

Statistical Methodology for Computed Tomography Scans of the Lung

Sarah M. Ryan, M.S.

Ph.D. Candidate
Department of Biostatistics and Informatics
Colorado School of Public Health
University of Colorado Anschutz Medical Campus

July 13, 2020



Committee Members

Advisor: Nichole Carlson, Ph.D., Colorado School of Public Health

Chair: Tasha Fingerlin, Ph.D., National Jewish Health

Member: Debashis Ghosh, Ph.D., Colorado School of Public Health

Member: John Muschelli, Ph.D., Johns Hopkins Bloomberg School of Public Health

Member: Fuyong Xing, Ph.D., Colorado School of Public Health

Outline

- 1 Chapter 1: Introduction
- 2 Chapter 2: Template Creation for High Resolution Computed Tomography Scans of the Lung in R Software
- 3 Chapter 3: An Eigenvector Spatial Filtering Model for Lung Imaging Data
- 4 Chapter 4: Cluster Activation Mapping with Applications to Medical Imaging Data
- 5 Chapter 5: Conclusion
- 6 Appendix

Chapter 1:

Introduction

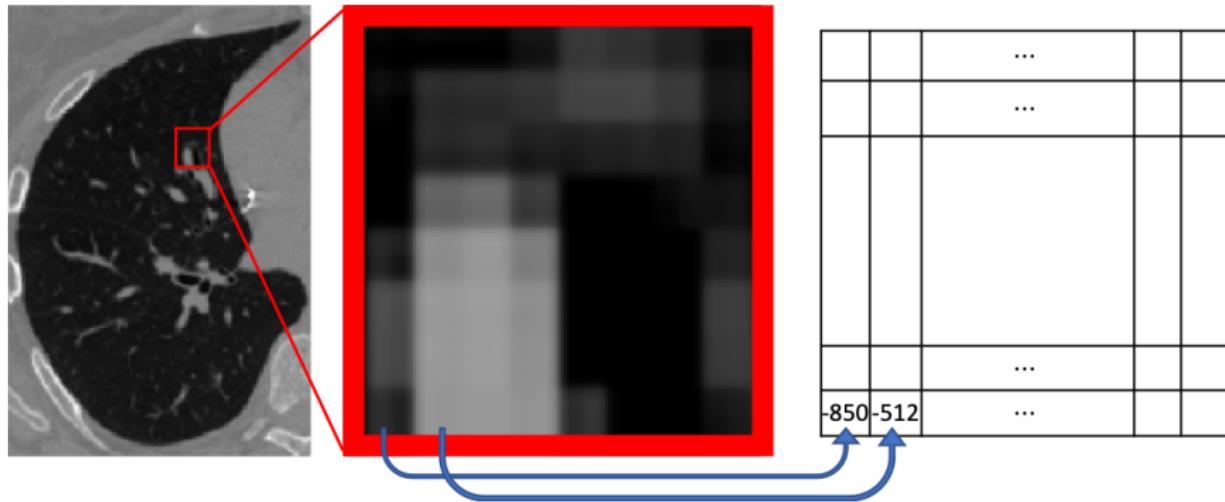
Computed Tomography (CT)

A computerized x-ray imaging procedure which generates cross-sectional images of the body that can be combined to form three-dimensional images



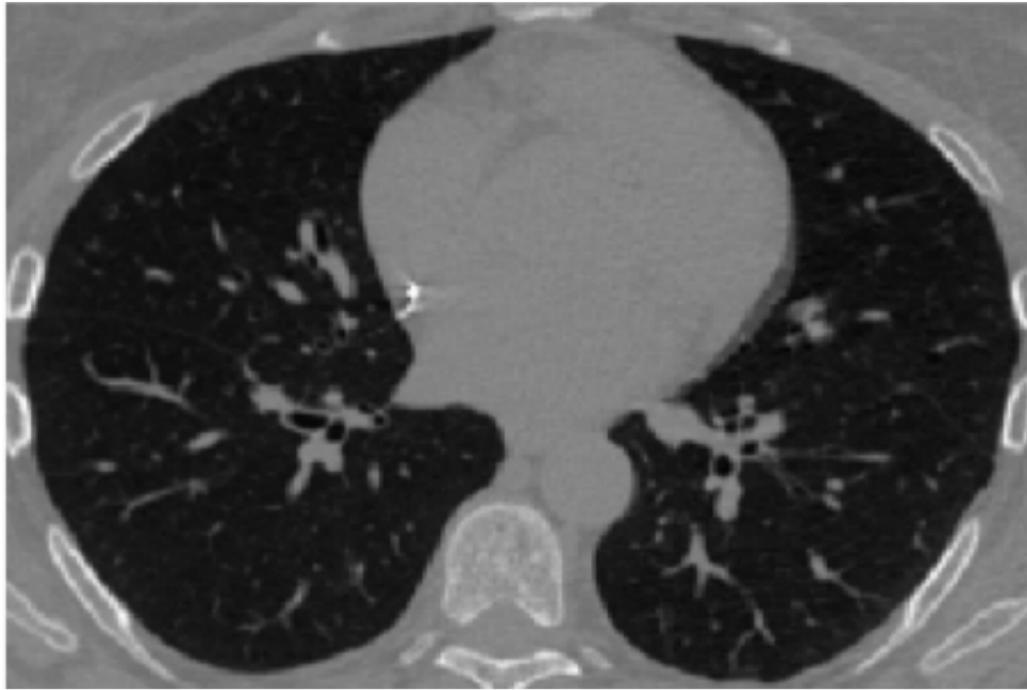
Hounsfield Unit (HU)

A measure of the radiodensity of a pixel from a CT scan

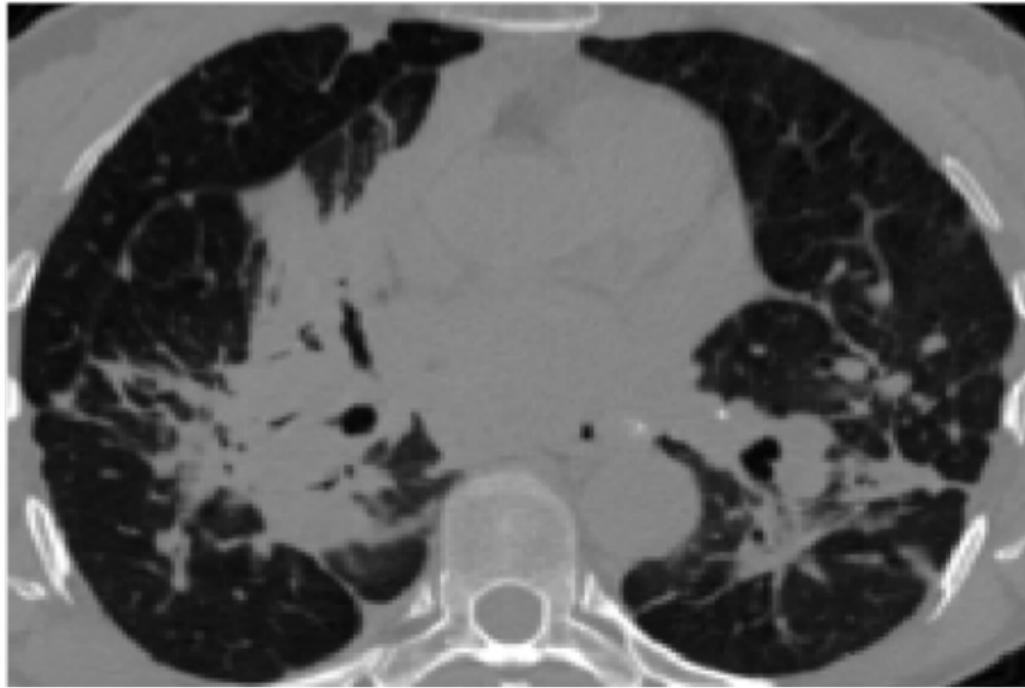


CT Scans of the Lung

Healthy Lung



Fibrotic Lung



An interstitial lung disease characterized by the formation of granulomas in the lung, resulting in compromised lung function and reduced quality of life [Nunes et al., 2005]

- Incidence rate: 8-18 per 100,000 [Baughman et al., 2016]
- Highest rates in Blacks, females

Table: Scadding staging for pulmonary sarcoidosis [Scadding, 1961]

Stage	Chest radiograph findings
0	No chest abnormality
I	Bilateral hilar lymphadenopathy (BHL)
II	BHL and parenchymal abnormality
III	Parenchymal abnormality
IV	Fibrosis with volume loss

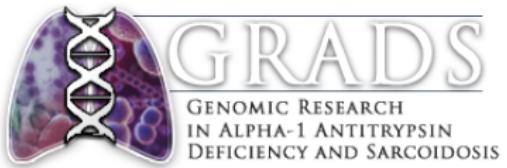


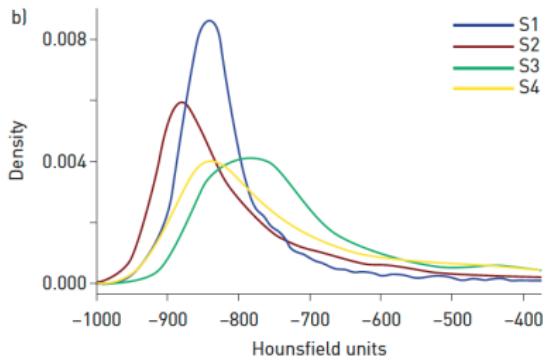
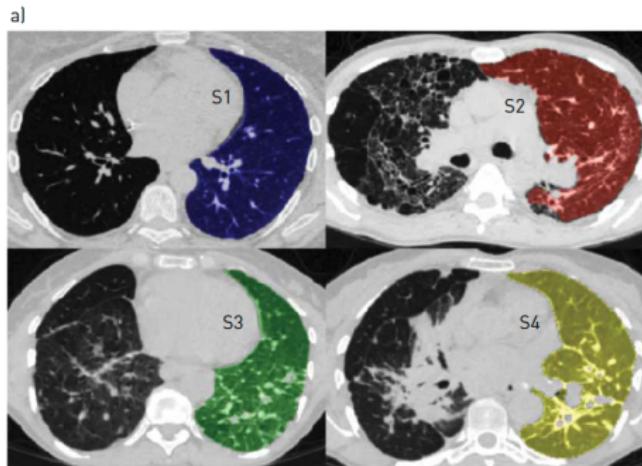
Table: GRADS characteristics

- Data: Demographics, pulmonary function testing, patient reported outcomes
- Research Chest HRCT: Siemens, standard B35f kernel, 0.75mm thickness, 0.5mm computed interval

	Overall
Sample Size	301
Male	142 (47.2)
White	218 (72.9)
Hispanic	14 (4.7)
Age (years)	52.87 (9.75)
Height (in)	67.08 (4.15)
FVC PRED	87.81 (16.75)
FEV1 PRED	84.90 (20.71)
Scadding	
0	40 (13.3)
1	59 (19.6)
2	87 (28.9)
3	43 (14.3)
4	72 (23.9)

Existing Methods for Image Analysis

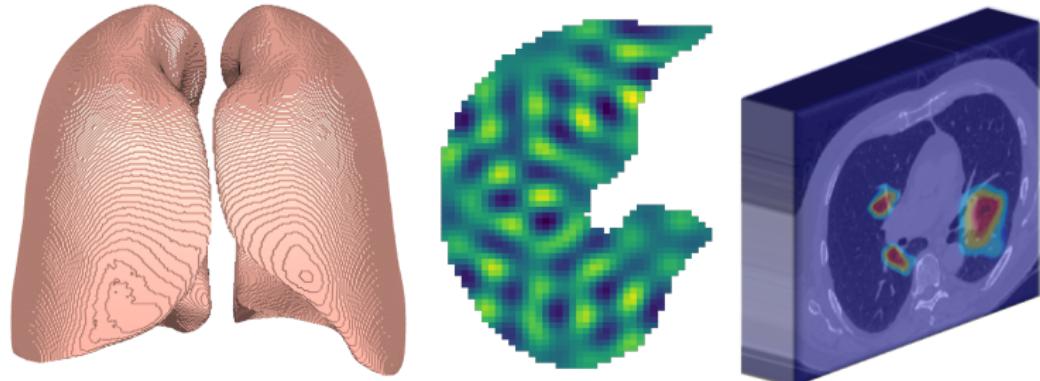
Radiomics: An emerging field in which large numbers of quantitative features are computed from medical images, providing a rapid, objective, and sensitive quantification of lung abnormalities [Ryan et al., 2019a]



Goal: Identify population-level spatial patterns for lung CT scans

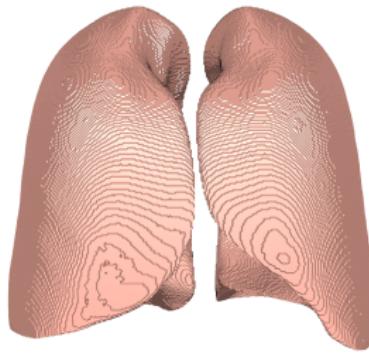
My three primary chapters are:

- ① Create a coordinate system to align voxels across scans
- ② Explain the variation in HU throughout the lung using a set of covariates
- ③ Identify novel clusters of CT scans for sub-typing



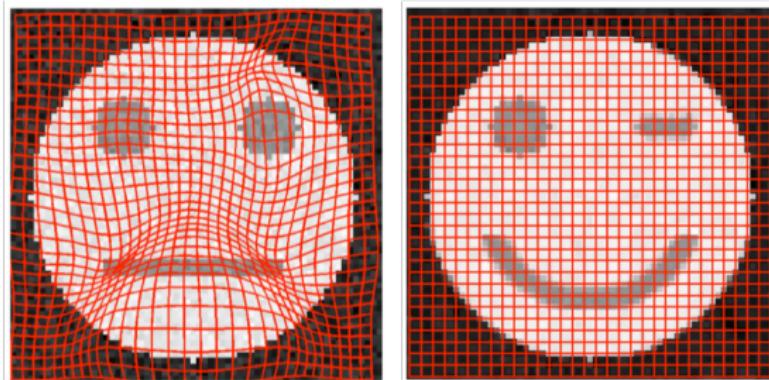
Chapter 2:

Template Creation for High Resolution Computed Tomography Scans of the Lung in R Software

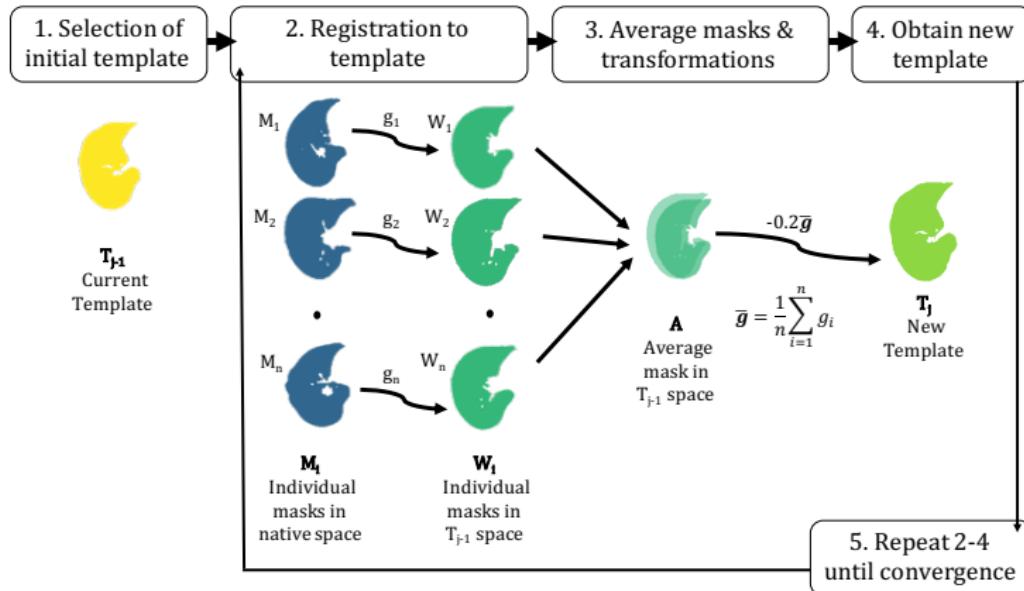


Template

- A **template** is a standardized 3D coordinate frame [Evans et al., 2012]
- By aligning each individual's image to the template, we can:
 - ▶ Identify anatomical regions that differ between groups
 - ▶ Compare findings across studies
 - ▶ Remove potential biases and technical issues related to image size and shape



Template Creation [Ryan et al., 2019b]



- Convergence is defined using the Dice similarity coefficient

Standard Lung Template

- Applied to $N = 62$ HRCT scans from a healthy non-smoking adult population [Moller et al., 2015]
- Converged after 14 iterations to average size and shape
- Right lung ≈ 3 L, left lung ≈ 2.6 L (at full-inspiration)



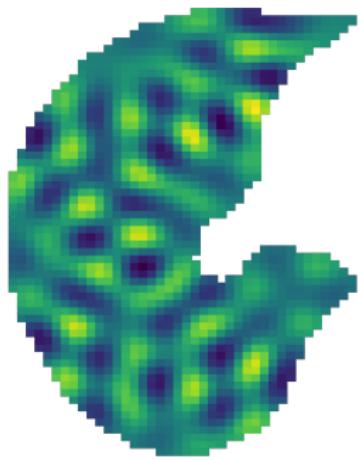
Conclusions of Chapter 2

- ① We create the first publicly available standard lung template using healthy adults, which is available for download via *lungct* [Ryan et al., 2019b]
- ② We develop a fully-automated and open-source image processing pipeline for lung CTs in R software

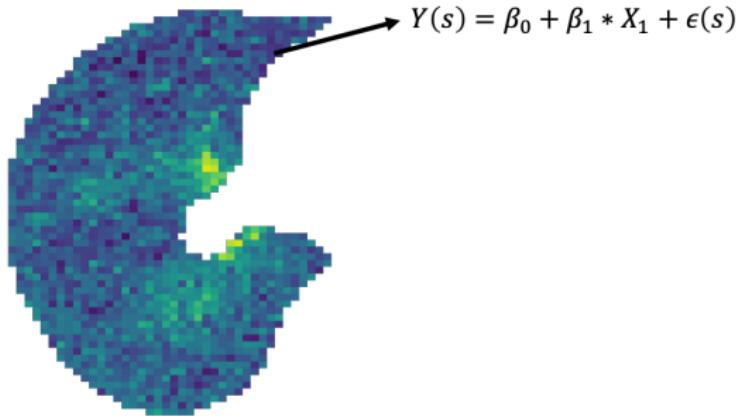


Note: For the remainder of this talk, we will assume all scans have been pre-processed, including the segmentation of the lung from the scans and registration to the lung template.

Chapter 3: An Eigenvector Spatial Filtering Model for Lung Imaging Data

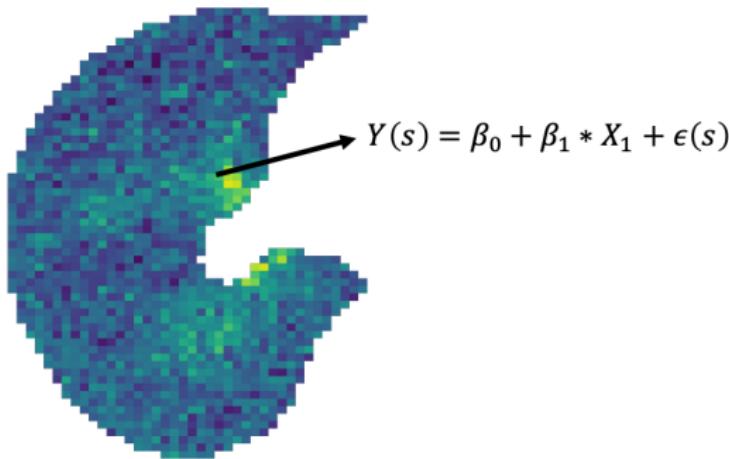


- **Voxel-based morphometry (VBM)** is an approach in neuroimaging to find associations between intensity values from structural MRI and various covariates [Ashburner and Friston, 2000]
- Traditionally, a separate statistical test is fit on every voxel

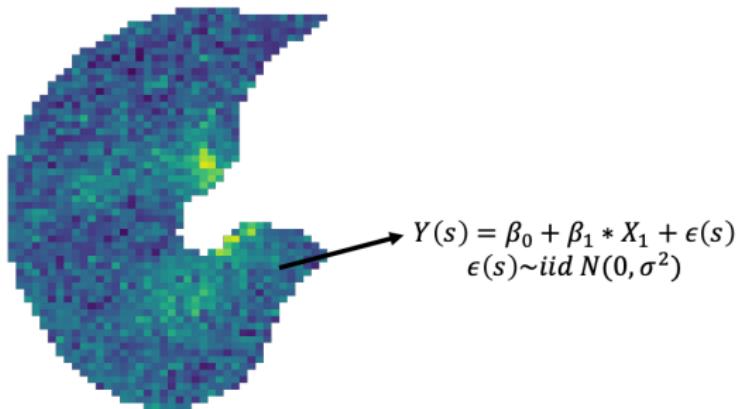


Classical Method for Imaging Data

- **Voxel-based morphometry (VBM)** is an approach in neuroimaging to find associations between intensity values from structural MRI and various covariates [Ashburner and Friston, 2000]
- Traditionally, a separate statistical test is fit on every voxel



- **Voxel-based morphometry (VBM)** is an approach, commonly used in neuroimaging, to find associations between intensity values from scans and various covariates [Ashburner and Friston, 2000]
- Traditionally, a separate statistical test is fit on every voxel



Fails to account for the spatial relationship between voxels
[Bookstein, 2001]

Existing spatial approaches

- Many spatial modeling approaches
[Friston and Penny, 2003, Smith and Fahrmeir, 2007,
Brown et al., 2014, Musgrove et al., 2016, Mejia et al., 2019]
- All rely on assumptions and approximations due to high-dimensionality
- Fast, fully Bayesian spatiotemporal inference for fMRI data
[Musgrove et al., 2016]
 - ▶ Uses MCMC for estimation
 - ▶ Partitions the brain into parcels
 - ▶ Within each parcel, models spatial correlation in residuals using Moran eigenvectors

Eigenvector Spatial Filtering (ESF)

- ESF is a type of low rank approximation, which describes spatial variation in an outcome using a linear combination of L basis functions where $L \ll N$
- These basis functions, also known as Moran eigenvectors, are related to Moran's I [Moran, 1950], a common spatial summary measure

$$MC(\mathbf{y}) = \frac{N}{\mathbf{1}'\mathbf{C}\mathbf{1}} \frac{\mathbf{y}'\mathbf{M}\mathbf{C}\mathbf{M}\mathbf{y}}{\mathbf{y}'\mathbf{M}\mathbf{y}} \quad (1)$$

- Traditionally used in high-dimensional geographic data and estimated using REML [Murakami et al., 2017]
- Good approximation and computationally efficient [Murakami and Griffith, 2019]

Moran Eigenvectors

- Distance matrix: ($N \times N$) dimension
- Eigenvector design matrix: ($N \times L$) dimension

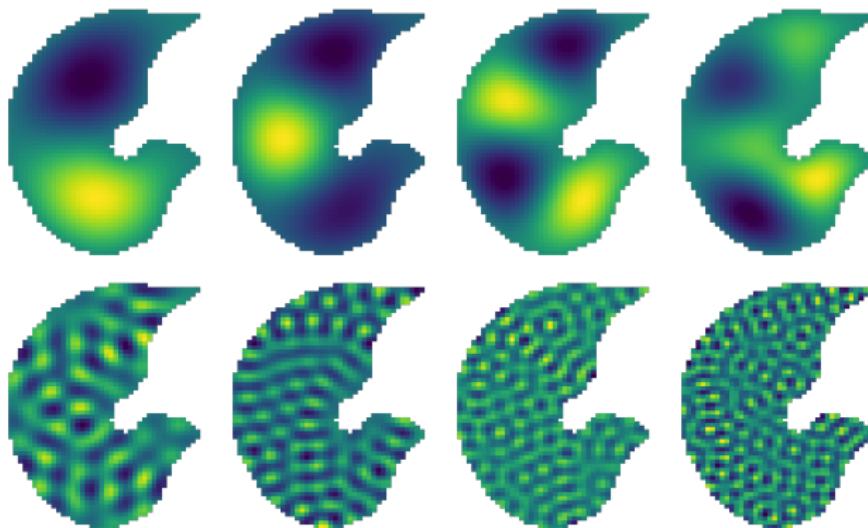
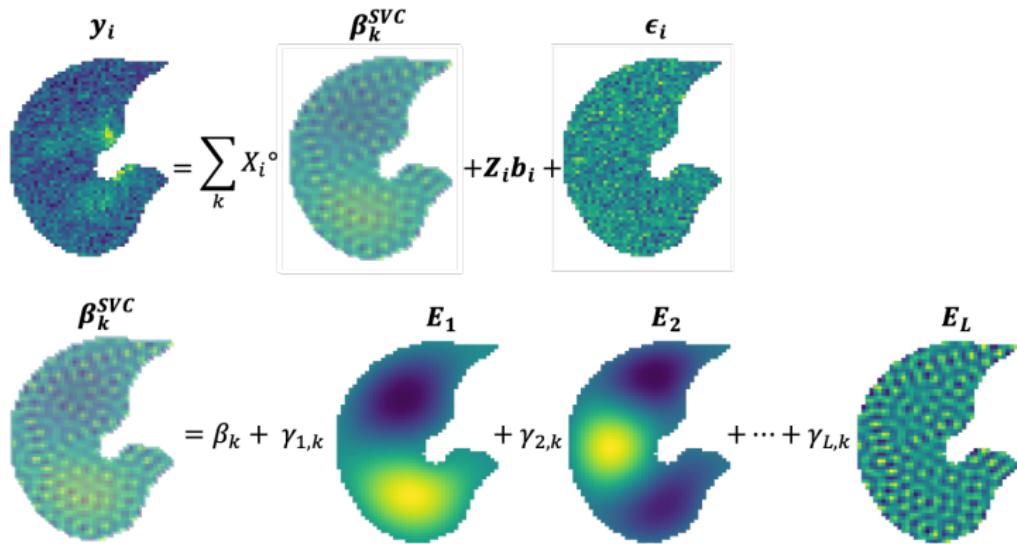


Figure: Moran eigenvectors based on the distance matrix from a 2D axial slice of the lung. The top row corresponds to the 1st - 4th eigenvectors. The bottom row corresponds to the 100th, 200th, 300th, and 397th eigenvectors.

Our spVBM Model

Our spatial voxel-based morphometry, or **spVBM**, for a single subject i is:



$$\mathbf{b}_i \sim N(\mathbf{0}, D), \quad \epsilon_i \sim N(0, \sigma_\epsilon^2 \mathbf{I}), \quad \gamma_k \sim N(\mathbf{0}, \sigma_k^2 \Lambda(\alpha_k))$$

spVBM: Reparameterization

We reparameterize so that the global term $\mathbf{X}\beta$ and spatial term $\tilde{\mathbf{E}}\gamma$ are separated, and we obtain a common variance term across the random variables following [Bates et al., 2014]:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\beta + \tilde{\mathbf{E}}\mathbf{V}(\theta)\mathbf{u} + \mathbf{Z}\Omega(\phi)\mathbf{w} + \varepsilon, \quad \mathbf{u}, \mathbf{w}, \varepsilon \sim N(\mathbf{0}, \sigma^2 \mathbf{I}) \\ \tilde{\mathbf{E}} &= [\mathbf{X}_1 \circ \mathbf{E} \cdots \mathbf{X}_K \circ \mathbf{E}] \\ \mathbf{V}(\theta) &= \begin{bmatrix} \mathbf{V}(\theta_1) & & \\ & \ddots & \\ & & \mathbf{V}(\theta_K) \end{bmatrix}, \quad \Omega(\phi) = \begin{bmatrix} \Omega(\phi_1) & & \\ & \ddots & \\ & & \Omega(\phi_H) \end{bmatrix} \end{aligned}$$

where $\mathbf{V}(\theta)\mathbf{u} = \gamma$ and $\Omega(\phi)\mathbf{w} = \mathbf{b}$

The log-likelihood of our model can be expressed as:

$$\log \text{lik}(\beta, \theta, \phi, \sigma^2) = -\frac{N}{2} \log(2\pi\sigma^2) - \log(|\mathbf{V}(\theta)|^2) - \log(|\Omega(\phi)|^2) - \frac{d(\theta, \phi)}{\sigma^2}$$

$$d(\theta, \phi) = \|\mathbf{y} - \mathbf{X}\beta - \mathbf{E}\mathbf{V}(\theta)\mathbf{u} - \mathbf{Z}\Omega(\phi)\mathbf{w}\|^2 + \|\mathbf{u}\|^2 + \|\mathbf{w}\|^2$$

Then, the best linear unbiased estimates of β , \mathbf{u} and \mathbf{w} are found by maximizing the log-likelihood:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\mathbf{u}} \\ \hat{\mathbf{w}} \end{bmatrix} = \mathbf{P}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \mathbf{V}(\theta)\tilde{\mathbf{E}}'\mathbf{y} \\ \Omega(\phi)\mathbf{Z}'\mathbf{y} \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\tilde{\mathbf{E}}\mathbf{V}(\theta) & \mathbf{X}'\mathbf{Z}\Omega(\phi) \\ \mathbf{V}(\theta)\tilde{\mathbf{E}}'\mathbf{X} & \mathbf{V}(\theta)\tilde{\mathbf{E}}'\tilde{\mathbf{E}}\mathbf{V}(\theta) + \mathbf{I} & \mathbf{V}(\theta)\tilde{\mathbf{E}}'\mathbf{Z}\Omega(\phi) \\ \Omega(\phi)\mathbf{Z}'\mathbf{X} & \Omega(\phi)\mathbf{Z}'\tilde{\mathbf{E}}\mathbf{V}(\theta) & \Omega(\phi)\mathbf{Z}'\mathbf{Z}\Omega(\phi) + \mathbf{I} \end{bmatrix}$$

Variance parameters θ, ϕ, σ^2 are estimated using residual maximum likelihood estimation (REML)

- Interested in β_k^{SVC}
- Voxel-level null hypothesis that $\beta_k^{SVC}(s) = 0$
- Wald statistic: $W(s) = \frac{(\beta_k^{SVC}(s)-0)^2}{Var(\beta_k^{SVC}(s))}$
- Multiple comparisons correction: Benjamini-Hochberg (BH)
[Benjamini and Hochberg, 1995, Mejia et al., 2019]

Simulation Study

- Simulated 2D lung slices, with varying disease abnormalities
- Also varied sample size, SNR
- Compared spVBM to VBM

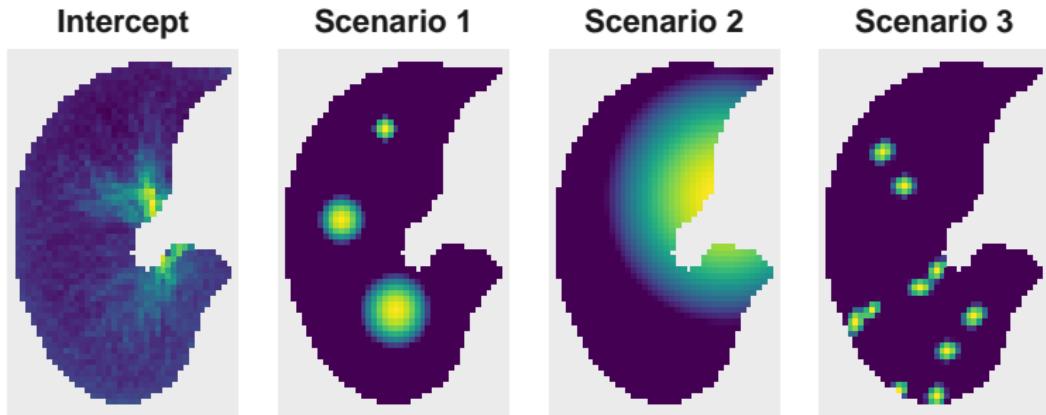


Figure: Simulated β^{SVC} for each scenario.

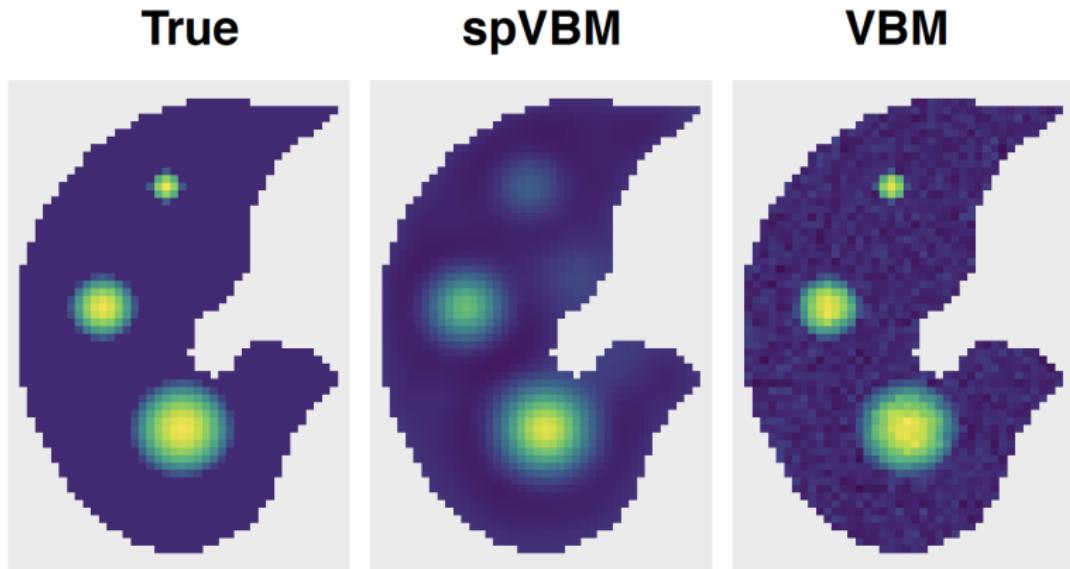


Figure: Estimated β_1^{SVC} for Simulation Scenario 1

$$\hat{\beta}_1^{SVC} > 0$$

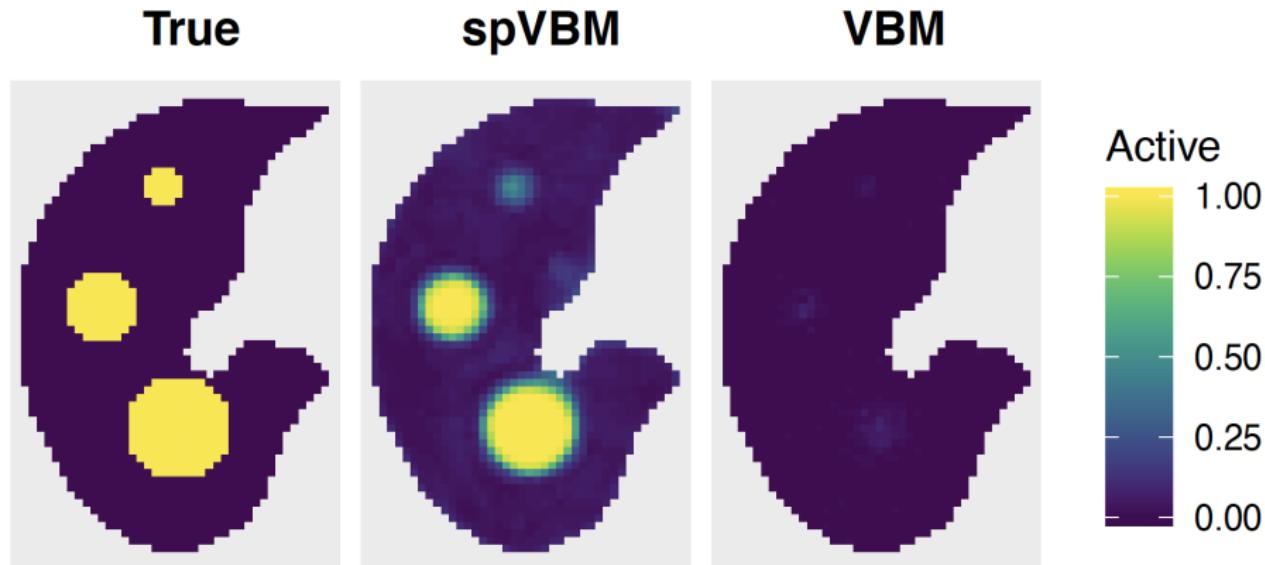


Figure: Proportion of simulations from Scenario 1 in which $\hat{\beta}_1^{SVC}(s)$ reached significance at $\alpha = 0.05$ after BH correction.

- **Data:** GRADS
- **Methods:**
 - ▶ Pre-processing: Segmentation, Registration
 - ▶ Resampled to 4mm^3 for computation (Right: 29K non-null voxels. Left: 24K non-null voxels)
 - ▶ Created eigenvector matrix (Right: 419 eigenvalues. Left: 385 eigenvectors)
- **spVBM analysis:**
 - ▶ Difference in HU between Scadding stage 4 (fibrosis) and stage 0 (healthy)
 - ▶ Association between HU and FEV1

Results

Compared to stage 0 patients (healthy), patients with stage 4 (fibrosis) have areas of significantly higher intensity near the hilar region.

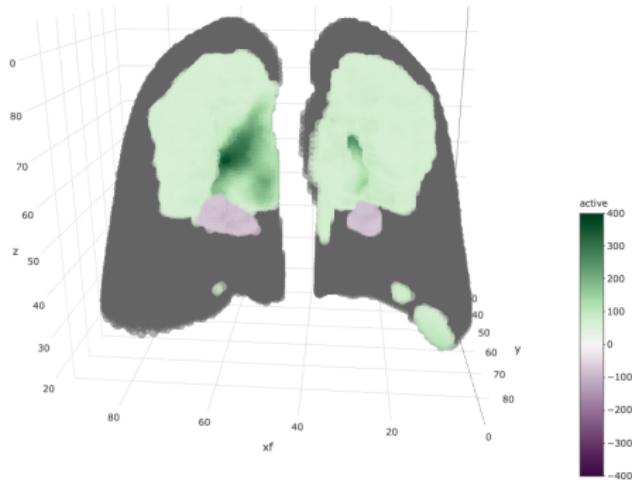


Figure: Statistical parametric map indicating the difference in HU between stage 0 and 4 Scadding, with colored values representing statistically significant effects at $|t| < 10$ after BH correction.

See in 3D

Results

FEV1 is significantly associated with HU throughout the whole lung.

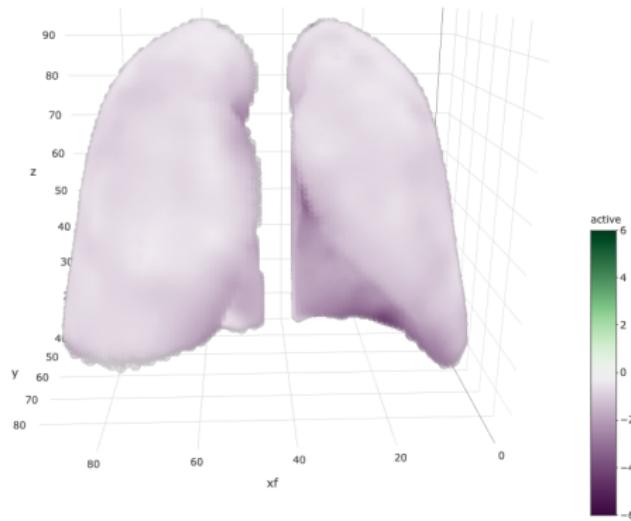
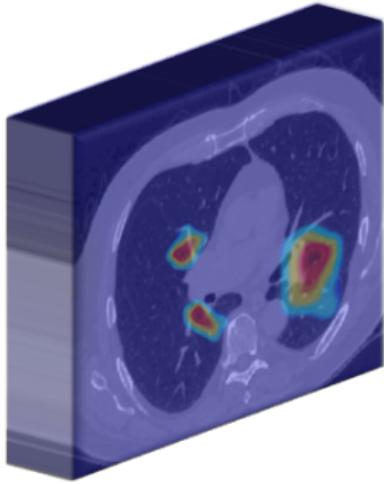


Figure: Statistical parametric map indicating the effect of FEV1 on HU, with colored values representing statistically significant effects at $p < 0.05$ after BH correction.

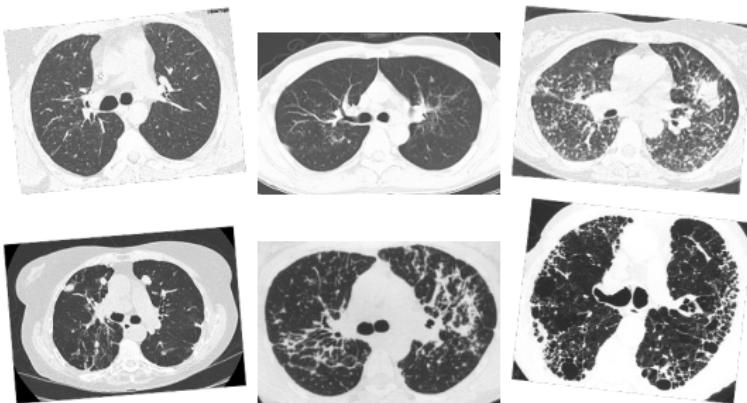
See in 3D

Chapter 4: Cluster Activation Mapping with Applications to Medical Imaging Data



Motivation

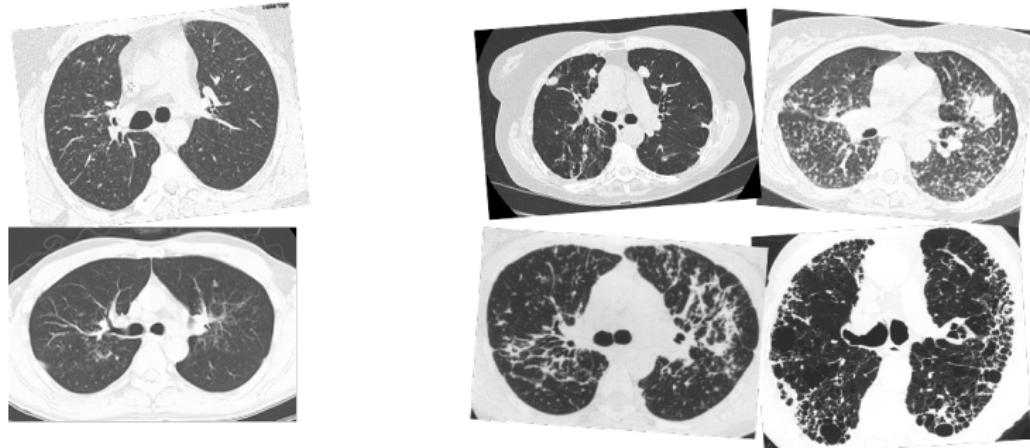
- Vast amounts of medical imaging datasets are becoming available, BUT annotation of these images is difficult
- Traditionally, we rely on visual assessment for annotation, which can be subjective, costly and slow to obtain
- This is especially true for sarcoidosis, which can have many disease manifestations on CT scans [Sluimer et al., 2006]



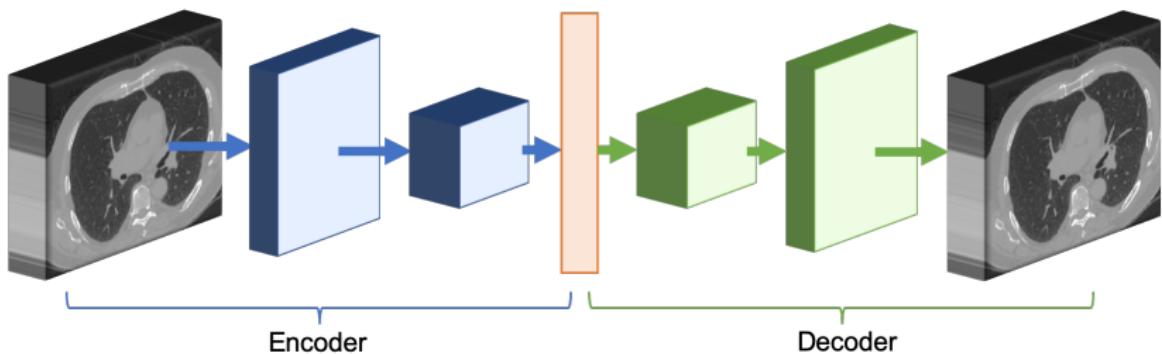
Motivation

For unlabeled or poorly labeled medical imaging datasets, such as sarcoidosis, we want to develop a data-driven method to:

- ① Assign labels to images to develop new disease subtypes
- ② Understand the reasoning behind the label, which is essential to its adoption in clinical practice

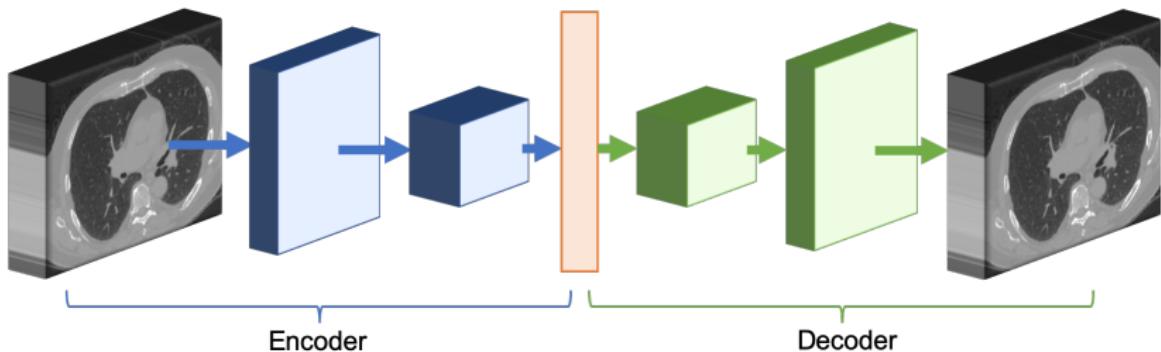


- Deep clustering is a recent methodology which assigns images into clusters using unsupervised machine learning networks, specifically autoencoders [Guo et al., 2017, Afshar et al., 2018, Masci et al., 2011, Xie et al., 2016]



Deep Clustering

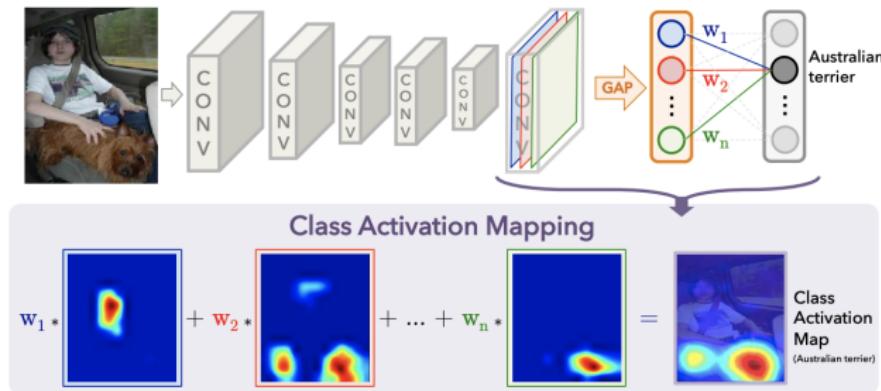
- Deep clustering is a recent methodology which assigns images into clusters using unsupervised machine learning networks, specifically autoencoders [Guo et al., 2017, Afshar et al., 2018, Masci et al., 2011, Xie et al., 2016]



- While these methods have proven successful for clustering images, they provide no understanding as to why an image is assigned to a particular cluster.

Class Activation Mapping (CAM)

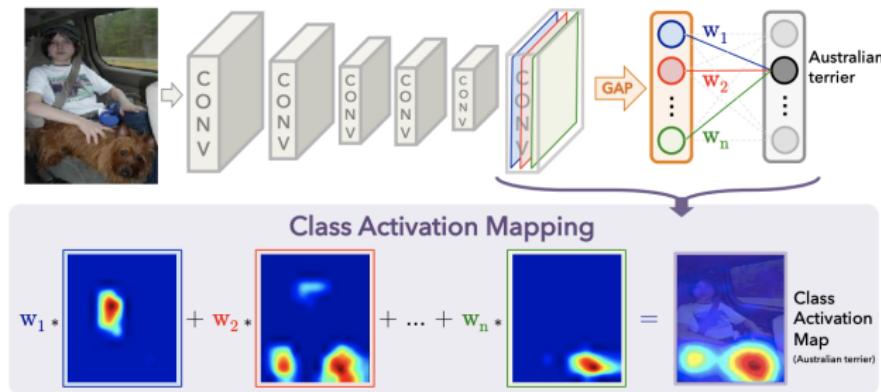
- In *supervised* machine learning, class activation maps are used to highlight the discriminative region from the original image used by the neural network to identify that category [Zhou et al., 2016]



- Other methods include: GradCAM, GradCAM++, Score-CAM [Selvaraju et al., 2017, Wang et al., 2019]

Class Activation Mapping (CAM)

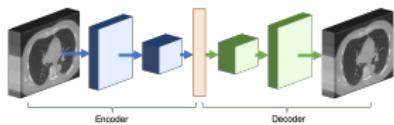
- In *supervised* machine learning, class activation maps are used to highlight the discriminative region from the original image used by the neural network to identify that category [Zhou et al., 2016]



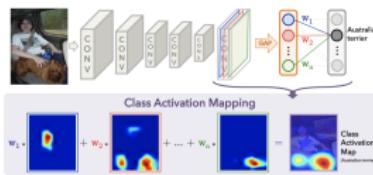
- No comparable methods exist in the *unsupervised* setting, such as in clustering

Our Methodology

Deep Clustering



Class Activation Mapping



Cluster Activation Mapping (CLAM)

CLuster Activation Mapping (CLAM)

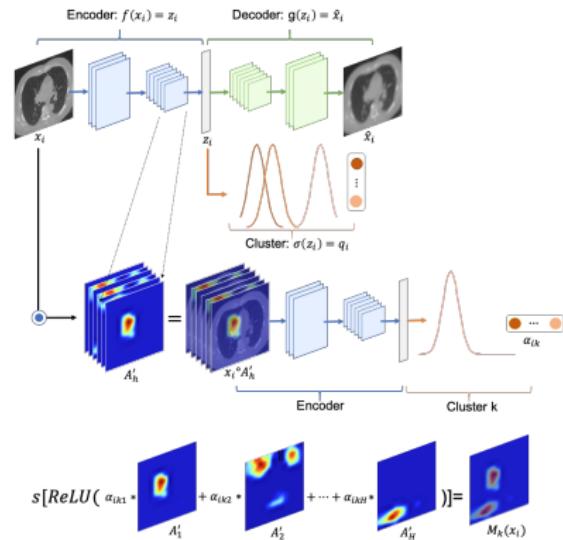
To create a CLAM for a particular image, we perform the following:

① Training

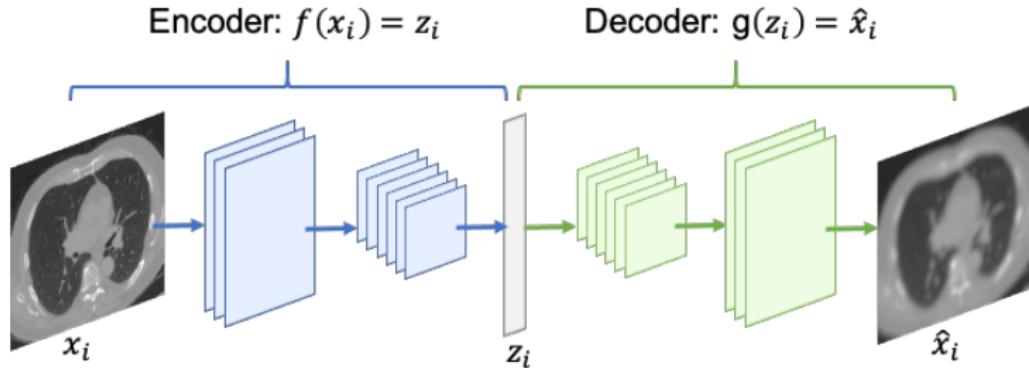
② Clustering

③ Weighting

④ Combining

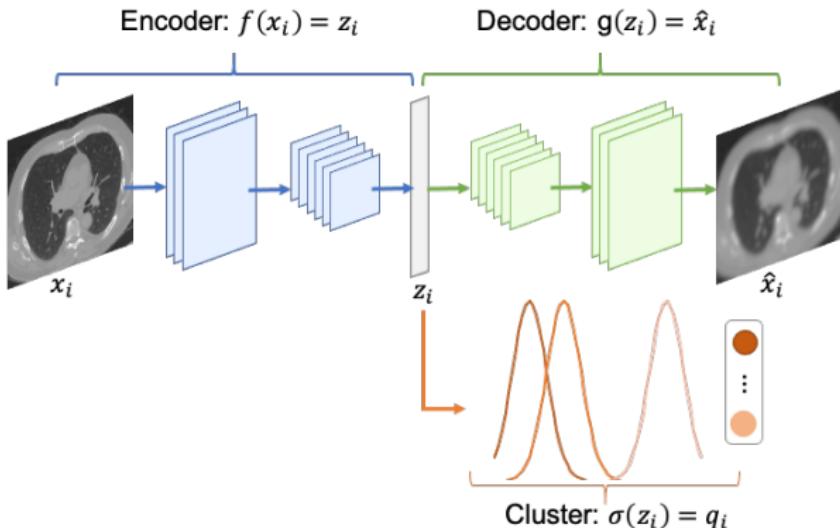


1. Training



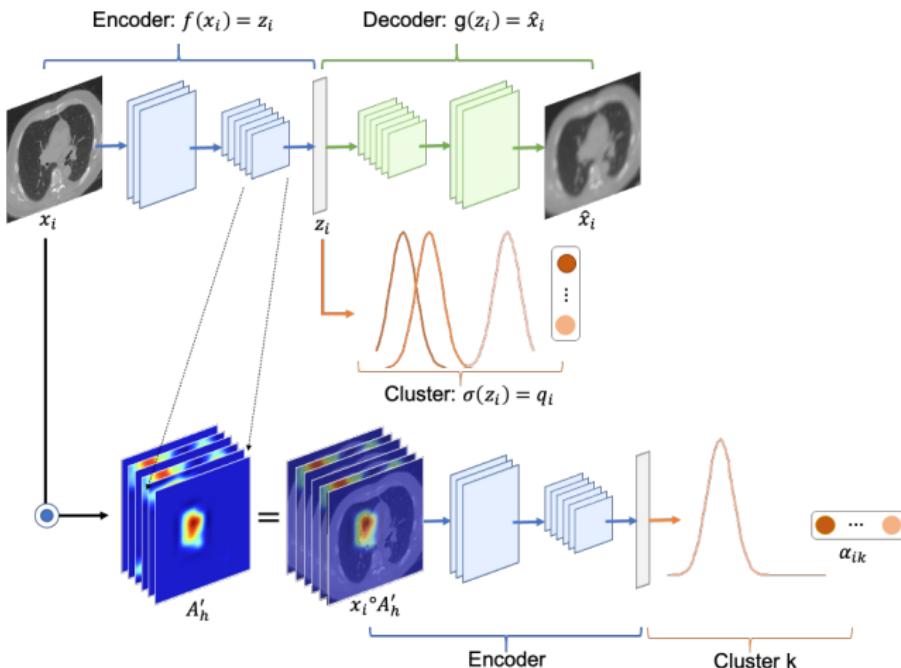
- Layers: Convolutional
- Loss function: $L_r = \sum_{i=1}^N \sum_{s=1}^S [x_i(s) - \hat{x}_i(s)]^2$
- Optimization function: Adaptive moment (Adam)

2. Clustering



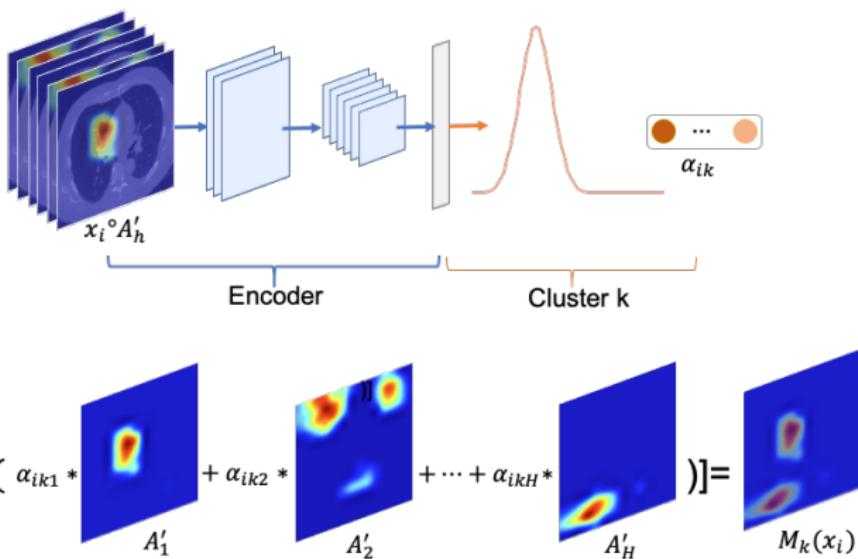
- Clustering layer: $q_{ik} = \frac{\left[1 + \sum_{j=1}^J (z_{i,j} - \mu_{k,j})^2\right]^{-1}}{\sum_{k=1}^K \left[1 + \sum_{j=1}^J (z_{i,j} - \mu_{k,j})^2\right]^{-1}}$
- Clustering loss: $L_c = KL(P||Q)$
- Total loss: $L = L_r + \gamma L_c$

3. Weighting



- Last convolutional layer used to create the activation map
- CLAM created for the cluster with the highest probability

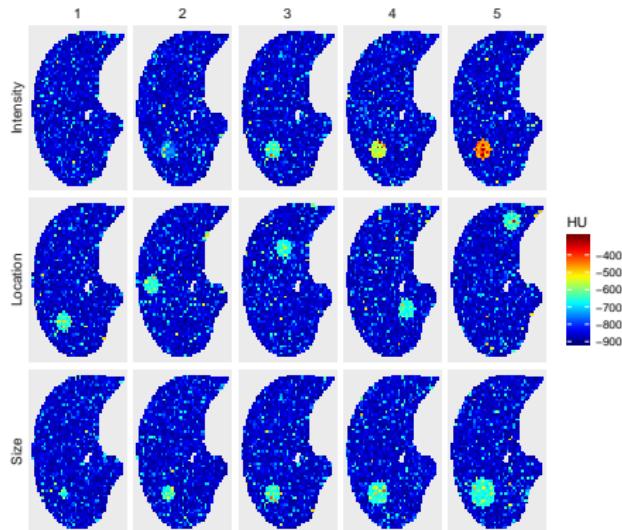
4. Linear Combination



- ReLU used to only keep sections with positive influence
- Normalized between [0, 1], so the most discriminative region is 1
- *Importance* - unit of measurement for the CLAMs

Simulation Study

- Simulated lung slices, varying either the intensity, location or size of the abnormalities
- $N=300$ scans, $k = 3, 4, 5$ clusters
- Ran for 200 epochs in training, then 500 epochs for clustering



Simulation Study

- All simulations resulted in 100% accuracy for cluster assignment, when the number of clusters is set at the true value
- CLAM successfully identified the region of the simulated abnormality in all cases.
- When no simulated abnormality is present, various portions of the *healthy* lung are highlighted

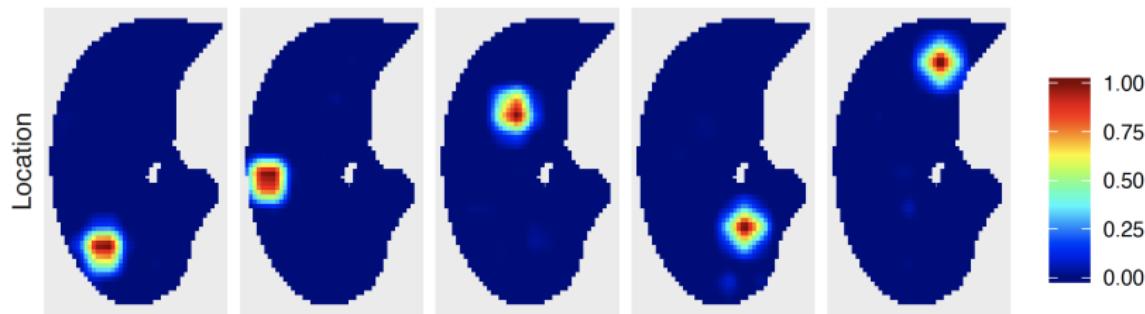


Figure: CLAMs from the simulation when $K = 5$ and location is varied

- Data: GRADS
- Methods:

- ▶ Pre-processing: Segmentation, Registration
- ▶ Resampling: $3mm^3$ voxel spacing
- ▶ CAE structure: $3D \ conv_{32}^2 \rightarrow conv_{64}^1 \rightarrow conv_{128}^2 \rightarrow FC_{60}$, with stride length of 2
- ▶ Ran for 200 epochs in training, then 500 epochs for clustering
- ▶ Estimated the number of clusters using the average silhouette method after training [Rousseeuw, 1987]

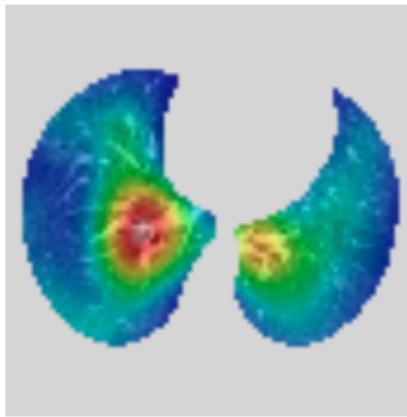
Results

Table: Demographic features from sarcoidosis population by cluster assignment

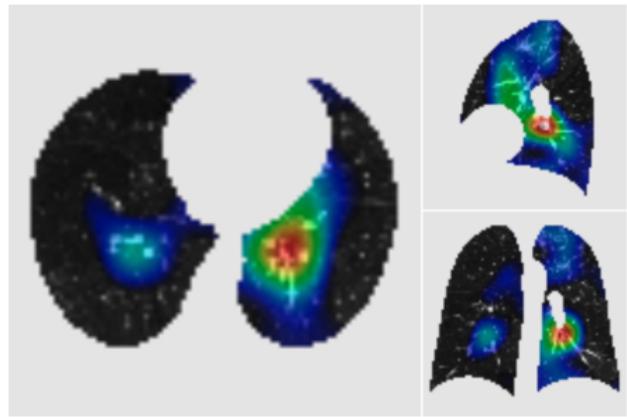
	Cluster 1	Cluster 2	P-value
Sample Size	126	175	
Male (%)	51 (40.5)	91 (52.0)	0.063
White (%)	77 (61.1)	141 (80.6)	<0.001
Hispanic (%)	8 (6.4)	6 (3.4)	0.355
Age (mean (sd))	53.58 (9.32)	52.36 (10.04)	0.287
BMI (mean (sd)))	31.05 (7.05)	30.20 (6.16)	0.267
Height (mean (sd))	66.52 (3.96)	67.49 (4.24)	0.045
Scadding (%)			<0.001
0	7 (5.6)	33 (18.9)	
1	17 (13.5)	42 (24.0)	
2	32 (25.4)	55 (31.4)	
3	16 (12.7)	27 (15.4)	
4	54 (42.9)	18 (10.3)	

CLAM superimposed on Images

Cluster 1



Cluster 2



See in 3D

Clinical Importance

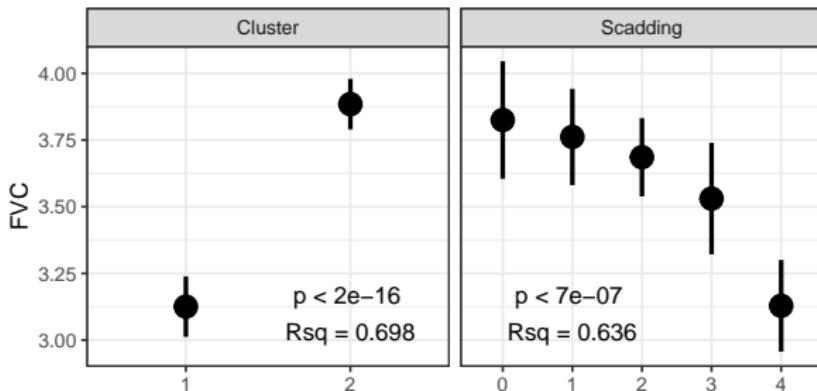


Figure: Associations between forced vital capacity (FVC) with our new clusters and Scadding stage. Results are presented mean (95% CI), adjusted for age, sex, race, BMI, height.

- FEV1, DLCO, SOBQ and SF-12 Physical scores also showed a significant association with our clusters, and explained more variation in these outcomes than Scadding stage.

Chapter 5:

Conclusion

We have developed three novel statistical methodologies for lung CT:



We created a publicly available standard lung template, along with an R software package called lungct.



We developed a population-level spatial modeling approach for lung CT scans, called spVBM.



We developed methodology to cluster lung CTs and to identify the cluster-specific discriminative regions from the images, which we call CLAM.

Together, these methodologies advance the field of lung imaging by providing novel methodologies and publicly available software to perform population-level inference and clustering for lung CT scans

Improved Clinical Understanding



We identified the average size and shape of the lungs from a healthy adult population in the U.S., enabled spatial comparisons across lung scans, and improved existing methodologies by removing biases related to differences in size and shape of the lung.



In pulmonary sarcoidosis, patients with Stage 4 (fibrosis) see increased HU near the hilar compared to patients with Stage 0 (no pulmonary involvement). HU is also significantly associated with lung function at a global level.



Applied to GRADS, we identified two clusters of patients with pulmonary sarcoidosis using just the CT scans, which are better predictors of PFTs and PROs than Scadding stage, the current classification

Acknowledgements

My committee members:

- Dr. Nichole Carlson
- Dr. Tasha Fingerlin
- Dr. Debashis Ghosh
- Dr. John Muschelli
- Dr. Fuyong Xing

My employers:

- Center for Innovative Design and Analysis
- National Jewish Health (Dr. Lisa Maier, Peggy Mroz, Bri Barkes, Dr. Shu-Yi Liao)

My family:

- My fiance, Gabriel
- My parents, Dave and Heidi
- My grandparents, Michael, Shirley, Marvin and Joyce
- My siblings, Jacob, Emily, Trey, Libby, Paul, Lauren, Henry and Michael

Friends in department:

- Harris Butler
- Alex Jensen
- Kevin Josey

- National Institutes of Health (R01 HL114587; R01 HL142049; U01 HL112695)
- GRADS study (NIH grant U01 HL112707, U01 HL112707, U01 HL112694, U01 HL112695, U01 HL112696, U01 HL112702, U01 HL112708, U01 HL112711, U01 HL112712)
- COPDGene study (NIH grants U01 HL089856 and U01 HL089897, COPD Foundation)

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, and Blood Institute or the National Institutes of Health.

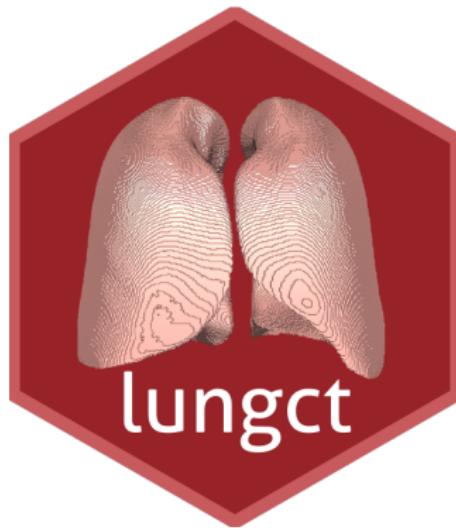
Contact

Email: Sarah.M.Ryan@cuanschutz.edu

Twitter: [@SarahBiostats](https://twitter.com/SarahBiostats)

GitHub: [ryansar](https://github.com/ryansar)

Website: www.SarahMRyan.com



References |

-  Afshar, P., Mohammadi, A., Plataniotis, K. N., Oikonomou, A., and Benali, H. (2018). From hand-crafted to deep learning-based cancer radiomics: Challenges and opportunities. *arXiv preprint arXiv:1808.07954*.
-  Ashburner, J. and Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage*, 11(6):805–821.
-  Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
-  Baughman, R. P., Field, S., Costabel, U., Crystal, R. G., Culver, D. A., Drent, M., Judson, M. A., and Wolff, G. (2016). Sarcoidosis in america. analysis based on health care use. *Annals of the American Thoracic Society*, 13(8):1244–1252.
-  Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, 57(1):289–300.
-  Bookstein, F. L. (2001). "voxel-based morphometry" should not be used with imperfectly registered images. *Neuroimage*, 14(6):1454–1462.
-  Brown, D. A., Lazar, N. A., Datta, G. S., Jang, W., and McDowell, J. E. (2014). Incorporating spatial dependence into bayesian multiple testing of statistical parametric maps in functional neuroimaging. *NeuroImage*, 84:97–112.

References II

-  Evans, A. C., Janke, A. L., Collins, D. L., and Baillet, S. (2012).
Brain templates and atlases.
Neuroimage, 62(2):911–922.
-  Friston, K. and Penny, W. (2003).
Posterior probability maps and spms.
Neuroimage, 19(3):1240–1249.
-  Guo, X., Liu, X., Zhu, E., and Yin, J. (2017).
Deep clustering with convolutional autoencoders.
In *International Conference on Neural Information Processing*, pages 373–382. Springer.
-  Masci, J., Meier, U., Cireşan, D., and Schmidhuber, J. (2011).
Stacked convolutional auto-encoders for hierarchical feature extraction.
In *International Conference on Artificial Neural Networks*, pages 52–59. Springer.
-  Mejia, A. F., Yue, Y., Bolin, D., Lindgren, F., and Lindquist, M. A. (2019).
A bayesian general linear modeling approach to cortical surface fmri data analysis.
Journal of the American Statistical Association, pages 1–26.
-  Moller, D. R., Koth, L. L., Maier, L. A., Morris, A., Drake, W., Rossman, M., Leader, J. K., Collman, R. G., Hamzeh, N., Sweiss, N. J., et al. (2015).
Rationale and design of the genomic research in alpha-1 antitrypsin deficiency and sarcoidosis (grads) study. sarcoidosis protocol.
Annals of the American Thoracic Society, 12(10):1561–1571.
-  Moran, P. A. (1950).
Notes on continuous stochastic phenomena.
Biometrika, 37(1/2):17–23.

References III

-  Murakami, D. and Griffith, D. A. (2019).
Eigenvector spatial filtering for large data sets: fixed and random effects approaches.
Geographical Analysis, 51(1):23–49.
-  Murakami, D., Yoshida, T., Seya, H., Griffith, D. A., and Yamagata, Y. (2017).
A moran coefficient-based mixed effects approach to investigate spatially varying relationships.
Spatial Statistics, 19:68–89.
-  Musgrove, D. R., Hughes, J., and Eberly, L. E. (2016).
Fast, fully bayesian spatiotemporal inference for fmri data.
Biostatistics, 17(2):291–303.
-  Nunes, H., Soler, P., and Valeyre, D. (2005).
Pulmonary sarcoidosis.
Allergy, 60(5):565–582.
-  Regan, E. A., Hokanson, J. E., Murphy, J. R., Make, B., Lynch, D. A., Beaty, T. H., Curran-Everett, D., Silverman, E. K., and Crapo, J. D. (2011).
Genetic epidemiology of COPD (copdgene) study design.
COPD: Journal of Chronic Obstructive Pulmonary Disease, 7(1):32–43.
-  Rousseeuw, P. J. (1987).
Silhouettes: a graphical aid to the interpretation and validation of cluster analysis.
Journal of computational and applied mathematics, 20:53–65.
-  Ryan, S. M., Fingerlin, T. E., Mroz, M., Barkes, B., Hamzeh, N., Maier, L. A., and Carlson, N. E. (2019a).
Radiomic measures from chest hrct associated with lung function in sarcoidosis.
European Respiratory Journal, page 1900371.

References IV

-  Ryan, S. M., Vestal, B., Maier, L. A., Carlson, N. E., and Muschelli, J. (2019b).
Template creation for high-resolution computed tomography scans of the lung in r software.
Academic Radiology.
-  Scadding, J. (1961).
Prognosis of intrathoracic sarcoidosis in england.
British medical journal, 2(5261):1165.
-  Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017).
Grad-cam: Visual explanations from deep networks via gradient-based localization.
In *Proceedings of the IEEE International Conference on Computer Vision*, pages 618–626.
-  Sluimer, I., Schilham, A., Prokop, M., and Van Ginneken, B. (2006).
Computer analysis of computed tomography scans of the lung: a survey.
IEEE transactions on medical imaging, 25(4):385–405.
-  Smith, M. and Fahrmeir, L. (2007).
Spatial bayesian variable selection with application to functional magnetic resonance imaging.
Journal of the American Statistical Association, 102(478):417–431.
-  Wang, H., Du, M., Yang, F., and Zhang, Z. (2019).
Score-cam: Improved visual explanations via score-weighted class activation mapping.
arXiv preprint arXiv:1910.01279.
-  Xie, J., Girshick, R., and Farhadi, A. (2016).
Unsupervised deep embedding for clustering analysis.
In *International conference on machine learning*, pages 478–487.

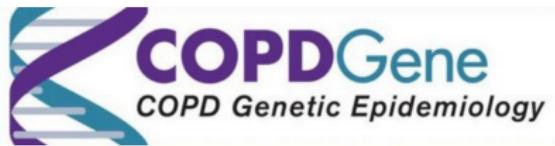
References V



- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016).
Learning deep features for discriminative localization.
In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2921–2929.

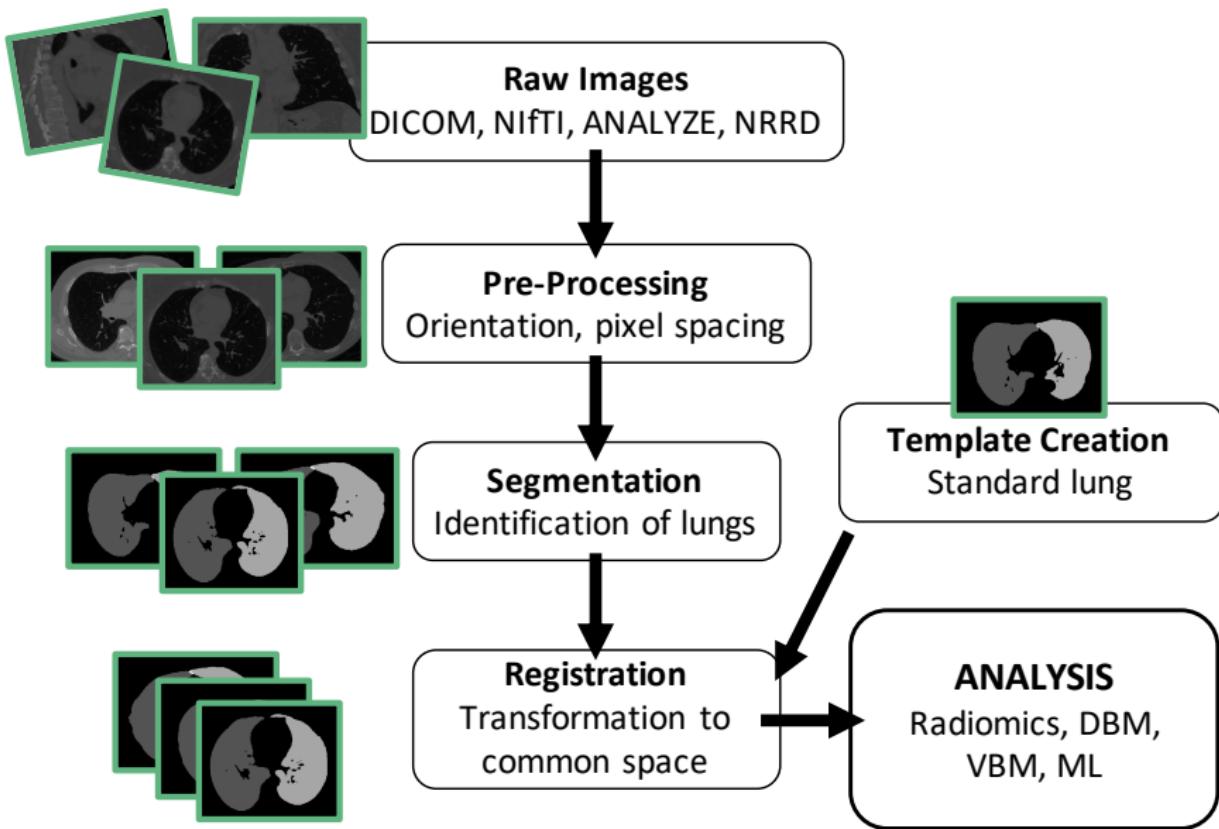
COVID-19 Lung





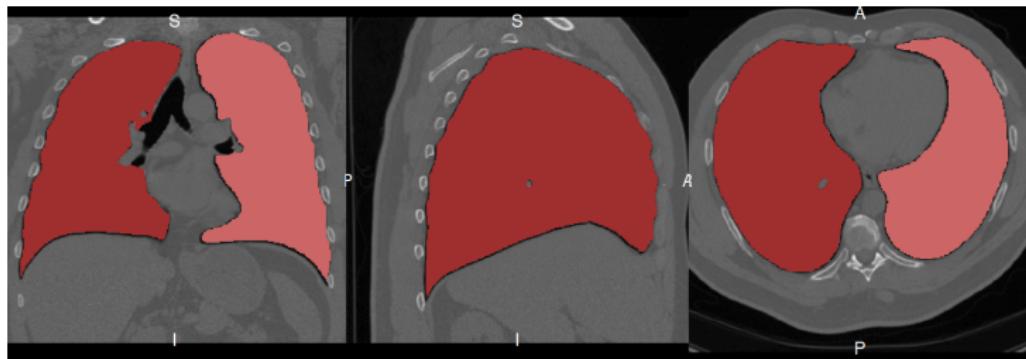
- Sample size: N = 62 adult subjects
 - ▶ Balanced across sex
 - ▶ Mean age 62 years
- Research Chest HRCT
 - ▶ Obtained in raw DICOM
 - ▶ Siemens
 - ▶ Standard B35f kernel
 - ▶ 0.75mm thickness, 0.5mm computed interval
 - ★ Pre-processing: converted to 1x1x1mm

Image Processing Pipeline



Segmentation

- Segmented the left and right lungs from the CT scan using our *lungct* R package
- Methodology includes a combination of thresholding, morphology and region-based methodology
- Compared to proprietary VIDA Diagnostics (Iowa City, IA) -> High level of overlap

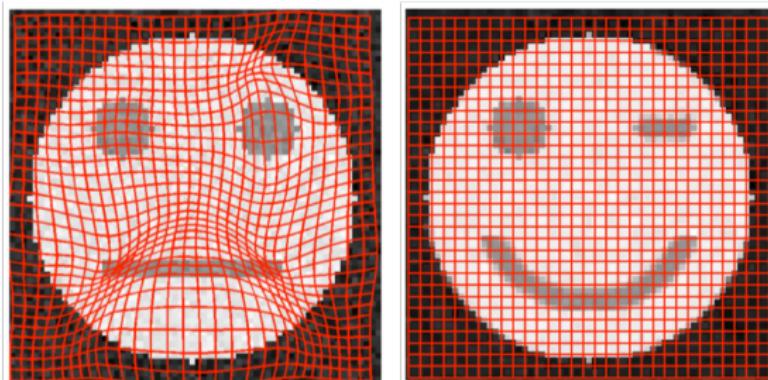


* All figures of the lung are in radiological convention, where the left side of the image is the right lung.

Registration

Registration is the procedure of transforming voxels from their original space into a common space

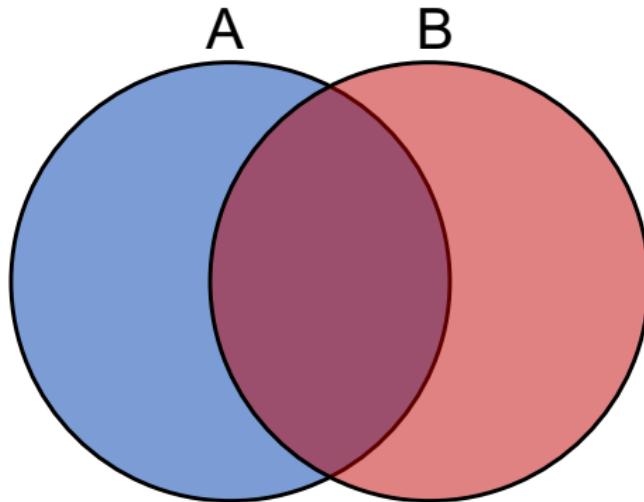
- Left and right lungs were registered separately to account for differences in lung shape and size
- We performed registration on lung masks to preserve biological variability in Hounsfield units (HU)
- Symmetric Normalization (SyN) non-linear registration was used due to its flexibility and success in EMPIRE10



Convergence

We define convergence using a dice similarity coefficient (DSC) between successive iterations of at least 0.99

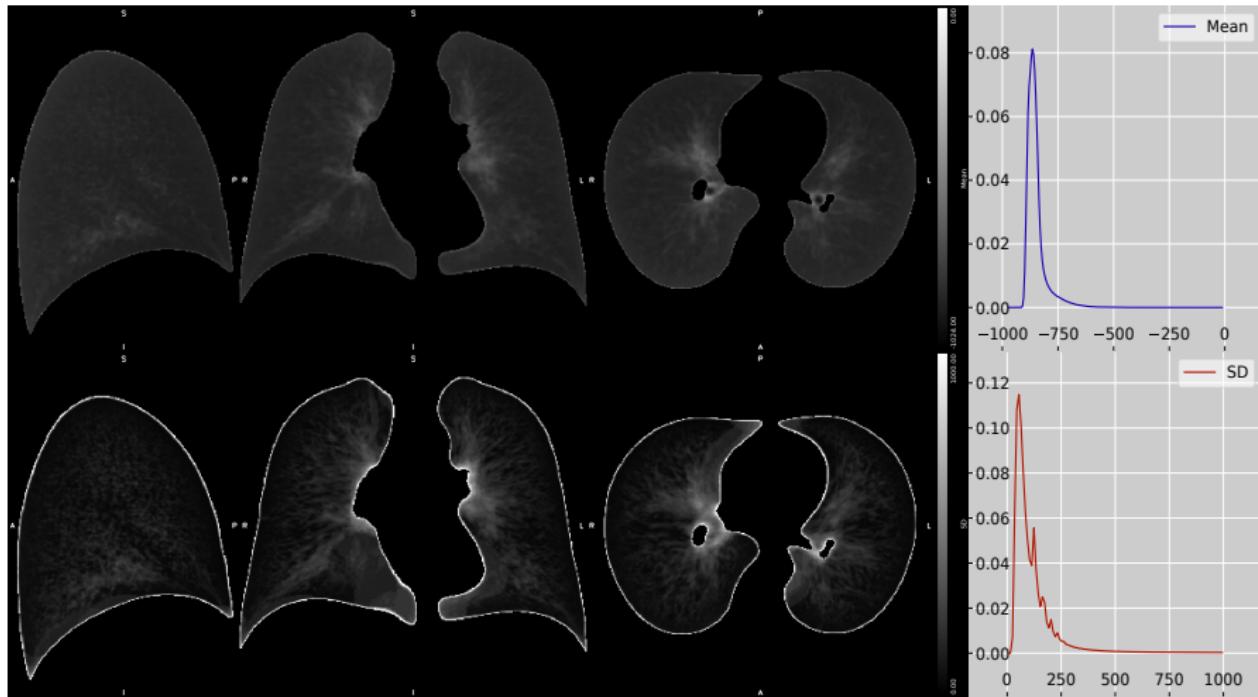
$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|}$$



Healthy Lung Template Characteristics

Number of iterations: 14

Volume: Right lung = 2967 cm³, Left lung 2623 cm³



Creation of a Lobe Template

Procedure:

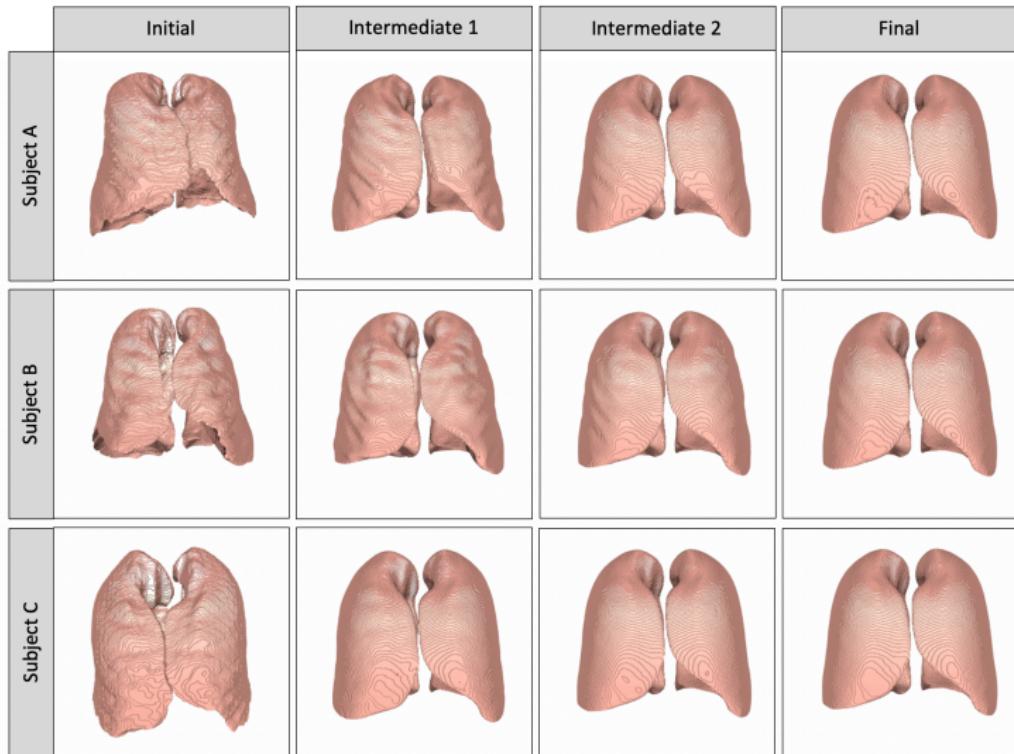
- Apply the final transformations, g_i , obtained from template creation to the lobe segmentations
- Using majority vote at each voxel, we obtain an average lobe mask in template space



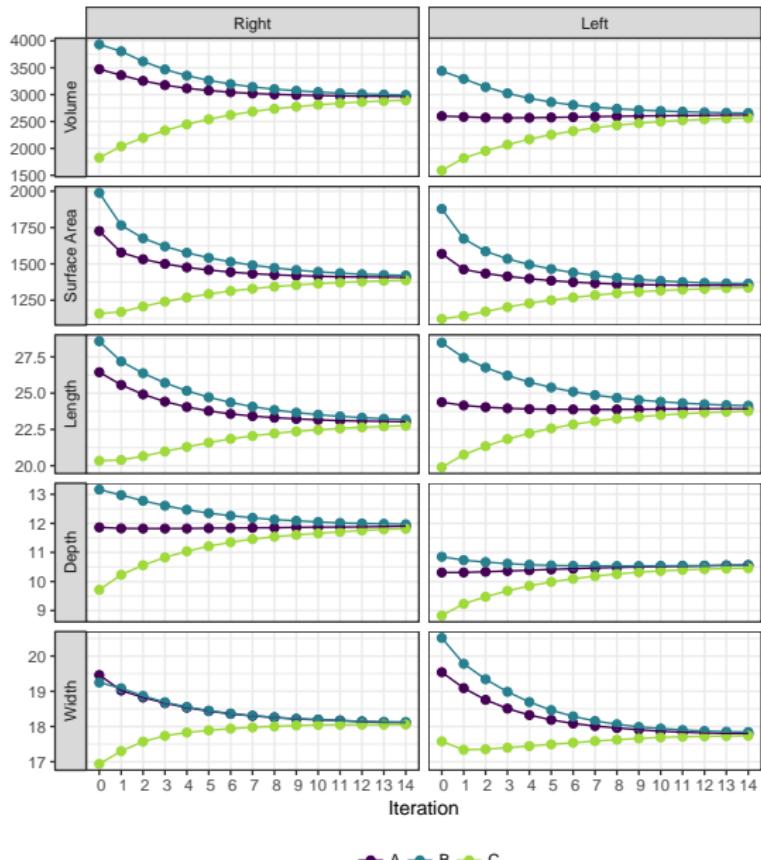
Sensitivity to Initial Template Choice

colorado school of
public health

UNIVERSITY OF COLORADO
COLORADO STATE UNIVERSITY
UNIVERSITY OF NORTHERN COLORADO



Sensitivity to Initial Template Choice



Effect of Registration on Radiomics

- **Data:** CT scans from $N = 102$ COPDGene healthy controls [Regan et al., 2011] and $N = 321$ GRADS sarcoidosis patients [Moller et al., 2015]
- **Methods:** Calculation of first-order and GLCM radiomic features, before and after registration



Figure: Example CT registered to the standard lung, showing the sagittal, coronal, and axial views with the Hounsfield unit (HU) histogram. Mean HU = -823, SD HU = 199, Skew = 3.36, Kurtosis = 10.7, Contrast = 5.26, Autocorrelation = 13.5, Energy = 0.19, Homogeneity = 0.75.

Effect of Registration on Radiomics

- Registration significantly affected first-order and GLCM radiomic features, by decreasing mean HUs and variability across voxels, and increasing smoothness.
- However, the registration improved detective ability of radiomics, reducing noise and enhancing signal [Ryan et al., 2019b].

Outcome	Registration Effect (Post- v. Pre-registration)		Group Effect (Control v. Sarcoidosis)	
	Controls	Sarcoidosis	Pre-Registration	Post-Registration
Mean	-6.22 (-7.97, -4.47)	-8.69 (-9.70, -7.69)	-12.37 (-20.24, -4.51)	-9.90 (-17.76, -2.04)
SD	-18.44 (-19.74, -17.13)	-21.93 (-22.68, -21.17)	-15.94 (-20.38, -11.49)	-12.45 (-16.89, -8.01)
Skew	0.51 (0.46, 0.55)	0.46 (0.44, 0.49)	0.49 (0.34, 0.64)	0.53 (0.38, 0.68)
Kurtosis	5.93 (5.52, 6.34)	4.49 (4.26, 4.73)	4.33 (3.01, 5.65)	5.77 (4.45, 7.08)
Contrast	-4.46 (-4.70, -4.22)	-5.17 (-5.31, -5.03)	0.36 (-0.51, 1.23)	1.07 (0.20, 1.94)
Autocorrelation	2.15 (2.00, 2.30)	3.14 (3.05, 3.23)	0.09 (-0.34, 0.52)	-0.90 (-1.33, -0.47)
Energy (x1000)	1.37 (1.28, 1.45)	1.72 (1.67, 1.77)	-0.05 (-0.26, 0.17)	-0.40 (-0.61, -0.18)
Homogeneity (x100)	4.44 (4.22, 4.66)	5.21 (5.09, 5.34)	-0.15 (-0.89, 0.60)	-0.92 (-1.67, -0.17)

Proof-of-Concept

- I fit a simplified two-stage deep clustering algorithm to sarcoidosis subjects (GRADS), where I first optimized an autoencoder, then clustered the extracted features using k-means

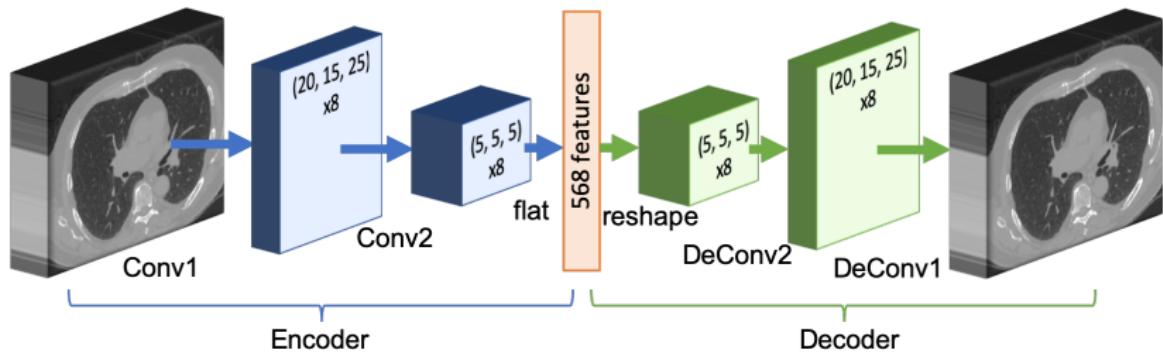


Figure: The autoencoder used in the proof-of-concept study, with 3D convolutional and deconvolutional layers, and a flattening layer prior to the feature extraction (colored by orange).

Proof-of-Concept: Results

- I compared the clusters I obtained from the learned radiomic features to classical radiomics and visual scores, to see which clusters would explain the most variation in lung function

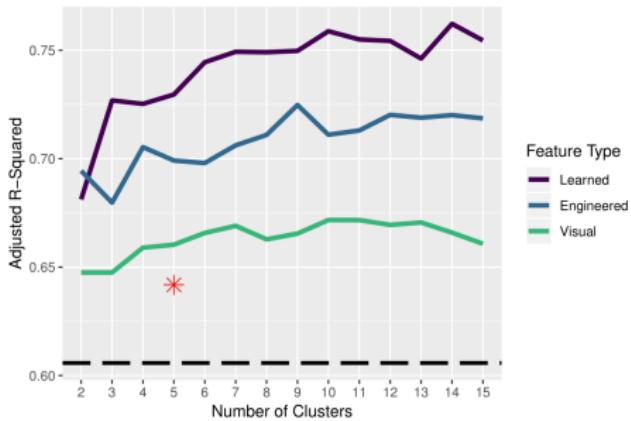


Figure: Amount of variance explained in forced vital capacity. The dashed horizontal black line indicates the base adjusted R^2 for the linear regression adjusted for gender, race, age, and height. The red star is indicative of the adjusted R^2 when the five-class Scadding stage is added to the base model.

Table: Amount of variation explained in the outcomes (i.e. adjusted r-square) for the base model (age, gender, race, BMI and height), base with the new cluster covariate, and base with the existing Scadding stage classification.

	Base	Clusters	Scadding
FEV1	0.484	0.616	0.583
FVC	0.596	0.698	0.636
DLCO	0.498	0.600	0.549
CD4 (log)	0.055	0.099	0.133
CRP (log)	0.024	0.021	0.024
FAS	0.034	0.036	0.014
GERDQ	0.011	0.013	0.003
CFQ	0.029	0.035	0.022
SOBQ	0.088	0.148	0.139
Promis	0.012	0.010	0.008
SF12 - Physical	0.095	0.167	0.130
SF12 - Mental	0.043	0.049	0.047