AUT Journal of Modeling and Simulation

AUT J. Model. Simul., 55(1) (2023) 53-70
DOI: 10.22060/miscj.2023.21791.5303

Classification and Evaluation of Muti Modal Medical Image Registration Methods and Similarity Measures

Z. Mohamadi, M. R. Keyvanpour*

Department of Computer Engineering, Faculty of Engineering, Alzahra University, Tehran, Iran

ABSTRACT: Image registration is a fundamental issue in medical image analysis. It refers to the matching process between two or more images using the optimization of a similarity metric to find an optimal transformation function. In recent decades, many studies have been done on the medical image registration topic. Therefore, this paper has investigated four main methodologies to solve the registration problem in medical applications. One of the most important topics in this area is the registration of multi-modal images. In this paper, we have reviewed various multimodal image registration techniques based on deep learning and proposed a classification for these methods. Also one of the essential components of the medical image registration framework is the similarity measure function. There are different similarity metrics in this area and choosing an appropriate measure according to the application is a challenging problem. This paper is to present a review of different similarity measures in medical applications and a classification of these methods. Based on this classification, techniques are investigated and each subclass is evaluated using performance criteria. Therefore, the main goals of this article are as follows: 1) Investigating the most significant image registration approaches. 2) Systematic review of deep learning-based multimodal medical image registration and classify them. 3) Providing classification for various similarity measure techniques according to registration applications. 4) Creating an appropriate platform for evaluating these approaches and introducing the main challenges.

Review History:

Received: Sep. 19, 2022
Revised: May, 27, 2023
Accepted: Jun. 01, 2023
Available Online: Oct. 10, 2023

Keywords:

image registration
medical image registration
multi-modal registration
learning-based approaches
similarity measure

1- Introduction

Image registration refers to the process of matching two input images that have been obtained in multi-time, multi-view, or multi-modal. The goal of this process is to achieve the optimal geometric transformation between similar structures from two input images. Image registration is a significant topic in the field of medical image processing. Image registration is used in surgery and treatment applications such as fusion of multimodal medical images, planning for treatment, diagnosis of diseases, and physician assistant in surgery. In addition to clinical application, image registration can be used in remote sensing and computer vision [1, 2].

In recent decades, the development of various imaging techniques in clinical applications has led to many studies in the field of multimodal medical image analysis. Researchers have proposed various methods of multi-modal image registration to increase matching accuracy. The most important challenge of multimodal registration methods is the difference in the nature of the images, which leads to the lack of one-to-one relationship between the intensities of similar structures in two images [3]. Therefore, in this article, we have reviewed various studies in learning-based multimodal medical image registration and classified them into three main categories, and we have also investigated the main challenges and limitations in this field.

As shown in Figure 1, the main components of the registration framework in medical applications include transformation function, similarity metric, and optimization algorithm. The core idea in registration approaches is an iterative optimization of the similarity metric to find the optimal geometric transformation. The optimization algorithm defines the search process and the similarity metric determines the degree of correspondence between two image contents [1, 4].

Choosing an effective metric is the most important component of the medical image registration process. Similar medical images that are matched with different metrics can achieve results with different efficiency and accuracy [5]. There are different types of similarity metrics in this field and selecting the suitable criteria is a challenging task. Considering the importance of this issue, in this article, an overview of similarity measurement approaches in medical applications is presented, and also a classification for these techniques is proposed.

The various parts of this article are organized as follows: Section II is dedicated to related works. In section III we have presented a definition of the registration process and studied

*Corresponding author's email: keyvanpour@alzahra.ac.ir

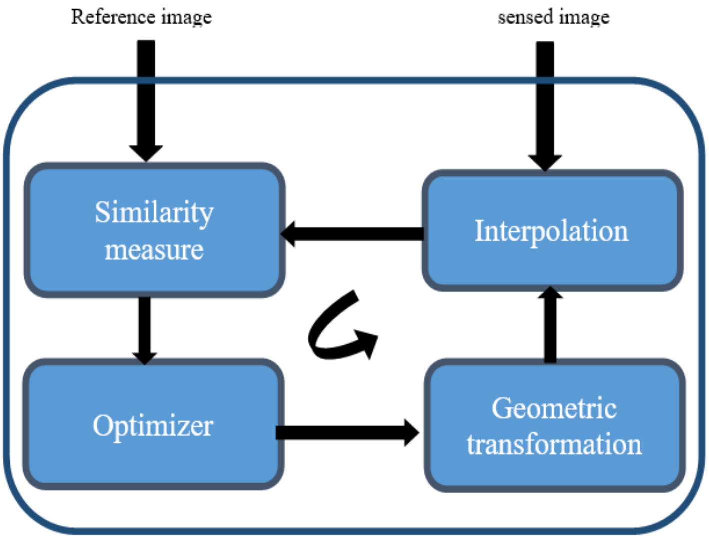

Fig. 1. Main components of image registration framework [1]

this problem based on four important methodologies in this area. In section IV, we reviewed different studies in the field of learning-based multi-modal medical image registration and classified them. In section V, different works of similarity measures in medical applications are collected and these techniques are studied based on the proposed classification. in section VI, a discussion on multimodal medical image registration methods and similarity criteria is presented, also the evaluation and limitations of these methods are investigated, and section VII is the conclusion of this paper.

2- Related Works

Medical image registration plays a significant role in clinical applications and many studies have been done in this field. In [1] they conducted a review paper on traditional medical image registration and discussed the main steps of the registration process in detail. In [6] they presented an overview of the general classification in medical image registration and studied recent developments in this field. This classical categorization is based on 9 groups, which include the dimensions of medical images, transformation methods, optimization, unimodality or multimodality, involved subject, etc. In this classification, the dimensionality of medical images is one of the subgroups. In [2] they presented a systematic review of traditional and newer registration techniques for 2D and 3D medical images. This paper is conducted to help the researchers design the efficient registration technique according to the specific scenario and application.

Another aspect of the classification mentioned above is the transformation type. While rigid transformation is used to match images of rigid body structures, most tissues use non-rigid transformation. In [7], they reviewed different types of deformation models and provided an appropriate reference for selecting the transformation function. This article reviews deformation models in four categories: elastic, viscous fluid flow, optical flow, and based on previous knowledge.

Subjects and modalities involved in image registration are other sub-categories of the mentioned classification. Multimodal image registration plays an important role in different fields, including clinical applications for diagnosis and treatment. In [8] they presented a comprehensive analysis of multimodal image registration from traditional to deep learning approaches based on image nature. In this paper, the main idea of methods and new trends in this field are investigated to help future works for designing more efficient multimodal registration approaches.

Although much research has been conducted in the medical image registration field, this subject can still be improved. In [9] they reviewed the main challenges of medical image registration in four classes including imaging technology, applied techniques, datasets, and human-related. In recent years, the development of learning-based methods has significantly improved the performance of registration algorithms for medical applications. In [10], they provided an overview of learning-based medical image registration techniques and current trends and limitations in this field. This paper can help future research to improve methods and find solutions for clinical needs.

3- Image Registration Problem

Image registration refers to the mapping process of two input images obtained from different times, viewpoints, or sensors. The goal of registration approaches is to find the spatial transformation (ϕ^*) between reference (I_R) and sensed (I_S) (input) images using similarity measure optimization. During the registration process, the sensed image is repeatedly warped using the transformation function ϕ [1, 11]. Equation 1 formulates this process:

$$\phi^* = argmax\ sim(I_R, \phi(I_S)) \quad (1)$$

Image registration is an important step in image analysis. There are various studies in the field of image registration and several approaches have been introduced to solve this problem. In this paper, we have investigated the registration problem based on four significant methodologies in this area: Registration based on Anatomical Landmarks, Extracted Features, Image Intensity Values, and Learning Methods.

Landmark-based registration is a common approach in clinical applications used to register corresponding anatomical landmarks in reference and sensed images. In these methods, the coordinates of the corresponding points in the two input images are determined, and then the spatial transformation parameters between these landmarks are calculated. In these approaches, anatomical landmarks can be determined as either intrinsic or extrinsic. In the extrinsic method, markers are connected to the patient's body. But in intrinsic techniques, landmarks are significant and extractable anatomical points that the treatment staff or expert user determines them. Manual feature extraction in these methods can make them less accurate than automatic methods[12, 13].

The feature-based registration methodology is based on the geometric information extracted from two input images. The main steps of feature-based methods usually include feature extraction, feature description, and feature matching. Feature-based methods are widely used in image processing because these approaches provide a simple representation of the input images by the few extracted features. Also, these methods have more flexibility and robustness against geometric deformation and noise. Feature-based approaches are suitable for registering images containing extractable structures such as remote sensing applications, while intensity-based methods are usually used for medical images [2, 12].

In the intensity-based method, the feature space can be the pixel's intensity value, intensity gradient, and statistical data about pixel intensity. Intensity-based methods can be formulated as an optimization problem. As shown in Figure 1, geometric transformation, similarity metric, and optimization of transformation parameters are the main components of this registration approach. In recent years, intensity-based image registration methods have been used more than feature-based. One of the important reasons for the growth of the intensity-based approach was the improvement of computing resources such as processing speed and available memory. Higher accuracy than feature-based approaches has motivated researchers to use the intensity-based method as a pre-processing step in image processing applications. In addition, the intensity-based process doesn't require the segmentation of input images, which can usually cause complexity. In intensity-based methods, the selection of similarity measures and overlap between input images can significantly affect the performance of these algorithms [8, 14, 15].

Learning-based registration methodology has grown significantly in recent years, and several studies have focused on medical images. Learning-based approaches can be replaced by the classical framework in feature extraction, similarity metric, and parameter mapping estimation. For example, in a learning-based manner, instead of utilizing a similarity measure based on intensity values, the similarity metric can be learned using pre-aligned images. Learning-based approaches have more accurate and efficient results in medical applications compared to traditional methods [16, 17]. Although learning-based methods have grown significantly in medical image registration, this field still remains an open topic and has challenges that will be discussed in the following sections.

4- Multi-Modal Medical Image Registration Approaches

In recent decades, the development of different imaging techniques and the use of these technologies in medical applications has increased the number of research in the field of multimodal image registration. The most important challenges of multi-modal registration methods are the difference in the nature of the images and the lack of one-to-one mapping between the image intensities of similar areas. These limitations reduce the accuracy of registration approaches, especially in traditional frameworks. In recent years, many researchers have developed deep learning-based methods in this field to overcome the aforementioned challenges [8, 18]. In this part, we have reviewed various studies of deep learning-based multimodal medical image registration, and categorized them into three main classes, as follows:

- Modality Transfer Methods
- Iteration-based Methods
- One Step Methods

Each class consists of two sub-classes. Figure 2 shows the proposed classification.

4- 1- Modality Transfer Methods

One strategy that aims to multi-modal medical image registration is modality transfer learning which reduces the multimodal transformation problem to the mono-modal one. These approaches map different modalities of input images to the same domain, and this makes it possible to use standard similarity measures for multi-modal image registration [19]. In recent years, several studies have been conducted on multimodal representation. Based on these studies, modality transfer methods can be divided into two general categories, including learning feature representation, and simulation learning based.

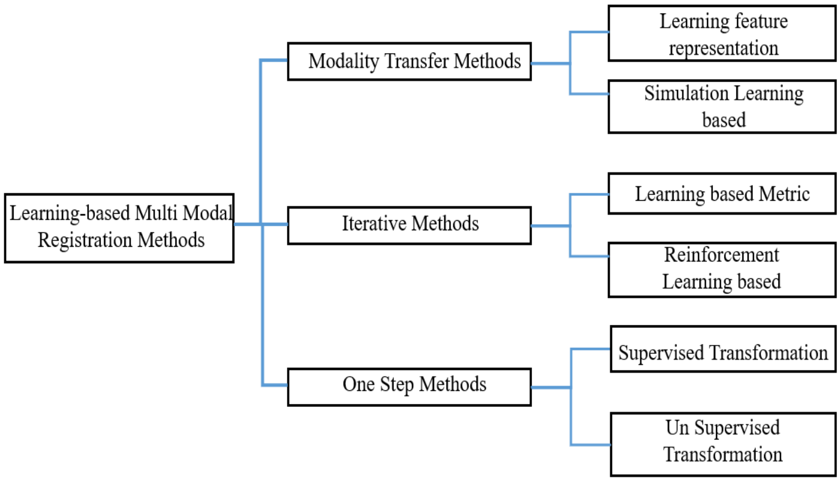

Fig. 2. Learning-based Multi-Modal Medical Image Registration

4- 1- 1- Learning Feature Representation

Due to the large intensity differences between the two modalities, manually engineered image features are not useful for multi-mode image registration. Therefore, some methods have been introduced to learn the common feature representation from two multi-modal input images [11]. In this section, methods based on structural representation (SR) are reviewed; which reduces the problem of multi-modal registration to a single-modal. These methods work on the assumption that there is similar structural information between input images. In these works, simple similarity criteria such as SSD and Euclidean distance can be used for image matching [8, 19].

In [19] they proposed a new feature representation technique for rigid and deformation multi-modal image registration. The structural image in this work is calculated based on the combination of entropy image and gradient-based information. After mapping the multimodal input images into structural images, traditional similarity metrics such as SSD and MI have been applied.

The aforementioned approaches are based on handcrafted extracted features, which have a high probability of error when dealing with complex medical images. To overcome this problem, learning-based methods can be used. In these methods, learning features of complicated medical images is very important for effective structural representation [20]. In [20] they presented a new deep learning-based structural representation registration (SRR) manner for multimodal medical images. In this work, a novel Laplacian Eigenmaps-based deep network is used to extract the features of input images with different modalities. Then the self-similarity of these features is utilized to build LDAD (Learning Data Adaptive Descriptor) for structural representation. The similarity measure between the structural representation of input images is the sum of squared differences (SSD). In [21] they proposed a novel structural representation method based on a deep learning network (PCANet) for non-rigid multimodal medical image registration. In this work, first, PCANet is trained on a large number of multimodal medical images to learn the convolutional kernel of the network. In the next step, medical images are given as input to the trained PCANet to be registered. The extracted features in different layers of the network are combined to make multi-level features. Finally, this multi-level information is used to produce an effective structure representation of medical images. In this work, Euclidean distance is used as a similarity metric between structural representations.

4- 1- 2- Simulation Learning based

As mentioned earlier, multi-modal image registration is more challenging due to the large appearance difference between modalities, which makes it difficult to apply existing similarity measures. Another way to convert a multi-modality problem into a single-modal one is to simulate one modality from another by considering to physical features of medical imaging. This section reviews some learning methods that have been used to learn a synthesis model between two input image modalities.

In [22], they proposed a synthesis method based on CycleGAN to reduce multi-modal registration to single-modal one. In this work, it was shown that the use of a CT synthesis derived from deep learning, instead of MRI in MRI-CT registration, improves the performance of head and neck image registration. In [23] they proposed a method based on bidirectional image synthesis dual core for multimodal medical image registration. In this paper, to improve the registration accuracy, a structured random forest is used to learn the synthesis model of CT from MRI as well as MRI from CT. Experimental results showed that bidirectional multi-modal image synthesis can effectively reduce registration bias and improve performance compared with one-directional synthesis [24]. In [25] they applied a FCN (ten-layer fully convolutional network) to learn transformation from one input modality to another. After image synthesis, the registration process can be performed by a simple similarity measure such as SSD.

Recently, some image representation methods have been introduced that convert different modalities of input images into a common third modality. After this mapping, the registration problem can be solved using a simple similarity measure. These methods work under the assumption that exists similar anatomical structures in two input images [19]. In [26] they presented a novel multi-modal deformation network by utilizing a two-channel registration and image simulation. In this work, they applied a probabilistic Cycle GAN to map multi-modal to single-modal registration in both CT and MRI channels. This paper utilizes a CT and MRI image synthesis sub-network with a two-channel registration sub-network and combines the two to estimate a single diffeomorphic deformation field.

4- 2- Iteration based Methods

Learning-based multi-modal image registration is an important research area and many papers have been conducted in this field. Early studies of learning-based methods in medical image registration are direct learning of similarity measures in the classical framework. Later, several works applied reinforcement learning to iteratively estimate a mapping function. In this section, firstly, the methods that use deep learning in the classical framework and then the registration based on deep reinforcement learning are reviewed.

4- 2- 1- Learning-based Metrics

Deep learning has been applied to iteratively learn improved similarity measures in an intensity-based framework. This group of deep learning-based methods is more applicable in multimodal registration, where traditional similarity measures don't achieve acceptable results. Traditional similarity measures can be replaced by superior learning-based metrics such as stack autoencoder or CNN [16].

In [27] they proposed a CNN-based similarity metric for multimodal medical image registration. This network is trained from pre-aligned image pairs and then iteratively estimates the transformation parameters in a classical matching framework. This learning-based metric shows improved registration performance compared to traditional metrics such as MI. In [28, 29] proposed deep learning-based similarity measure in the iterative framework for multi-modal medical image registration. Also in [30], they used CNN to learn the similarity measure for MR-TRUS registration. Although robust MR-TRUS registration is a challenging problem due to the large difference between the two modalities, it can be solved by using a learning-based similarity measure and appropriate optimization strategy. In this work, raw pixel data is used as input, and learned features are utilized to calculate the similarity metric.

4- 2- 2- Reinforcement Learning based

Estimating the mapping function iteratively using reinforcement learning is another iteration-based method in this field. In these approaches, an artificial agent is trained to learn policies using observation from the environment. The agent's decision-making about mapping the current state to the best action is done based on the control policy and environmental reward. In this section, works that have applied reinforcement learning to register multimodal medical images are reviewed. In these works, optimization methods are replaced by a trained agent in the registration process [31].

In Figure 3, a registration framework based on reinforcement learning is shown. This framework is more suitable for rigid transformation, although it can also be used for non-rigid registration. In [32] they converted a multi-modal image registration problem into a decision-making problem. In this work, the registration process is performed using an agent trained by reinforcement learning. In this method for Spatiotemporal feature extraction, they integrated CNN and conLSTM (convolutional short-term memory) in a reinforcement learning framework. Experiments on clinical datasets showed that this work achieved high performance in multimodal medical image registration.

In [33] they presented a multimodal registration method based on deep reinforcement learning instead of an optimization process. In this algorithm, the contextual features of medical images are provided by Q deep learning to reduce the appearance differences between the two modalities. Using contextual information of input images can improve robustness against noise and various data. In [34], instead of training a single agent, they used a multi-agent approach (with an automatic attention mechanism) in the RL framework to register multimodal spine images. This method can improve efficiency and robustness in the field of 2D/3D medical image registration.

In recent years, research has shown that reinforcement learning has more accurate results in registering multimodal medical images compared to previous methods. However, one of the major challenges of these methods is the high training time due to their iterative framework. Another challenge is the ability to deal with the complex deformation domain in non-rigid multi-modal image registration. Due to the importance of these methods in medical applications, it is expected that these topics will be more attention by researchers in the future [35].

4- 3- One Step Methods

Due to the high computational load of the iteration-based method, researchers are interested in using the end-to-end

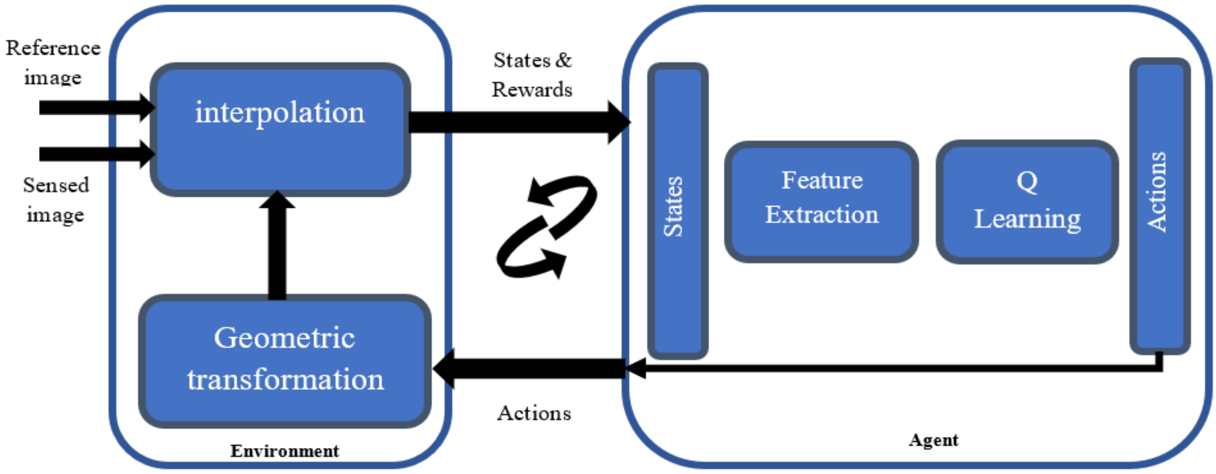

Fig. 3. Deep Reinforcement Learning-Based Registration

method in their works. In other words, the limitations of the iteration-based method can lead to a time-consuming registration process, especially when dealing with complex deformation registration. Therefore, in recent studies, the proposed networks can estimate the transformation parameters in one step [36]. This section first reviews approaches that use the one-step supervised transformation methods and then focuses on unsupervised transformation in the image registration method.

4- 3- 1- Supervised Transformation

In this part are reviewed studies that apply fully supervised and semi-supervised methods for one-step transformation estimation. In these methods, a deep network is used instead of an iterative optimization process to increase registration performance. In [37] they proposed a fully supervised regression convolutional network to generate rigid transformation parameters based on image features. In this work, the authors defined the loss function using a bivariate geodesic distance and also applied a residual network and a correction network to register T1 and T2 weighted MRIs. The residual network before the correction network is utilized for the initial registration of image volumes and can improve registration accuracy.

Fully supervised approaches need to ground truth data for network training. The biggest limitation is the lack of these samples with corresponding transformation parameters, which leads to the development of weakly supervised and dual-supervised registration. In weakly supervised methods, the similarity of segmented labels is used to train the network and estimate the transformation parameters. In [38] they proposed a novel deep learning-based framework called constrained affine network (CAN) to learn complex deformation between 3D multi-modal medical images. This work used a weakly supervised method to train a convolution network that can learn to predict a DVF (displacement vector field) between two fixed and moving image spaces.

Considering the limitations of fully supervised methods in the training phase, researchers proposed methods based on dual-supervised learning. In these approaches, both ground truth guidance and similarity metrics are used to estimate the deformation field. Metric-based guidance is utilized to solve the lack of ground-truth data and avoid the dependence on pre-aligned images [39]. In [40] they proposed a dual-supervised deformation estimation model (DDEM) to generate ultra-quality (UQ) 4D medical images from a set of low-quality data. MRI image volumes with T1-/T2 weight are given to DDEM, and UQ 4D-MRI images are obtained using a predicted deformation (DVF) between the reference and sensed images. Network training is performed using a dual-supervision strategy including pre-aligned images and the normalized cross-correlation (NCC) similarity measure. The experiment results show that the registration performance under a dual supervision network is significantly higher than in previous works.

4- 3- 2- UnSupervised Transformation

Considering the limitation of the supervised approach, several unsupervised methods are utilized to estimate the deformation model in an end-to-end framework. In unsupervised methods, the network structure doesn't change compared to the supervised one, but these methods are trained without any pre-aligned images. Unsupervised approaches can use fixed similarity metrics to define loss functions. In [41], they proposed an unsupervised learning-based approach in an end-to-end framework to register 3D multi-

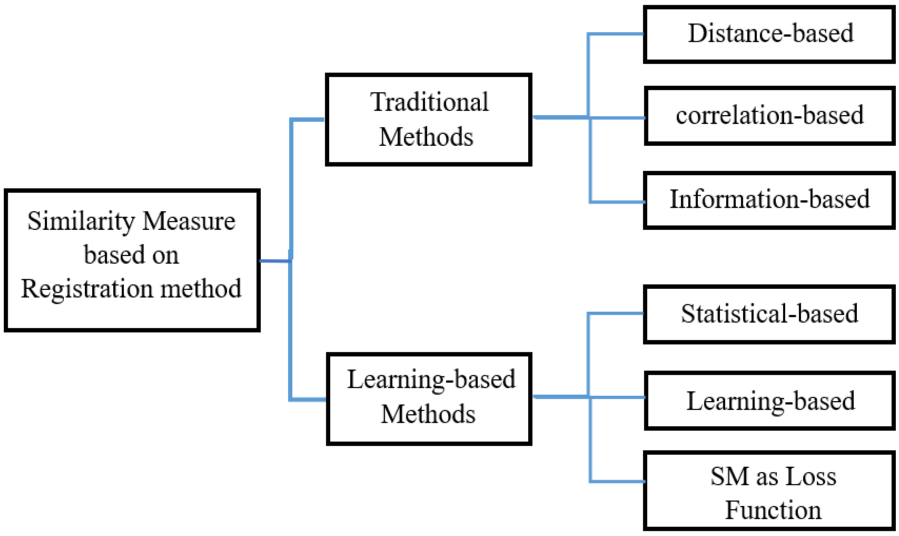

Fig. 4. Classification of similarity metrics in medical image registration

modal medical images. In this work, to predict the 3D DVF (displacement vector field) between two input images, they used a fixed similarity metric for the optimization process in the training phase. Also, in [42] they proposed a novel deep network named UDIR-Net to register 3D medical images. UDIRNet (Unsupervised Deformable Image Registration Network) is designed to directly predict the deformation model between two input images. In this work, they train the network without any ground truth data and based on a hierarchical loss function strategy such that the metric-based loss function is calculated at different levels of the network. This type of training improves registration performance and network optimization compared to the previous single loss function.

5- Similarity Metrics in Medical Image Registration

Choosing an effective similarity measure is the most important step in the medical image registration process and affects the matching accuracy. Similarity metrics are applied to assess the similarity between each region in the reference image and the homogeneous region in the sensed image. Different metrics are available in medical image registration, each of which has its benefits and disadvantages, and is utilized in various applications [43]. In this part, we have reviewed different studies on similarity measures in medical image registration and categorized them based on registration methodology. The two main groups of this classification are as follows:

- Traditional approach
- Learning-based approach

As shown in Figure 4, Each category of the proposed classification has three subcategories.

5- 1- Traditional Approach

Traditional similarity metrics refer to statistical methods that work on voxel intensities or spatial structures and calculate the correspondence between two input images [5]. Based on the registration problem and feature space, traditional methods can be categorized into three main groups: distance-based, correlation-based, and information-based.

5- 1- 1- Distance-based Metrics

The distance-based measure is a simple metric used to calculate the intensity difference of corresponding pixels or the distance of geometric features in mono-modal image registration. Sum of Absolute Difference (SAD), Sum of Absolute Difference (SSD), and Maximum of Absolute Differences (MAD) are examples of distance-based methods and work under the assumption that the intensity relationship between two input images is linear [44]. Geometric distance-based methods are used to calculate the distance between features extracted from two input images. These features can be two sets of points or lines extracted from two images, and the SSD or chamfer method is used to calculate the distance between them [1].

Recently, some papers applied a new approach based on image representation techniques, which transforms the complex registration problem into a simpler one. In these papers, distance-based similarity measures such as SSD and Euclidean distance can be used to solve the registration problem after multimodal to monomodal mapping [45]. The registration approaches based on the structure representation described in [21] are examples of these algorithms.

5- 1- 2- Correlation-based Metrics

The intensity correlation-based technique is another similarity measure that is applied in traditional medical image registration methods. The correlation-based methods calculate the statistical relationships between two sets of image intensity data. These methods are in their initial stages and are considered due to their simplicity and acceptable results. These metrics are used under the assumption that the intensity correlation between two images is linear. Cross-correlation (CC) is computed for a window pair from two input images, and this measure should be maximized during the registration process. Normalized Cross-Correlation (NCC), Correlation Ratio (CR), and Pearson's Correlation Coefficient are examples of correlation-based similarity measures. One of these measures applied for template matching is NCC, which is shown in equation 2 [46]:

$$NCC = \frac{\sum_{(i,j)}(X(i,j)-\bar{X})\sum_{(i,j)}(Y(i,j)-\bar{Y})}{\sqrt{\sum_{(i,j)}(X(i,j)-\bar{X})^2\sum_{(i,j)}(Y(i,j)-\bar{Y})^2}} \tag{2}$$

X(i,j) and *Y(i,j)* are the pixel intensity values at coordinates *(i,j)*. For example, in [47] they presented a 2D/3D medical image registration method based on the Normalized Cross Correlation metric. In [48] they proposed a multi-resolution manner for multi modal medical image registration. In this work, the input images are converted into several resolution levels and each level of the hierarchy is registered with a different similarity measure such as cross-correlation. This method has more accuracy and less calculation time compared to single resolution registration.

5- 1- 3- Information-based Metrics

Information-based metrics are applied to analyze the information flow between two signals. The most important measure in information-based approaches is mutual information, which is based on density functions. Histogram and kernel density estimation are techniques for calculating probability density functions between two signals [49]. If *X* and *Y* are two random variables, the Mutual information (MI) similarity measure can be defined as:

$$MI\ (X,Y) = H(X) + H(Y) - H(X,Y) \tag{3}$$

In this formula, the Shannon's entropy for variables *X* and *Y* are shown as *H(X)* and *H(Y)*, also *H(X,Y)* refers to the joint Shannon entropy.

Mutual information is suitable for multimodal image registration due to the use of statistical relationships between image intensities. Many researchers in this field have been interested in using mutual information in their works, and several advanced versions have been proposed [50], including correlation ratio-based MI [51], and normalized mutual information (NMI). In [52] they presented weighted mutual information for registration. In this work, they calculated a more accurate version of MI that allocated different weights to image patches.

5- 2- Learning based Approach

Traditional similarity measures are predefined and have some limitations such as dependence on initialization, time-consuming, and local minima. In addition, the degree of overlap between input images is an important factor in the performance of these measures. Therefore, these metrics have some challenges with different intensities or flat areas. The learning-based technique is an appropriate solution for the aforementioned limitations [53]. In this paper, learning-based similarity metrics in medical image registration are divided into three classes: statistics-based, learning-based, and similarity metric as a loss function.

5- 2- 1- Statistical based Metrics

Statistical-based similarity measures refer to a category of methods that use the pre-aligned dataset to learn the intensity distribution of images. Registration methods use this technique to calculate the similarity between two image intensity distributions. In [54] they presented a medical image registration approach that used learned intensity distribution and KLD (Kullback-Leibler distance) similarity metric. As shown in equation 4, KLD calculates the degree of difference between observed (P_o^T) and expected ($\hat{P}$) distributions:

$$D\left(P_o^T \| \hat{P}\right) = \sum_{i_f,i_r} P_o^T(i_f,i_r)\ log\frac{P_o^T(i_f,i_r)}{\hat{P}(i_f,i_r)} \tag{4}$$

The goal of this method is to minimize the distance between the P_o^T and $\hat{P}$ distributions to achieve an optimal transformation (T_0):

$$T_0 = \arg min_T\ \mathrm{D}\ (P_o^T \parallel \hat{P}) \tag{5}$$

As another statistical-based method can refer to [55], which applied Jensen-Shannon divergence (JSD) as a similarity measure for medical image registration. Jensen-Shannon is superior to KLD metric in terms of theoretical aspects and symmetry. In [56] they presented Bhattacharyya divergence (BD) to measure the difference between $P^{\wedge}(i_f i_r)$ (learned distribution) and $P_0^T(i_f i_r)$ (observed intensity distribution) for medical image registration. BD is defined in equation 6:

$$BD = -\log \int\int \sqrt{P_0^T(i_f,i_r)P^{\wedge}(i_f,i_r)}\ .d_{i_f}d_{i_r} \tag{6}$$

In the registration process, the Bhattacharyya distance is minimized according to the transformation function.

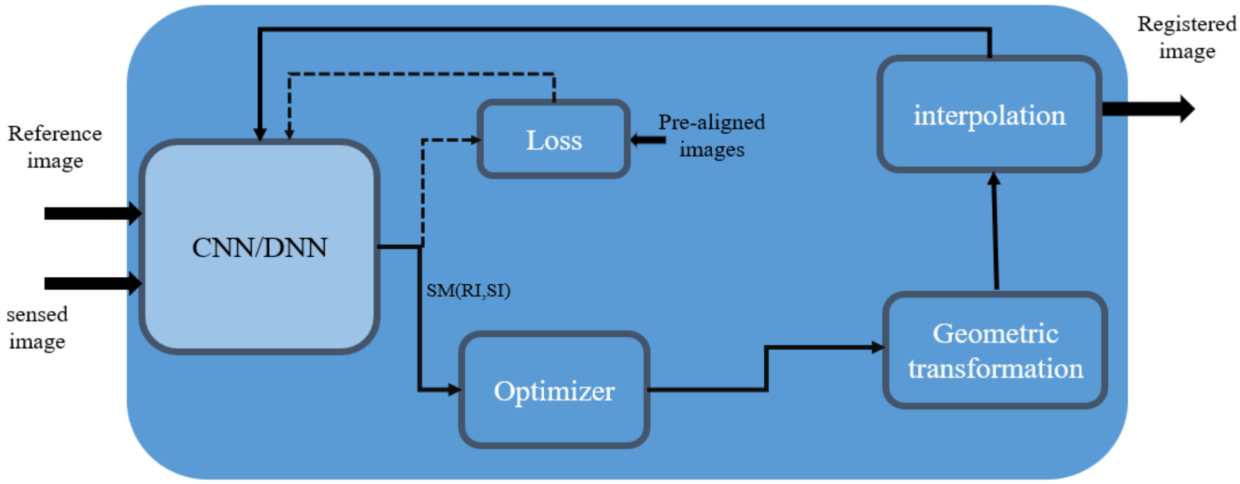

Fig. 5. Learning-Based Similarity Metric in the Traditional Framework

5- 2- 2- Learning-based Metrics

The introduced methods in this section directly learn the similarity criteria so that the input image pairs achieve a high degree of similarity. Unlike traditional methods that determine whether two images are similar or not, these metrics provide a continuous similarity score. In recent years, deep learning-based methods have grown significantly in the field of medical image registration, these approaches are also applied to learn similarity metrics directly. As shown in Figure 5, these metrics are applied in a traditional registration structure and work in interaction with other components such as geometric transformation, optimization algorithm, and interpolation techniques [57]. Regarding images 6 and 7, note that the dashed lines are related to the training stage of the neural network.

In [58] they proposed a deep learning-based similarity metric to improve the performance of mono modal intensity-based registration framework. This work applied a pre-trained convolutional neural network as a similarity measure for multi-temporal 3D ultrasound image registration and achieved higher accuracy compared to traditional methods.

In [29] they presented a CNN-based similarity metric for multimodal medical image registration. In this paper, they used a supervised learning-based method in a traditional intensity-based structure to achieve a high-performance similarity measure.

5- 2- 3- Fixed Similarity Metrics as Loss Function

The end-to-end approach refers to another group of methods that use deep learning to solve the registration problem. This approach estimates the deformation field between two input images in one step. In previous works, the feature extraction step was performed separately by feature matching, but in the end-to-end method, the transformation function is learned directly from the input images. End-to-end supervised methods have some limitations such as lack of medical datasets and overfitting of data, so researchers became interested in using semi-supervised and unsupervised methods in their works [59, 60].

The end-to-end registration process (deformation field prediction) is optimized in the training phase using information feedback. As shown in Figure 6, traditional similarity metrics such as NCC can be utilized in end-to-end unsupervised structures as a loss function to register images. Unlike previous learning-based works that used prior knowledge to train the network, in [60-62] they applied unsupervised CNN approaches for medical image registration. In these deep networks, the training phase is performed using a metric-based loss function strategy and by optimizing the correlation-based similarity measure. Also in [63], they presented an end-to-end unsupervised method that used a traditional metric to measure differences between medical images.

In [64] they proposed a novel unsupervised network called MS-DIRNet that used similarity metric as a loss function to register large data. In this paper, global and local training is used in a multi-scale strategy for accurate deformation estimation.

6- Discussion Similarity Metrics in Medical Image Registration

Considering the importance of multimodal image registration in clinical applications and the significant growth of deep learning-based methods, an overview of these approaches is presented in this article. In this section, we have evaluated these techniques and proposed a classification for the main challenges in this field. Also, the analysis of different similarity measure techniques based on performance

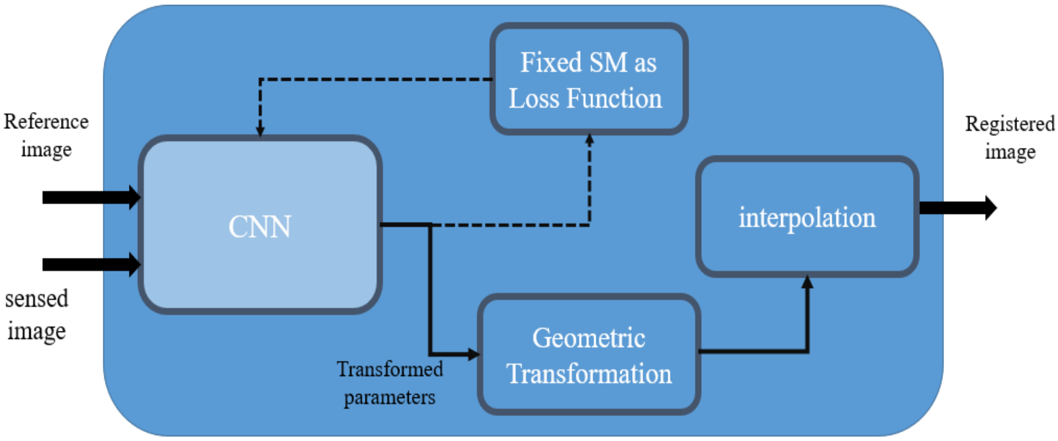

Fig. 6. Standard similarity metric as Loss Function in One Step Framework

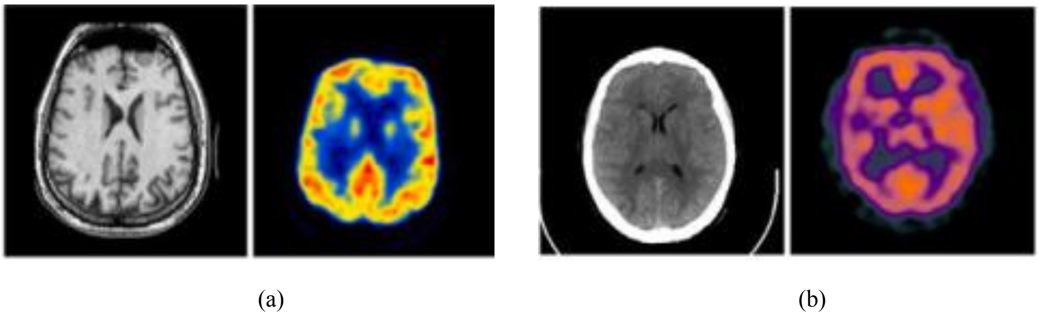

Fig. 7. Multi-modal medical images, a. MRI-PET b.CT-SPECT

criteria such as accuracy, speed, robustness, and complexity is presented in the second part.

6- 1- Evaluation of Techniques and Classification of Challenges in Multi-Modal Medical Image Registration

Multimodal image registration is one of the most important topics in medical imaging. The purpose of this type of registration is to align similar anatomical structures obtained from different imaging techniques. The two main categories of these image modalities in clinical applications are anatomical and functional images. As shown in Figure 7, anatomical images such as MRI and CT are used to present anatomical structures, but functional images such as PET and SPECT show functional deformation during the registration process. Therefore, these two modalities require to be combined using fusion and registration approaches to obtain complementary information.

In this section, firstly, we proposed a classification of the main challenges of multimodal medical image registration. Because of the different nature of images, large deformations, and technical limitations, the problem of multi-modal medical image registration has faced researchers with many challenges. Therefore, in this paper, a proposed classification for the main challenges is presented. As shown in Figure 8, the existing challenges are categorized into two main groups: data and techniques; and several subclasses that are explained in this section. The main challenges related to multi-modal medical image registration data are presented as follows:

- Heterogeneous and Various Intensity
- Need to PreProcessing
- Feature Detection
- Lack of Data Set

One of the most important challenges of multimodal registration is related to the nature of medical images. Considering that multi-modal images are achieved by different imaging approaches, these images have heterogeneous and various intensities, which makes some classical and even newer methods useless in solving the registration problem.

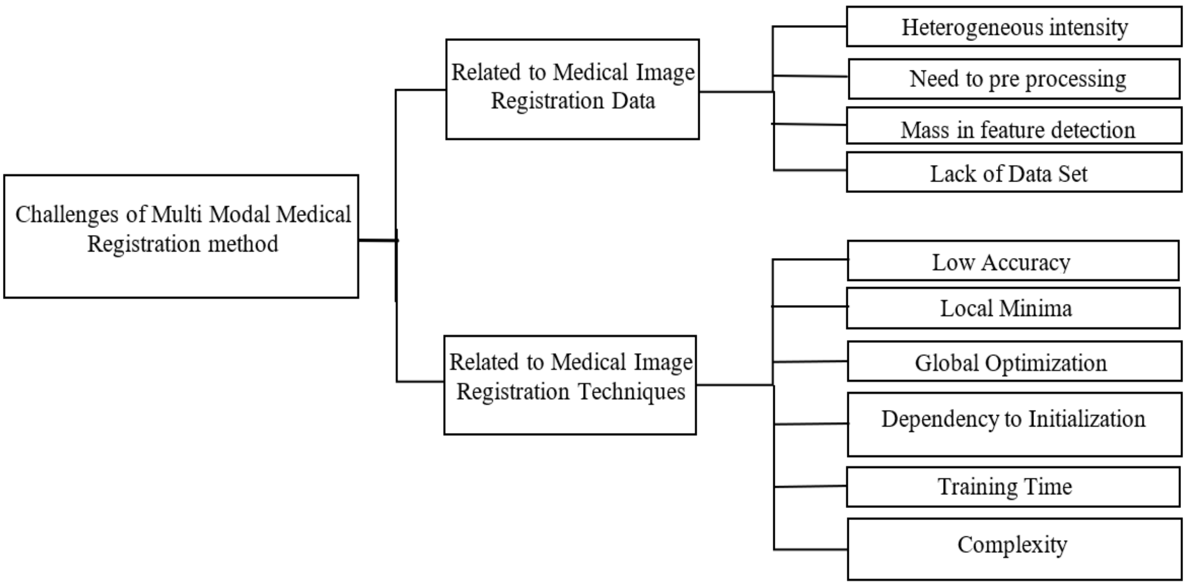

Fig. 8. Main challenges of Multi-Modal Medical Image Registration

Considering that multimodal medical images have low quality and differences in resolution and format, to reduce the existing gap, these images need a preprocessing step before analysis [65].

Feature detection from multimodal medical images is another limitation in this field, which is caused by the nonlinear intensity relationship between these image pairs. Since medical images have few detectable details and extractable features, many feature-based approaches cannot be efficient in clinical applications. On the other hand, feature extraction using manual or semi-automated methods can lead to inaccurate results [8].

Another challenge in this area is the availability of medical data sets. This limitation could be due to the sensitivity of the clinical data set and patient privacy. This problem causes the research works to be conducted with small data sets and restricts the methods' evaluation [9].

The main challenges related to multi-modal medical image registration techniques are presented as follows:

- Low Accuracy
- Local Minima
- Global Optimization
- Dependency to Initialization
- Training Time
- Complexity

According to previous research, the accuracy of methods is the most important challenge in clinical application. Although there are many proposed techniques in this field, achieving high accuracy, especially in multi-modal image registration, is still an open issue for researchers.

Global Optimization and Local Minima are other challenges in multimodal medical image registration, especially in traditional methods. Considering that the traditional registration method requires a local or global optimization process, global optimization for large search spaces in multimodal images has a high computational burden and is time-consuming. On the other hand, local optimization cannot find the optimal value when faced with different intensities in the images and falls into local minima [17].

Another source of drawbacks is the dependency on initialization in some traditional methods, which can affect the results of the registration process. In these algorithms, the starting point of similarity measure optimization is important to achieve optimal transformation parameters and registration efficiency.

The training step in learning-based registration methods can be time-consuming. In this phase, multiple iterations are needed to adjust the network parameters and find the optimal deformation model. The time requirements of these methods are due to the iterative nature of the training phase, which can cause the registration process to take longer [66].

Complexity in learning-based methods is another challenge in this field. Learning-based methods, especially deep learning, have complex structures compared with previous works. The complexity of these methods makes them more challenging to implement and understand [36].

Multimodal medical image registration based on deep learning approaches is an active research field and many papers have been conducted in this area. Learning-based methods are used in multimodal registration to overcome the gap between different image modalities. In this section, our proposed classification in deep learning-based multi-modal medical image registration is evaluated in terms of advantages and disadvantages in Table 1.

Table 1. Classification and Evaluation Learning-based Multi-Modal Medical Image Registration Approaches

Approaches		Methods	Core idea	Advantage	Disadvantage
Modality Transfer Methods	Learning Feature Representation	Combination of deep-learning-based Laplacian Eigenmaps network and self-similarity to extract features and structural representation [20]. In [21] they register multi-modal medical images using structural representation based on PCA deep network.	These methods are based on structural representation (SR) which reduces the problem of multi-modal registration to a single-modal. These methods work on the assumption that there is similar structural information between input images.	• SR method reduces the problem of multi-modal registration to a mono-modal • Fixed similarity metric can be used for these methods.	• These approaches work on the assumption that there is similar structural content between the two input images.
	Simulation Learning	In [22], they proposed a synthesis method based on CycleGAN to reduce multi-modal registration to single-modal. In [23] they proposed a method based on bidirectional image synthesis dual core. In [25] they applied a FCN to learn the transforming function from one input modality to another. In [26] they presented a novel multimodal deformation network by utilizing two channel registration and image simulation methods.	Another way to convert a multi-modality problem into a single-modal one is to simulate one modality from another considering to physical features of medical imaging.	• Reduce multi-modal registration problem into a monomodal one • Bi-directional image synthesis increases registration performance.	• The synthetic direction from each modality to another is not efficient and the synthesis direction should be from one modality with rich information to another.
Iteration based Methods	Learning-based Metric	In [27-29] they proposed a CNN-based similarity metric for multimodal medical image registration. Also in [30], they used CNN to learn the similarity metric for MR-TRUS registration.	Deep learning has been applied to iteratively learn improved similarity measures in an intensity-based framework.	• High accuracy • Similarity measures can be learned directly from pre-aligned images • significant Intensity differences of multi-modal images don't affect these metrics.	• Use of classic registration framework with iteration-based strategy.
	Reinforcement Learning	In [32,33] the registration process is performed using an agent trained by reinforcement learning. In [34], they used a multi-agent approach in the RL framework to register multimodal spine images.	Estimating the mapping function iteratively using reinforcement learning is another iteration-based method in this field. In this approach, an artificial agent is trained to learn policies using environment observation.	• The pre-trained agent is used to complete the registration process instead of optimization methods • This approach is appropriate for the registration of rigid organs.	• These methods are time-consuming due to the iteration-based strategy. • Non-rigid transformation estimation is a challenging problem for these methods.
One Step Methods	Supervised Methods	In [37] they proposed a fully supervised regression network to generate transformation parameters. In [40] they proposed a dual-supervised deformation estimation model to generate ultra-quality medical images. In [38] they proposed a weakly supervised learning-based framework (CAN) to learn complex deformation.	This group refers to approaches that use supervised networks to estimate transformation parameters in one step.	• This method is useful for dealing with large and various datasets. • High Speed of convergence • This registration approach is completed in one step.	• Training data can affect the quality of registration • Lack of training dataset. • Data overfitting
	Unsupervised Methods	In [41], they proposed an unsupervised learning-based approach in an end-to-end framework to register 3D multi-modal medical images. In [42] they proposed a novel deep network that uses a hierarchical loss function to register 3D medical images.	In unsupervised methods, the network structure doesn't change compared to the supervised one, but these methods are trained without any pre-aligned images. unsupervised approaches can use fixed similarity metrics to define loss functions.	• High accuracy • Suitable for large deformation estimation without training data • High efficiency and robustness • One step transformation estimation.	Complexity of methods

Table 1. Classification and Evaluation Learning-based Multi-Modal Medical Image Registration Approaches

Similarity Measure Techniques	Functional Metrics				Main Challenge
	Accuracy	Speed of convergence	Robustness	Complexity	
Distance-based	Low	Medium	Low	Low	Sensitive to difference in image Intensities
Correlation-based	Low	Low	Low	Low	The choose of corresponding windows size
Information-based	Medium	Low	Low	Low	Time Consuming
Statistical-based	Medium	Low	Medium	Medium	Dependency to previous knowledge
Learning-based	High	High	High	High	Lack of Pre-alignment Data
Fixed Similarity Metric as Loss Function	High	High	High	High	Complexity of Methods

In this classification, based on previous studies, the deep learning-based method to solve multi-modal registration is divided into three main categories including modality transfer methods, iteration-based methods, and one-step transformation estimation. In the first group of the proposed classification, modality transfer methods are considered, which map a complex multimodal registration to a simpler, unimodal problem. In these methods, the fixed similarity metric can be used to register the new representation of the images successfully. During the process of creating a modality simulation or image representation, it is very important that the generated image is similar to the original medical images. however, this problem depends on the image content and details may still be lost in these methods.

In the second group of deep learning-based multimodal registration, we refer to iteration-based methods, which include learning-based similarity measurement in the traditional framework and iteration-based transformation estimation using reinforcement learning. Although iteration-based methods have achieved some improvements in advanced similarity measures and more accurate transformation estimation, these methods are limited by their iterative nature and time requirements and are not suitable for real-time applications. Also, reinforcement learning-based transformation is more focused on rigid registration.

Due to the limitations in the previous methods, especially in non-rigid deformation estimation, deep learning-based methods have been developed to estimate the transformation parameters in one step, which are studied in the third group of the classification. Compared with iteration-based methods, one-step registration approaches significantly increase the speed in the field of multimodal medical image registration. Although these methods have some challenges, especially in supervised strategy, one-step deformation estimation is a major advance in this research area.

6- 2- Evaluation Similarity Metrics in Multi-Modal Medical Image Registration

Due to the importance of the similarity metric role in the image registration process, many researchers have worked in this field. In this paper, we reviewed different similarity measure techniques and proposed a new classification based on their applications. This classification can help researchers to choose the suitable metric based on their work.

In this section, we have evaluated the proposed classification of similarity measure techniques in medical applications with four important performance metrics including accuracy, speed, robustness, and complexity. This evaluation is presented in Table II.

Four performance criteria for evaluating the similarity measures are defined as follows:

Accuracy: is a significant metric for registration performance assessment, and refers to the degree of overlap between the reference and sensed images after the matching process. The most used criteria in this field include RMSE (Root Mean Square Error), PNSR (Peak to Noise Signal Ratio), and TRE (Target Registration Error).

Speed of convergence: is related to the number of iterations required to find the optimal transformation parameters.

Robustness: This metric determines the stability of the algorithm against noise, different intensities, or large deformation. Robustness is more critical in some applications, such as multimodal registration.

Complexity: refers to the computational cost needed to complete the registration process. This metric is more important in some algorithms.

Considering the analysis of different groups of similarity metrics in Table 2, the standard measures applied in the traditional framework are easy to understand and implement. However these methods are time-consuming and less accurate when faced with different intensities or large deformations. On the other hand, in learning-based approaches, the similarity metric is learned directly instead of using a fixed similarity measure. These methods, however, are more complex than conventional metrics but obtain results with greater accuracy, efficiency, and speed, and can also be more robust to noise and different intensities.

In this paper, we have studied two prominent topics in the field of medical image registration, i.e. deep learning-based multimodal image registration and similarity measures, and reviewed the latest studies in this research area. Also, in this work, classifications for these methods and their challenges are presented, which can help researchers design more efficient registration strategies for future studies.

7- Conclusion

Image registration is a fundamental issue in medical image analysis. Image registration refers to matching two or more input images in the same coordinate system. Due to the development of imaging technologies in recent years, researchers have conducted many studies in the field of multimodal medical image registration. In this article, an overview of multimodal registration techniques using deep learning is presented and a new classification for these methods is provided. In addition, an important component in intensity-based image registration is the similarity metric, which affects the accuracy and quality of the registered images. This article reviews the various studies conducted on similarity metrics in clinical applications and suggests a classification for different types of these methods. Finally, this paper presents the evaluation and challenges of these methods to help future work to improve the accuracy and efficiency of medical image registration.

References

[1] F.P. Oliveira, J.M.R. Tavares, Medical image registration: a review, Computer methods in biomechanics and biomedical engineering, 17(2) (2014) 73-93.

[2] J. Ma, X. Jiang, A. Fan, J. Jiang, J. Yan, Image matching from handcrafted to deep features: A survey, International Journal of Computer Vision, 129(1) (2021) 23-79.

[3] H. Hermessi, O. Mourali, E. Zagrouba, Multimodal medical image fusion review: Theoretical background and recent advances, Signal Processing, 183 (2021) 108036.

[4] G. Song, J. Han, Y. Zhao, Z. Wang, H. Du, A review on medical image registration as an optimization problem, Current Medical Imaging, 13(3) (2017) 274-283.

[5] Q.R. Razlighi, N. Kehtarnavaz, S. Yousefi, Evaluating similarity measures for brain image registration, Journal of visual communication and image representation, 24(7) (2013) 977-987.

[6] M.A. Viergever, J.A. Maintz, S. Klein, K. Murphy, M. Staring, J.P. Pluim, A survey of medical image registration–under review, in, Elsevier, 2016, pp. 140-144.

[7] M. Wang, P. Li, A review of deformation models in medical image registration, Journal of Medical and Biological Engineering, 39(1) (2019) 1-17.

[8] X. Jiang, J. Ma, G. Xiao, Z. Shao, X. Guo, A review of multimodal image matching: Methods and applications, Information Fusion, 73 (2021) 22-71.

[9] F.E.-Z.A. El-Gamal, M. Elmogy, A. Atwan, Current trends in medical image registration and fusion, Egyptian Informatics Journal, 17(1) (2016) 99-124.

[10] X. Chen, A. Diaz-Pinto, N. Ravikumar, A.F. Frangi, Deep learning in medical image registration, Progress in Biomedical Engineering, 3(1) (2021) 012003.

[11] X. Cao, J. Fan, P. Dong, S. Ahmad, P.-T. Yap, D. Shen, Image registration using machine and deep learning, in: Handbook of medical image computing and computer assisted intervention, Elsevier, 2020, pp. 319-342.

[12] F. Alam, S.U. Rahman, S. Ullah, K. Gulati, Medical image registration in image guided surgery: Issues, challenges and research opportunities, Biocybernetics and Biomedical Engineering, 38(1) (2018) 71-89.

[13] D. Rivas-Villar, Á.S. Hervella, J. Rouco, J. Novo, Color fundus image registration using a learning-based domain-specific landmark detection methodology, Computers in Biology and Medicine, 140 (2022) 105101.

[14] J. Öfverstedt, J. Lindblad, N. Sladoje, Fast and robust symmetric image registration based on distances combining intensity and spatial information, IEEE Transactions on Image Processing, 28(7) (2019) 3584-3597.

[15] Z. Mohammadi, M.R. Keyvanpour, Similarity Measures in Medical Image Registration A Review Article, in: 2021 12th International Conference on Information and Knowledge Technology (IKT), IEEE, 2021, pp. 89-95.

[16] G. Haskins, U. Kruger, P. Yan, Deep learning in medical image registration: a survey, Machine Vision and Applications, 31(1) (2020) 1-18.

[17] M.-R. Keyvanpour, S. Alehojat, Analytical comparison of learning based methods to increase the accuracy and robustness of registration algorithms in medical imaging, (2012).

[18] S. Suganyadevi, V. Seethalakshmi, K. Balasamy, A review on deep learning in medical image analysis, International Journal of Multimedia Information

Retrieval, 11(1) (2022) 19-38.

[19] K. Aghajani, Multi-modal image registration in the presence of spatially varying intensity distortion using structural representation, Multimedia Tools and Applications, 80(25) (2021) 33885-33909.

[20] F. Zhu, X. Zhu, Z. Huang, M. Ding, Q. Li, X. Zhang, Deep learning based data-adaptive descriptor for non-rigid multi-modal medical image registration, Signal Processing, 183 (2021) 108023.

[21] X. Zhu, M. Ding, T. Huang, X. Jin, X. Zhang, PCANet-based structural representation for nonrigid multimodal medical image registration, Sensors, 18(5) (2018) 1477.

[22] E.M. McKenzie, A. Santhanam, D. Ruan, D. O'Connor, M. Cao, K. Sheng, Multimodality image registration in the head-and-neck using a deep learning-derived synthetic CT as a bridge, Medical physics, 47(3) (2020) 1094-1104.

[23] X. Cao, J. Yang, Y. Gao, Y. Guo, G. Wu, D. Shen, Dual-core steered non-rigid registration for multi-modal images via bi-directional image synthesis, Medical image analysis, 41 (2017) 18-31.

[24] X. Cao, J. Yang, Y. Gao, Q. Wang, D. Shen, Region-adaptive deformable registration of CT/MRI pelvic images via learning-based image synthesis, IEEE Transactions on Image Processing, 27(7) (2018) 3500-3512.

[25] X. Liu, D. Jiang, M. Wang, Z. Song, Image synthesis-based multi-modal image registration framework by using deep fully convolutional networks, Medical & Biological Engineering & Computing, 57(5) (2019) 1037-1048.

[26] R. Han, C.K. Jones, J. Lee, P. Wu, P. Vagdargi, A. Uneri, P.A. Helm, M. Luciano, W.S. Anderson, J.H. Siewerdsen, Deformable MR-CT image registration using an unsupervised, dual-channel network for neurosurgical guidance, Medical image analysis, 75 (2022) 102292.

[27] M. Simonovsky, B. Gutiérrez-Becker, D. Mateus, N. Navab, N. Komodakis, A deep metric for multimodal registration, in: International conference on medical image computing and computer-assisted intervention, Springer, 2016, pp. 10-18.

[28] X. Cheng, L. Zhang, Y. Zheng, Deep similarity learning for multimodal medical images, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 6(3) (2018) 248-252.

[29] M. Blendowski, M.P. Heinrich, Combining MRF-based deformable registration and deep binary 3D-CNN descriptors for large lung motion estimation in COPD patients, International journal of computer assisted radiology and surgery, 14(1) (2019) 43-52.

[30] G. Haskins, J. Kruecker, U. Kruger, S. Xu, P.A. Pinto, B.J. Wood, P. Yan, Learning deep similarity metric for 3D MR–TRUS image registration, International journal of computer assisted radiology and surgery, 14(3) (2019) 417-425.

[31] S.K. Zhou, H.N. Le, K. Luu, H.V. Nguyen, N. Ayache, Deep reinforcement learning in medical imaging: A literature review, Medical image analysis, 73 (2021) 102193.

[32] J. Hu, Z. Luo, X. Wang, S. Sun, Y. Yin, K. Cao, Q. Song, S. Lyu, X. Wu, End-to-end multimodal image registration via reinforcement learning, Medical Image Analysis, 68 (2021) 101878.

[33] K. Ma, J. Wang, V. Singh, B. Tamersoy, Y.-J. Chang, A. Wimmer, T. Chen, Multimodal image registration with deep context reinforcement learning, in: International conference on medical image computing and computer-assisted intervention, Springer, 2017, pp. 240-248.

[34] S. Miao, S. Piat, P. Fischer, A. Tuysuzoglu, P. Mewes, T. Mansi, R. Liao, Dilated FCN for multi-agent 2D/3D medical image registration, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2018.

[35] H. Xiao, X. Teng, C. Liu, T. Li, G. Ren, R. Yang, D. Shen, J. Cai, A review of deep learning-based three-dimensional medical image registration methods, Quantitative Imaging in Medicine and Surgery, 11(12) (2021) 4895.

[36] S. Abbasi, M. Tavakoli, H.R. Boveiri, M.A.M. Shirazi, R. Khayami, H. Khorasani, R. Javidan, A. Mehdizadeh, Medical image registration using unsupervised deep neural network: A scoping literature review, Biomedical Signal Processing and Control, 73 (2022) 103444.

[37] S.S.M. Salehi, S. Khan, D. Erdogmus, A. Gholipour, Real-time deep pose estimation with geodesic loss for image-to-template rigid registration, IEEE transactions on medical imaging, 38(2) (2018) 470-481.

[38] X. Wang, L. Mao, X. Huang, M. Xia, Z. Gu, Multimodal MR image registration using weakly supervised constrained affine network, Journal of Modern Optics, 68(13) (2021) 679-688.

[39] J. Fan, X. Cao, P.-T. Yap, D. Shen, BIRNet: Brain image registration using dual-supervised fully convolutional networks, Medical image analysis, 54 (2019) 193-206.

[40] H. Xiao, R. Ni, S. Zhi, W. Li, C. Liu, G. Ren, X. Teng, W. Liu, W. Wang, Y. Zhang, A dual-supervised deformation estimation model (DDEM) for constructing ultra-quality 4D-MRI based on a commercial low-quality 4D-MRI for liver cancer radiation therapy, Medical Physics, 49(5) (2022) 3159-3170.

[41] B.D. De Vos, F.F. Berendsen, M.A. Viergever, H. Sokooti, M. Staring, I. Išgum, A deep learning framework for unsupervised affine and deformable image registration, Medical image analysis, 52 (2019) 128-143.

[42] Y. Ma, D. Niu, J. Zhang, X. Zhao, B. Yang, C. Zhang, Unsupervised deformable image registration network for 3D medical images, Applied Intelligence, 52(1) (2022)

766-779.

[43] S. Li, T. Zhang, D. Zhang, Y. Nie, J. Wang, Metric learning for patch-based 3-d image registration, IEEE Transactions on Automation Science and Engineering, 16(4) (2019) 1575-1583.

[44] M.I. Gandhi, Image registration quality assessment with similarity measures-a research study, in: 2015 International Conference on Communications and Signal Processing (ICCSP), IEEE, 2015, pp. 0084-0088.

[45] Z. Chen, Z. Xu, Q. Gui, X. Yang, Q. Cheng, W. Hou, M. Ding, Self-learning based medical image representation for rigid real-time and multimodal slice-to-volume registration, Information Sciences, 541 (2020) 502-515.

[46] A.A. Goshtasby, Image registration: Principles, tools and methods, Springer Science & Business Media, 2012.

[47] S. Liu, B. Yang, Y. Wang, J. Tian, L. Yin, W. Zheng, 2D/3D multimode medical image registration based on normalized cross-correlation, Applied Sciences, 12(6) (2022) 2828.

[48] M.A. Azam, K.B. Khan, M. Ahmad, M. Mazzara, Multimodal medical image registration and fusion for quality Enhancement, Cmc-Comput. Mater. Contin, 68 (2021) 821-840.

[49] H. Bakhshayesh, S.P. Fitzgibbon, A.S. Janani, T.S. Grummett, K.J. Pope, Detecting synchrony in EEG: A comparative study of functional connectivity measures, Computers in biology and medicine, 105 (2019) 1-15.

[50] D. Sengupta, P. Gupta, A. Biswas, A survey on mutual information based medical image registration algorithms, Neurocomputing, 486 (2022) 174-188.

[51] L. Gong, H. Wang, C. Peng, Y. Dai, M. Ding, Y. Sun, X. Yang, J. Zheng, Non-rigid MR-TRUS image registration for image-guided prostate biopsy using correlation ratio-based mutual information, Biomedical engineering online, 16(1) (2017) 1-21.

[52] H. Rivaz, Z. Karimaghaloo, D.L. Collins, Self-similarity weighted mutual information: a new nonrigid image registration metric, Medical image analysis, 18(2) (2014) 343-358.

[53] Y. Fu, Y. Lei, T. Wang, W.J. Curran, T. Liu, X. Yang, Deep learning in medical image registration: a review, Physics in Medicine & Biology, 65(20) (2020) 20TR01.

[54] R.W. So, A.C. Chung, Multi-modal non-rigid image registration based on similarity and dissimilarity with the prior joint intensity distributions, in: 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2010, pp. 368-371.

[55] S. Jabari, M. Rezaee, F. Fathollahi, Y. Zhang, Multispectral change detection using multivariate Kullback-Leibler distance, ISPRS Journal of Photogrammetry and Remote sensing, 147 (2019) 163-177.

[56] R.W. So, A.C. Chung, A novel learning-based dissimilarity metric for rigid and non-rigid medical image registration by using Bhattacharyya Distances, Pattern Recognition, 62 (2017) 161-174.

[57] X. Chen, X. Wang, K. Zhang, K.-M. Fung, T.C. Thai, K. Moore, R.S. Mannel, H. Liu, B. Zheng, Y. Qiu, Recent advances and clinical applications of deep learning in medical image analysis, Medical Image Analysis, (2022) 102444.

[58] N. Zhu, M. Najafi, B. Han, S. Hancock, D. Hristov, Feasibility of image registration for ultrasound-guided prostate radiotherapy based on similarity measurement by a convolutional neural network, Technology in cancer research & treatment, 18 (2019) 1533033818821964.

[59] M. Blendowski, L. Hansen, M.P. Heinrich, Weakly-supervised learning of multi-modal features for regularised iterative descent in 3D image registration, Medical image analysis, 67 (2021) 101822.

[60] B.D.d. Vos, F.F. Berendsen, M.A. Viergever, M. Staring, I. Išgum, End-to-end unsupervised deformable image registration with a convolutional neural network, in: Deep learning in medical image analysis and multimodal learning for clinical decision support, Springer, 2017, pp. 204-212.

[61] G. Balakrishnan, A. Zhao, M.R. Sabuncu, J. Guttag, A.V. Dalca, VoxelMorph: a learning framework for deformable medical image registration, IEEE transactions on medical imaging, 38(8) (2019) 1788-1800.

[62] Y. Zheng, S. Jiang, Z. Yang, Deformable registration of chest CT images using a 3D convolutional neural network based on unsupervised learning, Journal of Applied Clinical Medical Physics, 22(10) (2021) 22-35.

[63] Y. Lei, Y. Fu, T. Wang, Y. Liu, P. Patel, W.J. Curran, T. Liu, X. Yang, 4D-CT deformable image registration using multiscale unsupervised deep learning, Physics in Medicine & Biology, 65(8) (2020) 085003.

[64] G.A. Benvenuto, M. Colnago, W. Casaca, Unsupervised Deep Learning Network for Deformable Fundus Image Registration, in: ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2022, pp. 1281-1285.

[65] S.K. Zhou, H. Greenspan, C. Davatzikos, J.S. Duncan, B. Van Ginneken, A. Madabhushi, J.L. Prince, D. Rueckert, R.M. Summers, A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises, Proceedings of the IEEE, 109(5) (2021) 820-838.

[66] H. Siebert, L. Hansen, M.P. Heinrich, Learning a Metric for Multimodal Medical Image Registration without Supervision Based on Cycle Constraints, Sensors, 22(3) (2022) 1107.

HOW TO CITE THIS ARTICLE
Z. Mohamadi, M. R. Keyvanpour, Classification and Evaluation of Muti Modal Medical Image Registration Methods and Similarity Measures, AUT J. Model. Simul., 55(1) (2023) 53-70.

DOI: 10.22060/miscj.2023.21791.5303