4DCT DIR-Lab Registration Challenge using Elastix and Transformix

Abdelrahman HABIB
ESCOLA POLITÈCNICA SUPERIOR
Universitat de Girona
Girona, Spain
u1985258@campus.udg.edu

***Abstract*—This report presents a full lung segmentation, pre-processing and image registration methodologies that were applied on 4DCT DIR-Lab dataset challenge. The dataset comes with both raw intensity volumes and manually identified feature locations, forming a basis for the registration process. The project's objectives involve 3D CT lung image registration and subsequent evaluation using both Target Registration Error (TRE) and execution time. Among all experiments, we obtained a significant low mean TRE score among all of the normalized intensity images (1.5875) at a quick execution period (464 seconds) using a single parameter file with Elastix and Transformix.**

***Index Terms*—4DCT DIR-Lab, image registration, landmark transformation, Elastix, Transformix, lung segmentation, CT pre-processing**

I. INTRODUCTION

Image registration can be defined as the determination of a one-to-one mapping, known as transformations, between the coordinates in two spaces such that points in those spaces that corresponds to the same anatomical point are mapped to each other [1]. Many existing registration tools and algorithms can automatically register images that are related by a rigid body transformation (i.e. where the tissue deformation can be neglected), or by non-rigid registration algorithms that can compensate for tissue deformation [2].

In this project, 4DCT DIR-Lab dataset were given to perform image registration and keypoint transformation. The dataset consists of raw intensity volumes at two time frames, both maximum inhale and exhale, with different voxel sizes for each subject. Each subject intensity volume comes with a text file that contains a list of right/left (RL) - anterior/posterior (AP) - superior/inferior (SI) coordinate locations corresponding to 300 manually identified feature locations in its respective intensity volume. Thus for each subject, two intensity images and two keypoint feature files are given for maximum inhale and maximum exhale. The origin on the intensity images is located at (1, 1, 1) that represents the centroid of the right-anterior corner voxel of the volume, and the coordinate locations are given in units of voxel indices. The code developed for this project can be found in this repository https://github.com/abdel-habib/COPDgene-CT-registration.

II. PROJECT OBJECTIVES

The main objectives of this project are as following:

(a) To register the 3D CT lung images and transform the given landmarks.
(b) To evaluate the transformation using Target Registration Error (TRE), that is the euclidean distance between the transformed landmarks and the ground truth exhale landmarks.

III. CHALLENGE DESCRIPTION AND OBJECTIVES

The challenge of this course will be based on registering 2-3 new 4DCT test cases, that were not provided during the development stage. At the time of the challenge, intensity volumes with only inhale landmarks are given. The evaluation criteria for the challenge will be based on the accuracy of the registration and transformation, that is using TRE metric, the methodology used, and the computational time.

IV. METHODOLOGY & TOOLS

This section includes all the important steps that were used. Subsection IV-A includes the detail pre-processing steps experimented, and subsection IV-B includes all the experiments and versions experimented of lungs masks and improvements made along the project. The main registration and transformation approaches are explained in subsection IV-C.

As for the tools, different Elastix and Transformix versions were used, that were based on the model parameters files. The model parameters, along side with the elastix/transformix version used are listed below:

(a) Par0003 parameter file, elastix_windows64_v4.7.
(b) Par0007 parameter file, elastix_windows64_v4.7.
(c) Par0011 parameter file, elastix_windows64_v4.3 (4.301).
(d) Par0049 parameter file, elastix_windows64_v4.7.

Note that those parameter groups include multiple files, which all were experimented and reported. In addition, this project was developed using a Windows 11 Home, x64-based PC, 11th Gen Intel(R) Core(TM) i9-11900H @ 2.50GHz, 8 cores processor.

A. Pre-processing

All the pre-processing approaches, whether they are used in the final submitted result or not, that were experimented are explained in this subsection of the methodology.

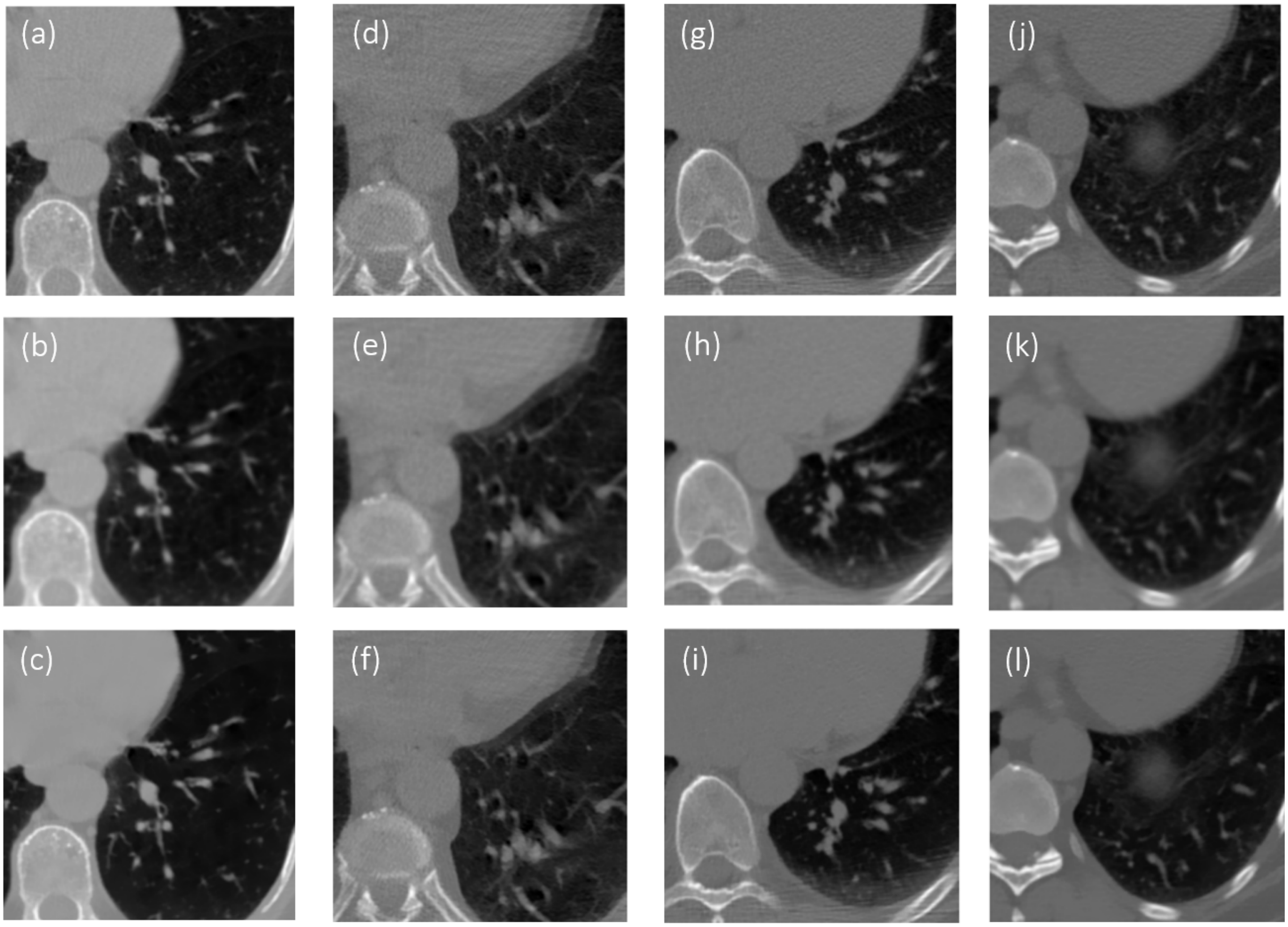


Fig. 1: Visual results of both filtering techniques applied on all given exhale intensity volumes. First row, that is (a), (d), (g), and (j) is for each subject copd1 to copd4 as in order. The second row, that is (b), (e), (h) and (k) is for the results of applying anisotropic filter. The final row, that is (c), (f), (i), and (l) is for the results of applying bilateral filter. All images were zoomed in the x- and y-coordinates [220:420] and in slice = 70.

1) Volume Denoising: Image de-noising in medical image analysis is a mandatory and essential pre-processing technique for many image analysis approaches. It is important to use it correctly to remove noise as noise can downgrade the overall quality and clarity of the visual information. It can be categorized into linear filtering, such as the Mean and the Gaussian filters, and non-linear filtering [3]. In the medical imaging field, several challenges could arise when it comes to de-noising, most importantly is edge-preserving. Linear filters cause blurring the images and simultaneously suppress the details as they blur sharp edges, destroy lines and other finer details [3], making them not suitable for medical images.

In this project, two de-noising methods were developed and experimented visually on the training images, which are Anisotropic diffusion and Bilateral filter. In the anisotropic diffusion method, the gradient magnitude is used to detect an image edge or a boundary as a step discontinuity in intensity [4]. In our implementation, we used $GradientAnisotropicDiffusionImageFilter$ from $SimpleITK$ library to filter the volumes with a $ConductorParameter = 3.0$, $TimeStep = 0.0625$, and $iterations = 5$ configurations.

The bilateral filter, that is as cited by [3] a non-iterative method, can be used to avoid such problems that can occur from linear filters as it preserves the edges while denoising, by means of a nonlinear combination of nearby image values. In our implementation, we used $BilateralImageFilter$ from $SimpleITK$ library to filter the volumes with a $DomainSigma = 2.0$ and $RangeSigma = 50.0$ configurations.

It can be seen clearly from Figure 1 that the bilateral filter results are more smoothed and edges are better preserved in all training subjects, compared to the anisotropic diffusion filter.

2) Intensity Normalization: For normalizing the image, we performed min-max normalization on the subject body, where as for that specific task, we create a mask for the body to

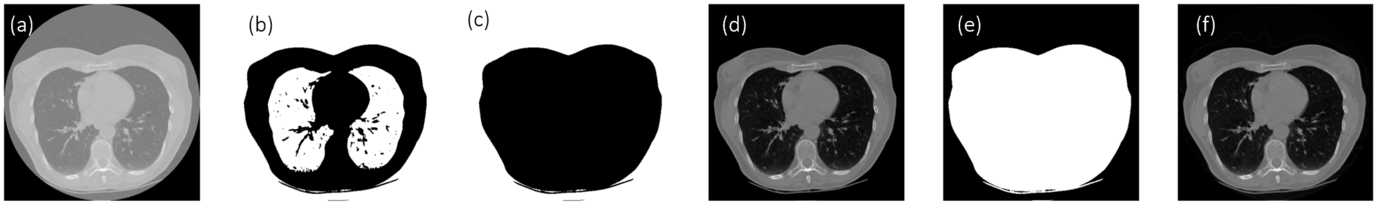


Fig. 2: Step by step, from left to right, of the intermediate results that were obtained to perform min-max normalization on the body, making sure to neglect the gantry and background. Sub-plot (a) is the original raw image, (b) is the threshold mask created, (c) is the largest mask obtained, (d) is the mask overlay-ed on top of the raw, segmenting the body, (e) is mask inverted, and finally (f) is the normalised result.

Fig. 3: Final pre-processing result obtained by applying all three pre-processing steps on subject copd1, slice = 70.

eliminate the gantry and the background using the mask. A detailed explanation of the segmentation approach is discussed in IV-B section, as the same code is used in both approaches to obtain the largest mask of the body using a threshold technique, $threshold = 700$ for subjects COPD1, COPD3, and COPD4, and $threshold = 470$ for COPD2. The reasoning for the difference of threshold value is also discussed in that section of the report. The key difference here is that we invert the largest mask to make the pixels with binary values = 1 as the region of interest. Figure 2 demonstrates the intermediate results obtained to produce a final normalized volume. In addition, the final normalized intensity image was clipped between 0 and the maximum data type value, that was 32767 for int16.

3) Contrast Enhancement: The final pre-processing approach experimented is Contrast Limited Adaptive Histogram Equalization (CLAHE), that is used from $SimpleITK$ library with $cliplimit = 0.01$ configuration. By applying the three pre-processing approaches, starting with min-max normalization, then bilateral filter, and finally CLAHE, the result obtained of the image is shown in Figure 3. This is to understand how the overall approach works together.

B. Lung Segmentation & Post-processing

Lung segmentation is an important step in this project as it assists in registering the region of interest, for our case, only the lung without the trachea, and obtain a low TRE score. The proposed segmentation approach consists of segmentation and post-processing, and explained below on raw images for easier understanding. In addition to that, multiple refined segmentation masks were obtained and the difference will be discussed briefly in this section as the refinement resulted in significant improvement in the TRE results obtained. All the results will be further discussed in Section VI.

The first step to segment the lung is to obtain an initial segmentation mask that we further process. This can be easily made using a global threshold method, as the histograms could be separated into two regions, one that includes the lung and trachea that is approximately between 100 to 700, 0 to 100 approximately is the gantry, and below 0 is the black background, approximately -2000 pixel value. Those values are approximations and depends on each subject case, where they were obtained during data analysis. For the cases of COPD1, COPD3, and COPD4, we used a threshold of 700 to create the mask. As for COPD2, we used a threshold of 430. It is the same mask that was used in the previous discussed normalization step in subsection IV-A2, where it is (b) in Figure 2. Those threshold values were experimental, and resulted in the best segmentation results. Subject COPD2 needed a specific threshold due to a minor difference in the lung intensity values in the lower lung regions, that caused some challenges in this segmentation task.

After obtaining a mask, we label connected components, with a $connectivity = 2$, that returns the labelled mask, and number of labels. We use the labelled masks to determine the largest regions based on their area. In our situation, we are interested to differentiate between the background and gantry together, and the both lungs, and the patient body, making the number of largest regions only 3, and ordered by their area sizes from largest (background for most of the case) to the smallest (trachea), making the lungs as the second largest

Fig. 4: Labelled mask obtained by measuring the area of the largest regions obtained from their connectivity property using copd1 raw image.

mask that we further work on. An example of a labelled mask is shown in Figure 4. We then use the second labelled mask to create a new mask only for it, making everything else out of the region of interest. In the first few slices, that second largest mask is usually the trachea as the lung is not introduced yet. However, as the slices go deeper, and the lung are visible, the second largest mask is always the lung. The trachea is eliminated in a further post-processing step.

As we mentioned before, subject COPD2 needed special threshold and processing steps to obtain an accurate segmentation. The reason behind it is due to the difference in some slices in the lung intensity, where regions of the lung has different intensities making them harder to segment and include in the global threshold we apply. Upon experimenting this in different approaches, we finalized the segmentation for this subject only to apply hole filling before we post-process and remove the trachea. The hole-filling is made using binary-closing morphological operation with a structure element size of (14, 14, 14) that we experimented. Note that the hole-filling is repeated after the post processing with a structure element (7, 7, 7) size to handle other cases.

The trachea elimination is an important post-processing step to eliminate the trachea as no key points are found there. We used the region properties of the largest masks such as the major and minor axis length, the area and the number of regions obtained from the largest mask. This means, from the largest mask, we count again the number of largest region as the trachea is included in the previous largest mask. Note that this post processing is applied on each slice of the volume individually. Several conditions were made and listed below:

(a) If there is only one largest region found, and the absolute difference between the major and minor axis length is above 30, we use this slice region.
(b) Else if there is only one region, and the previous value is less than 30, we return an empty slice to handle very small trachea.
(c) Else if there are 3 largest regions from the largest mask, this means that it found the two lungs (index 0 and 1 as we order by area) and the trachea as index 2, so we return both lungs by summing them into a single slice.
(d) Else if there are only 2 largest regions, this is where it gets challenging as we check if the difference in the area (area of the first region has to be at least 50 more than the second region) to indicate that it is a lung not a trachea, and also check if the minor axis of the second region (trachea) is less than 100. This condition happens when both lungs are touching each other as a region, and trachea as another region.
(e) Else if the previous condition is not yet met, it means it only captured the two lungs, thus we sum those regions as masks into a single slice and return it.
(f) Else if no condition is met, we return an empty slice.

This way, we make sure to eliminate the trachea in most of its conditions and have a more refined segmentation of the lung. Figure 5 demonstrates a slice example of removing the lung and applying hole-filling after to obtain a more refined segmentation mask for the lung.

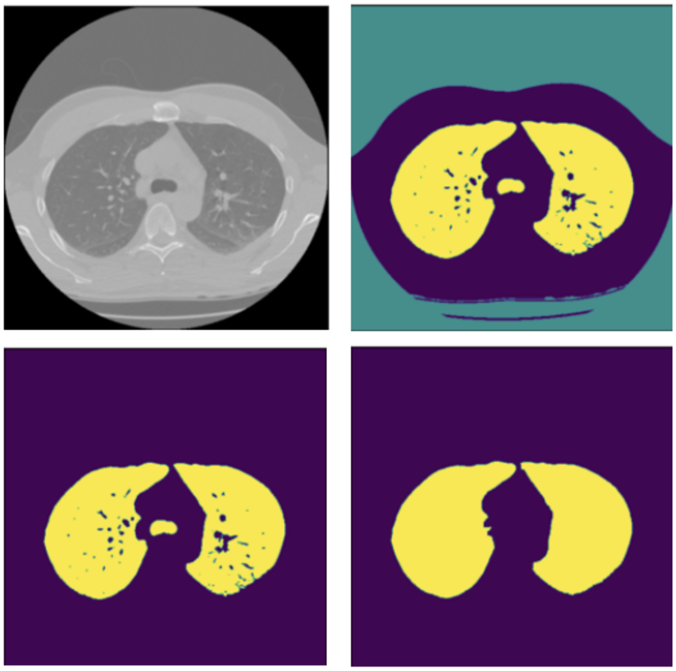


Fig. 5: Post-processing result to remove the trachea and hole-filling applied after to obtain a final lung segmentation mask.

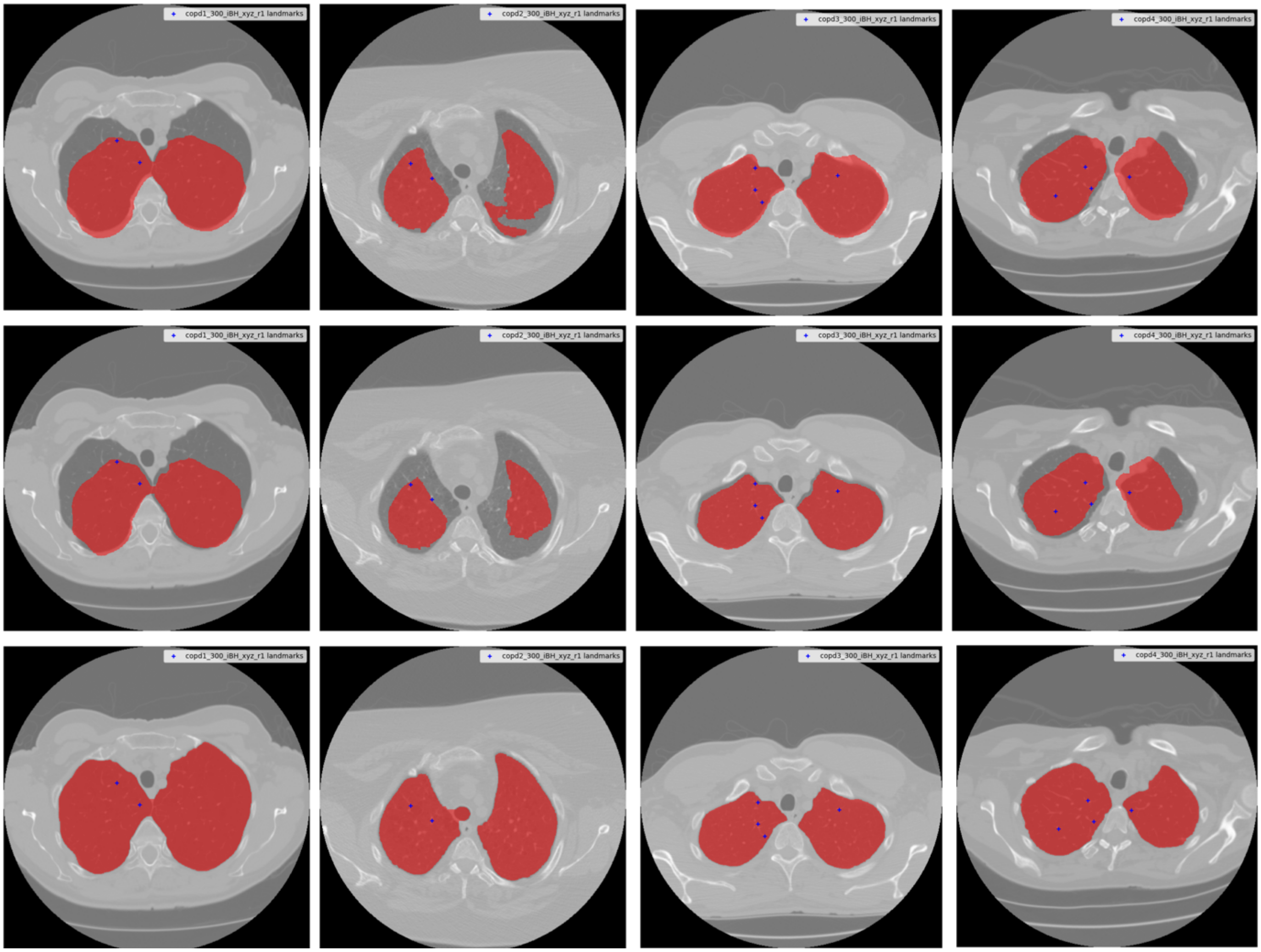


Fig. 6: Visual evaluation for the three segmentation trails results overlay-ed the intensity images at a difficult slice to segment, that is slice = 21, and with the keypoints on blue. From left to right are the four training subjects, and from top to bottom are the segmentation trials made, where the final row is the third segmentation trial (Mask 3) with the best results obtained.

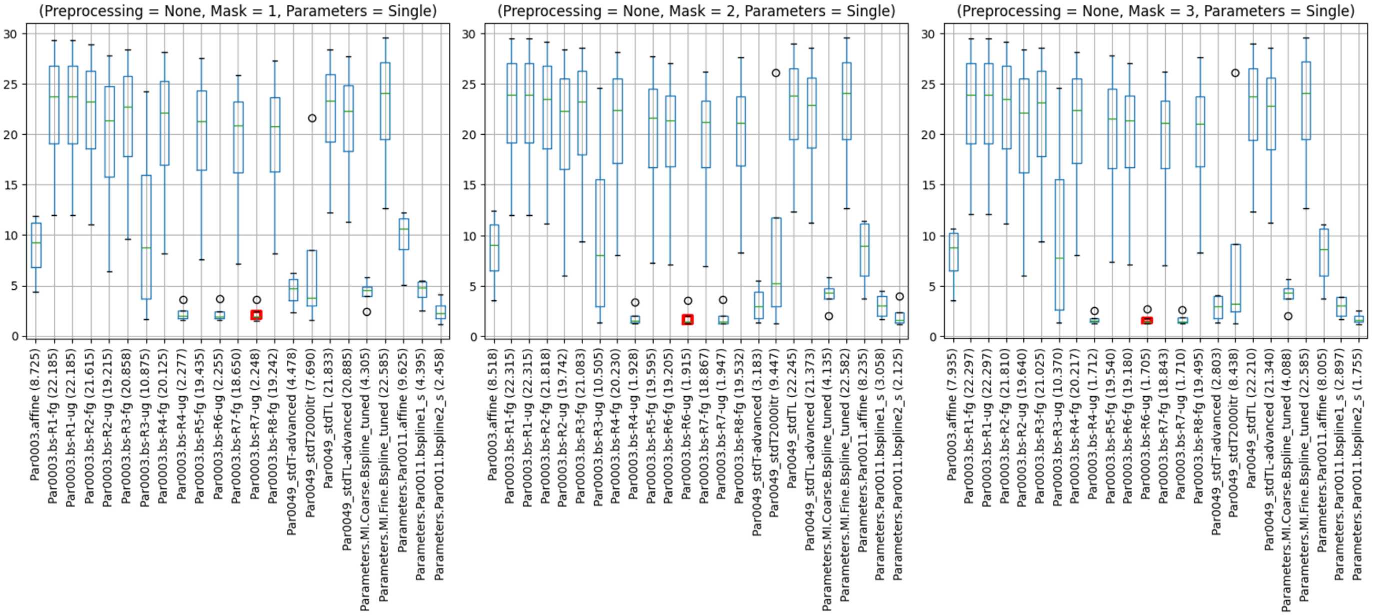
C. *Image Registration & Landmarks Transformation*

The intensity image registration is made using Elastix with its corresponding version of the parameter file that was mentioned previously in IV. One import decision to make is to decide which image is fixed, and which is moving between the inhale and exhale to pass them correctly when creating the executable script command file. This also is very related to landmark keypoint transformation using Transformix as we will be only given the inhale landmarks as testing landmarks, and will be asked to transform them.

The landmarks files are text files that contains 300 rows of coordinates, given in (x, y, z) of each landmark coordinate position in the intensity volume. Each slice could have none or more than one landmark, sharing the same z-axis value in the landmark file. As we are given inhale landmarks that we want to transform to exhale landmarks (considering the test case), the inhale landmarks are the one that should be given to Transformix for transformation.

In elastix manual, the transformation is defined as a mapping from the fixed image to the moving image, where the moving image $I_M(x)$ is deformed to fit the fixed image $I_F(x)$, where the registration task is to find this transformation $T(X) = x + u(x)$ that makes $I_M(T(x))$ spatially aligned to $I_F(x)$, where $u(x)$ is the displacement. It is also important to address a common confusion about the transformation is its direction. In elastix the transformation is defined as a coordinate mapping from the fixed image domain to the moving image domain, where one deforms the moving image, but the transformation is still defined from fixed to moving image.

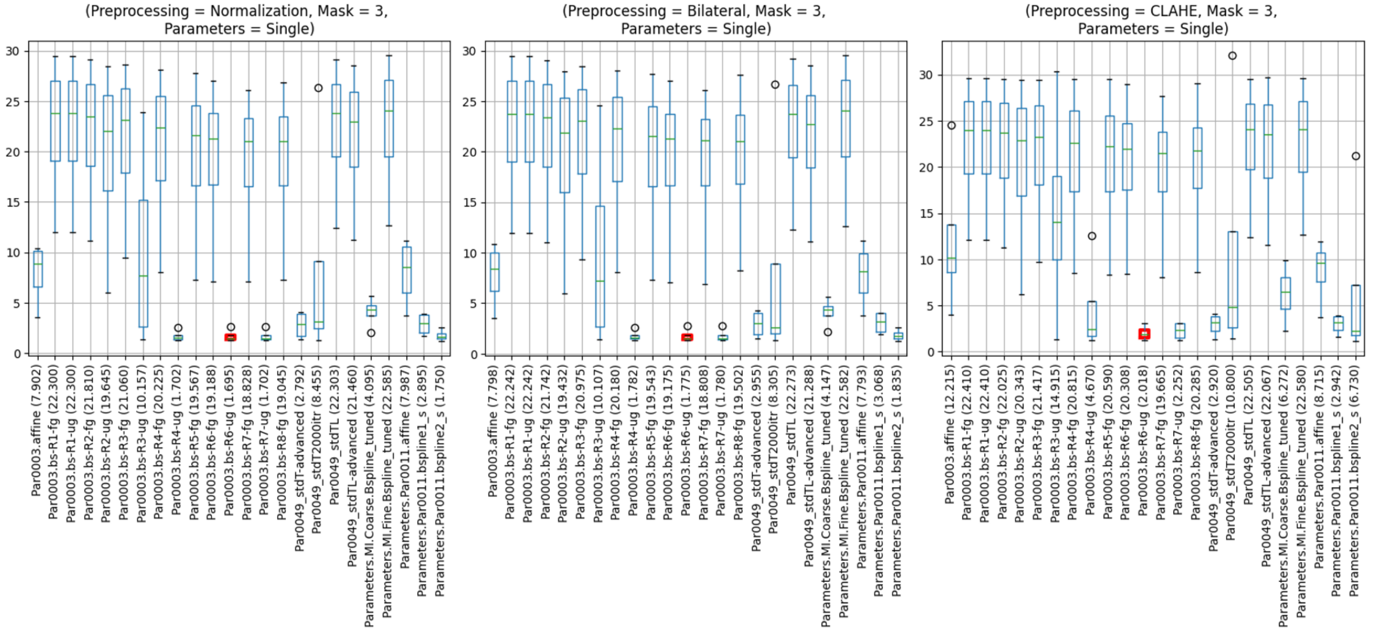
As for the given coordinate keypoints locations, as mentioned in Section I, the locations are given in the coordinate system in indices. Thus, based on elastix and transformix manual, to transform a keypoint file, some modifications has to be made to the file, which is to add as the first row of the file to transform $< indices, point >$, where for our case, it is $indices$, and to add in the second row the number of points, that is 300 for


(Preprocessing = None, Mask = 1, Parameters = Single)

(Preprocessing = None, Mask = 2, Parameters = Single)

(Preprocessing = None, Mask = 3, Parameters = Single)

Fig. 7: Overall mean (TRE_mean) results of the four training volumes obtained using raw data (not preprocessed) and the three mask segmentation iterations made using all single parameter files. Boxes in red are for the best experiments.


(Preprocessing = Normalization, Mask = 3,
Parameters = Single)

(Preprocessing = Bilateral, Mask = 3,
Parameters = Single)

(Preprocessing = CLAHE, Mask = 3,
Parameters = Single)

Fig. 8: Overall mean (TRE_mean) results of the four training volumes obtained using pre-processed data (separately) and the best mask segmentation iterations made using all single parameter files. Boxes in red are for the best experiments.

our case. Thus, this modification had to be made for all inhale keypoints to allow them to be transformed. It is also stated in the manual that the given points to be transformed has to be of the fixed image if they are indices, making it reverse the transformation direction.

Thus, in our implementation, the inhale intensity images were used as fixed images due to the requirement of the transformix, and the exhale were used as moving images. An example of a command line for elastix and transformix are shown below:

$$\begin{gathered}\text{elastix} - f\ \text{inhale\_img} - m\ \text{exhale\_img}\\ -fMask\ \text{inhale\_lung\_mask}\\ -mMask\ \text{exhale\_lung\_Mask}\\ -p\ \text{param\_file} - out\ \text{output}\end{gathered}$$

$$\text{transformix} - def\ \text{copd*\_300\_iBH\_xyz\_r1.txt}$$
$$-tp\ \text{path\_to\_transform\_file} - out\ \text{output}$$

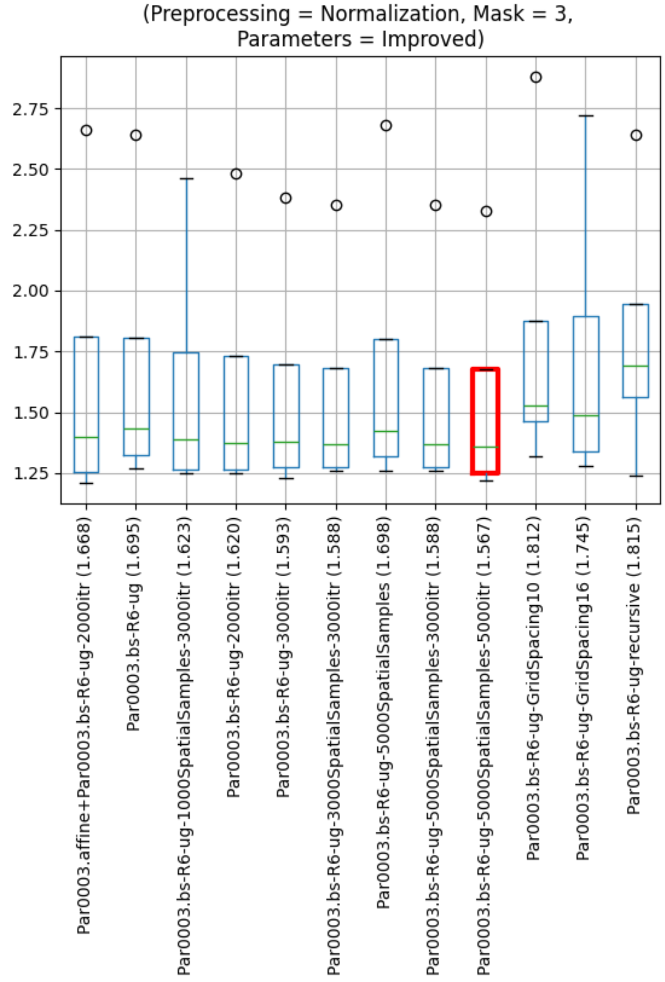


Fig. 9: Results obtained by modifying the best parameter file in order to improve the results.

V. EXPERIMENTS SETUP

A. Batch Script Automation

To handle experimenting different parameter files, all the process of pre-processing, segmentation, registration and transformation, as well as evaluation has been automated to run using an executable command line. We also automated the registration and transformation using Windows batch (.bat) files to run the registration and transformation together for a given data and parameter files. This way, it is easy to test any new set of volumes during the challenge data release.

B. Single Parameter File Experimenting

As we are using multiple parameter group files, those listed in Section IV, each of the sub-files has been tested separately. Some of them failed to register even with a given mask, some didn't improve the TRE compared to the ground truth, and some obtained a very low TRE error. This was performed on the intensity images with and without pre-processing, using different variations of segmentation mask improvements, which is discussed in more details in Section VI.

C. Improving Best Parameter

After evaluating and experimenting all parameters, we modified the best elastix parameter file to further improve the results and ran multiple experiments. The results with out modifications had improved slightly yielding lower TRE error, which is discussed in more details in Section VI.

VI. RESULTS AND DISCUSSION

A. Lung Mask Segmentation

The main objective for the lung segmentation is to assist in registering and transforming the keypoints inside the lungs. As we don't have a ground truth to evaluate how accurate is the segmentation, we need to insure that the segmentation in most, if not all of the slices, contains both lungs area, and has the keypoints inside it. We have evaluated that objective visually in all training subjects, by evaluating slice by slice, specially focusing on the harder segmentation regions of COPD2 subject. The mask segmentation was iterated three times, and the results of the registration improved on every iteration. The first iteration had the hole-filling only after the post-processing, with different sizes of structure elements for dilation and erosion, thus the mask was over-segmenting and not accurate for COPD2. The second iteration, hole-filling was applied only after the post-processing with a similar size of structure element (7, 7, 7), this made the segmentation more accurate, but not accurate enough to include all the keypoints. Another improvement was made was to use $connectivity = 2$ instead of $connectivity = 1$. The final iteration of the segmentation included the hole-filling before and after the post-processing as discussed in the previous sections. Figure 6 shows the segmentation results of the three trials, where row 1 is the first trail, till the third row of the best segmentation trial. Subjects are evaluated from left to right COPD1 to COPD4. As we can see clearly that the third iteration of segmentation improvement was very accurate in segmenting the lung in this difficult slice to segment, that was slice number 21.

B. Registration and Transformation

As mentioned in the subsection of V-B, experimented the different masks, using different parameter files, on different data (pre-processed and not pre-processed). As shown in Figure 7, we experimented the three masks improvement iterations on the data that was not pre-processed using all single parameter files of the parameter groups mentioned before. From the obtained results, the overall mean TRE obtained from calculating the mean from the four training volumes was improving significantly on every iteration segmentation improvement. The red box-plot of each of those plots highlights the best parameter file with the lowest obtained mean TRE values. Among all the experimented single parameter files on not pre-processed data, parameter file $Par0003.bs-R6-ug$ showed the best results on all iterations, and was improving significantly on every mask improvement (Mask1 = 2.248, Mask2 = 1,915, Mask3 = 1.705).

TABLE I: Sample results obtained using the best experiment for each class (Preprocessing=Normalization, Mask=3, Parameters=Par0003.bs-R6-ug)

Sample	***TRE_mean***	***TRE_std***	***Time (sec)***
COPD1	1.34	1.37	41
COPD2	2.64	3.57	38
COPD3	1.27	1.16	44
COPD4	1.53	1.20	42
Mean	1.695	1.825	Total = 165

TABLE II: Sample results obtained using the improved experiment for each class (Preprocessing=Normalization, Mask=3, Parameters=Par0003.bs-R6-ug-5000SpatialSamples-3000itr)

Sample	***TRE_mean***	***TRE_std***	***Time (sec)***
COPD1	1.26	1.18	117
COPD2	2.35	3.31	110
COPD3	1.28	1.19	121
COPD4	1.46	1.15	116
Mean	1.5875	1.7076	Total = 464

We also experimented the pre-processing using the best mask (Mask3) on each of the pre-processing steps discussed before separately. As in Figure 8, the results of using the pre-processing had a very slight improvement when we used min-max normalization using the same best parameter file $Par0003.bs-R6-ug$, where it improved from 1.705 (without pre-processing) to 1.695 (with pre-processing). We also experimented using the three pre-processing steps together but didn't observe any improvement more than that. As for the detailed TRE results of each sample using the best approach found, as well as the execution time, Table I summarizes the sample results where COPD2 still had the highest TRE result, yet the overall approach performed greatly. The total execution time was very quick too (165 seconds) for all subjects.

A final trial was made to further improve the results by modifying the best parameter file $Par0003.bs-R6-ug$ as explained in V-C. The file was analyzed and several values inside were modified and experimented to try to reduce the overall mean TRE from 1.695. The modifications made were to experiment more iterations from the original number (2000, 3000, 5000), different grid spacing values (10, 16), different spatial sample numbers (1000, 3000, 5000), different pyramid type, or combination as shown in Figure 9. The best results were observed when we increase the number of iterations, where it dropped to 1.588 at 3000 iterations and 1.567 at 5000 iterations instead of 1.695 using the default 1000 iterations. This improvement comes with a cost, that is the computational time, as the more iterations, the more it takes to register the images. In addition, the difference of improvement between 3000 and 5000 iterations is not very significant to increase the iterations more, this makes $Par0003.bs-R6ug$-$5000SpatialSamples3000itr$ more convenient to use, with 5000 spatial samples and 3000 iterations modified from the original file, resulting in 1.588 mean TRE among all train subjects. The TRE result for each subject using this parameter file is summarized in Table II.

CONCLUSION

In conclusion, different segmentation approaches and different parameter files had been experimented in this project to register and transform 4DCT data at inhale and exhale time frames. Among all the experiments performed, we found that pre-processing the data with min-max normalization, and using the third iteration (best segmentation) lung mask with $Par0003.bs-R6-ug-5000SpatialSamples-3000itr$ parameter file performed very greatly among all experiments in obtaining a very low TRE score with a descent execution time.

REFERENCES

[1] C. R. Maurer and J. M. Fitzpatrick, "A review of medical image registration," *Interactive image-guided neurosurgery*, vol. 1, pp. 17–44, 1993.

[2] D. L. Hill, P. G. Batchelor, M. Holden, and D. J. Hawkes, "Medical image registration," *Physics in medicine & biology*, vol. 46, no. 3, p. R1, 2001.

[3] D. Bhonsle, V. Chandra, and G. Sinha, "Medical image denoising using bilateral filter," *International Journal of Image, Graphics and Signal Processing*, vol. 4, no. 6, p. 36, 2012.

[4] Y. Yu and S. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, 2002.