Automated Classification of Pathological Gait after Stroke using Ubiquitous Sensing Technology

Elham Dolatabadi, Babak Taati and Alex Mihailidis

University of Toronto; Toronto Rehabilitation Institute, Toronto, ON, Canada

elham.dolatabadi@mail.utoronto.ca, babak.taati@uhn.ca, and alex.mihailidis@utoronto.ca

Abstract— **This study uses machine learning methods to distinguish between healthy and pathological gait. Examples of multi-dimensional pathological and normal gait sequences were collected from post-stroke and healthy individuals in a real clinical setting and with two Kinect sensors. The trajectories of rotational angle and global velocity of selected body joints (hips, spine, shoulders, neck, knees and ankles) over time formed the gait sequences. The combination of k nearest neighbor (kNN) and dynamic time warping (DTW) was used for classification. Leave one subject out cross validation was implemented to evaluate the performance of the binary classifier in terms of F1-score in the original feature space, and also in a reduced dimensional feature space using PCA. The pair of k = 1 in kNN and the warping window size 25% of gait sequences in DTW achieved maximum F1-score. Using PCA, pathological gait sequences were discriminated from healthy sequences with the F1-score = 96%.**

I. INTRODUCTION

Early diagnosis of pathological gait prompt early interventions, which can help prevent dysfunction and subsequent loss of independence [1], [2]. Early diagnosis could be facilitated if technology enables frequent gait analysis in the home or in any clinic, i.e., without the need to visit a specialized gait clinic. This would require the use of easy to use sensing devices for home uses, as well as algorithms capable of processing sensed data, e.g., machine learning (ML) methods to analyse patterns of walking [3].

Recent advances in sensor technologies and ML techniques have encouraged effort into the development of the intelligent and ubiquitous solutions to describe, predict, and classify a person's gait at different settings [4], [5], [6], [7], [8], [9], [10]. As a first step towards the development of a gait diagnostic tool, we focused on using ML to automatically discriminate between healthy and pathological gait, modeled here as a (binary) classification task [3]. Examples of pathological gait pattern were collected from individuals with stroke.

To date, much of the effort in this area has relied either on wearable technology [5] (which may cause user discomfort [8]), or on motion capture systems [6] (which are prohibitively expensive and require elaborate setup [10]). Many of recent studies using more affordable and unobtrusive sensors have only published preliminary analysis with healthy participants and data collected in a contrived lab setting [9], [11], [12], [13], [14] (Notable exceptions with in-home implementation and analysis include the work of Skubic et al., e.g., [15], [16]).

We have collected our data using an unobtrusive sensing technology and in a real clinical setting. We capture multi-dimensional gait sequences of two groups of participants: healthy and mobility impaired due to stroke. Specifically, we examine the applicability of the Microsoft kinect sensor for windows (K4W) v2 and ML methods to distinguish between healthy and stroke gait. The K4W v2 is affordable and convenient to use and tracks the human body without the need to wear any markers. Moreover, its accuracy has been analysed thoroughly by several research groups [17]. As such, K4W has attracted considerable attention in gait analysis.

The ML approach is motivated by the fact that all individuals with normal gait share a similar and harmonized pattern [18]. Deviations from this pattern could be used to detect pathological gait. Therefore, a k-nearest neighbour (kNN) classifier, with dynamic time warping (DTW) as a similarity measure, offers a simple but appealing approach to detect pathological gait sequences [19]. This combination benefits from the advantages of each approach. kNN is straightforward to implement and has been shown that its error ratio is at most twice the Bayes error ratio [20]. In addition, DTW has been used quite successfully in time series analysis applications, including gait recognition [21], word recognition [22], and speech recognition [23]. DTW measures the similarity between sequences by allowing them to be stretched or compressed [20]. As a similarity measure, DTW can handle local time shifting and is robust with respect to noise and misalignment in time between each pair of sequences. Once DTW score is obtained, a KNN is used to predict class labels for test sequences.

II. METHODS

A. Study Description

Gait sequences of two groups of participants (healthy and stroke) were collected. Group 1 included 5 healthy young adults (age: 28.8±7.1 years) and group 2 included 5 post-stroke adults with a variety of mobility status as listed in Table I. Walking sequences were collected at a gait clinic located at Toronto Rehabilitation Institute in Toronto, Canada. This clinic is equipped with a pressure mat to perform quantitative analysis of gait for eligible stroke in-patients during their stay at the hospital. Data collection for post-stroke participants took place at the same time they were examined at the clinic as a part of their required clinical

TABLE I
STROKE PARTICIPANTS INCLUDED IN THIS STUDY.

ID	Age	Aid	Mobility Status	More affected side
1	87	None	independent	Bilateral
2	72	Rollator	supervised outdoor	Left
3	55	Quade Cane	supervised indoor/outdoor	Left
4	49	None	independent	Right
5	79	Two wheeled walker	supervised indoor/outdoor	Right

Fig. 1. The orientation information and hierarchical orders of 12 body joints in reference to the hip centre.

care. For consistency, the same gait protocol was followed for participants in Group 1 as well. This study was approved by the Research Ethics Board at the Toronto Rehabilitation Institute - University Health Network. All participants from the two groups were completed informed consent forms.

Two K4W v2 sensors were installed in the clinic in such a way that no constraint was imposed on participants as they were walking along the mat. The sensors were mounted on each side and less than a meter (roughly 80 cm) away from the beginning of each walk. One sensor captured participants while walking in one direction, and the second sensor captured their walking steps after turning around and while walking back.

B. Feature Extraction

The human body pose during walking was represented by the rotational angle and the global velocity of selected body joints (see (Figure 1)). The absolute orientations of the body joints, as tracked by the Kinect, is expressed in the form of quaternion and in the sensor coordinate. Therefore, it is essential to implement the following preprocessing steps to represent hierarchical joint information in world coordinates, i.e., independent of the placement of the sensor. The hierarchical orders of human body joints used in this study are displayed in Figure 1.

In the first step, the absolute orientation of all joints were multiplied by a rotation-translation matrix, retrieved from the two orthogonal planes visible in the scene (e.g., two walls, or a wall and the floor) [24]. In the second step, the hierarchical orientation of all body joints with respect to the root joint (hip centre) were calculated as follows:

$$R_c^h = R_p^{a^{-1}} \times R_c^a \qquad (1)$$

where R_c^h is the hierarchical orientation of the child joint, R_p^a is the absolute orientation of the parent joint in the world coordinates and R_c^a is the absolute orientation of the child joint. Note that the orientations of the elbows and wrists are also tracked by the Kinect, but there are are not used in this study. In the third step, a median filter was used to smooth and stabilize the joint orientations over all collected frames for a given sequence. For the filter design, the length of the sliding window was set to three frames. First, The spherical interpolation (SLERP) of the corresponding quaternion was computed with each of its two neighbors. Then, the corresponding quaternion was replaced with the average of computed interpolants. The interpolation parameter in SLERP was set as a function of the angle between the corresponding quaternion and each of its
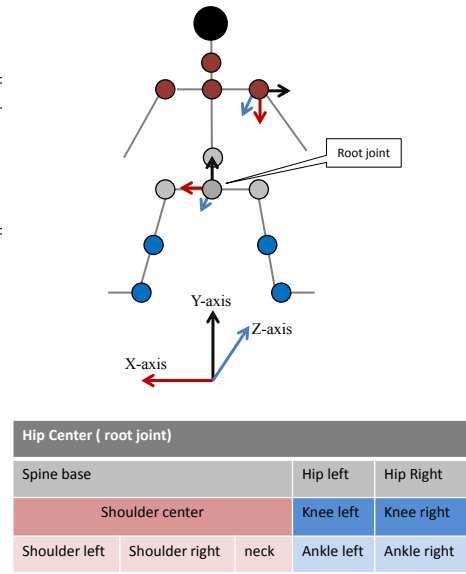
neighbors. The fourth (and final) preprocessing step involved converting the smoothed hierarchical orientations of selected body joints from quaternions to yzx Euler angles. The Euler angles were mapped to be between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$ and rescaled such that the variance of each dimension was one.

Given the above preprocessing steps, each gait sequence in this study included the trajectories of the following features over time:

- The hierarchical orientations of 10 body joints (all joints shown in Figure 1, except the root joint which is the hip centre), and
- Velocity and acceleration of the hip centre

Therefore the gait sequences are a $D \times T$ dimensional arrays, where D is the number of features and T denotes the number of frames. Note that (D = 36) is the same for all participants, while T varies depending on the duration of a walking sequence. All gait sequences in the database were segmented to contain one gait cycle only. Sequences that originally contained more than one gait cycle (the majority of them), were thus broken down to multiple single gait cycle sequences. In total, each participant completed eight to ten sequences of walks in front of the Kinect sensors and 194 gait cycles were collected.

C. sequence classification

All gait sequences from 10 participants were concatenated and considered to be independent for the task of classification. The binary classifier, therefore, predicts a label (c1 or c2) for each unseen gait sequence to be healthy or pathological.

The basic kNN assigns a class label to a new test sequence via majority voting among the k nearest neighbors in the training set. In this study, the distance between two walking sequences is their DTW score. An underlying assumption in

kNN is that the training set is balanced, i.e. data is equally distributed between the two classes. However, datasets associated with human subject research are often unbalanced, because it is easier to collect data from healthy participants, and this affects the performance of kNN. To address this issue, weighted kNN is used here; i.e. larger weights are assigned to neighbors from the class with smaller number of sequences (gait cycles in this study) [25]. Moreover, in order to further improve the performance of the kNN, as suggested in the studies [26], weights were parameterized in the way that nearer neighbors contribute more to the prediction than the more distant ones. In this study, the combination of kNN and DTW (kNN-DTW), assigns a test sequence Y_{test} into one of the two classes to maximize the following posterior probability:

$$p(c_i|Y_{test}) = \sum_{j \in nbd} \frac{1}{(c_i/min(c_1, c_2))^{\frac{1}{3}}} \omega_j \delta_{c_i} \qquad (2)$$

where k is the number of nearest neighbors used in prediction of Y_{test}, nbd is the k nearest neighbors to Y_{test}, c_i is the classifications of the training sequences (c1 or c2) in nbd and δ is 1 for c_i and 0 otherwise. ω_j is the parameterised weight value defined in the following way:

$$\omega_j = \begin{cases} \frac{S_k - S_j}{S_k - S_1} & \text{if } k \neq 1; \\ 1 & \text{if } k = 1. \end{cases}$$

where S is the DTW score (S_1 is the DTW score corresponds with the nearest neighbor and S_k corresponds with the most distant sequence in nbd).

DTW can be constrained by limiting the size of the warping window, w, as a percentage of the length of the longer of the two sequences. Note that the Euclidean distance between two sequences can be seen as a special case where the warping window equals zero. Constrained DTW would improve the speed and accuracy of similarity measures as extended warping may introduce incorrect matches between different time series and distort true similarities [20].

Performance of the classifiers was quantified using F1-score which is defined as:

$$F1 - score = \frac{2 \times Recall \times Precision}{Recall + Precision}, \qquad (3)$$

where Recall is counts of correctly classified pathological sequences/counts of total pathological sequences and Precision is counts of correctly classified pathological sequences/counts of total classified pathological sequences.

Leave one subject out Cross Validation (LOOCV) was used in all experiments where, at each iterations, one participant and all pertaining sequences were left out. Grid search was performed in LOOCV with k ranging from 1 to the square root of total number of sequences in the training set (on average 160 gait cycles at each repetition of LOOCV), and w as 10%, 25% and 100% of the length of the gait sequence.

Implementation of DTW in high dimensional feature spaces is computationally expensive [27]. Thus, a similar
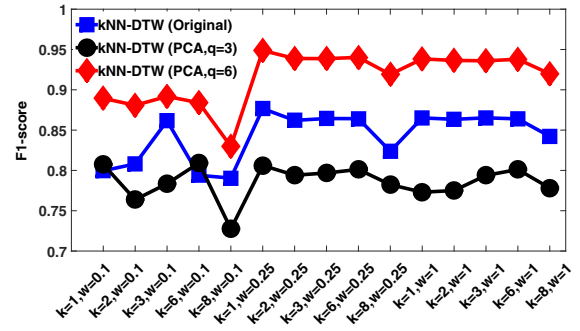

Fig. 2. Changes in F1-score for different values of parameter k (in kNN) and w (size of warping window in DTW) as a fraction of the length of the gait sequence. The plot also shows the performance of kNN-DTW in high dimensional feature space and low dimensional feature space using PCA.

TABLE II
CLASSIFICATION ACCURACY FOR POST-STROKE PARTICIPANTS.

ID	KNN-DTW (original)	KNN-DTW PCA(q=3)	KNN-DTW PCA(q=6)
1	0.43	0.43	0.85
2	0.8	0.6	1
3	1	1	1
4	0.4	0.2	0.6
5	1	1	1

experiment was also conducted to classify an unseen gait sequence in a reduced dimensional feature vector using the PCA.

III. RESULTS AND DISCUSSION

The performance of kNN-DTW on the original (higher dimensional) feature space was compared against reduced dimensional feature space using the PCA. The dimensionality of the latent space (q) was set to 3 or 6. The first 3 and 6 principal components covered 84% and 99% of the overall variance across all 10 participants. The F1-scores obtained for the method with different values of parameter k in nearest neighbor and w in dynamic time warping are shown in Figure 2.

In Figure 2, PCA with q = 6 achieves higher F1-scores than PCA with q = 3 or when using kNN-DTW in the original (higher dimensional) feature space. The highest F1-score occurred at k = 1 and w = 0.25. This is in line with previous findings which suggested to use k = 1 for small number of samples (less than 200 samples) regardless of the level of noise [27].

DTW with w equal to 25% of the sequence length obtained higher F1-score which is consistent with findings of [20], in which constrained DTW was proposed to be more accurate than the standard DTW (w = 100% of the sequence length).

All gait sequences pertaining to healthy participants were classified correctly. Table II shows the classification accuracy for each post-stroke participant. Accuracy is defined as the percentage of gait cycles that were classified correctly. The parameter k and w were selected according to best performance in its range shown in Figure 2.

A closer look at the Table I and Table II indicate that kNN-DTW with PCA (q = 6) was least successful in classifying the gait patterns of participants who were not using any aid and walked independently. The gait pattern of this group of participants is closer to the gait pattern of healthy individuals and it is more likely for a classifier to misclassify these participants as healthy. kNN-DTW with original features and kNN-DTW with PCA (q = 3) also failed in classifying a post-stroke participant who was using rollator and supervised for outside walks, but kNN-DTW with PCA (q = 6), however, correctly labled all gait cycles of this participant.

The results shown in Figure 2 and Table II strongly support the use of dimensionality reduction. This is understandable since the number of participants was small (10) and classification in the original feature space (with D = 38) was prone to overfitting.

IV. CONCLUSION AND FUTURE WORK

In this paper, we demonstrated successful integration of an affordable and unobtrusive sensing technology and a machine learning approach to discriminate stroke individuals with pathological gait from healthy individuals. We recorded gait sequences in a hospital clinic and during routine exams, using two Kinect sensors. High dimensional gait features including the orientations of the hips and spine (abdomen), shoulders and neck (upper limb), knees and ankles (lower limb) were shown to contain sufficient information to enable the automatic detection of stroke gait. A simple instance based ML algorithm (combination of kNN and DTW) provided good classification accuracy, particularly when combined with PCA for dimensionality reduction. In future work, we plan to analyse gait data from a larger cohort of participants with varying pathological conditions (e.g. stroke, acquired brain injury, etc.) and to investigate a more sophisticated ML approach which can incorporate dynamics of the human body.

REFERENCES

[1] A. H. Snijders, B. P. Van De Warrenburg, N. Giladi, and B. R. Bloem, "Neurological gait disorders in elderly people: clinical approach and classification," *The Lancet Neurology*, vol. 6, no. 1, pp. 63–74, 2007.
[2] F. Coutts, "Gait analysis in the therapeutic environment," *Manual therapy*, vol. 4, no. 1, pp. 2–10, 1999.
[3] D. T. Lai, R. K. Begg, and M. Palaniswami, "Computational intelligence in gait research: a perspective on current applications and future challenges," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 5, pp. 687–702, 2009.
[4] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *Biomedical Engineering, IEEE Transactions on*, vol. 52, no. 5, pp. 828–838, 2005.
[5] W. Tao, T. Liu, R. Zheng, and H. Feng, "Gait analysis using wearable sensors," *Sensors*, vol. 12, no. 2, pp. 2255–2283, 2012.
[6] N. M. Tahir and H. H. Manap, "Parkinson disease gait classification based on machine learning approach," *Journal of Applied Sciences*, vol. 12, no. 2, p. 180, 2012.
[7] D. T. Lai, P. Levinger, R. K. Begg, W. L. Gilleard, and M. Palaniswami, "Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, no. 5, pp. 810–817, 2009.
[8] S. Hagler, D. Austin, T. L. Hayes, J. Kaye, and M. Pavel, "Unobtrusive and ubiquitous in-home monitoring: a methodology for continuous assessment of gait velocity in elders," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 4, pp. 813–820, 2010.
[9] E. Cippitelli, S. Gasparrini, S. Spinsante, and E. Gambi, "Kinect as a tool for gait analysis: validation of a real-time joint extraction algorithm working in side view," *Sensors*, vol. 15, no. 1, pp. 1417–1434, 2015.
[10] J.-H. Yoo and M. S. Nixon, "Automated markerless analysis of human gait motion for recognition and classification," *Etri Journal*, vol. 33, no. 2, pp. 259–266, 2011.
[11] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, "Full body gait analysis with kinect," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*. IEEE, 2012, pp. 1964–1967.
[12] M. Parajuli, D. Tran, W. Ma, and D. Sharma, "Senior health monitoring using kinect," in *Communications and Electronics (ICCE), 2012 Fourth International Conference on*. IEEE, 2012, pp. 309–312.
[13] J. A. Garcia, Y. Pisan, C. T. Tan, and K. F. Navarro, "Assessing the kinects capabilities to perform a time-based clinical test for fall risk assessment in older people," in *Entertainment Computing–ICEC 2014*. Springer, 2014, pp. 100–107.
[14] A. Leu, D. Ristić-Durrant, and A. Gräser, "A robust markerless vision-based human gait analysis system," in *Applied Computational Intelligence and Informatics (SACI), 2011 6th IEEE International Symposium on*. IEEE, 2011, pp. 415–420.
[15] M. Rantz, M. Skubic, C. Abbott, C. Galambos, M. Popescu, J. Keller, E. Stone, J. Back, S. J. Miller, and G. F. Petroski, "Automated in-home fall risk assessment and detection sensor system for elders," *The Gerontologist*, vol. 55, no. Suppl 1, pp. S78–S87, 2015.
[16] E. Stone, M. Skubic, M. Rantz, C. Abbott, and S. Miller, "Average in-home gait speed: Investigation of a new metric for mobility and fall risk assessment of elders," *Gait & posture*, vol. 41, no. 1, pp. 57–62, 2015.
[17] J. Han, L. Shao, D. Xu, and J. Shotton, "Enhanced computer vision with microsoft kinect sensor: A review," *Cybernetics, IEEE Transactions on*, vol. 43, no. 5, pp. 1318–1334, 2013.
[18] M. P. Murray, "Gait as a total pattern of movement: Including a bibliography on gait." *American Journal of Physical Medicine & Rehabilitation*, vol. 46, no. 1, pp. 290–333, 1967.
[19] G. Kienast, D. Bachmann, G. Steinwender, E.-B. Zwick, and V. Saraph, "Determination of gait patterns in children with cerebral palsy using cluster analysis," *Gait & Posture*, vol. 10, no. 1, p. 57, 1999.
[20] H. Ding, G. Trajcevski, P. Scheuermann, X. Wang, and E. Keogh, "Querying and mining of time series data: experimental comparison of representations and distance measures," *Proceedings of the VLDB Endowment*, vol. 1, no. 2, pp. 1542–1552, 2008.
[21] R. Muscillo, S. Conforto, M. Schmid, P. Caselli, and T. D'Alessio, "Classification of motor activities through derivative dynamic time warping applied on accelerometer data," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*. IEEE, 2007, pp. 4930–4933.
[22] L. R. Rabiner, A. E. Rosenberg, and S. E. Levinson, "Considerations in dynamic time warping algorithms for discrete word recognition," *The Journal of the Acoustical Society of America*, vol. 63, no. S1, pp. S79–S79, 1978.
[23] B.-H. F. Juang, "On the hidden markov model and dynamic time warping for speech recognitiona unified view," *AT&T Bell Laboratories Technical Journal*, vol. 63, no. 7, pp. 1213–1243, 1984.
[24] E. Dolatabadi, B. Taati, and A. Mihailidis, "Vision-based approach for long-term mobility monitoring: Single case study following total hip replacement," *Journal of Rehabilitation Research & Development*, vol. 51, no. 7, 2014.
[25] S. Tan, "Neighbor-weighted k-nearest neighbor for unbalanced text corpus," *Expert Systems with Applications*, vol. 28, no. 4, pp. 667–671, 2005.
[26] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *Systems, Man and Cybernetics, IEEE Transactions on*, no. 4, pp. 325–327, 1976.
[27] D. Wettschereck, "A study of distance-based machine learning algorithms," Ph.D. dissertation, 1994.