Deep Breathing Intelligence: Inhaling Data, Exhaling Digital Respiratory Health Models.

Professor Dr habil.

Björn W. Schuller

FAAAC, FACM, FBCS, FDIRDI, FELLIS, FIEEE, FISCA

TUM University Hospital

Klinikum rechts der Isar
Technische Universität München

Imperial College London

IMPERIAL

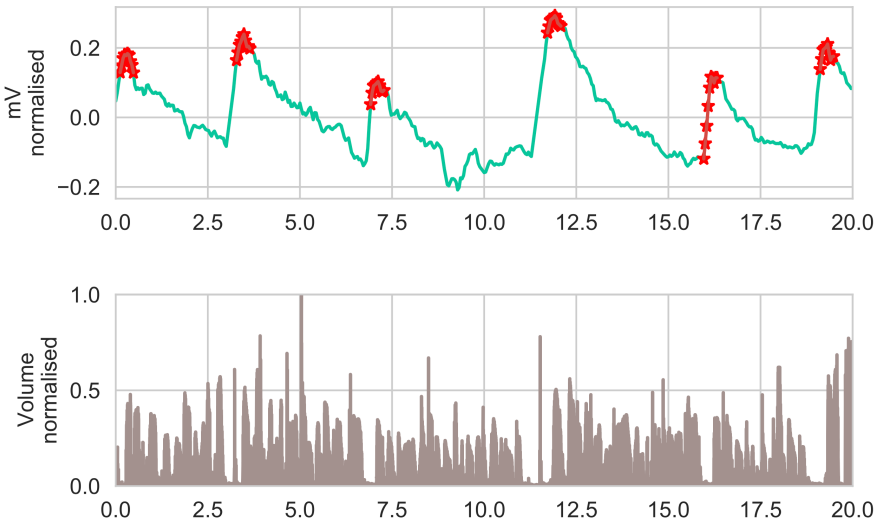
Audio based AI for Respiratory Health Monitoring, Cambridge, 2025

Cold & Snoring

UAR [%]	Addressee		Cold		Snoring	
	Dev	Test	Dev	Test	Dev	Test
M	**e2e: CNN + LSTM**					
2	59.8	60.1	59.1	60.0	37.0	37.9
3	60.9	59.1	58.6	59.6	40.3	40.3
C	**ComParE functionals + SVM**					
10^{-6}	55.8	65.8	62.9	63.9	29.3	48.4
10^{-5}	60.5	67.7	64.0	70.2	31.1	51.4
10^{-4}	61.8	67.6	61.7	66.5	40.6	**58.5**
10^{-3}	59.4	64.6	58.1	61.9	39.2	55.6
10^{-2}	57.4	60.9	58.8	59.5	39.2	55.6
10^{-1}	57.4	59.6	60.0	58.4	39.2	55.6
N	**ComParE BoAW + SVM**					
125/125	63.2	67.5	55.9	62.8	43.8	48.7
250/250	61.4	66.6	62.8	66.5	46.6	49.9
500/500	62.4	68.2	63.9	66.7	44.2	51.2
1000/1000	62.2	67.2	64.2	67.3	42.8	50.0
2000/2000	63.4	67.7	64.1	67.3	41.0	48.3
4000/4000	63.4	68.2	63.8	67.2	39.8	48.2
8000/8000	63.3	68.3	64.0	69.7	36.6	47.8
Models	**Late fusion**					
e2e+func	66.3	69.0	62.6	64.8	38.9	55.8
e2e+BoAW	67.8	68.4	62.7	62.5	45.1	46.0
func+BoAW	62.8	68.7	64.2	70.1	42.1	52.4
All (conf.)	66.4	**70.2**	66.1	70.7	43.5	53.0
All (maj.)	64.0	68.0	65.2	**71.0**	43.4	55.6

"The INTERSPEECH 2017 computational paralinguistics challenge: Addressee, cold & snoring", Interspeech, 2017.

Breathing & Masks



	E UAR [%]		B r		M UAR [%]	
	Dev (A/V)	Test (A/V)	Dev	Test	Dev	Test
C	**openSMILE: ComParE functionals+SVM**					
10^{-5}	39.1/33.3	47.9/33.3	.244	.442	56.8	59.8
10^{-4}	38.7/36.4	43.5/35.1	.234	.435	60.3	67.7
10^{-3}	34.1/40.4	42.4/41.7	.175	.333	62.3	67.8
10^{-2}	26.4/45.7	33.8/39.0	.081	.212	62.6	66.9
N	**openXBOW: ComParE BoAW+SVM**					
125	42.0/38.9	40.6/37.7	.185	.357	59.8	58.7
250	40.5/33.3	49.1/31.5	.201	.349	61.5	62.7
500	41.0/38.9	46.6/31.7	.209	.367	63.1	65.0
1000	39.0/38.7	42.2/32.4	.226	.366	63.6	66.1
2000	39.7/40.6	42.2/33.8	.215	.355	64.2	67.7
Network	**DeepSpectrum+SVM**					
ResNet50	35.0/31.6	**50.4**/40.3	–	–	63.4	70.8
X [dB]	**auDeep: S2SAE+SVM**					
-30	36.2/35.4	43.8/32.4	–	–	60.1	57.4
-45	34.9/36.7	44.3/33.8	–	–	61.3	60.3
-60	41.6/35.1	40.7/32.6	–	–	61.9	61.6
-75	40.4/32.7	42.9/33.4	–	–	61.6	62.2
Fused	36.3/29.2	43.5/32.0	–	–	64.4	66.6
Block	**LiFE: Transformer+SVM**					
GMax	39.6/54.2	37.9/41.3	–	–	–	–
BLAtt	40.6/49.2	44.0/**49.0**	–	–	–	–
BLAtt+POS	33.3/51.9	34.3/44.5	–	–	–	–
Fused	34.1/56.1	34.3/44.5	–	–	–	–
N_h RNN	**End2End: CNN+LSTM RNN**					
128	–	–	.498	.727	–	–
256	–	–	.507	**.731**	–	–
	Fusion of Best					
	–	47.9/39.8	–	.621	–	**71.8**

"The interspeech 2020 computational paralinguistics challenge: Elderly emotion, breathing & masks", Interspeech, 2020.

COVID-19

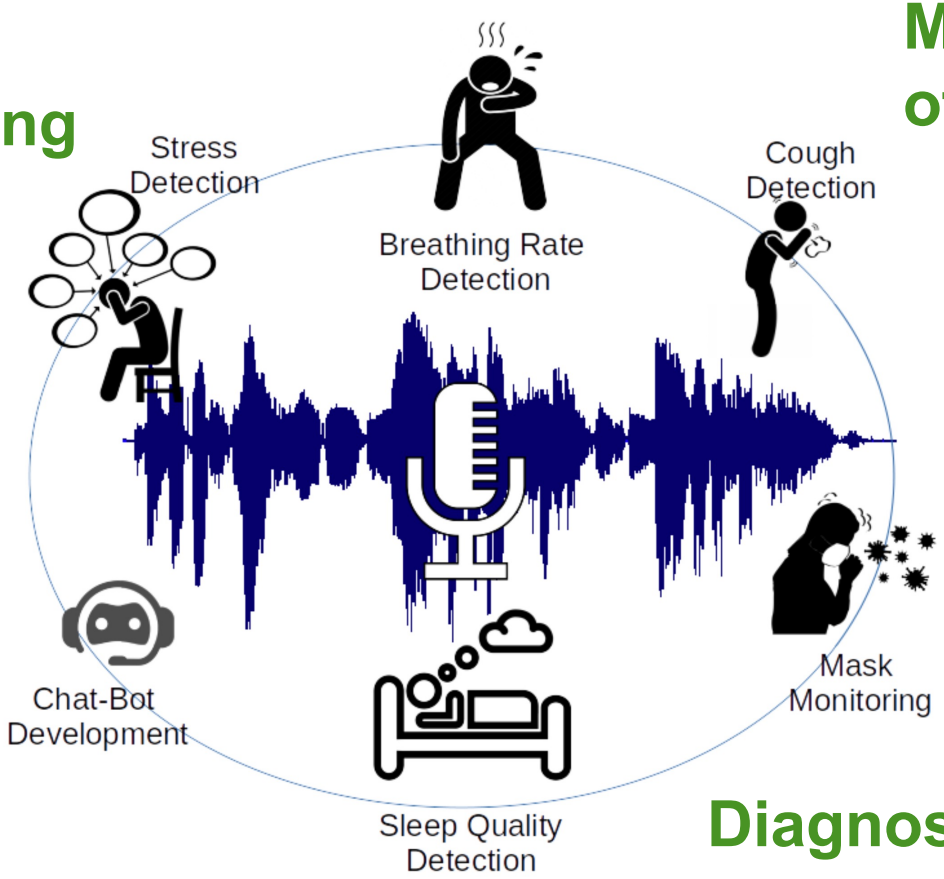
Monitoring of Social Distancing and Effects

Monitoring of Spread

Monitoring of Treatment & Recovery

Risk Assessment

Information Provision

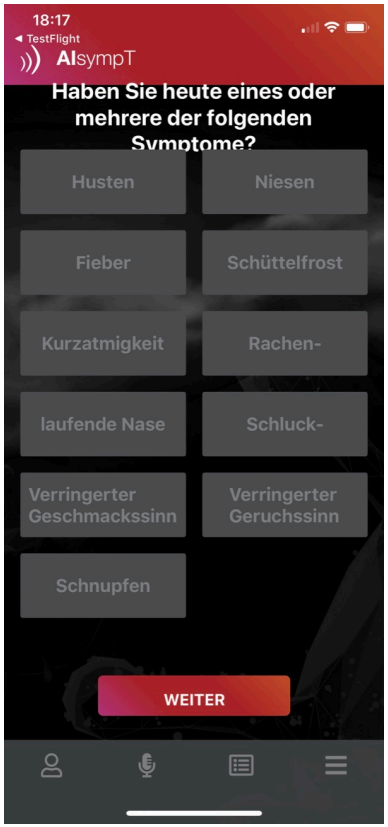


Diagnosis

"An Overview on Audio, Signal, Speech, & Language Processing for COVID-19", arXiv, May, 2020.

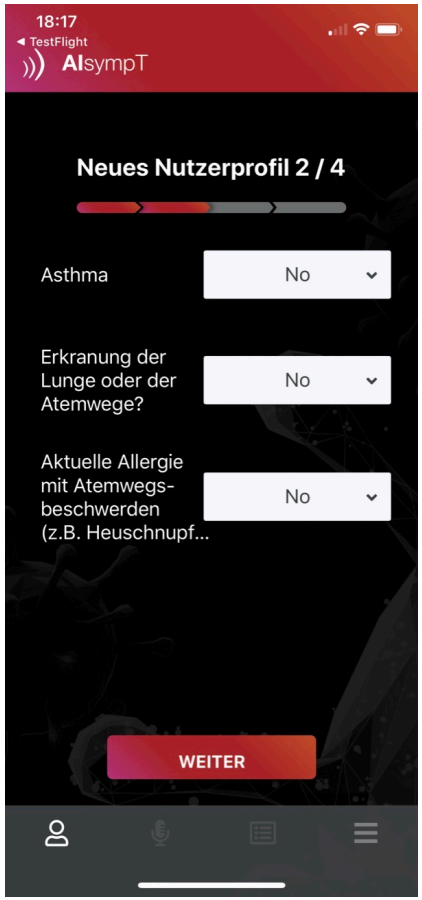
Audio-based?

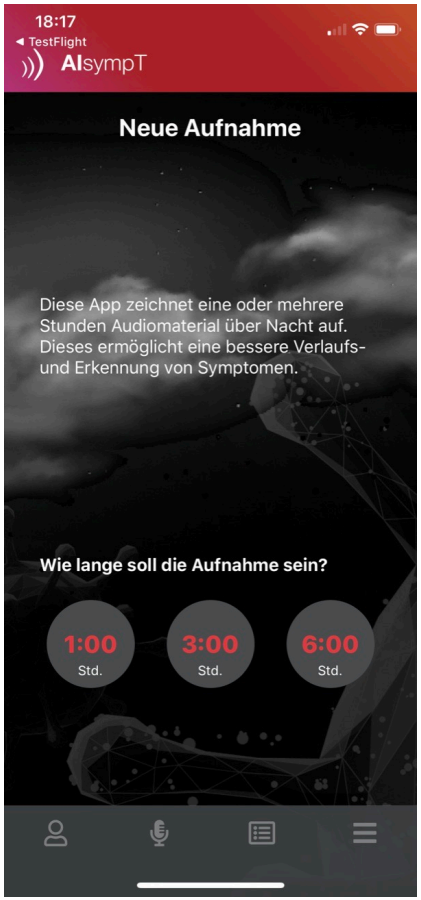
Symptom	COVID-19	Influenza	Cold
Breathing: Dypnea (Shortness)	+++	++	+
Breathing: Difficulty	+++	++	+
Rhinorrhea (Running nose)	+	++	+++
Nasal congestion	+	+	+++
Coughing	dry ++	dry ++	+
Sneezing	+	+	+++
Sore throat	+	++	+++
Pain: Body	+	+++	++
Pain: Head (Headache)	++	+++	+
Fatigue, Tiredness	mild ++	+++	+
Appetite loss	+	+++	+
Onset gradient	+	+++	+



"COVID-19 and Computer Audition: An Overview on What Speech & Sound Analysis Could Contribute in the SARS-CoV-2 Corona Crisis", Frontiers in Digital Health, 2021 / arXiv, March 2020.

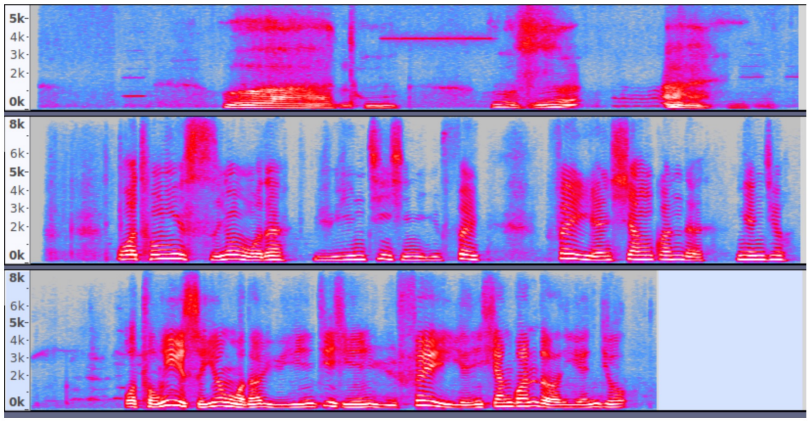
Diagnosis via Sympmtom Histograms?





Diagnosis via Speech?

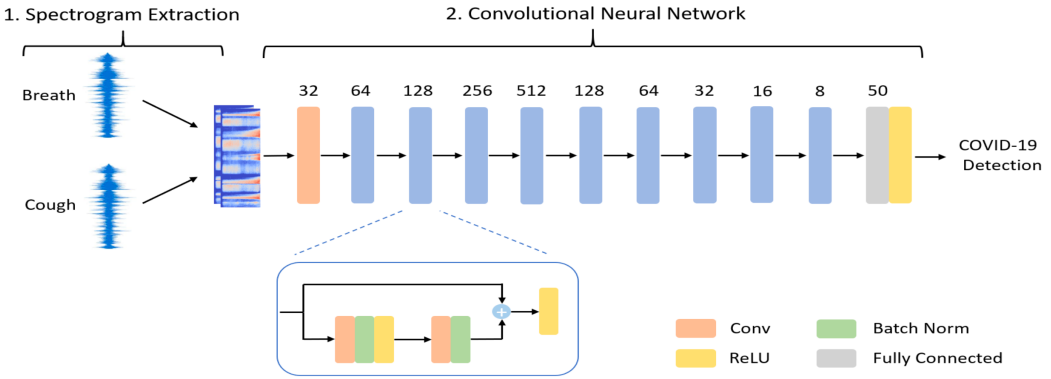
COVID-19	%AUC
2-way Diagnosis	84.6



+++

++

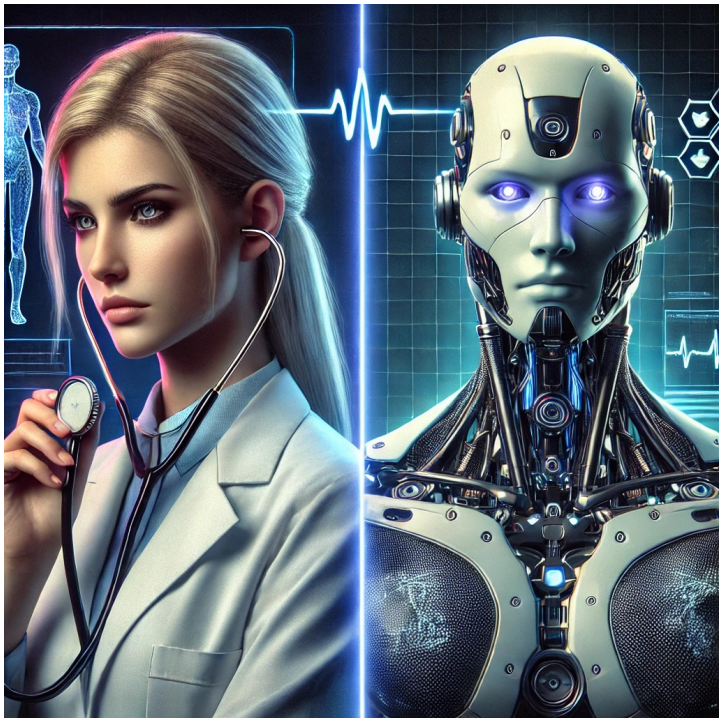
+

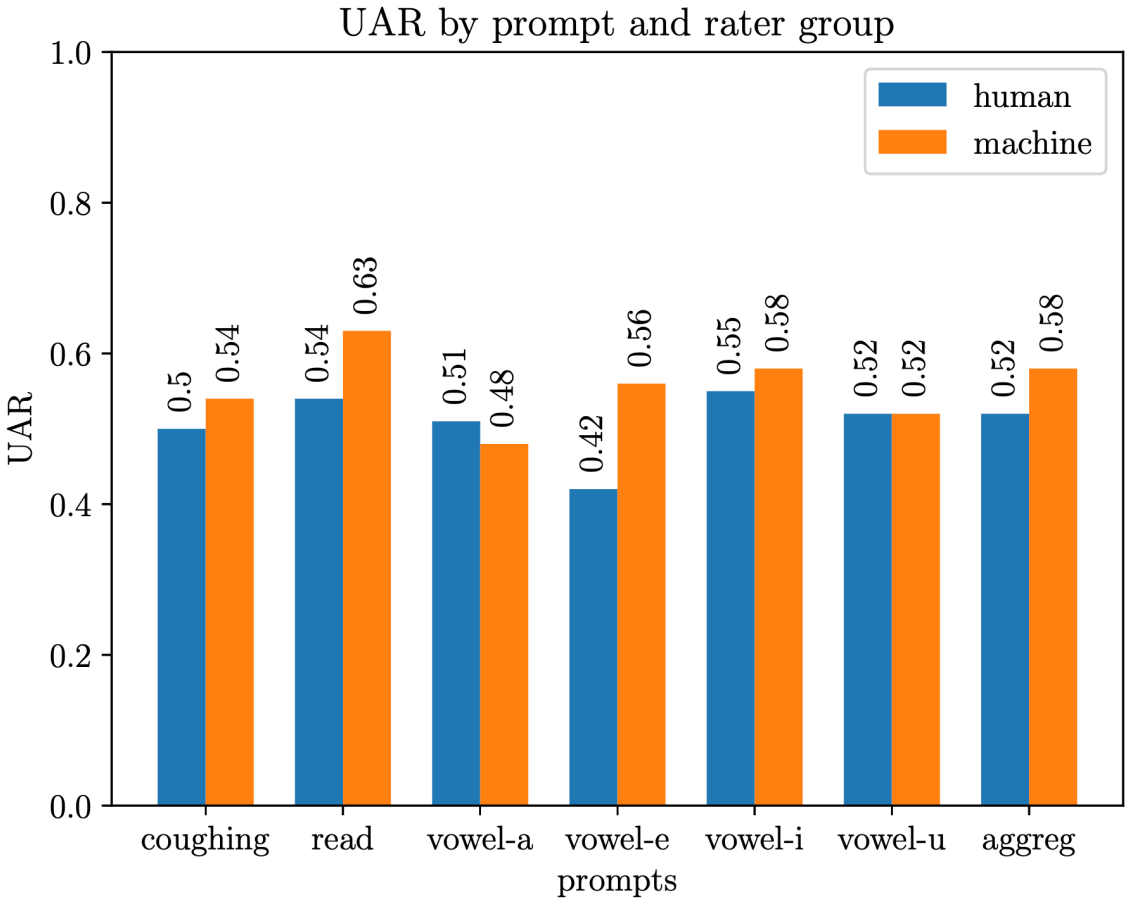


„我同意使用我的语音进行与肺炎相关的研究。“
(I agree to use my voice for coronavirus-related research purposes.)

“End-2-End COVID-19 Detection from Breath & Cough Audio“, BMJ Innovations, 2021.
“COVID-19 Detection from Audio: Seven Grains of Salt”, The Lancet Digital Health, 2021.

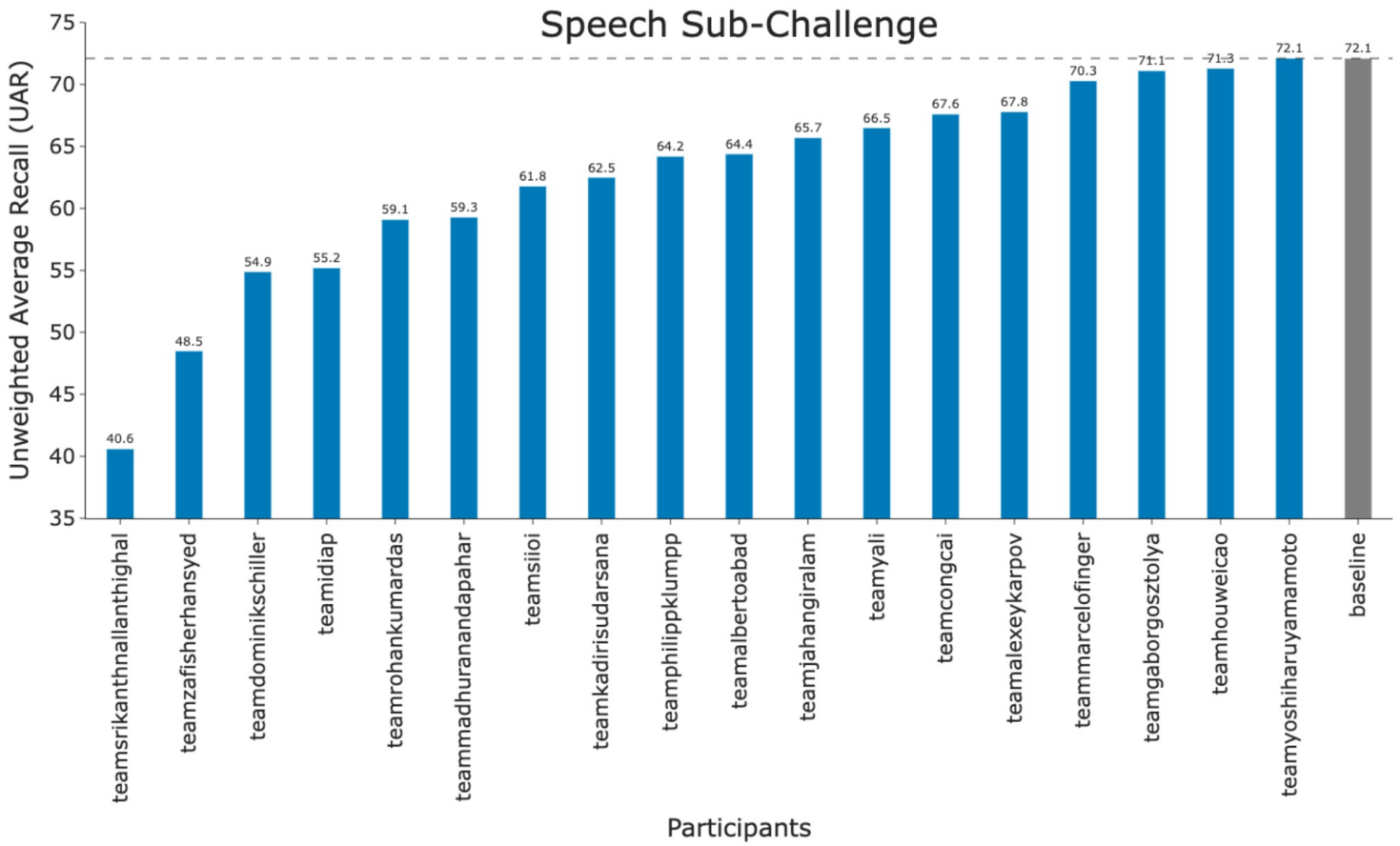
Can You Hear it?





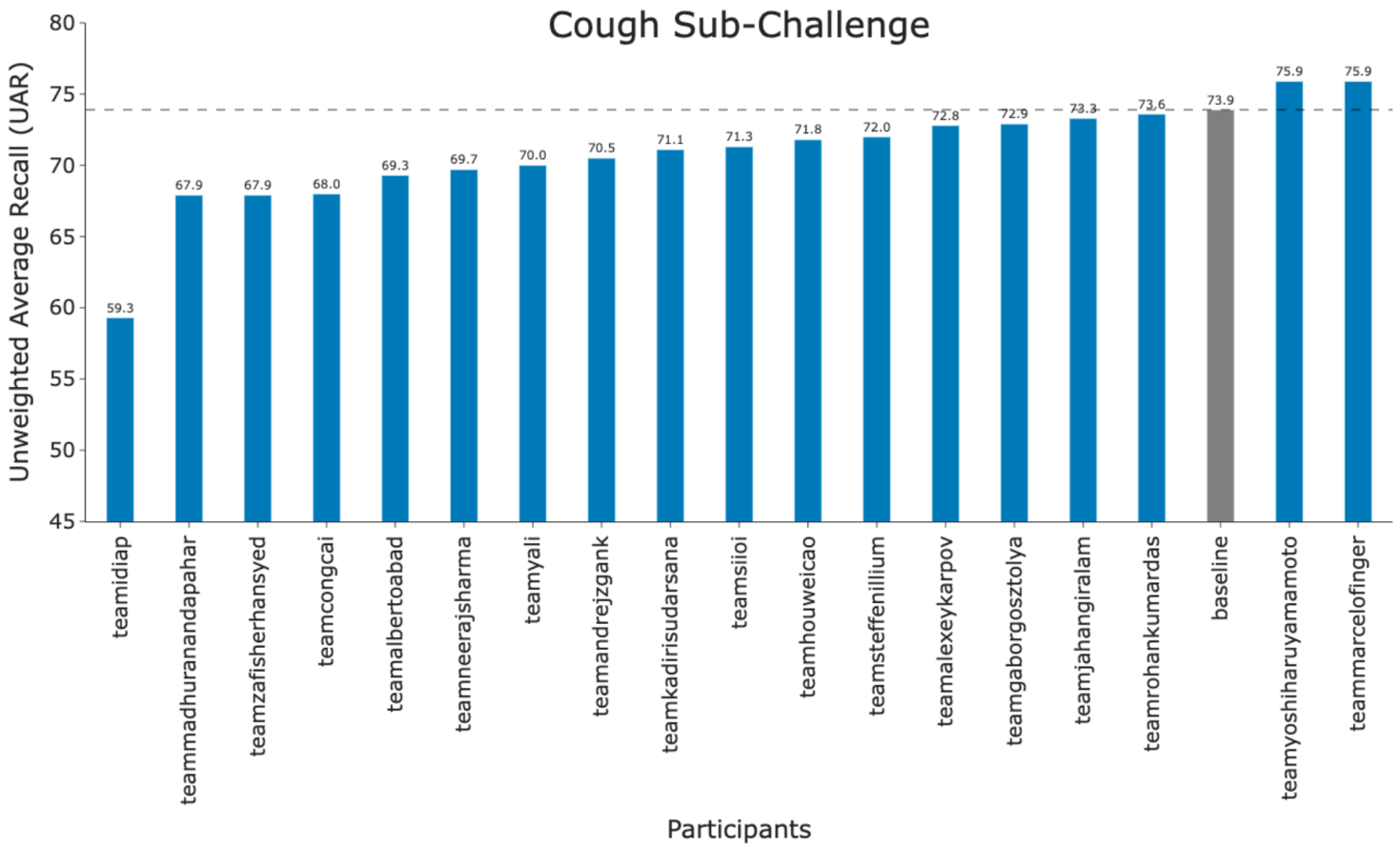
"Speaking Corona? Human and Machine Recognition of COVID-19 from Voice", Interspeech 2021.

ComParE



"A Summary of the ComParE COVID-19 Challenges", Frontiers in Digital Health 2023 / arXiv, 2022.

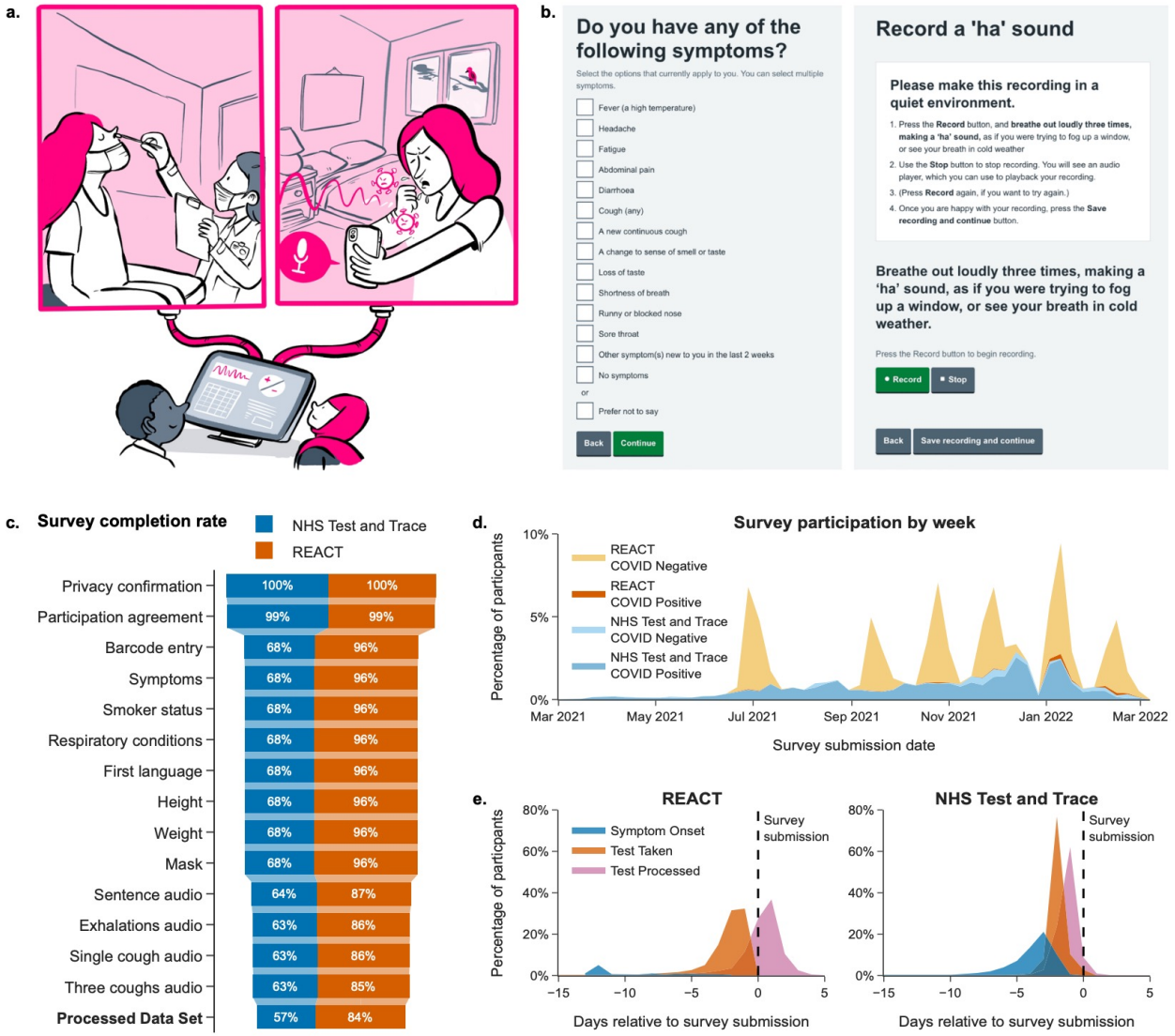
ComParE



"A Summary of the ComParE COVID-19 Challenges", Frontiers in Digital Health 2023 / arXiv, 2022.

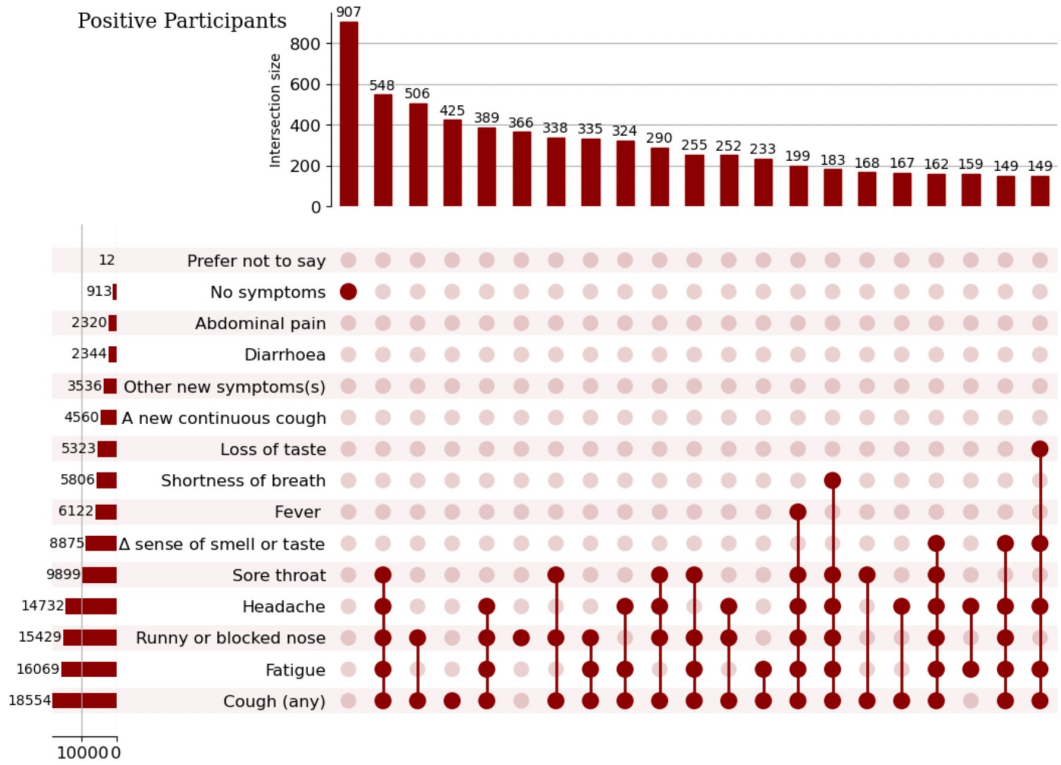
UKHSA.

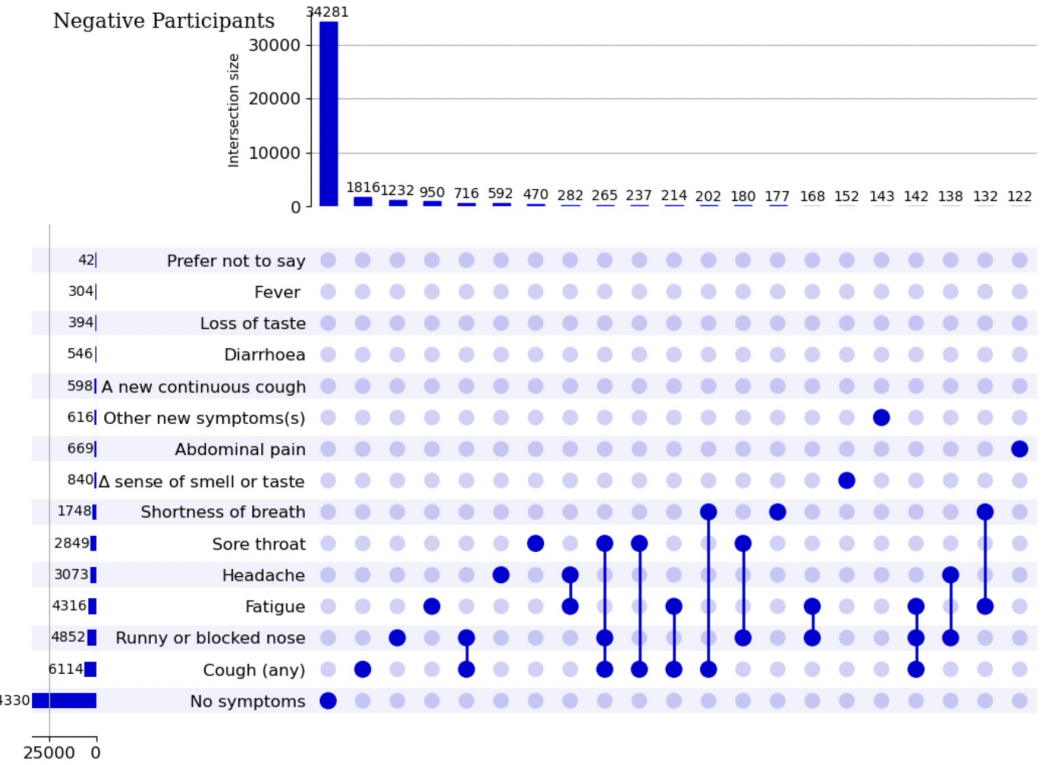
Providing COVID-19 expertise to the UK government



"A large-scale and PCR-referenced vocal audio dataset for COVID-19", Nature Scientific Data, 2024.

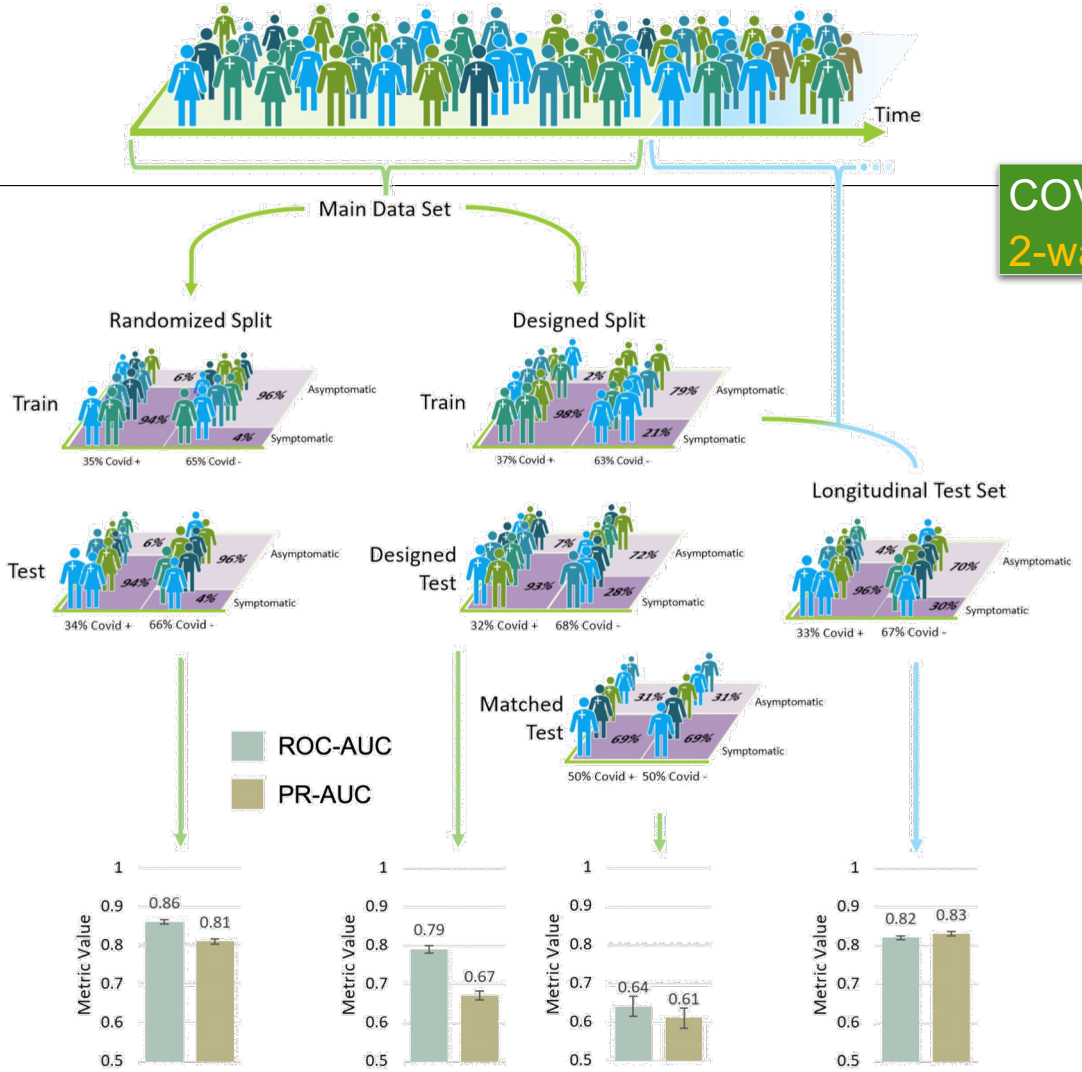
UKHSA.





"COVID-19 Detection from Respiratory Audio using Artificial Intelligence: Lessons learnt from a large scale study involving 67,895 participants", Award 3rd Imperial Computing Conference (ICC), 2022.

UKHSA.

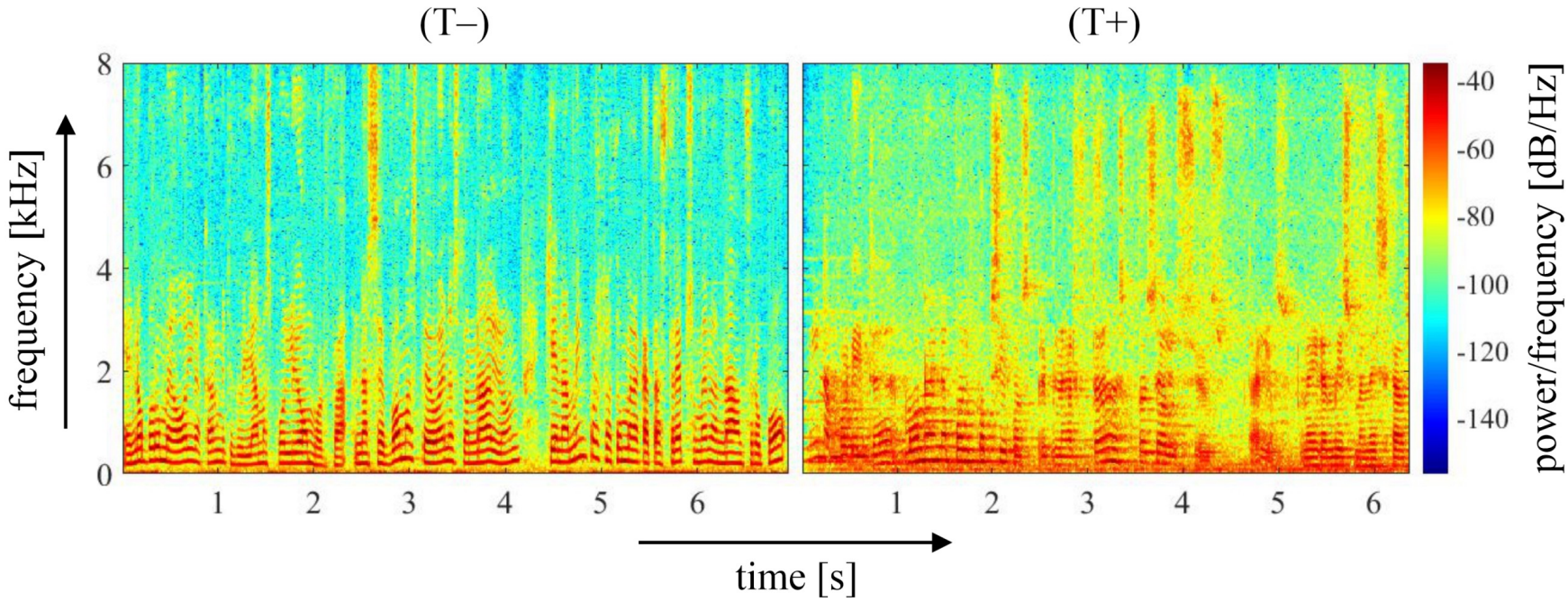


COVID-19	%AUC
2-way	82

"Audio-based AI classifiers show no evidence of improved COVID-19 screening over simple symptoms checkers", Nature Machine Intelligence, 2023.

"Bias and privacy in AI's cough-based COVID-19 recognition", The Lancet Digital Health, 2021.

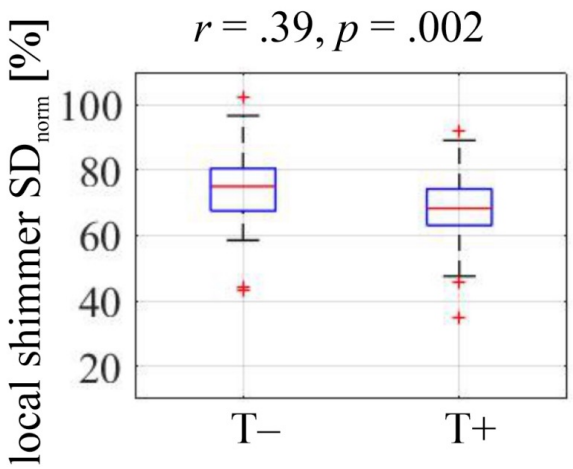
YouTube for COVID?

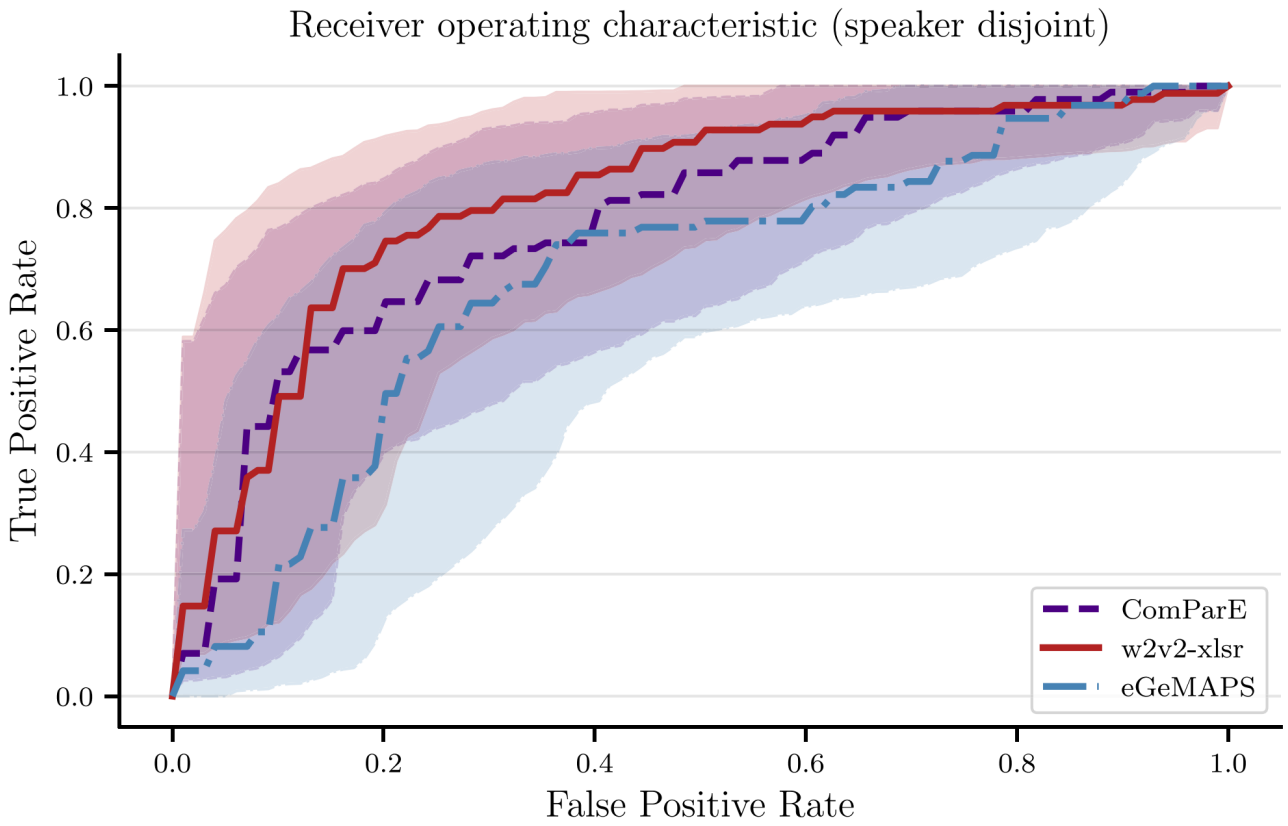


Spectrograms of utterances produced by one and the same speaker (gender: female, nationality: Greek, age at T+: 43) at a point in time without COVID-19 infection (T–) and at a point in time with COVID-19 infection (T+).

"COVYT: Introducing the Coronavirus YouTube and TikTok speech dataset featuring the same speakers with and without infection", Biomedical Signal Processing and Control 2024 / arXiv, 2022.

YouTube for COVID?

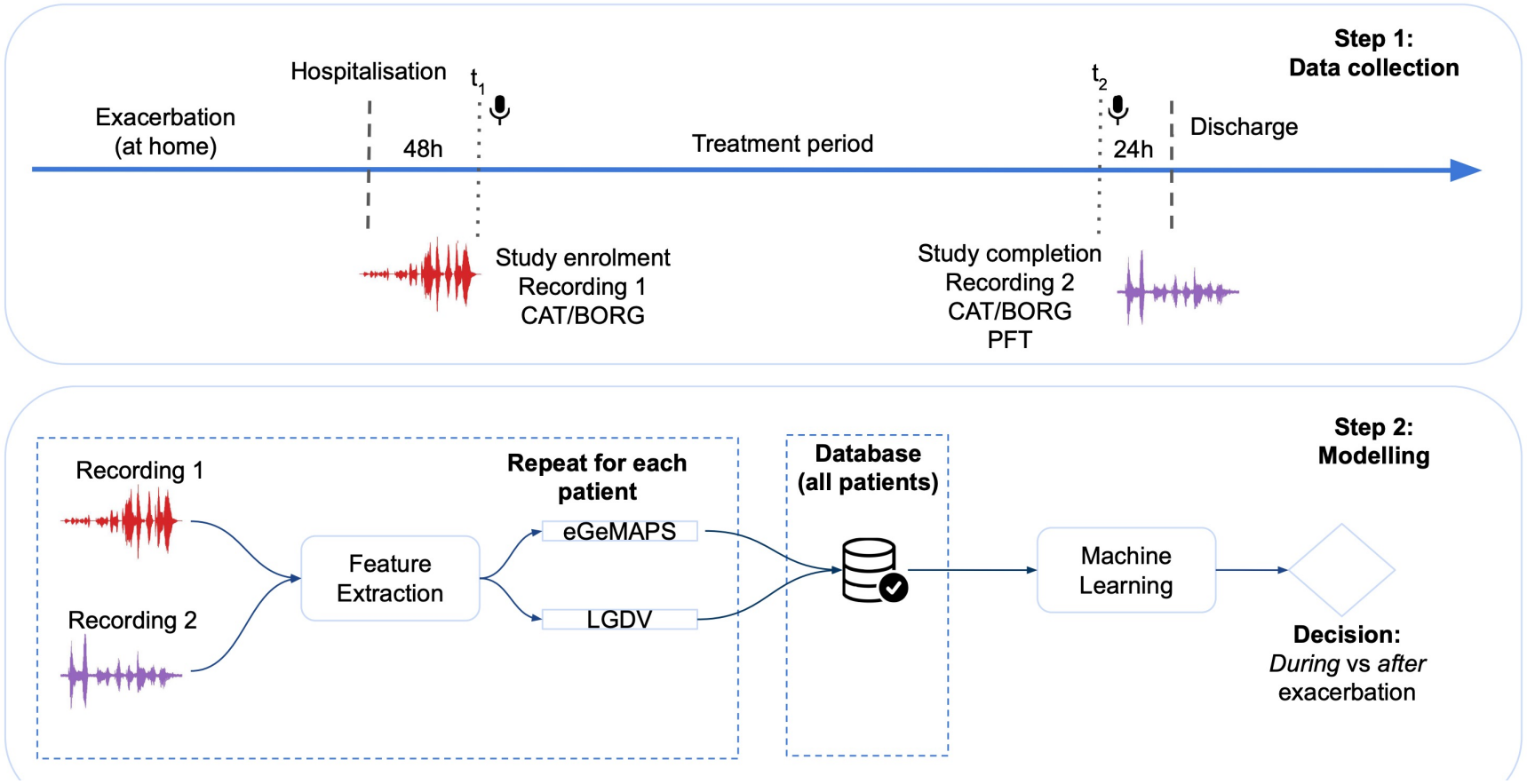




"COVYT: Introducing the Coronavirus YouTube and TikTok speech dataset featuring the same speakers with and without infection", Biomedical Signal Processing and Control 2024 / arXiv, 2022.

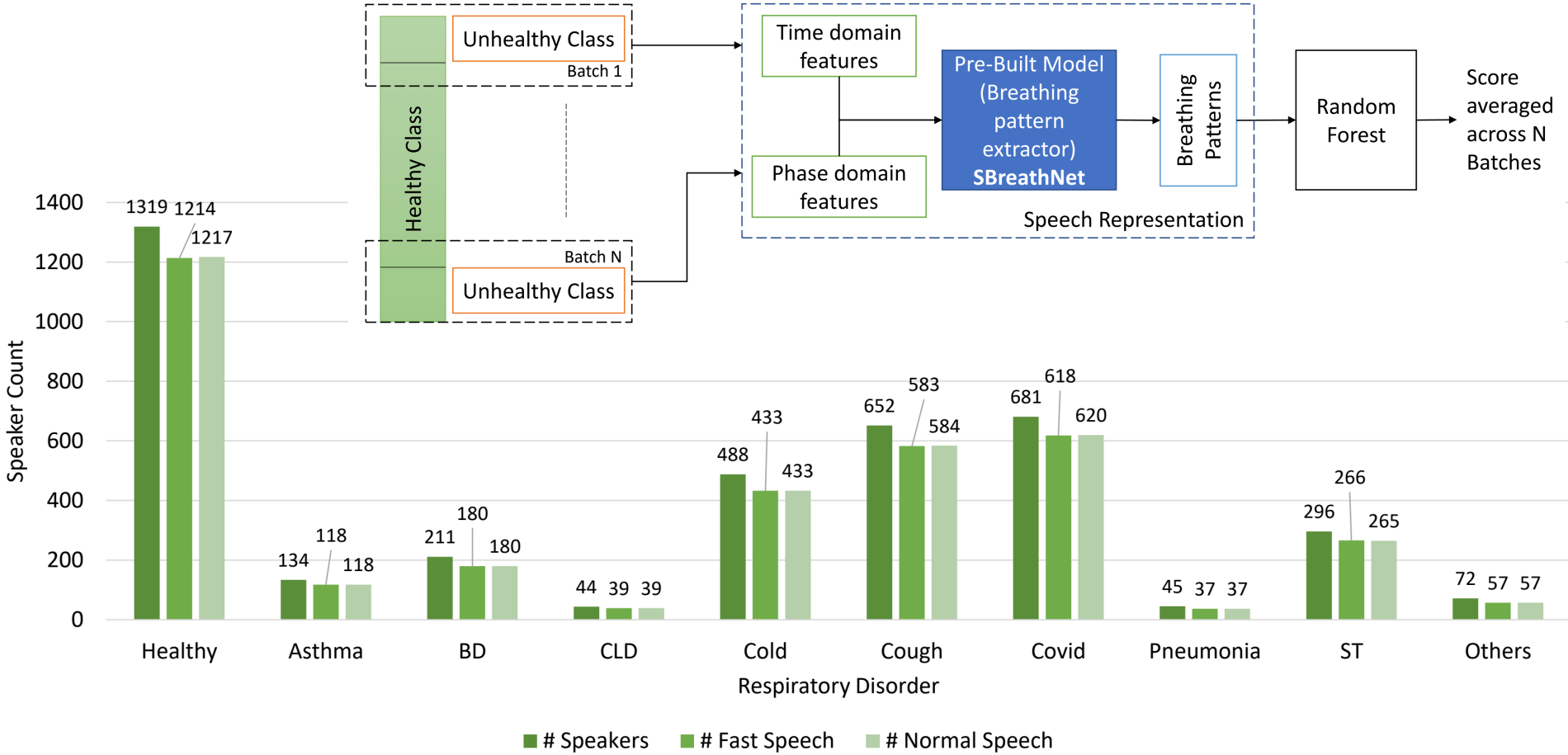

More?

COPD?



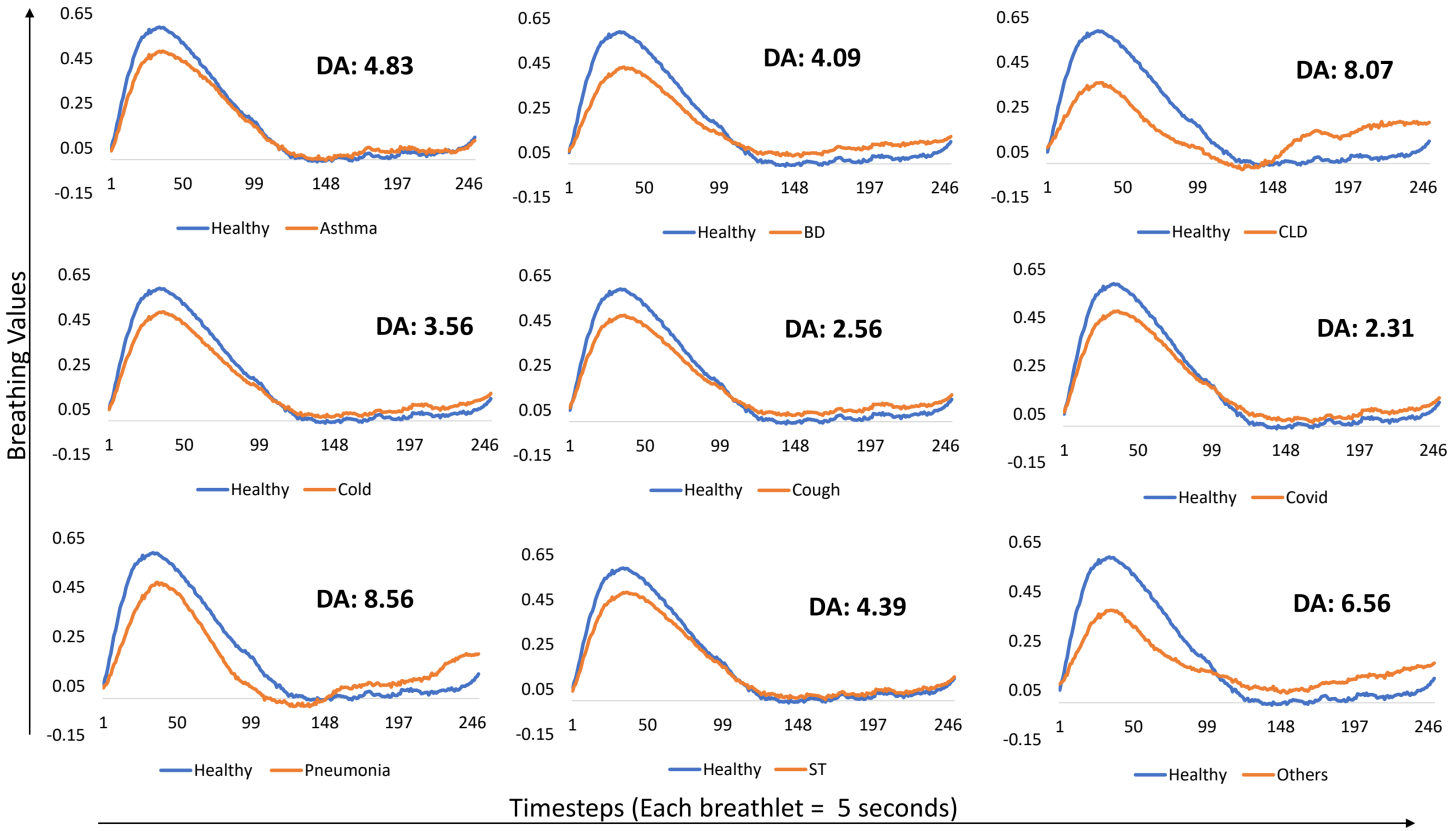
"Assessing the Clinical and Functional Status of COPD Patients Using Speech Analysis During and After Exacerbation", International Journal of Chronic Obstructive Pulmonary Disease, 2025.

Across the Board?



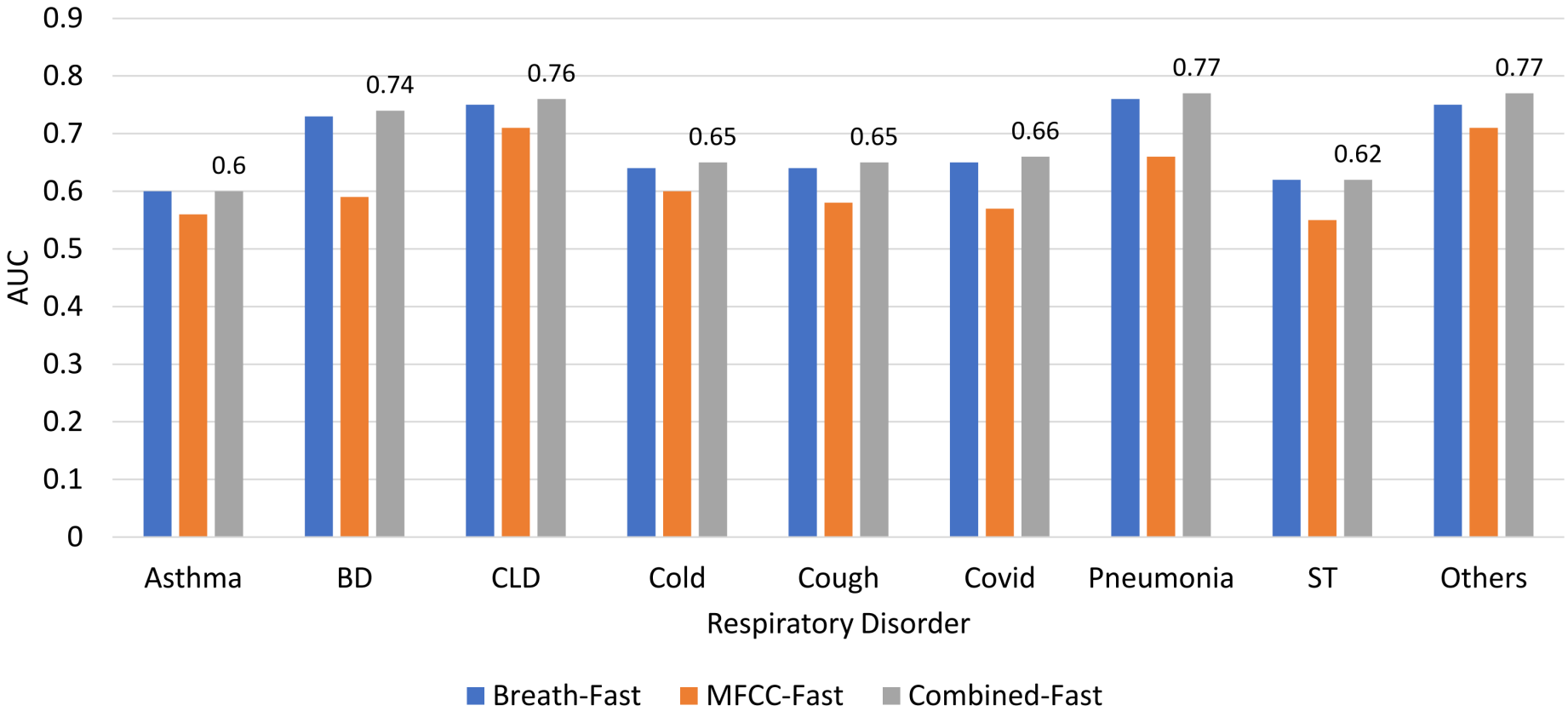
"Analysis of Respiratory Health Indicators in Speech-Breathing-Patterns", EUSIPCO, 2024.

Across the Board?



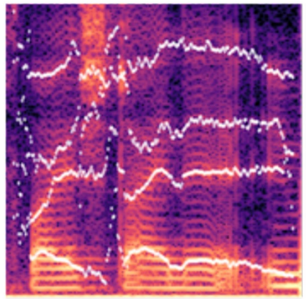
"Analysis of Respiratory Health Indicators in Speech-Breathing-Patterns", EUSIPCO, 2024.

Across the Board?

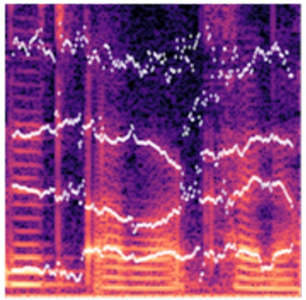


"Analysis of Respiratory Health Indicators in Speech-Breathing-Patterns", EUSIPCO, 2024.

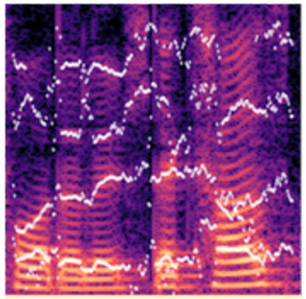
Facial Mask?



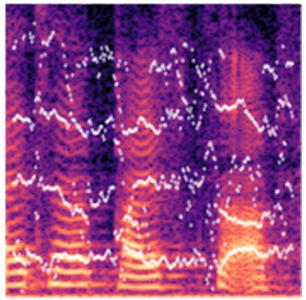
(a) *No Mask.*



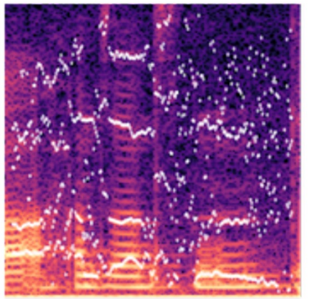
(b) *Surgical Partial.*



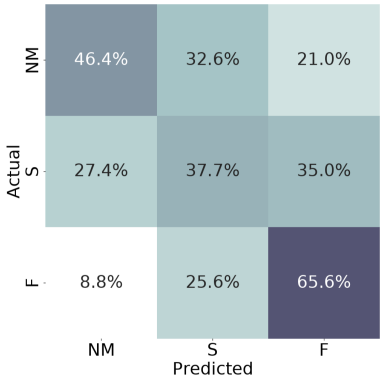
(c) *Surgical Total.*



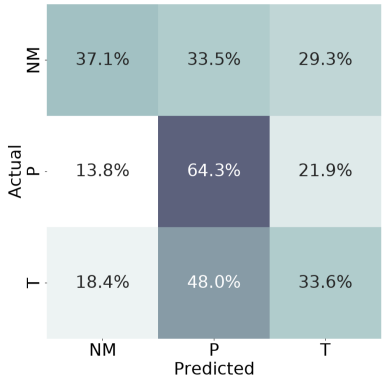
(d) *FFP2 Partial.*



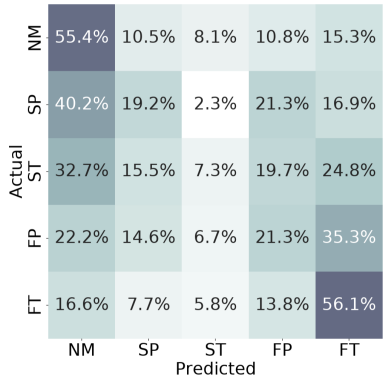
(e) *FFP2 Total.*



(a) *NM: No Mask, S: Surgical mask, F: FFP2 mask.*

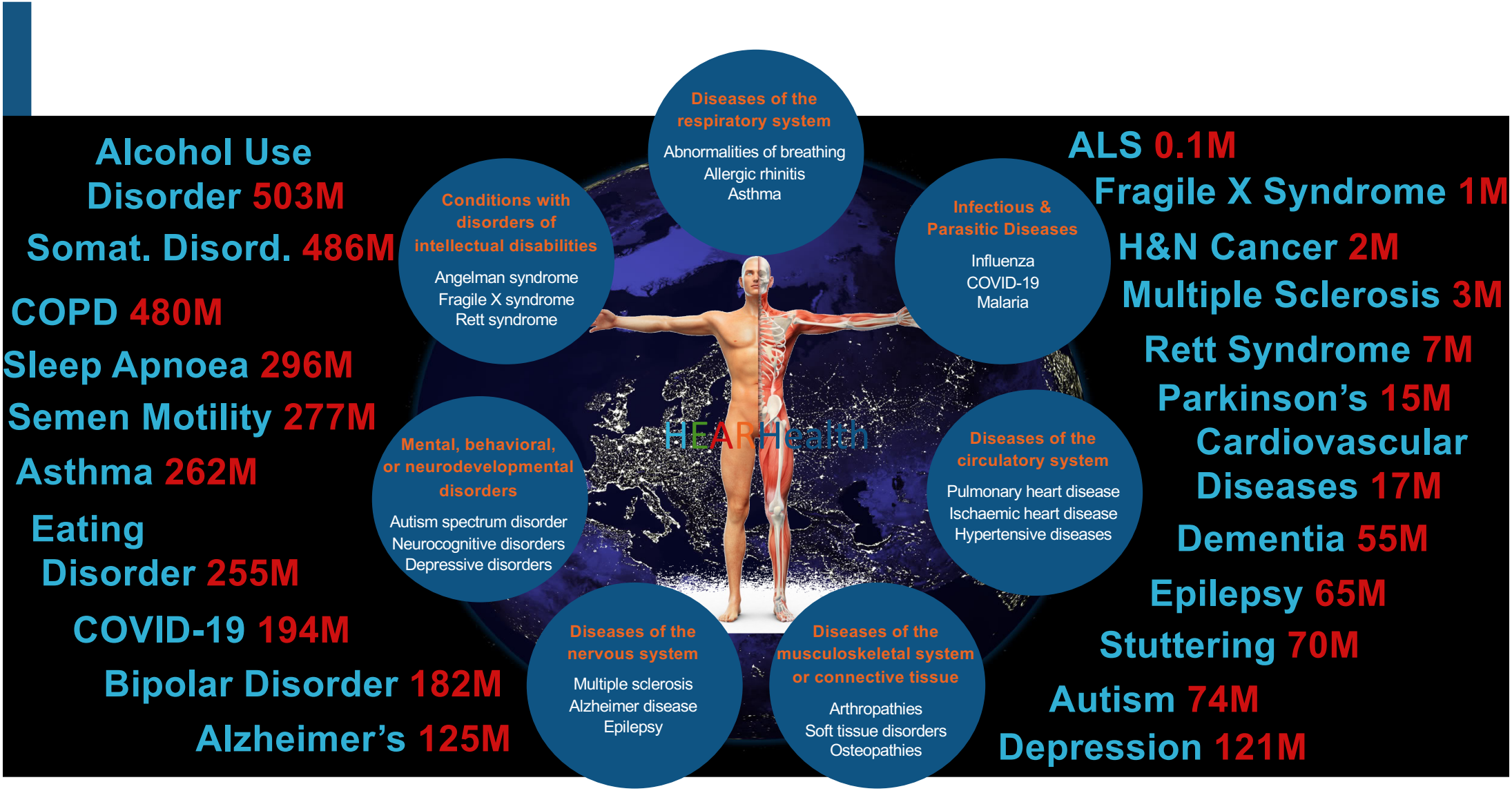


(b) *NM, P: Partial coverage (only mouth), T: Total coverage (mouth and nose).*

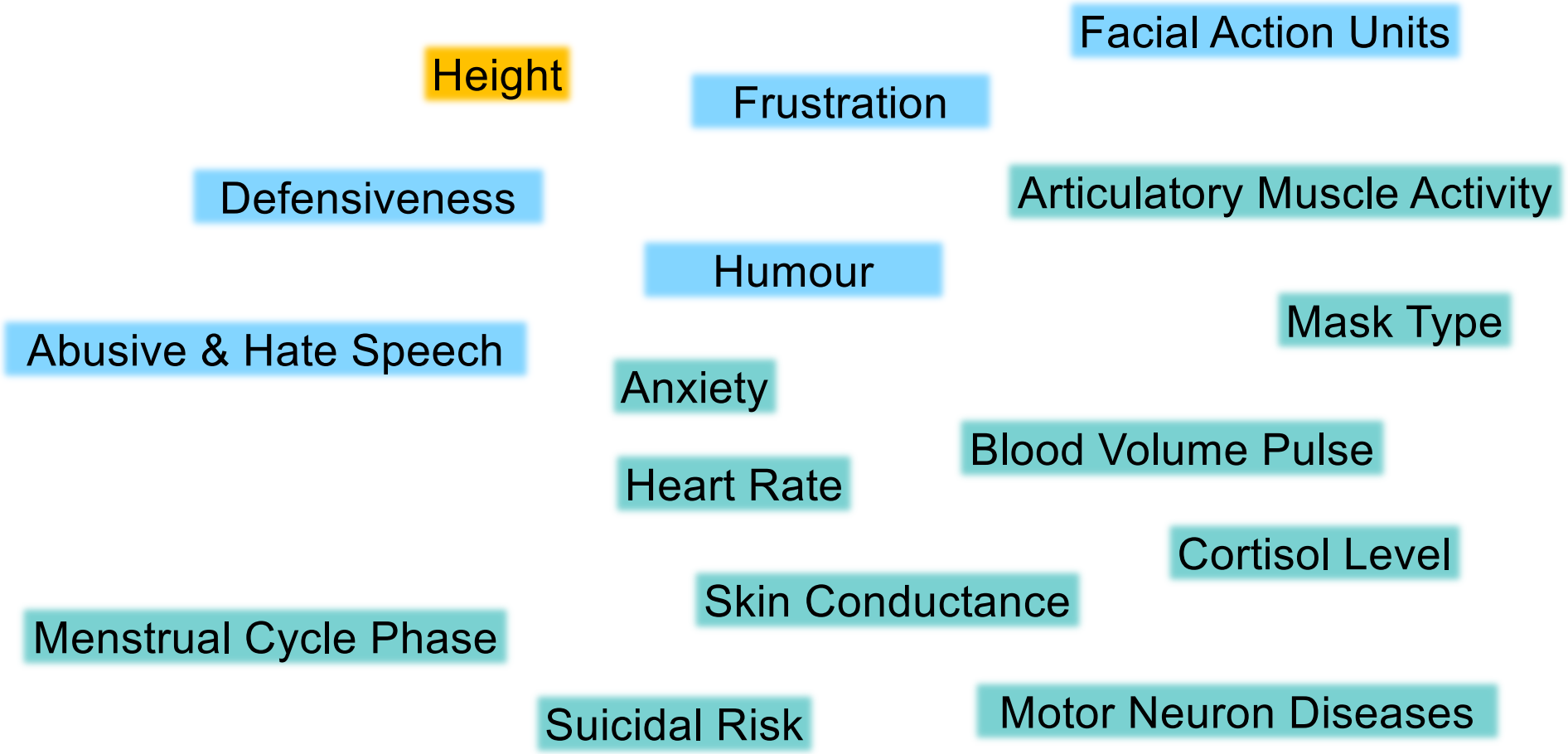


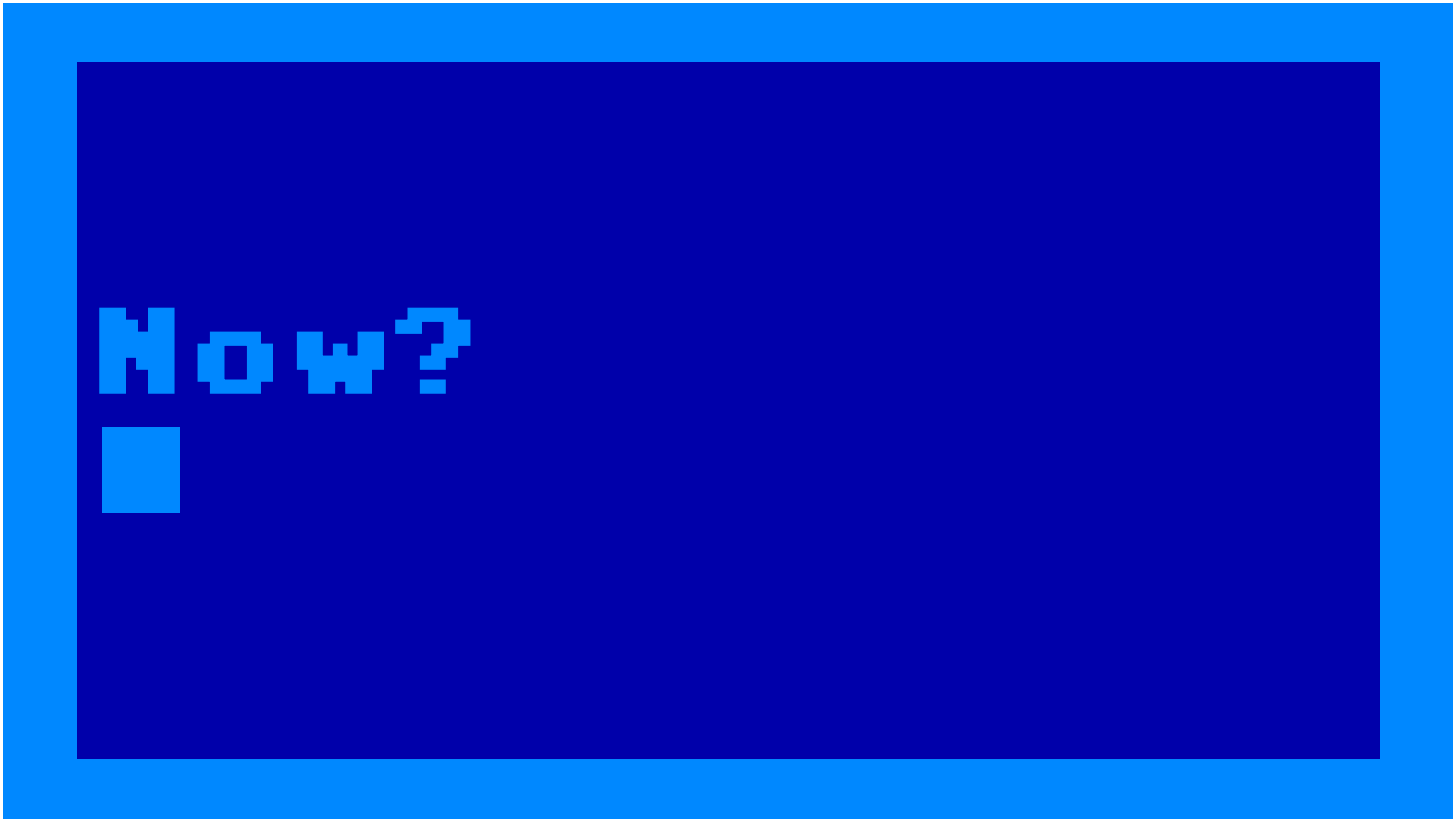
(c) *NM, SP: Surgical Partial, ST: Surgical Total, FP: FFP2 Partial, FT: FFP2 Total.*

"Face Mask Type and Coverage Area Recognition from Speech with Prototypical Networks", IberSPEECH, 2024.

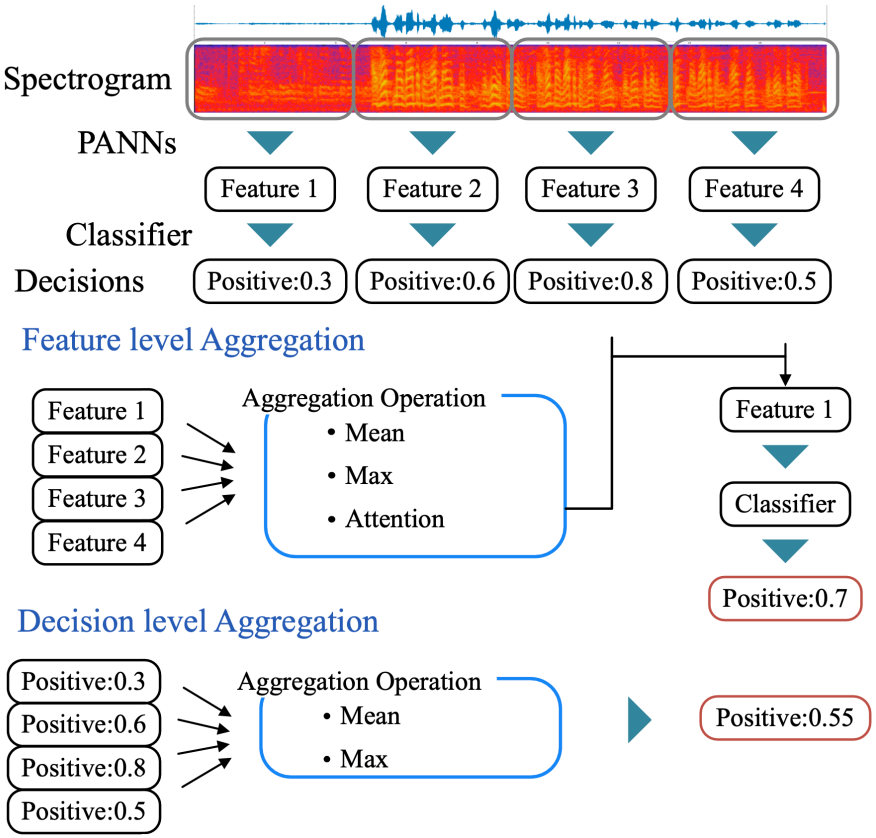

Alcohol Use Disorder 503M
Somat. Disord. 486M
COPD 480M
Sleep Apnoea 296M
Semen Motility 277M
Asthma 262M
Eating Disorder 255M
COVID-19 194M
Bipolar Disorder 182M
Alzheimer's 125M
ALS 0.1M
Fragile X Syndrome 1M
H&N Cancer 2M
Multiple Sclerosis 3M
Rett Syndrome 7M
Parkinson's 15M
Cardiovascular Diseases 17M
Dementia 55M
Epilepsy 65M
Stuttering 70M
Autism 74M
Depression 121M
Diseases of the respiratory system
Abnormalities of breathing
Allergic rhinitis
Asthma
Conditions with disorders of intellectual disabilities
Angelman syndrome
Fragile X syndrome
Rett syndrome
Infectious & Parasitic Diseases
Influenza
COVID-19
Malaria
Mental, behavioral, or neurodevelopmental disorders
Autism spectrum disorder
Neurocognitive disorders
Depressive disorders
Diseases of the circulatory system
Pulmonary heart disease
Ischaemic heart disease
Hypertensive diseases
Diseases of the nervous system
Multiple sclerosis
Alzheimer disease
Epilepsy
Diseases of the musculoskeletal system or connective tissue
Arthropathies
Soft tissue disorders
Osteopathies
HEARHealth

Many More…


Facial Action Units
Height
Frustration
Defensiveness
Articulatory Muscle Activity
Humour
Mask Type
Abusive & Hate Speech
Anxiety
Blood Volume Pulse
Heart Rate
Cortisol Level
Skin Conductance
Menstrual Cycle Phase
Suicidal Risk
Motor Neuron Diseases


How?

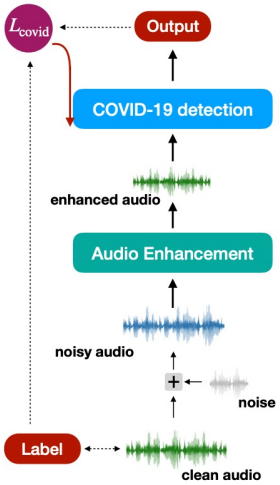
Augment



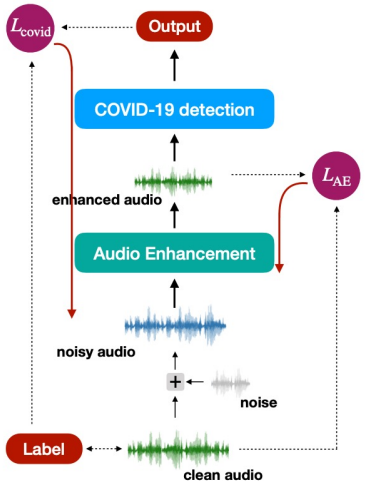
	CCS (CSS)		
%	UAR	Sensitivity	Specificity
STFT	65.6 (70.6)	50.0 (53.5)	86.0 (66.7)
Trimming	67.6 (74.6)	81.2 (92.9)	90.2 (78.7)
Augmentations			
Frequency mask	69.2 (74.3)	75.0 (99.9)	87.4 (90.4)
Time mask	68.5 (74.7)	86.2 (98.9)	87.2 (86.9)
White noise	68.4 (74.3)	85.4 (100.0)	88.1 (84.3)
Amplitude change	67.5 (73.7)	88.8 (94.4)	90.5 (81.8)
Random erase	68.5 (73.5)	83.3 (95.5)	87.4 (80.4)
Time stretch	**69.3** (73.1)	94.6 (99.6)	91.6 (86.2)
Pitch shift	65.9 (**75.3**)	87.1 (87.7)	93.8 (80.9)

"An Investigation on Data Augmentation and Multiple Instance Learning for Diagnosis of COVID-19 from Speech and Cough Sound", ICCE-Taiwan, 2023.

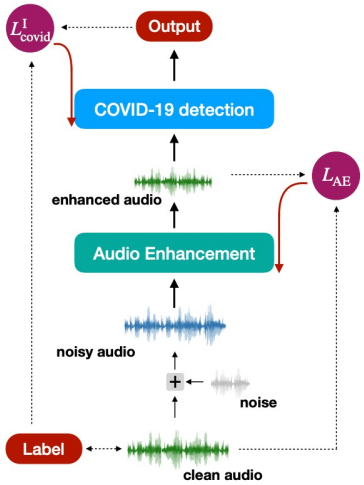
Noise?



(a) Cold cascade + data augmentation



(b) Multi-task learning



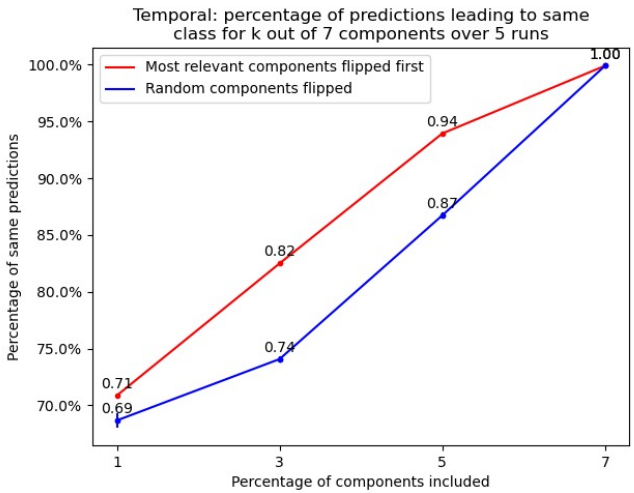
(c) Iterative optimisation

Methods	Inf	25dB	20dB	15dB	10dB	5dB	0dB	Average
Original	81.85	74.16	73.48	69.22	65.69	61.85	56.67	66.84
Cold Cascade	-	70.93	70.70	68.01	65.72	64.99	58.08	66.57
Cold Cascade + DA	-	78.42	76.33	73.65	70.02	68.48	66.74	72.27
MTL	-	**81.73**	80.62	**76.98**	**74.59**	74.45	71.15	76.59
Iterative Optimisation	-	81.35	**81.01**	76.49	74.48	**74.73**	**73.12**	**76.87**

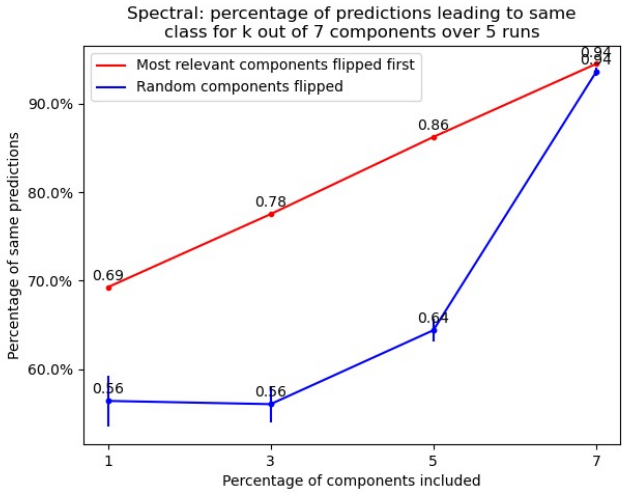
"COVID-19 Detection from Speech in Noisy Conditions", ICASSP, 2023.

Explain?

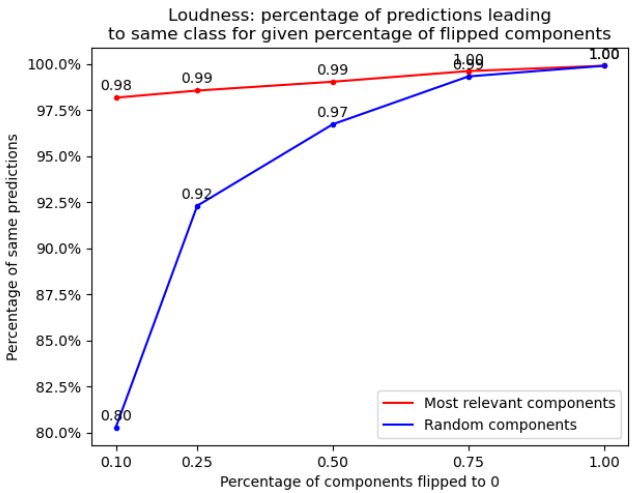
Temporal



Spectral



Loudness



"CoughLIME: Sonified explanations for the predictions of COVID-19 cough classifiers", EMBC, 2022.

Mobile?

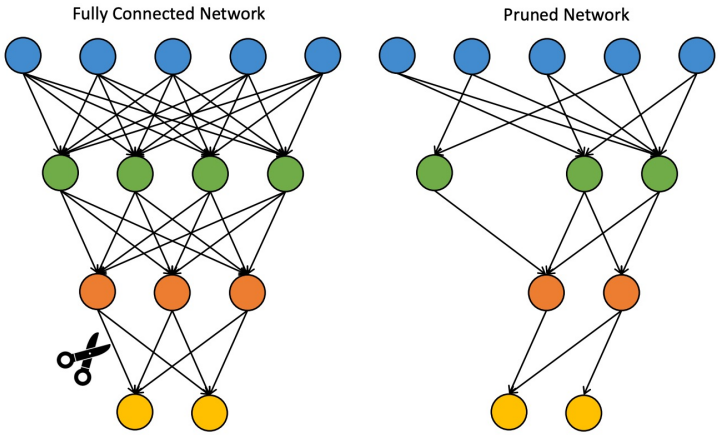
Squeezing by

Quantisation, Pruning, Teacher-Student Architectures

Equal UAR	Size/MB	%Rdtcn
Uncompressed	229	-
Quantised	55	76
Pruned	48	79
Quantised + pruned	12	95

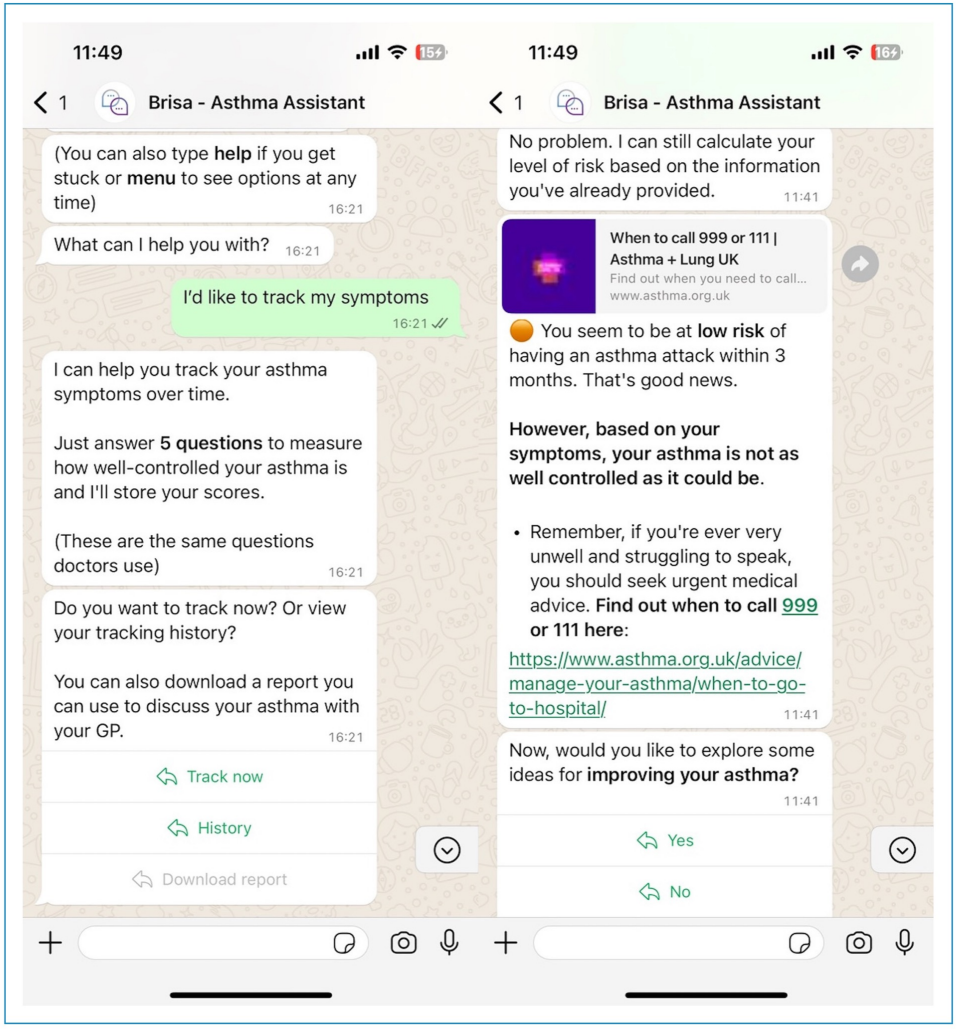
2.7831237	-2.8960736	4.000346
1.1066585	-4.8258014	-2.8570464
-3.9000013	-1.7565128	3.135015

2	-2	4
1	4	-2
-3	-1	3



"Squeeze for Sneeze: Compact Neural Networks for Cold and Flu Recognition", Interspeech, 2020.

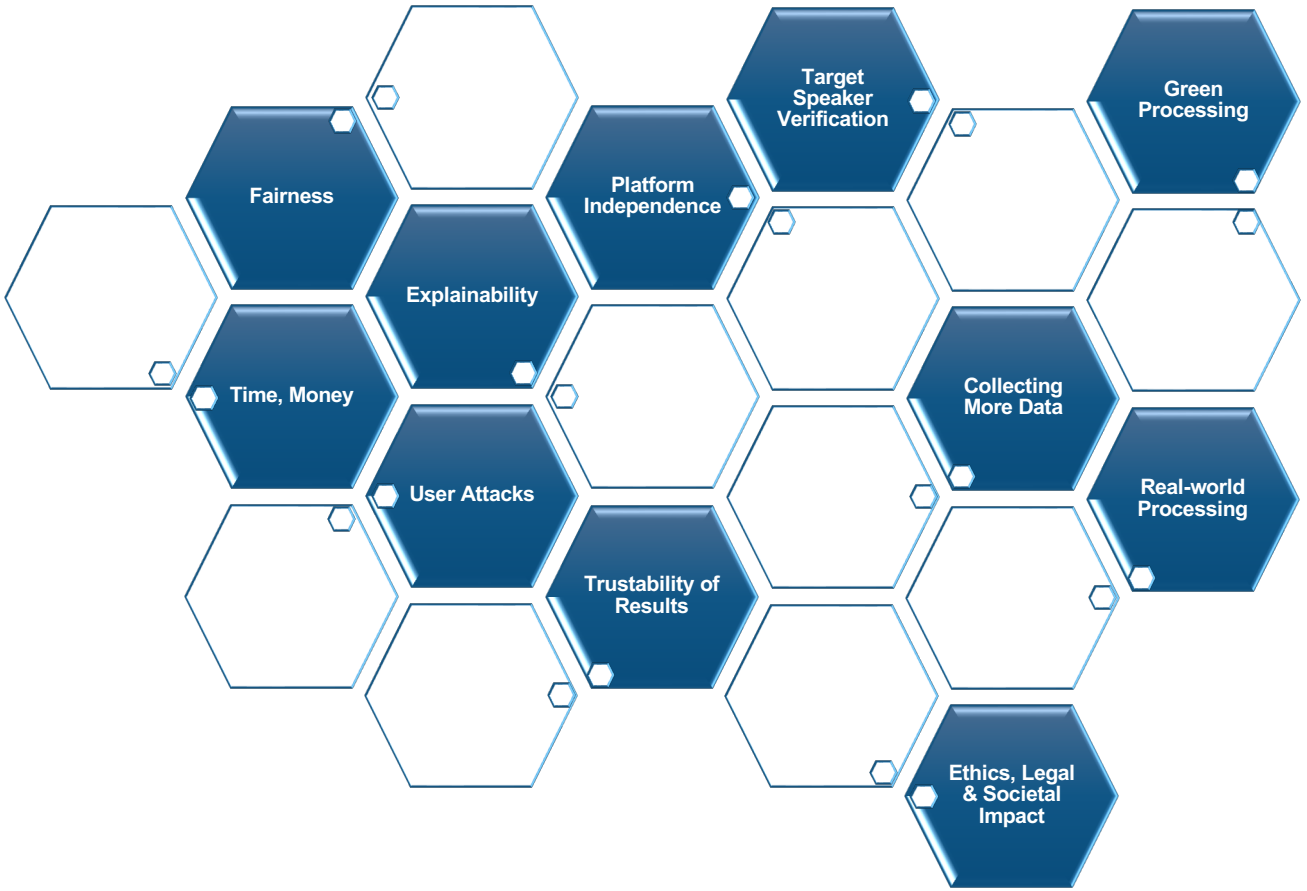
Deployment.



Demonstration of the Brisa dialogue during the asthma tracking function (left) and risk calculation (right) – WhatsApp version.

"A text-based conversational agent for asthma support: Mixed-methods feasibility study", SAGE Digital Health, 2014.

Useful and Mostly Ready…

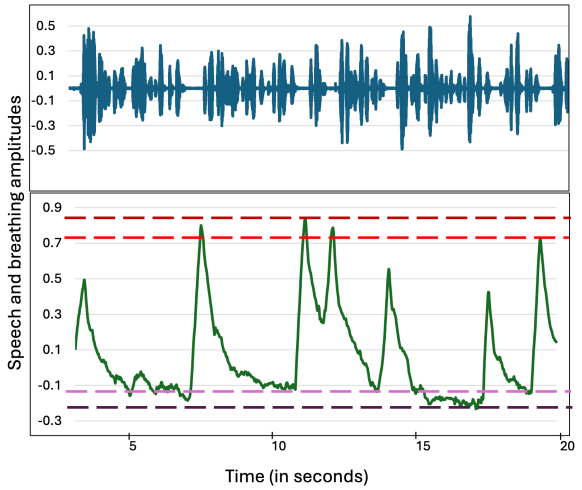


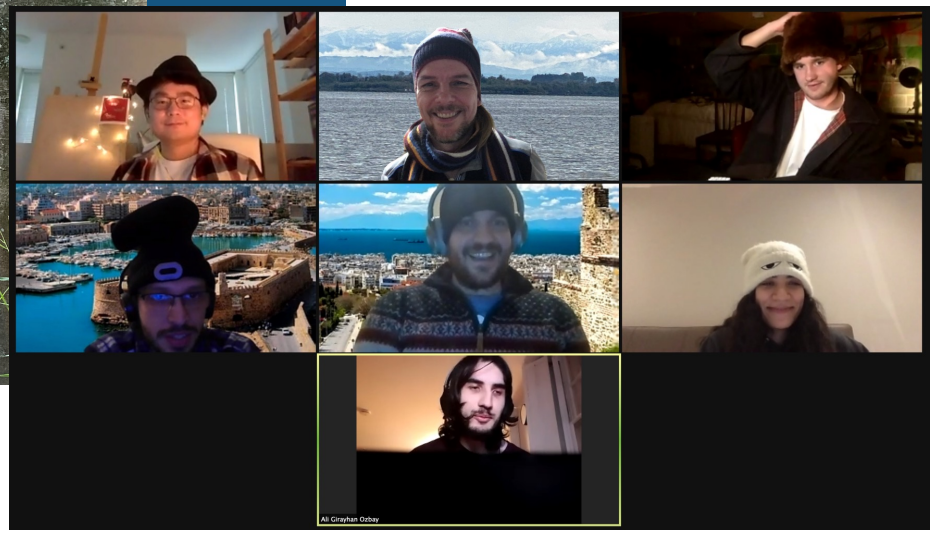
"An Overview on Audio, Signal, Speech, & Language Processing for COVID-19", arXiv, May, 2020.

audEERING
intelligent Audio Engineering

Klinikum rechts der Isar
Technische Universität München

IMPERIAL





Audio AI for Respiratory Health Monitoring :

Useful and Ready for the Next Pandemic.