IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 19 February 2023, accepted 15 March 2023, date of publication 24 March 2023, date of current version 30 March 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3261557

RESEARCH ARTICLE

Accelerometer-Based Bed Occupancy Detection for Automatic, Non-Invasive Long-Term Cough Monitoring

MADHURANANDA PAHAR[1], IGOR MIRANDA[2], ANDREAS DIACON[3], AND THOMAS NIESLER[1]

[1]Department of Electrical and Electronic Engineering, Stellenbosch University, Stellenbosch 7602, South Africa
[2]Division of Computer Engineering, Federal University of Recôncavo da Bahia, Cruz das Almas 44380-000, Brazil
[3]TASK Applied Science, Cape Town 7530, South Africa

Corresponding author: Thomas Niesler (trn@sun.ac.za)

This work was supported by the Lungenliga St. Gallen-Appenzell, Switzerland.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Stellenbosch Health Research Ethics Committee (HREC) 1 with the project ID 1666 and the ethics reference number M17/10/035.

ABSTRACT We present a new machine learning based bed occupancy detection system that uses only the accelerometer signal captured by a bed-attached consumer smartphone. Automatic bed occupancy detection is necessary for automatic long-term cough monitoring since the time that the monitored patient occupies the bed is required to accurately calculate a cough rate. Accelerometer measurements are more cost-effective and less intrusive than alternatives such as video monitoring or pressure sensors. A 249-hour dataset of manually-labelled acceleration signals gathered from seven patients undergoing treatment for tuberculosis (TB) was compiled for experimentation. These signals are characterised by brief activity bursts interspersed with long periods of little or no activity, even when the bed is occupied. To process them effectively, we propose an architecture consisting of three interconnected components. An occupancy-change detector locates instances at which bed occupancy is likely to have changed, an occupancy-interval detector classifies periods between detected occupancy changes and an occupancy-state detector corrects falsely-identified occupancy changes. Using long short-term memory (LSTM) networks, this architecture achieved an AUC of 0.94. To demonstrate the application of this bed occupancy detection system to a complete cough monitoring system, the daily cough rates along with the corresponding laboratory indicators of a patient undergoing TB treatment were estimated over a period of 14 days. This provides a preliminary indication that automatic cough monitoring based on bed-mounted accelerometer measurements may present a non-invasive, non-intrusive and cost-effective means of monitoring the long-term recovery of patients suffering from respiratory diseases such as TB and COVID-19.

INDEX TERMS Accelerometer, bed occupancy, cough monitoring, long short-term memory (LSTM), machine learning.

I. INTRODUCTION

Coughing is a symptom of many lung diseases including tuberculosis (TB) and COVID-19. Long-term cough monitoring, over periods ranging from days to months, may allow the progression of such conditions to be tracked in a cost-effective and non-invasive manner [1]. For example, TB is commonly assessed by the analysis of sputum samples. This is a time-consuming and costly clinical practice that requires the engagement of trained medical personnel as well as specialised laboratory facilities for the analysis [2]. However, experimental evidence suggests that TB patients who are responding to treatment also exhibit a reduction in cough frequency [3]. Therefore, cough monitoring would offer the

The associate editor coordinating the review of this manuscript and approving it for publication was Santosh Kumar.

advantages that it requires neither a laboratory nor medical personnel and is cost-effective since it could be implemented on a smartphone or similar device.

Although our proposed system has wider applications, we consider the particular scenario of long-term cough monitoring in the ward environment of a TB research hospital. The objective is to provide an alternative means of monitoring the success of treatment received by the patients in this facility. Previously, we have developed a system that can accurately detect coughs from the signal obtained from the tri-axial accelerometer on-board a consumer smartphone, where the smartphone itself is attached to the bed-frame of the patient under treatment [4]. By not relying on audio signals, as many alternative approaches do, the system sidesteps the privacy concerns that accompany such audio-based classifiers [5]. Furthermore, since this is not a wearable sensor, it is less intrusive and more convenient. However, since monitoring must take place continuously and over extended periods, it is necessary to know the times when the patient occupied the bed in order to reliably estimate a cough rate [6]. In this work, we consider such bed occupancy detection using the same acceleration signals. We note that eventually both the cough detector and the bed occupancy detector can be implemented on the same smartphone as a single integrated cough monitoring system. To demonstrate this application, we have measured the daily cough rates of a patient who was undergoing TB treatment over a period of 14 consecutive days. We report that in this case a decreasing cough rate was observed while the patient responded positively to the TB treatment. This provides a preliminary indication that our proposed bed occupancy detection system is a promising means to enable automatic long-term cough monitoring.

The remainder of this paper is structured as follows. First, we provide some background on bed occupancy detection and the use of accelerometers for human activity monitoring in Section II. Next, we describe the compilation of the dataset we use to train and evaluate our algorithms in Section III and the features extracted from this data in Section IV. The accelerometer-based bed occupancy detection strategy is presented in detail at Section V along with the classification strategy in Section VI. The subsequent application to the monitoring of long-term cough rates is presented in Section VII. Experimental results are shown in Section VIII and discussed in Section IX. Finally, Section X concludes this study.

II. BACKGROUND

A. BED OCCUPANCY DETECTION

One approach to bed occupancy detection is by means of automated video analysis [7]. However, video surveillance is intrusive, raises privacy concerns [8] and our own experience has shown strong resistance from patients to this type of monitoring. Another common way of determining bed occupancy is by means of pressure sensors placed under the mattress [9]. While this is a very direct way of establishing bed occupancy, it requires specialised and costly equipment. Furthermore, such pressure sensors have sometimes been found to be sensitive to factors such as the type of mattress and the weight of the patient, leading to incorrect measurements [10]. To our knowledge, bed occupancy detection based on accelerometer signals has not been reported in the literature before.

B. ACCELEROMETER-BASED PATIENT MONITORING

Accelerometers are well established as wearable sensors for human physical activity [11], [12], [13], [14], [15], [16]. For example, accelerometer signals have been used to successfully classify human movement such as walking, running, sitting, standing and climbing the stairs with sensitivities and specificities above 95% [17], [18], even at sampling rates as low as 5 Hz [19]. Similar success has been achieved when using accelerometer signals to distinguish between different walking styles [20], [21], [22] and for human fall detection [23].

Recently, wearable consumer devices such as smartphones with on-board tri-axial accelerometers have been used to classify human activity [24], [25], [26], [27], [28], [29] and vehicular motion [30]. While initial studies used simple classifiers such as logistic regression (LR) and multi-layer perceptrons (MLPs), more recently deep neural networks (DNNs) such as convolutional neural networks (CNNs) have offered better performance in recognising human activities from the wearable sensor data [31], [32]. Among deep approaches, CNNs have been shown to be well-suited to real-time human activity recognition using accelerometer measurements by offering the computational efficiency required by mobile platforms [33], [34], [35].

III. DATA

We use a manually annotated dataset of continuous accelerometer measurements obtained from seven patients to train and evaluate the classifiers used to detect bed occupancy. All data for this study has been collected at a TB research hospital in Cape Town, South Africa (TASK Clinical Trials Centre). This research hospital accommodates approximately 10 staff and up to 24 patients in six wards, each containing up to four beds. Typically, patients spend between 5 and 15 days at the research hospital, during which time they undergo treatment and are monitored. All patients for whom data was collected are adult males. No other patient information was gathered due to the ethical constraints of this study. The study was reviewed and approved by the Stellenbosch Health Research Ethics Committee (HREC) 1 with the project ID 1666 and the ethics reference number M17/10/035.

A. RECORDING SETUP

The data collection process is shown in Figure 1. Acceleration signals were captured by the on-board tri-axial accelerometer of a consumer smartphone (Samsung Galaxy J4). The smartphone was placed inside a rectangular plastic enclosure that had been firmly attached to the back of the headboard

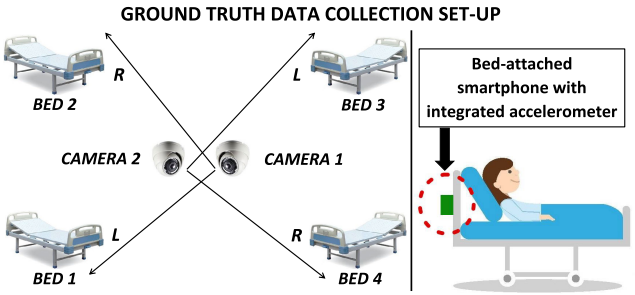


FIGURE 1. Data collection set-up. A plastic enclosure housing an inexpensive consumer smartphone (Samsung Galaxy J4) is firmly attached to the back of the headboard of each bed and an Android data recording application continuously monitors the accelerometer signal. To detect the bed occupancy, two ceiling-mounted cameras were used. The position of the beds, cameras and camera view angles are shown. Camera 1 monitors Beds 1 and 2, while Camera 2 monitors Beds 3 and 4.

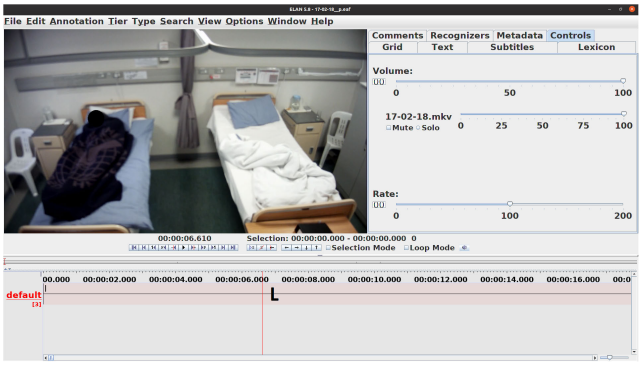


FIGURE 2. Annotation in ELAN. The bed on the left is occupied, hence the annotation label 'L'.

of each of the four beds in a ward. Within the enclosure, the smartphone was orientated horizontally on its side, with its back to the headboard [36]. Data capture software implemented on this device continuously monitored the accelerometer signals at a sampling frequency of 100 Hz. To reduce the volume of data captured, a simple energy threshold detector was implemented to exclude long periods with no measured acceleration. To further reduce the volume of data and also to remove dependence on the orientation of the smartphone, only the vector magnitude of the three tri-axial acceleration components was recorded, as indicated in Equation 1. Here $a_x(t)$, $a_y(t)$ and $a_z(t)$ are the measured tri-axial accelerations for a particular patient and t is the time from the start of monitoring. The total period of monitoring varied between 19 and 65 hours for the seven patients in our dataset. Finally, the respective acceleration signal $a(t)$ was normalised so that its maximum value corresponded to a value of 1 after completion of the recording for a patient.

$$a(t) = \sqrt{a_x^2(t) + a_y^2(t) + a_z^2(t)} \tag{1}$$

In addition to the acceleration signals captured by the smartphone, continuous simultaneous video recordings were made using two ceiling-mounted cameras, as shown in Figure 1. These video recordings were used only for the manual annotation of the captured acceleration signals and hence to provide accurate ground truth labels in terms of when each monitored bed was occupied.

B. DATA ANNOTATION

The accelerometer magnitude signals $a(t)$ were annotated using the ELAN multimedia software (shown in Figure 2), since this allowed consolidation with the video data from the ceiling-mounted cameras [37]. In this way the periods during which each bed was occupied could be accurately labelled, providing ground truth for our experiments. As every camera can view only two beds, the labels indicate whether the left bed (L), the right bed (R) or both beds (B) are occupied.

For a particular patient and therefore a particular acceleration signal $a(t)$, this ground truth annotation is represented by the signal $n(t)$, shown in Figure 3.

$$n(t) = \begin{cases} 1 & \text{if bed is occupied at time } t, \\ 0 & \text{if bed is not occupied at time } t. \end{cases} \tag{2}$$

We further note that, for all our patients, the bed is initially empty and hence $n(0) = 0$. Furthermore, $n(t)$ switches from 0 to 1 at the instant the patient begins to get into the bed, and switches from 1 to 0 at the instant the patient has completely left the bed. These conventions allowed the accurate determination of these bed occupancy changes from the joint inspection of the accelerometer and the video footage during the manual annotation process.

C. COMPILED DATASET

Data was captured from a total of seven patients, each of whom was continuously monitored for a period of between one and three days, as summarised in Table 1. In total, 249 hours of acceleration and video data were collected. All of this data was annotated as described in the previous section. Table 1 shows that patients occupied their beds for only 95.5 hours of the 249 hours of collected data. Furthermore, a total of only 104 occupancy changes (either getting into or getting out of bed) were identified. Thus, despite the many hours in our dataset, it remains very sparse in terms of bed occupancy changes. A sample accelerometer measurements and annotated data indicating a patient's bed occupancy pattern are shown in Figure 3.

IV. FEATURE EXTRACTION

The accelerometer magnitude signal $a(t)$ is split into overlapping frames, and features are extracted from each frame. The frame length (Ψ) and the number of frames (C) are hyperparameters that are optimised by varying the length of the frame skip. We calculate the frame skips by dividing the number of samples by the number of frames and taking the next positive integer. Extracted features include the power spectra, root mean square (RMS), moving average, kurtosis and crest factor [madhu: (Figure 4)]; as they have shown promising results in our previous studies using accelerometer signals [4], [38]. The power spectra [39] are a common

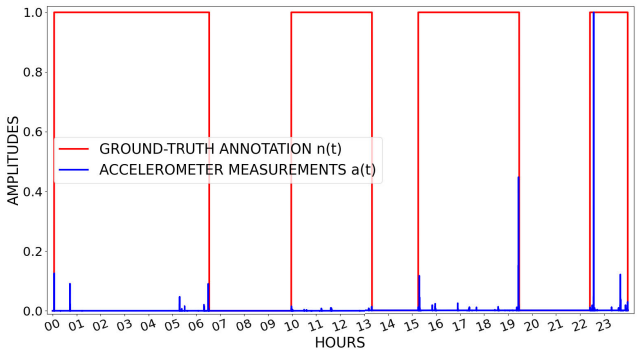


FIGURE 3. Example accelerometer signal $a(t)$ and ground-truth annotation $n(t)$ over a 24-hour period. The figure shows that the patient is often not very active while in the bed, indicated by very low almost-flat accelerometer signal amplitudes.

TABLE 1. Dataset summary. The number of continuous hours of data that were captured and annotated is indicated, as well as the portion of this period during which the bed was occupied. Occupancy changes refer to the total number of times the patient left or entered the bed during the observation period.

Patient	Total hours	Occupied hours	Occupancy changes
1	65	23.33	18
2	57	14.47	20
3	45	11.67	18
4	21	9.02	10
5	21	15.89	12
6	19	9.42	14
7	21	11.71	12
Total	**249**	**95.51**	**104**

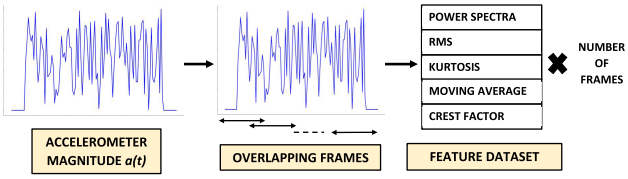


FIGURE 4. Feature extraction process: Power spectra, RMS, kurtosis, moving average and crest factor are extracted from the accelerometer measurements $a(t)$. The number of frames (C) is a feature extraction hyperparameter listed in Table 2.

feature to extract from sensor magnitudes [40], [41], [42], [43] as an input to the neural networks [44]. RMS [45], [46], and moving average [47] features are also well established in sensor signal analysis and can capture broad dynamic behaviour in the signal. Moving average is a simple way to measure the movement of a signal over a period and a useful feature for sensor analysis [47]. The kurtosis indicates the tailedness of a probability density and therefore is an indicator of the prevalence of higher amplitudes in the signal, while the crest factor measures the ratio between the peak and the RMS signal amplitude. Both have been found to be useful for machine learning applications [48], [49], [50].

For frames with Ψ samples, the power spectra have $(\frac{\Psi}{2}+1)$ coefficients, while RMS, moving average, kurtosis and crest factors are scalar. Hence, a $(\frac{\Psi}{2}+5)$ dimensional feature vector is extracted from each frame. According to Table 2, we consider frame lengths (Ψ) of 2^5 i.e. 32 and 2^6 i.e. 64 samples, corresponding to 320 and 640 millisecond intervals. These frames are shorter than those commonly used to extract features from audio for training and evaluating machine learning classifiers [51]. This is because the accelerometer integrated into the smartphone has a lower sampling rate (in our case 100 Hz). Using longer frames was observed to lead to deteriorated performance because the acceleration signal can then no longer be assumed to be approximately stationary.

Finally, we note that classification will generally consider a sequence of feature vectors extracted from successive frames, rather than individual feature vectors. Thus features will be arranged as a feature matrix of size $(C, \frac{\Psi}{2}+5)$, with the features themselves along one dimension and the frames along the other. We consider $C = 20$, 50 and 100, as listed in Table 2.

V. BED OCCUPANCY DETECTION STRATEGY

Bed occupancy detection based on bed-mounted accelerometer signals is challenging because these signals are characterised by bursts of activity (as the patient either moves or enters or leaves the bed) separated by often very long intervals of inactivity (Figure 3). These characteristics make, for example, the direct application of recurrent neural networks ineffective. In our preliminary experiments, we found that even structures designed specifically with long-term memory in mind, such as long short-term memory (LSTM) networks, are not able to learn effectively when presented with signals such as these.

Therefore, we have designed a system that incorporates three interconnected detectors. The first is trained to recognise specific portions of the acceleration signal that may be associated with a bed occupancy change, i.e. the short time interval during which a patient enters or leaves the bed. The second detector is trained specifically to determine whether the interval between two such occupancy changes corresponds to a period during which the bed is occupied or not. Finally, the third detector is presented with the outputs of the first two detectors and is designed to correct falsely detected bed occupancy changes. Figure 5 illustrates the overall structure of our proposed bed occupancy detector.

A. OCCUPANCY-CHANGE DETECTOR

The bed occupancy-change detector (Detector 1) classifies a five-second interval ($T_{oc} = 5$) as either containing an occupancy change or not. The acceleration signal $a(t)$ is divided into successive overlapping five-second (500 samples) frames, advancing in steps of 500 ms (50 samples). From each of these five-second frames, we extract C_{OC} sub-frames, each containing Ψ_{OC} samples. The number of samples by which these sub-frames overlap is determined to ensure that the C_{OC} sub-frames are extracted evenly across the full five-second interval. For each sub-frame, a feature vector is extracted, and these vectors are arranged as a

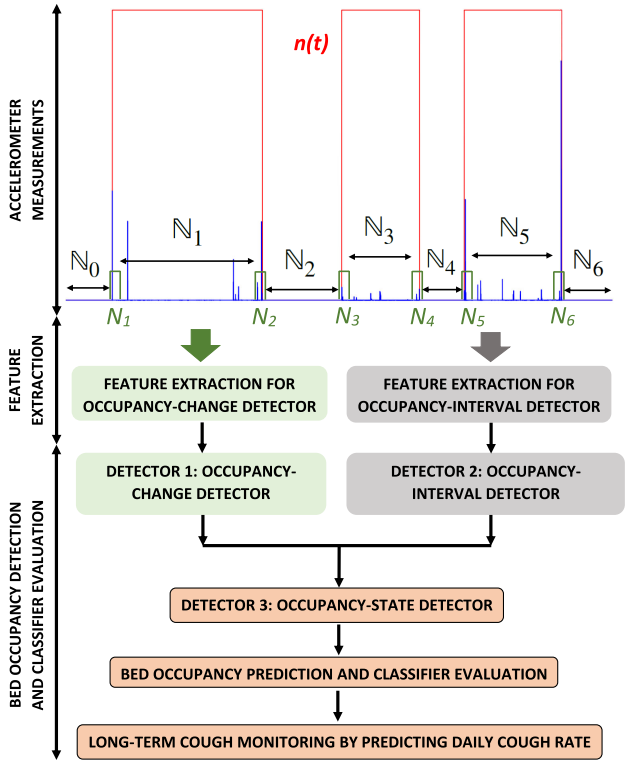


FIGURE 5. Bed Occupancy Detection Process: Features are extracted from the accelerometer signals to detect occupancy changes and occupancy intervals. Using these features, Detector 1 attempts to detect occupancy changes while occupancy intervals are classified by Detector 2. Detector 1 is seen to exhibit low specificity but high sensitivity, while Detector 2 displays high specificity but low sensitivity. Thus, Detector 3, the occupancy-state detector, attempts to correct falsely-detected occupancy changes based on the decisions by Detectors 1 and 2, as shown in Figure 6. In this particular case, there are $\mathcal{K} = 6$ occupancy changes, as the patient enters the bed and leaves the bed three times. Hence, the number of intervals is $\mathcal{K} + 1 = 7$.

feature matrix. Early experimentation showed that a five-second interval exhibits good performance, and this quantity was not optimised further. However the sub-frame length Ψ_{OC} and sub-frame skip (indirectly determined by C_{OC}), both of which influence the dimensionality of this feature matrix, are hyperparameters that are optimised.

We consider the instant τ_k at which the k^{th} occupancy change occurs to be the time at which the annotation signal $n(t)$ switches for the k^{th} time, where $k = 1, 2, 3 \ldots \mathcal{K}$ and where $\mathcal{K} = 104$ is the total number of occupancy changes in our dataset, as listed in Table 1. Since a bed is always empty at $t = 0$, τ_1 always indicates a transition from an empty to an occupied bed. Since in our data the bed is also empty at the end of the recording period, $\mathcal{K}$ is always an even integer.

Inspection of our data confirmed that the location of the instant τ_k within the five-second interval presented to the classifier should depend on whether the patient is entering or leaving the bed. When the patient enters the bed, there is almost no activity before τ_k, while when the patient leaves the bed, there is almost no activity after τ_k. Hence, during classifier training, the five-second interval N_k surrounding the time instant τ_k that the classifier is presented with to make its decision is specified differently for these two cases, as shown in Equation 3.

$$N_k = \begin{cases} t : \tau_k - \frac{T_{OC}}{5} \leq t \leq \tau_k + \frac{4T_{OC}}{5} & \text{for } k \text{ odd} \\ t : \tau_k - \frac{4T_{OC}}{5} \leq t \leq \tau_k + \frac{T_{OC}}{5} & \text{for } k \text{ even} \end{cases} \tag{3}$$

Here, the N_k are the time intervals from 1 sec before to 4 sec after the instants τ_k at which the bed occupancy changes from unoccupied to occupied and the time intervals from 4 sec before to 1 sec after the instants τ_k at which the bed occupancy changes from occupied to unoccupied. We will refer to the N_k as "occupancy changes" to distinguish them from the "occupancy intervals" considered in the next section.

Since most of the acceleration signal is not associated with an occupancy change, the data is highly unbalanced in terms of the two classification classes. In fact, Table 1 shows that there are only 104 occupancy changes. For our five-second analysis intervals, this corresponds to a total of only 8.67 minutes of the 249-hour dataset. Since such imbalance can affect machine learning detrimentally [52], [53], we have applied the synthetic minority over-sampling technique (SMOTE) to balance data during training [54], [55], [56]. This technique oversamples the minor class by generating synthetic samples, as an alternative to for example random oversampling. We have also implemented other extensions of SMOTE such as borderline-SMOTE [57], [58], [59] and adaptive synthetic sampling [60]. However, the best results were obtained by using SMOTE without any modification.

At classification time, the measured acceleration magnitude $a(t)$ is presented to the occupancy-change detector, which provides a sequence of hypothesised time instants $\hat{\tau}_k$ and associated intervals $\hat{N}_k$ at which the occupancy of the bed is likely to have changed. The early experimental evaluation revealed that this approach allows occupancy-change intervals to be identified with high sensitivity (above 99%) but low specificity. This means that, although most occupancy changes are predicted correctly, other activities, such as the movement of the patient while in the bed, have also (wrongly) been classified as occupancy changes.

B. OCCUPANCY-INTERVAL DETECTOR

We now consider the time interval between two consecutive occupancy changes and focus on the task of determining whether this interval is associated with the bed being occupied or empty. We will refer to these as "occupancy intervals" ($\mathbb{N}_k$) to distinguish them from the "occupancy changes" (N_k) considered in the previous section. The occupancy interval $\mathbb{N}_k$ between occupancy changes N_k and N_{k+1} is defined by Equation 4.

$$\mathbb{N}_k = \{t : N_k < t < N_{k+1}\} \tag{4}$$

where $k = 0, 1, 2, 3 \ldots \mathcal{K}$ and where $\mathbb{N}_0$ and $\mathbb{N}_\mathcal{K}$ indicate the start and the end of the signal respectively.

We note that since there are $\mathcal{K}$ occupancy changes, there are $\mathcal{K}+1$ occupancy intervals. Furthermore, $\mathbb{N}_0$ indicates the time interval before the first occupancy change and in our data always indicates an initial interval during which the bed is unoccupied, while $\mathbb{N}_{\mathcal{K}}$ is the interval following the last occupancy change, during which the bed is also empty.

For classification, we divide each occupancy interval $\mathbb{N}_k$ into ten-second (1000 sample) non-overlapping frames. From each of these ten-second frames, we extract C_{OI} sub-frames, each containing Ψ_{OI} samples. The number of samples by which these sub-frames overlap is determined to ensure that the sub-frames are extracted evenly across the full ten-second interval. For each sub-frame, a feature vector is extracted, and these vectors are arranged as a feature matrix. Finally, the feature matrices extracted from each ten-second frame in the occupancy interval $\mathbb{N}_k$ are averaged. In this way, information about the entire occupancy interval is encoded as a fixed-dimension feature matrix. This method of feature extraction is the result of extensive experimentation. In particular, approaches such as the direct application of LSTMs to features extracted from $\mathbb{N}_k$ in a more conventional way were, for example, not effective.

According to Table 1, there is a total of 104 occupancy changes, thus there are 105 occupancy intervals. This means that we generate a dataset containing only 105 feature matrices. This number is small, especially with a view to training DNN classifiers [61]. Thus, in order to provide justification for deeper architectures, two shallow classifiers (logistic regression and multilayer perceptron) were evaluated in addition to the CNN and LSTM for the occupancy-interval detector.

At classification time, the measured acceleration signal $a(t)$ as well as the hypothesised occupancy intervals $\hat{\mathbb{N}}_k$, which are derived directly from the hypothesised occupancy-change intervals $\hat{N}_k$ provided by Detector 1, are presented to the occupancy-interval detector, which in turn provides a classification decision $f_{OI}(\hat{\mathbb{N}}_k)$ for each occupancy interval $\hat{\mathbb{N}}_k$ as shown in Equation 5.

$$f_{OI}\left(\hat{\mathbb{N}}_k\right) = \begin{cases} 1 & \text{when occupied} \\ 0 & \text{when unoccupied} \end{cases} \tag{5}$$

Early experimental evaluation showed that this approach allows occupancy intervals to be identified with high specificity (above 99%) but lower sensitivity. Periods during which the bed is unoccupied are reliably identified, but for some patients, who are very quiet when sleeping, an occupied bed was identified as an empty bed.

The misclassifications by Detector 1 and Detector 2 are addressed by a third and final component, the occupancy-state detector.

C. OCCUPANCY-STATE DETECTOR

As pointed out in the previous two sections, the occupancy-change detector exhibits high sensitivity but low specificity while the occupancy-interval detector has high specificity but low sensitivity. A third detector, the occupancy-state detector, corrects some of the resulting errors by considering the outputs of both the occupancy-change and the occupancy-interval detectors. The process begins by viewing each occupancy change $\hat{N}_k$ hypothesised by Detector 1 as a true occupancy change, and labelling all occupancy intervals accordingly. This results in a labelling sequence for the occupancy intervals that alternate at every hypothesised occupancy change and where the first interval $\hat{\mathbb{N}}_0$ is always assumed to be unoccupied. Now the hypothesised changes are considered in turn, beginning with the first. In each case, if the occupancy intervals to the left and right of a hypothesised change are the same, indicating a false positive by Detector 1, the hypothesised occupancy change is discounted. This process is illustrated in Figure 6, where a false positive by Detector 1 at $\hat{N}_2$ is removed, thereby associating the accelerometer activity in this interval with patient movement rather than an occupancy change.

Hence, Detector 3 is based on a fixed decision rule. Experimentation indicated that it is so far not possible to improve on this simple rule by for example making Detector 3 a trainable architecture. This may be due to the very small training set for this detector, which consists of only 104 hypothesised occupancy changes and 105 occupancy intervals.

VI. CLASSIFICATION PROCESS

First, the accelerometer magnitude signal $a(t)$ is divided into successive five-second frames overlapping by 1 second and from which feature matrices are extracted, as described in Section V-A. Detector 1, the occupancy-change detector, classifies each such 5-second frame, and results in a set of time instants $\hat{\tau}_k$ and associated intervals $\hat{N}_k$ at which occupancy changes are likely to have occurred (Figure 6). From these intervals $\hat{N}_k$, the corresponding occupancy intervals $\hat{\mathbb{N}}_k$ are determined, and these are used to extract a new set of feature matrices as described in Section V-B. These are presented to the occupancy-interval detector, which labels each interval $\hat{\mathbb{N}}_k$ as either occupied or unoccupied. Finally, these labels for the $\hat{\mathbb{N}}_k$, together with the intervals $\hat{N}_k$ hypothesised by Detector 1, are processed by Detector 3, resulting in a final set of hypothesised occupancy instants $\tilde{\tau}_k$, which allow a bed occupancy signal $\hat{n}(t)$ to be determined. These can be compared with the ground truth annotation signal $n(t)$ in order to determine performance, as described in Section VI-C.

A. CLASSIFIER ARCHITECTURES

We have considered four classifier types to implement the bed occupancy detection strategy described in the previous section. These are logistic regression (LR), multilayer perceptron (MLP), convolutional neural network (CNN) and long short-term memory (LSTM) architectures. Of these, the first two are shallow architectures and have been used to classify occupancy intervals since in this case, the training dataset was especially small. For the other classification tasks, the training datasets were larger and thus deep architectures could be applied [62], [63].

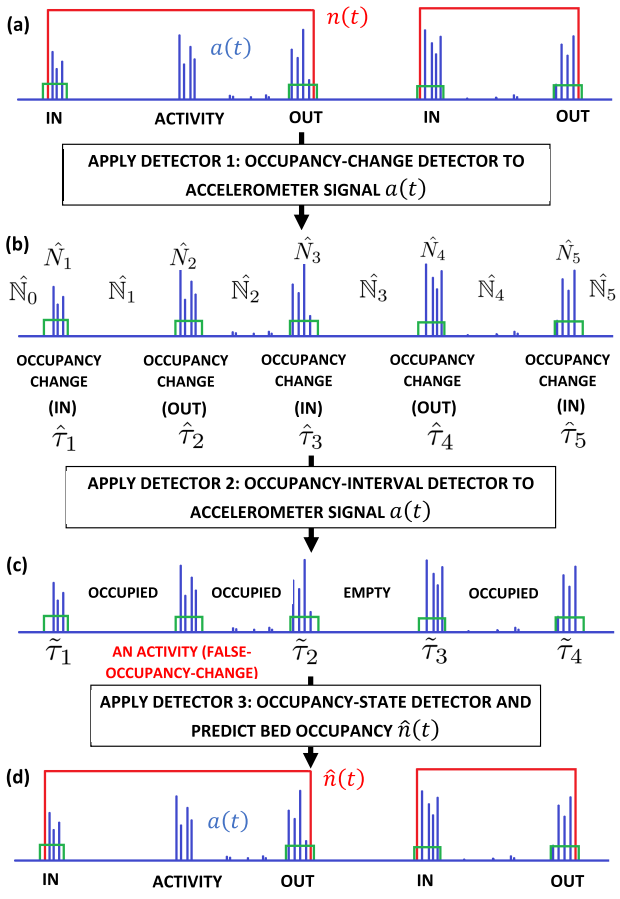


FIGURE 6. Bed Occupancy Detection Process. (a) The measured acceleration signal $a(t)$ and its ground-truth annotation $n(t)$. (b) The occupancy changes N_i are hypothesised by the occupancy-change detector (Detector 1) and labelled as 'in' and 'out' successively, while the occupancy-interval detector (Detector 2) decides whether each interval $\mathbb{N}_i$ between these occupancy changes is classified as 'occupied' or 'unoccupied'. In this illustration, $i = 1, 2, \ldots, 5$. (c) The occupancy-state detector (Detector 3) considers the decisions by Detectors 1 and 2 and reclassifies some occupancy changes asserted by Detector 1 from 'change' to 'activity', meaning the accelerometer signal was associated with patient movement in the bed and not with the patient entering or leaving the bed. (d) This automatically-determined bed occupancy $\hat{n}(t)$ is compared with the hand-annotated bed occupancy $n(t)$ in order to generate the results shown in Table 6.

LR models have been found to outperform other state-of-the-art classifiers such as classification trees, random forests, artificial neural networks and support vector machines in some clinical prediction tasks [62], [64], [65], [66]. We have used gradient descent weight regularisation as well as lasso ($l1$ penalty) and ridge ($l2$ penalty) estimators during training [67], [68]. The regularisation hyperparameters are optimised during cross-validation, as described in Section VI-B.

MLP models [69] are capable of learning non-linear relationships and have for example been shown to be effective when discriminating influenza coughs from other coughs [70]. MLPs have also been applied to tuberculosis coughs [62], [71], COVID-19 coughs [63] and to cough detection in general [72], [73], [74]. During training, we have applied stochastic gradient descent with the inclusion of an $l2$ penalty. This penalty, along with the number of hidden layers have been considered as hyperparameters.

CNNs are a deep neural network architecture which is popular especially in image classification [75]. CNNs have successfully solved complex tasks such as cough classification [4], [76], [77], [78], cough detection [36], wake-word detection [74], face recognition etc [79]. Our CNN architecture [80], [81] consists of α_1 2D convolutional layers with kernel size α_2, rectified linear unit (RELU) activation functions, a dropout rate α_3 and max-pooling. The convolutional layers are followed by two dense layers with α_4 and 8 units respectively, also using RELU activation functions. Finally, the network is terminated by a two-dimensional softmax layer, indicating occupancy changes for Detector 1 and bed occupancy for Detector 2. During training, feature matrices are presented to the classifier in batches of ξ_1 for a total of ξ_2 epochs.

LSTMs are a recurrent neural network (RNN) architecture which has been found to be effective for the modelling of sequential patterns in data. Originally suggested in [82], LSTMs have been successfully used in automatic cough detection and classification in our previous studies [4], [36], [63], [78], [83], [84], as well as acoustic feature detection in general [43], [85], [86]. Our LSTM architecture [87] uses β_1 units with RELU activation functions and a dropout rate α_3. The LSTM layer is followed by two dense layers with α_4 and 8 units, also using RELU activation functions. Finally, the network is terminated by a two-dimensional softmax layer, indicating occupancy changes for Detector 1 and occupancy intervals for Detector 2. During training, feature matrices are presented to the classifier in batches of ξ_1 for a total of ξ_2 epochs, where the feature vectors comprising each feature matrix are presented to the RNN in sequence.

The hyperparameters associated with these classifiers, as listed in Table 3, were optimised using a leave-one-out cross-validation scheme.

B. TRAINING AND HYPERPARAMETER OPTIMISATION

As our dataset contains only seven patients (Table 1), a leave-one-patient-out cross-validation scheme [88], [89] has been used to train and evaluate all classifiers. In this scheme, one patient is held out as a test patient in an outer loop. Among the remaining six patients, five are used in an inner loop to train the classifier while the sixth is used as a development set to optimise the hyperparameters listed in Table 3. The optimised hyperparameters are used to train the classifier using all six patients in the outer loop which is then evaluated on the seventh patient, which had been set aside for independent testing. This procedure is repeated so that all seven patients are used as an independent test set in turn. The area under the receiver operating characteristic (AUC) is used to optimise the hyperparameters. Final classifier performance is determined as the average sensitivity, specificity, accuracy and AUC over the seven outer-loop test sets.

TABLE 2. Feature extraction hyperparameters used for both occupancy change and occupancy interval detection. Frame lengths are varied between 320 and 640 milliseconds and the number of frames is varied between 20 and 100.

Hyperparameter	Description	Range
Ψ_{OC}	Number of samples per frame used by occupancy-change detector (Detector 1)	2^k where $k = 5, 6$
C_{OC}	Number of frames in feature matrix for occupancy-change detector (Detector 1)	20, 50
Ψ_{OI}	Number of samples per frame used by occupancy-interval detector (Detector 2)	2^k where $k = 5, 6$
C_{OI}	Number of frames in feature matrix for occupancy-interval detector (Detector 2)	50, 100

TABLE 3. Classifier hyperparameters optimised using the leave-one-out cross-validation (Section VI-B).

Hyperparameter	Description	Classifier	Range
ν_1	Regularisation strength of penalty ratios	LR	10^i where $i = -7, -6, \ldots, 6, 7$ (10^{-7} to 10^7)
ν_2	$l1$ penalty ratio	LR	0 to 1 in steps of 0.05
ν_3	$l2$ penalty ratio	LR	0 to 1 in steps of 0.05
η_1	No. of hidden layers	MLP	10 to 100 in steps of 10
η_2	$l2$ penalty ratio	MLP	10^i where $i = -7, -6, \ldots, 6, 7$ (10^{-7} to 10^7)
η_3	Learning rate	MLP	0 to 1 in steps of 0.05
ξ_1	Batch size	CNN, LSTM	2^k where $k = 6, 7, 8$
ξ_2	No. of epochs	CNN, LSTM	10 to 200 in steps of 20
α_1	No. of conv filters	CNN	3×2^k where $k = 3, 4, 5$
α_2	Kernel size	CNN	2 and 3
α_3	Dropout rate	CNN, LSTM	0.1 to 0.5 in steps of 0.2
α_4	Dense layer size	CNN, LSTM	2^k where $k = 4, 5$
β_1	LSTM units	LSTM	2^k where $k = 6, 7, 8$
β_2	Learning rate	LSTM	10^k where $k = -2, -3, -4$

The combination of hyperparameters for which the highest AUC has been obtained inside an inner loop, is defined as the 'best hyperparameters' in Table 4 and 5.

C. EVALUATION PROCESS

We evaluate the performance achieved by our bed occupancy detection system by comparing the true bed occupancy $n(t)$ with the predicted bed occupancy $\hat{n}(t)$ on a per-sample basis (Figure 6). The actual bed occupancy has been confirmed by manual annotation, as described in Section III-B, while $\hat{n}(t)$ is determined as described at the start of this section. Since the dataset is imbalanced in terms of the two classes (bed occupied and bed unoccupied), we evaluate classifier performance primarily in terms of the area under the curve (AUC) of the receiver operating characteristic (ROC). However, we also specify specificity, sensitivity and accuracy in our results.

In addition to evaluating the overall performance of the bed occupancy detection system, the individual performance of the occupancy-change detector (Detector 1) and the occupancy-interval detector (Detector 2) are also evaluated separately using the same method and performance indicators.

VII. LONG-TERM COUGH MONITORING

We now describe how the bed occupancy detection system (developed in this study) and the cough detection system (developed in our previous work [4]) can be integrated to allow the long-term cough monitoring of a patient who was undergoing TB treatment over a period of 14 days. The bed-mounted accelerometer signal was recorded for this patient, who was not part of the dataset compiled in Section III.

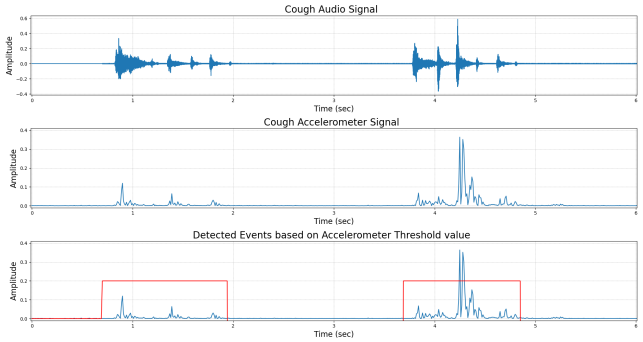


FIGURE 7. Threshold based event detection. A simple energy threshold detector is used to isolate portions of the accelerometer signal that are passed to the occupancy-change detector (Detector 1). The audio is shown for the purpose of illustration only and no acoustic information was used by the classifiers.

Furthermore, since the patient in question was undergoing treatment, the normal laboratory analyses used to assess the state of health were available.

A. COUGH COUNTING

We have previously developed a system that is able to reliably detect coughs from the same accelerometer signals $a(t)$ we are using for bed occupancy detection in this work [4]. Since the cough detector uses the accelerometer signal, it is insensitive to the coughs of other patients or visitors. This is especially useful in a multi-bed ward environment such as the TB research hospital at which we are attempting to accomplish automatic long-term cough monitoring. The simple threshold-based event detector, as used in [4] for manual event annotation and illustrated in Figure 7, was applied to

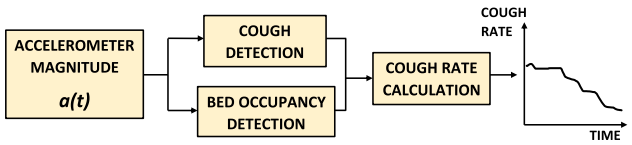


FIGURE 8. Long-term cough monitoring. A cough detection system provided the start and end times of all detected coughs in the accelerometer magnitude signal $a(t)$. The bed occupancy detection system provides the start and end times of all intervals during which the bed is believed to have been occupied. These two sources of information can be used to calculate a cough rate, which is the number of coughs per unit time, for a patient in a certain bed, as shown in Figure 10.

select the portions of the acceleration signal to pass through the cough detector. This algorithm extracted sections of the accelerometer signal for which the mean sample amplitude exceeded a small threshold (1% of the full-scale amplitude) for more than 0.5 seconds. Figure 8 provides a high-level diagram of the long-term monitoring system.

B. DAILY COUGH RATE

The average daily cough rate (coughs per 24-hour period) is a means of quantifying how much a patient coughs. It has been postulated that this figure is related to the state of the patient's health, and therefore can be used as a means of monitoring [3].

$$\mathbb{R} = \mathbb{C} \times \frac{\mathbb{B}}{24} \tag{6}$$

Let $\mathbb{C}$ be the number of coughs detected by the cough detector over a 24-hour period and $\mathbb{B}$ the total time (hours) within this same 24h period that the bed occupancy detection system believes the patient to have been present in the bed. Then, the daily cough rate ($\mathbb{R}$) is determined using Equation 6.

C. LABORATORY INDICATORS

For the patient whose cough rate we monitored, we also obtained the colony forming unit (CFU) and time to positivity (TTP) values for the same 14-day period. These are both indicators routinely used to monitor the effectiveness of TB treatment.

The colony forming unit (CFU) count is the number of TB bacterial colonies formed at a certain dilution and it is calculated by Equation 7.

$$\text{CFU} = \log_{10} \frac{\mathcal{P}_1 + \mathcal{P}_2}{2} \times 2 \times 5 \times 10^{\mathcal{D}} \tag{7}$$

The quantities $\mathcal{P}_1$ and $\mathcal{P}_2$ are the number of formed TB colonies in every 1 ml for the two plates used during culturing and $\mathcal{D}$ is the dilution strength, measured in ml [90].

The time to positivity (TTP) is the number of hours taken for the sputum samples to show signs of being TB-positive. When two plates are cultured, the TTP is calculated as the average:

$$\text{TTP} = \frac{\mathcal{H}_1 + \mathcal{H}_2}{2} \tag{8}$$

where $\mathcal{H}_1$ and $\mathcal{H}_2$ are the number of hours taken for TB samples to become positive for the two plates respectively [90].

Generally, a decrease in CFU is reflected as an increase in TTP, and therefore both can be used as a measure of successful TB treatment [90].

The estimated cost of calculating CFUs is around USD 100 and TTPs is around USD 90 and requires usually 3 to 4 weeks per patient to receive the results.

VIII. RESULTS

In the following, Sections VIII-A, VIII-B and VIII-C will present experimental results for Detectors 1, 2 and 3 respectively for the dataset described in Section III. Then, Section VIII-D presents results for the long-term cough monitoring described in Section VII.

A. DETECTOR 1: OCCUPANCY-CHANGE DETECTION RESULTS

For the occupancy-change classifier (Detector 1), only the two DNN architectures (CNN and LSTM) were considered, since the shallow classifiers exhibited substantially inferior performances in initial experiments. Both alternatives were trained and evaluated on the data introduced in Section III using the nested cross-validation procedure described in Section VI-B. The two best-performing (in terms of AUC) systems for each architecture are presented in Table 4. We note again that the hyperparameters listed in this table were optimised as part of the nested cross-validation, and that the classification performance indicators specificity, sensitivity, accuracy and AUC are averages over the seven outer loops of this process. In addition, the standard deviation of the AUC, also calculated over the outer loops, is presented and provides an indication of the robustness of the classifiers to variations in the training and testing data.

We find that the best classification performance for Detector 1 is achieved by the LSTM when features are extracted using a frame length $\Phi_{OC} = 64$ samples (640ms) and extracting $C_{OC} = 20$ sub-frames from each five-second classification frame. This system achieves a mean specificity of 71%, a mean sensitivity of 99%, a mean accuracy of 85% and a mean AUC of 0.87. We also see, as already pointed out in Section V, that all four systems in Table 4 exhibit a high sensitivity, meaning that very few occupancy changes are missed, but a lower specificity, meaning that activities such as movement by the patient while in bed are sometimes miss-classified as occupancy changes. Nevertheless, we note that the overall success of Detector 1 in identifying occupancy changes implies that the accelerometer signal for this type of event carries some distinguishing patterns that can be used for automatic classification.

B. DETECTOR 2: OCCUPANCY-INTERVAL DETECTION RESULTS

For the occupancy-interval classifier (Detector 2), two shallow (LR and MLP) and two DNN architectures (CNN and LSTM) were considered. All four were trained and evaluated on the data introduced in Section III using the nested cross-validation procedure described in Section VI-B.

TABLE 4. Leave-one-patient-out cross-validation results in detecting occupancy changes: The best two performing classifiers for each architecture are shown along with the associated optimised hyperparameters. The results show high sensitivity, but low specificity.

Classifier	Frame (Ψ_{OC})	Seg (C_{OC})	Mean Spec	Mean Sens	Mean Acc	Mean AUC	SD AUC	Best Hyperparameters
CNN	32	20	69%	99%	84%	0.86	0.0172	$\xi_1 = 2^7, \xi_2 = 100, \alpha_1 = 24, \alpha_2 = 2, \alpha_3 = 0.3, \alpha_4 = 2^4$
	64	50	68%	99%	83.5%	0.85	0.0278	$\xi_1 = 2^7, \xi_2 = 160, \alpha_1 = 48, \alpha_2 = 2, \alpha_3 = 0.5, \alpha_4 = 2^5$
LSTM	*64*	*20*	*71%*	*99%*	*85%*	*0.87*	*0.0179*	$\xi_1 = 2^8, \xi_2 = 180, \alpha_3 = 0.3, \alpha_4 = 16, \beta_1 = 2^7, \beta_2 = 10^{-3}$
	32	50	69%	99%	84%	0.86	0.0146	$\xi_1 = 2^7, \xi_2 = 140, \alpha_3 = 0.3, \alpha_4 = 16, \beta_1 = 2^7, \beta_2 = 10^{-2}$

The best-performing (in terms of AUC) two systems for each architecture are presented in Table 5. As before, the listed hyperparameters were optimised during cross-validation, and classification performance is indicated by the averages over the seven outer cross-validation loops.

We find that the best classification performance for Detector 2 is again achieved by an LSTM, in this case extracting features using a frame length of $\Phi_{OI} = 64$ samples (640 ms) and 50 sub-frames from each ten-second classification frame. This system achieves a mean specificity of 99%, a mean sensitivity of 85%, a mean accuracy of 92% and a mean AUC of 0.94. We also see, as already pointed out in Section V, that all eight systems in Table 5 exhibit a high specificity, meaning that intervals are rarely classified as "occupied" when in fact the bed was empty, but a lower sensitivity, meaning that in some cases the bed is classified as unoccupied when in fact it is not because its occupant exhibits very little movement.

C. DETECTOR 3: BED OCCUPANCY DETECTION

As described in Section V-C, the best performing occupancy-change and occupancy-interval classifiers presented in Sections VIII-A and VIII-B are considered by the occupancy-state classifier (Detector 3) in order to correct some false positive classification made by Detector 1. The resulting performance when using either CNN or LSTM classifiers for Detectors 1 and 2 is presented in Table 6, while Figure 9 shows the ROC curves for the two best systems. The results in this table reflect the overall per-sample classification performance of our accelerometer-based bed occupancy detection system. The best performance has been achieved by combining the outputs of the two LSTM classifiers used for Detectors 1 and 2, resulting in an AUC of 0.94, a specificity of 91.71% and a sensitivity of 94.51%. Hence the simple procedure implemented by Detector 3 has resulted in an overall system for which both sensitivity and specificity are high.

D. LONG-TERM COUGH MONITORING

The best LSTM-based bed occupancy detection system has been used to determine the daily cough rates $\mathbb{R}$ for a new patient over a continuous 14-day period, as described in Section VII. Even though further testing on a larger dataset of patients is necessary, this shows that the bed occupancy detection system is a promising component of a long-term automated cough monitoring system. Figure 10 shows the cough rate for this patient, together with the CFU and TTP values over the same period. We see that the CFU decreases

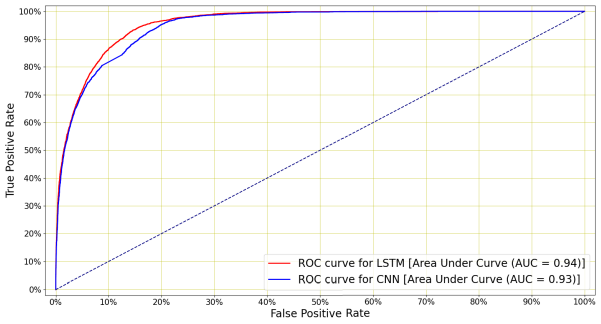


FIGURE 9. Mean ROC curve for bed occupancy detection, which shows the best performances for LSTM and CNN classifiers. The highest AUC of 0.94 has been achieved for the LSTM classifier, detailed in Table 6.

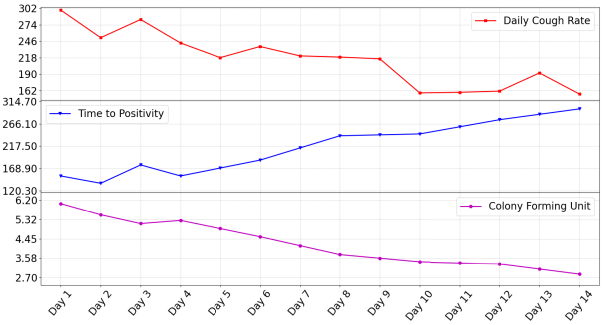


FIGURE 10. Daily cough rate ($\mathbb{R}$), time to positivity (TTP) and colony forming unit (CFU) for a patient undergoing TB treatment for TB over a continuous 14 day period. The graphs indicate that daily cough rate decreases while the TTP increases and CFU decreases over time. This suggests that the long-term cough monitoring system based on the bed occupancy detector presented in this study has the potential to be a useful means of monitoring the success of treatment for respiratory diseases such as TB and COVID-19.

over time, indicating that the number of colonies formed in the sample dilution on average decreases. We also see that the TTP increases over time, showing that the time taken for a TB sample to become positive is also increasing. Therefore, both microbiological indicators indicate that, in general, the TB treatment that the patient was receiving was successful. Finally, we see that the daily cough rate decreases over the same time interval. This suggests that automatic long-term cough monitoring, such as that implemented by the system we present in this study, may be an alternative means of monitoring the health of patients in a TB research hospital. We note that our observations agree with the observations made in [3], where the cough frequency was measured manually.

IX. DISCUSSION

The results in Section VIII demonstrate that bed occupancy detection is possible with a high accuracy based on only

TABLE 5. Leave-one-patient-out cross-validation results in detecting occupancy intervals: The best two performing classifiers for each architecture are shown along with the associated optimised hyperparameters. The results show high specificity, but low sensitivity.

Classifier	Frame (Ψ_{OI})	Seg (C_{OI})	Mean Spec	Mean Sens	Mean Acc	Mean AUC	SD AUC	Best Hyperparameters
LR	32	100	98%	81%	89.5%	0.91	0.0371	$\nu_1 = 10^{-4}, \nu_2 = 0.35, \nu_3 = 0.45$
	64	50	98%	80%	89%	0.90	0.0295	$\nu_1 = 10^{-3}, \nu_2 = 0.7, \nu_3 = 0.3$
MLP	64	100	99%	82%	90.5%	0.92	0.0309	$\eta_1 = 70, \eta_2 = 10^{-2}, \eta_3 = 0.55$
	32	50	98%	81.5%	90%	0.91	0.0205	$\eta_1 = 50, \eta_2 = 10^{-5}, \eta_3 = 0.2$
CNN	32	100	99%	84%	91.5%	0.93	0.0249	$\xi_1 = 2^8, \xi_2 = 140, \alpha_1 = 24, \alpha_2 = 2, \alpha_3 = 0.3, \alpha_4 = 16$
	64	100	99%	83%	91%	0.92	0.0239	$\xi_1 = 2^7, \xi_2 = 100, \alpha_1 = 48, \alpha_2 = 3, \alpha_3 = 0.1, \alpha_4 = 2^4$
LSTM	*64*	*50*	*99%*	*85%*	*92%*	*0.94*	*0.0288*	$\xi_1 = 2^8, \xi_2 = 180, \alpha_3 = 0.3, \alpha_4 = 2^4, \beta_1 = 2^7, \beta_2 = 10^{-3}$
	32	100	99%	83%	91%	0.93	0.0190	$\xi_1 = 2^8, \xi_2 = 180, \alpha_3 = 0.1, \alpha_4 = 16, \beta_1 = 2^8, \beta_2 = 10^{-3}$

TABLE 6. Final leave-one-patient-out cross-validation results in detecting bed occupancy: The best two performing classifiers for each architecture are shown along with the associated optimised hyperparameters. The results show both high specificity and high sensitivity and therefore also high accuracy and AUC. The highest AUC of 0.94 has been obtained by an LSTM classifier.

Classifier	Detector 1		Detector 2		Mean Specificity	Mean Sensitivity	Mean Accuracy	Mean AUC	SD AUC
	Frame (Ψ_{OC})	Seg (C_{OC})	Frame (Ψ_{OI})	Seg (C_{OI})					
CNN	32	50	64	50	93.09%	89.77%	91.43%	0.93	0.0205
	64	50	32	100	91.01%	89.91%	90.46%	0.92	0.0192
LSTM	*32*	*50*	*64*	*100*	*94.51%*	*91.71%*	*93.11%*	*0.94*	*0.0218*
	32	20	64	50	92.53%	90.88%	91.71%	0.93	0.0271

the signal captured by a bed-mounted smartphone-inbuilt accelerometer. Since cough detection is also possible using this signal, it allows the cough rate to be accurately determined by allowing periods during which the patient leaves the bed to be discounted from the calculation. Basing classification decisions only on the accelerometer signal allows a convenient, non-intrusive and privacy-preserving form of cough monitoring.

Since the accelerometer is attached to the bed, the system is insensitive to coughs from persons other than the patient in the bed. Therefore the system is perfectly suited, for example, for a multi-bed ward environment. The results in Section VII show that, for a patient undergoing standard TB treatment, the cough rate decreases over a period of 14 days while laboratory indicators such as CFU counts and TTP indicate that the treatment is effective. Even though we have shown this for only a single patient and therefore further validation is necessary, this is promising empirical evidence to suggest that cough rate can be used as a method of monitoring the effectiveness of treatment for TB patients, and perhaps also patients with other lung ailments. This could be of benefit because the proposed system is easier, quicker, less invasive and much less costly to implement than laboratory analyses.

X. CONCLUSION AND FUTURE WORK

We have described a machine learning based bed occupancy detection system that uses the accelerometer signal captured by a consumer smartphone attached to the patient's bed. Such bed occupancy detection is required to allow the implementation of automatic long-term cough monitoring using the same accelerometer signal, since the time at which the monitored patient is present in the bed must be known in order to accurately calculate a cough rate. Using a bed-mounted sensor is more convenient and less intrusive than wearable alternatives or video monitoring, and using only acceleration measurements intrinsically preserves privacy.

For experimental evaluation, we compiled a 249-hour dataset of manually-labelled acceleration signals gathered from seven patients undergoing treatment for tuberculosis (TB). Inspection of these acceleration signals revealed that they are challenging, since they are characterised by brief activity bursts interspersed with long periods of little or no activity, even when the bed is occupied. Initial experimentation revealed that recursive neural architectures, such as long short-term memory (LSTM) networks, which in other applications often deliver state-of-the-art performance for tasks that require the modelling of complex sequential data, are ineffective when presented with the acceleration signals in our dataset. Hence, to process this signal effectively, we propose a system based on three interconnected components. The first, termed as the occupancy-change detector, locates instances in time at which the occupancy of the bed is likely to have changed as a result of the patient entering or leaving the bed. The second, termed as the occupancy-interval detector, considers the periods between detected occupancy changes and classifies them as being associated with either an occupied or an unoccupied bed. The third and final component, termed the occupancy-state detector, uses the results of the first two detectors to correct some of the falsely-identified occupancy changes.

To implement the system, we consider two shallow classifiers such as liner regression (LR) and multilayer perceptron (MLP) and two deep architectures such as convolutional neural network (CNN) and LSTM. We employ nested cross-validation to train, optimise and evaluate these architectures and find that a system using LSTM networks for both the occupancy-change and the occupancy-interval detectors achieves the best performance, with an area under the ROC curve (AUC) of 0.94. For all considered combinations, we observed that the occupancy-change detector exhibits a high sensitivity but a lower specificity while the occupancy-interval detector exhibits a high specificity but a

lower sensitivity. The final occupancy-state detector was able to rectify many falsely-identified occupancy changes and thus achieve an overall system exhibiting both high sensitivity and high specificity.

As a final step, we implemented a complete cough monitoring system by integrating a previously-developed cough detector, which uses the same acceleration signal, with our proposed bed occupancy detection system. The cough detector provides the time instances of detected coughs, while the bed occupancy detector provides the time intervals during which the bed was occupied. Together, this information can be used to calculate an accurate estimate of the daily cough rate. We calculated this cough rate using acceleration signals gathered over a period of 14 days from a patient undergoing TB treatment and who was not present in our training dataset. The evolution of the resulting cough rate was compared with the evolution of the colony forming unit (CFU) counts as well as the time to positivity (TTP) determined for sputum samples from the same patient obtained by standard microbiological laboratory analyses. We were able to show that, as the CFU decreased with time and the TTP increased with time, indicating that TB treatment was effective, the measured cough rate decreased with time. This provides promising empirical evidence indicating that cough monitoring based on bed-mounted accelerometer measurements may present a quick, non-invasive, non-intrusive and cost-effective means of monitoring the long-term recovery of patients suffering from respiratory diseases such as TB and COVID-19.

As immediate future work, we aim to apply the presented cough monitoring system to a larger number of patients undergoing tuberculosis treatment, to verify whether the link between cough rate and the clinical indicators remains. We would also like to determine whether accuracy can be improved by replacing the occupancy-state detector with a trainable architecture. Although our first attempts at this have not met with success, the extension of our dataset, which is an ongoing process, may make this possible and, will also help us to investigate the effects of the various features extracted from the accelerometer signal on the classifier's performance. Finally, we would like to experiment with other more advanced neural architectures, such as residual networks (ResNets [91]) and transformers [92].

ACKNOWLEDGMENT

The authors would like to thank the South African Centre for High Performance Computing (CHPC) for providing computational resources on their Lengau cluster for this research, the Lungenliga St. Gallen-Appenzell, Switzerland, for funding parts of this work, and Amour Venter for assisting in the laboratory experiments.

REFERENCES

[1] I. Pinhas, A. Halperin, A. Averboukh, D. Lange, and Y. Gross, "Methods and systems for monitoring patients for clinical episodes," U.S. Patent 11 552 872, May 24, 2007.

[2] G. J. de Knegt, L. Dickinson, H. Pertinez, D. Evangelopoulos, T. D. McHugh, I. A. J. M. Bakker-Woudenberg, G. R. Davies, and J. E. M. de Steenwinkel, "Assessment of treatment response by colony forming units, time to culture positivity and the molecular bacterial load assay compared in a mouse tuberculosis model," *Tuberculosis*, vol. 105, pp. 113–118, Jul. 2017.

[3] A. Proaño, "Dynamics of cough frequency in adults undergoing treatment for pulmonary tuberculosis," *Clin. Infectious Diseases*, vol. 64, no. 9, pp. 1174–1181, May 2017.

[4] M. Pahar, I. Miranda, A. Diacon, and T. Niesler, "Deep neural network based cough detection using bed-mounted accelerometer measurements," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8002–8006.

[5] Y. Ren, R. Werner, N. Pazzi, and A. Boukerche, "Monitoring patients via a secure and mobile healthcare system," *IEEE Wireless Commun.*, vol. 17, no. 1, pp. 59–65, Feb. 2010.

[6] A. K. Mohiuddin, "Patient behavior: An extensive review," *Nursing Care Open Access J.*, vol. 6, no. 3, pp. 76–90, May 2019.

[7] M. Cournan, B. Fusco-Gessick, and L. Wright, "Improving patient safety through video monitoring," *Rehabil. Nursing*, vol. 43, no. 2, pp. 111–115, Mar./Apr. 2018, doi: 10.1002/rnj.308.

[8] P. T. Sanders, B. J. Cysyk, and M. A. Bare, "Safety in long-term EEG/video monitoring," *J. Neurosci. Nursing*, vol. 28, no. 5, pp. 305–314, 1996.

[9] M. H. Jones, R. Goubran, and F. Knoefel, "Identifying movement onset times for a bed-based pressure sensor array," in *Proc. IEEE Int. Workshop Med. Meas. Appl.*, Apr. 2006, pp. 111–114.

[10] M. Taylor, T. Grant, F. Knoefel, and R. Goubran, "Bed occupancy measurements using under mattress pressure sensors for long term monitoring of community-dwelling older adults," in *Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA)*, May 2013, pp. 130–134.

[11] H. J. Montoye, R. Washburn, S. Servais, A. Ertl, J. G. Webster, and F. J. Nagle, "Estimation of energy expenditure by a portable accelerometer," *Med. Sci. Sports Exerc.*, vol. 15, no. 5, pp. 403–407, 1983.

[12] M. S. Treuth, K. Schmitz, D. J. Catellier, R. G. McMurray, D. M. Murray, M. J. Almeida, S. Going, J. E. Norman, and R. Pate, "Defining accelerometer thresholds for activity intensities in adolescent girls," *Med. Sci. Sports Exerc.*, vol. 36, no. 7, p. 1259, 2004.

[13] C. E. Matthew, "Calibration of accelerometer output for adults," *Med. Sci. Sports Exerc.*, vol. 37, no. 11, pp. 512–522, 2005.

[14] S. G. Trost, K. L. Mciver, and R. R. Pate, "Conducting accelerometer-based activity assessments in field-based research," *Med. Sci. Sports Exerc.*, vol. 37, no. 11, pp. 531–543, 2005.

[15] D. S. Ward, K. R. Evenson, A. Vaughn, A. B. Rodgers, and R. P. Troiano, "Accelerometer use in physical activity: Best practices and research recommendations," *Med. Sci. Sports Exerc.*, vol. 37, no. 11, pp. 582–588, 2005.

[16] S. G. Trost, P. D. Loprinzi, R. Moore, and K. A. Pfeiffer, "Comparison of accelerometer cut points for predicting activity intensity in youth," *Med. Sci. Sports Exerc.*, vol. 43, no. 7, pp. 1360–1368, 2011.

[17] C. V. Bouten, K. R. Westerterp, M. Verduin, and J. D. Janssen, "Assessment of energy expenditure for physical activity using a triaxial accelerometer," *Med. Sci. Sports Exerc.*, vol. 26, no. 12, pp. 1516–1523, 1994.

[18] M. J. Mathie, B. G. Celler, N. H. Lovell, and A. C. F. Coster, "Classification of basic daily movements using a triaxial accelerometer," *Med. Biol. Eng. Comput.*, vol. 42, no. 5, pp. 679–687, Sep. 2004.

[19] C. Randell and H. Muller, "Context awareness by analysing accelerometer data," in *Proc. 4th Int. Symp. Wearable Comput.*, 2000, pp. 175–176.

[20] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," *Amer. Assoc. Artif. Intell.*, vol. 5, pp. 1541–1546, Jul. 2005.

[21] D. Gafurov, K. Helkala, and T. Søndrol, "Biometric gait authentication using accelerometer sensor," *J. Comput.*, vol. 1, no. 7, pp. 51–59, Nov. 2006.

[22] D. Gafurov, K. Helkala, and T. Soendrol, "Gait recognition using acceleration from MEMS," in *Proc. 1st Int. Conf. Availability, Rel. Secur. (ARES)*, 2006, p. 6.

[23] A. K. Bourke, J. V. O'Brien, and G. M. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait Posture*, vol. 26, no. 2, pp. 194–199, 2007.

[24] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proc. Int. Conf. Pervasive Comput.* Cham, Switzerland: Springer, 2004, pp. 1–17.

[25] T. Brezmes, J.-L. Gorricho, and J. Cotrina, "Activity recognition from accelerometer data on a mobile phone," in *Proc. Int. Work-Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, 2009, pp. 796–799.

[26] P. Casale, O. Pujol, and P. Radeva, "Human activity recognition from accelerometer data using a wearable device," in *Proc. Iberian Conf. Pattern Recognit. Image Anal.* Cham, Switzerland: Springer, 2011, pp. 289–296.

[27] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explor. Newslett.*, vol. 12, no. 2, pp. 74–82, May 2011.

[28] A. Bayat, M. Pomplun, and D. A. Tran, "A study on human activity recognition using accelerometer data from smartphones," *Proc. Comput. Sci.*, vol. 34, pp. 450–457, Jan. 2014.

[29] P. Siirtola and H. Röning, "Recognizing human activities user-independently on smartphones based on accelerometer data," *Int. J. Interact. Multimedia Artif. Intell.*, vol. 1, no. 5, pp. 38–45, 2012.

[30] S. Hemminki, P. Nurmi, and S. Tarkoma, "Accelerometer-based transportation mode detection on smartphones," in *Proc. 11th ACM Conf. Embedded Networked Sensor Syst.*, 2013, pp. 1–14.

[31] L. Zhang, X. Wu, and D. Luo, "Recognizing human activities from raw accelerometer data using deep neural networks," in *Proc. IEEE 14th Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2015, pp. 865–870.

[32] S.-M. Lee, S. M. Yoon, and H. Cho, "Human activity recognition from accelerometer data using convolutional neural network," in *Proc. IEEE Int. Conf. Big Data Smart Comput. (BigComp)*, Feb. 2017, pp. 131–134.

[33] I. Andrey, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Appl. Soft Comput.*, vol. 62, pp. 915–922, Jan. 2017.

[34] M. A. Rahman, Y. Mia, M. R. Masum, D. M. H. Abid, and T. Islam, "Real time human activity recognition from accelerometer data using convolutional neural networks," in *Proc. 7th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2022, pp. 1394–1397.

[35] M. M. Hassan, M. Z. Uddin, A. Mohamed, and A. Almogren, "A robust human activity recognition system using smartphone sensors and deep learning," *Future Gener. Comput. Syst.*, vol. 81, pp. 307–313, Apr. 2018.

[36] C. Leng, "Discriminating coughs in a multi-bed ward environment," MEng thesis, Dept. Elect. Electron. Eng., Univ. Stellenbosch, Stellenbosch, South Africa, 2021.

[37] P. Wittenburg, H. Brugman, A. Russel, A. Klassmann, and H. Sloetjes, "ELAN: A professional framework for multimodality research," in *Proc. 5th Int. Conf. Lang. Resour. Eval.*, 2006, pp. 1556–1559.

[38] M. Pahar, I. Miranda, A. Diacon, and T. Niesler, "Automatic non-invasive cough detection based on accelerometer and audio signals," *J. Signal Process. Syst.*, vol. 94, no. 8, pp. 821–835, 2022.

[39] C. Bingham, M. Godfrey, and J. Tukey, "Modern techniques of power spectrum estimation," *IEEE Trans. Audio Electroacoust.*, vol. AE-15, no. 2, pp. 56–66, Jun. 1967.

[40] B. Liang, S. D. Iwnicki, and Y. Zhao, "Application of power spectrum, cepstrum, higher order spectrum and neural network analyses for induction motor fault diagnosis," *Mech. Syst. Signal Process.*, vol. 39, nos. 1–2, pp. 342–360, Aug. 2013.

[41] L. Durak and O. Arikan, "Short-time Fourier transform: Two fundamental properties and an optimal implementation," *IEEE Trans. Signal Process.*, vol. 51, no. 5, pp. 1231–1242, May 2003.

[42] T.-P. Jung, S. Makeig, M. Stensmo, and T. J. Sejnowski, "Estimating alertness from the EEG power spectrum," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 1, pp. 60–69, Jan. 1997.

[43] M. Pahar and L. S. Smith, "Coding and decoding speech using a biologically inspired coding system," in *Proc. IEEE Symp. Ser. Comput. Intell. (SSCI)*, Dec. 2020, pp. 3025–3032.

[44] R. K. Sinha, "Artificial neural network detects changes in electroencephalogram power spectrum of different sleep-wake states in an animal model of heat stress," *Med. Biol. Eng. Comput.*, vol. 41, no. 5, pp. 595–600, Sep. 2003.

[45] R. L. Lux, C. T. Sower, N. Allen, S. P. Etheridge, M. Tristani-Firouzi, and E. V. Saarel, "The application of root mean square electrocardiography (RMS ECG) for the detection of acquired and congenital long QT syndrome," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e85689.

[46] J. Gilmore, M. Islam, J. Duncan, R. Natu, and R. Martinez-Duarte, "Assessing the importance of the root mean square (RMS) value of different waveforms to determine the strength of a dielectrophoresis trapping force," *Electrophoresis*, vol. 38, no. 20, pp. 2561–2564, Oct. 2017.

[47] X. Yuan, Q. Tan, X. Lei, Y. Yuan, and X. Wu, "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine," *Energy*, vol. 129, pp. 122–137, Jun. 2017.

[48] X. Zhang, P. E. Barkhaus, W. Z. Rymer, and P. Zhou, "Machine learning for supporting diagnosis of amyotrophic lateral sclerosis using surface electromyogram," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp. 96–103, Jan. 2014.

[49] J. Lepine, V. Rouillard, and M. Sek, "On the use of machine learning to detect shocks in road vehicle vibration signals," *Packag. Technol. Sci.*, vol. 30, no. 8, pp. 387–398, Aug. 2017.

[50] L. Ren, J. Cui, Y. Sun, and X. Cheng, "Multi-bearing remaining useful life collaborative prediction: A deep learning approach," *J. Manuf. Syst.*, vol. 43, pp. 248–256, Apr. 2017.

[51] G. Takahashi, T. Yamada, S. Makino, and N. Ono, "Acoustic scene classification using deep neural network and frame-concatenated acoustic feature," in *Proc. Workshop Detection Classification Acoustic Scenes Events (DCASE)*, Budapest, Hungary, Sep. 2016.

[52] J. Van Hulse, T. M. Khoshgoftaar, and A. Napolitano, "Experimental perspectives on learning from imbalanced data," in *Proc. 24th Int. Conf. Mach. Learn.*, Jun. 2007, pp. 935–942.

[53] B. Krawczyk, "Learning from imbalanced data: Open challenges and future directions," *Prog. Artif. Intell.*, vol. 5, no. 4, pp. 221–232, 2016.

[54] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, Jan. 2002.

[55] G. Lemaitre, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A Python toolbox to tackle the curse of imbalanced datasets in machine learning," *J. Mach. Learn. Res.*, vol. 18, no. 1, pp. 559–563, 2017.

[56] A. Windmon, M. Minakshi, P. Bharti, S. Chellappan, M. Johansson, B. A. Jenkins, and P. R. Athilingam, "TussisWatch: A smart-phone system to identify cough episodes as early symptoms of chronic obstructive pulmonary disease and congestive heart failure," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 4, pp. 1566–1573, Jul. 2019.

[57] R. Blagus and L. Lusa, "SMOTE for high-dimensional class-imbalanced data," *BMC Bioinf.*, vol. 14, no. 1, p. 106, 2013.

[58] H. Han, W.-Y. Wang, and B.-H. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning," in *Proc. Int. Conf. Intell. Comput.* Cham, Switzerland: Springer, 2005, pp. 878–887.

[59] H. M. Nguyen, E. W. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," in *Proc. 5th Int. Workshop Comput. Intell. Appl.*, 2009, pp. 24–29.

[60] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jun. 2008, pp. 1322–1328.

[61] J. Pan, C. Liu, Z. Wang, Y. Hu, and H. Jiang, "Investigation of deep neural networks (DNN) for large vocabulary continuous speech recognition: Why DNN surpasses GMMS in acoustic modeling," in *Proc. 8th Int. Symp. Chin. Spoken Lang. Process.*, Dec. 2012, pp. 301–305.

[62] M. Pahar, M. Klopper, B. Reeve, R. Warren, G. Theron, and T. Niesler, "Automatic cough classification for tuberculosis screening in a real-world environment," *Physiological Meas.*, vol. 42, no. 10, Oct. 2021, Art. no. 105014.

[63] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "COVID-19 cough classification using machine learning and global smartphone recordings," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104572.

[64] E. Christodoulou, J. Ma, G. S. Collins, E. W. Steyerberg, J. Y. Verbakel, and B. Van Calster, "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models," *J. Clin. Epidemiology*, vol. 110, pp. 12–22, Jun. 2019.

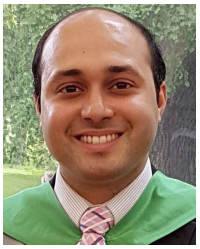
[65] G. H. R. Botha, G. Theron, R. M. Warren, M. Klopper, K. Dheda, P. D. van Helden, and T. R. Niesler, "Detection of tuberculosis by automatic cough sound analysis," *Physiological Meas.*, vol. 39, no. 4, Apr. 2018, Art. no. 045005.

[66] S. Le Cessie and J. C. Van Houwelingen, "Ridge estimators in logistic regression," *Appl. Statist.*, vol. 41, no. 1, pp. 191–201, 1992.

[67] Y. Tsuruoka, J. Tsujii, and S. Ananiadou, "Stochastic gradient descent training for L1-regularized log-linear models with cumulative penalty," in *Proc. Joint Conf. 47th Annu. Meeting ACL 4th Int. Joint Conf. Natural Lang. Process.*, 2009, pp. 477–485.

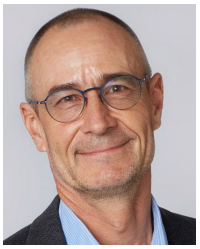
[68] H. Yamashita and H. Yabe, "An interior point method with a primal-dual quadratic barrier penalty function for nonlinear optimization," *SIAM J. Optim.*, vol. 14, no. 2, pp. 479–499, Jan. 2003.

[69] H. Taud and J. Mas, "Multilayer perceptron (MLP)," in *Geomatic Approaches for Modeling Land Change Scenarios*. Berlin, Germany: Springer, 2018, pp. 451–455.
[70] L. Sarangi, M. N. Mohanty, and S. Pattanayak, "Design of MLP based model for analysis of patient suffering from influenza," *Proc. Comput. Sci.*, vol. 92, pp. 396–403, Jan. 2016.
[71] B. H. Tracey, G. Comina, S. Larson, M. Bravard, J. W. Lopez, and R. H. Gilman, "Cough detection algorithm for monitoring patient recovery from pulmonary tuberculosis," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 6017–6020.
[72] J.-M. Liu, M. You, Z. Wang, G.-Z. Li, X. Xu, and Z. Qiu, "Cough detection using deep neural networks," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2014, pp. 560–563.
[73] J. Amoh and K. Odame, "DeepCough: A deep convolutional neural network in a wearable cough detection system," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2015, pp. 1–4.
[74] M. Pahar, M. Klopper, B. Reeve, R. Warren, G. Theron, A. Diacon, and T. Niesler, "Wake-Cough: Cough spotting and cougher identification for personalised long-term cough monitoring," in *Proc. 30th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2022, pp. 185–189.
[75] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 2, pp. 84–90, Jun. 2012.
[76] M. Pahar, M. Klopper, R. Warren, and T. Niesler, "COVID-19 detection in cough, breath and speech using deep transfer learning and bottleneck features," *Comput. Biol. Med.*, vol. 141, Feb. 2022, Art. no. 105153.
[77] I. Miranda, G. Cardoso, M. Pahar, G. Oliveira, and T. Niesler, "Machine learning prediction of hospitalization due to COVID-19 based on self-reported symptoms: A study for Brazil," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Jul. 2021, pp. 1–5.
[78] M. Pahar, M. Klopper, B. Reeve, R. Warren, G. Theron, A. Diacon, and T. Niesler, "Automatic tuberculosis and COVID-19 cough classification using deep learning," in *Proc. Int. Conf. Electr., Comput. Energy Technol. (ICECET)*, Jul. 2022, pp. 1–9.
[79] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 98–113, Jan. 1997.
[80] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6.
[81] X. Qi, T. Wang, and J. Liu, "Comparison of support vector machine and softmax classifiers in computer vision," in *Proc. 2nd Int. Conf. Mech., Control Comput. Eng. (ICMCCE)*, Dec. 2017, pp. 151–155.
[82] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[83] I. D. S. Miranda, A. H. Diacon, and T. R. Niesler, "A comparative study of features for acoustic cough detection using deep architectures," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 2601–2605.
[84] G. T. Frost, G. Theron, and T. Niesler, "TB or not TB? Acoustic cough analysis for tuberculosis classification," in *Proc. Interspeech*, Sep. 2022, pp. 2448–2452.
[85] E. Marchi, F. Vesperini, F. Weninger, F. Eyben, S. Squartini, and B. Schuller, "Non-linear prediction with LSTM recurrent neural networks for acoustic novelty detection," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–7.
[86] J. Amoh and K. Odame, "Deep neural networks for identifying cough sounds," *IEEE Trans. Biomed. Circuits Syst.*, vol. 10, no. 5, pp. 1003–1011, Oct. 2016.
[87] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306.
[88] C. Sammut and I. G. Webb, "Leave-one-out cross-validation," in *Encyclopedia of Machine Learning*. Boston, MA, USA: Springer, 2010, pp. 600–601.
[89] T.-T. Wong, "Performance evaluation of classification algorithms by *K*-fold and leave-one-out cross validation," *Pattern Recognit.*, vol. 48, no. 9, pp. 2839–2846, 2015.
[90] A. H. Diacon, J. S. Maritz, A. Venter, P. D. van Helden, R. Dawson, and P. R. Donald, "Time to liquid culture positivity can substitute for colony counting on agar plates in early bactericidal activity studies of antituberculosis agents," *Clin. Microbiology Infection*, vol. 18, no. 7, pp. 711–717, Jul. 2012.
[91] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[92] K. Han, A. Xiao, E. Wu, J. Guo, C. Xu, and Y. Wang, "Transformer in transformer," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 34, 2021, pp. 15908–15919.



MADHURANANDA PAHAR received the B.Sc. degree in mathematics from the University of Calcutta, India, and the M.Sc. degree in computing for financial markets and the Ph.D. degree in computational neuroscience from the University of Stirling, Scotland. He is currently a Postdoctoral Fellow with Stellenbosch University, South Africa. His research interests include machine learning and signal processing for audio signals and smart sensors in bio-medicine, such as the detection and classification of TB and COVID-19 coughs in a real-world environment. He is also involved in the application of deep learning to the detection and classification of TB and COVID coughs in real-world environments as well as the monitoring of patient behavior using smart sensors, such as an accelerometer.

IGOR MIRANDA received the B.Eng. and M.Eng. degrees in electronic engineering and the Ph.D. degree in industrial engineering from the Federal University of Bahia, Brazil, in 2007, 2009, and 2017, respectively. He joined the Division of Electrical and Computer Engineering, Federal University of Recôncavo da Bahia, as a Lecturer, in 2013, where he has been an Assistant Professor, since 2018. From 2018 to 2019, during a sabbatical year, he was a Postdoctoral Research Associate with Stellenbosch University. His research interests include signal processing, machine learning, and VLSI design.



ANDREAS DIACON was born in Switzerland. He received the Ph.D. degree from Stellenbosch University. He is currently leading TASK Applied Science, Cape Town, to advance innovative diagnostics, vaccines, and treatments for tuberculosis, and other diseases common in South Africa. He is also a Physician with specialist training in internal medicine and pulmonology. His research interest includes clinical tuberculosis research, particularly to bring new anti-tuberculosis agents to the patients that need them most quickly and safely.

THOMAS NIESLER received the B.Eng. and M.Eng. degrees in electronic engineering from Stellenbosch University, South Africa, in 1991 and 1993, respectively, and the Ph.D. degree from the University of Cambridge, U.K., in 1998. He joined the Department of Engineering, University of Cambridge, as a Lecturer, in 1998, and subsequently the Department of Electrical and Electronic Engineering, Stellenbosch University, in 2000, where he has been a Professor, since 2012. His research interests include signal processing, pattern recognition, and machine learning.