Запропоновано метод розпізнавання в реальному часі типів (шаблонів) дихання пацієнта з ціллю моніторингу його стану і загроз для здоров'я, що є частковим випадком проблеми розпізнавання людських активностей (HAR). Метод заснований на застосуванні глибинного машинного навчання з допомогою згорткової нейронної мережі (CNN) для класифікації швидкості руху його грудної клітки. Показано, що прийняті при цьому рішення узгоджуються з технологією мобільної медицини (mHealth) з використання натільних датчиків і смартфонів для оброблення їх сигналів в якості обчислювальних edge-вузлів, але CNN відкривають важливі додаткові можливості з підвищенні якості оброблення сигналів датчика-акселерометра в умовах наявності перешкоджаючих сигналів (шумів) від інших джерел та інструментальних похибок пристрою. Вхідні сигнали попередньо нормалізується щодо осі обертання, щоб зменшити вплив шуму на результати, оскільки акселерометр вимірює гравітаційне прискорення (g) і лінійне прискорення (a). Запропоновано спосіб перетворення одновимірних сигналів (1d) акселерометра в двовимірні (2d) графічні зображення, які оброблюються за допомогою CNN із декількома обробними шарами, завдяки чому точність визначення шаблону дихання в різних ситуаціях для різних фізичних станів пацієнтів зростає в порівнянні з випадком, коли двовимірні перетворення сигналів акселерометра не вживаються. При цьому зростання точності (або якості) визначення різних типів дихання відбувається при збереженні достатньої швидкості процедур запланованого методу, що дозволяє проводити класифікацію типів дихання в реальному часі. Дану методику було випробувано в якості компоненту Body Sensor Network (BSN) і встановлено високу точність (88 %) визначення стану дихання пацієнта, що в поєднанні з даними контексту, отриманими з інших вузлів BSN, дозволяє визначати стани пацієнтів і передбачати загострення їх респіраторних хвороб

Ключові слова: акселерометр, глибинне навчання, шаблони дихання, згорткові нейронні мережі, машинне навчання

UDC 004.93

DOI: 10.15587/1729-4061.2018.139997

DETECTION OF HUMAN RESPIRATION PATTERNS USING DEEP CONVOLUTION NEURAL NETWORKS

A. Petrenko
Doctor of Technical Sciences, Professor, Head of Department*
E-mail: tolja.petrenko@gmail.com

R. Kyslyi
Postgraduate student
E-mail: kvrware@gmail.com

I. Pysmennyi
Postgraduate student
E-mail: ihor.pismennyy@gmail.com

*Department of System Design
Institute of Applied Systems Analysis
National Technical University of Ukraine
"Igor Sikorsky Kiev Polytechnic Institute"
Peremohy ave., 37, Kyiv, Ukraine, 03056

1. Introduction

The need to diagnose and monitor various respiratory diseases or their exacerbations contributed to the development of various methods for respiration measuring. At the same time, recent advances in the IoT technology and deep learning [1] created an opportunity to combine them and create a system of continuous observation. Such a system could be used to detect or predict and prevent exacerbations of dangerous conditions at various everyday human activities (such as walking, sleeping and other activities that alter the physical state).

Since the amount of air pollution in cities is constantly increasing, so does the number of people with respiration problems. Therefore, the possibility of permanent respiratory monitoring is using the accelerometer, which does not interfere with everyday activities, is studied. The use of such a system will make it possible to recognize the current state of a human, as well as predict potential aggravations.

Automatic detection and recognition of human respiratory patterns for health monitoring without any uncomfortable sensors that make continuous measurements impossible was a key problem for technologies that use analysis of respiration and behavior of a human. As a result, there appeared a decision to detect respiration rate, based on signals, received from a variety of body sensors in real time. At continuous measurement of respiration rate, signals received from the sensors, wearable on the body, is much more preferable than signals received from external sensors (for example, a thermal sensor, a pressure sensor, etc.), even if the latter are more accurately. The main reasons are listed below.

– Cameras, spirometers or other sensors, not subject to wearing out, suffer from environmental influence and are complicated to use. At the same time, sensors worn on the body do not, so they can carry out measurements much more frequently and even continuously, which provides more accurate results.

– A body sensor receives only target signals, while the signals from external sensors can be distorted by information from other objects in the environment. This leads to the need for more complex preprocessing of a signal (which is not always possible).

– Signals from an accelerometer can be used in combination with signals of other sensors to achieve greater accuracy or to get additional contextual information that can be used to recognize more complex images. For example, a more frequent respiration rate while running should be considered normal, in contrast to the norm in a quiet state. Moreover, this context is critical to differentiate various chronic conditions. For example, normal respiration is about 12 breaths per minute, or 6 l/min, whereas people with respiratory diseases breathe faster and more deeply. Such a deviation is also apparent for many other common chronic diseases.

The problem of detecting respiration patterns can be considered as a special case of the problem of human activity recognition. A key factor that affects the quality of the solution of this problem is a good preprocessing of signals, collected from sensors. Usually this processing includes basic conversion of a signal (for example, wavelet transformation or a Fourier transformation) or an interpretation of unprocessed signals using statistical methods, such as variance and mean value [2]. Even though these procedures are widely used in many problems with time series, they do not depend on the problem and therefore can also be used to recognize respiration patterns.

Thus, the development of a method for continuous measurement of respiration rate and determining a respiratory pattern by combining the use of signal preprocessing and convolution neural networks (CNN) in real time is a relevant problem for exploratory research.

2. Literature review and problem statement

There are many different methods to measure the human respiration. For example, using a spirometer can be described as a perfect solution to this problem, as 100 % measurement accuracy is guaranteed [3]. However, a spirometer cannot be used permanently, since a patient must breathe into it, which it is impossible to do in everyday life.

There also exist many indirect methods for respiration analysis, for example, with determining of oxygen concentration in blood [4]. However, since oxygen concentration changes quite slowly, this method does not apply to detect a respiratory change in real time, and, therefore, a dangerous condition of a patient [5]. In addition, respiration patterns can be detected by the sounds of the trachea. This method is good for selection of deep respiration or respiration during such activity as running. However, this method is not very good to detect different patterns in quiet states, because a sound changes very slightly. Just because of the noise created by the environment or other human movements, it is very difficult to detect a respiration sound and its change [6].

Another way to measure respiration was proposed in [7]. Using a volume sensor based on a piezoelectric transducer, placed on the chest, a change in its circumference during respiration is measured. There are other known methods that require a large number of additional sensors or medical equipment, which are pretty accurate, but are not designed for continuous and daily use.

On the other hand, accelerometers, worn on the body, can measure angular changes during respiration, assessing acceleration of the thorax movement, and hence the respiration rate [8]. To do this, paper [8] applies the adaptive band-pass filter using the principal components analysis (PCA) combined with digital processing of signals. To improve the quality of a signal from most available accelerometers (for example on a user's smartphone or other devices available for carrying), the use of a Kalman filter is proposed in paper [9].

Based on the nature of a respiration signal, we consider it as a combination of several basic continuous motions, represented by signals of an accelerometer and a gyroscope (which did not influence accuracy of the experiment, and was not used after). The main problem with the use of this representation to detect respiration patterns is retrieving this pattern from a complex signal during the daily activities of a user.

Because computational power available on peripherals devices of the consumer market reached 16.6 GFLOPS, it is possible to use deep learning on intelligent sensors or a user's device and consequently to deliver messages and recommendations with a minimum delay and even offline, not depending on the server, or whether they are connected to the network.

At the same time, retrieving precise signals from noise polluted raw data of an accelerometer is still quite a complicated problem. A huge number of different routine actions, executed by people, make them difficult to detect.

3. The aim and objectives of the study

The purpose of this research is to determine 6 different types of respiration patterns, dependent on human activity.

To accomplish the set goal, the following tasks were set:

– generate a dataset for the experiment;

– undertake preliminary data preparation and cleaning;

– dentify respiration patterns, using CNN and collected data.

4. Data collection, their preprocessing and the construction of a neural network

To validate the proposed solutions and the hypotheses, first, it is necessary to generate a dataset. For this purpose, we developed an application for Android, which collects data from the accelerometer, attached to the chest of a user, and passes the data to the smartphone using Bluetooth with a discretization frequency of 30 Hz (30 values per second). To get different respiration patterns, data are updated after various physical actions of a patient that include quiet respiration, respiration after push-ups, respiration during running, deep respiration, fast respiration, respiration with a delay. An accelerometer was attached over the ribs under the nipple (Fig. 1) using a scotch tape. Previously, by changing the position of an accelerometer on the chest, it was found that this position gives the greatest amplitude of the moving chest and, as a consequence, the clearest respiration signal. In addition, the same place for an accelerometer was used in paper [10].

To obtain more balanced data, the test group of 8 users (4 people of each gender) was studied, four of them were 25 years old, two of them were 40 years old and two more were 55 years old. All users, included in the test group, had a medium level of physical activity. The measurements were made with the use of a smartphone with the developed application and a sensor, attached to the chest. Each action was recorded within 5 minutes (data after push-ups were recorded within 30 minutes and a series of 10 push-ups was repeated). 9,000 values for each type of activity for each person were recorded within 5 minutes when accelerometer data with the frequency of 30 Hz (30 values per second) were recorded. Thus, a complete set of data includes 432,000 values. Because the data, taken from one person, are fully used for dataset testing, it is supposed to be enough to test the methods and hypotheses.

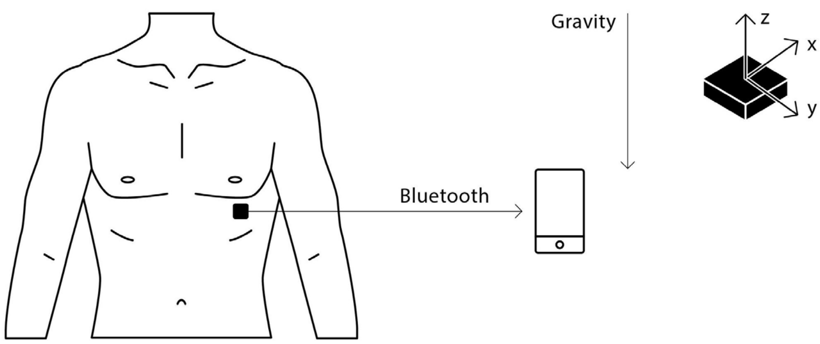

Fig. 1. General scheme of the system and the point where an accelerometer is attached

The data acquired from the sensor have the following structure: a timestamp (in the Unix format with milliseconds), their acceleration along the X axis, acceleration along Y axis, and acceleration along the Z axis.

1486291984848|1.61407470703125|–
–9.95855712890625|2.83917236328125.

After the recording session, the data were stored on a mobile phone, and then marked by adding an activity column. The resulting dataset was divided into parts for testing and training. Because measurements were conducted with generalized data, describing human activities, individual data, collected from one of the participants, were stored separately for subsequent verification of working capacity of the model. Because different people have different lungs volume, chest movement and other physiological parameters, this means that if the technique is valid for a person, not included in the test group, it can be used in relation to other people.

Convolution neural networks are used successfully for image recognition problems [11]. An accelerometer produces one-dimensional signals unlike two-dimensional images. It was therefore decided to convert a received signal in a 2D-image and compare the results with the traditional approach, using a 1D-signal.

Conversion of 1D-signal into a 2D-image can be performed in several ways, inter alia, by using normalization and transformation into the matrix. Using this approach, samples of a signal inside the observation window are converted into a grey color image with the scale of 0–255. At this image, a darker color means a greater amplitude in the original signal and pixel coordinates – (i, j) of the M×N matrix for each of n-th sample of a signal where

$$i=\frac{N}{M} \text{ and } j=\frac{N}{M}. \qquad (1)$$

Fig. 2 shows this approach.

This approach is simple and convenient to work with because it does not require great computing power, but at the same time, has a considerable drawback: it is very susceptible to noise and anomalies, so it cannot be used for our data.

As shown in [7], an alternative can be an approach, when signals are added together sequentially by stacking as an image, which allows each sequence of signals to correlate with other sequences. Then the 2D discrete Fourier transform (DFT) was performed, and its magnitude is an image that is used as the input signal of CNN. This approach is modified by inclusion of signals normalization procedure before stacking into a single image.

Because the accelerometer measures gravitational acceleration (g) and linear acceleration (a), as people move in different directions, the accelerometer rotates and each axis measures different acceleration (g or a). To continuously measure the motion on axis a and reduce noise (exclude irrelevant measurements), it is necessary to normalize the measurement values.

The θt angle shows how axis rotation changes over time from t–1 to t:

$$\theta_t=\cos^{-1}\left(a_t \times a_{t-1}\right), \qquad (2)$$

where a is the vector of values over time. To reduce the effect of noise on results, each observation is normalized by rotation angle θt, at time t. Thus, a normalized signal will take the form:

$$a_{xt}=a_x \times \theta_t,\ \ a_{yt}=a_y \times \theta_t,\ \ a_{zt}=a_z \times \theta_t. \qquad (3)$$

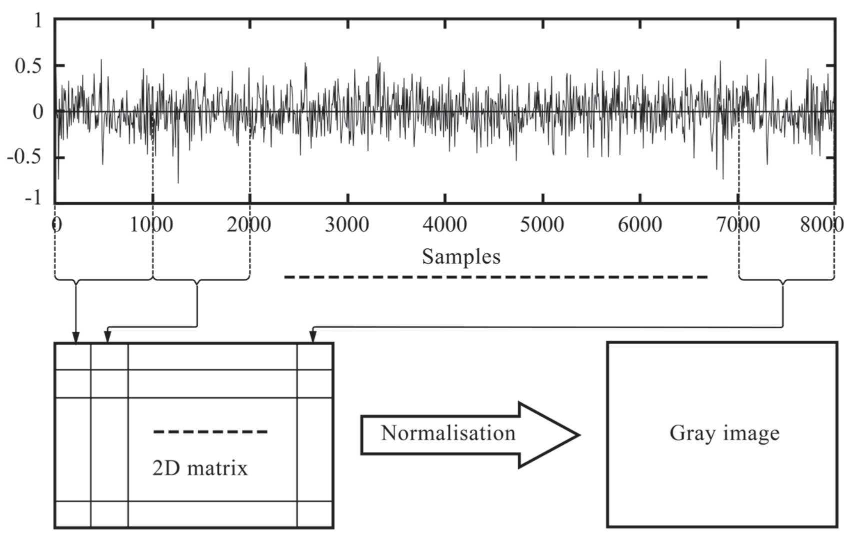

Fig. 2. Conversion of signal into a matrix [12]

Since the sliding window method is used, the Hamming function H(n) for the normalized signal was applied for the purpose of noise reduction.

As a result, the algorithm of preprocessing of an unprocessed input signal takes the following form:

1. To normalize the input signal relative to the rotation axis as described above.
2. To apply the Hamming window function for already normalized signal.
3. To convert the signal from our window using the algorithm from [7].
4. To apply DFT for the obtained image.

Upon implementing a given approach to the data from sensors, results are transmitted to CNN. The process of signal processing is shown in Fig. 3.

To evaluate how the proposed method of preprocessing improves the quality of pattern detection, it was compared to the case of using CNN for the normalized signal as input data [14].

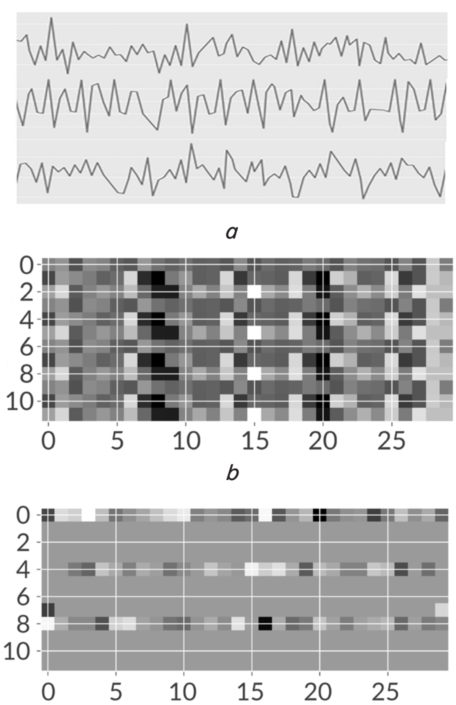

Fig. 3. Processing the accelerometer signal: *a* — incoming signal; *b* — signal image; *c* — amplitude spectrum

Convolution neural networks (CNN) are one of the most reliable, flexible and widely used methods for detection of signal patterns. At higher layers, different features of signals are detected (for example, in this case for characteristic of each chest movement). Deeper layers receive schemes of signals in high-level representation (for multiple movements). Each level can have some operators of convolutions or associations, therefore, several patterns, learnt from different aspects, can be used together. As a result, it is possible to detect more complex patterns compared to other traditional statistical methods such as Support Vector Machines (SVN), Random Forest or others [15].

For experiments, we used the sliding window strategy to process the signal both applied to the original signal and for the case of image generation from an input signal. The basic idea of a sliding window is to divide the signal of time series into short fragments.

When the data are acquired at a frequency of 30 observations per second, the sliding window size that is equal to 90 samples corresponds to 3 seconds of observations and a window step size that is equal to 32 samples is 1.06 seconds. A larger or smaller pitch size can be used either to increase accuracy (lower) or to decrease required computing power (larger), or energy consumption, which is a critical parameter. Therefore, a compromise between accuracy and the number of computations must be found.

Since the main objective of the study is to identify the respiration type in real time with minimal delay, a higher computational cost can be a problem, causing large delays (especially when applying not very powerful devices). For the first experiment with input data in the form of a normalized signal, window segments are generated and a third dimension is added to each component, so that the input vector for CNN should contain full segments, an input segment, a channel entrance (Fig. 4).

For the first experiment with input in the form of a normalized signal, window segments are generated and added to each component of the signal in the form of a third dimension, so that the input vector for CNN should take the following form: [complete segments, an input segment, an input channel].

Since the signal is one-dimensional and it implies the 1D convolution, it is necessary to change the generated windows that will be normalized to height 1.

For the second experiment with stacking and signal processing as a graphic image, changes are introduced. Since there are 3 sequences of input signals, in accordance with the used algorithm, each signal is repeated four times. As a result, the size of the input signal is equal to 12×32, where parameter 32 is determined by the selected window size.

The model consists of one convolution layer, followed by the maximum pooling layer, and another convolution layer. After that the model contains a fully-connected layer, which is connected to a Softmax layer (in problems of classification or pattern detection, the last CNN layer is usually is a Softmax layer – multilinear logistics regression). The described neural network architecture is shown in Fig. 4.

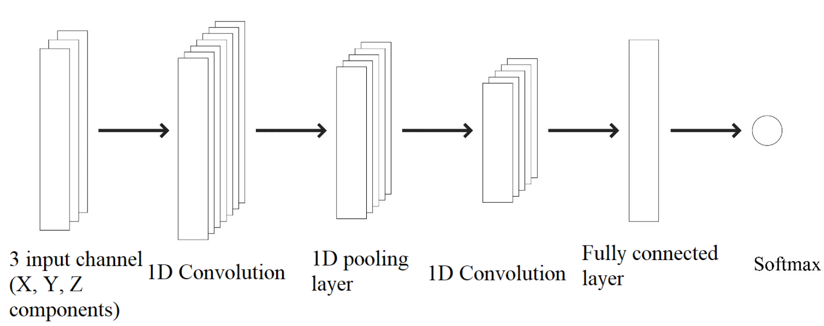

Fig. 4. CNN architecture for the first experiment

CNN are trained using iterative optimization with the help of the backpropagation algorithm. The most widely used method of optimization is stochastic gradient descent (SGD) and YellowFin optimizer [16]. In the experiment, YellowFin optimizer was used because it has the lowest learning error and dynamic learning speed, while SGD usually accumulates an error after 300 epochs [16].

The function of training cost of our CNN architecture is Softmax with L2 regularization. The rectified linear unit (ReLU) was used in the study as an activation function. CNN hyperparameters are:

– number and types of layers;

– size of filters for convolution and convolution pitch for each pooling layer;

– size of pooling domain and pooling pitch for each pooling layer;

– number of units for each fully connected layer.

The CNN, shown in Fig. 5, was constructed based on the above operators. All CNN layers can be grouped into five sections, as described below.

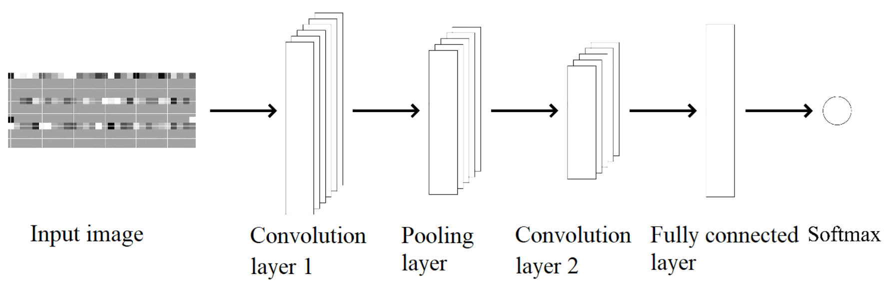

Fig. 5. CNN architecture for the preprocessed input signal

The first convolution layer has the filter size and depth 60 (the number of channels is obtained as the output from a convolution layer). The size of a pooling layer filter is set equal to 20 with the pitch of 2. Then a convolution layer takes an input layer with the maximum pooling level, applying the size 6 filter and it will have one tenth of the maximum depth of maximum pooling layer. After this the output is smoothed for the vector of entrance of a fully connected layer.

1,000 neurons are placed in a fully connected layer (this can be determined by a configuration). This layer uses the hyperbolic tangent function tanh for nonlinearity. Softmax layer is determined for output of probabilities of class marks. Function of negative log likelihood [17] is minimized by using stochastic gradient descent (SGD).

The use of the source input (with some preprocessing) in the CNN architectures is a normal situation for applications of deep learning in subject domain of computer vision. CNN, however, are generally not as effective with 1-dimensional signal and for the second experiment, the signal was converted into a 2D-image.

For the second experiment with the signal, converted to a 2D-image, the first and second CNN convolution layers (Fig. 5) perform two-dimensional convolution on their inputs. The output map is generated as follows:

$$y_i = \left(1+\exp\left(b_j + \sum_i k_{ij} \times x_i\right)\right)^{-1}, \tag{4}$$

where k_{ij} is the convolution kernel on the i-th input display x_i for generation of the j-th output display y_j, b_j.

The pooling layer is arranged as the sub-sampling method. Output y_i is calculated by accepting the mean values of non-overlapping regions x_i with filter m x m.

$$y_i(r,c) = \frac{\sum_{p=1}^{m}\sum_{q=1}^{m} x_i(r\times m + p, c\times m + q)}{m^2}, \tag{5}$$

where r, c are the coordinates of pixel y_i.

After completely connected layer, 1D-vector f at the output is obtained. Function Softmax is applied to determine the probability of each class, which corresponds to probability of each of 6 different respiration types:

$$p(s) = \frac{g_s}{\sum_{j=1}^{N_a} g_j},\ g_j = \max\left(0, \sum_i f_i \times w_{ij} + h_j\right), \tag{6}$$

where w and h are the coefficients of Softmax function, and s is one of the predictable classes.

5. Evaluation and comparison of accuracy of respiration patterns detection

The main problem for results evaluation was the problem of imbalance of classes. It is much easier to register a low physical activity for many, especially older people (who are the main target group of the developed solution). The result was obtaining much more data, recorded for quiet respiration than for other types. To reduce the effect of unbalanced classes, it is possible to use various unbalanced data groups for each epoch in YellowFin optimizer [18].

To assess the results of the constructed model, several indicators were calculated:

– Precision, recall;

– F1;

– Training loss;

– Learning accuracy.

Testing accuracy was verified using records from a person who was not added to the learning dataset, which helped avoid re-training on specific people.

Because in this research, 6 different classes (respiration patterns) with binary metrics are classified [19], it is necessary to display this classification as 6 different binary classifications, where each time one class is considered positive and all other classes are marked as negative (classification of one against all) [20]. It is facilitated by the construction of the confusion matrix, shown below in Fig. 6 in order to better understand efficiency of the constructed model.

Table 1 shows that after 40 learning epochs, the constructed model clearly started retraining, resulting in the level of 84 % accuracy of the model. The same can be seen in Table 2, 88 % accuracy was shown for the preprocessed signal, which is better than in the previous case, but still far from being perfect.

Table 1

Metrics using a normalized signal

Number of epochs	F1 score	Recall	Precision	Training Accuracy	Training Loss	Testing Accuracy
Epoch 0	0.615	0.67	0.6	0.67	7.82	0.45
Epoch 10	0.839	0.881	0.825	0.881	3.63	0.765
Epoch 20	0.942	0.947	0.941	0.947	2.653	0.818
Epoch 30	0.969	0.971	0.97	0.971	2.179	0.834
Epoch 40	0.976	0.977	0.977	0.977	1.81	0.843
Epoch 50	0.98	0.98	0.98	0.98	1.709	0.836

Table 2

Metrics using signal transformation into image

Number of epochs	F1 score	Recall	Precision	Training Accuracy	Training Loss	Testing Accuracy
Epoch 0	0.515	0.57	0.5	0.57	5.73	0.37
Epoch 10	0.727	0.732	0.749	0.76	4.02	0.65
Epoch 20	0.919	0.918	0.925	0.95	2.756	0.729
Epoch 30	0.959	0.951	0.939	0.961	2.217	0.8
Epoch 40	0.973	0.97	0.981	0.98	1.756	0.88
Epoch 50	0.983	0.983	0.983	0.983	1.69	0.873

Table 3

Confusion matrix of CNN for unprocessed signal

Respiration types	Quiet respiration	Respiration after push-ups	Deep respiration	Fast respiration	Respiration during running	Respiration with delay
Respiration with delay	0.97	0.028	0	0	0	0
Respiration during running	0.026	0.64	0.33	0	0	0
Fast respiration	0	0.2	0.75	0	0	0.055
Deep respiration	0	0	0	0.65	0.33	0.014
Respiration after push-ups	0	0	0	0.24	0.76	0
Quiet respiration	0	0.051	0.029	0	0	0.92

Table 4

Confusion matrix of CNN for a signal, converted into image

Respiration types	Quiet respiration	Respiration after push-ups	Deep respiration	Fast respiration	Respiration during running	Respiration with delay
Respiration with delay	0.95	0.041	0.014	0	0	0
Respiration during running	0.021	0.73	0.23	0.021	0	0
Fast respiration	0	0.19	0.72	0.017	0.069	0
Deep respiration	0	0.015	0.015	0.71	0.26	0
Respiration after push-ups	0	0.043	0	0.19	0.77	0
Quiet respiration	0	0	0.029	0	0	0.97

As can be seen from the confusion matrix, given in Tables 3, 4, the proposed approach can determine quite accurately the quiet respiration and the respiration with a delay, while a lot of erroneous starts are recorded when trying to classify other types of respiration. There is especially a lot of misclassification between running and fast respiration, as well as between push-ups and deep respiration. This can be explained by the fact that the signals from respiration with a delay and quiet respiration is less likely to be accompanied by other activities, associated with the thorax movement, and therefore are more unique. While during fast respiration and running (as well as push-ups and deep respiration) chest movement are more identical and, in addition, a received signal is interfered with by huge amount of data of physical activity, due to which a respiratory movement of the chest is harder to recognize.

6. Discussion of results of using CNN to determine the respiration patterns

The proposed method for recognition of respiration patterns and the proposed data preprocessing allows CNN to fix the patterns of an accelerometer signal in different situations and temporal scales with the use of 2D convolutions for obtaining high-level characteristics for the patterns detection. To prove that this method works better than the use of just an original signal as original data, these two approaches were compared when using the same dataset.

All determined patterns are standardized for prediction of the known respiration patterns. The key points, proposed in the present method, are:

– the use only one of body sensor to detect a respiration pattern;

– conversion of an accelerometer signal into a 2D-image using previous;

– normalization relative to the axis rotation;

– the use of deep learning methods to create a model that is independent of the characteristics extraction procedures;

– the possibility of using this approach to determine other types of human actions.

Using the proposed approach, it is possible to determine the respiration patterns with an 88 % accuracy, which can be considered a good result for using only one sensor.

At the same time, it is necessary to point out the limitations and shortcomings of the examined approach:

– when adding new patterns, it is necessary to modify the set of input data for neural network learning, and this will increase the size of the final model, which could make it difficult to use on mobile devices;

– if different respiration patterns are very similar to each other, the probability of recognition error increases.

It is also interesting to combine the obtained data with the data from other body sensors (for example, heat rate measuring) for forecasting more complicated human actions, which can become a promising subject-matter for subsequent studies. For example, the combination of the value of heart rate and current state of human respiration will enable an accurate prediction of the possibility of asthma attack. It is important to note that when adding data from other sensors, it is advisable to study the applicability of other architectures of neural networks excluding CNN.

7. Conclusions

1. To explore the possibility of automatic classification of respiration types (patterns) of patients, an input data array, taking into account different states of patients, is needed. This dataset was collected with the help of the body accelerometer, it contains 432,000 measurements of various types of respiration of 8 patients of all ages in their various states (quiet respiration, deep respiration, fast respiration, respiration with a delay, respiration after push-ups and after running). This is a decisive factor for the formation of the input data array, taking into account different states of patients. The input control data array, not involved in machine network learning, was additionally generated in order to validate the proposed method for the respiration patterns recognition.

2. To obtain a graphical representation of the generated dataset required for the effective use of CNN, it was proposed to convert 1D-signals of the accelerometer into the 2D graphical image using normalization and transformation into a matrix. An input signal is previously normalized relative to the rotation axis to reduce the effect of noise on the results because an accelerometer measures gravitational acceleration (g) and linear acceleration (a). It also uses the idea of sliding window, lying in separation of a signal of time series into short fragments. The values of a signal inside the window are processed by the discrete Fourier transform (DFT), with the help of which a grey image with the scale of 0–255, which is used as an input signal of SNN, is obtained.

3. The proposed method for respiration patterns classification using the CNN network can be realized in two versions. The first version is designed for processing 1D input data and is based on the interaction of layers with 1D convolution, and the second basic version – with interaction of layers with 1D and 2D convolutions through the pooling layer. Conducted experimental studies with the specified options have shown the advantages of the second option in terms of accuracy of recognition of respiration patterns, reaching 88 % on the selected arrays of indicators. This proves the usefulness of the proposed signal preprocessing when using deep learning in the problem of recognition of respiration patterns of patients.

References

1. Goodfellow I., Bengio Y., Courville A. Deep Learning. URL: http://www.deeplearningbook.org/
2. Huynh T., Schiele B. Analyzing features for activity recognition // Proceedings of the 2005 joint conference on Smart objects and ambient intelligence innovative context-aware services: usages and technologies – sOc-EUSAI '05. 2005. doi: https://doi.org/10.1145/1107548.1107591
3. SpiroSmart: Using a Microphone to Measure Lung Function on a Mobile Phone / Larson E. C., Goel M., Boriello G., Heltshe S., Rosenfeld M., Patel S. N. URL: https://homes.cs.washington.edu/~shwetak/papers/SpiroSmart.CR.Final.pdf
4. Shephard R. J. The oxygen cost of breathing during vigorous exercise // Quarterly Journal of Experimental Physiology and Cognate Medical Sciences. 1966. Vol. 51, Issue 4. P. 336–350. doi: https://doi.org/10.1113/expphysiol.1966.sp001868
5. Rakhimov A. Abnormal breathing pattern causes asthma and attacks. URL: https://www.worldwidehealth.com/health-article-Abnormal-breathing-pattern-causes-asthma-and-attacks.html
6. Movement analysis of the chest compartments and a real-time quality feedback during breathing therapy / Fekr A. R., Janidarmian M., Radecka K., Zilic Z. // In Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies. 2005.
7. Respiratory Rate and Flow Waveform Estimation from Tri-axial Accelerometer Data / Bates A., Ling M. J., Mann J., Arvind D. K. // 2010 International Conference on Body Sensor Networks. 2010. doi: https://doi.org/10.1109/bsn.2010.50
8. Phonospirometry for noninvasive measurement of ventilation: methodology and preliminary results / Que C.-L., Kolmaga C., Durand L.-G., Kelly S. M., Macklem P. T. // Journal of Applied Physiology. 2002. Vol. 93, Issue 4. P. 1515–1526. doi: https://doi.org/10.1152/japplphysiol.00028.2002
9. Estimation of Respiration Rate from Three-Dimensional Acceleration Data Based on Body Sensor Network / Liu G.-Z., Guo Y.-W., Zhu Q.-S., Huang B.-Y., Wang L. // Telemedicine and e-Health. 2011. Vol. 17, Issue 9. P. 705–711. doi: https://doi.org/10.1089/tmj.2011.0022

10. Improvement of Dynamic Respiration Monitoring Through Sensor Fusion of Accelerometer and Gyro-sensor / Yoon J.-W., Noh Y.-S., Kwon Y.-S., Kim W.-K., Yoon H.-R. // Journal of Electrical Engineering and Technology. 2014. Vol. 9, Issue 1. P. 334–343. doi: https://doi.org/10.5370/jeet.2014.9.1.334
11. Performance evaluation of a tri-axial accelerometry-based respiration monitoring for ambient assisted living / Jin A., Yin B., Morren G., Duric H., Aarts R. M. // 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2009. doi: https://doi.org/10.1109/iembs.2009.5333116
12. Uddin J., Van D. N., Kim J.-M. Accelerating 2D Fault Diagnosis of an Induction Motor using a Graphics Processing Unit // International Journal of Multimedia and Ubiquitous Engineering. 2015. Vol. 10, Issue 1. P. 341–352. doi: https://doi.org/10.14257/ijmue.2015.10.1.32
13. Wireless breathing system for long term telemonitoring of respiratory activity / Ciobotariu R., Adochiei F., Rotariu C., Costin H. // Advanced topics in electrical engineering, Proceedings of the 7th international symposium ATEE. 2011. P. 635–638.
14. Bulling A., Blanke U., Schiele B. A tutorial on human activity recognition using body-worn inertial sensors // ACM Computing Surveys. 2014. Vol. 46, Issue 3. P. 1–33. doi: https://doi.org/10.1145/2499621
15. Zhang J., Mitliagkas I. YellowFin and the Art of Momentum Tuning. URL: https://arxiv.org/pdf/1706.03471.pdf
16. Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition / Yang J. B., Nguyen M. N., San P. P., Li X. L., Krishnaswamy S. // Proceeding IJCAI'15 Proceedings of the 24th International Conference on Artificial Intelligence. 2015. P. 3995–4001.
17. Convolutional Neural Networks for Human Activity Recognition using Mobile Sensors / Zeng M., Nguyen L. T., Yu B., Mengshoel O. J., Zhu J., Wu P., Zhang J. // Proceedings of the 6th International Conference on Mobile Computing, Applications and Services. 2014. doi: https://doi.org/10.4108/icst.mobicase.2014.257786
18. Jiang W., Yin Z. Human Activity Recognition Using Wearable Sensors by Deep Convolutional Neural Networks // Proceedings of the 23rd ACM international conference on Multimedia – MM '15. doi: https://doi.org/10.1145/2733373.2806333
19. Ordóñez F., Roggen D. Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition // Sensors. 2016. Vol. 16, Issue 1. P. 115. doi: https://doi.org/10.3390/s16010115