computing,* vol. 50, pp. 1105-14, 2012.

[35] Das S. and Kundu M. K., "A neuro-fuzzy approach for medical image fusion," *IEEE Transactions on Biomedical Engineering,* vol. 60, pp. 3347-3353, 2013.

[36] Das S. and Kundu M. K., "MRI denoising using visual system response model simulating PCNN," in *2nd International Conference on Perception and Machine Intelligence*, 2015, pp. 56-63.

[37] Deng A., Wu J., and Yang S., "An image fusion algorithm based on discrete wavelet transform and canny pperator," in *International Conference on Advanced Research on Computer Education, Simulation and Modeling*, S. Lin and X. Huang, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 32-38.

[38] Deng X., Ma Y., and Dong M., "A new adaptive filtering method for removing salt and pepper noise based on multilayered PCNN," *Pattern Recognition Letters,* vol. 79, pp. 8-17, 2016.

[39] Dhawan A. P., *Medical image analysis*: John Wiley & Sons, 2011.

[40] Do M. N. and Vetterli M., "The finite ridgelet transform for image representation," *IEEE Transaction on Image Processing,* vol. 12, pp. 16-28, 2003.

[41] Draper N. R. and Smith H., *Applied regression analysis,* vol. 326: John Wiley & Sons, 2014.

[42] Easley G., Labate D., and Lim W.-Q., "Sparse directional image representations using the discrete shearlet transform," *Applied and Computational Harmonic Analysis,* vol. 25, pp. 25-46, 2008.

[43] Easley G. R., Labate D., and Lim W.-Q., "Optimally sparse image representations using shearlets," in *Fortieth Asilomar Conference on Signals, Systems and Computers* 2006, pp. 974-978.

[44] El-taweel G. S. and Helmy A. K., "Image fusion scheme based on modified dual pulse coupled neural network," *IET Image Processing,* vol. 7, pp. 407-414, 2013.

[45] Fowler J. E., "The redundant discrete wavelet transform and additive noise," *IEEE Signal Processing Letters,* vol. 12, pp. 629-632, 2005.

[46] Ganasala P. and Kumar V., "Multimodality medical image fusion based on new features in NSST domain," *Biomedical Engineering Letters,* vol. 4, pp. 414-424, 2014.

[47] Ganasala P. and Kumar V., "CT and MR image fusion scheme in nonsubsampled contourlet transform domain," *Journal of Digital Imaging,* vol. 27, pp. 407-418, 2014.

[48] Ganasala P. and Kumar V., "Feature-motivated simplified adaptive PCNN-based medical image fusion algorithm in NSST domain," *Journal of Digital Imaging,* vol. 29, pp. 73-85, 2016.

[49] Ganasala P., Kumar V., and Prasad A. D., "Performance evaluation of color models in the fusion of functional and anatomical images," *Journal of Medical Systems,* vol. 40, pp. 1-9, 2016.

[50] Geng P., Wang Z., Zhang Z., and Xiao Z., "Image fusion by pulse couple neural network with shearlet," *Optical Engineering,* vol. 51, pp. 1-7, 2012.

[51] Ghahremani M. and Ghassemian H., "Remote sensing image fusion using ripplet transform and compressed sensing," *IEEE Geoscience and Remote Sensing Letters,* vol. 12, pp. 502-506, 2014.

[52] Ghassemian H., "A review of remote sensing image fusion methods," *Information Fusion,* vol. 32, Part A, pp. 75-89, 2016.

[53] Greenspan H., Van Ginneken B., and Summers R. M., "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Transactions on Medical Imaging,* vol. 35, pp. 1153-1159, 2016.

[54] Guorong G., Luping X., and Dongzhu F., "Multi-focus image fusion based on non-subsampled shearlet transform," *IET Image Processing,* vol. 7, pp. 633-639, 2013.

[55] Gupta D., "Nonsubsampled shearlet domain fusion techniques for CT–MR neurological images using improved biological inspired neural model," *Biocybernetics and Biomedical Engineering,* vol. 38, pp. 262-274, 2018.

[56] Haghighat M. B. A., Aghagolzadeh A., and Seyedarabi H., "A non-reference image fusion metric based on mutual information of image features," *Computers & Electrical Engineering,* vol. 37, pp. 744-756, 2011.

[57] Hermessi H., Mourali O., and Zagrouba E., "Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain," *Neural Computing and Applications,* vol. 30, pp. 2029-2045, 2018.

[58] Hou B., Zhang X., Bu X., and Feng H., "SAR image despeckling based on nonsubsampled shearlet transform," *IEEE Journal of selected topics in applied earth observations and remote sensing,* vol. 5, pp. 809-823, 2012.

[59] Hou R., Zhou D., Nie R., Liu D., and Ruan X., "Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model," *Medical & Biological Engineering & Computing,* vol. 57, pp. 887-900, 2019.

[60] Hua T., Ya-nan F., and Pei-Guang W., "Image fusion algorithm based on regional variance and multi-wavelet bases," in *2nd International Conference on Future Computer and Communication*, 2010, pp. 792-795.

[61] Huang W. and Jing Z., "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognition Letters,* vol. 28, pp. 493-500, 2007.

[62] James A. P. and Dasarathy B. V., "Medical image fusion: A survey of the state of the art," *Information Fusion,* vol. 19, pp. 4-19, 2014.

[63] Jiang Q., Jin X., Lee S.-J., and Yao S., "A novel multi-focus image fusion method based on stationary wavelet transform and local features of fuzzy sets," *IEEE Access,* vol. 5, pp. 20286-20302, 2017.

[64] Jiang W., Yang X., Wu W., Liu K., Ahmad A., Sangaiah A. K., and Jeon G., "Medical images fusion by using weighted least squares filter and sparse representation," *Computers & Electrical Engineering,* vol. 67, pp. 252-266, 2018.

[65] Jin X., Zhou D., Yao S., Nie R., Yu C., and Ding T., "Remote sensing image fusion method in CIELab color space using nonsubsampled shearlet transform and pulse coupled neural networks," *Journal of Applied Remote Sensing,* vol. 10, p. 025023, 2016.

[66] Jin X., Jiang Q., Yao S., Zhou D., Nie R., Hai J., and He K., "A survey of infrared and visual image fusion methods," *Infrared Physics & Technology,* vol. 85, pp. 478-501, 2017.

[67] Johnson J. L. and Padgett M. L., "PCNN models and applications," *IEEE Transactions on Neural Networks,* vol. 10, pp. 480-498, 1999.

[68] Johnson K. A. and Becker J. A., "The Whole Brain Atlas," 1995.

[69] Kamble P. R., Keskar A. G., and Bhurchandi K. M., "A deep learning ball tracking system in soccer videos," *Opto-Electronics Review,* vol. 27, pp. 58-69, 2018.

[70] Kavitha C., Chellamuthu C., and Rajesh R., "Medical image fusion using combined discrete wavelet and ripplet transforms," *Procedia Engineering,* vol. 38, pp. 813-820, 2012.

[71] Khaleghi B., Khamis A., Karray F. O., and Razavi S. N., "Multisensor data fusion: A review of the state-of-the-art," *Information Fusion,* vol. 14, pp. 28-44, 2013.

[72] Kim M., Han D. K., and Ko H., "Joint patch clustering-based dictionary learning for multimodal image fusion," *Information Fusion,* vol. 27, pp. 198-214, 2015.

[73] Kinser J. M., "Pulse-coupled image fusion," *Optical Engineering,* vol. 36, pp. 737-742, 1997.

[74] Kong W. and Lei Y., "Technique for image fusion between gray-scale visual light and infrared images based on NSST and improved RF," *Optik,* vol. 124, pp. 6423-6431, 2013.

[75] Kong W. and Liu J., "Technique for image fusion based on NSST domain improved fast non-classical RF," *Infrared Physics & Technology,* vol. 61, pp. 27-36, 2013.

[76] Kong W. W., Lei Y. J., Lei Y., and Lu S., "Image fusion technique based on non-subsampled contourlet transform and adaptive unit-fast-linking pulse-coupled neural network," *IET Image Processing,* vol. 5, pp. 113-121, 2011.

[77] Kong W. W., "Multi-sensor image fusion based on NSST domain I^2CM," *Electronics Letters,* vol. 49, pp. 802-803, 2013.

[78] Kor S. and Tiwary U., "Feature level fusion of multimodal medical images in lifting wavelet transform domain," in *26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2004, pp. 1479-1482.

[79] Kumar B. S., "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform," *Signal, Image and Video Processing,* vol. 7, pp. 1125-1143, 2013.

[80] Lan X., Ma A. J., and Yuen P. C., "Multi-cue visual tracking using robust feature-level fusion based on joint sparse representation," in *IEEE conference on Computer Vision and Pattern Recognition*, 2014, pp. 1194-1201.

[81] Lan X., Ma A. J., Yuen P. C., and Chellappa R., "Joint sparse representation and robust feature-level fusion for multi-cue visual tracking," *IEEE Transactions on Image Processing,* vol. 24, pp. 5826-5841, 2015.

[82] Lan X., Zhang S., Yuen P. C., and Chellappa R., "Learning common and feature-specific patterns: a novel multiple-sparse-representation-based tracker," *IEEE Transactions on Image Processing,* vol. 27, pp. 2022-2037, 2018.

[83] Lang J. and Hao Z., "Novel image fusion method based on adaptive pulse coupled neural network and discrete multi-parameter fractional random transform," *Optics and Lasers in Engineering,* vol. 52, pp. 91-98, 2014.

[84] Lekadir K., Galimzianova A., Betriu A., Vila M. d. M., Igual L., Rubin D. L., Fernández E., Radeva P., and Napel S., "A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound," *IEEE Journal of Biomedical and Health Informatics,* vol. 21, pp. 48-55, 2017.

[85] Lewis J. J., Callaghan O. R. J., Nikolov S. G., Bull D. R., and Canagarajah N., "Pixel- and region-based image fusion with complex wavelets," *Information Fusion,* vol. 8, pp. 119-130, 2007.

[86] Li H., Manjunath B. S., and Mitra S. K., "Multisensor image fusion using the wavelet transform," *Graphical Models and Image Processing,* vol. 57, pp. 235-245, 1995.

[87] Li H., Jia X., and Zhang L., "Clustering based content and color adaptive tone mapping," *Computer Vision and Image Understanding,* vol. 168, pp. 37-49, 2018.

[88] Li M., Cai W., and Tan Z., "A region-based multi-sensor image fusion scheme using pulse-coupled neural network," *Pattern Recognition Letters,* vol. 27, pp. 1948-1956, 2006.

[89] Li S., Kwok J. T., and Wang Y., "Multifocus image fusion using artificial neural networks," *Pattern Recognition Letters,* vol. 23, pp. 985-997, 2002.

[90] Li S. and Yang B., "Multifocus image fusion by combining curvelet and wavelet transform," *Pattern Recognition Letters,* vol. 29, pp. 1295-1301, 2008.

[91] Li S. and Yang B., "Hybrid multiresolution method for multisensor multimodal image fusion," *IEEE Sensors Journal,* vol. 10, pp. 1519-1526, 2010.

[92] Li S., Yang B., and Hu J., "Performance comparison of different multi-resolution transforms for image fusion," *Information Fusion,* vol. 12, pp. 74-84, 2011.

[93] Li S., Yin H., and Fang L., "Group-sparse representation with dictionary learning for medical image denoising and fusion," *IEEE Transactions on Biomedical Engineering,* vol. 59, pp. 3450-3459, 2012.

[94] Li S., Kang X., and Hu J., "Image fusion with guided filtering," *IEEE Transactions on Image Processing,* vol. 22, pp. 2864-2875, 2013.

[95] Li S., Kang X., Hu J., and Yang B., "Image matting for fusion of multi-focus images in dynamic scenes," *Information Fusion,* vol. 14, pp. 147-162, 2013.

[96] Li S., Kang X., Fang L., Hu J., and Yin H., "Pixel-level image fusion: A survey of the state of the art," *Information Fusion,* vol. 33, pp. 100-112, 2017.

[97] Li T. and Wang Y., "Biological image fusion using a NSCT based variable-weight method," *Information Fusion,* vol. 12, pp. 85-92, 2011.

[98] Li W. and Zhu X.-f., "A new algorithm of multi-modality medical image fusion based on pulse-coupled neural networks," in *Advances in Natural Computation*: Springer, 2005, pp. 995-1001.

[99] Li W. and Zhu X.-F., "A new image fusion algorithm based on wavelet packet analysis and PCNN," in *IEEE International Conference on Machine Learning and Cybernetics*, 2005, pp. 5297-5301.

[100] Li X., Zhao L., Wei L., Yang M.-H., Wu F., Zhuang Y., Ling H., and Wang J., "Deepsaliency: Multi-task deep neural network model for salient object detection," *IEEE Transactions on Image Processing,* vol. 25, pp. 3919-3930, 2016.

[101] Li Z., Zheng J., Zhu Z., Yao W., and Wu S., "Weighted guided image filtering," *IEEE Transactions on Image Processing,* vol. 24, pp. 120-129, 2015.

[102] Liang J., He Y., Liu D., and Zeng X., "Image fusion using higher order singular value decomposition," *IEEE Transactions on Image Processing,* vol. 21, pp. 2898-2909, 2012.

[103] Liang Z., Xu J., Zhang D., Cao Z., and Zhang L., "A hybrid l1-l0 layer decomposition model for tone mapping," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4758-4766.

[104] Litjens G., Kooi T., Bejnordi B. E., Setio A. A. A., Ciompi F., Ghafoorian M., Van Der Laak J. A., Van Ginneken B., and Sanchez C. I., "A survey on deep learning in medical image analysis," *Medical Image Analysis,* vol. 42, pp. 60-88, 2017.

[105] Liu F., Li J., and Caiyun H., "Image fusion algorithm based on simplified PCNN in nonsubsampled contourlet transform domain," *Procedia Engineering,* vol. 29, pp. 1434-1438, 2012.

[106] Liu S., Wang M., and Fang Y., "A contourlet transform based fusion algorithm for nighttime driving image," in *Fuzzy Systems and Knowledge Discovery*, Berlin, Heidelberg, 2006, pp. 491-500.

[107] Liu S., Shi M., Zhu Z., and Zhao J., "Image fusion based on complex-shearlet domain with guided filtering," *Multidimensional Systems and Signal Processing,* vol. 28, pp. 207-224, 2017.

[108] Liu X., Mei W., and Du H., "Multimodality medical image fusion algorithm based on gradient minimization smoothing filter and pulse coupled neural network," *Biomedical Signal Processing and Control,* vol. 30, pp. 140-148, 2016.

[109] Liu Y., Liu S., and Wang Z., "A general framework for image fusion based on multi-scale transform and sparse representation," *Information Fusion,* vol. 24, pp. 147-164, 2015.

[110] Liu Y., Chen X., Ward R. K., and Wang Z. J., "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters,* vol. 23, pp. 1882-1886, 2016.

[111] Liu Y., Chen X., Cheng J., and Peng H., "A medical image fusion method based on convolutional neural networks," in *20th IEEE International Conference on Information Fusion*, 2017, pp. 1-7.

[112] Liu Y., Chen X., Peng H., and Wang Z., "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion,* vol. 36, pp. 191-207, 2017.

[113] Liu Y., Chen X., Wang Z., Wang Z. J., Ward R. K., and Wang X., "Deep learning for pixel-level image fusion: Recent advances and future prospects," *Information Fusion,* vol. 42, pp. 158-173, 2018.

[114] Liu Y., Chen X., Ward R. K., and Wang Z. J., "Medical image fusion via convolutional sparsity based morphological component analysis," *IEEE Signal Processing Letters,* vol. 26, pp. 485-489, 2019.

[115] Liu Z., Yin H., Chai Y., and Yang S. X., "A novel approach for multimodal medical image fusion," *Expert Systems with Applications,* vol. 41, pp. 7425-7435, 2014.

[116] Liu Z., Feng Y., Chen H., and Jiao L., "A fusion algorithm for infrared and visible based on guided filtering and phase congruency in NSST domain," *Optics and Lasers in Engineering,* vol. 97, pp. 71-77, 2017.

[117] Ma K., Li H., Yong H., Wang Z., Meng D., and Zhang L., "Robust multi-exposure image fusion: a structural patch decomposition approach," *IEEE Transactions on Image Processing,* vol. 26, pp. 2519-2532, 2017.

[118] Majumder S. and Pratihar D. K., "Multi-sensors data fusion through fuzzy clustering and predictive tools," *Expert Systems with Applications,* vol. 107, pp. 165-172, 2018.

[119] Mallat S. G. and Zhang Z., "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing,* vol. 41, pp. 3397-3415, 1993.

[120] Manchanda M. and Sharma R., "A novel method of multimodal medical image fusion using fuzzy transform," *Journal of Visual Communication and Image Representation,* vol. 40, Part A, pp. 197-217, 2016.

[121] Miao Q.-g., Shi C., Xu P.-f., Yang M., and Shi Y.-b., "A novel algorithm of image fusion using shearlets," *Optics Communications,* vol. 284, pp. 1540-1547, 2011.

[122] Miao Q. and Wang B., "The contourlet transform for image fusion," in *Defense and Security Symposium*, 2006, pp. 62420Z-62420Z-8.

[123] Mingjing L. and Yubing D., "Image fusion algorithm based on contrast pyramid and application," in *International Conference on Mechatronic Sciences, Electric Engineering and Computer*, 2013, pp. 1342-1345.

[124] Mitianoudis N. and Stathaki T., "Pixel-based and region-based image fusion schemes using ICA bases," *Information Fusion,* vol. 8, pp. 131-142, 2007.

[125] Mitianoudis N. and Stathaki T., "Optimal contrast correction for ICA-based fusion of multimodal images," *IEEE Sensors Journal,* vol. 8, pp. 2016-2026, 2008.

[126] Mohsen H., El-Dahshan E.-S. A., El-Horbaty E.-S. M., and Salem A.-B. M., "Classification using deep learning neural networks for brain tumors," *Future Computing and Informatics Journal,* vol. 3, pp. 68-71, 2017.

[127] Monica Subashini M. and Sahoo S. K., "Pulse coupled neural networks and its applications," *Expert Systems with Applications,* vol. 41, pp. 3965-3974, 2014.

[128] Na Y., Zhao L., Yang Y., and Ren M., "Guided filter-based images fusion algorithm for CT and MRI medical images," *IET Image Processing,* vol. 12, pp. 138-148, 2017.

[129] Naidu V., "Image fusion technique using multi-resolution singular value decomposition," *Defence Science Journal,* vol. 61, pp. 479-484, 2011.

[130] Nair V. and Hinton G. E., "Rectified linear units improve restricted boltzmann machines," in *27th International Conference on Machine Learning*, 2010, pp. 807-814.

[131] Nason G. P. and Silverman B. W., "The stationary wavelet transform and some statistical applications," in *Wavelets and statistics*: Springer, 1995, pp. 281-299.

[132] Nejati M., Samavi S., and Shirani S., "Multi-focus image fusion using dictionary-based sparse representation," *Information Fusion,* vol. 25, pp. 72-84, 2015.

[133] Nencini F., Garzelli A., Baronti S., and Alparone L., "Remote sensing image fusion using the curvelet transform," *Information Fusion,* vol. 8, pp. 143-156, 2007.

[134] Nirosha Joshitha J. and Selin R. M., "Image fusion using PCA in multifeature based palmprint recognition," *International Journal of Soft Computing and Engineering,* vol. 2, pp. 226-230, 2012.

[135] Pajares G. and Manuel de la Cruz J., "A wavelet-based image fusion tutorial," *Pattern Recognition,* vol. 37, pp. 1855-1872, 2004.

[136] Palsson F., Sveinsson J. R., Ulfarsson M. O., and Benediktsson J. A., "Model-based fusion of multi- and hyperspectral images using PCA and wavelets," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, pp. 2652-2663, 2014.

[137] Paramanandham N. and Rajendiran K., "A simple and efficient image fusion algorithm based on standard deviation in wavelet domain," in *International Conference on Wireless Communications, Signal Processing and Networking*, 2016, pp. 2207-2211.

[138] Patil U. and Mudengudi U., "Image fusion using hierarchical PCA," in *IEEE International Conference on Image Information Processing*, 2011, pp. 1-6.

[139] Perazzi F., Krähenbühl P., Pritch Y., and Hornung A., "Saliency filters: Contrast based filtering for salient region detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 733-740.

[140] Plataniotis K. N. and Venetsanopoulos A. N., *Color image processing and applications*: Springer Science & Business Media, 2013.

[141] Prabhakar S. and Jain A. K., "Decision-level fusion in fingerprint verification," *Pattern Recognition,* vol. 35, pp. 861-874, 2002.

[142] Pradhan P. S., King R. L., Younan N. H., and Holcomb D. W., "Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image

fusion," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, pp. 3674-3686, 2006.

[143] Qiguang M. and Baoshu W., "A novel image fusion method using contourlet transform," in *Proceedings: IEEE International Conference on Communications, Circuits and Systems*, 2006, pp. 548-552.

[144] Qu G. H., Zhang D. L., and Yan P. F., "Medical image fusion by wavelet transform modulus maxima," *Optics Express,* vol. 9, pp. 184-190, 2001.

[145] Qu X.-B., Yan J.-W., Xiao H.-Z., and Zhu Z.-Q., "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Automatica Sinica,* vol. 34, pp. 1508-1514, 2008.

[146] Rahman M. A., Liu S., Wong C. Y., Lin S. C. F., Liu S. C., and Kwok N. M., "Multi-focal image fusion using degree of focus and fuzzy logic," *Digital Signal Processing,* vol. 60, pp. 1-19, 2017.

[147] Ramlal S. D., Sachdeva J., Ahuja C. K., and Khandelwal N., "Multimodal medical image fusion using non-subsampled shearlet transform and pulse coupled neural network incorporated with morphological gradient," *Signal, Image and Video Processing,* vol. 12, pp. 1479-1487, 2018.

[148] Restaino R., Vivone G., Mura M. D., and Chanussot J., "Fusion of multispectral and panchromatic images based on morphological operators," *IEEE Transactions on Image Processing,* vol. 25, pp. 2882-2895, 2016.

[149] Rubinstein R., Bruckstein A. M., and Elad M., "Dictionaries for sparse representation modeling," *Proceedings of the IEEE,* vol. 98, pp. 1045-1057, 2010.

[150] Saeedi J. and Faez K., "A classification and fuzzy-based approach for digital multi-focus image fusion," *Pattern Analysis and Applications,* vol. 16, pp. 365-379, 2013.

[151] Sharma M., Sharma P., Pachori R. B., and Gadre V. M., "Double density dual-tree complex wavelet transform-based features for automated screening of knee-Joint vibroarthrographic signals," in *Machine Intelligence and Signal Analysis*, 2019, pp. 279-290.

[152] Shen C., Wang D., Tang S., Cao H., and Liu J., "Hybrid image noise reduction algorithm based on genetic ant colony and PCNN," *The Visual Computer,* vol. 33, pp. 1373-1384, 2017.

[153] Shreyamsha Kumar B. K., "Image fusion based on pixel significance using cross bilateral filter," *Signal, Image and Video Processing,* vol. 9, pp. 1193-1204, 2015.

[154] Singh R. and Khare A., "Fusion of multimodal medical images using Daubechies complex wavelet transform – A multiresolution approach," *Information Fusion,* vol. 19, pp. 49-60, 2014.

[155] Srivastava R., Prakash O., and Khare A., "Local energy-based multimodal medical image fusion in curvelet domain," *IET Computer Vision,* vol. 10, pp. 513-527, 2016.

[156] Starck J. L., Candes E. J., and Donoho D. L., "The curvelet transform for image denoising," *IEEE Transactions on Image Processing,* vol. 11, pp. 670-684, 2002.

[157] Sudheer A. and Bindu C. H., "Region based multi-focus image fusion using the spectral parameter variance," in *IEEE International Conference on Wireless Communications, Signal Processing and Networking* 2016, pp. 1306-1310.

[158] Sun X., Du J., Li Q., Li X., Xu L., and Li Y., "Improved energy contrast image fusion based on nonsubsampled contourlet transform," in *8th IEEE Conference on Industrial Electronics and Applications*, 2013, pp. 1610-1613.

[159] Tang L., Zhao F., and Zhao Z.-G., "The nonsubsampled contourlet transform for image fusion," in *International Conference on Wavelet Analysis and Pattern Recognition*, 2007, pp. 305-310.

[160] Tang L., Tian C., and Xu K., "IGM-based perceptual multimodal medical image fusion using free energy motivated adaptive PCNN," *International Journal of Imaging Systems and Technology,* vol. 28, pp. 99-105, 2018.

[161] Taruttis A., Rosenthal A., Kacprowicz M., Burton N. C., and Ntziachristos V., "Multiscale multispectral optoacoustic tomography by a stationary wavelet transform prior to unmixing," *IEEE Transactions on Medical Imaging,* vol. 33, pp. 1194-1202, 2014.

[162] Teng J., Wang S., Zhang J., and Wang X., "Neuro-fuzzy logic based fusion algorithm of medical images," in *3rd IEEE International Congress on Image and Signal Processing*, 2010, pp. 1552-1556.

[163] Teng J., Wang S., Zhang J., and Wang X., "Fusion algorithm of medical images based on fuzzy logic," in *Seventh IEEE International Conference on Fuzzy Systems and Knowledge Discovery*, 2010, pp. 546-550.

[164] Tian J., Chen L., Ma L., and Yu W., "Multi-focus image fusion using a bilateral gradient-based sharpness criterion," *Optics Communications,* vol. 284, pp. 80-87, 2011.

[165] Toet A., "Image fusion by a ration of low-pass pyramid," *Pattern Recognition Letters,* vol. 9, pp. 245-253, 1989.

[166] Toet A., van Ruyven J. J., and Valeton J. M., "Merging thermal and visual images by a contrast pyramid," *Optical Engineering,* vol. 28, p. 287789, 1989.

[167] Tropp J. A., Gilbert A. C., and Strauss M. J., "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing,* vol. 86, pp. 572-588, 2006.

[168] Vanmali A. V., Deshmukh S. S., and Gadre V. M., "Low complexity detail preserving multi-exposure image fusion for images with balanced exposure," in *2013 National Conference on Communications (NCC)*, 2013, pp. 1-5.

[169] Vanmali A. V. and Gadre V. M., "Visible and NIR image fusion using weight-map-guided Laplacian-Gaussian pyramid for improving scene visibility," *Sadhana,* vol. 42, pp. 1063-1082, 2017.

[170] Vishwakarma A., Bhuyan M. K., and Iwahori Y., "Non-subsampled shearlet transform-based image fusion using modified weighted saliency and local difference," *Multimedia Tools and Applications,* vol. 77, pp. 32013-32040, 2018.

[171] Wan T., Zhu C., and Qin Z., "Multifocus image fusion based on robust principal component analysis," *Pattern Recognition Letters,* vol. 34, pp. 1001-1008, 2013.

[172] Wan W., Yang Y., and Lee H. J., "Practical remote sensing image fusion method based on guided filter and improved SML in the NSST domain," *Signal, Image and Video Processing,* vol. 12, pp. 959-966, 2018.

[173] Wang G., Xu X., Jiang X., and Nie R., "A modified model of pulse coupled neural networks with adaptive parameters and its application on image fusion," *ICIC Express Letters,* vol. 6, pp. 2523-2530, 2015.

[174] Wang H.-h., "A new multiwavelet-based approach to image fusion," *Journal of Mathematical Imaging and Vision,* vol. 21, pp. 177-192, 2004.

[175] Wang L., Li B., and Tian L.-F., "Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients," *Information Fusion,* vol. 19, pp. 20-28, 2012.

[176] Wang N., Ma Y., Zhan K., and Yuan M., "Multimodal medical image fusion framework based on simplified PCNN in nonsubsampled contourlet transform domain," *Journal of Multimedia,* vol. 8, pp. 270-276, 2013.

[177] Wang N., Ma Y., and Wang W., "DWT-based multisource image fusion using spatial frequency and simplified pulse coupled neural network," *Journal of Multimedia,* vol. 9, pp. 159-165, 2014.

[178] Wang Q., Zheng W., and Piramuthu R., "Grab: Visual saliency via novel graph model and background priors," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 535-543.

[179] Wang Z. and Bovik A. C., "A universal image quality index," *IEEE Signal Processing Letters,* vol. 9, pp. 81-84, 2002.

[180] Wang Z. and Ma Y., "Medical image fusion using m-PCNN," *Information Fusion,* vol. 9, pp. 176-185, 2008.

[181] Wang Z., Ma Y., Cheng F., and Yang L., "Review of pulse-coupled neural networks," *Image and Vision Computing,* vol. 28, pp. 5-13, 2010.

[182] Wang Z., Ma Y., and Gu J., "Multi-focus image fusion using PCNN," *Pattern Recognition,* vol. 43, pp. 2003-2016, 2010.

[183] Wu J., Su Y., Cheng Y., Shao X., Deng C., and Liu C., "Multi-sensor information fusion for remaining useful life prediction of machining tools by adaptive network based fuzzy inference system," *Applied Soft Computing,* vol. 68, pp. 13-23, 2018.

[184] Xia J.-x., Duan X.-h., and Wei S.-c., "Application of adaptive PCNN based on neighborhood to medical image fusion," *Application Research of Computers,* vol. 10, pp. 3929-3933, 2011.

[185] Xia J., Chen Y., Chen A., and Chen Y., "Medical image fusion based on sparse representation and PCNN in NSCT domain," *Computational and Mathematical Methods in Medicine,* vol. 2018, pp. 1-12, 2018.

[186] Xiao X. and Wu Z., "Image fusion based on lifting wavelet transform," in *International Symposium on Intelligence Information Processing and Trusted Computing*, 2010, pp. 659-652.

[187] Xu J., Yang L., and Wu D., "Ripplet: A new transform for image processing," *Journal of Visual Communication and Image Representation,* vol. 21, pp. 627-639, 2010.

[188] Xu X., Shan D., Wang G., and Jiang X., "Multimodal medical image fusion using PCNN optimized by the QPSO algorithm," *Applied Soft Computing,* vol. 46, pp. 588-595, 2016.

[189] Xu X., Wang Y., and Chen S., "Medical image fusion using discrete fractional wavelet transform," *Biomedical Signal Processing and Control,* vol. 27, pp. 103-111, 2016.

[190] Xydeas C. S. and Petrovic V., "Objective image fusion performance measure," *Electronics Letters,* vol. 36, pp. 308-309, 2000.

[191] Yang B. and Li S., "Multifocus image fusion and restoration with sparse representation," *IEEE Transactions on Instrumentation and Measurement,* vol. 59, pp. 884-892, 2010.

[192] Yang B. and Li S., "Pixel-level image fusion with simultaneous orthogonal matching pursuit," *Information Fusion,* vol. 13, pp. 10-19, 2012.

[193] Yang B. and Li S., "Visual attention guided image fusion with sparse representation," *Optik,* vol. 125, pp. 4881-4888, 2014.

[194] Yang C., Liu P., Yin G., Jiang H., and Li X., "Defect detection in magnetic tile images based on stationary wavelet transform," *NDT & E International,* vol. 83, pp. 78-87, 2016.

[195] Yang L., Guo B., and Ni W., "Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform," *Neurocomputing,* vol. 72, pp. 203-211, 2008.

[196] Yang S., Wang M., Lu Y., Qi W., and Jiao L., "Fusion of multiparametric SAR images based on SW-nonsubsampled contourlet and PCNN," *Signal Processing,* vol. 89, pp. 2596-2608, 2009.

[197] Yang S., Wang M., Jiao L., Wu R., and Wang Z., "Image fusion based on a new contourlet packet," *Information Fusion,* vol. 11, pp. 78-84, 2010.

[198] Yang Y., Park D. S., Huang S., and Rao N., "Medical image fusion via an effective wavelet-based approach," *EURASIP Journal on Advances in Signal Processing,* vol. 2010, p. 579341, 2010.

[199] Yang Y., Tong S., Huang S., and Lin P., "Log-Gabor energy based multimodal medical image fusion in NSCT domain," *Computational and Mathematical Methods in Medicine,* vol. 2014, pp. 1-12, 2014.

[200] Yang Y., Tong S., Huang S., and Lin P., "Multifocus image fusion based on NSCT and focused area detection," *IEEE Sensors Journal,* vol. 15, pp. 2824-2838, 2015.

[201] Yang Y., Que Y., Huang S., and Lin P., "Multimodal sensor medical image fusion based on type-2 fuzzy logic in NSCT domain," *IEEE Sensors Journal,* vol. 16, pp. 3735-3745, 2016.

[202] Yang Y., Wan W., Huang S., Yuan F., Yang S., and Que Y., "Remote sensing image fusion based on adaptive IHS and multiscale guided filter," *IEEE Access,* vol. 4, pp. 4573-4582, 2016.

[203] Yang Y., Que Y., Huang S.-Y., and Lin P., "Technique for multi-focus image fusion based on fuzzy-adaptive pulse-coupled neural network," *Signal, Image and Video Processing,* vol. 11, pp. 439-446, 2017.

[204] Yang Y., Que Y., Huang S., and Lin P., "Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion," *IEEE Transactions on Instrumentation and Measurement,* vol. 66, pp. 691-703, 2017.

[205] Yang Y., Wu J., Huang S., Fang Y., Lin P., and Que Y., "Multimodal medical image fusion based on fuzzy discrimination with structural patch decomposition," *IEEE Journal of Biomedical and Health Informatics,* vol. 23, pp. 1647-1660, 2018.

[206] Yao P., Li J., Ye X., Zhuang Z., and Li B., "Iris recognition algorithm using modified log-gabor filters," in *18th International Conference on Pattern Recognition, 2006.* . vol. 4, 2006, pp. 461-464.

[207] Yin H. and Li S., "Multimodal image fusion with joint sparsity model," *Optical Engineering,* vol. 50, pp. 067007-067007-10, 2011.

[208] Yin H., Li Y., Chai Y., Liu Z., and Zhu Z., "A novel sparse-representation-based multi-focus image fusion approach," *Neurocomputing,* vol. 216, pp. 216-229, 2016.

[209] Yin M., Liu X., Liu Y., and Chen X., "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain," *IEEE Transactions on Instrumentation and Measurement,* vol. 68, pp. 49 - 64, 2019.

[210] Yin; M., Liu; X., Liu; Y., and Chen; X., "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain," *IEEE Transactions on Instrumentation and Measurement* vol. 68, pp. 49 - 64, 2019.

[211] Yu L. F., Zu D. L., Wang W. D., and Bao S. L., "Multi-modality medical image fusion based on wavelet analysis and quality evaluation," *Journal of Systems Engineering and Electronics,* vol. 12, 2001.

[212] Yu N., Qiu T., Bi F., and Wang A., "Image features extraction and fusion based on joint sparse representation," *IEEE Journal of Selected Topics in Signal Processing,* vol. 5, pp. 1074-1082, 2011.

[213] Zaidi H., Montandon M.-L., and Alavi A., "The clinical role of fusion imaging using PET, CT, and MR imaging," *Magnetic Resonance Imaging Clinics of North America,* vol. 18, pp. 133-149, 2010.

[214] Zhang D. and Nishimura T. H., "Pulse coupled neural network based anisotropic diffusion method for 1/f noise reduction," *Mathematical and Computer Modelling,* vol. 52, pp. 2085-2096, 2010.

[215] Zhang Z.-l., Sun S.-h., and Zheng F.-c., "Image fusion based on median filters and SOFM neural networks: a three-step scheme," *Signal Processing,* vol. 81, pp. 1325-1330, 2001.

[216] Zhao W. and Lu H., "Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation," *IEEE Transactions on Instrumentation and Measurement,* vol. 66, pp. 2283-2294, 2017.

[217] Zhenhua L., Zhongliang J., Hong W., and Shaoyuan S., "Multisensor image fusion algorithm using nonseparable wavelet frame transform," *Journal of Systems Engineering and Electronics,* vol. 16, pp. 728-732, 2005.

[218] Zhu J., Jin W., Li L., Han Z., and Wang X., "Multiscale infrared and visible image fusion using gradient domain guided image filtering," *Infrared Physics & Technology,* vol. 89, pp. 8-19, 2018.

[219] Zong J.-j. and Qiu T.-s., "Medical image fusion based on sparse representation of classified image patches," *Biomedical Signal Processing and Control,* vol. 34, pp. 195-205, 2017.