COVID-19 Sounds: A Large-Scale Audio Dataset for Digital Respiratory Screening

Tong Xia*, Dimitris Spathis*, Chloe Brownˆ, Jagmohan Chauhanˆ, Andreas Grammenosˆ, Jing Hanˆ, Apinan Hasthanasombatˆ, Erika Bondareva, Ting Dang, Andres Floto, Pietro Cicuta, Cecilia Mascolo.
*first authors, ˆalphabetical order. Contact: {tx229@, covid-19-sounds@cl.}cam.ac.uk

UNIVERSITY OF CAMBRIDGE

Highlights

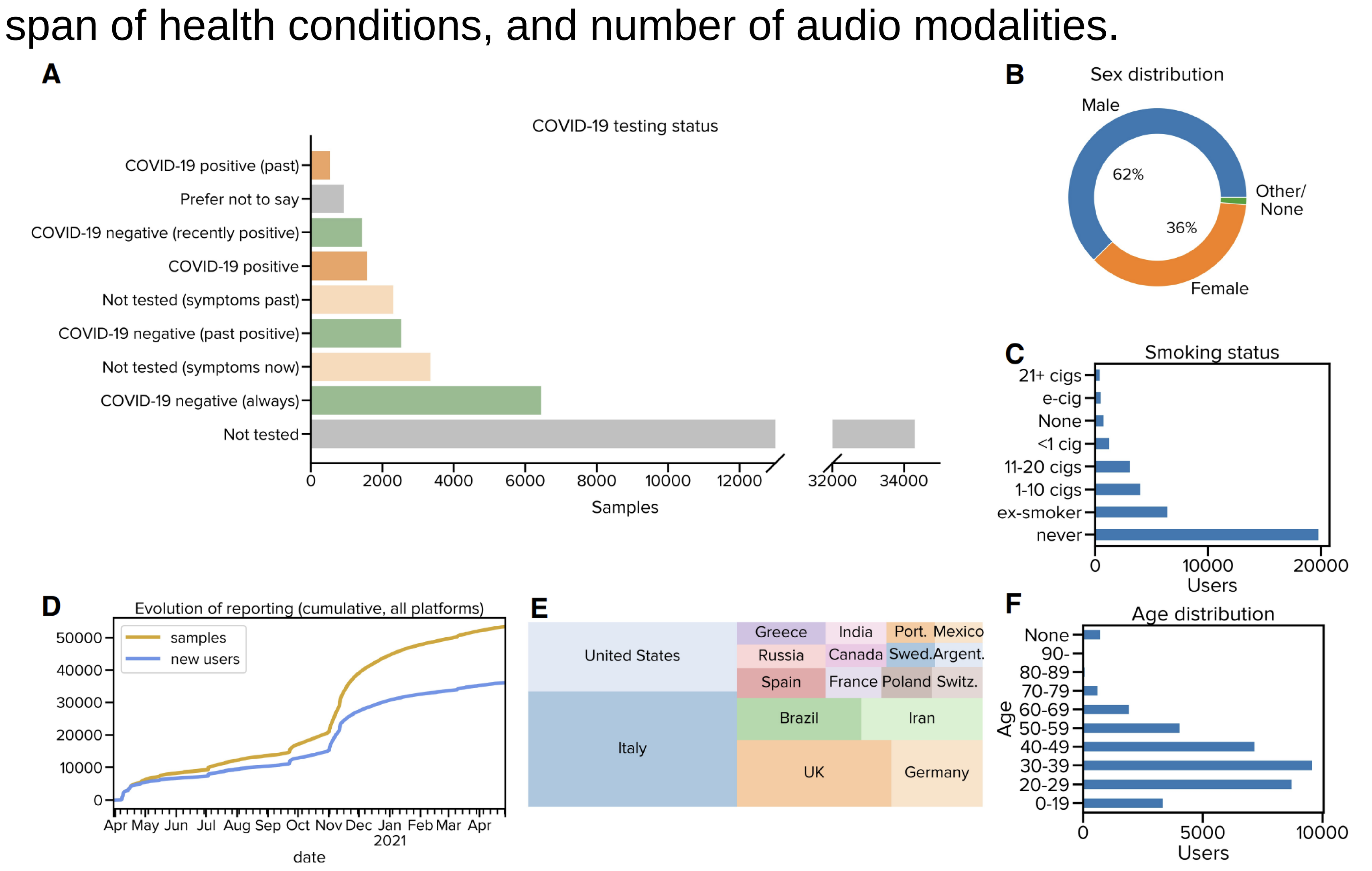
Audio signal based COVID-19 screening is non-invasive, affordable, and promising.

To facilitate the advancement and openness of audio-based machine learning for respiratory health, we release a dataset consisting of **53,449 audio samples** crowd-sourced from 36,116 participants through our ***COVID-19 Sounds app***.

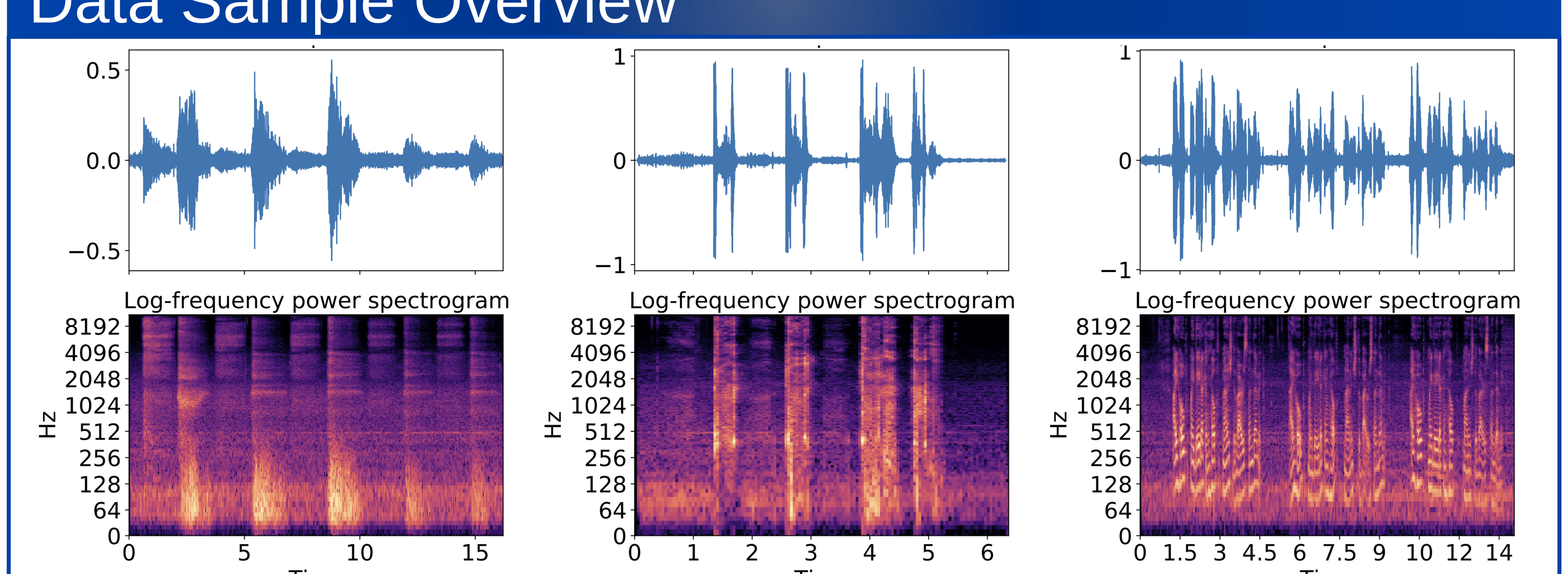
Crowd-sourced Data Collection

Early in the pandemic, we released apps for data collection, asking participants to report their ***COVID-19 test results*** along with other ***meta-data*** as well as record their ***breathing, cough, and voice***.

After a year of data collection, we release a comprehensive dataset in terms of global reach, number of languages, broad demographic range, span of health conditions, and number of audio modalities.



Data Sample Overview

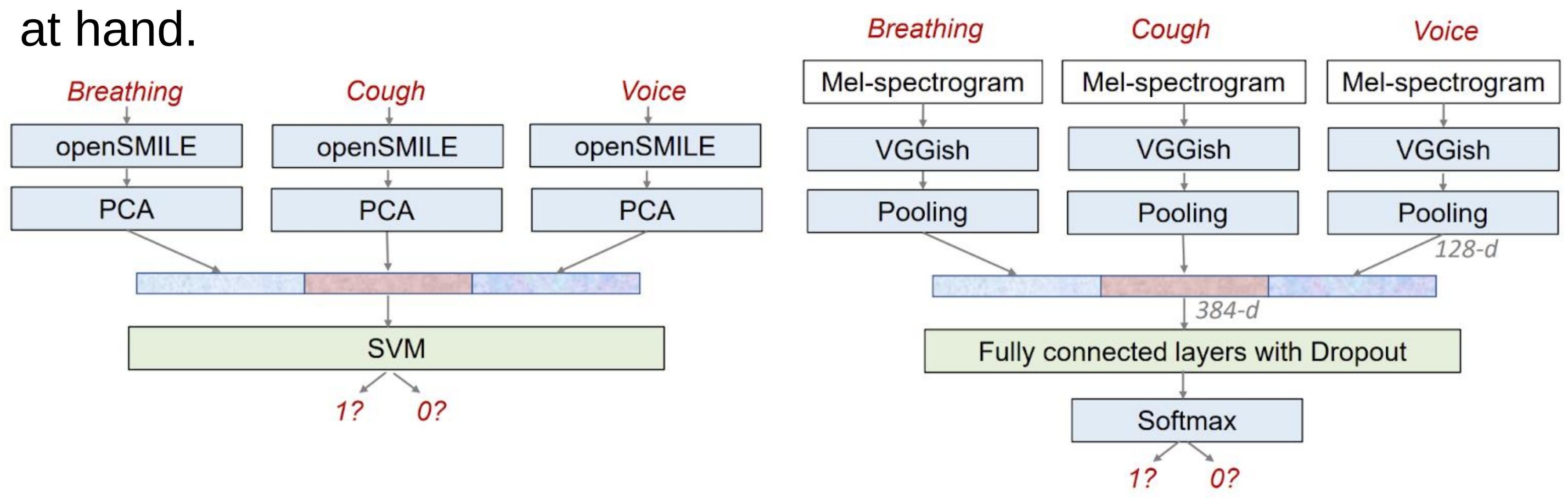


(a) Breathing recording (b) Cough recording (c) Voice recording

Sample data from a participant (a male aged over 30 with smoking history) who recently tested COVID-19 and displayed symptoms including wet cough, headache, and sore throat.

Models

We benchmarked both shallow and deep learning models. For the former, we follow an establish acoustic feature set which is then fed into SVM models. For the deep learning, we build upon a pre-trained neural network to extract generic audio features. We test two scenarios, a frozen encoder and a fine-tuned model tailored to the task at hand.



(a) Shallow model structure (b) Deep model structure

Tasks and Results

- ***Task 1 - Respiratory symptom prediction***: distinguish between healthy controls and generic respiratory abnormalities.
- ***Task 2 - COVID-19 prediction***: distinguish between tested-negative controls and tested-positive patients.

Modality&Method		Task 1 – 9,456 samples			Task 2 – 1,486 samples		
		ROC-AUC	Sensitivity	Specificity	ROC-AUC	Sensitivity	Specificity
Breathing	OpenSMILE+SVM	0.60(0.58-0.63)	0.47(0.44-0.50)	0.67(0.64-0.70)	0.56(0.50-0.61)	0.38(0.30-0.45)	0.70(0.64-0.77)
	Pre-trained VGGish	0.52(0.50-0.55)	0.07(0.05-0.08)	0.95(0.93-0.96)	0.59(0.52-0.65)	0.63(0.56-0.70)	0.49(0.41-0.56)
	Fine-tuned VGGish	0.65(0.63-0.67)	0.55(0.52-0.58)	0.66(0.63-0.69)	0.62(0.56-0.69)	0.64(0.56-0.71)	0.56(0.48-0.64)
Cough	OpenSMILE+SVM	0.70(0.67-0.72)	0.60(0.57-0.63)	0.65(0.62-0.68)	0.62(0.56-0.68)	0.56(0.48-0.63)	0.61(0.54-0.69)
	Pre-trained VGGish	0.66(0.63-0.68)	0.67(0.64-0.70)	0.53(0.50-0.56)	0.62(0.56-0.68)	0.69(0.61-0.76)	0.45(0.38-0.53)
	Fine-tuned VGGish	0.74(0.72-0.76)	0.70(0.67-0.73)	0.68(0.65-0.71)	0.66(0.59-0.71)	0.59(0.51-0.65)	0.66(0.59-0.73)
Voice	OpenSMILE+SVM	0.63(0.60-0.65)	0.56(0.53-0.59)	0.62(0.59-0.65)	0.52(0.45-0.58)	0.43(0.35-0.50)	0.62(0.54-0.69)
	Pre-trained VGGish	0.59(0.57-0.62)	0.56(0.53-0.59)	0.57(0.54-0.60)	0.61(0.54-0.67)	0.53(0.45-0.61)	0.66(0.59-0.74)
	Fine-tuned VGGish	0.69(0.66-0.71)	0.59(0.56-0.62)	0.67(0.64-0.70)	0.61(0.55-0.67)	0.57(0.49-0.65)	0.60(0.53-0.68)
Fusion	OpenSMILE+SVM	0.74(0.72-0.76)	0.65(0.62-0.68)	0.69(0.66-0.72)	0.64(0.58-0.70)	0.54(0.47-0.62)	0.67(0.60-0.75)
	Pre-trained VGGish	0.67(0.64-0.69)	0.61(0.58-0.64)	0.61(0.58-0.65)	0.64(0.58-0.70)	0.50(0.41-0.58)	0.63(0.56-0.70)
	Fine-tuned VGGish	0.75(0.73-0.77)	0.70(0.67-0.72)	0.70(0.67-0.72)	0.71(0.65-0.76)	0.65(0.57-0.72)	0.69(0.62-0.77)

Our results across both tasks show that using all modalities achieves the best performance through transfer learning, followed by the cough modality. These findings showcase the promise of audio-based models for respiratory prediction and the potential of detecting COVID-19 biomarkers (AUC>0.7).

Discussion

Based on the collected data, we have investigated a series of crucial research questions. Specifically, we have explored:

- ✓ The power of ***acoustic features*** from respiratory sounds as well as ***symptoms*** to detect COVID-19 [1,2];
- ✓ Transfer learning based end-to-end ***deep neural network*** to detect COVID-19 from limited labelled sounds [3];
- ✓ Leveraging ensemble learning to improve predictive confidence toward more ***robust COVID-19 detection*** [4];
- ✓ The impact of ***confounding factors*** on experimental results when utilizing sounds for COVID-19 screening [5];
- ✓ The potential of audio signals to capture ***COVID-19 progression*** [6].

Our Publications

1. Chole *et al. Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data*. KDD 2020. 'Better Future Award' by Computer Lab, University of Cambridge.
2. Jing *et al. Exploring Automatic COVID-19 Diagnosis via Voice and Symptoms from Crowdsourced Data*. ICASSP 2021.
3. Tong *et al. COVID-19 Sounds: A Large-Scale Audio Dataset for Digital Respiratory Screening*. NeurIPS Dataset Track 2021.
4. Tong *et al. Uncertainty-Aware COVID-19 Detection from Imbalanced Sound Data*. INTERSPPECH 2021.
5. Jing *et al. Sounds of COVID-19: exploring realistic performance of audio-based digital testing*. npj Digital Medicine 2022.
6. Ting *et al. Exploring Longitudinal Cough, Breath, and Voice Data for COVID-19 Disease Progression Prediction via Sequential Deep Learning: Model Development and Validation*. JMIR 2022, appear soon.