IPEM
Institute of Physics and
Engineering in Medicine

TOPICAL REVIEW • OPEN ACCESS

Algorithmic detection of sleep-disordered breathing using respiratory signals: a systematic review

View the article online for updates and enhancements.

Physiological Measurement

IPEM
Institute of Physics and
Engineering in Medicine

CrossMark

TOPICAL REVIEW

Algorithmic detection of sleep-disordered breathing using respiratory signals: a systematic review

Liqing Yang[1,2] (iD), Zhimei Ding[1,2], Jiangjie Zhou[2,3], Siyuan Zhang[1,2], Qi Wang[2,3], Kaige Zheng[2,3], Xing Wang[1,3] (iD) and Lin Chen[1,3,*] (iD)

[1] Key Laboratory of Biorheological Science and Technology of Ministry of Education, Chongqing University, Chongqing, People's Republic of China
[2] Chongqing Medical Electronics Engineering Technology Research Center, Chongqing University, Chongqing, People's Republic of China
[3] Chongqing Key Laboratory of Artificial Intelligence and Service Robot Control Technology, Chongqing, People's Republic of China
[*] Author to whom any correspondence should be addressed.

E-mail: clxyz@cqu.edu.cn

Abstract

Background and Objective. Sleep-disordered breathing (SDB) poses health risks linked to hypertension, cardiovascular disease, and diabetes. However, the time-consuming and costly standard diagnostic method, polysomnography (PSG), limits its wide adoption and leads to underdiagnosis. To tackle this, cost-effective algorithms using single-lead signals (like respiratory, blood oxygen, and electrocardiogram) have emerged. Despite respiratory signals being preferred for SDB assessment, a lack of comprehensive reviews addressing their algorithmic scope and performance persists. This paper systematically reviews 2012–2022 literature, covering signal sources, processing, feature extraction, classification, and application, aiming to bridge this gap and provide future research references. *Methods*. This systematic review followed the registered PROSPERO protocol (CRD42022385130), initially screening 342 papers, with 32 studies meeting data extraction criteria. *Results*. Respiratory signal sources include nasal airflow (NAF), oronasal airflow (OAF), and respiratory movement-related signals such as thoracic respiratory effort (TRE) and abdominal respiratory effort (ARE). Classification techniques include threshold rule-based methods (8), machine learning models (13), and deep learning models (11). The NAF-based algorithm achieved the highest average accuracy at 94.11%, surpassing 78.19% for other signals. Hypopnea detection sensitivity with single-source respiratory signals remained modest, peaking at 73.34%. The TRE and ARE signals proved to be reliable in identifying different types of SDB because distinct respiratory disorders exhibited different patterns of chest and abdominal motion. *Conclusions*. Multiple detection algorithms have been widely applied for SDB detection, and their accuracy is closely related to factors such as signal source, signal processing, feature selection, and model selection.

1. Introduction

Sleep-disordered breathing (SDB) is a group of disorders that disrupt normal breathing patterns and quality during sleep, manifesting as intermittent occurrences of respiratory apnea or hypopnea. According to the definition provided by the American academy of sleep medicine (AASM) (Iber C *et al* 2017), apnea is defined as a drop of more than 90% from the baseline airflow lasting at least 10 s. Hypopnea is defined as a drop of more than 30% from the baseline airflow lasting for at least 10 s, accompanied by either a desaturation of more than 3% from the pre-event baseline or an arousal from sleep. SDB can be classified into three types based on the underlying causes: obstructive sleep apnea (OSA), central sleep apnea (CSA), and complex sleep apnea (MIX), which is a combination of both OSA and CSA, each with different respiratory symptoms. Overall, SDB can have severe impacts on physical and mental function, resulting endothelial dysfunction, oxidative stress,

inflammation, glucose dysregulation, and brain and white matter pathological changes (Daulatzai 2015, Polsek *et al* 2018, Liguori *et al* 2021). Furthermore, it has been associated with various disorders, including hypertension, diabetes, metabolic syndrome, osteoporosis, and cardiovascular diseases (Floras 2015, Liguori C *et al* 2016, Sharma and Culebras 2016, Reutrakul and Mokhlesi 2017, Ryan 2017). Additionally, cognitive impairment and Alzheimer's disease (AD) have also been linked to SDB (Ju *et al* 2013, Polsek *et al* 2018, Shi *et al* 2018). Moreover, SDB can increase mortality in patients with heart failure, stroke, or coronary artery disease.

The prevalence of SDB overall remains uncertain. A recent review of 17 studies from 16 countries focusing on OSA, estimated that approximately 936 million adults aged 30–69 years worldwide have mild to severe OSA, with around 425 million adults identified as having moderate to severe OSA (Benjafield *et al* 2019). Despite the high prevalence, patients and their partners often overlook the symptoms of SDB, such as snoring, gasping, or choking, without recognizing their association with the condition. Consequently, patients typically do not seek medical attention for these symptoms. Moreover, the diagnosis of SDB presents challenges as it primarily occurs during sleep, making traditional clinical evaluation methods less effective. The gold standard method for SDB diagnosis is polysomnography (PSG) (Graco *et al* 2018), which involves utilizing specialized equipment in a dedicated laboratory with a minimum of 22 electrodes and 11 channels to continuously collect, measure, and analyze signals such as electroencephalography (EEG), electrocardiography (ECG), respiration, blood oxygen levels, and other relevant physiological parameters. Therefore, PSG is time-consuming, labor-intensive, and expensive with limited patient adherence. Overall, the underrecognized symptoms, limited diagnostic methods, and the challenges of PSG collectively contribute to underdiagnosis of SDB, with an estimated over 80% of SDB cases lacking accurate diagnosis and timely treatment (Jin and Sanchez-Sinencio 2015, Jaiswal *et al* 2017).
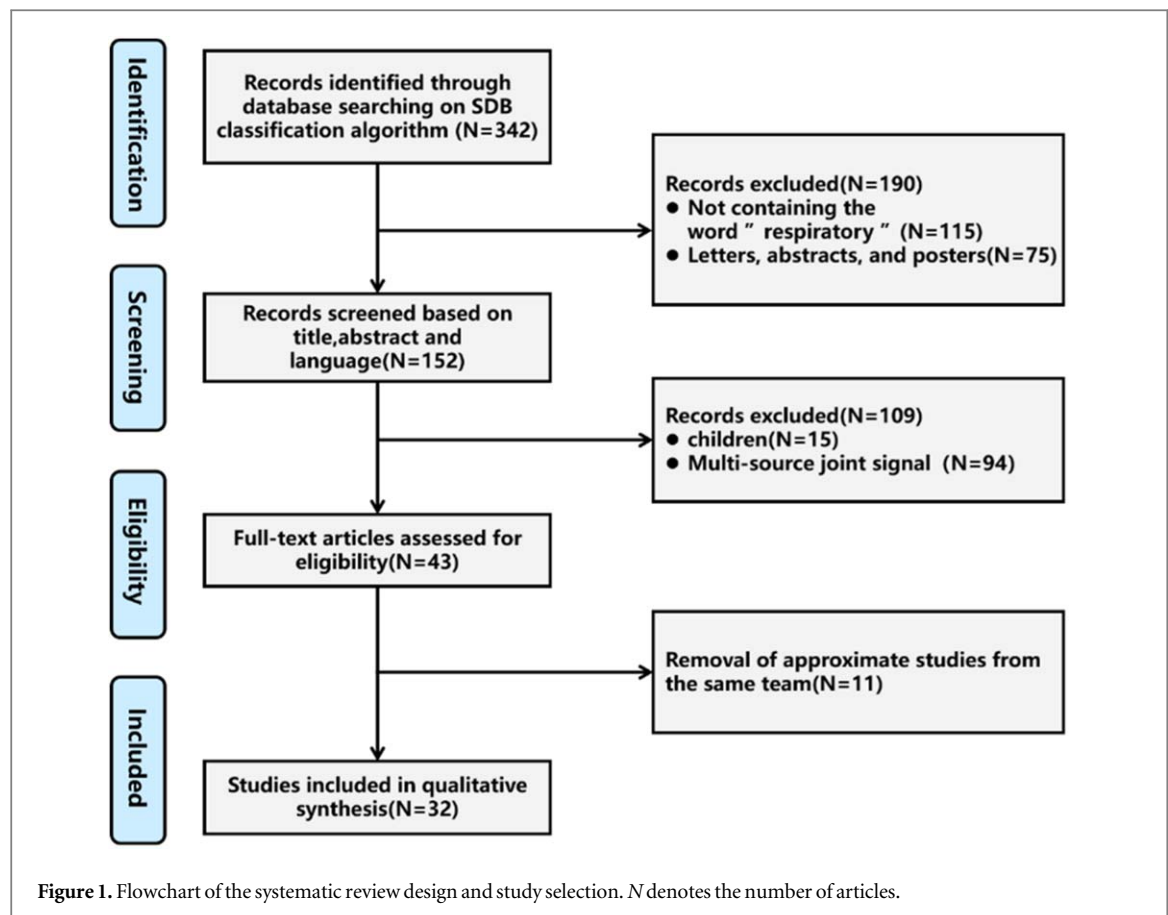
To overcome these challenges, researchers have proposed alternative methods for detecting SDB using various single-lead signals, such as respiratory signals (Van Steenkiste *et al* 2019), blood oxygen (Deviaene *et al* 2020), snoring (Hu *et al* 2022), and ECG (Shen *et al* 2021), etc. Among these, respiratory signals, including nasal airflow (NAF), oronasal airflow (OAF), and respiratory movement-related signals such as thoracic respiratory effort (TRE) and abdominal respiratory effort (ARE), are the preferred signal source for SDB detection according to the recommendations of the AASM. Over the past ten years, multiple studies (Nakano *et al* 2007, Makarie Rofail L *et al* 2010, Masa *et al* 2011, Crowley *et al* 2013, Morgenstern *et al* 2013, Masa *et al* 2014) have consistently showed the effectiveness and accuracy of single-channel respiratory signal detection in identifying and accurately diagnosing SDB, of which algorithms based on different methods, such as threshold rule, machine learning (ML) and deep learning (DL), have been developed. While several reviews have been conducted on SDB algorithms, including ECG-based algorithms (Faust *et al* 2016), respiratory and blood oxygen fusion-based algorithms (Alvarez-Estevez and Moret-Bonillo 2015, Uddin *et al* 2018), and multiple signals-based algorithms (Mendonca *et al* 2019, Serrano Alarcon *et al* 2021), there is currently a gap in the literature regarding a systematic review specifically focusing on algorithms for SDB identification using respiratory signals and/or respiratory movement-related signals. Thus, our review aims to fill this gap by conducting a systematic literature review of respiratory signal-based algorithms for SDB detection during the recent decade (2012–2022). The objectives are to compare the advantages and disadvantages of various algorithms based on respiratory signals and methods, and summarize and analyze current research trends in algorithm development, thus providing a comprehensive reference guide for new researchers developing SDB recognition algorithms.

2. Methods

This systematic review was conducted according to the protocol registered with PROSPERO (International Prospective Register of Systematic Reviews, CRD42022385130, https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022385130). The search strategy protocol followed guidelines from PRISMA statement (Liberati *et al* 2009) and its extension PRISMA-S (Rethlefsen *et al* 2021), and the following eligibility criteria were applied to define the structure of the current systematic review.

2.1. Search strategy
We conducted a comprehensive literature review of papers published between 2012 and 2022 via searching various databases, including Web of Science, IEEE explorer, and PubMed, as well as through the cited literature in the included articles and related journals. The keywords employed in the search were 'SDB detection algorithm', 'algorithm AND sleep Apnea', 'Respiration analysis AND apnea', and 'Apnea AND deep'. Articles were selected by 2 reviewers independently (blinded to each other's assessment) by applying the criteria to each title and abstract and then assessed fully. Divergent opinions were resolved through group discussion until reaching consensus. All processes were performed in Endnote, a bibliographic software used for managing references.

Figure 1. Flowchart of the systematic review design and study selection. *N* denotes the number of articles.

2.2. Selection criteria

The following criteria were used to select eligible studies in this review: (1) the signal must be a single-channel respiratory signal or respiratory movement-related signal, (2) a complete computer-based detection and/or prediction system for SDB must have been proposed, (3) detailed system evaluation data with comparisons to the gold standard (PSD) must have been included, which can verify system validity, and (4) preliminary or definitive results must have been formed. Additionally, these criteria also applied to studies obtained from reference tracking. Algorithms based on multi-source combined signals (e.g. respiration, combined blood oxygen) and indirect signals (e.g. respiratory variability extracted from ECG signals) were excluded. In addition, adults and children are two distinct entities in SDB detection. Sleep architecture, respiratory physiology, apnea definition, and apnea severity in adults differ from children/pediatric subjects (Alsubie and BaHammam 2017, Kljajić *et al* 2017). Therefore, the algorithms for SDB detection in pediatric subjects are quite different and usually need special consideration or criteria to obtain better detection results. Accordingly, this review focuses only on the adult population. Figure 1 presents flowchart of the systematic review design and study selection. Overall, a total of 342 studies were screened, of which 43 studies met the inclusion criteria and were eligible for data extraction.

2.3. Data extraction

We conducted detailed data abstraction from a total of 43 studies. The article information (title, author, publication date, journal), method (database, main decision, classification method) and evaluation metrics (accuracy, specificity, etc) were abstracted and recorded in the Excel table. Python packages (Pandas, NumPy, and Matplotlib) was used for data visualization. We classified and sorted the results in ascending order according to different algorithms (rule threshold-based, machine learning, deep learning), and further eliminated 11 papers with duplicated content or same approaches from the same author, including preprints or early-stage conference outcomes. In the final analysis, a total of 32 papers were included for final discussion.

3. Results and discussion

Based on the findings of the literature, the algorithms for SDB detection are varied by signal source, signal processing technique, feature extraction method, and the choice of model. Therefore, we present and discuss the results from these four aspects.

3.1. Respiratory signal acquisition sensors

According to the AASM guidelines (Iber C *et al* 2017), the primary sensors for measuring respiratory signals include oronasal thermal airflow sensors (mainly for detecting apnea) and nasal pressure sensors (mainly for detecting hypopnea). Another alternative sensor option is the thermal sensor. Thermal airflow sensors measure the temperature of the air flowing through the nasal and/or oral passages, while thermal sensors measure the temperature of the skin around the nose and mouth. These sensors are utilized to acquire respiratory signals by detecting changes in airflow and temperature associated with episodes of apnea and hypopnea. For measuring respiratory effort signals, commonly employed sensors include respiratory inductive plethysmography belts (RIP belts) and polyvinylidene fluoride belts (PVDF belts). RIP belts are devices worn around the chest and abdomen to measure respiratory movement-related signals, producing RIPsum (the sum of the chest and abdomen RIP belt signals) and RIPflow (the rate of airflow into and out of the lungs). Additionally, RIPsum can replace primary sensors (e.g. oronasal thermal airflow or nasal pressure sensors) when airflow signals are not available or unreliable (Farre *et al* 2004, Kaniusas and Kaniusas 2012). In addition, chest and abdominal signals related to respiration can also be acquired by PVDF belts, called PVDFsum.

Recent developments in sensor technology have introduced new methods for measuring respiratory signal during sleep, such as radar and imaging techniques. The radar method uses either Doppler radar or impulse radio ultra-wideband (IR-UWB) radar. Doppler radar employs the Doppler effect, which is the frequency shift of a signal due to the motion of the transmitter and receiver, to detect changes in the chest and abdomen (Lee *et al* 2014). On the other hand, IR-UWB radar employs wide bandwidth and high-frequency carrier waves to detect minute movements in the chest and abdomen (Kang *et al* 2020). Both methods have demonstrated promising results in the studies. However, Doppler radar is susceptible to interference from random body movements and encounters challenges with null-point detection (Sun and Matsui 2015, Park *et al* 2006). Similarly, IR-UWB radar indirectly determines the breathing state from chest movements, which may lead to discrepancies between the recorded data and the actual breathing state. Moreover, even small patient movements can significantly disrupt the measured signals, such as respiration rate and heart rate (Park *et al* 2019). To overcome these limitations, researchers have proposed the use of imaging techniques, such as Infrared Optical Gas Imaging (IR-OGI) (An *et al* 2022), which has been demonstrated to have a sensitivity of 96.0%. Additionally, the three-dimensional (3D) time-of-flight (TOF) Camera (Coronel *et al* 2019, Coronel *et al* 2020) has exhibited similar functionality to RIPsum.

Based on the signals obtained from these sensors, they can be roughly classified into NAF, OAF, and respiratory movement-related signals such as TRE and ARE. Among these signals, the NAF signal has been extensively studied in the literature, with a total of 18 studies (Selvaraj and Narasimhan 2013, Guijarro-Berdiñas *et al* 2012, Avci and Akbas 2015, Ciolek *et al* 2015, Gutierrez-Tobal *et al* 2012a, 2012b, 2013, 2016, Lee *et al* 2016, Haidar *et al* 2017, Choi *et al* 2018, McCloskey *et al* 2018, Gogus and Tezel 2019, Chen *et al* 2020, Gogus *et al* 2020, ElMoaqet *et al* 2020a, Wu *et al* 2021, Yue *et al* 2021). Subsequently, a combined total of 10 studies (Van Steenkiste *et al* 2019, Bianchi *et al* 2014, Kagawa *et al* 2016, Azimi *et al* 2018, Lin *et al* 2016, Adha and Igasaki 2021, Thommandram *et al* 2013, Hafezi *et al* 2020, Drzazga and Cyganek 2021, Nassi *et al* 2022) utilized ARE and TRE signals, while OAF was mentioned in 4 studies (Koley and Dey 2013a, 2013b, Kim *et al* 2019, ElMoaqet *et al* 2020b). The detailed list of signals and their corresponding literature sources can be found in table 1. The NAF signal has received widespread attention in research, potentially due to two primary reasons. Firstly, compared to the OAF signal, the NAF signal can be obtained using common and easy-to-use sensors such as oronasal thermal airflow sensors or nasal pressure sensors. In contrast, acquiring the OAF signal requires simultaneous measurement of airflow from both the nasal and oral cavities, making sensor selection and placement more complex. Additionally, most PSG devices routinely capture the NAF signal. Furthermore, compared to measuring TRE and ARE, these signals are susceptible to interference from body movements and other artifacts. This is likely another reason why the NAF signal has received relatively more attention in research.

3.2. Signal preprocessing

Respiratory signals are typically measured by air flow or monitoring chest and abdomen movement. However, they can be affected by various types of noise, such as baseline wander, power line interference, muscle artifacts, and electrode motion artifacts (Sankar *et al* 2010). This can mask the tiny features of the signal and lead to false diagnosis. To reduce power line interference, a notch filter with a bandwidth of 50–60 Hz was used in several studies (Ziarani and Konrad 2002, Łęski and Henzel 2005, Nassi *et al* 2022). For other types of noise, various filters such as Butterworth, finite impulse response (FIR) and infinite impulse response (IIR) have been applied to different frequency ranges of 0.01–30 Hz, as summarized in table 2. In general, there is no standardized frequency range for these signals in the studies. The potential reason, as stated by Várady *et al* (2002) could be

Table 1. Frequency of use of different respiratory signals in literature.

No	Signals	Frequency of use	Author
1	NAF	18	Selvaraj and Narasimhan (2013), Ciolek *et al* (2015), Lee *et al* (2016), Guijarro-Berdiñas *et al* (2012), Gutierrez-Tobal *et al* (2012a, 2012b, 2013, 2016), Avci and Akbas (2015), Gogus and Tezel (2019), Gogus *et al* (2020), Haidar *et al* (2017), McCloskey *et al* (2018), Choi *et al* (2018), ElMoaqet *et al* (2020a), Chen *et al* (2020), Wu *et al* (2021), Yue *et al* (2021)
2	OAF	4	ElMoaqet (Kim *et al* 2019, ElMoaqet *et al* 2020b), Koley and Dey (2013a, 2013b)
3	TRE, ARE	10	Bianchi *et al* (2014), Kagawa *et al* (2016), Azimi *et al* (2018), Adha and Igasaki (2021), Thommandram *et al* (2013), Lin *et al* (2016), Van Steenkiste *et al* (2019), Hafezi *et al* (2020), Drzazga and Cyganek (2021), Nassi *et al* (2022)

NAF: nasal airflow, OAF: oronasal airflow; TRE: thoracic respiratory effort, ARE: abdominal respiratory effort

Table 2. List of cutoff frequencies.

No	Author	Year	Filter	Cutoff frequency
1	Koley and Dey (2013b)	2013	Sixth-order band pass Butterworth filter	Pass band of 0.01–0.8 Hz
2	Selvaraj and Narasimhan (2013)	2013	low-pass filter	Passband corner frequency of 0.7 Hz
3	Gutierrez-Tobal *et al* (2016)	2015	Anti-aliasing FIR Butterworth low-pass filter;	Cutoff of 1.2 Hz
4	Choi *et al* (2018)	2018	Fifth-order IIR Butterworth filter	High-pass of 0.01 Hz, Low-pass of 3 Hz
5	Azimi *et al* (2018)	2018	FIR linear Phase band pass filter	Pass band of 0.07–0.8 Hz
6	Hafezi *et al* (2020)	2020	Bandpass filter	Pass band of 0.1–25 Hz
7	ElMoaqet *et al* (2020a)	2020	Low pass FIR filter	Cutoff of 0.5 Hz
8	Gogus *et al* (2020)	2020	Third-order Butterworth bandpass filter	Pass band of 0.01–0.7 Hz
9	Wu *et al* (2021)	2021	FIR band-pass filter	Pass band of 0.5–30 Hz
10	Adha and Igasaki (2021)	2021	Bandpass filter	Pass band of 0.07–0.8 Hz
11	Nassi *et al* (2022)	2022	Notch filter	Cutoff of 60 Hz
			Low-pass filter	Cutoff of 10 Hz

FIR: finite impulse response, IIR: infinite impulse response,

that airflow signals are specific to the applied sensor and they can change during measurements because of sensor or patient movements.

In addition to these common noises, motion artifacts are the most challenging to deal with in respiratory signals. These noises are both inevitable and unpredictable, arising from the subject's movements during sleep, such as head movement and body shifting (Liu *et al* 2013). Therefore, relying solely on simple filtering techniques with a fixed cutoff frequency proves to be inadequate in effectively addressing this problem. With the reviewed literature, four articles provided comprehensive explanations of motion artifacts processing. Among them, two articles employed adaptive filtering methods that dynamically adjust the cutoff frequency based on specific movement conditions. Keenan and Wilhelm utilized a least mean square (LMS) adaptive filter with a triaxial accelerometer (ACC) signal as reference signals (Keenan and Wilhelm 2005). Similarly, Fedotov *et al* introduced ACC as a reference signal and implemented an adaptive filtering based on the Winer-Hopf RLS algorithm (Fedotov *et al* 2018). However, these methods introduce new reference signals that interfere with respiratory signals. Subsequently, Liu *et al* proposed an algorithm based on mutual information and power criteria for automatically selecting appropriate intrinsic mode functions (IMFs) to remove tissue artifacts and reconstruct respiratory signal reconstruction (Liu *et al* 2013). However, complex algorithms increase computing power costs and hardware consumption. To address this challenge, Rosa *et al* proposed an energy-efficient Haar level 5 (Haar-5) wavelet transform architecture, which can save 38.19% of circuit area and reduce power dissipation and energy by 38.26% compared to other architectures (da Rosa *et al* 2021).

In conclusion, the different noise sources originating from various sensors and experimental environments make it difficult to scientifically evaluate the effectiveness of specific preprocessing methods. Furthermore, limited literature exists in this domain. In practical applications, researchers should analyze and select appropriate signal preprocessing methods based on the measurement technique, noise sources, and energy considerations to achieve optimal results.

3.3. Feature extraction

Through the literature review, respiratory signal features can be broadly classified into three categories: time domain features, frequency domain features, and nonlinear features. Commonly used methods for feature

Table 3. List of extraction methods and features.

No	Author	Year	Feature
1	Gutierrez-Tobal *et al* (2012a)	2012	Mf_{3PSD}^{RRV}, PA^{AF}, BP^{AF}
2	Gutierrez-Tobal *et al* (2013)	2013	CTM^{RRV}, WD^{AF}, M_{PSD}^{RRV}, S^{AF}, M_{PSD}^{AF}, S_{PSD}^{RRV}, $ApEn^{AF}$, CTM^{AF}, LZC^{RRV}
3	Koley and Dey (2013a)	2013	IRA_{MINI}, RPHA, LHP, IRA_M, LTP, LEN
4	Thommandram *et al* (2013)	2013	P2P stability, Peak stability, LP presence, FL
5	Gutierrez-Tobal *et al* (2016)	2015	BW^M, BW^{MA}, BW^{mA}, BW^{SD}, BW^{SpecEn}, BW^{MF}, CTM, LZC, SampEn
6	Avci and Akbas (2015)	2015	Mini, Max, Var, Ave, Energy, Mode, Entropy, S
7	Lin *et al* (2016)	2016	AR, FR, $COV_{TRE-ARE}$
8	Kaimakamis *et al* (2016)	2016	LLE, DFA, ApEn
9	Lakhan *et al* (2018)	2018	Num, Sum, Ave, Max, Mini, M, SD, Var, Weighted
10	Gogus and Tezel (2019)	2019	SF set, AF set, DMF set
11	Hafezi *et al* (2020)	2020	ATM, BDC, IDC, AUIC, SS, AUB, SD
12	Gogus *et al* (2020)	2020	GHE, min hq, VDH, KMS, AIdx, MSmin

PSD: Power spectral density; RRV: respiratory rate variability; Mf_{3PSD}^{RRV}: thirdly statistical moments of the amplitude values of PSDs from RRV; AF: airflow; PA: Peak amplitude; BP: Band power; CTM: Central tendency measure; WD: Wootters distance; M: Mean; S: Skewness; ApEn: approximate entropy; LZC: Lempel–Ziv complexity; IRA: Instantaneous Respiration Amplitude; RPHA: Relative power in HF and area; LHP: Logarithm of HF power; IRA_M: mean IRA; LTP: Logarithm of total power; LEN: Length; P2P stability: Stability of the peak-to-peak time; Peak stability: Stability of the heights of the peaks; LP presence: Presence of long pauses; FL: Flat-lining; BW^M: spectral band of interest mean; BW^{MA}: spectral band of interest maximum amplitude; BW^{mA}: spectral band of interest minimum amplitude; BW^{SD}: spectral band of interest standard deviation; BW^{MF}: spectral band of interest median frequency; BW^k: spectral band of interest kurtosis; SampEn: Sample entropy; Mini: minimum; Max: maximum; Var: variance; Ave: average; AR: Amplitude ratio; FR: Frequency ratio; Cov: covariance; LLE: Largest Lyapunov Exponent; DFA: Detrended Fluctuation Analysis; SF set: Statistics feature set; AF set: Amplitude feature set; DMF set: Descriptive model feature set; ATM: Amount of tracheal movement; BDC: Breathing duty cycle; IDC: Inspiratory duty cycle; AUIC: Area under inspiration curve; SS: Signal slope; AUB: Area under breath; GHE: Generalized Hurst exponent; min hq: Minimum singularity exponent value; VDH: Vertical distance between hqmin and hqmax; KMS: Kurtosis of multifractal spectrum; AIdx: Asymmetric index; MSmin: Multifractal spectrum corresponding to minimum;

extraction include statistical analysis, Fourier transform, and wavelet transform. Among the reviewed literature, 12 articles presented comprehensive explanations of features and extraction methods, which are outlined in table 3. In the subsequent paragraphs, we will provide a detailed exploration of each category, discussing their respective advantages and disadvantages.

(1) Time-domain feature

Time-domain analysis typically involves analyzing statistical properties and morphological features from a set of *N* discrete-time samples within a specified time window (Lutus 2008, Prahallad 2011). Time-domain analysis of respiratory signals and respirator effort signals can yield significant statistical and morphological features, which can be obtained by observing the waveform and performing simple calculations, such as the mean, median, standard deviation, skewness, and kurtosis of the signal. Additionally, morphological features, such as zero-crossings, peaks, valleys, signal slope, and signal amplitude, can also be derived from the waveform. However, due to the nonlinear and non-stationary nature of the respiratory signal and respiratory effort signal, their values vary with time, resulting in differences in their statistical properties at different time points (Fong *et al* 2013). Consequently, it becomes challenging for time-domain features to fully capture the characteristics of the signal, even when utilizing a random time window. To overcome this limitation, other signal analysis techniques like frequency domain analysis can be employed to acquire more detailed and accurate features.

(2) Frequency-domain feature

Frequency-domain analysis offers a comprehensive view into the respiratory system. By decomposing the signal into its frequency components, various features such as the frequency spectrum, spectral density, power spectral density (PSD), and spectral peak can be analyzed to identify the presence and characteristics of respiratory abnormalities, including Apnea and Hypopnea. The primary methods for frequency transformation include Fourier Transform and Wavelet Transform.

Fourier transforms commonly used in extracting frequency domain feature from respiratory signals encompass discrete fourier transform (DFT) (Gutierrez-Tobal *et al* 2012a, 2012b, 2013, Javaid *et al* 2015) and fast Fourier transform algorithm (FFT) (Koley and Dey 2013a, Diaz *et al* 2014, Ciolek *et al* 2015), an enhanced version of DFT. However, DFT assumes that the analyzed signal is stationary and lacks temporal or frequency variations (Sundararajan 2001). To capture the changing frequency characteristics of respiratory signals over time, researchers have started utilizing the short-time Fourier transform (STFT) for time-frequency conversion of respiratory signals (Wu *et al* 2021, Yue *et al* 2021). Additionally, to obtain the PSD feature of the frequency domain, which reflects the recurrent changes in air flow during night respiration (Krishnan and Athavale 2018), researchers such as Tobal *et al* (Gutierrez-Tobal *et al* 2016) and Koley *et al* (Koley and Dey 2013a) utilized DFT

and FFT, respectively, along with the nonparametric estimation Welch's method, which is very suitable for non-stationary signals, to obtain the PSD.

Wavelet transformation is commonly performed using Haar, Symmlet, and Daubechies wavelets to obtain coefficient features for classification purposes. For instance, Romero *et al* used the Symmlet wavelet family (Symmlet of order: O = 7) to obtained coefficients feature and achieved favorable Apnea classification results (Fontenla-Romero *et al* 2005). Maali *et al* decomposed the data using three levels of Haar wavelet transformation to extract wavelet coefficients, which were then used as inputs for an SVM classifier (Maali and Al-Jumaily 2012, Maali and Al-Jumaily 2011). Berdiñas *et al* applied discrete wavelet transformation and computed the average value of 16 wavelet coefficients for SDB detection (Guijarro-Berdiñas *et al* 2012). Avci *et al* employed the Daubechies wavelet to decompose the airflow signal and extract statistical features for classification purposes (Avci and Akbas 2015).

Although these feature extraction methods effectively reveal the frequency domain features of respiratory signals or respiratory movements, SDB is caused by a combination of anatomical upper airway predisposition and changes in neural activation mechanisms (Salisbury and Sun 2007), whose nighttime airflow signals contain rich nonlinear dynamical features which can provide important information for diagnosis and treatment. Therefore, nonlinear features are also commonly considered in feature extraction.

(3) Nonlinear feature

In the realm of nonlinear features, several measures are commonly employed to extract the nonlinear characteristics of respiratory signals, including Central Tendency Measure (CTM), Lempel-Ziv complexity, and Approximate Entropy (ApEn) (Alvarez *et al* 2010, Marcos *et al* 2012, Gutierrez-Tobal *et al* 2012a, 2013, 2016). CTM quantifies the variability degree within a time series (Cohen *et al* 1996), Lempel-Ziv complexity measures complexity in finite sequences (Lempel and Ziv 1976), and ApEn assesses the irregularity of a time series by assigning higher values to higher irregularity (Pincus 1991). The three nonlinear features can essentially serve as representatives of the nonlinear characteristics of respiratory signals.

Furthermore, researchers have also explored diverse combinations of methods to extract complex nonlinear features from respiration signals and respiration movement signals. For example, Kaimakamis *et al* used the monofractal scaling features obtained from detrended fluctuation analysis (DFA), combined with the maximum Lyapunov exponent (LLE), and incorporated them into ApEn to construct a comprehensive complexity index (Kaimakamis *et al* 2016). However, as Vaquerizo-Villar *et al* point out (Vaquerizo-Villar *et al* 2018), apnea and hypopnea events in patients with SDB can generate random spikes and irregular fluctuations in physiological signals. These variations and fluctuations align with a multifractal structure and cannot be entirely characterized by a single fractal provided by conventional DFA. Consequently, (Gogus *et al* 2020). employed multifractal detrended fluctuation analysis (MDFA) as a feature extraction technique on single-channel nasal cannula airflow signals and evaluated its performance using a random forest (RF) classifier.

In summary, feature extraction methods for respiratory signal can be categorized into three main types: time-domain, frequency-domain, and nonlinear features. Time-domain features involve analyzing statistical properties and morphological characteristics within a specific time window using discrete-time samples. Frequency-domain features are obtained by decomposing the signal into its frequency components. Nonlinear features are derived through various nonlinear analysis methods, such as CTM, approximate entropy, LLE, and Lempel-Ziv. It is crucial to emphasize that feature extraction plays a vital role in SDB detection. The extracted features should accurately capture the signal's characteristics to enhance performance with the classification algorithm. Researchers need to determine which feature is more effective by matching different threshold rules or classifiers.

3.4. Classification algorithm

In the systematic review, we found that SDB classification algorithms were primarily based on threshold rules (8), machine learning (13), and deep learning (11). As shown in figure 2(a), early algorithms were predominantly based on threshold rules and machine learning, while after 2017, deep learning became the predominant approach. However, there is no significant difference in the number of points of interest for the recognition targets among the different algorithms, as shown in figure 2(b). In this section, we will provide a detailed overview of the literature survey results focusing on the three algorithms.

(1) Threshold rule-based algorithms

The commonly employed methods for detecting sleep apnea include peak detection and envelope detection. The determination of judgment threshold is based on the regulations provided by the AASM (Iber C *et al* 2017). For instance, Selvaraj utilized piecewise cubic Hermite interpolation to interpolate local maxima and local minima points, allowing for the acquisition of upper and lower envelopes (Selvaraj and Narasimhan 2013). Then, set the threshold by obtaining the envelope width (E) and instantaneous amplitude base frequency variability feature value. MT Bianchi developed an adaptive envelope-tracking function that tracks the

Figure 2. Statistical visualization of the results after data analysis. (a) Comparison of the number of papers published utilizing three different algorithms over the years. (b) Number statistics of identified SDB types by three different algorithms. A: Apnea, H: Hypopnea. AH: the combination of apnea and hypopnea. AHI: Apnea hypopnea index.
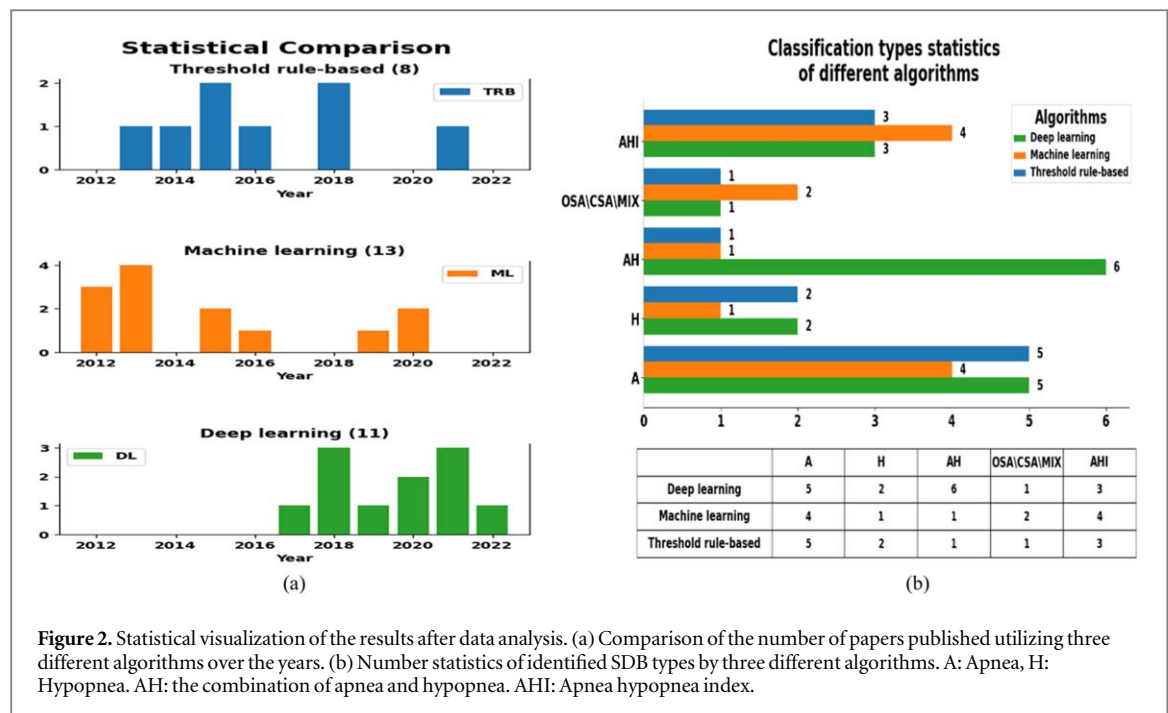
amplitude excursions of thoracic signals (Bianchi *et al* 2014), which adapts to changes in breath size or belt amplitude by defining a new threshold for each breath based on the height of the previous peak. Azimi proposed an algorithm based on an adaptive threshold power transformation of the RIPsum (Azimi *et al* 2018). This algorithm dynamically calculates the threshold as 20% of the highest power within each power segment, considering the 120 s preceding the current power segment. ElMoaqet *et al* proposed a method for continuous monitoring of air flow respiration, which consists of a 600 s baseline window (Wb) and 100 s detection window (Wm) (Kim *et al* 2019). In Wb, the detected amplitudes and time intervals are sorted in descending order to calculate the average value, establishing the baseline. In Wm, the detected amplitudes and time intervals are sorted in ascending order, and the average value is compared with the baseline to detect the occurrence of sleep apnea events. However, the presence of artifacts can lead to sudden changes in the respiratory rhythm, making it challenging to accurately characterize the baseline morphologically (Redline *et al* 2007, Otero *et al* 2011). Thus, Ciołek *et al* proposed an improved approach by replacing linear low-pass FIR filters L(ω)r used in Hilbert transform-based and square-law envelope-based detectors with a cascade of standard median (SM) and recursive median (RM) filters (Ciolek *et al* 2015). This improved approach addresses the issue of envelope distortion and phase shift caused by artifacts, resulting in a more robust envelope detection method.

Previous threshold rule-based literature mainly focuses on the identification of Apnea, even though the study of Hypopnea is mentioned, the Apnea and Hypopnea were still combined in the final identification process. Moreover, there has been limited research on specifically classifying the types of SDB. To address this issue, Lee *et al* introduced a median filter to obtain the AF amplitude and then distinguished A and H based on six rules, and the PPV of Hypopnea reached 65.7 (Lee *et al* 2016). Adha *et al* utilized the obstructive reciprocal divergence (ORD) algorithm, based on RIPsum signal to identify Hypopnea and the accuracy reached 73.34 (Adha and Igasaki 2021). To classify SDB types, Kagawa *et al* proposed a method for classifying SDB types based on the theory of paradoxical breathing motion (Lin *et al* 2016). This method involved using the Pearson correlation method to estimate the phase difference between chest and abdominal movement signals. Subsequently, amplitude drop and phase difference thresholds were utilized to classify OSA, CSA, and MIX (Kagawa *et al* 2016).

Threshold rule-based recognition algorithms provide clear explanations for labeling specific signal periods as containing sleep apnea events or not (Sannino *et al* 2014). These algorithms have demonstrated good performance, as shown in table 4. This white-box approach holds great value in the medical field. However, threshold rule-based algorithms, which rely on comparing simple features and experiment-derived thresholds, overlook the statistical distribution of input features and output categories (ElMoaqet *et al* 2020b). It becomes challenging to establish a universal threshold that can be applied to diverse individuals. To address this limitation, there is a need for more complex algorithms capable of learning from the data and adapting to individual differences, such as machine learning and deep learning.

Table 4. List of threshold rule-based algorithms and results.

Num	Year	Author	Signal	Subjects	Algorithm	Classification type	Results					
							ACC (%)	SPE (%)	SEN (%)	PPV (%)	AUC	CC
1	2013	Selvaraj et al (2013)	NAF\TRE	100	Hermite interpolation envelope detection	A	—	100	83.6	72.3	—	—
2	2014	Bianchi et al (2014)	TRE	116	Adaptive envelope tracking function	AHI	—	—	—	—	0.92 (cutoff = 5) 0.85 (cutoff = 15)	—
3	2015	Ciolek et al (2015)	NAF	309	SM-RM cascade envelope detection	SAH	95.0	90.0	96.0	—	—	—
						AHI	—	—	—	—	—	κ = 0.82
4	2015	Kagawa (Lin et al 2016)	BM from Doppler radars;	35	Time-varying amplitude baseline method	OSA\CSA\MIX	—	—	—	—	—	ICC = 0.94
						Severity	—	96 (cutoff = 15) 90 (cutoff = 30)	100 (cutoff = 15) 79 (cutoff = 30)	—	—	—
5	2016	Lee et al (2016)	NAF	50	Median filter peak detection	A\H	—	—	86.4	84.5	—	—
						AHI	—	—	—	—	—	PCC = 0.94
6	2018	ElMoaqet (Kim et al 2019)	OAF	30	AICPV	A	—	88.7	80.0	—	0.844	—
7	2018	Azimi et al (2018)	RIPsum	25	Adaptive threshold Power transformation	SAHS	59.62	—	72.06	—	—	—
8	2021	Adha and Igasaki (2021)	RIPsum	25	ORD continued excursion model	A	99.83	—	—	—	—	—
						H	73.34	—	—	—	—	—
						AHI	—	—	—	—	—	$R^2 = 0.98$

NAF:AF from nasal, OAF:AF from oronasal, TRE: Thoracic respiratory effort, ARE: Abdominal respiratory effort; OSA: Obstructive sleep apnea; CSA: Central sleep apnea; MSA: Mixed sleep apnea; SM: standard median filters; RM: recursive median filters; AICPV: The abbreviation used in the original text is not explicitly defined. For an explanation of the algorithm employing this abbreviation, please refer to section III, D, lines 20–25 of this paper. Additionally, for a more detailed algorithm explanation, you may consult the original paper; ORD: obstructive reciprocal divergence; AHI: Apnea hypopnea index; AI: Apnea index; HI: hypopnea index; A: Apnea; H: hypopnea; SAHS: sleep apnea-hypopnea syndrome, ACC: Accuracy; SPE: Specificity; SEN: Specificity; PPV: Positive predictive value; AUC: Area under the curve; CC: Correlation coefficient; cutoff: AHI threshold; —: Not mentioned;

(2) Machine learning algorithms

Machine learning algorithms employed in SDB detection discern patterns and relationships within the data for precise classifications. Initiating this process requires the expertise of medical professionals to meticulously annotate respiratory signals, including categories like apnea, hypopnea, normal, OSA, CSA, MIX, etc. Subsequently, algorithmic researchers, specializing in the field, undertake a crucial step known as feature engineering. This meticulous process involves manually extracting pertinent features from labeled data, detailed in section 3.3. In the final stage, algorithmic researchers apply feature selection algorithms meticulously. These selected features are then strategically integrated with machine learning algorithms to optimize results. Diverse machine learning algorithms play a pivotal role in recognizing SDB, including Support Vector Machines (SVM) (Koley and Dey 2013a, 2013b, Lin *et al* 2016), artificial neural networks (ANN) (Guijarro-Berdiñas *et al* 2012), Random Forest (RF) (Avci and Akbas 2015, Gogus and Tezel 2019, Gogus *et al* 2020), Gaussian processes (GP) (ElMoaqet *et al* 2020b), K-Nearest-Neighbors (KNN) (Thommandram et al 2013), and Logistic Regression (LR) (Gutierrez-Tobal *et al* 2012a).

Researchers customize feature combinations for different classifiers to attain optimal classification results. For instance, Berdiñas *et al* (Guijarro-Berdiñas *et al* 2012) utilized the SVM recursive feature elimination (SVM-REF) method to select the wavelet coefficient features of the DWT-transformed TRE signal and input them into the error-correction output code (ECOC) (Dietterich and Bakiri 1994) model of the ANN expert to detect SDB types. Koley *et al* employed SVM-RFE to select the optimal feature subset from the 36 power features extracted from OAF signal and applied it to the three binary SVM classifier to achieve the offline recognition of Apnea, Hypopnea, and Normal (Koley and Dey 2013a). The results of the three classifiers were evaluated using the 'one-against-all strategy' (Kim *et al* 2003) and the 'winner-takes-all rule' (Kim *et al* 2003) was used to determine the final result. Subsequently, online verification was carried out on 8 subjects (Koley and Dey 2013b). In another study, Lin utilized an SVM classifier based on three features of joint signals TRE and ARE, namely amplitude ratio (AR), frequency ratio (FR), and covariance between thoracic and abdominal movements, to achieve a classification of OSA and CSA (Lin *et al* 2016).

Other researchers have also employed different feature selection techniques and classifiers to improve the performance of machine learning algorithms in SDB detection. Avci *et al* and Gogus *et al* utilized feature subset selection methods such as Correlation-based Feature Subset Selection (CfsSubsetEval) (Avci and Akbas 2015), OneR Attribute Eval Feature Selection (OneRAttributeEval) (Gogus and Tezel 2019), and wrapper subset evaluation (WSE) (Gogus *et al* 2020) to select optimal feature sets. These selected features were then utilized in a RF classification model for the detection of Apnea. Thommandram *et al* developed a KNN model for Apnea identification (Thommandram *et al* 2013), based on four clinically observable features: peak-to-peak time stability, peak heights stability, presence of long pauses, and flat lining indication. The best performance was achieved with a *k* value of 44. Furthermore, ElMoaqet *et al* proposed a novel feature, which is the relative changes between the baseline window and the detection window used as feature input, and a Gaussian model was used to recognize the apnea (ElMoaqet *et al* 2020b). This model outperformed the threshold breath pause recognition algorithm proposed by the author in 2018 (Kim *et al* 2019).

Additionally, between 2012 and 2015, Tobal *et al* conducted extensive research on machine learning algorithms for sleep apnea-hypopnea syndrome (SAHS) recognition and apnea-hypopnea index (AHI) estimation based on AF and RRV (respiratory rate variability derived from AF) signal. To achieve optimal results, the researchers selected various feature combinations, including statistical features, spectral features, and nonlinear features, with particularly emphasis on the 'spectral bands of interest' feature. These feature combinations were chosen using techniques such as Forward stepwise logistic regression (FLSR) (Gutierrez-Tobal *et al* 2012a) and Fast correlation-based filter (FCBF) (Gutierrez-Tobal *et al* 2013, 2016). The selected feature combinations were then applied to different classifiers, including LR (Gutierrez-Tobal *et al* 2012a), multiple linear regression (MLR) (Gutierrez-Tobal *et al* 2012b), multi-layer perceptron (MLP) (Gutierrez-Tobal *et al* 2013), as well as adaptive boosting with linear discriminant analysis (AB-LDA) and adaptive boosting with classification and regression trees (AB-CART) (Gutierrez-Tobal *et al* 2016). Through continuous experimentation, the authors achieved an improvement in the recognition rate of SAHS, reaching an 91.5% accuracy, as reported in reference (Gutierrez-Tobal *et al* 2013).

These studies' results mentioned above have demonstrated promising results in the field of SDB algorithms based on machine learning recognition, as presented in table 5. However, machine learning techniques require manual engineering of features and can potentially miss interesting sleep apnea markers in the biometric signals due to human misinterpretations. To address this issue, researchers have increasingly turned to deep learning techniques, which are capable of relying on patterns and inferences to automatically learn the complex relationships between features and labels from data (LeCun *et al* 2015). Such techniques have been applied widely in the analysis of complex medical data (Ravi *et al* 2017).

(3) Deep learning algorithms

The commonly used deep learning technologies in SDB recognition algorithms include convolutional neural networks (CNN) (Haidar *et al* 2017, Choi *et al* 2018, McCloskey *et al* 2018, Wu *et al* 2021, Yue *et al* 2021),

Table 5. List of machine learning algorithms and results.

Num	Year	Author	Signal	Subjects	Algorithm	Feature selection	Classification type	Results				
								ACC (%)	SPE (%)	SEN (%)	AROC	CC
1	2012	Guijarro-Berdiñas et al (2012)	NAF\TRE	6	ANN+ECOC	SVM-RFE	OSA	94.62	—	—	—	—
							CSA	95.47	—	—	—	—
							MSA	90.45	—	—	—	—
2	2012	Gutierrez-Tobal et al (2012a)	NAF\RRV	148	LR	FLSR	SAHS	82.43	—	—	0.90	—
3	2012	Gutierrez-Tobal et al (2012b)	NAF	148	MLR	—	AHI	—	—	—	—	PCC = 0.809
							SAHS	69.7 (cutoff = 5) 75.39(cutoff = 10) 80.9 (cutoff = 15) 87.6 (cutoff = 30)	—	—	—	—
4	2013	Gutierrez-Tobal et al (2013)	NAF\RRV	148	MLP	FCBF	SAHS	91.5	89.5	92.5	—	—
							AHI	—	—	—	—	ICC = 0.849
5	2013	Koley and Dey (2013a)	OAF	56	SVM	SVM-RFE	AI	—	—	—	—	PCC = 0.986
							HI	—	—	—	—	PCC = 0.970
							AHI	—	—	—	—	PCC = 0.982
6	2013	Koley and Dey (2013b)	OAF	8(Online)	SVM	SVM-RFE	A	—	91.8	—	—	—
							H	—	94.9	—	—	—
							AH	—	96.5	—	—	—
7	2013	Thommandram et al (2013)	TRE	8	kNN	—	A	—	88.1	95.7	0.9604	—
8	2015	Gutierrez-Tobal et al (2016)	NAF	317	AB-LDA, AB-CART	FCBF	SAHS	86.5 (cutoff = 5) 86.5 (cutoff = 10) 81.0 (cutoff = 15) 83.3 (cutoff = 30)	—	—	—	—
9	2015	Avci and Akbas (2015)	NAF	8	RF	CfsSubsetEval	A	98.68	—	—	—	—
10	2016	Lin et al (2016)	TRE\ARE	34	SVM	—	OSA\CSA	81.8 ± 9.4	—	—	—	—
11	2019	Gogus and Tezel (2019)	NAF	5- Apnea-ECG MIT-BIH	RF	OneRAttributeEval	A	96.21 / 92.23	—	—	—	$\kappa = 0.92$ / $\kappa = 0.82$
12	2020	Gogus et al (2020)	NAF	120	RF	WSE	OSASH	95.83	93.75	96.88	—	—
							AHI	93.75	—	—	—	$\kappa = 0.91$
13	2020	ElMoaqet et al (2020b)	OAF	96	GMM	—	A	83.4	—	—	—	—

NAF:AF from nasal, OAF:AF from oronasal, TRE: Thoracic respiratory effort, ARE: Abdominal respiratory effort, ANN: Artificial neural networks; ECOC: Error correcting output code; LR: Logistic Regression; MLR: Multivariate linear regression; MLP: Multilayer perceptron; SVM: Support Vector Machines; kNN: K-Nearest-Neighbors; AB-LDA:(AdaBoost-linear discriminant); AB-CART:(AdaBoost-classification and regression trees); GMM: Gaussian Mixture Models; FLSR: Forward stepwise logistic regression; FCBF: Fast correlation-based filter; SVM-RFE: SVM-Recursive Feature Elimination; CfsSubsetEval: Correlation based feature subset selection; OneRAttributeEval: OneR Attribute Eval Feature Selection; WSE: Wrapper Subset Evaluation ; OSA: Obstructive sleep apnea; CSA: Central sleep apnea; MSA: Mixed sleep apnea; AHI: Apnea hypopnea index; AI: Apnea index; HI: hypopnea index; A: Apnea; H: hypopnea; SAHS: sleep apnea-hypopnea syndrome;OSASH: obstructive sleep apnea hypopnea syndrome; ACC: Accuracy; SPE: Specificity; SEN: Specificity; AROC: The area under the receiver operating characteristic curve; CC: Correlation coefficient; cutoff: AHI threshold; —: Not mentioned;

Waveform Neural Network (WaveNet) (Nassi *et al* 2022), Residual network (ResNet) (Wu *et al* 2021, Yue *et al* 2021), residual networks with extremely sparse factorized convolutions (ResNeXt) (Chen *et al* 2020), long short-term memory (LSTM) (Van Steenkiste *et al* 2019, Drzazga and Cyganek 2021), and their related combinations (Hafezi *et al* 2020) or variants (ElMoaqet *et al* 2020a).

A CNN commonly consists of convolutional, pooling, and fully connected layers (LeCun *et al* 2015). The structure for SDB recognition in CNN is conventionally structured into three stages: (1) signal preprocessing, where respiratory signals or respiratory movement-related signals are initially normalized and segmented. (2) Feature extraction, involving the automatic extraction of pertinent features in the convolutional and max-pooling layers of the network. (3) Event classification, wherein SDB events are classified by the fully connected layers. Researchers consistently achieve optimal results by refining the signal preprocessing method and adjusting critical parameters, including the number of convolutional layers, kernel size, number of filters, stride, activation function, and optimizer. For instance, Haidar *et al* utilized 30 filters with a [5×1] kernel size, 5 strides, and a Rectified Linear Unit (ReLU) activation function, concurrently employing the Adam optimizer. This configuration resulted in an accuracy of 74.7% with a 1D CNN on a 30 s non-overlapping NAF signal, surpassing the SVM model's accuracy of 72.0% (Haidar *et al* 2017). In contrast, McCloskey *et al* transformed original respiratory waveform images into wavelet spectrograms, implementing a 2D CNN architecture. This approach yielded an average precision of 79.8% (McCloskey *et al* 2018), marking a 5.3% improvement over the 1D CNN's accuracy of 74.5% (Haidar *et al* 2017). Choi *et al* utilized three convolutional layers and introduced overlapping sliding windows, enhancing the accuracy of SHAS recognition to 94.9% and enabling the estimation of AHI (Choi *et al* 2018).

Differing from conventional CNNs, WaveNet is a fully convolutional neural network that lacks fully connected layers (Oord *et al* 2016). Originally designed for generating audio waveforms, WaveNet utilizes a deep convolutional structure to capture long-term dependencies within audio signals. Given the resemblance between respiratory signal waveforms and audio waveforms, some researchers have leveraged WaveNet for SDB recognition. For example, Nassi enhanced WaveNe's feature extraction capabilities by utilizing its residual blocks to enlarge the receptive field. The architecture of WaveNet was further modified by incorporating non-causal convolutions, allowing the output nodes to depend on both past and future time steps, thereby improving recognition accuracy. Nassi's work demonstrated the effectiveness of WaveNet in recognizing respiratory Apnea-Hypopnea events using a single-effort belt and in multi-classification tasks, including OSA, CSA, and arousal-hypopnea (Nassi *et al* 2022).

There is another type based on the deep convolutional structure, ResNeXt, which adopts the concept of residual learning by introducing residual blocks to tackle issues such as gradient vanishing and exploding during the training of deep neural networks. Researchers have also improved the ResNeXt structure to enhance image feature extraction capabilities for improving SDB recognition rates. For example, Chen *et al* proposed an improved ResNeXt network called Multi-resolution ResNeXt (Mr-ResNeXt), which decomposes images into low-frequency and high-frequency features using octave convolution. Additionally, they upgraded the 3×3 filter in ResNeXt with a new block of multi-level group convolution. This enhancement led to an increase in the recognition accuracy of OSA by nearly 3% (from 91.02% to 94.23%) (Chen *et al* 2020). Wu *et al* introduced octave convolution and attention mechanisms based on the residual network to detect Apnea-Hypopnea, achieving an accuracy of 91.23% (Wu *et al* 2021). Subsequently, Yue *et al* verified the effectiveness of the multi-resolution residual network (Mr-ResNet) on two databases, obtaining satisfactory results (Yue *et al* 2021).

However, most of the neural network architectures based on convolutional structures mentioned above are commonly employed for image recognition, requiring substantial computational power (Urtnasan *et al* 2018). Recent developments in deep learning have introduced an alternative model: long short-term memory (LSTM) neural networks, known for their proficiency in capturing both long-term and short-term dependencies in temporal data. This capability makes LSTM effective in detecting patterns without relying on handcrafted features. In contrast to image-centric architectures, LSTM can significantly reduce the volume of data processing. The typical structure of an LSTM includes a cell (cell state), forget gate, input gate, and an output layer (commonly employing the Softmax or sigmoid function, depending on the functional requirement) (LeCun *et al* 2015). In the SDB recognition process, respiratory signals and movement-related signals undergo normalization and segmentation to obtain the input sequence initially. Subsequently, the cell, input gate, and forget gate play crucial roles in capturing and storing relevant features from the input sequence. The final step involves the output layer, which computes the probability distribution for each SDB event. To enhance recognition rates, researchers often explore adjustments such as varying the number of LSTM layers, activation functions, loss functions, or employing different combinations of LSTMs. For example, Steenkiste *et al* pioneered the application of LSTM to validate the ability to identify apnea based on single-channel respiratory signals—an important step towards a fully automated sleep apnea detection method (Van Steenkiste *et al* 2019), employing the adadelta optimizer. Drzazga *et al* proposed a structure consisting of two LSTM networks neural connected in series, combined with signal preprocessing and filtering, to demonstrate the ability of LSTM to

Table 6. List of deep learning algorithms and results.

Num	Year	Author	Signal	Subjects	Algorithm	Classification type	Results							
							ACC (%)	PRE (%)	SPE (%)	SEN (%)	NPV	RECALL	F1-score (%)	CC
1	2017	Haidar et al (2017)	NAF	100	1D CNN	AH	74.70 ± 1.43	74.50 ± 1.49	—	—	—	74.70 ± 3.75	74.60 ± 2.38	—
2	2018	McCloskey et al (2018)	NAF	1507	2D CNN	N	—	80.4	—	—	—	76.2	78.3	—
						H	—	68.8	—	—	—	73.6	71.1	—
						A	—	90.2	—	—	—	89.2	89.7	—
3	2018	Choi et al (2018)	NAF	179	CNN	AH	96.6	98.5	—	81.1	—	—	—	$\kappa = 0.82$
						SAHS	94.9	—	—	—	—	—	—	—
						AHI	—	—	—	—	—	—	—	PCC = 0.99
4	2018	Van Steenkiste et al (2019)	ARE	2100	LSTM	A	77.2 ± 1.4	—	—	—	91.1 ± 0.4	—	—	—
			TRE				75.0 ± 1.4	—	—	—	91.9 ± 0.4	—	—	—
5	2019	Hafezi et al (2020)	TM	69	CNN+LSTM	A	78.0 (cutoff = 5) 84.0 (cut-off = 15) 88.0 (cutoff = 30)	—	36.0 (cutoff = 5) 87.0 (cut-off = 15) 94.0 (cutoff = 30)	98.0 (cutoff = 5) 81.0 (cut-off = 15) 67.0 (cutoff = 30)	—	—	—	—
6	2020	ElMoaqet et al (2020a)	NAF	17	Bi-LSTM	A	—	—	83.7	90.3	—	—	—	—
7	2020	Chen et al (2020)	NAF	—	Mr-ResNeXt	AH	94.23	92.02	—	—	—	0.95	91.91	—
8	2021	Yue et al (2021)	NAF	17	Mr-ResNet	AH	91.2	—	90.50	90.80	—	—	—	—
9	2021	Wu et al (2021)	NAF	500	ResNet, Octave CNN, Spatial attention mechanism	AH	91.23	—	90.59	90.81	—	—	—	—
						AHI	—	—	—	—	—	—	—	PCC = 0.96

Table 6. (Continued.)

Num	Year	Author	Signal	Subjects	Algorithm	Classification type	Results							
							ACC (%)	PRE (%)	SPE (%)	SEN (%)	NPV	RECALL	F1-score (%)	CC
10	2021	Drzazga and Cyganek (2021)	OAF\TRE\ARE	SHHS1	LSTM	N	86.42	—	—	—	—	—	—	—
						A	68.2	—	—	—	—	—	—	—
						H	49.3	—	—	—	—	—	—	—
				physionet		N	84.35	—	—	—	—	—	—	—
						A	68.2	—	—	—	—	—	—	—
						H	58.28	—	—	—	—	—	—	—
11	2022	Nassi *et al* (2022)	ARE	9656	WaveNet	AH	—	65.0	98.0	68.0	—	0.71	67	—
						AHI	—	—	—	—	—	—	—	$R^2 = 0.98$
						OSA	51.0	—	—	—	—	—	—	—
						CSA	84.0	—	—	—	—	—	—	—
						Arousals-Hypopnea	40.0	—	—	—	—	—	—	—

NAF:AF from nasal, OAF:AF from oronasal, TRE: Thoracic respiratory effort; ARE: Abdominal respiratory effort; D: Dimension; CNN: Convolutional neural network;
LSTM: Long Short-Term Memory; Bi-LSTM: Bidirectional, ACC: Accuracy, SPE: Specificity, SEN: Specificity, NPV: Negative predictive value, AUC: Area under the curve;
CC: Correlation coefficient, ResNet: Residual network, cutoff: AHI threshold; —: Not mentioned;

recognize Apnea and Hypopnea based on the joint signal of OAF, TRE, and ARE (Drzazga and Cyganek 2021), utilizing the Adam optimizer. ElMoaqet *et al* utilized variants of LSTM (Bi-LSTM) to detect Apnea based on NAF, OAF and ARE signals respectively, and the results showed that the detection accuracy based on NAF was higher (ElMoaqet *et al* 2020a), employing the Adam optimizer. Hafezi *et al* combined the capability of CNN in extracting robust features (Krizhevsky *et al* 2017) with the proficiency of LSTM in encoding relative information across temporal measurements (Sepp Hochreiter 1997) to form a CNN+LSTM model. The RMSProp optimizer was employed for the recognition of Apnea (Hafezi *et al* 2020), yielding satisfactory results. Table 6 provides an overview of algorithms and results based on deep learning for recognizing SDB.

It is worth noting that utilizing DL techniques to deal with imbalanced data classification pose significant challenges (Buda *et al* 2018, Johnson and Khoshgoftaar 2019). Imbalanced datasets may bias the model towards the majority class, resulting in skewed outcomes. To address this issue, researchers have employed various algorithms to handle imbalanced SDB data. The most straightforward solutions include downsampling (Vluymans 2019, Singh and Mishra 2022), oversampling (ElMoaqet *et al* 2020a), and subsampling (Choi *et al* 2018, Lakhan *et al* 2018). However, these methods may result in the loss of valuable information and the risk of overfitting, thereby impacting the generalization ability of the new data. To overcome these disadvantages, an innovative procedure called balanced bootstrapping has been proposed (Wallace *et al* 2011). This method has been adopted in subsequent recognition algorithms to address the imbalance problem in SDB data (Van Steenkiste *et al* 2019, Nassi *et al* 2022).

4. Perspective

Due to variations in data collection methods, data volume, database sources, and evaluation criteria across the 32 papers identified in our systematic review, accurately determining the superiority of one algorithm over another is not feasible. However, it is possible to discuss the overall trends of various algorithms and their respective areas of expertise, analyze the common issues present in current algorithms, and explore potential research directions that may emerge in the future.

4.1. Trend of algorithm development

Early research predominantly relied on threshold rule-based algorithms, which provided a user-friendly and accepted approach for medical personnel. These algorithms can clearly explain why signals at certain moments are marked as containing sleep apnea events or not (Sannino *et al* 2014), which is more user-friendly and accepted for medical personnel. However, classical threshold rule-based detectors relied on experimentally derived thresholds and features, which posed challenges in establishing a generalized standard due to individual variations in physiological signals. Furthermore, these algorithms had limited applicability to small datasets and struggled to extract input features and statistical distributions effectively (ElMoaqet *et al* 2020b). The reviewed literature primarily focused on datasets of fewer than 120 people (Bianchi *et al* 2014).

In contrast, machine learning has emerged as a solution to address challenges mentioned above. Its strong learning ability, capacity to automatically detect complex patterns, higher model accuracy, and greater flexibility have made it indispensable in the recognition of SDB algorithms at all stages. Machine learning algorithms constitute a substantial proportion of the reviewed literature, accounting for 40.6%. However, the core of machine learning lies in feature selection and model selection to achieve optimal classification performance (Jordan and Mitchell 2015, Mahesh 2020), which also leads to some limitations. Firstly, researchers need to possess strong prior knowledge and a comprehensive understanding of the signals to extract effective feature sets, which can be challenging for non-medical personnel without expertise in SDB pathology. Secondly, the manually extracting of features for data acquisition is costly and subject to subjective influences. Additionally, machine learning algorithms have limited ability to capture temporal information, posing challenges when analyzing complex long-term vital sign signals like respiratory signals. Thus, there are inherent limitations to using machine learning for extracting hidden pathologic information from respiratory signals.

Since 2017, deep learning-based SDB recognition algorithms have gradually emerged as the dominant trend, primarily due to their ability to automatically extract features and recognize hidden information from time series data. In practical applications, researchers often employ CNN models, renowned for their prowess in image processing, and LSTM models, which excel at processing time series data. CNN technology primarily involves transforming respiratory signals into images through techniques such as Wavelet Transform and Fast Fourier transform, followed by enhancing features using various image processing techniques to improve SDB recognition accuracy. On the other hand, LSTM models focus on improving the model's ability to learn temporal features, thereby enhancing the accuracy of SDB recognition. To further improve recognition accuracy, researchers have also adopted a combined approach utilizing both CNN and LSTM models, leveraging the advantages offered by each (Hafezi *et al* 2020).

In summary, although our analysis did not reveal clear differences or trends among the various methods of signal acquisition, signal processing, and feature extraction, a significant trend emerged when examining classification models. This trend underscores the growing importance of deep learning in the field of SDB recognition. Its impact on SDB recognition is expected to persist until more advanced models emerge.

4.2. Techniques to enhance model performance

Although conclusive statistical evidence regarding the superiority of specific algorithms is lacking, there are effective techniques available to enhance the performance of SDB detection models. In terms of signal source, algorithms can be categorized into three categories: NAF, OAF, and respiratory movement-related signals (ARE, TRE). A comparison of algorithms based on different signal sources reveals that the overall recognition algorithm utilizing NAF signals achieves the highest accuracy at 94.11%, whereas the average accuracy for the other two signal types is only 78.19% (it is important to note that this comparison considers only the recognition results of the apnea algorithm, as different papers employ distinct classification targets and evaluation indicators). This statistical finding is consistent with the research result in Avci and Akbas (2015), which further emphasizes the significant impact of the signal source selection on model accuracy.

From a functional perspective, a simple respiratory signal alone can only provide recognition of apnea, hypopnea, or AHI, which possesses limited diagnostic capabilities. However, for a comprehensive SDB diagnosis, it is necessary not only to recognize apnea or hypopnea, but also to diagnose the type of SDB, such as OSA, MSA and MIX, to formulate effective treatment plans. In cases of OSA obstruction events, the direction of thoracic motion is opposite to the direction of abdominal motion when the subject tries to breathe, presenting a unique contradictory motion pattern (Tobin *et al* 1983, Staats *et al* 1984, Farre *et al* 2004). In CSA, when an obstruction occurs, the brain cannot produce or transmit signals to control respiratory muscles, leading to a complete cessation of breathing for a brief period (Ramachandran and Karuppiah 2021). Both situations will occur in the MIX pattern. Therefore, TRE and ARE can be added as two additional signals to increase the ability to identify SDB types (Lin *et al* 2016), thus expanding the model's function.

When dealing with SDB data, it is crucial to address the issue of data imbalance. This is primarily due to the limited positive instances of respiratory arrest for each patient, leading to an imbalanced dataset. When machine learning or deep learning models are trained using such datasets, they tend to exhibit a strong bias towards the majority class, resulting in skewed outcomes. Common methods for dealing with imbalanced data include Subsample, Bootstrapping, and Oversampling. In practical applications, researchers can use multiple techniques and evaluation metrics such as accuracy and specificity to compare the performance of models in multiple dimensions to identify the best-performing model. It is important to note that when dealing with imbalanced data, precision and recall (Hafezi *et al* 2020) should be added to evaluate the performance of the classifier.

4.3. Problems to be Improved and Solved

In general, algorithms relying on a single respiratory signal source are preferred due to their simpler hardware implementation and lower costs. However, the standard definition of a hypopnea event by the AASM requires a decrease in respiratory airflow of over 30% for at least 10 s, accompanied by a decrease in oxygen saturation of over 3% or an arousal (Iber C *et al* 2017). This definition poses limitations on detecting hypopnea using a single respiratory signal source. Most algorithms reviewed in the literature combined apnea and hypopnea into one event, and even those that recognized hypopnea have shown unsatisfactory overall performance. For instance, the highest recognition rate reported involved only 25 subjects at 73.34% (Adha and Igasaki 2021), while the lowest recognition rate was a mere 49.3% (Drzazga and Cyganek 2021). Currently, a common practice to improve hypopnea recognition rates is to incorporate blood oxygen signal sources, but this necessitates additional hardware resources. Hence, there is a need to develop algorithms solely based on a single respiratory signal source to enhance hypopnea recognition accuracy, which should be a new research direction.

Nowadays, deep learning algorithms are widely utilized in the medical and healthcare fields due to their ability to automatically extract features, capture long-term temporal dependencies, and handle large and unstructured datasets (Ravi *et al* 2017). However, these algorithms are often treated as a highly complex 'black box' system, making it challenging to explain their results, access the underlying mechanics, or make modifications in case of misclassifications. This lack of interpretability is not user-friendly for medical personnel who are not well-versed in algorithmic fields. Consequently, developing a new SDB classification model that is interpretable and human-machine coexistent will be a new research focus.

Over the years, SDB recognition technology has made significant advancements, transitioning from the initial gold standard of multi-lead PSG to various single-lead detection technologies, accompanied by cost reductions. However, precise measurement of respiratory signals still necessitates the installation of sensors at the mouth and nose, which must be measured continuously throughout the night. This approach is

inconvenient for patients and affects their compliance. With the emergence of new materials and the development of the Internet of Things, several contactless respiratory signal acquisition methods have been explored, such as infrared optical gas imaging (An *et al* 2022), radar (Javaid *et al* 2015, Kagawa *et al* 2016), and mattress of pressure sensor (Yizraeli Davidovich *et al* 2016). However, they are still in the early stages of research and are not yet applied clinically. Thus, developing a reliable contactless respiratory signal recognition algorithm for clinical applications is a promising research direction for the future.

5. Conclusion

This article presents a comprehensive investigation into the literature on the SDB detection algorithms based on respiratory and respiratory movement-related signals from 2012 to 2022. From the selected 32 articles, a detailed analysis and discussion were conducted, focusing on four main aspects: respiratory signal acquisition, signal processing, feature extraction, and classification algorithm.

The research findings indicate that NAF obtained from airflow thermal or thermal sensors demonstrates superior accuracy, making it the preferred signal for studying respiratory signal algorithms for majority of researchers. Additionally, due to the presence of respiratory paradoxical movement, ARE and TRE obtained through RIP belts or PVDF belts can serve as signals for recognizing different types of SDB. Regarding signal processing, the literature commonly employs techniques such as band-pass, median, notching, IIR, and FIR filters. However, given the substantial inter-individual variations in respiratory signals and the presence of artifacts, researchers should consider exploring adaptive filtering or more sophisticated algorithms to improve signal quality. Feature extraction is of paramount importance in rule-based and machine learning algorithms as it enables enhanced accuracy through use of suitable feature selection methods and algorithmic matching. Researchers are advised to conduct comparative experiments and carefully select suitable features and matching algorithms. Within the domain of classification models, deep learning has emerged as the predominant approach owing to its automatic feature extraction capabilities and superior learning performance. However, researchers must be aware of the 'black box' effect inherent in deep learning and strive to mitigate it through algorithmic enhancements or novel models.

In summary, this article provides a comprehensive overview of the SDB algorithm based on respiratory and respiratory-related movement signals. It aspires to serve as a quick reference guide for novice researchers and provide valuable insights for experienced researchers to help them adjust algorithms. Moreover, the article identifies and analyzes the current unsolved problems of the SDB algorithm, with the hope of helping researchers in formulating future research topics and developing new algorithms.

Acknowledgments

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary information files).

ORCID iDs

Liqing Yang ⊙ https://orcid.org/0000-0003-4467-8238
Xing Wang ⊙ https://orcid.org/0000-0002-0187-7408
Lin Chen ⊙ https://orcid.org/0000-0003-1397-6769

References

Adha M S and Igasaki T 2021 Three-stage breathing effort quantification for obstructive sleep apnea detection based on thoracic and abdominal movement signals *IEEE Access* **9** 72781–92
Alsubie H S and BaHammam A S 2017 Obstructive sleep apnoea: children are not little adults *Paediatric Respiratory Rev.* **21** 72–9

Alvarez D, Hornero R, Marcos J V and del Campo F 2010 Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis *IEEE Trans. Biomed. Eng.* **57** 2816–24

Alvarez-Estevez D and Moret-Bonillo V 2015 Computer-assisted diagnosis of the sleep apnea-hypopnea syndrome: a review *Sleep Disord.* **2015** 237878

An J Y, Shin H J, Yang M, Park D Y, Yang J and Kim H J 2022 'Non-contact diagnosis of sleep breathing disorders using infrared optical gas imaging: a prospective observational study *Sci. Rep.* **12** 21052

Avci C and Akbas A 2015 Sleep apnea classification based on respiration signals by using ensemble methods *Biomed. Mater. Eng.* **26** S1703–10

Azimi H, Gilakjani S S and Bouchard M 2018 Automatic apnea-hypopnea events detection using an alternative sensor *IEEE Sensors Applications Symposium (SAS)* pp 1–5

Benjafield A V *et al* 2019 Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis *Lancet Respir Med.* **7** 687–98

Bianchi M T, Lipoma T, Darling C, Alameddine Y and Westover M B 2014 Automated sleep apnea quantification based on respiratory movement *Int. J. Med. Sci.* **11** 796–802

Buda M, Maki A and Mazurowski M A 2018 A systematic study of the class imbalance problem in convolutional neural networks *Neural Netw.* **106** 249–59

Chen Q *et al* 2020 Mr-ResNeXt: a multi-resolution network architecture for detection of obstructive sleep apnea Singapore *Neural Computing for Advanced Applications: First International Conference, NCAA 2020 (Shenzhen, China, July 3–5, 2020)* pp 420–32

Choi S H, Yoon H, Kim H S, Kim H B, Kwon H B, Oh S M, Lee Y J and Park K S 2018 Real-time apnea-hypopnea event detection during sleep by convolutional neural networks *Comput. Biol. Med.* **100** 123–31

Ciolek M, Niedzwiecki M, Sieklicki S, Drozdowski J and Siebert J 2015 Automated detection of sleep apnea and hypopnea events based on robust airflow envelope tracking in the presence of breathing artifacts *IEEE J. Biomed. Health Inform.* **19** 418–29

Cohen M E, Hudson D L and Deedwania P C 1996 Applying continuous chaotic modeling to cardiac signal analysis *IEEE Eng. Med. Biol. Mag.* **15** 97–102

Coronel C 2020 Comparison of PSG signals and respiratory movement signal via 3D camera in detecting sleep respiratory events by LSTM models *2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)* pp 919–23

Coronel C *et al* 2019 Measurement of respiratory effort in sleep by 3D camera and respiratory inductance plethysmography *Somnologie* **23** 86–92

Crowley K E, Rajaratnam S M, Shea S A, Epstein L J, Czeisler C A, Lockley S W, Harvard Work Hours H and Safety G 2013 Evaluation of a single-channel nasal pressure device to assess obstructive sleep apnea risk in laboratory and home environments *J. Clin. Sleep. Med.* **9** 109–16

Daulatzai M A 2015 Evidence of neurodegeneration in obstructive sleep apnea: relationship between obstructive sleep apnea and cognitive dysfunction in the elderly *J. Neurosci. Res.* **93** 1778–94

Deviaene M, Borzee P, van Gilst M, van Dijk J, Overeem S, Buyse B, Testelmans D, Van Huffel S and Varon C 2020 Multilevel interval coded scoring to assess the cardiovascular status of sleep apnea patients using oxygen saturation markers *IEEE Trans. Biomed. Eng.* **67** 2839–48

Diaz J A, Arancibia J M, Bassi A and Vivaldi E A 2014 Envelope analysis of the airflow signal to improve polysomnographic assessment of sleep disordered breathing *Sleep* **37** 199–208

Dietterich T G and Bakiri G 1994 Solving multiclass learning problems via error-correcting output codes *J. Artif. Intell. Res.* **2** 263–86

Drzazga J and Cyganek B 2021 An LSTM network for apnea and hypopnea episodes detection in respiratory signals *Sensors* **21** 5858

ElMoaqet H, Eid M, Glos M, Ryalat M and Penzel T 2020a Deep recurrent neural networks for automatic detection of sleep apnea from single channel respiration signals *Sensors* **20** 5037

ElMoaqet H, Kim J, Tilbury D, Ramachandran S K, Ryalat M and Chu C-H 2020b Gaussian mixture models for detecting sleep apnea events using single oronasal airflow record *Appl. Sci.* **10** 7889

Farre R, Montserrat J and Navajas D 2004 Noninvasive monitoring of respiratory mechanics during sleep *Eur. Respiratory J.* **24** 1052–60

Faust O, Acharya U R, Ng E Y K and Fujita H 2016 A review of Ecg-based diagnosis support systems for obstructive sleep apnea *J. Mech. Med. Biol.* **16** 1640004

Fedotov A A, Akulov S A and Akulova A S 2018 Motion artifacts reduction in wearable respiratory monitoring device *Embec & Nbc 2017 IFMBE Proc.* pp 1121–4

Floras J S 2015 Hypertension and sleep apnea *Can. J. Cardiol.* **31** 889–97

Fong S, Lan K, Sun P, Mohammed S and Fiaidhi J 2013 A time-series pre-processing methodology for biosignal classification using statistical feature extraction *The Tenth IASTED International Conference on Biomedical Engineering (Biomed 2013)* (https://doi.org/10.2316/P.2013.791-100)

Fontenla-Romero O, Guijarro-Berdinas B, Alonso-Betanzos A and Moret-Bonillo V 2005 A new method for sleep apnea classification using wavelets and feedforward neural networks *Artif. Intell. Med.* **34** 65–76

Gogus F Z and Tezel G 2019 Apneic events detection using different features of airflow signals *Mehran Univ. Res. J. Eng. Technol.* **38** 1–16

Gogus F Z, Tezel G, Özşen S, Küççüktürk S, Vatansev H and Koca Y 2020 Identification of apnea-hypopnea index subgroups based on multifractal detrended fluctuation analysis and nasal cannula airflow signals *Trait. Signal* **37** 145–56

Graco M *et al* 2018 Diagnostic accuracy of a two-stage model for detecting obstructive sleep apnoea in chronic tetraplegia *Thorax* **73** 864–71

Guijarro-Berdiñas B, Hernández-Pereira E and Peteiro-Barral D 2012 A mixture of experts for classifying sleep apneas *Expert Syst. Appl.* **39** 7084–92

Gutierrez-Tobal G C, Alvarez D, Del Campo F and Hornero R 2016 Utility of adaboost to detect sleep apnea-hypopnea syndrome from single-channel airflow *IEEE Trans. Biomed. Eng.* **63** 636–46

Gutierrez-Tobal G C, Alvarez D, Marcos J V, del Campo F and Hornero R 2013 Pattern recognition in airflow recordings to assist in the sleep apnoea-hypopnoea syndrome diagnosis *Med. Biol. Eng. Comput.* **51** 1367–80

Gutierrez-Tobal G C, Hornero R, Alvarez D, Marcos J V and del Campo F 2012a Linear and nonlinear analysis of airflow recordings to help in sleep apnoea-hypopnoea syndrome diagnosis *Physiol. Meas.* **33** 1261–75

Gutierrez-Tobal G C, Hornero R, Alvarez D, Marcos J V, Gomez C and Del Campo F 2012b Apnea-hypopnea index estimation from spectral analysis of airflow recordings *IEEE 2012 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* vol 2012, pp 3444–7

Hafezi M, Montazeri N, Saha S, Zhu K, Gavrilovic B, Yadollahi A and Taati B 2020 Sleep Apnea Severity Estimation using a Deep Learning Model from Tracheal Movements *IEEE Access* **8** 22641–9

Haidar R, Koprinska I and Jeffries B 2017 Sleep apnea event detection from nasal airflow using convolutional neural networks *International Conference on Neural Information Processing, Neural Inf. Process. Lecture Notes Comput. Sci.* 83, 819–27

Hu X, Sun J, Dong J and Zhang X 2022 Auditory receptive field net based automatic snore detection for wearable devices *IEEE J. Biomed. Health Inform.* **27** 2255–63

Iber C A-I S, Chesson A and (Quan for the American Academy of Sleep Medicine) 2017 The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications *American Academy of Sleep Medicine* **V 2.4.0** 665

Jaiswal S J, Owens R L and Malhotra A 2017 Raising awareness about sleep disorders *Lung. India* **34** 262–8

Javaid A Q, Noble C M, Rosenberg R and Weitnauer M A 2015 Towards sleep apnea screening with an under-the-mattress IR-UWB radar using machine learning *IEEE 14th Int. Conf. on Machine Learning and Applications (ICMLA)* pp 837–42

Jin J and Sanchez-Sinencio E 2015 A home sleep apnea screening device with time-domain signal processing and autonomous scoring capability *IEEE Trans. Biomed. Circuits Syst.* **9** 96–104

Johnson J M and Khoshgoftaar T M 2019 Survey on deep learning with class imbalance *J. Big Data* **6** 2

Jordan M I and Mitchell T M 2015 'Machine learning: trends, perspectives, and prospects *Science* **349** 255–60

Ju Y E, McLeland J S, Toedebusch C D, Xiong C, Fagan A M, Duntley S P, Morris J C and Holtzman D M 2013 Sleep quality and preclinical Alzheimer disease *JAMA Neurol.* **70** 587–93

Kagawa M, Tojima H and Matsui T 2016 Non-contact diagnostic system for sleep apnea-hypopnea syndrome based on amplitude and phase analysis of thoracic and abdominal Doppler radars *Med. Biol. Eng. Comput.* **54** 789–98

Kaimakamis E, Tsara V, Bratsas C, Sichletidis L, Karvounis C and Maglaveras N 2016 Evaluation of a decision support system for obstructive sleep apnea with nonlinear analysis of respiratory signals *PLoS One* **11** e0150163

Kang S, Kim D K, Lee Y, Lim Y H, Park H K, Cho S H and Cho S H 2020 Non-contact diagnosis of obstructive sleep apnea using impulse-radio ultra-wideband radar *Sci. Rep.* **10** 5261

Kaniusas E and Kaniusas E 2012 *Physiological Phenomena and Biosignals* (Springer)

Keenan D B and Wilhelm F H 2005 Adaptive and wavelet filtering methods for improving accuracy of respiratory measurement *Biomed. Sci. Instrum.* **41** 37–42

Kim H-C, Pang S, Je H-M, Kim D and Yang Bang S 2003 Constructing support vector machine ensemble, *Pattern Recognit.* **36** 2757–67

Kim J, ElMoaqet H, Tilbury D M, Ramachandran S K and Penzel T 2019 Time domain characterization for sleep apnea in oronasal airflow signal: a dynamic threshold classification approach *Physiol. Meas.* **40** 054007

Kljajić Z, Roje Ž, Bečić K, Čapkun V, Vilović K, Ivanišević P and Marušić E 2017 Formula for the prediction of apnea/hypopnea index in children with obstructive sleep apnea without polysomnography according to the clinical parameters: Is it reliable? *Int. J. Pediatric Otorhinolaryngol.* **100** 168–73

Koley B L and Dey D 2013a Automatic detection of sleep apnea and hypopnea events from single channel measurement of respiration signal employing ensemble binary SVM classifiers *Measurement* **46** 2082–92

Koley B L and Dey D 2013b Real-time adaptive apnea and hypopnea event detection methodology for portable sleep apnea monitoring devices *IEEE Trans. Biomed. Eng.* **60** 3354–63

Krishnan S and Athavale Y 2018 Trends in biomedical signal feature extraction *Biomed. Signal. Process. Control.* **43** 41–63

Krizhevsky A, Sutskever I and Hinton G E 2017 Imagenet classification with deep convolutional neural networks *Commun. ACM* **60** 84–90

Lakhan P, Ditthapron A and Banluesombatkul N 2018 Deep neural networks with weighted averaged overnight airflow features for sleep apnea-hypopnea severity classification *IEEE TENCON 2018 - 2018 IEEE Region 10 Conference* (Korea (South)) pp 0441–5

LeCun Y, Bengio Y and Hinton G 2015 Deep learning *Nature* **521** 436–44

Lee H, Park J, Kim H and Lee K J 2016 New rule-based algorithm for real-time detecting sleep apnea and hypopnea events using a nasal pressure signal *J. Med. Syst.* **40** 282

Lee Y S, Pathirana P N, Steinfort C L and Caelli T 2014 Monitoring and analysis of respiratory patterns using microwave doppler radar *IEEE J. Transl. Eng. Health Med.* **2** 1800912

Lempel A and Ziv J 1976 On the complexity of finite sequences *IEEE Trans. Inf. Theory* **22** 75–81

Łęski J M and Henzel N 2005 ECG baseline wander and powerline interference reduction using nonlinear filter bank *Signal Process.* **85** 781–93

Liberati A, Altman D G, Tetzlaff J, Mulrow C, Gotzsche P C, Ioannidis J P, Clarke M, Devereaux P J, Kleijnen J and Moher D 2009 The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration *PLoS Med.* **6** e1000100

Liguori C, Maestri M, Spanetta M, Placidi F, Bonanni E, Mercuri N B and Guarnieri B 2021 Sleep-disordered breathing and the risk of Alzheimer's disease *Sleep Med. Rev.* **55** 101375

Liguori C M N *et al* 2016 Obstructive sleep apnoea as a risk factor for osteopenia and osteoporosis in the male population *Eur. Respir. J.* **47** 987–90

Lin Y Y, Wu H T, Hsu C A, Huang P C, Huang Y H and Lo Y L 2016 Sleep apnea detection based on thoracic and abdominal movement signals of wearable piezo-electric bands *IEEE J. Biomed. Health Inform.* **21** 1533–45

Liu S, Gao R X, John D, Staudenmayer J and Freedson P 2013 Tissue artifact removal from respiratory signals based on empirical mode decomposition *Ann. Biomed. Eng.* **41** 1003–15

Lutus P 2008 *Signal Processing Workshop* https://arachnoid.com/signal_processing/

Maali Y and Al-Jumaily A 2012 Hierarchical parallel PSO-SVM based subject-independent sleep apnea classification *Neural Inf. Process.* 500–7

Maali Y and Al-Jumaily A 2011 Automated detecting sleep apnea syndrome: a novel system based on genetic SVM *IEEE 2011 11th International Conference on Hybrid Intelligent Systems (HIS 2011)* (Melacca, Malaysia) pp 590–4

Mahesh B 2020 Machine learning algorithms-a review *Int. J. Sci. Res. (IJSR)* **9** 381–6

Makarie Rofail L W K, Unger G, Marks G B and Grunstein R R 2010 The role of single-channel nasal airflow pressure transducer in the diagnosis of OSA in the sleep laboratory *J. Clin. Sleep. Med.* **6** 349–56

Marcos J V, Hornero R, Alvarez D, Aboy M and Del Campo F 2012 Automated prediction of the apnea-hypopnea index from nocturnal oximetry recordings *IEEE Trans. Biomed. Eng.* **59** 141–9

Masa J F *et al* 2011 Effectiveness of home respiratory polygraphy for the diagnosis of sleep apnoea and hypopnoea syndrome *Thorax* **66** 567–73

Masa J F *et al* 2014 Effectiveness of home single-channel nasal pressure for sleep apnea diagnosis *Sleep* **37** 1953–61

McCloskey S, Haidar R, Koprinska I and Jeffries B 2018 Detecting hypopnea and obstructive apnea events using convolutional neural networks on wavelet spectrograms of nasal airflow *Advances in Knowledge Discovery and Data Mining, Lecture Notes in Computer Science* pp 361–72

Mendonca F, Mostafa S S, Ravelo-Garcia A G, Morgado-Dias F and Penzel T 2019 A review of obstructive sleep apnea detection approaches *IEEE J. Biomed. Health. Inform.* **23** 825–37

Morgenstern C, Randerath W J, Schwaibold M, Bolz A and Jane R 2013 Feasibility of noninvasive single-channel automated differentiation of obstructive and central hypopneas with nasal airflow *Respiration* **85** 312–8

Nakano H, Tanigawa T, Furukawa T and Nishima S 2007 Automatic detection of sleep-disordered breathing from a single-channel airflow record *Eur. Respir J.* **29** 728–36

Nassi T E, Ganglberger W, Sun H, Bucklin A A, Biswal S, van Putten M, Thomas R J and Westover M B 2022 Automated scoring of respiratory events in sleep with a single effort belt and deep neural networks *IEEE Trans. Biomed. Eng.* **69** 2094–104

Oord A *et al* 2016 *Cornell University* Wavenet: a generative model for raw audio https://doi.org/10.48550/arXiv.1609.03499

Otero A, Félix P and Álvarez M R 2011 Algorithms for the analysis of polysomnographic recordings with customizable criteria *Expert Syst. Appl.* **38** 10133–46

Park B K, Yamada S and Boric-Lubecke O 2006 Single-channel receiver limitations in Doppler radar measurements of periodic motion *IEEE 2006 IEEE Radio and Wireless Symposium (San Diego, CA, USA)* pp 99–102

Park J Y, Lee Y, Choi Y W, Heo R, Park H K, Cho S H, Cho S H and Lim Y H 2019 Preclinical evaluation of a noncontact simultaneous monitoring method for respiration and carotid pulsation using impulse-radio ultra-wideband radar, *Sci. Rep.* **9** 11892

Pincus S M 1991 Approximate entropy as a measure of system complexity *Proc. Natl. Acad. Sci.* **88** 2297–301

Polsek D, Gildeh N, Cash D, Winsky-Sommerer R, Williams S C R, Turkheimer F, Leschziner G D, Morrell M J and Rosenzweig I 2018 Obstructive sleep apnoea and Alzheimer's disease: In search of shared pathomechanisms *Neurosci. Biobehav. Rev.* **86** 142–9

Prahallad K 2011 Feature extraction in time and frequency domain https://archive.org/details/FeatureExtractionInTimeAndFrequencyDomain

Ramachandran A and Karuppiah A 2021 A survey on recent advances in machine learning based sleep apnea detection systems *Healthcare* **9** 914

Ravi D, Wong C, Deligianni F, Berthelot M, Andreu-Perez J, Lo B and Yang G Z 2017 Deep learning for health informatics *IEEE J. Biomed. Health Inform.* **21** 4–21

Redline S *et al* 2007 The scoring of respiratory events in sleep: reliability and validity *J. Clin. Sleep Med.* **03** 169–200

Rethlefsen M L, Kirtley S, Waffenschmidt S, Ayala A P, Moher D, Page M J, Koffel J B and P-S Group 2021 PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews *Syst. Rev.* **10** 39

Reutrakul S and Mokhlesi B 2017 Obstructive sleep apnea and diabetes: a state of the art review *Chest* **152** 1070–86

da Rosa M M, Seidel H B, Paim G, da Costa E A C, Almeida S and Bampi S 2021 An energy-efficient haar wavelet transform architecture for respiratory signal processing *IEEE Trans. Circuits Syst. Express Briefs.* **68** 597–601

Ryan S 2017 Adipose tissue inflammation by intermittent hypoxia: mechanistic link between obstructive sleep apnoea and metabolic dysfunction *J. Physiol.* **595** 2423–30

Salisbury J I and Sun Y 2007 Rapid screening test for sleep apnea using a nonlinear and nonstationary signal processing technique *Med. Eng. Phys.* **29** 336–43

Sankar A B, Kumar D and Seethalakshmi K 2010 Performance study of various adaptive filter algorithms for noise cancellation in respiratory signals *Signal Process.: Int. J. (SPIJ)* **4** 268–78

Sannino G, De F I and De P G 2014 An automatic rules extraction approach to support OSA events detection in an mhealth system *IEEE J. Biomed. Health Inform.* **18** 1518–24

Selvaraj N and Narasimhan R 2013 Detection of sleep apnea on a per-second basis using respiratory signals *IEEE 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp 2124–7

Sepp Hochreiter J S 1997 Long short-term memory *Neural Comput.* **1735–1780** 9

Serrano Alarcon A, Martinez Madrid N and Seepold R 2021 A minimum set of physiological parameters to diagnose obstructive sleep apnea syndrome using non-invasive portable monitors. A systematic review *Life* **11** 1249

Sharma S and Culebras A 2016 Sleep apnoea and stroke *Stroke Vasc. Neurol.* **1** 185–91

Shen Q, Qin H, Wei K and Liu G 2021 Multiscale deep neural network for obstructive sleep apnea detection using rr interval from single-lead ECG signal *IEEE Trans. Instrum. Meas.* **70** 1–13

Shi L, Chen S J, Ma M Y, Bao Y P, Han Y, Wang Y M, Shi J, Vitiello M V and Lu L 2018 Sleep disturbances increase the risk of dementia: a systematic review and meta-analysis *Sleep. Med. Rev.* **40** 4–16

Singh R S and Mishra A K 2022 Review of methods for handling class-imbalanced in classification problems arXiv:2211.05456

Staats B A, Bonekat H W, Harris C D and Offord K P 1984 Chest wall motion in sleep apnea *Am. Rev. Respiratory Dis.* **130** 59–63

Van Steenkiste T, Groenendaal W, Deschrijver D and Dhaene T 2019 Automated sleep apnea detection in raw respiratory signals using long short-term memory neural networks *IEEE J. Biomed. Health Inform.* **23** 2354–64

Sundararajan D 2001 *The Discrete Fourier Transform: Theory, Algorithms and Applications* (World Scientific)

Sun G and Matsui T 2015 Rapid and stable measurement of respiratory rate from Doppler radar signals using time domain autocorrelation model *IEEE 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Milan) 5985–8

Thommandram A, Eklund J M and McGregor C 2013 Detection of apnoea from respiratory time series data using clinically recognizable features and kNN classification *IEEE 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka) pp 5013–6

Tobin M J, Cohn M A and Sackner M A 1983 Breathing abnormalities during sleep *Arch. Intern. Med.* **143** 1221–8

Uddin M B, Chow C M and Su S W 2018 Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review *Physiol. Meas.* **39** 03TR01

Urtnasan E, Park J-U and Lee K-J 2018 Automatic detection of sleep-disordered breathing events using recurrent neural networks from an electrocardiogram signal *Neural Comput. Appl.* **32** 4733–42

Vaquerizo-Villar F, Alvarez D, Kheirandish-Gozal L, Gutierrez-Tobal G C, Barroso-Garcia V, Crespo A, Del Campo F, Gozal D and Hornero R 2018 Detrended fluctuation analysis of the oximetry signal to assist in paediatric sleep apnoea-hypopnoea syndrome diagnosis *Physiol. Meas.* **39** 114006

Várady P, Micsik T, Benedek S and Benyó Z 2002 A novel method for the detection of apnea and hypopnea events in respiration signals *IEEE Trans. Biomed. Eng.* **49** 936–42

Vluymans S 2019 Learning from imbalanced data *Dealing with Imbalanced and Weakly Labelled Data in Machine Learning using Fuzzy and Rough Set Methods* (Springer) (*Studies in Computational Intelligence*) pp 81–110

Wallace B C, Small K, Brodley C E and Trikalinos T A 2011 Class imbalance, redux *2011 IEEE 11th Int. Conf. on Data Mining* pp 754–63

Wu Y, Pang X, Zhao G, Yue H, Lei W and Wang Y 2021 A novel approach to diagnose sleep apnea using enhanced frequency extraction network *Comput. Methods Programs Biomed.* **206** 106119

Yizraeli Davidovich M L, Karasik R, Tal A and Shinar Z 2016 Sleep apnea screening with a contact: free under: the: mattress sensor *2016 Computing in Cardiology Conf. (CinC)*

Yue H *et al* 2021 Deep learning for diagnosis and classification of obstructive sleep apnea: a nasal airflow-based multi-resolution residual network *Nat. Sci. Sleep.* **13** 361–73

Ziarani A K and Konrad A 2002 A nonlinear adaptive method of elimination of power line interference in ECG signals *IEEE Trans. Biomed. Eng.* **49** 540–7