# TB or not TB? Acoustic cough analysis for tuberculosis classification

Geoffrey Frost[1], Grant Theron[2], Thomas Niesler[1]

[1]Department of Electrical and Electronic Engineering, Stellenbosch University, South Africa
[2]SAMRC Centre for Tuberculosis Research, DSI/NRF Centre of Excellence for Biomedical Tuberculosis Research, Stellenbosch University, South Africa

## Background

- In 2021, 10 million people were reported to have developed tuberculosis (TB).
- TB is typically a respiratory disease resulting in patients developing a chronic cough.
- TB cough classification is a potentially low-cost solution for rapid TB screening.

## Contributions

- Improve upon previous best methods.
- Analyze the distinct characteristics of the feature space learnt to be important for cough classification.
- Improve decision threshold generalization.

## Data

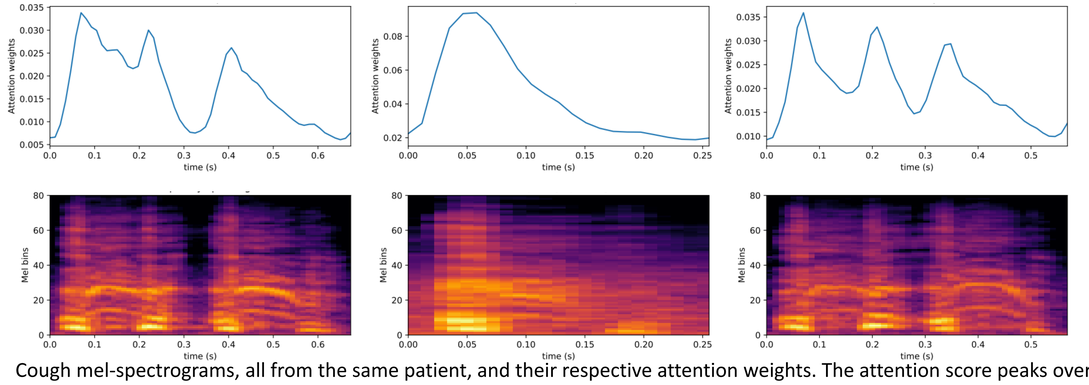
- Divided into train and test sets (49 and 25 patients respectively).
- 4-fold cross-validation is used during development

| | TB | $\overline{\text{TB}}$ | Total |
|---|---|---|---|
| Patients | 28 | 46 | 74 |
| Total coughs | 844 | 720 | 1564 |
| Mean cough length (s) | 0.60 | 0.64 | 0.62 |
| Std dev cough length (s) | 0.34 | 0.29 | 0.32 |

## Acoustic Signature of Cough

### Attention Weights

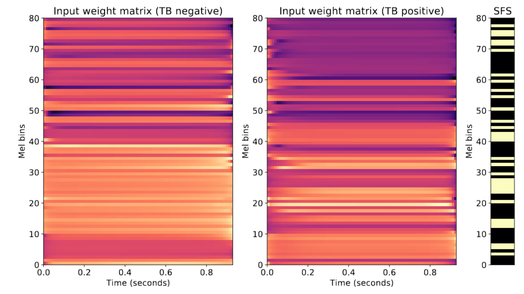
- Large importance for temporal regions where the signal has a high power and a large bandwidth.
- Coincide with initial voiced portions at the beginning of each coughing episode.
- This portion of the coughing sound originates from inside the lung itself (the bronchi)



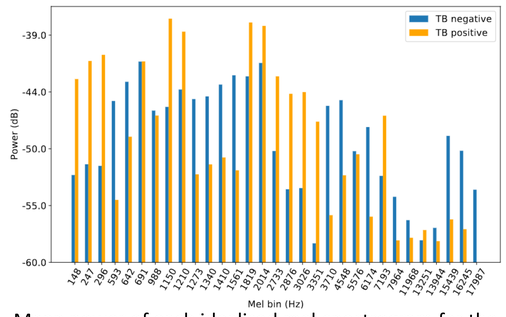

Cough mel-spectrograms, all from the same patient, and their respective attention weights. The attention score peaks over the initial burst of energy in each coughing episode

### Neural Style Transfer

- Clear differences between idealized TB and non-TB coughs.
- Higher power in TB cough for frequencies <500Hz and between 1.8kHz to 3.3kHz.
- Non-TB cough contains energy in bands far outside the range of the human speech (>8kHz).

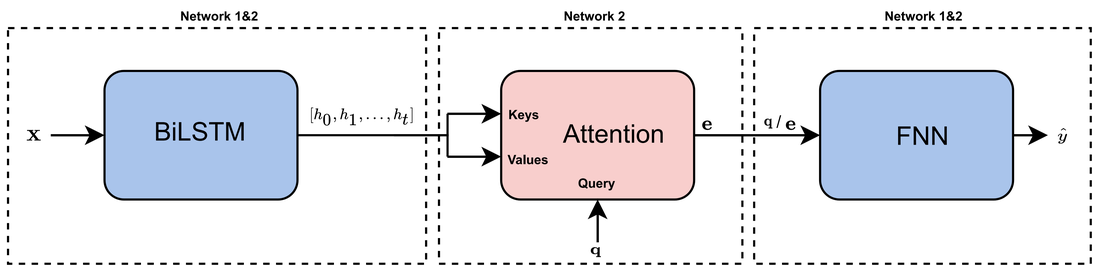



Idealised mel-spectrograms for TB negative (left) and positive (centre) coughs, with bins identified by SFS shown in yellow (right).




Mean power of each idealised mel-spectrogram for the frequency bins deemed most important by SFS.

## Architectures




Structure of the basic BiLSTM (Network 1) and its attention variant (Network 2), with shared components indicated.

Loss function: $$\mathcal{L} = -\frac{1}{B}\left(\sum_{b}^{B} \beta \cdot \mathbf{y_b} \cdot \log(\hat{\mathbf{y}}_\mathbf{b}) + \alpha \sum_{j,i} L_{GE2E}(\mathbf{e}_{ij})\right)$$

## Classification Performance

- BiLSTM-Att (SFS) achieves smallest decision threshold standard deviation, indicating better generalization.

Mean and standard deviation of the area under the ROC curve (AUC) and EER decision thresholds ($\bar{\gamma}$) observed during 4-fold cross validation.

| Model | $\bar{\gamma}$ | AUC |
|---|---|---|
| LR (*baseline*) | $0.272 \pm 0.103$ | $0.701 \pm 0.127$ |
| BiLSTM | $0.534 \pm 0.108$ | $0.777 \pm 0.094$ |
| BiLSTM (SFS) | $0.603 \pm 0.155$ | $0.919 \pm 0.081$ |
| BiLSTM-Att | $0.460 \pm 0.175$ | $0.873 \pm 0.054$ |
| BiLSTM-Att (SFS) | $0.568 \pm 0.070$ | $0.900 \pm 0.092$ |

- BiLSTM achieves sensitivity and specificity of 0.89 and 0.75 respectively

Test set performance for the investigated models evaluated using sensitivity, specificity, accuracy and AUC.

| Model | Sens | Spec | Acc | AUC |
|---|---|---|---|---|
| LR (*baseline*) | **0.889** | 0.625 | 0.720 | 0.769 |
| BiLSTM | **0.889** | 0.750 | **0.800** | 0.821 |
| BiLSTM (SFS) | 0.667 | 0.750 | 0.720 | **0.862** |
| BiLSTM-Att | 0.778 | 0.625 | 0.680 | 0.822 |
| BiLSTM-Att (SFS) | 0.778 | **0.813** | **0.800** | 0.850 |

The authors gratefully acknowledge funding from the EDCTP2 programme supported by the European Union (grant RIA2020I-3305, CAGE-TB)