IMPLEMENTATION OF VOXELMORPH AND ELASTIX BASED APPROACHES FOR COPD LUNGS DATA REGISTRATION

Zohaib Salahuddin and Abdullah Thabit

Universidad de Girona, Università degli studi di Cassino e del Lazio Meridionale and Université de Bourgogne

ABSTRACT

Image Registration is a fundamental step which can not only help in the direct diagnosis from the registered images but its performance has a direct impact on further algorithm's design to aid diagnosis in medical imaging. Inspiratory and Exipratory breath-hold CT image pairs acquired from the National Heart Lung Blood Institute COPDgene study archive have been used. Affine and BSpline Registrations were applied on unsegmented and segmented lung structures using Elastix. Mean(std) TRE of **1.11 (1.05), 1.98 (2.09) , 1.17 (0.99) and 1.45 (1.02)** was achieved using elastix for COPDgene cases 1,2,3 and 4 respectively. These results are better than the state of the art **NLR, LMP and SGM3D** registeration algorthims for these cases which is a testament to the efficiency of our approach and tunned parameters. A second approach using VoxelMorph was applied on the same dataset combined with the 4DCT dataset. Using VoxelMorph, we were able to get slight better TRE but the visual result showed better registration of the lungs. However, the nodules and fissures were not properly registered due to lack of data. However, Encouraging Results were obtained using VoxelMorph.

Index Terms— Chronic Obstructive Pulmonary Disease , Elastix , Transform , TRE, VoxelMorph, 4DCT/COPDgene, Affine Registration, BSpline Registration, Spatial Transform Network

1. INTRODUCTION

Image Registration refers to the process of aligning the images with respect to a fixed reference by the means of a global transformation applied to all the pixels together (Rigid Transformation) or local transformation applied to each region or pixel separately (Non Rigid Transformation). The Rigid Transformation is usually used as a precursor to Non Rigid Transformation. This ensures that a global alignment is performed first before focusing on the pixels locally. This may help in countering the unrealistic deformations that can be caused by non rigid registration.

Four Cases of COPDgene dataset were provided. Each case consisted of two 3D CT Lung volumes corresponding to inhale and exhale instances. 300 Pixel indices of the targets were also available. The aim of the project was to minimize the average and standard deviation between the inhale and exhale volumes of the targets after registration. The dataset consisted of varying voxel spacing, dimensions and displacement between the targets before the registration. Figure 1 shows the properties of the dataset that was utilized for training and tuning the registration algorithms.

CASE	DIMENSIONS	Voxel Spacing (mm)	Displacement - Mean (Std)
COPD1	512 x 512 x 121	0.625 x 0.625 x 2.5	25.90 (11.57)
COPD2	512 x 512 x 102	0.645 x 0.645 x2.5	21.77 (6.46)
COPD3	512 x 512 x 126	0.652 x 0.652 x 2.5	12.29 (6.39)
COPD4	512 x 512 x 126	0.59 x 0.59 x 2.5	30.9 (13.49)

Fig. 1: The COPDgene dataset provided for the project

2. PRE-PROCESSING

2.1. Segmentation of the Lungs

After the global Affine lung alignment, we need the registration process to focus on the nodules and fissures inside the lungs.Hence, using masks for the local registration process, helps to focus more on the nodules inside the lungs. The lungs segmentation process comprised of threshold, border removal to guide the seed placement for region growing. The output of region growing was refined by morphological operations.

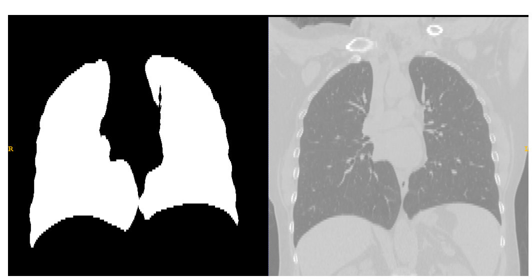

Fig. 2: The Segmentation of the Coronal Slice for COPD1 Lungs Inhale Volume

3. ELASTIX BASED REGISTRATION

In Elastix based implmentation, elastix was used to register the images and transformix was used to apply transformation from the moving image to the fixed image as well as transform the target pixel indices from the fixed image space to the moving image space.

3.1. Environment Setup

ITK-Snap was used to convert **.img** files to **.nii** files. Python was used as a front end in order to automate the elastix and trasformix registration process. The parameter files were populated through python. This automation helped in exploring the right parameter values for the registration. Similarly, registration and transformation were also applied by calling the elastix and transformix commands through python.

3.2. Affine Registration

Affine Registration is a type of global registration that is chosen to be the starting point of the registration process. The global Alignment of the lungs is done using affine registration. The following tuning was done in order ensure efficient and accurate initial global registration:

3.2.1. Masks

We use masks for the registration process for both the fixed and moving images because we don't want the influence of the rib structure on the registration process. The segmentation of the lungs is used as masks to guide this process. Since the boundary of the segmentation is also important, we set the option of **Erode Mask** to false. This ensures that the boundary information is taken into account.

3.2.2. Muti-Resolution Registration

Multiresolution Registration was employed because of the fact that it is not only faster but also helps in convergence. Several types of pyramid structures could be utilized.

1. **Smoothing Image Pyramid**: According to the number of resolutions defined, this option applies gaussian smoothing at each resolution. However, no downsampling is applied. As a result, it is very time consuming.

2. **Recursive Image Pyramid**: This option enables downsampling as well as smoothing to be applied at each resolution. Hence, multi-resolution registration using recursive image pyramid is fast. **This option was chosen as it also gives better results.** This fact is also highlighted by the figure 3

The number of resolutions were fixed at 6 after experimenting. While applying downsampling and smoothing, the pyramid schedule was defined taking into account the fact that the first two dimensions are equal and approximately four times greater than the third dimensions. Hence, all the investigated schedules were applied keeping this ratio in mind. At the end, the Pyramid Schedule was fixed at: **[15 15 3, 8 8 2, 4 4 2, 2 2 1, 1 1 1]**.

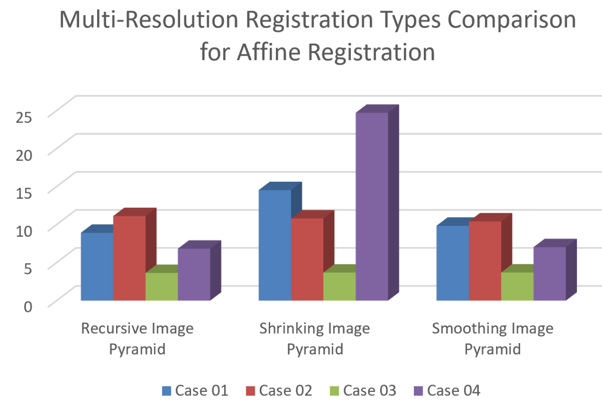

Fig. 3: The figure shows the effect of using different types of Image Pyramids in Multi-Resolution on the first 4 COPD cases in terms of **Mean TRE**.

3.2.3. Metric

After making choice of these basic initialization parameters, the choice of metric for affine registration was explored. Advanced Mattes Mutual Information was the best choice followed by Normalized Mutual Information and Normalized Cross Correlation. A multi metric system based on a weighted combination of Advanced Mattes Mutual Information and Normalized Cross Correlation was also tried. It gave results comparable to using Advanced Mattes Mutual Information Alone. Hence, the choice of metric was fixed as **Advanced Mattes Mutual Information**. The number of bins were fixed as 64.

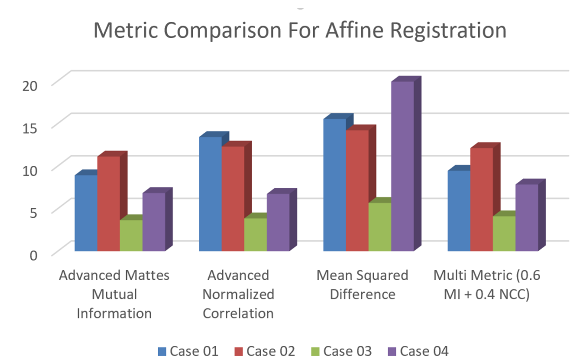

Fig. 4: The figure justifies the choice of Advanced Mattes Mutual Information as a metric during Affine Registration in terms of **Mean TRE**.

3.2.4. Optimizer

- **Standard Gradient Descent:** Elastix implements gradient descent at a considerably high speed due the fact that all the tune-able parameters for this optimizer have to be specified which include the gain factor and the step size. The last most important factor concerned with this optimizer is SP_a whose very high value renders the optimizer unstable causing image distortion and low value cause it to not recover from local minimum. These extra tune able parameters of gain factor and step size introduced additional complexity with inferior performance so this optimizer was dropped in the favour of Adaptive Gradient Descent Optimizer.

- **Adaptive Gradient Descent:** Adaptive Gradient Descent Optimizer estimates the parameters associated with the optimizer automatically. However, this comes at a price of additional computation cost. The time offset was not significant enough to undermine the superior performance of Adaptive Gradient Descent Algorithm. All the other parameters were left to be estimated automatically except SP_a. Tuning SP_a had an effect on the performance of this optimizer and a value of 100 was deemed to be the most appropriate.

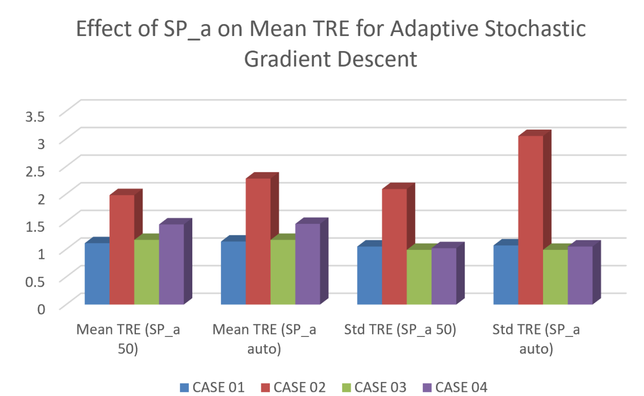

Fig. 5: The figure shows the effect of SP_a parameter of Adaptive Gradient Descent while tuning it for both Affine and BSpline Registration collectively.

3.2.5. Sampler and Interpolator

- **Sampler:** The use of Random Coordinate sampler is recommended with Standard and Adaptive Gradient Descent algorithms. Since, we chose to use Adaptive Gradient Descent, we followed this recommendation of using *Random Coordinate Sampler*.The number of spatial samples to be used per iterations are recommended to be more than 2000. Using more samples comes at the cost of more time consumption. The number of samples for the Random Coordinate sampler for each iteration was tuned to be at *10000*. In addition, it was observed that rather taking samples from the whole image, randomly selecting a voxel at each iteration and selecting samples in a region around its neighborhood worked better. This neighborhood around each voxel was maintained in the ratio of the dimension. It was tuned to be at *150,150,30*.

- **Interpolator:** Although linear and nearest neighbor interpolator take less time for interpolation, we chose to ignore the time constraints due to the superior performance of *BSpline Interpolation of order 3*.

3.3. BSpline Registration

Since the parameters of multi-registration were inherently chosen for affine registration because of the COPD image characteristics, we use the same parameters also for bspline registration. In an attempt to tune these parameters, it was established that previous parameters used for affine registration were the most suitable. Hence, same parameters for multi-resolution registration were used. Similarly, we proceed to use the same optimizer, **Adaptive Gradient Descent** for the bspline registration due to its superior performance. The same optimizer, interpolator and sampler were used for bspline registration as used in affine registration. The number of resolutions was fixed at 6 as in the case for affine registration.

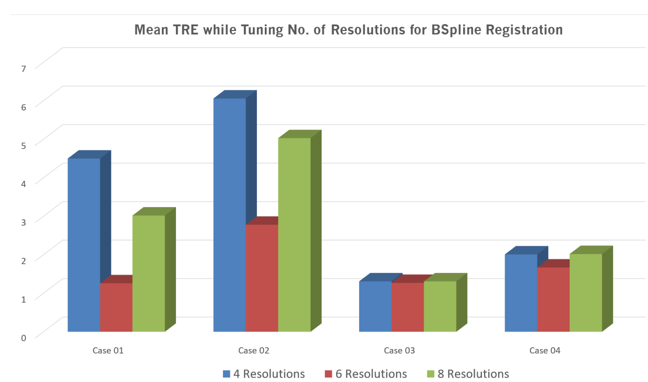

Fig. 6: The figure shows the effect of varying number of resolution for bspline registration while tuning the parameters in terms of Mean TRE.

3.3.1. Grid Spacing

BSpline Registration is defined by a uniform grid of nodes with a specified spacing in between them. This spacing may be defined in terms of voxels or physical units. We chose to use the unit of voxels to express the grid spacing. It's very tricky to choose the correct grid spacing as it involves a compromise between the alignment of small structures and maintaining the image homogeneity. If you make grid spacing too small, it comes at a cost of distortion and extra computational time. The grid spacing was maintained in the ratio of the images dimensions. Since the third dimension is the smallest with the ratio of 1:1:4, the grid spacing was maintained at this ratio. Different spacing were experimented with, however the final grid spacing was tuned at **20:20:4** in voxels.

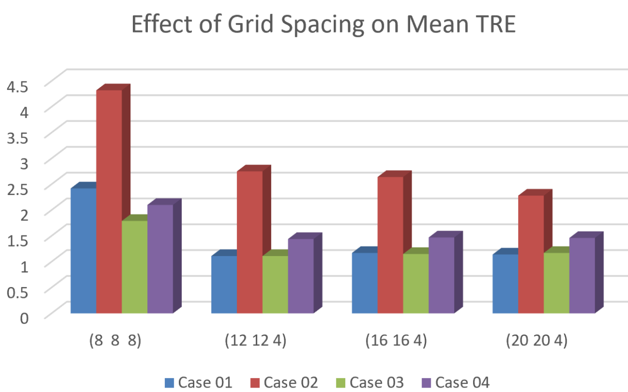

Fig. 7: The figure shows the effect of Grid Spacing on Mean TRE for the first four COPD cases while tuning the parameters.

3.3.2. *Regularization*

Transform Bending Energy regularization was performed. However, this regularization had little to no effect on the results. This can be attributed to the fact that we perform affine registration before bspline for global alignment. The Transform Bending Energy Penalty adds a lot of time for the registration process. Average Registration time for each image goes from **5.4 minutes** to **48 minutes**. Hence, we didn't proceed with the use of regularization for bspline registration.

	Bspline		Bspline + Transform Bending Energy	
	Mean TRE	Std TRE	Mean TRE	Std TRE
Case 01	1.11	1.05	1.31	1.49
Case 02	1.98	2.09	1.82	1.94
Case 03	1.17	0.99	1.26	1.31
Case 04	1.45	1.02	1.39	0.99

Fig. 8: The figure shows the effect of Transform Bending Energy Penalty Regularization on BSpline Registration.

3.4. Results

In order to quantify the final results, we present the target registration error for registered images with and without using segmentation masks in Figure 9. It is to be noted that results using segmentation masks are superior that half of the published results on the challenge website. This approach was tuned from scratch and due to implementation constraints of elastix, further improvement in the results is extremely difficult. The efficiency of our algorithm is further elaborated by Figure 10 and 11.

	With Segmentation Mask		Without Segmentation Mask	
	Mean TRE	Std TRE	Mean TRE	Std TRE
Case 01	1.11	1.05	1.61	2.15
Case 02	1.98	2.09	4.01	5.19
Case 03	1.17	0.99	1.87	2.49
Case 04	1.45	1.02	2.78	4.25

Fig. 9: The table shows that using the masks during the registration process enables better registration of the lung components hence reducing the TRE.

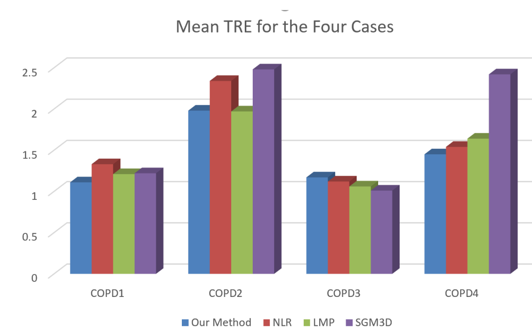

Fig. 10: The figure shows comparison of Mean of TRE for our Method against other published Methods for COPD data registration.

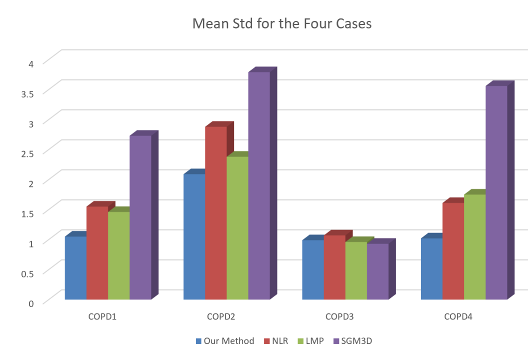

Fig. 11: The figure shows comparison of Standard Deviation of TRE for our Method against other published Methods for COPD data registration.

The time for registration taken by the COPD cases 1,2,3 and 4 according to our tuned parameters on an Intel Core i7 8750H CPU @ 2.20 GHz with 16 GHz of RAM are **4.24 minutes**, **6.2 minutes**, **3.5 minutes** and **5.6 minutes** respectively.

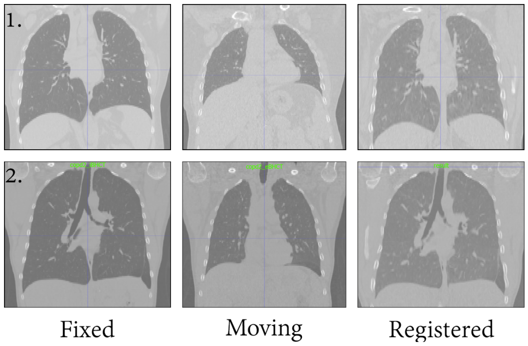

Fig. 12: This figure shows the result of our tuned registration algorithm on the two test images provided on the Challenge Day. The registered image corresponds very closely to the original fixed image qualitatively.

4. VOXELMORPH BASED REGISTRATION

In the contrary to non-learning based registration approaches, learning based approaches using neural networks gained recently the attention of many researchers, and different approaches were proposed for deep learning based medical image registration. VoxelMorph is a deep learning approach that does not require supervised information to perform the registration. It models the registration learning function using a convolutional neural network (CNN) and uses a spatial transform layer to reconstruct one image from another while constraining large displacements in the registration field [3]. Let F and M be a pair of fixed and moving 3D image volumes, VoxelMorph models a function $g_\theta(F, M) = \phi$ using a Unet-like architecture, where ϕ is a registration field and θ are learnable parameters of g. Hence, for each voxel $p \in \omega$, $\phi(p)$ is a location such that $F(p)$ and $M(\phi(p))$ represent the same anatomical landmarks. Figure 13 shows an overview of how VoxelMorph works.

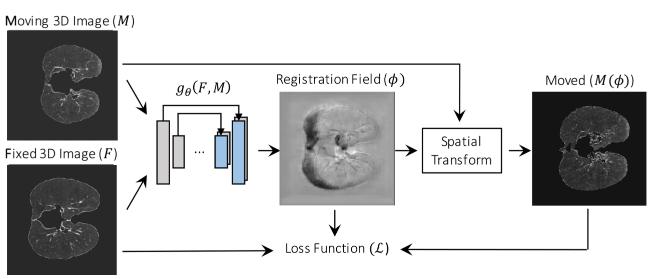

Fig. 13: The adopted VoxelMorph Unet architectures for Lung volume registration.

For VoxelMorph experiments, Keras with TensorFlow were used as the deep learning framework. All experiments were implemented on Google Colab with Tesla K80 GPU (13GB RAM). At first, a custom data generator were implemented that takes as input the fixed image volume (Inhale) and moving image volume (exhale) and yields a tensor concatenating the two volumes to be fed to our VoxelMorph network. VoxelMorph1 architecture was adopted in this work [3] (see Figure 14). The output of the VoxelMorph network is the registration field ϕ which is used to register the moving image with the fixed image by a spatial transformer network. It simply warps and interpolates the moving image to the fixed image space. The learning registration is optimized by a two-parts loss function, one part aims at maximizing a similarity metric between the fixed and the moving, and another smoothing part that tries to regularize the registration field to avoid unreasonable displacement. In our work we tried Mean Squared Error (MSE) and Normalized Cross Corelation (NCC) as a similarity metric.

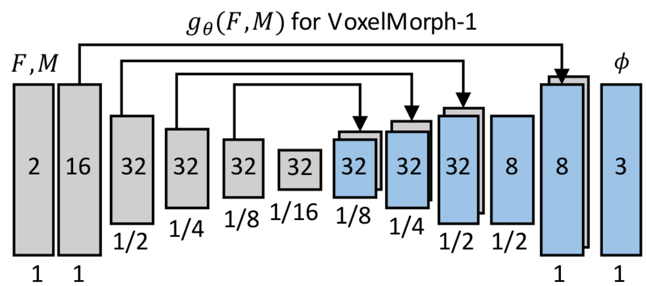

Fig. 14: The adopted VoxelMorph Unet architectures for Lung volume registration.

Two sets of experiments were performed using VoxelMorph: One applies VoxelMorph on the raw, unregistered image volumes, and another set that affinely align the moving to the fixed using Elastix before applying VoxelMorph local non-linear registration.

At start, 10 cases from the 4DCT dataset were added to the 4 COPDgene sample volumes since deep learning approaches need more data. meaning, 11 samples for training and 3 for testing. However, the resolution of the 4DCT dataset images was quite different than that of the COPDgene so adding this new data was rather confusing to the network than being helpful. We eventually only used the first three samples of COPDgene for training and the forth one for testing. All sample volumes were re-sampled to have the same size of 256x256x128 before feeding them to the VoxelMorph network.

Table 15 below shows the mean and standard deviation TRE for the two set of experiments. In the upper half of the table we show the results of performing VoxelMorph directly on unregistered images. Based on the TRE results we see that VoxelMorph reduced the TRE for two volumes (1 and 4) while it increased the TRE for the other two samples (2 and 3). This odd result can be better explained by taking a look at Figure 16.

Registration method	Subject			
	copd1	copd2	copd3	copd4
Unregestered	26.1 (6.4)	21.6 (6.4)	12.6 (6.4)	29.6 (12.9)
VoxelMorph	23.1 (10.8)	22.4 (6.04)	13.8 (6.3)	27.7 (9.9)
VoxelMorph on segmented	27.5 (12.0)	22.1 (6.0)	11.3 (5.8)	26.4 (9.9)
Affine (Elastix)	8.9 (4.4)	11.1 (6.9)	3.6 (2.0)	6.8 (3.6)
Affine + VoxelMorph	9.1 (4.2)	11.4 (7.1)	5.2 (2.6)	7.2 (3.4)
Affine + VoxelMorph on segmented	9.3 (4.4)	10.3 (6.1)	5.11 (2.7)	8.1 (3.8)

Fig. 15: Results of the two sets of experiments: VoxelMorph on unregistered images and VoxelMorph on affinly registered images

Figure 16 shows the fixed image, moving image before registration, moved image after registration, registration field, and moved image superimposed on the fixed image for both experiments, with and without using lungs segmentation masks for case 1. We can see that registration without using the lungs masks results in the network learning to register more the tissue outside the region of interest and therefore leading to sometimes higher TRE. While doing registration with the segmentation masks focus only on the lungs area, so we see the overall shape of the lungs matches, however, the network fails to learn the registration of the fine details of the lungs that compose the landmarks and resulting in a high TRE as well. This could be due to the fact that our dataset is only 4 samples making it hard to force the network to learn how to register the fine details.

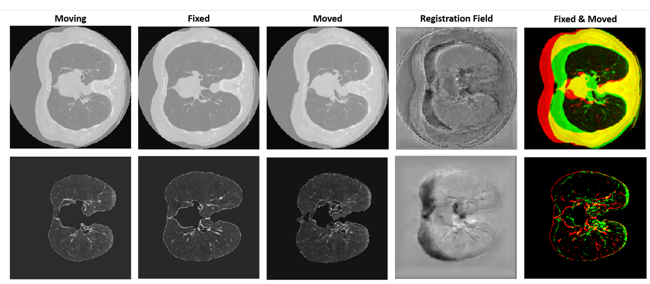

Fig. 16: Comparison between VoxelMorph applied on unmasked images and VoxelMorph applied on mask (segmented) images for case 1

The results of the experiments listed in table 15 is further visually compared in Figure 17 for case number 4, where we can see how affine registration before VoxelMorph helps aligning the lungs together but since voxelMorph has not been able to learn to register the fine details, we see differences in the lung nodules displacement (more clear for the registration applied on segmented images).

5. CONCLUSION

In this project, we exploited the use of traditional elastix approach for COPD lungs registration as well as VoxelMorph registration framework. We were able to achieve results comparable to the state of the art frameworks using Elastix. The

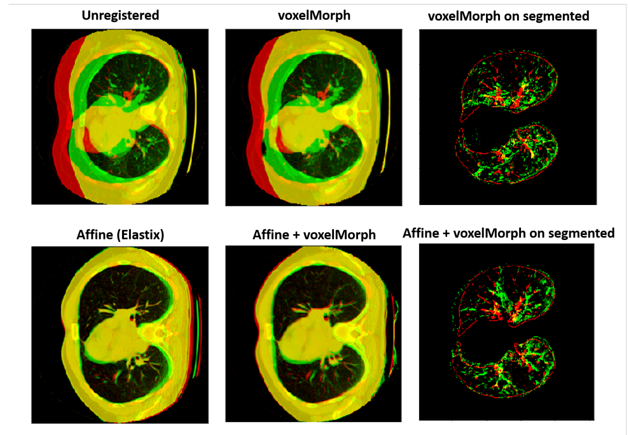

Fig. 17: Comparison between the different experiments in VoxelMorph showing moved image (green) superimposed on the fixed image (red) for case 4

generalization of our approach was further established by our second position in the MIRA Challenge. We also investigated the intricacies of the VoxelMorph Registration framework. Although we were able to achieve encouraging qualitative results with VoxelMorph registration, more data is needed in order to make this registration framework achieve comparable results for COPD lungs registration.

6. REFERENCES

[1] Castillo R, Castillo E, Fuentes D, Ahmad M, Wood AM, Ludwig MS, Guerrero T. 2013. A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive. Phys Med Biol 58 2861-2877.

[2] Elastix website. Accessed 2020-12-14

[3] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, John Guttag, Adrian V. Dalca. *An Unsupervised Learning Model for Deformable Medical Image Registration* CVPR 2018.