Lung Computed Tomography Image Registration: A Comparative Study between Intensity Based Method and Deep Learning

Muhammad Zain Amin
Universidad de Girona
Girona, Catalunya, Spain
u1985368@campus.udg.edu

Mohammad Imran Hossain
Universidad de Girona
Girona, Catalunya, Spain
u1984911@campus.udg.edu

ABSTRACT

The process of image registration not only aids in diagnosing directly from the aligned images but also significantly impacts the performance of algorithms supporting medical imaging diagnosis. This project used inspiratory and expiratory Breath-Hold CT image pairs from the COPDgene study repository. Our experimentation focused on various aspects of the registration framework, including the similarity metric, geometric transformation, interpolation, and resolutions. By preprocessing intensities, we reduced variations across images, enhancing comparability and facilitating easier image registration. Utilizing Elastix and Transformix computer software, we performed intensity-based registrations (rigid and non-rigid) on unsegmented and segmented lung structures. By defining the region of interest and eliminating non-essential areas, our approach yielded an average mean TRE (Target Registration Error) of 2.08 ± 2.26 mm, showcasing the effectiveness of our approach and fine-tuned parameters. Furthermore, using VoxelMorph, we obtained an average mean TRE of 2.18 ± 2.74 mm for the same cases.

Keywords: Image Registration · Lung CT · VoxelMorph · Elastix · Transformix

1 Introduction

Image registration is a technique for obtaining point-wise spatial correlation by aligning two or more images [1] [2]. It is considered as an essential stage in medical image analysis and the development of Computer-Aided diagnosis (CAD) systems. Over the last few decades, numerous researchers have actively conducted intensive research in image registration, particularly in the field of medical imaging [3] [4]. These researchers proposed a wide range of methodologies and personalized solutions to handle diverse difficulties.

Image registration is often formulated as an optimization issue that minimizes an appropriate cost function with respect to a spatial mapping and its solutions are estimated through the use of an iterative optimization algorithm. However, solving the optimization problems requires high computational resources and time. The conventional approach, such as intensity-based image registration, is widely used in medical image registration. As intensity-based methods directly utilize the image intensities, they are more flexible and robust than other approaches due to the usage of all the available intensity information from the image. However, it might be a computationally expensive and time-consuming approach for biomedical images due to its large sizes and dimensions. As a result, several studies have been conducted on intensity-based image registration to find out an optimum and faster solution [5].

Elastix, a publicly available computer software for intensity-based image registration, is extensively used for image registration in the biomedical imaging research area due to its faster and outstanding performance. It is compatible with several types of registration rigid (affine) and non-rigid (b-spline) parameters that help to achieve comparative results in image registration tasks [2].

Despite deep learning becoming the preferred method in numerous fields, there have been limited propositions for image registration algorithms based on deep learning [1]. This issue stems from the lack of ground truth and the wide range of possible deformations that can accurately align corresponding anatomical structures. Hence, the issue is notably less supervised compared to tasks like image classification or segmentation, for instance. Despite this, recent years have seen the introduction of several approaches that seek to replicate the functionality

Table 1: The COPDgene Dataset.

Case ID	Dimensions	Voxel Spacing (mm)	Displacement (mm) (mean ± std. dev.)
COPD1	512 × 512 × 121	0.625 × 0.625 × 2.5	25.90 ± 11.57
COPD2	512 × 512 × 102	0.645 × 0.645 × 2.5	21.77 ± 6.46
COPD3	512 × 512 × 126	0.652 × 0.652 × 2.5	12.29 ± 6.39
COPD4	512 × 512 × 126	0.59 × 0.59 × 2.5	30.9 ± 13.49

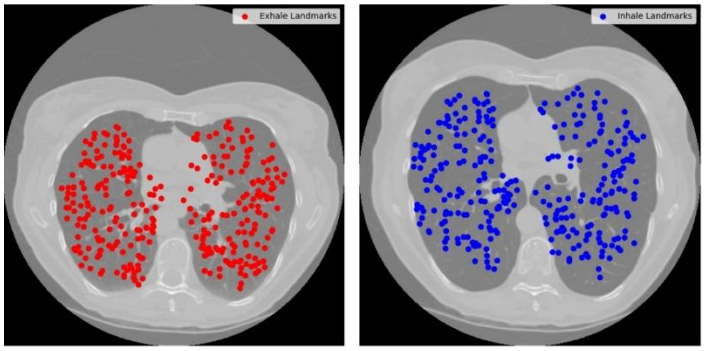

Fig 1: The exhale eBH-CT (left) slice and the inhale iBH-CT (right) slice of COPD1 image with corresponding landmark points.

of traditional image registration methods. These methods involve training a neural network to anticipate the non-linear deformation function when presented with two new unseen images. Given their ability to process images in real-time, trained neural networks hold immense potential for applications that require promptness, such as image guidance in radiotherapy, tracking, or conducting shape analysis using multi-atlas registration.

VoxelMorph stands as a rapid learning-driven framework designed for deformable, pairwise registration of medical images. Conventional registration methods involve optimizing an objective function for each image pair, a process that can become time-consuming with extensive datasets or complex deformation models. VoxelMorph tackles registration by framing it as a function that correlates an input image pair to a deformation field, aligning these images accordingly. This function is parameterized through a Convolutional Neural Network (CNN), with optimization of its parameters carried out across a collection of images [6].

In this project, we performed the lung computed tomography (CT) image registration using both the intensity (*elastix*) and deep learning (VoxelMorph) based approaches, utilizing the COPDgene dataset. Our experiments were conducted with and without image preprocessing to evaluate the impacts of preprocessing on the accuracy and efficacy of these registration techniques. For the intensity-based registration method using *elastix*, we used two both rigid and non-rigid parameters to observe the impact of using different parameters for registration. We computed the Target Registration Error (TRE) to evaluate the registration performance. The following sections of this article will explain the methodology and the analysis of the obtained results of the lung CT image registration project.

2 Materials and Methods

2.1. Dataset

The dataset, COPDgene, used for the lung CT image registration challenge is a publicly available dataset by the National Heart Lung Blood Institute in the United States. It consists of ten distinct cases: labeled COPD1-COPD10, however; we used COPD1-COPD4 images for the registration challenge. For each case, there are two raw CT images for exhale and inhale conditions with their corresponding landmarks [7]. Several properties of the dataset such as image dimensions, voxel-spacing, and displacement between the targets before the registration are mentioned in Table 1.

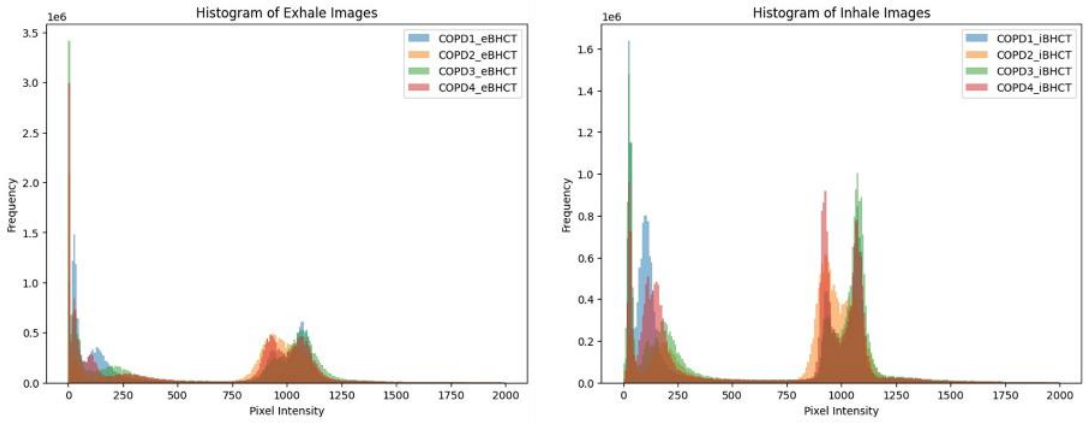

Fig 2: Histogram of voxel intensity values of the eBH-CT (left) images and iBH-CT (right) images of COPD1-COPD4.

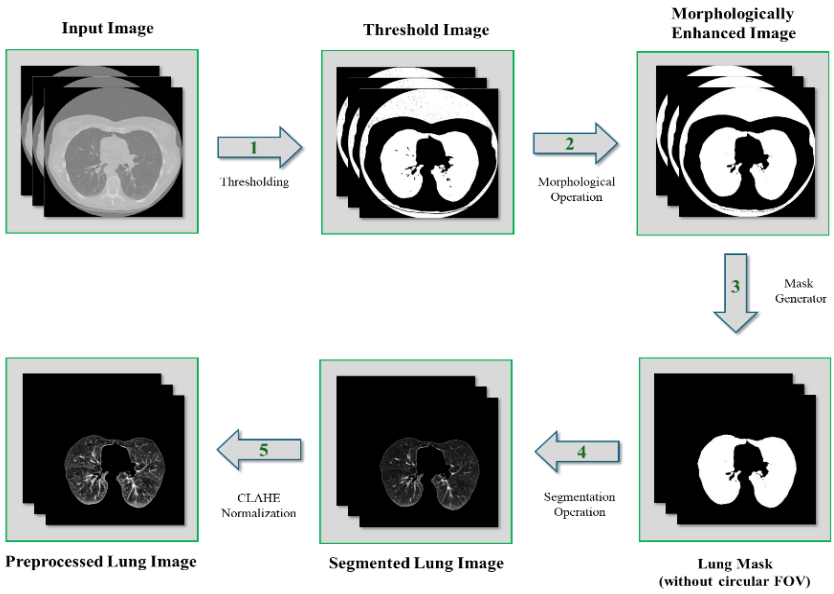

Fig 3: Proposed pipeline for image pre-processing and lung segmentation.

2.2. Pre-processing and Lung Segmentation

The pre-processing stage plays a crucial role in converting raw image files into the nifti format, involving several essential steps in the process. Firstly, this stage involves extracting detailed metadata from the raw images, which holds crucial information like image dimensions, voxel spacing, as well as the mean and standard deviation of displacement. Subsequently, the conversion to the Nifti format is carried out using the SimpleITK library [8], renowned for its effectiveness and adaptability in medical image processing.

Alongside the conversion process, the segmentation of lung images also plays a major role, particularly in image registration tasks. In the context of CT lung imaging, significant variability is observed due to breathing motion. This variability contrasts with the relatively static nature of structures such as the rib cage. Therefore, it is crucial that the alignment process does not disproportionately focus on these static structures at the cost of accurately capturing the lung regions. To tackle this issue, segmentation masks are employed to define Regions of Interest (ROIs) with irregular shapes. These masks are particularly valuable in focusing on the dynamic aspects of the lung, while considering the presence of static anatomical features. The lung segmentation itself is a multi-step process, designed to isolate and accentuate lung regions.

Our proposed lung segmentation pipeline involves a sequence of well-planned steps, each contributing to the precise isolation of the lung areas. These steps are comprehensively illustrated in Figure 3, offering a clear visual guide to the segmentation workflow.

- *Thresholding:* The procedure begins with an analysis of the histogram of voxel intensities found in both exhale (eBH-CT) and inhale (iBH-CT) images of COPD1-COPD4, as illustrated in Figure 2. It was observed that voxel intensities exceeding 800 did not correspond to the actual lung region. Instead, these intensities were associated with the Field of View (FOV) and, thus, deemed irrelevant for lung segmentation. To address this, a thresholding technique was applied with a range of 0-800, enabling the creation of a binary mask that isolates potential lung regions based on their intensity values.
- *Morphological Operations:* Upon the generation of the binary mask through thresholding, the presence of noisy voxels within and around the lung region was noted. To eliminate these noise elements and to further refine and smooth the lung region, a series of morphological operations, specifically closing and opening, were applied to the binary mask.
- *Region of Interest (ROI):* This step involved the extraction of the lung region, or ROI, from the morphologically enhanced binary image. Utilizing the dimensions of the image, midpoints were identified to define specific regions of interest. This step is crucial in creating an accurate lung mask for further detailed analysis.
- *Connected Components:* In this phase, the binary lung mask undergoes labeling to distinguish various connected components or regions within the image. This is an important step in identifying the largest connected components within specified regions, which are instrumental in pinpointing the primary lung areas.
- *Lung Masks and Segmentation:* Finally, lung masks are generated based on the identification of the largest connected components. These masks serve to create a distinct segmentation for the lung regions. An image exclusively containing the segmented lung regions is then produced, utilizing these masks.

Further steps involve normalizing intensities, enhancing contrast via techniques like Contrast-Limited Adaptive Histogram Equalization (CLAHE) [9], and rescaling the segmented lung regions back to their original intensity range. This method facilitates improved visibility and isolation of lung tissues, aiding medical professionals in analyzing and diagnosing various pulmonary conditions or diseases from CT scan data.

2.3. Intensity Based Registration

Registration methodologies that use voxel or pixel intensity values are commonly known as intensity-based [10]. The main idea of intensity-based image registration is to iteratively search for the geometric transformation that optimizes a similarity measure, between the fixed image and the transformed moving image. The similarity measure is related to voxel intensity. The optimizer defines the search strategy. The interpolator aims to resample the voxel intensity from the initial image into the new coordinate system according to the geometric transformation. All these components in image registration form a registration framework.

2.3.1. Similarity Metrics

The widely used similarity measures in image registration are based on intensity differences, intensity cross-correlation, and mutual information. In this section, we describe the similarity metrics we experimented with in our registration framework.

- *Normalized Cross-correlation (NCC):* Normalized cross-correlation (NCC) is a similarity measure that defines the correlation between pixels between images. The NCC between two images NCC *(f_1, f_2)* is at maximum when both are aligned. NCC requires the images to be registered to have a linear relationship between their intensity values.

$$NCC(\mu; I_F, I_M) = \frac{\sum_{x_i \in \Omega_F}(I_F(x_i) - \bar{I_F})\ (I_M\ (T_\mu(x_i)) - \bar{I_M})}{\sqrt{\sum_{x_i \in \Omega_F}(I_F(x_i) - \bar{I_F})^2\ \sum_{x_i \in \Omega_F}(T_\mu(x_i)) - \bar{I_M})^2}} \quad (1)$$

with the average grey-values $\bar{I_F} = \frac{1}{|\Omega_F|}\sum_{x_i \in \Omega_F} I_F(x_i)$ and, $\bar{I_M} = \frac{1}{|\Omega_F|}\sum_{x_i \in \Omega_F} I_M(T_\mu(x_i))$

- *Advanced Mattes Mutual Information (AMMI):* AMMI is based on the mutual information (MI) similarity measure, which measures the mutual dependence between two images. The AMMI measure is an extension of MI that is specifically designed for use in medical image registration. The only requirement of AMMI is that there exists a relationship between the probability distributions of the intensity values of the images being registered, and is, therefore, more general than NCC. It is a similarity measure really powerful for image registration. Mathematically, it can be formulated as:

$$MI(\mu; I_F, I_M) = \sum_{m \in L_M} \sum_{f \in L_F} p\,(f, m; \mu) \log_2 \left(\frac{p(f, m; \mu)}{P_f(f) P_M(m; \mu)}\right) \quad (2)$$

  where L_F and L_M are sets of regularly spaced intensity bin centers, p is the discrete joint probability, and P_f and P_m are the marginal discrete probabilities of the fixed and moving image, obtained by summing p over m and f, respectively. The joint probabilities $p(f, m; \mu)$ are estimated using b-spline parzen windows.

2.3.2. Optimizer

The optimal value of the similarity measure is supposed to correspond to the transformation parameters which best register the moving image to the fixed image. The optimization algorithm aims to search for the maximum or minimum value of the similarity or dissimilarity measure. The registration problem can be mathematically defined with the optimization function:

$$min_T D\left[f_1(x, y), T\left(f_2(x, y)\right)\right] \quad (3)$$

where D is the similarity measure function, f_1 and f_2 are the images or structures to be registered, and T is the transformation. We use Adaptive Stochastic Gradient Descent as an optimizer in our registration framework.

- *Adaptive Stochastic Gradient Descent (ASGD):* Adaptive Stochastic Gradient Descent Adaptive stochastic gradient descent optimization method performs image registration with adaptive step size prediction. The mechanism to adapt the step size γ_k is based on the inner product of the gradient $g\tilde{}_k$ and the previous gradient $g\tilde{}_{k1}$. Intuitively, if the gradients in two consecutive iterations point in (almost) the same direction, it is expected that larger steps can be taken. If the gradients point in opposite directions, the step size is reduced.

2.3.3. Geometric Transformation

The performance of the registration methodology is also dependent on the choice of the geometric transformation, which is linked to the nature of the images to be registered. We use Rigid, Affine, and B-spline geometric transformations.

2.3.4. Interpolation

During the registration process, when the geometric transformations map a point from one space into another, it may be mapped to a non-integer value. The goal of the interpolation step is to estimate the intensity at that new position. We use B-spline interpolator in our registration framework.

2.3.5. Multi-resolution

A multi-resolution registration approach can be defined by using n sub-sampled versions of the original image to aid in the registration process. A lower-resolution sub-sampled image is used in the single-resolution framework to output a transformation matrix to initialize the following single-registration process with a higher-resolution image. The parameters linked with the multi-resolution framework are the Number of Resolutions and the Image Pyramid Schedule.

- *Number of Resolutions*: It is the number of pyramids. Since we expect a large deformation across inhale and exhale volumes, a high number of resolutions would be suitable. However, it comes to finding a trade-off between processing time and registration accuracy. We choose the number of resolutions as six after certain experiments which are described in the next section.

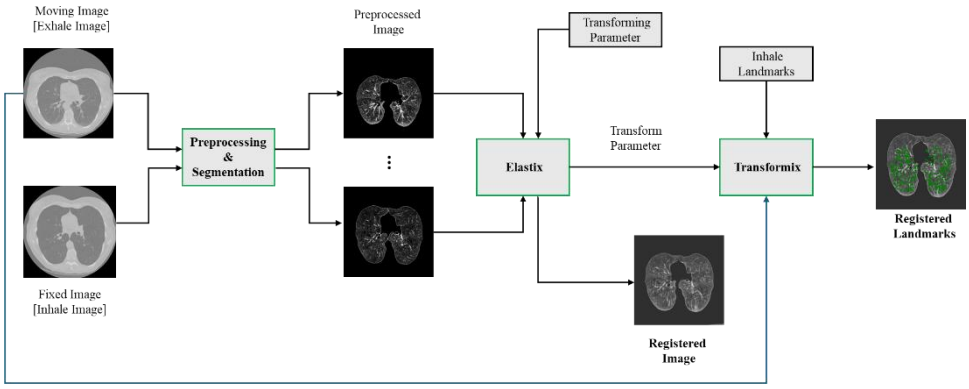

Fig 4: Proposed pipeline for intensity-based registration using *elastix-transformix*.

- *Image Pyramid Schedule:* It defines the down-sampling factor for the image pyramids. By default, the images are down-sampled by a factor of 2 compared to the next resolution. The number of elements of this parameter equals the number of resolutions times the image dimension.

2.3.6. Proposed Pipeline

There are several methods and tools used to perform the image registration. However, *elastix-transformix* based registration has become a popular approach in the medical imaging domain due to its outstanding performance as well as having enough resources such as transformation parameters. Our proposed pipeline of lung CT registration based on *elastix-transformix* consists of several steps shown in Figure 4. The main components of the proposed pipeline are image pre-preprocessing and segmentation, *elastix*, and *transformix*.

- *Elastix:* Elastix offers a platform for deploying intensity-based registration using various transformations. Built on a conventional approach, it seeks the optimal transformation by iteratively enhancing the similarity metric between fixed and moving images. Additionally, it allows for multi-resolution registration using image pyramids and to define many other components like interpolators, optimizers, metrics, iteration counts, and spatial samples, among others. Our proposed pipeline for intensity-based registration starts by pre-processing and segmenting the provided fixed (inhale) image and moving (exhale) image. After the segmentation of the lung region according to the proposed segmentation and preprocessing pipeline shown in Fig 3, both the segmented fixed and moving images were inserted into the *elastix* tool with the registration parameters. After performing the registration, *elastix* provided the registered image and transform parameters which were used for the *transformix* to generate registered landmarks points.

- *Registration Parameter:* In the context of elastix-based image registration, the parameter file is a critical element for successful outcomes. It encodes the specific transformation instructions required to register the moving (exhale) image with the fixed (inhale) image accurately. For chest CT images, particularly for lung registration, the Elastix Model Zoo offers a variety of parameter files [11]. In this project, three parameter files—Par007, Par0011, and Par0056—were identified as potentially suitable. These were selected based on their initial compatibility with the task requirements. A summary of the parameters used in these files is as follows:

  (i) Par007 uses the AdvancedMattesMutualInformation metric with a resolution of 5 levels, and a complex Image Pyramid Schedule starting with 16 and gradually going down to 1. It allows a maximum of 2000 iterations and 10000 spatial samples, with a final grid spacing in physical units for B-Spline starting at 16 and narrowing down to 4.0. The B-Spline grid spacing schedule begins at 16.0 and decreases to 1.0.

  (ii) Par0011 employs AdvancedNormalizedCorrelation with a combination of Affine and B-Spline transformations. It shares the same number of resolutions as Par007 and includes two Image Pyramid Schedules for different stages of the transformation. The maximum number of iterations is divided between the two stages, 1000 for Affine plus B-Spline S1 and 2000 for B-

Table 2: Registration parameters used in our experiments

Parameter	Paramter0007	Parameter0011	Parameter0056
Metric	AdvancedMattesMutualInformation	AdvancedNormalizedCorrelation (Affine + B-SplineS1 + BSplineS2), TransformBendingEnergyPenalty (B-SplineS1 + B-SplineS2)	AdvancedMattesMutualInformation
NumberOfResolutions	5	5	4
ImagePyramidSchedule	16 16 4 8 8 3 4 4 2 2 2 1 1 1 1	16 16 16 8 8 8 4 4 4 2 2 2 1 1 1(Affine + B-SplineS1), 16 16 4 8 8 3 4 4 2 2 2 1 1 1 1 1 (BSplineS1), 4 4 4 3 3 3 2 2 2 1 1 1 1 1 1 1 (B-SplineS2)	8 8 8 4 4 4 2 2 2 1 1 1
MaximumNumberOfIterations	2000	1000(Affine + B-SplineS1), 2000 (B-SplineS2)	2000
NumberOfSpatialSamples	10000	2000	5000
FinalGridSpacingInPhysicalUnits (B-Spline)	16.0 16.0 4.0	10.0 10.0 10.0 (B-SplineS1), 5.0 5.0 5.0 (B-SplineS2)	-
GridSpacingSchedule(B-Spline)	16.0 8.0 8.0 4.0 2.0 1.0	8.0 8.0 4.0 2.0 1.0 (B-SplineS1), 16.0 8.0 4.0 2.0 1.0 (B-SplineS2)	-

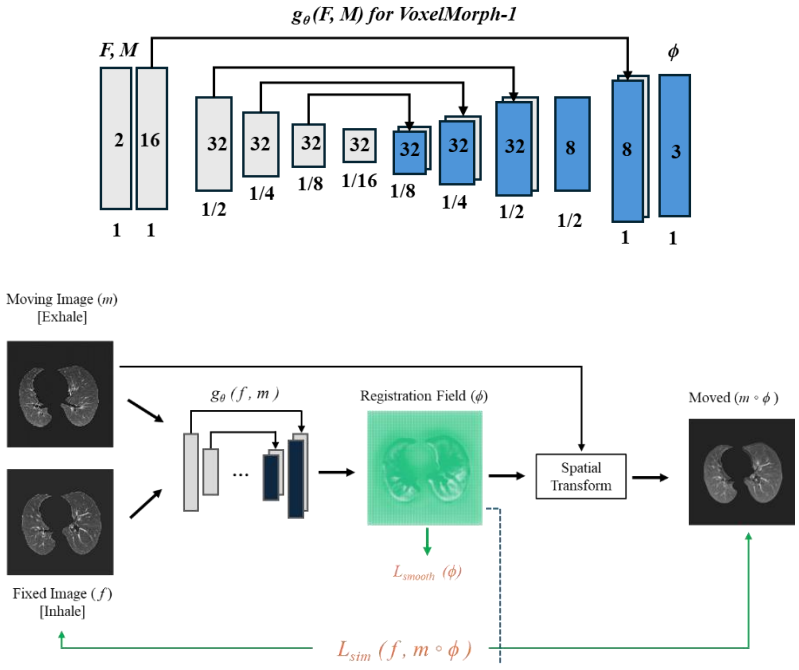

Fig 5: Pipeline for deep learning based registration using VoxelMorph.

Spline S2. The final grid spacing and grid spacing schedule are also divided accordingly, with a coarser adjustment for the Affine stage and a finer adjustment for the B-Spline stage.

(iii) Par0056 relies again on the AdvancedMattesMutualInformation metric but with 4 resolution levels. It has a simpler Image Pyramid Schedule and permits up to 2000 iterations with 5000 spatial samples. The final grid spacing and grid spacing schedule parameters for B-Spline transformations are not specified in this case. These parameters with different combinations were tested to determine which provided the best results for the specific task of lung registration in chest CT images. Table 2 provides a consolidated view of the different configurations explored to optimize the image registration process.

- *Transformix*: In the process of deformable image registration (DIR) using Elastix, landmark point pairs are pivotal for assessing accuracy. After executing the registration task with different parameter files for rigid and non-rigid transformations, the subsequent steps involve transforming the inhale-phase landmarks and comparing them to the exhale-phase landmarks. This transformation utilizes the parameters determined during the Elastix registration, guiding the adjustment of the inhale landmarks to align with the target, which could be an exhale-phase image or another point in a dynamic sequence. Notably, in this process, the moving exhale lung image is registered against the fixed inhale lung image to obtain the transformation parameters. Elastix internally projects the coordinates of the fixed image onto the moving image to ensure precise alignment, indicating that these transformation parameters are specifically designed for applying to the inhale landmarks to achieve accurate alignment with the exhale image.

2.4 Deep Learning Based Registration

Contrary to intensity-based registration methods, recent attention from researchers has been drawn towards learning-based approaches utilizing neural networks. These methods have introduced various strategies for deep learning-based medical image registration. One such approach is VoxelMorph, which operates as a deep learning model without requiring supervised information for registration. Employing a convolutional neural network (CNN), VoxelMorph establishes a registration learning function and utilizes a spatial transform layer to reconstruct one image from another, effectively managing significant displacements within the registration field [6]. Considering a pair of fixed and moving 3D image volumes denoted as F and M respectively, VoxelMorph employs a U-Net-like architecture to model a function $g_\theta(F, M) = \varphi$, where φ represents a registration field and θ denotes the learnable parameters of g. Thus, for each voxel $p \in \omega$, $\varphi(p)$ denotes a location such that $F(p)$ and $M(\varphi(p))$ correspond to identical anatomical landmarks.

The VoxelMorph experiments utilized Keras with TensorFlow as the underlying framework. These experiments were conducted on Google Colab using a V100 GPU with 16GB RAM. Initially, a custom data generator was implemented to process fixed (Inhale) and moving (Exhale) image volumes. This generator produced a concatenated tensor of the two volumes, which served as input for the VoxelMorph architecture. The output of this network was the registration field φ, utilized to register the moving image with the fixed image through a spatial transformer network. This process involved warping and interpolating the moving image to the fixed image space. Training the registration involved a two-part loss function: one part aimed at maximizing a similarity metric between the fixed and moving images, while the other focused on smoothing the registration field to prevent unreasonable displacement. Throughout our work, we experimented with similarity metrics such as Mean Squared Error (MSE) and Normalized Cross-Correlation (NCC). Two rounds of experiments were conducted with VoxelMorph. The first involved applying VoxelMorph directly on the unprocessed image volumes. In the second set, the moving images were first aligned to the fixed ones using Elastix before implementing VoxelMorph's local non-linear registration. For both experiments, we used COPD1-COPD4 datasets. We split the 4 datasets into 75:25 ratio, 3 cases for training and 1 case for testing. To standardize inputs, all sample volumes were resized to (256 × 256 ×128) before being utilized in the VoxelMorph network.

3 Experiments and Results

3.1. Evaluation Metric

To evaluate the efficacy of the developed registration framework, we employed the Target Registration Error (TRE) metric. For each image, 300 specific landmarks were identified, each assigned XYZ voxel coordinates. By leveraging the voxel spacing inherent in each CT image, we computed the 3D Euclidean distance in millimeter (mm) between the registered points during inhalation to exhalation and the corresponding ground truth exhale points. This distance served as a gauge for the accuracy of the registration process, with a lower value indicating better precision. Mathematically, TRE can be expressed as follows:

Table 3: Analysis of Results without Pre-processing and Measurement of Target Registration Error (TRE) in (mm)

Parameter	COPD1	COPD2	COPD3	COPD4	Average
Rigid (Par56)	26.96 ± 10.54	21.79 ± 06.46	12.64 ± 06.38	24.77 ± 11.27	21.54 ± 08.66
Affine (Par11)	25.36 ± 10.26	26.03 ± 05.40	07.56 ± 03.34	23.01 ± 08.52	20.54 ± 06.88
B-Spline-1 (Par07)	**05.52 ± 05.13**	**09.66 ± 06.49**	**03.36 ± 03.31**	**07.16 ± 05.00**	**04.43 ± 04.98**
B-Spline-2 (Par11)	08.17 ± 05.61	11.11 ± 06.50	04.86 ± 03.42	10.60 ± 05.54	08.69 ± 05.27
B-Spline-3 (Par11)	07.43 ± 04.87	11.82 ± 06.26	04.97 ± 03.27	08.55 ± 04.22	08.19 ± 04.65
Affine + B-Spline-1	05.49 ± 05.07	09.67 ± 04.45	03.21 ± 03.11	05.39 ± 04.18	05.94 ± 04.70
Rigid + Affine + B-Spline-1	05.53 ± 05.12	09.66 ± 06.43	03.22 ± 03.11	05.53 ± 04.22	05.98 ± 04.72

Table 4: Analysis of Results with Pre-processing and Measurement of Target Registration Error (TRE) in (mm)

Parameter	COPD1	COPD2	COPD3	COPD4	Average
Rigid	17.49 ± 09.23	18.40 ± 06.38	08.75 ± 03.71	43.45 ± 13.03	22.02 ± 08.10
Affine	11.64 ± 05.13	16.24 ± 05.49	05.93 ± 03.98	10.85 ± 04.95	11.17 ± 04.89
B-Spline-1	01.51 ± 01.84	04.16 ± 04.74	01.21 ± 01.02	01.83 ± 01.56	02.18 ± 02.29
Affine + B-Spline-1	**01.51 ± 01.82**	**03.87 ± 04.75**	**01.26 ± 01.43**	**01.67 ± 01.41**	**02.08 ± 02.26**
Rigid + Affine + B-Spline-1	01.45 ± 01.79	04.14 ± 04.72	01.25 ± 01.04	01.74 ± 01.59	02.15 ± 02.29
Affine + B-Spline-1 + B-Spline-2	01.75 ± 02.44	03.96 ± 05.43	01.29 ± 01.23	12.95 ± 11.31	04.99 ± 05.10

$$TRE = \sqrt{\sum_{i=1}^{n}(x_i * vspacing - \hat{x}_i * vspacing)^2} \quad (4)$$

Where n represents the number of dimensions, x_i signifies the actual position of the target in the i^{th} dimension, and $\hat{x}_i$ denotes the estimated position of the target in the i^{th} dimension and *vspacing* denote the voxel spacing, which might vary across different CT images.

3.1.2. Performance Analysis of Elastix Based Registration

We conducted an extensive experiment to perform image registration using the *elastix* software, focusing on the COPD1 to COPD4 dataset, which includes both inhale and exhale images along with corresponding landmarks. Our experiment was executed in two phases: initially, we carried out the registration using the original inhale and exhale images without any pre-processing. Subsequently, we repeated the registration process using segmented lung images. Additionally, we employed various parameter files from the Elastix Model Zoo, such as Par0007, Par0011, and Par0056, encompassing methods like Rigid, Affine, and B-Spline registration. We used different combinations of parameter files to perform a total of 13 experiments using both the original and segmented images. After performing registration, we computed the evaluation metric, mean and standard deviation of Target Registration Error (TRE) in mm, to evaluate our registration performance for all the experiments shown in Tables 3 and 4.

In the case of registration with the original image, the use of rigid and affine parameter files yielded suboptimal results. This is evident from the Target Registration Error (TRE) values across the COPD1 to COPD4 datasets. The TRE values for the rigid parameter file ranged from 12.64 ± 06.38 to 26.96 ± 10.54, with an average of 21.54 ± 08.66. Similarly, the affine parameter resulted in TRE values varying from 07.56 ± 03.34 to 26.03 ± 05.40, with an average of 20.54 ± 06.88. On the contrary, the use of B-Spline parameters significantly improved the registration accuracy. The B-Spline-1 parameter showed the most promising results with an average TRE of 04.43 ± 04.98, indicating a substantial reduction in registration error. The TRE values for B-Spline-1 ranged from 03.36 ± 03.31 to 09.66 ± 06.49 across the datasets. Similarly, B-Spline-2 and B-Spline-3 parameters also demonstrated improved performance over rigid and affine methods, with average TRE values of 08.69 ± 05.27 and 08.19 ± 04.65, respectively. Moreover, combining affine with B-Spline-1 parameters further reduced the TRE to an average of 05.94 ± 04.70, suggesting that a hybrid approach could be beneficial. A similar trend was

Table 5: Analysis of Results with VoxelMorph and Measurement of Target Registration Error (TRE) in (mm)

Method	COPD1	COPD2	COPD3	COPD4	Average
Preprocessing + VoxelMorph	07.47 ± 02.81	07.00 ± 03.45	03.88 ± 01.64	04.07 ± 01.91	05.60 ± 02.45
Preprocessing + Registration [Affine, B-Spline-1] + VoxelMorph	01.55 ± 01.87	04.21 ± 4.72	01.27 ± 00.99	01.71 ± 01.39	02.18 ± 02.24

Table 6: Analysis of Computational Time.

Image ID	Image Conversion (sec)		Preprocessing and Segmentation (sec)	Elastix Registration (sec)						
	Exhale	Inhale		Rigid	Affine	BSpline-1	BSpline-2	BSpline-3	Affine + BSpline-1	Rigid + Affine + BSpline-1
COPD1	8.89	9.05	21.23	76.76	28.71	549.91	551.20	1218.65	634.49	744.94
COPD2	6.77	7.34	16.28	87.60	58.52	544.51	497.42	1084.94	732.23	428.50
COPD3	8.50	9.73	20.19	83.60	52.42	613.12	552.83	1828.77	846.34	537.65
COPD4	8.53	9.49	20.03	83.90	54.38	604.06	517.75	2552.57	820.60	793.76
TOTAL (min)	0.545	0.594	1.296	5.531	3.234	38.527	35.320	111.416	50.561	41.748

observed with the combination of Rigid, Affine, and B-Spline-1 parameters, which resulted in an average TRE of 05.98 ± 04.72. These results indicate that while rigid and affine registrations were less effective for COPD datasets without pre-processing, the B-Spline methods, especially when combined with affine registration, significantly enhanced the registration accuracy, as reflected by the lower TRE values.

The analysis of registration performance with pre-processed and segmented images, as detailed in Table 4, shows a marked improvement in the accuracy of image registration compared to the results without pre-processing. This improvement is evident in the significantly lower Target Registration Error (TRE) values across various parameters and datasets. For the rigid parameter, the TRE values showed some reduction compared to the non-pre-processed images, ranging from 08.75 ± 03.71 to 43.45 ± 13.03, with an average of 22.02 ± 08.10. Although there was an improvement, the results still indicated relatively higher variability, especially in the COPD4 dataset.

The affine parameter, on the other hand, demonstrated a more notable improvement in performance. The TRE values were considerably reduced, ranging from 05.93 ± 03.98 to 16.24 ± 05.49, with an average of 11.17 ± 04.89. This suggests that affine registration is more effective with pre-processed images. Significant enhancements were observed with the B-Spline-1 parameter, where the average TRE dramatically decreased to 02.18 ± 02.29. The values ranged from a low of 01.21 ± 01.02 to 04.16 ± 04.74, indicating a high level of accuracy in the registration process. Combining affine with B-Spline-1 parameters further optimized the results, achieving an average TRE of 02.08 ± 02.26, closely comparable to B-Spline-1 alone. This combination seems to strike an effective balance between the methods. The combination of Rigid, Affine, and B-Spline-1 parameters also showed excellent results with an average TRE of 02.15 ± 02.29, underscoring the effectiveness of a multi-parameter approach. Lastly, the integration of Affine with both B-Spline-1 and B-Spline-2 parameters resulted in an average TRE of 04.99 ± 05.10. While this combination showed an increase in TRE compared to other combinations, it still represents a significant improvement over the non-pre-processed results. In summary, the pre-processing of images had a profound impact on the efficacy of the registration process. The B-Spline methods, particularly when combined with affine registration, yielded the most accurate results, significantly outperforming the rigid parameter and demonstrating the value of advanced pre-processing techniques in

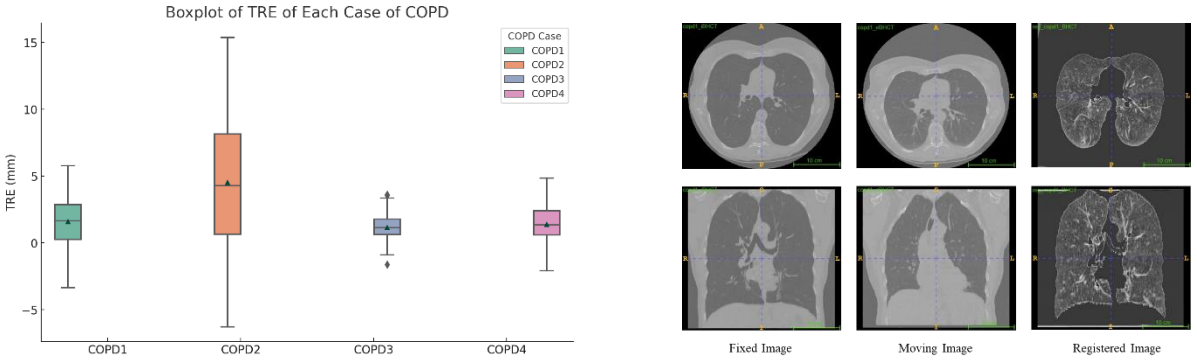

Fig 6: (a) Evaluation metric for the best model (b) Registration output (axial-coronal view) for the best model.

medical image registration. After analyzing the obtained results, it is clear that the performance of registration using segmented images was markedly better than the registration outcomes using the original images.

3.1.3. Performance Analysis of VoxelMorph Based Registration

In addition to the intensity-based registration, we also explored the deep learning approach such as the VoxelMorph model for performing the registration. The analysis is based on two different experimental setups with the VoxelMorph model, and the results are quantitatively evaluated using the Target Registration Error (TRE) as detailed in Table 5.

In the first experiment, the VoxelMorph model was trained using preprocessed segmented lung images, and then it was used to perform image registration. The TRE results for this method varied across the different COPD datasets, with TRE values of 07.47 ± 02.81 for COPD1, 07.00 ± 03.45 for COPD2, 03.88 ± 01.64 for COPD3, and 04.07 ± 01.91 for COPD4, yielding an average TRE of 05.60 ± 02.45. These figures indicate a decent reduction in registration errors, demonstrating the effectiveness of using deep learning for image registration tasks. In the subsequent experiment, a different approach was taken where, instead of using preprocessed images directly, the VoxelMorph model was trained with images that had already been registered using the *elastix* tool with Affine and B-Spline-1 parameters. The aim was to enhance the model's learning from more refined data and thus achieve better registration accuracy. However, the results from this experiment did not show a significant improvement. The mean and standard deviation of TRE values were almost similar to the elastix-based registration with Affine and B-Spline-1 parameters. The results were 01.55 ± 01.87 for COPD1, 04.21 ± 4.72 for COPD2, 01.27 ± 00.99 for COPD3, and 01.71 ± 01.39 for COPD4, resulting in an average TRE of 02.18 ± 02.24. It is clear that the registration performance using the deep learning-based model for example VoxelMorph did not perform outstanding in our case due to less number of training dataset. However, the increase of dataset and appropriate pre-processing techniques may provide better results than our obtained outcomes.

Besides evaluation metric, there are some other important factors such as generalization, robustness of the model, and the processing time that should be considered during the registration process. If a model is not generalized and requires more time to perform the registration, that model would not be considered as a recommended system even though the obtained results are good. For this reason, we observed the computational time during the entire registration process including image preprocessing and performing registration. Table 6 shows the analysis of computational time. From the analysis of computational time, it is evident that the registration using elastix with B-Spline parameters required more time than the registration with other parameters. Finally, the registration pipeline which provided the best result took approximately 45 minutes to complete the entire registration task for all the COPD1-COPD4 images.

4 Discussion and Conclusion

In our comprehensive evaluation of image registration techniques for lung CT images, we compared the traditional intensity-based method against modern deep learning approache. The results demonstrated that the intensity-based method, specifically the elastix-based registration utilizing affine and B-spline transformation parameters with pre-processed segmented lung images, exhibited superior performance. This method achieved remarkable precision, as evidenced by an average mean and standard deviation TRE (Target Registration Error) of 2.08 ± 2.26 mm, which underscores its robustness and reliability. Figure 6 (a) illustrates the statistical robustness of this approach through a box plot, reflecting the consistency and limited variability in TRE across multiple datasets of COPDgen. Furthermore, the visual comparison provided in Figure 6 (b) offers a clear depiction of the registration process, showcasing the fixed (inhale) image, moving (exhale) image, and the resulting aligned image post-registration. The images in both axial and coronal views reinforce the efficacy of the selected registration pipeline. While the deep learning model, voxelMorph, exhibited potential, its performance was hindered by the limited volume of training data. It is reasonable to surmise that with an expanded dataset for training, voxelMorph could achieve or even surpass the benchmark set by the intensity-based method. This underlines the critical importance of data availability in the training and success of deep learning models. Future work should therefore focus on amassing a more extensive and diverse set of training samples, which could enable deep learning methods to fully leverage their capacity for feature extraction and pattern recognition, potentially setting a new standard for image registration tasks in medical imaging.

References

[1] A. Hering, S. Häger, J. Moltz, N. Lessmann, S. Heldmann, and B. van Ginneken, "CNN-based lung CT registration with multiple anatomical constraints," *Med. Image Anal.*, vol. 72, p. 102139, Aug. 2021, doi: 10.1016/j.media.2021.102139.

[2] S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "elastix: a toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging*, vol. 29, no. 1, pp. 196–205, Jan. 2010, doi: 10.1109/TMI.2009.2035616.

[3] J. B. Maintz and M. A. Viergever, "A survey of medical image registration," *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, Mar. 1998, doi: 10.1016/s1361-8415(01)80026-8.

[4] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: a survey," *IEEE Trans. Med. Imaging*, vol. 32, no. 7, pp. 1153–1190, Jul. 2013, doi: 10.1109/TMI.2013.2265603.

[5] A. Myronenko and X. Song, "Intensity-based image registration by minimizing residual complexity," *IEEE Trans. Med. Imaging*, vol. 29, no. 11, pp. 1882–1891, Nov. 2010, doi: 10.1109/TMI.2010.2053043.

[6] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," *IEEE Trans. Med. Imaging*, Feb. 2019, doi: 10.1109/TMI.2019.2897538.

[7] R. Castillo *et al.*, "A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive," *Phys. Med. Biol.*, vol. 58, no. 9, pp. 2861–2877, May 2013, doi: 10.1088/0031-9155/58/9/2861.

[8] B. C. Lowekamp, D. T. Chen, L. Ibáñez, and D. Blezek, "The Design of SimpleITK," *Front. Neuroinformatics*, vol. 7, p. 45, 2013, doi: 10.3389/fninf.2013.00045.

[9] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: speed and effectiveness," in *[1990] Proceedings of the First Conference on Visualization in Biomedical Computing*, May 1990, pp. 337–345. doi: 10.1109/VBC.1990.109340.

[10] M.-E. Lee, S.-H. Kim, and I.-H. Seo, "Intensity-based registration of medical images," in *2009 International Conference on Test and Measurement*, Dec. 2009, pp. 239–242. doi: 10.1109/ICTM.2009.5412952.

[11] "SuperElastix/ElastixModelZoo." SuperElastix, Nov. 23, 2023. Accessed: Jan. 08, 2024. [Online]. Available: https://github.com/SuperElastix/ElastixModelZoo

.