Classification of Subjects With Balance Disorders Using 1D-CNN and Inertial Sensors

JAN ANDRZEJ NAPIERALSKI1, WOJCIECH TYLMAN1, RAFAŁ KOTAS1, PAWEŁ MARCINIAK1, MAREK KAMIŃSKI1, MAGDALENA JANC2, MAGDALENA JÓZEFOWICZ-KORCZYŃSKA3, AND EWA ZAMYSŁOWSKA-SZMYTEK2

1Department of Microelectronics and Computer Science, Lodz University of Technology, 93-005 Łódź, Poland
2Audiology and Phoniatrics Clinic, Nofer Institute of Occupational Medicine, 91-348 Łódź, Poland
3Department of Otolaryngology, Balance Disorders Unit, Medical University of Lodz, 90-153 Łódź, Poland

Corresponding author: Jan Andrzej Napieralski (jnapier@dmcs.pl)

This work has been supported by the Project STRATEGMED Innovative System for Evaluation and Rehabilitation of Human Imbalance from the National Centre for Research and Development under Grant 2/266299/19/NCBR/2016.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Nofer Institute under Protocol No. 17/2014, and performed in line with the Declaration of Helsinki.

ABSTRACT The article presents the concept of detecting subjects with balance disorders by the use of machine learning techniques. The proposed solution has been developed and tested based on a group of 40 subjects, the group included both patients with uncompensated dysfunction in the vestibular system and healthy volunteers. Presence of dysfunction was verified prior to the study by detailed clinical examination. The data for the study were collected with the use of miniature micromachine sensors, mounted on the body at selected locations. The task performed by the subjects consisted of free gait over a distance of three meters; the task was selected to make it easy to perform in any surroundings and not requiring additional equipment. The collected data was used as an input to an artificial neural network based on a one-dimensional convolution kernel. The proposed solution allows to classify subjects into healthy and non-healthy with an accuracy of 87.5%.

INDEX TERMS Machine learning, CNN, fall risk, balance disorder.

I. INTRODUCTION Balance disorders present a serious problem in the ageing human population. They cause a significant deterioration in the quality of life and, according to epidemiological studies, increase risk of an injurious fall [1]. This is especially evident in the elderly: 30% of those over 65 have experienced episodes of dangerous imbalance [1].

Multiple methods have been devised that can help in the detection of the imbalance disorder and thus counteract imbalance-related conditions, however, none of these are perfect. The approaches currently employed by physicians and physiotherapists can be roughly divided into three categories. The first group consists of methods utilising some sort of balance-measuring equipment. Here the strain-gauge based balance-plate is commonly used [2], [3], which can measure sway of a subject during quiet stance. This gives objective results, but the device is large, expensive and, due to its size and weight, can only be used in hospitals or doctor’s offices. Moreover, it does not allow to correctly detect all cases of imbalance disorders, the overall accuracy being about 85% [4].

The second group includes functional tests, in which the subject is asked to perform one or more tasks resembling everyday activities (walking, standing-up, turning around, etc.). According to the World Health Organization and Clinical Practice Guideline from the Academy of Neurologic Physical Therapy of the American Physical Therapy Association, it is recommended to use the International Classification of Functioning, Disability and Health [5].
It includes functional tests based on which the impact of dizziness and / or balance problems on the ability to perform activities and participation tasks is assessed. The tasks are assessed by the physician or physiotherapists using simple tools, such as a stopwatch or a tape-measure. These tests can be cheaply performed in any setting, but are subject to errors, for example, due to subjects not adhering to the test procedure or performing the procedure differently on the first and subsequent runs (effect of “learning” about how to perform it correctly). Clinical evaluation of functional tests involves observing the patient while performing the test and depends on the experience of the examiner. In the case of a test assessing gait, it is based on subjective feelings about the smoothness of movement of the lower limbs (right and left foot support time) and trunk (swaying, throwing), step length and height (asymmetry between right and left steps), gait width and consistency (deviation from the gait path without stopping), and time used to complete the task. If there is a need for even more detailed analysis, the gait strides should be decomposed into separate phases (stance phase, swing phase) and the occurring asymmetries in the work of the muscles of the lower limbs and trunk in individual elements of the phases such as initial contact, loading response, midstance, terminal stance, pre-swing, initial swing, midswing, terminal swing should be evaluated. The ubiquitous TimedUp-and-Go (TUG) test [6] an also be used to assess a patient’s functional status, as can the multi-task Berg Balance Score (BBS) [7]. Reported accuracy for detecting imbalances ranges from 76.8% [6] to 89.5% [7]. The BBS contains 14 simple tasks to perform, with a completion time of 20 minutes. It includes, among others, shifting, sitting, standing with eyes open, picking up an object from the floor, and a reaching test. The patient is scored on a 4-0 scale depending on the correctness of the task, where 0 indicates the inability to perform the task and 4 indicates the independent performance of the task. The maximum number of points is 56. Interpretation of the score allows determining the average risk of falls at 40-21 points, and high risk for values below 20 points. Tests based on multiple tasks can also be tiring for elderly people.

The methods of the third group are based on detailed examination of the patient in order to determine the presence or absence of particular diseases that may result in balance dysfunctions. Although these methods can unambiguously confirm or reject the presence of a disease, this not necessarily translates directly into balance dysfunctions. This is particularly evident in the case of rehabilitation, which in itself does not cure the disease, but may result in improved balance properties of the subject. Consequently, methods of this group are usually used in conjunction with one of the previous groups.

Given the lack of a “gold standard” in imbalance assessment [8], there is an ongoing research in this field. Modern hardware technologies allow to capture details of human movement, while software techniques make it possible to process and analyse the gathered data. Of the hardware approaches, micromachined Inertial Measurement Units (IMUs) are of particular interest, due to their diminutive size and low cost. However, their metrological properties are far from perfect, which means the processing stage is of particular importance. The research presented in this paper considers such a solution.

The use of artificial convolutional networks allowed to obtain satisfactory results on a relatively small set of input data. Conventional methods could prove insufficient to match the accuracy presented as a result of the conducted research. AI combines many input data (features) while conventional methods are based on programming a single feature. Prior research by authors showed that thanks to this AI in such applications gives better results.

The remainder of this paper is organised as follows, Section II presents a summary of the current state of the art in the field of the use of machine learning methods in the detection of imbalance disorders. Section III treats about the purpose and motivation for the research process described in this article. In Section IV the detailed description of the study group is presented. Section V evaluates the methods used in order to extract appropriate features which later are used as the input for the neural network together with the methods used in classification process - algorithms based on machine learning principles. In Section VI the obtained results are listed. The last Section VIII describes the final conclusion of the given research process.

II. STATE OF THE ART

Micromachined IMUs enjoy widespread use due to their inclusion in smartphones and automotive applications. Despite the fact, that it is possible to perform some analysis of the human motion, including gait classification, using only smartphone-based IMU [9], medical-grade solutions usually employ dedicated hardware, which enable multi-sensor approach, secure affixing to body segments, customised hardware preprocessing, etc.

The raw data coming from the IMU are of no use by itself. Many pre-processing and processing algorithms have been proposed to extract useful information and perform the ultimate task of pronouncing an opinion about the state of the subject’s balance system. Of these, artificial neural networks represent an important class of algorithms. Current research concerning the development of artificial neural networks for human motion recognition is essentially focused on the application of convolutional neural networks (CNN’s) [10], due to its superiority among other machine learning algorithms [11] in this field.

CNN’s are commonly associated with image processing. Indeed, their invention was inspired by the organisation of the animal visual cortex. Consequently, a significant portion of research concerning CNNs in human motion analysis is based on image processing [12], [13], [14]. Video frames acquired from a camera or a set of cameras are firstly decomposed by computer vision algorithms and essential body segments are extracted. Such segments with predefined labels are then used as the input of CNN. Another approach is to use 3D
skeleton data as the input of CNN [15]. However, using IMU data instead of visual data is also being investigated, due to its immunity to factors such as lighting conditions or clothes worn by the patients, its capability to record quantities unavailable in the video data (e.g., shocks resulting from foot impact on the ground) and usually lower cost of the equipment. IMUs are also easier to deploy in patient’s home, for example, in Ambient Assisted Living systems [16].

Use of IMUs to detect balance disorders based on gait or similar tasks is a current topic of research. For example in [17] the authors used a device similar to the one presented in this article in order to detect balance disorders based on the data gathered during the Dynamic Gait Index (DGI) test suite. After processing data with the support-vector machine algorithm (SVM) they achieved very high accuracy of 96.1%. However, it must be noted that their method requires performing a full DGI test suite, consisting of six separate tasks (e.g., walking at different speeds, walking with head movements, walking over an obstacle). The method presented in this article utilises only a five-second fragment of free gait.

Gait analysis may be also performed using different type of sensors. In [18] the authors present a method based on data obtained from a walkway equipped with pressure sensors. It allows to distinguish four different neurological disorders with sensitivity above 90% when analysed through artificial neural network or SVM algorithms. However, the measurement device used in the cited research was a 6.7 meter walking carpet, making the hardware part of the setup expensive and not suitable for many diagnosis scenarios (e.g., at patient’s home)

Recent research indicates that the use of a single sensor on the L4 vertebra enables the detection of balance problems with an accuracy similar to the clinical Tinetti test [19]. The classifier - SVM and k-NN used in this process shows compliance with professional medical examination on imbalance with 92% accuracy. The authors’ approach differs in the method of assessing imbalance in patients (TUG test) as well as the lack of young people in the data set - only the community-dwelling elderly participated in this study.

The detection of balance disorders, and in particular an attempt to predict whether a patient will fall or not, is proposed in the latest research with the use of cameras [20] or micro-Doppler radar with the use of SVM and CNN [21]. The latest research is also carried out using motion sensors such as Kinect, the image of which is used as the basis for classification process using a neural network technique [22]. Contrary to the approach proposed by the authors, these solutions do not use micromachine sensors.

With regard to research concerning the use of CNNs to analyse IMU data, Roshdibenam et al. proposed a solution for detecting risk of falls in elders using CNN model with the raw kinematics signal. The signal was obtained during the TUG test from proprietary IMU sensors, developed in MEMS technology, placed on the neck and feet [23]. The CNN model was using three seconds time series segments as the input, multiple deep layers and binary output distinguishing faller from non-faller. The constructed predictive model was compared to clinical assessments. The best results were obtained using neck angular velocity with a high sensitivity of 86% (in comparison to 56% sensitivity obtained using the traditional evaluation of the TUG test). Compared to our approach, the authors used different number of sensors and their configuration, as well as different structure and parameters of the CNN. Moreover, they used a database containing only elderly people and the aim was to strictly detect fall risk, not balance disorders in general, as in this study.

CNNs are also used to process other input signals usable for balance assessment. Savadkoohi et al. used proprietary One-One-One Deep Neural Network which was fed with data from open-source database containing force-plate measurements of 163 subjects [24]. Their study proved that the risk of falling can be derived based on balance metrics gathered from force-plate time series with the combination of the deep learning algorithms. Their results indicated that One-One-One Neural Network has the best ability to classify risk of falling with 99.9% accuracy. What is more, they achieve similar results using CNN (99.3% accuracy), RNN (96.9%) and LSTM (98.3%) which are by far the best results obtained on this dataset [25]. Contrary to the solution proposed in this article, their approach focused on static posturography with Falls Efficacy Scale (FES) [26] as the assessment method.

Compared to the solutions presented in this section, the approach presented by the authors of this article differs in the architecture of the neural network used, its parameters, input data and arrangement, and the number of sensors used.

III. AIM OF THE STUDY
The major objective of this research was to develop a robust method for detection of subjects with balance and gait abnormalities caused by vestibular dysfunction. The authors based research on data obtained from Medipost devices (see Section V). Together with the method proposed in this article they may be viewed as a complete system for automatic imbalance assessment. The system design goals included:

- high accuracy,
- objectivity,
- ability to be used out of the doctor’s office,
- low examination burden for the subject,
- low cost.

The first item on the above list cannot be attained without suitable processing algorithm, such as the one described in this article. The last two items relate to the number of Medipost devices that have to be placed on the subject’s body: smaller number of devices means simpler preparation for the examination and lower cost of the system. At the same time reduction of the number of devices should not impair the accuracy. Consequently, determining how the number of devices relates to the classification accuracy was also the aim of this research.
IV. STUDY GROUP
The study group consisted of 40 subjects divided into two groups:

- patients – group included 16 subjects with a vestibular impairment confirmed by the caloric test in video-nystagmography (VNG) assessment (canal paresis >20% according to the laboratory norms). They reported acute vertigo onset at least 3 weeks before the examination. The mean vertigo and imbalance was moderate in the study group, scored as 4 points in the 10 point Visual Analogue Scale (VAS). The mean value of the Dynamic Gait Index scoring was 20 in that group, however only 40% scored less than 21 points, which may suggest a higher risk of falling. All subjects in the study group presented subtle gait abnormalities in the gait-part of Tinetti Scale. The mean value of the Tinetti-gait scoring was 10 (out of 14 maximum). To summarize, the study group presented mild balance and gait abnormalities in clinical examination, caused by a not fully compensated vestibular dysfunction.

- healthy – usually young people, fully functional and mobile. Absence of balance dysfunction was determined based on the following data: - No history of dizziness, neurological diseases, circulatory system and musculoskeletal system diseases, diabetes or migraines. Any abnormalities results of physical examination, VNG test and dynamic posturography excluded the volunteer from the healthy group.

V. METHODS
A. DATA ACQUISITION SYSTEM
The sensor used in this study was a commercial off-the-shelf micromachined unit, consisting of three-axis accelerometer, gyroscope and magnetometer. This sensor was embedded in a custom microprocessor-based system, called Medipost. It is battery operated and small enough to be easily placed in various positions on the human body. The sensor readings were obtained at 200Hz frequency, then downsampled using a low-pass filter to 20 samples per second and wirelessly transmitted to a PC computer, which stored the data in a database. As the 1D-CNN algorithm is based on data coming from a number of sensors, it is not feasible to implement it in the Medipost device. Instead, the computations are performed by a PC computer, after completion of the exercise (i.e., free gait). Figure 1 presents the Medipost device which consists of the following components [27]:

- IMU sensor: LSM9DS1;
- Wi-Fi radio unit based on the ESP32 system;
- Power supply block based on a TPS63051 voltage converter;
- \( \pm 2/\pm 4/\pm 8/\pm 16 \) g linear acceleration full scale;
- \( \pm 245/\pm 500/\pm 2000 \) dps angular rate full scale;

For this study, six identical Medipost devices were used, placed in the following locations of the subject’s body:
- lumbar vertebra (L5),
- cervico-thoracic transition (C7-TH1),
- left shin,
- right shin,
- left thigh,
- right thigh.

Although the devices are attached to the body, using elastic bands, in a way that prevents any wobble, it is not possible to ensure that the device’s coordinate system matches the coordinate system of the body segment it rests on. This is mainly due to the curvature of the body in the place where the device is mounted. Consequently, it was necessary to use an appropriate calibration procedure that unambiguously transforms one coordinate system into the other. This procedure, performed by each subject prior to the actual examination, consisted of standing upright, then making a bow, and finally a shallow squat. It allowed to determine the “up” direction (during standing upright) and “forward” direction (by making all body segments rotate in the anterior-posterior plane during the bow or the squat). The result is a transformation matrix for each device.

B. EXAMINATION TASK
The task performed by each subject was free gait. This task was selected as it is very easy to perform, requires no additional equipment, and at the same time carries a large amount of information about the subject’s balance system [4]. The exact procedure was as follows:

- the task begins with the subject walking along a straight line 5 meters long,
- after passing the designated distance, the subject makes a U-turn,
- the subject returns along the same path to the place where the task began.

C. BIOMECHANICAL BODY MODEL
Based on the data obtained from the task, subjects’ movements were re-created in the software. To this end,
an appropriate body model has been developed [4]. The model represents the human body through a set of rigid segments, some of the segments are connected end-to-end to represent joints. Data from the IMU sensor, together with the transformation matrix obtained through the calibration procedure, allows to determine rotation in 3D space for each segment. The rotation is computed by one of many available IMU data fusion algorithm [28]. Moreover, given the assumption that one foot is on the floor at all times and the segments are connected at joints, it is possible to calculate translation of each segment. Moreover, the position of the center of mass (COM) of the whole body can be calculated as a weighted vector sum $\vec{x}_{w}$ of the COM’s of the segments, assuming that set of segments masses $w_1, w_2, \ldots, w_n \geq 0$ (1):

$$\vec{x}_{w} = \frac{\sum_{i=1}^{n} w_i \vec{x}_i}{\sum_{i=1}^{n} w_i}$$ (1)

Figure 2 presents visualisation of the body model. The dark grey ball represents the COM of the whole body, while the colored balls represent COMs of individual body segments. Incidentally, this figure is taken verbatim from the analysis and visualisation software, which is also part of the Medipost system [4].

**D. FEATURES EXTRACTION**

The body model allows to extract many features that describe the subject’s motion. Table 2 presents features extracted for each segment of the model. Additionally, the following features related to the body as a whole were extracted:

- COM position of the whole body,
- lengths of consecutive steps.

The length of the step is calculated as a maximum (for the duration of the step) Euclidean distance between the bottoms of the shin segments (2):

$$d(p, q)^2 = (q_1 - p_1)^2 + (q_2 - p_2)^2$$ (2)

Where $(q_1, q_2)$ are the coordinates of the bottom segment trailer, and $(p_1, p_2)$ are the coordinates of the top segment trailer.

**E. PREPROCESSING**

Before the learning process begins, the appropriate data preprocessing must be performed. The first stage of making the data suitable for the neural network is the normalization process – standard scores are computed for each parameter. After that, the data are trimmed so that each record contains actual gait (and not the initial standing phase or the U-turn performed in the middle of the task) and is of the same length (Figure 3 and Figure 4). The data trimming process is performed as follows:

- gait is assumed to be started, when the gyroscope reading from any of sensors situated on the subject’s shins is above predetermined threshold,
- gait is assumed to end, when five seconds from the beginning of the gait are gathered.

In the next step, the inputs are selected – a subset of the features from Table 2 and possibly also the whole-body features. This step is very important, as comparing classification accuracy for different combinations of features allows to determine which features are useful and which are not. In particular, it may lead to a conclusion that some of the sensors are not needed – and this will result in a system

---

**TABLE 2. Set of features extracted for each body model segment.**

| No | Feature                                | Type          |
|----|----------------------------------------|---------------|
| 1  | Total acceleration                     | 3-element vector |
| 2  | Interial acceleration                  | 3-element vector |
| 3  | COM of segment                         | 3-element vector |
| 4  | Top of segment                         | 3-element vector |
| 5  | Bottom of segment                      | 3-element vector |
| 6  | Angle between segment’s vertical axis and the zenith direction | scalar |
| 7  | Rotation in the Anterior- Posterior plane of the body | scalar |
| 8  | Rotation in the Left-Right plane of the body | scalar |
| 9  | Rotation in the horizontal plane of the body | scalar |
| 10 | Angle of rotation round the rotation axis | scalar |
| 11 | Vector from the bottom of the segment to the top of the segment | 3-element vector |
| 12 | Linear motion acceleration             | scalar |
| 13 | Linear motion speed                    | scalar |
| 14 | Angular speed                          | scalar |
that is simpler to use and cheaper to construct. The following combination of features were tested:

- features with similar meaning or source, for example, features describing only rotations or features based on accelerometer data for each body segment,
- same features buy for entire body model (all body segments)

If the value of the feature is a vector, all of its elements are selected. As an example, one of the features combination which was tested was the combination based on acceleration values. In such case features number $1 + 2$ from table 2 are combined together creating the input for the neural network. Finally, each recording is labelled as belonging to healthy or unhealthy subject, based on the medical examination performed by a qualified group of physicians. A detailed flow diagram is shown in Figure 5.

### F. ARTIFICIAL NEURAL NETWORK

In classification process the algorithms based on one dimensional convolutional neural networks (1D-CNN) are used. The designed sequential 1D-CNN model used for classification process is presented in Figure 6. One-dimensional convolutional networks use as an input the time series data and apply convolutional filters (kernels) by moving them along the axis of time. Number of epochs were determined using Keras EarlyStopping mechanism, configured in such way, that the CNN model stop learning after 20 epochs of no improvement of the cost function of the cross-validation data on the training set.

In order to obtain the best classification approach, different combinations of features were tested, as described in Section V-D. Moreover, hyperparameter of the network were changed in a process called hyperparameter tuning. The following hyperparameters and their values were taken into consideration:

- number of filters in Conv1D layers: 100, 300, 500,
- kernel size in Conv1D layers: 3, 8,
number of nodes in dense layer: 3, 8,
dropout layer rate: 0.2, 0.5,
activation function in the last layer: sigmoid, softmax
- treated as an additional hyperparameter. In the case of the sigmoid function, the output layer consisted of a single neuron: a threshold equal to half of the neurons output range discriminated between healthy and sick class. In the case of the softmax function, the output layer consisted of two neurons, corresponding to the healthy and sick class: the neuron with the higher output value determined the classification result.

In total, more than 3200 hyperparameter combinations were tested and evaluated.

As the hyperparameter tuning task is time consuming, it was not feasible to perform proper cross-validation when searching the combined space of the network hyperparameters and input features. For the same reason, analysis of the influence of the number of used sensors also could not be integrated into this process. Consequently, the whole task had to be divided into two separate stages:

1) hyperparameter tuning combined with a selection of the best input features. During this stage the input data set was divided into fixed training and test sets, containing 25 and 15 samples, respectively. The stage enabled identification of the best input features;

2) hyperparameter tuning combined with analysis of the influence of the number of used sensors. The input features identified in the previous step were used. During this stage 10-fold cross-validation was used. It was implemented using the appropriate functionality of the scikit-learn library. The stage enabled identification of best hyperparameters for each tested combination of sensors and also provided the final performance metrics.

In stage 2, the following procedure was used:

1) Set up a neural network with best input features as determined in stage 1,
2) Take head sensor as the single input,
3) Perform hyperparameter tuning,
4) Add next sensor from the list: trunk, right thigh, left thigh, right shin, left shin,
5) Go back to point 3.

The set of hyperparameters used for both stages was selected empirically using the brut-force method - each hyperparameter was tested with each remaining hyperparameter to cover all combinations.

VI. RESULTS

A. INPUT FEATURES AND NETWORK HYPERPARAMETERS

Results of stage 1 clearly indicate that the best classifier can be obtained by using both total and inertial acceleration as input features (entries 1 and 2 in Table 2). This combination was used by nine top-performing networks. Acceptable results were also obtained (in decreasing order of accuracy) using feature 6, feature 5 and a combination of features 6 and 10. All further work used a combination of features 1 and 2.

B. NUMBER OF SENSORS AND NETWORK HYPERPARAMETERS

As already stated in Section V-F, this stage searched the combined state space of hyperparameters and sensor combinations. The top performing instances are presented in Table 3. As can be seen, there are four entries with accuracy above 80%, with the best one reaching 87.5%. It is interesting to note that the best result was obtained using only two sensors, i.e. these located on the nape and trunk. The second-best instance (somewhat worse in terms of accuracy, but even lower in terms of the average loss) was also using
these sensors. The remaining top-four performers used four and six sensors, respectively.

Table 4 presents the best performers for each tested combination of sensors. The best three entries have been already present in Table 3. It is interesting to note that while using sensors on both thighs or both shins gives good results, limiting the sensors to only one leg somewhat lowers the performance. The worst results are obtained for a single, head-mounted, sensor.

VII. DISCUSSION
The presented study allowed to make interesting observations regarding a number of aspects of the classification solution.

A. DATA GATHERED FROM THE SENSORS AND FEATURES EXTRACTION
The sensors used in this study consist of three-axis accelerometer, gyroscope and magnetometer. The raw data may be given to the model directly, or they may form a basis for computation of some other features that are consumed by the model. In a traditional approach to machine learning, the second approach was used, as the models were not capable of extracting the features by themselves. This posed a serious limitation, as the performance of the solution could be heavily degraded by improper choice of the features, and this choice was performed by humans. The literature shows that the quest for determining the correct set of features in imbalance disorder detection has been long and tedious, but hardly conclusive, e.g., [29] provides analysis of no less than 130 features used in various studies. On the other hand, CNNs usually work with the raw data [30], [31], and the obtained results confirm their capabilities. The best results were obtained when the input vector contained raw acceleration data; excluding these data always resulted in worse performance, no matter which of the human-devised features were used. However, for the best results raw accelerometer data had to be supplemented by another acceleration vector, namely the inertial acceleration vector. It is comprised of the accelerations resulting from the movement of the body, as opposed to the former, which contains the sum of the accelerations resulting from the movement and from the gravity. The inertial acceleration vector is obtained by removing from the raw data the accelerations resulting from gravity, and in order to do so, the orientation of the body segment in 3D space must be known. This in turn is computed through fusing the IMU data by an appropriate algorithm (see Section V-C).

Ultimately, the study shows that a CNN for balance disorder detection works best with raw data. Nevertheless, due to the physical operation principle of the sensors, some of this “raw” data is not directly available and must be computed in the preprocessing stage [32].

B. PLACEMENT OF THE SENSORS
The obtained results indicate that increase in number of sensors beyond two does not improve classification accuracy. This may be due to the phenomenon known as overfitting, where the model learns some patterns that are not related to the classification task. In such situation the model performs perfectly on the data used in the learning process, but will fail on data never seen before, as the choice of the features and their values that is used for the classification is flawed. Overfitting may appear when the model is able to follow the learning data to closely, which may be caused by too complex model or too abundant (but not providing useful information) input features. Clearly, adding further sensors in the analysed case fulfills this second scenario. Similar observations are also present in literature, for example, in [Howcroft Plos One] authors found that increasing number of sensors did not lead to better model accuracy. Contrary to the current research, however, the top performing model was that based solely on the head sensor.

C. NETWORK STRUCTURE AND HYPERPARAMETERS
Based on the presented research it is not possible to indicate the overall best combination of network hyperparameters. However, some (inconclusive) relation may be observed between the number of sensors and the hyperparameters: in general configurations with a small number of sensors favour more elaborate networks, with more filters and more layers (cf. entries 1 and 2 in Table 3), while use all sensor configuration benefits of simpler networks (cf. entry 4 in Table 3). This is in accordance with the observation in Section VII-B.
D. OVERALL PERFORMANCE OF THE SOLUTION
The obtained accuracy places the solution in the midst of the of the contemporary research. Clearly, there are reports of better performance, however, the comparison should take into account the complete setup of the experiment. In the present research, the focus was placed on obtaining a solution that is cheap and simple to use, including examination outside doctor’s office. This aim has been achieved, as the employed IMU-base devices are cheap and compact, while the selected task (five-second free gait) is easy to perform and does not require any additional equipment. Better results are usually reported for experiments involving more intricate (and expensive) measurement devices [25] or more time-consuming or complicated tasks – e.g. [33], [34]. When compared with similar setups, the obtained accuracy is higher than reported in literature [35], [36], [37], [38].

VIII. CONCLUSION
The authors of the publication proved that it is possible to classify subjects into unhealthy (with balance and gait dysfunction caused by the uncompensated unilateral vestibular impairment) and healthy persons with the use of machine learning techniques and a set of simple IMU sensors. To achieve this goal, the algorithms based on Convolutional Neural Networks (1D-CNN) were used. The result of subjects classification gives the accuracy of 87.5%. These results are better than some of the results obtained with similar research [23], while using the smaller number of sensors. It should also be noted that the study group revealed only very subtle balance and gait abnormalities in the clinical examination.

The analysis of the influence of the number of sensors on the resulting accuracy, presented in this article, also confirms that high classification accuracy doesn’t require measurements from a large number of body segments. The research has indicated that the best results may be obtained from models based on inputs from sensors located on the head and trunk.

Moreover, there is a strong similarity in the results obtained by sensors containing two shins and both thighs. The results are clearly worse when using only a single sensor from the left-right thigh or left-right shin pair or adding only one of the elements from the sensors on the shin with the presence of both sensors on the thighs. This result may indicate learning problems in a neural network with incomplete information about a pair of sensors located in symmetrical sections of the body.

It is noteworthy that the results obtained in this article were based on one of the simplest tasks possible: the free gait. This means the classification procedure is easy to perform in any surroundings without additional equipment (apart from the sensors). This is of particular importance in the case of procedures performed outside hospital or doctor’s office, e.g., in the subject’s home.

REFRENCES
[1] L. Z. Rubenstein, “Falls in older people: Epidemiology, risk factors and strategies for prevention,” Age Ageing, vol. 35, no. 2, pp. i37–i41, 2006.
[2] B. R. Bloem, J. E. Visser, and J. H. Allum, “Posturography,” in Handbook of Clinical Neurophysiology, vol. 1. Amsterdam, The Netherlands: Elsevier, 2003, pp. 295–336.
[3] T. Łukaszewicz, D. Kania, Z. Kidof, and K. Pethe-Kania, “Posturographic methods for body posture symmetry assessment,” Bull. Polish Acad. Sci. Techn. Sci., vol. 63, no. 4, pp. 907–917, 2015.
[4] W. Tyłman, R. Kotas, M. Kamiński, P. Marciniak, S. Woźniak, J. Napieralski, B. Sakowicz, M. Janc, M. Józefowicz-Korzyńska, and E. Zamyłowska-Szmytke, “Fully automatic fall risk assessment based on a fast mobility test,” Sensors, vol. 21, no. 4, p. 1338, Feb. 2021.
[5] C. D. Hall, “Vestibular rehabilitation for peripheral vestibular hypo-function: An updated clinical practice guideline from the academy of neurologic physical therapy of the American physical therapy association,” J. Vestibul. Phys. Therapy, vol. 46, no. 2, p. 118, 2022.
[6] B. R. Greene, A. O. Donovan, R. Romero-Ortuno, L. Cogan, C. N. Scanaill, and R. A. Kenny, “Quantitative falls risk assessment using the timed up and go test,” IEEE Trans. Biomed. Eng., vol. 57, no. 12, pp. 2918–2926, Dec. 2010.
[7] H. Similä, J. Mäntyjärvi, J. Merilahni, M. Lindholm, and M. Ermes, “Accelerometer-based Berg balance scale score estimation,” IEEE J. Biomed. Health Informat., vol. 18, no. 4, pp. 1114–1121, Jul. 2014.
[8] H. Kingma, G. C. Gauchard, C. de Waele, C. van Nechel, A. Bisdorff, A. Yelnik, M. Magnusson, and P. P. Perrin, “Stocktaking on the development of posturography for clinical use,” J. Vestibular Res., vol. 21, no. 3, pp. 117–125, Jun. 2011.
[9] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, and Q. Li, “Deep learning-based gait recognition using smartphones in the wild,” IEEE Trans. Inf. Forensics Security, vol. 15, pp. 3197–3212, 2020.
[10] M. A. Loataibi and A. Mahmod, “Improved gait recognition based on specialized deep convolutional neural network,” Comput. Vis. Image Understand., vol. 164, pp. 103–110, Nov. 2017.
[11] M. Sayed, “Biometric gait recognition based on machine learning algorithms,” J. Comput. Sci., vol. 14, no. 7, pp. 1064–1073, Jul. 2018.
[12] S. Shuai, M. Kavitha, J. Miyao, and T. Kurita, “Action classification based on 2D coordinates obtained by real-time pose estimation,” in Proc. Int. Workshop Frontiers Comput. Vis., 2019, pp. 1–6.
[13] F. M. Castro, M. J. Marin-Jimenez, N. Gui, S. Lopez-Tapia, and N. P. de la Blanca, “Evaluation of CNN architectures for gait recognition based on optical flow maps,” in Proc. Int. Conf. Biometrics Special Interest Group (BISIG), Sep. 2017, pp. 1–5.
[14] A. Sokolova and A. Konshin, “Gait recognition based on convolutional neural networks,” Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci., vol. 42, pp. 1–6, May 2017.

| Number of sensors | Rank | Average accuracy | Average loss | Output activation function | Number of filters | Kernel size | Number of dense layer neurons | Number of internal layers | Dropout rate |
|-------------------|------|------------------|--------------|---------------------------|------------------|------------|-------------------------------|--------------------------|-------------|
| 1                 | 114  | 70               | 26,926       | Softmax                   | 500              | 8          | 2                             | 2                        | 0.5         |
| 2                 | 1    | 87.5             | 16,770       | Sigmoid                   | 500              | 8          | 3                             | 3                        | 0.2         |
| 3                 | 7    | 80               | 18,708       | Softmax                   | 100              | 3          | 3                             | 3                        | 0.2         |
| 4                 | 3    | 82.5             | 19,526       | Sigmoid                   | 500              | 8          | 3                             | 3                        | 0.2         |
| 5                 | 11   | 80               | 18,708       | Softmax                   | 500              | 8          | 3                             | 3                        | 0.2         |
| 6                 | 4    | 82.5             | 19,526       | Softmax                   | 100              | 3          | 8                             | 1                        | 0.2         |
