Automated Respiratory Pattern Analysis for Dynamic MRI of the Lung of Post COVID-19 patients at 0.55 T

THESIS

Submitted in Partial Fulfillment
of the Requirements for
the Degree of

MASTER OF SCIENCE (Electrical Engineering)

at the

NEW YORK UNIVERSITY
TANDON SCHOOL OF ENGINEERING

by

Prerna Luthra

May 2024

Automated Respiratory Pattern Analysis for Dynamic MRI of the Lung of Post COVID-19 patients at 0.55 T

THESIS

Submitted in Partial Fulfillment
of the Requirements for
the Degree of

MASTER OF SCIENCE (Electrical Engineering)

at the

NEW YORK UNIVERSITY

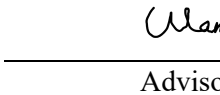
TANDON SCHOOL OF ENGINEERING

by

Prerna Luthra

May 2024

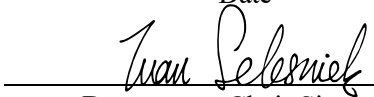
Approved:



Advisor Signature

12/24/2023

Date



Department Chair Signature

Dec 30, 2023

Date

University ID: N14832494
Net ID: pl2243

Approved by the Guidance Committee:

Major: Electrical Engineering

Yao Wang
Professor
NYU Tandon School of Engineering
12/24/2023
Date

Li Feng
Associate Professor
NYU Grossman School of Medicine
12/29/2023
Date

Chinmay Hegde
Associate Professor
NYU Tandon School of Engineering
12/29/2023
Date

Vita

Prerna Luthra was born in Omaha, NE, USA. She spent her childhood in India before coming back to the US to pursue her college education. In May 2016, she earned a Master of Science in Computer Science and a Bachelor of Science in Computer & Systems Engineering and Electrical Engineering from Rensselaer Polytechnic Institute, Troy, NY. Later, she started working as a software engineer at IBM. Her MS in Electrical Engineering at NYU began in September 2021. She spent March 2023 to December 2023 working on this thesis at NYU Video Lab and CAI^{2}R, NYU Grossman School of Medicine.

Acknowledgement

I extend my sincere gratitude to Prof. Yao Wang and Prof. Li Feng for their unwavering support and encouragement over the past couple of months. Their guidance has been instrumental in shaping both the research and writing aspects of this thesis, and I am truly appreciative of all the valuable insights they have shared.

Special gratitude is extended to Prof. Chinmay Hegde for generously serving on the thesis committee.

I also want to extend my appreciation to my lab colleague, Haoyang Pei, who not only assisted me with the research but also played a key role in helping me write various abstracts. He has consistently been available whenever I needed help, and his support and encouragement were instrumental in keeping me motivated during challenging times.

I am also thankful to Prof. Henry Rusinek, Artem Mikheev, Dr. Hersh Chandarana, Prof. Mehmet Kurt, and Tianyi Ren, for taking the time to meet and discuss various algorithms and techniques used in this project.

I am grateful to friends and lab members I encountered at CAI^{2}R, including Jingjia Chen, Lavanya Umapathy, and Nikola Janjusevic, who helped me with brainstorming solutions and reading research papers.

Lastly, I extend my deepest appreciation to my mom for her unwavering love, care, and constant encouragement, which played a crucial role in pushing me to successfully complete this thesis.

ABSTRACT

Automated Respiratory Pattern Analysis for Dynamic MRI of the Lung of Post COVID-19 patients at 0.55 T

by

Prerna Luthra

Advisor: Prof. Yao Wang, Ph.D.

Submitted in Partial Fulfillment of the Requirements for

the Degree of Master of Science (Electrical Engineering)

May 2024

Analyzing respiratory motion patterns in proton MRI for patients with lung diseases, including those with post-COVID-19 symptoms, poses a significant challenge, particularly in the context of non-invasive approaches. This study investigates the feasibility of classifying post-COVID-19 patients into categories of either long COVID-19 or no symptoms through the analysis of motion fields within diverse regions of lung MRIs. Further, it introduces an automated framework designed to assist radiologists in swiftly discerning not only the presence but also the severity of long COVID-19. Importantly, the versatility of this framework can be extended to address diverse lung diseases.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 Overview

After recovering from COVID-19, some patients continue to experience lingering symptoms, including dyspnea, a condition commonly associated with long COVID-19. Patients with long COVID-19 exhibit different breathing patterns, which is valuable for analyzing the severity of the long COVID-19 patients.

A study [1-2] previously examined the breathing patterns of long COVID-19 patients by introducing the concept of *incoherence* to quantitatively measure the periodicity of the change of lung's surface area. These incoherence values are then used to categorize these patients into several groups with different severity of symptoms. However, the area based breathing pattern analysis is not sufficient to capture the regional breathing pattern as it requires manual segmentation of lungs to different regions, which presents a barrier for the clinical settings.

In this study, we introduce an automated framework for automatic analysis of localized respiratory motion in lung MRI to allow efficient automated segmentation of the lungs into different regions while enabling localized breathing patterns analysis based on the incoherence calculated from the motion field.

1.2 Outline

The thesis is structured as follows:

Chapter 2 provides a summary of the dataset.

Chapter 3 presents an overview of the proposed automated pipeline for respiratory motion analysis.

Chapter 4 delves into two distinct lung segmentation techniques and their respective results.

Chapter 5 explores motion field estimation and regional segmentation of lung masks.

Chapter 6 focuses on quantitative analysis through incoherence computation.

Chapter 7 covers the discussion of all incoherence box plots and statistical analyses for various groups.

Chapter 8 presents additional results related to incoherence visualization.

Finally, Chapter 9 examines the potential use of machine learning to classify patients as having long COVID or no symptoms.

2. Dataset Summary

A total of 39 dynamic 2D lung MRI datasets acquired on a ramped down 0.55T MRI scanner (Aera, Siemens) were used. Each dataset has 250 motion frames with a matrix size of 256 x 256 or 92 x 112. For the sake of consistency and easier processing, these frames were either downscaled using nearest neighbor interpolation or upscaled by zero padding to a constant size of 128 x 128. It's important to note that ground truth lung masks were not available for any of the patients; only motion frames were accessible for analysis.

39 patients were divided into 3 groups: Group 0 consisted of 11 patients with no symptom, Group 1 consisted of 14 with mild long COVID-19 symptoms and Group 2 consisted of 14 patients with severe long COVID-19 symptoms.

Patient Severity	Total Patients
Group 0: No Symptom	11
Group 1: Mild	14
Group 2: Severe	14

Table 2.1: Data Summary

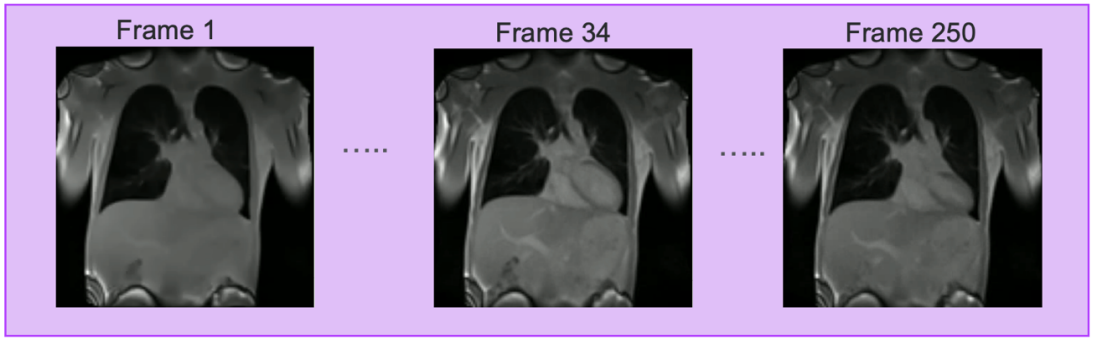


Figure 2.1: Visualization of Motion Frames for a patient

3. Pipeline Overview

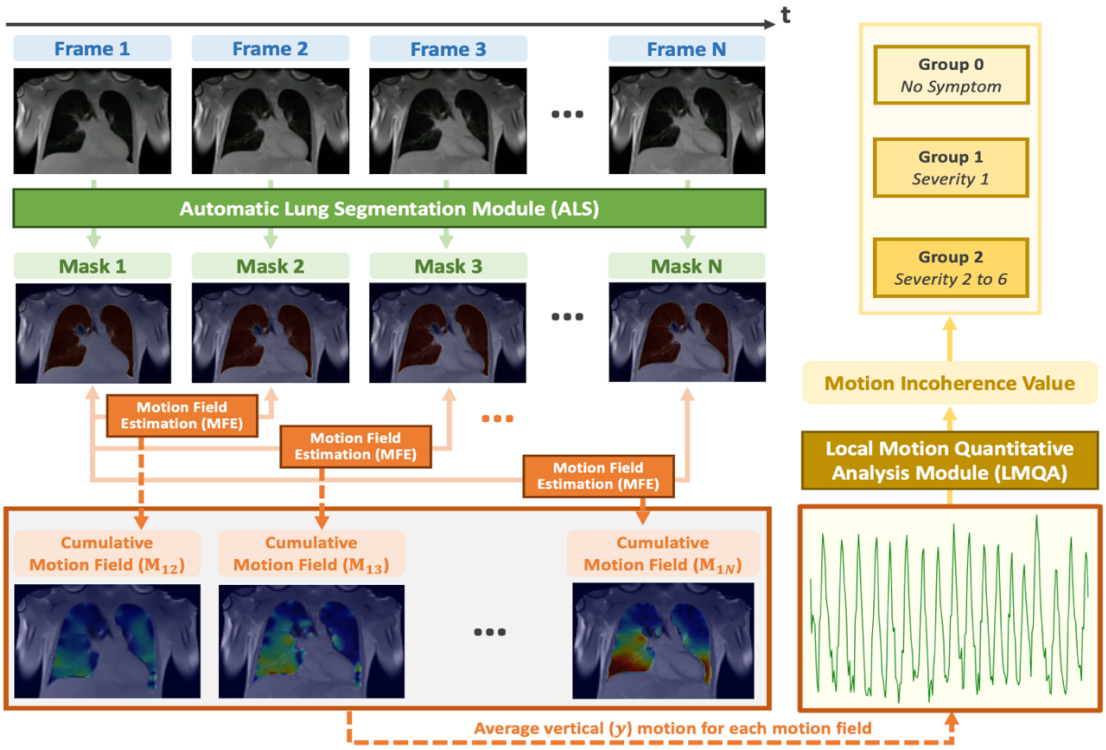


Figure 3.1: Automated Pipeline

The proposed framework comprises of the Automatic Lung Segmentation module (ALS), the Motion Field Estimation module (MFE), and the Local Motion Quantitative Analysis module (LMQA) as shown in Figure 3.1.

ALS utilized two distinct approaches for lung segmentation. In the first approach, a pre-trained deep learning model named UniverSeg [3] was employed to perform segmentation tasks on the dataset. This model utilized a limited number of annotated images and labels to accomplish the segmentation. Additionally, ALS employed Firevoxel [5] segmentation as a comparative method. Firevoxel is an algorithm that performs lung mask segmentation using thresholding techniques, providing an alternative to the deep learning model for the segmentation process.

In the MFE module, the Symmetric Diffeomorphism [4] technique was employed to estimate the mapping or transformation matrices between the reference frame (e.g., frame 1) and subsequent frames (e.g., 2, 3…). These mappings were then utilized to generate motion curves for different regions of the lung. To ensure consistent segmentation of the lung into upper and lower parts across frames, a line in the middle of the reference mask was estimated to divide the lung into two parts. The endpoints of this line were then tracked across frames by transforming the coordinates of the reference frame's endpoints to the destination frame using the corresponding mapping. The vertical (or Y) motion curve for different segmented regions is determined by calculating the average displacement of all Y-coordinate pixels between the reference frame and frame N.

In the LMQA module, the incoherence for average Y motion curve specific to a particular region of the lung is calculated. The resulting incoherence value facilitates the classification of patients into three distinct groups based on the severity levels of their condition.

4. Automatic Lung Segmentation Module

4.1 Deep Learning Based Approach: UniverSeg Model

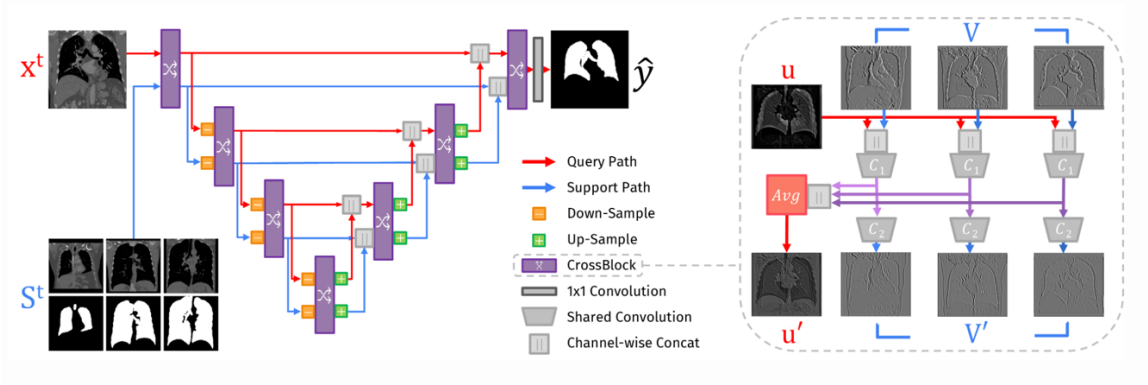


Figure 4.1: UniverSeg Model Architecture

Deep learning is a prevalent approach for organ segmentation in medical imaging, typically requiring a substantial amount of labeled ground truth data for training. In our unique situation, where labeled ground truth was not available, we had to resort to an alternative solution. Consequently, we opted to use a pre-trained model to address the segmentation challenge.

The UniverSeg model [3] is a pre-trained universal medical image segmentation model designed for organ segmentation tasks across various modalities such as CT, X-Ray, and MRI, covering multiple organs. Illustrated in Figure 4.1, the UniverSeg network (left) takes a query image along with a support set comprising image and label-maps (concatenated in the channel dimension) as input. It employs multi-scale CrossBlock features for segmentation. A CrossBlock (right) takes as input representations of the query u and support set V, and interacts u with each support entry to produce u' and V'.

In our specific application, each 2D lung frame is fed into the UniverSeg model, generating

lung segmentation masks for each frame. The model was trained using 17 2D lung MR support images and labels, each sized 128 x 128, which were distinct from the MR frames used for the final analysis of 39 patients. Manual segmentation of support labels was performed using Matrix User 2.2 [6].

The model was then tested on all 39 test patients (250 2D MR frames per patient). The model gave a soft prediction for these patients. These soft predictions were then clipped to a value of either 0 or 1 based on a certain threshold. These predictions were found to have some missing regions (or holes). These holes were filled by finding contours. The segmented lung masks were then used for motion estimation.

4.2 Algorithmic Approach: Firevoxel

The firevoxel [5] algorithm was introduced by Mikheev et al. It is a segmentation algorithm primarily developed for 4D lung segmentation. Setting certain initial parameters, such the radius, is necessary for this algorithm to work properly. The algorithm first applies an adaptive thresholding by setting radius. It then performs maximum component analysis on the binary mask to find connected components.

A BodyMask(t) for at a particular time 't' is generated using the above procedure. LungMask(t) is then constructed by applying a hole filling operator to BodyMask(t). Two largest connected components from LungMask(t) are extracted and all other components are removed. The BodyMask(t) is then updated by using following rule: BodyMask(t) := BodyMask(t) AND NOT LungMask(t).

The average signals, Slung and Sbody, are measured from the two updated binary masks LungMask(t) and BodyMask(t). The lung upper threshold is then calculated as (SLung+SBody)/2, and the auxiliary lung mask Aux(t) is extracted from gray-scale images as voxels with signal intensity below this threshold. The lung masks are augmented using set union LungMask(t) AND Aux(t). Finally, the mask is decomposed into two distinct regions for the left and the right lung using a custom wavefront propagation algorithm that tracks the ancestors of each voxel of the wavefront.

4.3 Discussion of Segmentation Results

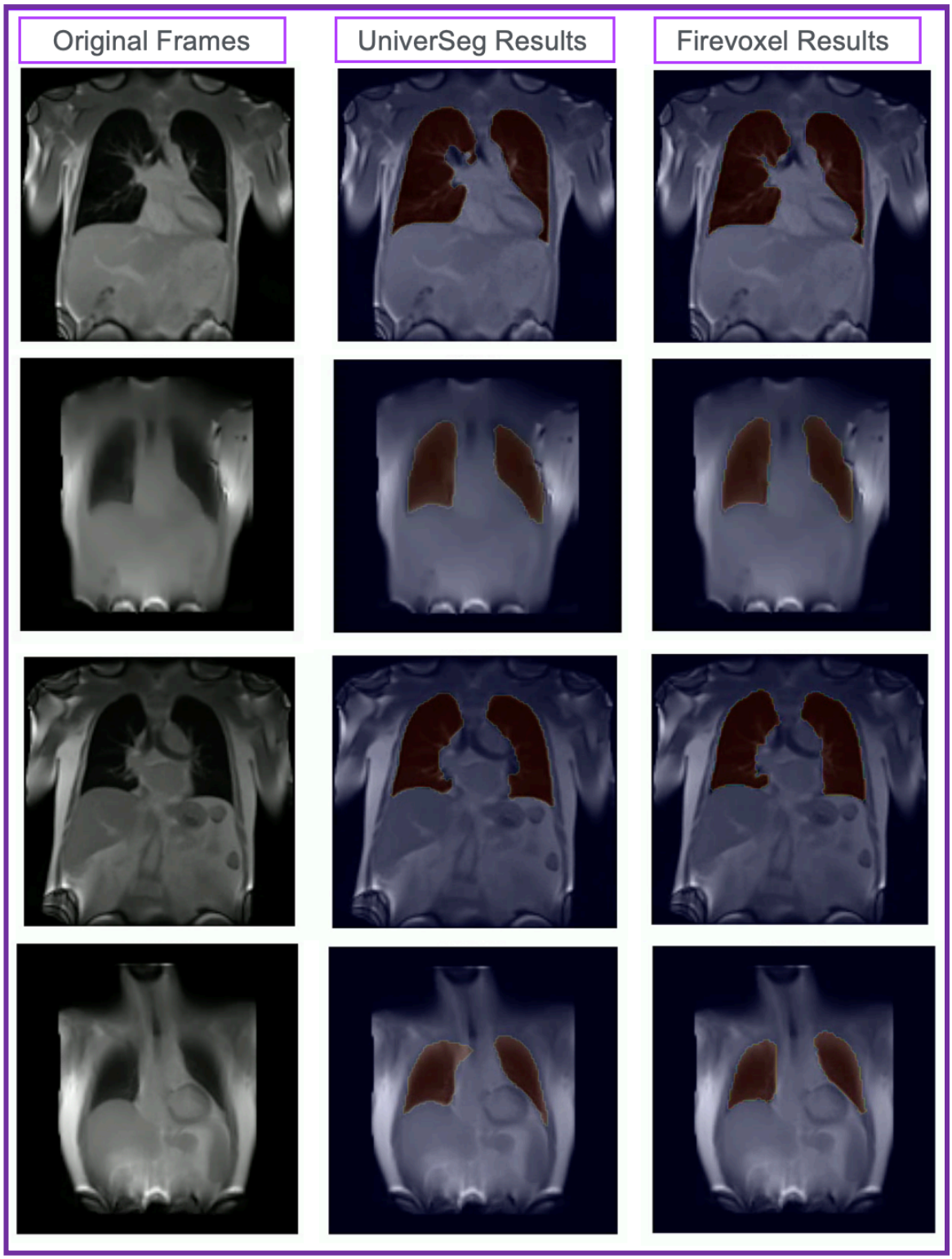


Figure 4.2: Good Segmentation Results for UniverSeg and Firevoxel

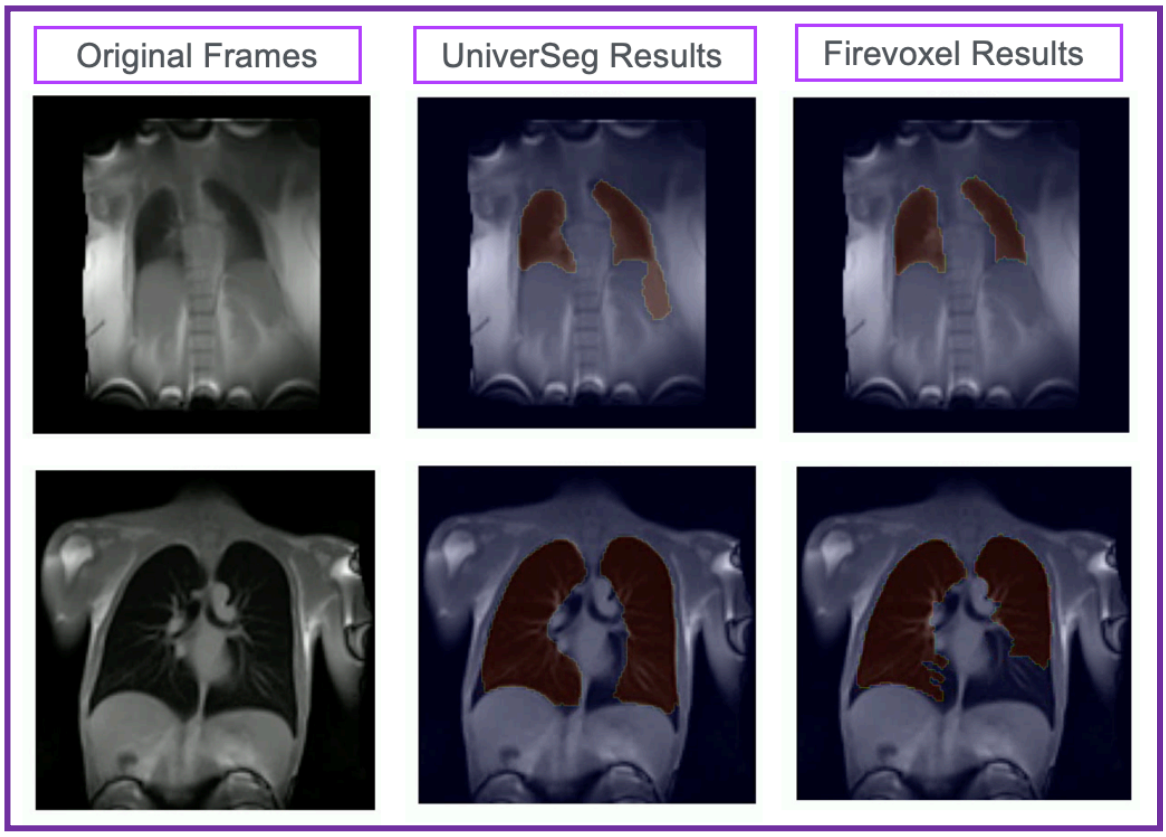


Figure 4.3: Bad Segmentation Results for UniverSeg and Firevoxel

The inability to get ground truth masks made it difficult to calculate DICE scores for the two methods. For every patient, however, a video was generated with masks superimposed on top. Upon manually examining the 39 videos for every patient, it was noted that the UniverSeg model did a decent job of segmenting the right lung. Firevoxel worked well on both the left and right lungs, but it is noted that when the lung region is rather vast in comparison to the entire body, it is unable to segment the lungs correctly. Figure 4.2 shows some of the good segmentation results for both methodologies. Figure 4.3 shows some poor results.

5. Motion Field Estimation Module

5.1 Symmetric Diffeomorphism

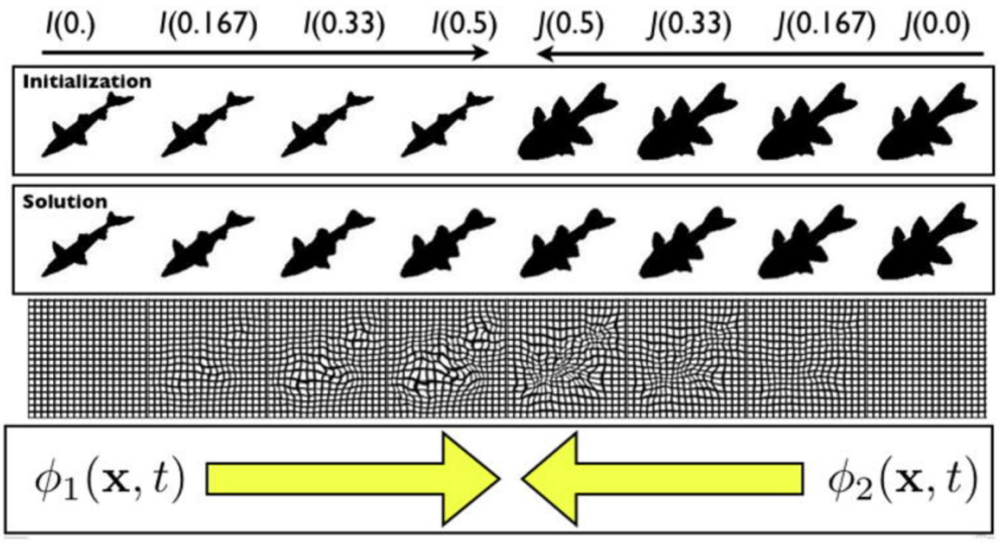


Figure 5.1: Symmetric Diffeomorphism

Symmetric diffeomorphic [4] registration is a technique used in medical imaging to align and match images of the same subject taken at different times.

It solves the following image matching problem: *Find a spatiotemporal mapping*, φ, *such that the cross correlation (a measure of similarity) between the image pair is maximized*.

$$\varphi_1(x,t)I = \varphi_2(z,1-t)J$$

In this equation, *I* and *J* are images that need to be matched, *φ*1 and *φ*2 are grid mappings, *x* and *z* are spatial coordinates and, *t* is time. Figure 5.1 illustrates the how images I and J converge to form grid mappings φ_1 and φ_2.

To determine the mapping between two segmented lung masks, we employed symmetric diffeomorphic registration. In particular, the mask taken from frame 1 was used as the reference frame for each patient. This algorithm was then used to align the masks from all

subsequent frames to match the mask from the first frame. Aligning a pair of masks helps in tracking the respiratory motion of the lungs.

5.2 Warping Results from Symmetric Diffeomorphism

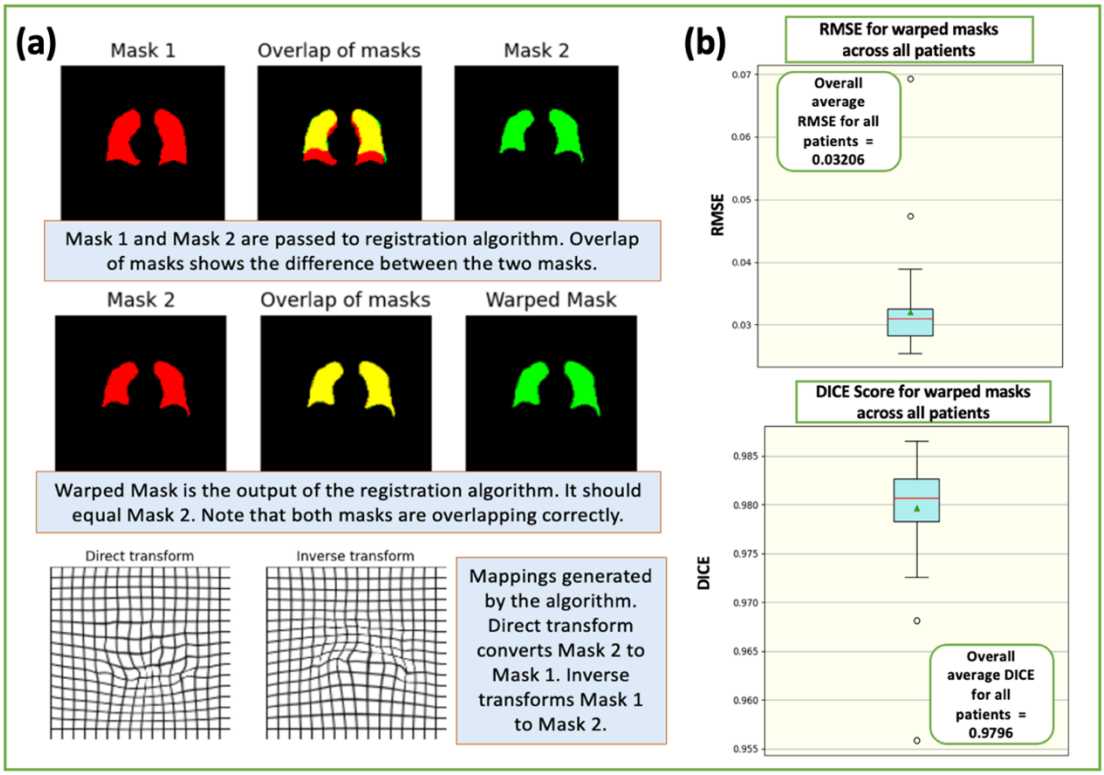


Figure 5.2: (a) Sample warping result for a lung mask, (b) Overall DICE and RMSE score for warped masks

Lung masks 2 through 250 were aligned to lung mask 1 for a specific patient in order to assess the efficacy of the symmetric diffeomorphism algorithm. To do this, a mapping between each pair of lung masks (φ_{1_2}, φ_{1_3}, φ_{1_4} .. φ_{1_250}) was generated. Next, masks 2 to 250 were transformed to mask 1 using these mappings.

For instance, *Mask30*(* φ_{1_30}*) = Warped Mask1*. Here, mapping φ_{1_30} is used to change Mask30 into Mask1. Note *Warped Mask 1* is the result generated by applying this mapping.

This mask should ideally equal *Mask 1*. Figure 5.2 (a) gives an example of generated warped mask as well as corresponding mapping or grid transformations.

Following the generation of the warped masks, the DICE score between two masks is calculated. The DICE score is given by 2 times the total number of pixels that overlap between Mask 1 and Mask N divided by total number of pixels in both masks. Note that the DICE score for a given patient is calculated as the average of all DICE scores across each pair of aligned masks. The DICE scores of the 39 patients vary from 0.97 to 0.99, with an average of 0.97988, as indicated in Figure 5.2 (b). In addition to DICE scores, the Root Mean Squared Error (RMSE) is computed by comparing the pixel values at corresponding locations in the warped mask and reference mask. The average RMSE across all patients and frames is determined to be 0.0306.

These results show the outstanding performance of symmetric diffeomorphism, demonstrating its effectiveness in achieving accurate and precise alignment between masks.

5.3 Motion Detection Using Symmetric Diffeomorphism

In motion detection workflow, Mask 1 serves as the reference frame for subsequent analysis. To establish a connection between Mask 1 and other masks (ranging from 2 to 250), a warped mapping is generated using symmetric diffeomorphism. This results in a series of mappings denoted as φ_{1_2}, φ_{1_3}, φ_{1_4} .. φ_{1_N}. Notably, φ_{1_N} represents a mapping or a transformation grid that gives us new coordinate locations for mask N so that it aligns with mask 1.

Our motion calculation methodology focuses predominantly on the Y coordinates, gauging the extent of their displacement from one image to the next. The Y Motion Displacement is specifically computed between Mask 1 and Mask N. This involves obtaining values such as: M_{12}, M_{13}, .. , M_{1N} where M_{1N} signifies the average Y motion or the average number of Y pixels moved (or displaced) between mask N and mask 1.

The dataset comprising of M_{12}, M_{13}, .. , M_{1N} is utilized to generate a curve. This curve represents the respiratory motion. Subsequently, the periodicity or incoherence of this curve is calculated, forming the basis for patient classification in our analysis.

5.4 Regional Segmentation: Splitting Lungs into left and right sides

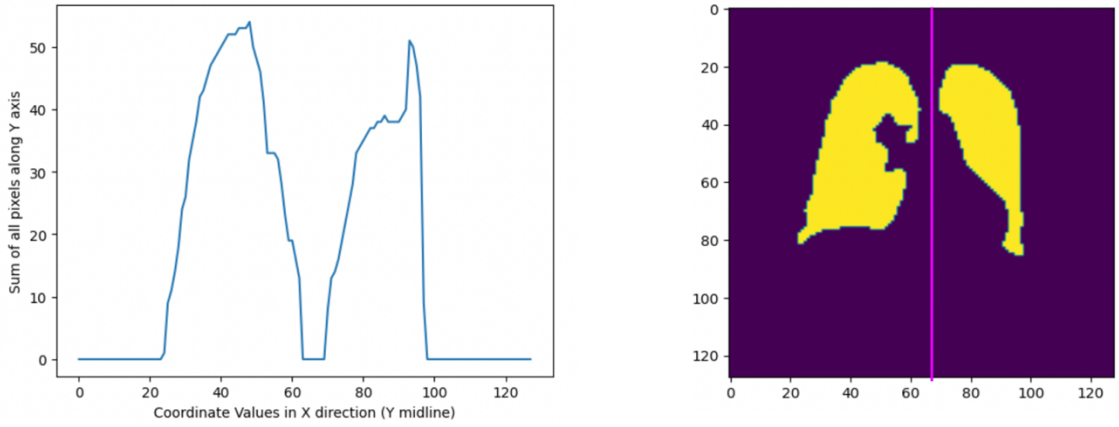


Figure 5.4: (left) Sum of pixels along Y axis, (right) Detected midpoint that separates left and right lung

To split the lungs into left and right, the algorithm follows a series of steps:

First, the midline along the Y-axis is found. This is done by summing the pixel values along the Y direction in the lung mask. Following that, within the array of summed pixels, the algorithm commences its search from the left side of the array and identifies the initial point where the sum becomes non-zero. The algorithm then locates the next point where the sum of pixels returns to zero. This point at which the sum of pixels is zero is the location of the midline or the line that splits the lungs into left and right.

5.5 Regional Segmentation: Splitting Lungs into upper and bottom region

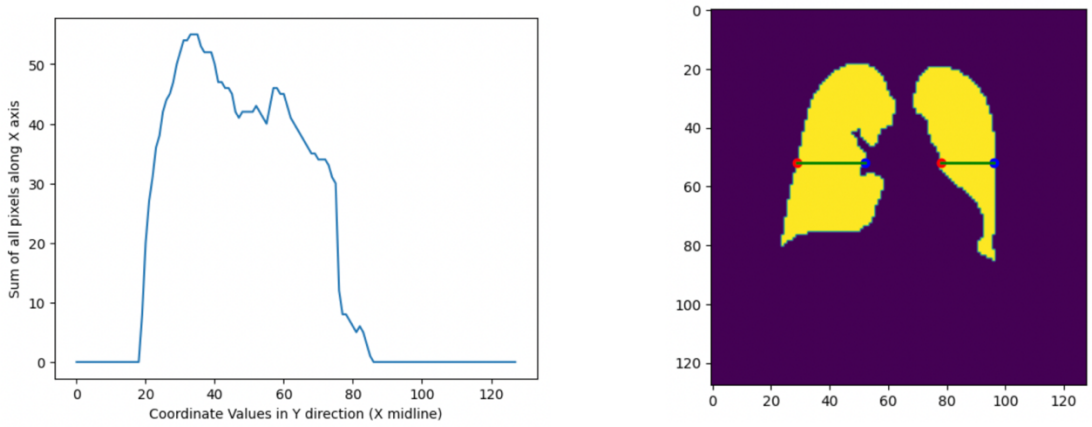


Figure 5.5: (left) Sum of pixels along X axis, (right) Detected midpoint that separates top and bottom lung

5.5.1 Method 1: Random Midline Split

To partition the lungs into top and bottom segments, the algorithm executes the following steps to find the midline along the X-axis:

First, all pixel values along the X direction are summed as illustrated in Figure 5.5 (left). Following that, within the array of summed pixels, the algorithm identifies the starting and ending points of non-zero values.

The midpoint between the determined start and end points serves as the division point for separating the top and bottom sections of the lungs. This approach is iteratively applied to all frames pertaining to a particular patient.

It is essential to recognize that due to the dynamic nature of the frames, characterized by upward and downward movements caused by respiratory motion, the X midline does not remain constant across all frames. In other words, the position of the midline undergoes changes relative to the first frame. The inconsistency in midline tracking using this method can pose challenges in monitoring respiratory motion, especially when the lungs are

unevenly split. Hence, it is crucial to adopt an approach that ensures a balanced division of the lungs across all frames.

5.5.2 Method 2: Two Point Midline Tracking

To ensure an even division of the lungs across all frames, the initial and terminal points of the midline are identified in frame 1 using method 1.

Subsequently, these two points are systematically tracked across successive frames through the application of Symmetric Diffeomorphism. In this process, warped mappings denoted as φ_{1_2}, φ_{1_3}, φ_{1_4} .. φ_{1_N} are generated where φ_{1_N} is a mapping or a transformation grid, that establishes new coordinate locations for mask N and aligns it with mask 1.

Here, φ_{1_N} is utilized to determine the updated coordinates for the start and end points of the midline, facilitating the identification of corresponding values in Mask N.

6. Local Motion Quantitative Analysis Module

6.1 Incoherence Computation

Incoherence serves as a metric for periodicity, indicating how far a curve is from being periodic. This measure of periodicity is particularly valuable in assessing the periodic nature of respiratory motion curves and is employed to differentiate between severity levels among long COVID-19 patients. Following equation defines incoherence:

$$Inc(t) = \frac{\sum_{i=0}^{Nsteps} | MF(i * step) - MF(i * step + t)|}{Nsteps * avg(MF)}$$

Please note, here, *MF* is motion field in vertical or Y direction for a specific region of the lung. *Nsteps* is total number of frames (i.e. 250). *step* is the step size which is set to 1. *t* is the time period of motion curve. *avg(MF)* is the average of motion field.

't' is determined through the following process:

An exhaustive search is conducted for the value of the inherent curve period, denoted as 'T' (e.g. ranging from 0 to 130 with a step of 1). For each assessed value of *t* within the interval [0, 130], an *Inc(t)* calculation is executed. The time *t* at which *Inc(t)* is minimum is chosen to be the inherent curve period *T*.

6.2 Assessing long COVID-19 Severity through Incoherence Measurement

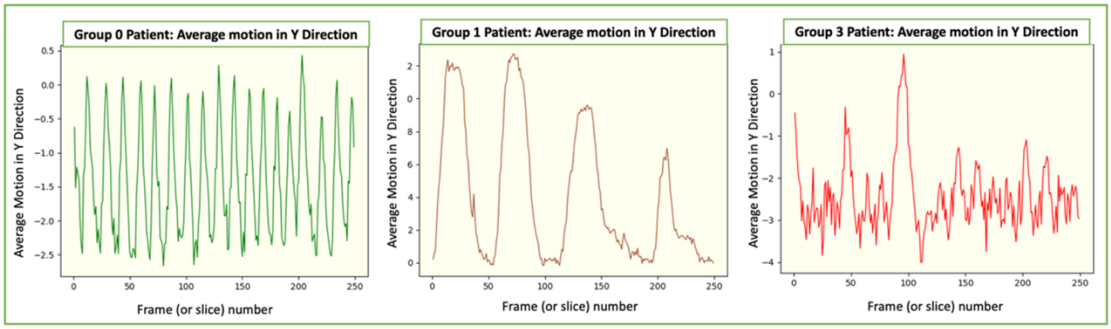


Figure 6.1: Comparison of respiratory Y motion curve for patients from different groups

Figure 6.1 above depicts the vertical (Y-axis) respiratory motion among patients belonging to three distinct groups. Group 0 with no symptoms, Group 1 includes those with mild symptoms, and Group 2 consists of patients with severe symptoms. Examining the figure reveals that the motion curve for patient from Group 0 exhibits a fairly periodic pattern, Group 1 displays a slightly irregular pattern, and Group 2 demonstrates a highly irregular pattern. This pattern is further confirmed by the corresponding incoherence values; specifically, the incoherence values increase with the level of irregularity. Therefore, incoherence serves as a valuable measure to assess the severity of long COVID-19 in patients. Elevated incoherence values signify an increased irregularity in respiratory motion. As a result, higher incoherence values suggest a heightened risk of the patient encountering severe long COVID-19.

7. Discussion of Incoherence Plots and Statistical Analysis

7.1 Introduction

The incoherence for vertical (or Y) motion is computed for various regions of the segmented lungs for patients in all three groups.

Further, statistical significance test or the t-test is done to compare the means of two groups and determine if there is a significant difference between them. The t-test calculates a t-statistic based on the sample data and its standard error, considering the means and variances of the two groups. It then assesses the p-value associated with the test statistic. If the p-value is below a predetermined significance level (commonly 0.05), it indicates that the observed difference is unlikely to have occurred by chance, and the null hypothesis of no difference between the groups is rejected.

7.2 Comparison between Y Motion Incoherence computed using UniverSeg and Firevoxel based right lung segmentations

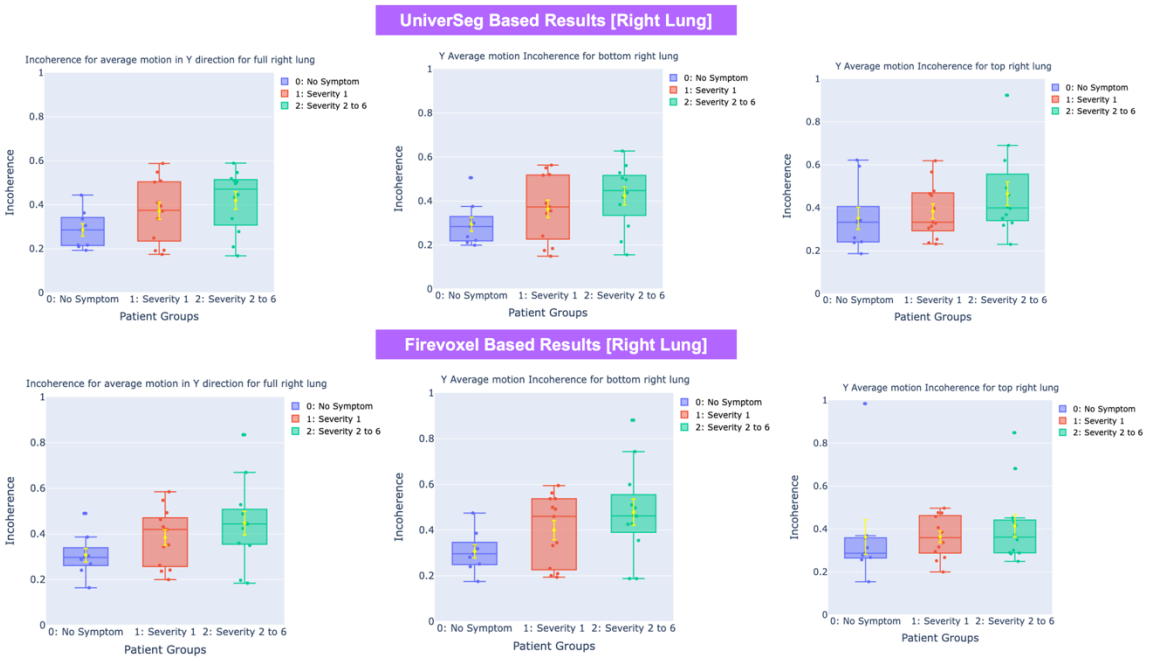


Figure 7.2: Comparison of Y motion incoherence for full, bottom and top right lung computed using UniverSeg (top row) and Firevoxel (bottom row) segmentation masks.

	Model Type	T-Statistic	P-Value	Is Significant ?
Group 0 vs. Group 1	USeg R Area	0.541	0.593	no
	USeg R Full Y	-1.809	0.0856	no
	USeg Bottom Y	-1.351	0.191	no
	USeg R Top Y	-0.5231	0.608	no
Group 0 vs. Group 2	USeg R Area	-2.014	0.0607	no
	USeg R Full Y	-2.734	0.0134	**yes**
	USeg Bottom Y	-2.452	0.024	**yes**
	USeg R Top Y	-1.497	0.1507	no
Group 1 vs. Group 2	USeg R Area	1.493	0.1509	no
	USeg R Full Y	0.844	0.407	no
	USeg Bottom Y	-1.2380	0.321	no
	USeg R Top Y	0.2311	0.2311	no

Table 7.2.1: Statistical Significance Results for UniverSeg based right lung segmentation

	Model Type	T-Statistic	P-Value	Is Significant ?
Group 0 vs. Group 1	FV R Area	-0.942	0.3575	no
	FV R Full Y	-1.686	0.1074	no
	FV R Bottom Y	-1.821	0.0839	no
	FV R Top Y	-0.009	0.9929	no
Group 0 vs. Group 2	FV R Area	-1.891	0.07510	no
	FV R Full Y	-2.444	0.0244	**yes**
	FV R Bottom Y	-2.784	0.0123	**yes**
	FV R Top Y	-0.5821	0.5700	no
Group 1 vs. Group 2	FV R Area	-1.1079	0.2810	no
	FV R Full Y	-1.043	0.308	no
	FV R Bottom Y	-1.109	0.278	no
	FV R Top Y	-0.980	0.339	no

Table 7.2.2: Statistical Significance Results for Firevoxel based right lung segmentation

The average Y motion for the right lung masks segmented using both UniverSeg and Firevoxel was calculated. Subsequently, incoherence was determined for each patient based on the average Y motion within a specific region of the lung. Figure 7.2 presents a comparative analysis through box plots, focusing on Y motion incoherence across the full, bottom, and top regions of the right lung segmented using UniverSeg and Firevoxel. Table 7.2.1 and 7.2.2 presents the corresponding statistical analysis. Please note that the comparison is exclusively conducted on the right lung mask since UniverSeg's segmentation performance is suboptimal for the left lung.

Examining both the plots and the results of the statistical significance test, we observe that the computed incoherence for the full and bottom right lung regions, using both UniverSeg and Firevoxel, effectively distinguishes between Group 0 and Group 2 patients.

However, for distinguishing between Group 0 and Group 1, the incoherence calculated for the entire and bottom right lung regions using both segmentation methods is not deemed significant, as indicated by p-values greater than 0.05. If a slightly more lenient threshold of 0.1 was employed, a discernible difference in incoherence for the entire and bottom right lung regions between Groups 0 and Group 1 might have been identified.

It is also crucial to note that the incoherence computed for the top right lung region does not exhibit statistical significance for any of the group comparisons using both segmentation methods.

Finally, it is essential to highlight that there are differences in the incoherence values when comparing masks segmented using UniverSeg and Firevoxel. Despite both methods exhibiting satisfactory performance in segmenting right lung masks, it is crucial to recognize that the accuracy of incoherence computation is contingent upon the quality of

lung segmentation. The effectiveness of the segmentation directly influences the accuracy of the derived incoherence values.

7.3 Comparison between Incoherence for Area and Y Motion using both UniverSeg and Firevoxel based right lung segmentations

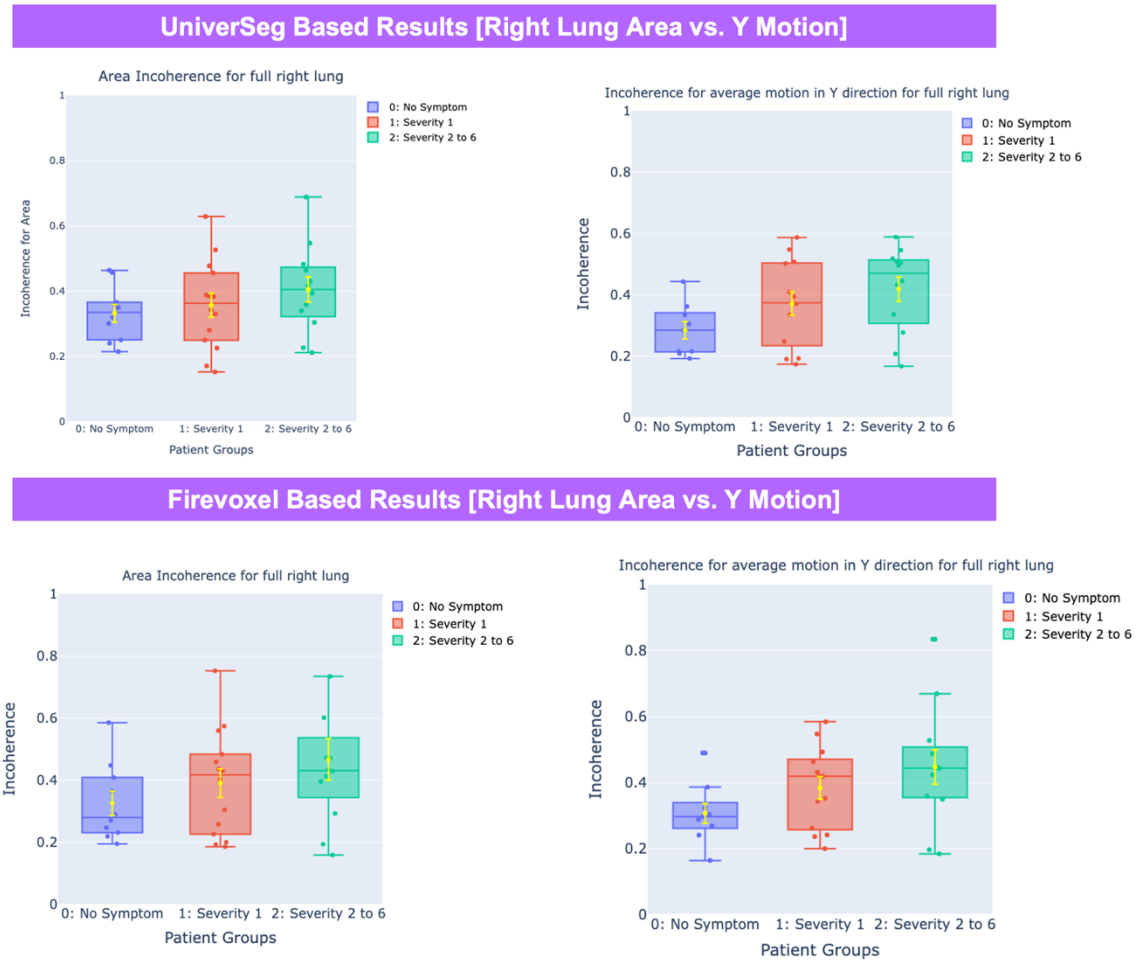


Figure 7.3: Comparison of Y motion and area incoherence for full right lung computed using UniverSeg (top row) and Firevoxel (bottom row) segmentation masks

Previous studies have demonstrated the feasibility of using lung surface area to calculate respiratory breathing pattern. In this section, we undertake a comparison between incoherence values derived from lung's surface area and those computed using average Y motion. Further, this comparison was done using the masks generated using both UniverSeg and Firevoxel. It is important to note that the area is simply the sum of all pixels within a specific lung region, in this case, the entire right lung mask.

It should be acknowledged that a regional analysis, particularly for the top and bottom regions, is not conducted using the area-based approach. This omission is attributed to the inherent challenges associated with splitting the lungs into top and bottom regions based on area. The difficulty arises from the fact that a greater total number of pixels in a region results in a smoother respiratory motion curve. Consequently, an uneven split, for instance, the bottom region having fewer pixels, leads to a more noisy curve. The noise in the curve, in turn, contributes to larger incoherence values, emphasizing the impact of uneven pixel distribution on the overall noisiness of the generated respiratory motion curve.

Figure 7.3 provides a comparative examination using box plots, specifically focusing on incoherence derived from area and Y motion across the entire right lung, segmented using UniverSeg and Firevoxel. The corresponding statistical analysis is presented in Figures 7.2.1 and 7.2.2 in the preceding section.

Upon scrutiny of both the plots and the outcomes of the statistical significance test, it becomes evident that, unlike Y motion incoherence, as elucidated in the previous section, the computed incoherence for area is not deemed significant in any of the three group comparisons using either UniverSeg or FireVoxel segmentations. This is indicated by p-values exceeding 0.05.

It is worth noting that if a more lenient threshold of 0.1 were applied, the area-based incoherence of the full right lung when comparing Group 0 and Group 2 might have reached significance. However, it is crucial to emphasize that Y motion serves as a more robust marker compared to area, as reflected by smaller p-values.

7.4 Comparison between Left and Right Lung Y Motion Incoherence using Firevoxel based lung segmentations

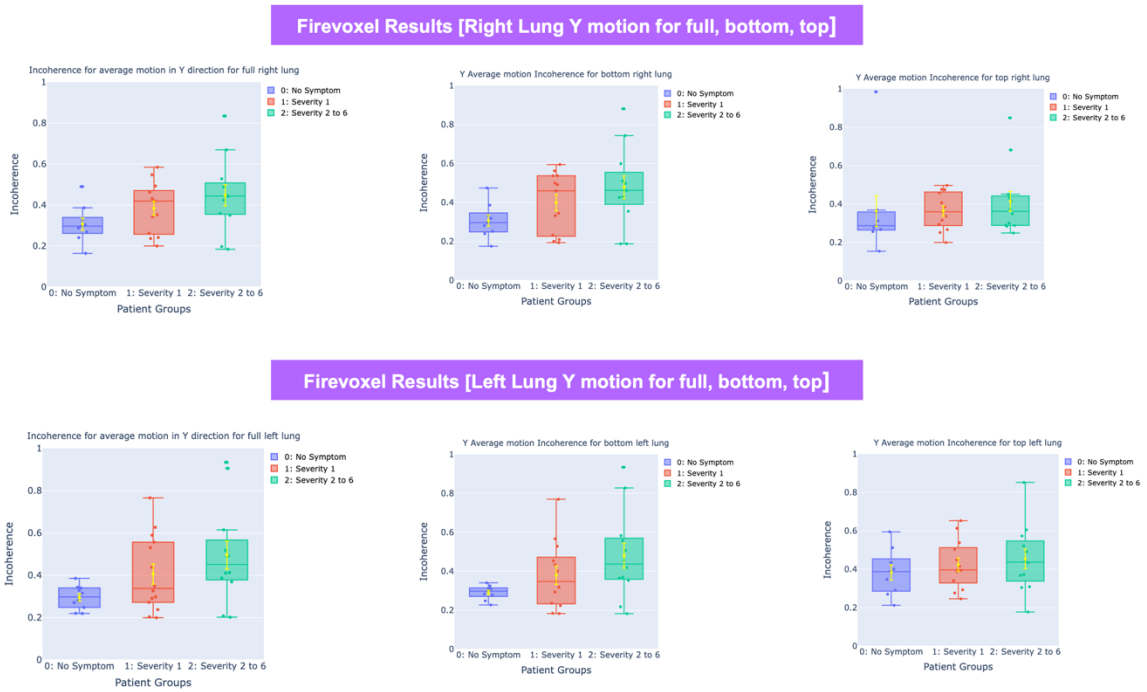


Figure 7.4: Comparison of Y motion incoherence for full, bottom and top right lung (top row) and left lung (bottom row) computed using Firevoxel segmentation masks

	Model Type	T-Statistic	P-Value	Is Significant ?
Group 0 vs. Group 1	FV R Area	-0.942	0.3575	no
	FV R Full Y	-1.686	0.1074	no
	FV R Bottom Y	-1.821	0.0839	no
	FV R Top Y	-0.009	0.9929	no
Group 0 vs. Group 2	FV R Area	-1.891	0.07510	no
	FV R Full Y	-2.444	0.0244	**yes**
	FV R Bottom Y	-2.784	0.0123	**yes**
	FV R Top Y	-0.5821	0.5700	no
Group 1 vs. Group 2	FV R Area	-1.1079	0.2810	no
	FV R Full Y	-1.043	0.308	no
	FV R Bottom Y	-1.109	0.278	no
	FV R Top Y	-0.980	0.339	no

Table 7.4.1: Statistical Significance Results for Firevoxel based right lung segmentation

	Model Type	T-Statistic	P-Value	Is Significant ?
Group 0 vs. Group 1	FV L Area	-0.255	0.8012	no
	FV L Full Y	-1.641	0.1205	no
	FV L Bottom Y	-1.811	0.0924	no
	FV L Top Y	-0.7572	0.4588	no
Group 0 vs. Group 2	FV L Area	-2.037	0.0575	no
	FV L Full Y	-2.781	0.01606	**yes**
	FV L Bottom Y	-2.894	0.0137	**yes**
	FV L Top Y	-1.094	0.2875	no
Group 1 vs. Group 2	FV L Area	-1.787	0.08976	no
	FV L Full Y	-1.483	0.1545	no
	FV L Bottom Y	-1.260	0.2214	no
	FV L Top Y	-0.495	0.625	no

Table 7.4.2: Statistical Significance Results for Firevoxel based left lung segmentation

In this section, we conduct a comparison of the average Y motion incoherence for both the left and right lungs, utilizing masks segmented with Firevoxel. It is worth noting that

Firevoxel demonstrates effectiveness in segmenting both left and right masks, unlike UniverSeg, which is why these masks are employed for this particular analysis. Figure 7.4 presents a comprehensive comparative assessment through box plots, emphasizing Y motion incoherence across the entire, bottom, and top regions of both the right and left lungs segmented using Firevoxel. Corresponding statistical analysis is detailed in Figure 7.4.1 and 7.4.2.

Upon examining both the plots and the outcomes of the statistical significance test, it is observed that the computed Y motion incoherence for the full and bottom regions of both the right and left lung effectively distinguishes between Group 0 and Group 2 patients.

However, when distinguishing between Group 0 and Group 1, the incoherence calculated for the bottom regions of both the left and right lungs is not considered significant, as indicated by p-values exceeding 0.05. If a slightly more lenient threshold of 0.1 were applied, a noticeable difference in incoherence for the bottom left and right lung regions between Groups 0 and Group 1 might have been identified.

Crucially, the incoherence computed for the top regions of both the left and right lungs does not exhibit statistical significance for any of the group comparisons.

In summary, the performance of Y motion incoherence is similar for both the left and right lungs, suggesting that they can be effectively utilized for comparing different groups in this context.

7.5 Summary of Various Comparisons

In this section, we provide a summary of the key findings from the various comparisons conducted.

In essence, it has been established that the incoherence of Y motion serves as a valuable indicator of significance, effectively distinguishing between Group 0 and Group 2. Notably, the incoherence of Y motion in the bottom lung region is deemed more significant when compared to the top lung region. Moreover, both the left and right lungs prove to be suitable for computing Y motion incoherence, exhibiting similar performance across various lung regions.

Contrary to this, the statistical analysis reveals that the incoherence of area does not exhibit significance in any of the three group comparisons, unless a slightly more lenient p-value threshold (0.1) is applied. Under this lenient threshold, it might have been capable of distinguishing between Group 0 and Group 2. Additionally, with a more lenient threshold, the Y motion incoherence for the bottom lung region could potentially distinguish not only between Group 0 and Group 2 (as previously indicated using a tighter threshold) but also between Group 0 and Group 1.

In conclusion, the findings strongly suggest that Y motion incoherence emerges as a superior marker for the analysis of respiratory breathing patterns compared to area-based incoherence. Moreover, Y motion incoherence proves versatile in performing regional analysis, distinguishing between top and bottom regions, a capability not shared by the area-based approach.

8. Pixel-wise Incoherence Map Visualization

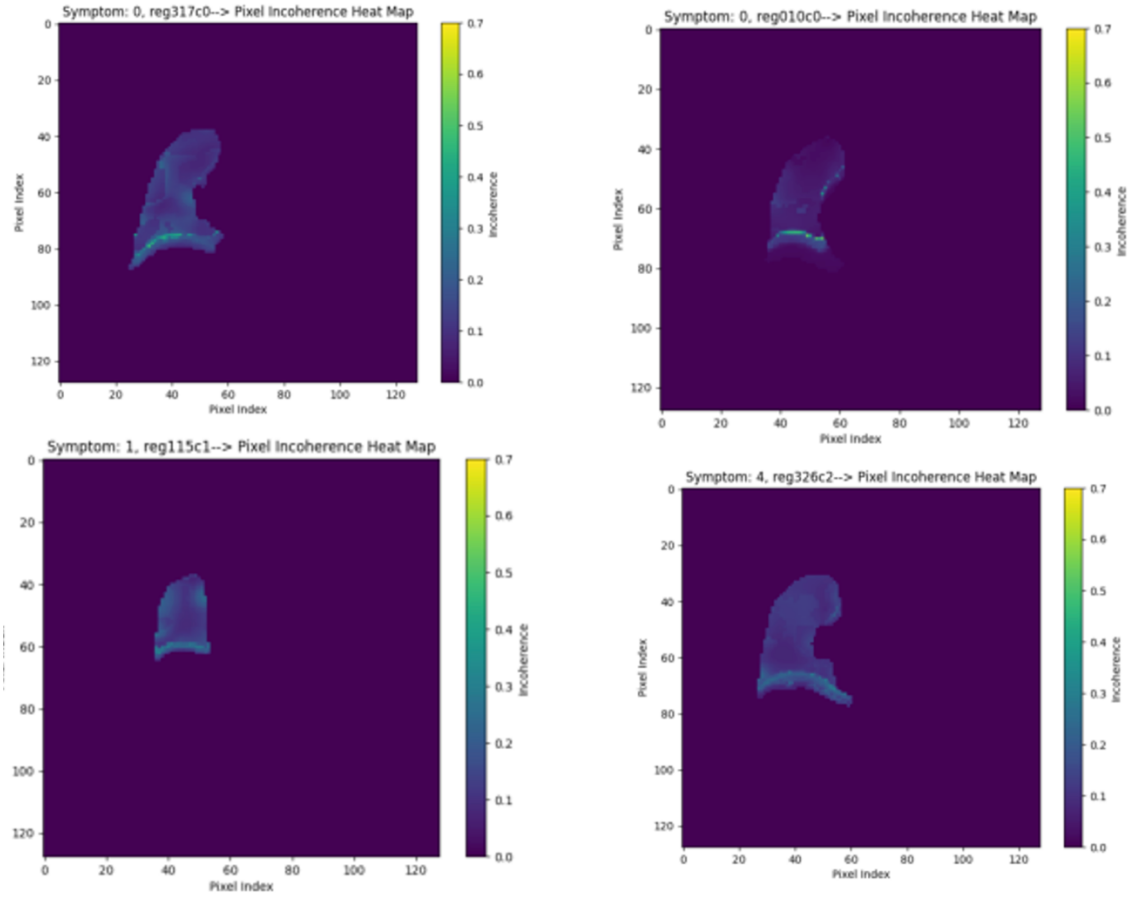


Figure 8.1: Visualization of pixel-wise incoherence maps for patients with different severities

In this section, visualizations of the incoherence of Y motion for each pixel are provided for patients with varying levels of severity. Although the expectation was to observe discernible differences in overall visualizations corresponding to increased patient severity, such distinctions are not readily apparent. However, a noteworthy observation is a prominent row of pixels near the bottom region of the lung, suggesting the region where incoherence is most pronounced. This further reinforces the notion, as illustrated in previous sections and reiterated here, that the bottom region of the lung can indeed be effectively utilized for computing Y motion incoherence.

The computation of incoherence for the Y motion of every pixel follows a methodology akin to the one elucidated in Chapter 5. Mask 1 acts a as the reference frame for subsequent analyses. To establish a coherent link between Mask 1 and other masks (ranging from 2 to 250), a warped mapping is generated through symmetric diffeomorphism. This results in a series of mappings denoted as φ_{1_2}, φ_{1_3}, φ_{1_4} .. φ_{1_N} where, φ_{1_N} signifies a mapping or transformation grid that provides new coordinate locations for mask N, aligning it with mask 1.

The Y motion of every pixel in mask 1 is tracked across 250 frames using the aforementioned warpings. To elaborate, assuming there are *p* pixels in mask 1, *p* Y motion curves are generated. Subsequently, the incoherence and time period of each of these *p* Y motion curves are computed. The incoherence's of all *p* curves are then averaged to yield the final incoherence value for a specific patient.

Please note that the pixel-wise Y motion for visualization purposes was computed by first normalizing the Y motion field. The normalization is described by the following equation:

$$MF' = \frac{MF' - min(MF'')}{max(MF'') - min\,(MF'')}$$

Please note here *MF'* is Y motion field of a particular pixel whose incoherence is to be computed, *MF''* is Y motion field of all pixels within the whole right lung. Note that, *max(MF'')* denotes the maximum Y motion field of a pixel (eg. p1) within the entire right lung; for instance, pixel p1 may have the highest Y motion field among all pixels in the right lung. Similarly, *min(MF'')* represents the minimum Y motion field of a pixel (eg. p2) within the entire right lung; for instance, pixel p2 may have the lowest Y motion field among all pixels in the right lung. It's important to note that p1 and p2 may or may not be the same.

Next, incoherence is computed using the following equation:

$$Inc(t) = \frac{\sum_{i=0}^{Nsteps} | MF'(i * step) - MF'(i * step + t)|}{Nsteps}$$

Please note, here, *MF'* is motion field in vertical or Y direction for a specific pixel of the lung. *Nsteps* is total number of frames (i.e. 250). *Step* is the step size which is set to 1. 't' represents the time period of the motion curve, and its computation follows a similar methodology as detailed in Chapter 6.

9. Machine Learning for Classification of long-COVID Patients

9.1 Feature Computation

In this analysis, the Firevoxel left lung segmentation results were utilized to compute the average Y motion, which is then used to compute various features. The features considered in the study included Y motion incoherence, Coefficient of Variation, and Power Spectral Entropy (PSE).

The computation of Y motion incoherence followed the established methodology outlined in previous sections. The Coefficient of Variation was specifically defined as the ratio of the standard deviation of the average Y motion to the mean of the average Y motion. To calculate Power Spectral Entropy, the average Y motion was transformed to the frequency domain through the application of the Fast Fourier Transform (FFT). The resulting complex Fourier coefficients had their magnitudes squared, generating the power spectrum that depicts power distribution across different frequencies. From this spectrum, the top seven power spectrals were chosen based on their magnitudes. Following this selection, Shannon Entropy was computed for the chosen power spectrals, offering a metric to assess the complexity or unpredictability of the Y motion curve.

Once these features were defined, a Decision Tree model was employed for classification purposes. The Decision Tree model utilizes the defined features to make informed classification decisions. After the classification process, the evaluation was conducted using the Leave One Out (LOO) method, a suitable choice due to the relatively small size of the dataset, comprising 39 patients. In LOO, each data point is systematically used as a test set, while the model is trained on the remaining data points. This process is iterated

until every data point has been utilized as a test set exactly once. LOO serves as a rigorous evaluation technique, particularly effective for small datasets, as it minimizes bias and provides a comprehensive assessment of the model's performance.

9.2 Discussion of Results

	Precision	Recall	F1
Group 0	0.5000	0.5455	0.5217
Group 1	0.4167	0.3571	0.3846
Group 2	0.3846	0.7143	0.6897
Overall	0.5299	0.5385	0.5328
Accuracy	0.538		

Table 9.2.1: Classification Results for Group 0, 1 and 2

	Precision	Recall	F1
Group 0	0.5455	0.5455	0.5455
Group 1+2	0.8214	0.8214	0.8214
Overall	0.7436	0.7436	0.7436
Accuracy	0.743		

Table 9.2.2: Classification Results for Group 0 and combined Groups 1 and 2

	Precision	Recall	F1
Group 0	0.9000	0.8182	0.8182
Group 2	0.8667	0.9286	0.8966
Overall	0.8813	0.8800	0.8792
Accuracy	0.88		

Table 9.2.3: Classification Results for Group 0 and Group 2. Note Group 1 has been excluded.

In addition to assessing overall accuracy, weighted precision, recall, and F1 score were computed for each distinct group. The classification involved Group 0, comprising 11 patients with no symptoms, Group 1 consisting of 14 patients with mild symptoms, and Group 2, which included 14 patients experiencing severe long COVID symptoms.

As detailed in Table 9.2.1, the overall accuracy of classifying patients into three different groups was approximately 0.54. The recall value for Group 1 patients was notably low at 0.36, contributing to the overall lower accuracy.

Following this, patients from Group 1 and Group 2 were consolidated into a single group, simplifying the problem to classify patients as either having long COVID or exhibiting no symptoms. As indicated in Table 9.2.2, the overall accuracy of classifying patients into two groups (no symptom or long COVID) was approximately 0.74. It is important to note that the recall value for Group 0 (no symptom) classification was somewhat low at 0.54. This could be attributed to the relatively small sample size for Group 0, consisting of a total of 11 patients, compared to the 28 patients in Group 1 and Group 2 combined. Overall, the classifier demonstrated effectiveness in classifying long COVID patients but was less proficient in classifying no symptom patients, suggesting a tendency to classify patients as having long COVID.

Finally, recognizing the low recall score observed previously for Group 1, it was removed to evaluate the classifier's performance in classifying Group 0 (no symptom) versus Group 2 (severe symptoms) patients. As outlined in Table 9.2.3, the overall accuracy was found to be 0.88, with a recall value of 0.82 for Group 0 and 0.93 for Group 2. This indicates that the classifier effectively classified patients as either having no symptoms or experiencing severe long COVID symptoms.

As decision trees evaluate the importance of various features, the analysis reveals that incoherence is the most crucial feature, followed by power spectral entropy, and then the coefficient of variation.

Bibliography

1. Dysfunctional Respiratory Patterns in Symptomatic Post-Acute Covid-19 Patients on Dynamic High Temporal Resolution Free-Breathing Lung MRI.
2. Quantitative Characterization of Respiratory Pattern on Dynamic High Temporal Resolution MRI to Stratify Patients by Cardiopulmonary Symptom Severity. *Journal of Magnetic Resonance Imaging. Under Review.*
3. Butoi*, V. I., Ortiz*, J. J. G., Ma, T., Sabuncu, M. R., Guttag, J., & Dalca, A. V. (2023). UniverSeg: Universal Medical Image Segmentation. *International Conference on Computer Vision*.
4. Avants, Brian B., et al. "Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain." *Medical image analysis* 12.1 (2008): 26-41.
5. Firevoxel. https://firevoxel.org/download/
6. Fang Liu (2023). MatrixUser v2.2 (https://www.mathworks.com/matlabcentral/fileexchange/43780-matrixuser-v2-2), MATLAB Central File Exchange. Retrieved August 7, 2023.