SCOMAE: Diagnosing Tuberculosis From Cough Acoustics Using Supervised COntrastive Masked AutoEncoders

Xu, Michael (School: The Taft School)
Nabulsi, Faisal (School: Acalanes High School)

Tuberculosis (TB), a leading cause of infectious disease death worldwide, disproportionately affects lower-and-middle-income countries (LMICs). When diagnosed early, TB is very curable, but when left undiagnosed for long periods of time, it can be deadly to the patient and spread throughout the community. Especially in LMICs, current diagnostic methods are slow, resource-intensive, and inaccessible, resulting in the prolific spread of TB within the host and within communities. In order to eradicate TB and ameliorate health disparities, more accessible and scalable diagnostic methods must be introduced. Cough, the most common symptom of TB, exhibits unique acoustic characteristics that vary across different respiratory diseases, suggesting potential for diagnosis based on cough sounds. Previous attempts to develop a machine learning-based algorithm to detect TB from cough audio have largely been unsuccessful due to limited availability of labeled data. In this work, we design and implement a novel self-supervised learning architecture optimized for TB detection from cough sounds, combining Masked Autoencoders and Supervised Contrastive Learning. We pretrain our network on a vast corpus of audio data and then fine-tune it on a dataset of labeled cough recordings from TB patients. Evaluated on a public benchmark, our model achieved an area under the receiver operating characteristic curve (AUROC) of 0.843 on the task of TB detection from cough sounds, outperforming all published techniques to the best of our knowledge. This breakthrough not only sets a new standard for cough-based TB diagnosis, but also marks a significant stride towards the global eradication of TB.