Design and Analysis of Robust Image Registration Schemes using Machine Learning Algorithms

A Thesis

submitted in fulfillment of the requirements
for the award of degree of

Doctor of Philosophy

by

Paluck Arora

(Roll No: 902003008)

Under the Guidance of

Dr. Rajesh Mehta

(Assistant Professor)

Dr. Rohit Ahuja

(Assistant Professor)

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY

PATIALA – 147004, INDIA

October 2024

Contents

List of Figures

List of Tables

Certificate

I hereby certify that the work which is presented in this thesis entitled " **Design and Analysis of Robust Image Registration Schemes using Machine Learning Algorithms**", in fulfilment of requirement for the award of the degree of "**Doctor of Philosophy**" submitted in Department of Computer Science and Engineering Thapar Institute of Engineering & Technology, Patiala, is an authentic record of my work carried out under the supervision of **Dr. Rajesh Mehta** and **Dr. Rohit Ahuja** refers other research works which are duly listed in the reference section.

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

(Paluck Arora)

Regn. No. 902003008

This is to certify that the above statement made by the candidate is correct and true to the best of my knowledge.

Dr. Rajesh Mehta
(Assistant Professor)

Dr. Rohit Ahuja
(Assistant Professor)

Department of Computer Science and Engineering
Thapar Institute of Engineering and Technology, Patiala

Acknowledgment

I am profoundly thankful for the guidance and support of many individuals who have been instrumental in the completion of this Ph.D. thesis. Their invaluable contributions have greatly influenced my research and significantly enriched my academic journey.

First and foremost, I extend my heartfelt gratitude to my supervisor, Dr. Rajesh Mehta, whose unwavering dedication, insightful guidance, and immense expertise have been instrumental in shaping the trajectory of my research. I am thankful to my co-supervisor, Dr. Rohit Ahuja, for his consistent encouragement, constructive feedback, and invaluable suggestions. Their collaborative efforts and vast knowledge have expanded the scope of my research and enhanced the quality of my work.

My sincere thanks to Dr. Shalini Batra, Professor and Head, DCSE, Thapar Institute of Engineering and Technology, Patiala, for providing me with the necessary administrative assistance and infrastructure that helped me in the completion of my research work. I extend my sincere appreciation to the members of my doctoral committee, Dr. Ashutosh Aggarwal, Assistant Professor, DCSE, Dr. Jyoti Maggu, Assistant Professor, DCSE, and Dr. Saurabh Bhardwaj, Professor, Department of Electrical and Instrumentation Engineering, Thapar Institute of Engineering and Technology, Patiala, for their thoughtful insights, critical feedback, and rigorous examination of my thesis.

On a personal note, I owe an immense debt of gratitude to my brother, Aggam, and his wife, Sahar, for their unwavering encouragement and support throughout this journey. Their belief in my abilities and constant motivation has been a source of strength, especially during challenging times.

I am also profoundly grateful to my brother, Ajit Dang, for his crucial assistance in securing the essential reports from doctors that were vital to my Ph.D. and thesis. His help was indispensable, and I am deeply appreciative of his efforts.

My heartfelt thanks to my best friend cum sister, Dr. Shruti Arora, who has been a guiding light throughout this process. Her timely advice, drawn from her own experience in completing her Ph.D., has been immensely helpful.

I would also like to express my sincere appreciation to my brother, Anmol, who generously helped me obtain essential reports that significantly contributed to my Ph.D. work. His support was crucial, and I am truly grateful.

My journey would not have been possible without the unwavering love and support of my parents, Sh. Tej Pal Arora and Smt. Usha Arora. The boundless encouragement, sacrifices, and belief in my potential have been the driving force behind my academic pursuits. I am also grateful to my sister, Priyanshi, for her belief and willingness to lend an ear, even during her busy schedule, which has been a source of strength.

Last but not least, I bow my head in front of God for showering His grace on me, without which none of this would have come to life.

Patiala (PALUCK ARORA)

October 2024

List of Abbreviations

MI	Mutual Information
PSO	Particle swarm optimization
GA	Genetic Algorithm
TLBO	Teaching Learning Based Optimization
SURF	Speeded Up Robust Feature
SIFT	Scale Invariant Feature Transform
ORB	Oriented Fast and Rotated Brief
RANSAC	Random Sample Consensus
CT	Computed Tomography
MRI	Magnetic Resonance Imaging
PET	Positron Emission Tomography
SPECT	Single-Photon Emission Computed Tomography
KPCA	Kernel Principal Component Analysis
PCA	Principal Component Analysis
VGG19	Visual Geometry Group
UEO	United Equilibrium Optimizer
U-Net	U-shaped convolution neural network
KNN	K-nearest neighbours
MI	Mutual Information
NCC	Normalized Cross-correlation
SSIM	Structural Similarity Index Measure
PSNR	Peak Signal-to-Noise Ratio
CC	Cross-Correlation
IGS	Image Guided Surgery
MIR	Medical Image Registration
ROI	Region Of Interest
CNN	Convolution Neural Network
SSD	Sum of Squared Differences
GWO	Grey Wolf Optimizer
BBO	Biogeography Based Optimization
NMI	Normalized Mutual Information
PROSAC	Progressive Sample Consensus
PI-SURF	Progressive SURF
E-SIFT	Entropy SIFT
SRC	Spatial Relation Correspondence
ATSRC	Alpha Trimmed Spatial Relation Correspondence
CRKEM-SIFT	Clustered Redundant Keypoints Elimination Method-Scale Invariant Feature Transform
KAZE	K-adaptive Scale-Invariant Feature Transform
AKAZE	Accelerated-KAZE
BRISK	Binary Robust Invariant Scalable Keypoints
3D	Three Dimension
2D	Two Dimension
SS-CSA	Salp Swarm-Crow Search Algorithm
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform

HOGD	Histogram of Oriented Gradient Distance
DDGWO	Data-Driven Grey Wolf Optimizer
LS-SVM	Least Square Support Vector Machines
SR	Spatial Reorganization
CR	Channel Refinement
DIR	Deformable Image Registration
CBCT	Cone-Beam CT
DCIGN	Deep Convolutional Inverse Graphics Network
UDIR-Net	Unsupervised Deformable Image Registration Network
JD	Jacobian determinant
ADMIR	Affine and Deformable Medical Image Registration
GANs	Generative Adversarial Networks
SCA	Sine Cosine Algorithm
ES	Evolution Strategies
ACO	Ant Colony Optimization
DE-DGCO	Dynamic Group Based Cooperative Optimization
LNMI	Logarithmic Normalized Mutual Information
MNMI	Modified Normalized Mutual Information
FIRE	Fundus Image Registration Dataset
ARO	Artificial Rabbits Optimization
SymReg-GAN	Symmetric Image Registration
RMSE	Root Mean Square Error
MSE	Mean Square Error
CC	Cross-Correlation
ATLAS	Anatomical Tracing Lesion
HPSO	Hybrid Particle swarm optimization
SRR	Single Rigid Registration
MRR	Multiple Rigid Registration
OASIS	Oasis Access Series of Imaging Studies
DS	Dice Score
HSD	Hausdorff Distance
Vxnet	Voxelmorph CNN network
UEO	United Equilibrium Optimizer
EO	Equilibrium Optimizer
IXI	Information Extraction from Images
TPS	Thin Plate Spline
RCV-Net	Robustly Configured View Network
ATSRC	Alpha Trimmed Spatial Relation Correspondence
MLESAC	Maximum Likelihood Estimation Sample Consensus
HM	Homography Matrix
RIRE	Retrospective Image Registration Evaluation
PCB	Phase Congruency Based
GBM	Gradient Based Metrics

ABSTRACT

Medical image registration plays a vital role in image-guided interventions to improve clinical decision-making, better visualization, and quantification of anatomical structures. Image registration is the process of transforming different sets of data into one coordinate system. Designing a universal framework for image registration is a challenging task due to the diversity of medical images and various forms of degradation that can occur during acquisition. A registration algorithm, rigid or deformable, aligns images within a common coordinate system to enhance their accuracy for diagnosis, monitoring, and treatment of medical conditions. This precise alignment facilitates clearer analysis and improved clinical decision-making. This research seeks to develop schemes that can effectively capture wide range of image registration scenarios by incorporating both rigid and deformable(non-rigid) registration techniques. Rigid registration is optimized using meta-heuristic methods to adjust rigid transformation parameters while deep learning is employed for feature extraction. In the deformable approach, a two-channel image from the moving and fixed images is processed through the U-Net model to generate a displacement field, while B-spline parameters predict and refine the warped image.

The primary objective of this research is to develop a novel approach for rigid and non-rigid medical image registration applicable to both monomodal and multimodal modalities, utilizing machine learning and meta-heuristic optimization techniques. This study includes a thorough review of existing methods and algorithms for medical image registration, aiming to identify and address current gaps and limitations in the field.

Firstly, area and feature-based methods are designed for monomodal and multimodal medical image modalities. The proposed approach employs teaching learning-based optimization (TLBO) to optimize rigid transformation parameters such as rotation and translation by considering mutual information (MI) maximization as an objective function. To address challenges in rigid transformation due to anatomical variations, a novel algorithm integrating both area-based technique such as TLBO and leveraging feature-based algorithm such as SURF are developed. This method combines the SURF framework for feature extraction with the RANSAC algorithm for feature refinement, followed by projective transformation for precise image registration.

Afterwards, a robust registered image is generated by estimating the rigid transformation parameters based on inliers, thereby enhancing the registration accuracy.

Secondly, a novel approach with hybridization of KPCA and TLBO for both monomodal and multimodal medical image registration is discussed. This approach starts with pre-processing such as noise removal and normalization to enhance image quality. Afterwards, contour extraction is carried out using Otsu's thresholding for multimodal images and a thresholding segmentation approach for monomodal images to create feature images and to determine ground truth translation parameters. Then, KPCA captures non-linear data mapping for rotation value determination, while TLBO optimizes the registration process by iteratively adjusting transformation parameters.

Next, a deformable medical image registration leveraging an unsupervised learning algorithm and deep learning framework visual geometry group (VGG19) with TLBO and unified equilibrium optimizer (UEO) is proposed. The Vxnet (U-Net) model processes two-channel images to generate non-linear displacement vector field, while TLBO optimizes the registration process by iteratively adjusting transformation parameters. However, this deformable registration focuses on monomodal medical images. Afterwards, a novel image registration scheme for both monomodal and multimodal modalities leveraging UEO in contrast to TLBO optimization techniques in combination with VGG-19 is discussed. Then, features are extracted using VGG-19 from both reference and UEO-registered image. Dynamic inlier selection is applied to enhance the matching accuracy by effectively filtering correct matches (inliers) while minimizing outliers. Finally, thin plate spline interpolation is employed to determine the optimal transformation matrix for accurate alignment between reference and UEO-registered image.

Lastly, an integration of the feature extraction technique SURF with VGG19 for anatomical and functional medical images is presented. The proposed scheme explores various feature extraction techniques namely SIFT, SURF, and ORB algorithms for detecting features. Then feature matching is performed using a brute force matcher, followed by outlier removal and projective transformation to generate the registered image. Despite the focus on feature extraction techniques, a novel image registration algorithm is introduced, combining deep learning with SURF for accurate image registration. The scheme enhances image quality through preprocessing and feature extraction using VGG19, followed by Euclidean distance calculations and dynamic inlier selection for refined alignment. Thin plate spline interpolation corrects misalignments, while the SURF algorithm and alpha-trimmed spatial correspondence minimize outliers. Finally, a homography matrix ensures robust alignment between sensed and VGG19-registered images.

To validate the effectiveness of the proposed methodologies, experimental evaluations are conducted on publicly available and in-house clinical datasets for both monomodal and multimodal medical images. The results demonstrate the superiority and robustness of proposed techniques in terms of similarity metrics as compared to state-of-the-art techniques.

Chapter 1
Introduction

This chapter provides a thorough overview of medical image registration, detailing its classification, stages, categorization, methods, components, and applications. It also presents the research objectives and emphasizes key contributions that address significant challenges in medical image registration, as well as identifying opportunities for improving existing methods.

1.1. BACKGROUND

Computer vision systems are designed to mimic the working of human visual systems. These systems play a vital role in several domains, such as medical imaging, satellite imaging, remote sensing, and surveillance applications [5]–[8]. Many of these applications depend on object recognition, object tracking, image retrieval, and classification. In the field of image processing, numerous applications require precise image alignment. In recent years, the alignment of images has become a significant concern, particularly in the realm of multi-sensor image registration which garners significant attraction from researchers around the globe. Image registration is the process of aligning two images of the same subject taken at different times, from different viewpoints, or by different sensors [9]. The goal is to overlay these images to create a composite image that combines the information from all the input images [10]. This process is essential in various fields such as medical imaging, remote sensing, etc. where it helps in comparing images, detecting changes, and combining information from different sources [11].

The challenges encountered in medical image registration lead to the misalignment of images. These challenges include difficulties in aligning images from different sensors due to the following factors:

- To determine the optimal geometric transformation parameter value.
- Achieving suitable correspondence among features remains challenging due to variations in spatial resolution across images from different modalities.

- To establish the association between intensity values of corresponding pixels is complex.
- The task of identifying and rejecting outliers presents a significant challenge in medical image registration, particularly in the context of image-guided surgery where numerous outliers are prevalent.
- Integrating contrasting information from various medical image types poses a significant challenge in multimodal image registration.

Accurate image alignment plays a vital role in medical applications including image fusion, multimodal segmentation, and image-guided surgery. However, above discussed challenges need to be confronted. The goal of image registration is to establish optimal correspondence between two images and determine a geometric transformation that aligns the reference and sensed images. Prior to exploring various approaches to medical image registration, certain terminologies are considered in this thesis:

Reference (Fixed) image (I_R): The image which remains unchanged and serves as a reference for the transformation process.

Moving (sensed) image (I_S): The image undergoes geometric transformation to align with the reference image.

Transformation function: It is employed to adjust the fixed image towards alignment with the sensed image.

Registration entails identifying the spatial transformation that correlates points in the target image with points in the source image. Moreover, by representing the images as two-dimensional arrays of a specific size, denoted as I_R and I_S, where $I_R(i,j)$ and $I_S(i,j)$ denote their respective intensity values at coordinates (i,j), mapping between images can be expressed as in Equation (1.1):

$$I_S(i,j) = I_R\big(f(i.j)\big) \tag{1.1}$$

where $f: I_R(i,j) \rightarrow I_S(i,j)$ performs spatial coordinate transformation between two images. The core of any registration problem typically involves determining the spatial or geometric transformation. This transformation can be approximated using various types of transformation functions such as rigid, affine, and projective transformation, etc.

1.2. Classifications

Medical image registration techniques are based on four fundamental criteria, each of which can be further divided into levels, as illustrated in Figure 1.1.

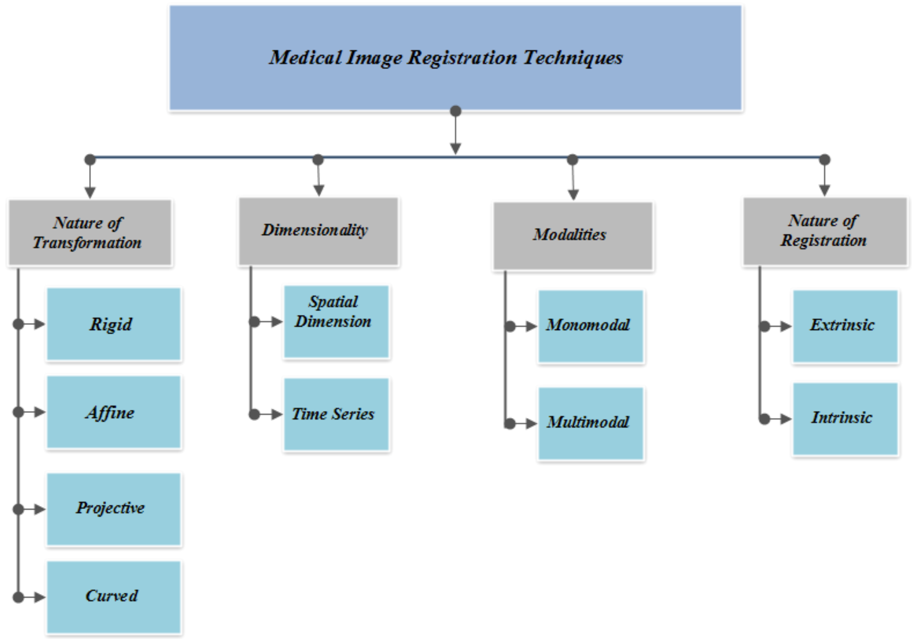


Figure 1.1: Classification of the Image Registration technique

- *Nature of the Transformation:* The nature of transformation can be categorized into rigid, affine, projective, and curved. A rigid transformation involves only translations and rotations, while an affine transformation includes scaling and shearing in addition to translations and rotations. Affine transformations preserve parallelism between lines, mapping straight lines to straight lines [12]. In contrast, projective transformations do not necessarily preserve parallelism. Elastic transformations, referred to as curved or non-rigid transformations, enable the mapping of straight lines to curves [13].
- *Dimensionality:* It encompasses spatial dimensions and temporal series, allowing for different types of image registration, such as 2D-2D, 2D-3D, and 3D-3D registrations [14]. Integrating both monomodal and multimodal 2D [15] and 3D images [16] obtained at various time points during preoperative and intraoperative procedures significantly enhances treatment delivery accuracy and precision.
- *Modalities:* Registration based on modalities involves aligning medical images acquired from different imaging modalities, including monomodal (intra-subject) and multimodal (inter-subject) approaches. Monomodal registration aligns images obtained from the same modality such as CT to CT or MRI to MRI, while

multimodal registration aligns images acquired from different modalities, such as PET to CT or MRI to CT scans[17].

- *Nature of Registration:* Registration can be categorized into extrinsic and intrinsic methods. Extrinsic utilizes external objects attached to the patient, ideal for image-guided surgery[18]. Intrinsic relies on the patient anatomy/image content only, extracting features directly from the image, offering flexibility and non-invasiveness in medical image processing.

1.3. Stages of Image Registration

Image registration is an elementary process involved in several applications to align images either from different viewpoints or captured at varying times to ensure accurate comparison and analysis. The image registration framework is illustrated in Figure 1.2 and consists of several key stages:

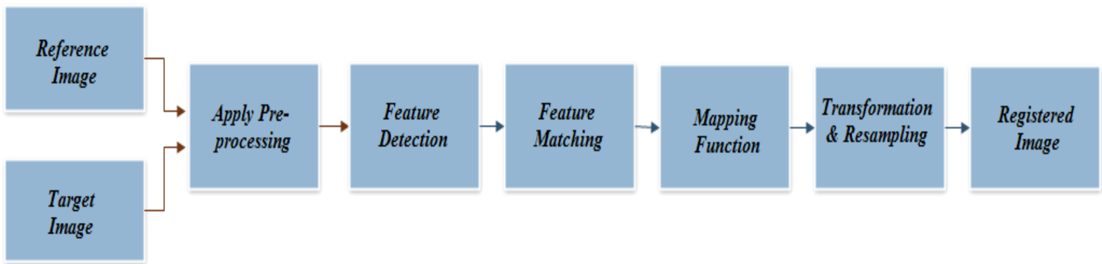


Figure 1.2: Generic Framework of image registration

- *Pre-processing:* In this step, pre-processing involves noise reduction, contrast enhancement, artifact removal, and intensity normalization to improve alignment[19]. These operations enhance image quality and standardize intensity levels across images, leading to more accurate registration outcomes.
- *Feature Detection*: From the pre-processed reference and target images, various features such as boundaries, edges, contours, and corners are selected and identified. These features create point representations known as feature points or interest points, which are used to establish correspondence between the images. Different feature detection methods namely scale-invariant feature transform (SIFT)[20], speeded-up robust features (SURF), oriented fast and rotated brief (ORB)**,** Harris corner detection, etc. [21] are employed to extract distinctive features from reference and target images.

- *Feature Matching:* During the feature matching step, mapping or correspondence is established between the detected features of the reference and target images. Various feature-matching algorithms, such as brute-force matcher [22] and K-nearest neighbours (KNN) [23], are employed to achieve accurate image matching.
- *Mapping Function*: The parameters of the mapping functions are derived from established feature correspondence in the previous step, enabling alignment between reference and target images [24]. These functions aim to determine optimal transformation parameters while minimizing or maximizing the similarity metric.
- *Transformation & Resampling*: Lastly, the target image is geometrically transformed based on the parameters estimated in the previous steps, aligning it with the reference image[25]. Resampling is used to interpolate pixel values and adjust for spatial distortions, ensuring that the registered image accurately reflects the spatial alignment between the images.

1.4. Categorization of Image Modalities

Image modalities employed for registering medical images can be broadly classified into structural and functional magnetic resonance imaging, as illustrated in Figure 1.3.

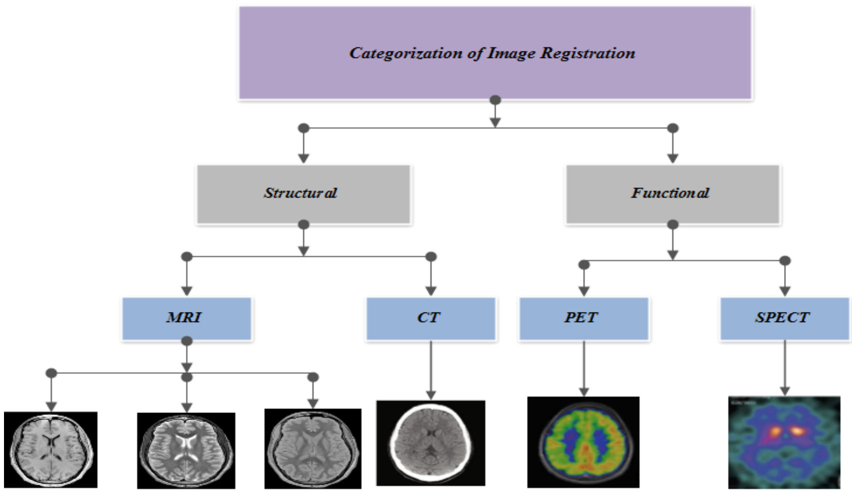


Figure 1.3: Categorization of image modalities

- *Structural imaging*: This technique provides detailed visualizations of anatomical structures such as organs, tissues, bones, and vessels within the body. Common modalities for structural images including CT and MRI [3].

CT offers detailed insights into hard tissue, while MRI provides comprehensive structural details of soft tissue[26].

- *Functional imaging*: It examines physiological processes and interactions within the body, offering real-time insights into organ and tissue function. Modalities such as SPECT and PET fall under functional imaging, offering valuable information about tissue functionality while lacking detailed structural and locational information [17], [27] Integrating functional and anatomical data is essential in image-guided surgery, assisting surgeons in diagnosis and treatment planning.

1.5. Image Registration Methods

Image registration methods are divided into four categories as shown in Figure 1.4.

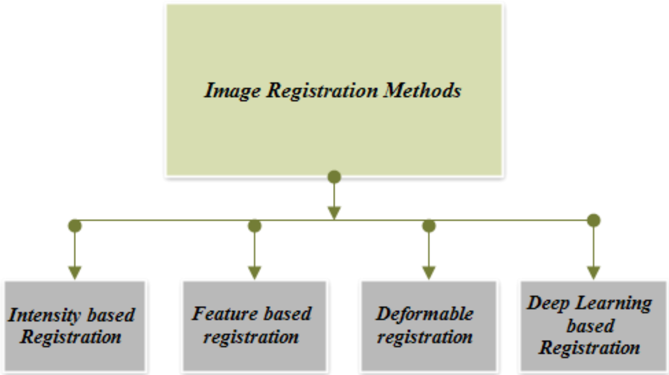


Figure 1.4: Image registration methods

- *Intensity-based method*: This approach utilizes image intensity values, prioritizing feature matching and optimization steps rather than distinct feature detection. Several area-based(intensity-based) methods are correlation methods, Fourier methods, mutual information, and optimization methods [28].

- *Feature-based method*: This method initially extracts salient features such as edges and corners, then matches these features to establish global correspondence and estimates geometric transformations, relying on accurate detection of image features. Various feature-based methods, including Harris Corner Detection, ORB, SIFT, and SURF, enhance registration robustness and accuracy by identifying essential features within the images[29].
- *Deformable registration*: Deformable registration is a non-rigid registration approach for aligning images with non-linear deformations or significant

anatomical differences. Unlike conventional rigid or affine transformations, deformable registration employs unsupervised learning techniques, with U-Net architecture being a prominent choice for learning complex deformation fields directly from input images [30].

- *Deep learning-based registration*: Deep learning techniques namely VGG16 and VGG19 etc., significantly improve deformable registration by extracting high-level features from input images, enhancing precision and robustness. Their flexibility allows adaptation to various imaging modalities, contributing to widespread applicability in medical tasks, and ultimately improving patient care [31].

1.6. Components of Image Registration

Image registration consists of the following components including similarity measure with an objective function, geometric (spatial) transformation with an interpolator, and optimization to optimize the similarity measure as illustrated in Figure 1.5.

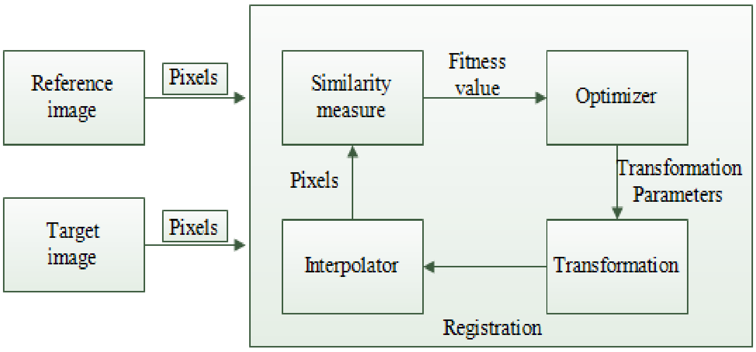


Figure 1.5: Components of image registration

- *Similarity Measure with an Objective Function*: Similarity measures determine image alignment quality and serve as an objective function in registration. They are divided into intensity-based measures (e.g., MI, NCC, SSIM)[17] that compare pixel values and feature-based measures that assess similarities between image features like edges and corners [24].
- *Geometric (Spatial) Transformation:* Geometric transformations involve modifying the spatial configuration of the image to achieve alignment with a reference image. These transformations can be rigid (involving only rotation, translation, and scaling) [7] or non-rigid (allowing for more complex deformations such as warping or bending) [32].

- *Optimization*: It is the process of finding the optimal transformation parameters that maximize the similarity measure. It involves iterative adjustments and refinements to minimize or maximize the objective function to ensure the best possible alignment between images. Methods include search-based (iterative parameter exploration) [33] and evolutionary (genetic algorithms) approaches [34].
- *Interpolation:* It is used to estimate pixel values in the transformed image, especially in areas where the transformation has created gaps or where pixel values need to be recalculated. It ensures that the final image is smooth and free of artifacts[32].

1.7. Applications

Medical image registration assists neurosurgeons and medical practitioners in addressing the difficulties of precisely identifying brain tumor and accurate diagnosis of disease in preoperative and intraoperative images, especially when relying solely on visual inspection. Applications for medical image registration include:

- *Radiation Therapy:* It aims to treat tumour while minimizing damage to surrounding tissues. Advanced medical image registration methods have enhanced radiotherapy by enabling precise patient positioning, treatment planning, and assessment. Both rigid and deformable registration techniques are employed, rigid registration is suitable for anatomically stable cases and deformable registration is necessary for cases with significant anatomical changes due to tumor growth or other factors, facilitating automatic segmentation, dose accumulation, mathematical modeling, and functional imaging in radiotherapy[35].
- *Cancer Detection:* Medical image registration is essential for cancer detection and therapy monitoring, enabling early detection by aligning segmented areas across different imaging modalities. It facilitates locating cancerous tissue, especially challenging due to low contrast in MR and CT images, and enhances accuracy through image registration, benefiting the diagnosis of cancers[36].
- *Template Atlas application:* It represents medical images to accommodate variations in organ shapes and sizes, commonly achieved through atlases[28]. Registration is crucial in constructing atlases, and aligning images to address variability, particularly in template-based medical image analysis for standardized

evaluation and interpretation in a reference atlas space, utilizing inter-subject registration in computational anatomy.

- *Image-guided intervention:* In image-guided surgery, registration aligns preoperative images with intraoperative patient data, enabling accurate localization of regions of interest (ROI). It ensures the integration of images from various time frames or scanners, enhancing image quality and surgical precision[5].

1.8. Research Motivation

To understand the complexities of image registration, a thorough investigation is conducted within the specialized domain of medical image registration. There is extensive literature available on various applications, such as medical imaging used for diagnosing and treating diseases affecting vital organs like the brain, heart, lungs, liver, and kidney, etc. Additionally, in computer vision, efforts are made to replicate human vision using advanced technologies to understand and interpret surroundings, which presents complex challenges. In medical image registration, significant challenges, i.e., optimizing geometric transformation parameters, establishing correspondence between features, aligning intensity values among corresponding pixels, and handling outliers, especially in contexts like image-guided surgery. The motivation of this research is to investigate area and feature-based methods for image registration across medical modalities, ultimately developing a robust model to effectively address registration challenges using various transformation techniques and signal/image processing methods. Unlike traditional rigid or affine transformations, deformable registration allows for a more flexible alignment of structures, accommodating complex spatial transformation. To further improve the accuracy and efficiency of image registration in medical imaging, especially for grayscale and colored images, advanced deep unsupervised medical image registration techniques like U-shaped convolution neural network (U-Net), alongside deep learning methods such as convolution neural network (CNN) are employed. These methods will be applied to both publicly available and in-house clinical datasets, facilitating comprehensive evaluation and validation.

1.9. Research Objective

The following research objectives are outlined:

1. To study and investigate area and feature-based methods of image registration for different modalities of images in the medical domain.
2. To develop a model, that addresses the challenges for image registration using transformation (Rigid/Non-Rigid, Affine, projective, etc.) or signal/image processing methods.
3. To enhance the degree of similarity between source and target image using machine learning as well as optimization methods.
4. Testing and validation of the developed schemes for image registration using standard (existing) and real-world datasets.

1.10. Research Contributions

The research contributions to fulfill the objectives in the domain of medical image registration include:

1. The research utilizes area-based methods to address anatomical variations and spatial transformations in medical image registration.
2. Enhancement of registration accuracy in complex anatomical scenarios through the incorporation of feature-based methods.
3. A novel approach that integrates kernel principal component analysis with teaching learning-based optimization to address the challenge of randomly initialized transformation parameters.
4. Integration of deformable unsupervised technique based on U-Net model with optimization methods and affine transformation to enhance the robustness of image registration.
5. A comparative analysis of various feature extraction methods, including Harris corner detection, SIFT, ORB, and SURF, along with alpha-trimmed spatial relation correspondence (ATSRC), has been conducted. Additionally, deep learning techniques like VGG16 and VGG19 have been integrated with SURF-based feature extraction to improve accuracy and robustness in medical image registration.

1.11. Thesis Organization

The thesis consists of a total of 7 chapters. The remaining sections of the thesis have been structured to offer in-depth information on the previously discussed ideas.

Figure 1.6: Chapter-wise organization of the thesis

Chapter 1 introduces the fundamentals of image registration and classification, as well as the various stages involved in the image registration process. This chapter also covers the key components and applications of image registration, outlines the research motivations and objectives, and highlights the contributions made by the research.

Chapter 2 presents a thorough literature review of rigid, non-rigid (deformable), and machine learning-based methods for monomodal and multimodal medical image registration. Additionally, it outlines the identified research gaps and discusses the quality measures relevant to medical image registration.

Chapter 3 extensively explores area-based methods for monomodal medical images, utilizing teaching learning-based optimization (TLBO) for optimizing rigid transformation parameters and maximizing similarity metrics such as MI. It addresses challenges posed by anatomical variations and spatial transformations. Subsequently, it introduces a novel image registration algorithm that integrates area and feature-based approaches to address these challenges in monomodal and multimodal medical images. Furthermore, incorporating the speeded-up robust features framework with the random sample consensus algorithm enhances feature extraction, and projective transformation further enhances the accuracy of the registered image, validated using publicly available datasets.

Chapter 4 addresses the challenge of randomly initialized transformation parameters, expanding its scope to include both clinical and publicly available datasets. It introduces a method that integrates kernel principal component analysis (KPCA) with TLBO, specifically designed to handle challenges associated with complex anatomical variations and spatial transformations. This approach encompasses three phases, including pre-processing, contour extraction followed by feature extraction through KPCA, and assessment of robust affine transformation parameters optimized by TLBO for accurate alignment.

Chapter 5 seeks to enhance the accuracy and efficiency of medical image registration for grayscale and colored images by integrating deformable unsupervised techniques, such as the U-Net model, with optimization methods and affine transformation. This objective encompasses work with various datasets, including 2-D and 3-D monomodal medical MR and CT images, to achieve robust and accurate registration.

Chapter 6 compares different feature extraction techniques, including Harris Corner Detection, SIFT, ORB, and SURF, ultimately determining that SURF offers superior performance compared to other traditional methods. Additionally, it addresses the challenge of improving image registration accuracy and robustness in medical imaging by integrating deep learning techniques like VGG16 and VGG19 with SURF-based feature extraction, resulting in more accurate and reliable registration outcomes for medical imaging applications.

Finally, the summary of findings and major contributions of the thesis are highlighted in Chapter 7. In this chapter, the scope of the work in future directions based on the studies conducted thus far is also discussed.

Chapter 2
Literature Review

This chapter provides a comprehensive literature review on image registration within the contexts of satellite, remote sensing, and medical imaging. It gathers detailed insights into current research and methodologies in this field. Intensity and feature-based rigid registration approaches are covered in Section 2.1. Leveraging deep learning approaches on intensity and feature-based registration is explored in Section 2.2. Non-rigid medical image registration is focussed in Section 2.3. Image registration using hybrid approaches is discussed in Section 2.4. Research gaps are presented in Section 2.5. An in-depth explanation of quality measures is employed in Section 2.6.

2.1. Approaches to Rigid Registration: Intensity and Feature-Based Registration

In this section, an extensive evaluation of medical image registration corresponding to rigid medical image registration using feature extraction and optimization techniques is described. Several researchers have discussed rigid registration methods with an emphasis on monomodal and multimodal image modalities, but existing approaches fail to capture features to precisely align the images.

a) Intensity-based registration

In this section, area-based or intensity-based techniques rely on pixel intensity values to align images, commonly using similarity metrics like MI and SSD. These methods are essential in medical imaging, remote sensing, and industrial inspection, leveraging optimization algorithms such as gradient descent and particle swarm optimization (PSO), etc.

Dida *et al* [37] illustrated the comparative study between grey wolf optimizer (GWO) and PSO for registering human brain magnetic resonance imaging (MRI) and computed tomography (CT) images revealed that GWO achieved superior performance in terms of robustness and precision. However, GWO is limited in its clinical dataset applicability and requires further validation.

Lin *et al*. [38] introduced an image registration algorithm based on an enhanced ant colony optimization (ACO) model to determine the rigid transformation parameter. This scheme lacks in providing the optimal geometric transformation parameter and its applicability is constrained to anatomical medical images. Moreover, this scheme addresses these limitations and enhances the algorithm's effectiveness and versatility across various image types and registration scenarios.

Abdel *et al.* [39] developed a hybrid approach for medical image registration by combining modified MI metrics with PSO. The modified MI incorporates both statistical and spatial image information, enhancing registration accuracy. This method is validated on diverse publicly available medical image datasets but encounters challenge due to its computational complexity, and sensitivity to parameters.

Ayatollahi *et al.* [40] employed PSO to maximize modified normalized mutual information (MNMI), ensuring robust results and preventing local optima traps. Additionally, a hybrid PSO algorithm, integrating subpopulation and crossover from genetic algorithm (GA)[41], enhances precision and efficiency. It offers smoother MNMI curves, improving accuracy and reducing computation time. However, it faces challenges with significant shear distortion, impacting accuracy.

Lu *et al.* [42] presented a reliable image registration method that leverages local features and geometric invariants, with progressive sample consensus (PROSAC) improving accuracy by eliminating unreliable matches. This approach enhances the number of reliable key points and reduces mismatch rates compared to SIFT and traditional ORB methods, ensuring precise image registration even in complex scenarios. However, challenges remain in optimizing efficiency for higher-resolution images and parallel processing.

Charif *et al*. [34] demonstrated an improved method for medical image registration using normalized mutual information and a biogeography-based optimization (BBO) algorithm, achieving high-precision and robust alignment of PET and MR images. This method outperforms the PSO algorithm in both accuracy and speed, effectively handling mono-modal and multi-modal registrations. However, challenges remain in optimizing computational efficiency for real-time applications and high-resolution images.

Wu *et al*. [43] demonstrated a multimodal continuous ant colony optimization (ACO) algorithm for multi-sensor remote sensing image registration, addressing the challenge of large differences between imaging sensors[44]. By incorporating an efficient local search operation, this method mitigates the risk of getting trapped in local optima, characteristic of complex multimodal problems.

Cocianu *et al*. [45] introduced a hybrid approach integrating auto-adaptive Evolution Strategies (ES) and the swift convergence ability of APSO[46]. This research focuses on evolving algorithms for image registration in signature recognition, presenting two firefly method variants for accuracy with moderate run times and two APSO variants for faster computation [47].

Ting *et al*. [48] developed a multi-resolution search optimization method, merging QPSO and Powell techniques to overcome Powell's limitations. It is applied to 2D MRI image registration and enhances precision and speed by addressing local extremum challenges in mutual information functions. Leveraging wavelet transform, the approach combines QPSO and Powell for robust registration, with applications extending to multi-mode and three-dimensional imaging.

From the above literature, it is observed that intensity-based techniques face challenges in significant computational resources and getting stuck in local minima. To address these issues, we employed feature-based techniques, which enhance efficiency and robustness by utilizing invariant features.

b) Feature-based registration

In this section, various feature extraction algorithms, including Harris corner detection, SIFT, SURF, and ORB, to detect features such as edges and corners are employed.

Zheng *et al*. [49] introduced progressive SURF (PI-SURF), a coarse-to-fine medical image registration method utilizing the SURF algorithm for enhanced accuracy. PI-SURF outperforms traditional methods like mutual information, and Demons in both unimodal and multimodal registration, significantly improving metrics such as MI, NMI, NCC, and MSD. Despite its effectiveness, PI-SURF faces challenges with clinical datasets due to variability in image quality, the presence of noise and artifacts, and significant anatomical differences.

Vishwakarma *et al*. [50] developed a hybrid approach for accurate remote sensing satellite image registration, crucial for various fields like remote sensing, medical imaging, and computer vision. This method incorporates a SURF detector for feature extraction, a random sample consensus (RANSAC) algorithm for outlier rejection, and a projective transformation function for image alignment. The assessment focuses on CC and RMSE of control points, demonstrating high accuracy.

Gu *et al.* [51] illustrated a medical image registration method that integrates the SURF algorithm for feature extraction and an improved R-RANSAC algorithm for eliminating erroneous matches. By efficiently extracting and matching feature points, this algorithm achieves robustness, speed, and accuracy in image registration, outperforming traditional methods in handling noise, intensity variations, and initial misalignments.

Swathi *et al.* [52] suggested an image registration scheme employing an optimization-driven weighted average model. This scheme utilizes entropy-based SIFT feature (E-SIFT) features extracted and matched using a point matching algorithm, followed by alignment of images through a similarity transformation model. The registration accuracy is further improved by a weighted average fusion model optimized by the Salp swarm-crow search algorithm (SS-CSA) algorithm.

Nejad *et al*. [20] presented a technique for MRI image registration focused on preserving the total number of initial matches while maximizing precision. The clustered redundant key points elimination method-scale invariant feature transform (CRKEM-SIFT) algorithm is introduced to enhance feature extraction, followed by the spatial relations correspondence (SRC) and alpha trimmed spatial relation correspondence (ATSRC) methods to improve correspondences based on spatial relationships. However, this algorithm faces challenges such as extended execution times and numerous inaccuracies in correspondences[53].

Nejad *et al.* [54] developed two adaptive RANSAC schemes for optimal threshold selection, one based on the mean distance between points and the other incorporating both mean and variance of distances, applied to the SIFT algorithm. This method enhances feature-matching accuracy and consequently enhances image registration performance compared to traditional RANSAC. However, it may face challenges with high variability in feature distribution and struggle with real-time processing due to increased computational complexity.

Tareen *et al.* [21] conducted an in-depth comparison of various feature detection and description algorithms, including SIFT, SURF, K-adaptive scale-invariant feature transform (KAZE), accelerated-KAZE (AKAZE), ORB, and binary robust invariant scalable key points (BRISK). It offers valuable insights into algorithm selection for vision-based applications, identifying SIFT and BRISK as the most accurate. However, while ORB and BRISK detect numerous features quickly their matching time can increase the overall duration of image registration.

Hyun *et al*. [55] presented a genetic programming (GP) approach to enhance feature-based image registration by effectively identifying features and eliminating outliers. This method initially employs SIFT to extract key points, followed by matching descriptors based on Euclidean distance. This approach improves registration accuracy by effectively eliminating outliers. However, the method may be computationally intensive due to the GP-based classification process.

Hamzehei *et al.* [56] developed a scheme for a three-dimensional (3D) image registration method for microscopy and medical imaging employing SIFT and ResNet50 for feature extraction along with MLESAC for outlier detection. This method demonstrates high precision, robustness, and flexibility across various imaging modalities. However, it struggles with severe distortions such as significant scaling, rotation, and non-rigid transformations, which can impact registration accuracy.

Although several researchers have made significant contributions to rigid image registration, challenges remain in determining optimal rigid transformation parameters, dealing with low-resolution images, finding sufficient matched points, and effectively rejecting outliers.

2.2. Leveraging Deep Learning Approaches on Intensity and Feature-Based Registration

In this section, various deep learning methods are specifically designed for image registration, utilizing architectures like Alexnet, ResNet, and VGG16. These algorithms have been adapted to directly learn and predict spatial transformations, enabling accurate image alignment by capturing complex patterns and features.

Gadde *et al.* [57] presented a method that combines SIFT in the spatial domain with either discrete cosine transform (DCT) or discrete wavelet transform (DWT) in the frequency domain, enhancing robust feature extraction and parameter estimation through nonlinear mappings[58]. Further, this approach refines neural network configurations and extensive testing across different scenarios to enhance the robustness and applicability of image registration.

Yan *et al.* [33] suggested a multimodal image registration approach using a histogram of oriented gradient distance (HOGD) as the similarity measure and a data-driven grey wolf optimizer (DDGWO) as the optimization algorithm[59]. This method was validated on real and four simulated image pairs, demonstrating high registration accuracy. Its key advantage lies in an efficient computational strategy, incorporating a trained SVM model, and the ability to avoid local optima. However, this approach faces challenges such as reliance on HOGD effectiveness and limitations in directly addressing scale differences.

Peng *et al.* [60] employed least square support vector machines (LS-SVM) to determine a mapping function based on control point coordinates. This approach demonstrates effectiveness in addressing geometric deformations and rectifying errors arising from inaccuracies in control point positions[61].

Liu *et al.* [62] introduced the registration via the Fourier transform with spatial reorganization and channel refinement network (RegFSC-Net) model, which utilizes Fourier Transform along with spatial reorganization (SR) and channel refinement (CR) to reduce computational costs and enhance performance. While effective, the model encounters difficulties when applied to low-quality images or cases with significant anatomical deformations.

Goswami *et al.* [63] developed a support vector machine (SVM) based approach for image registration, specifically effective for images with significant deformations, noise, and texture[61]. Moreover, this method accurately predicts relative transformations, outperforming existing techniques across various real-world datasets. Furthermore, it demonstrates robustness and efficiency in identifying image registrations, particularly in detecting homogeneous areas.

Wang *et al*. [64] addressed the challenge of registering heterogeneous optical and SAR images, essential for enhancing ground remote sensing observation capabilities through fusion applications. This approach leverages optical and SAR data, extracting correlation from limited information for precise and efficient automatic registration, particularly beneficial for high-resolution imagery.

Yang *et al*. [65] presented a feature-based image registration method leveraging a convolutional neural network (CNN) for robust feature extraction from remote sensing images(satellite and UAV)[66]. Additionally, leveraging a pre-trained VGG network generates robust multi-scale feature descriptors for enhancing both feature extraction and localization capabilities. This method addressed challenges like outlier detection and unevenly distributed feature points.

Zhao *et al.* [67] showcased an image registration approach that leverages both SIFT and ResNet34 features to improve the accuracy of matching features in multisource high-resolution remote sensing images[68]. This approach enhances registration accuracy by integrating ResNet features pre-trained on ImageNet with traditional SIFT features.

Kavitha *et al*. [69] utilized Alexnet, particularly emphasizing the fc6 and fc7 outputs for feature extraction from identified SIFT keypoints, to construct a CNN for image registration. However, identifying suitable distance measures and matching methods to achieve optimal results in image registration is a challenging task.

Ye *et al*. [70] presented an automated registration method by integrating SIFT with CNN features to extract both low and high-level features for remote sensing imaging[71]. The scheme is evaluated across different image modalities and employs various similarity metrics to improve alignment accuracy.

Elwan *et al*. [72] demonstrated the VGG-16 model for feature detection in SAR image registration, integrating intensity smoothing and scaling. Additionally, this scheme detects and extracts robust key points, strengthened by integrating the RANSAC approach, which improves outlier removal and registration accuracy[73].

Wang *et al.* [74] explored the CNN architecture, specifically VGGNet-16, for depth feature extraction to improve image registration accuracy from brain images, leading to significant improvements in similarity metrics[75]. This method is validated using different similarity metrics like mutual information (MI) and cross-correlation coefficient (CC), with a specific focus on monomodal medical images[68].

2.3. Non-rigid (Deformable) Image Registration

In this section, non-rigid image registration techniques focussed on aligning images with complex deformations by allowing local transformations. Unlike rigid registration, these methods account for anatomical or structural variations, making them particularly useful in applications like medical imaging, where flexible alignment of tissues or organs is essential for accurate analysis.

Kearney *et al.* [76] introduced a deep unsupervised learning method for deformable image registration (DIR) in cone-beam CT (CBCT) to CT imaging[77]. The deep convolutional inverse graphics network (DCIGN) model employs hierarchical feature learning and a sparse DIR algorithm[78]. However, it requires significant computational resources and needs optimization for handling anatomical complexities beyond head and neck regions.

Wang *et al.* [79] presented a method for registering mono-modal and multi-modal images using a differentiable mutual information metric with CNN to generate an accurate deformation field[80]. However, addressing the complexities of intensity relationships in multi-modal images remains challenging, requiring continuous refinement to improve accuracy.

Balakrishnan *et al*. [81] suggested an unsupervised learning-based method for deformable medical image registration, which employs T1-weighted brain MRI images[82]. However, challenges remain in fine-tuning parameters and adapting to diverse clinical datasets due to anatomical complexity and imaging variability.

Bhattacharjee *et al.* [83] employed a CNN-based unsupervised learning framework for multimodal registration of abdominal CT images, which involve significant anatomical variability and organ deformations[84]. In this method, high anatomical variability and significant organ movements require modification to address these specific issues.

Ma *et al.* [85] introduced an unsupervised deformable image registration network (UDIR-Net), which computes the flow field between moving and fixed images and generates warped images accordingly[82]. This technique was validated using a 3D MRI brain dataset suitable for mono-modality images.

Ho *et al.* [26] presented an unsupervised lung registration network (LRN) employing cycle consistency training without requiring ground-truth results, aimed at aligning lung CT images. However, this technique is limited to a single anatomical structure. Although this method shows potential for broader clinical applications, including the registration of other organs, it requires large training datasets and significant computational resources.

Zhu *et al*. [86] demonstrated a deformable medical image registration method integrating differential geometry and voxel morph CNN. To minimize dissimilarity between two images (fixed and warped) in terms of dice score and Jacobian determinant (JD), aiming to generate the deformation field. It is crucial to optimize scalability and computational efficiency, expand the method's applicability to different medical images, improve atlas construction, integrate it into clinical workflows, and refine algorithms to further enhance overall performance.

Tang *et al*. [87] introduced an affine and deformable medical image registration (ADMIR) method focusing on learning displacement vector fields. This method comprises an affine registration module, a deformation registration module for learning displacement vector fields, and a spatial transformer to obtain the warped image[88]. They conducted their evaluation by assessing results on similarity metrics such as the dice score and Hausdorff distance, with a specific focus on brain images[89].

Mahapatra *et al.* [90] developed deep learning methods for multimodal medical image registration using generative adversarial networks (GANs). By incorporating conditional and cyclic constraints in the GAN cost function, this method achieves more accurate, realistic, and smooth registration than other deep learning-based approaches.

Chen *et al*. [91] demonstrated a hybrid Transformer-ConvNet model for volumetric medical image registration. Moreover, they introduced diffeomorphic and Bayesian variants for topology-preserving deformations and uncertainty estimation. It is validated on 3D medical images across various applications[92].

Saygili *et al*. [93] presented random forest regressors with block-matching features to predict registration errors through independent directions, including horizontal, vertical, and diagonal axes. It provides accurate error predictions and useful feedback for improving registration quality. However, its effectiveness can be limited by significant image distortion or low-quality inputs[94].

2.4. Image Registration using Hybrid Techniques

In this section, hybrid techniques that integrate deep learning approaches with meta-heuristic algorithms or combine deformable models with meta-heuristic methods. It also explores different approaches, such as intensity-based and feature-based techniques, to improve alignment accuracy and robustness. By capitalizing on the strengths of these diverse methods, hybrid approaches provide superior performance, especially in complex tasks like multimodal image registration.

Lineta *et al*. [95] presented the double exponential-dynamic group-based cooperative optimization algorithm (DE-DGCO) to update weight and transformation parameters[96]. They validate their approach with a multi-similarity model utilizing Logarithmic Normalized Mutual Information (LNMI), Modified Normalized Mutual Information (MNMI), and Tversky's feature contrast model to assess the similarity of images[97].

Luo *et al*. [98] elaborated RL-mARO, amalgamating reinforcement learning with enhanced artificial rabbits optimization (ARO) for medical image registration [99]. They validate their approach using normalized mutual information (NMI) similarity metrics on various datasets, including retrospective image registration evaluation (RIRE) and fundus image registration dataset (FIRE).

Zheng *et al.* [100] presented a generative adversarial network (GAN) based approach for symmetric image registration (SymReg-GAN) for monomodal and multimodal images[94]. It addresses the limitations of traditional techniques by incorporating inverse-consistent geometric transformations and leveraging a semi-supervised strategy to utilize both labelled and unlabelled data.

Mahapatra *et al.* [101] presented a deep learning framework for medical image registration using unsupervised domain adaptation and generative adversarial networks (GANs)[102]. This approach transforms input image pairs into a latent feature space using autoencoders, enabling successful registration across different image types[103], [104].

Table 2.1: Review of image registration in different fields

S. No.	Reference	Registration class and modality used	Dataset	Evaluation measure	Pros and Cons
1	Nejad *et al*. [20], 2023	Rigid; Feature-based; Monomodal, Multimodal; (Brain, 2D, MR-T1 and MR-T2)	RIRE	MAE	CRKEM-SIFT with SRC technique improves MRI registration accuracy but struggles with execution time.
2	Charif *et al*. [34], 2019	Rigid; Area-based Monomodal (Brain, 2D, MR-T1, MR-T2 and PET)	Whole Brain ATLAS	MI, SSIM	The BBO-based method enhances precision and image registration but requires careful parameter tuning and struggles with noisy images.
3	Dida *et al* [37], 2020	Rigid; Area-based Monomodal (Brain, CT, 2D, MRI)	Real 2D Brain	MI	GWO shows superior performance but needs clinical validation and refinement for broader medical applications.
4	Lin *et al.* [38], 2016	Rigid; Area-based Monomodal (Brain, Knee, 2D, MRI)	National Taiwan University	SSD and CC	It accurately aligns images but lacks in optimizing geometric transformation parameters.
5	Abdel *et al*. [39], 2017	Rigid; Area and Feature-based Monomodal (Brain, 2D, CT, MRI)	Harvard ATLAS	MI, CC, and SSD	It enhances registration accuracy by integrating statistical and spatial information but is computationally complex.
6	Ayatollahi *et al.* [40], 2012	Rigid; Area-based Monomodal (Brain, 2D, CT, MR-T2 and MR-PD)	Whole Brain ATLAS	MI, NMI	By employing PSO to maximize modified NMI and avoid local optima, it struggles with accuracy under large shear distortions.
7	Lu *et al.* [42], 2016	Rigid; Area-based Monomodal	Harvard ATLAS Dataset	PSNR, MI	It achieves high precision and robust feature matching for complex image registration but

		(Brain, 2D, CT and MR-T2)			is time-consuming and sensitive to parameter tuning.
8	Wu *et al*. [43], 2019	Rigid Multimodal (Optical and SAR image, 2D).	Sardinia, island town, and Shuguang village dataset	MI, RMSE	ACO algorithm handles diverse sensors and avoids local optima but needs deep learning enhancement.
9	Ting *et al*. [48], 2016	Rigid Monomodal (CT, MR-T2, 2D)	Whole Brain ATLAS	RMSE	It achieves high accuracy and speed in 2D MRI image registration but encounters challenges with time optimization and local minima.
10	Zheng *et al*. [49], 2021	Rigid; Feature-base Multimodal (Brain, 2D, MR-T1 and CT)	Brain CT Images with Intracranial Hemorrhage Masks	MI, NMI, CC, and MSD	PI-SURF metric shows improvements but struggles with clinical datasets due to anatomical differences.
11	Vishwakarma *et al*. [50], 2018	Rigid; Feature-base (Satellite images, 2D)	MANIT, Bhopal campus	RMSE and NCC	It provides high accuracy in similarity metrics but faces challenges in robustness and adaptability.
12	Gu *et al*. [51], 2014	Rigid; Feature-base Monomodal, Multimodal (Brain, 2D, MR-T1 and MR-PD)	McConnell imitating brain database from Neurosciences Institute	NMI	It integrates SURF and improved R-RANSAC to enhance robustness in medical image registration but struggles with computational complexity.
13	Swathi *et al*. [52], 2020	Rigid; Area and Feature-base (Hyperspectral images, 2D)	Pavia Centre and University	RMSE and NCC	It achieves high precision with improved RMSE and NCC scores but encounters challenges with algorithmic complexity and parameter tuning.
14	Hyun *et al*. [55], 2017	Rigid; Feature-base; Monomodal (Remote sensing images, 2D)	Prof. Stewart's group at Rensselaer	MSE	The GP-based method improves registration accuracy but face challenges due to its computational complexity.

			Polytechnic Institute		
15	Hamzehei *et al.* [56], 2023	Rigid; Feature-base Monomodal (fluorescent and medical images, 2D, 3D)	CUMC12, Columbia University Medical Center	MI, PCB, and GBM	It enhances 3D image registration accuracy with SIFT, and ResNet50 but faces limitations in computational complexity.
16	Gadde *et al.* [57], 2016	Rigid; Optimization technique; Feature-base; Monomodal; (MR-T1, 2D)	IXI	RMSE	SIFT-DCT and SIFT-DWT improve image registration with robust feature extraction but fail problems in extensive testing for broader applicability.
17	Liu *et al.* [62], 2024	Non-Rigid; Monomodal; (MR-T1, 3D)	IXI, OASIS-T1	Dice score, Jacobian matrix	RegFSC-Net enhances registration accuracy using Fourier transform but struggles with low-quality images and significant anatomical deformations.
18	Goswami *et al.* [63], 2024	Non-Rigid Monomodal (MR-T2 ,2D)	ABIDE, OASIS	PSNR, SSIM	It enhances image registration accuracy against deformations using SVM-based methods but struggles with computational demands and extensive parameter tuning.
19	Yang *et al.* [65] , 2018	Satellite and UAV images	Multi-temporal satellite and UAV image dataset	RMSD, MAD, MED, STD	It improves feature-based multi-temporal image registration with robust CNN-generated descriptors but struggles with computational efficiency and large-scale feature matching.
20	Zhao *et al.* [67], 2021	Rigid Remote sensing images, 2D	Gao Fen Satellite, HR Satellite Datasets with Disaster	RMSE	To enhance multisource high-resolution remote sensing image registration accuracy but faces challenges with computational complexity and fine-tuning across diverse datasets.

			Information		
21	Kavitha *et al*. [69], 2018	Rigid Monomodal, 2D Graf, Leuven, Trees, and Wall	Middlebury Stereo Datasets	Euclidean, City block, Cosine, Minkowski, and CC	It employs Alexnet features for robust registration but faces challenges in selecting optimal distance metrics and matching strategies for diverse image deformations.
22	Ye *et al*. [70], 2018	Rigid Monomodal, 2D Remote sensing images	Landsat Thematic Mapper	RMSE	It improves remote sensing image registration by combining CNN features with SIFT but is limited by computational complexity.
23	Elwan *et al*. [72], 2022	Rigid Monomodal, 2D SAR images	Space Shuttle Imaging Radar, Terra-SAR-X	MSE, PSNR, and CC	It enhances SAR image registration robustness with the VGG-16 model but faces challenges with computational complexity.
24	Wang *et al.* [74], 2021	Rigid Multimodal (MR-T1, MR-T2, MR-PD, 2D)	Brain web dataset	MI, NMI, NCC, MSD	It improves CNN with higher registration accuracy but struggles with mismatched feature points.
25	Kearney *et al*. [76], 2018	Non-rigid Monomodal (MR-T1, 3D)	Synthetic CBCT dataset	NMI	CBCT-CT deformable image registration achieves high accuracy and efficiency despite noise but requires significant training time and computational resources.
26	Kearney *et al*. [79], 2021	Non-Rigid; Monomodal, Multimodal (MR-T1, MR-T2, 2D)	LPBA40, IXI	MI, SSIM, PSNR, Dice score	It effectively registers medical images using differentiable mutual information but struggles with complex intensity relationships in multi-modal images.
27	Balakrishnan *et al*. [81], 2018	Non-Rigid Monomodal (T1 weighted MRI, 2D)	ADNI, Harvard GSP, MCIC, PPMI OASIS,	Dice score	It provides accurate registration without requiring ground truth data but struggles with fine-tuning parameters.

			ABIDE, ADHD200, HABS.		
28	Bhattacharjee *et al.* [83], 2021	Non-Rigid Monomodal (Abdominal CT image, 2D)	LiTS dataset.	Dice score, Hausdorff distance	CNN method achieves accurate multimodal abdominal CT image registration but struggles with high anatomical variability and significant organ movements.
29	Ma *et al.* [85], 2022	Non-Rigid Monomodal (Abdominal CT image, 2D)	LONI Probabilistic Brain Atlas (LPBA40)	Dice score	It achieves accurate registration with feature adjustment and hierarchical loss but needs enhancement to handle varying fixed images.
30	Ho *et al.* [26], 2023	Non-Rigid Monomodal (Lung CT image,3D)	DIR-Lab dataset	Dice score, Target registration error, Jacobian determinant	Unsupervised lung CT registration provides accurate results but is limited to single anatomical structures and substantial computational resources.
31	Sheikhjafari *et al.* [86], 2019	Non- Rigid Monomodal (Brain MR-T2 MRI image, 2D).	ADNI and MRBrainS18	Dice score, Jacobian determinant	It improves MRI registration similarity metric but requires computational resources for generating differential geometry features.
32	Tang *et al.* [87], 2020	Non-Rigid Monomodal (Brain MR-T1 weighted MRI image, 3D).	ABIDE, OASIS	Dice score, Hausdorff distance, and average symmetric surface distance	It achieves accurate medical image registration without pre-alignment or supervised data but requires further refinement for different image types.
33	Mahapatra *et al.* [90], 2018	Non-rigid Multimodal (Retinal fundus and fluorescein angiography images, 2D)	Sunnybrook cardiac dataset	Dice score, Hausdorff distance, and MAD	The GAN-based method combines conditional and cyclic constraints for smooth alignment but struggles with modality robustness.
34	Chen *et al.* [91], 2022	Non-rigid Monomodal (Medical images, 3D)	IXI and OASIS dataset	Dice score, Hausdorff distance	Trans-Morph improves registration accuracy with Transformer for spatial relationships but struggles with

					resource constraints and high computational demands.
35	Saygili *et al.* [93], 2021	Non rigid Multimodal (Medical images)	DIR Lab, POPI, Hammers	MAE, TRE	It predicts registration errors accurately but struggles with significantly distorted or low-quality images due to block-matching reliance.
36	Luo *et al.* [98], 2024	Rigid Multimodal	RIRE, FIRE	NMI	It enhances registration accuracy and robustness but needs further validation on diverse datasets and clinical applications.
37	Zheng *et al.* [100], 2022	Non-rigid Multimodal (CT, MR-T1, MR-T2, 2D)	Alberts, LPBA40, IBSR18, CUMC12 and MGH10	Dice Similarity Coefficient	It enhances speed and accuracy in image registration using GANs but faces challenges in dissimilarity measurements and training.
38	Mahapatra *et al.* [101], 2020	Non-rigid Multimodal (Chest X-ray, retinal and brain MR images,2D)	NIH ChestXray 14 dataset	DS, Hausdorff Distance, MAD	It achieves dataset-independent medical image registration with unsupervised domain adaptation and GANs but faces GAN training challenges.
39	Gupta *et al.* [105], 2016	Rigid Monomodal (MR-T2, CT, 2D)	IXI	MI, CC, RMSE	The ACO-based method improves medical image registration but needs development for multimodal and scaled images.
40	Ahmed *et al.* [106]	Rigid Multimodal (CT, MR-T1, 2D)	Brain tumor MRI and CT scan dataset	LNMI, MNMI, and Tversky's feature contrast model	It enhances registration accuracy with advanced similarity metrics but may increase computational overhead and implementation difficulty.

2.5. Research Gaps

Based on a comprehensive analysis of the existing literature survey presented in the above sections, significant research gaps remain in the domains of image registration discussed. These approaches, categorized into rigid, and non-rigid, as well as machine learning and deep learning techniques, show significant gaps that impact overall performance and efficiency. The identified gaps are outlined below:

1. *Variability in Input Data:* Image registration becomes challenging due to significant variability in input data, often caused by distortions during the image capture process, such as scaling, rotations, and projective transformations, especially when multiple images are captured at different orientations.
2. *Geometric Transformation Parameters:* Determining the optimal geometric transformation parameters between images by maximizing similarity measures presents a significant challenge.
3. *Multimodal Feature Alignment:* Aligning features accurately is challenging due to variations in spatial resolution among images obtained from different modalities, impacting registration accuracy.
4. *Outlier Rejection:* Rejecting outliers in medical image registration is difficult, as they significantly affect the accuracy of the registration process.
5. *Intensity Relationships:* The complex relationship between intensity values of related pixels presents challenges in multimodal medical images, affecting the accurate computation of similarity measures.
6. *Contrast Mapping in Multimodal Images:* Accurately mapping contrast information in multimodal images, especially when organs are scanned with different scanners, poses a primary challenge, complicating the determination of patient location and orientation.
7. *Deep Learning Integration:* Implementing deep learning techniques to enhance the accuracy and efficiency of medical image registration algorithms, especially for complex anatomical structures is a significant research gap.
8. *Robustness and Adaptability:* There is a lack of robustness and adaptability in intensity-based registration methods and deep learning-based techniques for various image types and scenarios, limiting their generalizability and accuracy.

2.6. Quality Measures

The effectiveness of the proposed approach is evaluated across publicly available and in-house clinical datasets consisting of CT and MRI medical images as illustrated in Figure 2.1. The assessment employs various metrics, including RMSE, MSE, SSIM, PSNR, and CC to quantify image quality [107].

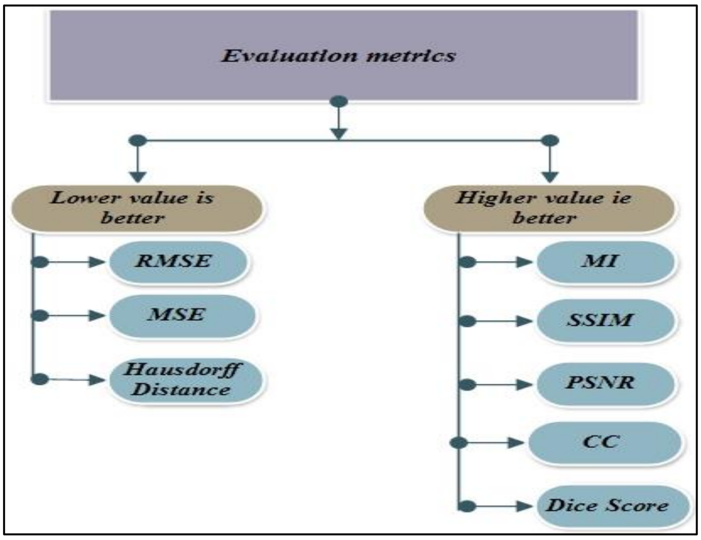


Figure 2.1: Evaluation metrics for image registration

1. *Mutual Information:* It remains unaffected by overlapping regions between two images [108]. It equals zero when images are independent and reaches its maximum value when they are properly aligned. MI between images A and B is determined by considering both their individual and joint entropy, as shown in Equation (2.1).

$$\mathrm{MI}(A, B) = H(A) + H\,(B) - H(A, B) \tag{2.1}$$

Shannon entropy quantifies the information content or uncertainty in an image. It is denoted as H(A) for the source image A and H(B) for the registered image B, as illustrated in Equations (2.2) and (2.3) respectively [109].

$$H(A) = -\textstyle\sum_a p_a(a) \log p_a(a) \tag{2.2}$$

$$H(B) = -\textstyle\sum_b p_b(b) \log p_b(b) \tag{2.3}$$

while, marginal probability distributions $p_a(a)$ and $p_b(b)$ correspond to the image intensities of A and B respectively. The joint entropy of the smoothed source and registered image is denoted as H (A, B) with joint histogram for both images A and B represented by $p_{ab}(a, b)$ as defined by Equation (2.4).

$$H(A, B) = -\textstyle\sum_a \sum_b p_{ab}(a, b) \log p_{ab}(a, b) \tag{2.4}$$

Maximizing the mutual information between two images involves identifying the complex overlapping regions by maximizing the individual entropy and minimizing the joint entropy.

2. *Root mean square error:* It serves as a measure of dissimilarity between source and registered images. A lower RMSE value signifies a more accurate registration process, indicating better alignment between the images. This

metric is calculated as the square root of mean square error, as defined by Equation (2.5).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}[M(i,j)-F(i,j)]}{M*N}} \tag{2.5}$$

where M and N represent the height and width of an image, respectively, with i and j representing the pixel locations. The registered image is denoted by $M(i,j)$ while source image is denoted by $F(i,j)$.

3. *Peak signal-to-noise ratio*: It quantifies the ratio between the optimal signal power and image noise power as expressed in decibels (dB). Higher PSNR values indicate a greater resemblance between source and registered image, as expressed in Equation (2.6).

$$PSNR = 10\log_{10}\frac{(2^b-1)^2}{MSE} \tag{2.6}$$

where, b indicates the number of bits per pixel and MSE refers to mean square error.

4. *Cross-correlation*: It determines the correlation coefficient between pixel values of source and registered image to measure the similarity. This coefficient ranges from -1 to +1, where values closer to +1 indicate a stronger similarity and lower values indicate dissimilarity between the images, as outlined in Equation (2.7).

$$\text{CC} = \frac{\sum_{x=1}^{P}\sum_{y=1}^{Q}(m(x,y)-\bar{m})(n(x,y)-\bar{n})}{\sqrt{\sum_{x=1}^{P}\sum_{y=1}^{Q}(m(x,y)-\bar{m})^2\ \sum_{x=1}^{P}\sum_{y=1}^{Q}(m(x,y)-\bar{n})^2}} \tag{2.7}$$

where, variables x and y represent the rows (x) and columns (y) of the matrices m and n respectively.

5. *Structural similarity index measure:* It evaluates the resemblance between two images by comparing them to source images of perfect quality. SSIM values range from -1 to 1, with higher values indicating a greater degree of similarity between source and registered images [33], as described in Equation (2.8).

$$\text{SSIM}=\frac{(2\bar{p}\bar{q}+k1)(2\sigma pq+k2)}{(\sigma p^2+\sigma q^2+k2)\left((\bar{p}^2)+(\bar{q})^2+k1\right)} \tag{2.8}$$

To compute the quality parameters, it is essential that both source and registered images have identical sizes. Let $p(a,b)$ and $q(a,b)$ denotes source and registered image, where b denotes the number of bits per pixel, so $(2^b - 1)$ represents the maximum gray value. The mean values for the source and register images are denoted by $\bar{p}$ and $\bar{q}$, respectively. The variance of the p and q images is represented by σp^2 and σq^2, while the covariance of p and q is denoted by σpq. Additionally, $k1$ and $k2$ are constants used in the computation.

6. *Dice score:* It quantitatively measures the accuracy of image registration by analyzing the similarity between fixed(source) and U-Net warped (registered) images, with values ranging from 0 to 1. Dice score ≈ 1 indicates a perfect overlap between the fixed and warped images, while Dice score ≈ 0 signifies no overlap, as described in Equation (2.9). It is calculated using the following formula:

$$\text{Dice Score} = \frac{2\,|W \cap F|}{|W| + |F|} \tag{2.9}$$

where W and F represent the number of pixels in the warped and fixed images, respectively. A high Dice score indicates accurate alignment between registered and fixed images, while a low score indicates poor alignment.

7. *Hausdorff distance:* It measures the maximum euclidean distance between points in subsets, like the fixed image (I_{fix}) and warped image (ω). Lower values of the Hausdorff distance indicate more accurate alignment between the images.

Let, $I_{fix} = \{f_1, f_{2,}, \ldots\ldots, f_m\}$ and $\omega = \{w_1, w_2, \ldots\ldots, w_k\}$ are two finite point sets, the hausdorff distance is computed as shown in Equation (2.10).

$$\text{HSD}(I_{fix}, \omega) = max\left(hsd(I_{fix}, \omega), hsd(\omega, I_{fix})\right) \tag{2.10}$$

where,

$$hsd(I_{fix}, \omega) = \max_{f \in I_{fix}} \min_{w \in \omega} \|f - w\|$$

$$hsd(\omega, I_{fix}) = \max_{w \in \omega} \min_{f \in I_{fix}} \|w - f\|$$

This equation computes the maximum of the minimum distances from each point in one set to the closest point in the other set, ensuring a symmetric measurement of the distance between the two-point sets.

2.7. Summary

In this chapter, we conducted a comprehensive literature review on image registration, categorizing the approaches into rigid, non-rigid, and deep learning approaches. Based on the existing literature, research gaps are identified and discussed. Additionally, various quality measures used in image registration are elaborated. In the next chapter, an area and feature-based method for monomodal and multimodal medical image registration is explored. Firstly, in Section 3.2, area-based methods for monomodal (isomodal) biomedical image registration using teaching learning-based optimization (TLBO) with MI maximization as an objective function are discussed. This section confronts the challenge of accurately determining the optimal values of rigid transformation parameters using optimization techniques in the presence of anatomical variations and complex spatial transformations. Furthermore, a novel image registration algorithm that integrates both area-based and feature-based approaches for monomodal and multimodal medical images is presented in Section 3.3. A speeded-up robust features (SURF) framework is incorporated with a random sample consensus (RANSAC) algorithm to enhance feature extraction, while projective transformation further improves registration accuracy.

Chapter 3
A TLBO-Driven Hybrid Approach for Adaptive Medical Image Registration using Affine Transformation and SURF

3.1. Introduction

The objective of this chapter is to introduce an area and feature-based method for monomodal and multimodal medical image registration. Firstly, in this section, area-based methods for monomodal (isomodal) biomedical image registration using TLBO with MI maximization as an objective function are explored. The visual quality of the registered image is assessed using MI and SSIM. This chapter confronts the challenge of accurately determining the optimal values of rigid transformation parameters using optimization techniques in the presence of anatomical variations and complex spatial transformations. Furthermore, a novel image registration algorithm that integrates both area-based and feature-based approaches for monomodal and multimodal medical images is presented. SURF framework is incorporated with the RANSAC algorithm to enhance feature extraction, while projective transformation further improves registration accuracy. Experiments conducted on publicly available datasets including Whole Brain ATLAS and Kaggle datasets confirm the effectiveness of proposed approaches across different similarity metrics.

This chapter is organized as follows. Firstly, area-based accurate alignment of monomodal modalities of medical images leveraging TLBO is described in Section 3.2. The performance of the scheme discussed in Section 3.2, is evaluated using MI and SSIM metrics with the consideration of the visual quality of registered images as presented in experimental results in Subsection 3.2.1.1. Then, in Subsection 3.2.1.2, the analysis and effectiveness of the scheme are measured by performing a fair comparison using monomodal modalities of images, with recently developed meta-heuristic optimization-based techniques. The conclusion and limitations of the scheme are discussed in Section 3.2.2. Secondly, to overcome the limitations of the scheme discussed in Section 3.2, a novel area and feature-based scheme using TLBO with affine transformations and SURF followed by projective transformation on both monomodal and multimodal medical image modalities is elaborated in Section 3.3. Experimental and comparison results with state-of-art techniques in terms of robustness are described in Subsection 3.3.1.1 and 3.3.1.2 respectively. The summary of the chapter including advantages and limitations is presented in Section 3.4.

3.2. Teaching Learning-Based Optimization-Driven Approach for Accurate Alignment of Medical Images

To achieve novel medical image registration, gaussian filtering is utilized for noise reduction. Subsequently, rigid transformation aligns the original and floating image, with TLBO optimizing the values of rigid transformation parameters (RTP) (θ, t_x, t_y). MI and SSIM serve as similarity metrics to assess the correlation between the original and registered image, with TLBO maximizing MI as an objective function [110]. The optimal transformation $(\hat{T})$, as per Equation (3.1), yields the highest value of MI, defining the optimization algorithm objective function to determine the optimum values of transformation parameters for enhancing similarity metrics.

$$\hat{T} = \arg\max_{T} MI\,[f_s(x,y), f_t(x,y)] \tag{3.1}$$

The original (reference) image is denoted by $f_s(x,y)$ where x and y represent the image coordinates, and the floating image is denoted by $f_t(x,y)$. $\hat{T}$ represents the optimum transformation, with T representing the transformation as well as its parameters. The proposed scheme framework is illustrated in Figure 3.1, with the following steps involved:

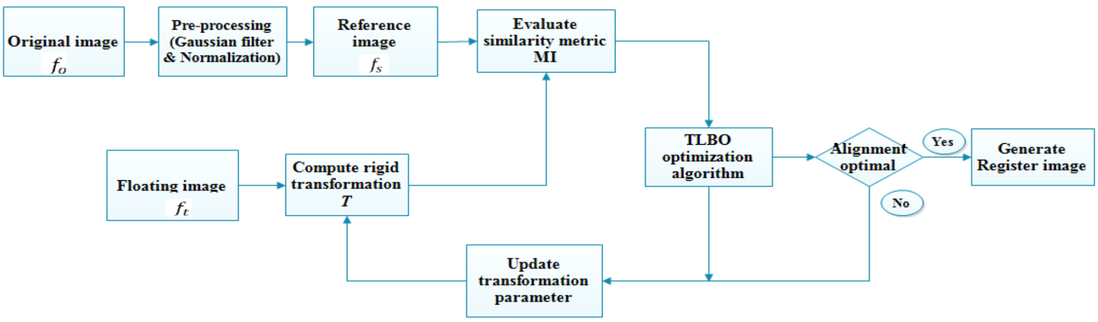
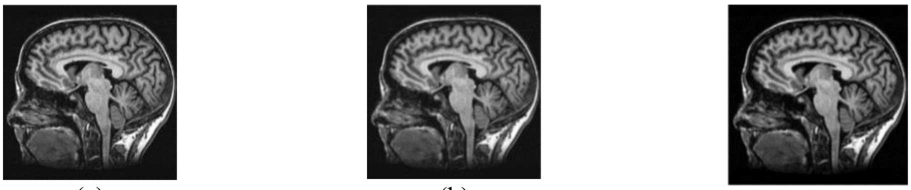


Figure 3.1: Framework of the proposed scheme

Step 1: The source (reference) image undergoes preprocessing using a Gaussian filter for denoising before normalization. Although noise removal during preprocessing results in slight degradation [111], normalization is applied to standardize the image data on a uniform scale. The output of Step 1 is illustrated in Figure 3.2.

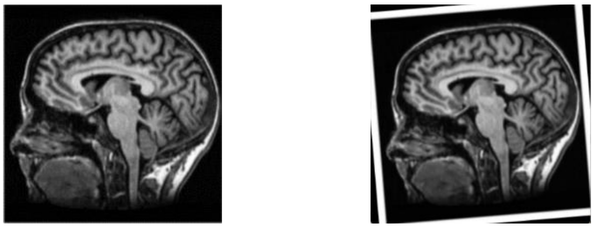


(a) (b) (c)

Figure 3.2: Brain image from ATLAS Database [112]

(a) Source Image (b) Gaussian Blur Image (c) Normalized Image

Step 2: The output of the reference image obtained in Step 1 undergoes rigid transformation, involving rotation preceded by translation, to generate the floating image depicted in Figure 3.3.

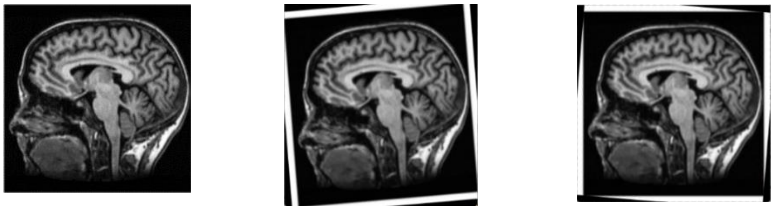


(a) (b)

Figure 3.3: (a) Reference Image (b) Floating Image

Step 3: The similarity metrics SSIM and MI are utilized to quantify the similarity between the reference and transformed floating images. TLBO is then employed to determine the optimal values of RTP, with MI maximization serving as an objective function for the similarity metric.

Step 4: The optimal rotation (θ) and translation (tx, ty) values derived from TLBO are utilized to transform the floating image, generating the registered image as depicted in Figure 3.4 Subsequently, MI computation validates the correlation between the reference and registered images, with a high MI value signifying precise alignment, as depicted in Figure 3.1. TLBO optimizes the RTP to maximize or minimize similarity metrics, ensuring image registration reliability. If the objective function validates the results, the registered image is achieved otherwise parameters are modified and the process is repeated from Step 2 as depicted in Figure 3.1. However, if the results are not accurate, the transformation parameters are updated, and the procedure is repeated from Step 2, as described in Figure 3.1.



(a) (b) (c)

Figure 3.4: (a) Reference image (b) Floating image with transformation parameters (θ, t_x, t_y) as (4.7,0,0) (c) Registered Image

3.2.1. Experimental and Comparison Results

This section evaluates the performance of the proposed scheme using MI and SSIM metrics as well as the visual quality of registered images. The effectiveness of the scheme is showcased through comparisons with state-of-the-art methods [37], [40], [113].

3.2.1.1. Experimental Results

All the experiments are conducted using the PYTHON 3.9.4 platform on a system equipped with 16.0 GB RAM and a 2.80 GHz Intel(R) Core TM i7-1165G7 CPU. The experimental results utilize various isomodal medical database images as illustrated in Figure 3.5 (a1-a6) sourced from the Whole Brain ATLAS database [112]. These medical brain images comprising CT and MRI scans with dimensions $(256 * 256)$ (Figure 3.5, a1-a5) and 354×353 (Figure 3.5, a6) are employed to assess and validate the effectiveness of the scheme. The reference (source) image undergoes a rigid transformation, involving rotation at an angle θ followed by a translation $\mathrm{T} = (tx, \mathrm{ty})$, resulting in the floating (template) image. This transformation occurs after applying preprocessing to the source image, and TLBO optimizes the transformation parameters to register the floating image with the reference image.

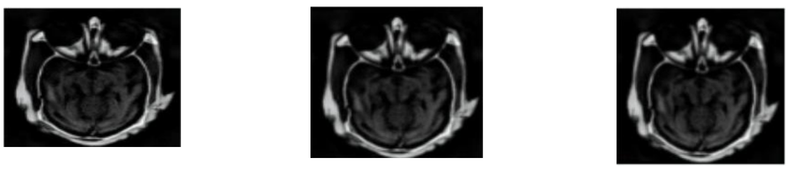
Initially, the floating image is registered with source image followed by observing the impact of rotating an isomodal brain image at a rotation $\theta = 2^{\circ}$ preceded by $\mathrm{T} = (\mathrm{tx}, \mathrm{ty}) = (2,2)$, as depicted in Figure 3.5 (b1-b4). The TLBO-derived optimal rotation and translation parameters undergo a rigid transformation, resulting in the formation of the registered image as illustrated in Figure 3.5 (Col.3, c1-c4), corresponding to the floating images (Col. 2, b1-b4). In the second test case, the same procedure is repeated with varying transformation parameters, such as rotation $\theta = 15^{\circ}$ and translation $\mathrm{T} = (tx, \mathrm{ty}) = (17,17)$. The registration result for these parameter values is displayed in Figure 3.5 (c5), corresponding to the floating image in Figure

3.5 (Col. 2, b5), along with their MI values in Table 3.1. In the third test case, brain image of dimension (354×353) is considered with different rotation and translation parameters (rotation $\theta = 4.7^{\circ}$ and translation $T = (tx, ty) = (0,0)$) The resulting registration is depicted in Figure 3.5 (Col. 3, c6), corresponding to the floating image in Figure 3.5 (Col.2, b6). MI and SSIM metric values across all isomodal images are summarized in Table 3.1.

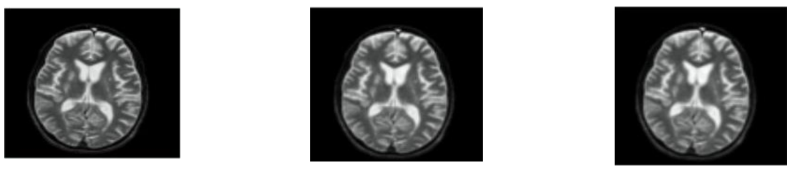
Table 3.1: Experimental results of the proposed scheme using different test images

Test images	Initial value	RTP			$(SSIM/MI)$
	(tx, ty, θ)	tx	ty	θ	
(a1)	(2,2,2)	-1.9634	-2.0316	-1.9848	(0.9643/12.087)
(a2)	(2,2,2)	-1.9636	-2.0338	-1.9866	(0.9868/7.5647)
(a3)	(2,2,2)	-1.9378	-2.0486	-2.0003	(0.9982/5.4848)
(a4)	(2,2,2)	-1.9716	-2.0326	-1.9939	(0.9961/4.5934)
(a5)	(17,17,15)	-16.1619	-17.9579	-14.6466	(0.9749/5.7530)
(a6)	(0,0,4.7)	-0.0290	-0.0314	-4.69861	(0.9561/13.2613)

In each case, higher MI and SSIM values across all test images as outlined in Table 3.1, explicitly state that the registered images are perfectly aligned with the reference image with significantly reduced error rates. These findings demonstrate the superior performance of the TLBO algorithm in isomodal medical image registration as compared to other evolutionary algorithms [37], [40], [113]. The efficacy of TLBO in obtaining the RTP is further illustrated through the experimental results with the scheme described in this work.



(a1) (b1) (c1)



(a2) (b2) (c2)

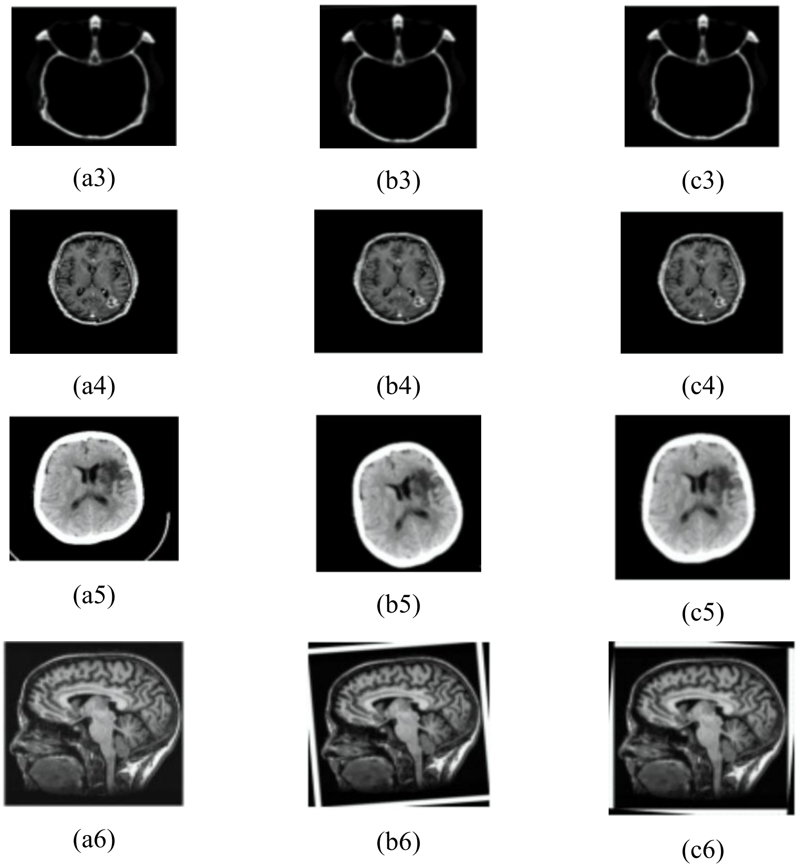


Figure 3.5: MRI & CT modalities of test images (i) MRI (a1, a2 & a6); (ii) CT (a3, a4, a5), RTP (θ, t_x, t_y) for a1-a4 (2,2,2) , for a5 (15,17,17) and for a6 (4.7,0,0)

3.2.1.2. Comparative Analysis

The effectiveness of the proposed scheme in isomodal medical image registration is evaluated by comparing it with recently developed meta-heuristic optimization-based techniques. This comparison ensures a fair assessment as identical images and parameters are used as those considered by [37], [40], [113]. The presented scheme is contrasted with existing approaches presented by Dida *et al.* [37], Ayatollahi *et al.* [40], and Maddaiah *et al*. [113]. Isomodal test images, namely MR1, MR2, CT1, CT2, CT, and brain MR images obtained from the aforementioned database [112], are depicted in Figure 3.5 (Col.1, a1-a6), with dimensions (256×256). The comparison results reveal that the proposed scheme achieves better similarity metrics SSIM and RMSE as compared to existing meta-heuristic approaches, indicating more accurate registration of CT and MRI images with enhanced visual quality, as demonstrated in Figure 3.5. The improved outcomes are a result of TLBO effectively optimizing transformation parameters and effectively reducing image noise across all experiments. All experiments are conducted using TLBO for image registration, with a population size of 30 particles and a maximum of 60 iterations. Future research will focus on extending the proposed method to multi-modal medical image registration by integrating machine learning algorithms to improve the alignment of source and target images.

Table 3.2: Comparison results of rigid transformation parameters for proposed scheme via existing scheme [37] on a1-a4 images in Figure 3.5, where IV is initial value and OA is optimization algorithm

Modalities	IV	OA / RTP	GWO [27]	PSO [27]	Proposed Scheme	GWO [27]	PSO [27]	Proposed Scheme
MR1 & MR2	2	tx	-1.931	-1.930	-1.963	-1.936	-1.948	-1.963
	2	ty	-2.074	-2.074	-2.031	-2.082	-2.076	-2.033
	2	θ	-1.982	-1.978	-1.984	-1.986	-1.991	-1.986
		SSIM	0.9106	0.9196	**0.9653**	0.9007	0.9006	**0.9838**
CT1 & CT2	2	tx	-1.930	-1.936	-1.937	-1.933	-1.933	-1.9716
	2	ty	-2.074	-2.073	-2.048	-2.079	-2.068	-2.0326
	2	θ	-1.978	-1.978	-2.000	-1.983	-1.977	-1.9939
		SSIM	0.9196	0.9195	**0.9852**	0.9191	0.9190	**0.9741**

Table 3.3: Comparison results of rigid transformation parameter of proposed scheme with existing scheme [40] on a5 image in Figure 3.5

RTP	tx	ty	$\boldsymbol{\theta}$	**Similarity Measure (RMSE)**
Initial value / Algorithm	17	17	15	
Hybrid PSO	-16.0480	-17.0133	-14.4844	0.4953
Proposed scheme	-16.1619	-17.9579	-14.7543	**0.1306**

Table 3.4: Comparison results of transformation parameters of the proposed scheme with existing scheme [113] on a6 images of Figure 3.5

RTP	tx	ty	$\boldsymbol{\theta}$	**Similarity Measure**	
Initial value / Algorithm	0	0	4.7	RMSE	RMSE
Hybrid PSO	-0.7240	-0.3521	-4.3768	3.7858	3.7858
Proposed scheme	-0.0290	-0.0314	-4.69861	**0.18091**	**0.18091**

3.3. Robust Medical Image Registration Leveraging Hybrid Approach: TLBO, SURF, Affine and Projective Transformations

In this section, effective area and feature-based methods for aligning both monomodal and multimodal medical images are discussed. This technique initiates noise reduction through Gaussian filtering and proceeds to normalization, rigid transformation, and an optimization strategy. To find the optimal values of rigid transformation parameters (θ, t_x, t_y) using maximization of MI as an objective function of TLBO. Afterwards, the image obtained through the TLBO process and source image undergoes further enhancement to extract robust features using SURF. Then, similar feature points are matched using the K-nearest neighbours (KNN) algorithm, followed by the RANSAC algorithm to eliminate mismatches (outliers). Subsequently, a projective transformation is applied to obtain the accurately registered image. The similarity between source and final registered image is evaluated using various similarity metrics, including MI, RMSE, and SSIM. A high MI and SSIM value, along with a low RMSE value, indicate the robustness of proposed scheme. Additionally, accurate registration results are achieved through the extraction of a higher number of key points (features) and minimizing mismatches. The proposed scheme discussed in this section comprising of several steps as depicted in Figure 3.6.

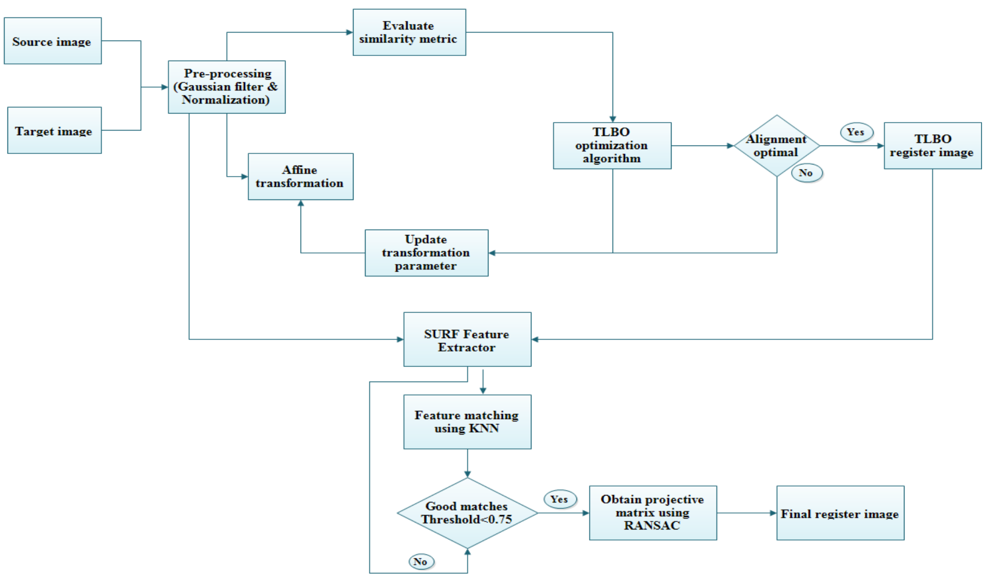
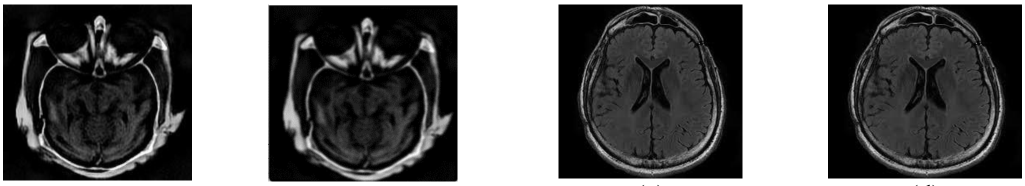


Figure 3.6: Framework of the proposed scheme

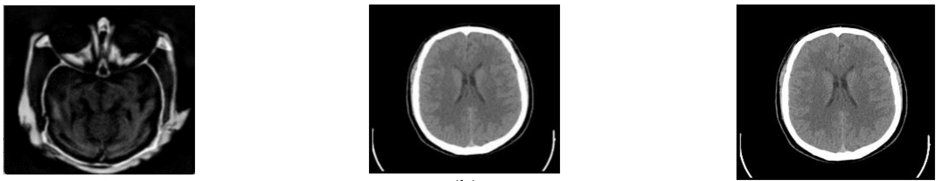
Step 1: Initially, source image undergoes pre-processing, where the Gaussian filter is applied to reduce noise before proceeding to normalization. Subsequently, noise is effectively reduced during the noise removal process, leading to a significant reduction in contrast between anatomical features of an image. Finally, normalization is applied to standardize the image data to a uniform scale, as depicted in Figure 3.7, illustrating the outcome of Step 1.



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

Figure 3.7: (a) Monomodal source MR image [112] (b) Pre-processed image corresponding to (a) (c) Multimodal source MR image [114] (d) Pre-processed image corresponding to (c).

Step 2: In monomodal image registration, the pre-processed source image undergoes affine transformation to produce the template image. Conversely, in multimodal image registration, an affine transformation is applied to the template image first, followed by a rigid transformation. The results of Step 2 on both monomodal and multimodal medical images are illustrated in Figure 3.8 below.



(a) (b) (c)

Figure 3.8: (a) Monomodal template image from Figure 3.7 (b) with rigid transformation parameters $(\theta = 2°, t_x = 2, t_y = 2)$ (b) Multimodal Template CT image (c) CT image from Fig. 3.8 (b) with rigid transformation parameters $(\theta = 2°, t_x = 2, t_y = 2)$.

Step 3: TLBO is utilized to determine the optimal values for rigid transformation parameters (θ, t_x, t_y) with a fitness function as maximizing MI. Subsequently, similarity measures such as MI, RMSE, and SSIM are employed to assess the similarity between source and TLBO registered image which is not accurately aligned. The results for monomodal and multimodal images, depicted in Figure 3.9 (a) and Figure 3.9 (b) respectively, are tabulated in column 3 of Table 3.5.

Figure 3.9: (a) TLBO registered monomodal image (b) TLBO registered multimodal image.

Step 4: To improve the registered image, the SURF feature extraction algorithm is applied to both monomodal and multimodal images to extract features like key points, lines, edges, and corners in both the source and TLBO registered image as depicted in Figure 3.10.

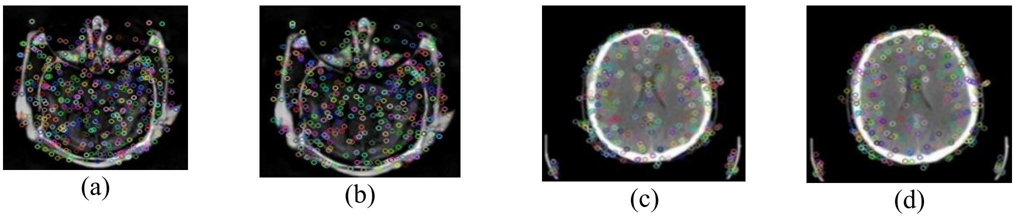
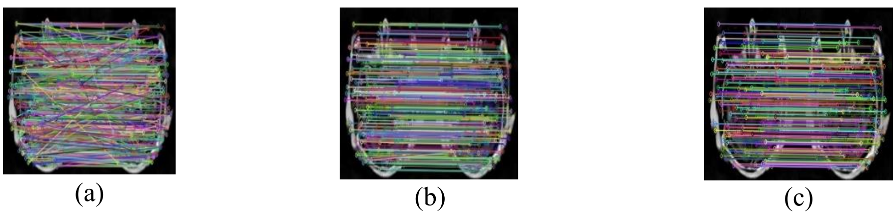


Figure 3.10: (a) Features detected from pre-processed image referring to Fig. 3.7(b) (b) Features detected from TLBO registered image referring to Figure 3.9 (a)) (c) Features detected from pre-processed template image (d) Features detected from TLBO registered image referring to Figure 3.9(b).

Step 5: KNN is employed to match similar points between feature-extracted points from step 4 in both source and TLBO registered image, as illustrated in Figure 3.11(a) and Figure 3.11(d). Matching is considered successful when the distance to the first neighbour falls below a specified threshold value (0.75), as shown in Figure 3.11(b) and Figure 3.11(e). Subsequently, RANSAC is utilized to eliminate mismatches (outliers) between KNN feature-matching images, as depicted in Figure 3.11(c) and Figure 3.11(f). The objective is to achieve the best homography matrix, which results in accurate registration by maximizing the number of inliers and minimizing outliers.



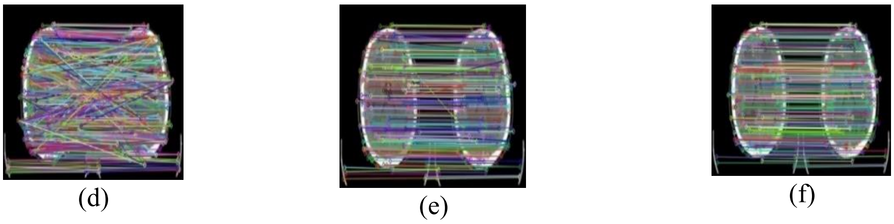


Figure 3.11: (a) Total matches detected between monomodal images (Figure 3.10 (a), Figure 3.10 (b)) (b) Good matches derived from Figure 3.11(a) (c) Removal of outliers using RANSAC (Figure 3.11(b)) (d) Total matches identified between multimodal images (Figure 3.10(c), Figure 3.10(d)) (e) Good matches (Figure 3.11(d)) (f) Removal of outliers using RANSAC (Figure 3.11(e)).

Step 6: Finally, a combination of area and feature-based registered image is generated, as depicted in Figure 3.12(a) and Figure 3.12(c). The registration accuracy is assessed by examining the difference between the source and registered image, illustrated in Figure 3.12(b) and Figure 3.12(d). These illustrations reveal precise alignment between the source and template images. The comparison results for monomodal and multimodal images, evaluating MI for TLBO-Affine registered image and hybridization of TLBO-Affine followed by SURF-Projective transformation. Quantitative analysis of the datasets [112], [115] are used for experimental work as described in Table 3.5.

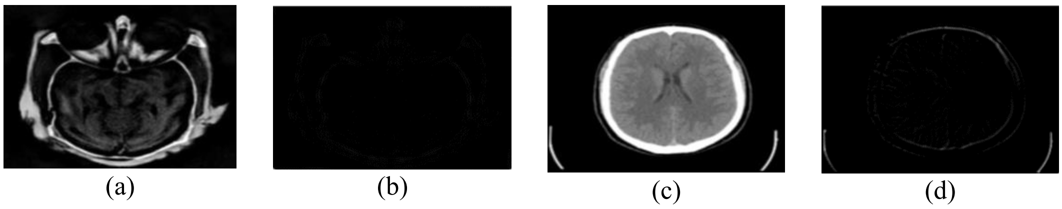


Figure 3.12: (a) Final registered monomodal image (b) Difference between source and final registered monomodal image (c) Final registered multimodal image (d) Difference between source and final registered multimodal image.

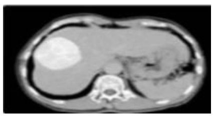
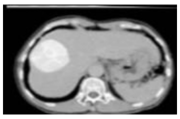
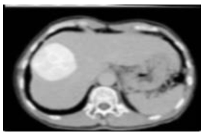
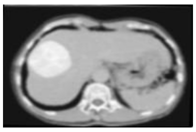
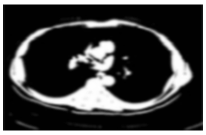
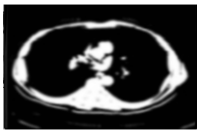
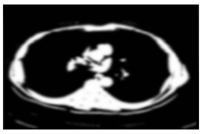
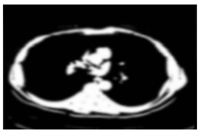
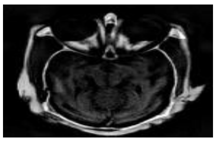
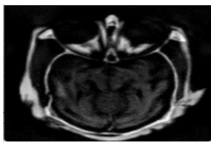
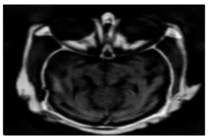
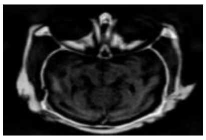
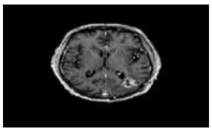
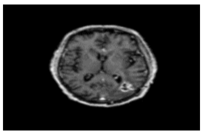
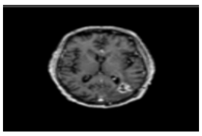
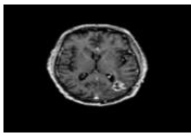
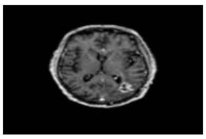
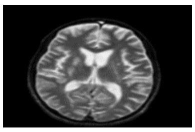
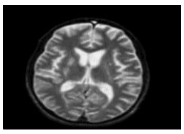
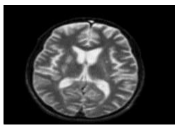
3.3.1. Experimental and Comparative Analysis

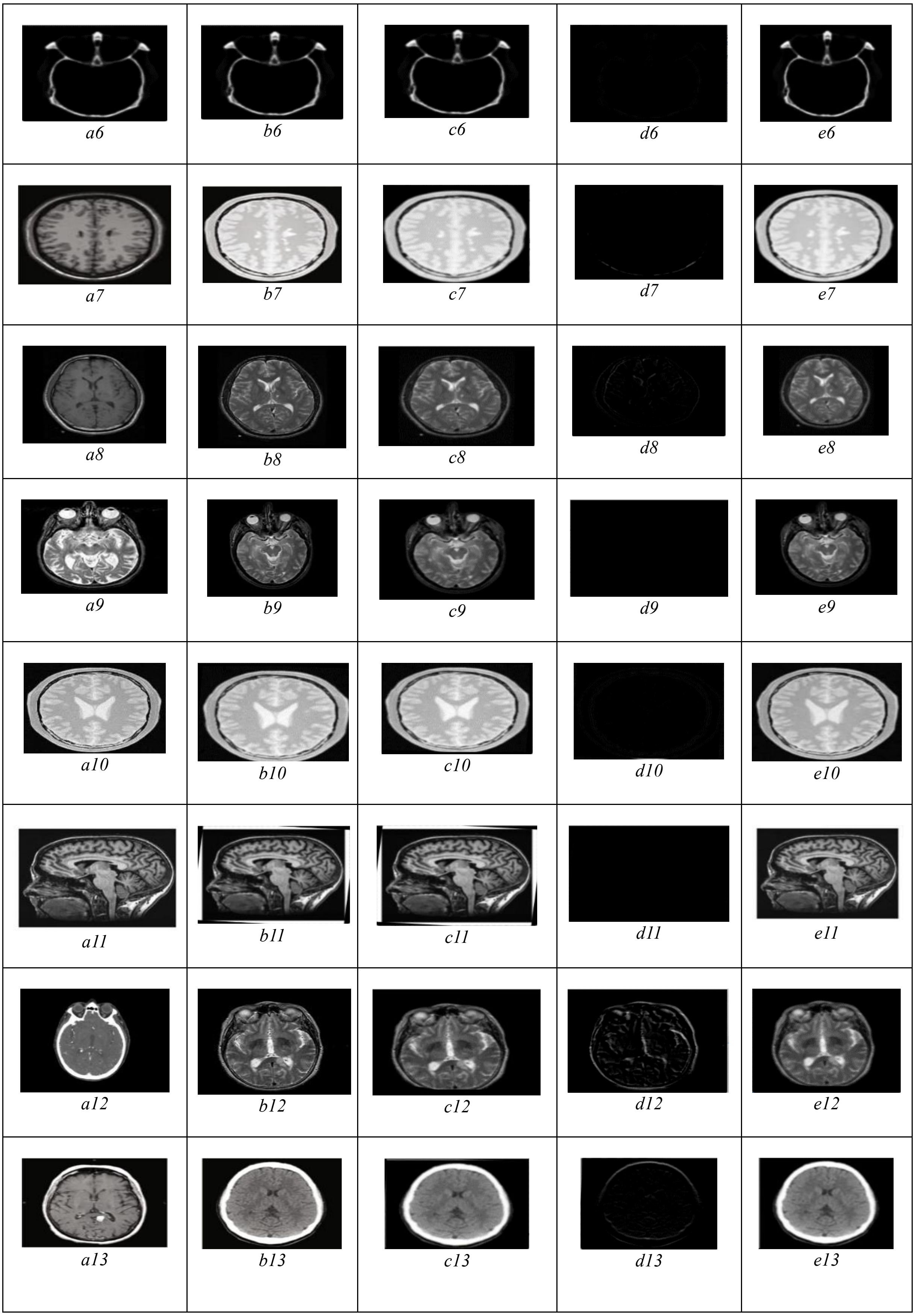
This section illustrates the effectiveness of the scheme through an evaluation of similarity metrics (MI, RMSE, and SSIM) on both monomodal and multimodal medical images. Additionally, we analyse the number of matches and the visual quality of the registered image.

3.3.1.1. Experimental Results

The experiments are conducted on a Windows 11 system equipped with an Intel(R) Core TM i7-1165 G7 CPU running at 2.80 GHz and 16.0 GB RAM, using the

PYTHON 3.9.4 platform. The experiments involved a variety of monomodal medical images sourced from the Whole Brain ATLAS Dataset [112], as depicted in Figure 3.13 (a1-a11). The monomodal and multimodal images used in this study, along with the transformation parameters considered, are summarized in Table 3.6. The CT and MRI monomodal brain images, with dimensions 256×256 (Figure 3.13, a1-a10), 354×353 (Figure 3.13, a11), respectively, as well as the multimodal medical images sized at 256×256 (Figure 3.13, a12-a15), served as the reference images for the experiments. To generate the template (target) image, a rigid transformation is applied to the source (reference) image, involving rotation by an angle θ and translation by $\mathrm{T} = (\mathrm{t_x}, \mathrm{t_y})$ as specified in Table 3.5. In monomodal medical image registration, this transformation is applied to the source image, while in multimodal medical image registration, it is applied to the template image. The alignment of the template image with reference image is achieved using TLBO algorithm, configured with a population of 30 particles and 50 iterations to determine the optimal values for rigid transformation parameters. Next, features are extracted from both monomodal and multimodal medical images using SURF algorithm between pre-processed source image and TLBO-registered image. Subsequently, these features are matched using KNN matcher algorithm. False match key points(features) are then filtered out to enhance the repeatability of feature matching using RANSAC algorithm. Afterwards, the best homography matrix is determined by optimizing for the highest number of inliers (IS) and minimal outliers (OS). Subsequently, a robust registered image is generated by estimating the rigid transformation parameters based on the remaining matches, thereby enhancing the registration accuracy. Experimental results demonstrate this scheme yields more accurate results, achieving maximum values of MI and SSIM and minimum values of RMSE in both monomodal and multimodal medical image registration as compared to traditional approaches. In the initial test case, template image is aligned with reference image. The results of rotating the monomodal image at an angle $\theta = 2^{\circ}$followed by translation $\mathrm{T} = (\mathrm{t_x}, \mathrm{t_y}) = (2,2)$ are illustrated in Figure 3.13 (b1-b9). TLBO registered image is subsequently generated as depicted in Figure 3.13 (Column 3, (c1-c9)), matched with the template image (Column 2, (b1-b9)) employing the optimized rotation and translation parameters. The difference image between the source and final registered image is displayed in Figure 3.13 (Column 4, (d1-d9)). Subsequently, final registered image between the pre-processed source and TLBO registered image is shown in Figure 3.13 (Column 5, (e1-e9)). To further enhance TLBO registered image, features are extracted using SURF in the source image, as illustrated in Figure 3.14 (Column 1, (a1-a9)), while the keypoints of the TLBO registered image are depicted in Figure 3.14 (Column 2, (b1-b9)). Then, features are matched between the source and TLBO registered image using KNN, as depicted in Figure 3.14 (Column 3, (c1-c9)). Subsequently, good matches are extracted, as shown in Figure 3.14 (Column 4, (d1-d9)), and RANSAC is utilized to reduce mismatches and enhance match repeatability, as illustrated in Figure 3.14 (Column 5, (e1-e9)). The alignment quality of the TLBO registered image is further improved through source and SURF-Projective transformation. In various test cases, different values of translation and rotation parameters are adjusted to simulate different scenarios. The higher MI and SSIM values, along with lower MSE and RMSE values, indicate the accurate alignment of registered images in all these scenarios, as detailed in Table 3.6. The scheme demonstrates superior performance as compared to traditional evolutionary algorithms [37], [40], [42], [49], [113], [116].

Source image	Template image	TLBO registered image	Difference between source and final registered image	Final registered image
 *a1*	 *b1*	 *c1*	 *d1*	 *e1*
 *a2*	 *b2*	 *c2*	 *d2*	 *e2*
 *a3*	 *b3*	 *c3*	 *d3*	 *e3*
 *a4*	 *b4*	 *c4*	 *d4*	 *e4*
 *a5*	 *b5*	 *c5*	 *d5*	 *e5*


a6
b6
c6
d6
e6
a7
b7
c7
d7
e7
a8
b8
c8
d8
e8
a9
b9
c9
d9
e9
a10
b10
c10
d10
e10
a11
b11
c11
d11
e11
a12
b12
c12
d12
e12
a13
b13
c13
d13
e13

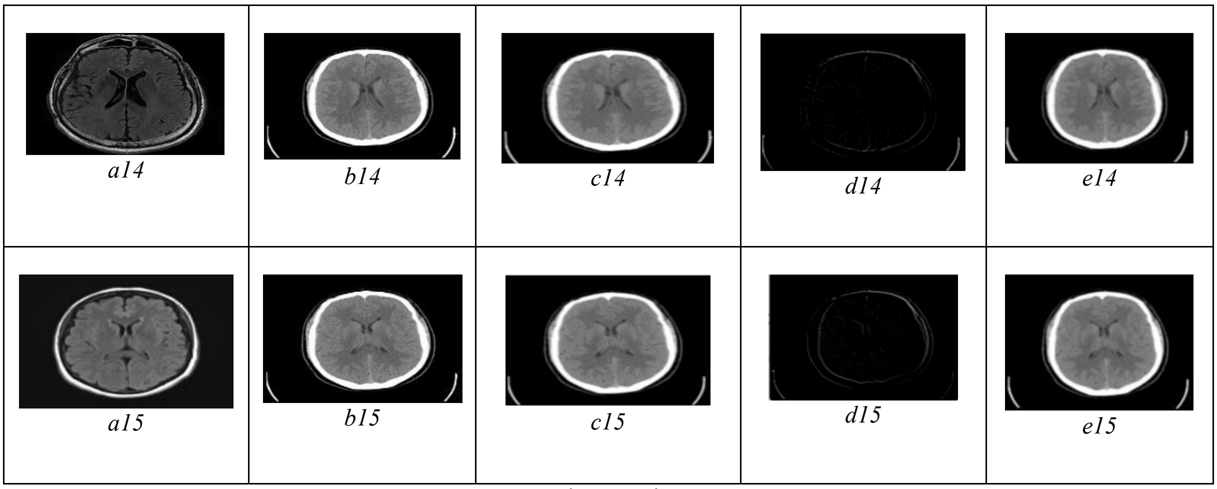
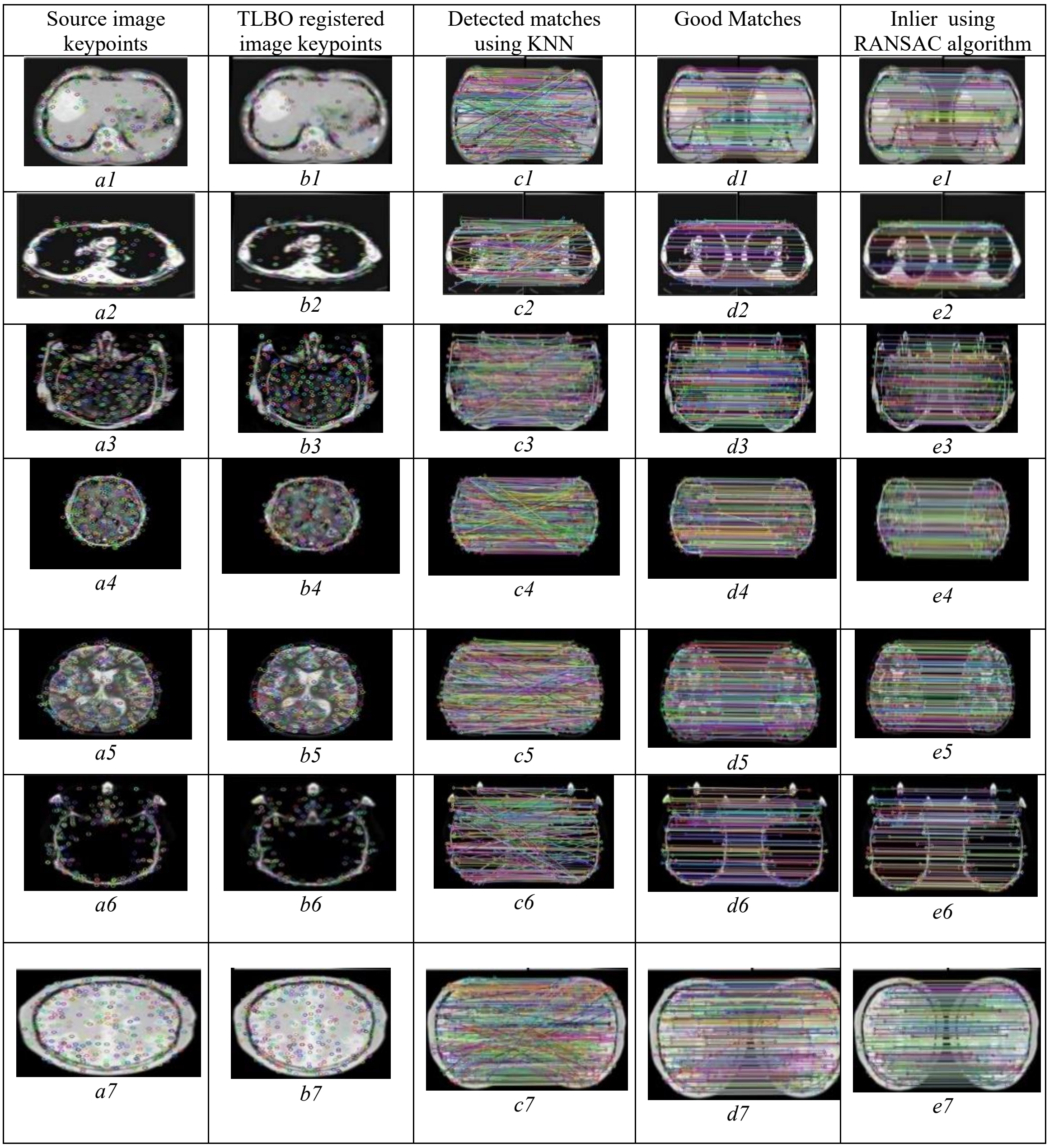


Figure 3.13: Rigid transformation parameters (θ, t_x, t_y) for a1-a9 are (2,2,2) , for a10 are (3,5,7) and for a11 are (4.7,0,0), for a12 are (15,17,17) and a13-a15 are (2,2,2)



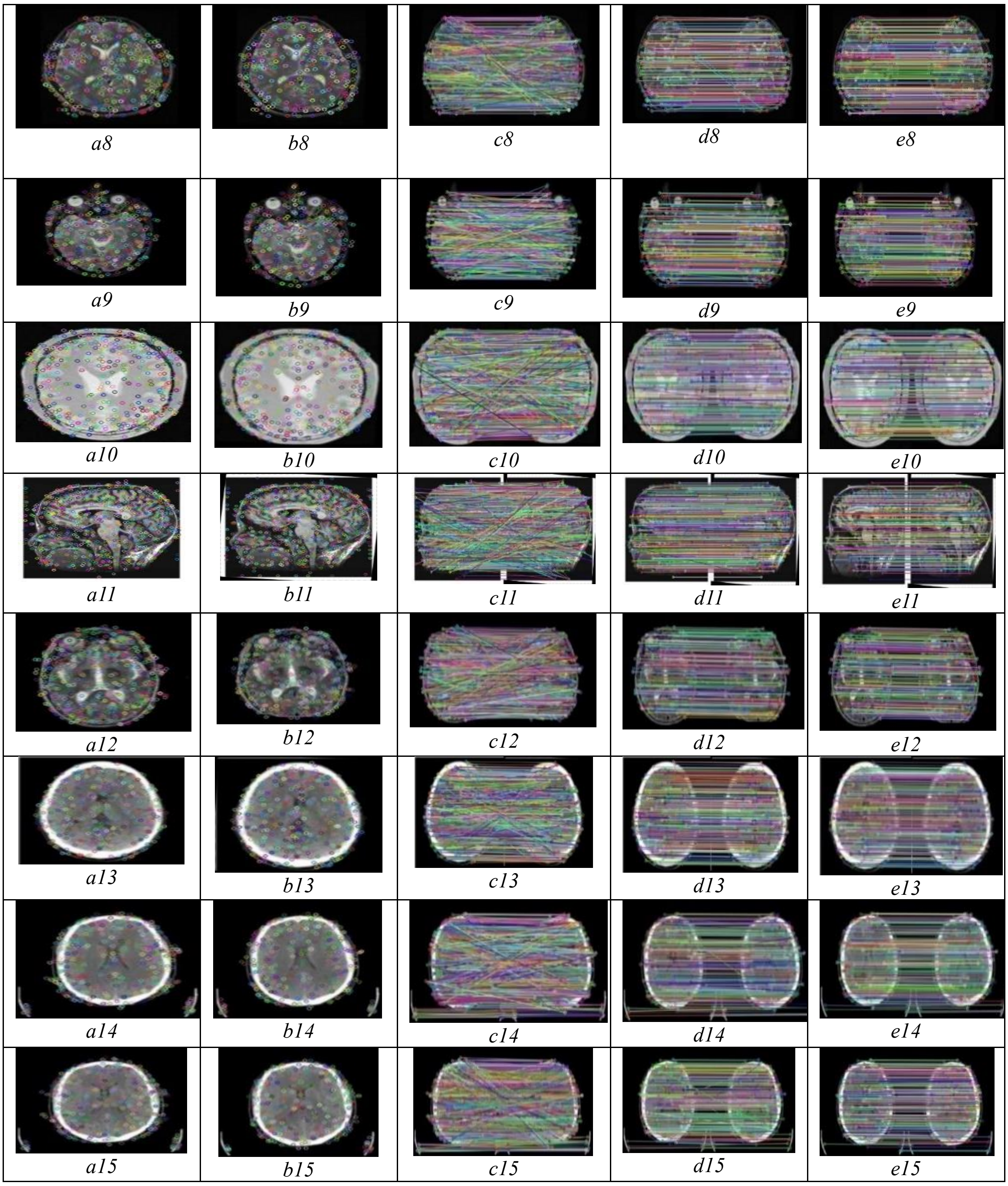


Figure 3.14: Experiment outcome of different monomodal and multimodal images: (a) keypoints extraction of source image(Figure a1- Figure a15) (b) keypoints extraction of TLBO registered image (Figure b1- Figure b15) (c) Total matches using KNN (d) Good matches (e) Inliers using RANSAC

Table 3.5: Experimental results on fifteen different monomodal/multimodal images taken from Whole Brain ATLAS and Kaggle Dataset

Images	Ground Truth	Registration Parameters			$\left(\frac{MI}{SSIM}\right)$	$\left(\frac{RMSE}{MSE}\right)$	$\left(\frac{MS}{GM}\right)$	$\left(\frac{IS}{OS}\right)$
	(θ, t_x, t_y)	θ	t_x	t_y				
a1	(2,2,2)	-1.995246	-1.953835	-2.024725	$\left(\frac{11.8616}{0.99326}\right)$	$\left(\frac{0.05992}{0.003591}\right)$	$\left(\frac{291}{242}\right)$	$\left(\frac{236}{6}\right)$
a2	(2,2,2)	-1.877150	-1.968412	-1.973259	$\left(\frac{7.68657}{0.975127}\right)$	$\left(\frac{0.06843}{0.00469}\right)$	$\left(\frac{208}{167}\right)$	$\left(\frac{166}{1}\right)$
a3	(2,2,2)	-1.984798	-1.961969	-2.026555	$\left(\frac{12.47402}{0.99560}\right)$	$\left(\frac{0.00725}{0.0000526}\right)$	$\left(\frac{455}{374}\right)$	$\left(\frac{367}{7}\right)$

a4	(2,2,2)	-1.993975	-1.959465	-2.038087	$\left(\frac{4.838813}{0.99831}\right)$	$\left(\frac{0.0066655}{0.0000443}\right)$	$\left(\frac{310}{259}\right)$	$\left(\frac{252}{7}\right)$
a5	(2,2,2)	-1.986545	-1.958860	-2.031092	$\left(\frac{7.84023}{0.998380}\right)$	$\left(\frac{0.005882}{0.0000346}\right)$	$\left(\frac{372}{319}\right)$	$\left(\frac{314}{5}\right)$
a6	(2,2,2)	-2.000038	-1.929636	-2.076320	$\left(\frac{5.38299}{0.999254}\right)$	$\left(\frac{0.00435988}{0.00001900}\right)$	$\left(\frac{252}{201}\right)$	$\left(\frac{196}{5}\right)$
a7	(2,2,2)	-2.023572	-0.843061	-2.967087	$\left(\frac{8.79876}{0.45759}\right)$	$\left(\frac{0.097752}{0.009956}\right)$	$\left(\frac{437}{349}\right)$	$\left(\frac{337}{12}\right)$
a8	(2,2,2)	-2.591633	-1.160894	-2.097307	$\left(\frac{8.99190}{0.706819}\right)$	$\left(\frac{0.1220083}{0.014420}\right)$	$\left(\frac{337}{263}\right)$	$\left(\frac{255}{8}\right)$
a9	(2,2,2)	-2.589592	-2.955757	-1.036038	$\left(\frac{5.14818}{0.474578}\right)$	$\left(\frac{0.25798}{0.0665544}\right)$	$\left(\frac{310}{237}\right)$	$\left(\frac{234}{3}\right)$
a10	(3,5,7)	-3.007931	-4.623836	-7.218201	$\left(\frac{12.8284}{0.97447}\right)$	$\left(\frac{0.031026}{0.0009626}\right)$	$\left(\frac{480}{390}\right)$	$\left(\frac{380}{10}\right)$
a11	(4.7,0,0)	-4.698647	0.025615	0.026671	$\left(\frac{12.9876}{0.96367}\right)$	$\left(\frac{0.18900}{0.0357210}\right)$	$\left(\frac{832}{652}\right)$	$\left(\frac{641}{11}\right)$
a12	(15,17,17)	-13.999654	-15.903154	-17.038512	$\left(\frac{5.30944}{0.588887}\right)$	$\left(\frac{0.20526}{0.042134}\right)$	$\left(\frac{391}{277}\right)$	$\left(\frac{272}{5}\right)$
a13	(2,2,2)	-2.010362	-1.959127	-0.033599	$\left(\frac{7.3538}{0.605810}\right)$	$\left(\frac{0.23395}{0.0547327}\right)$	$\left(\frac{325}{268}\right)$	$\left(\frac{264}{4}\right)$
a14	(2,2,2)	-0.794920	-2.032515	-0.979775	$\left(\frac{5.2202}{0.471868}\right)$	$\left(\frac{0.284271}{0.08081}\right)$	$\left(\frac{341}{231}\right)$	$\left(\frac{219}{12}\right)$
a15	(2,2,2)	-1.400124	-2.966063	-0.976563	$\left(\frac{5.75569}{0.595064}\right)$	$\left(\frac{0.241492}{0.058318}\right)$	$\left(\frac{355}{240}\right)$	$\left(\frac{222}{18}\right)$

Table 3.6: Summary of different modalities of images shown in Fig. 3.9 along with transformation parameters

Sr. No	Modality		Rigid transformation parameters (θ, t_x, t_y)	Dimension
1	Mono-modality	Liver CT/Liver CT	(2,2,2)	$(256 * 256)$
2		Brain CT/ Brain CT		
3		MR I1 / MR I1		
4		CT I2/ CT I2		
5		MR I2 / MR I2		
6		CT I1/ CT I1		
7		PD- weighted MR/ T1-weighted MR		
8		MR T1 / MR T2		
9		MR T2 / MR T2		
10		MRI-PD / MRI-PD	(3,5,7)	
11		MR / MR	(4.7,0,0)	$(354 * 353)$
12		CT / MR T2	(15,17,17)	

13	Multi-modality	MR / CT	(2,2,2)	$(256 * 256)$
14		MR / CT		
15		MR / CT		

3.3.1.2. Comparative Analysis

This section provides a comparison between proposed scheme and recently developed image registration methods [37], [40], [42], [49], [113], [116]. All comparisons are conducted on similar images with the same number of parameters, considering their respective ground truth values. Table 3.7 outlines the methodologies employed by various authors in the state-of-the-art approaches [37], [40], [42], [49], [113], [116] that are compared to proposed scheme.

Table 3.7: Summary of state-of-the-art methods used for comparison with proposed scheme

Reference	Methodology
Dida *et al.* [37]	The method has been assessed using medical images, including CT and MRI, and is suitable for area-based techniques.
Samantaray *et al.* [116]	Implementation of an effective image registration algorithm incorporating SURF and BRISK feature extraction, enhanced with upgraded GMS and RANSAC algorithms.
Ayatollahi *et al.* [40]	Hybrid PSO is employed on multimodal images, utilizing RMSE to measure similarity.
Zheng *et al.* [49]	Employing the progressive image method and SURF algorithm for medical image registration, based on the Brain Web database.
Maddaiah *et al.* [113]	Implementation of a monomodal image registration algorithm specifically designed for aligning T1-weighted MR images using PSO.
Gupta *et al.* [105]	A global optimization algorithm such as ACO is applied to medical images using various similarity metrics.

The scheme is evaluated against existing meta-heuristic optimization strategies presented by Dida *et al.* [37] for monomodal medical image registration. The comparison is conducted using test images MR I1, CT I2, MR I2, and CT I1, each with dimension 256×256, sourced from the whole brain Atlas database [112] as shown in $(\text{Figure 3.13, Col. 1, a3} - \text{a6})$. A template image is created by applying an affine transformation with θ=2° followed by translation $T = (t_x, t_y) = (2,2))$ to serve as the ground truth as depicted in $(\text{Figure 3.13, Col. 2, b3} - \text{b6})$. TLBO method is utilized to find optimal transformation parameters and reduce noise. The quantitative evaluation, employing SSIM metric as presented in Table 3.8 and Figure 3.15. The proposed scheme demonstrates superior SSIM values as compared to the method outlined in [37] for monomodal image registration as illustrated in Table 3.8. The higher SSIM values indicate that the CT and MRI images are aligned accurately, and the visual quality of the images is improved, as depicted in Figure 3.13. All experiments are conducted across the specified images as shown in $(a3-a6, \text{Col. } 1, \text{Fig. } 3.13)$ for alignment using TLBO with a population size of 30 particles and a maximum of 50 iterations.

Table 3.8: Comparison of SSIM for proposed approach with existing method [37] corresponding to images a3-a6 as shown in Fig. 3.13

Images	**Ground Truth**	**RTP \ Algorithm**	**GWO [37]**	**PSO [37]**	**Proposed TLBO**	**GWO [37]**	**PSO [37]**	**Proposed TLBO**
MR I1 and MR I2	2	t_x	-1.9824	-1.9997	-1.986545	-1.9863	-1.9774	-2.03109
	2	t_y	-1.9311	-1.9514	-1.958860	-1.9367	-1.9333	-1.9865
	2	θ	-2.0741	-2.0954	-2.031092	-2.0822	-2.0686	-1.9865
		SSIM	0.9106	0.9103	**0.9956**	0.9007	0.9190	**0.99833**
CT I1 and CT I2	2	t_x	-1.9788	-1.9789	-2.000038	-1.9835	-1.9774	-1.986545
	2	t_y	-1.9306	-1.9360	-1.929636	-1.9335	-1.9333	-1.958860
	2	θ	-2.0744	-2.0734	-2.076320	-2.0797	-2.0686	-2.031092
		SSIM	0.9196	0.9195	**0.99779**	0.9191	0.9190	**0.998332**

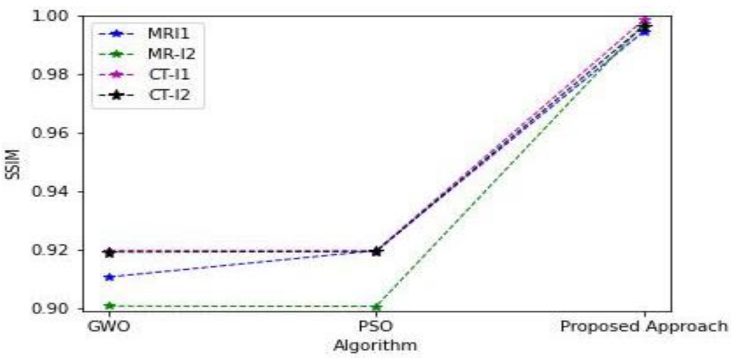


Figure 3.15: Comparison of SSIM between proposed approach and existing method based on GWO [37] and PSO [37] on monomodal MR-I1, MR-I2, CT-I1,CT-I2 images.

The proposed scheme undergoes comparison with scheme by Samantaray *et al.* [116]. For the experimental evaluation, CT images of the liver and brain, both monomodal images are sized at (256×256), sourced from the aforementioned database [112] as depicted in (Fig. 3.13, Col.1, a1 and a2). An affine transformation is then applied to generate target images with ground truth values (rotation $\theta = 2^{\circ}$ and translation $T = (t_x, t_y) = (2,2)$) as depicted in Fig. 3.13 (Col.1, b1and b2). Matching between reference and target images are accomplished using KNN matcher, illustrated in Fig. 3.14 (Col. 3, c1 and c2). To enhance accuracy while minimizing errors, RANSAC algorithm is deployed to identify inliers in the images, as shown in Fig. 3.14 (Col. 5, e1 and e2). A quantitative assessment of the discussed scheme on liver and brain CT images is summarized in Table 3.9. The results tabulated in Table 3.9 and Fig. 3.16 with higher MI and lower RMSE values reveal precise registration. This is due to the optimal values of rigid transformation parameters obtained through TLBO and these results still lack the precise alignment of source and target images. Then, SURF is applied to extract robust features followed by KNN matcher to find the best match and RANSAC to remove the outliers between source and TLBO register image. The results corresponding to each step are shown in Figure 3.14. The authors of [14] did not claim the results corresponding to multimodal medical images in their study.

Table 3.9: Comparison of proposed approach with [116] based on MI and RMSE similarity metrics corresponding to liver and brain CT image

Modalities	**Technique** / **Similarity metrics**	**SURF + BRISK + GMS + RANSAC**	**TLBO +Affine + SURF+ Projective Transformation**
Liver CT	MI	9.2814	**11.8616**
	RMSE	0.9271	**0.05992**
Brain CT	MI	8.6573	**9.89657**
	RMSE	0.8124	**0.04702**

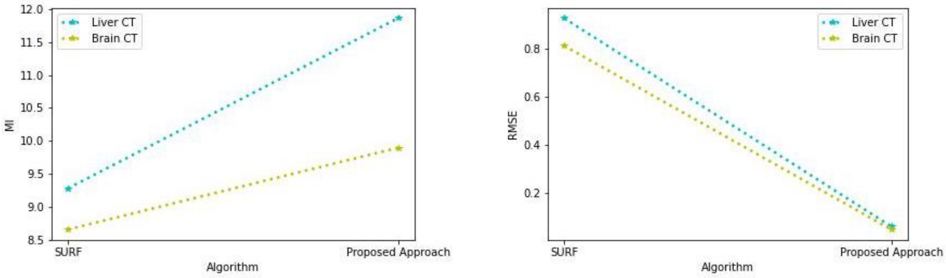


Figure 3.16: Comparison of MI and RMSE between proposed approach and existing method [116] based on SURF for monomodal images.

The method presented by Ayatollahi *et al*. [40] , which employs hybrid particle swarm optimization (HPSO) for multimodal brain image registration is contrasted with proposed scheme. We employed a source image of CT modality with dimensions of (256×256) from the Kaggle dataset [115], as depicted in Figure 3.13 (Col.1, a12). Subsequently, a template image corresponding to MR-T2 modality is generated by applying rigid transformation with rotation ($\theta = 15^{\circ}$ and translation $T = (t_x, t_y) = (17{,}17)$), establishing ground truth values, as shown in Figure 3.13 (Col.2, b12).The comparative analysis between discussed scheme and [40] based on the average RMSE value is presented in Table 3.10. The lower RMSE value achieved by the proposed scheme utilizing a population size of 30 particles and a maximum of 50 iterations, indicates more accurate alignment between source and template images, as shown in Table 3.10. Furthermore, the enhanced performance of the proposed scheme is due to the optimal rigid transformation parameters achieved through TLBO, which initially lack precise alignment. Subsequently, SURF is then used to extract robust features, with KNN matcher identifying the best matches and RANSAC is used for eliminating outliers between the source and TLBO-registered image. Figure 3.14 shows step-by-step outcomes.

Table 3.10: Transformation parameters comparison of proposed approach with [40] on multimodal image a12 as shown in Figure 3.14

Rigid Transformation Parameters	$\boldsymbol{\theta}$	$\boldsymbol{t_x}$	$\boldsymbol{t_y}$	**Similarity Metric (RMSE)**
Ground Truth / Optimization Algorithm	15	17	17	
HPSO	-15.1919	-17.4194	-17.6803	0.4953
Proposed Approach	**-13.9996**	**-15.903154**	**-17.038512**	**0.2053**

The proposed scheme undergoes comparison with the method discussed by Zheng *et al*. [49], which employs progressive-SURF. To conduct the comparison, we utilize a monomodal PD-weighted MR source image with dimension $(256 \text{ x } 256)$ from the aforementioned database [112], as depicted in Figure 3.13 (Col.1, a7). The affine transformation applied to create the floating image (T1-weighted MR) has ground truth values of rotation ($\theta = 2^{\circ}$ and translation $T = (t_x, t_y) = (2{,}2)$) as depicted in Figure 3.13 (Col.2, b7). The comparison involves a multimodal MR source image with dimensions of $(256 \text{ x } 256)$ from the Kaggle dataset [115], as depicted in Figure 3.13 (Col.1, a13). The ground truth values for the affine transformation used to construct the template image (CT image) include rotation ($\theta = 2^{\circ}$ and translation $T = (t_x, t_y) = (2{,}2)$), shown in Figure 3.13 (Col.2, b13). The comparison outcomes between the discussed scheme and [49], based on NCC and MI values, are summarized in Table 3.11 and Figure 3.17. A higher MI value along with lower NCC and MSD values between the source and final registered image indicate that the proposed scheme outperforms [49]. This is achieved through TLBO-derived optimal rigid transformation parameters even with initial misalignment. SURF is used to extract robust features, followed by KNN matcher and RANSAC for robust matching and outlier removal. The results at each stage are illustrated in Figure 3.14.

Table 3.11: Comparison of proposed approach with [49] based on MI, NCC and MSD on a7 and a13 images as shown in Figure 3.14

Type	Modalities	Method / Similarity metric	PI-SURF	TLBO with surf and projective transformation
Monomodal	PD-weighted MR & T1-weighted MR	MI	2.2577	**8.79876**
		NCC	0.9219	**0.98224**
		MSD	0.0082	**0.0047**
Multimodal	MR & CT	MI	1.5861	**7.3538**
		NCC	0.7334	**0.7543**
		MSD	0.0521	**0.05021**

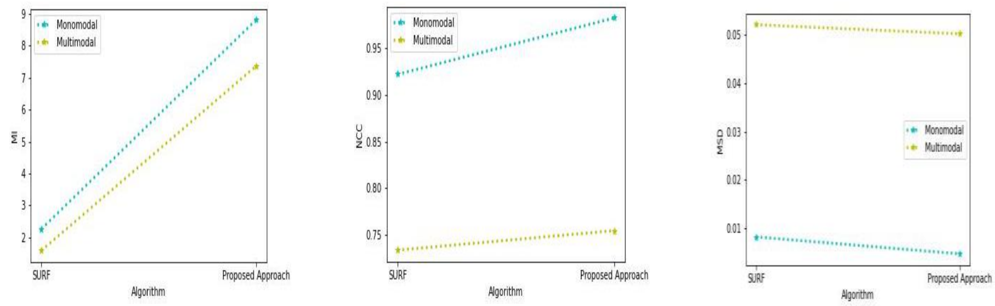


Figure 3.17: Comparison of MI, NCC and MSD between proposed approach and existing method [40] based on SURF for monomodal as well as multimodal images

The proposed scheme is compared to the method by Maddaiah *et al*. [113], which employs PSO to estimate the transformation parameters. For fair comparison, a monomodal brain MR image as used by authors [113] with dimensions (354×353) utilized from the database [112], as illustrated in Figure 3.13 (Col.1, a11). Subsequently, an affine transformation is applied to create the template image, with the ground truth values set at rotation ($\theta = 4.7^{\circ}$and translation $T = (t_x, t_y) = (0,0)$) as shown in Figure 3.13 (Col.2, b11). Quantitative results of the discussed scheme and the method presented in [113], utilizing RMSE and MI as similarity metrics, are summarized in Table 3.12. The improved outcomes indicated by higher MI and lower RMSE values are achieved through TLBO-derived optimal transformation parameters overcoming initial alignment issues. The method uses SURF for feature extraction, KNN for matching, and RANSAC for outlier removal, as shown in Figure 3.13.

Table 3.12: Transformation parameters and similarity metric comparison for discussed scheme with [113] on monomodal image a11 as shown in Figure 3.14

Rigid Transformation Parameters	$\boldsymbol{\theta}$	$\boldsymbol{t_x}$	$\boldsymbol{t_y}$	**Similarity Metric**	
Ground Truth / Optimization Algorithm	4.7	0	0	RMSE	MI
HPSO	-4.6868	-0.7410	-0.3611	78.8358	5.3642
TLBO with SURF and projective transformation	-4.69864	0.02561	0.02667	**0.17489**	**13.5823**

A comparison between the proposed scheme and Gupta *et al*. [42], which employs ACO for medical image registration is carried out. We utilize a source image (MRI-PD) from database [112], sized at (256×256), as shown in Figure 3.13 (Col.1, a10). To construct the template image, a rigid transformation with rotation ($\theta = 3^{\circ}$ and translation $T = (t_x, t_y) = (5,7)$) is applied, as illustrated in Figure 3.13 (Col.2, b10).

The comparative results of proposed scheme and [42], based on similarity metrics such as MSE and MI values, are summarized in Table 3.13. The smaller MSE value achieved with proposed scheme results from TLBO-derived optimal transformation parameters addressing initial alignment challenges. Utilizing SURF for feature extraction, KNN for matching, and RANSAC for outlier removal as illustrated in Figure 3.14, ensures accurate alignment of the source and template images. This was achieved with a population size of 30 particles and 50 iterations, culminating in a robust registered image (e10, Col.5, Figure 3.13).

Table 3.13: Comparison of transformation parameters for proposed approach with existing approach [42] on a10 image in Figure 3.14

Rigid Transformation Parameters	θ	t_x	t_y	**Similarity Metric**	
Ground Truth / Optimization Algorithm	3	5	7	MSE	MI
ACO	-2.8976	-4.2567	-6.68532	45.88	0.96
TLBO with surf and projective transformation	-3.00793	-4.623836	-7.21820	**0.000962**	**12.828**

Comparative analysis across various monomodal and multimodal medical images indicates that the proposed scheme outperforms [37], [40], [42], [49], [113], [116] for several reasons:

- It achieves precise alignment by accurately identifying optimal values for rigid transformation parameters.
- Utilizing SURF provides robust alignment by producing images with fewer outliers as compared to SIFT.
- Achieves superior feature quality and ensures precise registration accuracy.
- Through effective feature extraction leveraging SURF, KNN for best matching keypoints, and RANSAC for removal outliers ensures the accurate alignment between source and template images.

3.3.2. Conclusion

With a notion to proffer a novel and robust image registration method, we employ an optimization technique with rigid transformation and enhance the registered image through feature extraction followed by projective transformation. TLBO optimizes rigid transformation parameters to maximize MI between the source and TLBO-registered images. The output image produced by TLBO and affine transformation is not precisely aligned, SURF and projective transformation further improve this issue. SURF is used to extract robust feature points and KNN is employed to match the extracted feature points. Mismatched features between the source and template image are reduced using RANSAC algorithm. Finally, projective transformation is applied to generate the robust registered image. Extensive experiments on monomodal and multimodal medical images demonstrate significant improvements: SSIM increases by 8%, MI by 12.71% on monomodal images, and reduces RMSE by 41.44% on multimodal images, confirming its robustness as compared to state-of-the-art methods.

3.4. Chapter Summary

This chapter outlines the process of generating precisely registered images for isomodal medical images, highlighting the visual quality of aligned images using TLBO and rigid transformation. The proposed scheme employs TLBO to find optimal rotation and translation parameters with MI maximization as an objective function. To assess the proposed scheme performance, extensive experiments are conducted on diverse image modalities, including CT and MRI images with noise and artifacts. The experimental results indicate an 8% increase in MI and SSIM values, signifying better quality of registered images. To address challenges such as the optimal value of rigid transformation parameters due to anatomical variations and spatial transformations, we proposed a novel image registration algorithm that integrates both area and feature-based approaches for monomodal and multimodal medical images. This method employs the SURF framework for feature extraction along with the RANSAC algorithm to refine the features. Subsequently, projective transformation is used to achieve a more accurate register image. Experiments are conducted on whole brain ATLAS and Kaggle datasets demonstrating the robustness of this scheme, with an improvement of 8% in SSIM, a 12.71% increase in MI for monomodal images, and a 41.44% reduction in RMSE for multimodal images as compared to recent state-of-the-art methods. In the next chapter, we present a method that integrates kernel principal component analysis (KPCA) with TLBO specifically designed to handle challenges related to complex anatomical variations and spatial transformations.

Chapter 4 Multimodal Image Registration Leveraging Kernel PCA and TLBO

4.1. Introduction

In this chapter, an integration of KPCA and TLBO-based medical image registration techniques for monomodal and multimodal modalities is presented. This scheme comprises three phases pre-processing, extracting contours using Ostu's thresholding method for finding ground truth values of translation parameters, feature extraction using KPCA to determine rotation value, and evaluating optimal rigid transformation parameters using TLBO for precise alignment. This process involves contour extraction to create a boundary feature image with centroid localization used to compute translation parameters $\left(t_x, t_y\right)$ for spatial alignment. KPCA enhances feature representation by capturing non-linear relationships and determining rotation value (θ) while TLBO optimizes rigid transformation parameters for accurate alignment. Extensive experiments are conducted on monomodal and multimodal medical images, including CT and MRI from the Harvard Brain ATLAS [112], Kaggle [114], [115], [117]–[119] and an in-house clinical dataset from Fortis Hospital Mohali, India. The proposed scheme significantly outperforms state-of-the-art methods with improvements in RMSE from 64.4034 to 59.1837, MSE from 4147.797 to 3502.710 and SSIM from 0.72456 to 0.75321.

This chapter is structured as follows. The proposed scheme based on the combination of KPCA and TLBO is described in Section 4.2. The proposed scheme effectiveness is validated through experimental results on various datasets is summarized in Subsection 4.3.1, using quality metrics like RMSE, PSNR, SSIM, and CC for different modalities of images. The efficiency of proposed scheme with respect to image alignment is evaluated in Subsection 4.3.2 through comparisons with state-of-the-art methods. The chapter summary is outlined in Section 4.4.

4.2. Multimodal Medical Image Registration Based on KPCA and TLBO

The proposed CT-MRI registration approach comprises a set of steps aims to achieve accurate image alignment is illustrated in Figure 4.2. Consider the image registration as an optimization problem, in proposed method maximization of SSIM between source and target images is considered as an objective function of TLBO to obtain optimal value of transformation parameters, mathematically depicted in Equation (4.1).

$$\hat{T} = \arg\max_{T} SSIM\left[f_p(x,y), f_q(x,y)\right] \tag{4.1}$$

The source image $\left(f_p(x,y)\right)$ and target image $\left(f_q(x,y)\right)$ defined by coordinates x and y undergo an optimal transformation denoted by $\hat{T}$. The proposed scheme framework includes the steps outlined in Figure. 4.1.

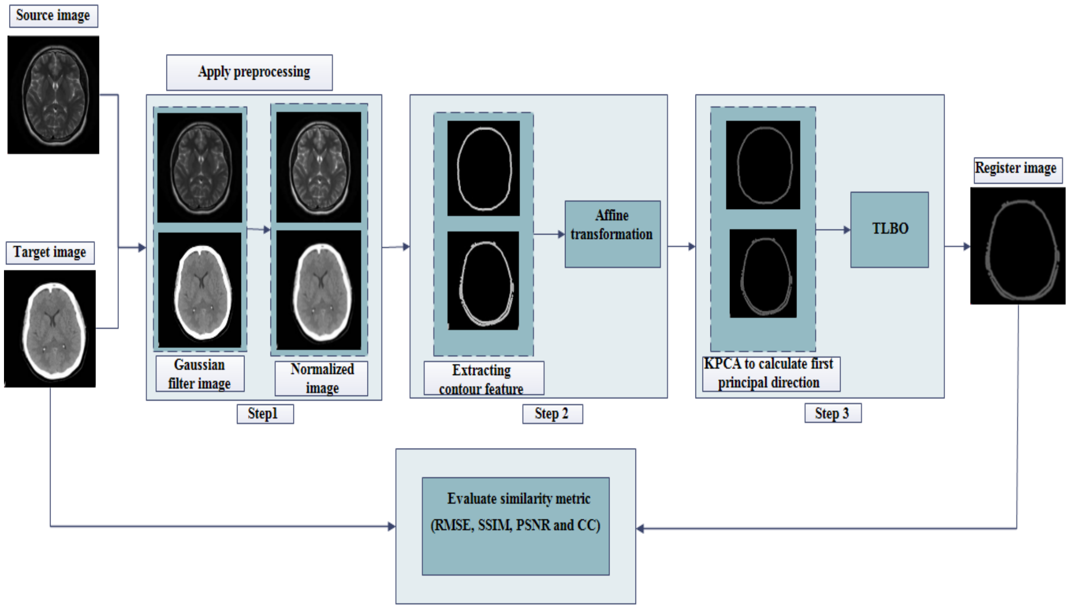


Figure 4.1: Framework of the proposed approach

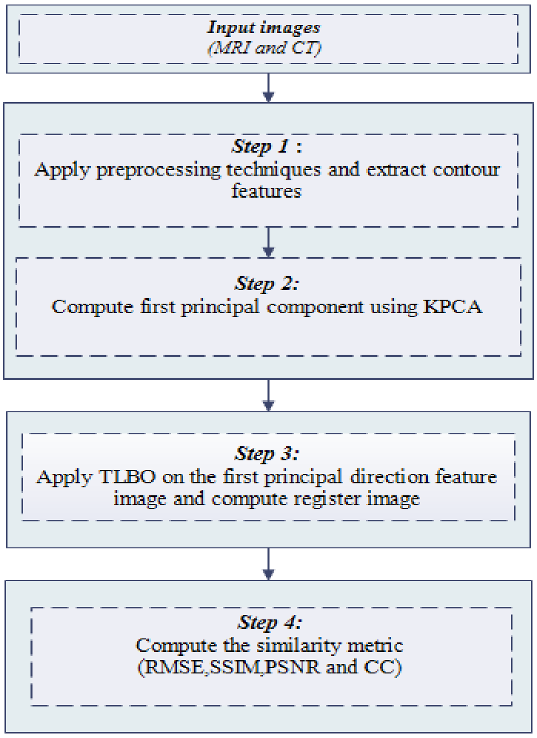
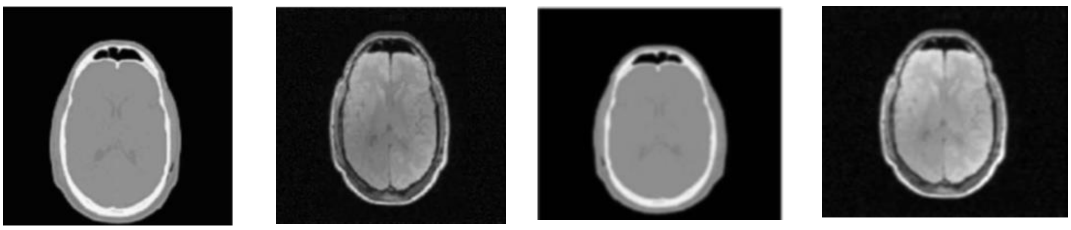


Figure 4.2: Workflow of proposed approach

Step 1: *Pre-processing*: The source (reference) and target images, illustrated in Figure 4.3 (a_1) and Figure 4.3 (b_1) respectively, undergo initial pre-processing using a Gaussian filter in CT-MRI registration. Subsequently, the pre-processed images undergo min-max normalization, resulting in the reference and target images depicted in Figure 4.3 (c_1) and Figure 4.3(d_1) respectively. The integration of Gaussian filtering and min-max normalization improves image quality by diminishing noise levels and enhancing overall contrast. This process ensures that image pixel values are scaled within a specific range, facilitating accurate comparison, and matching of the CT and MRI images during the registration process.



(a_1) (b_1) (c_1) (d_1)

Figure 4.3: (a_1) Reference image (b_1) Target image (c_1) Pre-processed source image (d_1) Pre-processed target image

Step 2: *Feature extraction*: The contour tracking approach(Moore Neighbour Tracing) [120] is employed to extract contour information from binary image obtained from Otsu's thresholding, which subsequently serves as the feature image for registration, depicted in Figure 4.4 (c_1) and Figure 4.4 (d_1) respectively. Otsu's thresholding approach is utilized to determine a threshold value of 40 and convert the grayscale image into a binary image, as illustrated in Figure 4.4 (a_1) and Figure 4.4 (b_1). In contrast, feature images for monomodal MRI-MRI are generated through threshold image segmentation as shown in Figure 4.10.

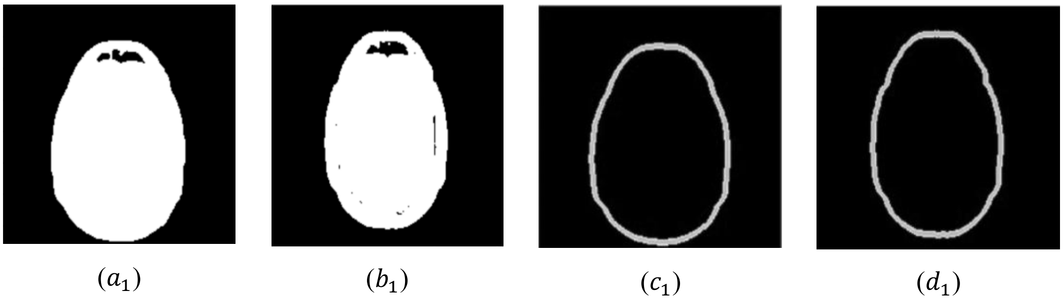


Figure 4.4: (a_1) Source image using Otsu's threshold method (b_1) Target image using Otsu's threshold method (c_1) Contour source image (d_1) Contour target image

Step 3: *Translation computation*: The proposed scheme computes translation by subtracting the centroid of the target feature image from the centroid of the source feature image, as detailed in Table 4.1.

Table 4.1: Determine translation from centroid of source and target image

Centroid source feature image	Centroid target feature image	t_x	t_y
[126.42619198 127.10039494]	[133.81266718 128.90327034]	1.802875402379641	7.386475209125692

Step 4: *KPCA*: Figure 4.5 (a_1) and Figure 4.5 (b_1) illustrate the first principal direction of both the source and target feature images obtained after applying KPCA.

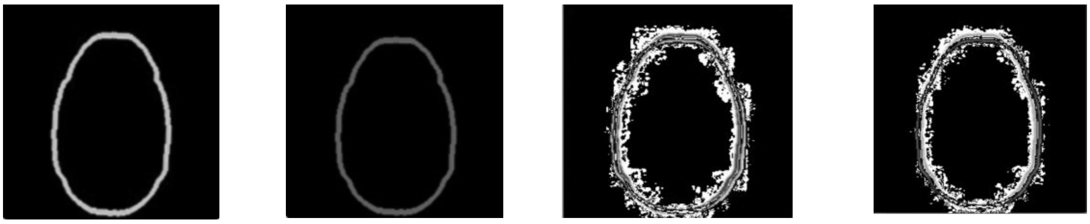
Figure 4.5: (a_1) First principal direction source feature image (b_1) First principal direction target feature image.

Step 5: *Rotation angle computation*: The rotation angle measured in degrees, is determined by calculating the angle between the first principal direction of the source and target feature images, as demonstrated in Table 4.2.

Table 4.2: Determine rotation angle using first principal direction of source and target image

First principal direction of source image	First principal direction of target image	Rotation angle
[-0.6851638, -0.14675629, 0.13050633, 0.70141376]	[-0.66689827, -0.22749482, 0.21968467, 0.67470842]	7.142097993499146 degrees

Step 6: To ensure accurate image alignment, the proposed method employs the TLBO algorithm to optimize registration parameters, including rotation angle and translations as illustrated in Figure 4.6 (a_1) and Figure 4.6 (b_1). The difference image highlights the inconsistency between the target and registered image for PCA and KPCA, as depicted in Figure 4.6 (c_1) and Figure 4.6 (d_1).



(a_1) (b_1) (c_1) (d_1)

Figure 4.6: (a_1) Register image related to PCA (b_1) Register image related to KPCA (c_1) Difference image between reference and register image related to PCA (d_1) Difference image between reference and register image related to KPCA

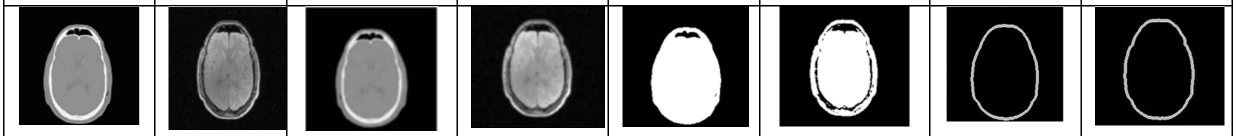
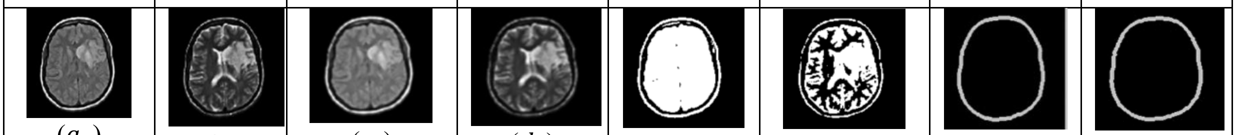
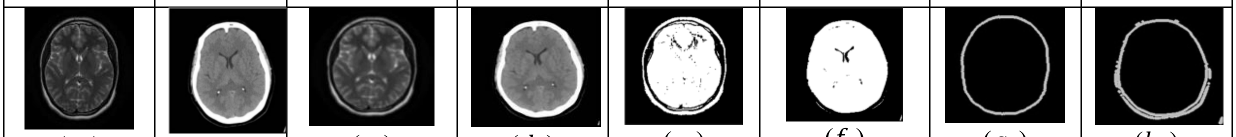
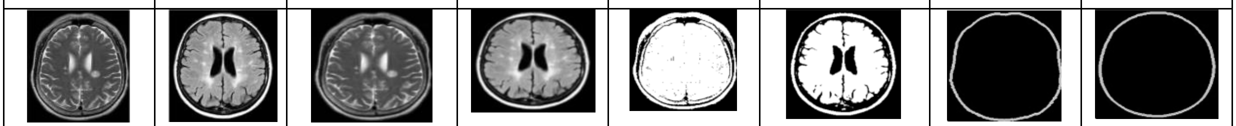
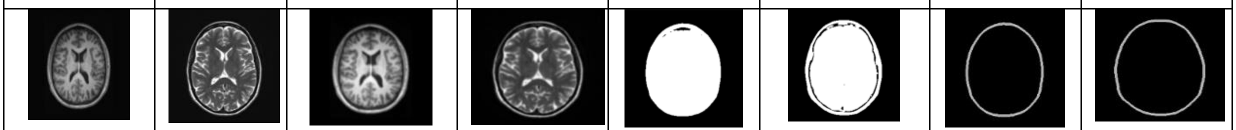
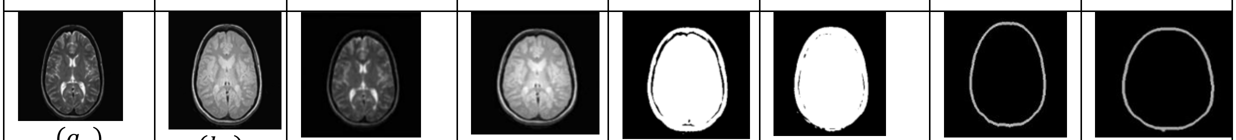
4.3. Experimental Outcomes and Comparative Study

In this section, the effectiveness of the proposed scheme across different datasets [112], [114], [115], [117]–[119], demonstrating its performance through experimental results is presented. Quality assessment metrics RMSE, PSNR, SSIM, and CC are utilized to validate the performance against various modalities of images.

4.3.1. Experimental Outcomes

Numerous experiments are conducted to assess the proposed scheme employing Keras with a TensorFlow backend. These experiments are performed on a system featuring an NVIDIA GeForce GTX GPU, an Intel (R) Core TM i7-1165G7 CPU operating at

2.80 GHz, and 16.0 GB of RAM. The implementation is conducted on the PYTHON 3.9.4 platform, operating within the Windows 11 environment. The proposed approach is validated using Harvard Brain ATLAS [112], Kaggle datasets [114], [115], [117]–[119] and in-house clinical dataset collected from Fortis Hospital Mohali, India including various modalities such as CT-MRI. Evaluation of proposed scheme is performed on medical brain images with dimension (256×256) pixels. First, preprocessing is applied to the source and target images, involving Gaussian filtering and min-max normalization for quality enhancement and ensuring pixel values fall within a specific range, as shown in Figure 4.7 (Column3, c1-c9 and Column4, d1-d9). A threshold of 40, determined via Otsu's thresholding [120], is used to convert grayscale image into binary image as illustrated in Figure 4.7 (Column 5, e1-e9 and Column 6, f1-f9) for CT and MRI images, respectively. Subsequently, a contour tracking method is applied to extract contour details from both the CT and MR images, serving as feature images for registration as visualized in Figure 4.7 (Column 7, g1-g9 and Column 8, h1-h9) for CT and MR images respectively. Following this, translations are determined by calculating the difference between the centroids of the target and source feature images. Then, KPCA is utilized to compute the first principal directions of the feature images, facilitating the calculation of the rotation angle between them, as depicted in Figure 4.8 (Column 1, a1-a9 and Column 2, b1-b9). Finally, the TLBO algorithm is utilized to optimize the registration parameters, including rotation angle and translations, ensuring precise image alignment. This algorithm iteratively refines the parameters to obtain the optimal rotation and translation values. The resulting transformation yields the registered image as depicted in Figure 4.8 (Column 3, c1-c9 and Column 4, d1-d9), aligned with the target image. The difference image highlights any misalignment between the source and target images, with a smaller difference indicating improved alignment for PCA and KPCA, as demonstrated in Figure 4.8 (Column 5, e1-e9 and Column 6, f1-f9) respectively.

Source image	Target image	Pre-processed source image	Pre-processed target image	Threshold source image	Threshold target image	Contour source image	Contour target image
 (a_1)	(b_1)	(c_1)	(d_1)	(e_1)	(f_1)	(g_1)	(h_1)
 (a_2)	(b_2)	(c_2)	(d_2)	(e_2)	(f_2)	(g_2)	(h_2)
 (a_3)	(b_3)	(c_3)	(d_3)	(e_3)	(f_3)	(g_3)	(h_3)
 (a_4)	(b_4)	(c_4)	(d_4)	(e_4)	(f_4)	(g_4)	(h_4)
 (a_5)	(b_5)	(c_5)	(d_5)	(e_5)	(f_5)	(g_5)	(h_5)
 (a_6)	(b_6)	(c_6)	(d_6)	(e_6)	(f_6)	(g_6)	(h_6)
 (a_7)	(b_7)	(c_7)	(d_7)	(e_7)	(f_7)	(g_7)	(h_7)
 (a_8)	(b_8)	(c_8)	(d_8)	(e_8)	(f_8)	(g_8)	(h_8)

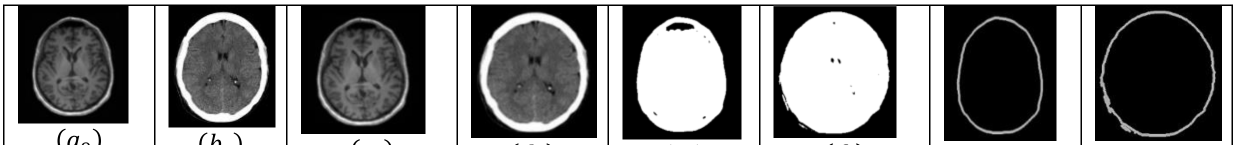
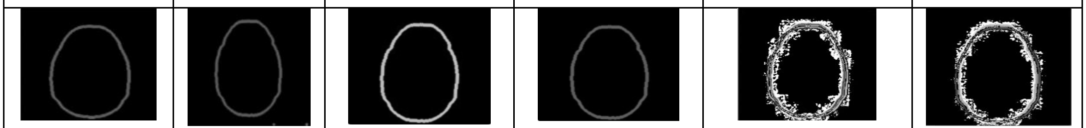
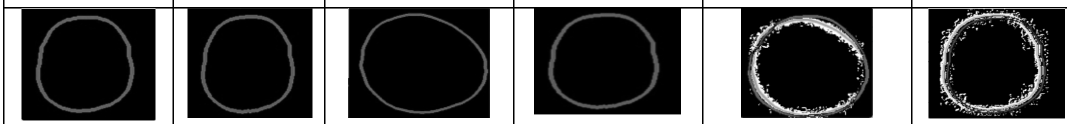
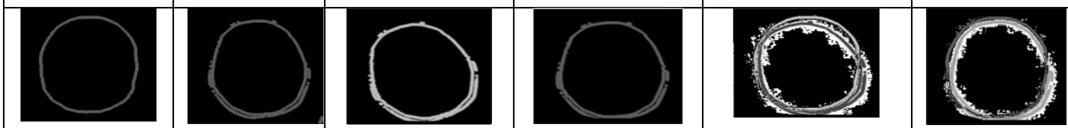
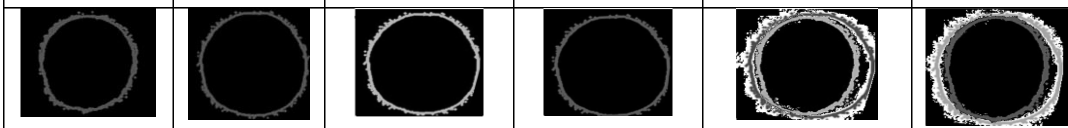
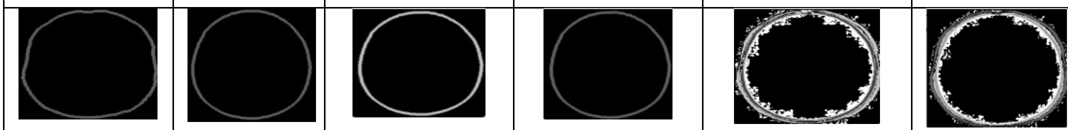
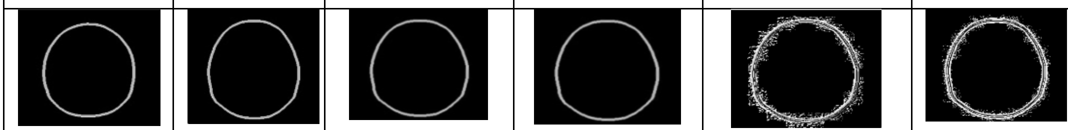
							
(a_9)	(b_9)	(c_9)	(d_9)	(e_9)	(f_9)	(g_9)	(h_9)

Figure 4.7: Evaluation along with multimodal medical images (a) Source image(Fig. a1- Fig. a9) (b) Target image(Fig. b1- Fig. b9) (c) Pre-processed source image(Fig. c1- Fig. c9) (d) Pre-processed target image(Fig. d1- Fig. d9) (e) Threshold source image(Fig. e1- Fig. e9) (f) Threshold target image(Fig. f1- Fig. f9) (g) Contour source image(Fig. g1- Fig. g9) (h) Contour target image(Fig. h1- Fig. h9)

KPCA source feature image	KPCA target feature image	Register image (PCA)	Register image (KPCA)	Difference image (PCA)	Difference image (KPCA)
					
(a_1)	(b_1)	(c_1)	(d_1)	(e_1)	(f_1)
					
(a_2)	(b_2)	(c_2)	(d_2)	(e_2)	(f_2)
					
(a_3)	(b_3)	(c_3)	(d_3)	(e_3)	(f_3)
					
(a_4)	(b_4)	(c_4)	(d_4)	(e_4)	(f_4)
					
(a_5)	(b_5)	(c_5)	(d_5)	(e_5)	(f_5)
					
(a_6)	(b_6)	(c_6)	(d_6)	(e_6)	(f_6)
					
(a_7)	(b_7)	(c_7)	(d_7)	(e_7)	(f_7)

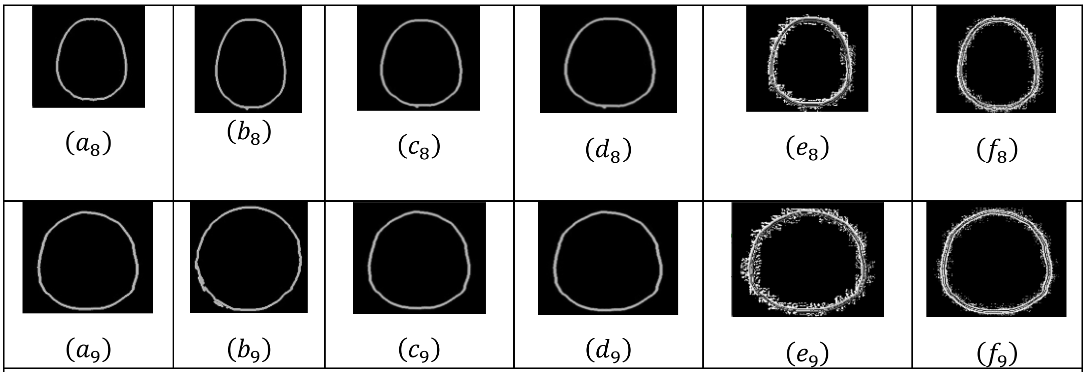


Figure 4.8 Experimental outcomes through multimodal medical images (a) KPCA source feature image (Fig. a1- Fig. a9) (b) KPCA target feature image (Fig. b1- Fig. b9) (c) PCA register image (Fig. c1- Fig. c9) (d) KPCA register image (Fig. d1- Fig. d9) (e) PCA difference image (Fig. e1- Fig. e9) (f) KPCA difference image (Fig. f1- Fig. f9)

For monomodal medical images, pre-processing steps are applied to both the source and target images, as depicted in Figure 4.9 (Column1, a1-a3, and Column2, b1-b3), involving Gaussian filtering and min-max normalization. This enhances image quality and maintains pixel values within a specific range, as displayed in Figure 4.9 (Column 3, c1-c3 and Column 4, d1-d3). Subsequently, feature images are generated through threshold segmentation. In this procedure, the threshold value (θ) is set within the range of 0 to 255, determined through a series of experiments. Intensity values exceeding θ are marked as 1, while those below are assigned 0, as shown in Figure 4.9 (Column5, e1-e3, and Column6, f1-f3). Subsequently, translations are calculated by subtracting the centroids of the target and original feature images. KPCA is utilized to determine the first principal directions of the feature images, followed by the determination of the rotation angle between them as depicted in Figure 4.9 (Column 7, g1-g3 and Column 8, h1-h3). To achieve precise image alignment, the TLBO method optimizes registration parameters, including rotation angle and translations, determining their most effective values. This process generates registered images for PCA and KPCA, depicted in Figure 4.10 (Column 3, c1–c3 and Column 4, d1–d3), aligning with the target image. Variations in the difference images between the target and registered images for PCA and KPCA are illustrated in Figure 4.10 (Column 5, e1-e3 and Column 6, f1-f3) respectively.

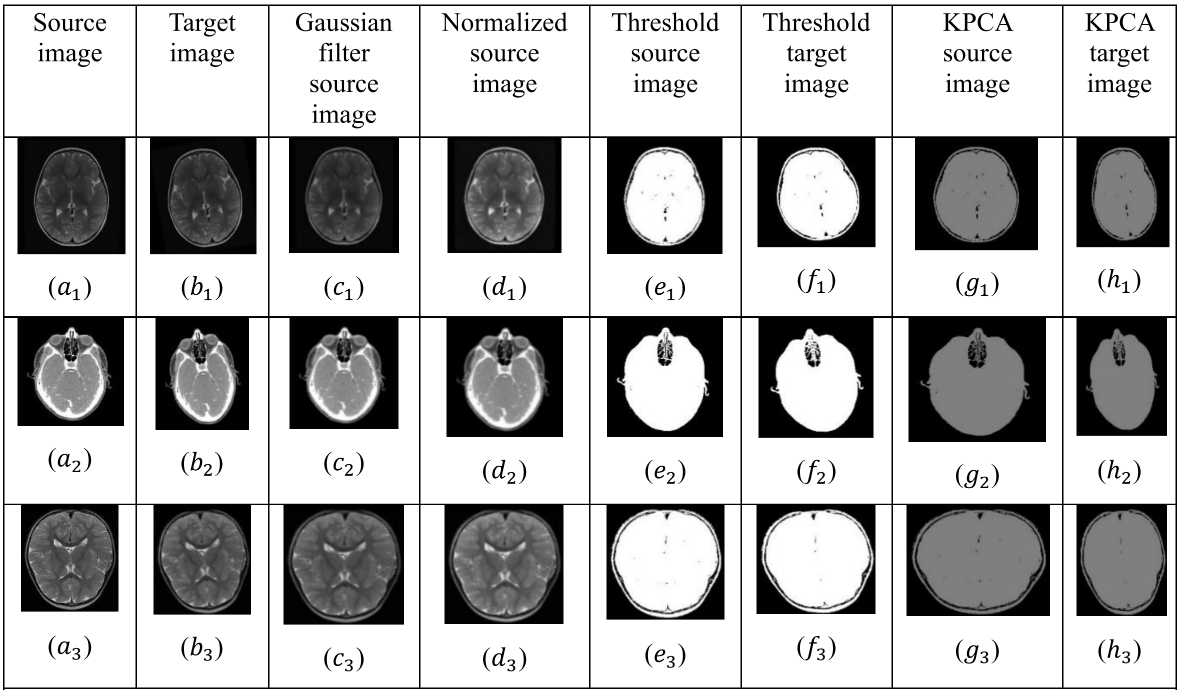


Figure 4.9: Analysis of monomodal medical images (a) Source image(Fig. a1- Fig. a3) (b) Target image(Fig. b1- Fig. b3) (c) Gaussian filter image(Fig. c1- Fig. c3) (d) Normalized image(Fig. d1- Fig. d3) (e) Threshold source image(Fig. e1- Fig. e3) (f) Threshold target image(Fig. f1- Fig. f3) (g) KPCA source image(Fig. g1- Fig. g3) (h) KPCA target image(Fig. h1- Fig. h3)

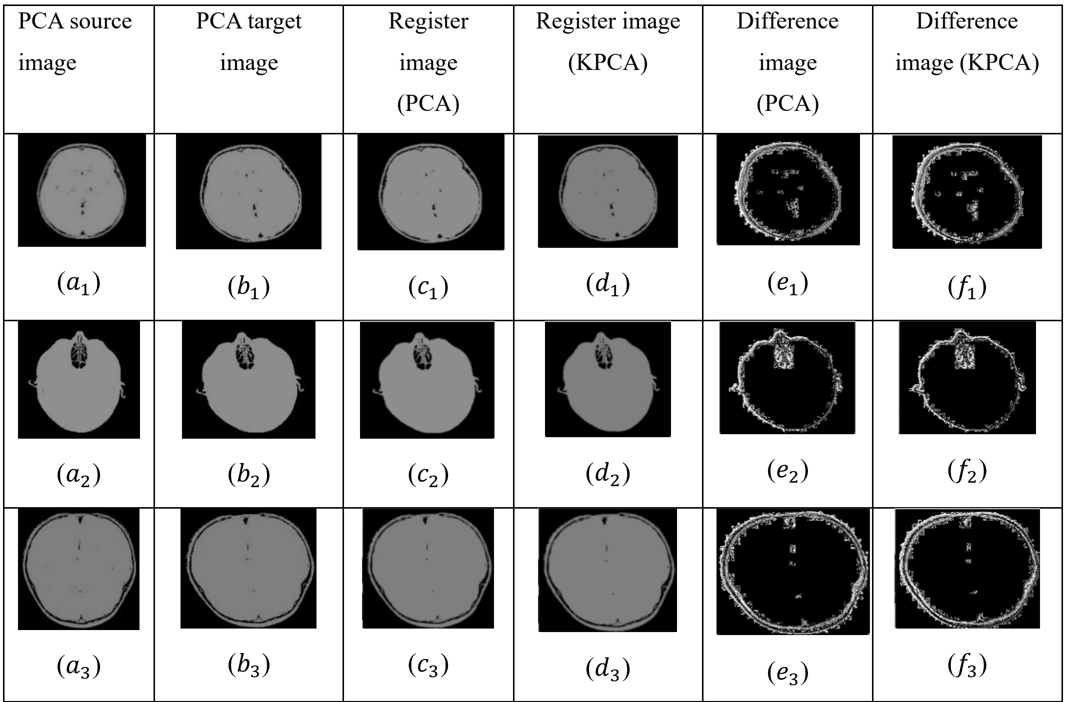


Figure 4.10: Assessment incorporating monomodal medical imaging (a) PCA source image (Fig. a1- Fig. a3) (b) PCA target image (Fig. b1- Fig. b3) (c) PCA register image (Fig. c1- Fig. c3) (d) KPCA register image (Fig. d1- Fig. d3) (e) PCA difference image (Fig. e1- Fig. e3) (f) KPCA difference image (Fig. f1- Fig. f3)

4.3.2. Comparative Analysis

In this section, the effectiveness of proposed scheme for image alignment is compared to the state-of-the-art approaches [1]–[4]. Evaluations are conducted on similar test images, assessing image alignment quality including SSIM, RMSE, PSNR, and CC. Robust image alignment is indicated by high SSIM, PSNR, and CC values, along with low RMSE values. In this study, experiments are carried out using multimodal images by combining PCA and TLBO to verify the efficiency of KPCA with PCA and the results of both with TLBO are shown in Figure 4.11 and Figure 4.12. The results using PCA and KPCA in image registration domain are tabulated in Table 4.3. The results for multimodal images across all test data show that the registered image achieves lower RMSE and higher SSIM values, indicating precise alignment and reduced error rates. Similarly, the superior efficiency of the registered image in terms of SSIM and RMSE across all monomodal test data, with bold values highlighting the performance of the proposed scheme as illustrated in Table 4.4.

Table 4.3: Quantitative multimodal evaluation of proposed approach with PCA based on various quality assessment metric

S. No	Test data	Ground data	Rigid transformation parameters			RMSE	MSE	SSIM
		(θ, t_x, t_y)	θ	t_x	t_y			
1.	CT and MR-T2	(1.7,1.8, 7.3)	-3.438160	0.508370	8.000000	88.7611	7878.54	0.82922
		(7.14,1.8,7.3)	**-5.947361**	**0.489140**	**8.000000**	**44.3892**	**1970.40**	**0.84660**
2.	MR-T1 and MR-T2	(20.23, 13.67, -16.61)	-20.7843	-17.6223	20.37187	70.1808	4925.34	0.79213
		(6.25, 13.67, -16.61)	**-17.00000**	**-17.6223**	**20.37187**	**43.9439**	**1931.06**	**0.80268**
3.	MR-T1 and CT	(23.4, 7.5, 15.7)	-9.000000	-8.76648	-9.00000	105.0097	11027.04	0.718113
		(15.57, 7.50, 15.7)	**-16.0000**	**-3.06**	**-16.00000**	**49.41230**	**2441.575**	**0.768216**
4.	MR-T2 and MR-T1	(11.67, 4.98, 3.47)	-12.58782	-6.65866	-3.585819	122.4458	14992.98	0.662219
		(8.27,4.98,3.47)	**-4.369748**	**6.700710**	**-2.32389**	**64.40342**	**4147.801**	**0.697403**
5.	MR-T2 and MR-Flair	(10.27,0.55, -1.723)	-9.488335	3.516067	3.007994	98.67542	9736.840	0.745325
		(12.52,0.55, -1.723)	**-10.660063**	**4.006762**	**3.494542**	**49.46119**	**2446.409**	**0.7654439**
6.	MR-T1 and MR-T2	(1.064, 1.862, 0.433)	-1.322375	-2.46558	-1.04286	60.4562	3654.952	0.7523
		(2.365, 1.862, 0.433)	**-2.322375,**	**-1.86558**	**-0.44286**	**52.2130**	**2726.197**	**0.7843**
7.	CT and MR-T1	(10.21, -1.64, 9.601)	-11.43121	2.670056	-10.95737	76.4585	5845.902	0.7582
		(13.43, -1.64, 9.601)	**-13.4123**	**3.900567**	**-8.9573**	**53.4821**	**2860.441**	**0.77127**

8.	MR-T2 and MR-PD	(10.35, 2.72, -3.349)	-11.32152	4.31364	-3.52462	58.3425	3403.847	0.7453
		(7.654, 2.72, -3.349)	**-7.65465**	**-3.21776**	**2.91464**	**51.3230**	**2634.050**	**0.76217**
9.	MR-T1 and CT	(9.43, -1.25, 6.1468)	-10.7658	1.05673	-5.3264	64.4034	4147.797	0.72456
		(8.432, -1.25,6.1468)	**-8.454307**	**2.108806**	**-5.89666**	**59.1837**	**3502.710**	**0.75321**

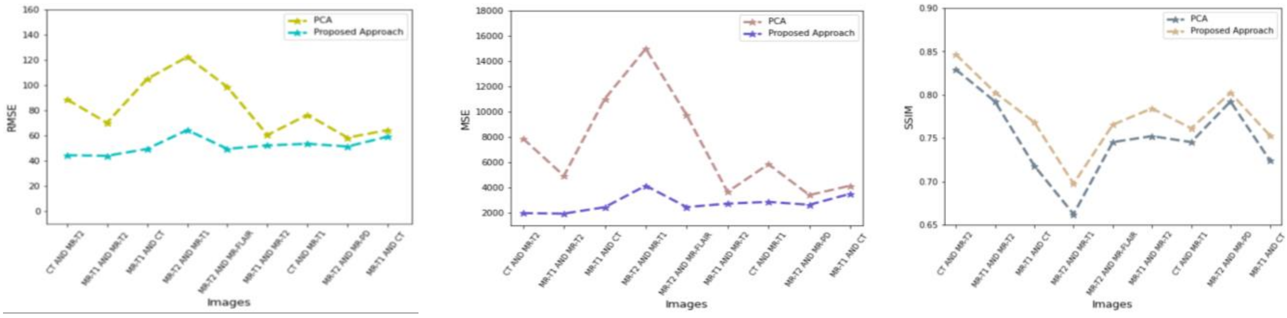


Figure 4.11: Experimental evaluation of proposed method with PCA employing different quality assessment metrics for multimodal medical images.

Table 4.4: Experimental outcomes for three distinct monomodal medical images taken from Harvard dataset[112]

S. No	Test data	Ground data	Rigid transformation parameters			RMSE	MSE	SSIM
		(θ, t_x, t_y)	θ	t_x	t_y			
1.	MR-T1	(10.45,2.78,1.99)	-11.000000	-6.70814	-4.28567	41.10013	1689.220	0.848974
		(17.50,2.78,1.99)	**-17.999014**	**-5.65553**	**-4.362010**	**37.59951**	**1413.723**	**0.852570**
2.	CT	(9.30,3.90,1.41)	-10.000000	-5.711361	-3.24576	49.87451	2487.467	0.838184
		(4.73,3.90,1.41)	**-5.000000**	**-4.741481**	**-2.675474**	**45.39408**	**2060.623**	**0.838848**
3.	MR-T2	(5.10,1.88,1.62)	-3.449475	-2.791043	-2.766178	43.0594	1854.12	0.801655
		(12.95,1.88,1.62)	**-13.000000**	**-3.302510**	**-3.467804**	**42.21010**	**1781.692**	**0.812868**

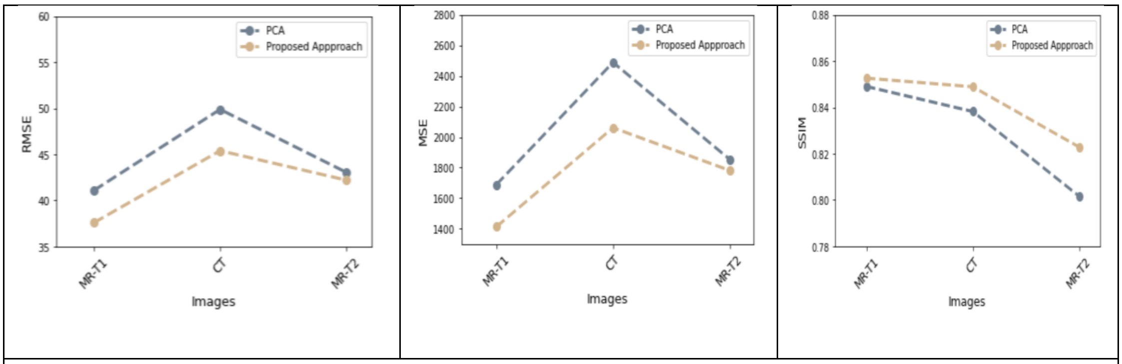
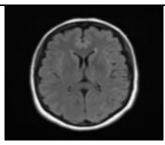
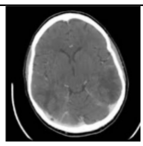
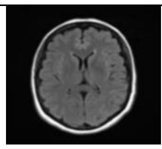
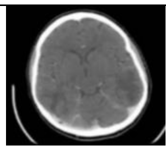
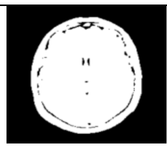
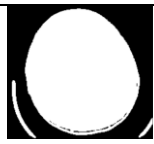
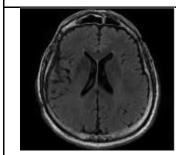
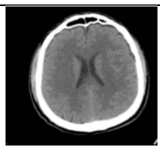
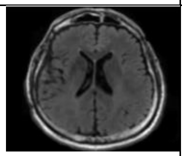
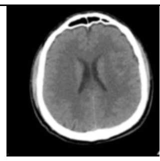
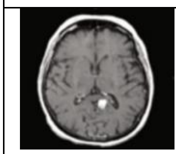
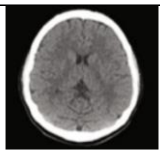
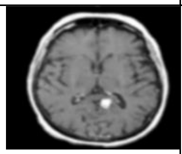
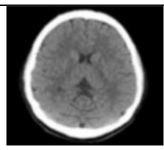
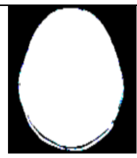
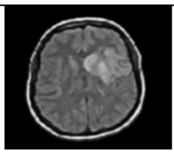
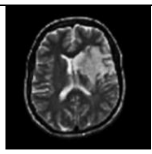
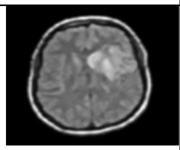
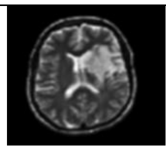
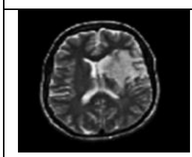
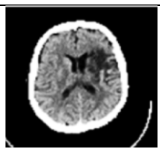
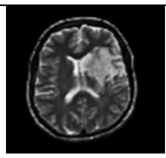
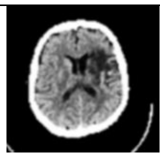
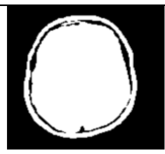
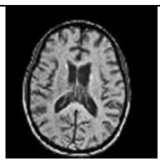
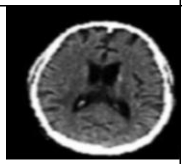
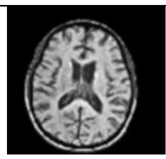
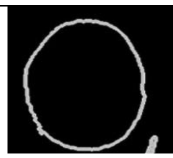
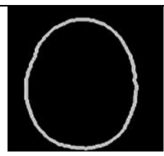


Figure 4.12: Computational evaluation of the proposed approach in comparison to PCA using a variety of similarity measures for monomodal medical images.

S. No.	Source image	Target image	Pre-processed source image	Pre-processed target image	Threshold source image	Threshold target image
1.	 (a_1)	 (b_1)	 (c_1)	 (d_1)	 (e_1)	 (f_1)
	 (a_2)	 (b_2)	 (c_2)	 (d_2)	 (e_2)	 (f_2)
	 (a_3)	 (b_3)	 (c_3)	 (d_3)	 (e_3)	 (f_3)
2.	 (a_4)	 (b_4)	 (c_4)	 (d_4)	 (e_4)	 (f_4)
	 (a_5)	 (b_5)	 (c_5)	 (d_5)	 (e_5)	 (f_5)
	 (a_6)	 (b_6)	 (c_6)	 (d_6)	 (e_6)	 (f_6)

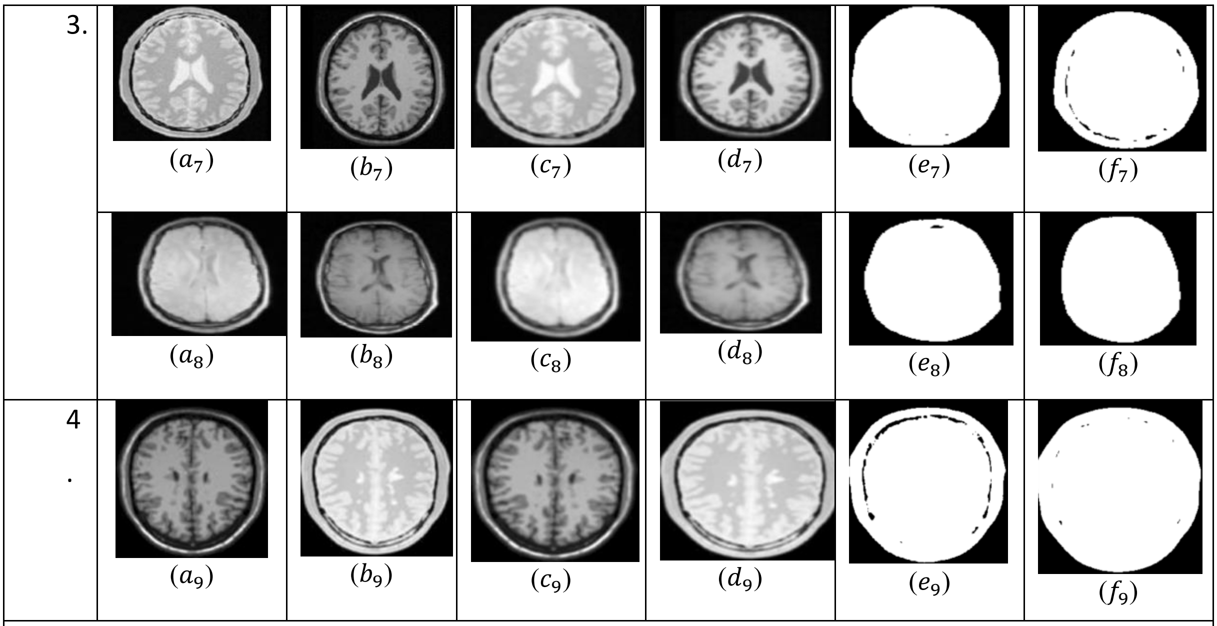
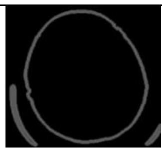
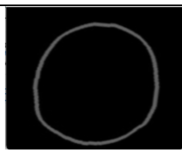
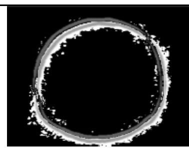
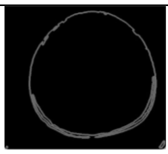
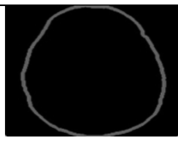
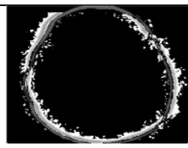
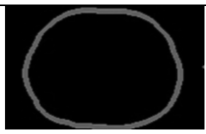
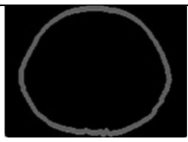
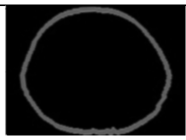
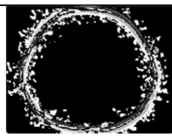
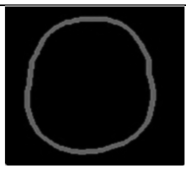
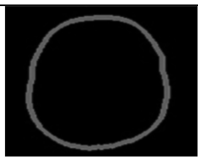
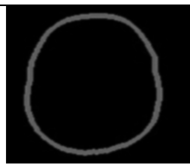
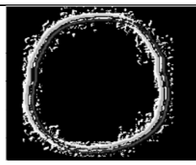
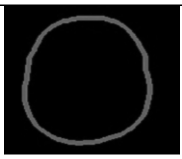
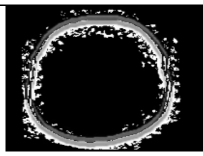
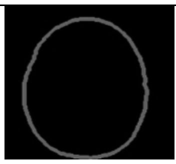
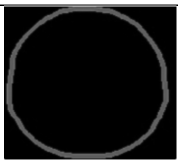
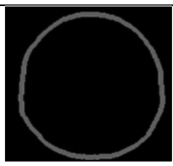
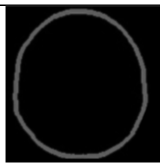
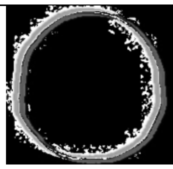
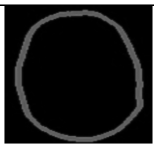
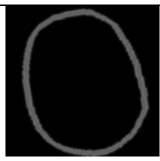
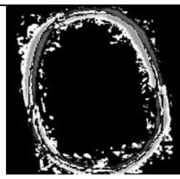


Figure 4.13: Evaluation utilising multimodal imaging (a) Source image (Fig. a1- Fig. a9) (b) Target image (Fig. b1- Fig. b9) (c) Pre-processed source image (Fig. c1- Fig. c9) (d) Pre-processed target image (Fig. d1- Fig. d9) (e) Threshold source image (Fig. e1- Fig. e9) (f) Threshold target image (Fig. f1- Fig. f9)

S. No.	KPCA source feature image	KPCA target feature image	Register image (KPCA)	Difference image (KPCA)
1.	 (a_1)	 (b_1)	 (c_1)	 (d_1)
	 (a_2)	 (b_2)	 (c_2)	 (d_2)
	 (a_3)	 (b_3)	 (c_3)	 (d_3)
2.	 (a_4)	 (b_4)	 (c_4)	 (d_4)
	 (a_5)	 (b_5)	 (c_5)	 (d_5)
	 (a_6)	 (b_6)	 (c_6)	 (d_6)
3.	 (a_7)	 (b_7)	 (c_7)	 (d_7)
	 (a_8)	 (b_8)	 (c_8)	 (d_8)

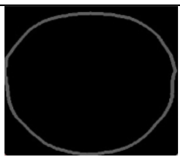
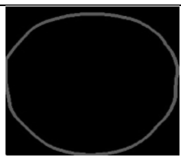
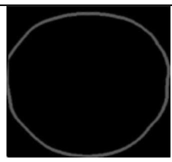
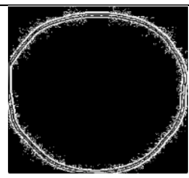
4.	 (a_9)	 (b_9)	 (c_9)	 (d_9)

Figure 4.14: Evaluation utilising multimodal imaging (a) KPCA source feature image (Fig. a1- Fig. a9) (b) KPCA target feature image (Fig. b1- Fig. b9) (c) Register KPCA image (Fig. c1- Fig. c9) (d) Difference KPCA image (Fig. d1- Fig. d9)

The method demonstrated by Arora *et al.* [1] employs SURF-TLBO for multimodal brain image registration. To validate the proposed approach, a comparison is conducted between [1] and the proposed scheme using similar multimodal images from the Harvard dataset [112] as depicted in Figure 4.13 (Column 1: a1-a3, Column 2: b1-b3). Afterwards, source and target images undergo a pre-processing phase, as shown in Figure 4.13 (Column 3: c1-c3 and Column 4: d1-d3) respectively which aims to smooth and enhance the quality of an image. Subsequently, contour extraction is applied to both source and target images to generate feature images and determine translation parameters. The first principal directions are computed from these feature images using the KPCA method, as shown in Figure 4.14 (Column 1: a1-a3 and Column 2: b1-b3) are used to find the rotation parameter of rigid transformation. In addition, TLBO is utilized to optimize the parameters of the rigid transformation, thereby improving the registration process, as depicted in Figure 4.14 (Column 3: c1-c3). The difference image visually evaluates registration quality, indicating misalignment and errors between the target and registered images, with a smaller difference implying better registration as depicted in Figure 4.14 (Column 4: d1-d3). In Table 4.5 and Figure 4.15 (a) show the comparison of the proposed scheme with [1] and PCA-based registration using the SSIM as a similarity measure. From Table 4.5 and Figure 4.15 (a), it is observed that the proposed scheme outperforms both SURF-TLBO [1] and PCA on different image pairs, including MR-T1 and CT, MRI-PD and CT, and MR-T1 and CT. The optimal SSIM values achieved by employing 50 data samples (particles) for TLBO optimization and a maximum of 100 iterations, ensuring accurate alignment of reference and target images.

Table 4.5: Quantitative SSIM-based analysis of proposed approach in comparison to [1] for multimodal images

Images	Similarity metric / Method	SSIM
MR-T1 and CT	SURF+TLBO	0.595064
	PCA+TLBO	0.654481
	Proposed Approach	**0.794376**
MR-PD and CT	SURF+TLBO	0.471868
	PCA	0.736781
	Proposed Approach	**0.772339**
MR-T1 and CT	SURF+TLBO	0.605810
	PCA	0.72653
	Proposed Approach	**0.843610**

The proposed scheme is evaluated against single rigid registration (SRR) and multiple rigid registration (MRR) methods presented by Azam *et al*. [3]. To assess their performance, several multimodal CT and MRI brain images, each with dimension (256×256), were obtained from [121], as depicted in Figure 4.13 (Column 1: a4-a6) and Column 2: b4-b6). Then, contour extraction is applied to both source and target images to generate a feature image. The rotation angle is determined by comparing the first principal directions of the source and target images as shown in Figure 4.14 (Column 1: a4-a6 and Column 2: b4-b6). The proposed scheme employs TLBO to determine the optimal values for rigid transformation parameters, including rotation angle and translations, aiming for accurate image alignment, as depicted in Figure 4.14 (Column 3: c4-c6). The difference image visually represents the misalignment between the reference and registered images, as shown in Figure 4.14 (Column 4: d4-d6). In Table 4.6 and Figure 4.15 (b-d), the proposed scheme is compared with [3] using SSIM, PSNR, and cross-corelation as a similarity measure. The results demonstrate that proposed scheme achieves better outcomes across different image pairs, including MR-PD and MR-T2, MR-T2 and CT, and CT and MR-T1.

Table 4.6: Comparative analysis between proposed approach and [3] leveraging several quality assessment metrics

Images	Similarity metric / Method	SSIM	PSNR	Cross correlation
MR-PD and MR-T2	SRR	0.7945	34.5266	0.8998
	MRR	0.7997	50.7112	0.9073
	Proposed Approach	**0.86511**	**57.3883**	**0.93203**
MR-T2 and CT	SRR	0.4755	56.22	0.47709
	MRR	0.5252	53.35	0. 66671
	Proposed Approach	**0.7679**	**60.2503**	**0.87928**
CT and MR-T1	SRR	0.5287	34.8154	0.4232
	MRR	0.5799	45.680	0.5568
	Proposed Approach	**0.90017**	**69.6563**	**0.7954**

The method presented by Su *et al.* [4], which employs a multimodal brain image registration technique using optical flow and SURF is contrasted with proposed scheme. The process involves preprocessing, contour extraction, and feature image creation from reference and target images. Principal directions are determined using KPCA, followed by optimization of rigid transformation parameters using TLBO method. The difference image highlights inconsistencies between reference and registered images. The performance of the proposed scheme is compared with [4] employing SSIM, PSNR, and cross correlation as a similarity measure across different image pairs CT and MR-T1, CT and MR-T2 as illustrated in Table 4.7 and Figure 4.15 (e-f).

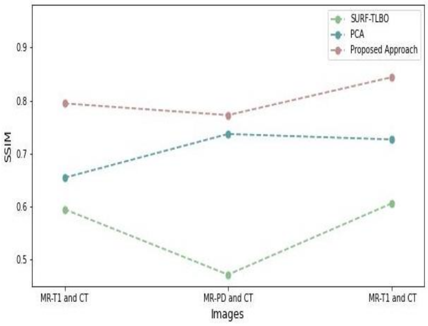
Table 4.7: Performance evaluation of proposed approach with [4] based on PSNR and cross correlation metrics for multimodal medical images.

Images	Similarity metric / Method	PSNR	Cross correlation
CT and MR-T1	Demons	53.83	0.77
	SURF	63.11	0.83
	Proposed Approach	**83.123**	**0.9921**
CT and MR-T2	Demons	56.02	0.91
	SURF	68.26	0.79
	Proposed Approach	**76.079**	**0.9578**

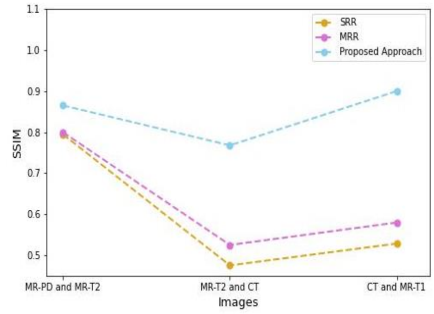
The proposed scheme undergoes comparison with the method discussed by Zheng *et al*. [49], which employs Progressive-SURF. For fair comparison, similar source and target images with dimensions of $(256 * 256)$ from the dataset [114] are considered as illustrated in Figure 4.13 (Col.1, a9) and Figure 4.13 (Col.2, b9). Contour extraction generated feature images from both source and target images, followed by computing their first principal directions using KPCA as illustrated in Figure 4.14 (Col. 1, a9) and Figure 4.14 (Col. 2, b9). The TLBO optimizes rigid transformation parameters such as rotation angle and translations to achieve accurate image alignment, as shown in Figure 4.14 (Col.3, c9). The difference image highlights variations between the reference and registered images as illustrated in Figure 4.14 (Col.4, d9).The performance comparison between the proposed scheme and [49], utilizing MSD and cross-correlation as similarity measures across image pair such as MR-T1 and CT, is presented in Table 4.8 and Figure 4.15 (g).

Table 4.8: Proposed approach comparison of proposed approach with [49] employing MSD and cross correlation on multimodal medical images for evaluating performance.

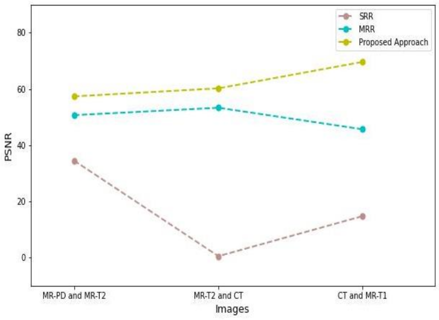
Images	Similarity metric / Method	MSD	Cross correlation
MR-T1 and CT	Demons	0.1054	0.8389
	SURF	0.0197	0.8142
	P1-SURF	0.0082	0.9219
	Proposed Approach	**0.00423**	**0.9421**



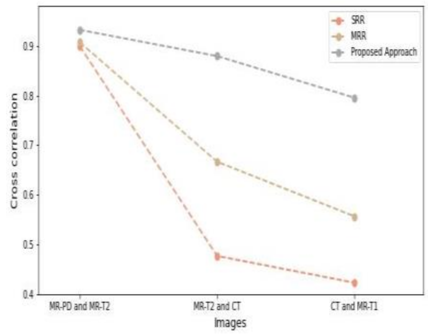
(a)



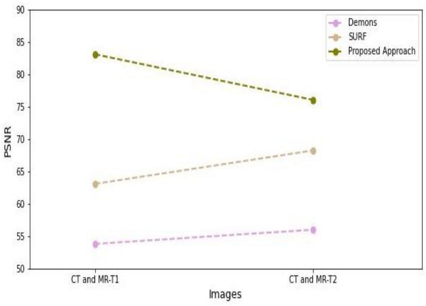
(b)



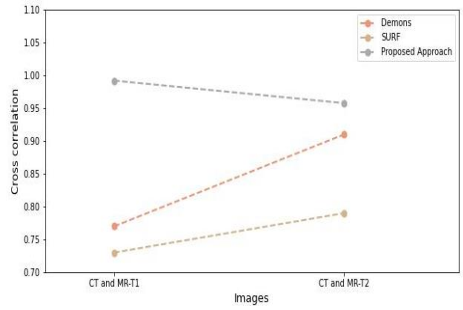
(c)



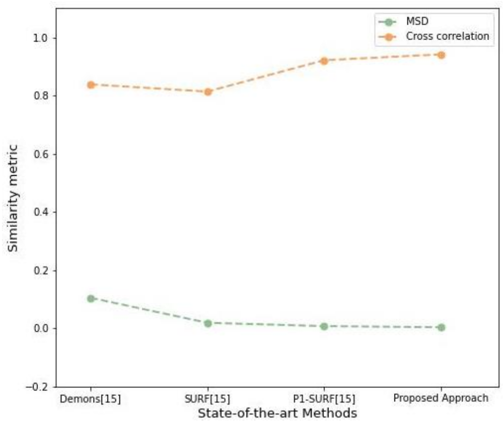
(d)



(e)



(f)



(g)

Figure 4.15: Comparative computational analysis of the proposed method with state-of-the-art approaches[1]–[4] which employ numerous kinds of quality assessment metrics on multimodal medical images.

4.4. Chapter Summary

In this chapter, a novel approach with the hybridization of KPCA and TLBO for both monomodal and multimodal medical image registration is discussed. The approach starts with pre-processing steps such as noise removal and normalization to enhance image quality. In the next step, contour extraction is performed using Otsu's thresholding for multimodal images and a thresholding segmentation approach for monomodal images to create feature images and determine ground truth translation parameters. Then, KPCA captures non-linear data mapping for rotation value determination, while TLBO optimizes the registration process by iteratively adjusting transformation parameters. The proposed scheme has been evaluated using various similarity metrics for both monomodal and multimodal medical images from publicly available and in-house clinical datasets. This chapter focuses on rigid medical image registration on both standard as well as real-time medical image datasets but however lacking in deformable image registration. In the next chapter, a deformable medical image registration leveraging unsupervised learning algorithm and visual geometry group (VGG19), a deep learning framework to enhance the accuracy and efficiency for both 2-D and 3-D monomodal and multimodal images will be elaborated.

CHAPTER 5 Deformable and Rigid Medical Image Registration on different modalities using the combination of Deep Learning and Meta-Heuristic Approaches

5.1. Introduction

In this chapter, an integration of deep learning approaches including U-shaped convolution neural network (U-Net) and VGG-19 with meta-heuristic techniques including TLBO and unified equilibrium optimizer (UEO) is outlined. Firstly, an unsupervised learning algorithm U-Net model is integrated with TLBO for the registration of monomodal image modalities. In this scheme, a two-channel image created from moving and fixed images are input into the U-Net model encoder and decoder phases to extract image features and generate a displacement field. Moreover, B-spline transformation parameters such as control point positions and their associated coefficients are predicted resulting in the warped image. To enhance the quality of the warped image, TLBO with rigid transformation parameters is applied to both fixed and U-Net warped images to detect the optimum transformation parameters (θ, t_x, t_y). This scheme is evaluated on five publicly available datasets [122]–[126] of 2-D and 3-D mono-modal medical brain MR and chest CT images, employing transfer learning with pre-trained weights from voxel morph U-Net model for 3-D images. It demonstrated superior results in deformable medical image registration, with dice scores increasing from 0.7421 to 0.9542 and from 0.7102 to 0.9024, and Hausdorff distances decreasing from 2.0521 to 1.5607 and from 2.04463 to 1.0923, outperforming symmetric image normalization (SyN) and voxel morph.

This chapter is structured as follows. Firstly, an unsupervised learning algorithm U-Net model integrated with TLBO and affine transformation for monomodal medical images is presented in Section 5.2. The performance of proposed scheme is examined in Section 5.2, evaluated with dice score and hausdorff distance as presented in experimental results Subsection 5.2.1.1. Then, in Subsection 5.2.1.2, the effectiveness of proposed scheme is assessed through a fair comparison with recent state-of-the-art techniques using monomodal image modalities. The conclusion outlining its benefits and limitations is demonstrated in Section 5.2.2. Secondly, to address the limitations discussed in Section 5.2, a deep learning method VGG19 integrated with UEO optimization for monomodal and multimodal medical image modalities is elaborated in Section 5.3 and its algorithm is discussed in Section 5.3.1. Experimental results along with comparative analysis with state-of-the-art techniques in terms of robustness are detailed in Subsections 5.3.2.1 and 5.3.2.2 respectively. The summary of the chapter including both advantages and limitations is discussed in Section 5.4.

5.2. U-Net with TLBO: A Deformable Medical Image Registration

The objective of deformable image registration is to identify the optimal displacement field (∅) over an n-dimensional spatial domain $\Omega \subset R^n$, where R^n is Euclidean space to ensure the warped image is aligned precisely with fixed image. Image registration is characterized by the minimization of a cost function, as formulated in Equation. (5.1), which aims to achieve optimal alignment between images. The loss function used to quantify alignment quality is mathematically expressed in Equation (5.2).

$$\widehat{\emptyset} = \arg\min_{\emptyset} \mathcal{L}_{loss}\left(I_{fix}, I_{mov}, \emptyset\right) \qquad (5.1)$$

$$\mathcal{L}_{loss}\left(I_{fix}, I_{mov}, \emptyset\right) = \mathcal{L}_{sim}\left(I_{fix}, I_{mov} \circ \emptyset\right) + \lambda\, \mathcal{L}_{reg}(\phi) \qquad (5.2)$$

In the context of deep learning methods, the deformation field is represented as $\emptyset = v_\theta\left(I_{mov}, I_{fix}\right)$, where θ denotes the trainable parameter. The term λ, a user-defined weight, adjusts the impact of $\mathcal{L}_{reg}$, which serves as a regularizer, while $\mathcal{L}_{sim}$ stands for the similarity loss.

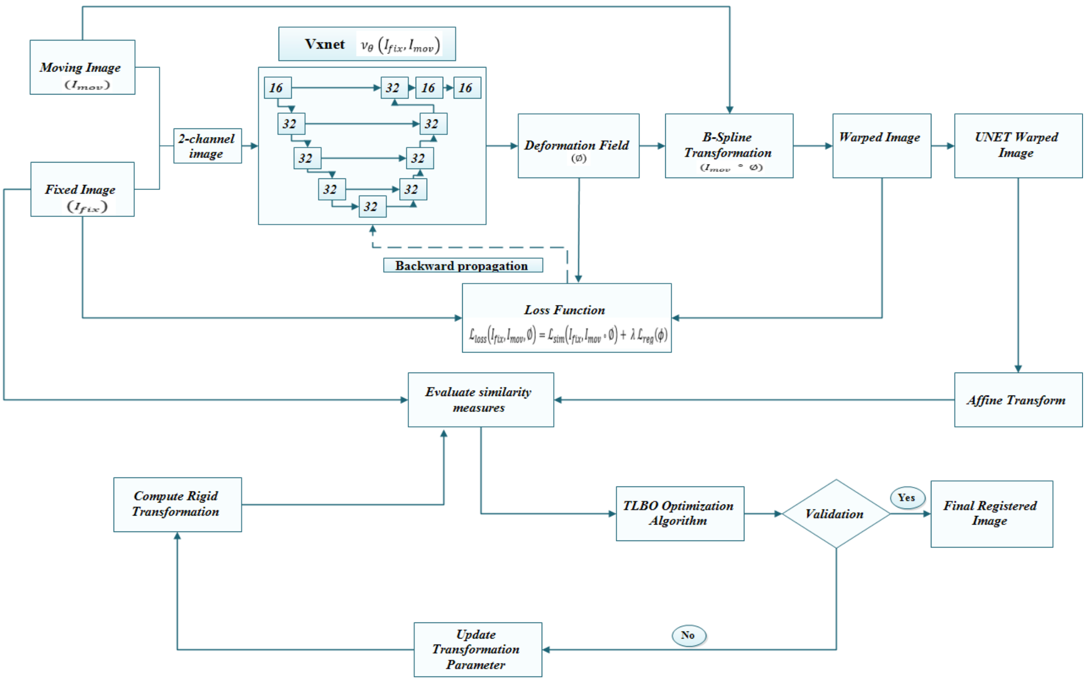


Figure 5.1: Workflow of Proposed Approach

This work includes 2-D and 3-D medical MR and CT image modalities from datasets [122]–[127]. The proposed scheme utilizes the voxel morph CNN architecture (Vxnet) with parameter θ, expressed as $v_\theta\left(I_{mov}, I_{fix}\right) = \emptyset$, to compute the displacement field ($\emptyset$) using image pairs I_{mov} and I_{fix}. Figure 5.1 depicts the workflow diagram of proposed scheme which consists of voxel morph CNN architecture for predicting the deformation field (registration field) and a B-spline spatial transformer to achieve the U-Net warped image with minimized loss function. The detailed steps are provided as follows:

Step 1: Voxel morph CNN architecture

Voxel morph (Vxnet) CNN architecture employs U-Net structure to extract image features through an encoder-decoder network with a bridge, creating a deformation field ($\emptyset$) based on fixed $\left(I_{fix}\right)$ and moving (I_{mov}) images. This architecture has an encoding path that accepts a single 2-channel input image and includes a convolution block with Leaky ReLU activation, as shown in Figure 5.3 and Figure 5.4. Afterwards, max pooling is employed for down-sampling, which reduces the spatial resolution of the feature maps from the convolutional layers by half. The base block (bridge) shown in Figure 5.4 includes only a convolution block. The decoder block needs three inputs: the output from the previous block, a skip connection or features from the corresponding encoder block, and the number of filters, as illustrated in Figure 5.3. The last layer of the decoder phase uses a convolution block to convert the combined three inputs into the deformation field, as depicted in Figure 5.1. Algorithm 5.1 provides a comprehensive outline of the CNN architecture, deformation field, and spatial transformation.

Step 2: **B-spline transformation**

The B-spline transformation technique aligns images using B-spline interpolation, which determines the values at transformed locations based on neighbouring control points through B-spline basis functions. The main goal of this transformation is to compute $(I_{mov} \circ \emptyset)$, applying a nonlinear warp to the moving image with the displacement field (ϕ). The displacement vector flow field (u) defines the warp by correlating pixel intensity between the output and moving image. Consequently, the grid spacing between control points in the flow field (u) is one, and the number of control points matches the image size. The intensity $\left(I_{mov} \circ \emptyset(px)\right)$ for each pixel position (px) in the output image is depicted in Equation (5.3)

$$\left(I_{mov} \circ \emptyset(px)\right) = \left(I_m \left(px - u(px)\right)\right) \quad (5.3)$$

where, $\left(p - u(px)\right)$ is non-integer and pixel intensity is defined only at integer positions, $\left(I_m \circ \emptyset(px)\right)$ is found by interpolating the nearest pixel associated with $\left(px - u(px)\right)$, as represented in Equation (5.4).

$$\left(I_{mov} \circ \emptyset(px)\right) = \sum_{w \in \mathbb{Z}(px')} I_m(w) \prod_{j \in \{a,b\}} t\left(px'_j - w_j\right) \quad (5.4)$$

Here, $px' = \left(px - u(px)\right)$, and $\mathbb{Z}(px')$ represents the neighbouring pixel of px'. The variable j iterates over the spatial dimensions of the domain Ω and $t()$ denotes a generic sampling kernel used for image interpolation. This process ensures the spatial transformation function is differentiable, enabling the backpropagation of error gradients (derivatives of the loss function that guide parameter adjustments) during optimization for effective learning.

***Step 3*: Loss Function**

The loss function $(\mathcal{L})$, as outlined in Equation (5.5), incorporates both the similarity measure $(\mathcal{L}_{sim})$ and the regularization function $\left(\mathcal{L}_{reg}\right)$.

$$\mathcal{L}_{loss}\left(I_{fix}, I_{mov}, \emptyset\right) = \mathcal{L}_{sim}\left(I_{fix}, I_{mov} \circ \emptyset\right) + \lambda \, \mathcal{L}_{reg}(\phi) \quad (5.5)$$

where, $\mathcal{L}_{sim}$denotes the similarity loss, while $\mathcal{L}_{reg}$ refers to the regularization applied to the deformation field. The displacement field is represented as $\emptyset = v_\theta\left(I_{fix}, I_{mov}\right)$, where θ is a learnable parameter and λ is a weighted parameter that adjusts the impact of $\mathcal{L}_{reg}$. The parameter v_θ is updated by backpropagating the loss. This work employs MSE as the similarity measure, which calculates the pixel-wise difference between images as outlined in Equation (5.6).

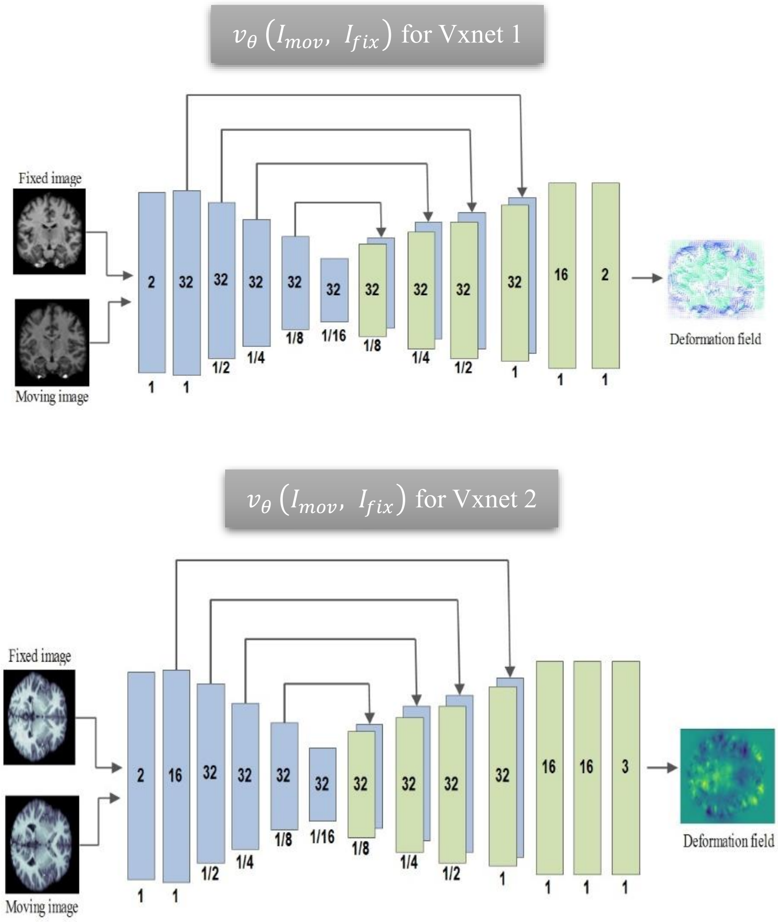
$$MSE\left(I_{fix}, I_{mov} \circ \emptyset\right) = \frac{1}{|\Omega|} \sum_{p \in \Omega} \left[I_{fix}(p) - \left[I_{mov} \circ \emptyset(p)\right]\right] \quad (5.6)$$

In the spatial domain Ω, images must have similar intensity distributions to be considered identical. Minimizing $\mathcal{L}_{sim}$ encourages the warped image $(I_{mov} \circ \emptyset)$ to resemble I_{fix}, which can cause a discontinuous $\emptyset$. Therefore, Equation (5.7) includes an additional regularization term that penalizes non-smooth deformation fields $(\emptyset)$ through the norm of their spatial derivatives.

$$\mathcal{L}_{reg}(\emptyset) = \sum_{p\in\Omega}\|\nabla\emptyset(p)\|^2 \tag{5.7}$$

The loss function includes this regularization term to produce anatomically consistent outcomes, thereby avoiding sharp discontinuities in the predicted displacement field ($\emptyset$).

The Voxelmorph (Vxm) framework includes Vxnet1 for 2-D datasets and Vxnet2 for 3-D datasets, both utilizing a U-Net structure. Figures 5.2(a) and 5.2(b) depict these structures as rectangles, representing 2-D and 3-D volumes. The numbers within each rectangle show the number of channels in an n-dimensional volume n$\in$ {2,3} and the corresponding spatial resolution of the input. During encoding, each channel employs max pooling operations to down-sample feature maps, reducing their spatial resolution by half. In the decoding phase, each channel uses 32 filter convolutions and an up-sampling layer, restores the original resolution. Arrows represent skip connections between encoder and decoder features. The final decoder layer converts the outcome into a deformation (displacement) field with a convolution block.



a.

b.

Figure 5.2:Voxelmorph architecture $v_\theta\,(I_{mov},\ I_{fix})$ to generate deformation field. (a) for 2D dataset (b) 3D dataset

Algorithm 5.1: Vxm CNN Architecture Registration	
Input: Initialize fixed and moving image as I_{fix} and I_{mov} respectively	
Output: To compute the loss and deformation(registration) field.	
1. Procedure Vxm_CNN_WarpImg(I_{fix}, I_{mov})	■ $Input\ I_{fix}\ and\ I_{mov}$
2. $v_\theta = Initializer(Vxnet)$	
3. while $j < no\_of\_iteration$ do	
4. $\emptyset = v_\theta(I_{fix}, I_{mov})$	■ $Compute\ registration\ field\ \emptyset$
5. $\omega = I_{mov} \circ \emptyset$	■ $Warped\ \ image\ (\omega)$
6. $\mathcal{L}_{loss}(\omega, \emptyset) = \mathcal{L}_{sim}(\omega, I_{fix}; \emptyset) + \lambda\, \mathcal{L}_{reg}(\phi)$	■ $Compute\ loss$
7. $v_\theta = BackPropogate\_to\_Converge(\, v_\theta, \mathcal{L}_{loss}\,)$	■ $Update\ Vxm\_CNN$
8. $j = j + 1$	
9. return ω , $\emptyset$	

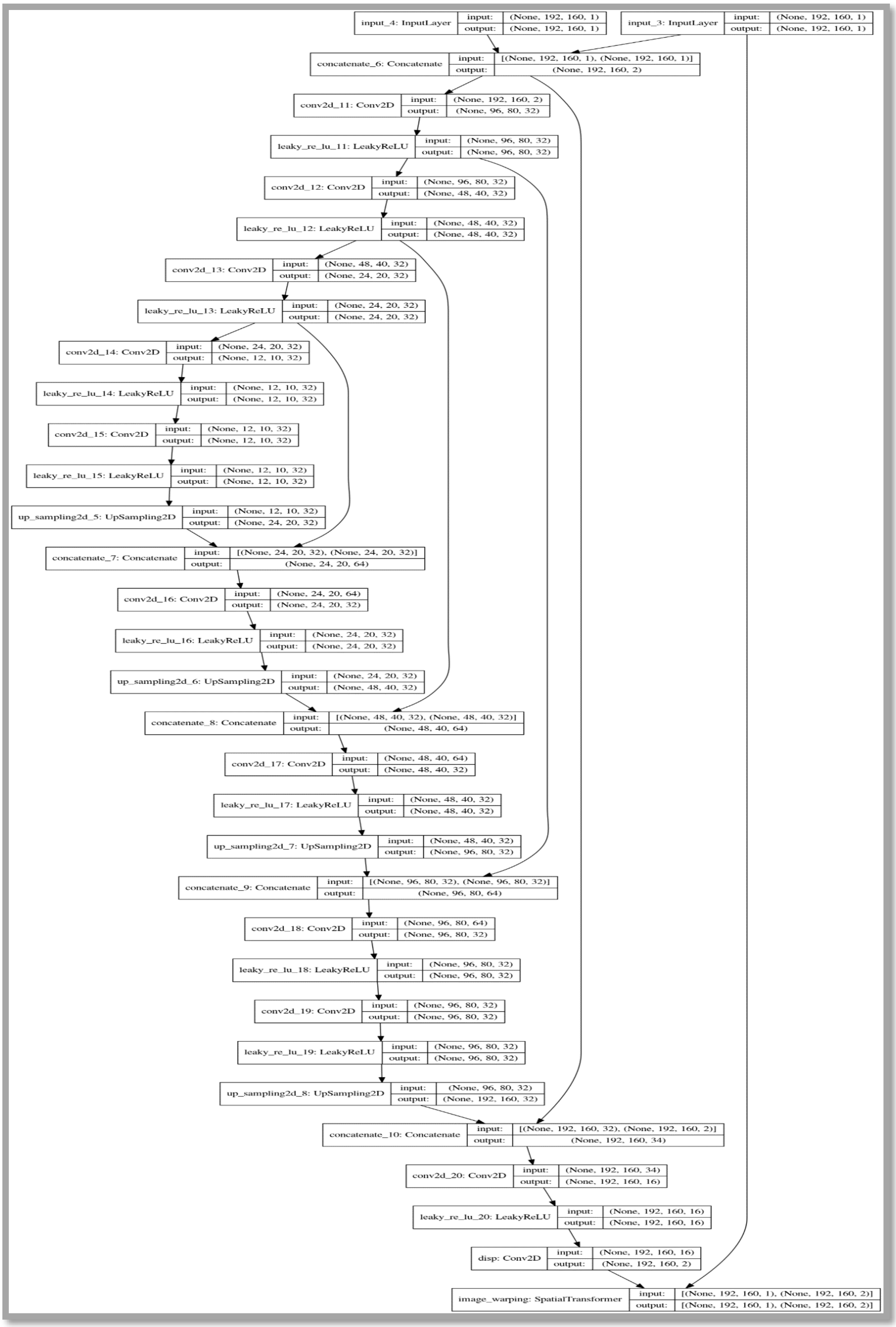


Figure 5.3: Structure representation of 2-D U-Net architecture.

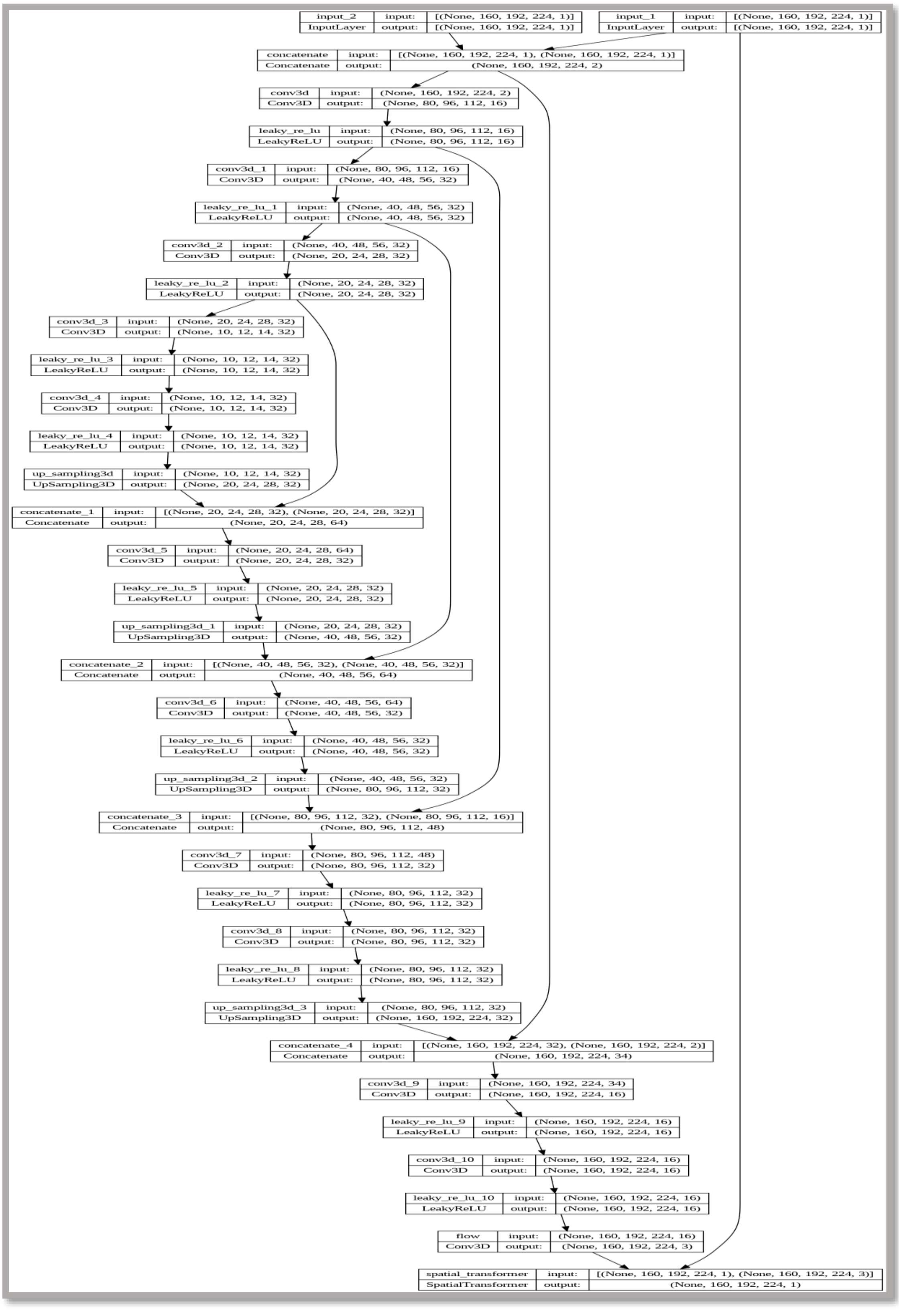


Figure 5.4: Structure representation of 3-D U-Net architecture

5.2.1. Experimental Results and Comparison Analysis

The effectiveness of the proposed scheme across different datasets [122]–[127] is demonstrated through the experimental and comparison results by considering dice score and Hausdorff distance as a similarity measure.

5.2.1.1. Experimental Results

To evaluate the proposed scheme, experiments are conducted using Keras with TensorFlow backend on an NVIDIA GEFORCE GTX GPU. The implementation is carried out on a 2.80 GHz Intel(R) Core (TM) i7-1165G7 CPU with 16.0 GB RAM, running Python 3.9.4 on Windows 11. The unsupervised Voxelmorph CNN for deformable medical image registration is trained with a batch size of 1, learning rate of 10^{-4}, and a regularization parameter (λ)=0.25. Moreover, the Adam optimizer is employed to optimize the learning model. Performance evaluation of the proposed registration model is carried out using five different datasets namely D1-D5: Atlas_norm_2D [122], 2D oasis [123], Atlas_norm_3D [126], ADNI [124] , CT scan: Before and After[127]. The 2-D dataset includes the ATLAS and 2D OASIS datasets, comprising 349 and 414 MRI images with dimensions 256×256 pixels. These images are cropped to 192×160 which reduces the training and computational time of Voxelmorph CNN architecture. For the ATLAS dataset, 208 images are used for training and 141 for validation, whereas the OASIS dataset consists of 138 images for training, validation, and testing individually. The model is trained using MSE similarity metric with a regularization parameter λ= 0.25 over 500 epochs. Intermediate results of the loss function after 100 and 500 epochs for the Atlas_norm_2D (D1) and 2D OASIS (D2) datasets are depicted in Figure 5.5 (a) and Figure 5.5 (b), respectively, with λ= 0.25. Accurate alignment of moving and fixed images is determined by exploring the value of λ ranging from 0 to 1 as depicted in Figure 5.7 through various experiments. The 3-D datasets Atlas Norm 3D and ADNI contain 320 and 1961 T1-weighted MRI brain images respectively, while CT scan includes 514 chest CT images sized at $256 \times 256 \times 256$. Similarly, the 3-D datasets are down-sampled to $160 \times 192 \times 224$ for computational efficiency. Experiments are conducted using pre-trained weights of the Vxnet CNN architecture on these 3-D datasets. For experimentation mono-modal medical images, including MRI image (Figure 5.8 (a1-a4, Col. 1)) and CT image (Figure 5.8 (a5, Col. 1)) are used. Two randomly selected images moving and fixed from the training dataset are shown in

Figure 5.8 (a1-a5, Col. 1) and Figure 5.8 (b1-b5, Col. 2) respectively. These images are concatenated into two-channel input for the Vxnet architecture, resulting in a deformation field as visualized in Figure 5.8 (c1-c5, Col. 3). The CNN output is evaluated on the test dataset, resulting in a vector representation of a moving image called deformation field is obtained from the trained model. This deformation field undergoes transformation through three components of B-spline transformation, i.e., localization network, grid generator, and sampler. These components process the deformation field and features of the moving image to generate the warped image as shown in Figure 5.8 (d1-d5, Col. 4). The localization network predicts spatial transformation parameters including rotation, scaling, and translation. The grid generator creates control points that define the deformation field. Through B-spline interpolation, the sampler applies this deformation field to the input feature map producing a transformed output feature map. However, the alignment of the warped image produced by Vxnet is suboptimal. Therefore, TLBO with rigid transformation parameters (RTP) is applied to enhance the alignment between CNN-generated warped and fixed image. By incorporating TLBO with the output (warped image) from Vxnet for alignment with the fixed image can effectively enhance the registration process. This integration results in a significant enhancement in the quality of the registered image as illustrated in Figure 5.8 (e1-e5, Col. 5). High value of dice score and low value of Hausdorff distance reveal the improved accuracy of the registration process.

Initially, the warped image is aligned with fixed image by rotating it at an angle $\theta = 2^{\circ}$ and then translating it using $T = (t_x, t_y) = (4,8)$. TLBO with a population of 100 particles and 500 iterations is utilized to determine the optimal value of rigid transformation parameter (RTP). This determination involves evaluating different objective functions: (i) Dice score, (ii) MI, and (iii) Hausdorff distance as specified in Table 5.1. These experiments include various image modalities defined in Equation (5.8) to Equation(5.10). The final registered image is created using effective translation and rotation parameters as depicted in Figure 5.8 ((e1), Col. 5). This process is applied to various test scenarios by altering the translation and rotation parameters as illustrated in Figure 5.8 ((e2-e6), Col. 5). The significant reduction in the difference between fixed and final registered images indicates that the deformation field effectively addresses misalignment as illustrated in Figure 5.8 ((f1-f5), Col. 6). The flow field (displacement vector field) generated by a spatial transformer illustrates the direction and magnitude of displacement at each pixel as shown in Figure 5.8 ((g1-g5), Col. 7). Experimental results show that the proposed scheme outperforms the state-of-the-art (SOTA) method in deformable medical image registration by more than 20% in dice score and less than 0.9% in Hausdorff distance (HD), as tabulated in Table 5.1.

$$MI = max\left(MIF\left(I_{fix}, \omega\right)\right) \tag{5.8}$$

$$Dicescore = max\left(DS\left(I_{fix}, \omega\right)\right) \tag{5.9}$$

$$Hausdorffdistance = min\left(HSD\left(I_{fix}, \omega\right)\right) \tag{5.10}$$

Here, ω represents the warped image, denoted as $I_{mov} \circ \emptyset$, whileI_{fix} is the fixed image. MIF stands for mutual information, DS for dice score, and HSD for hausdorff distance.

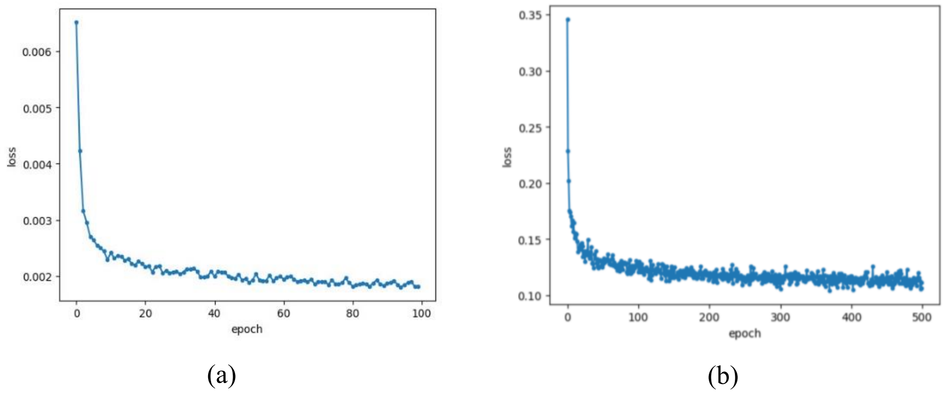


Figure 5.5: Training of Vxnet for 2-D oasis dataset showing loss function vs number of epochs (a) D1 (b) D2 dataset

Initially, the Voxelmorph CNN network (Vxnet) is used to evaluate results on the five datasets [122]–[127]. Although Vxnet demonstrated superior results as compared to [81], [128]–[130] the moving and fixed images are still not properly aligned. This is demonstrated by the similarity metrics in Table 5.1 and the visual representations in Figure 5.8. To enhance alignment accuracy, TLBO is applied to the U-Net warped image and the fixed image to find the optimal transformation parameters. The outcomes for Vxnet and the proposed approach (Vxnet-TLBO) across datasets D1-D5 are presented in Table 5.1. The dice score results for these datasets using Vxnet and the proposed Vxnet-TLBO approach as illustrated in Figure 5.6.

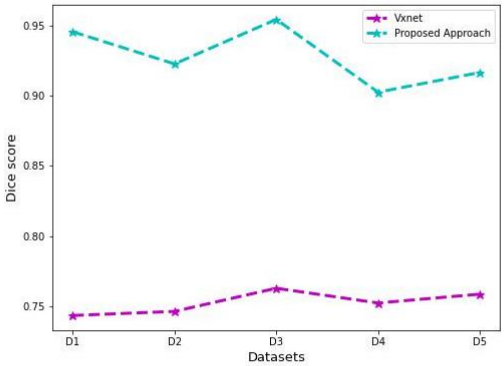


Figure 5. 6: Dice score analysis of proposed approach with Vxnet for five different datasets(D1-D5)

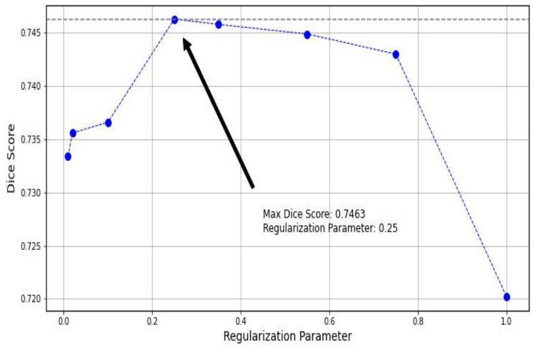


Figure 5.7: Significance of regularization weighing parameter (λ) on 2D atlas-based registration

Table 5.1: For monomodal medical images, experimental results between proposed approach and state-of-the-art approaches [81], [128]–[130] with different similarity metrics

Images	Ground Truth	Registration Parameters			Dice Score (Vxnet)	Dice Score (PA)	HD	MI	SSIM
	(θ, t_x, t_y)	θ	t_x	t_y					
a1	(2,4,8)	-1.93979	-3.85162	-8.00326	0.7434	0.9454	2.4518	2.1315	0.9345
a2	(3,6,9)	-2.80325	-5.90263	-8.73835	0.74632	0.9225	1.5607	2.9508	0.9570
a3	(3,4,8)	-2.72203	-4.18827	-7.91516	0.7628	0.95421	2.3809	2.4380	0.9553
a4	(1,2,2)	-1.00367	-1.89093	-2.19094	0.7523	0.9024	1.0923	3.8280	0.9654
a5	(2,4,8)	-1.73510	-4.18120	-7.96743	0.7586	0.9164	3.9442	1.8963	0.92633

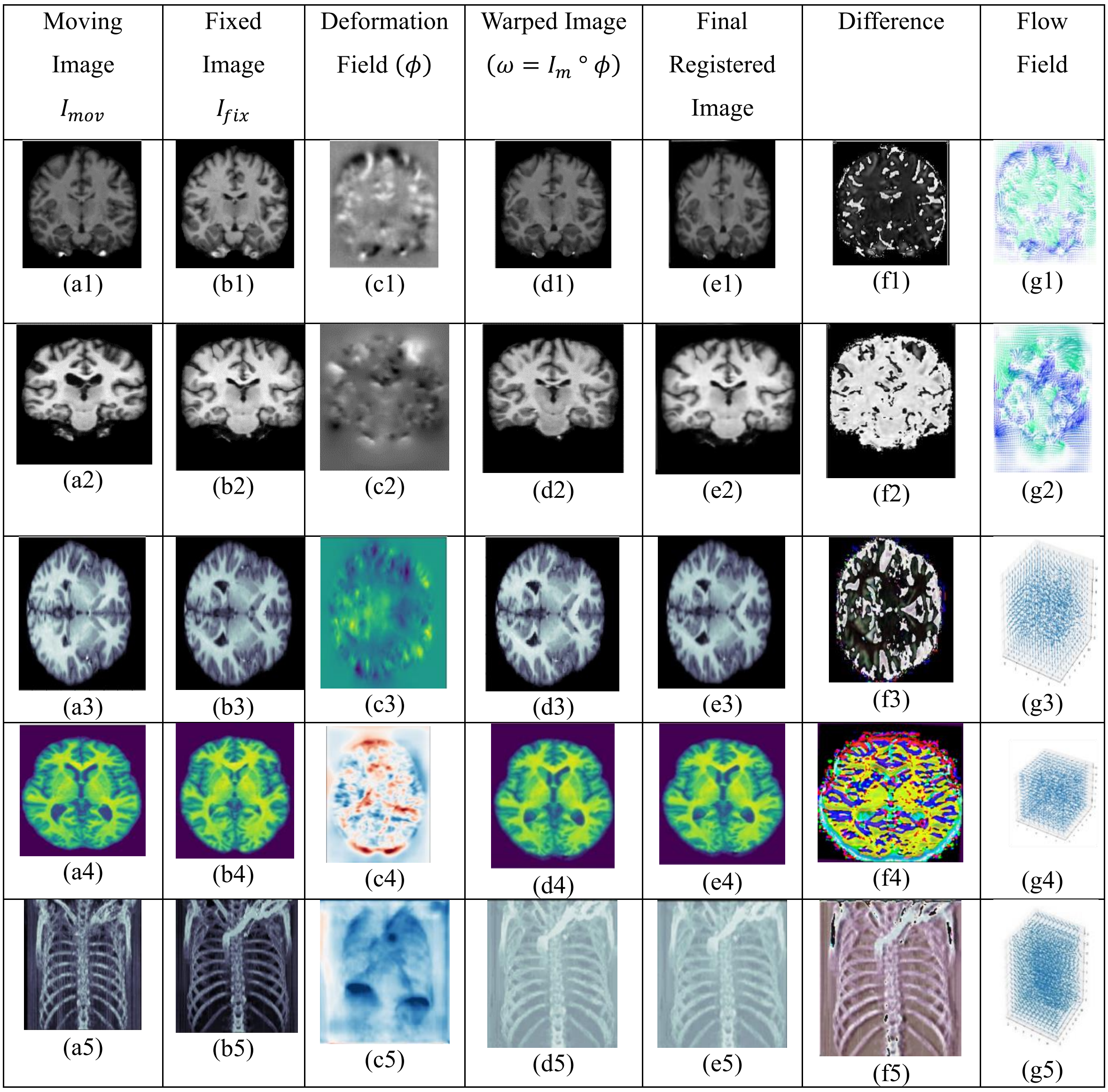


Figure 5.8: Experimental evaluation of monomodal medical images (a) Moving image $(a1 - a5)$ (b) Fixed image $(b1 - b5)$ (c) Deformation field $(c1 - c5)$ (d) Warped image $(d1 - d5)$ (e) Final registered image $(e1 - e5)$ (f) Difference $(f1 - f5)$ (g) Flow field $(g1 - g5)$

5.2.1.2. Comparative Analysis

In this section, a comparative analysis between proposed scheme and state-of-the-art approaches [81], [128]–[130], for 2-D and 3-D medical image registration is presented. The minimum dice score for affine implementation is 0.516 as tabulated in Table 5.2. The Vxnet framework alone (without TLBO) achieved dice scores from 0.64310 to 0.7463 across datasets. With the integration of TLBO, Vxnet showed notable improvements in dice scores on 2-D and 3-D datasets from 0.7463 to 0.9225. The proposed scheme achieved dice scores from 0.7628 to 0.9542, 0.7523 to 0.9024, and 0.7463 to 0.9225. This demonstrates that deep learning and optimization algorithm outperforms the evaluated baselines [81], [128], [129] indicating accurate alignment. Vxnet achieves superior results as compared to traditional methods [81], [128]–[130] . The proposed scheme extends Vxnet by utilizing TLBO with optimal transformation parameters outperforming both Vxnet and conventional methods. The effectiveness of the proposed approach is quantitatively evaluated against [81], [128]–[130] for MRI image registration between fixed and moving images as detailed in Table 5.2. The proposed method is compared with an unsupervised learning model [81], spatial attention (SA-Voxelmorph) [128], and multi-view feature slicing attention module [129] for deformable medical image registration on 2-D and 3-D datasets [122]–[127] for monomodal MRI. The comparison involves T1-weighted brain MR images with similarity metrics as summarized in Tables 5.2 and 5.3. Higher dice scores and lower Hausdorff distances accurately align the moving and fixed images. Additionally, SSIM and MI metrics detailed in Table 5.1 quantify registration, indicating that proposed scheme offers superior accuracy. Registration performance of traditional methods [81], [128], [129], Vxnet, and the proposed scheme are compared based on similarity metrics and Hausdorff distance in Table 5. 2, Table 5.3, and Figure 5.9.

Table 5.2: Quantitative comparison of dice score for proposed approach with state-of-the-art methods [81], [128], [130]

Method	3-D MRI brain [81]	ADNI [128]	OASIS [130]
Affine	0.567	0.486	0.516
SyN	0.742	0.601	0.599
Voxelmorph	0.749	0.710	0.643
Vxnet	0.7628	0.7523	0.74632
Proposed scheme	**0.9542**	**0.9024**	**0.9225**

Table 5.3: Comparison of hausdorff distance similarity metric of proposed approach with [129] for MRI image.

Method	OASIS [129]	ADNI [129]
Affine	3.5300	3.5068
SyN	2.0521	1.9650
Voxelmorph	2.0900	2.0446
Proposed scheme	**1.5607**	**1.0923**

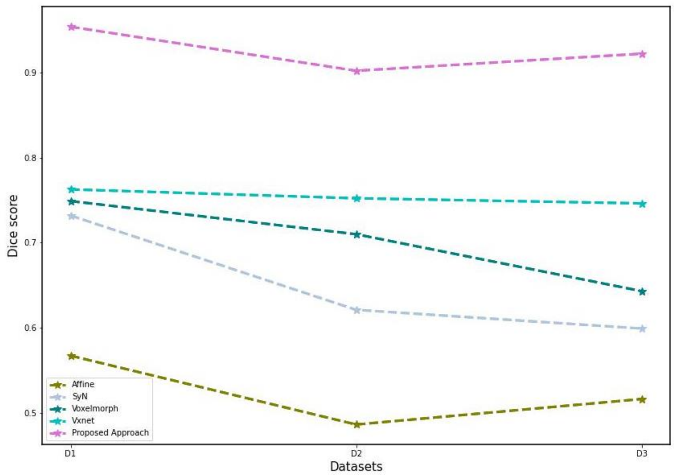


Figure 5.9:Quantitative comparison between proposed approach and traditional method [81], [128], [130] corresponding to different [123], [124], [126].

5.2.2. Conclusion

In this section, an unsupervised learning U-Net model followed by TLBO along with affine transformation for monomodal medical image registration is discussed. Vxnet(U-Net) model takes a two-channel image (moving and fixed) through its encoder and decoder phases to extract features and create a displacement field. The displacement field provides direction vectors for each pixel in the moving image to align with the fixed image, followed by a B-Spline transformation for improved alignment. Despite its capabilities, Vxnet alone does not provide efficient registration as measured by dice score and Hausdorff distance. To improve the warped (registered) image quality, TLBO with RTP is applied to detect the optimum transformation parameters (θ, t_x, t_y) for accurate alignment of both fixed and U-Net warped image. Extensive experiments on 2-D and 3-D images from CT and MRI datasets indicate that proposed scheme achieves superior registration accuracy. However, this work of deformable (non-rigid) image registration is limited to monomodal medical images which can be further enhanced to multimodal modalities of images. In the next section our objective is to improve rigid transformation parameters for precise alignment using a novel UEO, a meta-heuristic approach along with VGG19. This novel approach outperforms the techniques TLBO with VGG19, KPCA, and TLBO [131] in terms of qualitative and quantitative parameters.

5.3. VGG-19 with United Equilibrium Optimizer for Medical Image Registration

In this section, a robust method for aligning both monomodal CT-CT, MRI-MRI and multimodal CT-MRI, MRI-CT medical images leveraging UEO and VGG19 is presented. This approach consists of a series of steps as detailed in Figure 5.10, which are designed to ensure accurate image alignment. Image registration is characterized as an optimization process which aims to maximize an objective function, as defined in Equation (5.11).

$$\hat{T} = \arg\max_{T} SSIM\, [r_i(p, q), t_i(p, q)] \tag{5.11}$$

The reference image is represented by $r_i(p, q)$, where p and q are the image coordinates. The target image is denoted by $t_i(p, q)$. The optimal transformation is indicated by $\hat{T}$, while $\hat{T}$ representing the transformation and its parameters.

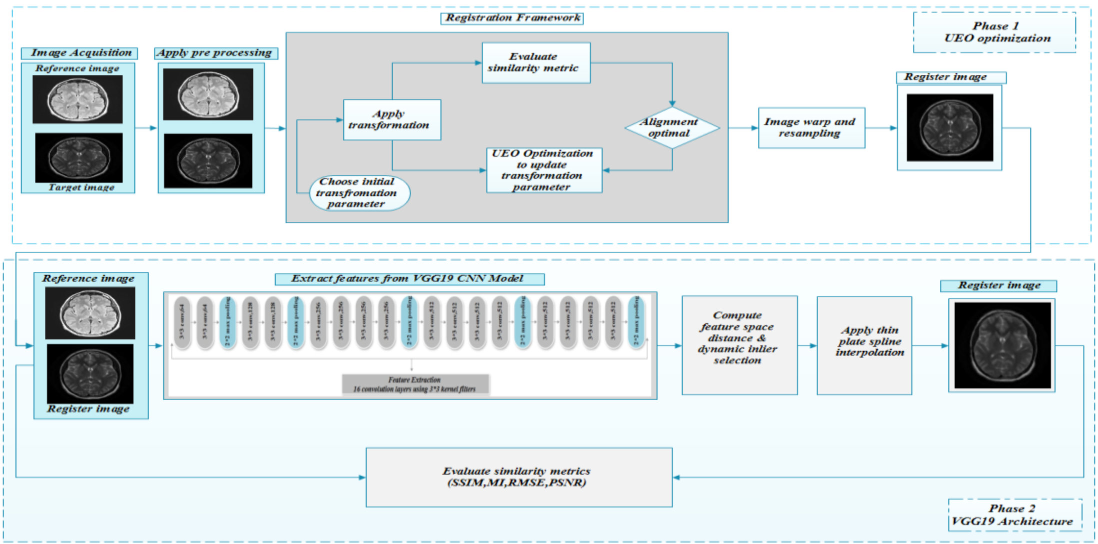


Figure 5.10: Framework of proposed strategy

5.3.1. United Equilibrium Optimizer Algorithm

The UEO algorithm dynamically fine tunes its equilibrium pool with elite candidates $(p_{eq1}, p_{eq2}, p_{eq3})$ to enhance the optimization process for medical image registration. This iterative approach involves refines the equilibrium pool (p_{pool}) by incorporating the best performing particles, evaluating each fitness particle based on mutual information. Furthermore, the equilibrium pool is enhanced by generating a united candidate (U_{avg}) through aggregating the top candidates (p_{avg}, t_{avg}). This strategy aims to continually enhance the quality of candidates within the equilibrium pool, thereby improving the overall optimization process.

Dynamic parameters (a_1, a_2) strike a balance between exploration and exploitation, guiding the search process. This iterative process persists until maximum iterations $(t_{maxiter})$ are reached, which utilized newly generated vectors (rand, λ) and evaluating fitness $fit(P_i)$. The transformation parameters (P$[i]$) undergo iterative refinement to ensure precise image alignment. By avoiding local optima, UEO efficiently traverses the search space by initially exploring with random vectors and progressively concentrating on refining solutions around the optimal ones are outlined in Algorithm 5.2.

Algorithm 5.2: UEO optimization

Input: Set the maximum iterations as $t_{maxiter}$, where population is denoted by $P_i(i = 1,2,3 \ldots\ldots, n)$, and dimension as D. Set parameter value a_2 =1 and GP = 0.5

$(p_{eq1}, p_{eq2}, p_{eq3})$: Best equilibrium positions (equilibria) found by the particles during the search.

$fit(P_i)$: A measure of the effectiveness of a particle's position in aligning images.

$(p_{avg}, t_{avg}, u_{avg})$: Average position calculated from the top particle and their dimensions.

$(\mathrm{C_i}, \mathrm{T_i}) =$ Indicates the current and maximum iterations

$(\mathrm{a_1}, \mathrm{a_2}) =$ Control the exploration and exploitation magnitude of algorithm

$e^{-\lambda(1-t)}$ = Exponential decay function reduces F over time, shifting from exploration to exploitation as the algorithm progresses.

P_{eq} = Candidate equilibrium position.

G = Attraction vector influenced by the global candidate position

$(\lambda \; and \; V)$ = Vectors controlling adaptation and velocity, respectively

Output: The objective is to optimize the registration of medical images by finding the optimal transformation parameters that maximize the mutual information between reference and target images.

while $(iter < t_{maxiter})$
    for each i particle
        Calculate the fitness of particle $fit(p_i)$
        if $fit\ (p_i) < fit\ (p_{eq1})$
            Replace(p_{eq1}) with (p_i) and $fit\ (p_{eq1})$ with $fit\ (p_i)$ # Update the best particle position if the current fitness is better.
        elseif $fit\ (p_i) > fit\ (p_{eq1})$ & $fit\ (p_i) < fit\ (p_{eq2})$
            Replace (p_{eq2}) with(p_i) and $fit\ (p_{eq1})$ with $fit\ (p_i)$ # Update the second-best particle position if the current fitness is better.

8 ***elseif*** $fit\ (p_i) > fit\ (p_{eq1})$ & $fit\ (p_i) > fit\ (p_{eq2})$ & $fit\ (p_i) < fit\ (p_{eq3})$

9 Replace(p_{eq3}) with(p_i) and $fit\ (p_{eq1})$ with $fit\ (p_i)$ # Update the third-best particle position if the current fitness is better

10 ***endif***

11 ***end for***

12 $(p_{avg})= ((p_{eq1}) + (p_{eq2}) + (p_{eq3}))/3$ # candidate average is the average of the top three equilibrium positions.

13 ***for*** each dimension d_m

14 Compute $p_{united} = \sum_1^{d_m} randomselect(p_{eq1,d_m}, p_{eq2,d_m}, p_{eq3,d_m})$ # Randomly select values from $p_{eq1,d_m}, p_{eq2,d_m}, p_{eq3,d_m}$ for each dimension d_m

15 ***end for***

16 $(t_{avg})= \left((p_{eq1}) + (p_{eq2})\right)/2$ # elite average is the average of the top two equilibrium positions

17 $(U_{avg})= \left((p_{eq1}) + (p_{eq2}) + (p_{eq3}) + (p_{united})\right)/4$ # united average value is average of the top three equilibrium positions and united positions.

18 $p_{pool} = \{(p_{eq1}), (p_{eq2}), (p_{eq3}), (p_{avg}), (t_{avg}), (U_{avg})\}$ # pool of equilibrium positions is the combination of candidate, elite and united average

19 Compute $t = (1 - C_i/M_i)^{a2(C_i/M_i)}$ # t represents iteration constraint

20 ***for*** each i particle

20 choose one candidate from p_{pool}

21 Generate two vectors $\overrightarrow{rand}, \vec{\lambda}$ # $\overrightarrow{rand}$ vector is equal to $\vec{\lambda}$ that is between 0 and 1.

22 Construct $\vec{F} = a_1 sign\left(\overrightarrow{rand} - random(0,1)\right).\left(e^{-\lambda(1-t)} - 1\right)$ # $\vec{F}$ represents the scaling factor.

23 Construct $GCP = \lceil 2.random() \rceil . r1.1.(random \geqq GP)$ # GCP is the Global Convergence Point parameter.

24 Update the concentration $\mathrm{P}[i] = p_{eq} + (p[i] - p_{eq}) \times F + \left(\frac{G}{\lambda} \times V\right) \times (1 - F)$ # $\mathrm{P}[i]$ *is* current position of the i^{th} particle.

25 Compute $G0 = GCP.(P_{eq} - \lambda.\mathrm{P}[i])$ # G0 is the initial attraction calculation.

26 Compute G= G0.F # G is the final attraction calculation.

27 ***end for***

28 Iter = Iter++

29 ***end while***

Algorithm 5.3: VGG19 based multimodal medical image registration

Input: (I_{Ref}, I_{Ueo}): reference image I_{Ref} and UEO register image I_{Ueo} from dataset D_S

Output: (FI_{Reg}): final register image

1 ***for*** **each image pair** $(\boldsymbol{I_{Ref}}, \boldsymbol{I_{UEO}}) \in \boldsymbol{D_S}$**:**

2 $I'_{Ref} = F_E\left(I_{Ref}(a,b)\right)$, $I'_{Ueo} = F_E(I_{Ueo}(a,\mathrm{b}))$ # Apply feature extraction on reference

and UEO register image

3 $\boldsymbol{I_{Cn}} \xleftarrow{\boldsymbol{concatenate}} (\boldsymbol{I'_{Ref}}, \boldsymbol{I'_{Ueo}})$ # concatenate the feature extracted images

4 $\boldsymbol{F_{Cn}} \xleftarrow{\boldsymbol{apply\ CNN}} \boldsymbol{VGG_{19}}\ (\boldsymbol{I_{Cn}})$ # apply VGG19 on concatenate images

5 $d_i(k, l) = \sqrt{\sum_{i=1}^{n} \left(D_i(k) - D_i(l)\right)^2}$ # calculate the euclidean distance for feature points

6 $d(k, l) = \sqrt[2]{d_1(k, l) + d_2(k, l) + d_3(k, l)}$ # compute the weighted sum of distance

7 $d(i, l) < \theta . d(k, l)$ # check matching criteria for computing inliers

8 $f(k, l) = a_1 + a_2\, k + a_3 l + \sum_{t=1}^{f} T_t\, df_k^2 \ln df_l$ # apply transformation function

9 $m_{fp} = m_{fp}(f(k, l), dynamic_{threshold} = \theta)$ # match feature point using dynamic threshold

9 $tps_{matrix} = tps_{matrix}(m_{fp})$ # compute TPS matrix from matched feature points

10 $optimized\_tps_{coff} = optimized\_tps_{coff}\ (tps_{matrix})$ # optimized TPS coefficient

11 $I_{reg} \xleftarrow{tps_{coff}} (I_R, optimized\_tps_{coff})$ # Generate register image

12 ***end for***

Step1: In this step, randomly two images are selected from publicly accessible repository and clinical dataset. Afterwards, gaussian filtering is applied to both reference and target images to decrease noise. To further improve the contrast of the filtered images and minimize the influence of noise, min-max normalization is conducted. Finally, these images are resized to dimensions of $224 * 224$ pixels, aiming to reduce computation time and improve the efficiency of the registration process.

Step 2: In this step, the pre-processed reference image is aligned with target image using a transformation T. This transformation ensures that the coordinates of target image match the reference image. Algorithm 5.2 is employed to optimize the parameters of the transformation including rotation (θ) as well as translation (t_x, t_y). This optimization process involves a population of 30 particles and maximum of 60 iterations to find the optimal transformation parameters that maximize or minimize the similarity metrics and test the reliability of IR process. This process is complete if the registered image matches the desired results. If not, the transformation parameters are modified, and the procedure is repeated until the desired outcome is achieved.

Step 3: After determining the optimal rigid transformation parameters value, the similarity metric MI is employed to measure the degree of resemblance between pre-processed reference and registered image. A higher MI value signifies a more precise alignment between these images as outlined in Table 5.5.

Step 4: During this step, our goal is to extract both high and low-level features from reference and UEO-registered images acquired from step 3. These images are fed into VGG-19 model, which consists of 16 convolution layers and 5 max-pooling layers as illustrated in Figure 5.11. Each convolution layer employs a kernel size of 3x3 pixels. The 2x2 size of the max-pooling layers notably improves the extraction of image feature points, as depicted in Figure 5.11.

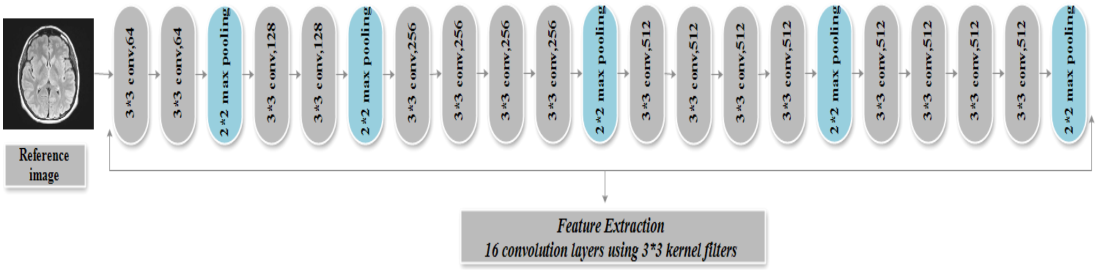


Figure 5.11: Step-by-Step Methodology for Feature Extraction Using VGG19

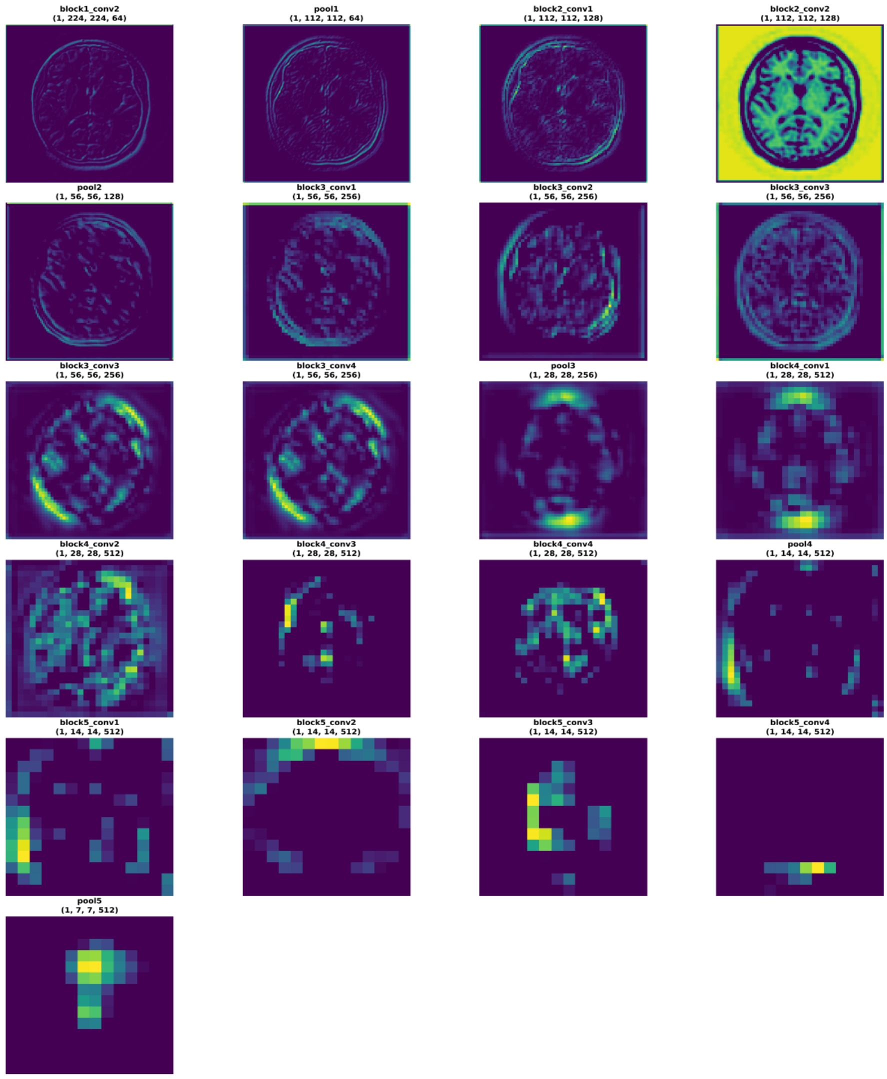


Figure 5.12: Layer-by-Layer Feature Extraction of Target Image Using VGG19

Step 5: In step 4, the feature points obtained from VGG19 model are then utilized to compute feature space distances. These distances are calculated through weighted sum of three distance values for feature points k and l within the pre-processed reference and UEO registered image, as depicted in Equation (5.11). This computation aims to quantify the dissimilarity between feature points, expressed mathematically in Equation(5.12).

$$d(k,l) = 1.414 * d_1(k,l) + d_2(k,l) + d_3(k,l) \tag{5.11}$$

$$d_p(k,l) = \sqrt{\sum_{i=1}^{m} \left(D_p^i(k) - D_p^i(l)\right)^2} \tag{5.12}$$

The feature space distance, denoted as $d(k,l)$, is computed as the square root of the weighted sum of three distances: $d_1(k,l), d_2(k,l)\ and\ d_3(k,l)$. Each $d_1(k,l)$ or $d_p(k,l)$ represents the euclidean distance between corresponding feature vectors.

Subsequently, a matching criterion $\mathrm{d}(i,l) < \theta.d(k,l)$ is used to align reference feature point k and UEO register feature point l. First, $\mathrm{d}(i,l)$ is computed using gaussian radial basis function $G_{ij} = exp\left(-\frac{\|x_i - x_j\|^2}{2\beta^2}\right)$ to construct (G_G) with euclidean distances. Apply a non-rigid transformation using (G_G) and transformation parameters a_j to compute transformed positions z_i of reference points: $z_i = x_i + \sum_{j=1}^{M} G_{ij} a_j$. Finally, compute $\mathrm{d}(i,l)$ as the euclidean distance between z_i and the target feature point y_l.

Step 6: In this step, dynamic inlier selection enhances feature point alignment between reference and UEO register images by iteratively adjusting distance thresholds. Thresholds can be adjusted randomly or based on a heuristic measure such as the ratio $\frac{d(i,l)}{d(k,l)}$, which indicates how closely l matches i relative to the initial pair (k,l).This approach enhances matching accuracy by effectively filtering correct matches (inliers) while minimizing outliers, thereby improving alignment and registration processes.

Step 7: After identifying inliers with dynamic inlier selection(X_i, Y_i), TPS is used to correct local misalignments among selected feature points, mapping points in the register image(Y_i) to corresponding points in the reference image (X_i). The TPS transformation coefficients (a_1, a_2, a_3, T_t) are extensively adjusted, as depicted in Equation (5.13). Moreover, TPS method is employed to compute a non-linear transformation, enabling complex warping to align features between images with non-uniform distortions. This facilitates precise alignment between the reference and UEO registered images. The TPS transformation function is given by:

$$f(k,l) = a_1 + a_2\, k + a_3 l + \sum_{p=1}^{q} T_t\, d_{qk}^2 \ln dq_l \qquad (5.13)$$

Here, a_1, a_2, a_3 are the TPS interpolation coefficients, T_t represents the transformation coefficients for each feature points p and q, and d_{qk} is the distances between corresponding feature points in the reference and registered images. TPS uses a combination of linear terms (a_1, a_2k, a_3l) and non-linear terms $\left(\sum_{p=1}^{q} T_t\, d_{qk}^2 \ln dq_l\right)$ to create a smooth, flexible transformation. Final registered image generated by these transformation guarantees accurate alignment between reference and UEO register image, as shown in Figure 5.13 (Col 5, e_1- e_8).

5.3.2. Experimental Findings and Comparative Analysis

The effectiveness of proposed scheme across different datasets [28–31] is demonstrated through experimental findings. The performance of the proposed scheme is evaluated and compared with recently developed methodologies [10–13], utilizing PSNR, RMSE, CC, and SSIM as a similarity measure.

5.3.2.1. Experimental Findings

The proposed scheme effectiveness is evaluated through a series of experiments using Keras with a TensorFlow backend. The experiments are performed on a system equipped with an Intel Core i7-1165G7 processor at 2.80 GHz, an NVIDIA GeForce GTX GPU, and 16 GB of RAM. The experiments are carried out on different datasets taken from public repositories [28–31] and clinical sources collected from Fortis Hospital Mohali and Paras Hospital Panchkula, India. The experiments are conducted on Python 3.9.4 running on Windows 11. Initially, reference and target medical brain images of sized 224×224 pixels undergo gaussian filtering and min-max normalization as depicted in Figure 5.13 (Column 3, c1-c8) and Figure 5.13 (Column 4, d1-d8). UEO optimization employs 30 particles over 100 iterations, optimizing the rigid transformation parameters (θ, t_x, t_y) to enhance similarity metrics and validate IR reliability as illustrated in Table 5.4. To further enhance the quality of UEO based registered image, features are extracted from reference and UEO registered images using VGG-19. The feature space distance quantifies the dissimilarity between feature points, ensuring reliable matching. TPS interpolation determines the transformation matrix, ensuring precise alignment as illustrated in Figure 5.13 (Column 5, e_1-e_8). UEO outperforms EO and TLBO in optimizing transformation parameters (θ, t_x, t_y) as shown in Table 5.4. The experimental results of VGG19 with optimization methods namely UEO, EO, and TLBO are illustrated in Figure 5.13 (Column 5, e1-e8), Figure 5.13 (Column 6, f1-f8), and Figure 5.13 (Column 7, g1-g8) respectively. Experimental results on different medical image modalities using VGG19 with UEO, EO, and TLBO are detailed in Table 5.4 and Figure 5.13. The integration of VGG19 with UEO generates more accurate registered images as compared to VGG19 with EO and VGG19 with TLBO, highlighting the effectiveness of combining deep learning with optimization techniques as tabulated in Table 5.4.

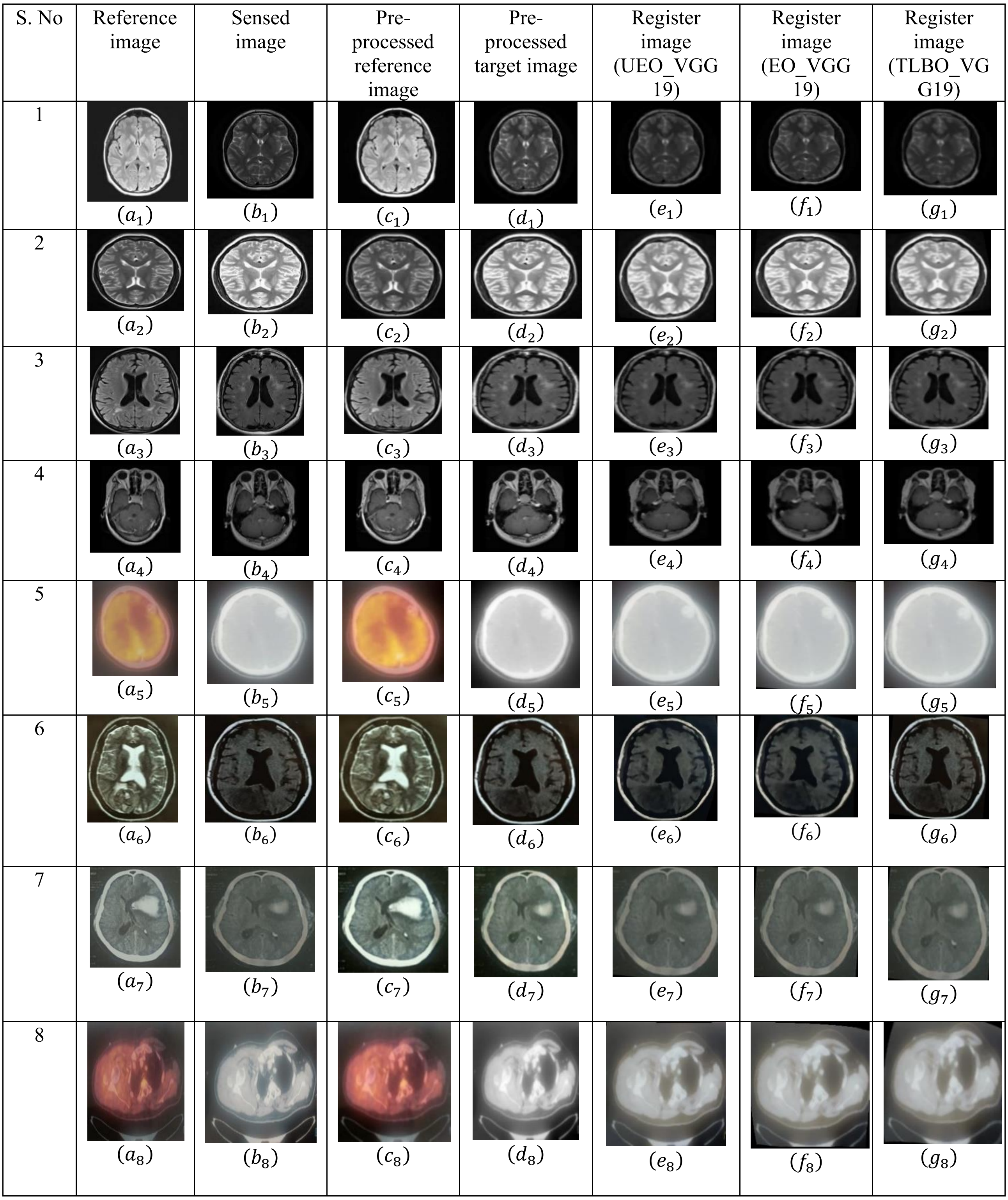


Figure 5.13: Experimental outcomes for intermodal images sourced from different dataset (a) Reference image $(a_1 - a_8)$ (b) Target image $(b_1 - b_8)$ (c) Pre-processed reference image $(c_1 - c_8)$ (d) Pre-processed target image $(d_1 - d_8)$ (e) UEO-VGG19 register image $(e_1 - e_8)$ (f) EO-VGG19 register image $(f_1 - f_8)$ (g) TLBO-VGG19 register image $(g_1 - g_8)$ through extensive experiments.

UEO based image registration outperforms TLBO and EO due to its dynamic adaptation and superior balance between exploration (global search) and exploitation (local search) for registering images. Moreover, UEO is used to avoid local optima and converge faster significantly improves the registration process, as illustrated in Table 5.4.

Table 5.4: Quality assessment of the proposed strategy through optimization and multimodal quantitative metrics

S. No	Modalities	Optimization technique	Baseline data	Rigid transformation parameters			Runtime
			(θ, t_x, t_y)	θ	t_x	t_y	
1	MR-PD & MR-T2	TLBO	(6,3,3)	-5.995860	-2.674944	-3.281899	693.194184
		EO		-5.294570	-2.726899	-3.259100	438.221639
		UEO		**-5.938140**	**-2.971266**	**-3.045696**	**322.407385**
2	MR-T1 & MR-PD	TLBO	(6,2,4)	-6.100289	-1.265337	-3.986796	976.2715
		EO		-5.991806	-1.557665	-4.128676	369.00653
		UEO		**-6.029381**	**1.913722**	**-4.085147**	**318.31214**
3	MR-T1 & MR-Flair	TLBO	(5,1,3)	-5.205508	-0.746337	-2.785595	828.07057
		EO		-4.995493	-0.748790	-3.023298	435.296716
		UEO		**-5.089314**	**-1.147896**	**-2.948231**	**327.439110**
4	MR-T1 & MR-PD	TLBO	(8,4,6)	-7.676082	-3.049796	-6.463344	945.136919
		EO		-7.888079	-3.068525	-6.495980	472.780036
		UEO		**-8.078014**	**-4.023437**	**-5.940318**	**365.566450**
5	PET & CT	TLBO	(15,5,10)	-15.28855	-4.349520	-10.927449	866.66941
		EO		-15.135263	-4.248680	-10.671824	572.911726
		UEO		**-15.071168**	**-4.990295**	**-10.099779**	**420.259864**
6	MR-T2 & MR-T1	TLBO	(5,2,6)	-5.014824	1.465775	-6.107085	732.321653
		EO		-5.002278	-1.456214	-6.108630	574.2576
		UEO		**-5.002684**	**-1.966390**	**-6.077750**	**409.035870**
7	CT & CT	TLBO	(10,4,6)	-9.979010	-2.786063	-6.916888	783.234812
		EO		-10.002417	-3.428292	-6.592560	437.45875
		UEO		**-10.003808**	**-3.949715**	**-6.091766**	**312.56289**
8	PET & CT	TLBO	(18,6,12)	-18.044082	-4.966658	-11.619130	763.458621
		EO		-18.044082	-5.366658	-11.969130	438.825986
		UEO		**-18.033999**	**-6.077473**	**-12.061225**	**326.644915**

Evaluations are conducted across different monomodal and multimodal medical images as shown in Table 5.4, confirm the superior efficacy of UEO as compared to TLBO and EO. The amalgamation of UEO with VGG19 significantly enhance the performance as indicated by bold values in the proposed scheme (UEO+VGG19) as tabulated in Table 5.5 in terms of higher values of SSIM, PSNR, CC, MI and lower values of RMSE and MSE.

Table 5.5: Analysis of experimental outcomes for monomodal and multimodal medical images from clinical and kaggle datasets [112], [114], [117], [132]

S. No.	Algorithm	RMSE	MSE	SSIM	PSNR	CC	MI
1	TLBO+VGG19	20.19463300	407.8232020	0.7146996	32.021682	0.855337	1.427260
	EO+VGG19	15.09262571	227.7873508	0.78379395	33.624235	0.896324	1.918590
	Proposed strategy	**11.452526**	**131.1603517**	**0.8161808**	**36.174729**	**0.9382891**	**2.422579**
2	TLBO+VGG19	17.0498321	290.6967746	0.7865309	35.196428	0.9031239	2.315678
	EO+VGG19	11.642091	135.5382828	0.8309431	37.926781	0.9389123	2.762875
	Proposed strategy	**7.609321**	**57.901766**	**0.8605672**	**39.24501**	**0.9862340**	**3.247109**
3	TLBO+VGG19	23.000324	529.0149041	0.7531738	30.925705	0.8839589	1.739446
	EO+VGG19	18.057839	326.0855493	0.8060965	32.126743	0.9069066	2.143929
	Proposed strategy	**12.9845223**	**168.5978193**	**0.8487474**	**34.052307**	**0.9353055**	**2.803540**
4	TLBO+VGG19	19.824610	393.0151616	0.75426711	30.562123	0.9256713	2.345781
	EO+VGG19	13.6902431	187.4227561	0.84980126	33.254315	0.9437231	2.642146
	Proposed strategy	**6.984213**	**48.7792312**	**0.89244126**	**37.424567**	**0.9964214**	**3.478126**
5	TLBO+VGG19	16.981702	288.3782028	0.79413419	32.214612	0.931264	2.256373
	EO+VGG19	12.56021	157.75887	0.84209124	34.59812	0.9623176	2.562301
	Proposed strategy	**7.7432192**	**59.95744357**	**0.88941267**	**36.421456**	**0.9836431**	**3.156238**
6	TLBO+VGG19	19.83928069	393.5970582	0.8236046	29.387712	0.9017364	2.107127
	EO+VGG19	16.48160260	271.643224	0.84325269	31.740490	0.9514732	2.647973
	Proposed strategy	**10.4636015**	**109.4869563**	**0.88802740**	**33.246257**	**0.9823816**	**3.029268**
7	TLBO+VGG19	19.374176044	375.3586973	0.74922152	34.643961	0.943835	1.933757
	EO+VGG19	11.97264764	143.3442915	0.81353570	36.775699	0.9653631	2.983099
	Proposed strategy	**8.5775874**	**73.57500560**	**0.86166029**	**39.964495**	**0.9915824**	**3.039142**
8	TLBO+VGG19	22.21094	493.32585	0.7190832	30.176044	0.9304633	1.670321
	EO+VGG19	13.369012	178.7304818	0.8203985	32.046321	0.9573210	2.214320
	Proposed strategy	**10.983654**	**120.6406551**	**0.8409321**	**34.095471**	**0.9809432**	**2.986021**

The comparison results illustrating the performance of proposed scheme using UEO and VGG19 with TLBO and VGG19 and EO and VGG19 on monomodal and multimodal medical images based on similarity metrics are shown in Figure 5.14.

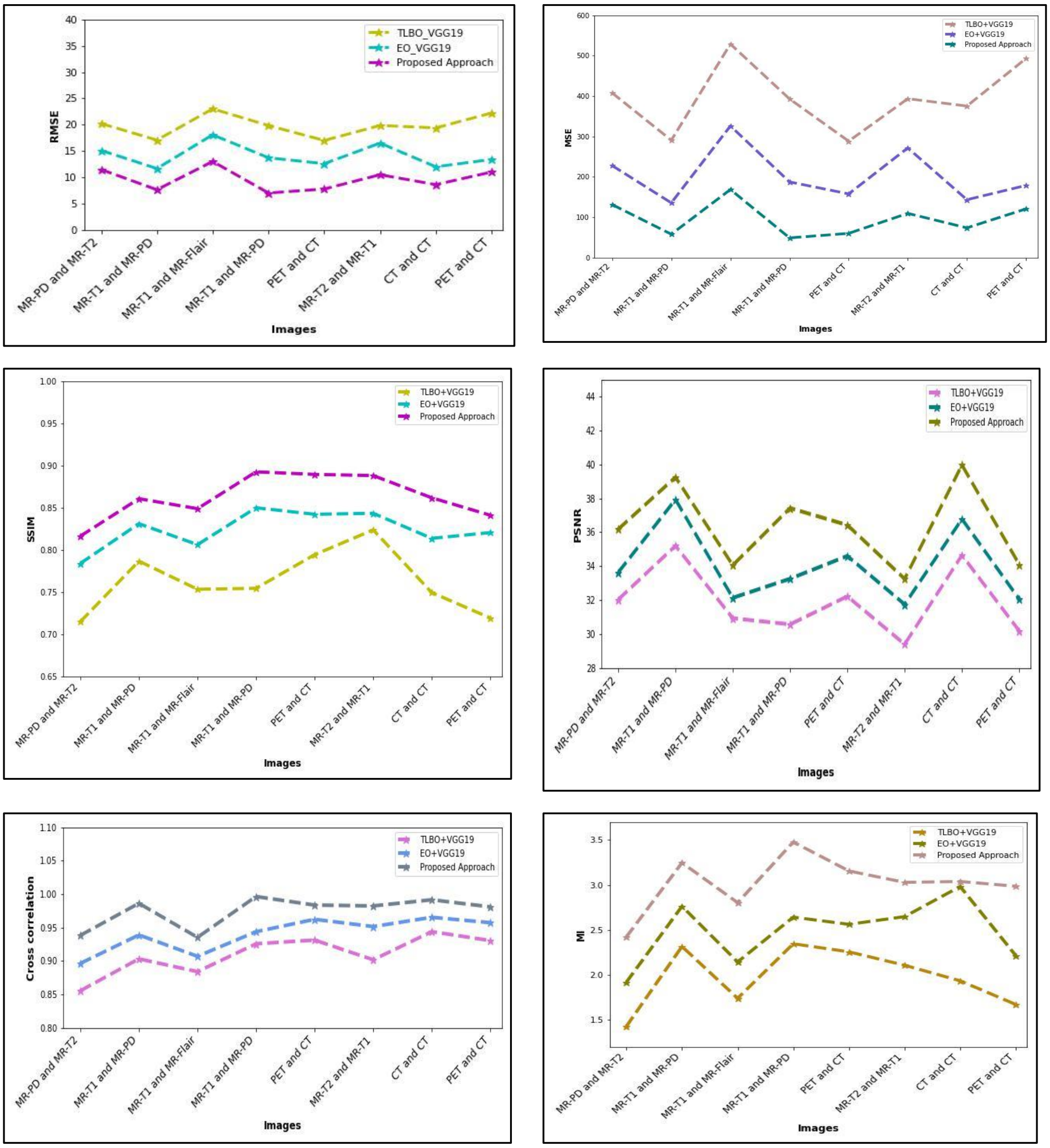


Figure 5.14: Analysing the effectiveness of proposed strategy against conventional methods with various similarity metrics for anatomical and functional images.

5.3.2.2. Comparative Analysis

The proposed scheme has been evaluated against multimodal image registration technique presented by Kori *et al.* [133] which leverages VGG19 for feature extraction on MR-T1 and MR-T2 images from the IXI dataset [134]. The quantitative results of MI and SSIM measures, as tabulated in Table 5.6 and shown in Figure 5.15 which clearly reveals that the proposed scheme with the combination of UEO and VGG19 outperforms the method in [133]. This superior performance is due to the optimal value of rigid transformation parameters and robust feature extraction capabilities of the VGG19 architecture, resulting in precise and accurate registration.

Table 5.6: Assessment of the proposed scheme versus [133] based on MI and SSIM for intermodal medical images.

Method \ Evaluation metrics	MI	SSIM
Simple Elastic [133]	0.221	0.404
SIFT Based [133]	0.15	0.383
CNN(VGG19)[133]	0.373	0.431
Proposed scheme	**1.3214**	**0.8432**

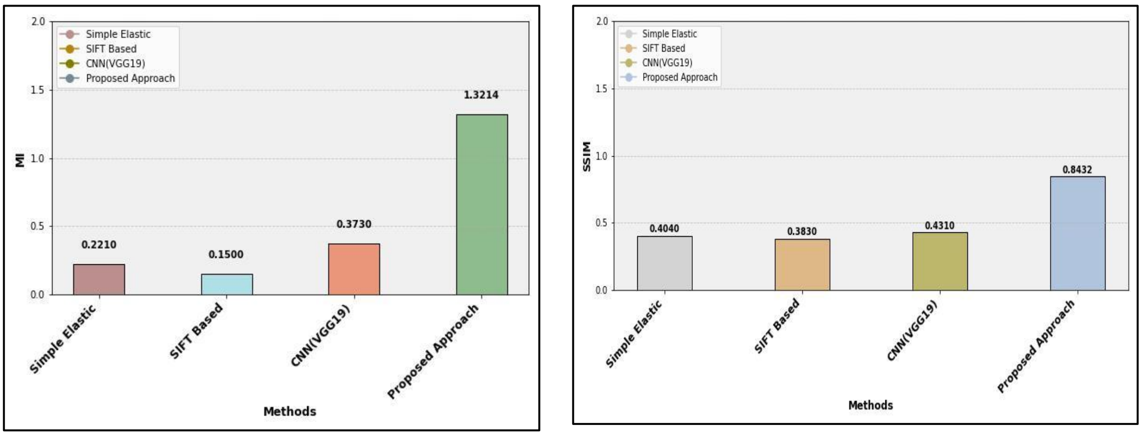


Figure 5.15: Comparative evaluation of assessment metrics between proposed strategy versus the conventional method for intermodal medical images.

The technique outlined by Arora *et al.* [1], [131] employs SURF with TLBO and KPCA with TLBO for multimodal medical image registration as compared to proposed scheme on similar images. The proposed scheme outperforms [1], [131] with different similarity metric namely SSIM, RMSE, and MSE as tabulated in Tables 5.7 and 5.8. The integration of UEO and VGG19 followed by TPS interpolation, results in better image registration accuracy than [1], [131] as indicated in Table 5.7 and 5.8. The results demonstrate that proposed scheme achieves better outcomes across different image pairs, including MR-T1 and CT, MR-PD and CT, CT and MR-T2, MR-T1 and MR-T2, and MR-T1 and CT as illustrated in Figure 5.16 (a), and Figure 5.16 (b).

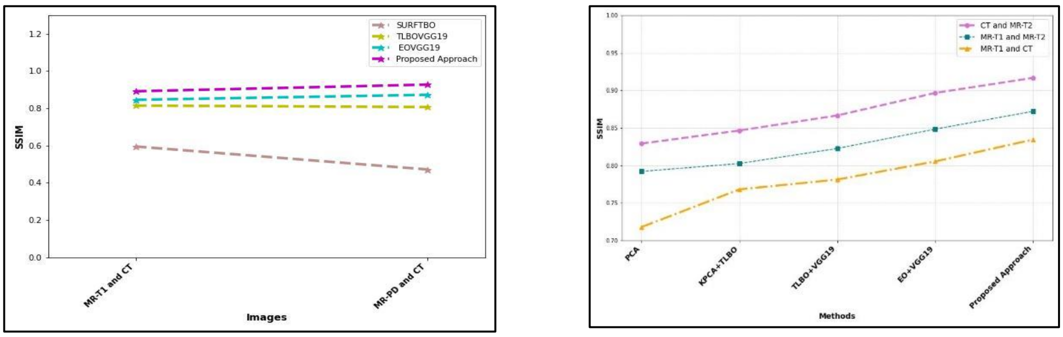
Table 5.7: Comparative quantitative analysis using SSIM of proposed strategy versus [1] for multimodal modalities.

Modalities	Technique \ Assessment metric	SSIM
MR-T1 and CT	TLBO+SURF [1]	0.595064
	TLBO+VGG19	0.714481
	EO+VGG19	0.805423

	Proposed Strategy	**0.831324**
MR-PD and CT	TLBO+SURF [1]	0.471868
	TLBO+VGG19	0.606781
	EO+VGG19	0.692341
	Proposed Strategy	**0.8271432**

Table 5.8: Quantitative assessment of proposed strategy in comparison to [131] for both monomodal and multimodal medical images.

Modalities	Assessment metric / Technique	RMSE	MSE	SSIM
CT and MR-T2	PCA+TLBO [131]	88.7611	7,878.5328	0.82922
	KPCA+TLBO [131]	44.3892	1,970.4010	0.84660
	TLBO+VGG19	31.3892	985.281876	0.86660
	EO+VGG19	26.3892	696.3898	0.89660
	Proposed Strategy	**18.3892**	**338.16267**	**0.91660**
MR-T1 and MR-T2	PCA+TLBO [131]	70.1808	4,925.3446	0.79213
	KPCA+TLBO [131]	43.9439	1,931.06634	0.80268
	TLBO+VGG19	33.1658	1,099.970289	0.82268
	EO+VGG19	29.5471	873.03111	0.84843
	Proposed Strategy	**20.0653**	**402.61626**	**0.87213**
MR-T1 and CT	PCA +TLBO [131]	105.0097	11027.04	0.718113
	KPCA+TLBO [131]	49.41230	2441.575	0.768216
	TLBO+VGG19	35.09432	1,231.611296	0.78132
	EO+VGG19	28.96532	838.98976	0.805423
	Proposed Strategy	**16.6532**	**277.32907**	**0.834512**



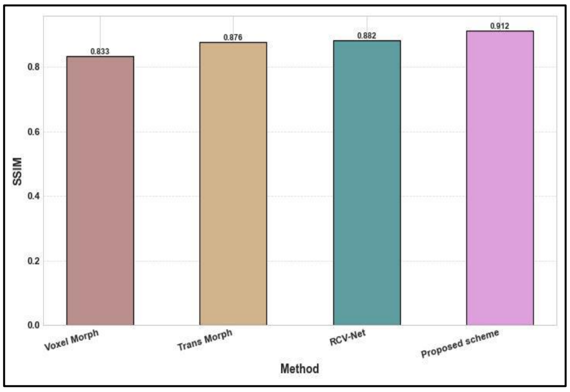
(a) (b)

Figure 5.16: Analysing the proposed scheme in contrast to conventional method for anatomical images, leveraging various similarity metrics.

The proposed scheme is evaluated against Deng *et al.* [135] multimodal image registration approach which integrates ResNet and Voxel Morph for feature extraction. The proposed scheme integrates UEO with VGG19, followed by TPS interpolation, results in better image registration accuracy than [135]. The performance of proposed scheme is compared with [135] employing PSNR and SSIM as a similarity measure across an image pair MR-T2 and MR-T1 as illustrated in Table 5.9 and Figure 5.17.

Table 5.9: Quantitative analysis of the proposed scheme with unsupervised deep learning strategies [135]

Method \ Evaluation metrics	SSIM	PSNR
Voxel Morph	0.833	22.725
Trans Morph	0.876	24.379
RCV-Net	0.882	25.753
Proposed scheme	**0.91234**	**31.3841**



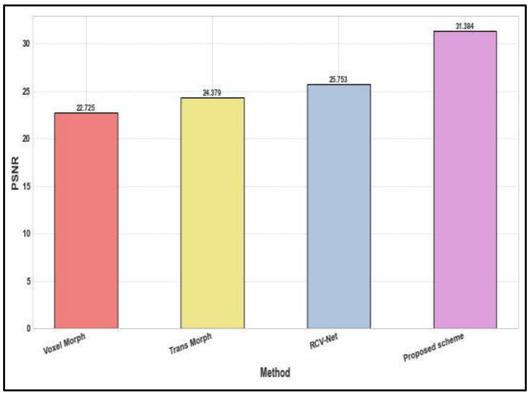


Figure 5.17: Comparative assessment of the proposed scheme against unsupervised deep learning methodology [135].

5.3.3. Conclusion

In this section, a dual-phase scheme integrating UEO with VGG-19 for monomodal and multimodal medical images is presented. This scheme starts by improving the quality of input images through preprocessing, followed by UEO to optimize rigid transformation parameters through maximizing MI between reference and target images. To further enhance registration accuracy, features are extracted using VGG-

19 from both reference and UEO-registered images. Then, feature space distance quantifies the dissimilarity between feature points, ensuring reliable matching. Afterwards, dynamic inlier selection is applied to enhance the matching accuracy by effectively filtering correct matches (inliers) while minimizing outliers. Finally, thin plate spline interpolation is employed to determine the optimal transformation matrix for accurate alignment between reference and UEO-registered images. The proposed scheme has been validated on monomodal and multimodal medical images through extensive experiments using in-house clinical and standard datasets, demonstrating robustness and reliability based on evaluation metrics such as MI, RMSE, SSIM, and PSNR.

5.4. Chapter Summary

In this chapter, deep learning and meta-heuristic approaches for both monomodal and multimodal medical image registration is presented. Deformable image registration scheme employs an unsupervised U-Net model followed by TLBO optimization and affine transformation for monomodal medical images. The Vxnet (U-Net) model processes two-channel images to generate a non-linear displacement vector field, while TLBO optimizes the registration process by iteratively adjusting transformation parameters. Although this deformable registration work focuses on monomodal medical images, which can be further enhanced to multimodal modalities of images. In the next section, we improve rigid transformation parameters using a novel UEO meta-heuristic approach and VGG19, considering both qualitative and quantitative measures. This scheme employs UEO optimization technique to optimize rigid transformation parameters by maximizing MI. Then, features are extracted using VGG-19 from both reference and UEO-registered images. Afterwards, dynamic inlier selection is applied to enhance the matching accuracy by effectively filtering correct matches (inliers) while minimizing outliers. Finally, thin plate spline interpolation is employed to determine the optimal transformation matrix for accurate alignment between reference and UEO-registered images. This approach is validated using both in-house clinical and publicly available standard datasets, demonstrating robustness and reliability through evaluation metrics such as MI, RMSE, SSIM, and PSNR. This scheme struggle with capturing complex anatomical and functional variations in multimodal modalities of images. In the next chapter, we present a method that integrates feature extraction technique SURF with VGG19 for anatomical and functional medical images. 268400

CHAPTER 6 Anatomical and Functional Medical Image Registration using VGG19-SURF and Alpha Trimmed Spatial Relation Correspondence

6.1. Introduction

The aim of this chapter is to discuss the hybridization of deep learning and feature-based methods for registering both monomodal and multimodal medical images. Firstly, this section employs SURF algorithm for feature point extraction followed by brute force matcher to determine total matches between reference and target image. To eliminate mismatches of reference and target image alpha trimmed spatial relation correspondence (ATSRC) algorithm is utilized which outperforms efficiently than the conventional RANSAC method. Finally, a homography matrix is computed using projective transformation to generate the registered image. Extensive experiments are conducted on monomodal and multimodal medical images, including CT and MRI images taken from Kaggle dataset [117], [119]. The proposed scheme demonstrates significant improvements by enhancing MI from 2.53 to 3.654 and CC from 0.9219 to 0.95475 outperforming conventional approaches [49], [56]. This chapter confronts the challenge of accurately alignment of medical images. Furthermore, a novel image registration algorithm that integrates both deep learning and feature based approaches for monomodal and multimodal medical images is presented. This scheme initiates by improving input image quality through preprocessing steps, followed by extracting both low and high-level features using VGG19. Subsequently, euclidean distance is computed for each pair of feature points across different feature spaces. Dynamic inlier selection improves feature point alignment between pre-processed fixed and sensed images by iteratively adjusting distance thresholds to identify inliers. Then, thin plate spline interpolation corrects local misalignments among selected feature points and computes a non-linear transformation. This process enables complex warping to align features between images with non-uniform distortions, resulting in precise alignment. To enhance the reliability of registered image obtained using CNNs (VGG16 and VGG19), SURF algorithm is employed to detect and match features with a BF matcher. Additionally, outliers are minimized using the alpha-trimmed spatial resolution correspondence algorithm. Finally, a homography matrix is applied to generate the final registered image, achieving robust alignment between the sensed and VGG19 registered images. Experiments conducted on publicly available [112],

[114], [117], [119], [132], [134] and clinical datasets taken from Fortis Hospital Mohali and Paras Hospital Panchkula, India confirm the effectiveness of the proposed approaches across different similarity metrics.

The structure of this chapter is outlined as follows. Firstly, feature based method with ATSRC for monomodal and multimodal modalities are described in Section 6.2. The performance of the scheme is outlined in Section 6.2, evaluating MI and CC metrics as presented in experimental results Subsection 6.2.1.1. The effectiveness of proposed scheme in image alignment is evaluated in Subsection 6.2.1.2 by comparing with state-of-the-art methods [49], [56]. Conclusion and limitations of the scheme is discussed in Section 6.2.2. Secondly, to address the limitations of the scheme discussed in Section 6.3, a novel feature and deep learning-based scheme on both monomodal and multimodal medical image modalities is elaborated. Experimental results and comparative analysis with state-of-art techniques in terms of robustness is described in Subsection 6.3.1.1 and 6.3.1.2 respectively. The summary of the chapter including advantages and limitations is presented in Section 6.4.

6.2. Optimizing Image Registration with SURF and Alpha Trimmed Spatial Relation Correspondence

The proposed CT-MRI registration approach involves a series of steps designed to achieve an accurate image alignment as illustrated in Figure 6.1.

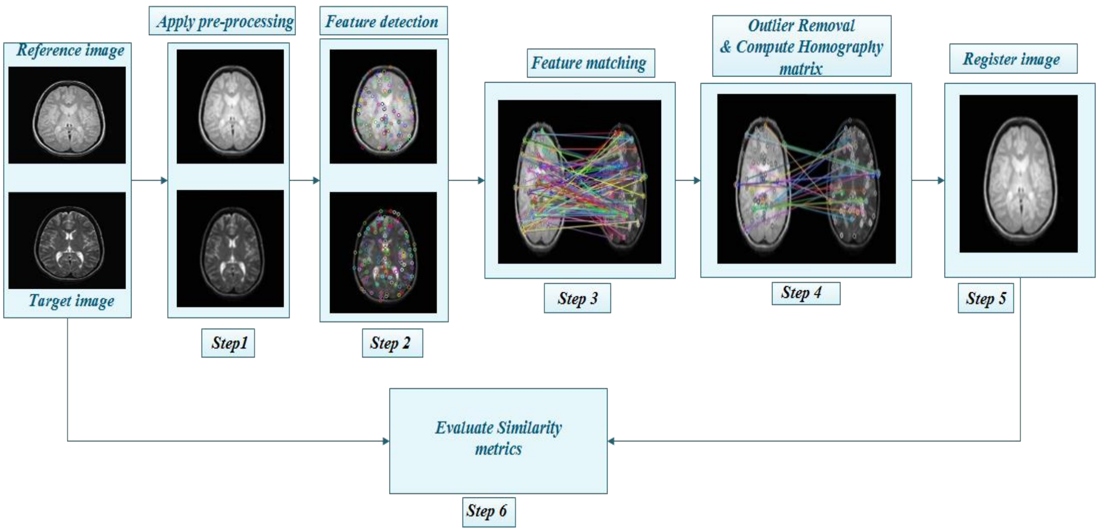


Figure 6.1: Framework of proposed scheme

Step 1: The reference and target images are first pre-processed with gaussian filter to reduce noise, and then normalization is carried out to standardize the data consistently.

Step 2*:* Following pre-processing, different feature extraction algorithms namely SIFT, SURF, and ORB are employed to detect distinctive features in the images.

Step 3*:* In this step, feature matching is performed using a brute force matcher to match the similar features in reference and target images.

Step 4*:* After feature matching in Step3, the subsequent step involves computing homography estimation and employing outlier removal methods such as alpha trimmed spatial relation correspondence to eliminate mismatches (outliers).

Step 5: Next, a projective transformation is applied to generate the enhanced register image. The aim is to obtain the optimal homography matrix that maximizes inliers and minimizes outliers for accurate registration.

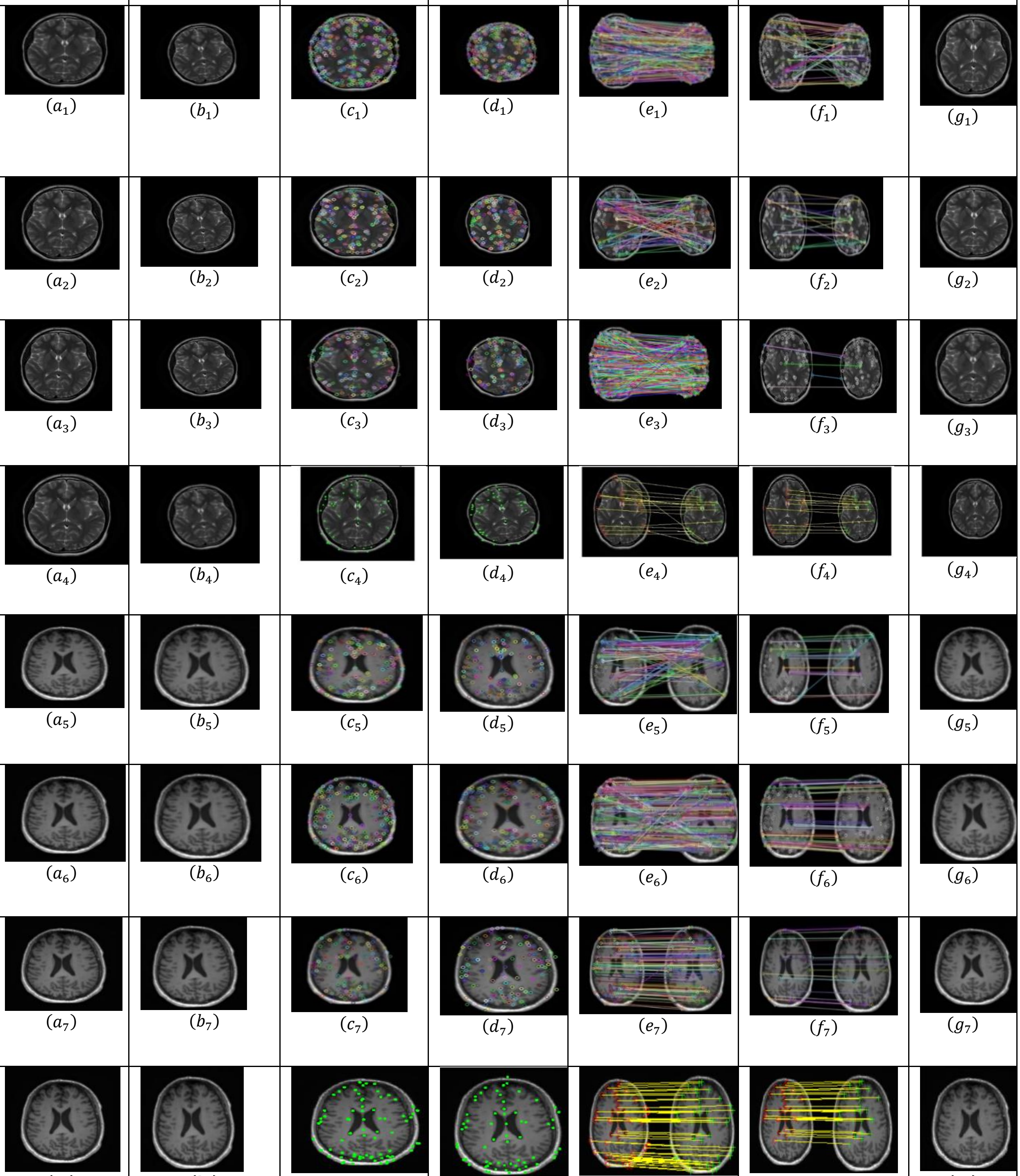
Step 6: In this step, evaluating similarity metrics namely SSIM, PSNR, MI, and CC to measure the resemblance between the reference and registered images. Higher MI and SSIM value along with lower RMSE value, indicate the robustness of proposed scheme.

6.2.1. Experimental and Comparison Results

In this section, the effectiveness of proposed scheme across various datasets [117], [119], illustrating its performance through experimental results. The performance is validated using RMSE, PSNR, SSIM, and CC metrics on various image modalities.

6.2.1.1. Experimental Results

The effectiveness of proposed scheme is evaluated through a series of experiments using Keras with a TensorFlow backend on a system equipped with an Intel Core i7-1165G7 processor at 2.80 GHz, an NVIDIA GeForce GTX GPU, and 16 GB RAM. The experiments are conducted on Python 3.9.4 on Windows 11, utilize standard datasets [117], [119]. CT and MRI brain images of dimension 256×256 are used, as illustrated in Figure 6.2 (a1-a8) for monomodal and Figure 6.2 (a9-a16) for multimodal medical images. Initially, reference and target brain images undergo min-max normalization and Gaussian filtering as depicted in Figure 6.2 (a1-a16) and Figure 6.2 (b1-b16), followed by SURF key point extraction as illustrated in Figure 6.2 (c1-c16) and Figure 6.2 (d1-d16). Then, matches are identified using brute force matcher, and outliers are removed by alpha-trimmed spatial relation correspondence as illustrated in Figure 6.2 (e1-e16) and Figure 6.2 (f1-f16). Finally, homography matrix is determined by maximizing the number of inliers and minimizing outliers, resulting in a robust registered image as depicted in Figure 6.2 (g1-g16).



S. No	Reference image	Target image	Keypoints reference image	Keypoints target image	Total matches	Matches after ATSRC	Register image
1	(a_1)	(b_1)	(c_1)	(d_1)	(e_1)	(f_1)	(g_1)
2	(a_2)	(b_2)	(c_2)	(d_2)	(e_2)	(f_2)	(g_2)
3	(a_3)	(b_3)	(c_3)	(d_3)	(e_3)	(f_3)	(g_3)
4	(a_4)	(b_4)	(c_4)	(d_4)	(e_4)	(f_4)	(g_4)
5	(a_5)	(b_5)	(c_5)	(d_5)	(e_5)	(f_5)	(g_5)
6	(a_6)	(b_6)	(c_6)	(d_6)	(e_6)	(f_6)	(g_6)
7	(a_7)	(b_7)	(c_7)	(d_7)	(e_7)	(f_7)	(g_7)
8	(a_8)	(b_8)	(c_8)	(d_8)	(e_8)	(f_8)	(g_8)

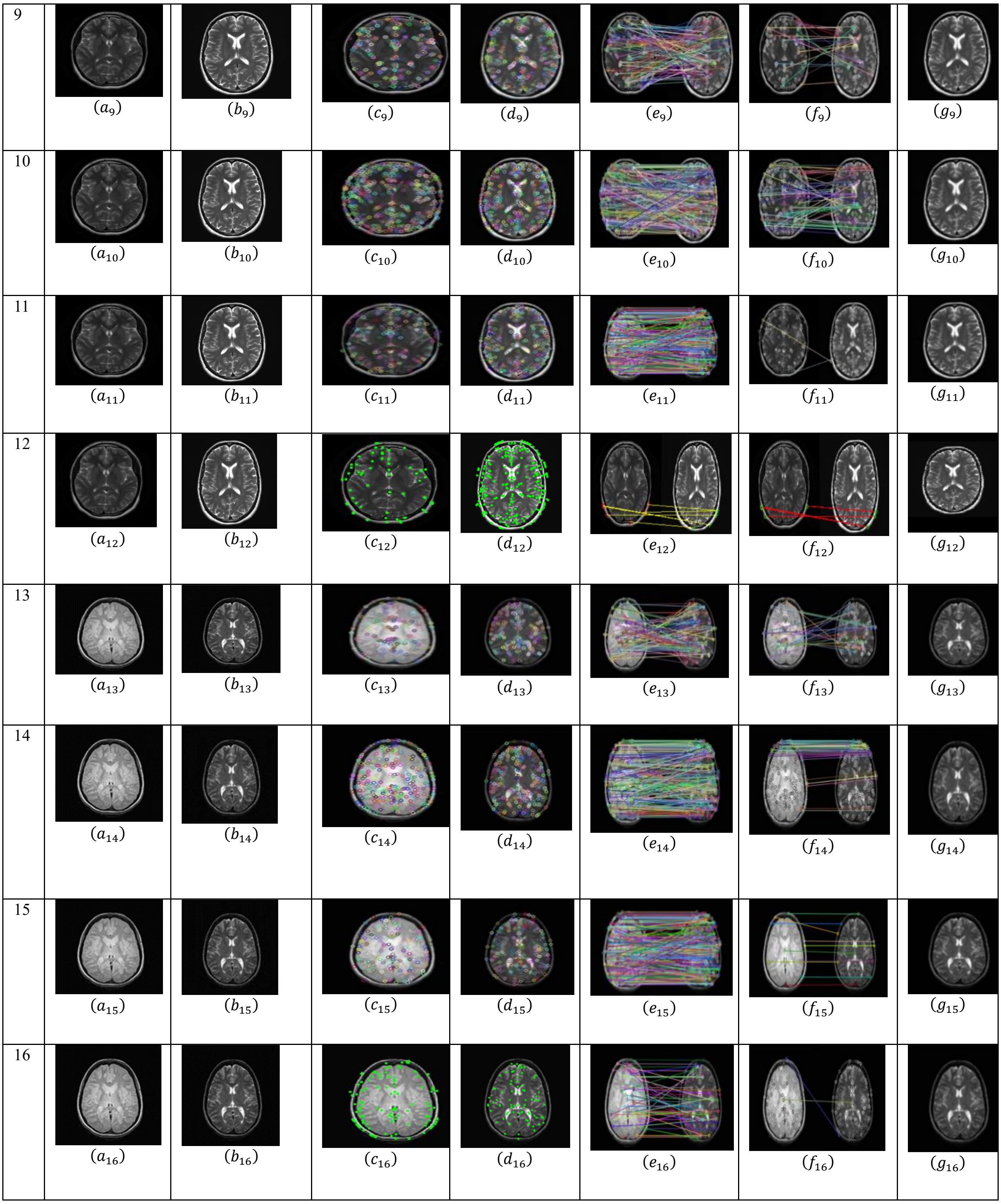


Figure 6.2: Experiment conducted on various monomodal/multimodal medical images (a) Reference image (b) Target image (c) Keypoints of reference image (d) Keypoints of target image (e) Total matches (f) Matches after ATSRC (g) Register image

The quantitative analysis of proposed SURF-ATSRC methodology against traditional methods [49], [56] such as Harris corner detection, ORB, and SIFT is depicted in Table 6.1.

Table 6.1: Experimental outcomes on various monomodal/multimodal medical images taken from Kaggle dataset [81], [83]

S. No.	Algorithm	Images	Reference image keypoints	Target image keypoints	Total matches	Matches after ATSRC
1	Harris corner detection	MRI-Flair and MRI-Flair	467	423	350	68
2	ORB		372	396	126	25
3	SIFT		158	130	630	5
4	**Proposed Methodology**		53	48	17	12
5	Harris corner detection	MRI-T1 and MRI-T1	291	191	143	25
6	ORB		241	182	279	55
7	SIFT		137	142	110	22
8	**Proposed Methodology**		78	65	52	27
9	Harris corner detection	MRI- Flair and MRI-T2	372	394	131	22
10	ORB		467	334	366	69
11	SIFT		158	252	327	2
12	**Proposed Methodology**		372	394	131	22
13	Harris corner detection	CT and MRI-T2	270	394	111	19
14	ORB		239	229	245	48
15	SIFT		131	160	230	10
16	**Proposed Methodology**		104	74	10	4

The proposed scheme outperforms other feature extraction methods across monomodal and multimodal medical images evaluated by MI, RMSE, SSIM, CC, and PSNR as depicted in Table 6.2 as highlighted in bold values and Figure 6.3.

Table 6.2: Computed evaluation metric for monomodal and multimodal medical images

S. No	Algorithm	RMSE	MSE	SSIM	PSNR	Cross correlation	MI
1	Harris corner detection	25.4532	647.86539	0.83652	29.8765	0.87654	2.0987
2	ORB	21.4321	459.33491	0.85142	31.3245	0.89643	2.31432
3	SIFT	18.4322	339.745997	0.86754	33.4562	0.91342	2.5432
4	**Proposed Methodology**	**16.5643**	**274.376034**	**0.87654**	**35.6754**	**0.93452**	**2.76532**
5	Harris corner detection	21.3574	456.14130	0.83212	31.53923	0.93924	2.39973
6	ORB	20.3428	455.51745	0.85295	34.54575	0.939354	2.69345
7	SIFT	20.6343	385.50573	0.86143	37.42513	0.95143	3.024265
8	**Proposed Methodology**	**17.5432**	**307.763866**	**0.89765**	**40.32456**	**0.96543**	**3.23452**
9	Harris corner detection	58.3421	3403.80063	0.76543	22.4568	0.7689	1.1532
10	ORB	54.6325	2984.71006	0.786543	24.6754	0.78654	1.4998
11	SIFT	50.9843	2599.39885	0.819872	26.7843	0.79834	1.8321
12	**Proposed Methodology**	**44.6543**	**1994.00651**	**0.832465**	**28.9765**	**0.81987**	**1.9876**
13	Harris corner detection	38.5821	1488.57844	0.70253	26.60862	0.80753	1.4547
14	ORB	36.8765	1359.87625	0.72431	27.89765	0.83245	1.5432

15	SIFT	35.4536	592.55730	0.75326	29.75321	0.85423	1.85231
16	**Proposed Methodology**	**33.5432**	**1125.14627**	**0.76549**	**30.8765**	**0.86573**	**1.98654**

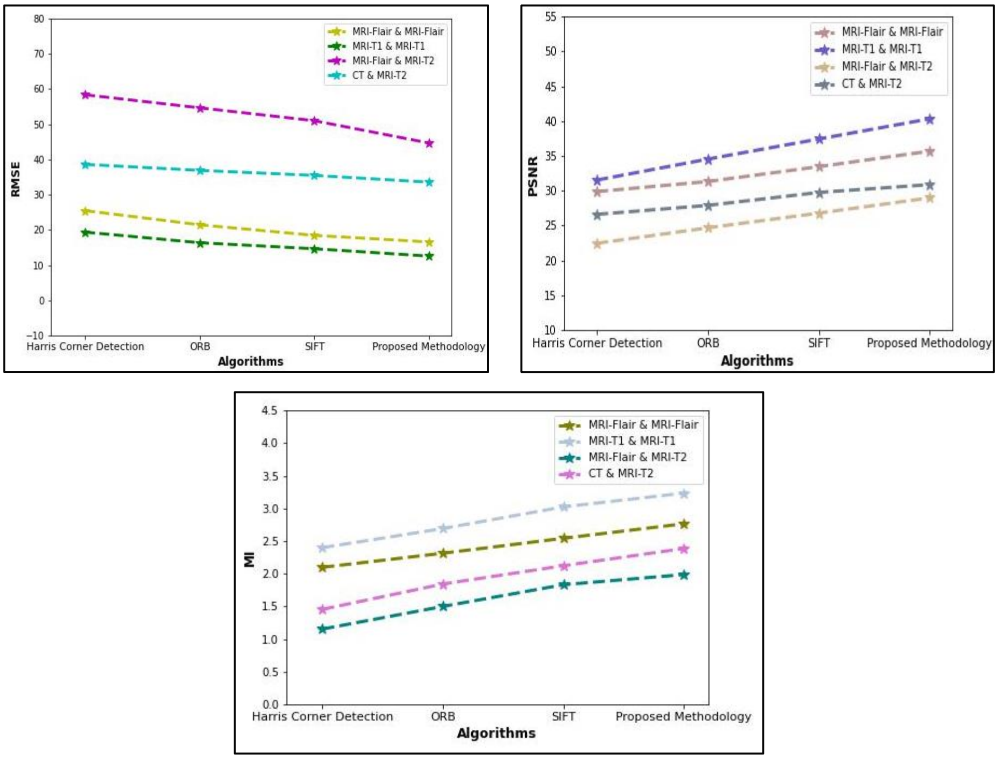


Figure 6.2: Comparison of various similarity measures for medical images with proposed methodology

6.2.1.2. Comparison Results

The proposed scheme has been evaluated against monomodal and multimodal image registration techniques by Zheng *et al.* [49] and Hamzehei *et al.* [56] . Their method utilized SIFT and SURF feature extraction algorithm integrates with maximum likelihood estimation sample consensus (MLESAC) on MR-T1and CT images from Kaggle datasets [114], [115] as illustrated in Figure 6.4. The proposed scheme employs SURF feature extraction with ATSRC outlier, yielding higher MI and CC values as shown in Tables 6.3 and 6.4. The proposed scheme achieves superior performance due to the robust feature extraction capabilities of the SURF architecture and effective outlier elimination using ATSRC, resulting in precise and accurate registration.

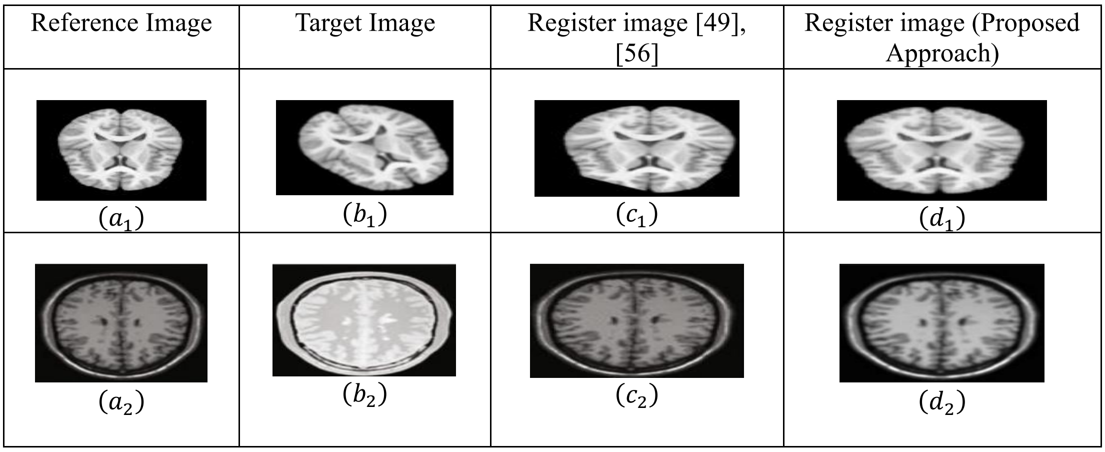


Figure 6.3: Experimental comparison of [49], [56] and proposed approach for monomodal images from dataset [115]

Table 6.3: Comparison of MI for proposed methodology with existing scheme [56]

Transformation \ Evaluation metric	MI
MLESAC (Affine)	2.35
MLESAC (Rigid)	1.53
MLESAC (Rotated)	1.54
MLESAC (Translated)[56]	2.53
Proposed Approach	**3.654**

Table 6.4: Evaluation of proposed methodology with existing scheme [49] corresponding to MI and CC

Methods \ Evaluation metric	MI	CC
P1-SURF[49]	2.2577	0.9219
Proposed Methodology	**3.17389**	**0.95475**

6.2.2. Conclusion

In this section, an integration of SURF with alpha trimmed spatial relation correspondence method for robust image registration is presented. In this analysis, different feature detection algorithms namely Harris corner detection, ORB, SIFT, and SURF along with the alpha trimmed spatial relation correspondence outlier detection method are discussed. Experimental results demonstrate that SURF integrated with alpha trimmed spatial relation correspondence algorithm outperforms traditional techniques for monomodal and multimodal medical image modalities. The effectiveness of proposed scheme indicates significant improvements in different similarity metrics including MI, RMSE, PSNR, and SSIM leading to optimal image alignment. This is due to the robust feature extraction from reference and target images using SURF and removal of outliers by leveraging ATSRC. Despite its advantages, the proposed scheme does not work on clinical images. In the next section, an integration of deep learning with feature extraction algorithm in monomodal and multimodal modalities on both publicly and clinical datasets will explore.

6.3. VGG19-SURF with Projective Transformation for Medical Image Registration

In this section, hybridization of VGG19 with SURF feature extraction algorithm for monomodal and multimodal medical image registration is explored. This scheme initiates with gaussian filtering for noise reduction followed by normalization leads to maintain smoothing uniformly. The pre-processed input images are fed into VGG19 to extract both low and high-level features. Next, euclidean distance between each pair of feature points is computed across different feature spaces. Dynamic inlier selection enhances feature point alignment between pre-processed fixed and sensed images by iteratively adjusting distance thresholds, either randomly or based on a heuristic measure. After identifying the inliers, TPS interpolation corrects local misalignments among the selected feature points and computes a non-linear transformation. This enables complex warping to align features between images with non-uniform distortions, thus ensuring precise alignment. To improve the reliability of the registered image obtained through CNNs (VGG16 and VGG19), SURF is used to detect and match features with a BF matcher. Additionally, outliers are minimized with the alpha-trimmed spatial resolution correspondence algorithm. Finally, a homography matrix is employed to generate the final registered image, achieving robust alignment between the sensed and VGG19 registered images. This approach consists of a series of steps as detailed in Figure 6.5, which are designed to ensure accurate image alignment.

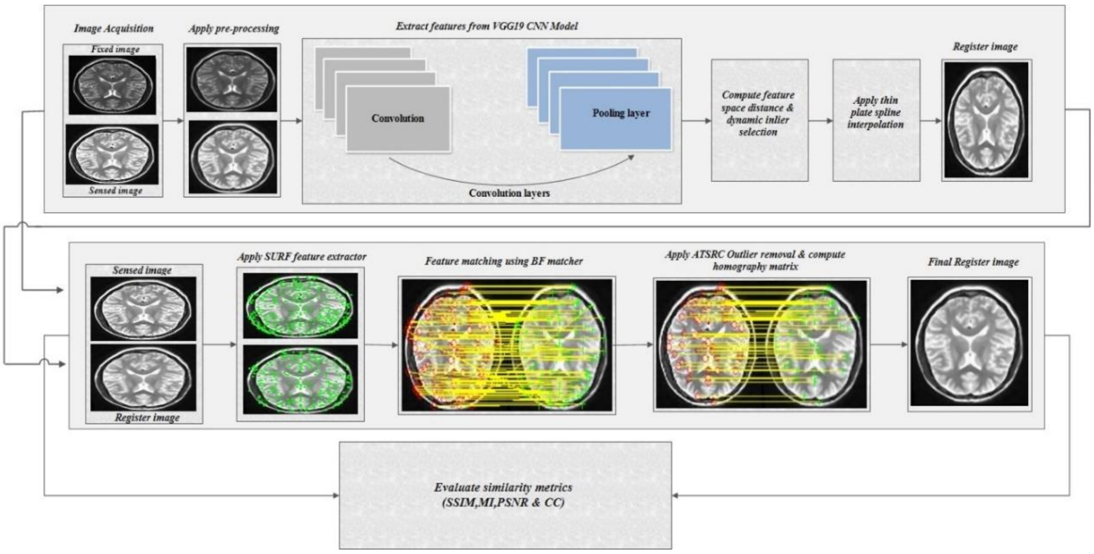


Figure 6.5: Framework of proposed scheme

Algorithm 6.1: Deep learning induced multimodal medical image registration

Input: (I_R, I_T): fixed image I_R and sensed image I_T from dataset D_K

Output: (FI_{Reg}): final register image

Begin

 # Deep learning (CNN)

for each image pair $(\boldsymbol{I_R}, \boldsymbol{I_T}) \in \boldsymbol{D_K}$:

$$I'_R = G(I_R(x,y)), \; I'_T = G(I_T(x,y))$$

$$\widehat{\boldsymbol{I_R}} = \frac{\boldsymbol{G(I_R(x,y)-\min(G(I_R)))}}{\boldsymbol{max(G(I_R))-min(G(I_R))}}, \quad \widehat{\boldsymbol{I_T}} = \frac{\boldsymbol{G(I_T(x,y)-min(G(I_T)))}}{\boldsymbol{max(G(I_T))-min(G(I_T))}}$$

$$\boldsymbol{I_C} \xleftarrow{\boldsymbol{concatenate}} (\widehat{\boldsymbol{I_R}}, \widehat{\boldsymbol{I_T}})$$

$$\boldsymbol{F_C} \xleftarrow{\boldsymbol{apply\ CNN}} \boldsymbol{VGG_{19}}\ (\boldsymbol{I_C})$$

$$d_i(x,y) = \sqrt{\sum_{i=1}^{n} (D_i(x) - D_i(y))^2}$$

$$d(x,y) = \sqrt[2]{d_1(x,y) + d_2(x,y) + d_3(x,y)}$$

$$f(m,n) = a_1 + a_2\, m + a_3 n + \sum_{t=1}^{f} T_t\, df_m^2 \ln df_m$$

$$m_{fp} = m_{fp}(f(m,n), dynamic_{threshold} = \theta)$$

$$tps_{matrix} = tps_{matrix}(m_{fp})$$

$$optimized\_tps_{coff} = optimized\_tps_{coff}\ (tps_{matrix})$$

$$I_{reg} \xleftarrow{tps_{coff}} (I_R, optimized\_tps_{coff})$$

End for

 # SURF feature extraction

$$(keyI_{Reg}, desI_{Reg}) \xleftarrow{Apply_{SURF\_REG}} (I_{reg})$$

$(keyI_T, desI_T) \xleftarrow{Apply_{SURF\_T}} (I_T)$

$matches \xleftarrow{BF_{matcher}} (desI_{Reg}, desI_T)$

$d = \left[\sqrt{\sum_{t=1}^{k}(desI_{Reg}[m] - desI_T[n])^2}\right]$

$\alpha = 0.1$

$best_{matches} \xleftarrow{trimmed_{matches}} (matches, d, \alpha)$

for match in $best_{matches}$:

$mpI_{Reg} = keyI_{Reg}(m)$

$mpI_T = keyI_T(n)$

End for

$H = comp_{hy}(mpI_{Reg}\ ,\ mpI_T)$

for each point $(\boldsymbol{i}, \boldsymbol{j})$ $\boldsymbol{in\ I_{Reg}}$:

$m = \frac{(H[0,0]*i+H[0,1]*j+H[0,2])}{(H[2,0]*i+H[2,1]*j+1)}, n = \frac{(H[1,0]*i+H[1,1]*j+H[1,2])}{(H[2,0]*i+H[2,1]*j+1)}$

$tr_{cor} = tr_{cor}(m, n)$

$FI_{reg} = gt_{reg}(tr_{cor}, I_{Reg})$

End for

End

Step 1: Pre-processing

In this step, two images are randomly selected from both public repository or clinical dataset. Afterwards, gaussian filtering is applied to fixed and sensed images. To enhance contrast and reduce noise, min-max normalization is performed. Then images are resized to 224×224 pixels to reduce computation time, as illustrated in Figure 6.6.

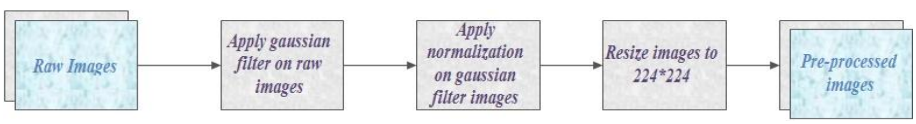


Figure 6.6: Step-by-step image preprocessing flow

Step 2: Extract feature points form CNN model

In this step, extracting high and low-level features from images obtained in Step-1 by feeding them into VGG-19 model as outlined in Algorithm 6.1. This model consists of 5 max pooling layers and 16 convolution layers with a kernel size of 3×3 pixels as shown in Figure 6.7 and Figure 6.8. The max pooling layers of size 2×2 which significantly enhances the extraction of image feature points as illustrated in Figure 6.9 and Figure 6.10.

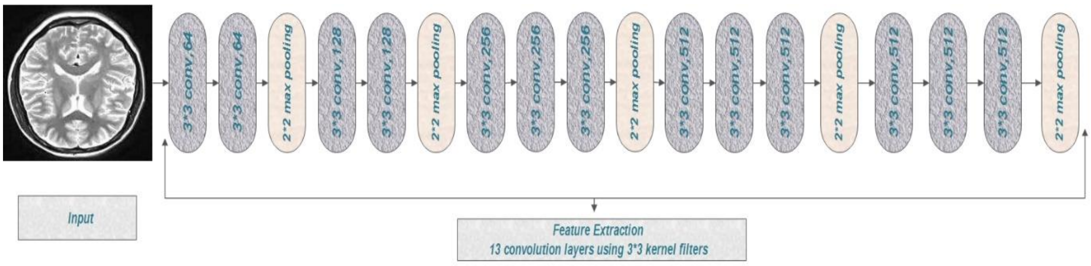


Figure 6.7: Process of feature extraction utilizing VGG16 CNN

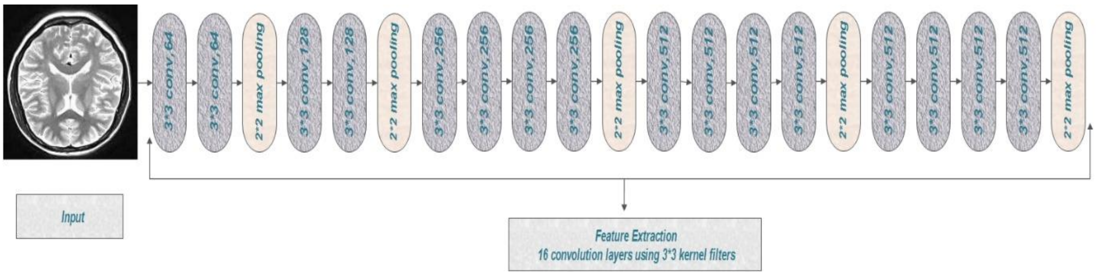


Figure 6.8: Stepwise mechanism for feature extraction using VGG19 CNN

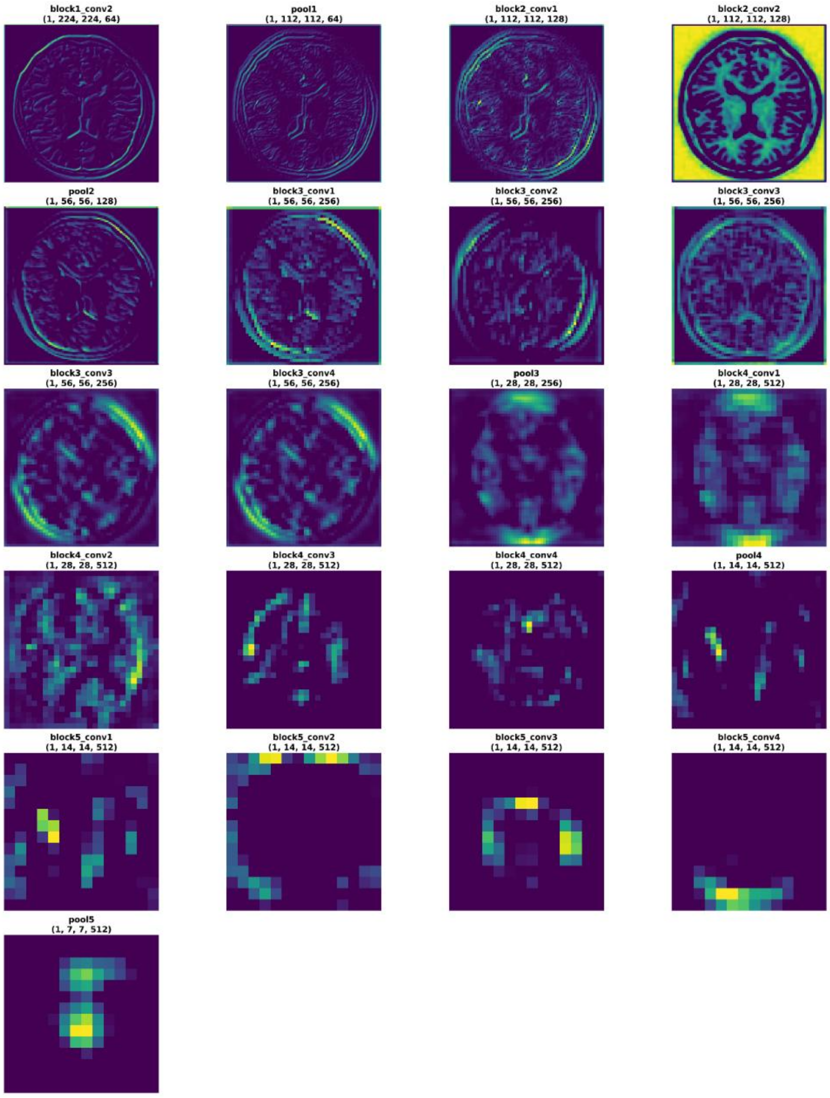


Figure 6.9: Layer wise feature extraction representation of fixed image using VGG19

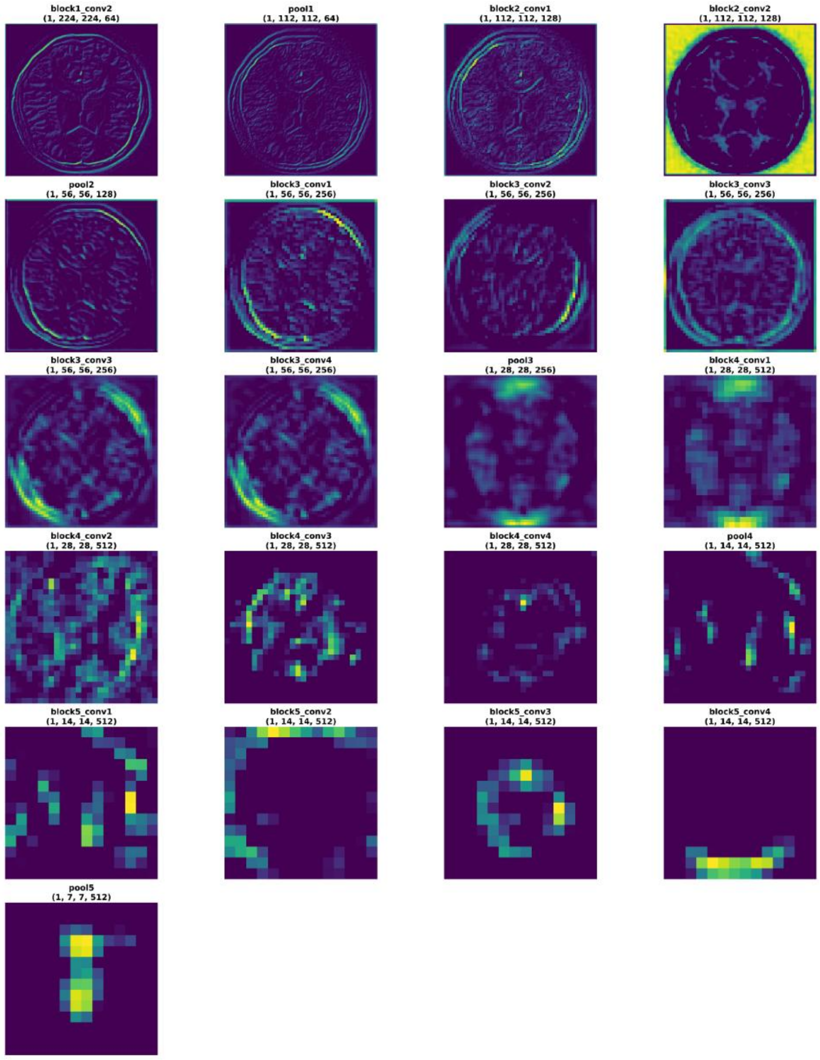


Figure 6.10: An in-depth exploration of extracting features from sensed image across all layers

Step 3: Compute feature space distance

After extracting feature points from VGG19 model in Step 2, the feature space distances $\left(d_1(m,n), d_2(m,n), d_3(m,n)\right)$ are computed using euclidean distance as mathematically described in Equation (6.2). Then, weighted sum of three distance values $d(m,n)$ for feature points m and n in the pre-processed fixed and sensed images as shown in Equation (6.1).

$$d(m,n) = 1.414 * d_1(m,n) + d_2(m,n) + d_3(m,n) \tag{6.1}$$

$$d_t(m,n) = \sqrt{\sum_{t=1}^{k}\left(D_t(m) - D_t(n)\right)^2} \tag{6.2}$$

The feature space distance $d_t(m,n)$ is computed as the square root of the weighted sum of three distances namely $d_1(m,n), d_2(m,n)\ and\ d_3(m,n)$. Each $d_1(m,n)$ or

$d_t(m,n)$ represents the euclidean distance between corresponding feature vectors. Following this, fixed feature point m and sensed feature point n are matched based on the criterion $\mathrm{d}(k,n) < \theta . d(m,n)$, where θ is an adaptively changing dynamic threshold during pre-matching. Here, $\mathrm{d}(k,n)$ denotes the feature space distance between the transformed feature point (k) in the fixed image and the feature point (n) in the sensed image, while $d(m,n)$ represents the distance between matching feature points (m) and (n).

Step 4: Perform dynamic inlier selection

In this step, dynamic inlier selection enhances the alignment of feature points between sensed and fixed points by iteratively adjusting distance thresholds, either randomly or using a heuristic measure. This approach improves matching accuracy by effectively filtering inliers and reducing outliers, thus enhancing alignment of images.

Step 5: Generate registered image with TPS

After identifying inliers through dynamic inlier selection, TPS interpolation is used to correct local misalignments among the selected feature points, mapping points in the sensed image to the corresponding points in fixed image. The TPS transformation coefficients (a_1, a_2, a_3, T_t) are adjusted, as depicted in Equation (6.3). Furthermore, TPS method computes a non-linear transformation, enabling complex warping to align features between images with non-uniform distortions. The TPS transformation function $f(m,n)$ is given by:

$$f(m,n) = a_1 + a_2\, m + a_3 n + \sum_{t=1}^{f} T_t\, df_m^2 \ln df_m \tag{6.3}$$

Here, a_1, a_2, a_3 and T_t are the TPS interpolation coefficients, with f representing the total number of feature points, and df_m denoting the distance between corresponding feature points in the fixed and sensed images. This method enhances registration robustness and accuracy by fine-tuning adjustments based on feature point alignment, yielding a registered image that accurately aligns the fixed and sensed images as displayed in Figure 6.11(Col 5, $\mathrm{e}_1 - \mathrm{e}_{10}$).

Step 6: Feature detection

After obtaining the registered image through VGG19 method, further refinement is achieved using SURF feature extraction. This process identifies different keypoints (features) in both sensed and VGG19 registered images as illustrated in Figure 6.13(Col 1, $a_1 - a_{10}$) and Figure 6.13 (Col 2, $b_1 - b_{10}$) respectively.

Step 7: Feature matching

In this step, brute force matcher is employed for feature matching between sensed and registered features images obtained from Step 6 as shown in Figure 6.13 (Col 3, c_1-c_{10}). The determination of optimal matches involves calculating the Euclidean distance (D) as depicted in Equation (6.4), quantifying the spatial separation of their feature vectors.

$$D = (V1, V2) = \sqrt{\sum_{m=0}^{P}(v_1[m] - v_2[m])^2} \tag{6.4}$$

Here, the dimensionality of the feature vectors is represented by P, while $v_1[m]$ and $v_2[m]$ denote the m^{th} elements of the feature vectors for sensed image $V1$ and register image $V2$ respectively.

Step 8: Homography matrix computation

After detecting the total matches between the images in Step 7, an alpha-trimmed spatial relation correspondence scheme is utilized for outlier elimination to enhance the accuracy of feature point correspondences as shown in Figure 6.13 (Col 4, d_1-d_{10}). Then, homography matrix is employed to generate final registered image between the sensed and VGG19 registered images as shown in Figure 6.13 (Col 5, e_1-e_{10}). Moreover, homography matrix requires eight degrees of freedom (DOF) to operate with homogeneous coordinates and the last parameter of the homography matrix (HM) is typically set to one as shown in Equation (6.5).

$$HM = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \tag{6.5}$$

Determining homography matrix requires at least four points are (i_1, j_1) , (i_2, j_2), (i_3, j_3), (i_4, j_4) as outlined in Equation (6.6).

$$\begin{bmatrix} i_1 & j_1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & i_1 & j_1 & 1 \\ i_2 & j_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i_2 & j_2 & 1 \\ i_3 & j_3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i_3 & j_3 & 1 \\ i_4 & j_4 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & i_4 & j_4 & 1 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \end{bmatrix} = \begin{bmatrix} i'_1 \\ j'_1 \\ i'_2 \\ j'_2 \\ i'_3 \\ j'_3 \\ i'_4 \\ j'_4 \end{bmatrix} \tag{6.6}$$

Equation (6.7) and Equation (6.8) are employed in the final step to derive the image coordinates (m'_i, n'_i).

$$m'_i = \frac{(h_{11i} + h_{12j} + h_{13})}{(h_{31i} + h_{32j} + 1)} \tag{6.7}$$

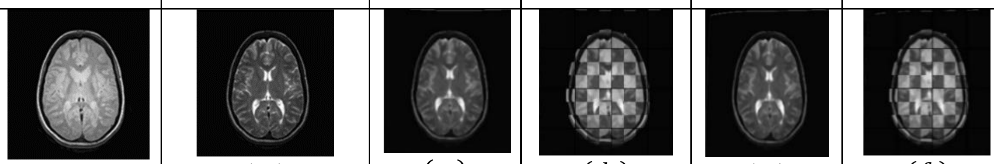
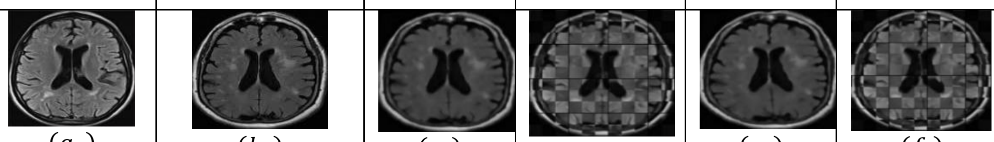
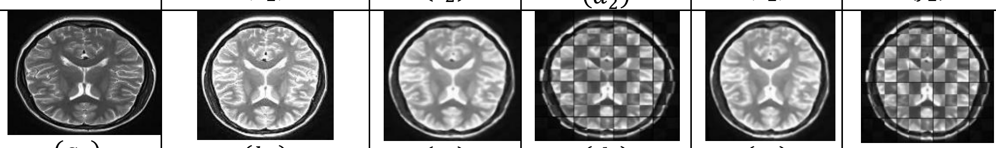
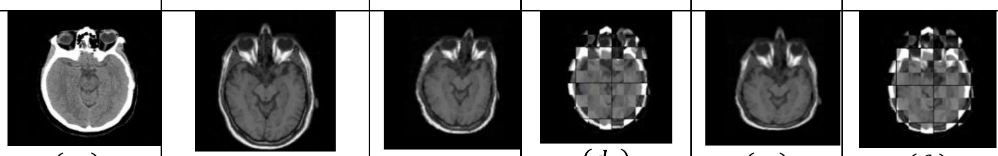
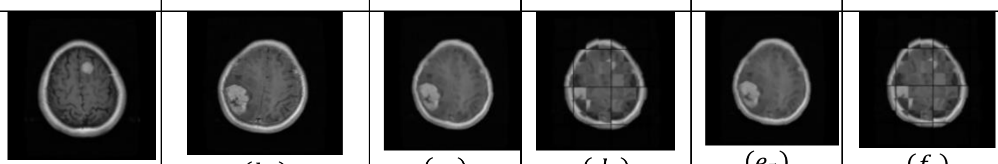
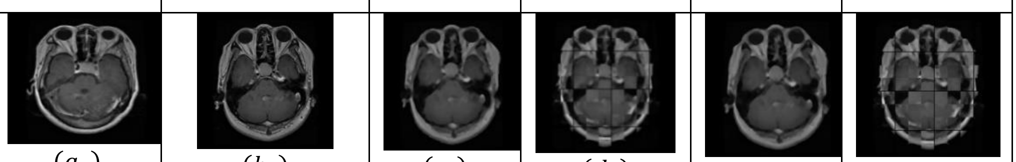
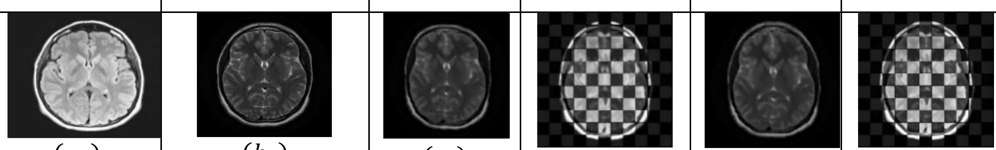
$$n'_i = \frac{(h_{11i} + h_{12j} + h_{13})}{(h_{31i} + h_{32j} + 1)} \tag{6.8}$$

6.3.1. Experimental and Comparative Analysis

In this section, the effectiveness of proposed scheme is analysed using Kaggle [112], [114], [117], [132], [134] and in house clinical datasets, with different image modalities and several evaluation metrics namely RMSE, CC, SSIM, and PSNR.

6.3.1.1. Experimental Analysis

The effectiveness of proposed scheme is assessed through experiments using Keras with TensorFlow backend on a system featuring an NVIDIA GeForce GTX GPU, 16.0 GB RAM, and an Intel Core i7-1165G7 CPU at 2.80 GHz. The implementation is carried out on Python 3.9.4 with Windows 11, utilizing datasets from public repositories [112], [114], [117], [132], [134] and clinical sources collected from Fortis Hospital Mohali and Paras Hospital Panchkula, India for validation. Initially, fixed and sensed medical brain images of sized 224×224 pixels undergo min-max normalization and gaussian filtering as depicted in Figure 6.11 (Col 1, a_1-a_{10}) and Figure 6.11 (Col 2, b_1-b_{10}) respectively. Then, pre-processed input images are fed into VGG19 model for extracting high and low-level features, as illustrated in Figure 6.9 and Figure 6.10 respectively. The feature space distance measures dissimilarity between feature points, yielding a reliable matching metric. Consequently, the transformation matrix is determined using TPS interpolation, ensuring precise alignment between fixed and sensed images for the VGG16 and VGG19 models as demonstrated in Figure 6.11 (Col 3, c_1-c_{10}) and Figure 6.11 (Col 5, e_1-e_{10}). The accuracy and effectiveness of the registration process are validated using a checkerboard pattern, which highlights distortions or misalignments for VGG16 and VGG19 in Figure 6.11 (Col 4, d_1- d_{10}) and Figure 6.11 (Col 6, f_1- f_{10}) respectively.

S. No	Fixed image	Sensed image	Register image (VGG16)	Checkboard image (VGG16)	Register image (VGG19)	Checkboard image (VGG19)
1	 (a_1)	(b_1)	(c_1)	(d_1)	(e_1)	(f_1)
2	 (a_2)	(b_2)	(c_2)	(d_2)	(e_2)	(f_2)
3	 (a_3)	(b_3)	(c_3)	(d_3)	(e_3)	(f_3)
4	 (a_4)	(b_4)	(c_4)	(d_4)	(e_4)	(f_4)
5	 (a_5)	(b_5)	(c_5)	(d_5)	(e_5)	(f_5)
6	 (a_6)	(b_6)	(c_6)	(d_6)	(e_6)	(f_6)
7	 (a_7)	(b_7)	(c_7)	(d_7)	(e_7)	(f_7)

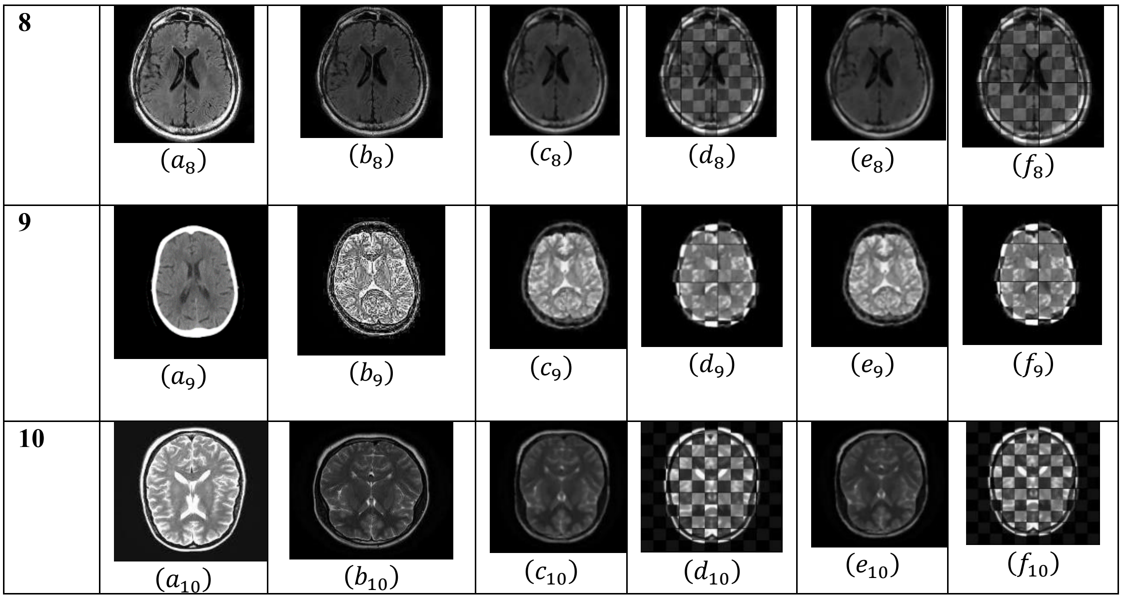


Figure 6.11: Results of multimodal medical images sourced from Kaggle dataset (a) Fixed image $(a_1 - a_{10})$ (b) Sensed image $(b_1 - b_{10})$ (c) VGG16 register image $(c_1 - c_{10})$ (d) VGG16 checkboard image $(d_1 - d_{10})$ (e) VGG19 register image $(e_1 - e_{10})$ (f) VGG19 checkboard image $(f_1 - f_{10})$ through extensive experiments

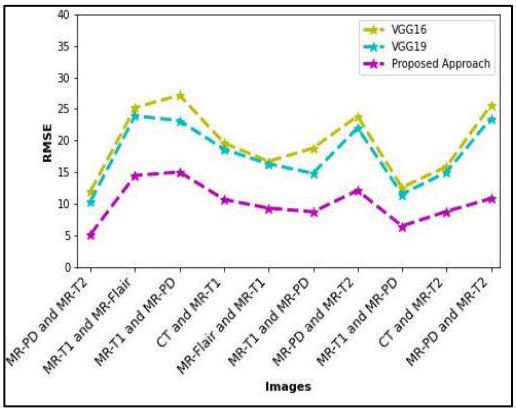
Evaluation conducted across multimodal medical images demonstrate the superior performance of proposed approach as compared to VGG16 and VGG19, as depicted in Table 6.5. The integration of SURF with VGG19 significantly improves its effectiveness as highlighted by the bold values in the proposed scheme as tabulated in Table 6.5 in terms of higher values of SSIM, PSNR, CC and lower values of RMSE and MSE.

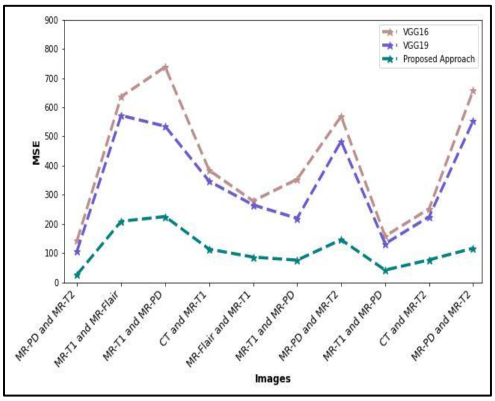
Table 6.5: Comparative evaluation of numerous similarity indices across three techniques: VGG16, VGG19 and proposed scheme

S. No	Images	Algorithm	RMSE	MSE	SSIM	PSNR	CC
1.	MR-PD and MR-T2	VGG16	11.994569	143.86497	0.860807	34.03808	0.9727739
		VGG19	10.3929664	108.0137	0.886991	34.6968	0.9797634
		Proposed scheme	**5.146179**	**26.4831583**	**0.8963374**	**38.437298**	**0.994838**
2.	MR-T1 and MR-Flair	VGG16	25.22458988	636.27993	0.8066257	31.238862	0.93041205
		VGG19	23.914663	571.91111	0.8191049	32.28359	0.9474801
		Proposed scheme	**14.4584**	**209.045331**	**0.83128647**	**33.975340**	**0.97198594**
3.	MR-T1 and MR-PD	VGG16	27.1730781	738.37617	0.76442818	30.87104	0.941904
		VGG19	23.13293	535.132712	0.791750	31.179722	0.956929
		Proposed scheme	**15.0112**	**225.336125**	**0.8374371**	**33.50120124**	**0.9635435**
4.	CT and MR-T1	VGG16	19.5981	384.085524	0.8562362	32.7041249	0.9613452

		VGG19	18.61293	346.44116	0.8731245	33.345216	0.9743632
		Proposed scheme	**10.6502**	**113.42676**	**0.898432**	**34.93516**	**0.983421**
5.	MR-Flair and MR-T1	VGG16	16.6921857	278.629065	0.82770910	34.6691737	0.9592526
		VGG19	16.2853740	265.21340	0.834141373	34.837625	0.961142
		Proposed scheme	**9.289468**	**86.2942157**	**0.8999728**	**35.726003**	**0.9710693**
6.	MR-T1 and MR-PD	VGG16	18.7956332	353.275829	0.82504288	32.771237	0.9339187
		VGG19	14.80797447	219.27610	0.8751107	33.380234	0.95937383
		Proposed scheme	**8.69846**	**75.6632064**	**0.89615193**	**35.800669**	**0.9773345**
7.	MR-PD and MR-T2	VGG16	23.8567293	569.14353	0.84293566	32.737780	0.858698
		VGG19	21.9901945	483.568654	0.85364008	33.512361	0.8694104
		Proposed scheme	**12.08201**	**145.974966**	**0.8886187**	**34.100217**	**0.884712**
8	MR-T1 and MR-PD	VGG16	12.5682789	157.9616	0.82997994	33.341547	0.96973974
		VGG19	11.48999519	132.019989	0.84361977	34.4854853	0.9701106
		Proposed scheme	**6.453017**	**41.6414284**	**0.89070639**	**36.3452907**	**0.9858878**
9	CT and MR-T2	VGG16	15.87534	252.026420	0.8653271	35.24516	0.9512343
		VGG19	14.98632	224.589787	0.8763124	36.186423	0.965843
		Proposed scheme	**8.76543**	**76.832763**	**0.8973252**	**38.63214**	**0.9823453**
10.	MR-PD and MR-T2	VGG16	25.6402277	657.421277	0.8832141	32.508742	0.83760417
		VGG19	23.5186719	553.127929	0.8953241	33.6591657	0.86357087
		Proposed scheme	**10.79826**	**116.602419**	**0.9135261**	**34.17617**	**0.8987133**

The comparison results illustrating the performance of proposed scheme using VGG19 and SURF with VGG16 and VGG19 on multimodal medical images based on similarity metrics are shown in Figure 6.12.





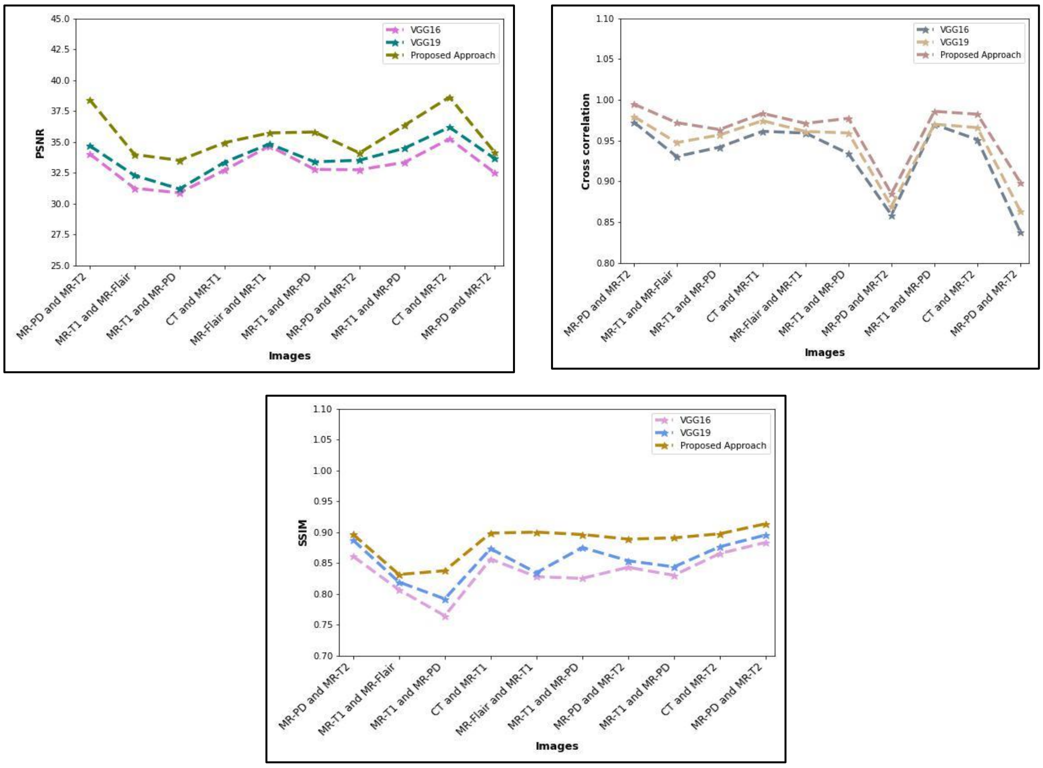
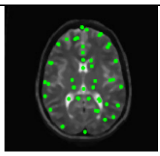
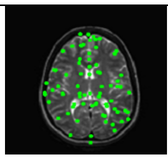
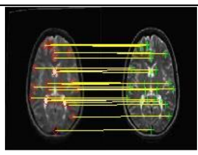
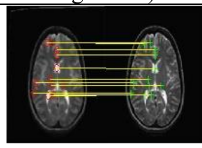
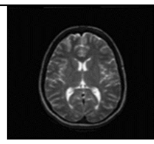
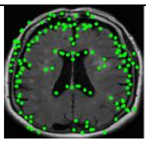
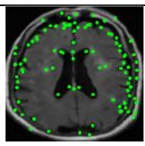
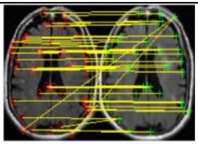
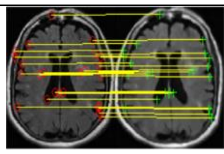
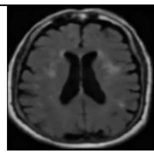
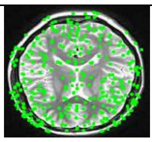
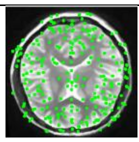
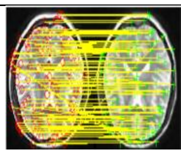
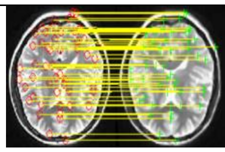
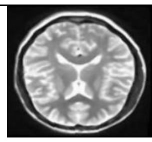
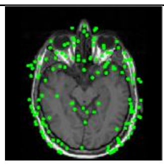
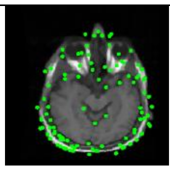
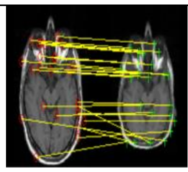
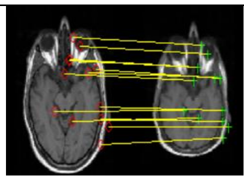
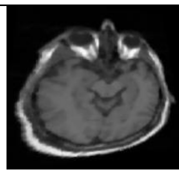
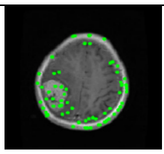
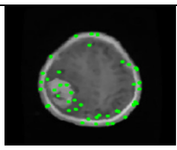
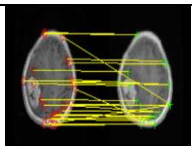
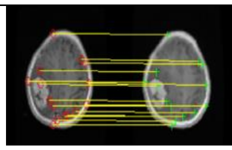
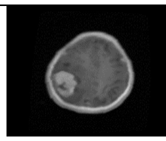
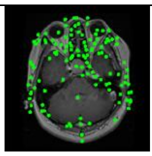
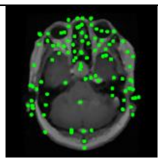
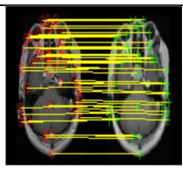
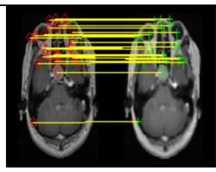
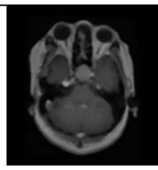
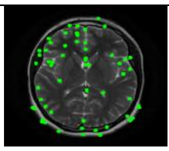
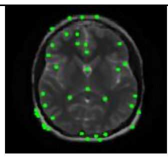
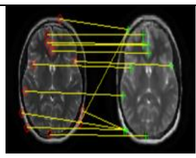
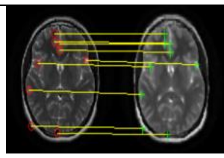
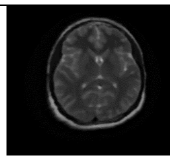
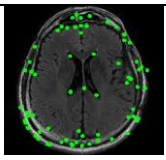
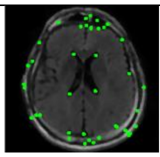
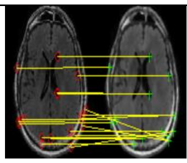
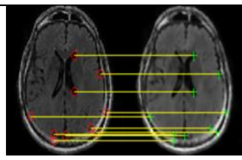
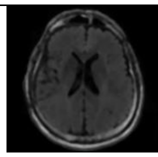


Figure 6.12: Comparative computational assessment of the proposed scheme against VGG16 and VGG19 employing various similarity metrics on multimodal medical images.

To improve the registration performance of VGG19, SURF is utilized to detect features in both sensed and registered images as illustrated in Figure 6.13 (Col 1, a_1-a_{10}) and Figure 6.13 (Col 2, b_1-b_{10}) respectively. Subsequently, brute force matcher is employed to establish correspondences between the identified keypoints as shown in Figure 6.13 (Col 3, c_1-c_{10}). Additionally, an alpha-trimmed spatial relation correspondence method is utilized to remove outliers and improve the accuracy of feature point correspondences as depicted in Figure 6.13 (Col 4, d_1- d_{10}). Then, homography matrix is applied to generate the registered image, ensuring an accurate alignment between fixed and registered images as displayed in Figure 6.13 (Col 5, e_1-e_{10}). The quantitative analysis of multimodal images is tabulated in Table 6.6.

S. No	Sensed image key points	Register image key points	Total matches	Good matches (Inliers using ATSRC algorithm)	Register image
1.	 (a_1)	 (b_1)	 (c_1)	 (d_1)	 (e_1)
2	 (a_2)	 (b_2)	 (c_2)	 (d_2)	 (e_2)
3	 (a_3)	 (b_3)	 (c_3)	 (d_3)	 (e_3)
4	 (a_4)	 (b_4)	 (c_4)	 (d_4)	 (e_4)
5	 (a_5)	 (b_5)	 (c_5)	 (d_5)	 (e_5)
6	 (a_6)	 (b_6)	 (c_6)	 (d_6)	 (e_6)
7	 (a_7)	 (b_7)	 (c_7)	 (d_7)	 (e_7)
8	 (a_8)	 (b_8)	 (c_8)	 (d_8)	 (e_8)

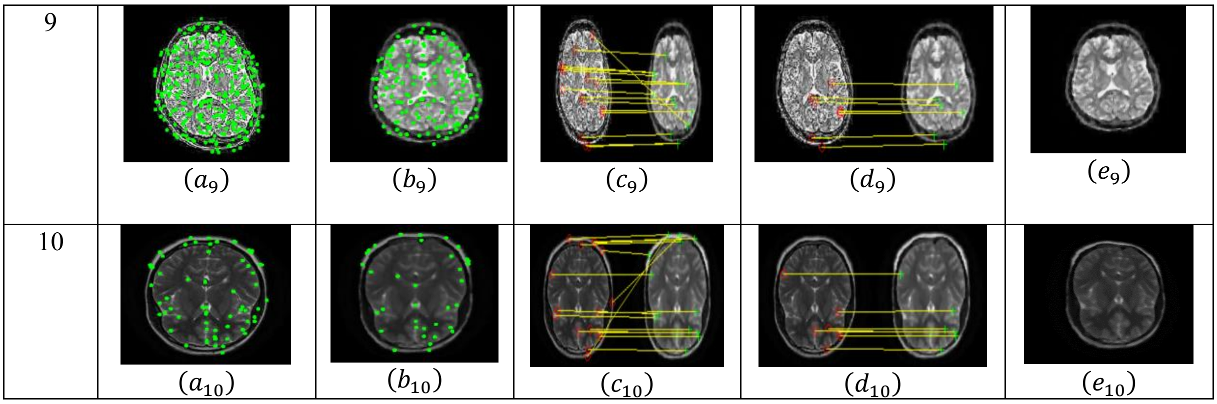


Figure 6.13: Experimental outcome on multimodal medical images include (a) Extraction of key points from fixed image $(a_1 - a_{10})$ (b) Extraction of key points from sensed image $(b_1 - b_{10})$ (c) Total matches achieved employing BF matcher $(c_1 - c_{10})$ (d) Identification of good matches using ATSRC $(d_1 - d_{10})$ (e) Register image $(e_1 - e_{10})$

Table 6.6: Experiment outcomes performed on various intermodal medical images gathered from Kaggle dataset

S. No	Sensed image key points	Register image key points	Total matches using BF matcher	Good matches after ATSRC
1	45	74	24	12
2	135	98	46	24
3	304	216	103	51
4	138	102	25	13
5	64	43	25	15
6	126	99	54	30
7	59	37	13	9
8	69	41	19	10
9	295	143	15	7
10	65	35	16	7

The difference image between target and registered image visually demonstrates the spatial transformations required for accurate alignment, thereby evaluating the effectiveness of the registration process, as depicted in Figure 6.14.

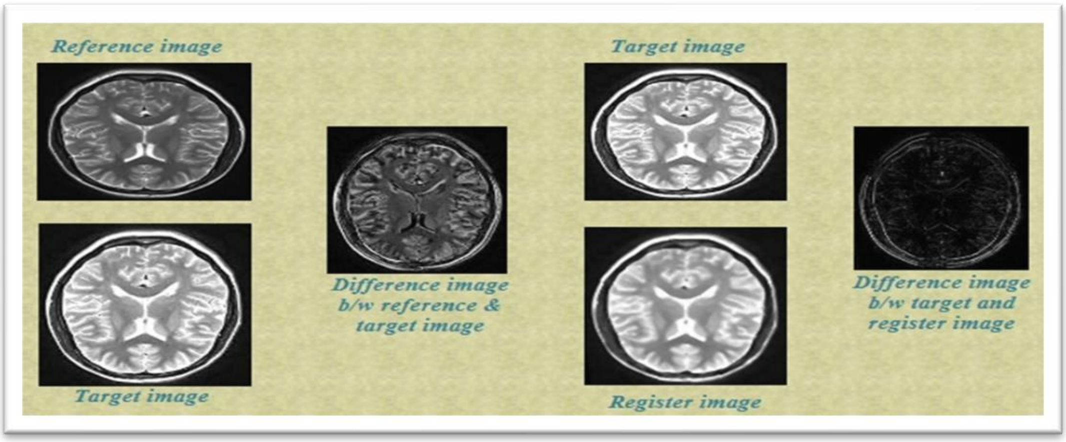
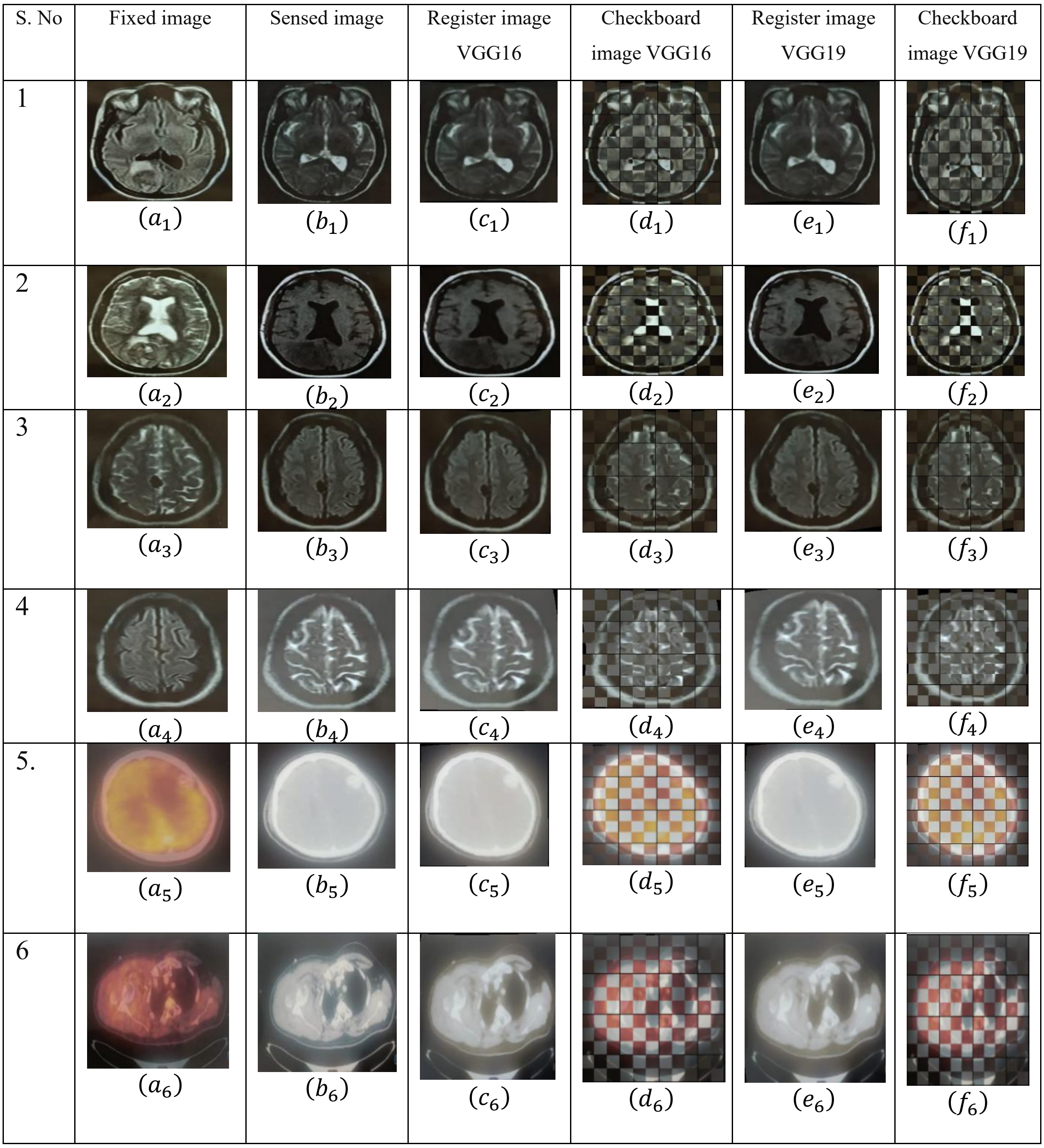


Figure 6.14: Registration and difference of brain images acquired through multiple modalities

The effectiveness of proposed scheme has been evaluated on monomodal and multimodal medical images obtained from clinical datasets from Fortis Hospital

Mohali and Paras Hospital Panchkula, India. The quality of these images is enhanced by pre-processing the fixed and sensed images, as depicted in Figure 6.15 (Col 1, a_1-a_9, and Col 2, b_1-b_9). These pre-processed images are fed into VGG19 model to capture both high and low-level features, as shown in Figure 6.9 and Figure 6.10. The feature space distance measures dissimilarity between feature points, thereby improving matching accuracy. Thin plate spline interpolation ensures precise alignment between fixed and sensed images for both VGG16 and VGG19 models, as illustrated in Figure 6.15 (Col 3, c_1-c_9, and Col 5, e_1-e_9). This process is validated by a checkboard pattern to detect distortions or misalignments, as shown in Figure 6.15 (Col 4, d_1-d_9, and Col 6, f_1-f_9).



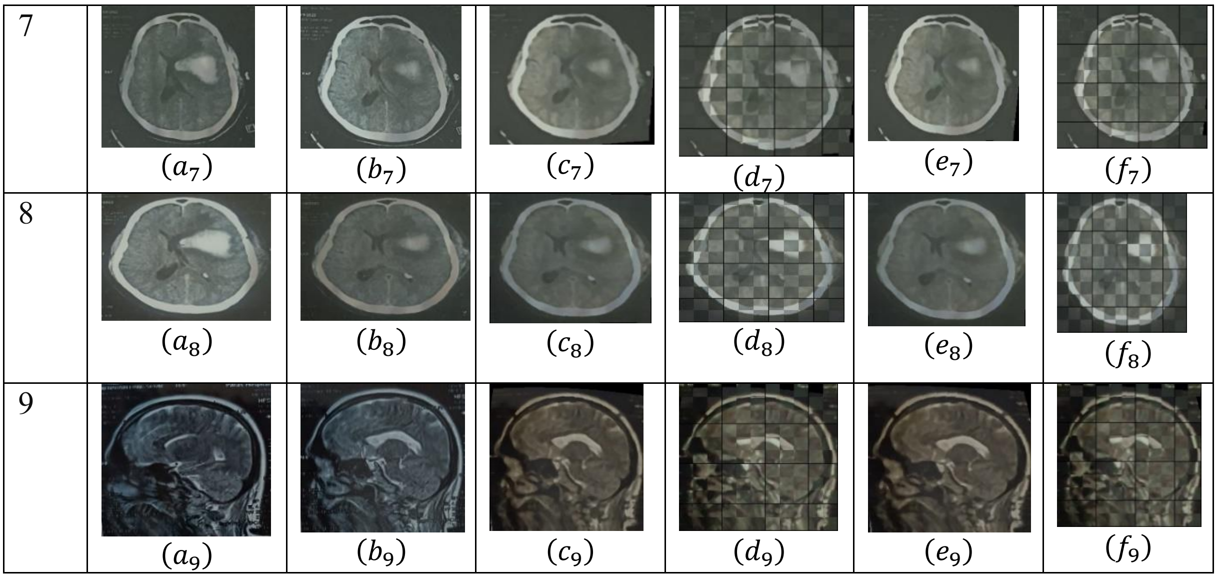


Figure 6.15: Experiments conducted on intramodal and intermodal medical images sourced from clinical datasets (a) Fixed image $(a_1 - a_9)$ (b) Sensed image $(b_1 - b_9)$ (c) VGG16 register image $(c_1 - c_9)$ (d) VGG16 checkboard image$(d_1 - d_9)$ (e) VGG19 register image $(e_1 - e_9)$ (f) VGG19 checkboard image $(f_1 - f_9)$

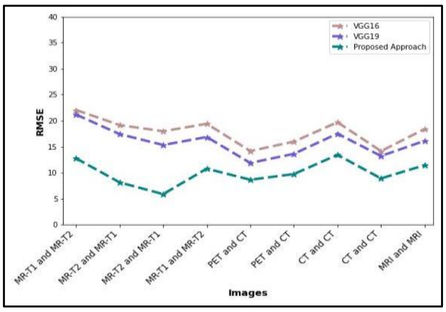
The proposed scheme shows better performance as compared to VGG16 and VGG19 on both multimodal and monomodal medical images numbered from one to six and seven to nine respectively in column 1 of Table 6.7. Slices $(a_1 - a_8)$ are coronal, and slice (a_9) is sagittal

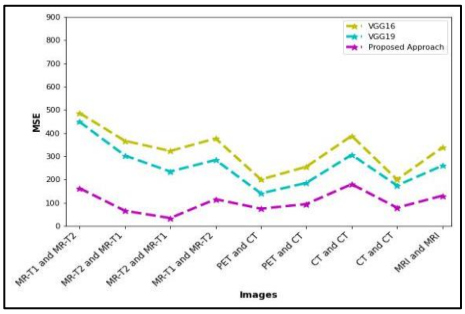
Table 6.7: Experimental outcome on intramodal and intermodal medical images obtained from clinical dataset employing different similarity metric

S. No	Images	Algorithm	RMSE	MSE	SSIM	PSNR	CC
1	MR-T1 and MR-T2	VGG16	22.055361	486.4389	0.8389496	30.19423	0.916321
		VGG19	21.182854	448.7133	0.8460740	31.5939	0.921984
		Proposed scheme	12.767959	163.0207	0.8621166	32.852484	0.946302
2	MR-T2 and MR-T1	VGG16	19.154667	366.90126	0.855877	32.61350	0.93434
		VGG19	17.4209570	303.489742	0.887156	33.440380	0.958798
		Proposed scheme	8.12345	65.990439	0.9173271	34.94380	0.964762
3	MR-T2 and MR-T1	VGG16	17.969750	322.911915	0.8643999	34.825872	0.96739
		VGG19	15.3239924	234.8247430	0.8827692	35.544528	0.971511
		Proposed scheme	5.8732145	34.4946485	0.9293415	37.281345	0.987321
4	MR-T1 and MR-T2	VGG16	19.412314	376.837934	0.842134	31.76432	0.932652
		VGG19	16.865243	284.436421	0.863245	32.98501	0.955421

		Proposed scheme	10.76324	115.8473352	0.904321	34.864301	0.972163
5	PET and CT	**VGG16**	14.157400	200.43197943	0.8728082	33.7288895	0.9964075
		VGG19	11.8628823	140.72797751	0.8848065	34.1935445	0.9974507
		Proposed scheme	8.6421341	74.6864818	0.9123145	35.765432	0.9987643
6	PET and CT	**VGG16**	15.9760843	255.23527184	0.75887961	32.3346512	0.9926628
		VGG19	13.6306303	185.79408482	0.7745100	32.4935472	0.9946475
		Proposed scheme	9.7345621	94.761699	0.812453	34.672134	0.9987231
7	CT and CT	**VGG16**	19.7027086	388.196727	0.7201682	33.60530533	0.9820879
		VGG19	17.509906	306.59681	0.7400097	35.86255801	0.9858754
		Proposed scheme	13.426781	180.278448	0.8054321	37.896453	0.997689
8	CT and CT	**VGG16**	14.1468692	200.1339086	0.79196303	33.39277	0.9870018
		VGG19	13.2082967	174.4591039	0.81431548	34.7053785	0.9886942
		Proposed scheme	8.892141	79.070171	0.8423413	35.34521	0.9976321
9	MRI and MRI	**VGG16**	18.395753824	338.4037587	0.8219534	30.8438971	0.966542
		VGG19	16.12452652	260.000355	0.8465321	32.7632542	0.976213
		Proposed scheme	11.43627145	130.78830467	0.874352	34.8752912	0.989321

The comparison results illustrating the performance of proposed scheme using VGG19 and SURF with VGG16 and VGG19 on both monomodal and multimodal medical images based on similarity metrics are shown in Figure 6.16.





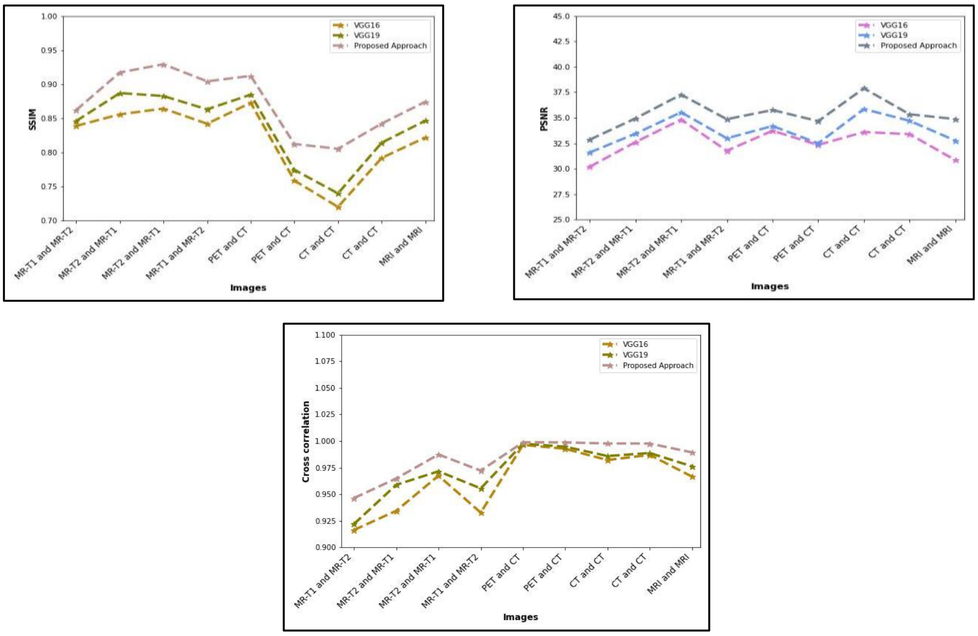
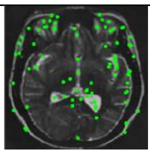
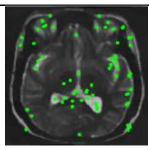
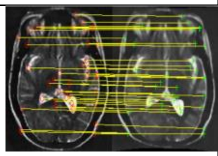
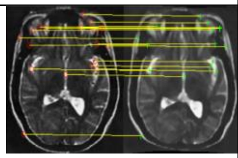
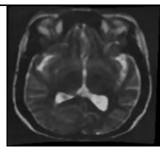
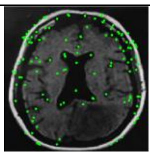
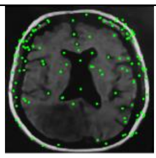
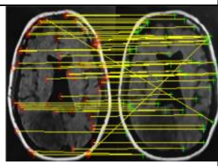
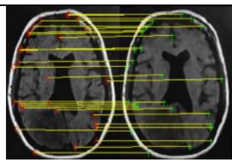
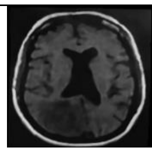
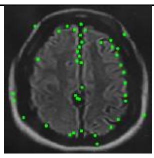
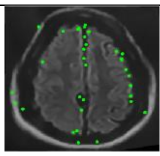
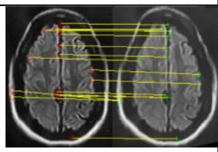
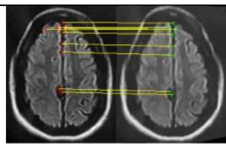
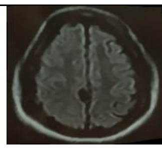
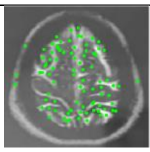
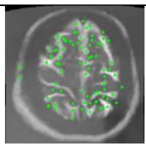
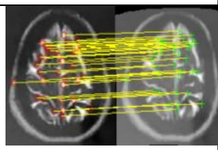
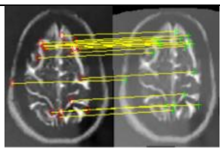
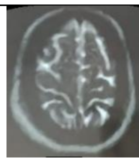
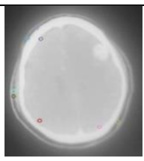
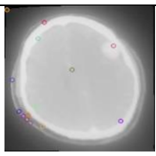
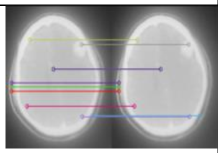
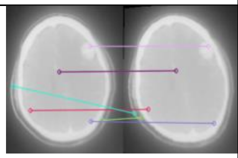
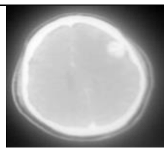
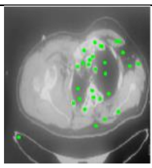
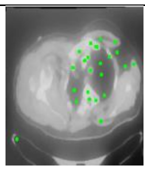
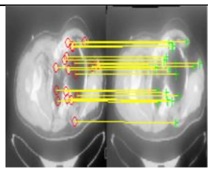
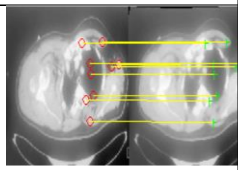
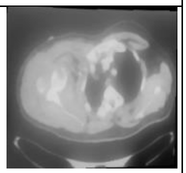
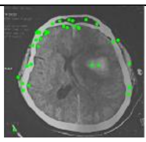
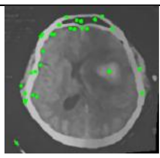
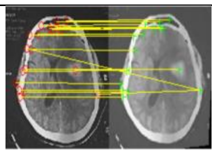
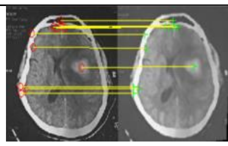
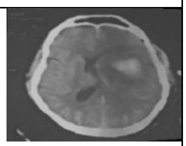
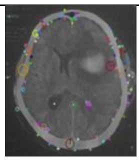
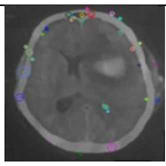
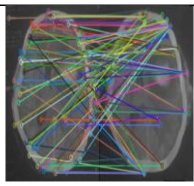
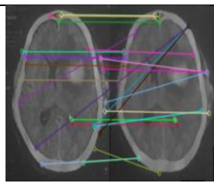
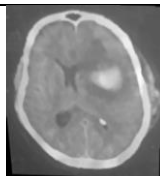


Figure 6.16: Comparative assessment of proposed scheme against traditional method employing various similarity metrics for anatomical and functional images

To enhance the VGG19 registration performance, SURF is utilized to identify features in both sensed and registered images, as shown in Figure 6.17 (Col 1, a_1-a_9, and Col 2, b_1-b_9). Moreover, a brute force matcher is then used to establish correspondences between detected key points as demonstrated in Figure 6.17 (Col 3, c_1-c_9). An alpha-trimmed spatial relation correspondence scheme eliminates outliers, enhancing feature point correspondence accuracy, and resulting in more precise matches, as shown in Figure 6.17(Col 4, d_1-d_9). Finally, the homography matrix achieves geometrically accurate alignment between fixed and registered images, as illustrated in Figure 6.17 (Col 5, e_1-e_9). Quantitative analysis of both monomodal and multimodal medical images are tabulated in Table 6.8 and graphical representation is depicted in Figure 6.17.

S. No.	Sensed image key points	Register image key points	Total matches	Good matches (Inliers using ATSRC algorithm)	Register image
1.	 (a_1)	 (b_1)	 (c_1)	 (d_1)	 (e_1)
2.	 (a_2)	 (b_2)	 (c_2)	 (d_2)	 (e_2)
3.	 (a_3)	 (b_3)	 (c_3)	 (d_3)	 (e_3)
4.	 (a_4)	 (b_4)	 (c_4)	 (d_4)	 (e_4)
5	 (a_5)	 (b_5)	 (c_5)	 (d_5)	 (e_5)
6	 (a_6)	 (b_6)	 (c_6)	 (d_6)	 (e_6)
7	 (a_7)	 (b_7)	 (c_7)	 (d_7)	 (e_7)
8	 (a_8)	 (b_8)	 (c_8)	 (d_8)	 (e_8)

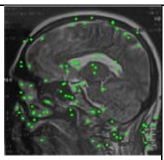
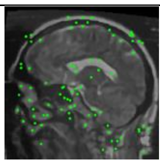
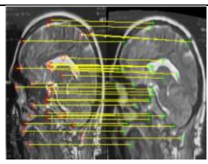
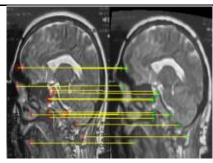
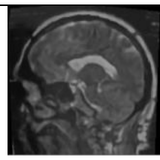
9	 (a_9)	 (b_9)	 (c_9)	 (d_9)	 (e_9)

Figure 6.17: Experiments carried out on intramodal and intermodal medical images include (a) Key points extraction from fixed image $(a_1 - a_9)$ (b) key points extraction from sensed image $(b_1 - b_9)$ (c) Total matches obtained employing BF matcher $(c_1 - c_9)$ (d) Identification of good matches using ATSRC $(d_1 - d_9)$ (e) Register image $(e_1 - e_9)$

Table 6.8: Experiment outcomes obtained from various intermodal medical images sourced from two in house clinical dataset

S. No	Sensed image key points	Register image key points	Total no of matches	Matches after ATSRC
1	80	66	41	19
2	119	121	57	29
3	42	34	17	9
4	86	82	36	19
5	9	16	9	6
6	29	24	19	9
7	35	31	19	11
8	85	36	76	23
9	90	81	32	15

6.3.1.2. Comparative Analysis

The effectiveness of the proposed scheme in monomodal and multimodal medical image registration is evaluated against meta-heuristic optimization-based techniques using identical images and parameters from [3], [49], [74], [133], [136].

The proposed scheme is compared to multimodal brain image registration methods using VGG16 model and a modified SURF method, as presented by Wang *et al*. [74] and Zheng *et al.* [49]. The test images used for comparison are MR and CT images from the Kaggle dataset [114], [115], each with dimension of $256 * 256$. The proposed scheme is compared to [49], [74] with different similarity metrics namely MI and CC. The higher value of MI and CC for proposed scheme is tabulated in Table 6.9 and Figure 6.18.

Table 6.9: Quantitative evaluation of proposed scheme with [49], [74] in terms of MI and CC

Evaluation metric	MI	CC
P1-SURF[49]	1.5861	0.7334
CNN[74]	2.0342	0.8300
Proposed scheme	**3.964369**	**0.993543**

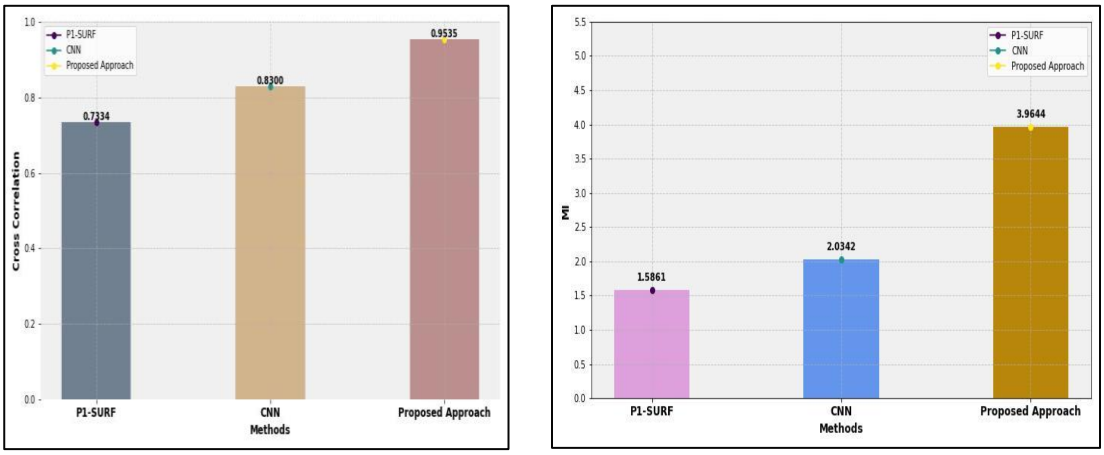


Figure 6.18: Comparative analysis of MSE and CC between proposed scheme and traditional methods [74], [137] applied to intermodal medical images.

The proposed scheme has been evaluated against multimodal image registration technique presented by Kori *et al.* [133],which leverages SIFT and VGG19 for feature extraction on MR-T1 and MR-T2 images from the IXI dataset [134]. The quantitative results of MI and SSIM measures indicate that proposed scheme provides superior performance than [133] as illustrated in Table 6.10 and Figure 6.19.

Table 6.10: Comparative analysis of proposed scheme and [133] based on MI and SSIM for intermodal medical images.

Evaluation metrics / Method	MI	SSIM
Simple Elastic [133]	0.221	0.404
SIFT Based [133]	0.15	0.383
CNN(VGG19)[133]	0.373	0.431
Proposed scheme	**1.5387**	**0.8643**

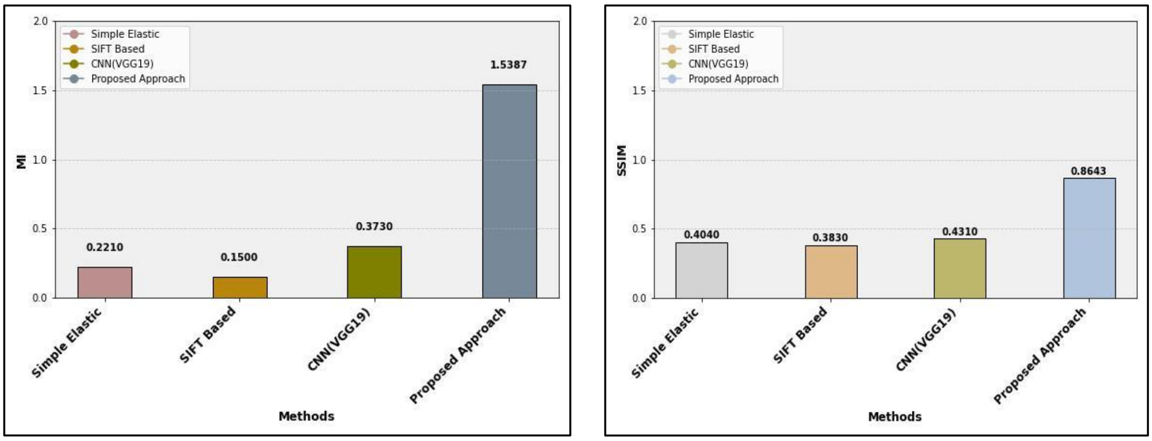


Figure 6.19: Comparative analysis of the proposed approach versus [133] employing various similarity metrics.

The method presented by Mohanty *et al*. [136], which employs unsupervised deep learning framework is contrasted with proposed scheme. To assess their performance, several multimodal CT and MRI brain images, each with dimensions of 256 × 256, are obtained from ADNI dataset [132]. The proposed scheme outperforms the multimodal medical image registration technique [136] in terms of MI and CC. The higher value of MI and CC for proposed scheme is tabulated in Table 6.11 and Figure 6.20.

Table 6.11:Comparing the proposed scheme with unsupervised deep learning methods [49]

Evaluation metric / Method	CC	MI
SyN	0.9355	0.6120
Voxelmorph	0.9822	0.8620
Diffeomorphism scheme[136]	0.9962	0.8830
Proposed scheme	**0.9982**	**1.3982**

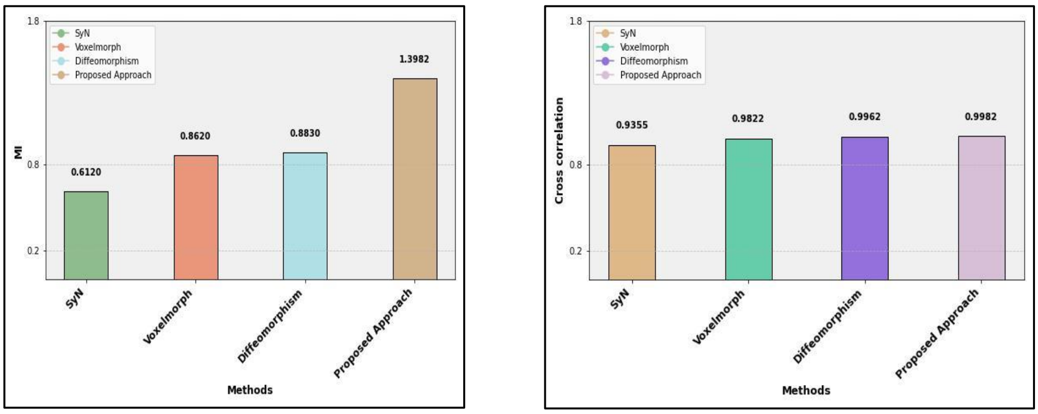


Figure 6.20: Comparative analysis of assessment metric between our scheme and conventional method [136] for intermodal medical images.

The proposed scheme has been compared to the single rigid registration (SRR) and multi-resolution rigid registration (MRR) techniques presented by Azam *et al.* [3]. The test images used for comparison are CT and MR of dimension 256×256 taken from Kaggle dataset [114], [117]. The proposed scheme demonstrates superior performance over multimodal image registration strategy [3] across different similarity metrics namely MI, CC, SSIM, and PSNR as tabulated in Table 6.12.

Table 6.12: Evaluation of proposed scheme against traditional method [3] for intramodal and intermodal medical images

S. No	Images	Methods	MI	CC	SSIM	PSNR
1	MR-PD and MR-T2	SRR [3]	1.1658	0.8819	0.7770	30.9114
		MRR [3]	1.1860	0.8917	0.7955	40.4342
		Proposed scheme	**2.5123**	**0.9472**	**0.9125**	**42.3781**
2	CT and MR-T1	SRR [3]	0.9185	0.3232	0.5287	14.8154
		MRR [3]	1.1877	0.5568	0.5799	25.680
		Proposed scheme	**2.1456**	**0.80345**	**0.79754**	**30.4562**
3	MR-T1 and MR-T2	SRR [3]	1.1634	0.7176	0.5861	24.2732
		MRR [3]	1.2287	0.7449	0.6022	36.7877
		Proposed scheme	**2.65321**	**0.86532**	**0.88965**	**40.6421**
4	CT and MR-T2	SRR [3]	1.0503	0.4755	0.5622	17.3709
		MRR [3]	1.1495	0.5252	0.5335	36.0671
		Proposed scheme	**1.9864**	**0.80324**	**0.78965**	**38.98653**

The comparison results illustrating the performance of proposed scheme using VGG19 and SURF with SRR [3] and MRR [3] on both monomodal and multimodal medical images based on similarity metrics are shown in Figure 6.21.

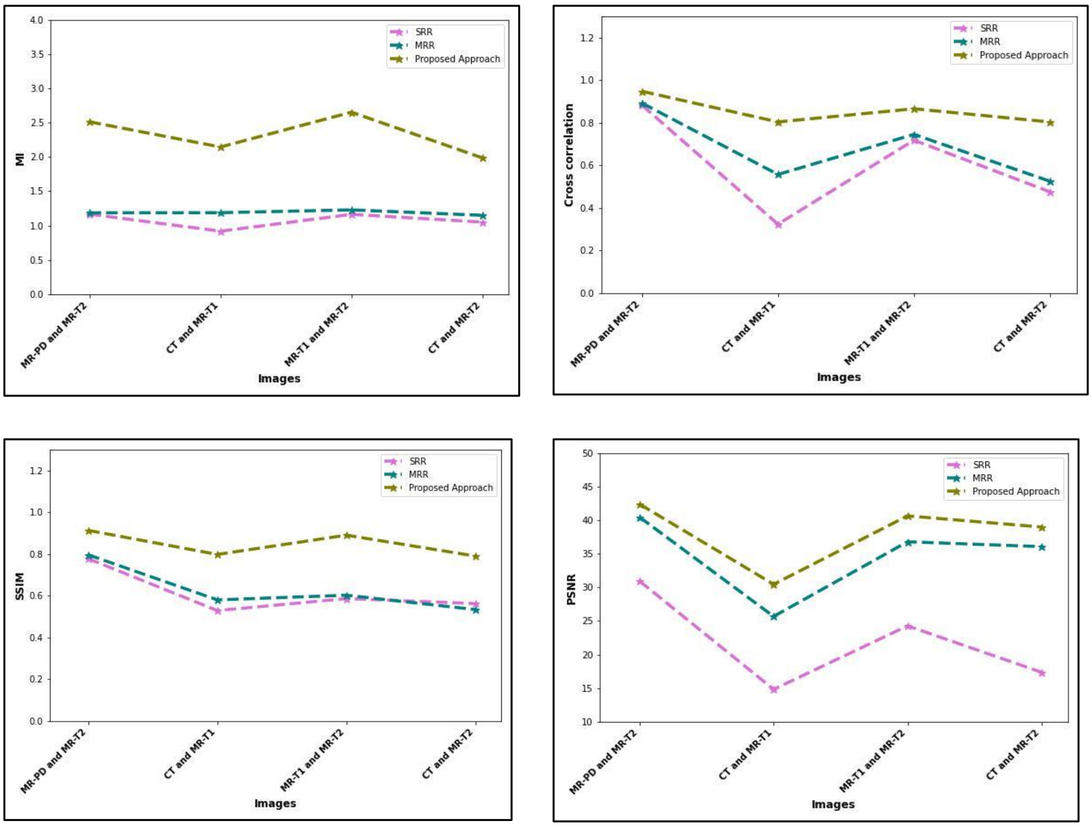


Figure 6.21: Comparative assessment of proposed scheme against conventional methods for anatomical and functional images employing various similarity metrics.

6.3.2. Conclusion

In this scheme, an integration of VGG19 with SURF for monomodal and multimodal medical image registration is discussed. This scheme initiates by improving input image quality through preprocessing steps, followed by extracting both low and high-level features using VGG19. Subsequently, euclidean distance is computed for each pair of feature points across different feature spaces. Dynamic inlier selection improves feature point alignment between pre-processed fixed and sensed images by iteratively adjusting distance thresholds to identify inliers. Then, thin plate spline interpolation corrects local misalignments among selected feature points and computes a non-linear transformation. This process enables complex warping to align features between images with non-uniform distortions, resulting in precise alignment. To enhance the reliability of registered image obtained using CNNs (VGG16 and VGG19), SURF algorithm is employed to detect and match features with a BF matcher. Additionally, outliers are minimized using the alpha-trimmed spatial resolution correspondence algorithm. Finally, a homography matrix is applied to generate the final registered image, achieving robust alignment between the sensed and VGG19 registered images.

6.4. Chapter Summary

In this chapter, a deep learning approach combined with feature extraction techniques for both monomodal and multimodal medical image registration is presented. The proposed scheme explores various feature extraction technique namely SIFT, SURF and ORB algorithm for detecting features. Then, feature matching is performed using a brute force matcher to align similar features in reference and target images. After matching features, outlier removal methods such as alpha trimmed spatial relation correspondence are employed to eliminate outliers. Additionally, projective transformation is applied to pre-processed reference and target images for generating registered image. This is achieved through robust feature extraction from reference and target images using SURF, along with effective outlier removal by leveraging ATSRC. Despite this research focus on feature extraction techniques, it has several disadvantages, such as time-consuming iterative methods and limited accuracy in complex scenarios. To address these limitations, we introduced a novel image registration algorithm that combines deep learning approaches with SURF feature extraction technique for accurate registration of images. The scheme begins by enhancing input image quality through preprocessing and extracting features using VGG19. Then, euclidean distances between feature points are calculated, and dynamic inlier selection adjusts distance thresholds to determine inliers and refine alignment. Thin plate spline interpolation then corrects local misalignments for precise alignment. To enhance the reliability of registered image obtained using CNNs (VGG16 and VGG19), SURF algorithm is employed to detect and match features with a BF matcher, while outliers are minimized using the alpha-trimmed spatial resolution correspondence algorithm. Finally, a homography matrix is applied to achieve robust alignment between the sensed and VGG19 registered images.

CHAPTER 7 Conclusion and Future Scope

This chapter concludes the thesis by summarizing the research contributions, outcomes, and scope of future work in the domain of medical image registration. The main objective of thesis is to align reference and target images of the same subject taken at different times, from different viewpoints, or by different sensor. The thesis mainly consists of (i) exploring area and feature-based methods for image registration across various medical imaging modalities, (ii) developing a model to address image registration challenges using transformation methods (Rigid/Non-Rigid, Affine, Projective, etc.) or signal/image processing techniques, (iii) enhancing alignment between source and target images through machine learning and optimization methods, (iv) testing and validating the developed schemes with standard and real-world datasets, and (v) assessing the effectiveness of proposed scheme against state-of-the-art techniques. Based on the research contribution as discussed from Chapter 3 to Chapter 6, several conclusions are drawn. The research contributions and outcomes are detailed in Sections 7.1 to 7.4, with future directions outlined in Section 7.5.

7.1. A TLBO-Driven Hybrid Approach for Adaptive Medical Image Registration Using Affine Transformation and SURF

In the first phase, TLBO is employed to determine the optimal rotation and translation parameters with MI maximization as an objective function. To assess the performance of proposed scheme, extensive experiments are conducted on diverse image modalities, including anatomical images such as CT and MRI images with noise and artifacts. The experimental results show an 8% improvement in MI and SSIM values as compared to state-of-the-art methods, indicating better quality of registered images. To address challenges such as optimal value of rigid transformation parameters due to anatomical variations and spatial transformations, we proposed a novel image registration algorithm that integrates both area and feature-based approaches for monomodal and multimodal medical images. This method employs SURF framework for feature extraction along with RANSAC algorithm to refine the features. Subsequently, projective transformation is used to achieve more accurate register image. Experiments are conducted on Whole Brain ATLAS and Kaggle datasets demonstrating the robustness of the proposed scheme, with an improvement of 8% in SSIM, 12.71% in MI for monomodal and multimodal images as compared to state-of-the-art methods.

7.2. Multimodal Image Registration leveraging Kernel PCA and TLBO

In the second phase of this research work, a novel approach with the hybridization of KPCA and TLBO for registering monomodal and multimodal medical images is presented. The approach starts with pre-processing steps including noise removal and normalization to improve image quality. In the next step, contour extraction is performed using Otsu's thresholding for multimodal images and a thresholding segmentation approach for monomodal images to create feature images and determine ground truth translation parameters. Then, KPCA is used to capture non-linear data mapping for determining rotation values, while TLBO optimizes the registration process by iteratively refining transformation parameters. The proposed scheme has been assessed using various similarity metrics on both monomodal and multimodal medical images from publicly available and in-house clinical datasets. This scheme focuses on rigid medical image registration for both standard and real time (clinical) medical image datasets.

7.3. Deformable and Rigid Medical Image Registration on Different Modalities Using the Combination of Deep learning and Meta Heuristic Approaches

In the third phase, deep learning and meta-heuristic approaches for both monomodal and multimodal medical image registration is presented. Deformable image registration scheme employs an unsupervised U-Net model followed by TLBO optimization and affine transformation for monomodal medical images. The Vxnet (U-Net) model processes two-channel images to generate a non-linear displacement vector field, while TLBO optimizes the registration process by iteratively adjusting transformation parameters. Although this deformable registration work focuses on monomodal medical images, which can be further enhanced to multimodal modalities of images. After that, a novel medical image registration using UEO meta-heuristic approach and VGG19 on both monomodal and multimodal modalities. This scheme employs UEO optimization technique to optimize rigid transformation parameters by maximizing MI. Then, features are extracted using VGG-19 from both reference and UEO-registered images. Afterwards, dynamic inlier selection is applied to enhance the matching accuracy by effectively filtering correct matches (inliers) while minimizing outliers. Finally, thin plate spline interpolation is employed to determine the optimal transformation matrix for accurate alignment between reference and UEO-registered images. This approach is validated using both in-house clinical and publicly available standard datasets, demonstrating robustness and reliability through evaluation metrics such as MI, RMSE, SSIM, and PSNR.

7.4. VGG19-SURF with Projective Transformation for Medical Image Registration

In the fourth phase, a deep learning approach combined with feature extraction techniques for both monomodal and multimodal medical image registration is presented. The proposed scheme involves various feature extraction techniques like SIFT, SURF and ORB for detecting features. Then, feature matching is performed using a brute force matcher to align similar features in reference and target images. After matching features, outlier removal methods such as alpha trimmed spatial relation correspondence is employed to eliminate outliers. Additionally, projective transformation is applied to pre-processed reference and target images for generating registered image. Even though this research focuses on feature extraction techniques, we introduced a novel image registration algorithm that combines deep learning approach with SURF feature extraction technique for accurate registration of images. The scheme begins by enhancing input image quality through preprocessing and extracting features using VGG19. Then, euclidean distances between feature points are calculated, and dynamic inlier selection adjusts distance thresholds to determine inliers and refine alignment. Thin plate spline interpolation then corrects local misalignments for precise alignment. To enhance the reliability of registered image obtained using CNNs (VGG16 and VGG19), SURF algorithm is employed to detect and match features with a BF matcher, while outliers are minimized using the alpha-trimmed spatial resolution correspondence algorithm. Finally, a homography matrix is applied to achieve robust alignment between the sensed and VGG19 registered images.

7.5. Future Scope

The objective of proposed scheme in this thesis is to develop robust accurate alignment of medical images through feature extraction and machine learning algorithms. There are some research issues that can be resolved as an extension of current research work for medical image registration. The possible enhancements are discussed below.

1. In this thesis, the proposed scheme primarily focuses on rigid and non-rigid medical image registration techniques using CT-MRI brain and lung anatomical images. This approach can be further extended to different modalities, such as PET, SPECT, and applied to other anatomical structures, including heart and kidneys etc.
2. In this research, non-rigid algorithms such as unsupervised learning is employed for mono modal medical images that can be extended to handle multi modal modality and further analysis can be conducted using supervised and weakly supervised learning.
3. In future work, the proposed scheme can be enhanced by incorporating hybrid optimization techniques and advanced feature extraction methods, such as vision transformers and Generative Adversarial Networks (GAN), to improve the accuracy and efficiency of image alignment.
4. In this thesis, we focused on 2-D rigid registration techniques for multi-sensor, multi-temporal, and multi-viewpoint images. In future, the aim is to work on multi-resolution rigid registration techniques and extend these methods to 3-D, enabling more detailed diagnostic evaluations.

References

[1] P. Arora, R. Mehta, and R. Ahuja, “An adaptive medical image registration using hybridization of teaching learning-based optimization with affine and speeded up robust features with projective transformation,” *Cluster Computing*, vol. 3, 2023, doi: 10.1007/s10586-023-03974-3.

[2] J. Jin, X. Bi, M. Jiang, J. Chang, and J. Cui, “Medical image registration based on PCA and M _ PSNR,” *Journal of Complexity in Health Sciences*, vol. 3, no. 1, pp. 62–72, 2020, doi: 10.21595/chs.2020.21266.

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

[4] M. Su, C. Zhang, Z. Chen, and S. Jiang, “Registration of multimodal brain images based on optical flow,” *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017*, vol. 2017, no. 10, pp. 1–5, 2017, doi: 10.1109/CISP-BMEI.2017.8302198.

[5] H. R. Boveiri, R. Khayami, R. Javidan, and A. R. MehdiZadeh, “Medical image registration using deep neural networks: A comprehensive review,” *Computers & Electrical Engineering*, vol. 87, no. 2, pp. 1–45, 2020, doi: 10.1016/j.compeleceng.2020.106767.

[6] X. Liu, Z. Tian, C. Chai, and H. Fu, “Multiscale registration of remote sensing image using robust SIFT features in Steerable-Domain,” *Egyptian Journal of Remote Sensing and Space Science*, vol. 14, no. 2, pp. 63–72, 2011, doi: 10.1016/j.ejrs.2011.08.001.

[7] J. Hu *et al.*, “Towards accurate and robust multi-modal medical image registration using contrastive metric learning,” *IEEE Access*, vol. 7, no. 2, pp. 132816–132827, 2019, doi: 10.1109/ACCESS.2019.2938858.

[8] Y. Zhuang, K. Gao, X. Miu, L. Han, and X. Gong, “Infrared and visual image registration based on mutual information with a combined particle swarm optimization - Powell search algorithm,” *Optik*, vol. 127, no. 1, pp. 188–191,

2016, doi: 10.1016/j.ijleo.2015.09.199.

[9] G. Song, J. Han, Y. Zhao, Z. Wang, and H. Du, "A Review on Medical Image Registration as an Optimization Problem," *Current Medical Imaging Reviews*, vol. 13, no. 3, pp. 274–283, 2017, doi: 10.2174/1573405612666160920123955.

[10] X. Jiang, J. Ma, G. Xiao, Z. Shao, and X. Guo, "A review of multimodal image matching: methods and applications," *Information Fusion*, vol. 73, no. 2, pp. 22–71, 2021, doi: 10.1016/j.inffus.2021.02.012.

[11] H. O. Velesaca, G. Bastidas, M. Rouhani, and A. D. Sappa, "Multimodal image registration techniques: a comprehensive survey," *Multimedia Tools and Applications*, no. 1, pp. 63919–63947, 2024, doi: 10.1007/s11042-023-17991-2.

[12] R. Panda, S. Agrawal, M. Sahoo, and R. Nayak, "A novel evolutionary rigid body docking algorithm for medical image registration," *Swarm and Evolutionary Computation*, vol. 33, no. 8, pp. 108–118, 2017, doi: 10.1016/j.swevo.2016.11.002.

[13] M. S. Hosseini, M. H. Moradi, M. Tabassian, and J. D'hooge, "Non-rigid image registration using a modified fuzzy feature-based inference system for 3D cardiac motion estimation," *Computer Methods and Programs in Biomedicine*, vol. 205, p. 106085, 2021, doi: 10.1016/j.cmpb.2021.106085.

[14] O. Cordón, S. Damas, and J. Santamaría, "A fast and accurate approach for 3D image registration using the scatter search evolutionary algorithm," *Pattern Recognition Letters*, vol. 27, no. 11, pp. 1191–1200, 2006, doi: 10.1016/j.patrec.2005.07.017.

[15] P. Kotsas and T. Dodd, "Rigid registration of medical images using 1D and 2D binary projections," *Journal of Digital Imaging*, vol. 24, no. 5, pp. 913–925, 2011, doi: 10.1007/s10278-010-9352-z.

[16] E. Saiti and T. Theoharis, "An application independent review of multimodal 3D registration methods," *Computers and Graphics (Pergamon)*, vol. 91, pp. 153–178, 2020, doi: 10.1016/j.cag.2020.07.012.

[17] S. Shehanaz, E. Daniel, S. R. Guntur, and S. Satrasupalli, "Optimum weighted multimodal medical image fusion using particle swarm optimization," *Optik*,

vol. 231, no. 1, p. 166413, 2021, doi: 10.1016/j.ijleo.2021.166413.

[18] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, “A survey of medical image registration – under review,” *Medical Image Analysis*, vol. 33, no. 1, pp. 140–144, 2016, doi: 10.1016/j.media.2016.06.030.

[19] H. Mahmoud, F. Masulli, and S. Rovetta, “Feature-based medical image registration using a fuzzy clustering segmentation approach,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 7845 LNBI, pp. 37–47, 2013, doi: 10.1007/978-3-642-38342-7_4.

[20] Z. Hossein-Nejad and M. Nasri, “Clustered redundant keypoint elimination SURF method in MRI image registration based on alpha-trimmed relationship,” *Frontiers in Biomedical Technologies*, vol. 10, no. 2, pp. 120–131, 2023, doi: 10.18502/fbt.v10i2.12216.

[21] S. A. K. Tareen and Z. Saleem, “A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK,” *International Conference on Computing, Mathematics and Engineering Technologies: Invent, Innovate and Integrate for Socioeconomic Development, iCoMET Proceedings*, pp. 1–10, 2018, doi: 10.1109/ICOMET.2018.8346440.

[22] A. Oad, K. Kumari, I. Hussain, F. Dong, B. Hammad, and R. Oad, “Performance comparison of ORB, SURF and SIFT using Intracranial Haemorrhage CTScan Brain images,” *International Journal of Artificial Intelligence & Mathematical Sciences*, vol. 1, no. 2, pp. 26–34, 2023, doi: 10.58921/ijaims.v1i2.41.

[23] Syed Abdul Hanan Ahmed Shah, Shahnawaz Talpur, and Bushra Naz, “A Discriminative Image Registration Model with Hybrid Feature Detector,” *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 3307, pp. 272–278, 2023, doi: 10.32628/cseit2390369.

[24] W. Zhang and Y. Zhao, “An improved SIFT algorithm for registration between SAR and optical images,” *Scientific Reports*, vol. 13, no. 1, pp. 1–8, 2023, doi: 10.1038/s41598-023-33532-1.

[25] J. M. Calvin, C. J. Gotsman, and C. Zheng, "Global optimization for image registration," *AIP Conference Proceedings*, vol. 2070, no. 1, 2019, doi: 10.1063/1.5089975.

[26] T. T. Ho, W. J. Kim, C. H. Lee, G. Y. Jin, K. J. Chae, and S. Choi, "An unsupervised image registration method employing chest computed tomography images and deep neural networks," *Computers in Biology and Medicine*, vol. 154, no. 1, pp. 1–11, 2023, doi: 10.1016/j.compbiomed.2023.106612.

[27] M. Zhai *et al.*, "Generating CT images in delayed PET scans using a multi-resolution registration convolutional neural network," *Biomedical Signal Processing and Control*, vol. 78, no. 3, p. 103853, 2022, doi: 10.1016/j.bspc.2022.103853.

[28] A. Valsecchi, S. Damas, J. Santamaría, and L. Marrakchi-Kacem, "Intensity-based image registration using scatter search," *Artificial Intelligence in Medicine*, vol. 60, no. 3, pp. 151–163, 2014, doi: 10.1016/j.artmed.2014.01.006.

[29] M. Engin, H. Ogul, M. Agildere, and E. Sumer, "An evaluation of image registration methods for chest radiographs," *IntelliSys 2015 - Proceedings of 2015 SAI Intelligent Systems Conference*, pp. 822–827, 2015, doi: 10.1109/IntelliSys.2015.7361237.

[30] H. Ramadan, D. El Bourakadi, A. Yahyaouy, and H. Tairi, "Medical image registration in the era of Transformers: A recent review," *Informatics in Medicine Unlocked*, vol. 49, no. 6, p. 101540, 2024, doi: 10.1016/j.imu.2024.101540.

[31] K. T. Islam, S. Wijewickrema, and S. O'Leary, "A deep learning based framework for the registration of three dimensional multi-modal medical images of the head," *Scientific Reports*, vol. 11, no. 1, pp. 1–13, 2021, doi: 10.1038/s41598-021-81044-7.

[32] L. He, "Non-rigid multi-modal medical image registration based on improved maximum mutual information PV image interpolation method," *Frontiers in Public Health*, vol. 10, no. 6, pp. 1–10, 2022, doi: 10.3389/fpubh.2022.863307.

[33] X. Yan, Y. Zhang, D. Zhang, and N. Hou, "Multimodal image registration using

histogram of oriented gradient distance and data-driven grey wolf optimizer," *Neurocomputing*, vol. 392, pp. 108–120, 2020, doi: 10.1016/j.neucom.2020.01.107.

[34] F. Charif, A. Benchabane, and Z. Bebboukha, "Multimodal medical images registration using biogeography-based optimization algorithm," *Proceedings - 2019 6th International Conference on Image and Signal Processing and their Applications, ISPA 2019*, no. 11, pp. 6–10, 2019, doi: 10.1109/ISPA48434.2019.8966862.

[35] M. Zou *et al.*, "Rigid medical image registration using learning-based interest points and features," *Computers, Materials and Continua*, vol. 60, no. 2, pp. 511–525, 2019, doi: 10.32604/cmc.2019.05912.

[36] I. Nazir, I. U. Haq, S. A. Alqahtani, M. M. Jadoon, and M. Dahshan, "Machine learning-based lung cancer detection using multiview image registration and fusion," *Journal of Sensors*, vol. 1, 2023, doi: 10.1155/2023/6683438.

[37] H. Dida, F. Charif, and A. Benchabane, "A comparative study of two meta-heuristic algorithms for MRI and CT images registration," *3rd International Conference on Information and Communications Technology, ICOIACT*, pp. 411–415, 2020, doi: 10.1109/ICOIACT50329.2020.9332126.

[38] T. X. Lin and H. H. Chang, "Medical image registration based on an improved ant colony optimization algorithm," *International Journal of Pharma Medicine and Biological Sciences*, vol. 5, no. 1, pp. 17–22, 2016, doi: 10.18178/ijpmbs.5.1.17-22.

[39] M. Abdel-Basset, A. E. Fakhry, I. El-henawy, T. Qiu, and A. K. Sangaiah, "Feature and intensity based medical image registration using particle swarm optimization," *Journal of Medical Systems*, vol. 41, no. 12, pp. 1–15, 2017, doi: 10.1007/s10916-017-0846-9.

[40] F. Ayatollahi, S. B. Shokouhi, and A. Ayatollahi, "A new hybrid particle swarm optimization for multimodal brain image registration," *Journal of Biomedical Science and Engineering*, vol. 05, no. 04, pp. 153–161, 2012, doi: 10.4236/jbise.2012.54020.

[41] R. Mishra and M. K. Bajpai, "Hybrid priority-based genetic algorithm for limited-view computed tomography," *Multimedia Tools and Applications*, pp.

1–21, 2024, doi: 10.1007/s11042-024-19506-z.

[42] Y. Lu *et al.*, “A New Optimization Approach Using Smoothed Images Based on ACO for Medical Image Registration,” *Applied Mathematical Modelling*, vol. 77, no. 1, pp. 30–36, 2018, doi: 10.1016/j.apm.2019.07.046.

[43] Y. Wu, W. Ma, Q. Miao, and S. Wang, “Multimodal continuous ant colony optimization for multisensor remote sensing image registration with local search,” *Swarm and Evolutionary Computation*, vol. 47, no. 7, pp. 89–95, 2019, doi: 10.1016/j.swevo.2017.07.004.

[44] R. Kaliski, “Green NOMA M2M,” *Proceedings - IEEE Global Communications Conference, GLOBECOM*, pp. 1–7, 2021, doi: 10.1109/GLOBECOM46510.2021.9685420.

[45] C. L. Cocianu and A. Stan, “New evolutionary-based techniques for image registration,” *Applied Sciences (Switzerland)*, vol. 9, no. 1, p. 176, 2019, doi: 10.3390/app9010176.

[46] R. Kaliski and H. Y. Wei, “Dynamic Resource Allocation and Advertisement Revenue Optimization for TV over eMBMS,” *IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 579–597, 2016, doi: 10.1109/TBC.2016.2570013.

[47] K. K. Singh, M. K. Bajpai, and R. K. Pandey, “A novel approach for edge detection of low contrast satellite images,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, vol. 40, no. 2, pp. 211–217, 2015, doi: 10.5194/isprsarchives-XL-3-W2-211-2015.

[48] P. Ting-Ting and Z. Ji, “Research on medical image registration based on QPSO and powell algorithm,” *Proceedings - 14th International Symposium on Distributed Computing and Applications for Business, Engineering and Science, DCABES*, pp. 316–319, 2016, doi: 10.1109/DCABES.2015.86.

[49] Q. Zheng, Q. Wang, X. Ba, S. Liu, J. Nan, and S. Zhang, “A Medical Image Registration Method Based on Progressive Images,” *Computational and Mathematical Methods in Medicine, Hindawi*, vol. 2021, no. 7, pp. 1–10, 2021, doi: 10.1155/2021/4504306.

[50] H. Vishwakarma and S. K. Katiyar, "Accuracy assessment of projective transformation based hybrid approach for automatic satellite image registration," *International Journal of Civil Engineering and Technology*, vol. 9, no. 13, pp. 1514–1523, 2018.

[51] Z. Gu, L. Cai, Y. Yin, Y. Ding, and H. Kan, "Registration of Brain Medical Images Based on SURF Algorithm and R-RANSAC Algorithm," *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 12, no. 3, pp. 2290–2297, 2014, doi: 10.11591/telkomnika.v12i3.4500.

[52] R. Swathi and A. Srinivas, "An Improved Image Registration Method Using E-SIFT Feature Descriptor with Hybrid Optimization Algorithm," *Journal of the Indian Society of Remote Sensing*, vol. 48, no. 2, pp. 215–226, 2020, doi: 10.1007/s12524-019-01063-w.

[53] R. Mishra and M. K. Bajpai, "A priority based genetic algorithm for limited view tomography," *Applied Intelligence*, vol. 51, no. 10, pp. 6968–6982, 2021, doi: 10.1007/s10489-021-02192-x.

[54] Z. Hossein-Nejad and M. Nasri, "An adaptive image registration method based on SIFT features and RANSAC transform," *Computers and Electrical Engineering*, vol. 62, no. 8, pp. 524–537, 2017, doi: 10.1016/j.compeleceng.2016.11.034.

[55] I. Hyun and M. Tariq, "Adaptive outlier elimination in image registration using genetic programming R," *Information Sciences*, vol. 421, pp. 204–217, 2017, doi: 10.1016/j.ins.2017.08.098.

[56] S. Hamzehei, J. Bai, G. Raimondi, R. Tripp, L. Ostroff, and S. Nabavi, "3D biological/biomedical image registration with enhanced feature extraction and outlier detection," *ACM-BCB 2023 - 14th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pp. 1–10, 2023, doi: 10.1145/3584371.3612965.

[57] P. Gadde and X. H. Yu, "Image registration with artificial neural networks using spatial and frequency features," *Proceedings of the International Joint Conference on Neural Networks*, vol. 12, no. 2, pp. 4643–4649, 2016, doi: 10.1109/IJCNN.2016.7727809.

[58] R. Kaliski and Y. H. Han, "Socially-Aware V2X QoS for NOMA Dual-

Connectivity," *IEEE Vehicular Technology Conference*, pp. 1–5, 2021, doi: 10.1109/VTC2021-Fall52928.2021.9625542.

[59] R. Sharma, A. M. Joshi, C. Sahu, and S. J. Nanda, "Detection of false data injection in smart grid using PCA based unsupervised learning," *Electrical Engineering*, vol. 105, no. 4, pp. 2383–2396, 2023, doi: 10.1007/s00202-023-01809-3.

[60] D. Peng, J. Liu, J. Tian, and S. Zheng, "Transformation model estimation of image registration via least square support vector machines," vol. 27, pp. 1397–1404, 2006, doi: 10.1016/j.patrec.2006.01.012.

[61] D. Li, D. V. Vargas, and S. Kouichi, "Universal Rules for Fooling Deep Neural Networks based Text Classification," *2019 IEEE Congress on Evolutionary Computation, CEC 2019 - Proceedings*, pp. 2221–2228, 2019, doi: 10.1109/CEC.2019.8790213.

[62] C. Liu, K. He, D. Xu, H. Shi, H. Zhang, and K. Zhao, "RegFSC-Net : Medical Image Registration via Fourier Transform with Spatial Reorganization and Channel Refinement Network," *IEEE Journal of Biomedical and Health Informatics*, pp. 1–12, 2024, doi: 10.1109/JBHI.2024.3376334.

[63] S. Goswami, A. Sharma, and P. Mahajan, "Efficient Image Registration Using Support Vector Machines," *2024 International Conference on Optimization Computing and Wireless Communication (ICOCWC)*, pp. 1–6, 2024, doi: 10.1109/ICOCWC60930.2024.10470766.

[64] M. Wang, "Fast Registration of High-Resolution Optical-SAR Images Based on Grayscale Normalized Mutual Information," *2023 8th International Conference on Control, Robotics and Cybernetics (CRC)*, pp. 320–325, 2023, doi: 10.1109/CRC60659.2023.10488537.

[65] Z. Yang, T. Dan, and Y. Yang, "Multi-temporal remote sensing image registration using deep convolutional features," *IEEE Access*, vol. 6, pp. 38544–38555, 2018, doi: 10.1109/ACCESS.2018.2853100.

[66] F. Zhao, Y. Hori, and K. Sakurai, "Two-servers PIR based DNS query scheme with privacy-preserving," *Proceedings The 2007 International Conference on Intelligent Pervasive Computing, IPC 2007*, pp. 299–302, 2007, doi: 10.1109/IPC.2007.27.

[67] X. Zhao, H. Li, P. Wang, and L. Jing, “An image registration method using deep residual network features for multisource high-resolution remote sensing images,” *Remote Sensing*, vol. 13, no. 17, p. 3425, 2021, doi: 10.3390/rs13173425.

[68] R. Mehta, N. Rajpal, and V. P. Vishwakarma, “A robust and efficient image watermarking scheme based on Lagrangian SVR and lifting wavelet transform,” *International Journal of Machine Learning and Cybernetics*, vol. 8, no. 2, pp. 379–395, 2017, doi: 10.1007/s13042-015-0331-z.

[69] K. Kavitha, B. Thirumala Rao, and B. Sandhya, “Evaluation of distance measures for feature based image registration using AlexNet,” *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, pp. 284–290, 2018, doi: 10.14569/IJACSA.2018.091034.

[70] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, “Remote sensing image registration using convolutional neural network features,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 2, pp. 232–236, 2018, doi: 10.1109/LGRS.2017.2781741.

[71] J. Su, D. V. Vargas, and K. Sakurai, “One Pixel Attack for Fooling Deep Neural Networks,” *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 5, pp. 828–841, 2019, doi: 10.1109/TEVC.2019.2890858.

[72] M. Elwan, A. S. Amein, A. Mousa, A. M. Ahmed, B. Bouallegue, and A. S. Eltanany, “SAR image matching based on local feature detection and description using convolutional neural network,” *Security and Communication Networks*, vol. 2022, no. 1, p. 5669069, 2022, doi: 10.1155/2022/5669069.

[73] K. L. Kashyap, M. K. Bajpai, and P. Khanna, “Globally supported radial basis function based collocation method for evolution of level set in mass segmentation using mammograms,” *Computers in Biology and Medicine*, vol. 87, pp. 22–37, 2017, doi: 10.1016/j.compbiomed.2017.05.015.

[74] Q. Wang, S. Liu, X. J. Ba, and B. Yuan, “Brain MR image registration based on an improved CNN model,” *ACM International Conference Proceeding Series*, pp. 66–72, 2021, doi: 10.1145/3510513.3510524.

[75] Y. Feng, Y. Hori, K. Sakurai, and J. Takeuchi, “A behavior-based detection method for outbreaks of low-rate attacks,” *Proceedings - 2012 IEEE/IPSJ 12th*

International Symposium on Applications and the Internet, SAINT 2012, pp. 267–272, 2012, doi: 10.1109/SAINT.2012.50.

[76] V. Kearney, S. Haaf, A. Sudhyadhom, G. Valdes, and T. D. Solberg, "An unsupervised convolutional neural network-based algorithm for deformable image registration," *Physics in Medicine and Biology*, vol. 63, no. 18, p. 185017, 2018, doi: 10.1088/1361-6560/aada66.

[77] Z. Lin, H. Liu, X. Chu, Y. W. Leung, and I. Stojmenovic, "Constructing Connected-Dominating-Set with Maximum Lifetime in Cognitive Radio Networks," *IEEE Transactions on Computers*, vol. 73, no. 4, pp. 1165–1179, 2024, doi: 10.1109/TC.2013.77.

[78] B. Sajja and A. Singh, "Detection of Violent Content in Videos Using 2D Attention-Augmented Convolutional Networks and Gated Recurrent Unit," *2023 IEEE 20th India Council International Conference, INDICON 2023*, pp. 967–972, 2023, doi: 10.1109/INDICON59947.2023.10440672.

[79] X. Wang, G. Ning, N. Yang, X. Zhang, H. Zhang, and H. Liao, "An unsupervised convolution neural network for deformable registration of mono/multi-modality medical images," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 3455–3458, 2021, doi: 10.1109/EMBC46164.2021.9630731.

[80] G. Khosla, N. Rajpal, and J. Singh, "Evaluation of euclidean and manhanttan metrics in content based image retrieval system," *2015 International Conference on Computing for Sustainable Global Development, INDIACom 2015*, pp. 12–18, 2015.

[81] G. Balakrishnan, A. Zhao, M. R. Sabuncu, A. V. Dalca, and J. Guttag, "An unsupervised learning model for deformable medical image registration," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 9252–9260, 2018, doi: 10.1109/CVPR.2018.00964.

[82] A. Singh, P. Singh, V. Sharma, D. Tyagi, N. Pandey, and B. Vaid, "Palanam: A Deep Learning Based Childhood Defense System," *2024 16th International Conference on COMmunication Systems and NETworkS, COMSNETS 2024*,

pp. 273–275, 2024, doi: 10.1109/COMSNETS59351.2024.10427109.

[83] R. Bhattacharjee, F. Heitz, V. Noblet, S. Sharma, and N. Sharma, “Evaluation of a Learning-based Deformable Registration Method on Abdominal CT Images,” *IRBM*, vol. 42, no. 2, pp. 94–105, 2021, doi: 10.1016/j.irbm.2020.04.002.

[84] O. Roeva, S. Fidanova, and M. Paprzycki, “Influence of the population size on the genetic algorithm performance in case of cultivation process modelling,” *2013 Federated Conference on Computer Science and Information Systems, FedCSIS 2013*, pp. 371–376, 2013.

[85] Y. Ma, D. Niu, J. Zhang, X. Zhao, B. Yang, and C. Zhang, “Unsupervised deformable image registration network for 3D medical images,” *Applied Intelligence*, vol. 52, no. 1, pp. 766–779, 2022, doi: 10.1007/s10489-021-02196-7.

[86] Y. Zhu, Z. Zhou, G. Liao, and K. Yuan, “Deformable Registration Using Average Geometric Transformations for Brain MR Images,” pp. 1–9, 2019, [Online]. Available: http://arxiv.org/abs/1907.09670.

[87] K. Tang, Z. Li, L. Tian, L. Wang, and Y. Zhu, “ADMIR-Affine and deformable medical image registration for drug-addicted brain images,” *IEEE Access*, vol. 8, no. 4, pp. 70960–70968, 2020, doi: 10.1109/ACCESS.2020.2986829.

[88] F. Lo, S. M. Cheng, and R. Kaliski, “Optimization of Lightweight Malware Detection Models for AIoT Devices,” *2023 IEEE World Forum on Internet of Things: The Blue Planet: A Marriage of Sea and Space, WF-IoT 2023*, pp. 1–6, 2023, doi: 10.1109/WF-IoT58464.2023.10539588.

[89] K. K. Singh, M. K. Bajpai, and R. K. Pandey, “Edge Detection in Low Contrast Images,” no. 1, pp. 355–359, 2015, [Online]. Available: https://doi.org/10.5281/zenodo.1099850.

[90] D. Mahapatra, “Elastic Registration Of Medical Images With GANS,” vol. 7, 2018, [Online]. Available: http://arxiv.org/abs/1805.02369.

[91] J. Chen, E. C. Frey, Y. He, W. P. Segars, Y. Li, and Y. Du, “TransMorph: Transformer for unsupervised medical image registration,” *Medical Image Analysis*, vol. 82, no. 6, p. 102615, 2022, doi: 10.1016/j.media.2022.102615.

[92] N. Chugh, M. Kumar, M. Bhattacharya, and R. S. Gupta, “RF Performance comparison of Dual Material Gate (DMG) and Conventional AlGaN/GaN High Electron Mobility Transistor,” *Proceedings of the 4th International Conference on Devices, Circuits and Systems, ICDCS 2018*, pp. 137–142, 2019, doi: 10.1109/ICDCSyst.2018.8605122.

[93] G. Saygili, “Predicting medical image registration error through independent directions,” *Signal, Image and Video Processing*, vol. 15, no. 1, pp. 223–230, 2021, doi: 10.1007/s11760-020-01784-3.

[94] I. Kadyan and M. Kumar, “Design and analysis of a low phase noise, wide tunable CMOS based low power VCO with active inductor,” *Analog Integrated Circuits and Signal Processing*, vol. 119, no. 2, pp. 319–329, 2024, doi: 10.1007/s10470-024-02266-z.

[95] I. Lineta and P. J. J. Carol, “An approach for medical image registration using a double exponential - dynamic group based cooperative optimization algorithm,” *Multimedia Tools and Applications*, no. 2, pp. 1–25, 2024, doi: 10.1007/s11042-024-18429-z.

[96] R. Kaliski, C. C. Chou, H. Y. Meng, and H. Y. Wei, “Dynamic Resource Allocation Framework for MooD (MBMS Operation On-Demand),” *IEEE Transactions on Broadcasting*, vol. 62, no. 4, pp. 903–917, 2016, doi: 10.1109/TBC.2016.2590829.

[97] N. Chugh, M. Kumar, M. Bhattacharya, and R. S. Gupta, “Extraction of admittance parameters of symmetrically doped AlGaN/GaN/AlGaN DH-HEMT for microwave frequency applications,” *Microsystem Technologies*, vol. 27, no. 11, pp. 4065–4072, 2021, doi: 10.1007/s00542-020-04805-w.

[98] X. Luo *et al.*, “Neurocomputing Synergistic registration of CT-MRI brain images and retinal images : A novel approach leveraging reinforcement learning and modified artificial rabbit optimization,” vol. 585, no. 2, p. 127506, 2024, doi: 10.1016/j.neucom.2024.127506.

[99] V. Sharma, A. Singh, and S. Gaito, “Video Captioning using Spatio-temporal Graphs: An Encoder-Decoder Approach,” *2024 16th International Conference on COMmunication Systems and NETworkS, COMSNETS 2024*, pp. 424–426, 2024, doi: 10.1109/COMSNETS59351.2024.10427136.

[100] Y. Zheng *et al.*, “SymReg-GAN: Symmetric Image Registration with Generative Adversarial Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 9, pp. 5631–5646, 2022, doi: 10.1109/TPAMI.2021.3083543.

[101] D. Mahapatra and Z. Ge, “Training data independent image registration using generative adversarial networks and domain adaptation,” *Pattern Recognition*, vol. 100, p. 107109, 2020, doi: 10.1016/j.patcog.2019.107109.

[102] K. Kumar, M. Kumar, and M. Jhamb, “Performance Comparison of CMOS Based Voltage Reference Circuits in 180nm Technology,” *9th International Conference on Signal Processing and Communication, ICSC 2023*, pp. 156–161, 2023, doi: 10.1109/ICSC60394.2023.10441359.

[103] M. Singh, M. Kumar, and V. Nath, “Comparative Analysis of Reversible Logic Gates and Reversible Decimal to Binary Encoders using Verilog,” *2023 International Conference on Integration of Computational Intelligent System, ICICIS 2023*, pp. 1–6, 2023, doi: 10.1109/ICICIS56802.2023.10430307.

[104] R. Singh, N. Rajpal, and R. Mehta, “An empiric analysis of wavelet-based feature extraction on deep learning and machine learning algorithms for arrhythmia classification,” *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 6, pp. 25–34, 2021, doi: 10.9781/ijimai.2020.11.005.

[105] S. Gupta, N. Grover, and Zaheeruddin, “A New Optimization Approach Using Smoothed Images Based on ACO for Medical Image Registration,” *International Journal of Information Engineering and Electronic Business*, vol. 8, no. 2, pp. 30–36, 2016, doi: 10.5815/ijieeb.2016.02.04.

[106] A. S. Ahmed and H. A. Salah, “The IoT and registration of MRI brain diagnosis based on genetic algorithm and convolutional neural network,” vol. 25, no. 1, p. 273, 2022, doi: 10.11591/ijeecs.v25.i1.pp273-280.

[107] B. Handa, G. Singh, R. Kamal, A. S. Oinam, and V. Kumar, “Evaluation method for the optimization of 3D rigid image registration on multimodal image datasets,” *International Journal of Engineering and Advanced Technology*, vol. 9, no. 1, pp. 5539–5545, 2019, doi: 10.35940/ijeat.A2078.109119.

[108] M. Bhattaeharya and A. Das, “Multi resolution medical image registration using maximization of mutual information & optimization by Genetic algorithm,” *IEEE Nuclear Science Symposium Conference Record*, vol. 4, pp. 2961–2964, 2007, doi: 10.1109/NSSMIC.2007.4436755.

[109] W. Kosiński, P. Michalak, and P. Gut, “Robust Image Registration Based on Mutual Information Measure,” *Journal of Signal and Information Processing*, vol. 3, no. 2, pp. 175–178, 2012, doi: 10.4236/jsip.2012.32023.

[110] A. K. Shukla, P. Singh, and M. Vardhan, “An adaptive inertia weight teaching-learning-based optimization algorithm and its applications,” *Applied Mathematical Modelling*, vol. 77, no. 1, pp. 309–326, 2020, doi: 10.1016/j.apm.2019.07.046.

[111] N. Kumar and M. Nachamai, “Noise Removal and Filtering Techniques used in Medical Images,” *Oriental journal of computer science and technology*, vol. 10, no. 1, pp. 103–113, 2017, doi: 10.13005/ojcst/10.01.14.

[112] Keith A. Johnson and J. Alex Becker, “The Whole Brain ATLAS,” *Harvard University*, 2008. https://www.med.harvard.edu/aanlib/home.html. Accessed January 2024.

[113] P. N. Maddaiah, P. N. Pournami, and V. K. Govindan, “Optimization of Image Registration for Medical Image Analysis,” *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 3394–3398, 2014.

[114] VBOOKSHELF, “Brain CT Images with Intracranial Hemorrhage Masks,” 2019. https://www.kaggle.com/vbookshelf/computed-tomography-ct-images. Accessed February 2024.

[115] N. Chakrabarty, “Brain MRI Images for Brain Tumor Detection,” *Kaggle*, 2019. https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection. Accessed December 2023.

[116] M. Samantaray, M. Panigrahi, and K. C. Patra, “Enhanced Feature Matching Algorithm for Medical Image Registration,” in *Smart and Sustainable Technologies: Rural and Tribal Development using IoT and Cloud Computing*, 2022, pp. 357–363, doi: 10.1007/978-981-19-2277-0_33.

[117] Fernando Feltrin, “Brain Tumor MRI images 44 classes,” *Kaggle*, 2022.

https://www.kaggle.com/datasets/fernando2rad/brain-tumor-mri-images-44c. Accessed January 2024.

[118] Sartaj, "Brain Tumor Classification (MRI)," *Kaggle*, 2019. https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri. Accessed July 2023.

[119] A. Simeon, "Brain Tumor Images Dataset," *Kaggle*, 2019. https://www.kaggle.com/datasets/simeondee/brain-tumor-images-dataset. Accessed August 2023.

[120] L. Shang, J. Cheng Lv, and Z. Yi, "Rigid medical image registration using PCA neural network," *Neurocomputing*, vol. 69, no. 13–15, pp. 1717–1722, 2006, doi: 10.1016/j.neucom.2006.01.007.

[121] J. Hazra, A. R. Chowdhury, K. Dasgupta, and P. Dutta, "A hybrid structural feature extraction-based intelligent predictive approach for image registration," *Innovations in Systems and Software Engineering*, 2022, doi: 10.1007/s11334-022-00436-8.

[122] ADALCA, "learn2reg_mri_2d," 2019. https://www.kaggle.com/datasets/adalca/mri-2d.

[123] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young , Middle Aged , Nondemented , and Demented Older Adults," pp. 1498–1507, 2007.

[124] S. G. Mueller *et al.*, "Ways toward an early diagnosis in alzheimers disease : the alzheimers disease neuroimaging initiative (adni)," *Alzheimer's & Dementia*, vol. 1, pp. 55–66, 2005, doi: 10.1016/j.jalz.2005.06.003.

[125] A. H. and A. Dalca, "Open access series of imaging studies (oasis)," 2021. https://surfer.nmr.mgh.harvard.edu/pub/data/voxelmorph/.

[126] A. Dalca, "3-D brain MRI," 2019. https://www.kaggle.com/datasets/adalca/learn2reg-mri-3d Accessed July 2023.

[127] K. S. MADER, "CT scan:Before and After," *2019*. https://www.kaggle.com/datasets/kmader/ct-scans-before-and-after Accesssed January 2022.

[128] Z. Zhou, "A new unsupervised learning method for 3D deformable medical image registration," *IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, no. 4, pp. 908–912, 2021.

[129] J. Zhang, Z. Liu, Y. Ma, X. Zhao, and B. Yang, "Part-and-whole: A novel framework for deformable medical image registration," *Applied Intelligence*, no. 12, pp. 1–18, 2022, doi: 10.1007/s10489-022-04329-y.

[130] Z. H. and A. C. S. Chung and Lo, "Unsupervised end-to-end groupwise registration framework without generating templates," *IEEE International Conference on Image Processing (ICIP)*, no. 10, pp. 375–379, 2020.

[131] P. Arora, R. Mehta, and R. Ahuja, "An integration of meta-heuristic approach utilizing kernel principal component analysis for multimodal medical image registration," *Cluster Computing*, vol. 1, 2024, doi: 10.1007/s10586-024-04281-1.

[132] Madhu Charan, "Alzheimers-Disease-5-Class-Dataset-ADNI," *Kaggle*, 2021. https://www.kaggle.com/datasets/madhucharan/alzheimersdisease5classdatasetadni/data Accessed January 2024.

[133] A. Kori and G. Krishnamurthi, "Zero Shot Learning for Multi-Modal Real Time Image Registration," pp. 1–9, 2019.

[134] Mohrezaeib, "Brain MRI classification IXI Dataset," *Kaggle*, 2022. https://www.kaggle.com/code/mohrezaeibtest/brain-mri-classification-ixi-dataset/input Accessed March 2024.

[135] L. Deng, Q. Lan, Q. Zhi, S. Huang, J. Wang, and X. Yang, "Deep learning-based 3D brain multimodal medical image registration," *Medical & Biological Engineering & Computing*, vol. 62, no. 2, pp. 505–519, 2024, doi: 10.1007/s11517-023-02941-9.

[136] S. Mohanty and S. P. Dakua, "Toward Computing Cross-Modality Symmetric Non-Rigid Medical Image Registration," *IEEE Access*, vol. 10, pp. 24528–24539, 2022, doi: 10.1109/ACCESS.2022.3154771.

[137] H. Anzid, G. le Goic, A. Bekkari, A. Mansouri, and D. Mammass, "A new SURF-based algorithm for robust registration of multimodal images data," *Visual Computer*, vol. 39, no. 4, pp. 1667–1681, 2023, doi: 10.1007/s00371-

022-02435-z.

LIST OF RESEARCH PUBLICATIONS

SCI/SCIE

1. Paluck Arora, Rajesh Mehta and Rohit Ahuja. "An adaptive medical image registration using hybridization of teaching learning-based optimization with affine and speeded up robust features with projective transformation." *Cluster Computing* (2023), Springer Nature, pp: 1-21. *(IF- 3.6)* ***(Published)***
2. Paluck Arora, Rajesh Mehta and Rohit Ahuja. " An integration of meta-heuristic approach utilizing kernel principal component analysis for multimodal medical image registration." *Cluster Computing* (2024), Springer Nature, pp: 1-21. *(IF- 3.6)* ***(Published)***
3. Paluck Arora, Rajesh Mehta and Rohit Ahuja. "Deep-UEO: Empowering Medical Image Registration with Hybrid Strategy based on Deep Learning and United Equilibrium Optimizer." Computers and Electrical Engineering, Elsevier *(IF- 4.0)* ***(Published)***
4. Paluck Arora, Rohit Ahuja and Rajesh Mehta. "Deep VGG19-SURF Feature Extraction with Projective Transformation for Anatomical and Functional Medical Image Registration" Soft Computing, Springer *(IF- 3.1)* ***(Under Review)***
5. Paluck Arora, Rajesh Mehta and Rohit Ahuja. "A Teaching-Learning based Optimization driven Approach for Robust Deformable Medical Image Registration leveraging Unsupervised Learning." Concurrency and Computation: Practice and Experience, Wiley *(IF- 1.5)* ***(Under Review)***

Conferences

1. Paluck Arora, Rajesh Mehta and Rohit Ahuja. "A Novel Image Alignment Technique Leveraging Teaching Learning-Based Optimization for Medical Images." Lect. Notes in Networks, Syst., Vol. 785, Springer Nature, pp: 1-18.
2. Paluck Arora, Rajesh Mehta and Rohit Ahuja. Enhancing Image Registration leveraging SURF with Alpha Trimmed Spatial Relation Correspondence" 24th International Conference on Computational Science and its Applications, LNCS series, Scopus index.