Learning Motion Primitives for the Quantification and Diagnosis of Mobility Deficits

Fujian Yan, Jiaqi Gong, Qiang Zhang, and Hongsheng He

***Abstract*—The severity of mobility deficits is one of the most critical parameters in the diagnosis of Parkinson's disease (PD) and rehabilitation. The current approach for severity evaluation is clinical scaling that relies on a clinician's subjective observations and experience, and the observation in laboratories or clinics may not suffice to reflect the severity of motion deficits as compared to daily living activities. This paper presents an approach to modeling and quantifying the severity of mobility deficits from motion data by using nonintrusive wearable physio-biological sensors. The approach provides a user-specific metric that measures mobility deficits in terms of the quantities of motion primitives that are learned from motion tracking data. The proposed method achieved 99.84% prediction accuracy on laboratory data and 93.95% prediction accuracy on clinical data. This approach presents the potential to supplant traditional observation-based clinical scaling, providing an avenue for real-time feedback to fortify positive progression throughout the course of rehabilitation.**

***Index Terms*—Machine learning, motion tracking, wearable sensors.**

I. Introduction

PARKINSON'S disease (PD) is a neurodegenerative disorder that causes mobility deficits, including tremor, bradykinesia, limb rigidity, and gait-balance problems. This mobility deficit greatly limits people' s ability to maintain independent living and severely depresses quality of life. Nearly 1.2 million in the United States will be living with PD by 2030, which is more than the combined number of people diagnosed with multiple sclerosis, muscular dystrophy, and Lou Gehrig's disease [1]. More than 10 million people worldwide are living with PD. PD is common for senior citizens, and an estimated 96% of patients are aged 50 and over. The requirement of healthcare services and rehabilitation facilities is tremendous for patients to recover independent mobility capabilities. The combined direct and indirect cost of PD, including treatment, social security payments, and lost income, is estimated to be nearly $25 billion per year in the United States alone.

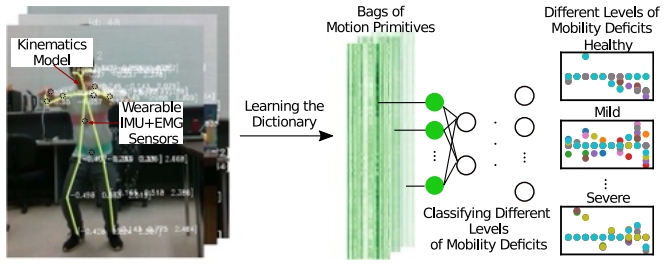

Fig. 1. Quantification of mobility deficits by using wearable sensors. Raw sensory data are collected by IMU and EMG sensors. Different levels of motion deficits are predicted.

The severity of mobility deficits is the most critical parameter in the diagnosis and rehabilitation of PD. The current approach for severity evaluation is clinical scaling when a potential patient performs predefined movements in a constrained environment. The Unified Parkinson's Disease Rating Scale (UPDRS) has been a clinical standard in the study and diagnosis of PD [2]. These scales, however, rely on a clinician's subjective observations and experience. The first and second parts of the UPDRS use questionnaires to evaluate the patient's status [3]. Moreover, the observation in laboratories or clinics may not suffice to reflect the severity of motion deficits compared to daily living activities [4]. Accurate nonintrusive quantification methods and evaluation metrics are required for diagnosing PD and timely feedback on rehabilitation procedures. Therefore, there is a great and vital need for an effective approach that reliably tracks body motion and quantifies mobility deficits.

In this paper, we quantify the severity of mobility deficits from daily activities by analyzing motion data measured by nonintrusive wearable physio-biological sensors. The framework is shown in Fig. 1, where wearable sensors are placed on limbs and the torso. Signature body movements are captured and analyzed. We developed a calibration-free motion tracking system and designed a multi-block neural-network model to discover unique motion primitives from time series. It is hypothesized that the motion primitives (small segments of motion patterns) in the motion tracking data significantly correlate with the severity of mobility deficits. The hypothesis was tested and the effectiveness of the motion primitives was evaluated in terms of the correlation with physician evaluations.

Manuscript received 9 September 2023; revised 1 April 2024 and 27 April 2024; accepted 10 May 2024. Date of publication 22 May 2024; date of current version 22 November 2024. This work was supported by NSF under Grant 2327313 and Grant 2427895. *(Corresponding author: Hongsheng He.)*

Fujian Yan is with the School of Computing, Wichita State University, USA.

Hongsheng He is with the Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487 USA (e-mail: hongsheng.he@ua.edu).

Jiaqi Gong is with the Department of Computer Science, The University of Alabama, USA.

Qiang Zhang is with the Department of Mechanical Engineering, The University of Alabama, USA.

Digital Object Identifier 10.1109/TBME.2024.3404357

Quantification of mobility deficits for the diagnosis of PD has attracted many researchers' interests. Studies have been performed to quantify PD in terms of finger-tapping severity by extracting and synthesizing finger motion properties [5]. The typical work to detect and quantify tremor includes the detection and assessment of resting tremor [6], [7], discrimination and quantification of resting/action tremor [8], [9], [10], and estimation of tremor severity [11]. Recent research has promoted motion tracking during non-standardized activities of daily living tasks [10], [12]. Diagnosis based on long-term monitoring is definitely valuable to understand the impact on mobility, and it is critical for home-based rehabilitation [13]. After-treatment rehabilitation is essential for PD patients to recover their motor functions. It is a dynamic process to correct undesired motions by facilities and experienced clinicians. A better way to evaluate the rehabilitation is to continuously monitor patients' movements, identify, and rectify problematic motion patterns at an early time. The state-of-art methods, however, could not monitor body motions in a free-living environment throughout the day, due to complex calibration procedures and drifting in long-term tracking.

Wearable motion-tracking systems are built on robotic platforms, visual tracking and wearable inertial sensors. In robotic-platform-based motion tracking, exoskeletons are installed to the participants' limbs to track their movements. Despite high accuracy, the cost and cumbersome procedures hinder the application into ambulatory environments. Visual tracking system can achieve highly accurate 3D human models by detecting and reconstructing reflective markers attached to limbs from different cameras, but they are expensive and usually complex to use. Motion tracking using wearable inertial sensors is convenient, low-cost, and accurate for short-term tracking. Traditional motion tracking systems with inertial sensors focus on accurately estimating a segment's orientation and position, such as the estimation of the body segment's orientation and position by combining the acceleration and gyroscope [14] and estimation of limb's orientation [15] using gyroscopes, accelerometers and magnetometers. Biomechanics constraints and optimization technologies were also utilized to achieve body segment's orientation with high accuracy [16]. The analysis of knee joint kinematics with inertial sensors has been fully addressed and explored, such as the knee's flexion and extension rotation with two Inertial Measurement Unit (IMU) sensors mounted on the thigh and shank [17], and knee joint's rotation by aligning two inertial sensors on adjacent segments [18]. Though the tracking of a joint angle has been thoroughly studied, full body tracking with flexible placement of sensors for long-term data recording is still challenging. Soft-tissue artifacts are a main source of errors either when the wearable sensors are mounted on a garment or directly attached to skin, and multiple sensor measurements have been integrated to increase the reliability of sensor placement [19]. One challenge in wearable body motion tracking is to remediate soft-tissue artifacts while reducing the complexity of sensor calibration.

To address the challenging problems, the system of body motion tracking is developed to simplify calibration and alleviate soft-tissue artifacts. From the data collected by the motion tracking system, a deep model of dictionary learning is designed to discover and associate motion primitives with mobility deficits. Specific contributions of the paper include:

1) The paper proves the existence of signature motion primitives that are associated with mobility deficits in motion tracking data;
2) The paper presents an effective approach to assisting diagnosis of diseases that may cause mobility deficits;
3) The paper designs a calibration-free motion tracking method and a novel deep learning model for mobility analysis.

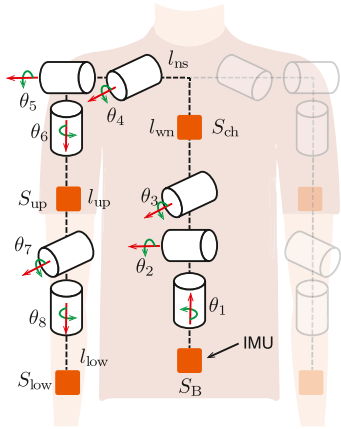

Fig. 2. Human kinematic model and placement of wearable physical-biological sensors.

II. Wearable Motion Sensing

Electromyography (EMG) and Inertial Measurement Unit (IMU) sensors are complementary in motion analysis, where IMU tells isometric or dynamic contractions, and EMG measures whole biomechanics. Applying EMG and its combination with inertial sensors for PD diagnosis can provide an objective assessment of tremors through muscular fluctuations [7]. In the early stages of PD tremor, the fluctuation of body segments is mild, which is hardly noticed visually from acceleration or angular velocity data, whereas EMG can detect such mild fluctuations from the muscular signal. Consequently, the fusion of EMG and IMU sensors would help early-stage detection and severity assessment. Motion tracking using wearable inertial sensors has been well explored and studied [20], and commercial products are available, such as Delsys, Nexonar, and Notch systems. In this section, we focus on the modeling of IMU sensors in relative configurations for reliable tracking.

A. Human Body Model

A total of eight sensors are placed on the limbs and trunk of the human body to track joint angles, acceleration of body parts, and electromyograph. The placement of the wearable sensors is illustrated in Fig. 2. The structure of human bodies is symmetric for the limbs, so we use the upper limb for modeling, and the other limbs are obtained by analogy due to similar

anatomical structures. The human upper limb consists of the upper arm, forearm, and hand, which has ten degrees of freedom (DoFs) in total. The shoulder is regarded as a ball-and-socket joint with three DoFs. The rotation angles are composed of flexion-extension, internal-external and abduction-adduction rotations. The elbow joint is regarded as two hinge joints with two DoFs: flexion-extension and pronation-supination rotations. The orientation of spinal movements would be considered as some rotations in waist joints with three DoFs.

As shown in Fig. 2, the upper limb includes four rigid bodies, i.e., chest, upper arm, lower arm, and pelvis. It has eight DoFs in total, which are respectively represented by θ_i for $i = \{1, 2, \ldots, 8\}$. The terms l_{wn}, l_{ns}, l_{up} and l_{low} denote each segment's length. Body frame B is assigned to the pelvis' center as a global reference frame, and its direction is defined to be aligned with sagittal and coronal planes.

Four IMU sensors S_{ch}, S_{up}, S_{low} and S_B sharing the same global frame determined by the gravity and earth magnetic fields are mounted on body segments to monitor their orientations. The term S_{ch} indicates the sensor that is attached to the chest, and the term S_{B} indicates the sensor that is attached to the pelvis's center. The terms S_{up} and S_{low} denote the sensors that are attached to the upper arm and lower arm. These IMU sensors can record angular acceleration and rotation angles that can be used to compute the transformation from body frames to the global frame. This paper used two IMU sensors for S_{low}: one is at the front arm (brachioradialis), and the other is at hand (abductor pollicis brevis). The global frame is a stationary reference frame, which serves as an absolute reference for spatial measurements, and the body frame is a reference frame located at a predefined point on the body. The IMU sensor and its corresponding body segment are regarded as one rigid body whose orientation is represented by quaternions. The equivalent rotation angle to a quaternion is $\theta = 2 \times \arccos(q_0)$, and the unit axis vector is $\omega = [q_1, q_2, q_3]^T / \sin(\theta/2)$ when $\theta \neq 0$. As frame B is the global reference frame, all the representations of motions in the global frame need to be converted into the body frame. The initial configuration of the body frame is determined when one stands straight and holds their arms vertically to the ground and then performs shoulder joint's pronation-supination rotation. The axis of rotation relative to the initial posture will be set as the y axis of frame B. The opposite direction of gravity is selected as its z axis. After achieving the initial Body frame $[X_b^g, Y_b^g, Z_b^g]^T$, the initial rotation matrix of global frame R_{g0}^b relative to body frame is calculated by $R_{g0}^b[X_b^g, Y_b^g, Z_b^g]^T = I$. Representing the rotation matrix R_{g0}^b as a quaternion q_{g0}^b, the relative quaternion between the body frame and its corresponding sensor S_B is

$$q_{S_B}^b = q_{g0}^b \otimes q_{s_B}^g \tag{1}$$

where $q_{S_B}^b$ and $q_{S_B}^g$ respectively denotes sensor S_B's quaternion with respect to body frame and global frame, and $\otimes$ denotes Hamilton product. Quaternion $q_{S_B}^b$ updates the body frame's orientation q_g^b corresponding to human body motion. The term q_{g0}^b represents the initial orientation of the body frame. The difference of the quaternion is computed as $\dot{q}_g^b = q_{g0}^b \otimes \omega$. The next timestep's quaternion is computed as $q_g^b = q_{g0}^b + \dot{q}_g^b$. Since the global frame of the sensor S_B is the same as its body frame, all other sensors that are attached to other body segments can use the S_B's quaternion q_g^b. The sensor's reference frame can be transferred to the body frame by $q_*^b = q_g^b \otimes q_*^g$, where $*$ represents anybody segments.

Contemporary IMU sensors are capable of automatic calibration, but they cannot address the problem of sensor drift and accumulated biases over a long period of use. The proposed motion-tracking method uses relative values from sensor readings instead of absolute values, i.e., relative joint angles instead of absolute angular velocities of each body segment. As such, the approach does not require calibration after initialization unless sensor placement changes.

B. Kinematics

At initial positions, all body segments are assumed to be aligned with the direction of gravity during the initial posture. As such, the initial positions of the joints p_{s0}, p_{e0} and p_{w0} with reference to the body frame are $[0, -l_{\text{ns}}, l_{\text{wn}}]^T$, $[0, -l_{\text{ns}}, (l_{\text{wn}} - l_{\text{up}})]^T$ and $[0, -l_{\text{ns}}, (l_{\text{wn}} - l_{\text{up}} - l_{\text{low}})]^T$. Let q_{ch0}^b, q_{up0}^b and q_{low0}^b denote the initial orientations of body segments, i.e., the initial configuration of forward kinematics. The change in sensor's orientation ${}^r q_*^b$ with respect to the initial configuration is ${}^r q_*^b = q_*^b \otimes ({}^r q_*^b)^{-1}$, from which the unit rotation axis ω_*^r and angle θ_*^r are computed.

By representing angular velocity ω and linear velocity $v \in \mathbb{R}^3$ by a screw axis $T = [\omega, v]$, the pose (including position and orientation) of a part of the human body is determined by the forward kinematics of all consecutive joints linking the part and the global reference

$$p' = \prod_{i=1}^{N} e^{[T_i]\theta_i} p_0 \tag{2}$$

where N is the number of consecutive joints, $[T] \in \mathbb{R}^{4\times4}$ is the matrix representation for the screw axis T, and θ_i is the angle of rotation around the screw axis. Exponential mapping is employed to delineate rotations within three-dimensional space through exponential coordinates. This approach entails a transformation from the rotation matrices to skew-symmetric matrices, which encapsulate infinitesimal rotations. We used $e^{[T_i]\theta_i}$ to represent the matrix exponents, and it can represent poses and incremental motions of the body segment's end-effector using homogeneous transformation matrices. Leveraging the properties of matrix exponential, exponential mapping provides a unified framework for describing both rotational and transnational components of motion. We developed the kinematic model on the anatomical characteristics of the human upper limb, incorporating the actual degrees of freedom (DoF) present in each segment. Accordingly, we devised the shoulder model with three interconnected joints, the upper arm with six such joints, and the lower arm with eight joints. While more joints increase computational demand, they also enhance the model's representability. The proposed kinematic model emphasizes the representation of DoFs for each joint, particularly in the shoulder case. This aligns with the objective of our study, which identifies

mobility deficits through the analysis of motion primitives. As such, the forward kinematics for the orientation of the shoulder joint p_s is achieved by

$$e^{[T_{\text{ch}}]\theta_{\text{ch}}} = \prod_{i=1}^{3} e^{[T_i]\theta_i} \tag{3}$$

where $q_{\text{ch}} \Rightarrow \{\theta_{\text{ch}}, \omega_{\text{ch}}\}$. Since their rotation axes intersect at the origin of the body frame, there would be no translation of the trunk's movements. The orientation of the upper arm is the product of the rotation of the waist and shoulder joints

$$e^{[T_{\text{up}}]\theta_{\text{up}}} = \prod_{i=1}^{6} e^{[T_i]\theta_i} \tag{4}$$

where $q_{\text{up}} \Rightarrow \{\theta_{\text{up}}, \omega_{\text{up}}\}$. Following the same idea, the orientation of the lower arm can be achieved by combining the connecting joints

$$e^{[T_{\text{low}}]\theta_{\text{low}}} = \prod_{i=1}^{8} e^{[T_i]\theta_i} \tag{5}$$

where $q_{\text{low}} \Rightarrow \{\theta_{\text{low}}, \omega_{\text{low}}\}$.

C. Recovering Joint Angles and Limb Dynamics

The angles $q \Rightarrow \{\theta, \omega\}$ corresponding to rotation quaternions represent a composite rotation in the 3D space. The composite rotation can be converted to joint angles with respect to the degree-of-freedom of the joint. The relative orientation caused by a joint's rotation is computed by utilizing the measured orientation of two arbitrary segments. Considering the active range of joint motion and the degree of freedom (DOF), joint angles could be eventually determined [21]. For example, waist joints θ_1, θ_2 and θ_3 are directly associated with chest's orientation, and the relative rotation matrix between pelvis and trunk can be calculated from (3) with

$$e^{[T_{\text{ch}}]\theta_{\text{ch}}} = \begin{bmatrix} c_1c_2 & -c_3s_1 - c_1s_2s_3 & s_1s_3 - c_1c_3s_2 \\ c_2s_1 & c_1c_3 - s_1s_2s_3 & -c_1s_3 - c_3s_1s_2 \\ s_2 & c_2s_3 & c_2c_3 \end{bmatrix} \tag{6}$$

where c_i denotes $\cos(\theta_i)$ function and s_i denotes $\sin(\theta_i)$ function. By analogy, the shoulder joint is resolved by

$$(e^{[T_{\text{ch}}]\theta_{\text{ch}}})^{-1} e^{[T_{\text{up}}]\theta_{\text{up}}} = \prod_{i=4}^{6} e^{[T_i]\theta_i} \tag{7}$$

which represents the orientation of the upper arm and trunk. The elbow joints is solved by

$$(e^{[T_{\text{up}}]\theta_{\text{up}}})^{-1} e^{[T_{\text{low}}]\theta_{\text{low}}} = \prod_{i=7}^{8} e^{[T_i]\theta_i} \tag{8}$$

Quaternions offer several advantages over traditional rotation matrices when representing orientation in three-dimensional space, including offering faster operations, seamless interpolation, and overcoming issues like singularities and gimbal lock.

Likewise, the acceleration of adjacent body parts $\mathbf{a}_{\text{B}}$, $\mathbf{a}_{\text{ch}}$, $\mathbf{a}_{\text{up}}$, and $\mathbf{a}_{\text{low}}$ are measured by individual IMU sensors. In addition to limb dynamics, the measured acceleration contains whole-body dynamics, gravity, and soft-tissue artifacts. As a person moves, distinct body segments have different orientations and translations, thus taking the average of these segments can approximate the body's overall acceleration. A mass-weighted average approach would be more precise; however, it is difficult to obtain mass of each body segment. The mean acceleration is removed as a bias, and the gravity component is filtered out by a high-pass Butterworth filter. The gravity component in IMU readings typically manifests as a low-frequency signal that can be removed by the high-pass filter. The Butterworth high-pass filter works in different environments with different gravitational conditions when the gravity is constant.

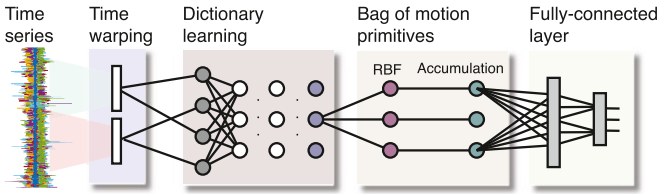

Fig. 3. The structure of bag-of-motion-primitives and dictionary learning. The inputs are raw sensory data collected by the IMU and EMG sensors, and the outputs are different levels of motion deficits.

To reduce the soft-tissue artifacts, which are typically in any direction along skin surfaces, we project the relative acceleration onto a plane perpendicular to the screw axis, which between adjacent quaternions. The presence of soft-tissue artifacts introduces unpredictable movements orthogonal to the screw axis, thus impacting the readings of the IMUs. Projecting the acceleration vector onto the plane perpendicular to the screw axis, which effectively reduces a portion of these extraneous motions, because the captured motion caused by soft-tissue artifacts in the directions that are not supported by the degree of freedom, e.g., motion along the arms is eliminated by the model. The motion of soft-tissue artifacts aligning with the degree of freedom is, however, not removed. For instance, the normalized relative acceleration $\bar{\mathbf{a}}_{\text{ch-up}} = \bar{\mathbf{a}}_{\text{ch}} - \bar{\mathbf{a}}_{\text{up}}$ is projected onto the plane perpendicular to $\omega_{\text{ch-up}}$, where $q_{\text{ch}} \ominus q_{\text{up}} \Rightarrow \{\theta_{\text{ch-up}}, \omega_{\text{ch-up}}\}$ with $\ominus$ as quaternion division. The components that does not satisfy the constraint of body kinematics are therefore filtered out to mitigate soft-tissue artifacts.

III. Learning the Dictionary of Motion Primitives

We developed a machine-learning structure to learn motion primitives and classify mobility deficits. As shown in Fig. 3, the structure consists of four major blocks of layers for time warping, dictionary learning, bag-of-motion primitives, and fully connected layers for classification. The time-warping block accepts time series with different lengths as input, and the bag-of-motion-primitives block classifies the severity of mobility deficits.

A. Dictionary of Motion Primitives

Let $X = [\mathbf{x}_i]_{i=1}^{N}$ be the time series of the collected data, where N is the number of motion data $\mathbf{x}_i = [x_{i,1}, \ldots, x_{i,m}]^T \in \mathbb{R}^m$,

where m stands for the number of the features, including joint angles, angular velocities, acceleration, and muscular electromyography. We utilize dictionary learning to learn a dictionary of representative motion primitives. The dictionary $D = \{\mathbf{d}_i\}_{i=1}^{K}$ is designed to contain K time series of motion primitives $\mathbf{d}_i \in \mathbb{R}^{n_i}$. The length of motion time series $\mathbf{x}_i$ depends on the time of motion tracking and segmentation. The sparse coding of motion tracking with time warping based on the dictionary of motion primitives $D = \{\mathbf{d}_j\}_{j=1}^{K}$ is formulated as

$$E(\mathbf{x}_i, D) = \min_{\boldsymbol{\alpha}_i, D} \|\mathbf{x}_i - F_i(D)\boldsymbol{\alpha}_i\|_2^2 \quad \text{s.t.} \|\boldsymbol{\alpha}_i\|_0 \leq \tau \tag{9}$$

where $F_i(D) = [f_i(\mathbf{d}_1), \ldots, f_i(\mathbf{d}_K)] \in \mathbb{R}^{n_i \times K}$ is the transformation of motion primitives D to a time-warped dictionary of motion primitives $f_i(\mathbf{d}_j) \in \mathbb{R}^n$, which are aligned to original motion tracking time series $\mathbf{x}_i$. The sparse codes $\boldsymbol{\alpha}_i = [\boldsymbol{\alpha}_{1,i}, \ldots, \boldsymbol{\alpha}_{K,i}]^T$ represent the coordinates of $\mathbf{x}_i$ with respect to the dictionary $F_i(D)$.

The dictionary learning problem is designed to extract a group of representative vectors for data with a fixed length; therefore, the time warp invariant function $F_i(D_l)$ is designed to map the dictionary to time series with different lengths, as the length of the time series vary when recorded during a long time span. The transformation function $F_i(D_l)$ is defined as an affine transformation of D, $f_i(\mathbf{d}_j) = \Delta_{ij}\mathbf{d}_j$, where the projection function $\Delta_{ij} \in \{0, 1\}^{m_i \times n_j}$ specifies the temporal alignment between $\mathbf{x}_i$ and $\mathbf{d}_j$.

The elements of motion tracking $\mathbf{x}_i$ and the elements of the dictionary $\mathbf{d}_j$ are aligned by a mapping

$$\boldsymbol{\pi} = [((\pi_1(1), \pi_2(2)), ((\pi_1(2), \pi_2(2)), \ldots, (\pi_1(p), \pi_2(p))] \tag{10}$$

where the applications π_1 and π_2 are defined from $\{1, \ldots, p\}$ to $\{1, \ldots, n\}$ and $\{1, \ldots, m\}$. Denote $\mathcal{A}$ as the set of all alignments between two time series. The alignment function can be computed by a cosine maximization of the time series with different lengths,

$$\boldsymbol{\pi}^* = \arg\max_{\boldsymbol{\pi} \in \mathcal{A}} \mathcal{L}(\boldsymbol{\pi}) \tag{11}$$

where $\mathcal{L}(\boldsymbol{\pi})$ is the cost function of the alignment $\boldsymbol{\pi}$ [22]. The alignment function $\boldsymbol{\pi}$ is modeled using a multiple-layer neural network with a linear activation function for the linear transformation Δ_{ij}, as given in the Fig. 3. Neural networks are frequently selected for temporal alignment tasks considering their proficiency in discerning intricate patterns and relationships within datasets. In contrast, traditional motion-tracking techniques, including optical flow algorithms and feature-based methods, often hinge on explicit assumptions on motion types or feature structures. Although these methods may exhibit computational efficiency under specific circumstances, they can encounter difficulties in handling complex motion patterns, occlusions, or variations in lighting conditions. Moreover, traditional approaches often necessitate manual parameter adjustment and may exhibit limited generalization capabilities across diverse datasets or real-world scenarios. In contrast to conventional motion tracking methods such as motion tracking with Extended Kalman Filter (EKF) [23], which present challenges in developing systems with high nonlinearities for human activity monitoring [24]. Neural networks demonstrate capability in addressing nonlinearities. The transformation layer allows the model to take time series with different lengths and therefore long-time tracking is possible for diagnosis and data collection.

One level of dictionary learning finds the optimal representation of data and the dictionary atoms. We intend to discover motion primitives that represent the mobility features of the time series instead of the time series themselves. To represent latent features of the time series, we further represent the time series using a multi-level dictionaries

$$X = D_1\varphi_1\left(D_2\left(\cdots\varphi_N\left(F\left(D_N\right)\boldsymbol{\alpha}\right)\right)\right) \tag{12}$$

where φ_i for $i = 1 \cdots N$ represents nonlinear activation functions. The optimization problem (9) can be explicitly written out as

$$\min_{\boldsymbol{\alpha}_i, D_l} \left\|\mathbf{x}_i - D_1\varphi\left(D_2\left(\cdots\varphi\left(F_i\left(D_N\right)\boldsymbol{\alpha}_i^l\right)\right)\right)\right\|_2^2$$
$$\text{s.t.} \left\|\boldsymbol{\alpha}_i^l\right\|_0 \leq \tau \tag{13}$$

The optimization problem can be solved in an iterative manner from $\varphi(D_2(\cdots\varphi(F(D_N)\boldsymbol{\alpha})))$ to $F(D_N)\boldsymbol{\alpha}$ by repeatedly applying dictionary learning as in [25].

The learned dictionaries D_N represents latent motion primitives that can be used to construct the tracked motion X. For patients with PD, their motion exhibits certain unique features, such as stiffness, imbalanced movement, or slowness. The movement of a non-PD person does not contain these features. Through the dictionary learning, we discover and extract the dictionary of motion primitives D_N to represent the movement of PD and non-PD people. By inspecting the way the movements are reconstructed by the dictionary D_N, we differentiate the movements and evaluate the severity of mobility deficits.

B. Bag of Motion Primitives

With the learned dictionary of motion primitives, we propose a concept of the Bag of Motion Primitives (BoMP) to model the construction of body motion in terms of motion primitives. The BoMP is inspired by the concepts of Bag of Words (BoW) [26] and Bag of Features (BoF) [27] models, which have been successfully used to solve computer vision and natural language processing problems. As compared to BoW and BoF, Bag of Motion Primitives (BoMP) introduces a data structure tailored to specifically capture motion deficits, and these primitives act as signatures indicative of these deficits. The motion primitives are not used as features or words but representative signatures from which we can quantify the severity of mobility deficits. The BoMP model generates a fixed-length histogram for each segment of the tracked motion by compiling motion primitives into codewords. The length of the generated histogram vector does not depend on the number of available motion primitives in the dictionary, so the BoMP model can handle time series with different lengths.

We formulate the BoMP model shown in Fig. 3 by using two layers: a Radial Basis Function (RBF) layer and an accumulation layer. The RBF layer measures the similarity of the time series

to the learned dictionary of motion primitives, and the accumulation layer computes the histogram of the motion primitives. The output of the k-th RBF neuron is the distance between time series and the k-th motion primitives. The RBF neurons

$$\phi_k(x) = \exp\left(-\left\|x - \mathbf{d}_k\right\|_2 / \sigma_k\right) \tag{14}$$

measures the similarity to the motion primitives and span the feature space, where σ is a scaling factor that adjusts the width of the Gaussian function. The output of the RBF neurons is accumulated to generate the final representation of each time series

$$s_i = \frac{1}{N_i} \sum_{j=1}^{N_i} \phi(x_{ij}) \tag{15}$$

where $\phi(x, \mathbf{d}) = [\phi_1(x, \mathbf{d}), \ldots, \phi_N(x, \mathbf{d})]^T \in \mathbb{R}^N$ is the output vector of the RBF layer. The output s_i defines a histogram distribution over the RBF neurons that describes the motion primitives. The activation of RBF neurons can be selected with a threshold such that only RBF neurons with maximal response is counted.

The severity of mobility deficits are classified and quantified by using the histograms. The features are designed to be representative, so we design deep learning models after the BoMP model as a classifier, considering the complexity of the whole architecture, as given in Fig. 3. A fifteen-layers convolutional neural network (CNN) model has been designed for classifying motion primitives. To avoid overfitting, we added two dropout layers. The first one is added after the input layer with a 0.5 dropout rate. The second one is added after the second convolutional layer with a 0.3 dropout rate. The ReLU has been adopted as the activation function for each layer except the output layer. The Sigmoid has been used as the activation function for the output layer.

The severity of mobility deficits is quantified by the weights of the motion primitives in whole-body motion. The motion primitives are, however, not explainable, i.e., we cannot distinguish the corresponding dysfunctional body segments from a motion primitive. Though we may be able to focus on one particular body segment, we cannot use motion primitives to identify the root cause of mobility deficits that may be helpful for the intervention for muscular or biomechanical dysfunctions.

IV. Experiment

We evaluated the effectiveness of the proposed method in detecting the level of mobility deficits caused by PD. We approximated motion primitives on the laboratory dataset that contains EMG signals and the clinical dataset. Due to the limited dataset, we simplified the approximation of (12) and (13) with the K-SVD method. We compared the reconstructed results from different sizes of approximated dictionaries. To assess the classification accuracy of the proposed models, we have applied the designed deep-learning models to the laboratory and clinical datasets. A comparison study between the proposed method and two other PD detection methods has been done. To investigate the generalization of the proposed method, we also applied the method to unknown human subjects. The datasets are shared online.[1]

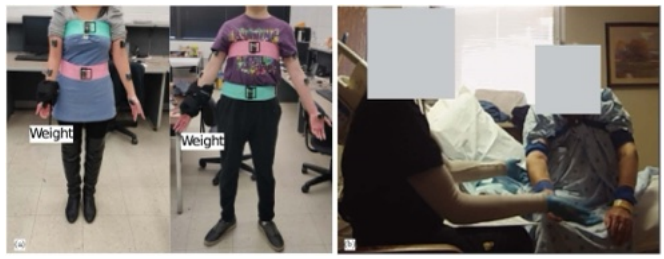

Fig. 4. Data collection. (a) Configuration during laboratory dataset collection. (b) Configuration of clinical dataset collection.

A. Data Collection

Both laboratory data and clinical data are collected from human subjects, in different configurations as shown in Fig. 4. During laboratory data collection, subjects wore weights to simulate motion deficits. During clinical data collection, both PD patients and healthy individuals were recruited.

1) Laboratory Data: Ten healthy subjects (eight males and two females) performed six activities (walking, squatting, picking up objects, drawing, yoga, and building toys) wearing eight Dysystem sensors–two on the torso and three on each arm. Arm sensors are placed on three major muscles on each arm: abductor pollicis brevis, brachioradialis, and biceps brachii. Arm sensors measured EMG and acceleration data along three axes (X, Y, and Z axes), while torso sensors only measured accelerations. Each participant performed each activity continuously for at least two minutes.

To simulate the mobility deficits, male participants were equipped with two five-pound sandbags, and female participants were equipped with one five-pound sandbag. Adding weights to the body causes mobility deficits. Participants were instructed to wear different sandbags, starting with a 1-pound weight and increasing incrementally to 2 lbs, 2.5 lbs, 5 lbs, 10 lbs, and 20 lbs. The aim was to identify the optimal balance weight that emulates motion instability similar to those experienced by individuals with PD. To compare data collected from healthy subjects with that collected from patients, each participant performed each activity twice: once with the weights labeled as positive (illness) and once without the weights labeled as negative (health). An illustration of subjects wearing sandbags is shown in Fig. 4. Although weights were equipped on only one arm, they affected the entire body's movement.

2) Clinical Data: Clinical data were collected from eight participants who had different levels of Parkinson's disease (including level 0 means health). Five sensors were placed on the subjects, each collecting accelerations in the X, Y, and Z axes. Two sensors were placed on each subject's arm, and one was placed on the chest of each subject. An illustration of the locations of each sensor is shown in Fig. 4. Participants are instructed to sit in a bed in front of a table. Each participant was

[1]https://github.com/hhelium/Quantification-and-Diagnosis-of-Mobility-Deficits.git

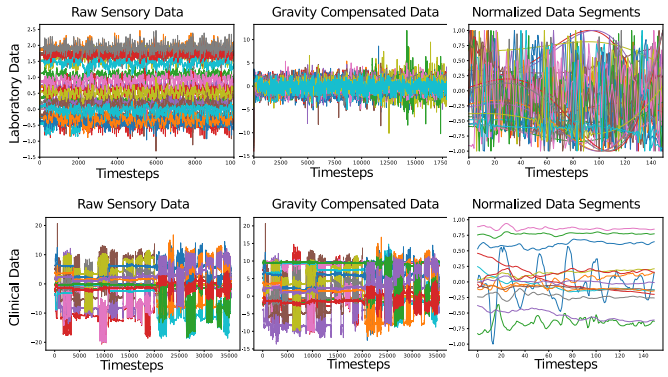

Fig. 5. Data pre-processing results. The top row shows the laboratory data, which has 30 features, and the bottom row shows the clinical data, which has 15 features. Each line represents a single feature.

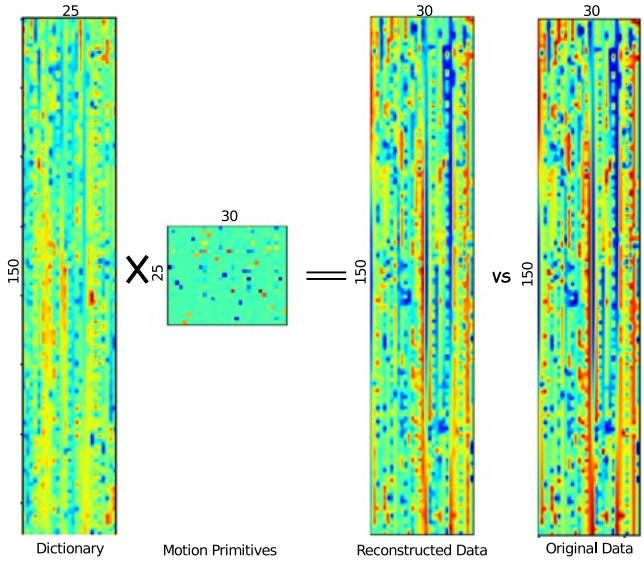

Fig. 6. Visualization of the learned dictionary, approximated motion primitives, the reconstructed data, and the corresponding ground truth. Colors nearing red indicate values approaching positive one, while those nearing blue signify values approaching negative one.

required to complete eight actions, including raising arms, moving arms from horizontal to up, moving arms from horizontal to down, moving arms from horizontal to left, moving arms from horizontal to right, stacking objects, grasping items, and passing objects from left hand to right hand.

3) Data Pre-Processing: To learn the motion primitives, we pre-processed the data through re-sampling, gravity compensation, normalization, and segmentation. Since the EMG signal and acceleration have different sampling frequencies, we re-sampled the EMG signal based on the acceleration signal's frequency. We compensated for gravity using a bandpass Butterworth filter [28] with a lower cutoff frequency of 0 Hz and an upper cutoff frequency of 0.168 Hz. The component of acceleration caused by the local gravity was removed by the high-pass filter. We then normalized the gravity-compensated signal in the region of $[-1, 1]$ using the MinMax method. To ensure uniform input shape for the designed deep learning models, we segmented the gravity-compensated data in 150 s lengths with 50 s stepping time. The clinical data underwent the same pre-processing steps, with the exception of re-sampling and stepping time. As the sensors used in clinical data collection does not include EMG signals, re-sampling was not necessary, and the stepping time was selected as 1 s due to limited data. To segment sensor readings in the laboratory data, a 50-second step time was chosen to provide adequate data for model training. In contrast, the sensors used for laboratory and clinical data collection have discrepancies in sampling rates. Thus, a one-second step time was employed to gather requisite data for model training from clinical sources. These time intervals divided sensor readings into segments of consistent length, ensuring each segment contains a suitable amount of data points for training purposes. The pre-processing results for the laboratory and clinical data are shown in Fig. 5, with the upper row displaying each pre-processing result for one segment of the laboratory data and the lower row shows the same for one segment of the clinical data.

Due to the limitations in clinical data collection, laboratory and clinical data were not collected in an identical setting. Eight sensors were placed on subjects in the laboratory data collection, while only five were placed on subjects in the clinical setup. The sensors used in laboratory data collection can collect four features, including EMG and accelerations along the X, Y, and Z axes. However, the sensors used in the clinical data collection only collect the accelerations along the X, Y, and Z axes. Due to physical and practical limitations, the laboratory and clinical data were collected in different locations and with different groups of subjects. The clinical dataset was collected before as supported by a different project, when EMG data were not required. Based on experiment results, the classification accuracy for motion deficits utilizing laboratory data, which include both IMU and EMG data, surpasses that based on clinical data, which solely comprises IMU data. This discrepancy may attribute to the absence of EMG data during clinical data collection. The application of the approach in the future will involve the collection of both EMG and IMU data, as contemporary sensors commonly integrate both functionalities.

B. Discovered Motion Primitives

We conducted two experiments to determine the best size for the approximated dictionary. These experiments focused on evaluating the textures of the reconstructed results and the reconstruction error. After analyzing the results, we found that the optimal size for the dictionary is 30, which provides the most accurate classification accuracy for mobility deficits.

1) The Texture of Reconstructed Data With Different Sizes of Dictionary: The first experiment was designed to assess the impact of different dictionary sizes on the texture difference between the reconstructed data $\hat{Y}$ derived from approximated motion primitives and the original data Y, measured through the reconstruction error ϵ computed as $\epsilon = \|Y - \hat{Y}\|_2$. An example of dictionary representation and discovered motion primitives is shown in Fig. 6, the first left one represents the learned dictionary D from the input data, the second left one

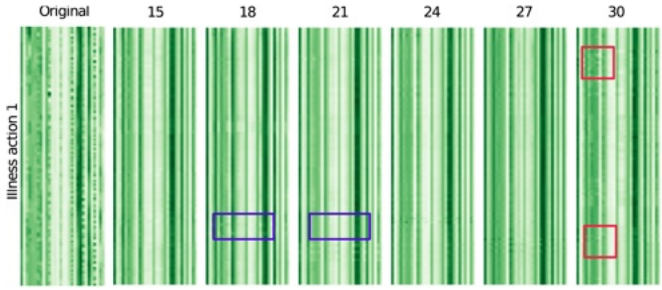

Fig. 7. Reconstruction results with different dictionary sizes for the laboratory data. The original column is the corresponding ground truth. The dictionary sizes are listed on the top of each reconstructed result. The red rectangles are textures difference, and the blue rectangles are missing textures.

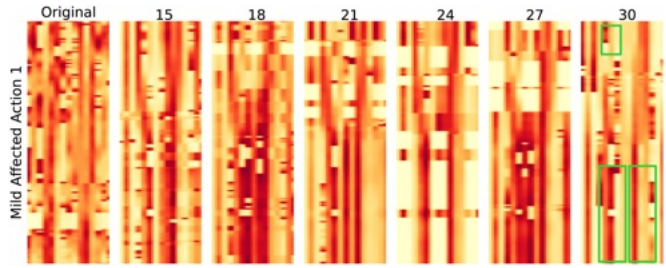

Fig. 8. Reconstruction results with different dictionary sizes for the clinical data. The original column is the corresponding ground truth. The dictionary sizes are listed on the top of each reconstructed result.

portrays the approximated motion primitives of the input data, α. The two right columns display the reconstructed data and their corresponding original data.

In this experiment, we combined the data from each subject into one dataset and learned the dictionary. The feature-length was 150, so the vocabulary size was also 150. We adjusted the vocabulary numbers to 15, 18, 21, 24, 27, and 30 while keeping the non-zero coefficient fixed at three. We selected multipliers of five, six, seven, eight, nine, and ten, as anything below five could not approximate a valid dictionary, and the algorithm would not converge, while those above ten dramatically increased the computation time. Additionally, we investigated different dictionary sizes' impact on laboratory and clinical data. We found that using larger multipliers to approximate the dictionary resulted in lower PD detection accuracy than using smaller multipliers.

We used learned motion primitives from a dictionary and corresponding ground truth to reconstruct laboratory data. The reconstructed data and their corresponding original data are shown in Fig. 7. The "illness" action involved subjects wearing weights to simulate mobility deficits, while the "healthy" the action involved the same motion without weights. From Fig. 7, the major features of the original motion data were retrieved in the reconstructed data. As the dictionary size increased, the textures of the reconstructed results became finer, as highlighted by the red rectangles in the "illness" and "healthy" actions. Moreover, some parts of the reconstructed results were finer with smaller dictionary sizes than with larger ones, as boxed by the blue rectangles in Fig. 7. Specifically, the reconstructed result using a dictionary size of 18 was finer than that using a dictionary size of 21.

In Fig. 8, we compared reconstructed clinical data with its original data. The figure uses darker colors to show higher data values. From Fig. 8, we can see that larger dictionary sizes lead to more detailed reconstructions of different parts, as indicated by the green rectangles. The dictionary sizes we selected for reconstructing clinical data are 15, 18, 21, 24, 27, and 30. We chose to demonstrate results from one mildly affected action from one subject, as similar patterns were observed in other results.

TABLE I
RECONSTRUCTION ERRORS WITH DIFFERENT DICTIONARY SIZES

Dictionary Size	15	18	21	24	27	30
Laboratory	16.3 %	15.9 %	15.4 %	15.1 %	14.8 %	14.5 %
Clinical	3.6%	2.9 %	2.2%	1.7%	1.7 %	1.6 %

2) Reconstruction Error of Dictionaries with Different Sizes: We analyzed the reconstruction errors of different dictionary sizes and presented the results in Table I. The columns labeled 'Laboratory' indicate the reconstruction error for laboratory data, while the columns labeled 'Clinical' indicate the reconstruction error for clinical data. Table I shows that the reconstruction errors decrease as the vocabulary size of the dictionary increase. The smallest dictionary size (15) has the highest reconstruction error (16.47%) for laboratory data and (3.66%) for clinical data.

In contrast, the largest dictionary size (30) has the smallest reconstruction error (14.48%) for laboratory data and (1.59%) for clinical data. Furthermore, the reduction in reconstruction errors was more gradual for clinical data than laboratory data. This is because laboratory data has more features than clinical data, making it harder to reconstruct. Based on the analysis, we selected the dictionaries with the smallest reconstruction errors for both laboratory data (30) and clinical data (30).

C. Quantification of Mobility Deficit for Known Subjects

To evaluate the efficacy of the proposed methods, we investigated the classification accuracy for mobility deficit severity on the laboratory and clinical data. Twelve actions were evaluated for each participant from the laboratory dataset: six positive actions simulating actions under PD influence and six negative actions considered healthy counterparts. We formulated the problem as a binary classification, with outputs indicating either a positive (mobility deficit) or negative (no mobility deficit) result. We divided 3,660 inputs into 75% (2745) training data and 25% (915) testing data.

The UPDRS categorizes the severity of Parkinson's Disease into three distinct levels [29], [30]. In alignment with this classification system, we structured the classification results into three categories accordingly. We approached the clinical data as a multi-label classification problem with outputs indicating

TABLE II
COMPARISON OF MOBILITY DEFICIT PREDICTION WITH DIFFERENT METHODS FOR LABORATORY DATA

		Precision	Recall	F-1 Score	Accuracy
This paper	Healthy	100%	88%	93%	93.95%
	Illness	89%	100%	94%	
Linear Model	Healthy	64%	65%	64%	64.86%
	Illness	65%	65%	65%	
SVM	Healthy	92%	91%	91%	91.35%
	Illness	92%	91%	91%	

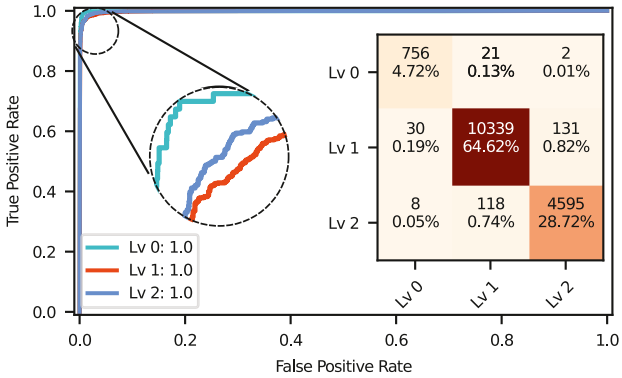

Fig. 9. Performance of the mobility deficit classification on known clinical data. The plot contains ROC curves, AUC scores, and a confusion matrix for each label.

health, mild PD affection, and severe PD affection. In this paper, the motion primitive is a scale to differentiate the severity of mobility deficits. From the extracted motion primitives, we classified the severity of the motion deficits by the amount of primitives representing different mobility deficits. There are 64,000 data in the clinical data, 75% (48,000) of the total data is used as training data, and 25% (19,200) is used as testing data.

The performance of the proposed methods has been investigated using both laboratory and clinical data. The proposed CNN-based classifier achieved a mobility deficit classification accuracy of 93.95% on laboratory data. 5% of the test data with healthy labels were classified as an illness, and 4% of the test data with illness labels were classified as healthy.

We have compared the proposed method with other methods, and the comparison results are shown in Table II. The accuracy of the proposed model achieved 93.95%, which is 29.09% higher than the linear model with logistic regression [31] and 2.6% higher than the support vector machine (SVM) method [32]. These results indicate that the proposed CNN-based model can effectively classify PD-related mobility deficits on the laboratory data and outperforms other methods in classifying PD-affected mobility deficits.

The proposed MLP-based model achieved 99.84% accuracy when classifying three levels of mobility deficits affected by PD. The confusion matrix in Fig. 9 indicates that the majority of data are mildly affected by PD. Because the unbalanced training

TABLE III
COMPARISION OF MOBILITY DEFICIT PREDICTION WITH DIFFERENT METHODS FOR CLINICAL DATA

		Precision	Recall	F-1 Score	Accuracy
This paper	Healthy	95%	97%	96%	99.84%
	Mild	99%	98%	99%	
	Severe	97%	97%	97%	
Linear Model	Healthy	0	0	0	69.48%
	Mild	72%	88%	79%	
	Severe	60%	40%	48%	
SVM	Healthy	92%	64%	76%	90.35%
	Mild	91%	95%	93%	
	Severe	88%	85%	87%	

data, most falsely classified results were misclassified as mildly affected data.

We compared the proposed classifier with other methods [31], [32], and the comparision results are shown in Table III. The proposed model achieved 99.84%, outperforming the linear model with logistic regression [31] by 30.36% and SVM [32] model by 9.49%. The proposed models achieved the highest accuracy in classifying mobility deficits that due to the PD for both laboratory data and clinical data. Therefore, using motion primitives that are learned from the approximated dictionary can enhance the ability to classify the mobility deficits level caused by PD.

To investigate the difference in classification accuracy between using motion primitives and without using motion primitives as inputs. We conducted an experiment to compare the effectiveness of using motion primitives versus raw sensory data for detecting mobility deficits. Both experiments used CNN-based classifiers with the same structures. However, one experiment used motion primitives as training data, while the other used raw sensory data as input. The classifier trained with motion primitives achieved a 95% accuracy rate on known test data and a 61.69% accuracy rate on unknown test data in 92.41 seconds. On the other hand, the classifier trained with raw sensory data had a classification accuracy of 48.08% on known test data and 46.56% on unknown test data. Both known and unknown test data were not included in the training dataset. Known test data was obtained by splitting the total dataset, while unknown test data was gathered from participants who were not part of the total dataset.

D. Quantification of Mobility Deficit for New Subjects

To evaluate the efficacy of the proposed method on new subjects, we applied the models to new subjects that were not included in the training data for both laboratory and clinical data. The accuracy of the classifiers was 59.25% (proposed method), 48.88% [31], and 64.00% [32]. The SVM classifier [32] achieved the highest classification accuracy. The recall of the classifiers on unseen test subjects is shown in Fig. 11. The proposed method achieved the highest recall, which is 83.27%, and it is 35.8%

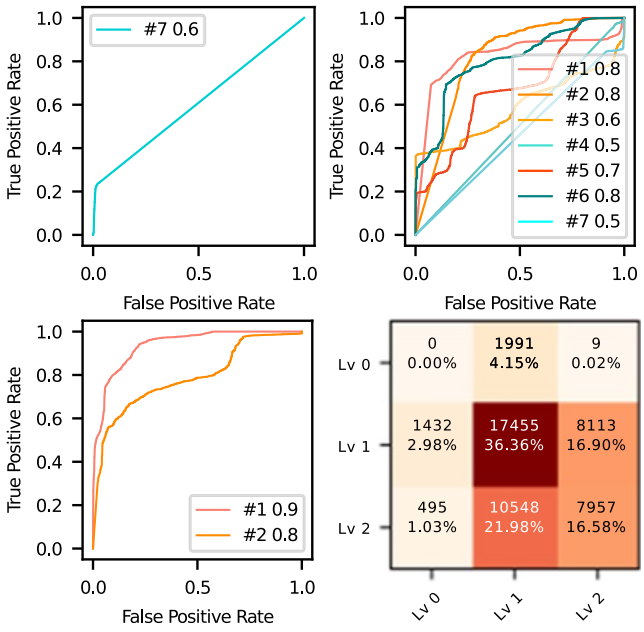

Fig. 10. Performance of the mobility deficit classification on unknown clinical data. The plot contains ROC curves, AUC scores, and a confusion matrix for each label.

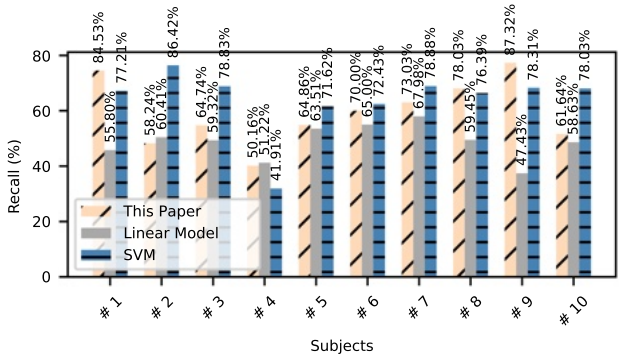

Fig. 11. Diagnosis rate of different methods on laboratory data. Different methods are applied to the laboratory data, which include the linear model [31], SVM [32], and the proposed method.

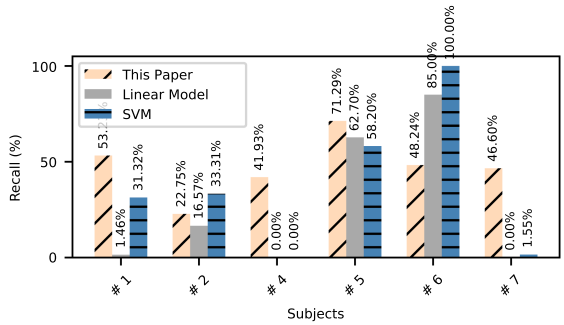

Fig. 12. Diagnosis rate of different methods on clinical data. Different methods are applied to the laboratory data, which include the linear model [31], SVM [32], and the proposed method.

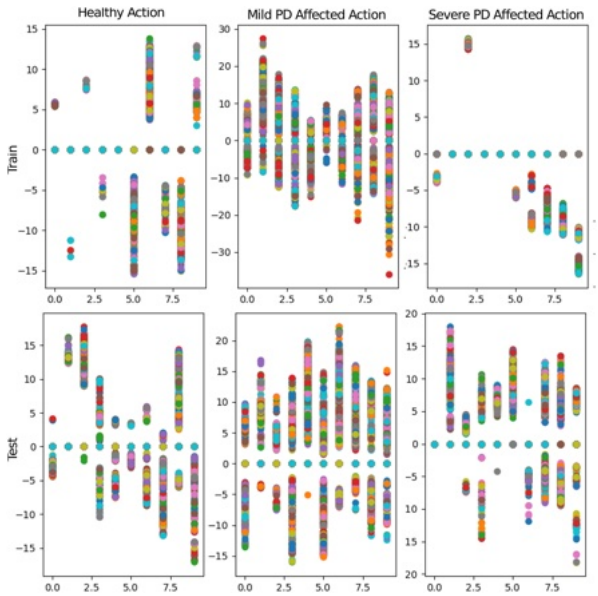

Fig. 13. Comparison of the training motion primitives and the testing motion primitives for affected action by the same level of motion deficits. Each dot denotes a value in the corresponding motion primitives.

higher than the linear model with logistic regression [31], and 14.22% higher than the SVM model [32].

Similar experiments were performed on the clinical data, where 798 (4%) samples were from healthy subjects, 12725 (66.28%) were from mildly affected subjects, and 5677 (29.57%) were from the severely affected subjects. Because the percentage of healthy subjects is low, we removed all such data to evaluate the proposed method's efficacy. The classification accuracy achieved by each model was: the proposed method (68.4%), the linear model with logistic regression [31] (72.47%), and SVM model [32] (67.36%). The linear model with logistic regression [31] achieved the highest average classification accuracy. The recalls of classification models are shown in Fig. 12. The proposed model achieved the highest recall, which is 38%. It is 10.38% higher than the linear model with logistic regression [31], and 0.61% higher than SVM model [32]. Fig. 10 presents the ROC curves and AUC values for each participant in the clinical dataset. The level of mobility deficits varies across tasks for the same participants. Consequently, the number of ROC curves and AUC values differ for each level of mobility deficit. Additionally, the ROC curve and AUC results for the same participant vary depending on the level of mobility deficit.

In the previous section, we compared the average classification accuracy of different models using both laboratory and clinical data. Tables II and III show that the proposed method performed better than other methods on known subjects. However, it did not achieve the highest accuracy on unknown subjects. The proposed method is based on deep learning, which excels at approximating data with features within the distribution, but deep learning methods are not as effective when features fall outside the distribution. Fig. 13 illustrates that the motion primitives of the unknown patient fall outside the distribution. The first row in the figure displays features of the training data while the second row shows features of the testing data from the

same group. Each dot represents a value of an element in the vocabulary.

Based on the results shown in Figs. 11 and 12, the proposed method achieved the highest recall for both laboratory and clinical data. The proposed model can achieve better recall than classification accuracy because the motion patterns of healthy subjects can vary. However, the motion patterns are similar if subjects'. mobility deficits are similar or close. Therefore, the classification model can more easily detect mobility deficits than a healthy subjects' motion patterns. We need data from a broad spectrum of patients to achieve high accuracy.

V. Conclusion

The paper revealed that motion primitives exist in simple mobility measurements and demonstrated the effectiveness of the primitives in quantifying mobility deficits. This paper has provided a user-specific metric that quantifies the mobility deficits in terms of the quantities of motion primitives that are learned from motion tracking data. The approach can be used to replace observation-based clinical scaling and provide feedback in real-time to reinforce positive recovery during rehabilitation. The proposed method outperformed two other PD detection methods on known subjects and the proposed method performed better in diagnosing unknown subjects with mobility deficits.

References

[1] Parkinson's Foundation. [Online]. Available: https://www.parkinson.org/understanding-parkinsons/statistics

[2] C. G. Goetz et al., "Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): Scale presentation and clinimetric testing results," *Movement Disord., Official J. Movement Disorder Soc.*, vol. 23, no. 15, pp. 2129–2170, 2008.

[3] C. Ramaker et al., "Systematic evaluation of rating scales for impairment and disability in Parkinson's disease," *Movement Disord., Official J. Movement Disorder Soc.*, vol. 17, no. 5, pp. 867–876, 2002.

[4] F. M. Ivey et al., "The unified Parkinson's disease rating scale as a predictor of peak aerobic capacity and ambulatory function," *J. Rehabil. Res. Develop.*, vol. 49, no. 8, 2012, Art. no. 1269.

[5] Y. Sano et al., "Quantifying Parkinson's disease finger-tapping severity by extracting and synthesizing finger motion properties," *Med. Biol. Eng. Comput.*, vol. 54, no. 6, pp. 953–965, 2016.

[6] S. Rissanen et al., "Surface EMG and acceleration signals in Parkinson's disease: Feature extraction and cluster analysis," *Med. Biol. Eng. Comput.*, vol. 46, pp. 849–858, 2008 doi: 10.1007/s11517-008-0369-0.

[7] S. Askari, M. Zhang, and D. S. Won, "An EMG-based system for continuous monitoring of clinical efficacy of Parkinson's disease treatments," in *Proc. IEEE 2010 Annu. Int. Conf. Eng. Med. Biol.*, 2010, pp. 98–101.

[8] G. Rigas et al., "Assessment of tremor activity in the Parkinson's disease using a set of wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 3, pp. 478–487, May 2012.

[9] S. M. Rissanen et al., "Analysis of EMG and acceleration signals for quantifying the effects of deep brain stimulation in Parkinson's disease," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 9, pp. 2545–2553, Sep. 2011.

[10] D. A. Heldman et al., "Essential tremor quantification during activities of daily living," *Parkinsonism Related Disord.*, vol. 17, no. 7, pp. 537–542, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1353802011001118

[11] S. Patel et al., "Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.

[12] A. Salarian et al., "Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 313–322, Feb. 2007.

[13] C. Anderson et al., "Home or hospital for stroke rehabilitation? Results of a randomized controlled trial I: Health outcomes at 6 months," *Stroke*, vol. 31, no. 5, pp. 1024–1031, 2000.

[14] D. Giansanti, G. Maccioni, and V. Macellari, "The development and test of a device for the reconstruction of 3-D position and orientation by means of a kinematic sensor assembly with rate gyroscopes and accelerometers," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 7, pp. 1271–1277, Jul. 2005.

[15] X. Yun and E. R. Bachmann, "Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking," *IEEE Trans. Robot.*, vol. 22, no. 6, pp. 1216–1227, Dec. 2006.

[16] H. Zhou et al., "Use of multiple wearable inertial sensors in upper limb motion tracking," *Med. Eng. Phys.*, vol. 30, no. 1, pp. 123–133, 2008.

[17] G. Cooper et al., "Inertial sensor-based knee flexion/extension angle estimation," *J. Biomech.*, vol. 42, no. 16, pp. 2678–2685, 2009.

[18] J. Favre et al., "Ambulatory measurement of 3D knee joint angle," *J. Biomech.*, vol. 41, no. 5, pp. 1029–1035, 2008.

[19] J. Lin and D. Kulic, "Human pose recovery using wireless inertial measurement units," *Physiol. Meas*, vol. 33, no. 12, pp. 2099–2115, 2012.

[20] A. Filippeschi et al., "Survey of motion tracking methods based on inertial sensors: A focus on upper limb human motion," *Sensors*, vol. 17, no. 6, 2017, Art. no. 1257.

[21] C. C. Norkin and D. J. White, *Measurement of Joint Motion: A Guide to Goniometry*. Philadelphia, PA, USA: F.A. Davis, 2009.

[22] S. Yazdi et al., "Time warp invariant dictionary learning for time series clustering: Application to music data stream analysis," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discov. Databases*, 2018, pp. 356–372.

[23] M. El-Gohary and J. McNames, "Human joint angle estimation with inertial sensors and validation with a robot arm," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 7, pp. 1759–1767, Jul. 2015.

[24] S. Majumder, T. Mondal, and M. J. Deen, "Wearable sensors for remote health monitoring," *Sensors*, vol. 17, no. 1, 2017, Art. no. 130.

[25] J. Yang et al., "Linear spatial pyramid matching using sparse coding for image classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 1794–1801.

[26] H. M. Wallach, "Topic modeling: Beyond bag-of-words," in *Proc. 23rd Int. Conf. Mach. Learn.*, 2006, pp. 977–984.

[27] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 490–503.

[28] S. S. Daud and R. Sudirman, "Butterworth bandpass and stationary wavelet transform filter comparison for electroencephalography signal," in *Proc. IEEE 2015 6th Int. Conf. Intell. Syst., Modelling Simul.*, 2015, pp. 123–126.

[29] M. Skorvanek et al., "Differences in MDS-UPDRS scores based on Hoehn and Yahr stage and disease duration," *Movement Disord. Clin. Pract.*, vol. 4, no. 4, pp. 536–544, 2017.

[30] R. Balestrino et al., "Applications of the European Parkinson's disease association sponsored Parkinson's disease composite scale (PDCS)," *NPJ Parkinson's Dis.*, vol. 5, no. 1, 2019, Art. no. 26.

[31] T. Arroyo-Gallego et al., "Detection of motor impairment in Parkinson's disease via mobile touchscreen typing," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 1994–2002, Sep. 2017.

[32] M. Zhang and A. A. Sawchuk, "Motion primitive-based human activity recognition using a bag-of-features approach," in *Proc. 2nd ACM SIGHIT Int. Health Inform. Symp.*, 2012, pp. 631–640.