
Article:
Romero, H.E., Ma, N., Brown, G. orcid.org/0000-0001-8565-5476 et al. (1 more author) (2022) Acoustic screening for obstructive sleep apnea in home environments based on deep neural networks. IEEE Journal of Biomedical and Health Informatics, 26 (7). pp. 2941-2950. ISSN 2168-2194

https://doi.org/10.1109/jbhi.2022.3154719

Acoustic Screening for Obstructive Sleep Apnea in Home Environments Based on Deep Neural Networks

Hector E. Romero, Ning Ma, Guy J. Brown and Elizabeth A. Hill

***Abstract*— Obstructive sleep apnea (OSA) is a chronic and prevalent condition with well-established comorbidities. However, many severe cases remain undiagnosed due to poor access to polysomnography (PSG), the gold standard for OSA diagnosis. Accurate home-based methods to screen for OSA are needed, which can be applied inexpensively to high-risk subjects to identify those that require PSG to fully assess their condition. A number of methods that analyse speech or breathing sounds to screen for OSA have been previously investigated. However, these methods have constraints that limit their use in home environments (e.g., they require specialised equipment, are not robust to background noise, are obtrusive or depend on tightly controlled conditions). This paper proposes a novel method to screen for OSA, which analyses sleep breathing sounds recorded with a smartphone at home. Audio recordings made over a whole night are divided into segments, each of which is classified for the presence or absence of OSA by a deep neural network. The apnea-hypopnea index estimated from the segments predicted as containing evidence of OSA is then used to screen for the condition. Audio recordings made during home sleep apnea testing from 103 participants for 1 or 2 nights were used to develop and evaluate the proposed system. When screening for moderate OSA the acoustics based system achieved a sensitivity of 0.79 and a specificity of 0.80. The sensitivity and specificity when screening for severe OSA were 0.78 and 0.93, respectively. The system is suitable for implementation on consumer smartphones.**

***Index Terms*— Obstructive sleep apnea, screening, acoustic analysis, deep learning, smartphone.**

I. Introduction

OBSTRUCTIVE sleep apnea (OSA) is a chronic and prevalent condition that results from the reversible collapse of the upper airway during sleep, resulting in a full cessation (apnea) or significant reduction (hypopnoea) of airflow [1]. Left untreated, OSA has well-established cardiovascular [2], cerebrovascular [3], neurocognitive [4], and metabolic [5] comorbidities [6], [7]. OSA affects approximately 1 billion people worldwide [8], and its prevalence is increasing [9], though many individuals remain undiagnosed due to poor access to appropriate healthcare [10], [11].

H. E. Romero, N. Ma, and G. J. Brown are with the Department of Computer Science, University of Sheffield, Sheffield S1 4DP, UK. E-mail: {h.e.romero.ramirez, n.ma, g.j.brown}@sheffield.ac.uk. E. A. Hill is with the Sleep and Circadian Neuroscience Institute (SCNi), Nuffield Department of Clinical Neurosciences, University of Oxford, and the Sleep Research Unit, Centre for Clinical Brain Sciences, University of Edinburgh. Email: lizzie.hill@ndcn.ox.ac.uk.

Level I attended polysomnography (PSG) [12] is the reference-standard test for objective measurement of sleep and wake, and for diagnosis of a number of sleep disorders, including OSA. Multiple channels of physiological data, including brain activity (electroencephalography; EEG), eye movements (electro-oculography; EOG) and muscle activity (electromyography; EMG), are recorded overnight, allowing experienced sleep technologists to delineate sleep and wake stages in line with recognised international guidelines [13]. However, PSG is a resource-heavy procedure with significant patient burden. The procedure is generally conducted in an in-patient sleep laboratory setting, requiring expensive specialist equipment and highly trained staff – usually Registered Polysomnographic Technologists (RPSGTs) – to set up, monitor and score the study.

Testing can be performed in the home using Home Sleep Apnea Testing (HSAT). This is a level III study [12], which records a limited number of channels of cardio-respiratory data, not usually including EEG/EOG/EMG. Therefore, HSAT allows assessment of apnoeas, hypopnoeas and oxygen desaturations, but does not quantify sleep *per se*. The reduced number of sensors (which can be described using the SCOPER classification [14]) and portability of equipment means that HSAT can be conducted in the patient's usual sleeping environment, without the inconvenience of an attended in-patient stay, and with a significant cost saving for the sleep service [15].

Severity of OSA is assessed using the apnea-hypopnea index (AHI), defined as the average number of apneas and hypopneas per hour of sleep on PSG, or per time in bed on HSAT. Current international scoring guidelines [13] define an apnea as a $\geq$90% reduction in nasal/oral airflow for $\geq$10 seconds. A hypopnea is defined as a $\geq$30% reduction in nasal airflow for $\geq$10 seconds, and, on PSG, must be associated with either a $\geq$3% oxygen desaturation and/or EEG arousal. Scoring guidelines for HSAT do not include an arousal criterion given the absence of electrophysiologically-derived sleep staging [13]. As a result of the differing scoring criteria and the use of time in bed rather than total sleep time as a denominator, a dilution effect on the AHI between PSG and HSAT is well-documented, i.e. HSAT systematically underscores the AHI [16]–[18]. Despite this, current international guidelines recommend the use of HSAT as an acceptable alternative to PSG [19] in uncomplicated adult patients with a high clinical suspicion of moderate-to-severe OSA, and HSAT is the first line test in many sleep services, particularly in the

UK and Europe. The generally-accepted standard AHI cut-offs for classification of OSA severity in adults, based on expert consensus [20] are shown in Table I.

TABLE I: OSA Severity and AHI

OSA Severity	AHI (events/hour)
Normal	< 5
Mild	5 – 15
Moderate	15 – 30
Clinically relevant	> 25
Severe	> 30

PSG and HSAT suffer from a number of disadvantages. In both cases – but particularly in the case of PSG – patients are required to sleep with multiple sensors attached to their head and body, which limit movement and can be uncomfortable. PSG is known to be particularly susceptible to a 'first-night effect', in which the first night of testing displays more sleep fragmentation, longer initial sleep latency, less total sleep time, and more wakefulness when compared to successive nights [21], [22]. Although OSA presence and severity can differ substantially from night to night, due to limited diagnostic resources and the high cost of PSG most laboratory sleep studies are limited to one night, which might not be representative of the actual condition of the subject under study [23], [24]. A further concern is that availability of sleep laboratories has been adversely affected by demands that the COVID-19 pandemic has placed on respiratory wards, a situation that is likely to persist for some time [25]. HSAT is better in this regard, but still requires time and expense to thoroughly clean equipment between each study. Screening methods for OSA are therefore needed, which can be deployed in the home and use inexpensive equipment (such as the patient's own smartphone), in order to identify those that require PSG to fully assess their condition [11], [26].

Here, we focus on methods for OSA screening based on sound recordings made in the home. OSA displays symptoms with particular acoustic characteristics, for example, snores, chokes, loud gasps and absence of breathing. Furthermore, breathing sounds and speech have similarities: both are time series signals produced in the vocal tract with a similar frequency range. Drawing inspiration from speech technology tools, a number of methods that analyse breathing sounds or speech in order to screen for OSA have been previously published. These are summarised in Table II. It is worth noting that direct performance comparison between these is not possible, as the methods are not evaluated on the same kind of data (e.g., tracheal sound recordings vs. ambient sound recordings). Rule-based approaches have been proposed in many studies. Such approaches are unlikely to effectively capture the high variability of breathing sounds during sleep and to be robust enough in uncontrolled acoustic conditions. For instance, the performance of a system that detects apneas by simply looking for silent segments in sleep audio recordings will degrade in the presence of noise. Al-Mardini et al. [27] screened for OSA using sound energy, oxygen saturation and body movement collected with a smartphone and specialised equipment (a tracheal microphone). Sound sample entropy and additional non-acoustic data, oxygen saturation, were used to predict AHI values by Castillo et al. [28] from audio recordings made with a smartphone attached to the subject's chest. Saha et al. [26] estimated AHI values using oxygen saturation, tracheal sounds and respiratory related movements recorded with a device placed on the suprasternal notch.

Other studies have implemented systems based on machine learning techniques. Yadollahi et al. [29] predicted AHI values with a fuzzy algorithm using tracheal sound energy and oxygen saturation. Goldshtein et al. [30] developed a Gaussian mixture model (GMM) classifier using time and frequency features from speech to screen for OSA. Using speech might be a limitation, since breathing sounds are not directly analysed. Logistic regression with acoustic features and additional non-acoustic data – sleep stage from scored PSG – were used by Kim et al. [31] to predict the OSA severity from ambient sound recordings.

Two previous studies are notable for their use of deep learning. Nakano et al. [32] detected apnea events and predicted sleep status to compute the AHI using spectrogram images from tracheal sound recordings as input to convolutional neural networks (CNNs). Tiron et al. [33] estimated AHI values and screened for OSA by detecting snoring from ambient sound recordings with a CNN, and analysing active sonar reflections recorded with a smartphone.

All of these studies, with the exception of the one by Castillo et al. [28], used data collected in the controlled conditions of a sleep clinic. This limits their applicability in typical sleep environments – a bedroom at home – where a range of different room acoustics and higher levels of background noise will be found, which will negatively impact upon the performance of screening systems.

Detecting OSA events using acoustics is a challenging task, since they are caused by a collapse of the upper airway that results in an absence of airflow [28]. Directly detecting such periods of relative silence is generally not possible in the presence of non-stationary noise. An alternative to detecting individual OSA events is to analyse the temporal pattern of sleep breathing sounds in large analysis windows, as the events surrounding an apnea or hypopnea typically provide salient cues that indicate its occurrence (e.g., the recovery breath or loud gasp after an apnea followed by regular breathing). Here, we implement this approach within a deep neural network (DNN) and show that it is an effective strategy for OSA screening from whole-night sleep audio recordings.

II. Data

Data collection from home environments was undertaken, as no suitable dataset was available from previous studies. Data collection and storage protocols were subjected to the ethical review procedures of the University of Sheffield.

A. Home sleep apnea testing and audio recordings

103 participants (67 males and 36 females) were included in this home-based study. The average age of the participants was

TABLE II: Parts of related acoustics based OSA screening studies

Task	Study	Data	Reference	Performance
OSA screening	Goldshtein et al. [30]	Speech	PSG	Sensitivity: 0.83 Specificity: 0.81
	Al-Mardini et al. [27]	Tracheal sounds with a smartphone	PSG	Sensitivity: 1.00 Specificity: 0.86
OSA severity prediction	Kim et al. [31]	Ambient sound at clinic	PSG	Sensitivity: 0.88 Specificity: 0.88
AHI estimation	Yadollahi et al. [29]	Tracheal sounds at clinic	PSG	Sensitivity: 0.90 Specificity: 0.90
	Nakano et al. [32]	Tracheal sounds at clinic	PSG	Sensitivity: 0.98 Specificity: 0.76
	Castillo et al. [28]	Ambient sound at home	HSAT	Sensitivity: 0.76
	Saha et al. [26]	Tracheal sounds at clinic	PSG	Sensitivity: 0.91 Specificity: 0.89
OSA screening and AHI estimation	Tiron et al. [33]	Ambient sound with a smartphone at clinic	PSG	Sensitivity: 0.88 Specificity: 0.80

45±13 years and the average BMI was 31±7. Each participant took HSAT for one or two nights while audio was recorded at the same time. In total 157 nights of recordings were collected and the average recording duration per night was 7±1.4 hours. The detailed demographic information is shown in Table III.

TABLE III: Demographics of the participants included in this study. The percentages of data groups, data ranges, and averages with standard deviations are also given.

Total Participants	103	
Males	67	65%
Females	36	35%
Age (years)	45 ± 13	25 – 71
BMI (kg/m^2)	31 ± 7	19 – 48
AHI (events/hour)	25 ± 25	1 – 114
Recording Duration (hours/night)	7.0 ± 1.4	3.0 – 9.8
Total Nights	157	
Normal AHI $<$ 5 (nights)	13	8%
Mild 5 $\leq$ AHI $<$ 15 (nights)	67	43%
Moderate 15 $\leq$ AHI $<$ 30 (nights)	38	24%
Severe AHI $\geq$ 30 (nights)	39	25%

HSAT was carried out using a SOMNOmedics SOMNOtouch RESP [34], which consists of several channels with attached sensors that record oxygen saturation (SpO_2), heart rate, nasal airflow, respiratory effort, snore (derived from airflow and sampled at 256 Hz), sleep-wake status, and body position. Participants were instructed to sleep on their own. During each HSAT session, audio recordings were made simultaneously using a smartphone (iOS or Android) placed next to the bed at head level. Audio recordings were collected using a purpose-built Sleep Study App, which records sound with a sampling frequency of 16 kHz and 16 bit resolution, and sends 2-minute blocks of audio data continuously to a central server throughout the HSAT session.

HSAT data was annotated for apnea events by a technologist with the RPSGT qualification certified by the Board of Registered Polysomnologists (USA, AASM). Audio recordings were checked for quality. Data was excluded from this study if the HSAT recordings could not be scored due to corrupted sensor data (e.g., missing flow channel due to misplaced nasal cannula), the audio recording was missing, or the data was too short (less than 4 hours). Data from 17 participants had issues with HSAT device sensors and could not be synchronised with the audio data. Data from 8 participants had issues with audio recordings, which had corruptions or interruptions (caused by incorrect use of the app and poor placement or orientation of the device). 5 participants had audio recordings missing (caused by streaming problems). Furthermore, the first 20 and the last 2 minutes from the audio recordings of each night were excluded, as these normally correspond to moments where the participants were awake. The final OSA corpus consisted of 103 participants, and 157 nights amounting to over 1,094 hours of data.

We note that the incidence of problems due to audio corruption or interruption (8) was half that of problems due to the HSAT equipment. This in part justifies the proposed acoustic-based method, which minimises the number of sensors used. In practice, we believe the acoustic-based method would have even higher usability when users do not have to wear the HSAT sensors and there is no audio streaming.

An important difference between the present study and previous ones is that the audio recordings were collected in a real home setting where the proposed OSA screening system is intended to be used. The proposed system must deal with several challenges in home environments: varying room acoustics (reverberation), complex acoustic environments at home (e.g., some participants slept with the TV on or a window open), smartphones with distinct microphone characteristics (iPhones, different Android brands), and suboptimal recording procedures (e.g., because the participant placed the smartphone too far away, or with the microphone facing in the wrong direction). Therefore, there are significant variabilities in the audio recordings that are representative of the intended use case. This is in contrast with several other relevant studies where acoustic signals were recorded in a controlled quiet environment (e.g., in a sleep laboratory [26], [27], [29], [31]–

[33] or using tracheal sound recordings [35]).

B. Synchronisation between HSAT and audio recordings

The scored HSAT data (inhalations, exhalations, desaturations, snores, and apnea-hypopnea events) provides references for the acoustic recordings. However, because the SOMNOtouch device clock may not be accurate or tightly synchronised with the smartphone, the timestamps in the data could not be used on their own to synchronise the HSAT and audio recordings. An algorithm for automatic synchronisation was, therefore, developed.

The proposed synchronisation algorithm uses a 20-minute segment of the audio recordings after the HSAT recording started. The *snore channel* signal from HSAT (sampled at 256 Hz) was correlated with the audio segment (sampled at 16 kHz) as follows:

1) Downsample the audio signal to the HSAT sampling frequency of 256 Hz;
2) Calculate an approximate initial time difference between the audio signal and the snore channel using their timestamps;
3) Use the initial time difference to identify corresponding 20-minute segments from the snore channel and audio signal;
4) Half-wave rectify the snore and audio segments and scale them to the range [0,1];
5) Compute the cross-correlation function (C) of the snore (s) and audio (a) segments as

$$C(\tau) = \sum_{n=0}^{N-1} s(n)a(n-\tau) \tag{1}$$

where $\tau \in [0, N)$ is the time delay and N is the number of samples in the segment. In this study, $N = 307,200$ for a 20-minute segment sampled at 256 Hz;
6) Adjust the initial time difference using the time delay indicated by the peak of the crosscorrelation function;
7) Add the adjusted time difference to the timestamps of the HSAT data.

In a small number of cases the audio recordings and the snore channel from HSAT became desynchronised towards the end of the night. This is likely caused by the clock in the HSAT device not having an exact frequency of 256 Hz (a range of 255.995 Hz – 256.007 Hz was reported by the manufacturer [36]). Resampling the audio recordings at a slightly lower rate (255.997 Hz) resolved this desynchronisation issue.

III. Acoustic-Based OSA Screening

We propose a deep learning-based automatic method to screen for OSA, based on the analysis of breathing sounds during sleep. We hypothesised that the temporal pattern of respiration can be exploited to robustly screen for OSA using ambient sound recordings made with readily available hardware in typical sleep conditions. This was achieved by predicting the presence or absence of apnea-hypopnea events in large audio recording segments with a DNN, and using the percentage of predicted apnea-hypopnea segments in a night to screen for OSA, and estimate the AHI.

An overview of the proposed system is given in Figure 1. The whole-night sleep audio recordings were divided into overlapping segments with a 10-s window shift. A long temporal window (30-s to 40-s) was used in order to effectively capture the temporal pattern of respiration, since an apnea/hypopnea event typically lasts for 20–30 seconds. The 10-s window shift allowed the system to have a relatively high temporal resolution to capture an apnea/hypopnea event. Each audio segment was labeled as either having an apnea/hypopnea event in it or not according to the scored HSAT data. After this, a time-frequency representation based on an auditory model [37] was computed for every segment and provided as input features to the system. The DNN classified each segment as either containing OSA events or not. Since adjacent segments (10-s window shift) could contain the same OSA event, their labels were grouped together to form the same OSA event if they were all classified as containing OSA. Finally, the predicted OSA event number was used to compute the AHI for the whole night.

A. Acoustic signal processing

Auditory-motivated signal representations have been successfully used for speech recognition and sound classification [37]. Here, Mel-frequency analysis was used which is based on human perception experiments. Audio signals were sampled at 16 kHz. The power-spectrogram was first computed by applying a short-time Fourier transform (STFT) to a 50 ms frame with 20 ms frame shift and a Hann window. A bank of 64 Mel-filters, which have overlapping pass-bands and are logarithmically spaced in frequency between 70 Hz and 7.5 kHz, was applied to the power-spectrogram. This process creates Mel-spectrograms, commonly known as Mel-filterbank features [37]. They have an expanded low-frequency representation compared to the linear frequency scale of the STFT.

The filterbank features were then divided into overlapping segments with a 10-s shift window. In this study segment windows in the range of 30-s to 40-s were considered which correspond to 1,500 to 2,000 frames in each segment, respectively. The two-dimensional (64×1,500 to 64×2,000) features were finally normalized to have zero mean and unit variance, and used as the input to the DNN. The Mel-filterbank feature representations of various breathing events are illustrated in Fig. 2.

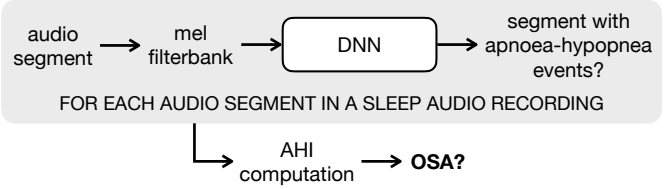
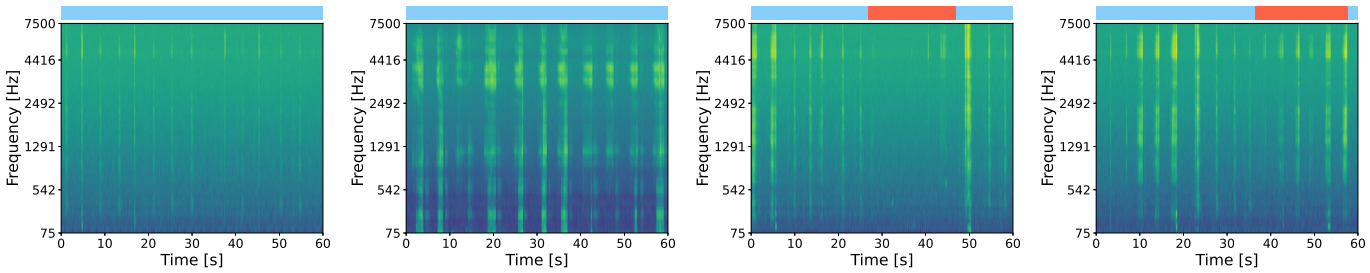


Fig. 1: Schematic diagram of the proposed system to screen for OSA.



(a) Health breathing (b) Snoring (c) Apnea (d) Hyponea

Fig. 2: Mel-filterbank features for 60-second audio segments from the corpus. (a) Healthy breathing. Periodic low energy events are observed throughout the segment. (b) Snoring. Periodic high energy events are observed throughout the segment. (c) Apnea. The apnea corresponds to a region of very low energy (silence) marked by the orange patch at the top of the figure, followed by a gasping noise. (d) Hypopnea. Low energy events corresponding to shallow breathing are seen, in the region marked by the orange patch.

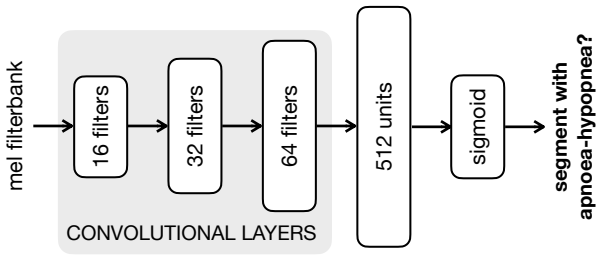


Fig. 3: Proposed CNN architecture for OSA screening.

B. Deep neural network

The DNN architecture is a convolutional neural network [38]. It consists of three convolutional layers with a kernel size of 3×3 and 16, 32, and 64 filters. A 4×3 max-pooling layer is applied to the output of each convolutional layer, as well as a batch normalization layer. The 'relu' activation function was used for convolutional layers. A dropout rate of 0.3 was used to help regularize the CNN. The output of the convolutional layer was flattened and passed to a fully connected layer with 512 'relu' activation units. Finally, the classification is carried out by a fully connected layer having one 'sigmoid' activation unit with a binary output: 'segment with apnea-hypopnea events' or 'segment with no apnea-hypopnea events'. This is illustrated in Fig. 3. The network was trained with a learning rate of 0.001, and a batch size of 128. Binary cross-entropy was used as the loss function with the 'Adam' optimizer. Convergence was commonly reached within 50 epochs. The system was developed using TensorFlow [39], and the hyperparameters were selected heuristically based on previous sound classification experiments. Different number of filters, number of convolutional layers, and kernel sizes were considered while trying to keep the network small so it would properly generalise, and be sufficient compact for deployment on a smartphone.

C. AHI estimation

The DNN does not directly detect individual apnea/hypopnea events. To estimate the AHI for a whole night, the system first groups adjacent segments that are predicted to contain an apnea/hypopnea event into a single event. This is based on the observation that an OSA event tends to last around 30 seconds, and thus if two adjacent segments with 10-s shift are both predicted to contain an OSA event, it is likely that they belong to the same event. The number of apnea/hypopnea events is then counted over the whole-night recordings and AHI is computed by dividing the number of events by the duration of the recording. As noted in the introduction, the use of recording time rather than total sleep time as a denominator has a diluting effect on the predicted AHI, which is also a known limitation of HSAT [16]–[18].

IV. Experiments

Experiments were conducted on the OSA corpus using 10-fold cross-validation. The 103 participants were randomly divided into 10 folds; 10 participants per fold, except the last fold which had 13 participants. For each cross-validation run, data from the participants of the fold (10 or 13 participants) was used as the evaluation dataset, and the data from the remaining 9 folds (90 or 93 participants) was split into training and validation datasets based on a 9:1 ratio. Cross-validation was performed across all 10 folds, ensuring that evaluation data was not used for training the system. Note that since participants recorded one or two nights, each fold included different numbers of nights depending on the data available. Cross-validation was performed across participants rather than recording nights to increase generalizability to unseen patients.

A. Evaluation framework

The system estimated the AHI for each night and was evaluated using different AHI cut-off points; 5, 10, 15, 20, 25 and 30 events/hour. The number of nights below and above each AHI cut-off point is displayed in Table IV. The generally accepted AHI cutoff thresholds (5, 15, and 30) [19] are highlighted in bold.

Sensitivity, specificity, receiver operating characteristic (ROC) curve, and area under the ROC curve (AUC) were reported for the different system configurations, as these reflect the diagnostic or screening capability of a test. Bland-Altman plots [40] were generated to evaluate the performance on the AHI estimation task.

B. OSA screening systems

1) Acoustic-based system: The proposed acoustic-based system employed Mel-filterbank features as described in Section III. To assess the importance of context in the OSA screening task, we experimented with different segment window sizes: 30-s and 40-s. The same 10-s segment shift was used in both experiments.

2) SpO2-based system: To provide a benchmark for the performance of the proposed acoustic-based approach, a system using blood oxygen saturation (SpO_2) was developed and evaluated. Oxygen saturation data from pulse oximetry alone can be used for screening for sleep apnea [41]. Here, the temporal changes of SpO_2 from the corresponding HSAT channel, measured as SpO_2 deltas, were used as input to a separate DNN that had a similar architecture to the one used by the acoustic-based system. Deltas (Δ) were computed as

$$\Delta x(n) = \frac{\sum_{t=1}^{T} t\,(x(n+t) - x(n-t))}{2\sum_{t=1}^{T} t^2} \tag{2}$$

where x is the SpO_2 data, and T is the window size. Here, $T = 51$ samples for a 199 ms window. These delta features approximate the first derivative of the SpO_2 signal, and provide information on its temporal dynamics.

3) Sample entropy-based system: The sample entropy approach proposed by Castillo et al. [28] is a state-of-the-art acoustic-based system for OSA screening which was implemented to provide a baseline for comparison. Castillo et al. also employed audio signals recorded with a smartphone during HSAT. However, they attached the smartphone to the participant's chest; in this study the smartphone was placed on a bed-side table. The sample entropy approach is a rule-based method that detects apnea/hypopnea events by extracting audio recording segments close to desaturations. Silent regions are found within those segments with a duration of >6 seconds using sample entropy. The AHI is directly calculated from the number of silent regions and the duration of the audio recording. We note that Castillo et al.'s approach is not purely based on acoustics, since it requires oxygen saturation data to identify desaturated segments.

V. Results and Discussion

Table IV lists the sensitivity, specificity and AUC of various OSA screening systems at different AHI cut-off points. The generally accepted AHI cut-off points (5, 15, and 30) are highlighted in bold, but for completeness the results for all cut-off points between 5 and 30 in steps of 5 are given. At each cut-off point, the number of nights in each class group is also shown.

The SpO_2-based system provides a benchmark performance for the corpus used in this study. Using SpO_2 data from HSAT, the system achieved an AUC of 0.88 at the AHI cut-off point of 15 (moderate OSA). At this threshold the two classes (below and above AHI 15) were fairly balanced, and the SpO_2-based system achieved similar sensitivity (0.81) and specificity (0.78). At the AHI cut-off point of 30 (severe OSA), all the metrics improved compared to those at the cut-off point of 15, with more improvement in sensitivity (0.91, 12% absolute improvement) than in specificity (0.82, 4% absolute improvement). The AUC increased to 0.93. This suggests that the SpO_2 data provided more reliable information for detecting severe OSA cases, with a higher proportion of true positives being identified. It is also noted that there were significantly fewer nights in the severe OSA class (39 nights) than those in the non-severe OSA class (118 nights), but the SpO_2-based system performed well in both sensitivity and specificity.

Compared to the SpO_2-based benchmark, the proposed acoustic-based OSA screening system performed well across different AHI cut-off thresholds. The best performing acoustic-based system was the one using 30-s long segments. In general, the performance of acoustic-based systems increased when a higher AHI cut-off was used, achieving an AUC of 0.92 at the AHI cut-off of 30, just slightly below the SpO_2 AUC of 0.93. The main contribution of the improvement was the specificity. In most cases, the specificity of the acoustic-based system was higher than that of the SpO_2-based system, especially when the AHI cut-off was higher than 15. Although the SpO_2-based system had high sensitivity, it also produced more false positives than the acoustic-based system. It is possible that the SpO_2-based system gives a false positive result for patients who have desaturations that are unrelated to sleep apnea: more HSAT sensors (such as nasal flow) would be needed to exclude these cases. The acoustic-based system was able to exploit breathing sounds in this case to achieve a higher specificity. Its sensitivity (0.78 at the AHI cut-off point of 30) was lower than the SpO_2 system (0.93), which reflects the challenge of using ambient sound recordings at home for the OSA screening task. It is likely that in some cases the audio was too quiet (patient facing away from the smartphone) or too noisy (TV on, traffic noise), causing the acoustic-based system to fail to detect sleep apnea events.

The AUC is not reported for the sample entropy-based system because it directly estimates the AHI. The use of acoustic and desaturation data by the sample entropy-based system resulted in an improved specificity in comparison to the system using acoustic data only, as the number of false positives was reduced by labelling segments with high oxygen saturation as non-apnea [28]. However, this strategy of using desaturation information also caused low sensitivity, since the number of true positives could also be reduced. The system failed to identify around half of the true positives at most of the AHI cut-off points. It is possible that some of the apnea segments, especially at the start of an apnea/hypopnea event, may not be accompanied by a desaturation in blood oxygen as apnea/hypopnea could have a delayed effect on desaturation. There were also differences in acoustic data collection between the study of Castillo et al. [28] and ours. Their rule-based silence detection method using sample entropy requires quiet breathing sounds to be recorded by a microphone close to the patient's chest. Here, audio recordings were collected using a smartphone placed on a bedside table. It is therefore likely that some breathing sounds were not audible in the recordings used in this study, causing the sample entropy method to misclassify them as silence.

This highlights a challenge in comparing our work to other state-of-the-art studies, since audio recordings used in

TABLE IV: OSA screening results using different AHI cut-off points. Generally accepted thresholds are listed in bold.

	AHI cut-off points c	**5**	10	**15**	20	25	**30**
	AHI $< c$ | AHI $\geq c$ (157 nights)	**13 | 144**	51 | 106	**80 | 77**	92 | 65	107 | 50	**118 | 39**
SpO_2-based	**Sensitivity**	**0.75**	0.76	**0.81**	0.86	0.85	**0.93**
	Specificity	**0.82**	0.67	**0.78**	0.82	0.78	**0.82**
	AUC	**0.89**	0.82	**0.88**	0.91	0.93	**0.93**
Acoustic-based 30-s segment	**Sensitivity**	**0.86**	0.79	**0.79**	0.79	0.79	**0.78**
	Specificity	**0.59**	0.74	**0.80**	0.91	0.90	**0.93**
	AUC	**0.73**	0.79	**0.84**	0.90	0.91	**0.92**
Acoustic-based 40-s segment	**Sensitivity**	**0.87**	0.73	**0.77**	0.79	0.77	**0.73**
	Specificity	**0.71**	0.70	**0.80**	0.86	0.86	**0.95**
	AUC	**0.75**	0.77	**0.81**	0.87	0.90	**0.91**
Sample entropy + desaturations	**Sensitivity**	**0.51**	0.39	**0.40**	0.39	0.48	**0.52**
	Specificity	**1.00**	1.00	**1.00**	1.00	1.00	**1.00**

previous studies were collected differently: e.g. using different types of microphones (high quality microphones [32], PSG-microphones [31], smartphone microphones [27], [28], [42]) and at different recording locations (tracheal sound recording [27], [29], [32], ceiling mount [31], chest [28], bed-side [42]). As the comparison with Castillo et al. [28] shows, the techniques employed often depend on assumptions made for the audio collection. Failing to meet the assumption could lead to degraded performance. Broadly, our approach compares favourably to previous approaches because it makes few assumptions about the audio recordings and our machine learning approach tends to be more robust than rule-based methods in an acoustically complex home environment.

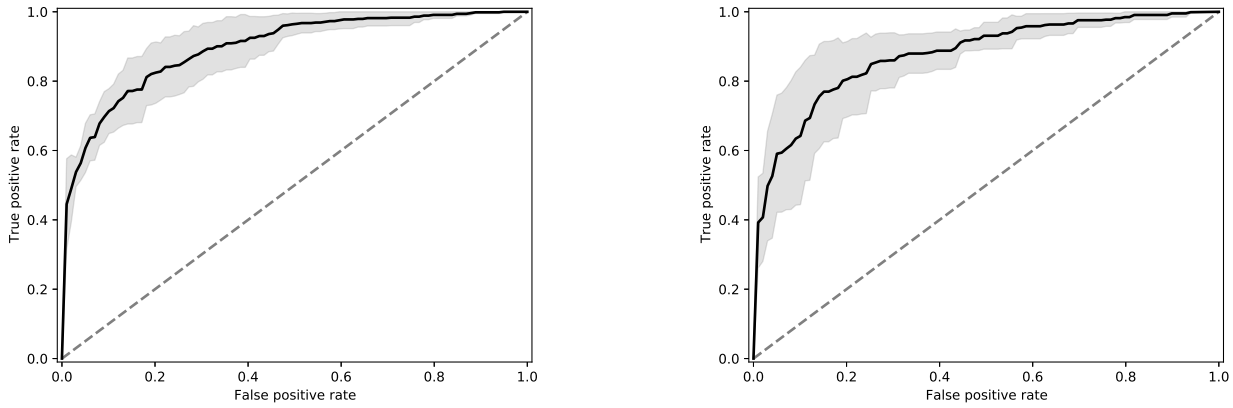
Fig. 4 shows the average ROC curves of the SpO_2 based system and the acoustic-based system using Mel-filterbank features with 30-s segments, at different AHI cut-off points considered for screening. A ROC curve is not shown for the sample entropy baseline because it directly computes the AHI. The grey area corresponds to the standard deviation, and the dashed line indicates random performance. All ROC curves were clearly above the random guess, which confirms that the network architecture and feature representations allowed the network to effectively learn from the data. Comparing the two average ROC curves, it can be seen that the SpO_2-based system yielded slightly better diagnostic capability than the acoustic-based system, and had a smaller standard deviation across different AHI cut-off points. Nonetheless the proposed acoustic-based OSA screening system exhibited a performance comparable to the SpO_2 benchmark.

Fig. 5 presents the reference AHI (solid line) for each night in the dataset computed from the reference HSAT scoring. The AHI estimated by the benchmark SpO_2-based system is shown as a dashed line with stars, and the AHI predicted by the acoustic-based system is shown as the dotted line with triangles. Both the benchmark SpO_2 predictions and the acoustic-based predictions show underestimations and overestimations across different nights, but overall it can be seen that the predictions matched the reference AHI patterns. These observations are consistent with the quantitative screening results previously discussed.

Fig. 6 plots the Bland-Altman analysis of the AHI estimation task using SpO_2 deltas, Mel-filterbank features, and sample entropy plus desaturations. The mean difference between the estimated and reference AHI is shown as a solid line, and the agreement limits (i.e., ± 1.96 times the standard deviation) are shown as dashed lines. 94% of the nights were within the agreement limits when using SpO_2 deltas, 95% when using Mel-filterbank features, and 93% when using sample entropy plus desaturations. Most of the nights outside the agreement limits are above 30 events/h (i.e., severe OSA), for which the screening systems tend to underestimate the AHI. Similar behaviour has been reported by related studies [32], [33], [43]. This is likely caused by the use of a large analysis window, since for participants with very high AHIs (e.g., 115 events/h), multiple apnea-hypopnea events could be present in a 30-second segment. However, any AHI above 30 events/h is considered severe.

The proposed acoustic-based system to screen for OSA relies only on breathing sounds recorded unobtrusively at home with a smartphone. This might improve accessibility to sleep-disordered breathing (SDB) diagnosis by allowing a screening service at home and better use of limited PSG resources (for example, reserving PSG to confirm diagnosis in complex cases). Also, as the screening system is inexpensive and easy to use, it can be used for long-term monitoring to investigate inter-night variability, assess OSA progression, evaluate treatment effectiveness and adherence.

Although the experiments reported here were performed on a computer, the proposed DNN was also deployed on smartphones to investigate the feasibility of on-device processing. 9.6 hours of data were classified in approximately 313 seconds on an Apple iPhone 6, in 100 seconds on an Apple iPhone 7 Plus, in 6 seconds on an Apple iPhone XS, and in 4 seconds on an Apple iPhone SE (2020) thanks to the small network size (about 6 million parameters), and hardware acceleration for machine learning in newer devices. This demonstrates that the whole screening system (i.e., audio recording, classification, results reporting, etc.) can be run in real-time on a modern smartphone. On-device machine learning [44] preserves user privacy, as the data strictly remains on the user's device. This might also contribute to reducing the psychological effects of being under observation during a sleep test [22].



(a) SpO_2-based system (b) Acoustic-based system (30-s Mel-filterbank)

Fig. 4: Average ROC curves across different AHI cut-off points

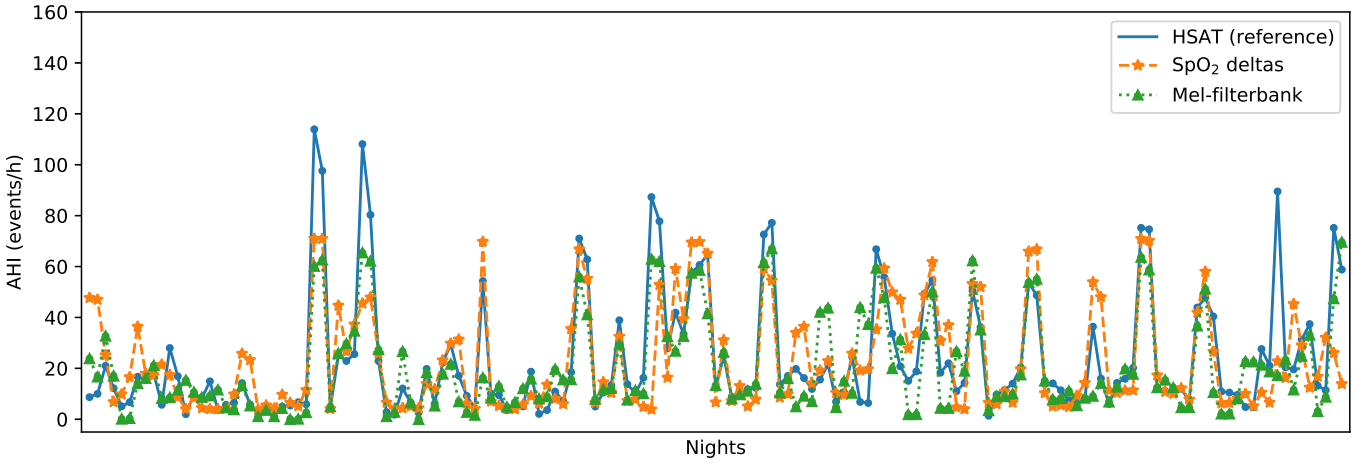
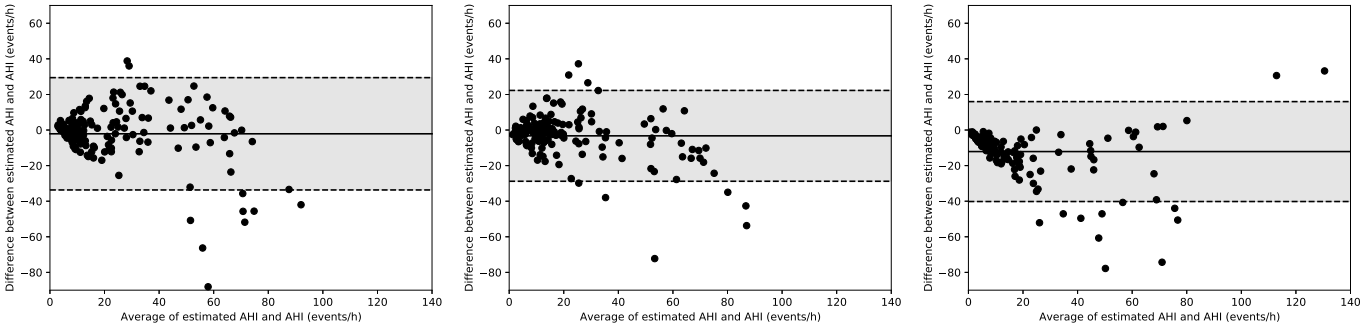


Fig. 5: AHI estimation for each night in the dataset using Mel-filterbank features. The reference AHI is calculated from the manually scored HSAT data.



(a) SpO_2-based (b) Acoustic-based (30-s Mel-filterbank) (c) Sample entropy and desaturation-based

Fig. 6: Bland-Altman plots for estimated AHI

VI. Conclusion

Screening for OSA based on ambient sound recordings made in a home environment is a challenging task. Two main problems have to be addressed: (1) OSA events are usually quiet or even silent, and (2) lack of physiological data. This paper has described a solution to these problems by using deep learning to exploit the temporal pattern within a long context window (30-s) of breathing sounds. The proposed solution allows a person to be screened for OSA at home using breathing sounds recorded with a smartphone. In this study, 103 participants (157 nights) were tested using HSAT while audio recordings were collected using a smartphone. The manually-scored HSAT data was used as reference for the AHI predictions made by the proposed acoustic-based system.

The results showed that 78% (sensitivity) of the participants with severe OSA and 93% (specificity) of the participants without severe OSA were correctly identified by the acoustic-based system. An AUC of 0.92 was achieved and 95% of the estimated AHIs were within the agreement limits in a Bland-Altman plot. This was compared to a benchmark system based on SpO2, which achieved a sensitivity of 0.93, a specificity of 0.82, and an AUC of 0.93. These results demonstrate the potential of the proposed screening system, and highlight the capability of smartphone-based approaches in healthcare.

This study has some limitations. First, the applicability of the proposed screening system to the general population remains to be investigated, since a limited amount of data obtained from 103 participants was used for training and testing. Second, data was collected in a country (UK) in which air conditioning is not routinely used. Therefore, the robustness of the screening system to this kind of background noise at home was not evaluated. Furthermore, the data was collected from participants sleeping on their own, and interference from a bed partner was not tested. Third, sleep stages/status were not considered here, as no EEG was used. As a result, the total recording time rather than total sleep time was used as a denominator for computing AHI, which is likely to systematically underscore the AHI [16]–[18].

In a world adapting to the lasting effects of the COVID-19 pandemic, there is a growing demand for novel contactless methods to screen for OSA [33], [45]. Practically, the pandemic is putting clinical respiratory services under pressure, meaning that investigations of OSA are not being prioritised. There is also considerable overlap between the risk factors for OSA and COVID-19 (e.g., age, male, asthma, diabetes, excess weight and hypertension). Although there is currently no evidence identifying OSA as an independent risk factor for COVID-19, it is possible that OSA exacerbates respiratory complications due to COVID-19 [46]. The wide availability of a low-cost screening tool, as proposed here, could therefore play an important role in determining the course of treatment after COVID-19 infection.

In the future, robustness to noise from a bed partner will be investigated, building upon our previous work on snore analysis [47] and snorer diarisation [48]. A possible approach would be to extend this work to detect other breathing events in addition to snores, for example breaths, and analyse the temporal pattern of each subject's breathing. This would most likely require collecting audio recordings with a smartphone while both the subject under study and their bed partner undergo HSAT, in order to have a reference for the breathing events of both.

VII. Acknowledgements

This research was funded by Innovate UK Open Grant 26767 and Innovate UK Knowledge Transfer Partnership (KTP) 10008165. HER was funded by a PhD studentship from Passion For Life Healthcare Limited (PFLH) and the University of Sheffield. We thank Iain Spray, Stelios Giannoulis and Amna Rahim from PFLH for collecting the OSA data, Philip Gillespie from PFLH for helpful suggestions, and Stevie Williams from the University of Edinburgh Sleep Research Unit / Western General Hospital for scoring the HSAT data.

References

[1] M. P. Mansukhani, B. P. Kolla, E. Kent, and T. I. Morgenthaler, *Conn's Current Therapy*. Philadelphia, United States: Elsevier, 2020, ch. Sleep Disorders, pp. 739–753.

[2] G. Lee, L. Lee, C. Wang, N. Chen, T. Fang, C. Huang, W. Cheng, and H. Li, "The frequency and energy of snoring sounds are associated with common carotid artery intima-media thickness in obstructive sleep apnea patients," *Nature Scientific Reports*, 2016.

[3] E. R. Cho, H. Kim, H. S. Seo, S. Suh, S. K. Lee, and C. Shin, "Obstructive sleep apnea as a risk factor for silent cerebral infarction," *Journal of Sleep Research*, vol. 22, pp. 452–458, 2013.

[4] C. Lal, C. Strange, and D. Bachman, "Neurocognitive impairment in obstructive sleep apnea," *Chest*, vol. 141, no. 6, pp. 1601–1610, June 2012.

[5] A. Castaneda, E. Jauregui-Maldonado, I. Ratnani, J. Varon, and S. Surani, "Correlation between metabolic syndrome and sleep apnea," *World Journal of Diabetes*, vol. 9, no. 4, pp. 66–71, April 2018.

[6] P. Lévy, M. Kohler, W. McNicholas, F. Barbé, and et al, "Obstructive sleep apnoea syndrome," *Nature Reviews Disease Primers*, vol. 1, no. 1, pp. 1–21, June 2015.

[7] M. R. Bonsignore, P. Baiamonte, E. Mazzuca, A. Castrogiovanni, and O. Marrone, "Obstructive sleep apnea and comorbidities: a dangerous liaison," *Multidisciplinary Respiratory Medicine*, vol. 14, no. 8, Feb 2019.

[8] A. V. Benjafield, N. T. Ayas, P. R. Eastwood, R. Heinzer, and et al, "Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis," *The Lancet*, vol. 7, no. 8, pp. 687–698, August 2019.

[9] K. K. Motamedi, A. C. McClary, and R. G. Amedee, "Obstructive sleep apnea: a growing problem," *Ochsner Journal*, vol. 9, pp. 149–153, 2009.

[10] T. Young, L. Finn, P. E. Peppard, M. Szklo-Coxe, D. Austin, F. J. Nieto, R. Stubbs, and K. M. Hla, "Sleep disordered breathing and mortality: Eighteen-year follow-up of the wisconsin sleep cohort," *Sleep*, vol. 31, no. 8, pp. 1071–1078, 2008.

[11] J. L. Hossain and C. M. Shapiro, "The prevalence, cost implications, and management of sleep disorders: an overview," *Sleep Breath*, vol. 6, pp. 85–102, 2002.

[12] R. Ferber, R. Millman, M. Coppola, J. Fleetham, and et al, "Portable recording in the assessment of obstructive sleep apnea," *Sleep*, vol. 17, no. 4, pp. 378–392, 06 1994.

[13] American Academy of Sleep Medicine, *AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications (Version 2.6)*, January 2020.

[14] N. Collop, S. Tracy, V. Kapur, R. Mehra, D. Kuhlmann, and et al, "Obstructive sleep apnea devices for out-of-center (OOC) testing: Technology evaluation," *Journal of Clinical Sleep Medicine*, vol. 7, no. 5, pp. 531–48, Oct 2011.

[15] J. Corral-Penafiel, J. Pepin, and F. Barbe, "Ambulatory monitoring in the diagnosis and management of obstructive sleep apnoea syndrome," *European Respiratory Review*, vol. 22, no. 129, pp. 312—24, Sep 2013.

[16] P. Escourrou, L. Grote, T. Penzel, W. Mcnicholas, and et al, "The diagnostic method has a strong influence on classification of obstructive sleep apnea," *Journal of sleep research*, vol. 24, no. 6, pp. 730—8, Dec 2015.

[17] M. Bianchi and B. Goparaju, "Potential underestimation of sleep apnea severity by at-home kits: Rescoring in-laboratory polysomnography without sleep staging," *Journal of clinical sleep medicine*, vol. 13, no. 4, pp. 551—5, Sep 2017.

[18] R. Malhotra, D. Kirsch, D. Kristo, E. Olson, and et al, "Polysomnography for obstructive sleep apnea should include arousal-based scoring: An american academy of sleep medicine position statement," *Journal of clinical sleep medicine*, vol. 14, no. 7, pp. 1245—7, Jul 2018.

[19] K. V., D. Auckley, S. Chowdhuri, and et al, "Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: An american academy of sleep medicine clinical practice guideline," *Journal of Clinical Sleep Medicine*, vol. 13, no. 3, pp. 479–504, Mar 2017.

[20] A. Malhotra, I. Ayappa, N. Ayas, and et al, "Metrics of sleep apnea severity: beyond the apnea-hypopnea index," *Sleep*, vol. 44, no. 7, Jul 2021.

[21] H. W. Agnew Jr., W. B. Webb, and R. L. Williams, "The first night effect: An EEG study of sleep," *Psychophysiology*, vol. 2, no. 3, pp. 263–266, 1966.

[22] W. Lu, J. Cantor, R. N. Aurora, M. Nguyen, T. Ashman, L. Spielman, A. Ambrose, J. W. Kerllman, and W. Gordon, "Variability of respiration and sleep during polysomnography in individuals with TBI," *Neurorehabilitation*, vol. 35, pp. 245–251, 2014.

[23] E. Sforza, F. Roche, C. Chapelle, and V. Pichot, "Internight variability of apnea-hypopnea index in obstructive sleep apnea using ambulatory polysomnography," *Frontiers in Physiology*, vol. 10, no. 849, pp. 1–8, July 2019.

[24] N. Ahmadi, G. K. Shapiro, S. A. Chung, and C. M. Shapiro, "Clinical diagnosis of sleep apnea based on single night of polysomnography vs two nights of polysomnography," *Sleep and Breathing*, vol. 13, pp. 221–226, 2009.

[25] A. Voulgaris, L. Ferini-Strambi, and P. Steiropoulos, "Sleep medicine and COVID-19. Has a new era begun?" *Sleep Medicine*, vol. 73, pp. 170–176, September 2020.

[26] S. Saha, M. Kabir, N. M. Ghahjaverestan, M. Hafezi, B. G. anf K. Zhu, H. Alshaer, and A. Yadollahi, "Portable diagnosis of sleep apnea with the validation of individual event detection," *Sleep Medicine*, vol. 69, pp. 51–57, 2020.

[27] M. Al-Mardini, F. Aloul, A. Sagahyroon, and L. Al-Husseini, "Classifying obstructive sleep apnea using smartphones," *Journal of Biomedical Informatics*, vol. 52, pp. 251–259, 2014.

[28] Y. Castillo-Escario, I. Ferrer-Lluis, J. M. Montserrat, and R. Jane, "Entropy Analysis of Acoustic Signals Recorded With a Smartphone for Detecting Apneas and Hypopneas: A Comparison with a Commercial System for Home Sleep Apnea Diagnosis," *IEEE Access*, vol. 7, pp. 128 224–128 241, 2019.

[29] A. Yadollahi and Z. Moussavi, "Acoustic obstructive sleep apnea detection," in *Proceedings of the 31st Annual International Conference of the IEEE EMBS*. IEEE, 2009, pp. 7110–7113.

[30] E. Goldshtein, A. Tarasiuk, and Y. Zigel, "Automatic detection of obstructive sleep apnea using speech signals," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 5, pp. 1373–1382, May 2011.

[31] T. Kim, J. W. Kim, and K. Lee, "Detection of sleep-disordered breathing severity using acoustic biomarker and machine learning techniques," *Biomedical Engineering Online*, vol. 17, 2018.

[32] H. Nakano, T. Furukawa, and T. Tanigawa, "Tracheal sound analysis using a deep neural network to detect sleep apnea," *Journal of Clinical Sleep Medicine*, vol. 15, no. 8, pp. 1125–1133, 2019.

[33] R. Tiron, G. Lyon, H. Kilroy, A. Osman, N. Kelly, N. O'Mahony, C. Lopes, S. Coffey, S. McMahon, M. Wren, K. Conway, N. Fox, J. Costello, R. Shouldice, K. Lederer, I. Fietze, and T. Penzel, "Screening for obstructive sleep apnea with novel hybrid acoustic smartphone app technology," *Journal of Thoracic Disease*, vol. 12, no. 8, pp. 4476–4495, 2020.

[34] SOMNOmedics, "SOMNOtouch RESP," https://somnomedics.de/en/solutions/sleep_diagnostics/polygraphy-devices/somnotouch-resp/, accessed: 2021-06-11.

[35] T. Penzel and A. Sabil, "The use of tracheal sounds for the diagnosis of sleep apnoea," *Breathe*, vol. 13, no. 2, pp. e37–e45, June 2017.

[36] Personal communication, SomnoMedics Ltd, 4th November 2021.

[37] D. O'Shaughnessy, *Speech communication: human and machine*. Addison-Wesley, 1987.

[38] S. Albawi, T. A. Mohammed, and S. Al-Azawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1–6.

[39] M. Abadi et al., "TensorFlow: a system for large-scale machine learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation*. USENIX, 2016, pp. 265–283.

[40] J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *The Lancet*, vol. 327, no. 8476, pp. 307–310, 1986.

[41] D. E. Jonas, H. R. Amick, C. Feltner, R. P. Weber, M. Arvanitis, A. Stine, L. Lux, and R. P. Harris, "Screening for Obstructive Sleep Apnea in Adults: Evidence Report and Systematic Review for the US Preventive Services Task Force," *JAMA*, vol. 317, no. 4, pp. 415–433, 01 2017. [Online]. Available: https://doi.org/10.1001/jama.2016.19635

[42] S. Narayan, P. Shivdare, T. Niranjan, K. Williams, J. Freudman, and R. Sehra, "Noncontact identification of sleep-disturbed breathing from smartphone-recorded sounds validated by polysomnography," *Sleep and Breathing*, 2018.

[43] A. Alakuijala and T. Salmi, "Predicting obstructive sleep apnea with periodic snoring sound recorded at home," *Journal of Clinical Sleep Medicine*, vol. 12, no. 7, pp. 953–958, 2016.

[44] G. A. Kaissis, M. R. Makowski, and D. Ruckert, "Secure, privacy-preserving and federated machine learning in medical imaging," *Nature Machine Intelligence*, vol. 2, pp. 305–311, June 2020.

[45] L. Williamson, "Obstructive sleep apnoea: latest surgical advances and considerations during the COVID-19 pandemic," *The Lancet Respiratory Medicine*, 2020.

[46] C. M. Suen, D. S. C. Hui, S. G. Memtsoudis, and F. Chung, "Obstructive sleep apnea, obesity, and noninvasive ventilation: Considerations during the coronavirus disease 2019 pandemic," *Anesthesia and Analgesia*, vol. 131, no. 2, pp. 318–322, Apr 2020.

[47] H. E. Romero, N. Ma, G. J. Brown, A. V. Beeston, and M. Hasan, "Deep learning features for robust detection of acoustic events in sleep-disordered breathing," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2019, pp. 810–814.

[48] H. E. Romero, N. Ma, and G. J. Brown, "Snorer diarisation based on deep neural network embeddings," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 876–880.

Hector E. Romero is a Bioengineer and Computer Scientist. He obtained his PhD at the University of Sheffield in 2022 investigating deep learning methods for the acoustic analysis of sleep-disordered breathing. Hector is currently a Machine Learning and Sensor Integration Engineer at the University of Sheffield and Passion for Life Healthcare.

Ning Ma is currently a Research Fellow at the University of Sheffield. He obtained a MSc (distinction) in Advanced Computer Science in 2003 and a PhD in speech and hearing technology in 2008, both from the University of Sheffield. He has been a Research Fellow at the MRC Institute of Hearing Research. His research interests include speech technology, machine learning in acoustics for healthcare applications, and hearing impairment. He has authored or coauthored over 50 papers in these areas.

Guy J. Brown received a BSc Applied Science from Sheffield City Polytechnic in 1984 and a PhD in Computer Science from the University of Sheffield in 1992. He is currently Professor of Computer Science, University of Sheffield. His research interests include computational auditory scene analysis, hearing impairment and acoustic monitoring for medical applications.

Elizabeth A. Hill is a Clinical Physiologist with over 20 years' experience in sleep medicine, holding the Registered Polysomnographic Technologist (RPSGT) & European Somnologist-Technologist (EST) credentials. She gained her PhD by research at The University of Edinburgh in 2016, examining prevalence and treatment of obstructive sleep apnoea in adults with Down syndrome. Lizzie is a Departmental Lecturer in Sleep Medicine on the Oxford Online Programme In Sleep Medicine, University of Oxford. She is also President of the European Society of Sleep Technologists & Education Lead of the British Sleep Society.