CLASSIFICATION OF FALLERS VERSUS NON-FALLERS FROM GAIT ANALYSIS OF OLDER ADULTS USING LONG SHORT-TERM MEMORY (LSTM)

[1]TOSHIYA MATSUNASHI, [2]TAM Q NGUYEN, [3]DONALD Y. C. LIE

Tokushima University, Tokushima, Japan
Texas Tech University and Texas Tech University HealthSciences Center, Lubbock, TX, USA
E-mail: [1]matsunashi@ee.tokushima-u.ac.jp, [2]tam.nguyen@ttuhsc.edu, [3]Donald.Lie@ttu.edu

Abstract - Gait analysis using a lightweight body-worn sensor is a convenient and inexpensive method that potentially allows for easy and quantitative assessment of fall risks for the geriatric population. In this study, older adults who each worn a gait belt on the lower back were evaluated using data obtained from lightweight sensors including a triaxial accelerometer and a triaxial gyroscope. Algorithms such as artificial neural network (ANN), support vector machine (SVM) and Long Short-Term Memory (LSTM) were used to classify fallers vs. non-fallers. The accuracy obtained by focusing on the temporal change of gait patterns using LSTM is 72.72% for the 1-min lab walk and 81.81% for the 3-day recording of daily movement, which are comparable to state-of-the-art prior arts reported in the literature. These results suggest that long-term monitoring on the temporal changes in gait behaviors can be a rather effective and novel technique for quantitively evaluating fall risks on the community-dwelling older adults performing their daily living activities.

Keywords - Artificial Neural Network (ANN); Support Vector Machine (SVM); Long Short-Termmemory (LSTM).

I. INTRODUCTION

Walking is a natural means of movement in our lives, and it is one of the easiest exercises for human beings. If a person cannot perform stable walking, it may be difficult for him/her to lead an ordinary life and the risk of various illness rises [1]. Falls arevery common and yet dangerous among the elderly population and are a major health concern. With the increase of elderly population all over the world, objective fall risks evaluations from various walking gaits for the geriatrics can be very attractive and useful.In addition, medical expenses increase year by year, and are expected to rise more in the future. Many previous studies have assessed fall risks by patients'self-reporting, and numerous studies evaluated patients' walking gaits performed in specific places and times, and thus they maynot be very practical or clinically relevant. Many research groups such as, ours, have reported the use of accelerometers for fall detection. We have presented a simple solution for effective fall detection using both an accelerometer and two gyroscopes placed along the subject's thoracic vertebrae[2] Using a simple rule-based multi-thresholds algorithm that utilizes the recorded resultant gravitational acceleration, angular change, angular velocity, and angular acceleration, we were able to successfully detect all 60 falls and differentiate between falls and activities of daily living (ADL) with no false positives on young volunteers.

This leads to our follow-up study where we developed a real-time fall detection system using our improved wireless gait analysis sensor (WGAS). This WGAS is capable of differentiating falls vs. various ADL and the Dynamic Gait Index (DGI) tests performed by young volunteers using a Back Propagation Artificial Neural Network (BP ANN) algorithm. The custom WGAS we built, which includes a tri-axialaccelerometer, 2 gyroscopes, a wireless module and a Texas Instruments (TI) MSP430 microcontroller can be worn by the subjects at either T4 (at back) or on a belt-clip (in front at the center of the waist) for performing the various falls, ADL, and DGItests. The raw gait data is wirelessly transmitted from the WGAS to a nearby PC for real-time fall classification, where six features were extracted for the BP ANN. Overall fall classification accuracies of 97.0% and 97.4% have been achieved for the data taken at the T4 and at the belt positions, respectively. The preliminary data demonstrates an overall sensitivity of 97.0% and overall specificity of 97.2% for this WGAS fall detection system, suggesting this can be a good candidate as a real-time low-cost and effective fall detection device for wireless acute-care and wireless assisted-living [3]-[6]. Recently, we have extended our earlier works to make them a bit more clinically relevantby examining feature extraction methodology on the DGI tests and used machine learning classifiers to differentiate actual patients with balance problems versus normal subjects on an expanded cohort of 60 patients. All data was obtained using our custom designed low-cost WGAS containing a basic inertial measurement unit (IMU) worn by each subject during the DGI tests. The raw gait data is wirelessly transmitted from the WGAS for real-time gait data collection and analysis. We have demonstrated predictive classifiers that achieve high accuracy, sensitivity, and specificity in distinguishing abnormal from normal gaits. These results show that gait data collected from ourwearable wireless gait sensor can effectively differentiate patients with balance disorders from normal subjects in real-time

using various classifiers, with accuracy reaching above 96% [7].Oneof our ultimate goals is to be able to use a remote sensor such as the WGAS and the DGI tests and machine learning to accurately stratify an individual's risk for falls.

However, since many fall accidents occur while older adults are performing their average daily living activities, it would be more practical to evaluate the subjects' fall risks by analyzing their walking gaits during patients' long-term everyday living, rather than just performing short laboratory gait testing (such as DGI tests) to achieve a more complete assessment of a patient's gait patterns for accurate evaluation of his/her fall risks. Therefore, in this work we decided to evaluate the fall risks from the gait data recorded in older adults' daily living. We decided to use the high-quality recording obtained from the "Long Term Movement Monitoring (LTMM) Database" in Physio Net, where The LTMM database contains 3-day 3D accelerometer recordings of 71 elder community residents, that can be used to study gait, stability, and fall risks [8]. The acceleration sensor worn by the subjects was held in place at the lower back by a belt with a customized pocket [8]-[9]. Previous studies in the literature on fall risks evaluations using LTMM have often not paid enough attentions tothe time sequence nature of the gait data, where the previous gait pattern can affect the subsequent ones. For this study, we tried to improve the accuracy of fall risks assessment by emphasizingon the temporal change of the gait data.

Therefore, we used the long short-term memory (LSTM) method [10]-[11] to evaluate the risk of falling. LSTM is an evolved version of recurrent neural network (RNN) [12]-[13] developed for learning sequential data. RNN was developed to learn continuous time series data that previous algorithms such as Support Vector Machine (SVM)[14]-[15] and Artificial Neural Network (ANN) [15] could not learn. We will show LSTM does provide higher accuracy than SVM and ANN, and can objectively evaluate the fall risks on community-living older adults from their IMU-recorded gait data during their average daily living.

II. METHOD

In the dataset we used for this work, 71 community-living elderly people (mean age = 78.3 ± 4.71 years, range = 65 - 87 years, mean height = 1.62 ± 0.07m, mean weight: 71.98±12.88kg) were tested. Subjects are divided into faller or non-faller based on self-reporting of past fall experience. If the subject has had more than two falls in the past year, that person is classified as a faller, otherwise they were considered non-fallers. Subjects attached the gait belt on lower part of waist as shown in Fig. 1 and took various tests while walking. Among them, the data from the following two experiments were used for this work:

a. 1-minute walk in lab: Walk in the lab at a personal selected comfortable speed with the gait belt on lower back.
b. 3-day long term recording: After a 1-minute walk test in the lab, subjects were required to spend 3-days wearing the gait belt, except for taking shower, etc.but the reason for the removal must be reported.)

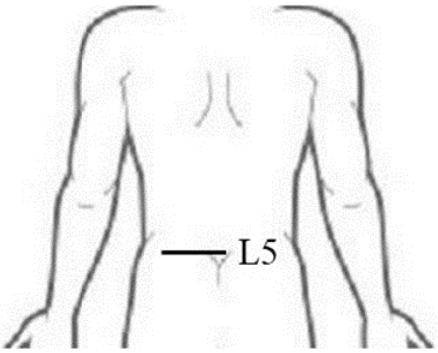


Figure 1.Indicating wherethe gait belt was attached to subjects (at L5; lower back)during gait data recording.

The gait belt is a light-weight and small sensor (74g, 87×45×14mm) and includes a triaxial accelerometer(sensor range and resolution: ±6 g and ±mg for each axis)and a triaxial gyroscope (sensor range and resolution:±100° /second and ±0.0069°/second for each axis). The obtained data are three acceleration axes, vertical acceleration(V), mediolateral acceleration (ML), anterior posterior acceleration (AP), and 3 angular velocity axes, yaw, pitch and roll.

Figure 2 shows the 1-min lab walk data obtained from one subject (non-faller). From top to bottom, vertical acceleration (V), mediolateral acceleration (ML), anterior posterior acceleration (AP), yaw angular velocity, pitch angular velocity and roll angular velocity are displayed, respectively. There are a lot of sequential data (text, audio, video, etc.) all over the world. Since ANN and SVM use the data at a given moment to make predictions, it was not possible to classify sequential data well.Therefor, Recurrent Neural Network (RNN) was developed. RNN predicts the current state from the previous state, and predicts the next state from current state. LSTM is one of the models for sequential data analysis that appeared as an extension of RNN. LSTM is realized by replacing the unit in the middle layer of RNN with a memory,and a block with three gates that are called LSTM block as shown in Fig. 3.In RNN, in order to maintain the storage ofinput data, the middle layer has a loop. This allowed RNN to include the previous data in the judgment. However, the basic RNN had various issues such as the gradient loss problem, and its learning is not very effective.

The LSTM block includes a forget layer, a retention layer, and so on[16]. The feature of LSTM is that it can learn long-term dependencies that were difficult to

learn in the RNN and can solve the gradient loss problem. In other words, it can bedivided into necessary information and unnecessary information, then proceed with learning using only the necessary information. Thus, LSTM has attracted the attention of its learning superiority on long-term sequential data that was difficult to predict in RNN.

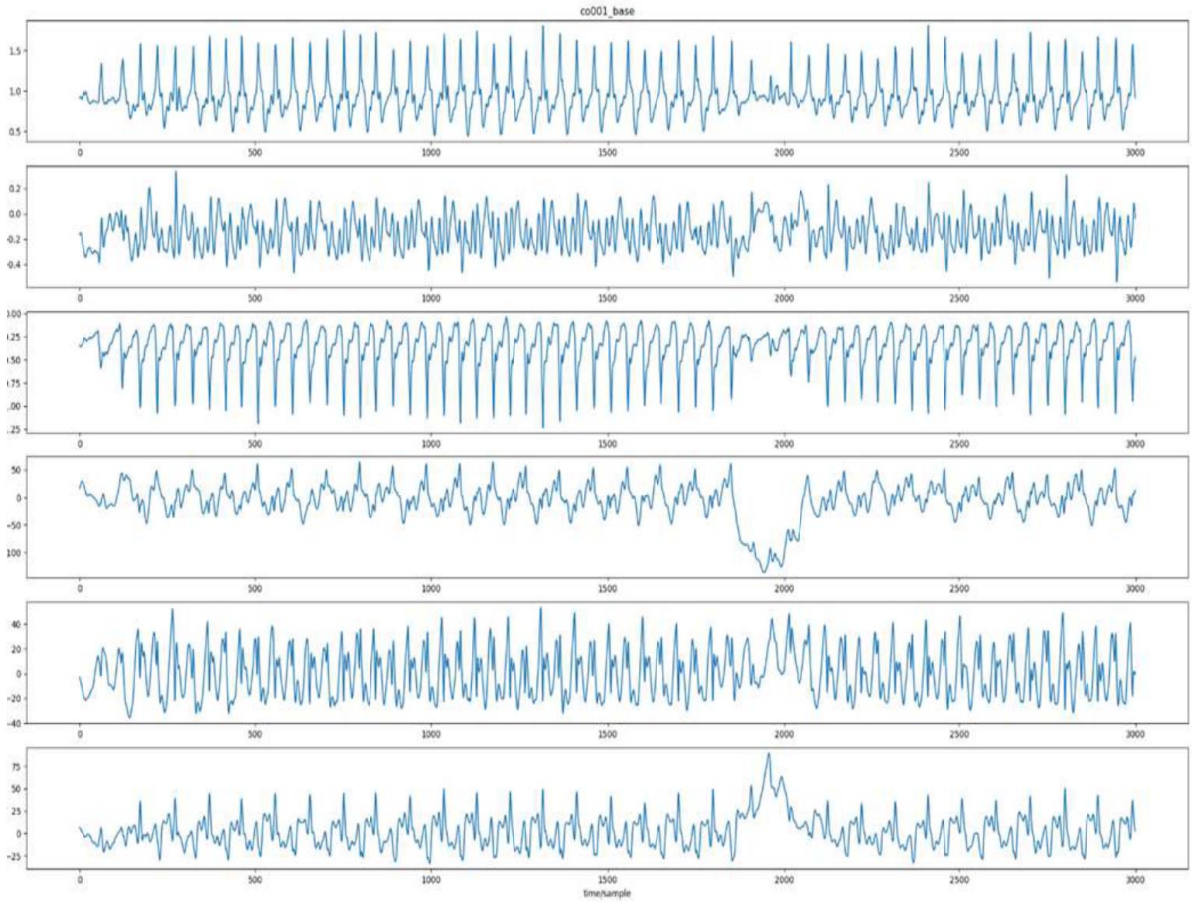


Figure 2.An example of recorded data from a non-faller subject (1-min lab walk).

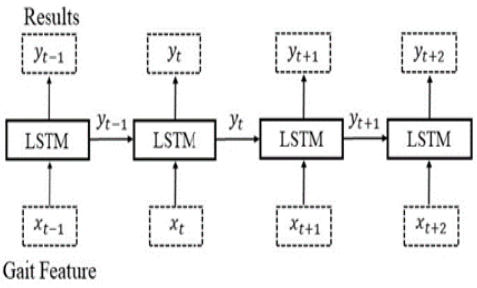


Figure 3.Learning Structure of LSTM.

III. RESULTS AND DISCUSSIONS

In this study, we classified fallers vs. non-fallers using SVM, ANN and LSTM as algorithms. However, since SVM and ANN are not directly suitable for classification of sequential data, the data obtained as input data was modified and used.Range, standard deviation (STDEV) and Range + standard deviation (STDEV) were used as input data when we classify using SVM and ANN. Range is the maximum value minus the minimum value of the signal obtained on each axis. The reason for using the STDEV is that the fallers'movement in their daily living is likely to be unstable, and the variation in data might become large, and it is thus reasonable that larger STDEV could differentiate themselves from the non-fallers [2-7].Table 1 shows the results when we use SVM and ANN as classifying algorithms; 70% was used as training data and 30% as test data.As can see from the Table1, not all features have achieved very high accuracy. In particular, when only STDEV was used in the input data, it showed the same or lower accuracy than when using Range alone. This indicates that the variability of the data may not be significantly different between fallers and non-fallers. In the case of Range + STDEV, a high accuracy of 71.67% was obtained when 1-min lab walk was classified using ANN, but thisdid not improve accuracy for other cases. Forthe 1-min walk, ANN had higher accuracy than SVM in all features by ~2% to 16%. For the 3-day recording, ANN has higher accuracy than SVM in two features by 5%-6%, but in the case of STDEV it has 4% lower accuracy. Although the overall accuracy is not so high as shown in Table 1, it is suspected that the cause is at least partially due to that classification was performed by ignoring the time factor despite of the

time series sequential gait data we are given. In the case of Range, in particular, it uses the maximum and minimum values in tests that can be for a very long time, so it may not be very suitable for accurate classification.

Next, analysis was performed using LSTM. Note in the case of the above ANN and SVM, the obtained gait data was transformed (features extracted) to make it easier to classify, but in the case of LSTM, it was classified using the obtained raw gait data (all six axes data were used as input data) without deformation. This is done to preserve the continuous time series component of the recorded data by not transforming it into range, etc.

Figure 4 shows the accuracy and loss when learning from the 1-min lab walking and the 3-day recording using LSTM. In each figure, the vertical axis is accuracy, and the horizontal axis is the learning process. The blue line is the training data, and the orange line is the test data. Again, 70% was used astraining data and 30% as test data. The upper figure is a 1-minute laboratory walk, and the lower figure is a 3-day record of daily activities. In either case, the accuracy is not initially as high as 50%, but we can see that the accuracy gradually increases as learning progresses. Also, in the loss it can be seen that it decreases gradually with time, andwecan observe that learning is smooth without over-learning. The final classification accuracy was 72.72% for the 1-minlab walking, and a very high 81.81% for the 3 days of recording, as shown in Table 2. Compared to ANN and SVM, LSTM gave that by using LSTM, it can help to predict future results from past data.

Table1.Overall classification accuracy of fallers vs. non-fallers using ANN and SVM for both1-min and 3-day recordings

Feature	Overall Classification Accuracy			
	1-min Walk		3-day Recording	
	ANN	SVM	ANN	SVM
Range	67.27%	56.52%	67.27%	61.10%
STDEV	60.76%	58.79%	57.14%	61.91%
Range + STDEV	71.67%	55.91%	57.14%	52.38%

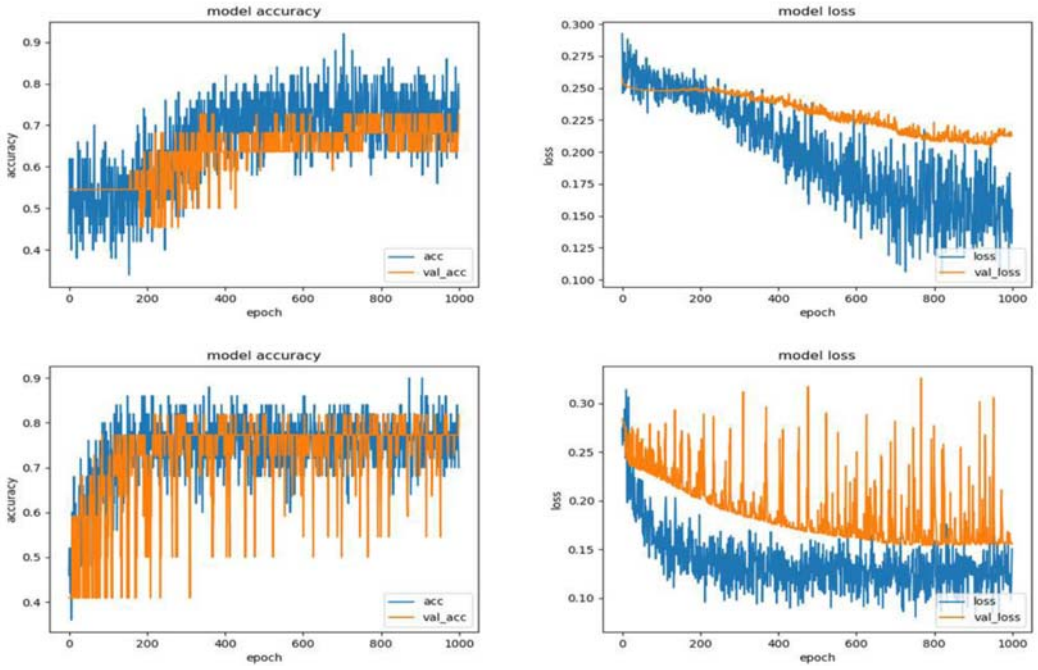


Figure 4.LSTM classification results: (Upper left) accuracy from the 1-min lab walk data; (upper right) loss from the 1 minlab walk data; (lower left) accuracy from the 3-day recording data; (lower right) lossfrom the 3-day recording data.

Table2.Final accuracy with LSTM.

Classifier	1-min Walk	3-day Recording
LSTM	72.72%	81.81%

From the above analysis and results, it is clear that we can classify fallersvs. non-fallers with high accuracy by classifying using LSTM rather than using ANN or SVM.The final accuracy after learning is summarized in Table 2. One goal of this study was to use the temporal characteristics from activities in daily living and to assess the fall risks. Human beings are likely to do the same routines every day in their daily lives, especially for the elderly people who live at home all day long. The walking patterns and gait activity patterns obtained differed among subjects, but it was shown LSTM is useful for predicting fallers in the 1-min lab walk and more so from the 3-day recording. This is likely because themuch longer-term 3-day gait

recordings of daily living activities are simply more realistic and representative forcapturing patients' unstable gait patterns to evaluate fall risks,compared with the somewhat artificial1-min lab walk recording. However, the increased accuracy may also be contributed by the fact that the amount of 3-day recording data is simply much larger for the LSTM to learn more effectively. Note when dealing with the prediction accuracy of sequential data, naturally, when the amount of data is large, the prediction accuracy will often be higher if the sequential data are strongly correlated with time.

Table3.Summary of each classifier's speed when different input features are used.

Classifier	Overall Classifier Speed (s)	
	1-min Walk	3-day Recording
ANN	7.681	7.379
SVM	1.079	1.128
LSTM	27249	39815

However, LSTM is not full of good points only. Of course, when forecasting using LSTM, the more data we have, the longer it will take to calculate. Accuracy is important for fall risks prediction, but the speed of calculation may also be important for performing fall risks assessment. Even if high accuracy is obtained, real-time analysis is not feasible if classification takes more than a few minutes.

Therefore, it is necessary to compare the calculation speed of each classifier. Fast calculation speeds lead to quick feedback from the doctor to the patient, for example. Table 3 summarizes the calculation speed of each classifier. SVM takes only about 1 second to calculate on a typical PC (Windows 10 Pro, Intel core i5-7300U, 2.60GHz, 2.71GHz, 8.00GB, 64 bit,64-bit operating system), and ANN takes about 7 seconds. In the case of LSTM, the 1-mi labwalk classification takes 27249 seconds and the 3-day recording classification takes 39815 seconds. As a result, although LSTM has higher accuracy than SVM and ANN, it has the disadvantage that it takes much more calculation time. Even though with higher accuracy.It turned out that SVM with a simple structure is very fast and suitable for real-time gait analysis. Which classifier is best for fall risks evaluation is likely to vary depending on the specific application. If we don't need immediate feedback and can use a fast server to calculate, we shouldopt to spend more time making predictions with LSTM as it is considerably more accurate.

Table4.Our LSTM classification vs, the literature.

Reference	Algorithm	Analysis Method	Accuracy
[17]	Binary Logistic Regression	Consideration of Average Step Duration, Mediolateral Amplitude, Age and Median Bout Duration	71.6%
[18]	NIPALS	Initial Perturbation at 0%, 20%, 40%, 60% and 80% of the Step Cycle	82%
This work	LSTM	Treat Obtained 3 Acceleration Axes and 3 Angular Velocity Axes as Time-Series Data	81.81%

Table 4 shows our LSTM classification results compared with previously published state-of-the-art data in the literatures using the same recorded data on PhysioNet.For example, Weiss et al. [16] explored added value from the 3-day recording by using binary logistic regression. The results of gait tests to evaluate the gait conditions of subjects such as DGI, Berg Balance Scale (BBS), Timed Up and Go (TUG), Four Square Step Test (FSST), etc. were used. In performance-based classification, 67.2%subjects were successfully identified fallers or non-fallers. In addition to the test results, classification was performedby adding 3-day recording of acceleration measurements to the previous feature, and then 71.6% subjects were correctly classified. The features remained in the final step were Average Step Duration, Mediolateral Amplitude, Age and Median Bout Duration.

Lhlenet al. [17] compared the ability to distinguish between phase-dependent local dynamic stability measurements and conventional gait function. Using the Initial Perturbation at 0%, 20%, 40%, 60% and 80% of the Step Cycle, the non-fallersbecome significant different compared to the 0%, 20% and 60% fallers of the step cycle. It had a large local dynamic stability, median. And when classified by nonlinear iterative partial least square (NIPALS) using a total of 46 gait features, such as phase-dependent measures at 0% and 60%, an accuracy of 82% was obtained.

We, instead, tried to classify the differences in movement from the time-domain data among non-fallers and fallersfrom their daily living using long-term sequential data (3-axis acceleration and 3-axis angular velocity). However, it can be very difficult to classify the fallersvs. the non-fallers from the large and general gait recording of their daily living, as there are various noises in data from their daily lives; it is expected that the specific movement tests such as the DGI gait test is more likely to show

the difference between the fallers vs. non-fallers. Therefore, one limitation in this study is that it may be difficult to determine which movement or gait patterns in the subjects' daily livingare needed for improving the classification accuracy. Finding out which section of the large datasetcan be important to not only reduce the calculation time but also increase the classification accuracy, rather than blindly using all the long-term recording data.

CONCLUSION

The results presented in this study using ANN, SVM, and LSTM suggest that we can classify fallers vs. non-fallers with considerably higher accuracy using LSTM than ANN or SVM. In this study, we used data collected from the 1-min walking and the 3-day daily living activities for 71 subjects and achieved a classification accuracy of 72.72% and 81.81%, respectively. These results suggest that long-term monitoring and the analysis on the temporal changes in gait behaviors can be a rather powerful and novel technique for quantitively evaluating fall risks on the geriatric patients performing their average daily living activities. Furthermore, we suspect thatmachine learning algorithms such as LSTM combined long-term gait "big-data" monitoring might also help to evaluate the effectivenessof certain fall prevention measures for these older adults of higher fall risks as well.

REFERENCES

[1] J. Jacob, T. Nguyen, S. Zupancic and D. Y. C. Lie, "A Fall Detection Study on the Sensor Placement Locations and the Development of a Threshold-Based Algorithm Using Both Accelerometer and Gyroscope", IEEE Int'l conf. on Fuzzy Logics (Fuzz'11), pp. 666-671, Jun 2011.

[2] D. Y. C. Lieet al., "The Design of Real-Time Wearable Fall Detection Systems Aiming for Fall Prevention", "Activities of Daily Living (ADL): Cultural Differences, Impacts of Disease and Long-Term Health Effects", Chapter 3, Edited by Scott T. Lively, Series: "Public Health in the 21st Century", Nova Science Publishers, Inc., NY, 3rd Quarter, pp.51-76, 2015, ISBN 978-163463-9132.

[3] B.T. Nukala et al., "An Efficient and Robust Fall Detection System using Wireless Gait Analysis Sensor with Artificial Neural Network (ANN) and Support Vector Machine (SVM) Algorithms", Open Journal of Applied Biosensor, pp.29-39, 2014.

[4] N. Shibuyaet al., "A Real-Time Fall Detection System Using a Wearable Wireless Gait Analysis Sensor and a Support Vector Machine (SVM) Classifier", Proc. IEEE 8th Int'l Conf. on Mobile Computing and Ubiquitous Networking, pp 66-67, Jan 2015.

[5] B.T. Nukalaet al.,"Comparing Nape vs. T4 Placement for a Mobile Wireless Gait Analysis Sensor Using the Dynamic Gait Index Test", Proc. IEEE 8th Int'l Conf. on Mobile Computing and Ubiquitous Networking, pp. 68-69, Jan 2015.

[6] B.T. Nukalaet al.,A Wireless Gait Analysis Sensor for Real-Time Robust Fall Detection Using an Artificial Neural Network", Proc. IEEE Point-of-Care Technologies Conf. pp. 219-222, Oct 2014.

[7] T Q. Nguyen, J. H. Young, A. Rodriguez, S. Zupancicand D. Y.C. Lie,"Differentiation of Patients with Balance Insufficiency

[8] (Vestibular Hypofunction) vs. Normal Subjects Using a Low-Cost Small Wireless Wearable Gait Sensor", Biosensors, Special

[9] Issue on "Feature Papers: State-of-the-Art Biosensors Technology 2018", Biosensors 2019, 9, 29

[10] A. L. Goldbergeret al., PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation101(23):e215-

[11] e220[2000(June13). https://physionet.org/physiobank/database/ltmm/

[12] S. A. England and K. P. Granata, "The influence of gait speed on local dynamic stability of walking,",Gait&Posture, Volume25, Issue2, pp. 172-178, Feb2007.

[13] X.Du, R. Vasudevan and M. Johnson-Roberson, "Bio-LSTM: A Biomechanically Inspired Recurrent Neural Network for 3-D Pedestrian Pose and Gait Prediction", IEEE Trans. IEEE Robotics and Automation Letters, Volume4, Issue2,pp. 1501-1508, Jan 2019.

[14] Y. Feng, Y. Li and J. Luo, "Learning Effective Gait Feature Using LSTM", IEEE Trans. 23rd International Conference on Pattern Recognition (ICPR), pp.325-330, Dec 2016.

[15] Y. Yang, D. Yu and G. Li, "Gait Recognition Using Flow Histogram Energy Image", IEEE Trans. 22nd International on Pattern Recognition, pp. 444-449, Stockholm, Sweden, Aug24-28, 2014.

[16] D. Liu, X. Wu, W. Du, C. Wang, C. Chen and T. Xu, "Deep Spatial-Temporal Model for rehabilitation gait: optimal trajectory generation for knee joint of lower-limb exoskeleton", Assembly Automation, Vol. 37 Issue: 3, pp. 369-378, Feb 2017.

[17] H. Jeon, J. Han, W. Yi, B. Jeon, K. S. Park, "Classification of Parkinson gait and normal gait using Spatial-Temporal Image of Plantar pressure", Proc. 30th Annual Int'l Conf. of the IEEE Engineering in Medicine and Biology Society (EMBS), pp.4672-4675, Aug 2008.

[18] B. T. Nukalaet al., "Real-Time Classification of Patients with Balance Disorder vs. Normal Subjects Using a Low-Cost Small Wireless Wearable Gait Sensor", Biosensors, 6(4), 58, Nov 2016.

[19] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", Neural Computation, Volume9, Issue8, pp. 1735-1780, Nov 1997.

[20] A. Weiss, M. Brozgol, M. Dorfman, T. Herman, S. Shema, Nir Giladi and J. M. Hausdorff, "Does the Evaluation of Gait Quality During Daily Life Provide Insight Into Fall Risk? A Novel Approach Using 3-day Accelerometer Recording",Neurorehabilitation and Neural Repair, Volume27,Issue8, pp. 742-752, Oct 2013.

[21] E. A. F. Ihlen, A. Weiss, J. L. Helbostad and J. M. Hausdorff, "The Discriminant Value of Phase-Dependent Local Dynamic Stability of Daily Life Walking in Older Adult Community-Dwelling Fallers and Nonfallers", BioMed Research International, Volume 2015, Article ID 402596, 11pages, Sep 2015.

★★★