Unsupervised Discovery of Novel Emphysema Subtypes

Andrew Francis Laine, D.Sc.
Percy K. and Vida L. W. Hudson Professor of Biomedical Engineering and
Professor of Radiology (Physics)
Department of Biomedical Engineering
Columbia University
New York, NY 10027

Presented at
New Jersey Public Health Annual Conference:
“Pressing Public Health Challenges: Systems Thinking, Opioid Crisis, and Use of Artificial Intelligence”
Rutgers Continuing Education Center at Atrium, Division of Continuing Education
October 4, 2019, Somerset, NJ
Acknowledgements

Collaborators

Elsa D. Angelini, Ph.D.
R. Graham Barr, M.D., Ph.D.

Heffner Biomedical Imaging Lab

Jie Yang (Ph.D. Student, 2019)
Xinyang Feng (Ph.D. student, 2019)
Yrjö Häme (Ph.D. student)
Yu Gan (Post-doc)
Thomas Vetterli (M.Sc. Intern)

Additional Collaborators

Columbia University Medical Center:
Pallavi P. Balte, Ph.D.
John H.M. Austin, M.D.
Benjamin M. Smith, M.D.
Yifei Sun, Ph.D.
Wei Shen, M.D.

Iowa University:
Eric A. Hoffman, Ph.D.

University of Virginia:
Ani Manichaikul, Ph.D.

And the MESA, SPIROMICS Investigators!
Emphysema

- Emphysema + COPD is 3rd leading cause of death in USA.
- Defined by loss of interalveolar septa
- Predicts mortality in patients with and without COPD

Leopold/Gough, Thorax, 1957
Johannessen, AJRCCM, 2013
Oelsner, Ann Intern Med, 2014
Computed tomography (CT) used to analyze lung structure:

**Resolution** = 0.5×0.5×0.75 mm

**Matrix size** = 512×512×500 pixels

**Intensity range** = [-1024 1024] HU

40 megavoxels of the lung:
- Enable *in vivo* study of lung structure and disease patterns.
Lung tissue abnormalities on CT:
- Characterized by **localized texture patterns**.

<table>
<thead>
<tr>
<th>COPD and Pulmonary Emphysema</th>
</tr>
</thead>
<tbody>
<tr>
<td>❑ 4th leading cause of death in the world [1]:</td>
</tr>
<tr>
<td>❑ Affects 16 millions of subjects in the US.</td>
</tr>
<tr>
<td>❑ Emphysema: Lower attenuation abnormality.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lung Cancer and Nodule</th>
</tr>
</thead>
<tbody>
<tr>
<td>❑ 1st leading cause of cancer-related death in the US [2].</td>
</tr>
<tr>
<td>❑ &gt; 150,000 deaths in the US in 2018.</td>
</tr>
<tr>
<td>❑ Lung nodule: Higher attenuation abnormality.</td>
</tr>
</tbody>
</table>

---

Lung texture learning to characterize emphysema subtypes:

COPD and Emphysema

- Exact mechanism of developing COPD remains unknown;
- Three standard emphysema subtypes defined at autopsy [1]:
  - Limited inter-rater agreement.

Lung texture learning for emphysema subtyping can advance disease understanding

Emphysema Subcategories

- Centrilobular Emphysema (CLE)
- Panlobular Emphysema (PLE)
- Paraseptal Emphysema (PSE)

N = 140
N = 2

Classic Emphysema Subtypes

- Panlobular
- Centrilobular
- Paraseptal
Limitations of existing CT-based texture analysis:

⚠️ Limited to **supervised** learning;

⚠️ Limited to **texture** features, without considering **spatial** locations;

---

**Background**

Emphysema subtypes:
- Different in spatial prevalence [1].

Lung nodules:
- Location predicts malignancy [2].

---


Limitations of existing CT-based texture analysis:

- Limited to \textit{supervised} learning;
- Limited to \textit{texture} features, without considering \textit{spatial} locations;
- Limited to \textit{high-resolution full-lung} CT scans.

\textbf{Large dataset} of cardiac CT scans are available:

- 10,000s longitudinal cardiac CT scans + gold standard full-lung CT scans in MESA \cite{Bild2002};
- \textasciitilde2/3 of the lung region + reproducible \%\textit{emphasis} \cite{Hoffman2009};
- Enable large-scale \textit{longitudinal study}.

\cite{Bild2002, Hoffman2009}
SPIROMICS and MESA

- **SubPopulations and Intermediate Outcome Measures In COPD Study**
  - COPD case-control study
  - 2,983 participants with CT scans (~6,000 scans)
  - Whole genome sequencing, multi-omics

- **Multi-Ethnic Study of Atherosclerosis Lung Study**
  - Population-based, prospective cohort study
  - 3,205 participants with full-lung CT scans (~50,000 scans)
  - Whole genome sequencing, multi-omics
Hypothesis

• Unsupervised learning of spatial lung texture patterns on research CT scans will yield novel emphysema subtypes.
  – Reproducible
  – Distinct symptoms
  – Specific histology and genetic basis
Aimed to tackle the problem of CT-based lung texture learning exploiting spatial localization, using unsupervised / weakly-supervised learning.

**Aim 1:** Develop an algorithm for unsupervised learning of localized texture patterns for emphysema.

**Aim 2:** Label the discovered localized texture patterns on large datasets of cardiac CT scans.

**Aim 3:** Examine possible correlations / hits with GWAS genomic information in MESA and SPIROMICS.
Data

<table>
<thead>
<tr>
<th>Study</th>
<th># of subjects</th>
<th># of CT scans</th>
<th># of scanner types</th>
<th>Longitudinal follow-up (years)</th>
<th>Pulmonary disease to study</th>
</tr>
</thead>
<tbody>
<tr>
<td>MESA</td>
<td>6,814</td>
<td>31,228 cardiac</td>
<td>11</td>
<td>12</td>
<td>Emphysema &amp; COPD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3,131 full-lung</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MESA COPD</td>
<td>321</td>
<td>317 full-lung</td>
<td>3</td>
<td>5</td>
<td>Emphysema &amp; COPD</td>
</tr>
<tr>
<td>SPIROMICS</td>
<td>3,200</td>
<td>3,200 full-lung</td>
<td>9</td>
<td></td>
<td>Emphysema &amp; COPD</td>
</tr>
<tr>
<td>LIDC-IDRI</td>
<td>1,010</td>
<td>1,018 full-lung</td>
<td>17</td>
<td></td>
<td>Nodule &amp; Lung Cancer</td>
</tr>
<tr>
<td>Kaggle DSB2017</td>
<td>2,101</td>
<td>2,101 full-lung</td>
<td>unknown</td>
<td></td>
<td>Nodule &amp; Lung Cancer</td>
</tr>
</tbody>
</table>

- **Lung masks:**
  - VIDA Diagnostics APOLLO®;
  - **Emphysema masks** for full-lung scans:
    - Hidden Markov measure field (HMMF) based model \([1]\): \(\text{emp}_{HMMF}\);
    - Intensity thresholding: \(\text{emp}_{-950}\);
  - **Lung masks**:
    - Intensity thresholding < -400 HU \([2]\) + closed space dilation \([3]\).

---

Background: Processing Overview

Unsupervised machine learning defined novel emphysema patterns on CT.

From 3 standard subtypes (1950’s) to 10 novel spatial lung texture patterns (sLTP) for emphysema.

Yang et al., Unsupervised Discovery of Spatially-Informed Lung Texture Patterns for Pulmonary Emphysema: The MESA COPD Study. MICCAI, 2017
### Contributions

- Applied a method to standardize **lung shape spatial mapping**;
- Developed a two-stage **unsupervised** framework combining **spatial and texture** information;
- Discovered emphysema patterns on **large COPD cohorts** with compelling clinical significance.

### Challenges

- Learning **localized texture** patterns in an **unsupervised** manner;
- **Homogeneity** vs. **redundancy** of the learned patterns.

---

[1] Jie Yang et al., Unsupervised Machine Learning to Define Quantitative Subtypes of Pulmonary Emphysema on CT. *Science* (*to submit in December 2018*).

Method
Unsupervised learning of localized texture patterns for pulmonary emphysema

1 Lung Shape Spatial Mapping

\[(x, y, z) \rightarrow (r, \theta, \phi)\]

- **Poisson Distance Map** \(^{[1]}\) (PDM): \(r = 1 - U\)

  \[
  \Delta U(x, y, z) = -1, \quad \text{for } (x, y, z) \in V
  \]

  \[
  \text{subject to } U(x, y, z) = 0, \quad \text{for } (x, y, z) \in \partial V
  \]

  \[
  \Delta U = U_{xx} + U_{yy} + U_{zz}
  \]

  Lung mask boundary

- **Poisson Distance Conformal Map** (PDCM): \(\theta \& \phi\)

Preliminary Evaluation of PDCM
Unsupervised learning of localized texture patterns for pulmonary emphysema

PDCM for Studying Emphysema Spatial Locations over Populations

MESA COPD Study:
- N = 317 CT scans;
- Angular and radial projections of intensity or relative intensity values;
- Agree with definitions or observations of standard emphysema subtypes.

PDCM = a useful tool for population study of spatial patterns.
Method
Unsupervised learning of localized texture patterns for pulmonary emphysema

Learning Stage 1:
- Augmented Lung Texture Patterns (LTPs)

Region of interest (ROI) = $25 \times 25 \times 25$ mm$^3$

Texture feature $FT$ = texton$^1$ - based feature

Spatial feature $FS$ = location in 36 lung sub-regions $(r/3, \theta/4, \phi/3)$

$FT_x / FS_x = \text{texture / spatial feature of ROI } x$

$FT_k / FS_k = \text{texture / spatial centroid of } LTP_k$

Learning Stage 1:  
- Augmented Lung Texture Patterns (LTPs)

Iteratively update ROI assignment $\Lambda_k^{(t)}$ of LTP$_k$, by \textit{minimizing a dedicated cost function} \[^1\]:

$$
\chi^2 \left( FT_x, \overline{FT_k}^{(t-1)} \right) + 
\omega \cdot W \cdot \left\| FS_x, \overline{FS_k}^{(t-1)} \right\|^2 + 
\gamma \cdot \mathbb{I} \left( \chi^2 \left( FT_x, \overline{FT_k}^{(t-1)} \right) > \text{thresh}_{\chi^2} \right)
$$

Texture distance  
Spatial regularization  
Texture penalty

$FT_x / FS_x$ = texture / spatial feature of ROI $x$  
$\overline{FT_k} / \overline{FS_k}$ = texture / spatial centroid of LTP$_k$
Learning Stage 2:

> Spatially-informed lung texture patterns (sLTPs)

**Similar LTP:** can be replaced by each other.

\[ N_{i \rightarrow j} = \# \text{ of ROIs labeled with } LTP_j \text{ when removing } LTP_i. \]

**Infomap** \[^1\] **Graph Partitioning of LTP similarity:**

\[ G_{i,j} = \frac{N_{i \rightarrow j} + N_{j \rightarrow i}}{N_i + N_j} \cdot \mathbb{1} \left( \sum_k N_{i \rightarrow k} > \text{thresh}_{N \rightarrow} \right) \cdot \mathbb{1} \left( \sum_k N_{j \rightarrow k} > \text{thresh}_{N \rightarrow} \right) \]

Learning Pipeline and Results in SPIROMICS and MESA Lung Study

Unsupervised learning of localized texture patterns for pulmonary emphysema

DATA: SPIROMICS (N=2,922)

Learning Reproducibility

- Regional level:
  - Proportion of labeling overlap of test ROIs = 0.82.

- Individual level:
  - Spearman’s correlation of sLTP histograms over all subjects (N = 2,922)
  - All sLTP > 0.95
Learning Pipeline and Results in SPIROMICS and MESA Lung Study
Unsupervised learning of localized texture patterns for pulmonary emphysema

DATA: SPIROMICS (N=2,922) and MESA Lung Study (N=3,128)

Unsupervised Learning of sLTPs

Population Clustering

Heatmap of %sLTP label histograms over populations

10 sLTP to 6 QES

Quantitative Emphysema Subtype (QES) 1-6

Population-based sLTP Merging

High-Resolution CT Images

Feature Extraction

Textons

Textures

Spatial Feature

Frequency

Anterior

Superior

Lateral

Core

Medial

Posterior

Interior

Apical

Vanishing Lung

Obstructive CPFE

Restrictive CPFE

Diffuse

Senile

SPIROMICS N = 2,922

sLTP 1

sLTP 2

sLTP 3

sLTP 4

sLTP 5

sLTP 6

sLTP 7

sLTP 8

sLTP 9

sLTP 10

MESA Lung N = 3,128

sLTP histogram

sLTP histogram

All SPIROMICS (N = 2,922)
Population Distribution of the QES

### Summary of Clinical Significance in SPIROMICS and MESA:
- Association of QES with respiratory symptoms, physiology, and prognosis

### Experimental Results in SPIROMICS and MESA Lung Study

Unsupervised learning of localized texture patterns for pulmonary emphysema

#### SPIROMICS (n=2,853)

<table>
<thead>
<tr>
<th></th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRC-Dyspnea</td>
<td>0.1</td>
<td>-0.001</td>
<td>0.1</td>
<td>0.2</td>
<td>-0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>SGRQ Score</td>
<td>1.1</td>
<td>0.8</td>
<td>-0.4</td>
<td>4.2</td>
<td>-1</td>
<td>0.4</td>
</tr>
<tr>
<td>Resting O2 saturation (%)</td>
<td>0.004</td>
<td>-0.4</td>
<td>0.1</td>
<td>-0.9</td>
<td>0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>6 Minute Walk Test Distance (m)</td>
<td>-0.3</td>
<td>-2.3</td>
<td>5.2</td>
<td>-22.9</td>
<td>5.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.06</td>
<td>0.06</td>
<td>0.03</td>
<td>0.15</td>
<td>-0.05</td>
<td>0.07</td>
</tr>
<tr>
<td>Exacerbations</td>
<td>0.1</td>
<td>0.2</td>
<td>0.03</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

#### MESA Lung (n=2,694)

<table>
<thead>
<tr>
<th></th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
<th>β Est.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRC-Dyspnea</td>
<td>0.1</td>
<td>-0.01</td>
<td>-0.002</td>
<td>0.2</td>
<td>-0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Resting O2 saturation (%)</td>
<td>-0.5</td>
<td>-0.3</td>
<td>0.2</td>
<td>-0.8</td>
<td>0.3</td>
<td>-2.2</td>
</tr>
<tr>
<td>FEV1 (mL)</td>
<td>-20.0</td>
<td>18.8</td>
<td>14.6</td>
<td>-126.5</td>
<td>-82.9</td>
<td>817.6</td>
</tr>
<tr>
<td>FVC (mL)</td>
<td>-150.2</td>
<td>153.7</td>
<td>102.8</td>
<td>-100.3</td>
<td>-56.0</td>
<td>1149.6</td>
</tr>
<tr>
<td>FEV1/FVC (%)</td>
<td>-6.9</td>
<td>-2.6</td>
<td>-2.6</td>
<td>-0.9</td>
<td>-2.0</td>
<td>8.1</td>
</tr>
<tr>
<td>Total Lung Volume (mL)</td>
<td>-64.2</td>
<td>485.3</td>
<td>333.8</td>
<td>-378.4</td>
<td>145.9</td>
<td>1781.7</td>
</tr>
</tbody>
</table>

#### MESA (n=6,683)

<table>
<thead>
<tr>
<th></th>
<th>HR</th>
<th>HR</th>
<th>HR</th>
<th>HR</th>
<th>HR</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLRD Hospitalization</td>
<td>2.9</td>
<td>1.5</td>
<td>0.8</td>
<td>1.0</td>
<td>1.4</td>
</tr>
<tr>
<td>CLRD Mortality</td>
<td>2.2</td>
<td>1.5</td>
<td>0.9</td>
<td>0.99</td>
<td>1.8</td>
</tr>
<tr>
<td>All-cause Mortality</td>
<td>1.6</td>
<td>0.9</td>
<td>0.96</td>
<td>1.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Red shading = statistically significant worsening;
Blue/green shading = statistically significant “improvement”.

MRC= Medical Research Council; O2=Oxygen; SGRQ=St. George's respiratory questionnaire; 6MW=Six minute walk; FEV1= Forced expiratory volume in one second; FVC=Forced expiratory volume in one second; HR=Hazards ratio; CLRD=Chronic lower respiratory disease

β estimates compared to normal lung from multivariate linear regression models adjusted for age, sex, race, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV1, other QES
Experimental Results in SPIROMICS and MESA Lung Study
Unsupervised learning of localized texture patterns for pulmonary emphysema

Associated with
Dyspnea
Desaturation on exertion
↓ 6MWT
Exacerbations, COPD death

↓↓ FEV₁
↓ FVC
↓ FEV₁/FVC
- TLV on CT

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV₁, other QES.
Experimental Results in SPIROMICS and MESA Lung Study

Unsupervised learning of localized texture patterns for pulmonary emphysema

-associated with

Hypoxemia at rest
Desaturation on exertion
Exacerbations, COPD death

FEV₁, FVC, FEV₁/FVC, TLV on CT

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV₁, other QES.
Experimental Results in SPIROMICS and MESA Lung Study

Unsupervised learning of localized texture patterns for pulmonary emphysema

Associated with

- FEV₁
- FVC
- FEV₁/FVC
- TLV on CT

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV₁, other QES.
Experimental Results in SPIROMICS and MESA Lung Study

Unsupervised learning of localized texture patterns for pulmonary emphysema

Associated with
Dyspnea
Hypoxemia at rest
Desaturation on exertion
↓↓6MWT
Exacerbations

↓ FEV₁
FVC
FEV₁/FVC
↓↓ TLV on CT

Restricted CPFE

Apical  Diffuse  Senile  Restrictive CPFE  Obstructive CPFE  Vanishing Lung

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV1, other QES.
Experimental Results in SPIROMICS and MESA Lung Study
Unsupervised learning of localized texture patterns for pulmonary emphysema

Associated with

- Desaturation on exertion
- COPD death

Obstructive CPFE

- FEV$_1$
- FVC
- FEV$_1$/FVC
- TLV on CT

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV1, other QES.
Vanishing Lung

Associated with
Dyspnea
Desaturation on exertion

▲▲ FEV₁
▲▲ FVC
FEV₁/FVC
▲▲ TLV on CT

Adjusted for age, sex, race/ethnicity, height, weight, smoking status, pack-years, COPD, scanner manufacturer, FEV₁, other QES.
• GWAS results:
  • 5 genetic variants for four QES
  • Apical QES: DRD1
Summary
Unsupervised learning of localized texture patterns for pulmonary emphysema

- Novel **unsupervised learning** of emphysema patterns on CT:
  - A standardized lung shape **spatial mapping**;
  - A **two-stage** learning framework.

- Applied on large COPD and controls yielded:
  - 10 **highly-reproducible** sLTPs;
  - Six quantitative emphysema subtypes, associated independently with distinct **symptoms, lung function changes** and **mortality**.

- Enables:
  - **Novel definitions** of emphysema subtypes;
  - CT-based **emphysema-specific signatures (biomarkers)** of the lungs;
  - May facilitate future study for understanding COPD and emphysema, and the design of personalized / gene / drug therapies.

---

**Other Evaluations:**
- GWAS; 3 “hits” reproducible
- Extensive evaluation of sLTP reproducibility in MESA COPD (N = 317);
- Linking sLTP and standard emphysema subtypes in MESA COPD (N=317).
Labeling the Unsupervised Localized Texture Patterns on Large Datasets of Cardiac CT Scans

Overview of Method

- Introduced a robust **emphysema segmentation framework** on **cardiac** CT scans;
- Proposed a deep learning based **domain adaptation** \([2]\) for robust lung texture learning.
- Labeled emphysema texture patterns on 17,039 longitudinal cardiac and full-lung CT scans.

**Challenges working with cardiac CT Scans**

- Missing apical regions;
- Degraded texture quality;
- Heterogeneous scanner types.

---


Method
Labeling the unsupervised localized texture patterns on large datasets of cardiac CT scans

1. Adaptation of HMMF-based method for emphysema segmentation on cardiac CT scans [1]

HMMF-based segmentation [2]:
1. Parametric models $P(I|q, \theta)$ of intensity distributions for emphysema and normal tissue classes.
2. Enforces spatial coherence via $P_q(q)$:

$$P(q, \theta|I) = \frac{1}{R} P(I|q, \theta) P_q(q) P_\theta(\theta)$$

$R = \text{constant}$
$q = \text{measure field}$
$\theta = \text{likelihood parameter}$
Method
Labeling the unsupervised localized texture patterns on large datasets of cardiac CT scans

Adaptation of HMMF-based method for emphysema segmentation on cardiac CT scans

1. **Parametric models**
   
   \[ P(I|q, \theta) \]

   - Enforces **spatial coherence** via \( P_q(q) \);

2. **Scanner-specific**
   - **Subject-specific**

   \[ \theta_E = [\mu_E, \sigma_E] \]

   **Emphysema:**
   - Normal distribution

   **Normal lung tissue:**

   - \( \mu_N \)

**HMMF-based segmentation** [2]:

- **C = clique**: 8-connected in 2D.
- **\( \lambda = \text{Markovian weight} \):**
  - scanner-specific.
  - Use longitudinal scans of healthy population to tune.
Method
Labeling the unsupervised localized texture patterns on large datasets of cardiac CT scans

2 CNN model to classify ROIs from synthetic cardiac CT scans
## Method

Labeling the unsupervised localized texture patterns on large datasets of cardiac CT scans

2. **CNN model to classify ROIs from synthetic cardiac scans**

### Feature output:
- **kernel number** @ feature map size

### Operators:
- **Convolution (Conv)**
  - kernel size = 3x3x3
  - stride = 1
- **Max-pooling**
  - kernel size = 3x3x3
  - stride = 2
- **Fully-connected (FC)**
- **Rectified linear unit (ReLU)**
- **Softmax**
Method
Labeling the unsupervised localized texture patterns on large datasets of cardiac CT scans

Unsupervised domain adaptation with adversarial learning (UDAA) to classify ROIs from real cardiac scans

Source domain $S$
- ROIs ! : from synthetic cardiac scans
- ROIs ! #: from real cardiac scans

Target domain $T$

Loss function:

\[
L_{\text{total}} = L_{\text{class}} - \alpha \cdot L_{\text{domain}}
\]

\[
\begin{align*}
L_{\text{class}} &= - \sum_{c=1}^{N_{\text{S-LTP}}} y_c \cdot \log(\hat{y}_c) \\
L_{\text{domain}} &= - y_d \cdot \log(\hat{y}_d) - (1 - y_d) \cdot \log(1 - \hat{y}_d)
\end{align*}
\]
Experimental Results: HMMF-based Emphysema Segmentation
Labeling the unsupervised lung texture patterns on large datasets of cardiac CT scans

Data - MESA Lung Study:
• N = 6,814 subjects in Exam 1 – Exam 5 (2000 – 2012)
• Exam 1-4: two repeated cardiac CT scans per visit;

Higher intra-class correlation (ICC) on repeated cardiac scans

Higher Dice overlap on repeated cardiac scans in Exam 1-4 (N_{\text{disease}}=471)

Example of emphysema spatial overlap:
HMMF => less FN and less FP.
Data - MESA Lung Study:
• N = 6,814 subjects in Exam 1 – Exam 5 (2000 – 2012)
• Longitudinal cardiac scans in Exam 1-4 and full lung scan in Exam 5.

Experimental Results: HMMF-based Emphysema Segmentation
Labeling the unsupervised lung texture patterns on large datasets of cardiac CT scans.

Higher pairwise Pearson’s correlation $r$ on longitudinal cardiac scans within 18 months ($N_{normal}=478$).

Steadier emphysema longitudinal progression ($N_{normal}=87; N_{disease}=238$)
• $\Delta(t) = %\text{emph}(t) - %\text{emph}(\text{baseline})$
Experimental Results – UDAA-based Lung Texture Learning

Fig. 1: (A) Illustration of the UDAA framework: (a) Generation of synthetic cardiac CT scans; (b) CNN architecture for sLTP labeling on synthetic cardiac CT scans; (c) Domain adaptation component to learn discriminative image features between synthetic and real cardiac scans.

### 3. RESULTS

#### 3.1. Experimental Setting

We use all full-lung HRCT scans and cardiac CT scans in MESA that have ULN values of percent emphysema [11].

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td># of Helical scans (train/val/test)</td>
<td>-/-/-</td>
<td>-/-/-</td>
<td>-/-/-</td>
<td>-/-/-</td>
<td>1,146/422/414</td>
</tr>
<tr>
<td># of MDCT scans (train/val/test)</td>
<td>934/329/2,245</td>
<td>408/150/983</td>
<td>423/151/699</td>
<td>152/54/241</td>
<td>-/-/-</td>
</tr>
<tr>
<td># of EBT scans (train/val/test)</td>
<td>748/260/2,167</td>
<td>271/102/967</td>
<td>459/152/772</td>
<td>245/83/380</td>
<td>-/-/-</td>
</tr>
<tr>
<td>Total # of scans evaluated</td>
<td>6,683</td>
<td>2,881</td>
<td>2,656</td>
<td>1,155</td>
<td>1,982</td>
</tr>
</tbody>
</table>

### Table 1: MESA cardiac CT exams along with splits used to train the domain adaptation module.

### 2.3.2. Domain Discrimination in Longitudinal Setting

In our longitudinal setting, lung mask bounding box (hence not using ACC stages. The weight $\epsilon_{\text{domain}}$ is set to 10, and $p$ is the training progress, linearly increasing from 0 to 1. This strategy allows the $NN_d$ to be less sensitive to noisy signal at the early training stages. The weight $\epsilon_{\text{max}}$ is determined by maximizing the sLTP classification accuracy on synthetic ROIs $\epsilon_{\text{class}}$.

\[
\epsilon = 1 + \exp(-\gamma \cdot p)
\]

where $\gamma$ was set to 10, and $p$ is the training progress, linearly increasing from 0 to 1. This strategy allows the $NN_d$ to be less sensitive to noisy signal at the early training stages. The weight $\epsilon_{\text{max}}$ is determined by maximizing the sLTP classification accuracy on synthetic ROIs $\epsilon_{\text{class}}$.

\[
\epsilon_{\text{class}} = \frac{|d_i - d_j|}{\epsilon_{\text{dom}}} \leq \epsilon_{\text{class}}
\]

where $\epsilon$ is a positive weight that defines the relative importance of the domain-adaptation task for the sLTP classifier. During training, $\epsilon$ is initiated at 0 and is gradually increased up to $\epsilon_{\text{max}}$ using the following schedule [12]:

\[
\epsilon = \frac{\epsilon_{\text{max}}}{1 + \exp(-\gamma \cdot p)}
\]

This excludes pairing ROIs with drastic changes in emphysema progression, which may introduce some bias when training $NN_d$.
### Individual-Level Evaluation of sLTP Labeling in MESA Exam 1-5

\( N_p \) = number of scan pairs;
\( N_k \) = number of sLTPs present;
\( \chi^2 \) Distance = average \( \chi^2 \) distance between sLTP histograms in pairs of scan;
Correlation = Spearman’s correlation between %sLTP, reporting mean, min and max among sLTPs.

- UDAA has generally higher consistency:
  - Significantly better when scanner type changes.

### Individual-level reproducibility of sLTP labeling on longitudinal scan pairs in MESA acquired within a time lapse \( \leq 48 \) months

<table>
<thead>
<tr>
<th></th>
<th>( N_p )</th>
<th>( N_k )</th>
<th>( \chi^2 ) Distance</th>
<th>Correlation: Mean [Min, Max]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CNN</td>
<td>UDAA</td>
<td>CNN</td>
<td>UDAA</td>
</tr>
<tr>
<td>Ex5-Syn.</td>
<td>369</td>
<td>8</td>
<td>2.81</td>
<td>0.73 [0.46, 0.90]</td>
</tr>
<tr>
<td>Ex5-Helical</td>
<td></td>
<td></td>
<td></td>
<td>0.79 [0.52, 0.90]</td>
</tr>
<tr>
<td>Ex5-Syn.</td>
<td>51</td>
<td>6</td>
<td>4.61</td>
<td>0.50 [0.19, 0.79]</td>
</tr>
<tr>
<td>Ex4-MDCT</td>
<td></td>
<td></td>
<td></td>
<td>0.60 [0.28, 0.81]</td>
</tr>
<tr>
<td>Ex5-Syn.</td>
<td>73</td>
<td>6</td>
<td>5.10</td>
<td>0.57 [0.14, 0.81]</td>
</tr>
<tr>
<td>Ex4-EBT</td>
<td></td>
<td></td>
<td></td>
<td>0.59 [0.17, 0.82]</td>
</tr>
<tr>
<td>ExB-MDCT</td>
<td>1,812</td>
<td>7</td>
<td>1.81</td>
<td>0.80 [0.67, 0.86]</td>
</tr>
<tr>
<td>ExF-MDCT</td>
<td></td>
<td></td>
<td></td>
<td>0.82 [0.68, 0.90]</td>
</tr>
<tr>
<td>ExB-EBT</td>
<td>1,839</td>
<td>10</td>
<td>2.15</td>
<td>0.76 [0.55, 0.90]</td>
</tr>
<tr>
<td>ExF-EBT</td>
<td></td>
<td></td>
<td></td>
<td>0.77 [0.59, 0.92]</td>
</tr>
<tr>
<td>ExB-EBT</td>
<td>171</td>
<td>8</td>
<td>4.69</td>
<td>0.53 [0.15, 0.84]</td>
</tr>
<tr>
<td>ExF-MDCT</td>
<td></td>
<td></td>
<td></td>
<td>0.67 [0.42, 0.85]</td>
</tr>
</tbody>
</table>

**Bold** = significantly better performance (\( p < 0.05 \)).
Robust **emphysema segmentation** on *cardiac* CT scans:
- Scanner-specific and subject-specific parameterization.

Consistent **lung texture learning** on *cardiac* CT scans:
- Unsupervised domain adaptation with adversarial training;
- Leads to domain invariant feature learning.

Enables:
- **Large-scale multi-site longitudinal** studies over 10 years of follow-up;
- May facilitate future study for better understanding of the disease progression.
Summary - Contributions and Impact

- **Unsupervised learning of localized emphysema texture patterns:**
  - Novel lung shape spatial mapping = a useful tool to study spatial patterns on lung CT.
  - Novel discovery of 10 highly-reproducible sLTPs and 6 clinically-significant QES:
    - May facilitate disease understanding and personalized therapy.

- **Labeling emphysema texture on cardiac CT scans:**
  - Robust emphysema segmentation on cardiac CT scans;
  - Novel lung texture labeling with domain adaptation on cardiac CT scans:
    - Enable usage of widely available cardiac CT scans.
Association Between Long-term Exposure to Ambient Air Pollution and Change in Quantitatively Assessed Emphysema and Lung Function

Overview of attention for article published in JAMA: Journal of the American Medical Association, August 2019

The Harmful Effects of Poor Air Quality
A recent study published in The Journal of the American Medical Association (JAMA) found that long-term exposure to even…

August 2019 Briefing - Pulmonology
Physician’s Briefing, 03 Sep 2019
Here are what the editors at HealthDay consider to be the most important developments in Pulmonology for August 2019.

August 2019 Briefing - Radiology
Physician’s Briefing, 03 Sep 2019
Here are what the editors at HealthDay consider to be the most important developments in Radiology for August 2019.

Polusi udara kota ternyata tidak cuma berdampak pada pernapasan
EPA-Express, 15 Sep 2019
Hak atas foto: Getty Images/Kaiten antara polusi perkotaan dan penyakit pernapasan mendorong banyak aerian untuk memberikan…

EPA-Funded Research: Climate Change Will Worsen Lung Disease for Americans
Replay News, 01 Aug 2019
E.A. Gundersen Environment, World By as much as a pack of cigarettes a day. Air pollution, especially one type that is worsening…
References


