Real-Time Gait Anomaly Detection Using 1D-CNN and LSTM

Jakob Rostovski[1(✉)], Mohammad Hasan Ahmadilivani[2], Andrei Krivošei[1], Alar Kuusik[1], and Muhammad Mahtab Alam[1]

[1] TJS Department of Electronics, Tallinn University of Technology, Tallinn, Estonia
{jakob.rostovski,andrei.krivosei,alar.kuusik,muhammad.alam}@taltech.ee

[2] Computer Systems Department, Tallinn University of Technology, Tallinn, Estonia
mohammad.ahmadilivani@taltech.ee

Abstract. Anomaly detection and fall prevention represent one of the key research areas within gait analysis for patients suffering from neurological disorders. Deep Learning has penetrated into healthcare applications, encompassing disease diagnosis and anomaly prediction. Connected wearable medical sensors are emerging due to computationally expensive machine learning tasks, which traditionally require use of remote PC or cloud computing. However, to reduce needs for wireless communication channel throughput, for data processing latency, and increase service reliability and safety, on device machine learning is gaining attention. This paper presents an innovative approach that leverages one dimensional convolutional neural network (1D-CNN) and long-short term memory (LSTM) neural network for the real-time detection of abnormal gait patterns during the step. Real-time anomaly detection pertains to the algorithm's ability to promptly detect true gait abnormality occurrence during the swing phase of an ongoing step.

For the experiments, we have collected eight different common gait anomalies, simulated by 22 persons, using motion sensors containing multidimensional inertial measurement units (IMUs).

Results have demonstrated that the proposed 1D-CNN-AD algorithm achieves an average accuracy of 95% and an average F1-score of 88% for all gait types and can run in true real-time. Average earliness for 1D-CNN-AD algorithm was 0.6 s, which is mid-swing phase of the step. Proposed LSTM-AD algorithm achieved average accuracy of 87% and average F1-score of 70% for all gait types.

Keywords: Human gait · Anomaly detection · Gait analysis · Machine learning · Real-time · 1D-CNN · LSTM · Wearable sensors

This work has been supported by Estonian Research Council, via research grant No PRG424, by the Center of Excellence (TK) project TAR16013 (EXCITE) and Estonian IT Academy project "Sustainable Artificial Internet of Things (SAIoT)".

M. Särestöniemi et al. (Eds.): NCDHWS 2024, CCIS 2084, pp. 260–278, 2024.
https://doi.org/10.1007/978-3-031-59091-7_17

1 Introduction

According to the World Health Organisation (WHO) report about one billion persons are affected by neurological disorders worldwide [3]. Neurological diseases ranging from migraine to stroke, and Alzheimer are the leading causes of Disability Adjusted Life Years (DALY) loss [7]. For instance, there is a substantial risk of falling for patients with gait impairments from neurological diseases [23]. It is especially true for patients suffering from neuromuscular diseases, because high variability and deviations from the optimal gait pattern can be seen in their gait [13]. Therefore, it is challenging to analyze patients' gait patterns in real-time. The gait of a person can be described by a set of parameters such as: step length, duration of individual step phases, muscle force, etc. [19]. Wearable motion sensors, containing multidimensional Inertial Measurement Units (IMUs), are the most widely used gait assessment devices in recent years for supporting daily activities [25]. For example, motion sensors are used to detect initial and final contact events of the gait cycle for different persons - healthy, with stroke, and with other neurological disorders, and select the best algorithms and sensor placements for correct classification between them [10]. Motion sensors can be employed to detect activities of daily life, fall events and their directions [9], to determine rehabilitation progress and analyze gait normalcy index [2,36]. Also such devices can be used to discover environment dependent differences in gait, which will help with context-aware decisions [29]. Finally, in combination with Neural Networks (NNs), identify if person has balance disorder [20], to track rehabilitation progress for broken limbs [4] etc.

It is shown that Functional Electrical Stimulation (FES) can be used to assist walking and help with fall prevention [12] as well as for generic gait improvements [17]. Long-term gait deviation analysis and efficient run-time control of FES devices require automated real-time recognition of gait deviations. Average swing phase of a step is 300–400 ms long [8], and the time of full contraction of the muscle using electrical stimulation is 100–200 ms long [5], thus the detection time of step pattern deviations should be under 100 ms. Considering that the incoming signal must be processed, a correct decision made, and stimulation actuation started, a detection time of 50 ms is required since the gait abnormality has started.

Connected wearable medical sensors are emerging due to computationally expensive machine learning tasks, which traditionally require use of remote PC or cloud computing [14]. Nowadays, it is common to offload such data analysis from wearable sensors to wirelessly connected smartphones [11]. For example, data processing unit, sensors and muscle stimulator shall be wireless for gait correction system, i.e. based on Bluetooth or SmartBAN standard. However, to reduce needs for wireless communication channel throughput, for data processing latency, and increase service reliability and safety, on device machine learning is gaining attention [31]. Existing real-time algorithms are used in gait analysis for identification by gait [15]; detecting of gait events like heel-strike and toe-off for elderly healthy subjects; stroke patients and patients with Parkinson disease [35], as well as with other impairments [24,37]; haptic biofeedback devices are implemented using inertial measurement units (IMUs), to correct toe-in or toe-out during walking in real-time [32].

Notably, there are not found state-of-the-art solutions in gait analysis for real-time anomaly detection of realistic gait deviations during the ongoing step, caused by neurological diseases.

In our prior research work [27,28] we proposed a base method for real-time anomaly detection in gait during the ongoing step, with an algorithm based on Support Vector Machines (SVM), which is one of the most popular algorithms used in gait analysis. On the other hand, NNs are widely adopted in gait analysis [30]. They are capable of solving complex tasks in time-series data. Nonetheless, to the best of our knowledge, there is no research exploiting NNs for real-time anomaly detection during the ongoing step in gait analysis. In this paper, for the first time, we leverage Convolutional Neural Network (CNN) and Long Short-Term Memory NNs for real-time anomaly detection during the ongoing step in human gait.

The contributions of this work are:

- Estimation of the performance of One Dimensional-Convolutonal Neural Network-Anomaly Detection algorithm (1D-CNN-AD) and Long Short-Term Memory Neural Network-Anomaly Detection algorithm (LSTM-AD) on the collected simulated gait deviation dataset in comparison to the Real-time tsSVM Anomaly Detection algorithm (RTtsSVM-AD).
- Exploiting hyperparameters for the neural networks to optimize performance on simulated gait dataset for real-time in-step anomaly detection.

This paper consists of six sections: after the introduction, in Sect. 2 data acquisition and gait types are described, as well as metrics used for analysis in addition to presenting the proposed 1D-CNN-AD and LSTM-AD algorithms, then in Sect. 3 we briefly describe evaluation metrics and the SVM-based algorithm – RTtsSVM-AD, which is continued with experimental setup in Sect. 4; this is followed by the results and discussion in the Sect. 5 and the paper is concluded in Sect. 6.

2 Methodology

2.1 Dataset

Data Acquisition. The dataset in our experiments is collected from twenty-two healthy persons of different genders, ages, heights and weights (Table 1), while walking in a straight line and simulating abnormalities. Simulations are recreating actual patients' video recordings of gait deviations in collaboration and guidance from a professional physiotherapist of Tallinn East Central Hospital. We have included the most frequent human gait abnormalities, regarding reference [1]: Ataxic, Diplegic, Hemiplegic, Hyperkinetic, Parkinonian, Slap, Steppage, and Trendelenburg (lurch). Table 2 shows eight under-study gait types and the number of collected gait recordings per gait type. Collected data is labeled

Table 1. Persons' Information Used in This Study (Mean ± *Standard Deviation)*

No. of subjects	Age (years)	Height (cm)	Mass (kg)
15 (Male)	32.1 ± 11.1	177.7 ± 5.5	76.8 ± 15.1
7 (Female)	26.3 ± 5.5	169.5 ± 6.2	62.7 ± 8.9

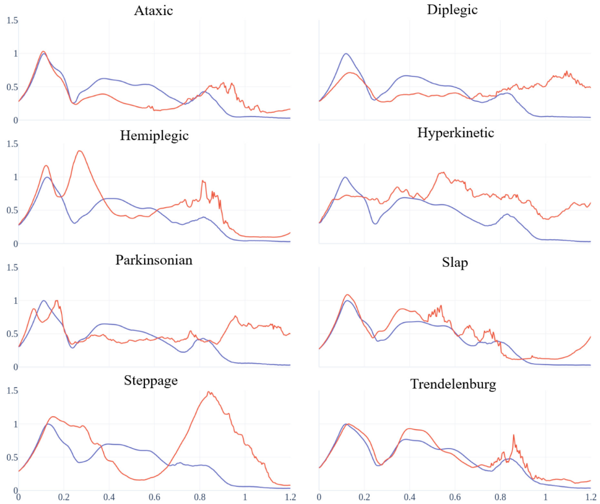

Fig. 1. Example of the typical shape of simulated step of studied gait types in comparison to normal step shape, from the data used in this study. Blue line is normal step shape and red line is corresponding typical shape for this gait type. On X-axis is time in seconds and on Y-axis is normalized magnitude of angular velocities of gyroscope. (Color figure online)

Table 2. Labeled data collected for this study.

Gait type	Total number of recordings for all persons
Ataxic	32
Diplegic	25
Hemiplegic	17
Hyperkinetic	6
Parkinsonian	29
Slap	8
Steppage	32
Trendelenburg	6

step-wise, thus all steps are annotated as *normal* or *abnormal*. Figure 1 illustrates the patterns of each gait type in comparison with a normal step.

Such dataset to the best of authors knowledge is first to have combination of normal and abnormal steps in one dataset. Other datasets are focusing on normal gait patterns; have only abnormal steps in the dataset; compare separate normal gait datasets and abnormal gait datasets, etc. [6,18,26,33].

Data Preprocessing. The collected data is in a form of time-series including a three-axis gyroscope and their calculated magnitude (1).

$$\mathbf{Mag}(X, Y, Z) = \sqrt{\mathbf{X}^2 + \mathbf{Y}^2 + \mathbf{Z}^2}, \tag{1}$$

where $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ are gyroscope axes data vectors, $\mathbf{X} = [x_0, x_1, \ldots, x_i]^T$, $\mathbf{Y} = [y_0, y_1, \ldots, y_i]^T$ and $\mathbf{Z} = [z_0, z_1, \ldots, z_i]^T$, sample index $i \in \mathbb{Z}$. And the $\mathbf{Mag}(X, Y, Z)$ is the magnitude vector of these axes.

To address future works with embedded devices in regard to data transmission and data gathering, data is collected into chunks. One chunk contains M samples for each gyroscope axis. The collected data sample rate is $256 Samples/s$ in the current study. Collected data is labeled stepwise as *"normal"* step or *"abnormal"* step.

Data Preparation for Real-Time Anomaly Detection. For 1D-CNN-AD and LSTM-AD algorithms each person's data is assessed separately. Data for one gait type is prepared by separating training and validation datasets. One gait recording is used as a validation dataset in real-time step anomaly detection estimation, and all other recordings are combined into one training dataset. The ratio between the training and validation datasets can change depending on the person, gait type and available gait recordings for particular gait type.

To enable real-time abnormality detection in the swing phase of the ongoing step, training dataset is divided into overlapping sliding windows. Figure 2 depicts how the windowing of the dataset is designed. As it is shown, each window contains P chunks (i.e., window factor), and each chunk includes M samples and the overlap is N chunks.

Labeling of the windows is conducted according to the labels of the steps. In edge cases, where one step is ending and new step is begging, label is assigned by the proportion of samples of abnormal steps in the window. If this proportion is less than *abnormality proportion threshold* then the window is labeled as *normal*, if more, then it is labeled as *abnormal*.

One of the key advantages of the sliding windows for this study is independence of the anomaly detection algorithms from gait phases.

As a part of hyperparameters optimization, hyperparameters, which affect sizes, overlaps and labels of the sliding windows are investigated. These hyperparameters are a) chunk duration – time in milliseconds, where number of samples M in one chunk is calculated from chunk duration as $M = round(Chunk\ duration * Sample\ rate)$; b) window factor P – determines window size and is proportional to P chunks; c) Abnormality proportion threshold – fraction of the window, which should contain abnormal samples, to consider the label of the window to be abnormal.

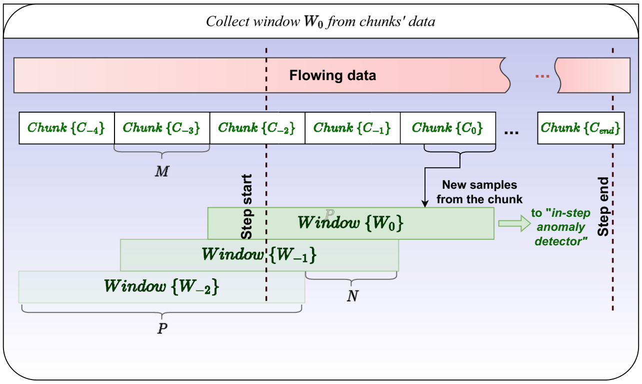

Fig. 2. Windowing of the data for training and for real-time anomaly detection performance estimation. Ongoing gait data is incoming as flowing data, which is split into chunks. From these chunks sliding windows are collected and used in real-time in-step anomaly detector. Step start can be misaligned with sliding window. Chunks are aligned with the sliding windows If abnormality is detected during the chunk C_0, then earliness is time between step start and end of the chunk C_0.

2.2 Proposed Neural Networks

One Dimensional-Convolutonal Neural Network-Anomaly Detection Algorithm. The hypothesis of the 1D-CNN-AD algorithm is following: if real-time gait data could be collected in the form of sliding windows, and neural network could be trained on the dataset using same form of sliding windows with known labels, then it is possible to detect abnormalities in gait during the ongoing step.

The CNN in this study consists of two 1D convolutional layers, max pooling layer, and two fully-connected (dense) layers to provide a binary classification. The 1D-CNN-AD algorithm has the following hyperparameters: i) number of filters; ii) kernel size; iii) batch size; iv) and number of epochs. These hyperparameters would be optimized in this study to achieve the best performance for 1D-CNN-AD algorithm.

In the explorations, the CNN is initialized with a fixed seed of parameters (i.e., weights and bias). The neural network is trained on training dataset with Adam optimizer and cross-entropy loss function. Moreover, a 20% dropout is also considered between the convolutional layer and dense layers.

Long Short-Term Memory Neural Network-Anomaly Detection Algorithm. Hypothesis of the LSTM-AD algorithm is identical to the hypothesis of the 1D-CNN-AD algorithm.

The LSTM-AD algorithm in this work consists of one layer of LSTM followed by two fully-connected (dense) layers to provide a classification probability. The number of cells in the LSTM layer is equal to the number of neurons in the first dense layer. The LSTM-AD algorithm has the following hyperparameters: i) number of LSTM cells; ii) batch size; iii) and number of epochs. These hyperparameters would be optimized in this study to achieve the best performance for LSTM-AD algorithm.

In the explorations, the LSTM is initialized with a fixed seed of parameters (i.e., weights and bias). The LSTM-AD algorithm is trained on training dataset with Adam optimizer and cross-entropy loss function.

2.3 Anomaly Detection

To estimate performance of the real-time anomaly detection of the algorithms, validation dataset is processed in online-fashion. It means, that data is arriving sample by sample. Each sample is collected into chunks. Chunks are collected into windows, as was described in the Sect. 2.1. Algorithms return anomalous class probability for each window, which is collected to the buffer. After the real-time estimation, collected probabilities are analyzed. Different thresholds for anomalous class probability are estimated to achieve best results. This results in the binary classification. These classification results are compared to the labels of the validation dataset, resulting in confusion matrix. Accuracy and F1 score are calculated from confusion matrix.

3 Baseline and Evaluation

3.1 Real-time tsSVM Anomaly Detection Algorithm

RTtsSVM-AD algorithm is based on a *tslearn* [34] Python library. Optimization of hyperparameters is done by dividing training data into two datasets: training and testing with ratio of 70%:30%. Trained classifier with best results for test dataset is used in real-time performance estimation. Model step is calculated as normal step ensemble average from test dataset, which consist of normal steps that have been classified correctly.

The hypothesis of the algorithm is following: if full time-series step pattern could be collected in real-time by combining the average normal step from training phase with the ongoing step data, then anomaly could be detected during the swing phase of the ongoing step by the RTtsSVM-AD algorithm.

We have adopted the RTtsSVM-AD algorithm in our prior work [28] as a baseline for comparing the results of the proposed 1D-CNN-AD and LSTM-AD algorithms in this paper.

Brief overview of the algorithm. Data is collected chunk wise, when step start is detected. Step start and end events are detected if the step detection threshold crosses 20% of the gyroscope magnitude range. Hyperparameter γ is optimized on training and testing datasets. This hyperparameter is used by the global alignment kernel (GAK), where γ is the hyperparameter controlling soft dynamic time warping (softDTW) smoothness [34]. Multiple classifiers with different values of γ could have same performance. Average normal step is created from correctly classified normal steps from training dataset. In real-time in-step gait anomaly detection performance estimation, if step start is detected, data is collected into a chunk. This chunk is replacing corresponding chunk in the model step. Such chunkwise replacement converts regular time-series SVM into the real-time anomaly detection algorithm.

3.2 Evaluation Metrics

For evaluation, several metrics are exploited: Accuracy, F1-score, *earliness*, and real-time factor (RTF). *Earliness* in this paper is defined as – time between the beginning of a step and the moment in time when anomaly is detected in this step. The minimal achievable earliness naturally depends on the gait deviation type. Such a measure has been introduced, because the concrete moment when anomaly starts to occur can fluctuate, depending on a gait type.

3.3 Score and Alarm

For estimation of the performance of the algorithms, anomalous class probability is collected from the classifier. Binary decision is performed later in post-processing of the results. *Score* is the resulting anomalous class probability. For RTtsSVM-AD algorithm *Score* is average score from used classifiers in estimation, because multiple classifiers could be used simultaneously. *Score* is compared with the selected *threshold*, giving alarm signal in (2), finalizing the anomaly detection.

$$Alarm = \begin{cases} 1, & \text{if } S > threshold \\ 0, & \text{if } S \leq threshold \end{cases} \tag{2}$$

If *Alarm* is triggered, then earliness is the time duration from the beginning of the step to the current moment in time.

4 Experimental Setup

For 1D-CNN-AD and LSTM-AD algorithms, the considered hyperparameters are presented in the Table 3 and Table 4.

For the RTtsSVM-AD algorithm parameters used in this work are as follows: a) one chunk is $M = 12$ samples; b) predefined γ values are in the range from 100 to 1000 with an increase of 100 and in the range from 5 to 100 with an increase of 10; c) the step detection threshold is 200°/s.

All training and validation experiments are implemented in Python 3.10.13, tslearn 0.6.2, and TensorFlow 2.9.1 and performed on a prebuilt HP computer

Table 3. Global hyperparameters for 1D-CNN-AD and LSTM-AD algorithms

Hyperparameter	Values
Window factor (P)	6 to 10. Default 8
Chunk size	25 ms to 100 ms. Default 50 ms
Samples in a chunk (M)	6 to 25. Default 12
Sliding window overlap (N)	1
Abnormality proportion threshold	50% to 90%. Default 70%
Batch size	2^n where n is from 3 to 8. Default n is 5
Number of epochs in training	1 to 30. Default 20

Table 4. Algorithm-Specific Hyperparameters

Algorithm	Specific Hyperparameters
LSTM-AD	Number of LSTM cells: 20, 25, 30. Default: 25
1D-CNN-AD	Number of filters in convolutional layer: 2^n where n is from 3 to 8. Default n is 6 Kernel size in convolutional layer: 2, 3, 5, 7, 9, 11. Default 5 Dense layer with 100 neurons

with Intel Core i7 and 16Gb of DDR4 memory. We conducted CPU experiments to model the execution on the embedded devices in future works.

5 Experimental Results and Discussion

Results for the 1D-CNN-AD, LSTM-AD and RTtsSVM-AD algorithms are presented in this section.

5.1 Optimization of 1D-CNN-AD and LSTM-AD Algorithms Hyperparameters

In this paper, optimization is performed by one parameter at a time, while the other parameters are set to their default values.

Chunk Length. The first hyperparameter to consider is the length of the chunk. Table 5 shows the best mean F1 scores with corresponding chunk sizes. It is observed that the best results are achieved with chunk sizes of 75 and 100 ms for all gait types for LSTM-AD and most of the gait types for 1D-CNN-AD. Chunk size of 40 and 50 ms performed better for Steppage, and Trendelenburg gait types for 1D-CNN-AD algorithm. Despite the better performance with longer chunks for some gait types, chunk size is set to 50 ms, with consideration of fast anomaly detection. Larger chunk sizes would lead to slow anomaly detection.

Table 5. Best mean F1 scores for different chunk sizes (CS)

Gait type	LSTM-AD		1D-CNN-AD	
	F1	CS, ms	F1	CS, ms
Ataxic	62.63%	100	79.28%	75
Diplegic	72.36%	100	87.49%	100
Hemiplegic	81.03%	100	83.52%	75
Hyperkinetic	75.95%	100	96.3%	75
Parkinsonian	75.24%	100	84.65%	100
Slap	57.45%	75	78.7%	75
Steppage	75.09%	100	84.17%	40
Trendelenburg	59.65%	75	81.3%	50

Window Factor and Abnormality Proportion. These hyperparameters should be considered in correlation with each other because both of them change the number of samples in the window, which can change the final label of the window. Table 6 presents the best mean F1 scores for combination of window factor and abnormality proportion. It could be seen, that 1D-CNN-AD algorithm is performing best with shorter windows for most of the gait types, whereas LSTM-AD algorithm is performing best with longer windows for most of the gait types. In terms of abnormality proportion threshold, for most of the gait types for both 1D-CNN-AD and LSTM-AD algorithms higher threshold is needed. Only for Hyperkinetic and Steppage gait types it was 70% for LSTM-AD and 60% for 1D-CNN-AD algorithms respectively. It means, that for Hyperkinetic and Steppage gait types edge cases are important for correct anomaly detection. Thus, in general, most of the windows should contain mostly abnormal samples to be labeled abnormal for best performance. With the default settings for other parameters, 1D-CNN-AD algorithm achieves mean F1 scores of 96.3% for Hyperkinetic gait type. On the other hand, LSTM-AD algorithm achieves best mean F1 score of 73.78% for Hemiplegic gait type.

Diplegic and Hyperkinetic gait types have anomalies in the middle and end of the step, thus short windows should be best suited for them to detect abnormality early, as can be seen in 1D-CNN-AD algorithm results. Both Ataxic and Parkinsonian gait types have multiple abnormal steps in a row, which can be similar to normal steps, thus requiring well defined long abnormal windows during the training phase. Slap gait is usually characterized by the sharp short peak at the end of the step, whereas the rest of the step can be similar to normal, thus making it more critical to have a correct classification in edge cases. Steppage gait type have different amplitudes from the normal step for its peaks when the knee is raised up to compensate for lack of movement in the forefoot. Hemiplegic gait type can be similar to a normal gait, which makes it more difficult to differentiate from normal steps which require well-defined shorter windows.

Table 6. Best mean F1 scores for different window factor (WF) and abnormality proportion threshold (AP)

Gait type	LSTM-AD			1D-CNN-AD		
	F1	WF	AP	F1	WF	AP
Ataxic	58.16%	9	90%	78%	10	90%
Diplegic	58.31%	8	90%	82.81%	7	80%
Hemiplegic	73.78%	10	90%	84.03%	6	90%
Hyperkinetic	69.05%	10	70%	95.15%	7	90%
Parkinsonian	63.25%	9	90%	84.38%	10	80%
Slap	62.55%	8	80%	86.9%	6	80%
Steppage	64.83%	10	80%	88.39%	6	60%
Trendelenburg	64.75%	10	80%	83.75%	6	90%

Number of Filters and Kernel Size in the Convolutional Layer for 1D-CNN-AD Algorithm and Number of LSTM Cells for LSTM-AD Algorithm. As presented in Table 7, the best scores for 1D-CNN-AD algorithm are generally achieved with a higher number of filters of 128 and 256, except for Diplegic gait type with 32 filters. This means that extracting more features from the data improves the performance of the 1D-CNN-AD algorithm demonstrating the complexity of the human gait. For Diplegic gait type a smaller network is best suited, meaning that extracting too many features can confuse the 1D-CNN-AD algorithm, because the shapes of the abnormal steps for them are more defined than the ones in other gait types.

Best performance is achieved for 1D-CNN-AD algorithm with medium kernel size of 7 except for Hyperkinetic and Steppage gait types with a kernel size of 11 and for Parkinsonian and Trendelenburg gait types with kernel size of 9. For Hyperkinetic, Steppage, Parkinsonian and Trendelenburg gait types bigger kernel size is needed to neglect the variance between individual abnormal steps in the data.

For LSTM-AD algorithm larger number of LSTM cells results in a better performance, due to the complexity of the gait signal. For Ataxic, Diplegic and Slap gait types algorithm performs best with 25 cells showing, that they have simpler shapes, compared to other gait types. For Parkinsonian gait type the best performance was with 20 cells, meaning, that this gait type, has more pronoun shape, compared to other gait types.

Batch Size and Number of Epochs in Training. As presented in Table 8, the best scores are generally achieved with a bigger batch size of 128 and 256, except for Trendelenburg gait type with a size of 32 for 1D-CNN-AD algorithm, and Slap and Trendelenburg gait types with size of 16 and 64 respectively for LSTM-AD algorithm. This means, that a more accurate training gradient of the neural network is needed for these gait types.

Table 7. Best mean F1 scores for different numbers of LSTM cells (#C) and 1D-CNN kernel size (KS) and number of filters (#F)

Gait type	LSTM-AD		1D-CNN-AD			
	F1	#C	F1	KS	F1	#F
Ataxic	52.4%	25	75.31%	7	75.04%	128
Diplegic	56%	25	81.67%	7	80.3%	32
Hemiplegic	61.37%	30	82.3%	7	87.87%	256
Hyperkinetic	69.5%	30	92.45%	11	86.7%	128
Parkinsonian	57.49%	20	85.45%	9	87.74%	128
Slap	56.95%	25	87.1%	7	81.8%	256
Steppage	60.17%	30	85.23%	11	87.43%	256
Trendelenburg	59.85%	30	83.05%	9	81.75%	128

Table 8. Best mean F1 scores for different batch size (B) and number of epochs (#E)

Gait type	LSTM-AD				1D-CNN-AD			
	F1	B	F1	#E	F1	B	F1	#E
Ataxic	56.11%	128	55.95%	5	77.52%	256	78.95%	3
Diplegic	68.77%	256	71.06%	5	85.92%	256	89.33%	5
Hemiplegic	77.08%	256	81.32%	2	85.08%	256	85.45%	4
Hyperkinetic	60.7%	256	73.2%	2	92.45%	256	98.1%	5
Parkinsonian	68.43%	256	66.38%	4	86.81%	256	88.36%	2
Slap	59.8%	16	55.8%	10	83.9%	256	90.8%	4
Steppage	64.89%	128	67.8%	5	87.7%	256	88.1%	2
Trendelenburg	55.4%	64	57.05%	10	81.3%	32	81.3%	20

In terms of the amount of training required by the algorithms, it is clear, that more than 5 epochs could lead to overfitting, thus reducing classification quality in this study. Only LSTM-AD algorithm performed better with 10 epochs for Slap and Trendelenburg gait types, and 1D-CNN-AD algorithm performed better with 20 epochs for Trendelenburg gait type. This could be due to similarities between normal step and typical step shape for Trendelenburg gait type, thus needing more time to properly fit the network. Considering the overall performance of 55.8% for Slap gait type for LSTM-AD algorithm in epoch optimization, algorithm struggled with this gait type. Training dataset usually contains around 5000 windows, thus every epoch has around 39 iterations with batch size of 128. Normal and abnormal steps have mostly consistent shapes in one gait type. Thus, smaller number of epochs can fit such data better. Larger number of epochs could lead to lower performance due to overfitting of the training data and would trigger anomaly detection while classifying unknown data. Therefore, better results are generally achieved with 2 to 5 epochs.

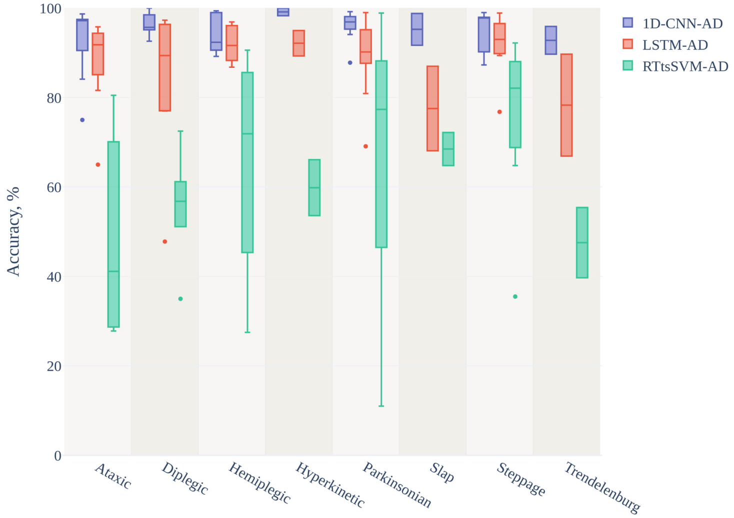

Fig. 3. Distribution of accuracy across different algorithms for all persons for different gait types. On y-axis is accuracy in percents or time in seconds, on x-axis are different gait types.

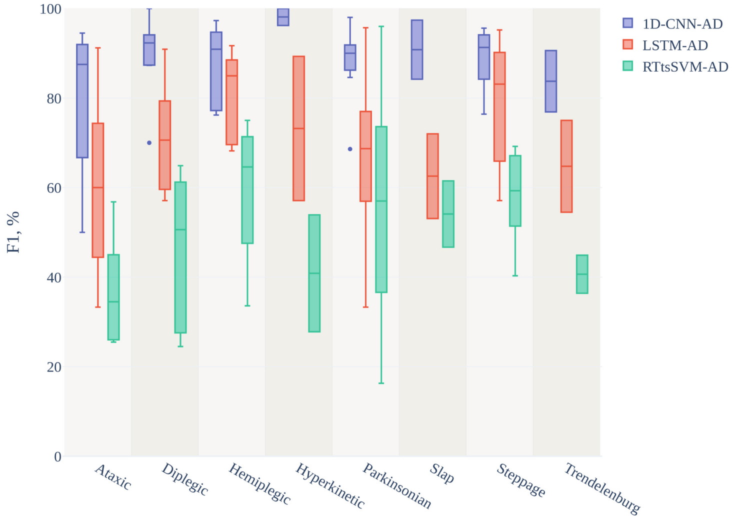

Fig. 4. Distribution of F1 scores across different algorithms for all persons for different gait types. On y-axis is F1 score in percents, on x-axis are different gait types.

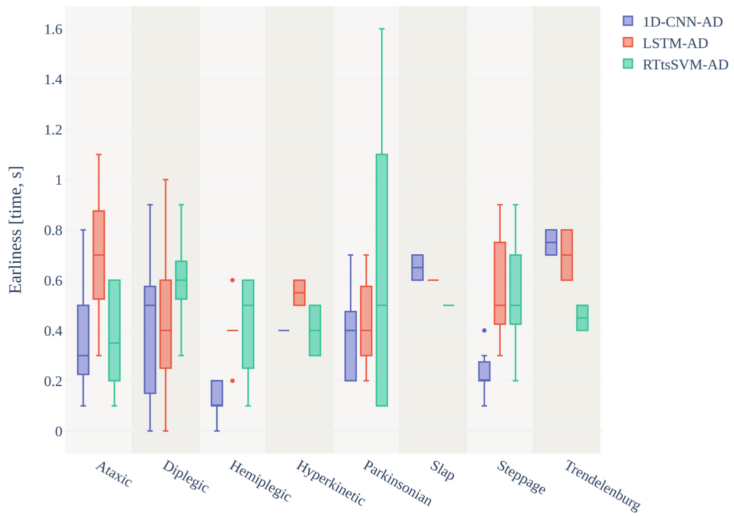

Fig. 5. Distribution of Earliness across different algorithms for all persons for different gait types. On y-axis is time in seconds, on x-axis are different gait types.

Comparison of Algorithms. In Fig. 3 and Fig. 4 could be seen, that both 1D-CNN-AD and LSTM-AD algorithms are outperforming the RTtsSVM-AD base comparison algorithm. The best scores for all gait types are achieved by 1D-CNN-AD algorithm with an average accuracy of 95% and average F1-score of 88%. LSTM-AD algorithm achieved an average accuracy of 87% and average F1-score of 70%. Best results for 1D-CNN-AD algorithm are for Hyperkinetic and Slap gait types with F1 scores of $98.1 \pm 2.7\%$ and $90.8 \pm 9.3\%$ respectively. It could be observed that for Ataxic, Hemiplegic, Slap, Steppage, and Trendelenburg gait types there are some deviations in results from person to person, that could be improved with additional optimization. Best result for LSTM-AD algorithm is achieved for Hemiplegic gait type with average F1 score of $81.32 \pm 9.96\%$. 1D-CNN-AD algorithm is achieving accuracies over 92.6% for all gait types and F1 scores of over 83% for all gait types, except for Ataxic gait type with F1 score of $78.95 \pm 15.43\%$. LSTM-AD algorithm achieved accuracies over 78.3% for all gait types with F1 scores of $71.06 \pm 12.47\%$, $73.2 \pm 22.77\%$ and $79.61 \pm 14.92\%$ for Diplegic, Hyperkinetic and Steppage gait types respectively. Lowest F1 score of $60.24 \pm 18.65\%$ is achieved for Ataxic gait type.

Time of detection is relevant, when classification accuracy is high. Typical normal step length in this study is ranging from 1 to 1.2 s depending on the person, whereas abnormal step duration ranges from 1 to 1.7 s, depending on the person and gait type. Mid-swing phase of the step is starting at around 0.2–0.4 s from the step beginning. Therefore, for the earliness metric depicted in Fig. 5, it could be observed that for most gait types the earliness is less than one second. For Steppage gait type the most common earliness measure is around 0.6 s for RTtsSVM-AD and LSTM-AD algorithms and 0.2 s for 1D-CNN-AD algorithm which is in the middle or at the beginning of a step. For other gait types it could be observed that detection was mainly in the middle of a step, which shows, that algorithms can detect anomalies early, during the mid-swing phase of a step. For some gait types RTtsSVM-AD and LSTM-AD are detecting abnormality earlier than 1D-CNN-AD, but in combination with quality of prediction, 1D-CNN-AD is outperforming other presented algorithms.

Table 9. Average real-time factor for all algorithms

Algorithm	RTF
1D-CNN-AD	0.09 ± 0.03
LSTM-AD	1 ± 0.07
RTtsSVM-AD	9.13 ± 6.54

In Table 9, it can be observed that the main issue of RTtsSVM-AD algorithm is *computational real-time factor*. It means that for every second of incoming data, it takes 9.13 ± 6.54 s to classify it, which is 3 to 15 times longer than the amount of collected data in real-time. The main reason for this is the usage of prediction probability in *tslearn* classifier, which uses an expensive 5-fold cross-validation method to calculate probability. Using regular class prediction is not possible due to inaccurate results from the classifier, as it outputs only zero or one as class identification, drastically reducing classification quality. LSTM-AD algorithm is performing classification in near real-time but not faster than it, because recurrent operations of the algorithm are computationally expensive. Thus, 1D-CNN-AD algorithm is most suitable for real-time applications, for example, to operate in real-time on a real gait assistive device.

This work have several limitations: a) Simulated gait deviation could differ from the real patient's gait with neurological disorders. However, the main goal in this study is to classify the step as normal or abnormal during the mid-swing phase of the step. If patient's normal step pattern after the rehabilitation is sufficiently different from the patient's abnormal step pattern (i.e. because of fatigue or other reasons), then algorithms will be able to detect gait abnormalities during the mid-swing phase of the step as they are able to detect them in this study with simulated gait. Also, as it was stated in the Sect. 2.1: simulations are recreating actual patients' video recordings of gait deviations in collaboration and guidance from a professional physiotherapist of Tallinn East Central Hospital. Thus, such simulated gait types are representing real gait types as close as possible. b) Neural networks in this study are not aware of the gait phase, thus multiple alarms could be triggered during one abnormal step, thus they would be optimized further. Cross-correlation of different hyperparameters could improve classification performance and would be studied in future work.

6 Conclusion

Proposed in this study real-time in-step anomaly detection algorithms are at the very beginning of the research towards context aware assistive devices, which will help to improve gait quality and reduce falling risk for patients suffering from neurological disorders.

Results of this study shows that 1D-CNN-AD algorithm is suitable for real-time anomaly detection in realistic gait deviations during the ongoing step with average earliness of 0.4 s. An average accuracy of 95% and average F1 score of 88% across different studied gait types is achieved for 1D-CNN-AD algorithm, with best F1 score of $98.1 \pm 2.7\%$ for Hyperkinetic gait type. Benefits of this algorithm are, that it is not dependent on gait phases, resistant to the non-optimal hyperparameters and can run in real-time. Second proposed LSTM-AD algorithm achieved average accuracy of 87% and average F1-score of 70% across different studied gait types and best result is achieved for Hemiplegic gait type with F1 score of $81.3 \pm 9.96\%$.

Future gait correction systems and assistive devices will benefit from context awareness in a form of real-time anomaly detection algorithms, leading to more tailored approach for patients suffering from neurological disorders. This will help them to maintain better gait quality, which they obtained after rehabilitation, giving higher chance to continue daily living activities without major restrictions. Main benefit of context aware assistive devices compared to regular assistive devices would be less muscle fatigue from using it. Considering, that FES is used in current assistive devices [16,21,22], where electrical stimulation is given every step, context aware FES would be used only, when step deviation is detected and stimulation is necessary.

Future work will be focusing on further optimization of the presented algorithms, in-step abnormality estimation with more persons and real-time in-step abnormality detection tests with embedded devices running proposed in this study algorithms.

References

1. Stanford Medicine 25: Gait abnormalities. https://stanfordmedicine25.stanford.edu/the25/gait.html
2. Anwary, A.R., Arifoglu, D., Jones, M., Vassallo, M., Bouchachia, H.: Insole-based real-time gait analysis: feature extraction and classification. In: 2021 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL), pp. 1–4 (2021). https://doi.org/10.1109/INERTIAL51137.2021.9430482
3. Bertolote, J.M.: Neurological disorders affect millions globally: WHO report. World Neurol. **22**(1), 1 (2007). https://worldneurologyonline.com/wp-content/uploads/2013/03/WFN-March-2007-Issue.pdf
4. Boompelli, S.A., Bhattacharya, S.: Design of a telemetric gait analysis insole and 1-D convolutional neural network to track postoperative fracture rehabilitation. In: 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), pp. 484–488 (2021). https://doi.org/10.1109/LifeTech52111.2021.9391975

5. Cameron, M.H.: Physical Agents in Rehabilitation: From Research to Practice, 4 edn. Elsevier/Saunders, St. Louis, Mo (2013)
6. Chang, C.W., Yan, J.L., Chang, C.N., Wen, K.A.: IMU-based real time four type gait analysis and classification and circuit implementation. In: 2022 IEEE Sensors, pp. 1–4 (2022). https://doi.org/10.1109/SENSORS52175.2022.9967269
7. Feigin, V.L., et al.: Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the global burden of disease study 2016. Lancet Neurol. **18**(5), 459–480 (2019). https://doi.org/10.1016/S1474-4422(18)30499-X
8. Hollman, J.H., McDade, E.M., Petersen, R.C.: Normative spatiotemporal gait parameters in older adults. Gait & Posture **34**(1), 111–118 (2011). https://doi.org/10.1016/j.gaitpost.2011.03.024
9. Hsieh, C., Shi, W., Huang, H., Liu, K., Hsu, S.J., Chan, C.: Machine learning-based fall characteristics monitoring system for strategic plan of falls prevention. In: 2018 IEEE International Conference on Applied System Invention (ICASI), pp. 818–821 (2018)
10. Hsu, W.C., et al.: Multiple-wearable-sensor-based gait classification and analysis in patients with neurological disorders. Sensors **18**(10), 3397 (2018)
11. Huan, J., et al.: A wearable skin temperature monitoring system for early detection of infections. IEEE Sens. J. **22**(2), 1670–1679 (2022). https://doi.org/10.1109/JSEN.2021.3131500
12. Kluding, P.M., et al.: Foot drop stimulation versus ankle foot orthosis after stroke: 30-week outcomes. Stroke **44**(6), 1660–1669 (2013)
13. Kuusik, A., Gross-Paju, K., Maamägi, H., Reilent, E.: Comparative study of four instrumented mobility analysis tests on neurological disease patients. In: 2014 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops, pp. 33–37. IEEE (2014)
14. Lavado, D.M., Vela, E.A.: A wearable device based on IMU and EMG sensors for remote monitoring of elbow rehabilitation. In: 2022 E-Health and Bioengineering Conference (EHB), pp. 1–4 (2022). https://doi.org/10.1109/EHB55594.2022.9991526
15. Li, R., Song, C., Wang, D., Meng, F., Wang, Y., Tang, Q.: A novel approach for gait recognition based on CC-LSTM-CNN method. In: 2021 13th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), pp. 25–28. IEEE, Hangzhou, China, August 2021. https://doi.org/10.1109/IHMSC52134.2021.00014
16. Matsumoto, S., et al.: Effect of functional electrical stimulation in convalescent stroke patients: a multicenter, randomized controlled trial. J. Clin. Med. **12**(7) (2023). https://doi.org/10.3390/jcm12072638. https://www.mdpi.com/2077-0383/12/7/2638
17. Miller, L., et al.: Functional electrical stimulation for foot drop in multiple sclerosis: a systematic review and meta-analysis of the effect on gait speed. Arch. Phys. Med. Rehabil. **98**(7), 1435–1452 (2017)
18. Moura Coelho, R., Gouveia, J., Botto, M.A., Krebs, H.I., Martins, J.: Real-time walking gait terrain classification from foot-mounted inertial measurement unit using convolutional long short-term memory neural network. Expert Syst. Appl. **203**, 117306 (2022). https://doi.org/10.1016/j.eswa.2022.117306
19. Murray, M.: Gait as a total pattern of movement. Am. J. Phys. Med. **46**(1), 290–333 (1967)
20. Napieralski, J.A., et al.: Classification of subjects with balance disorders using 1D-CNN and inertial sensors. IEEE Access **10**, 127610–127619 (2022). https://doi.org/10.1109/ACCESS.2022.3225521

21. O'Dell, M.W., et al.: Response and prediction of improvement in gait speed from functional electrical stimulation in persons with poststroke drop foot. PM&R **6**(7), 587–601 (2014). https://doi.org/10.1016/j.pmrj.2014.01.001. https://onlinelibrary.wiley.com/doi/abs/10.1016/j.pmrj.2014.01.001
22. Peishun, C., Haiwang, Z., Taotao, L., Hongli, G., Yu, M., Wanrong, Z.: Changes in gait characteristics of stroke patients with foot drop after the combination treatment of foot drop stimulator and moving treadmill training. Neural Plast. **2021**, 1–5 (2021). https://doi.org/10.1155/2021/9480957
23. Pirker, W., Katzenschlager, R.: Gait disorders in adults and the elderly. Wien. Klin. Wochenschr. **129**(3), 81–95 (2017)
24. Pérez-Ibarra, J.C., Siqueira, A.A.G., Krebs, H.I.: Real-time identification of gait events in impaired subjects using a single-IMU foot-mounted device. IEEE Sens. J. **20**(5), 2616–2624 (2020). https://doi.org/10.1109/JSEN.2019.2951923
25. Ramdhani, R.A., Khojandi, A., Shylo, O., Kopell, B.H.: Optimizing clinical assessments in Parkinson's disease through the use of wearable sensors and data driven modeling. Front. Comput. Neurosci. **12**, 72 (2018)
26. Robles, D., et al.: Real-time gait pattern classification using artificial neural networks. In: 2022 IEEE International Workshop on Metrology for Living Environment (MetroLivEn), pp. 76–80 (2022). https://doi.org/10.1109/MetroLivEnv54405.2022.9826927
27. Rostovski, J., Krivošei, A., Kuusik, A., Ahmadov, U., Alam, M.M.: SVM time series classification of selected gait abnormalities. In: Ur Rehman, M., Zoha, A. (eds.) BODYNETS 2021. LNICS, vol. 420, pp. 195–209. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95593-9_16
28. Rostovski, J., Krivošei, A., Kuusik, A., Alam, M.M., Ahmadov, U.: Real-time gait anomaly detection using SVM time series classification. In: 2023 International Wireless Communications and Mobile Computing (IWCMC), pp. 1389–1394 (2023). https://doi.org/10.1109/IWCMC58020.2023.10182666
29. Roth, N., et al.: Do we walk differently at home? A context-aware gait analysis system in continuous real-world environments. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp. 1932–1935 (2021). https://doi.org/10.1109/EMBC46164.2021.9630378
30. Saboor, A., et al.: Latest research trends in gait analysis using wearable sensors and machine learning: a systematic review. IEEE Access **8**, 167830–167864 (2020)
31. Sayeed, M.A., Nasrin, F.: An edge-computing platform for low-latency and low-power wearable medical devices for epilepsy. In: 2023 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS), pp. 1–4 (2023). https://doi.org/10.1109/WMCS58822.2023.10194265
32. Shull, P.B., Xia, H., Charlton, J.M., Hunt, M.A.: Wearable real-time haptic biofeedback foot progression angle gait modification to assess short-term retention and cognitive demand. IEEE Trans. Neural Syst. Rehabil. Eng. **29**, 1858–1865 (2021). https://doi.org/10.1109/TNSRE.2021.3110202
33. Singh, Y., Vashista, V.: Gait classification with gait inherent attribute identification from Ankle's kinematics. IEEE Trans. Neural Syst. Rehabil. Eng. **30**, 833–842 (2022). https://doi.org/10.1109/TNSRE.2022.3162035
34. Tavenard, R., et al.: Tslearn, a machine learning toolkit for time series data. J. Mach. Learn. Res. **21**(118), 1–6 (2020)
35. Wang, F.C., Li, Y.C., Kuo, T.Y., Chen, S.F., Lin, C.H.: Real-time detection of gait events by recurrent neural networks. IEEE Access **9**, 134849–134857 (2021). https://doi.org/10.1109/ACCESS.2021.3116047

36. Wang, L., Sun, Y., Li, Q., Liu, T., Yi, J.: IMU-based gait normalcy index calculation for clinical evaluation of impaired gait. IEEE J. Biomed. Health Inform. **25**(1), 3–12 (2021). https://doi.org/10.1109/JBHI.2020.2982978
37. Zhang, M., Wang, Q., Liu, D., Zhao, B., Tang, J., Sun, J.: Real-time gait phase recognition based on time domain features of multi-MEMS inertial sensors. IEEE Trans. Instrum. Meas. **70**, 1–12 (2021). https://doi.org/10.1109/TIM.2021.3108174