DeepMotion: A Deep Convolutional Neural Network on Inertial Body Sensors for Gait Assessment in Multiple Sclerosis*

Jiaqi Gong, *Member, IEEE*, Myla D. Goldman and John Lach, *Senior Member, IEEE*

Abstract— **Walking impairment resulted by various chronic diseases, disorders and injuries have been investigated using recent emerging wearable technology, for instance, gait assessment using inertial body sensors in 6-minute walk (6MW) for persons with Multiple Sclerosis (PwMS) to identify spatiotemporal features useful to assess MS progression. However, most studies to date have investigated the features extracted from movements of the lower limbs and do not provide a holistic gait assessment. A recent pilot study demonstrated that the holistic gait assessment such as evaluating the associations among lower and upper limbs provided better discrimination between healthy controls and PwMS. This paper is motivated by this and further aim to answer the following question: can we identify the temporal gait patterns in terms of the holistic gait assessment? Traditionally this suffers from the statistical property of the causality inference method adopted by previous study. We proposed a deep convolutional neural network (CNN) to learn the temporal and spectral associations among the time-series motion data captured by the inertial body sensors. A simulated model was developed to train the CNN, and then the trained CNN was adopted to assess the gait performance from a pilot dataset with 41 subjects (28 PwMS and 13 healthy controls). Experimental results are reported to illustrate the performance of the proposed approach.**

I. INTRODUCTION

Walking impairment is one of the most ubiquitous symptoms of various chronic diseases, such as Multiple Sclerosis (MS) [1], amyotrophic lateral sclerosis [2], Parkinson's disease [3], Alzheimer's disease [4] and so on, therefore, gait assessment regarding walking impairment contributes highly in tracking the disease progression, generates patient-centric outcomes and enhances the persons' quality of life. For instance, assessment of walking ability contributed heavily in clinical research of persons with MS (PwMS), especially in the expanded disability status scale (EDSS) which is one of the most commonly used clinical measures.

Previous research in gait assessment for quantifying walking impairment in PwMS includes ambulatory performance measures (walking speed and endurance) [6], physiological measures (energetic cost of walking) [7], and kinematic measures [8] (spatiotemporal features extracted from motion capture systems or body sensors). The popular ambulatory performance measures include the timed 25-foot walk (T25FW) and 6-minute walk (6MW); walking speed and distance traveled during the tests are extracted as an estimate of walking endurance. Physiology measures include the energy expenditure of the PwMS to calculate the energetic efficiency of walking. Motion capture system and inertial body sensors are recent emerging techniques for kinematic measures in gait assessment. Especially due to the convenience, comfort and affordable cost of the system, inertial body sensors are getting popular in clinical research for extracting more precise and objective measures in kinematics and/or the quality of life.

Despite the previous measures having been proven to be of clinical relevance, the impact of MS disease on walking impairment is not comprehensively understood. Especially most of the measures are extracted from the motion data of lower limbs, however, the upper limb dysfunctions that nearly 75% of PwMS experience [9]. Since the impact from MS disease could be reflected by the variability of movement throughout the body, including upper and lower limbs, holistic gait assessment to reveal the mechanism underlying the associations between limbs is needed.

Recent advances in causality modeling of inertial body sensors for quantifying the holistic gait performance, for instance, evaluating the strength of the associations among lower and upper limbs, demonstrated better discrimination between healthy controls and PwMS. However, the statistical characteristics of causality analysis limited the intuitive explanation of the results, especially cannot represent the temporal gait cycles as previous kinematic measures did. This paper is motivated by this and proposed a method based on a deep convolutional neural network (CNN) which is designed to learn the temporal and spectral associations among the inertial body sensor data. The proposed method was tested in a pilot dataset with 41 subjects (28 PwMS and 13 healthy controls). Compared with existing methods, the proposed method provided better discrimination between healthy controls and PwMS and an intuitive representation of holistic gait patterns. The main contributions of this paper are:

- Development of rationales for the deep CNN on multichannel inertial sensor data processing based on a holistic model of gait assessment;
- An approach incorporating the deep CNN to learn the holistic representation of the gait patterns in high-dimensional time-series data (i.e., 15-dimensional signal from 5 inertial body sensors);
- First attempt with strong evidence to show that the deep CNN learned the temporal and spectral associations

*This work was supported in part by the U.S. National Science Foundation under grant numbers IIS-1231712 and EEC-1160483.

Jiaqi Gong and John Lach are with Department of Electrical and Computer Engineering, UVA Center for Wireless Health, University of Virginia, Charlottesville, VA, USA 22904 (email: {jgong, jlach}@virginia.edu).

Myla D. Goldman is with Department of Neurology, UVA Center for Wireless Health, University of Virginia, Charlottesville, VA, USA 22904 (email: mdg3n@ virginia.edu).

among the time-series motion data and provided better discrimination between healthy controls and PwMS;

II. Related Work and Motivation

Since mobility assessment for tracking MS progression is a quantification task of motion quality [11] rather than a recognition task of motion types [12], here we investigate the previous work in two categories: mobility assessment techniques for MS study and recent emerging technique; deep learning for inertial sensor data processing.

A. *Mobility Assessment*

Previous research in mobility assessment has been conducted to assess the mobility impairment regarding lower and upper limbs. Particularly, motor assessment of limbs functions of PwMS using inertial body sensors is getting popular in recent years. Traditional clinical assessment instrumenting inertial body sensors demonstrated better performance to discriminate between healthy controls and PwMS in terms of lower and upper limbs functions. Recently, Carpinella et al. [9] adopted inertial body sensors in a traditional mobility assessment of upper limbs functions, named Action Research Arm Test (ARAT), to reveal subtle arm alterations not detectable from previous ARAT scores. Otherwise, numerous literatures [11, 12] have been published to develop walking assessment using inertial body sensors with many types of protocols, such as 6MW, T25FW and so on.

Previous signal processing methods for inertial body sensors decompose the sensor data into cycles or patterns and then further segments the cycles or patterns into phases, such as swing phase and stance phase for lower limbs, reaching phase and returning phase for upper limbs. Based on the temporal segmentation of the sensor data, feature, such as stride length, cadence or rhythm, individual time for each phase, joint angles, momentum existence of tremors, etc., are extracted from the sensor data [13, 14]. Then, extracted features may be combined or transformed using a particular function and are then passed to statistical tools to generate clinical measures [15].

Although these separate measures from inertial body sensors regarding lower or upper limbs have demonstrated better performance, the whole-body mobility, especially the strong associations between upper and lower limbs, which may be important factors underlying mobility impairment in PwMS, have not been revealed yet. Herman et al. [16] studied the muscle power relationship between upper and lower limbs in mobility-limited older adults and then concluded that the mechanisms underlying the associations between upper and lower limbs maybe important factors in elderly persons. This discovery motivated our recent research which conducted causality model to evaluate the strength of the interactions among the limbs and demonstrated strong evidence to prove the hypothesis that the strength of associations among limbs in healthy control subjects during 6MW is stronger than in MS subjects.

B. *Deep Learning for Inertial Data Processing*

Deep learning is a new name for large-scale multiple-layer backpropagation neural networks, which is revolutionizing many aspects of data processing, such as, image processing, computer vision, speech recognition, and natural language processing and so on [17]. The developing infrastructures of neural networks for deep learning are divided into two categories; unsupervised learning (deep Boltzmann machines and its varieties) and supervised learning (convolutional neural networks (CNN), multiple perception layers (MLP), and recurrent neural networks (RNN)).

Recent researches have been conducted to explore the performance of deep learning on motion data processing. Lane et al. [18] claimed the revolution of mobile sensing based on deep learning in terms of accelerometer data and audio data from smart phone, but lacked of detailed explanation how the algorithm worked for the different data format. Yang et al. [19] adopted CNN on multichannel time series captured by inertial body sensors for human activity recognition, and Rad et al. [20] applied CNN for high-dimensional data processing for stereotypical motor movement detection in terms of autism spectrum disorders. Otherwise, RNN with long short-term memory was developed for gesture and activity recognition using inertial body sensors [21].

C. *Motivation*

Although the deep learning performed better than state-of-the-art, the mismatch between the characteristics of networks' structures and the characteristics of the inertial signals is one of the major concerns. For instance, the structure of CNN is designed for modeling the spectral associations in the signal, which means it may not be suitable for directly processing the multichannel inertial sensor data as an image like previous studies conducted [18, 19, 20], because the associations between the inertial sensors across time and frequency domains cannot be well represented in a single image.

In order to solve the concern of mismatch between characteristics of networks' structures and the characteristics of the signals, this paper proposed to design the structure of the neural networks based on previous studies on associations between multichannel inertial sensor data. First, previous causality model had been studied to reveal the mechanism that the spectral associations between multichannel time-series data captured by inertial body sensors represent the strength of interactions among the limbs. Second, previous studies which focused on reversely decomposing the CNN [25] demonstrated that the structures of CNN are suitable for modeling the spectral associations. Therefore, designing CNN to learn the associations between multichannel inertial sensor data provides rationales to learn better structures of CNN for matching the characteristics of the signals. This is also the fundamental principle of this paper, and section III discusses it in detail.

III. Problem Statement

This section gives a formal description of the holistic gait assessment using inertial body sensors in terms of walking impairment, including the rationales of walking impairment in MS, problem formulation and description of the holistic model for gait assessment.

A. Rationales of Walking Impairment in MS

Walking impairment in chronic disease is a degeneration process along with physical inactivity and physiological deconditioning. The influence cycle between physical inactivity, physiological deconditioning and walking impairment had been discussed in previous studies on MS [22], which described that physical inactivity might lead to physiological deconditioning while physiological deconditioning might influence the onset and progression of walking impairment.

The degeneration process of walking impairment in MS can be explained in a way of probability; in the early stage of the walking impairment, the PwMS demonstrates the physiological deconditioning in low probability, but along with the progression of the walking impairment, the PwMS demonstrates physiological deconditioning in higher probability or eventually walking disabilities. Therefore, the physiological deconditioning during the progression of walking impairment is revealed by the inertial body sensors, such as the difficulty of swinging arms or legs during walk, and its probability is the focus of the data processing.

B. Problem Formulation

Based on previous description of the rationales of walking impairment, here we formulate the problem to infer the probability of physiological deconditioning which the PwMS demonstrate in the inertial sensor data captured during the walking tests (i.e. 6MW) in clinical research.

Data acquisition. Consider holistically the body motion, human walks by alternatively and repetitively swing their left and right lower/upper limbs. Since MS affects the movements of all the limbs, collecting data from lower/upper limbs is needed for inferring the physiological deconditioning. Therefore, the subjects participated in this research were asked to wear 5 inertial sensors on the left/right wrists, left/right ankles, and sacrum to undergo an in-clinic 6MW. The University of Virginia Institutional Review Board approved the study and data collection. All subjects gave written informed consent prior to assessments. Figure 2 (a) illustrates the locations of the inertial body sensors on the body. Subjects were asked to walk as far and as fast as possible (without running) back and forth a 75-foot hallway.

Equipment. The wearable motion analysis system consisted of five inertial body sensors, named Technology-Enabled Medical Precision Observation (TEMPO) version 3.2 with Bluetooth [23], each housing a 3-dimentional gyroscope and tri-axial accelerometer sampling at 128Hz, which is sufficient to capture the frequency band of the body motion in walking action. Gyroscopes measured rotational velocity in roll, pitch and yaw planes with a ±2000 degree per second range. Tri-axial accelerometers measured linear acceleration with a ±16 g range. The inertial sensor data was wirelessly transmitted to a laptop for post-processing. The operator of the data collection system is required to make timestamp annotations in order to indicate the beginning and end of the data collection.

Time-series segmentation. The aim of this study is to infer the probability of physiological deconditioning during 6MW. It means the inertial body sensor data in some time windows might be different with others. Therefore, in order to infer the probability, we segment the time-series data into window-sized data pieces. Previous studies on deep learning of inertial sensor data processing chose the window size as 1s, 2s, 4s or 10s, which seems there is no justified strategy for the selection of window size. However, since this study is to learn the associations among limbs during each gait cycles, ideally the window-sized data pieces at least should contain one gait cycle which includes alternatively swing their left and right lower/upper limbs. Based on this criterion, this study adopted the window size as 2s in order to include at least 1 gait cycles in the window. In addition, it is noteworthy that the objective of segmentation is to ensure each data piece has at least 1 gait cycle so that there are somewhat overlap between the segments, which is different with previous data processing methods based on deep learning.

Mathematical formulation. Consider the multiple inertial sensors take samples of the form

$$\mathrm{r_n} = \mathrm{r_n^*} + \mathrm{w_n}, \qquad \mathrm{n} = 1, 2, \dots \mathrm{N}\ (\mathrm{N} \approx \mathrm{sr} * 360) \quad (1)$$

Where $\mathrm{r_n} = [\mathrm{r_n^x}, \mathrm{r_n^y}, \mathrm{r_n^z}]$ is a 3D gyroscope data sample and composed of 3-axis angular velocity. $\mathrm{w_n} \in \mathbb{R}^3$ is a noise vector of independent zero-mean Gaussian random variables with variance σ_w^2 such that $\mathrm{w_n} \sim \mathcal{N}(0, \sigma_\mathrm{w}^2 \mathbb{I}_3)$. N is the number of total samples during 6MW, equals around sampling rate multiplying 360 seconds. The reason for selecting gyroscope data instead of accelerometer data is because of the noise (e.g. random spikes and other artifacts) in accelerometer data. Although previous studies adopted filtering methods to eliminate the noise in accelerometer data, it might rise up another issue that filtering methods may eliminate the subtle movements as well. Otherwise, previous research in causality modeling for evaluating the associations between multichannel inertial sensor data proved that gyroscope data worked for representing the associations among limbs, therefore, this paper continue the same setting as before.

The samples are segmented into 2-second time windows and each window-sized data piece includes 1 gait cycle. We adopted a robust segmentation method proposed in [24] with considerations of human factors during the real-world deployment. The window-sized data pieces is in the form

$$\mathrm{s_k} = [\mathrm{r_k} \ \dots \ \mathrm{r_{k+sr*2}}] \qquad \mathrm{k} = 1, 2, \dots \mathrm{M} \quad (2)$$

While $\mathrm{s_k}$ might present the physiological deconditioning, such as slow swing of arms or legs.

Consider the probability of physiological deconditioning (PD) of the i[th] PwMS as $\mathrm{p_i}$, and then the emergence of physiological deconditioning during 6MW can be represented as

$$\textstyle\sum_{\mathrm{s_k\ with\ PD}} 1 \ \propto\ \mathrm{p_i} * \mathrm{M} \qquad (3)$$

Therefore, the problem of inferring the probability of the physiological deconditioning for evaluating walking impairment is casted as quantifying the window-sized data pieces with physiological deconditioning or not.

C. Holistic Model for Gait Assessment

In order to quantify the window-sized data pieces segmented from the inertial sensor data captured during 6MW, a model is proposed here to holistically represent the physiological deconditioning. The motor symptoms of MS had been described in clinical research as that the emergence of physiological deconditioning might appear as the difficulty of moving arms or legs. This description can be explained in another way; the physiological deconditioning can be evaluated in the scale of the difficulty of moving arms or legs. If the difficulty of moving arms or legs were defined in 4-demensional space, then, a model can be designed in the space and each dimension represents the physiological deconditioning on left/right arm, left/right leg, respectively. The quantification of window-sized data pieces can be casted as mapping each data piece into the 4-dimensional space in terms of the physiological deconditioning on left/right arm, left/right leg. Each dimension is normalized, and the thresholding for the space is used to decide whether there is emerging physiological deconditioning in the data piece or not. Figure 1 illustrates the modeling process and presents two examples of the data pieces mapping into the space.

Based on this holistic model for gait assessment, the quantification of the window-sized data pieces is described in the form

$$y_k^i = f_{mapping}(s_k),\ y_k^i = [y_k^1, y_k^2, y_k^3, y_k^4] \tag{4}$$
$$y_k^1 \in (0,1), y_k^2 \in (0,1), y_k^3 \in (0,1), y_k^3 \in (0,1)$$
$$Y_k = \begin{cases} 0 & if\, y_k^{i=1,2,3,4} \le \alpha \\ 1 & Otherwise \end{cases}$$

Where $f_{mapping}$ represents the function or tool which is used to mapping the window-sized data piece into the 4D space, and y_k^i represents the mapping point in the 4D space where $y_k^{i=1,2,3,4}$ represents the physiological deconditioning on each limb. Y_k represents the thresholding result to determine whether there is physiological deconditioning in the specified data piece or not.

And then equation (3) can be modified as

$$\sum\ Y_k \propto p_i * M \tag{5}$$

It is noteworthy that since this study is the first attempt to explore better structures of deep learning for evaluating the associations among limbs, therefore, the objective of the algorithm is to infer the deconditioning probability rather than classify the disorder types in limbs.

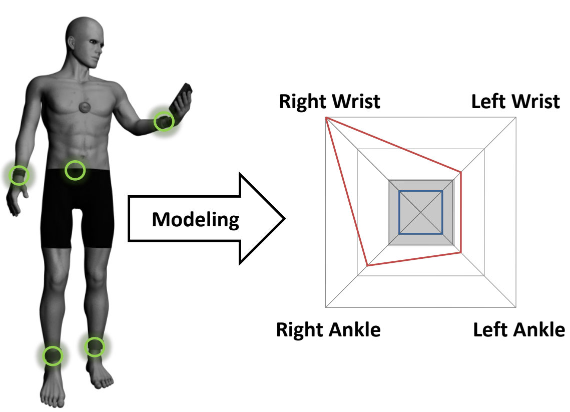

Figure 1. Modeling the physiological deconditioning in a 4-dimensional space. The grey area is used as threshold to decide where there is emerging physiological deconditioning or not, for instance, the red line represents a data piece with physiological deconditioning on right arm, while the blue line represents a data piece without physiological deconditioning.

IV. Methods

This section presents the deep CNN for inferring the probability of physiological deconditioning in each window-sized data pieces. As discussed as before, the deep CNN is designed to obtain the associations between multichannel sensor data across all time and frequency domains. In order to achieve that, the multichannel sensor data was preprocessing to generate multiple spectrograms. Then the structure of CNN was designed to extract the spectral and temporal associations between the sensor data.

A. *Multichannel Spectrograms*

A spectrogram of an inertial signal is a three dimensional representation of changes in the motion energy content of a signal as a function of frequency and time. Previously, spectrograms of speech waveforms are widely used as distinguishable features in acoustic modeling (e.g., the Mel-frequency cepstral). The multichannel spectrograms fully represent the frequency and time components so that representation learning from these spectrograms contains subtle information regarding the associations among multichannel sensor data.

Each data piece in 2-second window was being calculated to generate the multichannel short term Fourier transformation (STFT) spectrograms (15 channels in this study including 15D gyroscope data from 5 inertial sensors). Each data piece contained 256 data samples, and the size of each STFT spectrogram was 32×48, where 32 is the frequency resolution and 48 is the time resolution.

B. *Structure of CNN*

A typical convolutional network is shown in Figure 2. We use the multichannel spectrograms calculated from the multichannel inertial sensor data pieces as inputs to our network. $V = \Re^{t*f*D}$, where t and f are the input feature dimension in time and frequency respectively, and D is the number of signal channels. For instance, for color image recognition, D equals 3 (red, green, blue), while in this study, D equals 15 (15D spectrograms from 5 inertial sensors). As shown in Figure 2, our network is composed of convolutional layers, rectified linear units (ReLUs), max-pooling layers and fully connected layers.

Convolutional layers convolved the full input V using a weight matrix to generate feature maps. According to previous efforts in shining light into the “black box” of CNN, it is proved that convolutional layers extracted the preliminary associations among the multichannel data. After performing convolution, ReLUs and max-pooling layers help to remove variability in the time-frequency domain that exists due to noise. After the lower convolutional layers, the higher layers are always fully connected. This is the general architecture of CNN in previous studies. Although the CNN performed better than state-of-the-art in many domains, such as computer vision, speech recognition and natural language processing, but the design of the detail structure for different signal is always different. Yet there is no general principle for the structure design of CNN, which is one of the major challenges for deep learning technique.

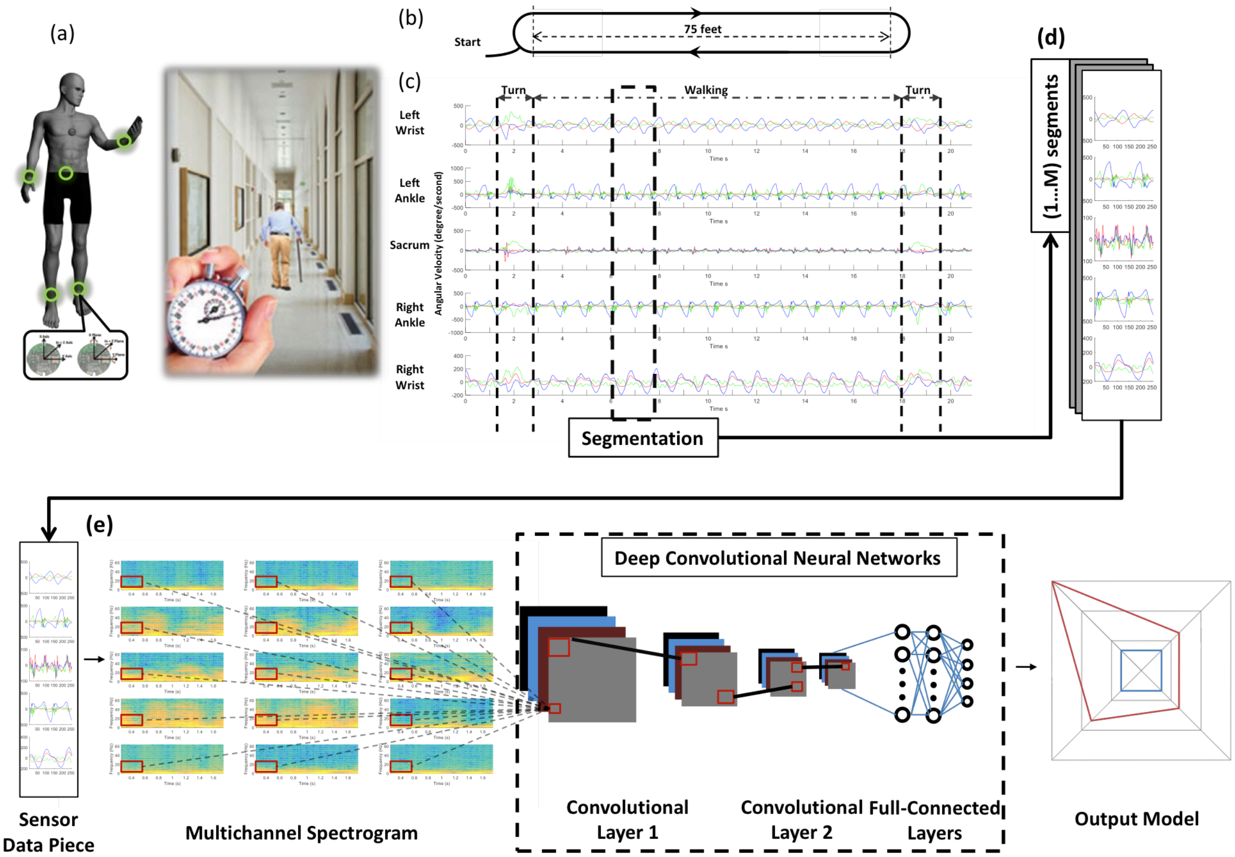

Figure 2: (a) A subject performing 6MW in the clinic hallway; (b) schematic of the task divided into walking and turning phases; (c) an example of the motion signal from the walking task showing the angular velocity of the inertial body sensors. (d) segmenting the walking period into 2-second windows and each window-sized data piece contains 2 gait cycles. (e) an illustration of the architecture of our CNN. The input to the network is a 2-second STFT spectrogram. Network consists of 2 convolutional layers, 2 fully connected layers and a output layer. Each convolutional layer has 64 filters, with rectified linear unit (ReLU) activations. Filter and pooling sizes are chosen to encourage learning of patterns across both temporal and spectral domain.

Thanks to recent advances in reverse engineering of CNN for representation learning from image recognition [25], the general guideline of structure design for multichannel spectrograms was revealed. Based on that, we designed the CNN for this study. Table I shows the hyperparameters which were tuned and selected on the training set. Otherwise, we discussed the tuning strategy in detail as below.

TABLE I. MODEL HYPERPARAMETERS

Hyperparameters	Values
#Convolutional Layers	{2,3}
#Convolutional hidden units	{32, 64, 128}
ReLUs at each convolutional layer	{1}
Max-pooling at each convolutional layer	{1,0}
#Fully-connected layer	{2,3}
#Fully-connected layer hidden units	{128}

Numbers of convolutional and fully connected layers: Determining better structures of CNN for image feature extraction is still a challenging task in most current CNN work. Recent studies in visualization of deep learning algorithm revealed that convolutional layers are able to model spectral associations with robustness regarding spectral variation, while fully connected layers extract high-level associations among the local information learned in the convolutional layers.

It is noteworthy that spectrograms are slightly different with images. The associations among pixels in images mostly span all the spatial domain, however, the associations across the spectrograms does span so wide in the time and frequency domains. Otherwise, regarding the training task in this study, it is better to train a small model specifically for gait patterns. Therefore, we adopted only a small number (2 or 3) of convolutional layers and fully connected layers in this study.

Number of hidden units: There is an ongoing argument in the research field about the number of hidden units. Since the goal of this paper is not to get involved with this argument, therefore, we tuned the hyperparameters in the training process and then chose 64 hidden units for convolutional layers and 128 for fully connected layers in less computation complexity without much performance loss.

Filter sizes: The filtering sizes in the convolutional layers should be related to the spectral associations in multichannel sensor data. Since there is no adequate prior knowledge about the spectral associations in inertial sensor data regarding walking, different sizes of filters were tuned during the training process, and then were chosen in, 9×7 (frequency-time) for the first convolutional layer and 6×5 for the second convolutional layer.

Pooling strategies: Table I shows that each convolutional layer can be tuned to have a max-pooling layers or not. The tuned result is that the first convolutional layer uses a pooling size of 2×2, and the second convolutional layer has no pooling layer.

Output: As mentioned before, this method is an immediate step to infer the probability of physiological deconditioning rather than to classify the disorder types. Therefore, according to equation (5), the probability is calculated as the ratio between the identified data pieces with physiological deconditioning and overall data pieces during the 6MW. Figure 2 illustrates the processing pipeline of the deep CNN for inertial sensor data in this study.

V. EXPERIMENTAL SETUP

A. Testing Dataset

41 subjects (28 PwMS and 13 healthy controls) participated in the testing data collection. EDSS score and other related measures for performance comparison, including 12-item Multiple Sclerosis Walking Scale (MSWS-12), Distance Traveled during 6MW, and Ratio between Double Stance Time and Single Stance Time (DST/SST), were collected, and EDSS scores were used as gold standard to diagnose the disease states of PwMS. For more information about the measures and correlation studies, please read reference [8]. The data collection protocol during the 6MW exactly followed the description in reference [1].

Each subject participated at least once 6MW in the data collection period, and 132 data sessions were collected. The calibration of inertial sensors had been done before the data collection [14]. There is no general normalization in the data preprocessing. However, due to the technical issues of our custom data collection system and human factors in the real-world deployment, 11 data session failed in the calibration process. Finally, 121 data sessions were included in the testing dataset; 36 data sessions were collected from 13 healthy controls while 85 data sessions were collected from 28 MS subjects, successfully.

Accelerometer data is susceptible to random spikes and other artifacts, which may cause inaccurate causality inferences. In contrast, short-term calibrated gyroscope data has better stationary properties. Therefore, calibrated gyroscope data was used for the causality analysis, with each data session containing 15 dimensional gyroscope data – 3 planes from each of the 5 inertial body sensors.

All the data sessions are calibrated with the recorded calibration parameters that had been determined before the data collection [26]. There is no general normalization or signal filtering in the data preprocessing. Therefore, the frequency of the data is up to 64Hz, which is half of the sampling rate (128Hz) according to Nyquist theory.

After segmenting the available 121 data sessions into 2-second windows, we have 17642 testing data sessions. Average number of the data pieces for each data session is around 145.

B. Training Dataset

In order to train the proposed CNN with clear ground truth that can exactly identify the window-sized data piece w/o physiological deconditioning, we collected another training dataset.

8 healthy controls participated in the training data collection. Each subject was asked to perform 6MW in 5 types of conditions; normal 6MW, 6MW with 3-lbs right wrist weight, 6MW with 3-lbs left wrist weight, 6MW with 3-lbs right ankle weight, and 6MW with 3-lbs left ankle weight. The weights on the wrists and ankles were used to affect the gait patterns of the healthy controls. Although the effect is not the same as the MS attack on the PwMS, the learning process of CNN might be able to extract similar information regarding the associations among the lower and upper limbs. The experimental results in next section demonstrated the efficacy of this assumption.

After segmenting the available 40 data sessions into 2-second windows, we have 6415 training data sessions. Average number of the data pieces for each data session is around 160.

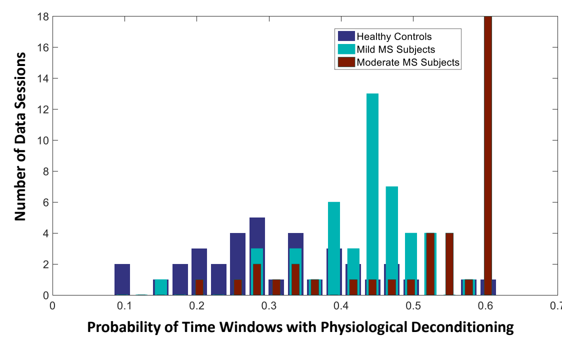

Figure 3: Experimental results of probability with physiplogical deconditioning caculated from 121 data sessions. The blue bars indicate the 36 data sessions are collected from healthy controls, while green bars indicate the data sessions collected from persons with mild MS, and the read bars indicate the data sessions collected from persons with moderate MS.

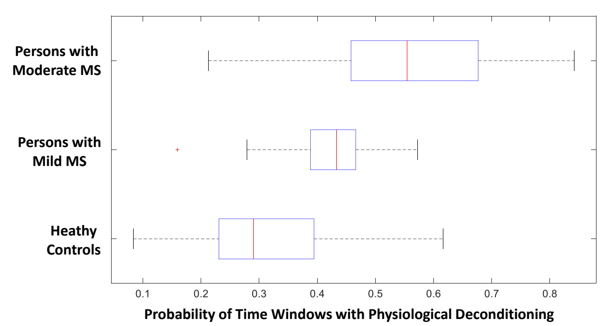

Figure 4: Comparision of probability with physiplogical deconditioning between healthy control and MS subjects.

VI. EXPERIMENTAL RESULTS

Figure 3 shows the probability calculated from the ratio between the identified 2-second data pieces with physiological deconditioning and overall data pieces of each 6MW data session. The blue bars indicate the 36 data sessions are collected from healthy controls, while green bars indicate the data sessions collected from persons with mild MS, and the read bars indicate the data sessions collected from persons with moderate MS. Apparently, the healthy control subjects have lower probability with physiological deconditioning than persons with MS.

Figure 4 illustrates the comparison of the probability with physiplogical deconditioning from the three groups: healthy control, persons with mild MS, and persons with moderate MS. The average value of the distribution in healthy control group is lower than in MS groups.

We adopt two statistical evaluation methods, Cohen-D (Effect Size) [27] and t-test (p value), to compare the performance of different features in separability of the three groups in the testing dataset. The groups of healthy controls, persons with mild MS, and persons with moderate MS were distinguished using the measures extracted from 6MW. Figure 5 shows the performance comparison of five measures in separability between healthy gait and MS-affected gait: effect size (p value) of MSWS-12, distance walked, (DST/SST), causality index and proposed method are 0.62 ($p<0.01$), 0.74 ($p<0.01$), 0.96 ($p<0.01$), 1.12 ($p<0.0001$) and 1.20 ($p<0.0001$), respectively. The performance of the proposed method is comparably better than recent causality-based method.

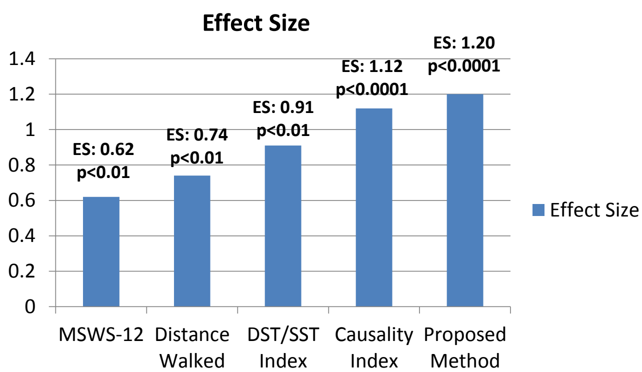

Figure 5: Performance comparison of different features in separability between three groups; healthy controls, persons with mild MS, and persons with moderate MS in the testing dataset.

VII. Discussion

A previous study on associations between lower and upper limbs has been reported that mechanism underlying the holistic measurement of body motion contains much richer information in terms of walking impairment in MS. However, the limitation of previous study is lacking of intuitive meaning of the causality model. This study is motivated by the limitation with the goal of enhancing the utility of inertial body sensors for assessing walking impairment of early-stage PwMS and tracking the disease progression which was traditionally measure by EDSS scale. In order to reveal the temporal gait patterns with physiological deconditioning, this study reformed the gait assessment problem as a probability inference and adopted a supervised method to learn the representation of the gait patterns regarding physiological deconditioning.

The proposed deep CNN was trained by the training dataset and then was used to evaluate the testing dataset. With tuning the structure of CNN, the performance of the proposed method performed comparably better in discrimination between healthy controls and PwMS than previous studies. The reason for this is that the deep CNN method somewhat extracted the spectral and temporal associations between multichannel signals which was captured by the inertial body sensors on lower and upper limbs, and the extracted information underlying the holistic model for gait assessment performed well in distinguishing the healthy control from the PwMS.

Limitations of this study are the relatively small sample size, the mismatch in age and gender in the participated populations, the fact that training dataset might not include enough information to discriminate the motor symptoms which was resulted by the MS attacks. Although previous studies demonstrated strong evidence to prove the fact that the associations among limbs may include richer information of the mechanisms underlying the walking impairment in MS, but it is still a challenging task to identify the effect specifically caused by MS attacks. The training dataset for the supervised learning method has small number of participants and the age and gender of them were not well matched to the participants in testing dataset. This mismatch and small sample size might be one of major limitations of this study.

Otherwise, although walking impairment in MS is one of the most ubiquitous symptoms, but the discrimination accuracy based on walking impairment should have a limitation as well, since the cause of MS is still unknown and the walking impairment only belongs to one of superficial symptoms in MS. Therefore, this might be also a reason for the slightly improvement in the discrimination performance of the proposed method. Furthermore, the substantial overlap between healthy controls and individuals with mild and moderate MS explained the heterogeneity of MS symptoms and presentations. We looked into the outlier of the healthy control subject in Figure 3, and found out in the clinic records this healthy control subject got a surgery between the two moments of 6MW data collection. This observation inspired another research direction in correlated data mining from electronic health records and wearable sensor data.

VIII. Conclusion

In summary, deep CNN for inertial body sensors provides another angle of view into gait assessment, enhances the performance of discriminating PwMS from healthy controls, and has the potential to track MS progression in early stage. The proposed holistic model of gait assessment which represents the associations between lower and upper limbs during 6MW provides advantages in intuitive interpretation of emergence of walking impairment in MS.

Future work focuses on reverse decomposition of the deep CNN to retrieval the information of the associations among limbs, and intuitive interpretation for clinicians, modeling of correlated data or symptoms which could be directed at individuals with disease progression associated with primarily gait disturbance such as Parkinson's or musculoskeletal problems involving the lower extremities.

References

[1] M. D. Goldman, R. A. Marrie, and J. A. Cohen, "Evaluation of The Six-Minute Walk in Multiple Sclerosis Subjects and Healthy Controls", Multiple Sclerosis, vol. 14, no. 3, pp. 383-390, 2007.

[2] S. B. Rutkove, James B. Caress, Michael S. Cartwright, Ted M. Burns, Judy Warder, William S. David, and Namita Goyal, "Electrical Impedance Myography Correlates with Standard Measures of ALS Severity" Muscle & Nerve, vol. 49, no. 3, pp. 441-443, 2014.

[3] L. Palmerini, Sabato Mellone, Guido Avanzolini, Franco Valzania, and Lorenzo Chiari, "Quantification of Motor Impairment in Parkinson's Disease Using An Instrumented Timed Up and Go Test", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 21, no. 4, pp. 664-673, 2013.

[4] S. W. Muir, Mark Speechley, Jennie Wells, Michael Borrie, Karen Gopaul, and Manuel Montero-Odasso, "Gait Assessment In Mild Cognitive Impairment And Alzheimer's Disease: The Effect Of Dual-Task Challenges Across The Cognitive Spectrum", Gait & Posture, vol. 35, no. 1, pp. 96-100 2012.

[5] J. F. Kurtzke, "Rating Neurologic Impairment in Multiple Sclerosis an Expanded Disability Status Scale (EDSS)", Neurology, vol. 33, no. 11, pp. 1441-1444, 1983.

[6] R. I. Spain, RJ St George, A. Salarian, M. Mancini, J. M. Wagner, F. B. Horak, and Dennis Bourdette, "Body-Worn Motion Sensors Detect Balance And Gait Deficits In People With Multiple Sclerosis Who Have Normal Walking Speed", Gait & Posture, vol. 35, no. 4, pp. 573-578, 2012.

[7] B. M. Sandroff, Rachel E. Klaren, Lara A. Pilutti, and Robert W. Motl, "Oxygen Cost Of Walking In Persons With Multiple Sclerosis: Disability Matters, But Why?", Multiple Sclerosis International, 2014.

[8] J. Gong, Matthew M. Engelhard, Myla D. Goldman, and John Lach, "Correlations Between Inertial Body Sensor Measures And Clinical Measures In Multiple Sclerosis", In Proceedings of the 10th EAI International Conference on Body Area Networks, pp. 18-24, 2015.

[9] I. Carpinella, D. Cattaneo, and M. Ferrarin, "Quantitative Assessment of Upper Limb Motor Function in Multiple Sclerosis using An Instrumented Action Research Arm Test", Journal of Neuroengineering and Rehabilitation, vol. 11, no. 1, pp: 1, 2014.

[10] Gong, Jiaqi, John Lach, Yanjun Qi, and Myla D. Goldman. "Causal analysis of inertial body sensors for enhancing gait assessment separability towards multiple sclerosis diagnosis." In Wearable and Implantable Body Sensor Networks (BSN), 2015 IEEE 12th International Conference on, pp. 1-6. IEEE, 2015.

[11] R. I. Spain, M. Mancini, F. B. Horak, and D. Bourdette, "Body-Worn Sensors Capture Variability, But Not Decline, of Gait and Balance Measures in Multiple Sclerosis Over 18 Months", Gait & Posture, vol. 39, no. 3, pp: 958-964, 2014.

[12] B. R. Greene, S. Rutledge, I. McGurgan, C. McGuigan, K. O'Connell, B. Caulfield, and N. Tubridy, "Assessment and Classification of Early-Stage Multiple Sclerosis With Inertial Sensors: Comparison Against Clinical Measures of Disease State", IEEE Journal of Biomedical and Health Informatics, vol.19, no. 4, pp: 1356-1361, 2015.

[13] W. Tao, T. Liu, R. Zheng, H. Feng, "Gait Analysis using Wearable Sensors", Sensors, vol.12, no.2, pp:2255-83, 2012.

[14] A. Muro-de-la-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait Analysis Methods: An Overview of Wearable and Non-Wearable Systems, Highlighting Clinical Applications", Sensors vol. 14, no. 2, pp: 3362-3394, 2014.

[15] V. Cimolin, and M. Galli, "Summary Measures for Clinical Gait Analysis: A Literature Review", Gait & Posture vol. 39, no. 4 pp: 1005-1010, 2014.

[16] S. Herman, D. K. Kiely, S. Leveille, E. O'Neill, S. Cyberey, and J. F. Bean, "Upper and Lower Limb Muscle Power Relationships in Mobility-Limited Older Adults", The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, vol. 60, no. 4, pp: 476-480, 2005

[17] Y. LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436-444, 2015.

[18] N. D. Lane, and Petko Georgiev, "Can Deep Learning Revolutionize Mobile Sensing?", In Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications, pp. 117-122. ACM, 2015.

[19] J. B, Yang, Minh Nhut Nguyen, Phyo Phyo San, Xiao Li Li, and Shonali Krishnaswamy, "Deep Convolutional Neural Networks On Multichannel Time Series For Human Activity Recognition", In Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI), Buenos Aires, Argentina, pp. 25-31. 2015.

[20] N. M. Rad, Andrea Bizzego, Seyed Mostafa Kia, Giuseppe Jurman, Paola Venuti, and Cesare Furlanello, "Convolutional Neural Network for Stereotypical Motor Movement Detection in Autism", arXiv preprint arXiv:1511.01865, 2015.

[21] G. Lefebvre, Samuel Berlemont, Franck Mamalet, and Christophe Garcia, "Inertial Gesture Recognition with BLSTM-RNN", In Artificial Neural Networks, pp. 393-410. Springer International Publishing, 2015.

[22] R. W. Motl, Myla D. Goldman, and R. H. Benedict. "Walking Impairment in Patients With Multiple Sclerosis: Exercise Training As A Treatment Option", Neuropsychiatr Disease Treatment, vol. 6, pp. 767-774, 2010.

[23] A. T. Barth, M. A. Hanson, H. C. Powell Jr, and J. Lach, "TEMPO 3.1: A Body Area Sensor Network Platform For Continuous Movement Assessment", In Procedings of Sixth International Workshop on Wearable and Implantable Body Sensor Networks (BSN) pp. 71-76. IEEE, 2009.

[24] J. Gong, P. Asare, J. Lach, and Y. Qi, "Piecewise Linear Dynamical Model for Actions Clustering from Inertial Body Sensors with Considerations of Human Factors", BodyNets: 9th International Conference on Body Area Networks, London, UK, Sep. 2014

[25] A. Mahendran, and Andrea Vedaldi, "Understanding Deep Image Representations By Inverting Them", In Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on, pp. 5188-5196. IEEE, 2015.

[26] S. Chen, J. S. Brantley, T. Kim, and J. Lach, "Characterizing And Minimizing Synchronization and Calibration Errors in Inertial Body Sensor Networks", 5th International Conference on Body Area Networks, pp. 138-144, 2010.

[27] H. Hentschke, and M. C. Stüttgen, "Computation of Measures of Effect Size for Neuroscience Data Sets", European Journal of Neuroscience, vol. 34, no. 12, pp. 1887-1894, 2011.