Image Registration of 4D Chest CT Volumes Using DIR-Lab Dataset

Mahdi Islam[1] and Musarrat Tabassum[1]

[1] *University of Girona, Erasmus Mundus Joint Master's Program in Medical Imaging and Applications (MAIA).*

January 11, 2025

Abstract

Medical image registration is critical for aligning images into a common reference frame, enabling comparative analysis, disease monitoring, and treatment planning. This study evaluates traditional intensity-based methods using ITK Elastix and deep learning-based VoxelMorph for chest CT volume registration on the COPDgene dataset. Traditional methods, augmented with preprocessing techniques like segmentation and mask guidance, outperformed deep learning approaches. The controlled registration using the `Par11 Affine + Bspline` configuration achieved the lowest Target Registration Error (TRE) of 2.63 ± 2.77 mm, highlighting its robustness in managing deformations. Automated segmentation with the Lung Mask Library provided comparable accuracy (2.80 ± 3.02 mm) with minimal manual intervention. Conversely, VoxelMorph struggled with large anatomical deformations, yielding a high TRE of 22.22 ± 8.21 mm. These results underscore the effectiveness of traditional methods and the potential for hybrid frameworks that integrate deep learning. Future work should explore domain-specific optimizations, expanded datasets, and advanced training strategies to improve deep learning-based models, paving the way for more accurate and reliable registration in clinical applications.

Keywords—Chest CT Volumes, Affine, TRE, Image Registration, Voxelmorph, Elastix

1 Introduction

Image registration is a fundamental process in medical imaging, facilitating the spatial alignment of two or more images into a shared reference frame. This alignment is critical for various applications, such as tracking changes in patient scans over time, integrating information from different imaging modalities, and mapping individual scans to standardized anatomical atlases. By employing transformations such as rigid, affine, or non-rigid, image registration corrects translational, rotational, and scaling differences between images. In thoracic imaging, this task becomes particularly challenging due to the dynamic nature of thoracic structures, including the diaphragm, lungs, and heart, which undergo positional and shape changes as a result of respiration and other bodily movements. The accurate registration of chest CT volumes is essential in monitoring disease progression, planning and evaluating thoracic treatments, and conducting detailed lung analyses.

Traditional image registration methods, particularly intensity-based approaches, have been extensively studied and widely utilized in medical imaging due to their flexibility and robustness. These methods leverage the full range of intensity information available in the images, making them well-suited for aligning biomedical scans. However, intensity-based registration often involves solving complex optimization problems, which can be computationally expensive and time-consuming, especially when dealing with large image sizes and high-dimensional data. To address these challenges, tools like Elastix, a publicly available software for intensity-based registration, have been developed. Elastix supports both rigid (affine) and non-rigid (B-spline) transformations, offering a versatile framework for achieving accurate and efficient image registration.

In recent years, advancements in deep learning have introduced new possibilities for medical image registration. Unlike traditional methods that optimize a cost function for each image pair, deep learning-based approaches, such as VoxelMorph, train neural networks to predict deformation fields directly. These methods circumvent the need for iterative optimization during testing,

Table 1: **The COPDgene Dataset.**

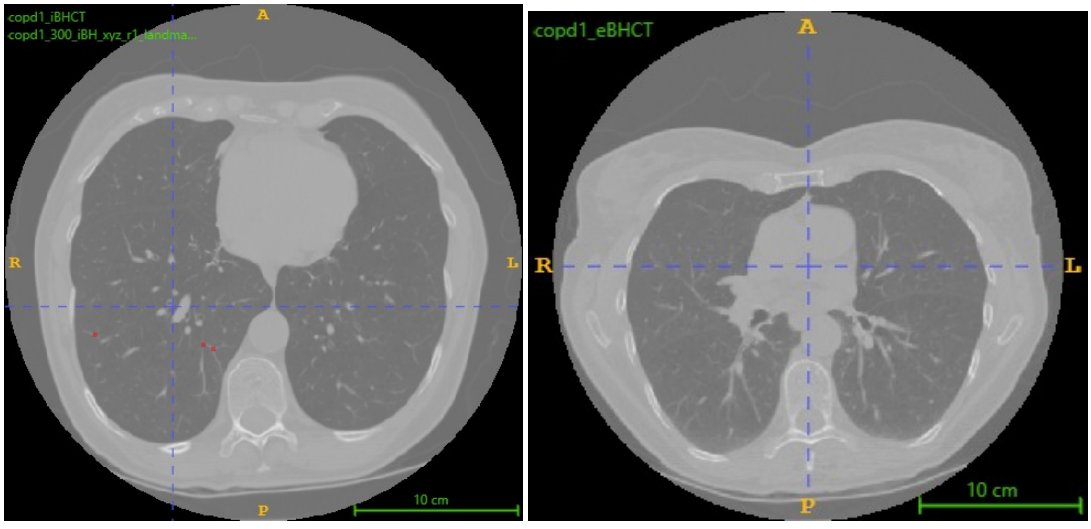
Case ID	Dimensions	Voxel Spacing (mm)	Displacement (mm) (mean ± std. dev.)
COPD1	$512 \times 512 \times 121$	$0.625 \times 0.625 \times 2.5$	25.90 ± 11.57
COPD2	$512 \times 512 \times 102$	$0.645 \times 0.645 \times 2.5$	21.77 ± 6.46
COPD3	$512 \times 512 \times 126$	$0.652 \times 0.652 \times 2.5$	12.29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

enabling real-time registration. However, the development of learning-based registration algorithms faces challenges such as the absence of ground truth and the complexity of capturing diverse anatomical deformations. Despite these limitations, deep learning models like VoxelMorph have shown significant promise in applications requiring rapid and reliable registration, such as image guidance in radiotherapy and multi-atlas shape analysis.

In this study, we performed chest CT image registration using both intensity-based and learning-based methods. The intensity-based approach utilized Elastix with rigid and non-rigid parameter configurations, while the deep learning-based approach employed VoxelMorph, a convolutional neural network framework for deformable registration. To assess the accuracy of the registration methods, the Target Registration Error (TRE) was computed. This paper presents a detailed analysis of the methodologies, results, and implications of our findings, providing insights into the effectiveness of traditional and modern approaches for chest CT image registration.

2 Dataset

The lung CT image registration challenge utilized the publicly available COPDgene dataset, provided by the National Heart, Lung, and Blood Institute in the United States. This dataset can be found in the widely renowned Deformable Image Registration Laboratory website [1]. This dataset comprises ten unique cases, labeled COPD1 through COPD10. For this study, we focused on the COPD1-COPD4 cases. Each case includes two raw CT scans captured during exhalation and inhalation, accompanied by corresponding anatomical landmarks. Key characteristics of the dataset, such as image dimensions, voxel spacing, and the initial displacements between target landmarks prior to registration, are summarized in Table 1. Figure 2 shows the images of the dataset.



(a) Inhale (b) Exhale

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

3 Methodology

3.1 Preprocessing

To enhance the quality of the input chest CT images and improve the accuracy of registration, we employed two primary preprocessing techniques: Normalization and Contrast Limited Adaptive Histogram Equalization (CLAHE).

- **Normalization**: This step was performed to standardize the intensity values of the CT images across all respiratory phases. By rescaling the pixel intensities to a common range, normalization ensures uniformity in the dataset, reducing the influence of varying intensity distributions caused by differences in scanning protocols or respiratory phases. Normalization has been performed following the segmentation. Here, min-max normalization was chosen. The normalization equation used can be found in Equation 1:

$$x_{\text{normalized}} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

- **Contrast Limited Adaptive Histogram Equalization (CLAHE)**: CLAHE is a technique used to enhance the contrast of images by applying histogram equalization to small, localized regions (tiles) of an image rather than the entire image. This approach prevents over-amplification of noise while improving the visibility of structures, particularly in low-contrast regions like the lungs. The enhanced contrast facilitates better alignment of thoracic structures during registration.

3.2 Registration

We conducted an ablation study under two main categories of image registration approaches:

1. Traditional registration using ITK Elastix.
2. Deep learning-based registration using VoxelMorph.

The trials systematically explored the effects of preprocessing, segmentation, and mask usage on registration performance. These are discussed in detail in the following sections.

3.2.1 Elastix-Based Registration

In the context of chest CT volume registration, we explored a range of trials under the umbrella of Elastix-based registration [3]. This approach involved applying the ITK Elastix framework to align fixed and moving images using rigid and non-rigid transformations. For each of the transformations, multiple parameter sets are available. Here, the publicly available parameter set from modelzoo2 has been selected. The parameter files were selected after emperical and rigorous study of all the parameter files in modelzoo. We tested a few other parameter files including BSpline Fine Tune of Parameter 7, Parameter 56 files. But these are not included here as these did not demonstrate promising rseults.

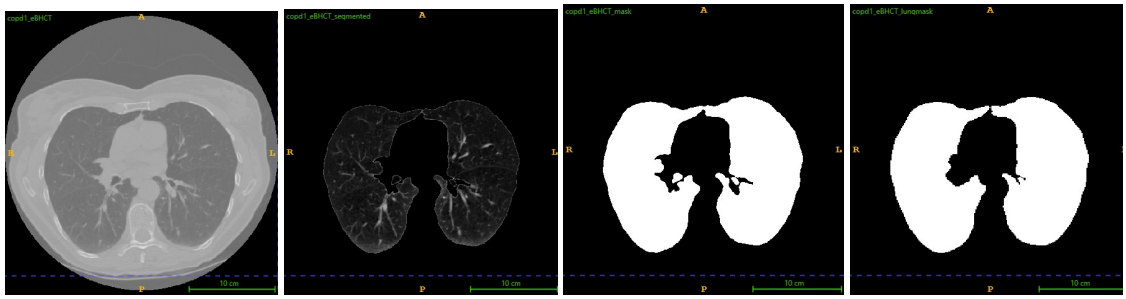
This set of parameters was designed specifically for intra-patient lung CT registration, which is also applicable to this challenge. The registration trials employed three distinct parameter configurations; Parameter 11 BSpline, Parameter 11 Affine followed by Parameter 11 BSpline, and Parameter 7 Coarse, to investigate the impact of varying registration strategies on chest CT volume alignment.

The Parameter 11 BSpline configuration utilized a B-Spline transformation, optimized for non-rigid deformations, making it well-suited for handling intricate anatomical variations in thoracic structures. This setup also included multi-metric optimization with Advanced Normalized Correlation and bending energy penalty for enhanced accuracy. This is especially important for this registration case, as the expected transformation is non-rigid.
The Parameter 11 Affine followed by Parameter 11 BSpline approach combined an initial affine transformation to address global scale, rotation, and translation differences. Equation 2 shows the three-dimensional affine transformation.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & t_x \\ a_{21} & a_{22} & a_{23} & t_y \\ a_{31} & a_{32} & a_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{2}$$

In this equation:

- $\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$ is the original point in 3D space.
- $\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix}$ is the transformed point.
- The 4×4 matrix represents the affine transformation, where a_{ij} elements are responsible for rotation, scaling, and shearing, while t_x, t_y, and t_z are the translation components in the x, y, and z directions, respectively.



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

Figure 2: A visual comparison of the segmentation masks of two approaches; (a) Original Image (b) Segmented Image Using [2] (c) Segmented Mask Using [2] (d) Segmented Mask using Lung Mask Library.

Secondly a B-Spline transformation for finer, localized deformations has been performed. This hybrid configuration aimed to achieve precise alignment across multiple resolutions.
Lastly, the Parameter 7 Coarse setup focused on coarse-grained registration using a multi-resolution pyramid strategy with larger grid spacings, facilitating rapid convergence for larger-scale deformations. These configurations collectively provided insights into the efficacy of traditional and hybrid registration techniques. In this challenge, the test sets only consist of inhale landmarks. Due to the fact that itk-elastix uses the forward registration for point sets rather than the normally used inverse registration for images, the fixed image is chosen as the inhale and the moving image as the exhale. Additionally, it has been tested for use on slices with spacings of up to 2.5 mm, a characteristic shared by all the COPD images.This parameter set has also been used as a starting point to develop custom parameter files for this specific challenge.

The trials were systematically designed to evaluate the impact of preprocessing, segmentation, and mask-guided registration on alignment accuracy. The primary categories of these trials are described below:

1. Registration Using the Original Fixed and Moving Images This trial utilized the original fixed and moving CT images without any preprocessing or segmentation. It served as a baseline for assessing the effectiveness of advanced preprocessing and segmentation techniques in improving registration accuracy.

2. Registration Using the Segmented Fixed and Moving Images To focus on the alignment of specific thoracic structures, such as the lungs, segmentation was applied to isolate relevant regions in the fixed and moving images. This segmentation-driven approach aimed to enhance the precision of the registration process by reducing the influence of irrelevant anatomical features.

3. Registration Using Fixed and Moving Images Along with Fixed and Moving Masks This trial introduced masks corresponding to both fixed and moving images as additional inputs to guide the registration process. The inclusion of masks provided anatomical constraints, facilitating more accurate alignment of key structures and improving overall registration performance.

   The experiments were further extended into a detailed ablation study to systematically analyze the effect of various preprocessing, segmentation, and registration strategies. The ablation studies included:

   - Baseline Registration
   - Baseline Registration with Preprocessing
   - Registration with Segmented Images
     - Using the code found here[2]
   - Controlled Registration with Masks:
     - Using the code found here[2]
     - Using the Lung Mask Library

3.2.2 Deep Learning-Based Registration

Recent advancements in medical image registration have shifted focus from intensity-based methods to learning-based approaches leveraging neural networks. Among these, VoxelMorph [4] has gained significant attention as a deep learning model that performs registration without requiring labeled data. Utilizing a convolutional neural network (CNN), VoxelMorph learns a registration function and employs a spatial transformation layer to align images by reconstructing one image from another. This approach is particularly effective in managing substantial displacements within the registration field [**6**]. Given a pair of fixed and moving 3D image volumes, denoted as F and M, VoxelMorph models the function $g_\theta(F, M) = \phi$, where ϕ represents the registration field and θ refers to the network's learnable parameters. For each voxel $p \in \omega$, $\phi(p)$ indicates a location such that $F(p)$ and $M(\phi(p))$ align anatomically. Figure 3 shows the architecture of VoxelMorph.

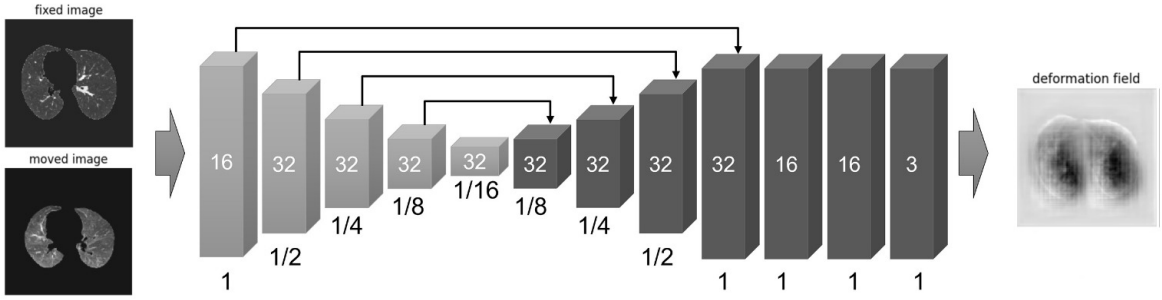


Figure 3: VoxelMorph Architecture

The VoxelMorph experiments were conducted using Keras with TensorFlow as the backend on Google Colab, utilizing a V100 GPU with 16GB RAM. A custom data generator was designed to preprocess the fixed (inhale) and moving (exhale) image volumes, producing concatenated tensors of the two volumes as input to the VoxelMorph network. The network's output, the registration

field ϕ, was then applied to register the moving image to the fixed image using a spatial transformer layer. This involved warping and interpolating the moving image into the fixed image's spatial domain. Training the network employed a two-component loss function: the first component maximized the similarity between the fixed and moving images, while the second ensured smoothness in the registration field to prevent unrealistic displacements. Similarity metrics, including Mean Squared Error (MSE) and Normalized Cross-Correlation (NCC), were explored throughout the study.

Two sets of experiments were carried out using the COPD1-COPD4 datasets. The first applied VoxelMorph directly to unprocessed image volumes, while the second involved pre-aligning the moving images to the fixed ones using Elastix, followed by VoxelMorph's local non-linear registration. The dataset was divided into a 75:25 split, with three cases used for training and one for testing. To standardize the input, all image volumes were resized to $256 \times 256 \times 128$ before being fed into the VoxelMorph network. To compare traditional methods with modern approaches, we implemented the VoxelMorph framework for deep learning-based registration. The model was trained to predict non-rigid deformation fields, with trials conducted on both original and preprocessed data to ensure a fair comparison. Despite its potential, VoxelMorph struggled to handle the large respiratory deformations in the dataset.

4 Results & Discussion

To evaluate the performance of the described methods, the target registration error (TRE) is used. Equation 3 shows the computation.

$$\text{TRE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|\mathbf{p}_i' - T(\mathbf{p}_i)\|^2} \tag{3}$$

- N is the number of target points.
- $\mathbf{p}_i$ is the i-th point in the original space.
- $\mathbf{p}_i'$ is the corresponding point of $\mathbf{p}_i$ in the registered space.
- T is the transformation function applied for registration.
- The notation $\|\cdot\|$ denotes the Euclidean distance.

The Table 2 presents the results of the Target Registration Error (TRE) analysis in detail for different registration strategies applied to COPD1–COPD4 cases. The analysis emphasizes the significant improvements achieved through mask-guided approaches in minimizing TRE. Among the various configurations, the `Controlled Masks` approach using `Par11 Affine + Bspline` achieved the best overall performance with a TRE of **2.63 ± 2.77 mm**, demonstrating its superior precision and robustness. For individual cases, the same configuration achieved the lowest TRE values for COPD2 (**3.76 ± 5.05 mm**) and COPD4 (**1.83 ± 1.51 mm**), highlighting its effectiveness in handling both global and localized deformations.

Furthermore, the use of automated masks from the Lung Mask Library maintained a high level of performance, achieving an overall TRE of **2.80 ± 3.02 mm**, with the lowest TRE for COPD3 (**1.38 ± 0.88 mm**). This underscores the potential of automated segmentation techniques in achieving accurate alignment, even with minimal manual intervention.

In contrast, the deep learning-based `VoxelMorph` approach yielded the highest TRE values, with an overall TRE of **22.22 ± 8.21 mm**. This indicates the challenges faced by neural network-based methods in accurately handling large deformations without domain-specific optimization or additional constraints. Figure 5 shows the box plot of the TREs of the best model for the four data samples.

The results are consistent with the theory which are: Controlled registration with mask provided the best results as it specifically focuses on the lung region using the lung mask to filter out the unnecessary parts. A similar effect happens implicitly in case of segmented images since the registration algorithm focuses on the lung region as the rest are zerod out. For this reason these two approaches were providing the best results.

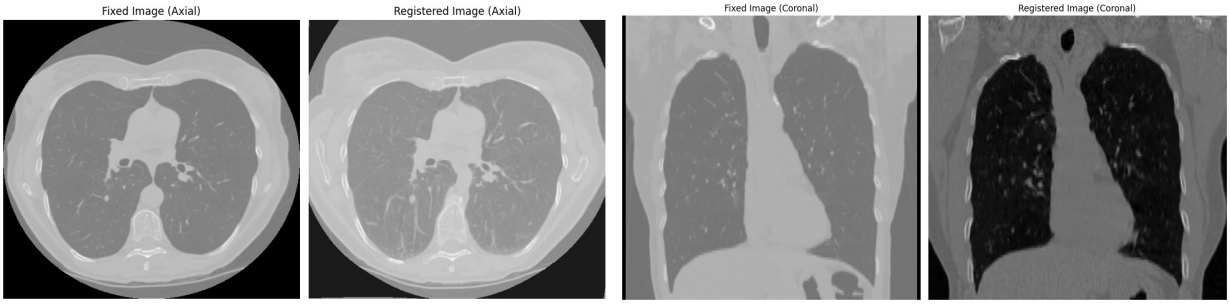
Table 2: TRE Results for Different Registration Experiments. All TRE values are reported as Mean ± Std (in mm).

Experiment	**Parameter Files**	**COPD1**	**COPD2**	**COPD3**	**COPD4**	**Overall**
Baseline	Par11 Bspline	8.23 ± 5.59	14.39 ± 6.90	5.02 ± 3.15	8.61 ± 4.17	9.06 ± 5.45
	Par11 Affine + Bspline	8.11 ± 5.57	14.47 ± 7.07	4.53 ± 2.93	7.44 ± 3.72	8.64 ± 4.82
	Par7 Coarse Bspline	8.72 ± 4.74	13.11 ± 6.26	4.80 ± 2.92	9.77 ± 3.98	9.10 ± 4.97
Baseline + Preproc	Par7 Coarse Bspline	8.69 ± 5.58	11.69 ± 6.24	4.44 ± 2.75	7.81 ± 3.52	8.16 ± 4.52
	Par11 Bspline	7.52 ± 5.97	8.65 ± 6.54	3.93 ± 2.82	6.60 ± 4.34	6.68 ± 4.92
	Par11 Affine + Bspline	7.33 ± 5.69	8.27 ± 6.39	3.72 ± 2.64	5.43 ± 3.76	6.69 ± 4.62
Segmented	Par11 Affine + Bspline	5.39 ± 4.86	4.48 ± 4.78	1.59 ± 1.06	5.68 ± 4.91	4.28 ± 3.90
	Par11 Bspline	6.38 ± 6.46	4.72 ± 4.88	1.66 ± 1.13	8.30 ± 6.74	5.76 ± 4.80
	Par7 Coarse Bspline	6.94 ± 3.92	9.09 ± 5.50	2.40 ± 1.30	6.90 ± 2.90	6.83 ± 3.90
Segmented + Preproc	Par11 Affine + Bspline	5.13 ± 4.86	4.31 ± 4.66	1.51 ± 1.02	10.28 ± 8.27	5.31 ± 4.70
	Par11 Bspline	5.60 ± 5.53	4.41 ± 4.94	1.51 ± 1.04	12.03 ± 9.45	5.89 ± 5.24
	Par7 Coarse Bspline	5.56 ± 3.61	7.24 ± 4.82	2.16 ± 1.17	6.61 ± 3.69	5.89 ± 3.32
Controlled Masks	Par7 Coarse Bspline	4.21 ± 3.33	6.36 ± 5.84	1.89 ± 1.27	4.50 ± 2.88	4.24 ± 3.33
	Par11 Affine + Bspline	**3.54 ± 3.62**	**3.76 ± 5.05**	1.40 ± 0.89	**1.83 ± 1.51**	**2.63 ± 2.77**
Controlled Masks (Lung Mask Library)	Par11 Affine + Bspline	3.98 ± 3.87	3.83 ± 5.20	**1.38 ± 0.88**	2.02 ± 2.14	2.80 ± 3.02
	Par7 Coarse Bspline	5.10 ± 4.04	7.27 ± 4.87	1.71 ± 1.25	13.93 ± 12.74	7.00 ± 5.73
Voxel Morph	-	28.84 ± 9.31	25.60 ± 7.19	10.46 ± 3.91	23.98 ± 12.43	22.22 ± 8.21

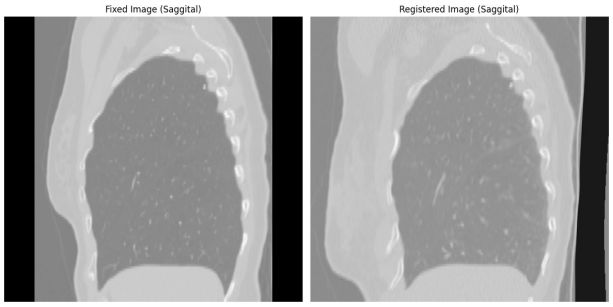
Overall, the results highlight the superiority of traditional intensity-based methods, particularly when guided by precise masks, in achieving accurate and reliable registration for chest CT volumes. Figure 4 shows the comparison between the registered image generated by the best model and the fixed image.

Table 3: Analysis of Computational Time.

Image ID	**Exhale**	**Inhale**	**Affine**	**BSpline-1**	**Affine+BSpline-1**	**Coarse BSpline-2**	**Voxel Morph**
COPD1	7.92	8.14	18.45	81.23	188.24	143.34	912.43
COPD2	6.43	7.21	15.32	72.56	201.43	156.52	836.58
COPD3	8.14	9.22	20.78	85.64	230.68	185.69	905.86
COPD4	8.01	9.18	19.89	83.21	238.51	162.45	935.39
Total (min)	0.513	0.567	1.247	5.382	14.308	10.800	59.821



(a) (b)



(c)

Figure 4: A visual comparison between registered image with the fixed image. (a) Axial (b) Coronal (c) Sagittal

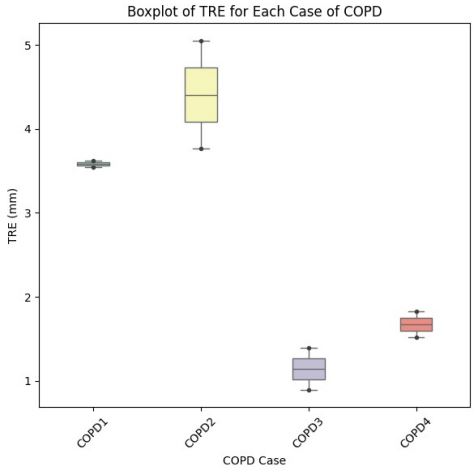


Figure 5: Box Plot

5 Conclusion

This study presents a detailed comparison of registration methodologies for chest CT volumes, emphasizing the strengths and limitations of traditional intensity-based techniques and deep learning-based approaches. The analysis highlights the effectiveness of incorporating preprocessing steps, such as segmentation and mask guidance, to enhance registration accuracy by focusing on relevant anatomical regions. Among the methods evaluated, the controlled registration approach with the `Par11 Affine + Bspline` configuration demonstrated superior performance, achieving the lowest TRE and underscoring the value of targeted preprocessing strategies. While automated segmentation methods, such as the Lung Mask Library, showed potential in streamlining the registration process with minimal manual intervention, the performance of deep learning-based approaches, particularly VoxelMorph, fell short in managing the complex deformations inherent in chest CT volumes. This disparity underscores the need for further development in neural network-based models, focusing on the integration of domain-specific priors, enhanced training datasets, and advanced optimization strategies.

Looking ahead, the integration of traditional intensity-based methods with the flexibility and scalability of deep learning holds promise for advancing registration techniques. Hybrid frameworks that leverage the strengths of both approaches could enable more precise and efficient solutions for complex clinical tasks. Future research should also explore the application of these methods to broader datasets, encompassing diverse anatomical and pathological variations, to validate their robustness and generalizability. In conclusion, this work not only reaffirms the critical role of preprocessing in traditional registration methods but also identifies opportunities for innovation in deep learning-based approaches. These insights provide a foundation for future advancements in medical image registration, aiming for higher accuracy and reliability in clinical practice.

6 Reference

1. R. Castillo, E. Castillo, R. Guerra, V. Johnson, T. McPhail, A. Garg, and T. Guerrero, "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," Physics in Medicine and Biology, vol. 54, no. 7, pp. 1849–1870, 2009, accessed: March 5, 2009.

2. Frederik Hartmann, Yusuf Baran Tanriverdi. "Lung Segmentation: Three-Dimensional Lung Segmentation from CT Scans." Advanced Image Analysis Course Project, taught by Prof. Alessandro Bria. April 2023. Available online at: Github Repository

3. Stefan Klein et al. "elastix: A Toolbox for Intensity-Based Medical Image Registration." In: *IEEE Transactions on Medical Imaging* **29.1** (2010), pp. 196–205. **DOI:** 10.1109/TMI.2009.2035616.

4. G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag and A. V. Dalca, "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," in IEEE Transactions on Medical Imaging, vol. 38, no. 8, pp. 1788-1800, Aug. 2019,**DOI:**: 10.1109/TMI.2019.2897538