

# Medical Image Registration and Applications - Final Project: Lung Registration

Juan Pablo Betancur Rengifo , Jose Joaquin Rodriguez Rojas and Hector Vadillo Valero

## 1 Introduction and Problem Definition

Image registration is a critical technique in medical imaging, enabling the alignment of images acquired under different conditions or from different patients. In the context of lung CT imaging, accurate registration facilitates the analysis of inspiratory and expiratory phases, crucial for studying diseases such as Chronic Obstructive Pulmonary Disease (COPD). Registration allows the transformation of landmarks and intensity volumes to align corresponding anatomical structures between phases, improving the analysis of lung mechanics and pathology progression.

This project focuses on the registration of 3D lung CT images from the 4DCT DIR-Lab Challenge dataset[2]. The primary goal of this challenge is to register inspiratory and expiratory scans using intensity-based methods and evaluate performance through the Target Registration Error (TRE), a metric quantifying the Euclidean distance between transformed landmarks[1].

The objective of this project is to develop a registration pipeline capable of robustly addressing three unseen cases presented during an in-person challenge. To achieve this, several tasks were undertaken, including analyzing and preprocessing intensity-based 3D images, implementing registration methods learned during the Medical Imaging and Registration course, and integrating additional techniques identified through literature review and research. The pipeline was optimized to balance the trade-off between registration accuracy and computational efficiency. Evaluation metrics such as Target Registration Error (TRE), similarity metrics, and computational time were utilized, with a strong emphasis on methodological rigor and practical application. These metrics guided the selection of the most effective methodology to apply to the unseen test dataset, ensuring readiness for the challenge evaluation.

## 2 Methodology

The methodology for this project is structured into distinct phases: preparation, preprocessing (including normalization and mask extraction), registration implementation, and evaluation. Each step aims to ensure the accurate alignment of lung CT volumes, with minimal computational overhead and robust performance across the dataset. This is depicted by figure Figure 1.

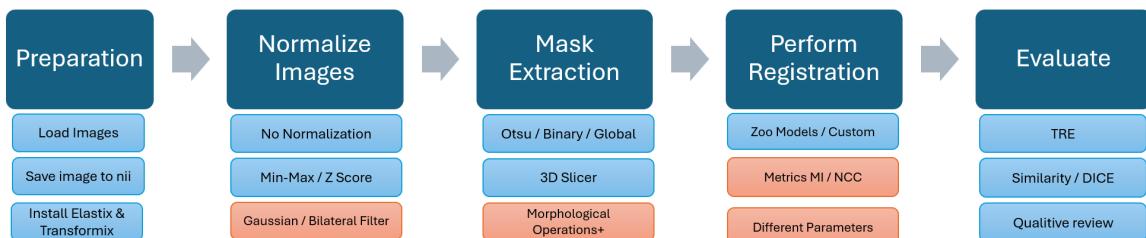


Figure 1: Overview of the image registration process from preparation to evaluation.

### 2.1 Data Preparation

The preparation stage involves loading the CT images and ensuring the computational environment is properly configured. Two methods were used to manage the raw binary image files provided for evaluation:

**1. Using ITK-SNAP:** The framework was used to properly load the raw binary images, and the workflow for this step is shown in Figure 2. The process starts by loading the data as raw binary since the images are in the .img format. Next, it is necessary to configure the image dimensions, voxel spacing, and voxel type based on the information provided for each image. After this, the images are ready to use. However, as shown in the third step of the workflow, the coronal and axial views are flipped. While this does not significantly affect registration, we decided it was more visually practical to work with the Right-Anterior-Superior (RAS) orientation instead of the Right-Anterior-Inferior (RAI) orientation. In the final step of the workflow, ITK-SNAP provides an easy way to change the orientation from RAI to RAS, making the images more intuitive to view. This method would generate .nii.gz files.

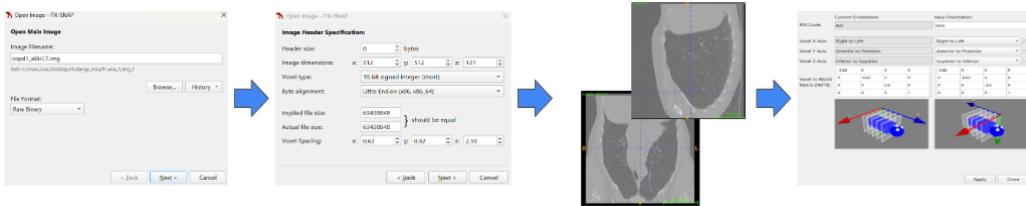


Figure 2: Workflow for loading and configuring raw binary .img files in ITK-SNAP.

**2. Using Python SimpleITK:** The SimpleITK library in Python was utilized to generate .nii files by encoding the raw binary data with the necessary metadata (spacing, dimensions, and data type). This automated method streamlined the preparation process.

This process was applied to the four images provided in the dataset, ensuring consistency and reproducibility. The image specifications are shown in Table 1.

Label	Image Dims	Voxel Size (mm)
COPD1	512 x 512 x 121	0.625 x 0.625 x 2.5
COPD2	512 x 512 x 102	0.645 x 0.645 x 2.5
COPD3	512 x 512 x 126	0.652 x 0.652 x 2.5
COPD4	512 x 512 x 126	0.590 x 0.590 x 2.5

Table 1: Specifications of the 4 images used for evaluation.

## 2.2 Preprocessing

Normalization was applied to standardize intensity values across scans. Different normalization techniques, including Z-Score normalization and Min-Max normalization, were evaluated to identify the most effective method for improving registration performance. As part of the evaluation process, histograms of both respiratory images were plotted to analyze intensity distributions before and after normalization.

Additionally, smoothing techniques were incorporated to assess their impact on registration results between lung inhalation and exhalation images. The filtering methods tested included Gaussian smoothing and bilateral filtering. For each filtering approach, experiments were conducted using different sigma values ranging from 1 to 3. These preprocessing combinations, involving normalization and smoothing, were applied and tested iteratively. Each variation was subsequently evaluated by performing registrations using the best-known parameters available at that stage of the pipeline.

### 2.2.1 Manual Mask Extraction

Since the Elastix framework includes a functionality that enhances registration accuracy by utilizing not only the intensity image but also a mask of the regions of interest, a mask extraction was performed across different views to isolate the lungs, which are the regions of interest in this application. This approach encourages the registration process to focus primarily on these anatomical structures, improving its effectiveness and avoiding irrelevant structures. Thresholding methods, including Otsu's algorithm and global thresholding, were applied, followed by morphological operations such as closing to refine the

masks by filling holes and eliminate artifacts. Figure 3 presents the visual outputs of different approaches tested for generating a binary mask. Image (a) highlights various methods explored: image 1 shows the result of Otsu's thresholding, image 2 demonstrates the outcome of global thresholding, and image 3 depicts a binary mask generated without the application of appropriate morphological operations. In contrast, image (b) illustrates the final optimized binary mask, created using the best morphological parameters, which will be explained in section 3.2. This approach ensures improved segmentation clarity and accuracy while addressing issues such as missing regions (e.g., holes) or oversegmentation of structures like the diaphragm.

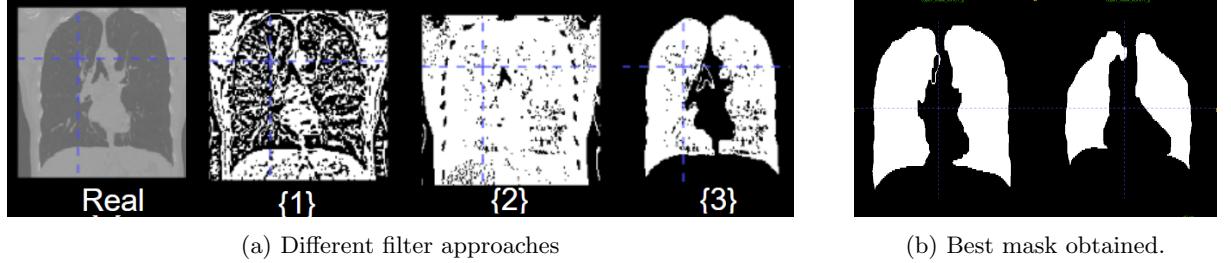


Figure 3: Image (a) contains the different approaches: 1 for otsu's method, 2 for global thresholding method and 3 for binary thresholding. Image (b) shows the best mask obtained using a binary method

### 2.2.2 3D Slicer Mask Extraction

Another approach employed for mask extraction was the use of 3D Slicer, an open-source software platform widely used for medical image analysis and visualization. It provides advanced tools for segmentation and mask extraction, using techniques like thresholding, region growing, or manual editing. Figure 4 shows the workflow for this approach.

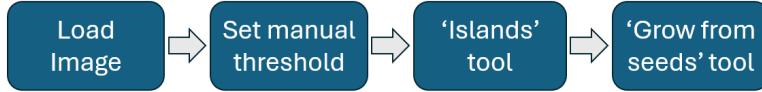


Figure 4: Workflow for 3D Slicer approach

To perform the segmentation process, the Threshold effect, available from the list of tools, was used. Adjusting the lower and upper threshold sliders an attempt was made to capture the intensity range corresponding to the desired structure (in this case the lungs). However, using only a manual threshold did not provide a sufficiently fine segmentation, so it was necessary to refine the mask further using other tools offered by 3D Slicer.

As the initial segmentation included sections outside the lungs, the ‘Islands’ tool was used to remove small segmented regions, leaving only those that meet a specified minimum size. Once the unwanted regions outside the lungs had been removed, the ‘Grow from seeds’ tool was used to further define the region of interest and to fill in any gaps that might remain. Figure 5 shows an example of one of the final masks obtained in 3D Slicer.

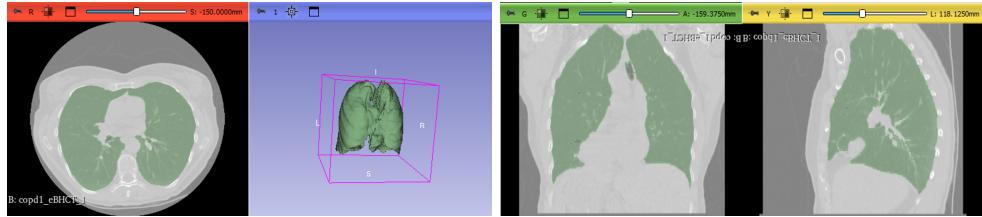


Figure 5: Visualization of mask COPD1 exhalation in 3D Slicer

Figure 6 shows in ITK-SNAP the evolution of the mask after the use of the above mentioned tools, where (1) is the mask with only threshold applied, (2) is the mask after ‘Islands’ and (3) is the mask after ‘Grow from seeds’.

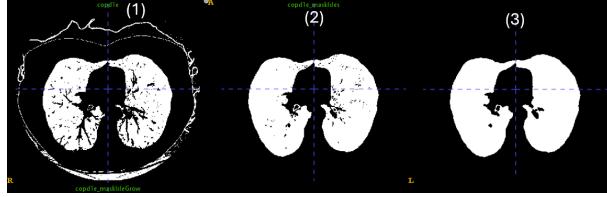


Figure 6: Mask evolution in ITK-SNAP

Although the use of 3D Slicer offers good results, it was discarded and other approaches were used mainly because it is not generalisable to any set of images since the segmentation must be done manually and supervised for each image, modifying parameters when necessary, thus making this a time consuming approach.

### 2.3 Image Registration

The registration process utilized Elastix and transformix for its robust and flexible framework. The process was divided into the following stages:

- Affine Registration:** Initial alignment to correct for global translation, rotation, and scaling discrepancies between inspiratory and expiratory scans.
- Deformable Registration:** Fine-tuning of the alignment using non-rigid transformations to account for anatomical differences.

Parameters for registration were downloaded from the Elastix Model Zoo repository [3], specifically prioritizing models applicable to lung CT registration. Models designed for inhale-to-exhale transformations were selected after reviewing the repository, as shown in Figure 7. Additionally, a default BSpline parameter file was modified to create a custom configuration for this project.

Par0054 - elastix	Par0077 - elastix	Par0015 - elastix	Par0011 - elastix
Registration Description intra-subject respiratory motion; B-spline transformation	Registration Description expiratory & systole transformation; mutual information	Registration Description intrapatient; B-spline transformation; several similarity metrics	Registration Description intrapatient (sometimes intra-sleep); B-spline transformation; normalized correlation
Image data • 3D thoracic CT (inhale/exhale phases from 4DCT) • Voxel size 1 x 1 x 3 mm • Dimension 512 x 512 x 114	Image data • 3D chest CT • Lung • Peak inhalate and peak exhalate • Voxel size resampled to 1.0 x 1.0 x 1.0 mm • Dimension 304 x 304 x 300 • Format DICOM or Analyze format Screen shot:	Image data • 3D chest CT • Lung • 21 scans • Voxel size mostly 0.7x0.7x0.5 mm. • Dimension on average about 400 x 300 x 350	Image data • 3D chest CT • Lung • 20 scans • Voxel size mostly 0.7x0.7x0.7 mm, four scans with thicker slices 1,2,5,2,2.5 mm. • Dimension on average about 400 x 300 x 350
The sample CT dataset is available for <a href="#">download</a> .			
Par0016 - elastix	Par0003 - elastix	Par0008 - elastix	
Registration Description intrapatient; Multi B-spline transformation; sliding motion	Registration Description intrapatient; after	Registration Description intrapatient; B-spline transformation; mutual information	
Image data • 3D chest CT • Lung • 16 scans (10 dir-lab, 6 extended popi-model) • Voxel size mostly 0.976x0.976x5 mm for dir-lab patients // 1x1x2 for extended popi-model. • Dimension on average about 256 x 256 x 100 for 5 first dir-lab patients // 512 x 512 x 165 for extended popi-model	Image data • 3D chest CT • Lung • Full inspiration, without contrast agent, follow-up scans • Voxel size resampled to 0.7x0.7x0.7 mm to 0.8x0.8x0.7 mm • Dimension 512 x 512 x 256 - 529 • Acquired with a Philips 16-detector row CT scanner (MD4000 IDT or Brilliance 16FD)	Image data • 3D chest CT • Lung • Breath-hold inspiration scans • Voxel size resampled to 0.7x0.7x0.7 mm • Dimension 512 x 512 x ~500 (sub-sampled to 256 x 256 x ~200 such that voxels were isotropic)	

Figure 7: Extracted models from Elastix Model Zoo used in the project.

After performing several iterations with the different parameters shown in Figure 7 for registration steps 1 and 2, it was observed that the training images were already in similar spatial coordinates. This indicated that affine registration was more of an optional step rather than a mandatory process. This observation was further supported when using parameters that applied a BSpline approach, which combines rigid and non-rigid registration steps. This method proved to be robust to slight variations in spatial coordinates.

During the development of our registration pipeline, we encountered challenges related to the integration of both intensity images and their corresponding landmarks. Initially, our approach focused on transforming a moving intensity image to the space of a fixed intensity image using elastix [8]. While this method is effective for intensity-based registration, the inclusion of landmark points as our primary target introduced additional complexity. Upon consulting the documentation of elastix and SimpleElastix, we

gained crucial insights into how the software generates an inverse transform file rather than a direct 1-to-1 transformation parameter. This realization, as supported in Figure 9, made us reconsider our problem formulation.



Figure 8: Initial approach to the challenges.

Consequently, we redefined our approach: we set the inspiratory data (both image and landmarks) as the fixed reference, while designating the expiratory image as the moving data. This configuration aligns with the inverse nature of the transformation parameters produced by elastix. Essentially, the resulting parameters, which are apply using transformix, describe the transformation necessary to convert inspiratory points into their corresponding expiratory positions. This strategy is particularly advantageous for our challenge scenario, where we will only have access to inspiratory landmarks on the challenge day. Using inspiratory data as our fixed reference ensures the registration pipeline remains valid and applicable in the absence of expiratory landmarks. Figure 10 provides a visual representation of this refined pipeline, illustrating how the inverse transformation allows us to effectively map inspiratory landmarks to their estimated expiratory positions, even without explicit expiratory landmark data during the registration process.

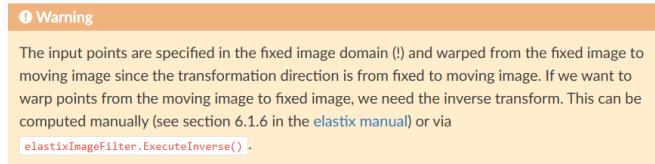


Figure 9: Insights on which direction the registration had to be done



Figure 10: Registration pipeline showing transformation from inspiratory to expiratory phase.

## 2.4 Evaluation

### 2.4.1 TRE

Evaluation of the registration process involved computing the Target Registration Error (TRE), defined as the Euclidean distance between corresponding transformed landmarks in the fixed (inspiratory) and moving (expiratory) scans. This metric provides a quantitative assessment of the alignment accuracy (1).

$$\text{TRE} = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_i - \mathbf{T}(\mathbf{x}'_i)\|_2 \quad (1)$$

Where:

- $\mathbf{x}_i$ : Coordinate of the point  $i$  in the fixed image (o ground truth).
- $\mathbf{x}'_i$ : Coordinate of the point  $i$  in the moving image before the register.

- $\mathbf{T}(\mathbf{x}'_i)$ : Coordinate transformed point  $i$  after applying the transformation  $\mathbf{T}$ .
- $N$ : Total number of points.

#### 2.4.2 Intensity based similarity metrics

In addition to TRE, similarity metrics such as Mutual Information (MI) (2), which measures the statistical dependence between the intensities of two images, and Normalized Cross-Correlation (NCC) (3), which measures the linear similarity between two images (intensities), were computed to evaluate the intensity alignment, as well as time consumption. Qualitative assessments, including visual inspection of overlayed scans and mask alignment, were also performed to validate the registration outcomes.

The formulas for MI and NCC are shown below:

$$MI(A, B) = \sum_{a \in A} \sum_{b \in B} P(a, b) \log \left( \frac{P(a, b)}{P(a)P(b)} \right) \quad (2)$$

Where:

- $P(a, b)$ : Combined probability of intensities  $a$  in image  $A$  and  $b$  in image  $B$ .
- $P(a)$ : Marginal probability of intensity  $a$  in image  $A$ .
- $P(b)$ : Marginal probability of intensity  $b$  in image  $B$ .

$$NCC(A, B) = \frac{\sum_i (A_i - \mu_A)(B_i - \mu_B)}{\sqrt{\sum_i (A_i - \mu_A)^2 \sum_i (B_i - \mu_B)^2}} \quad (3)$$

Where:

- $A_i, B_i$ : Corresponding pixel intensities in the images  $A$  and  $B$ .
- $\mu_A, \mu_B$ : Average intensities in the images  $A$  and  $B$ .

#### 2.4.3 Dice mask similarity

After registration, the DICE score (4), a metric that quantifies the similarity between the predicted segmentation and the ground truth, was employed. This was achieved by applying the transformation parameters obtained from the registration process to the mask of the inhalation moment. Since the exhalation (Ground truth) mask has been computed, evaluating the DICE score of the registered inhalation mask ("segmentation") provides additional insight into the accuracy of the process.

$$DICE = \frac{2 \cdot |A \cap B|}{|A| + |B|} \quad (4)$$

Where:

- $A$ : Reference target mask (*ground truth*).  $B$ : Registered mask to target.
- $|A \cap B|$ : It is the number of elements at the intersection between  $A$  and  $B$ .
- $|A|$  and  $|B|$ : These are the numbers of elements in  $A$  and  $B$ , respectively.

#### 2.4.4 Qualitative Evaluation

The qualitative evaluation focused on visually assessing the quality of the registration results. This included reviewing the resulting images from the registration process and observing how well the masks were mapped from inhalation to exhalation, including checkerboard visualization.

Additionally, a method was implemented to plot the corresponding points from the landmark files, comparing the original inhalation to exhalation transformation, as well as the registered exhalation points against the ground truth exhalation points. This approach provided a visual means to assess the registration accuracy and alignment qualitatively, highlighting any discrepancies and confirming the alignment of anatomical structures between the transformed images.

#### 2.4.5 Time

The evaluation pipeline included validation of TRE calculations using fixed landmark points. Execution time for each result was recorded to assess the computational performance of different methods. This additional metric was crucial in selecting the most suitable approach for the challenge day. The included code provides the timing implementation for further elaboration if necessary.

```
import time
start_time = time.time()
elapsed_time = time.time() - start_time
```

## 3 Experimental Section and Result Analysis

This section details the experimental execution based on the methodology explained above with its results, including TRE values, computational time, and visual inspections of registered volumes. Comparative analyses of different parameter settings and methods are included here.

### 3.1 Computing the most accurate masks

Before conducting the experiments described below, the optimal parameters were selected to generate the best possible masks. This step was crucial as it significantly enhanced the algorithm's performance, leading to improved results. To produce masks such as the one shown in 11, a binary thresholding method was applied with a lower threshold of 100 and an upper threshold of 500. These values were determined by analyzing the intensity of the images and through several iterations. The goal was to make this step as robust as possible for all the training images.

Following this step, it was necessary to determine the number of connected components to retain. In some cases, the segmented lungs appeared as a single connected component, while in others, it was necessary to extract the two largest connected components. This decision was made through a trial-and-error approach.

Subsequently, a binary closing operation with a structuring element of dimensions  $7 \times 7$  was applied to the mask in order to fill small gaps that might have been present. Finally, the segmented images were saved as .nii files.

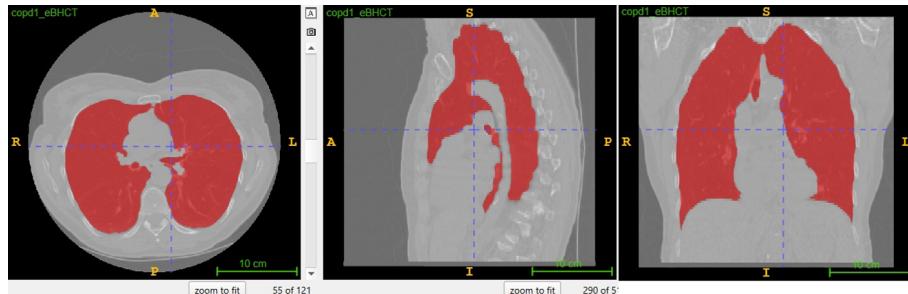


Figure 11: Example of the best overlayed mask for image COPD1

### 3.2 Parameter selection

Prior to computing metrics and transformation parameters, a subset of relevant parameters was selected based on information from the Elastix Model Zoo GitHub repository. These parameters were compared against a custom B-spline parameter file (A5) using Target Registration Error (TRE) as the primary metric. Table 2 summarizes the results for parameters 003, 007, 016, 054, and the custom B-spline.

The custom B-spline parameter file (A5) achieved the best overall registration accuracy, with the lowest average TRE of 1.433 and the highest average DSC of 0.766, while maintaining a processing time of approximately 2-3 minutes. Parameter 007 (A2) was originally designed for regional pulmonary function comparisons and yielded an average TRE of 22.587 and DSC of 0.616, with significantly higher computational times averaging 20.328 seconds. Parameter 054 (A4), designed for intra-patient thoracic CT scans addressing respiratory motion, demonstrated lower performance with an average TRE of 19.228 and DSC of 0.524.

The custom parameter (A5) exhibited consistent accuracy across all patients, with particularly strong results for patients 1, 2, and 3. Notably, patient 4 showed a higher TRE of 1.2907 compared to the other patients, which highlights a potential edge case. Despite this, the custom parameter consistently outperformed both standard Zoo Model parameters and default configurations in accuracy and processing efficiency, as shown in Table 2. This indicates that fine-tuning a parameter file for specific applications can significantly enhance registration outcomes.

Parameter	I1 TRE	I1 Time	I1 DSC	I2 TRE	I2 Time	I2 DSC	I3 TRE	I3 Time	I3 DSC	I4 TRE	I4 Time	I4 DSC	Avg TRE	Avg Time	Avg DSC
A1:Par_003	26.2406	16.75	0.639	21.4121	14.07	0.525	12.1174	17.83	0.547	29.4405	16.55	0.533	22.303	16.300	0.561
A2:PAR007 MI FINE BSpline	26.3342	20.06	0.632	21.786	18.13	0.632	12.6392	22.16	0.728	29.5868	20.96	0.471	22.587	20.328	0.616
A3:Par007MI.RPBsplinetunedM	1.0257	183.24	0.825	<b>2.2581</b>	199.35	0.701	1.1498	200.26	0.792	33.4504	201.19	0.337	9.471	196.010	0.664
A3:Par0016.multibsplines	1.4025	166.72	0.64	3.4543	178.17	0.502	1.2812	184.67	0.525	2.2778	177.68	0.536	2.104	176.810	0.551
A4:Par0054_ssstd	24.158	86.99	0.608	20.066	72.28	0.49	5.417	88.97	0.511	27.2696	82.86	0.488	19.228	82.775	0.524
A5:Custom BSpline Param	<b>1.0009</b>	182.05	0.825	2.3355	142.31	0.7	<b>1.1053</b>	168.11	0.793	<b>1.2907</b>	157.31	0.746	<b>1.433</b>	162.445	0.766

Table 2: Comparison of Registration Parameters ZooModel and Custom

### 3.3 Intensity Normalization

The histogram analysis, presented in Figure 12, highlights the impact of different preprocessing methods on the intensity distributions of the inhalation and exhalation images. These histograms provide insights into the effectiveness of normalization and filtering techniques in aligning intensity ranges, an essential step for successful image registration.

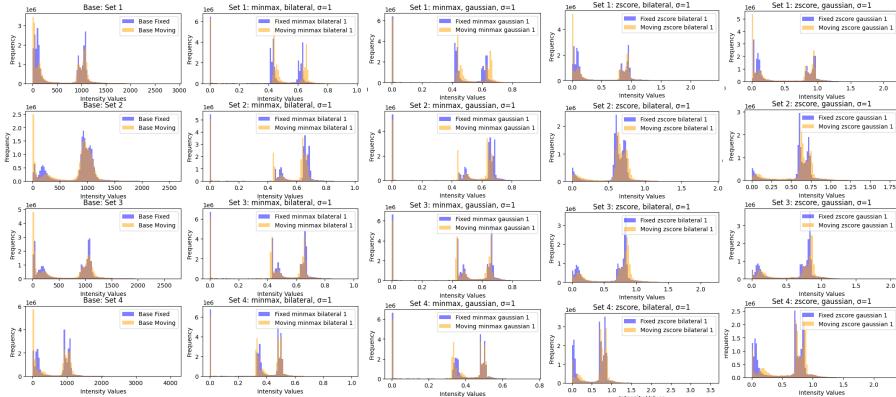


Figure 12: Histogram comparison base against combination of pre-processing

The registration results of the tested preprocessing combinations are summarized in Table 2. The best average TRE of 1.386 was achieved with the combination of Z-Score normalization and bilateral filtering with a sigma of 1 (B4), accompanied by an acceptable average DSC of 0.769. However, this configuration required the highest computational time (1327.345 seconds), indicating a trade-off between accuracy and efficiency.

Table 3: Comparison of Preprocessing Methods

Parameter	I1 TRE	I1 Time	I1 DSC	I2 TRE	I2 Time	I2 DSC	I3 TRE	I3 Time	I3 DSC	I4 TRE	I4 Time	I4 DSC	Avg TRE	Avg Time	Avg DSC
A5:No Normalization	1.0009	182.05	0.825	2.3355	142.31	0.7	<b>1.1053</b>	168.11	0.793	1.2907	157.31	0.746	1.433	162.445	0.766
B1: ZScore+Gaussian1	1.2108	149.93	0.829	2.4538	306.91	0.719	1.2809	329.38	0.793	1.697	289.53	0.75	1.661	268.938	0.774
B2: ZScore+Gaussian2	2.147	185.21	0.831	4.1915	157.03	0.716	1.9785	281.77	0.799	3.2202	245.96	0.746	2.884	217.493	0.773
B3: ZScore+Gaussian3	3.903	140.83	0.825	7.1658	140.62	0.685	3.1408	149.46	0.789	5.6077	148.78	0.734	4.954	144.923	0.758
B4: ZScore+Bilateral1	<b>0.9899</b>	151.31	0.826	2.178	253.82	0.703	1.1091	4693.27	0.796	<b>1.2667</b>	210.98	0.75	<b>1.386</b>	1327.345	0.769
B5: ZScore+Bilateral2	0.9717	212.04	0.826	2.1904	142.99	0.707	1.1266	239.45	0.796	1.2867	227.05	0.75	1.394	205.383	0.770
B6: ZScore+Bilateral3	0.9804	200.24	0.827	<b>2.1672</b>	213.36	0.707	1.1293	242.56	0.797	1.3102	229.87	0.75	1.397	221.508	0.770
B7: MinMax+Gaussian1	1.2016	219.96	0.829	2.4774	231.25	0.719	1.2873	227.71	0.799	1.6555	225.61	0.75	1.655	226.133	0.774
B8: MinMax+Gaussian2	2.1352	213.07	0.831	4.0602	231.27	0.716	1.9785	236.77	0.799	3.231	222.51	0.746	2.851	225.905	0.773
B9: MinMax+Gaussian3	3.8794	216.24	0.825	7.1715	226.18	0.685	3.1421	236.6	0.789	5.6222	222.94	0.733	4.954	225.490	0.758

In contrast, methods involving Min-Max normalization with Gaussian smoothing (B7, B8, B9) generally resulted in higher TRE values, reflecting lower registration accuracy. The absence of normalization (A5) yielded a moderate average TRE of 1.433 but lacked improvements in DSC or computational efficiency seen with other approaches. Overall, bilateral filtering demonstrated consistent performance across varying sigma values, outperforming Gaussian smoothing in terms of registration accuracy.

The variations presented in the tables correspond to parameter changes in the Elastix parameter file. These variations include adjustments to the number of iterations (*iter*), resolution levels (*res*), voxel size (*vox*), and similarity metrics (*MI* for AdvancedMattes Mutual Information and *NCC* for Normalized Cross Correlation). These changes were aimed at exploring the impact of parameter tuning on registration performance.

From Table 4, it can be observed that the lowest average TRE of 1.3553 mm was achieved using parameter file *C2:Iter20k, res 4, vox 16 MI*, albeit with the longest average execution time of 408.05 seconds. This highlights the trade-off between computational time and registration accuracy when increasing the number of iterations. Conversely, the shortest average execution time (145.34 seconds) was observed with *A5:Iter10k, res 4, vox 16 MI*, which also delivered a competitive average TRE of 1.3859 mm, making it an efficient choice in time-constrained scenarios.

Table 4: BSpline Custom Parameter Tuning Performance Evaluation

Parameter File	I1 TRE	I1 Time	I1 DSC	I1 MI	I2 TRE	I2 Time	I2 DSC	I2 MI	I3 TRE	I3 Time	I3 DSC	I3 MI	I4 TRE	I4 Time	I4 DSC	I4 MI
A5:Iter10k, res 4, vox 16 MI	0.9899	145.23	0.826	0.3144	2.178	154.53	0.703	0.3057	1.1091	145.71	0.796	0.4007	1.2667	135.9	0.75	0.2748
C1:Iter10k, res 4, vox 16 NCC	1.8648	140.75	0.841	0.3425	2.9004	146.05	0.743	0.355	1.1789	163.85	0.804	0.4154	1.2964	210.48	0.761	0.2847
C2:Iter20k, res 4, vox 16 MI	0.9648	408.26	0.826	0.3139	<b>2.1134</b>	418.02	0.709	0.3051	1.1321	454.24	0.796	0.3953	<b>1.2109</b>	351.67	0.75	0.2713
C3:Iter10k, res 5, vox 16 MI	1.0165	151.27	0.825	0.2695	2.1808	164.24	0.708	0.2781	1.1405	218.27	0.796	0.3794	1.2511	264.15	0.75	0.2493
C4:Iter10k, res 4, vox 8 MI	<b>0.9589</b>	358.78	0.836	0.3887	3.0758	462.24	0.722	0.3995	<b>0.9607</b>	426.21	0.799	0.4686	20.7274	345.17	0.54	0.3526
C5:Iter10k, res 3, vox 16 MI	1.3727	163.52	0.832	0.3888	3.5588	168.36	0.714	0.3998	1.4173	182.24	0.8	0.469	20.367	163.75	0.542	0.3528

Table 5 further summarizes the average performance metrics. While *C1:Iter10k, res 4, vox 16 NCC* produced the highest average DSC (0.7873) and MI (0.3494), its TRE (1.8101 mm) was higher compared to *C2:Iter20k, res 4, vox 16 MI* and *A5:Iter10k, res 4, vox 16 MI*. The use of a smaller voxel size, as in *C4:Iter10k, res 4, vox 8 MI* and *C5:Iter10k, res 3, vox 16 MI*, resulted in significantly higher TRE values (6.4307 mm and 6.6790 mm, respectively), indicating poorer registration performance, despite having high MI values (0.4024 and 0.4026, respectively). These results suggest that smaller voxel sizes might not necessarily improve alignment but may instead introduce instability in the registration process.

Table 5: Average Performance Summary for Custom variation

Parameter File	Avg TRE (mm)	Avg Time (s)	Avg DSC	Avg MI
A5:Iter10k, res 4, vox 16 MI	1.3859	<b>145.34</b>	0.769	0.3239
C1:Iter10k, res 4, vox 16 NCC	1.8101	165.28	<b>0.7873</b>	<b>0.3494</b>
C2:Iter20k, res 4, vox 16 MI	<b>1.3553</b>	408.05	0.77	0.3214
C3:Iter10k, res 5, vox 16 MI	1.3972	199.48	0.7698	0.2941
C4:Iter10k, res 4, vox 8 MI	6.4307	398.10	0.7243	0.4024
C5:Iter10k, res 3, vox 16 MI	6.6790	169.47	0.722	0.4026

Additionally, increasing the resolution levels, as in *C3:Iter10k, res 5, vox 16 MI*, provided a balance between execution time and TRE (199.48 seconds and 1.3972 mm, respectively), suggesting that higher resolutions might offer improved performance at a reasonable computational cost. Overall, these findings highlight the importance of parameter tuning in balancing accuracy, efficiency, and computational cost for optimal registration results.

### 3.4 Qualitative Analysis

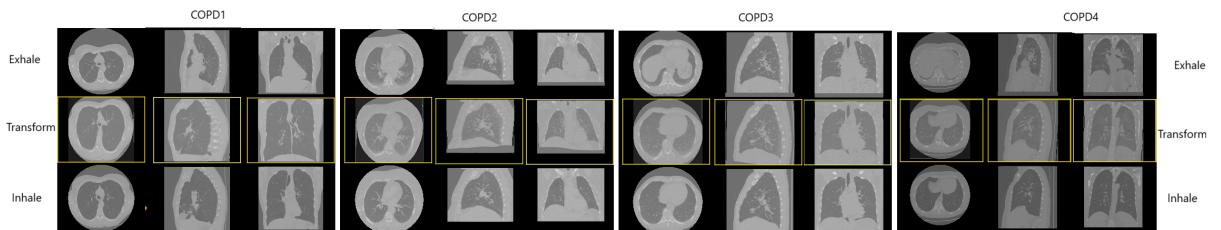


Figure 13: Visual Review of Images Original, Transformed and Target

Qualitative review of the best parameter as shown in image below confirms the good performance of the best model obtained from pipeline described in previous section, which included normalization with ZScore, with Bilateral filtering of sigma 1 and BSpline custom parameter with 20k iterations, FinalGridSpacingInVoxels of 16, using Mutual Information metric and multiresolution of 4.

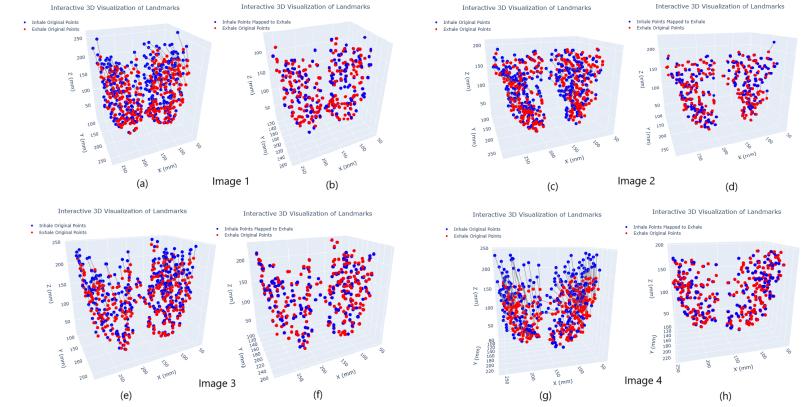


Figure 14: 3D Plot of points with correspondence original points and registered to GT.

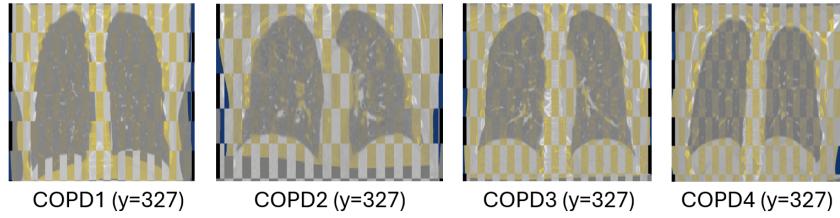


Figure 15: Checkerboard Representation of images

## 4 Project Management Details

The project was structured into three main tasks: data preprocessing, registration implementation, and evaluation, with balanced contributions from team members. Work was conducted using local resources, with code shared and reviewed collaboratively via GitHub. Weekly meetings facilitated progress reviews and pipeline refinements, with an average of 8 hours per member per week over 4.5 weeks, totaling approximately 115 hours for the team.

## 5 Conclusions

The project successfully demonstrated the use of advanced preprocessing, registration, and evaluation techniques tailored for inspiratory and expiratory CT lung scans. Key steps included normalization, mask refinement using morphological operations, and intensity-based registration using Elastix. This systematic approach ensured not only accurate anatomical alignment but also computational efficiency, making the pipeline robust for unseen test cases in the 4DCT DIR-Lab Challenge.

Furthermore, the described approach highlighted the critical importance of parameter tuning for registration performance. Through iterative testing of configurations with different transformation parameters, the custom BSpline model achieved the best results, with an average TRE of 1.433 mm and DICE scores of 0.766. This finding underscores the value of tailored parameterization in achieving high accuracy while maintaining computational feasibility, particularly in medical image registration tasks.

Finally, the pipeline's robustness was validated using a diverse set of evaluation metrics, including TRE, DICE similarity, and intensity-based metrics such as MI and NCC. In addition to quantitative assessments, qualitative visual inspections—such as checkerboard overlays and mask alignments—provided critical insights into the alignment accuracy.

## References

- [1] Pablo Alvarez et al. “Lung deformation between preoperative CT and intraoperative CBCT for thoracoscopic surgery: a case study”. In: *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*. Ed. by Baowei Fei and Robert J. Webster III. Vol. 10576. International Society for Optics and Photonics. SPIE, 2018, p. 105761D. doi: [10.1117/12.2293938](https://doi.org/10.1117/12.2293938). URL: <https://doi.org/10.1117/12.2293938>.
- [2] Richard Castillo et al. “A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive”. In: *Physics in Medicine Biology* 58.9 (Apr. 2013), p. 2861. doi: [10.1088/0031-9155/58/9/2861](https://dx.doi.org/10.1088/0031-9155/58/9/2861). URL: <https://dx.doi.org/10.1088/0031-9155/58/9/2861>.
- [3] Elastix Model Zoo. *Elastix Model Zoo Repository*. <https://github.com/SuperElastix/ElastixModelZoo>. Accessed: 2025-01-10. 2023.