University of Southampton

Leveraging Smartphones and Speech for Lung Health Monitoring

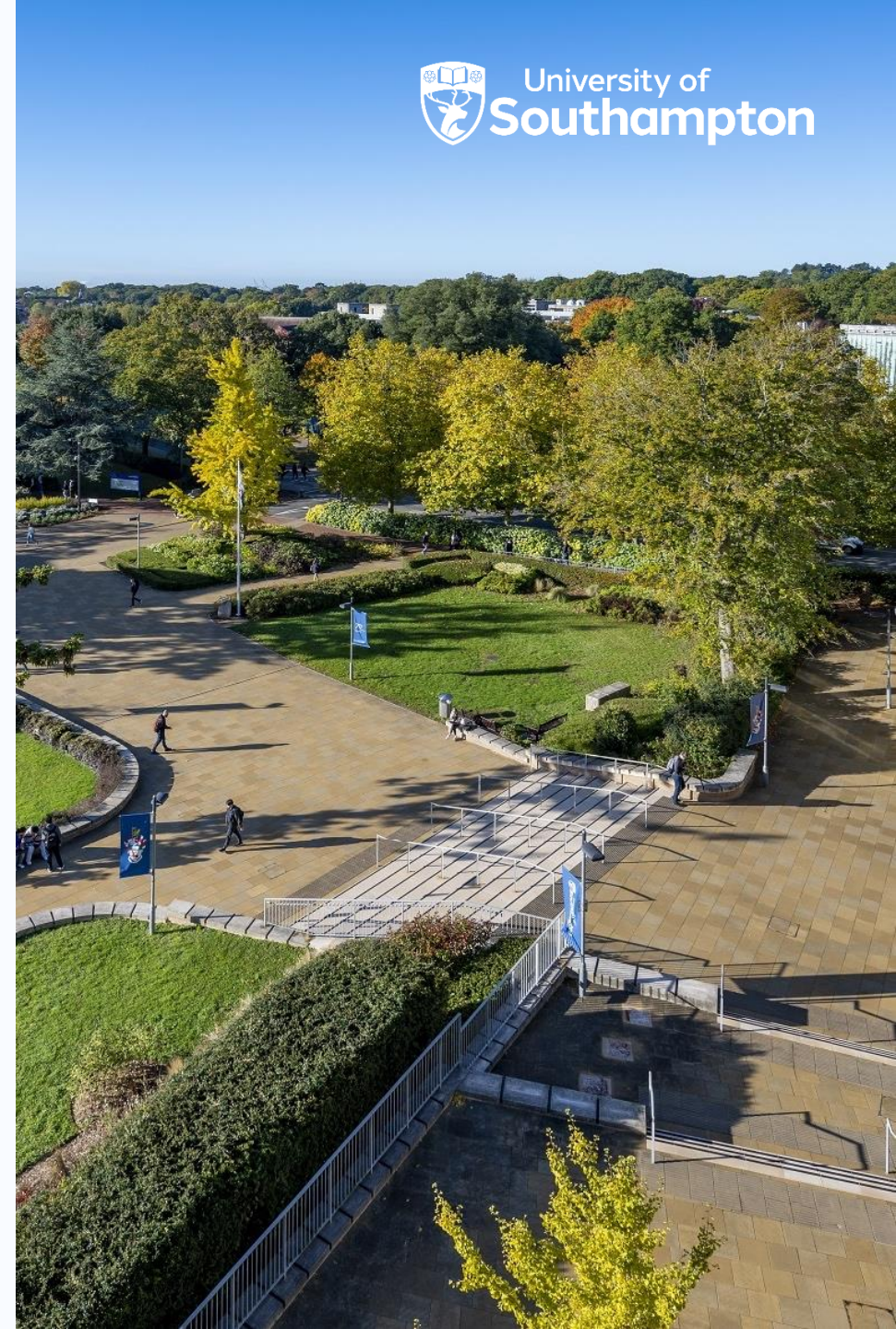Presenter: Mohammed Mosuily (PhD Student)

Supervisors: Dr Jagmohan Chauhan, Prof Paul White
15/03/2025

University of Southampton

Outline

– Motivation behind this research

– LungLDM: Prompt-Based Synthetic Lung Sound Generation Using Latent Diffusion Models

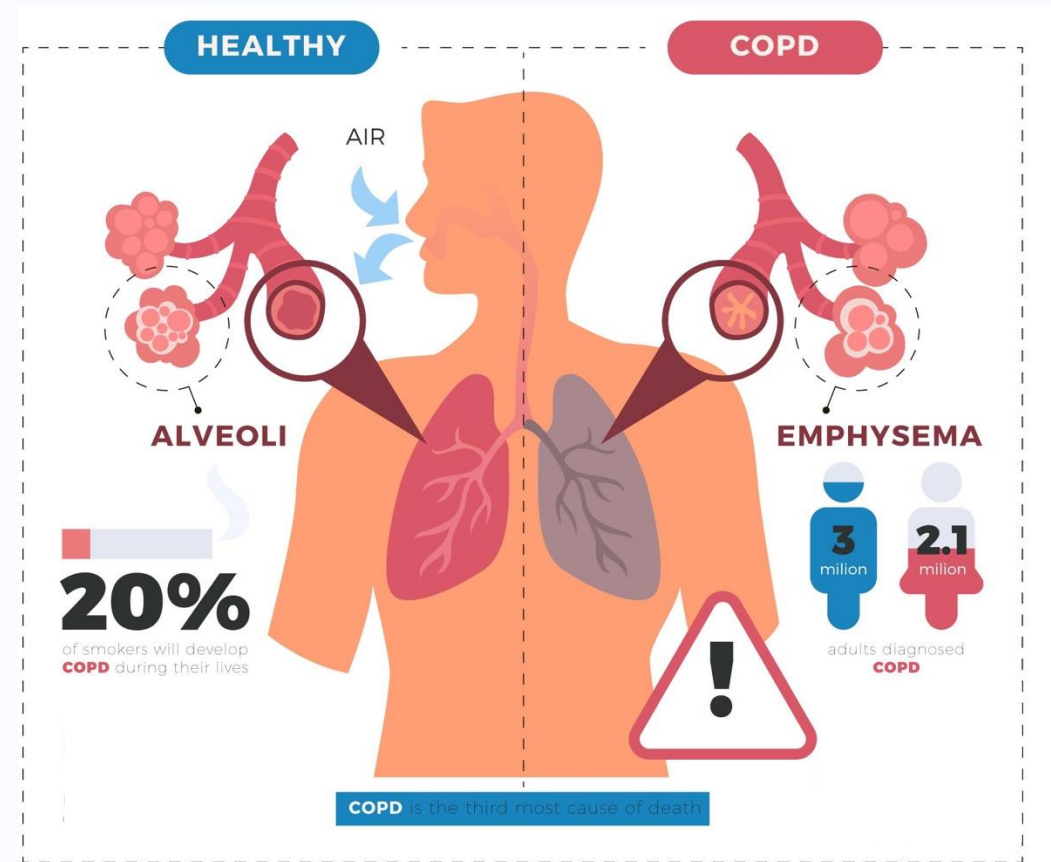– MMLung: Moving Closer to Practical Lung Health Estimation using Smartphones

– Conclusion

Motivation

Pulmonary diseases such as COPD and Asthma are the 3rd leading cause of mortality in the world [1].

[1] Soriano, J. B., et al. (2017). Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet Respiratory Medicine,* 5 (9), 691–706. https://doi.org/10.1016/S2213-2600(17)30293-X

Healthy & COPD lung illustration

LungLDM: Prompt-Based Synthetic Lung Sound Generation Using Latent Diffusion Models
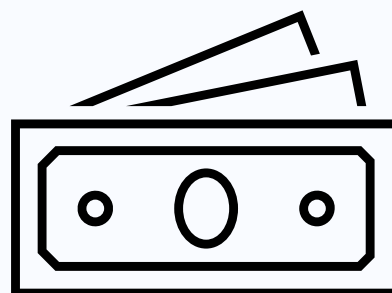
Motivation (LungLDM)

Ethics

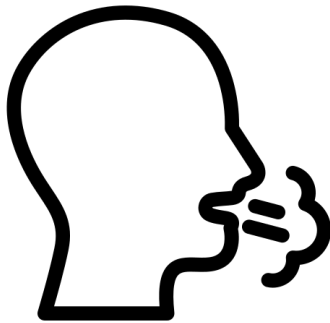Healthcare Professionals
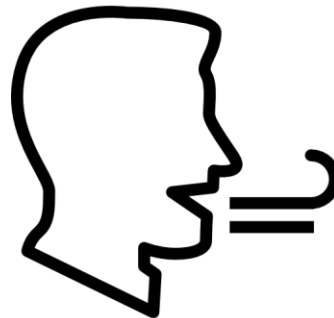
Time

Data Scarcity

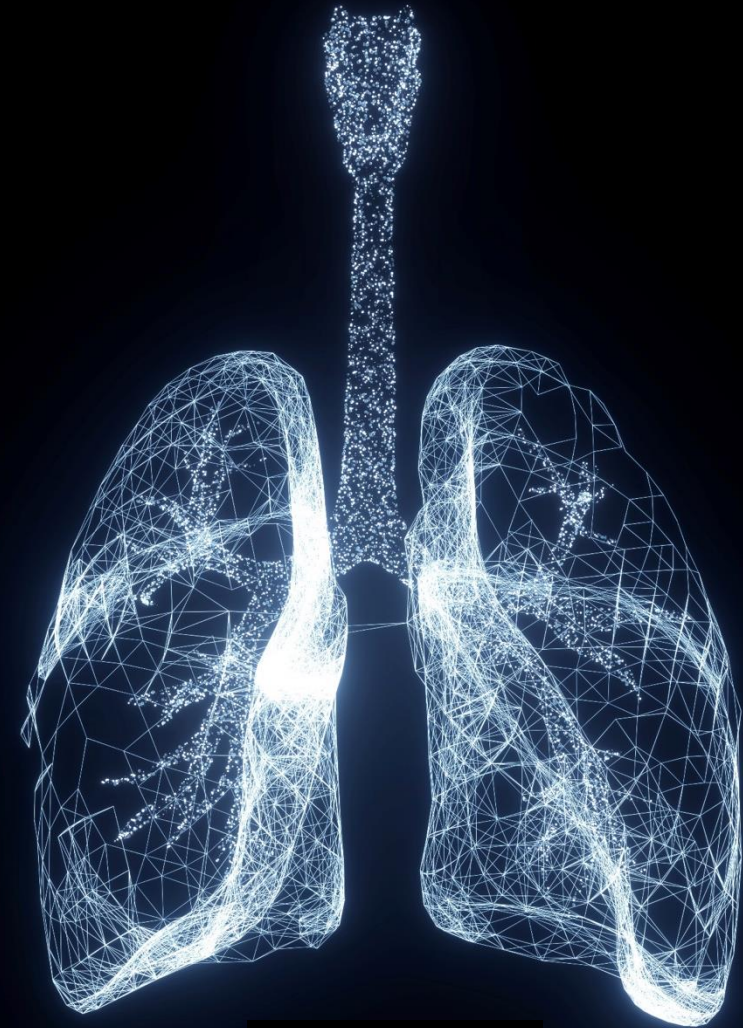Cost

Coordination with Public

LungLDM

- Can we generate realistic lung sound like coughs?

- To address this, we developed **LungLDM**, a diffusion model designed to generate **synthetic lung sounds**.

Cough sounds

Breath sounds

LungLDM (Latent Diffusion Model)

Figure: LungLDM Pipeline

LungLDM (Demo)
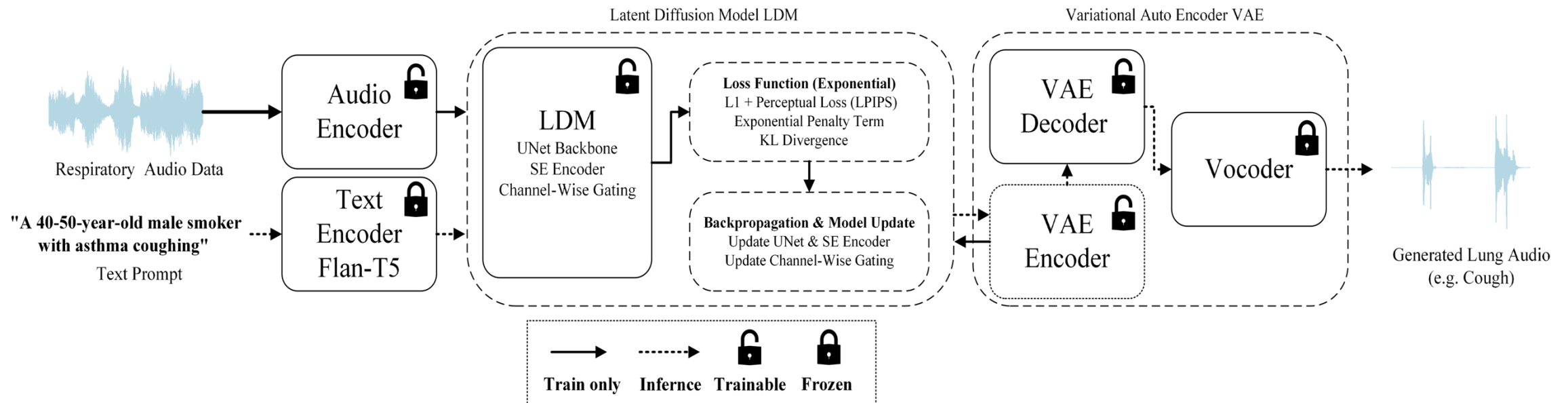

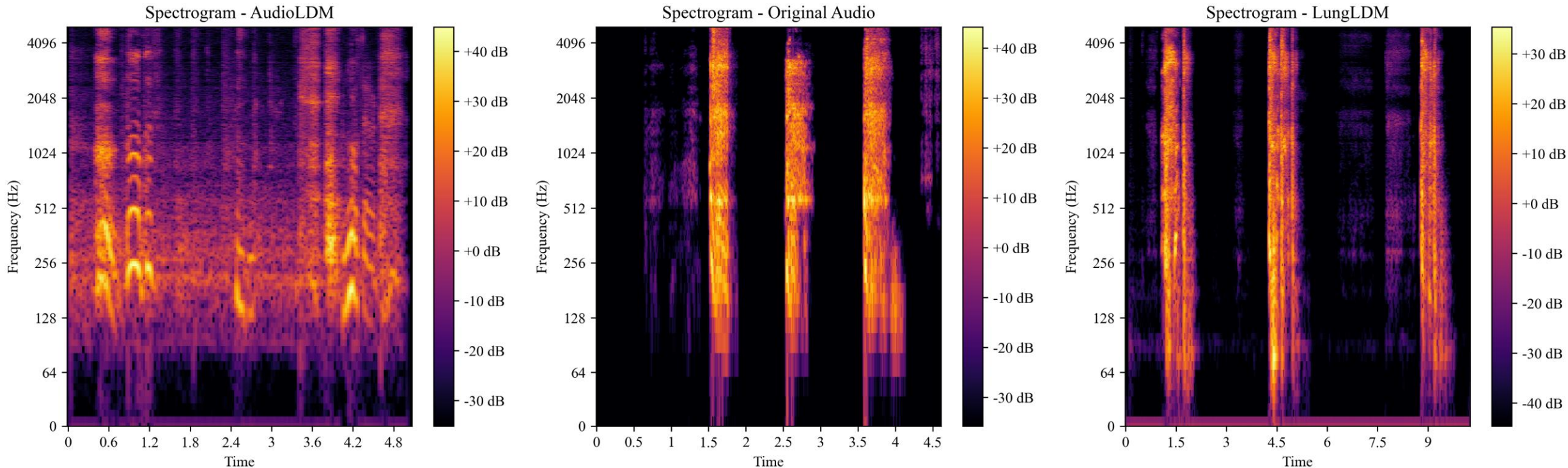Spectrogram Comparison: A person coughing by a male aged 40 50 no smoker.wav

Figure: Spectrogram comparison of a person coughing

AudioLDM (baseline)   Original sound   LungLDM (Ours)

MMLung: Moving Closer to Practical Lung Health Estimation using Smartphones

Motivation (MMlung)
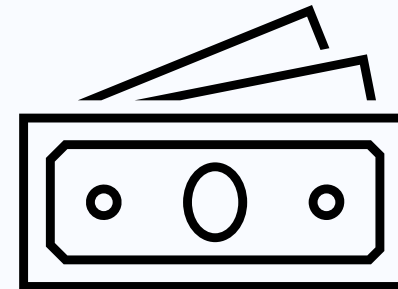

Specialized equipment

Healthcare Professionals

Time

Hospital Appointment

Cost

Motivation (Spirometry Tests)

1. The spirometry test is the **gold standard** for diagnosing pulmonary diseases [1].

2. Spirometry Parameters:

   - **FVC** = Forced vital capacity.

     - **Total air** that you blow out in one breath

   - **FEV1** = Forced expiratory volume in 1 second.

     - Amount of air you can blow out within one second.

   - **FEV1/FVC** ratio, normal value is above 0.75-0.85

Traditional Spirometry Test

[1] GOLD. (2022). *Global Strategy for Prevention, Diagnosis And Management Of COPD*. https://goldcopd.org/wp-content/uploads/2021/12/GOLD-REPORT-2022-v1.1-22Nov2021_WMV.pdf

University of
Southampton

MMLUNG: Multi-Modal Lung

Make lung function tests more accessible

Figure: Traditional Spirometry Test

Figure: Spirometry Test Using Smartphone

Mosuily, M., Welch, L., Chauhan, J. (2023) MMLung: Moving Closer to Practical Lung Health Estimation using Smartphones. Proc. Interspeech 2023, 2333-2337, doi: 10.21437/Interspeech.2023-721

Visualising Exhaling Audio Recording (Phone Speech Task)

Figure 1: Healthy Participant

Figure 2: Participant with Chronic obstructive pulmonary disease (COPD)

MMLUNG Methodology

From 40 Participants:

(Cough, Vowels, Spirometer, Speech)

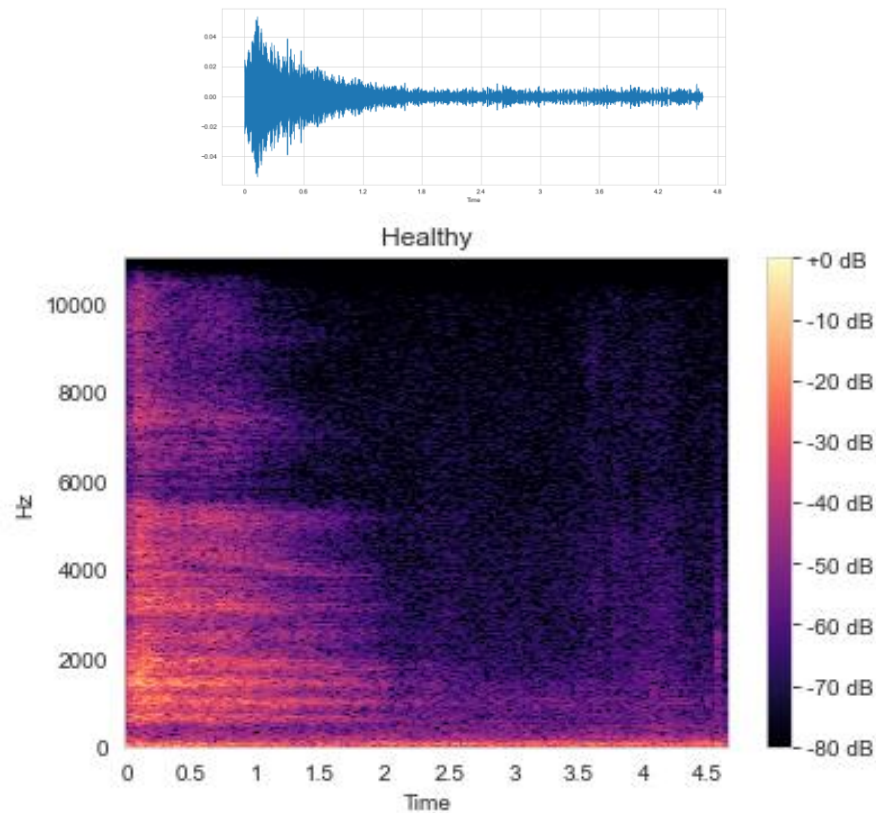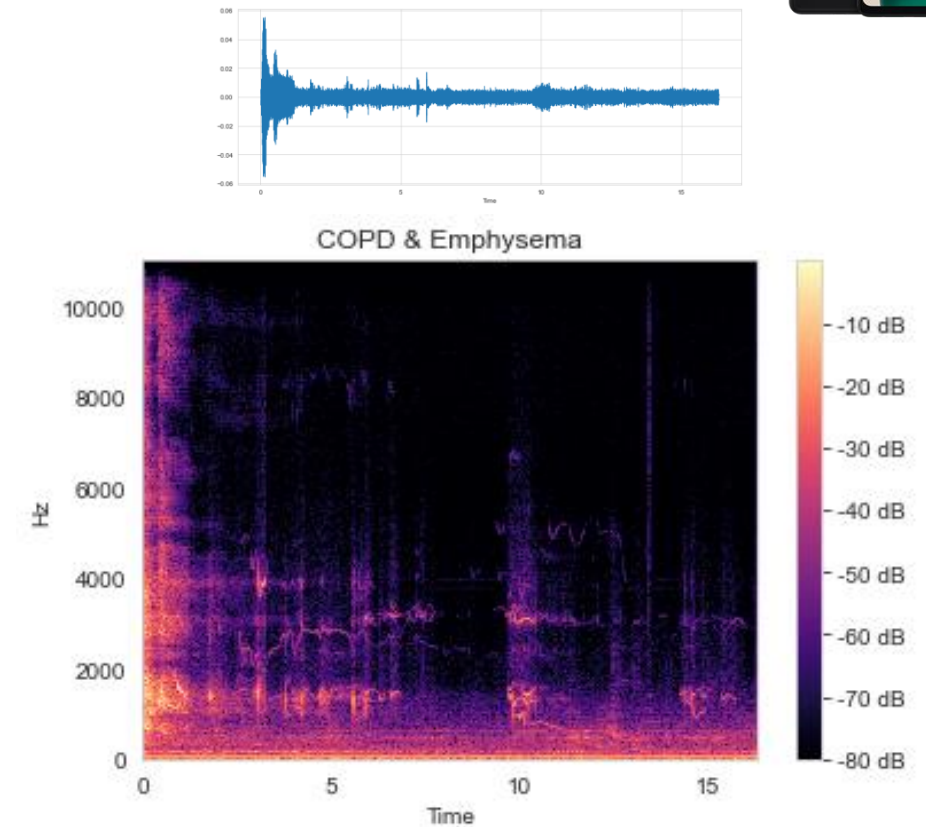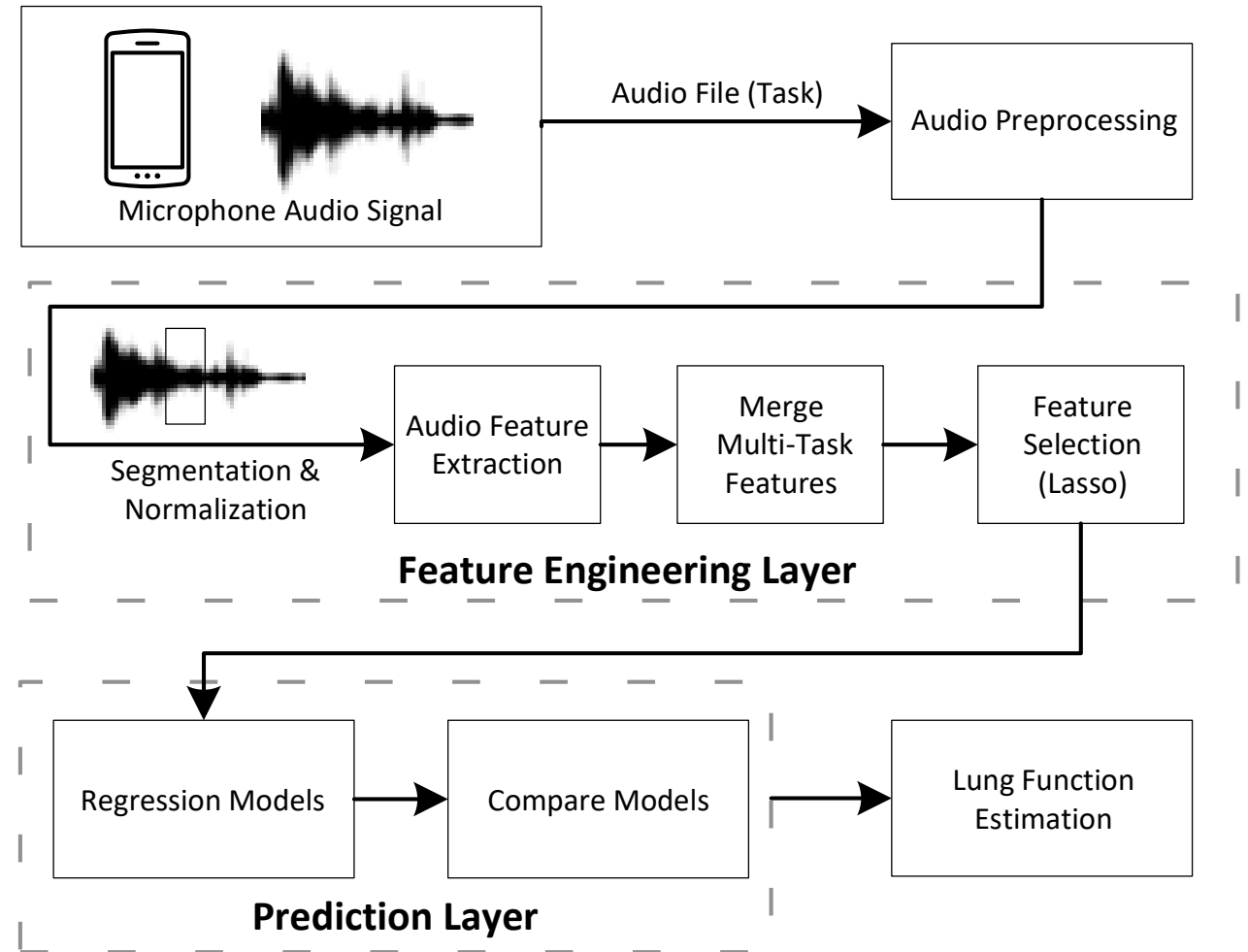Figure: Data collection setup

Figure: MMLUNG Pipeline

15

MMlung Dataset Description

Attribute	Description
Participants	40 individuals (20 male, 20 female), aged 18–85 years, English-speaking UK residents
Health Conditions	12 healthy, 7 with COPD, 7 with asthma, 14 with other long-term conditions
Devices Used	Google Pixel 6 Smartphone, Easy on-PC Ultrasonic Spirometer
Audio Sampling Rate	44,100 Hz, stereo mode, .wav format
Environment	Silent room conditions
Ground Truth	FEV_1/FVC ratio measured via medical-grade spirometer (ATS/ERS standards)
File Format	Audio: .wav files Ground Truth: Spirometry FEV_1, FVC, and (FEV_1/FVC ratio)
Total Recordings	560 recordings (40 participants × 14 tasks)

MMLung Tasks

No.	Task Name	Description
1	Forceful Coughing	Up to 10 forceful coughs
2	Vowel Pronunciations	Sustained vowels [a, e, i, o, u] in one breath
3	Smartphone Spirometry	Deep breath and blow into the microphone
4	Rainbow Passage Reading	Reading a standardized text aloud
5	Picture Description	Describing a picture for 2 minutes
6	Short Sentence	Reading: "London is the capital and largest city of England and the United Kingdom."
7	Long Sentence	Reading: "The weather today is sunny, with lots of wind and a lot of clouds in the sky which will bring heavy rain and thunderstorms in the afternoon."
8	Room Description	Describing the room where the recording takes place
9	Action Words Text	Reading a text rich in action verbs
10	Non-Action Words Text	Reading a text with minimal action verbs

MMLUNG Results

Table: MMLUNG Results

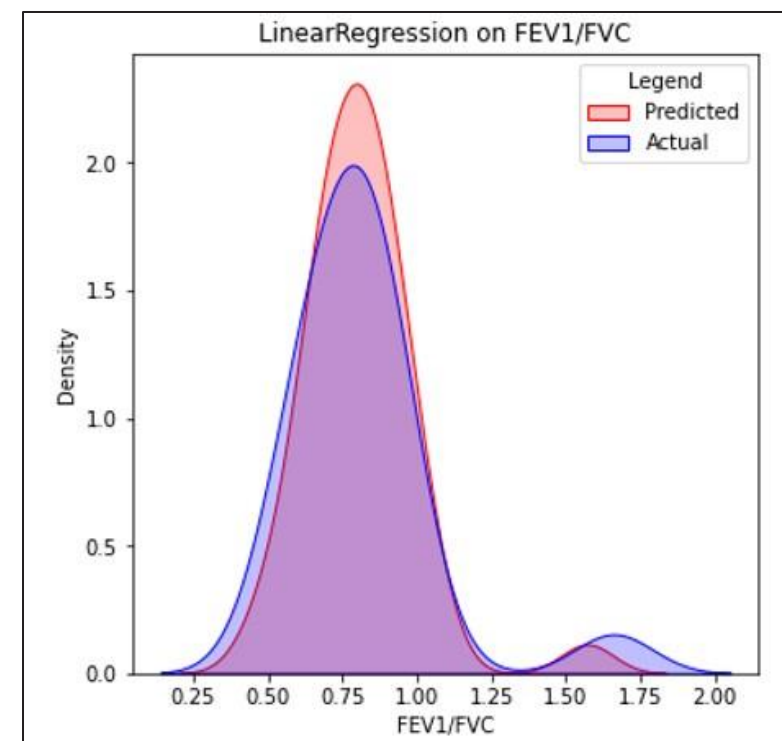| #	Task	feat.	Model	MSE	MAPE
1	**All Tasks**	**35**	**Linear**	**0.01%**	**1.13%**
2	All Vowels	39	NuSVR	0.02%	1.15%
3	Cough,Spiro., Long,[i]	39	Linear	0.02%	1.22%
4	Cough,[i],Long	39	NuSVR	0.03%	1.55%
5	All (No Spiro.)	39	NuSVR	0.03%	1.68%
6	Cough,Long,Short	39	NuSVR	0.10%	2.18%
7	Spiro.,Long,[i]	39	NuSVR	0.05%	2.48%
8	All Speech	39	Ridge	0.07%	3.03%
9	All (No Speech)	39	Ridge	0.14%	3.04%
10	Cough,Spiro.,Long	38	Ridge	0.16%	3.72%
11	Cough, Spiro, [i]	37	Ridge	0.20%	3.87%
12	Long, [i]	38	Ridge	0.34%	5.42%

Figure: Linear Regression Prediction Performance

In summary, this research aims to:

- Push the boundaries of **synthetic lung sound generation**
- Make lung health monitoring more accessible through **smartphones**
- Future works:
  - Expand MMLUNG datasets in languages like **Arabic**
  - Explore **long-term monitoring** through longitudinal data

YOUR QUESTIONS

Email: mtm1g19@soton.ac.uk