Image registration of chest CT volumes

MIRA: Final Project 2023 MAIA/MIC

Agustín CARTAYA
Master in Medical Imaging and Applications (MAIA)
Universitat de Girona
Girona, Spain
u1985500@campus.udg.edu

Micaela RIVAS
Master in Medical Image Computing (MIC)
Universitat de Girona
Girona, Spain
u1962546@campus.udg.edu

***Abstract*—Image registration is one of the most important steps in the treatment of medical images. This process is essential for accurate analysis and interpretation of medical images, where precise alignment of different datasets is necessary for effective diagnosis and treatment. In this project, we focused on comparing two different approaches to register 4D CT images from the inhalation phase to the exhalation phase. The comparison includes a traditional approach using Elastix software, incorporating affine and BSpline registrations, contrasted with an alternative method utilizing the VoxelMorph deep learning algorithm. The evaluation was performed using the Target Registration Error (TRE) criterion and achieved an average TRE of 4.12mm in the training images.**

***Index Terms*—Image registration, affine, BSpline, Elastix, 4D CT, inhalation, exhalation, VoxelMorph, Target Registration Error**

I. Introduction

The registration of medical images is a fundamental step in numerous image analysis pipelines [4]. This process primarily involves applying various transformations to one image, known as the moving image, to align it with another, the fixed image. Such alignment is key for accurate comparison and analysis, particularly in situations where precise anatomical changes need to be tracked over time or under different conditions.

The aim of our project is to develop a registration pipeline tailored for chest CT scans within the context of a Chronic Obstructive Pulmonary Disease (COPD) dataset [5]. The images to be aligned represent inhalation and exhalation phases. This alignment is needed for understanding structural changes in the lungs and thoracic area during distinct respiratory phases, with potential implications for diagnosis and treatment planning in COPD patients.

To develop this algorithm, we compared two approaches: one using the well-known software Elastix [1], a traditional method, and the other employing the widely recognized deep learning framework VoxelMorph [2].

In the performance assessment of the registration methods, we utilized landmarks within the lung region. Such comparative analysis is essential to determine the most effective method for this specific application in medical imaging.

II. Dataset description

The dataset we employ for this project is taken from the COPD dataset, with four cases selected for the training set. Each of these cases contains raw format scans captured at maximum inhalation and exhalation phases.

Accompanying the images, a .txt file is provided for each image, containing 300 landmarks (one set for inhalation and another for exhalation). These landmarks consist of a list of Right-Left, Anterior-Posterior, Superior-Inferior coordinates.

Also, a table with useful information about the images is included, this table contains:

1) Image dimensions: provided in voxel units.
2) Voxel spacing: specified in millimeters.
3) Number of Features: indicating full point sets identified in each case
4) Displacement (in millimeters): representing the TRE and the standard deviation without any form of registration. The primary objective of this project is to get an error smaller than this.
5) Number of Repeats: denoted as N_m/N_{obs}, where N_m represents the number of repeat registration measurements conducted by each N_{obs} independent observer.
6) Observers (in millimeters): indicating the difference, in millimeters, between the actual point and the point marked by the doctor.

The testing set consisted of 3 cases with the same characteristics described above for the training set, but only the inhale landmarks were provided.

III. Methodology

A. Data preparation

As the images were given in a raw format and in a different orientation than the landmarks, we used the software ITK-Snap to read each file (using the data from the information table given) and correct its orientation, from Right-Left, Anterior-Posterior, Inferior-Superior to Right-Left, Anterior-Posterior, Superior-Inferior. We also converted the images to a NIFTI format to make it easier to work with them in python.

B. Data analysis

Initially, we conducted an analysis of the images to compare various similarity metrics between inhale and exhale images without any registration. The computed metrics included TRE, Root Mean Square (RMS), Normalized Cross-Correlation (NCC), and Mutual Information (MI).

Table I presents the values of these metrics and Fig. 1 illustrates the joint histogram for all cases before registration.

TABLE I
DIFFERENT METRICS OF BETWEEN INHALE AND EXHALE IMAGES BEFORE REGISTRATION

Patients	**Similarity metric**			
	TRE (mm)	*RMS*	*NCC*	*MI*
1	26.3342	0.0912	0.5675	0.6775
2	21.7859	0.0649	0.7118	0.6238
3	12.6391	0.0464	0.8910	1.1562
4	29.5835	0.0831	0.6220	0.7175
Mean	22.5856	0.0714	0.6980	0.7937

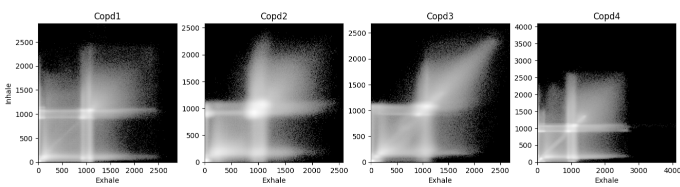

Fig. 1. Joint histogram between inhale and exhale images of all patients before registration

C. Lung Segmentation

Since the landmarks were located in the lung region, we created a threshold base algorithm to segment them. The masks obtained were used to specifically target the registration process on the lung regions, thereby excluding other irrelevant parts.

The generation of these masks involved a thresholding approach on the pixel values of the images to identify lung tissues. Pixels with intensity values greater than 0 and less than 800 were considered part of the lung tissue. This strategy was effective in differentiating lung tissue from other anatomical structures. Following the initial thresholding, binary morphological operations were employed to refine the masks. The purpose of this refinement was to eliminate noise and close small gaps within the lung regions.
The final step in the segmentation process involved extracting a centered cube from each mask, within a predefined range. In this central region, connected components were identified. The two most prevalent labels in the mask, which were present in the central region and excluding the background, were used to create the final masks.
A segmentation of the lungs with the landmarks can be seen in Fig. 2.

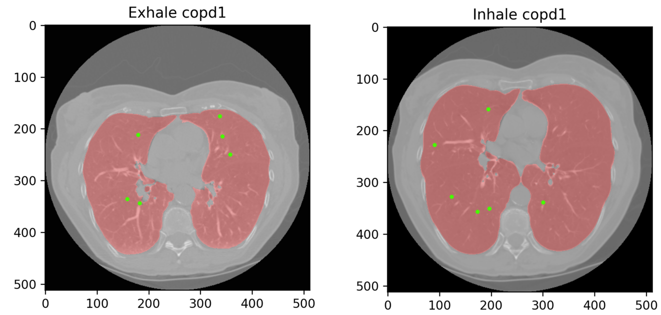

Fig. 2. Inhale and exhale images of the patient 1 with the lung mask in red and the landmarks in blue

D. Preprocessing

To enhance the registration process, various preprocessing techniques were explored:

a) Transform negative values to zero:
Given that the voxels values of the images were not in the range of Housfield units, as an first preprocessing, we eliminated the negative values that were part of the background. By transforming these negative values to zero, we effectively reduced the standard deviation of the image and decreased the importance of the background.

b) Image normalization with masks:
In order to have the voxel values of the inhale and exhale images in the same range, we normalize them taking into account only the region inside the masks. This method aimed to mitigate high variations in pixel intensity and facilitate more robust registration outcomes.

c) Histogram matching:
To ease the registration process, we implemented a histogram matching technique to match the histogram of the exhale image to the one from the inhale one. This process was performed pair-wise for all patients of the dataset.

E. Registration using Elastix

As first approach, we utilized the renowned registration software Elastix. We chose this software for its robust capabilities in precisely aligning medical images and transforming the landmarks with the same transformation of the image.

a) Parameters files:
In the context of registering CT scans from different respiratory phases, rigid registrations are typically used to align the images, while non-rigid registrations are employed to align elements that undergo deformations at different points during breathing. Those different types of registrations are possible in Elastix by changing the parameters files. We conducted experiments with these two types of registrations both individually and in combination to compare their results.

Table II shows some of the main parameters used in the parameters files to register with Elastix.

TABLE II
MAIN PARAMETERS IN THE ELASTIX PARAMETERS FILES

Metric	Registration type	
	Rigid	*Non-rigid*
Metric	NormalizedMutualInformation	AdvancedMattesMutualInformation
Interpolator	BSplineInterpolator	BSplineInterpolator
Optimizer	AdaptiveStochasticGradientDescent	StandardGradientDescent
Registration	MultiResolutionRegistration	MultiResolutionRegistration
NumberOfResolutions	4	4

All these parameters were chosen since they were the ones that produced the best TRE results in all cases, based on prior tests experimenting with different values. To combine rigid and non-rigid transformations, we added these two parameters files to the registration using the method compose (specified in the parameters files as HowToCombineTransforms).

F. Registration using VoxelMorph

We implemented a second approach using deep learning. The main idea was to create a specific model to register each pair of images and then, with the obtained deformation field, transform the landmarks of the moving image. This was achieved using the Voxelmorph library [1].

a) Model:
The model constructed was based on a U-Net architecture with 5 encoding layers and 7 decoding layers with (32, 48, 96, 128, 256) and (256, 128, 96, 48, 32, 16, 16) features, respectively. The model takes as input two 3D volumes corresponding to the fixed and moving images. At the end of the network, two additional convolutional layers were added to generate the network's output. The first layer corresponds to a 3x3x3 convolutional layer to obtain the deformation field, and the second layer was a transformation layer that takes the deformation field and the moving image as input, returning the transformed image.

b) Loss Functions:
The implemented loss function includes two main components. The first is a similarity metric between the resulting images from the network (transformed) and the fixed image. The second is a penalty parameter for the obtained deformation. For the similarity metric, experiments were conducted with Normalized Cross Correlation (NCC) and Mean Square Error (MSE). The penalty metric was implemented with L2 regularization for the Jacobian of the deformation field, penalizing large values provided by the network. This regularization aimed to control the smoothness and consistency of the deformation to avoid unrealistic or non-smooth solutions in the image space.

c) Optimizers:
Different types of optimizers were used to find which one allowed the models to converge faster and have fewer fluctuations. The implemented optimizers were Adadelta, Adam, SGD, and Adagrad.

IV. EXPERIMENTS

A. Using Elastix

a) Comparing parameters files:
We conducted experiments involving Affine, BSpline, and a combination of Affine followed by BSpline registration. All these tests were performed using the original image to compare their TRE and determine which registration method produced the best results.

Table III presents the TRE results, showing a substantial variation in TRE across different registration methods. Affine registration tends to exhibit higher errors, whereas a combination of Affine and BSpline yields notably lower errors.

TABLE III
TRE OF THE REGISTERED LANDMARKS WITH DIFFERENT REGISTRATIONS ORIGINAL

Patients	TRE (mm)		
	Affine	*BSpline*	*Affine + BSpline*
1	70.9289	28.3524	4.6182
2	136.1411	21.7451	6.6572
3	62.3602	15.3322	1.5890
4	69.8942	24.4608	4.6393
Mean	84.8318	22.4726	4.3759

b) Comparing preprocessings:
We experiment with the different preprocessings explained in the methodology section, which are transform negative values to 0, image normalization with masks and histogram matching. All these tests were performed with the best registration obtained in the previous experiments (Affine + BSpline). Table IV shows the TRE of the different of the landmarks after registering the images with the different preprocessings. Histogram matching gave us the worst results, being even worse than the results without any kind of registration, while transforming the negative values to 0 led us to the best performance

TABLE IV
TRE OF THE REGISTERED LANDMARKS WITH DIFFERENT PREPROCESSINGS

Patients	TRE (mm)			
	Original	*Remove <0 values*	*Normalization*	*Histogram matching*
1	4.6182	3.0842	3.1404	70.0679
2	6.6572	8.2429	8.7303	127.3280
3	1.5890	1.4922	1.4692	58.3221
4	4.6393	3.6744	3.6335	70.1286
Mean	4.3759	4.1234	4.2433	81.4616

B. Using VoxelMorph

All experiments conducted with Voxelmorph were done by resizing the images to a size of 128x128x32 due to computational limitations. To calculate the resulting Target Registration Error (TRE), landmarks were reduced to the dimensions allowed by the deformation field, transformed, and then brought back to their original dimensions. We are aware that the obtained results are approximations of the real ones and that there is significant information loss for the models when resizing the images to a smaller scale.

a) Comparing Loss Functions: We performed all initial experiments comparing the NCC and MSE similarity functions used as loss functions in the model. Upon obtaining results for all patients, we noted that models trained with NCC exhibited superior results compared to MSE. Table V showcases the outcomes achieved by the models trained for each patient using both loss functions.

TABLE V
TRE OF THE REGISTERED LANDMARKS WITH DIFFERENT LOSS FUNCTIONS

Patients	**TRE (mm)**	
	Loss func. NCC	*Loss func. MSE*
1	14.6493	19.7363
2	16.7237	20.6539
3	9.2823	10.3582
4	15.7693	17.7218
Mean	14.1061	17.1175

b) Comparing Preprocessing Methods: A second experiment was conducted, comparing different preprocessing techniques for the model inputs. The applied preprocessing methods mirrored those used for Elastix. After comparing the results obtained with the different preprocessing approaches, it became evident that the best approach was normalization, followed by histogram matching. Models trained with the original images or those with background intensities set to 0 yielded random results with extremely large TREs. Table VI presents a comparison of the outcomes obtained with the two best preprocessing techniques.

c) Optimizing Optimizers: A third experiment was performed to select the best optimizer. The results of Adadelta, SGD, and Adagrad optimizers with learning rates of 1.0, 0.01, and 0.001 were evaluated, along with Adam using a default learning rate of 0.001. After multiple iterations for each patient, it was observed that the Adam optimizer converged faster and exhibited considerably fewer fluctuations during the initial iterations compared to the others.

TABLE VI
TRE OF THE REGISTERED LANDMARKS WITH DIFFERENT PREPROCESSINGS

Patients	**TRE (mm)**	
	Normalization	*Histogram matching*
1	14.6493	26.3342
2	16.7237	21.7859
3	9.2823	12.6391
4	15.7693	29.5835
Mean	14.1061	22.5856

V. RESULTS AND DISCUSSION

The results obtained indicate better performance in terms of TRE using Elastix compared to VoxelMorph. This difference may be attributed to the resize of the image, the resize of the image made us lose relevant information. However, in other metrics, VoxelMorph yields better results. This could be due to the fact that the registration process in Elastix primarily focuses on the lungs region (where all the landmarks are), whereas in VoxelMorph, the lung masks are only employed for image normalization.

Table VII displays the resulting similarity metrics of the registrations using Elastix and VoxelMorph. Additionally, Fig. 3 illustrate the joint histograms in the original images and after registration with Elastix and VoxelMorph.

Additionally, we computed the difference images, as represented in Fig. 4, between the fixed and moving images before and after registration with Elastix and VoxelMorph. These images show that, in the lung region, the difference is significantly reduced with Elastix registration compared to VoxelMorph registration, overall. While VoxelMorph produces better results in the rest of the image.
These results can be caused by the fact that in VoxelMorph the registration is not specifically focused on the lungs, while in Elastix it is.

TABLE VII
SIMILARITY METRICS OF THE BEST REGISTRATIONS

Patients	**Similarity metric**			
	TRE (mm)	*RMS*	*NCC*	*MI*
1 Elastix	3.0842	0.1090	0.4047	0.7582
1 VoxelMorph	14.6493	5.7881×10^{-5}	0.6219	0.9029
2 Elastix	8.2429	0.0730	0.6187	0.7095
2 VoxelMorph	16.7237	4.9374×10^{-5}	0.7072	0.8969
3 Elastix	1.4922	0.0567	0.8338	1.2331
3 VoxelMorph	9.2823	3.8871×10^{-5}	0.8183	1.3882
4 Elastix	3.6744	0.0964	0.4558	0.7283
4 VoxelMorph	15.7693	5.8886×10^{-5}	0.6101	0.8869
Mean Elastix	**4.1234**	0.0837	0.5782	0.8572
Mean VoxelMorph	14.1061	$\mathbf{5.1246 \times 10^{-5}}$	**0.6893**	**1.0187**

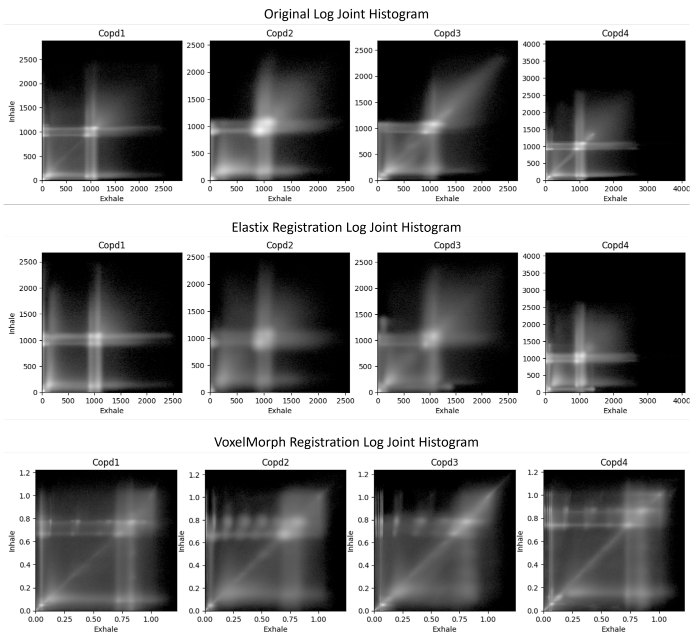

Fig. 3. Joint histogram between inhale and exhale images of all patients before and after registration with Elastix and VoxelMorph

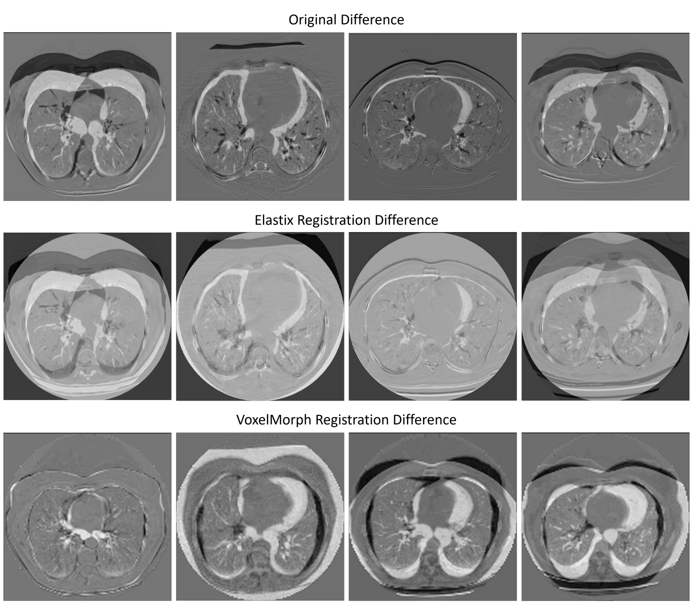

Fig. 4. Difference images between inhale and exhale before and after registration with Elastix and VoxelMorph

VI. Conclusion

In this study, we showcase the differences between two currently widely used methodologies for medical image registration. We conducted multiple experiments and optimized parameters for our methods. It was observed that non-rigid registrations are necessary for aligning inhalation and exhalation CT images, as evidenced by Elastix using B-spline and VoxelMorph utilizing the deformation field. Additionally, preprocessing data significantly aids in image registration, given the differing intensity scales among images that might otherwise lead models to fail to converge or produce random outcomes with a substantial Target Registration Error (TRE). Ultimately, it was evident that when employing Deep Learning for registration, maximizing information intake is crucial, thus reducing image dimensions may yield less reliable results

References

[1] Klein S et al. “Elastix: a toolbox for intensity-based medical image registration.” IEEE Trans Med Imaging. 2010 Jan;29(1):196-205.

[2] Guha Balakrishnan et al. “VoxelMorph: A Learning Framework for Deformable Medical Image Registration.” CVPR 2018. arXiv:1802.02604

[3] Michihiro Nakamura, Tsuneo Yamashiro, et al. “Assessment of Registration Accuracy in Chest CT using Anatomical Landmarks”

[4] Wang D et al. “PLOSL: Population learning followed by one-shot learning pulmonary image registration using tissue volume preservation and vesselness constraints.” Med Image Anal. 2022 Jul;79:102434

[5] https://www.cdc.gov/copd/index.html