Artificial Intelligence In Medicine 130 (2022) 102328

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

A meta-learning algorithm for respiratory flow prediction from FBG-based wearables in unrestrained conditions

Mariangela Filosa [a,b,*], Luca Massari [a,b], Davide Ferraro [a,b], Giacomo D'Alesio [a,b], Jessica D'Abbraccio [a,b], Andrea Aliperta [a,b,e], Daniela Lo Presti [c], Joshua Di Tocco [c], Martina Zaltieri [c], Carlo Massaroni [c], Maria Chiara Carrozza [a,b,d], Maurizio Ferrarin [d], Marco Di Rienzo [d], Emiliano Schena [c], Calogero Maria Oddo [a,b,*]

[a] The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy
[b] Department of Excellence in Robotics & A.I., Scuola Superiore Sant'Anna, Pisa, Italy
[c] Department of Engineering, Università Campus Bio-Medico di Roma, Rome, Italy
[d] IRCCS Fondazione don Carlo Gnocchi, Milan, Italy
[e] ARTES4.0 Competence Center on Advanced Robotics and enabling digital TEchnologies and Systems, Pontedera, Italy

ARTICLE INFO

Keywords:
Artificial intelligence
meta-learning
Deep learning
Wearable systems
Fiber Bragg Grating sensors
Breathing activity

ABSTRACT

The continuous monitoring of an individual's breathing can be an instrument for the assessment and enhancement of human wellness. Specific respiratory features are unique markers of the deterioration of a health condition, the onset of a disease, fatigue and stressful circumstances. The early and reliable prediction of high-risk situations can result in the implementation of appropriate intervention strategies that might be lifesaving. Hence, smart wearables for the monitoring of continuous breathing have recently been attracting the interest of many researchers and companies. However, most of the existing approaches do not provide comprehensive respiratory information. For this reason, a meta-learning algorithm based on LSTM neural networks for inferring the respiratory flow from a wearable system embedding FBG sensors and inertial units is herein proposed. Different conventional machine learning approaches were implemented as well to ultimately compare the results. The meta-learning algorithm turned out to be the most accurate in predicting respiratory flow when new subjects are considered. Furthermore, the LSTM model memory capability has been proven to be advantageous for capturing relevant aspects of the breathing pattern. The algorithms were tested under different conditions, both static and dynamic, and with more unobtrusive device configurations. The meta-learning results demonstrated that a short one-time calibration may provide subject-specific models which predict the respiratory flow with high accuracy, even when the number of sensors is reduced. Flow RMS errors on the test set ranged from 22.03 L/min, when the minimum number of sensors was considered, to 9.97 L/min for the complete setting (target flow range: 69.231 $\pm$ 21.477 L/min). The correlation coefficient r between the target and the predicted flow changed accordingly, being higher ($r = 0.9$) for the most comprehensive and heterogeneous wearable device configuration. Similar results were achieved even with simpler settings which included the thoracic sensors (r ranging from 0.84 to 0.88; test flow RMSE = 10.99 L/min, when exclusively using the thoracic FBGs). The further estimation of respiratory parameters, i.e., rate and volume, with low errors across different breathing behaviors and postures proved the potential of such approach. These findings lay the foundation for the implementation of reliable custom solutions and more sophisticated artificial intelligence-based algorithms for daily life health-related applications.

Abbreviations: FBG, Fiber Bragg Grating; M-IMU, magneto-inertial measurement unit; LSTM, long-short term memory; AI, artificial intelligence; RR, respiratory rate; RV, respiratory volume; RMSE, root mean square error; IQR, interquartile range; GUI, Graphical User Interface.

* Corresponding authors at: The BioRobotics Institute and Department of Excellence in Robotics & A.I., Scuola Superiore Sant'Anna, Pisa, Italy.
E-mail addresses: mariangela.filosa@santannapisa.it (M. Filosa), calogero.oddo@santannapisa.it (C.M. Oddo).

https://doi.org/10.1016/j.artmed.2022.102328
Received 26 September 2021; Received in revised form 23 May 2022; Accepted 25 May 2022
Available online 29 May 2022

1. Introduction

1.1. Motivation and background

The monitoring of the respiratory activity is paramount for the health and safety assessment and to improve the quality of life since it's highly correlated with many psychophysical factors [1]. In the clinical field, continuous breathing tracking systems may help identifying the onset of lung diseases, as well as potential failures or changes occurring to the organs [2]. In addition, treatment follow-up [3] and condition exacerbation detection may be achieved through remote measurement and warning systems [4], consequently prompting intervention and therapy adaptation strategies [2,5,6]. Besides the clinical field, as an example, in the occupational health and safety framework, both the onset or the monitoring of work-related diseases and mental stress [7] and accidents forecasting, could be achieved by evaluating the parameters of breathing [8–10]. Consequently, strategies for work environment enhancement and injury prevention may be implemented. Therefore, in these scenarios, the continuous recording of an individual's breathing and robust real-time event detection might be lifesaving. Mainstream breathing monitoring techniques usually provide one-time, although accurate, measurements, since the device usage is limited to the clinical setting because of their high cost, encumbrance and expertise requirements. The clinical gold standard procedure for breathing evaluation is spirometry [11]. However, mouthpiece-induced discomfort and natural breathing conditioning [12–14] mainly prevent spirometers to be used in long-term continuous monitoring. More portable solutions rely on chest wall movement detection, which is currently the most popular technique to monitor respiratory activity [15], since the compartmental model of the chest wall has been proposed [16]. However contact-based conventional method measurements [17–19] are strongly affected by posture changes, slight motion-induced displacements of the sensors and electromagnetic interference [13]. System encumbrance and potential difficulties in the device positioning make them unsuitable for a remote continuous monitoring in everyday life. Tighter and stabler wearable sensing systems embedding instrumented garments, may be seamlessly integrated in the user's daily activities, regardless of the setting, without being neither invasive nor obtrusive. Among the most common wearable technologies, inertial sensors are widely used in combination with either similar units positioned elsewhere on the body [20] or in addition to other types of sensors serving as motion artifact compensation tools [21]. According to [15], most of the recently developed wearables for breathing activity monitoring are equipped with optical fiber sensors. These systems rely on either the intensity modulation through macro-/micro-bending effects [22,23] or the wavelength modulation, as for Fiber Bragg Grating (FBG) sensors [24]. The advantages of such technology lie in its small size, flexibility, robustness and immunity to electromagnetic fields, which are a major concern in many environments. In addition, FBGs are inherently safe and very sensitive to strain (about 1 μm/με), thus achieving the detection of even very small chest movements. Besides that, the most outstanding feature of these sensors is the multiplexing possibility, since many gratings can be inscribed within a single fiber wire, thus minimizing the number of needed links. Lastly, given their form-factor, optical fibers could be integrated in a straightforward manner into smart textiles by being attached, woven or knitted in fabric [25]. Most of the wearable existing systems only focus on retrieving the respiratory rate (i.e., the number of breaths per minute), without accounting for other parameters or overall characteristics [15]. As an example, in [22,23,25,26], wearable optical sensors captured the respiratory frequency with very low errors. Although the rate itself has been proven to be a valuable parameter for identifying adverse events [27], the monitoring of the respiratory pattern variability, as well as the minute ventilation (i.e., the product between the respiratory rate and the tidal volume, i.e., the amount of inhaled and exhaled air per breath during quiet respiration), could provide a deeper insight on the user ventilatory status, revealing even underlying damages [2,28,29]. Nevertheless, the accurate measurement of volume is an open challenge when dealing with wearables, since artifacts arising from either whole-body movements or unstable sensors are a major concern [30,31]. Few studies reported optical sensors in comprehensive breathing monitoring applications, and the ones that do, besides the frequency, usually focus only on relative volume estimation [32–34]. Along with these, FBGs have been used for chest volume computation via a calibration procedure based on an optoelectronic plethysmography reference [35–37]. Results were promising but still preliminary, since only the quiet breathing activity during standing stationary postures was evaluated. In order to extrapolate the absolute respiratory volume, a calibration against a reference device is usually required. The outcomes of conventional methods strongly depend on the calibration procedure itself [13]. Traditional approaches are based on linear models [38] which are quite accurate when dealing with constant or quasi-constant tidal volume [30,39,40]. However, the respiratory system is considered to be inhomogeneous and chaotic even at rest [41], and the thoracoabdominal interactions are complex [42] to such an extent that non-linear models have been proposed and often outperformed the linear counterparts [43,44]. Deep learning algorithms have been proposed as well, but mainly for event detection and recognition. In this field, sleep studies have been carried out in order to identify abnormal breathing patterns by means of one-dimensional convolutional neural networks (CNN) [45], or to detect sleep apnea through different approaches (1D CNN, Long Short-Term Memory (LSTM)-based and fully connected networks) [46,47]. Other examples reported the recognition of sleep stages from ECG and respiratory signals via five deep-learning algorithms (CNN and LSTM-based) [48], and the *end-to-end* respiratory signal prediction from sleep photoplethysmography data by applying a CNN-based encoder-decoder solution [49]. A general-purpose application is addressed in [50], where the combination of the accelerometer data and a 2D CNN is intended for the recognition of different breathing patterns. Nevertheless, the presented approaches do not deal with the reconstruction of respiratory activity-related signals, and they have been validated mostly in very stationary conditions (e.g., sleep). Furthermore, the potentiality of such algorithms combined with chest wearable sensor data is still under investigation.

1.2. Main study contribution

In view of the above-mentioned criticalities regarding the continuous and comprehensive monitoring of the respiratory activity, in this work, a LSTM-based deep learning algorithm for the prediction of the respiratory flow, starting from a chest wearable system, is presented. Optical sensors, specifically FBGs, along with an inertial unit, provided information about the chest motion without hindering the user's activities. The primary aim of this work was overcoming the limitations of both linear models and more complex algorithms. Hence, the proposed approach was validated by considering both static postures and dynamic activities. In addition, the deep learning model was enriched with the implementation of a meta-learning solution [51,52], thus facing the major concern of the across-subject variability in the respiratory activity. Therefore, the possibility of adapting the model to new subjects through a short calibration session is a substantial added value of such approach. Furthermore, the performance of the model has been compared with additional benchmarking machine learning algorithms, as in [53] and tested for different sensor configurations, thus providing a deeper insight into chest FBG-based wearables for respiratory flow predictions via an artificial intelligence algorithm. All things considered, the possibility of retrieving a continuous and prolonged respiratory signal by means of lightweight and comfortable sensors paves the way to the further implementation of more complex and reliable well-being evaluation systems.

In the following sections, the wearable system, as well as the experimental session for data collection, is presented. Then, the meta-

learning algorithm is detailed and compared with four different approaches, henceforth the benchmarking machine learning algorithms. Further, the computation of the respiratory parameters is addressed. Finally, the algorithm performance results are presented and discussed.

2. Materials and methods

2.1. FBG-based wearable system and experimental setup

The prototype of the wearable device consisted of two adjustable elastic belts designed to be tightened around the thorax and the abdomen (Fig. 1A). The single belt embedded two flexible sensors, each composed of a polymeric substrate (2 mm thick) and encapsulating an FBG sensor (10 mm length; reflectivity > 90%; AtGrating Technologies; further details about the respiratory sensors and their working principle in [54]). Given the FBG sensitivity to strain, which is encoded in the wavelength shifts ($\Delta\lambda$) from the central value (λ_0) of the back-reflected spectrum (Fig. 1B–C), when placed on the chest, these sensors can track its periodic inward and outward motion [55]. In particular, the FBGs are elongated when the thorax enlarges (Fig. 1C), i.e., during the air inhalation phase, and they shorten when the chest shrinks (Fig. 1B), i.e., during the air exhalation phase. Since the FBGs are sensitive to any kind of motion-induced strain, breathing-unrelated movements may affect the quality of the signal of interest. In order to decouple different motion effects, a M-IMU (LSM9DS1 by STMicroelectronics), including an accelerometer and a gyroscope, was added and worn at the sternum level (Fig. 1A). The temperature influence on the FBG wavelengths, instead, was considered negligible since the current wearable prototype was intended for experiments in a controlled environment. Besides the wearable sensors, a commercial variable orifice flowmeter (SpiroQuant P by EnviteC, Honeywell) was used as a reference device. The respiratory airflow, collected by means of a facemask and causing a pressure drop in the flow sensor, was then encoded into a voltage signal (sampling rate 100 Hz). Both the M-IMU and the flowmeter electronics were encased in a box (Fig. 1A) and their data were sent to a laptop via Bluetooth. Conversely, the FBG absolute wavelength data were transduced by an optical interrogator (Hyperion si255, Luna, Inc.; Fiber Fabry-Pérot Tunable Filter laser-based [56,57]; 4 measurement channels; wavelength range: 1460–1620 nm; wavelength accuracy: 1 pm; dynamic range: 25 dB peak) with a sampling rate of 100 Hz (Fig. 1A) and transmitted to the laptop via Ethernet, through a TCP/IP communication protocol. Extension cords from the sensors to the interrogation unit were appended to facilitate the recruited subjects' movements with negligible signal loss, since the FBG wavelength measurement is not affected by power fluctuations along the optical path [58]. Data collection, synchronization, visualization and storage were managed through a custom Graphical User Interface (LabVIEW, National Instruments, TX, USA). Proprietary LabVIEW libraries for recording the optical interrogator outputs were integrated in the acquisition code, as well. Furthermore, an additional GUI, namely the virtual illustrator, was developed to show the subject the specific actions he/she had to execute during the experimental session (Fig. 1A).

2.2. Experimental protocol

The wearable system and the reference flowmeter were used to capture the respiratory activity of thirteen healthy volunteers (seven females; age 27 ± 4; height 168 ± 10 cm; weight 63.1 ± 12.0 kg). Recruited subjects were required to be mentally and physically healthy and to not have any respiratory difficulty or disorder. No exclusion criteria derived from the hardware thanks to its adaptability. The experiments were approved by the ethical committee of Università Campus Bio-Medico di Roma and volunteers signed an informed consent prior to being recruited. Each subject donned the wearable system while

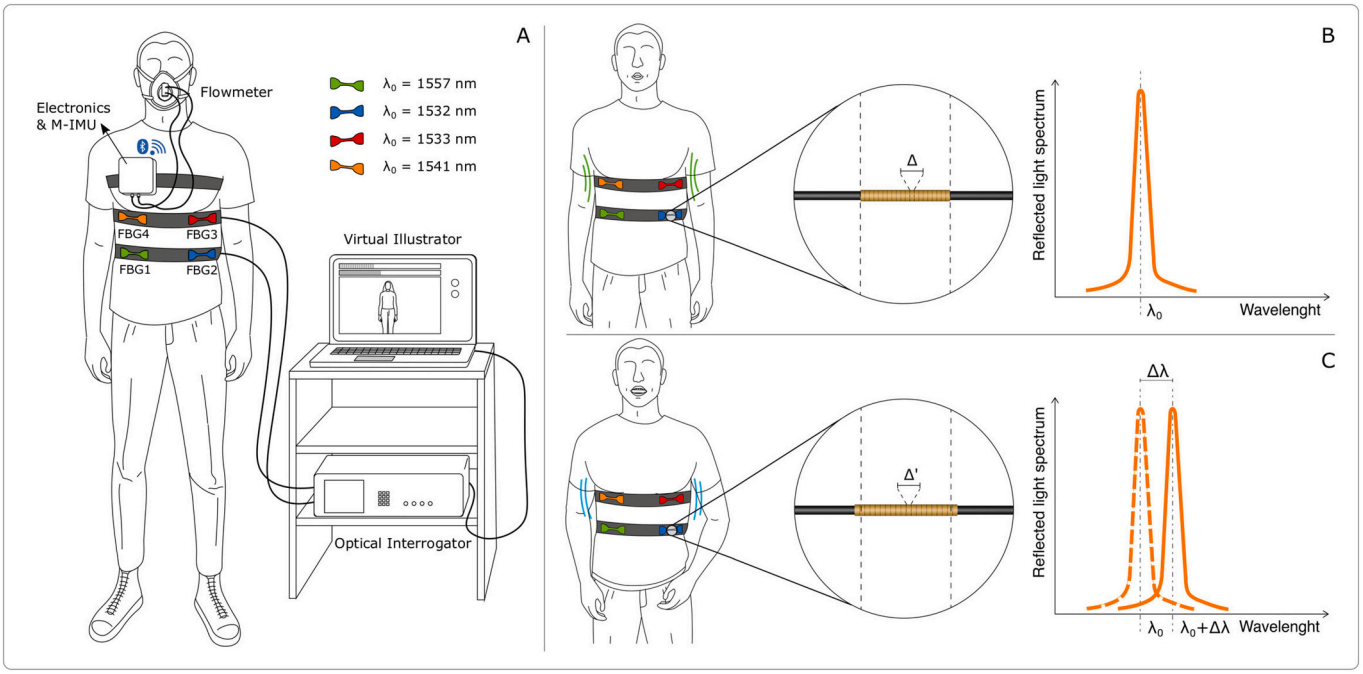

Fig. 1. A) Experimental setup. A volunteer wearing the wearable device and the facemask for the reference respiratory flow collection is portrayed. The wearable system was made up of 4 FBG sensors with different nominal central wavelengths λ_0 and a M-IMU. The M-IMU was enclosed into the depicted box and fastened around the upper thorax. The output fibers are then connected to the desktop optical interrogator (FBG1–2 and FBG 3–4 connected to two separate channels), which illuminated the fibers and read the $\Delta\lambda$. In its turn, the interrogation unit is wired to the laptop, which showed the virtual illustrator for the experimental session. B and C) Functioning principle of the FBG-based respiratory sensors. B) During the exhalation, the FBGs shorten up to their nominal length. The pitch distance, that varies with the applied strain and contributes to the wavelength shift, is Δ. Accordingly, the FBG wavelength is centered at λ_0 (graph on the right). Each FBG λ_0 slightly differed across subjects, according to the initial fiber prestress. C) During the inhalation the FBGs are stretched, causing a pitch increase, thus obtaining Δ'. The back-reflected spectrum shows a peak which is shifted of $\Delta\lambda$ from the central wavelength (graph on the right).

being supervised by an experimenter so he/she could correctly place the sensors attached to the elastic belts on the thorax and the abdomen. Any repositioning of the device was avoided during each session. During the experiment, the subjects performed three series of activities, henceforth the *protocols*, as illustrated in Fig. 2. In particular, they were asked to quietly breath (*eupnea*), hold the breath (*apnea*) and rapidly breath (*tachypnea*) while *standing*, *sitting* and *lying on their backs* (in Fig. 2, P1, P2 and P3 the postures, respectively). Transitions from a posture to another, namely the *adaptations*, were included as well. Besides these static activities, there was a self-paced walk while spontaneously breathing (from now on, *walk*) in the second *protocol* (in Fig. 2, referred to as W). For the entire experimental session, the subject was instructed to refer to the monitor, where the virtual illustrator (Fig. 1A) showed the specific action he/she had to perform and its timing.

2.3. Meta-learning algorithm for respiratory flow prediction

Data processing was implemented in MATLAB (Mathworks, Inc.). Collected data were conveyed to a meta-learning algorithm in order to estimate the respiratory flow. In particular, the wearable system data (Fig. 2), i.e., the FBG wavelength shifts ($\Delta\lambda$) along with their derivatives and the M-IMU three-axes accelerations (a) and angular velocities (ω), were the algorithm input features. The flowmeter voltages (V_O, raw signals in Fig. 2) were the algorithm target outputs that had to be predicted. Four benchmarking machine learning algorithms were also trained for comparing the results of the proposed approach with conventional regression methods. Finally, the respiratory benchmarking parameters, i.e., the respiratory rate (RR) and volume (RV), were computed for evaluating the robustness of the algorithm in potential applications. Fig. 3 shows the data processing workflow. The raw signals were first preprocessed. Abnormal FBG wavelength values, which originated from temporary fiber cables bending or bad connections, were detected and deleted, and missing signal portions were reconstructed when possible (the maximum admitted gap length for interpolation was 500 samples). All data were also resampled at a frequency of 50 Hz to guarantee evenly spaced data points. Then, signals were smoothed by means of a first-order low-pass Butterworth filter with a cut-off frequency of 2 Hz, which is suitable for capturing breathing dynamics. Furthermore, the FBG wavelengths were de-offset differently for each subject-related dataset by subtracting their initial values λ_0, thus obtaining the FBG $\Delta\lambda$. Time dependence was preserved by splitting *protocols* in about 10 s sequences in order to have short and equal length timeseries. Only *apnea*, *eupnea*, *tachypnea* and *walk* sequences were kept. Afterward, these timeseries were conveyed to the regression meta-learning algorithm.

Meta-learning means *learning to learn* [51], hence, as it happens for human learning strategies, a model can learn new tasks from the experience of similar and familiar ones. A prior knowledge of the problem helps gaining new skills more easily than starting from scratch. Herein, the proposed few-shot learning approach reproduced the serial version of the Reptile algorithm [59], which, specifically, approximates the Model-Agnostic Meta-Learning procedure (MAML [60]). Differently from the MAML approach, the Reptile algorithm does not deal with differentiation, but it is based on a linear updating rule for retrieving the initial testing model parameter setting. In detail, the neural network learns to regress the air flow from the chest movements of new subjects from only a few respiratory sequences. An initial setting of the model parameters is firstly obtained by training the algorithm on an extensive dataset of similar examples. This stage is called *meta-training*. Then, when a new, but familiar, example is presented, i.e., a new subject, the model parameters are adjusted via a few training iterations in order to ultimately find the tailored parameter setting for it. Starting from previously established parameters, instead of random ones, helps making this last tuning phase as fast as possible and to generalize well from the few sequences related to the new subject. This stage is referred to as *meta-testing*.

Firstly, this algorithm progressively computes the parameter values in the *meta-training* phase by moving them in the direction of those computed on randomly sampled sequences of a subject belonging to the *meta-training* dataset. The *meta-training* dataset consisted of the above-mentioned sensor data collected from 10 (*S* in Fig. 3) out of 13 randomly selected subjects, for a total of 834 input sequences. Each one was about 10 s long, for a total duration of 2.3 h. In the following pseudo-code, the *meta-training* stage is summarized.

Meta-training algorithm

ϕ initialization
for e = 1: *meta-epochs*
  for s = 1: *subjects*
    N sequences random sampling → *meta-training* set
    Training set normalization
    ϕ_s = ***model*** (ϕ, k)
  $\phi_{update} = \phi + \varepsilon(\phi_s - \phi)$
  $\phi \leftarrow \phi_{update}$
$\phi_{meta\text{-}training} \leftarrow \phi$

In other words, the parameters (ϕ), i.e., the *model* weights and biases, were at first randomly initialized. Then, the algorithm was run for 50 iterations (*meta-epochs*). At the *e*-th *meta-epoch*, each subject *s* dataset was considered for randomly sampling *N* (30) sequences. This data pool was then normalized to facilitate the algorithm convergence to the global minimum. Hence, the *model* was trained on these randomly selected sets for *k* (50) iterations, starting from the previous parameter setting (ϕ). By doing so, the new parameters (ϕ_s) were obtained and they reflected the features of the *s*-th subject. Finally, both the previous parameter configuration (ϕ) and the new one (ϕ_s) contributed to the parameter update. In particular, the new settings (ϕ_{update}) were computed through a linear rule (in the pseudo-code) which was governed by the ε factor, that changed linearly at each *meta-epoch* as follows:

$$\varepsilon = 1 - e/meta\_epochs \tag{1}$$

The updated parameters (ϕ_{update}) became the ones (ϕ) that were then serially fed into the algorithm, as reported in the pseudo-code. Once the last *meta-epoch* was reached, the final *meta-training* parameters ($\phi_{meta\text{-}training}$) were retrieved. These became the initial settings for the following *meta-testing* stage, where weights and biases for each new subject were adjusted accordingly. The *meta-testing* and *testing* datasets consisted of the sensor data collected from the 3 (*S1*) out of 13 remaining subjects. In particular, the *meta-testing* one was made up of the randomly selected 70% of the total number of sequences related to each activity, i.e., *apnea*, *eupnea*, *tachypnea* and *walk*. Therefore, for each *meta-testing* subject, the training set was made up of 58 10 s input sequences, for a total duration of 9.67 min. The residual sequences (25, i.e., 4.17 min of recordings) were considered for the next *testing* phase. The main steps of the *meta-testing* algorithm are reported below, and they are intended to be run separately for each new *meta-testing* subject.

Meta-testing algorithm (one subject)

$\phi_{s1} \leftarrow \phi_{meta\text{-}training}$
meta-testing set normalization
for e_{s1} = 1: *epochs*
  M sequences random sampling → training set
  ϕ_{s1}' = ***model*** (ϕ_{s1}, $k1$)
  $\phi_{s1\_update} = \phi_{s1} + \varepsilon(\phi_{s1}' - \phi_{s1})$
  $\phi_{s1} \leftarrow \phi_{s1\_update}$
$\phi_{subject} \leftarrow \phi_{s1}$

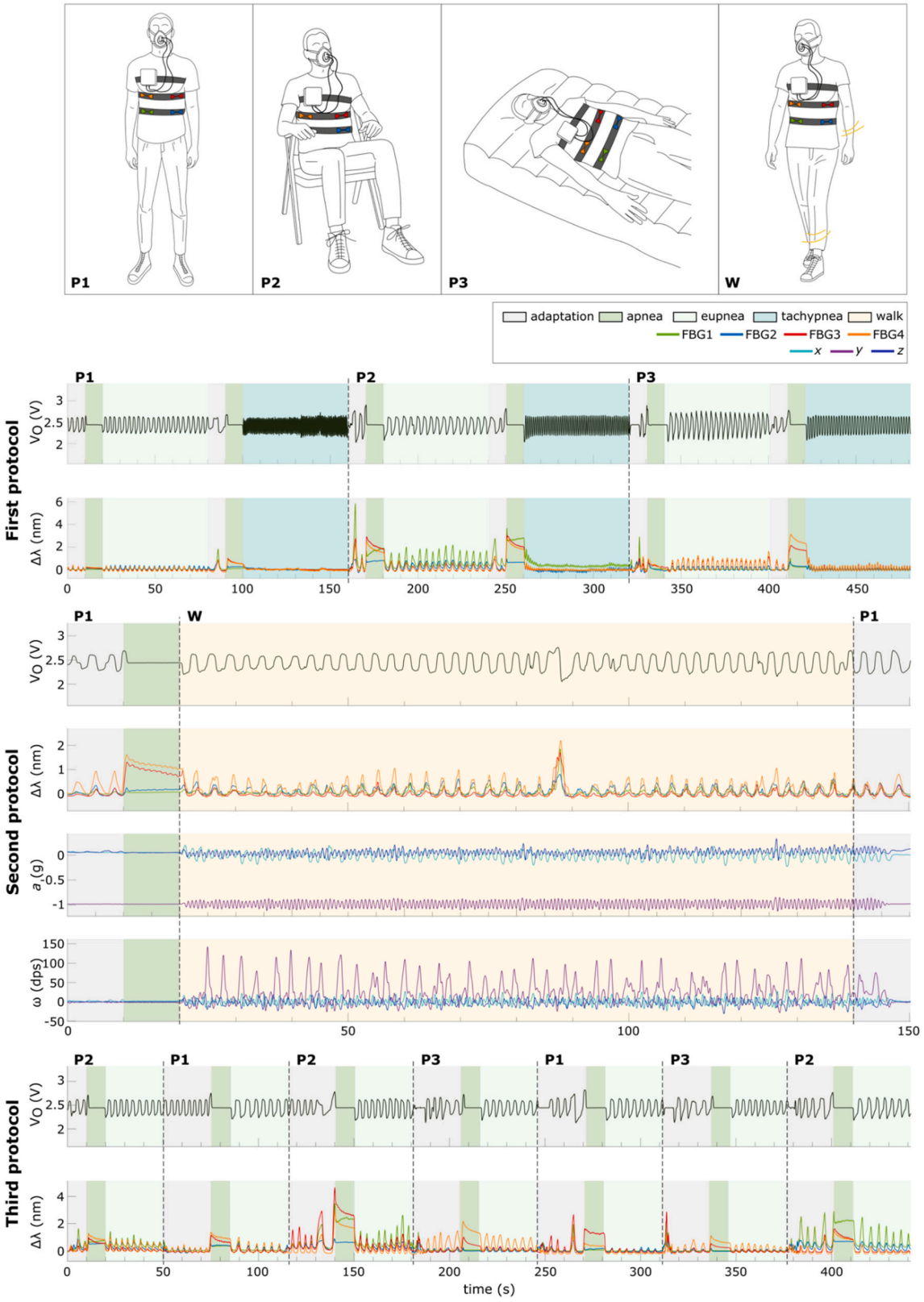

Fig. 2. Experimental session and raw signals. On top, the postures (P1, P2 and P3, i.e., *standing*, *sitting* and *lying on the back*, respectively) and the activity (W, i.e., *walk*) to be performed during the experiment. Each of them is reported accordingly in the *protocol* sequences. For each *protocol* the collected raw data are reported as well. M-IMU (accelerations (a) and angular velocities (ω)) data are illustrated for the second *protocol* only, since it included the dynamic *walk* task. In the other *protocols*, variations of M-IMU data were significantly smaller. Respiratory tasks (i.e., *apnea*, *eupnea* and *tachypnea*) are illustrated in the *protocols* through the colored background areas.

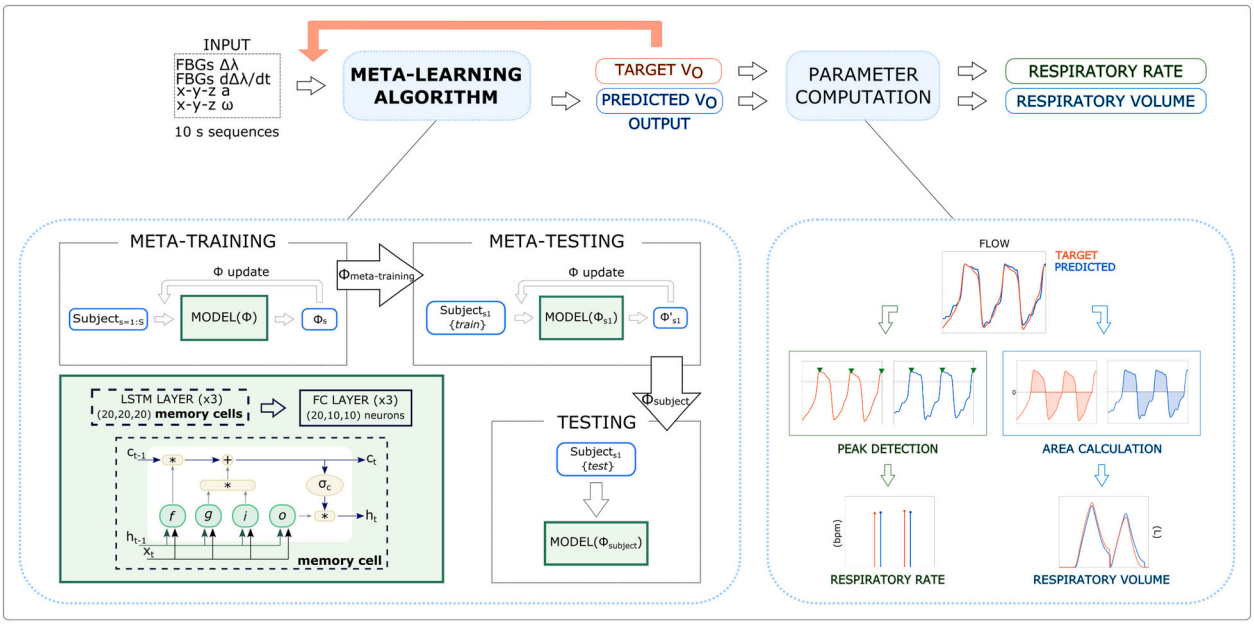

Fig. 3. Meta-learning algorithm and data processing workflow. The input features were conveyed to the meta-learning algorithm (bottom-left), which output the prediction of the flow voltages (V_O). The target V_O was fed into the algorithm for the *meta-training* and the *meta-testing* stages only. The *meta-training* block output the parameters ($\phi_{meta\text{-}training}$) computed for S subjects, whereas the *meta-testing* and the *testing* stages were run separately for each of the $S1$ subjects. The LSTM model and the *memory cell* structure are represented as well (green box). Both the estimated and target *testing* voltages were processed for the computation of the respiratory parameters (bottom-right), thus obtaining the respiratory rate and volume. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Firstly, the parameters were initialized with the previous extrapolated *meta-training* settings ($\phi_{meta\text{-}training}$), thus providing ϕ_{s1}. After the dataset normalization, at each epoch (e_{s1}), M (30) sequences were randomly selected from the *meta-testing* set. The *model* was trained for $k1$ (30) iterations, starting from the previous parameter settings (ϕ_{s1}). Then, the parameters were updated (ϕ_{s1_update}) and fed sequentially into the algorithm. When the number of epochs (e_{s1}) reached *epochs* (20), the *meta-testing* algorithm stopped, and the final parameter set ($\phi_{subject}$) was retrieved. Finally, the *testing* stage relied on the $\phi_{subject}$ obtained during the *meta-testing*. Hence, the input features, i.e., the FBG Δλ, their derivatives and the M-IMU data, were conveyed to the model (*model* ($\phi_{subject}$)) for further predicting the reference flowmeter signal for that subject. In Fig. 3, on the left, the workflow of the proposed *meta-learning* algorithm is illustrated. The implemented deep-learning *model* consisted in a Long-Short Term Memory (LSTM)-based network [61,62] structure made up of 3 stacked LSTM and 3 fully connected (FC) layers (Fig. 3, bottom-left). LSTM neurons, *memory cells* or units, are provided with multiple internal gates intended to manage long-term memory. In more detail, the information flows through the *memory cells*, which are as many as the number of time steps to consider backwards. Each *memory cell* has its own *cell state* (c), which is regulated by means of the *forget* (f), the *input* (i) and the *output* (o) gates (Fig. 3). At each time step t, the current input x_t and the previous cell and hidden states, respectively c_{t-1} and h_{t-1}, enter the LSTM unit to finally output the current updated states as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \tag{2}$$

$$h_t = o_t \odot \sigma_c(c_t) \tag{3}$$

where $\odot$ denotes the element-wise product, σ_c is the hyperbolic tangent activation function, and f_t, i_t, o_t and g_t are respectively the *forget*, the *input*, the *output* gates and the *cell candidate* at time t. Thoroughly, the forget gate [63] controls what to retain and reject for the *cell state* computation by looking at both the previous hidden state and the given input. The forget gate output is:

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \tag{4}$$

The σ_g is the sigmoid activation function, W, R and b indicate, in general, respectively the input weights, the recurrent weights and the biases, here pertained to the forget gate as the subscripted f denotes. The other gate outputs are computed as follows:

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \tag{5}$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \tag{6}$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \tag{7}$$

The proper combined opening and closure of the gates protect the *cell state* from potential changes originating from irrelevant inputs or noise, thus preserving the memory ability of the network.

The implemented LSTM recurrent network was devised to gather temporal correlations in the signals in order to facilitate the respiratory flow prediction. Each LSTM layer included 20 hidden units (*memory cells* in Fig. 3). The same amount was considered for the first FC layer, whereas the last two were provided with half the neurons. The *model* learnable parameters (ϕ_s or ϕ_{s1}) were iteratively updated within the k (or $k1$) iterations, with the adaptive moment estimation (*Adam*) [64] approach, whose decay rates were set to 0.9 and 0.999 for the first and the second moments respectively. The learning rate was initialized at each epoch (e or e_{s1}) to 0.005 and was then dropped by 0.1 factor every 10 iterations (out of k or $k1$). The *model* gradient evaluation function was the half mean squared error, as in the following *loss* expression:

$$loss = (1/2T)\sum_{i=1}^{T}(\hat{y}_i - y_i) \tag{8}$$

where T denotes the maximum sequence length, $\hat{y}$ the predicted output and y the expected one. Once being computed, the gradients were clipped at 0.5 with a Euclidean norm-based scaling factor. The whole algorithm, comprehensive of the *meta-training*, the *meta-testing* and the *testing* stages, was run for different input data configurations. In

particular, the M-IMU, the thoracic and the abdominal FBG data were considered both alone and in combination with the other ones. Hence, seven configurations were investigated, as depicted in Fig. 4, and named M-IMU, $2FBG_{AB}$, $2FBG_{TX}$, $2FBG_{AB}$ + M-IMU, $2FBG_{TX}$ + M-IMU and 4FBG + M-IMU. The *meta-training* and *meta-testing* dataset arrangements were kept across all the configurations.

The *testing* algorithm output the predicted flowmeter voltages which were then compared with the target signals. The data analysis provided overall considerations, without accounting for the differences between individual subjects. At first, both the root mean square errors and the Pearson correlation coefficient (r) of target and predicted values were computed for each configuration, in order to evaluate the overall algorithm performance. Then, both the signals were processed for the estimation of the respiratory parameters, as reported in the following section.

2.4. Benchmarking machine learning algorithms

In order to compare the LSTM-based meta-learning results with other conventional machine learning approaches, four additional models have been trained. In detail, a regression ensemble of trees, a shallow neural network, a recurrent neural network and an LSTM model were implemented. The choice of these benchmarking learners aimed at comparing the performance of the simplest regression approaches, i.e., regression trees and shallow networks, and the ones dedicated to timeseries prediction, i.e., basic recurrent networks and LSTM models. In particular, each algorithm was trained more than once while changing some relevant parameters. Given the relatively high amount of data, bagging regression tree ensembles were considered to limit overfitting, through averaging out the results of each tree in the collection [65]. Different numbers of weak learners (20, 50, 100 and 200; minimum leaf size = 10) were grown by picking random samples in the training dataset with replacement. The validation regression errors were then computed from the out-of-bag data, which did not contribute to the tree growth. Then, a shallow neural network was trained with different number of neurons in the hidden layer (20, 50, 100 and 200). The scaled conjugate gradient method for backpropagation and the early stop criterion (6 validation checks) were implemented. The dataset was randomly sorted into training (80%) and validation (20%) sets. Concerning the timeseries dedicated approaches, firstly, a recurrent network, as a generalization of the Elman model [66] was designed. Three hidden layers, with 20 neurons each, were included and different delays (10, 20 and 50 samples) were considered for training (gradient descent with momentum (0.7) and adaptive learning rate (0.01 and drop factor 0.5) backpropagation). The same dataset partition as the shallow network was devised. Finally, the LSTM model was implemented as described in the previous section but with the exclusion of the meta-learning procedure. Each algorithm input features were normalized and the datasets, i.e., the training sets, encompassing randomly sampled validation sets or out-of-bag data, and the test sets, were kept the same as for the meta-learning algorithm so they could be comparable. As for the meta-learning approach, the benchmarking method results were assessed for each configuration of sensors.

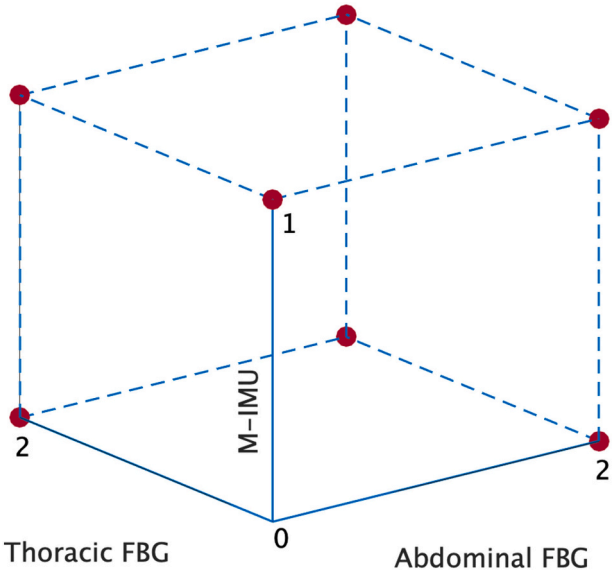

Fig. 4. Device configurations. On the axes, the number of abdominal and thoracic FBGs (0 or 2) and M-IMU (0 or 1). All the cube vertices, with the exclusion of the (0,0,0) point, represent a possible device arrangement. The dark red spots are the investigated sensor combinations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. Benchmarking respiratory parameters

The benchmarking RR and RV parameters were extrapolated from both the target signals and the meta-learning *testing* outputs, i.e., the predicted flowmeter voltages, for each input feature configuration. Firstly, voltage signals were converted into flows (L/min), being the sensor sensitivity 0.0069 V·min/L, and the offset an ad hoc value computed from target *apneas* for each subject-related test set. Afterward, the respiratory flow signals were smoothed by means of a moving-window Savitzky-Golay filter (window size: 27 samples). Then, the above-mentioned parameters were extrapolated.

RR is expressed in bpm, i.e., breaths per minute, as follows:

$$RR = 60/\Delta t \tag{9}$$

where Δt (s) is the time period between two consecutive flow peaks (Fig. 3, bottom-right). Specifically, the above threshold peaks in the signals were identified and checked to be the global maxima in each single breath. In detail, the threshold changed across the subjects, and it was one third of the interquartile range of all the target *testing* flow sequences. The median RR values were computed for each target and predicted 10 s sequence and further used for statistical analyses.

RV (L) refers to the inhaled and exhaled air quantity and it is defined as the flow signal subtended area. Therefore, the target and predicted flow approximate integrals were computed for each complete breath, which was in turn identified from the signal zero-crossing points (Fig. 3, bottom-right). Furthermore, at the end of each breath, RV was forced to 0 in order to avoid outcome drifts caused by asymmetries in the flow signal. Finally, the RV mean values from each 10 s sequence were considered for the next statistical analysis. Differences between the target and each configuration predicted parameters (i.e., the RR medians and the RV averages) were investigated by applying paired sample t-tests or Wilcoxon signed-rank tests, according to the normality assessment of each data pair. Then, the errors (RR and RV errors, in the following) for both the parameters were computed as the difference between the target values and each configuration predicted ones, thus obtaining 7 condition samples. The repeated-measures one-way ANOVA (or the Friedman test, for not normally distributed data) was applied to investigate differences between the 7 configurations. If significant p-values were found, post hoc pairwise comparisons (paired sample t-test or Wilcoxon signed-rank) with the relevant correction were carried out. Before the ANOVA, the sphericity assumption was assessed through the Mauchly's test and, when it was violated, the Greenhouse-Geisser correction was considered. Both the analyses on absolute and error data were run separately for each respiratory behavior (*apnea*, *eupnea*, *tachypnea*) and posture (*standing*, *sitting*, *lying on the back* and *walk*) dataset, to ultimately investigate potential differences in relevant parameter estimation across the considered breathing patterns and body positions and motion. The overall alpha significance level was 0.05. The statistical analysis was implemented in SPSS Statistics 26 (IBM, New

York, NY, USA).

3. Results

The meta-learning algorithm regression performance for each device configuration is depicted in Fig. 5 along with the corresponding Pearson correlation coefficients (r). In addition, the Root Mean Square Errors for all the algorithm stages, i.e., the meta-training, the meta-testing and the testing ones, are listed in Table 1. For a qualitative evaluation of the test results, some examples of the 4FBG + M-IMU superimposed target and predicted flows are showed in Fig. 6.

The benchmarking algorithm test results are presented in Fig. 7 and compared with the meta-learning method performance, in terms of across-subject median and percentiles RMSE. In addition, aggregated-subject flow root mean square errors for each tested algorithm and each condition are reported in Table 2. The average and the standard deviation of the test target flow range were 69.23 ± 21.48 L/min.

For what concerns the parameters, the analysis of the target vs. each configuration predicted RR medians resulted in significant differences in the *eupnea* (target vs. $2FBG_{AB}$ $p < .005$, $2FBG_{AB}$ + M-IMU $p = .005$), *standing* (target vs. M-IMU $p = .006$, $2FBG_{AB}$ + M-IMU $p = .041$), *lying on the back* (target vs. $2FBG_{AB}$ $p = .024$, M-IMU $p = .006$) and *walk* (target vs. $2FBG_{TX}$ $p = .02$) groups. The comparison among the device configurations through the RR errors resulted in differences for the *lying on the back* case ($\chi^2(6) = 19.66$, $p < .005$). However, post-hoc comparisons with the Bonferroni correction did not reveal paired differences. The Root Mean Square values of the RR errors are reported in Fig. 8. The paired comparisons of target and configuration-related RV averages revealed significant differences for the *eupnea* (target vs. $2FBG_{AB}$ $p = .02$, M-IMU $p < .001$, $2FBG_{TX}$ + M-IMU $p = .017$), *tachypnea* (target vs. $2FBG_{AB}$, M-IMU, $2FBG_{AB}$ + M-IMU, $2FBG_{TX}$ + M-IMU $p \leq .005$, 4FBG $p = .043$), *lying on the back* (target vs. $2FBG_{AB}$ $p = .002$, M-IMU $p = .001$, $2FBG_{TX}$ + M-IMU $p = .006$) and *walk* (target vs. all the configurations, $p \leq .01$, with the exclusion of $2FBG_{AB}$ + M-IMU $p = .051$ and 4FBG + M-IMU $p = .075$) groups. The Friedman test, for the configuration comparisons on RV errors, output significant differences for the *eupnea* ($\chi^2(6) = 36.32$, $p < .001$), *tachypnea* ($\chi^2(6) = 19.13$, $p < .005$), *standing* ($\chi^2(6) = 13.31$, $p = .038$) and *sitting* ($\chi^2(6) = 18$, $p = .006$) cases. Post hoc pairwise comparisons with the Bonferroni correction revealed that the M-IMU configuration was statistically different from all the other ones ($p \leq .001$), except for the $2FBG_{AB}$ one ($p = .248$) for the *eupnea* case, whereas no differences were found for the other groups. For the *lying on the back* and *walk* tasks, the ANOVA with the Greenhouse-Geisser correction determined differences between the device configurations ($F(2.88,48.99) = 5.44$, $p < .005$; $F(2.62,18.28) = 3.73$, $p = .034$ respectively). Post hoc tests using the Bonferroni correction revealed that paired conditions did not differ for *walk,* whereas significant differences were found for *lying on the back* ($2FBG_{AB}$ vs. $2FBG_{AB}$ + M-IMU ($p < .001$), $2FBG_{TX}$ ($p < .005$), 4FBG + M-IMU ($p < .005$); $2FBG_{TX}$ + M-IMU vs. 4FBG ($p = .034$), $2FBG_{TX}$ ($p < .005$)). The root mean square values of the RV errors are illustrated in Fig. 8.

4. Discussion and conclusion

4.1. Main achievements

In this work, a meta-learning algorithm for the prediction of the respiratory flow, starting from FBG-based wearable sensor data, has been presented. Currently, well-established wearables for respiratory purposes usually do not provide a comprehensive insight into the user status [15]. Hence, the ultimate goal of this study was to train the wearable sensors to estimate the respiratory flow, that was measured with a gold standard flowmeter, as a starting point for future applications in cardiorespiratory monitoring. In addition, most of the state-of-the-art approaches which combine wearables and AI, do not deal with dynamic activities, since whole body movements usually overwhelm the respiratory ones, the latter being significantly smaller. Therefore, achieving robustness in breathing assessment during free activities is a stated challenge. In this scenario, the feasibility of the proposed approach was tested with both static and dynamic experimental conditions. Besides the main novelty of predicting a reference respiratory flow from wearables under different circumstances, the meta-learning algorithm addresses the major concern of the inter-subject breathing variability [67,68], through the fine-tuning of the model. This requires a one-time calibration procedure which is very short, since only a few minutes are needed to record the respiratory activity (about 10 min per subject, as the duration of the proposed *meta-testing* dataset) and adjust the model. This approach also overcomes the need of very large datasets in order to achieve a fair generalization. The comparison with conventional machine learning algorithms for regression data prediction demonstrated that the proposed meta-learning approach achieved the lowest test errors (Fig. 7 and Table 2), thus confirming that tuning procedures might improve subject-dependent generalization properties. Furthermore, the meta-learning lowest error interquartile ranges (Fig. 7) suggested a fair across-subject repeatability of the results, unlike the other approaches. Comparable test results might be achieved in traditional machine learning approaches through very large datasets and extensive training sessions. Among the benchmarking methods, the ones for timeseries forecasting, i.e., the LSTM and the Recurrent NN, performed better than those without memory, i.e., Tree Ensembles and Shallow NN, with the LSTM-based approach being the most accurate (Fig. 7, Table 2). This evidence endorsed the application of the LSTM model to the meta-learning implementation. All the algorithms were tested in different wearable sensor configurations, i.e., both complete and simplified variations, to ultimately evaluate the meta-learning approach robustness in even more unobtrusive conditions. The meta-learning model showed better results than the conventional methods almost for every combination of wearable sensors, except for the $2FBG_{AB}$ configuration, since comparable or even lower errors were obtained through the benchmarking algorithms. Nevertheless, the corresponding flow error for that configuration (Table 2) across them was always remarkable with respect to the average target flow and considerably higher than the best meta-learning algorithm achievements. Further

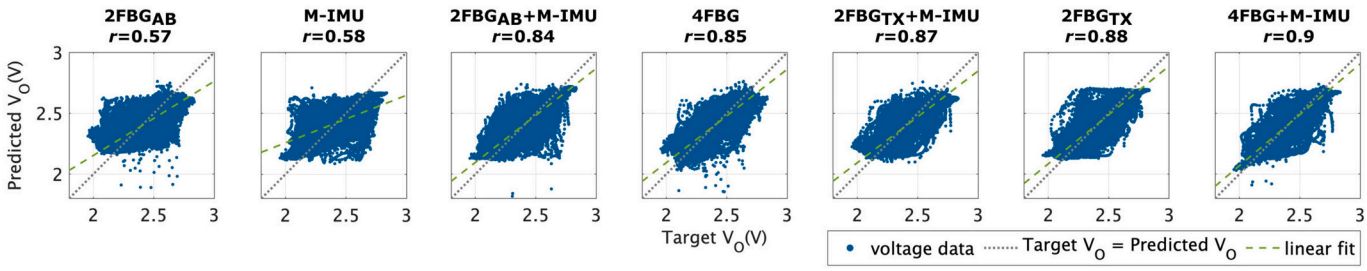

Fig. 5. Meta-learning target vs. predicted voltage correlation. For each device configuration, the data point cloud, as well as its regression line (in green) are showed. The desired regression line (in gray) is reported as well. Configurations are sorted from the worst to the best in terms of the achieved correlation coefficient r (in the headings). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Meta-Learning algorithm errors. Meta-testing and testing errors are across-subject (*S1*) median and interquartile range (in parentheses) values.

		Root mean square errors [V]						
		$2FBG_{AB}$	M-IMU	$2FBG_{AB}$ + M-IMU	4FBG	$2FBG_{TX}$ + M-IMU	$2FBG_{TX}$	4FBG + M-IMU
Meta-training	Training	0.089	0.096	0.048	0.072	0.043	0.054	0.060
	Validation	0.082	0.102	0.066	0.074	0.056	0.070	0.050
Meta-testing	Training	0.076 (0.017)	0.091 (0.023)	0.055 (0.002)	0.068 (0.003)	0.046 (0.010)	0.063 (0.011)	0.044 (0.006)
	Validation	0.106 (0.036)	0.133 (0.017)	0.077 (0.014)	0.062 (0.003)	0.072 (0.014)	0.081 (0.027)	0.062 (0.023)
Testing		0.131 (0.088)	0.131 (0.009)	0.081 (0.027)	0.085 (0.007)	0.077 (0.011)	0.075 (0.009)	0.063 (0.014)

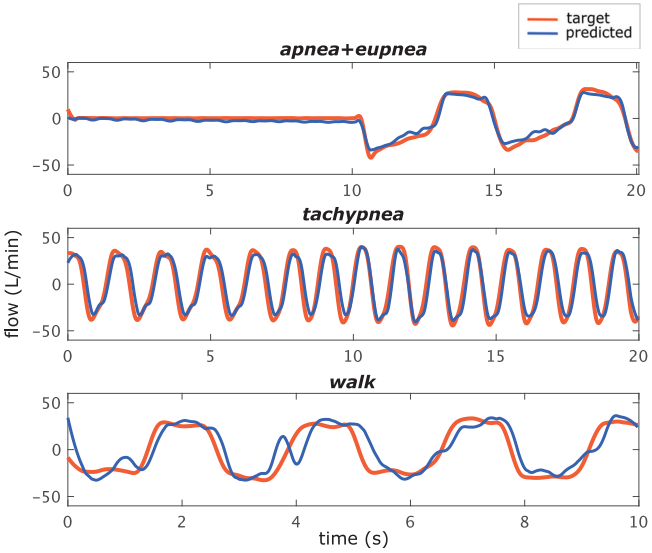

Fig. 6. Superimposed target and predicted signals for each activity (*apnea*, *eupnea*, *tachypnea* and *walk*) in the 4FBG + M-IMU configuration. An example of the transition from *apnea* to *eupnea* is reported in the top graph.

evidence is provided by the test regression plots and the training and the validation RMSEs for all the benchmarking algorithms (Supplementary Figs. 1 and 2, Supplementary Table 1). Considering the worse benchmarking outcomes, the following discussion mainly focuses on the meta-learning prediction accuracy. As expected, the 4FBG + M-IMU arrangement turned out to be more robust than the other ones, with a correlation against the reference flow of 0.9 (Fig. 5). Similar results are listed in Tables 1 and 2. The performance of the air flow estimation was almost the best for the complete set of the input features across all the meta-learning stages, especially in the *meta-testing* and *testing* ones. In the latter, the 4FBG + M-IMU flow error resulted to be about half the worst $2FBG_{AB}$ and M-IMU RMSEs (Table 2). Qualitatively, the signal fitting was very good for the static tasks (i.e., *apnea*, *eupnea* and *tachypnea*), whereas the prediction in the *walk* case exhibited slightly higher error but it was still satisfactory (Fig. 6). Furthermore, transitions between two different respiratory patterns, *apnea* and *eupnea* (Fig. 6), were accurately predicted. Comparable results were achieved with the $2FBG_{TX}$ configuration ($r = 0.88$, Fig. 5; RMSE = 10.639 L/min, Table 2), with slightly higher algorithm errors (Table 1). This outcome suggested that the usage of only two FBG sensors in the thoracic area may provide fair flow predictions. This might originate from the rigid structure of the thorax, which could be less sensitive to artifacts than the abdomen. Furthermore, the effect of gravity determines changes in the respiratory behaviors and lung capacity as a consequence of different postures [69], but the thorax, encompassing the rib cage, is less affected by this phenomenon. Conversely, shape changes are more common for the abdomen, which is mainly made up of soft tissues. Accordingly, the results showed that the $2FBG_{AB}$ configuration performance was the worst one ($r = 0.57$, algorithm errors ranging from 0.76 V to 0.131 V, Table 1). Nevertheless, when the abdominal FBGs were used in combination with the inertial sensors, the correlation coefficient resulted to be one and half time ($r = 0.84$) more than the $2FBG_{AB}$ case, as they jointly compensated for artifacts while perceiving the breathing movements as well. The algorithm errors reported in Table 1 bolstered this hypothesis, being similar to the 4FBG + M-IMU case ones. The 4FBG configuration resulted in $r = 0.85$, which was close to the $2FBG_{AB}$ + M-IMU correlation with the target. This suggests that the two configurations are equivalent, and this deduction is supported by the position of the M-IMU, which was close to the sternal region, thus limiting the lack of thoracic FBGs for the $2FBG_{AB}$ + M-IMU condition. In most cases, the combination of the inertial unit with other sensors resulted in better performance than the M-IMU configuration one. The correlation of the M-IMU signals with the target one was very low ($r = 0.58$) and similar to $2FBG_{AB}$ thus confirming that such positions for wearable sensors, if uniquely considered, are largely affected by non-respiratory components. Overall, the algorithm performance-related errors, as reported in Table 1, were consistent with what suggested hitherto, being similar across all the different settings of the input features, except for the $2FBG_{AB}$ and M-IMU cases. In fact, their *testing* median errors (0.131 V) and aggregated-subject flow RMSE (22.03 L/min, 18.84 L/min, respectively) revealed that noisier and less informative data from isolated sensors deteriorate the actual flow prediction and the algorithm generalization capability, as well. It is worth mentioning that this severe effect was remarkable mainly for one subject. In this respect, a deeper investigation on the *testing* results revealed that the M-IMU and $2FBG_{AB}$ correlation coefficients were 0.26 and 0.43 respectively, whereas they were fair for the other subjects (between 0.6 and 0.8). These results suggested that, given the high inter-subject variability of both respiratory patterns and anatomical features [68,70], some sensor configurations could be more suitable for a specific subject than other ones. In this scenario, the proposed method might help in finding the optimal device arrangement with a one-time recording session and just a few fast *meta-testing* stages.

In order to further investigate the robustness of the algorithm results in the different proposed device configurations, two relevant respiratory parameters, i.e., RR and RV, from both the target and the estimated signals, were calculated and further compared. The estimated RR vs. the target RR comparisons revealed differences mainly with the configurations including the $2FBG_{AB}$ and the M-IMU sensors alone or in combination with each other, confirming the above reported considerations. Conversely, the configuration errors resulted to be statistically similar. Nevertheless, the RR error root mean squares (Fig. 7) showed remarkable differences between the high-correlation configurations and the other ones, for the *apnea*, *eupnea*, *tachypnea*, *standing*, *sitting* and *lying on the back* groups. The *walk* outcomes were a special case, since the $2FBG_{TX}$ and $2FBG_{TX}$ + M-IMU showed on average a substantial error (13.2 bpm and 11.9 bpm, respectively), against the lower 4FBG + M-IMU, 4FBG and $2FBG_{AB}$ + M-IMU ones (3.2 bpm, 5.9 bpm and 4.5 bpm, respectively). This suggested that the positioning of multiple sensors in different trunk areas could help in decoupling walk-related artifacts and respiratory movements for robust RR computation. Nevertheless, the rate estimation strongly depended on the calculation procedure, which was quite simple. However, the more accurate the signal fitting is the more reliable the RR computation might be, as it happened for the 4FBG + M-IMU configuration. In the other cases, ripples on the peaks, arising

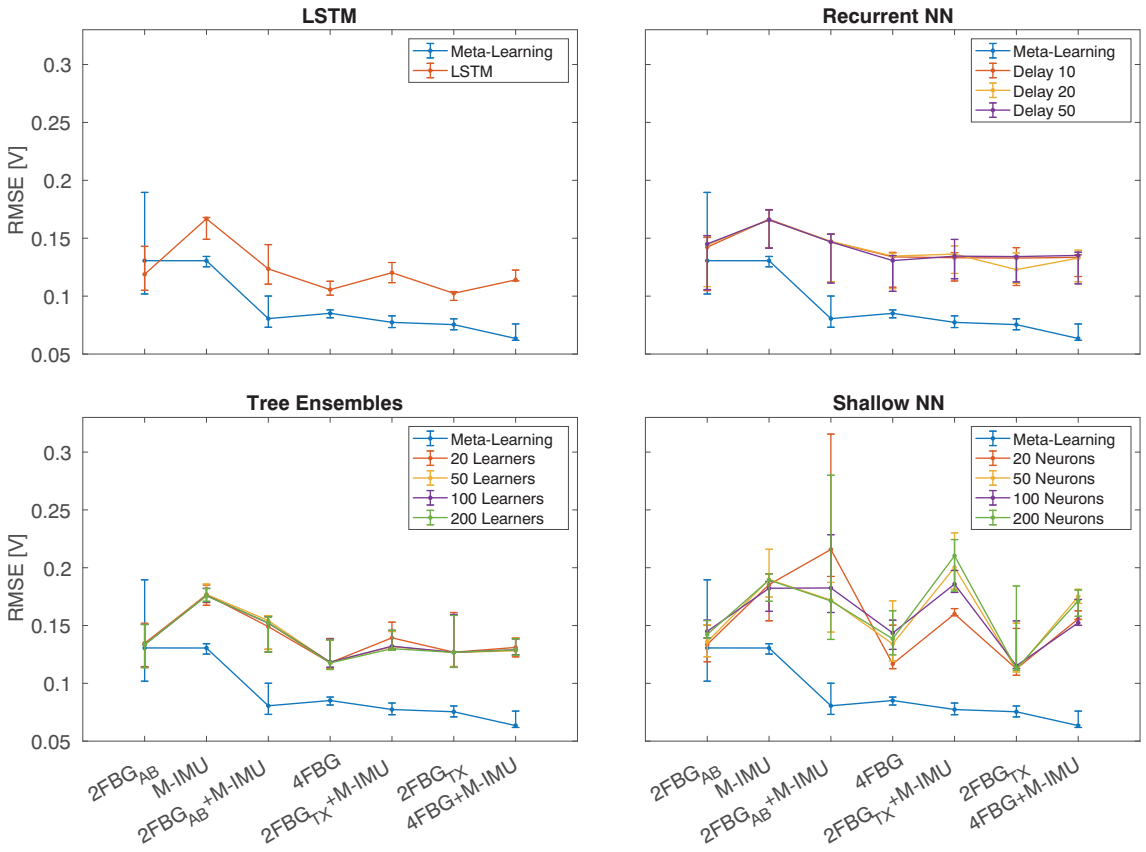

Fig. 7. Benchmarking algorithm performance. For each approach, the testing across-subject median and interquartile range RMSE [V] across all the sensor configurations is illustrated in comparison with the meta-learning results.

Table 2
Meta-learning and test errors achieved with benchmarking algorithms. The testing aggregated-subject flow RMSEs are reported for each approach.

		Root mean square errors [L/min]						
		$2FBG_{AB}$	M-IMU	$2FBG_{AB}$ + M-IMU	4FBG	$2FBG_{TX}$ + M-IMU	$2FBG_{TX}$	4FBG + M-IMU
Meta-learning		22.03	18.84	12.65	12.29	11.31	10.99	9.97
LSTM		17.24	22.45	17.94	14.85	16.17	13.57	16.32
Recurrent NN	Delay 10	19.14	23.23	19.83	18.15	18.40	18.50	18.49
	Delay 20	19.43	23.17	19.84	18.04	19.23	18.10	18.51
	Delay 50	19.36	23.18	19.75	17.71	19.36	18.15	18.34
Tree ensembles	20 learners	19.55	25.60	21.07	18.25	20.56	20.14	19.04
	50 learners	19.43	25.80	21.19	18.08	19.78	20.01	19.07
	100 learners	19.43	25.55	20.95	18.26	19.88	20.00	19.05
	200 learners	19.41	25.58	20.94	18.09	19.83	19.93	19.04
Shallow NN	20 neurons	19.65	25.68	37.64	19.10	23.40	18.53	23.00
	50 neurons	19.95	28.41	24.38	21.31	29.99	19.05	25.42
	100 neurons	21.30	25.59	28.57	20.70	27.27	19.40	23.32
	200 neurons	21.23	26.67	31.92	20.97	29.64	21.87	24.64

from dynamic tasks, may compromise the RR computation robustness. For what concerns the RV analysis, the target vs. predicted significant differences changed across the activities, but the complete configuration was always comparable to the reference. The RV error analysis revealed that the M-IMU and the $2FBG_{AB}$ configurations were similar in the case of *eupnea* and that the *lying on the back* condition resulted in differences between configurations of sensors that were placed in different trunk areas. This outcome was consistent with the above-mentioned hypotheses about the changes of lung capacity depending on postures [69]. Post hoc comparisons with the Bonferroni correction for the other tasks did not result in any differences. However, the error root mean square values were lower for the high-correlation configurations, as expected, being the volume directly derived from the flow. Fig. 8 reports analogous RV *standing* error values across all the configurations, whereas more evident differences between the best and the worst sensor configurations (i.e., $2FBG_{AB}$ and M-IMU) occurred for the *sitting* and the *lying on the back* postures, confirming again their effect on lung capacity [69]. Overall, the errors for each activity group were comparable across the best configurations, thus suggesting that different arrangements could be suitable for the volume estimation from the predicted flow.

4.2. Limits and perspectives

The achieved results of the meta-learning algorithm were promising when dealing with some respiratory patterns and different tasks. Nevertheless, the tested experimental conditions did not include more

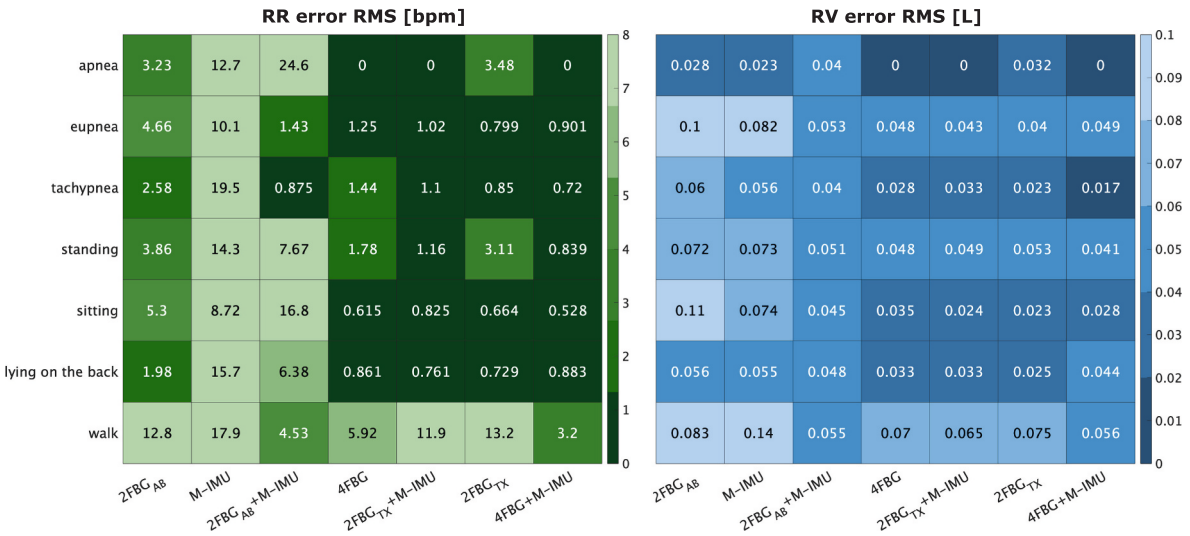

Fig. 8. RR (on the left) and RV (on the right) root mean square errors values for each device configuration (x axis) and each activity (y axis). Cell color intensities are scaled according to each RMS weight. Darker cells represent lower error values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

challenging activities which might affect the accuracy of the results. In particular, wider limb movements, trunk twisting or flexion, and posture changes which elicit additional strains on the FBGs were not taken into account. The *meta-training*, as well as the *testing* datasets should include more noisy information in order to ultimately evaluate the actual robustness of the flow prediction. The proper positioning of inertial units or even other FBG sensors for differential data capture may help to enhance the algorithm. Actually, the M-IMU was placed close to the sternum and, as it has been proven, it was sensitive to both whole body and chest movements. In order to more effectively decouple the two effects, its repositioning should be devised. Different solutions could be investigated through the implementation of this algorithm, with just one short data collection session. Along with the limited dynamic conditions explored, only a few respiratory rates and depths were investigated. However, the present results suggested that different patterns can be reliably reconstructed, since transitions between different rates were precisely tracked (Fig. 6). Further experimental conditions are hence required and tests with longer time series and multiple device donning sessions for repeatability evaluation should be performed as well. Larger *meta-training* dataset, encompassing more and heterogeneous subjects and more respiratory behaviors, are expected to improve the overall performance and speed up the calibration phase. Furthermore, starting from a well-established knowledge of the problem, generalization properties of each specifically tuned model may be enhanced. Hence, brand-new patterns could be predicted through inference. This overcomes the major concern of the tiresome procedures of the traditional calibration-dependent approaches [30,44]. Nonetheless, although the proposed algorithm may be robust to new conditions, the user's physical or respiratory behavioral changes (because of a disease onset, for example) may require further model adjustments. Additional evidence of the potentiality of this meta-learning algorithm results was bolstered via the estimation of the respiratory parameters. Referring to the complete sensor configuration, i.e., 4FBG + M-IMU, which reached the highest correlation with the target, RR and RV were computed with low errors. However, further improvements are needed since the measurements for the dynamic task were less accurate. Volume estimates presented errors which on average ranged from 5% to 11% (considering the target median absolute volume values, i.e., 0.6 L, 0.33 L and 0.5 L for *eupnea*, *tachypnea* and *walk*, respectively). This performance-related outcome might be satisfactory for other applications besides the most crucial clinical settings, where the maximum accuracy is not required (e.g., telemedicine, homecare, remote monitoring). Nevertheless, these preliminary results suggested that an overall assessment of the user's status could be however reliably provided, even though more robust outcomes might be achieved through more extensive model training. Further improvements would primarily deal with the implementation of more sophisticated intelligent solutions for accurate real-time estimation of the respiratory parameters of interest, even training *end-to-end* neural networks. Then, in the perspective of real-life application scenarios, the enhancement of the wearable system robustness and portability would be devised. Firstly, the targeted future enhancement of the technology readiness level would require the implementation of temperature-compensation strategies. As an example, the positioning of additional FBG sensors in a way to not make them sensitive to mechanical effects, could be a solution to retrieve temperature-induced variations of the signals. In addition, more reproducible positioning of the required sensors should be implemented through subject-specific reference points on the instrumented garments. More importantly, given the remarkable constraint resulting from the bulky interrogation unit, portable and wireless devices [71] should be considered. Moreover, the brand-new on-chip FBG interrogators [72] pave the way to even more integrated and less expensive systems, which would be completely provided with edge computing hardware. In this way, the implementation of feedback and warning strategies for safety, health preservation and rehabilitation purposes, relying on FBG-based wearable devices, would be enabled for varied real-life applications.

Funding

This work was supported in part by INAIL, the Italian National Institute for Insurance against Accidents at Work, under the BRIC 2018 SENSE-RISC Project (ID10/2018), in part by the Ministry of Health, Italy (Fondi IRCCS Ricerca Corrente), and in part by the UCBM (Università Campus Bio-Medico di Roma) under the University Strategic HOPE (HOspital to the PatiEnt) Project.

Declaration of competing interest

The authors declare no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.artmed.2022.102328.

References

[1] Tipton MJ, Harper A, Paton JFR, Costello JT. The human ventilatory response to stress: rate or depth? JPhysiol 2017;595:5729–52.
[2] Scott JB, Kaur R. Monitoring breathing frequency, pattern, and effort. Respir Care 2020;65:793–806.
[3] Jaana M, Pare G, Sicotte C. Home telemonitoring for respiratory conditions: a systematic review. AmJManagCare 2009;15:313–20.
[4] Lynn LA, Curry JP. Patterns of unexpected in-hospital deaths: a root cause analysis. Patient Saf Surg 2011;5:3.
[5] Estève D, Chan M, Fourniols J-Y, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012;56:137–56.
[6] Risso NA, et al. A cloud-based mobile system to improve respiratory therapy services at home. J Biomed Inform Oct. 2016;63:45–53.
[7] Javorka M, et al. Role of respiration in the cardiovascular response to orthostatic and mental stress. Am J PhysiolRegul Integr Comp Physiol 2018;314:R761–9.
[8] Joyce S, et al. Workplace interventions for common mental disorders: a systematic meta-review. Psychol Med 2016;46:683–97.
[9] Gatti UC, Schneider S, Migliaccio GC. Physiological condition monitoring of construction workers. Autom Constr 2014;44:227–33.
[10] Can YS, Arnrich B, Ersoy C. Stress detection in daily life scenarios using smart phones and wearable sensors: a survey. J Biomed Inform Apr. 2019;92:103139.
[11] Miller MR, et al. Standardisation of spirometry. Eur Respir J 2005;26:319–38.
[12] Bloch KE, Barandun J, Sackner MA. Effect of mouthpiece breathing on cardiorespiratory response to intense exercise. Am J Respir Crit Care Med 1995; 151:1087–92.
[13] Mannée DC, de Jongh F, van Helvoort H. Telemonitoring techniques for lung volume measurement: accuracy,artifacts and effort. Front Digit Health 2020;2.
[14] Folke M, Cernerud L, Ekström M, Hök B. Critical review of non-invasive respiratory monitoring in medical care. MedBiolEngComput 2003;41:377–83.
[15] Vanegas E, Igual R, Plaza I. Sensing systems for respiration monitoring: a technical systematic review. Sensors (Switzerland) 2020;20:1–84.
[16] Konno K, Mead J. Measurement of the separate volume changes of rib cage and abdomen during breathing. J Appl Physiol 1967;22:407–22.
[17] Cohn MA, et al. The respiratory inductive plethysmograph: a new non-invasive monitor of respiration. Clin RespirPhysiol 1982;18:643–58.
[18] Mead J, Peterson N, Grimby G, Mead J. Pulmonary ventilation measured from body surface movements. Science (80-.) 1967;156:1383–4.
[19] Henderson RP, Webster JG. An impedance camera for spatially specific measurements of the thorax. IEEE Trans Biomed Eng 1978;25:250–4.
[20] Karacocuk G, et al. Inertial sensor-based respiration analysis. IEEE Trans Instrum Meas 2019;68:4268–75.
[21] Chu M, et al. Respiration rate and volume measurements using wearable strain sensors. npj DigitMed 2019;2:8.
[22] Hu Hfeng, Sun Sjia, Lv Rqing, Zhao Y. Design and experiment of an optical fiber micro bend sensor for respiration monitoring. Sensors Actuators A Phys 2016;251: 126–33.
[23] Presti DLo, et al. Wearable system based on flexible FBG for respiratory and cardiac monitoring. IEEE Sensors J May 2019;PP:1.
[24] Massaroni C, Saccomandi P, Schena E. Medical smart textiles based on fiber optic technology: an overview. J Funct Biomater 2015;6:204–21.
[25] Gong Z, et al. Wearable fiber optic technology based on smart textile: a review. Materials 2019;12:3311.
[26] Koyama Y, Nishiyama M, Watanabe K. Smart textile using hetero-core optical fiber for heartbeat and respiration monitoring. IEEE Sensors J 2018;18:6175–80.
[27] Cretikos MA, Bellomo R, Hillman K, Chen J, Finfer S, Flabouris A. Respiratory rate: the neglected vital sign. Med J Aust 2008;188:657–9.
[28] Whited L, Graham DD. Abnormal respirations. 2018.
[29] Braun SR. Respiratory rate and pattern. 1990.
[30] Sharma P, Zhou J, Conroy TB, Kan EC. Wearable radio-frequency sensing of respiratory rate, respiratory volume, and heart rate. npj DigitMed 2020;3:98.
[31] Ramos-Garcia RI, Imtiaz MH, Sazonov E, Tiffany ST. Evaluation of RIP sensor calibration stability for daily estimation of lung volume. In: Proceedings of the International Conference on Sensing Technology. ICST; 2018.
[32] Petrović MD, et al. Non-invasive respiratory monitoring using long-period fiber grating sensors. Biomed Opt Express 2014;5:1136–44.
[33] Witt J, et al. Medical textiles with embedded fiber optic sensors for monitoring of respiratory movement. IEEE Sensors J 2012;12:246–54.
[34] Kam W, et al. Compact and low-cost optical fiber respiratory monitoring sensor based on intensity interrogation. J Light Technol 2017;35:4567–73.
[35] Massaroni C, et al. Design and feasibility assessment of a magnetic resonance-compatible smart textile based on fiber Bragg grating sensors for respiratory monitoring. IEEE Sensors J 2016;16:8103–10.
[36] Ciocchetti M, et al. Smart textile based on fiber Bragg grating sensors for respiratory monitoring: design and preliminary trials. Biosensors 2015;5:602–15.
[37] Massaroni C, et al. Smart textile for respiratory monitoring and thoraco-abdominal motion pattern evaluation. J Biophotonics 2018;11:e201700263.
[38] Houssein A, Ge D, Gastinger S, Dumond R, Prioux J. Estimation of respiratory variables from thoracoabdominal breathing distance: a review of different techniques and calibration methods. Physiol Meas 2019;40:03TR01.
[39] Reyes BA, Reljin N, Kong Y, Nam Y, Chon KH. Tidal volume and instantaneous respiration rate estimation using a volumetric surrogate signal acquired via a smartphone camera. IEEE J Biomed Health Informa 2017;21:764–77.
[40] De Groote A, Paiva M, Verbandt Y. Mathematical assessment of qualitative diagnostic calibration for respiratory inductive plethysmography. J Appl Physiol 2001;90:1025–30.
[41] Donaldson GC. The chaotic behaviour of resting human respiration. Respir Physiol 1992;88:313–21.
[42] Suki B, et al. Fluctuations, noise and scaling in the cardio-pulmonary system. Fluct Noise Lett 2003;03:R1–25.
[43] Raoufy MR, Hajizadeh S, Gharibzadeh S, Mani AR, Eftekhari P, Masjedi MR. Nonlinear model for estimating respiratory volume based on thoracoabdominal breathing movements. Respirology 2013;18:108–16.
[44] Dumond R, et al. Estimation of respiratory volume from thoracoabdominal breathing distances: comparison of two models of machine learning. Eur J Appl Physiol 2017;117:1533–55.
[45] McClure K, Erdreich B, Bates JHT, McGinnis RS, Masquelin A, Wshah S. Classification and detection of breathing patterns with wearable sensors and deep learning. Sensors (Switzerland) 2020;20:6481.
[46] Hafezi M, et al. Sleep apnea severity estimation from tracheal movements using a deep learning model. IEEEAccess 2020;8:22641–9.
[47] Biswal S, Sun H, Goparaju B, Westover MBrandon, Sun J, Bianchi MT. Expert-level sleep scoring with deep neural networks. J Am Med Informa Assoc Dec. 2018;25 (12):1643–50.
[48] Sun H, et al. Sleep staging from electrocardiography and respiration with deep learning. arXiv 2019;43:zsz306.
[49] Ravichandran V, et al. RespNet: a deep learning model for extraction of respiration from photoplethysmogram. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBS; 2019.
[50] Petrenko A, Kyslyi R, Pysmennyi I. Detection of human respiration patterns using deep convolution neural networks. EastEurJEnterpTechnol 2018;4:6–13.
[51] Vanschoren J. Meta-learning: a survey. arXiv; 2018.
[52] Hospedales T, Antoniou A, Micaelli P, Storkey A. Meta-learning in neural networks: a survey. arXiv; 2020.
[53] Aguirre A, Pinto MJ, Cifuentes CA, Perdomo O, Díaz CAR, Múnera M. Machine learning approach for fatigue estimation in sit-to-stand exercise. Sensors 2021;21 (15):5006.
[54] Di Tocco J, et al. A wearable system based on flexible sensors for unobtrusive respiratory monitoring in occupational settings. IEEE Sensors J 2020;21:14369–78.
[55] Quandt BM, et al. Body-monitoring and health supervision by means of optical fiber-based sensing systems in medical textiles. Adv Healthc Mater Feb. 2015;4(3): 330–55.
[56] Bao Y, et al. High-speed liquid crystal fiber Fabry-Perot tunable filter. IEEE Photonics Technol Lett 1996;8:1190–2.
[57] Miller CM. Characteristics and applications of high performance, tunable, fiber Fabry-Perot filters. ProcElectronComponents Conf 1991:489–92.
[58] Majumder M, Gangopadhyay TK, Chakraborty AK, Dasgupta K, Bhattacharya DK. Fibre Bragg gratings in structural health monitoring-present status and applications. SensorsActuatorsAPhys 2008;147:150–64.
[59] Nichol A, Schulman J. Reptile: a scalable metalearning algorithm. arXiv; 2018.
[60] Nichol A, Achiam J, Schulman J. On first-order meta-learning algorithms. arXiv; 2018.
[61] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9: 1735–80.
[62] Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int J UncertainFuzinessKnowlBased Syst 1998;06:107–16.
[63] Hochreiter S, Schmidhuber J. LSTM can solve hard long time lag problems. In: Advances in neural information processing systems; 1997.
[64] Kingma DP, Ba JL. Adam: a method for stochastic optimization. In: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings; 2015.
[65] Berk RA. An introduction to ensemble methods for data analysis. Sociol Methods Res Feb. 2006;34(3):263–95.
[66] Elman JL. Finding structure in time. Cognit Sci Apr. 1990;14(2):179–211.
[67] Chen A, et al. Machine-learning enabled wireless wearable sensors to study individuality of respiratory behaviors. Biosens Bioelectron 2021;173:112799.
[68] Lutfi MF. The physiological basis and clinical significance of lung volume measurements. Multidiscip Respir Med 2017;12:3.
[69] Kera T, Maruyama H. The effect of posture on respiratory activity of the abdominal muscles. J Physiol Anthropol Appl Human Sci 2005;24:259–65.
[70] Benchetrit G. Breathing pattern in humans: diversity and individuality. Respir Physiol 2000;122:123–9.
[71] Mendoza EA, Esterkin Y, Andreas T. Low power, low cost, lightweight, multichannel optical fiber interrogation system for structural health management of rotor blades. In: AIAC18: 18th Australian International Aerospace Congress (2019): HUMS-11th Defence Science and Technology (DST) International Conference on Health and Usage Monitoring (HUMS 2019): ISSFD-27th International Symposium on Space Flight Dynamics (ISSFD); 2019. p. 861–6.
[72] Marin YE, et al. Integrated dynamic wavelength division multiplexed FBG sensor interrogator on a silicon photonic chip. J Light Technol 2019;37(18):4770–5.