Personalized Estimation of Intended Gait Speed for Lower-Limb Exoskeleton Users via Data Augmentation using Mutual Information

Roopak M. Karulkar[1], Patrick M. Wensing[1]

***Abstract*—This letter presents a method for data-driven user-specific gait speed estimation for people with Spinal Cord Injuries (SCIs) walking in lower-limb exoskeletons. The scarcity of training data for this population is addressed by leveraging common patterns across users that relate gait changes to speed changes. To bootstrap the process, widely available walking data from uninjured individuals was used as a base dataset. The distribution of this data was first transformed to match smaller user-specific training sets from walking trials of subjects with SCIs. User-specific trials were then selected based on the mutual information between gait speed and features for the combined dataset. The resulting selected data was finally used to build a model for estimating the user's intended gait speed. The performance of this approach was evaluated using data from two users with SCIs walking in an EksoGT exoskeleton with a walker or crutches. Estimation trials were compared when using the base data alone versus when providing personalization via the addition of novel data. The average successful estimation of speed-up and slow-down changes increased from 52% to 67% with personalization using only 8 to 12 steps' worth of user-specific data, with a best-case improvement of 32%, from 48% to 80%. Overall, the proposed method uses the mutual information between gait features and speed to provide a reliable alternative to manual data selection while pooling data from healthy and injured individuals.**

***Index Terms*—Prosthetics and Exoskeletons, Intention Recognition, Rehabilitation Robotics**

I. Introduction

A. Motivation & Previous Work

THE US has over a quarter-million existing cases of SCIs [1] with the cost of care per patient exceeding half a million dollars [2]. One of the main mechanisms of recovery from SCIs is accomplished through the reorganization of a person's intact neuronal pathways [3]. In order to take advantage of neural plasticity and support this reorganization, gait rehabilitation strategies often involve repeatedly moving the patient's legs through prescribed walking trajectories. In recent years, robotic exoskeletons have been cleared by the FDA for use in gait rehabilitation due to their ability to consistently track the necessary joint trajectories, which may result in accelerated rehabilitation [4]. There are multiple commercial exoskeletons currently available, including the Ekso GT [5], Indego [6], and ReWalk Personal System [7].

As these devices increase the patient's level of autonomy, fluent Human-Robot Interaction (HRI) is desired to maintain the safety and efficacy of the treatment. HRI fluency is an abstract notion but can be defined roughly as the reliability with which a human and robot can predict each other's future actions [8]. It may be quantified by the inverse of the time it takes to complete desired tasks [9]. This time may be minimized if the robot can anticipate changes to the user's intent and assist as necessary. The overall goal of this work is to increase the fluency in HRI for lower-extremity exoskeletons by estimating user intent. Intent itself is difficult to quantify, so an exoskeleton user's desired forward speed is considered as representing their intent in this study.

Some available exoskeletons use rudimentary methods to infer user intent. The HAL exoskeleton uses force sensors under the feet to detect weight transfer to initiate a step [10] and the ReWalk system uses a combination of ground reaction force sensors and torso tilt [11]. Using additional gait features, such as step length, frequency, and joint angles in intent estimation may provide further insight into the user's desired motion through their gait patterns.

An exoskeleton user's intended gait speed may be inferred using gait feature information, and this inference is often pursued through data-driven strategies. Gait features of steady-state walking can be qualitatively modeled using physics-based models of locomotion such as the Bipedal Spring-Loaded Inverted Pendulum (B-SLIP) [12], [13]. However, aspects of human decision-making and intent realization are not well modeled using first principles, so it is challenging to use them to anticipate gait changes. As a result, data-driven strategies such as Convolutional Neural Networks (CNNs) [14], Gaussian Processes (GPs) [15], gradient boosted decision trees [16], and Gaussian methods [17] have been more commonly used to infer user intent. These methods require a large amount of training data, and this requirement may further increase when attempting to train user-independent models due to the inter-subject gait variability seen in human walking. The work presented herein (Fig. 1) describes a method to estimate intended gait speed while addressing data requirements and gait variability.

Gait patterns of people with iSCI may exhibit higher inter-

Manuscript received: February, 24, 2022; Revised May, 23, 2022; Accepted June, 21, 2022.

This paper was recommended for publication by Editor Angelika Peer upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the National Science Foundation under Grant IIS-1734532.

The human subject trial data used in this work was collected as part of a study approved by the IRB of the University of Notre Dame (Protocol 18-04-4650).

[1]Authors are with the Department of Aerospace and Mechanical Engineering, University of Notre Dame, Notre Dame, IN, USA {rkarulka,pwensing}@nd.edu

Digital Object Identifier (DOI): see top of this page.

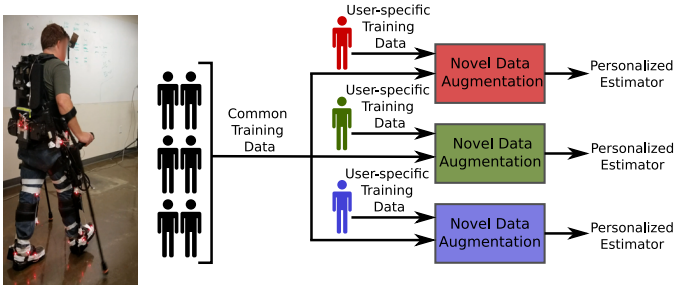

Fig. 1. Estimator personalization by using base and user-specific novel data from walking trials in the EksoGT exoskeleton.

subject variability than uninjured people [18], which creates the need for personalized exoskeleton assistance to better suit each user. Additionally, personalization may better capture the increased intra-subject variability because of spastic disturbances resulting from iSCIs [19]. Tucker et al. [20] further highlight the benefits of personalizing exoskeleton assistance by giving the user pairwise options to modify exoskeleton control parameters and increase comfort. While controller personalization is important for comfort, the personalization of intent estimation represents a complementary area of need for HRI fluency.

One of the challenges of effective personalization is the prohibitive amount of training data required. As there is a limited amount of data for exoskeleton-assisted walking, the work presented in this letter aims to address that scarcity by fusing trial data from uninjured users with data from novel users (Fig. 1). The main idea of the approach is that gait feature trends exhibit similarities across subjects, e.g., step length and frequency increase with speed. The estimator personalization developed in this work seeks to exploit these similarities and create a base dataset from uninjured subjects. As people with iSCIs are expected to show similar gait feature trends, this dataset may then be transformed to provide additional training data for them. Similar ideas of exploiting gait feature commonalities have previously been used to develop user-independent gait mode estimation approaches for healthy individuals [21], [22].

User-independent intent detection strategies for powered prostheses are also relevant to the goals of this study, since prosthesis control strategies must address similar HRI challenges. There is very little work regarding user-independent intent recognition for powered prostheses and even less for exoskeleton-assisted walking for individuals with iSCIs. User-independent prediction of gait mode for prosthesis users has been implemented using Linear Discriminant Analysis (LDA) [23], Dynamic Bayesian Networks (DBNs) [24], and gradient boosted decision trees [25] with each method showing improvement over the last.

A majority of the state-of-the-art work referenced previously considers the problem of activity classification, i.e., detecting walking on flat ground, or ascent/descent on ramps and stairs. The objective of the work presented in this letter is to capture intent changes through changes in intended gait speed. This objective, coupled with the gait variability seen in exoskeleton walking, means there is difficulty in obtaining labeled training data about changes in intended speed as compared to changes in activity. This shortage of training data increases the difficulty of developing estimators that can identify changes in the intended gait speed. Further, while there is some initial previous work regarding continuous speed estimation for prostheses [26], no such work is available regarding exoskeleton-assisted walking.

B. Contribution

Despite inter-subject gait variability, common patterns relating gait changes to speed changes are observed across users. For example, step length and frequency, pitch and roll motions of the torso, and joint angle trajectories all show qualitatively similar trends relating to speed changes across individuals. The main contribution of this work is a method to personalize the estimation of the gait speed for subjects with SCIs walking in a lower-limb exoskeleton. The proposed method addresses training data scarcity by supplementing novel user data with transformed gait data collected from trials of uninjured users.

The new method was evaluated on an experimental dataset containing speed change trials of subjects with and without iSCIs walking in an Ekso GT lower-limb exoskeleton (Fig. 1) using an estimation framework based on a Buttressed Kalman Filter (BKF) [27]. In the original BKF, the models used in the estimator were trained using data from a single subject. In contrast, the work presented herein explores how to incorporate data from multiple subjects to generate the necessary models, while retaining personalization for new users.

This study also explores the effect that the user's choice of assistive ambulatory device, a walker or crutches, has on the personalized estimator. On average, estimator personalization resulted in increased success in estimating the subjects' desire to speed-up (SU)/slow-down (SD), with changes detected before they were physically realized.

C. Overview

The remainder of the paper is organized as follows. Section II details the training data selection and augmentation, and the estimator framework. The performance of the method for crutch and walker-assisted walking was evaluated on experimental data, as discussed in Section III. Concluding remarks and future work are given in Section IV.

II. Methods

One of the difficulties in using learning-based strategies is the shortage of training data. There are multiple datasets of walking trials of uninjured individuals [28]–[30] however there is a lack of data from walking trials of exoskeleton users. This problem is further complicated by the coupling present in human-robot dynamics, as each user may interact differently with the robot [31]. This variability increases when considering different ambulatory devices (e.g., crutches) [32] or injury severity [33], [34]. Therefore, it is important to develop a method that may address the data scarcity in exoskeleton-assisted walking by enabling the re-use of training data across multiple users.

A. Exoskeleton Dataset

The reliance of many state-of-the-art intent inference approaches on external sensors like EMGs may be problematic in practical applications, as EMGs need consistency in placement and may slip during usage due to perspiration [35], [36]. Therefore, this work strives to exclusively use measurements from sensors onboard the exoskeleton as they may offer a more reliable option [17] in addition to being cost-effective.

The work herein was applied with trial data [17] from users walking in the EksoGT exoskeleton developed by Ekso Bionics (Fig. 1). Data were acquired from three uninjured and two injured subjects. The exoskeleton has two modes of operation, free and adaptive. Free mode is similar to gravity compensation, whereas, in adaptive mode, the exoskeleton follows a predefined trajectory and corrects any deviations from it. Uninjured users underwent trials in both modes and injured users only underwent trials in adaptive mode.

Sensors onboard the exoskeleton provide hip pitch, knee pitch, and torso pitch and roll angles, and are fused to estimate the height and fore-aft position of the hip in a global frame. These readings were used to approximate the location of the subject's CoM with respect to the stance foot. Since the position of the CoM is considered relative to the stance foot, the drift that may be present in the global position estimate does not affect step-to-step calculations. The subject's height, thigh, and shank lengths were recorded and the location of their CoM was approximated to be at the centroid of the pelvis. The remaining dimensions, such as ankle height and hip width, were computed using anthropometry relationships defined by Winter [37]. The CoM velocities and angular velocities of the joints and torso were computed with finite-difference approximations.

The two main gait events to be identified for the estimator were midstance (MS) and touchdown (TD). A zero-crossing event between the left and right hip angles was used to detect MS. TD was detected when force sensor readings from both feet were above a threshold of 5% of the maximum sensor value. Eighteen gait features ($S = 18$) listed in Table I were considered for use in the estimator.

TABLE I
GAIT FEATURES CONSIDERED FOR DESIRED GAIT SPEED ESTIMATION

Gait Feature	Description
Step Length (m)	Step length as computed at TD
RMS Swing Current - Hip (A)	Swing leg hip motor - MS to TD
RMS Swing Current - Knee (A)	Swing leg knee motor - MS to TD
Time-to-TD (s)	Time from MS to TD - proxy for step frequency
Hip Angle - Swing (rad)	Hip angle of the swing leg at TD
Knee Angle - Swing (rad)	Knee angle of the swing leg at TD
Hip Angular Velocity - Swing (rad/s)	Hip joint velocity - swing leg at TD
Knee Angular Velocity - Swing (rad/s)	Knee joint velocity - swing leg at TD
Hip Angle - Stance (rad)	Hip angle of the stance leg at TD
Knee Angle - Stance (rad)	Knee angle of the stance leg at TD
Hip Angular Velocity - Stance (rad/s)	Hip joint velocity - stance leg at TD
Knee Angular Velocity - Stance (rad/s)	Knee joint velocity - stance leg at TD
Torso Pitch Angle (rad)	Angle with the vertical in the sagittal plane
Torso Pitch Angular Velocity (rad/s)	Angular velocity in the sagittal plane
Torso Roll Angle (rad)	Angle with the vertical in the frontal plane
Torso Roll Angular Velocity (rad/s)	Angular velocity in the frontal plane
Leg Angle (rad)	The angle made with the vertical by the line connecting the estimated CoM and leading foot position at TD
Leg Angle (rad)	The angle made with the vertical by the line connecting the estimated CoM and leading foot position at TD
Current gait speed (m/s)	The gait speed measured at MS

B. Novel User Data Augmentation

1) Transforming Data from Uninjured Users to Match Novel Data: Data from healthy users walking in an exoskeleton may be easily obtained to satisfy the requirements of data-driven methods. These data still retain high-level similarities in gait feature trends (e.g., changes in step length with changes in gait speed [27]). This commonality between gait patterns may be exploited to augment the amount of available training data for injured users.

Gait feature and gait speed measurements were found to be well approximated with Gaussian distributions [38], which motivated transforming the data from healthy user trials to match the mean and standard deviation of data from an injured user. A transformation was performed on a vector of measurements $\mathbf{q}_s \in \mathcal{R}^N$ of an individual gait feature, where N is the number of measurements and $s \in \{1 \dots S\}$. A vector containing measurements of a single gait feature is also denoted by $\mathbf{q}$, with the subscript s omitted for readability. Its mean and standard deviation are $\bar{q}$ and σ respectively. Subscripts b and n denote base and novel data respectively and n/b represents base data that has been transformed to match the distribution of novel data from a single user via:

$$\mathbf{q}_{n/b} = (\mathbf{q}_b - \bar{q}_b)\sigma_n\sigma_b^{-1} + \bar{q}_n \quad (1)$$

$$\mathbf{q} \leftarrow [\mathbf{q}_{n/b}^T \ \mathbf{q}_n^T]^T \quad (2)$$

The features are then collected in a matrix $\mathbf{Q} \in \mathcal{R}^{N \times S}$ such that $\mathbf{q} = [\mathbf{q}_1 \dots \mathbf{q}_S]$.

It is important to choose appropriate novel data to ensure that the gait feature data carries a sufficiently high amount of information about the subject's desired gait speed.

2) Choosing Appropriate Novel Data: Steady-state walking in trials of subjects with iSCIs had a standard deviation of up to 0.18 m/s for their walking speed compared to 0.1 m/s seen in healthy users walking without robot assistance [39]. In addition to the severity of the iSCI, variability may be affected by user fatigue, discomfort, or misfit orthoses. As a result, some walking trials may better represent the exoskeleton user's gait patterns than other trials performed on the same day. Therefore, choosing the appropriate training datasets from injured users is important to reduce noise in the data and accurately capture their gait patterns. As shown in Fig. 2, the accuracy of the estimator in predicting gait speed changes differs based on the novel data used for customization, so these data must be chosen carefully. This choice may be increasingly difficult to make as the number of trials to consider increases.

One way to make this choice is to consider the Mutual Information (MI) between the measured variables (gait features) and those to be inferred (desired speed). MI is a measure of the information obtained about one random variable by observing

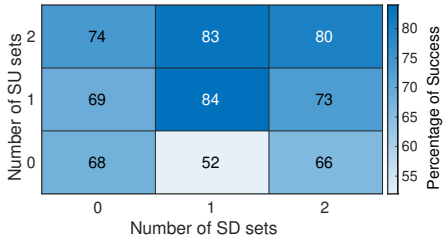

Fig. 2. Estimator performance for a user with SCI for chosen novel data.

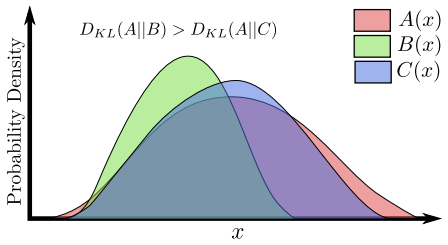

Fig. 3. The difference between two distributions may be quantified using the KL divergence between them.

another variable [40]. The MI between two variables may be computed using the Kullback-Leibler (KL) divergence, D_{KL}, between their joint and marginal distributions. For example, consider two distributions A and B of an arbitrary variable x, then the KL divergence $D_{KL}(A||B)$ is a measure of how different the two distributions are, as illustrated in Fig. 3. The mutual information between the two variables x and y is $I(x;y) = D_{KL}(p_{(x,y)}||p(x)p(y))$ where $p_{(x,y)}$ is their joint probability and $p(x)$ and $p(y)$ are their marginal probabilities. Roughly, the MI provides a scalar measure of the correlation between the two variables.

MI was used to measure the utility of the measured gait features for estimating the desired gait speed. Since the true desired gait speed is difficult to determine, the speed at the next step, denoted v', was assumed as its proxy [27]. The distributions of gait speed and features were approximated as Gaussian and the MI $I(V;Q)$ was computed, where V is the distribution of the step-to-step changes in desired speed estimated via $\Delta v'$, and Q is the distribution of the corresponding changes in gait feature measurements $\Delta\mathbf{Q}$. The intuition behind using distributions of the changes in gait speed and features is to incorporate the knowledge of their evolution through intent changes into the selection process.

To avoid training the estimator on all available speed change data and leave some data for testing, the novel training dataset was limited to at most three out of all available trials for each injured user. Combinations of available novel trial data were generated by choosing two or three out of the available number of trials and collected in a set W. The MI was computed for the novel/base data pairing for each combination in the set, stored in a vector $\boldsymbol{\iota} \in \mathcal{R}^{\text{len}(W)}$. The pairing with the highest MI was chosen. Algorithm 1 details this overall process to select the appropriate novel trial data for augmenting the base data.

C. Estimating the Desired Gait Velocity

Gait features and desired speed were assumed to follow Gaussian distributions during estimation as well. Let the desired gait speed v^d be rewritten as z to simplify notation:

$$\begin{bmatrix}\mathbf{q}\\ z\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\bar{\mathbf{q}}\\ \bar{z}\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{qq}} & \boldsymbol{\Sigma}_{\mathbf{q}z}\\ \boldsymbol{\Sigma}_{z\mathbf{q}} & \Sigma_{zz}\end{bmatrix}\right) \tag{3}$$

where the means and covariances were computed using the training data that includes both base and novel data. Given measurements $\tilde{\mathbf{q}}$ of the gait features, the estimated mean and variance of the desired gait speed were determined using standard conditional probability equations

$$\hat{z} = \bar{z} + \boldsymbol{\Sigma}_{z\mathbf{q}}\boldsymbol{\Sigma}_{\mathbf{qq}}^{-1}(\tilde{\mathbf{q}} - \bar{\mathbf{q}}) \tag{4}$$

$$\hat{\Sigma}_{zz} = \Sigma_{zz} - \boldsymbol{\Sigma}_{z\mathbf{q}}\boldsymbol{\Sigma}_{\mathbf{qq}}^{-1}\boldsymbol{\Sigma}_{\mathbf{q}z} \tag{5}$$

The estimate $\hat{z}$ is driven by the error between the training mean $\bar{\mathbf{q}}$ and gait feature measurements $\tilde{\mathbf{q}}$. Along with $\hat{z}$, the resulting estimator outputs an SU/SD signal at TD as the difference between $\hat{z}$ at the current and previous TD. A speed change threshold for a SU/SD was determined by recording the standard deviation in step-to-step speed changes observed from three steady-state walking steps from each trial in the training data. This threshold is important due to the minimal detectable change (MDC) or the minimal change in the measured gait speed required to distinguish between a true change and noise. The MDC for people with iSCIs was shown to be around 0.17 m/s [41].

Algorithm 1 Training set selection

Require: $\mathbf{v}'_n$, $\mathbf{v}'_b$, $\mathbf{Q}_n$, $\mathbf{Q}_b$

Note: W denotes a set where each element is a combinations of novel trials to be considered
for $m = 1$ to $\text{len}(W)$ **do**
  $\mathbf{v}'_{n/b} = (\mathbf{v}'_b - \bar{v}'_b)\sigma_{v'_n}\sigma_{v'_b}^{-1} + \bar{v}'_n$
  $\mathbf{v}' \leftarrow [\mathbf{v}'^T_{n/b}\ \mathbf{v}'^T_n]^T$
  for $s = 1$ to S **do**
    $\mathbf{q}_{n/b} = (\mathbf{q}_b - \bar{q}_b)\sigma_n\sigma_b^{-1} + \bar{q}_n$
    $\mathbf{q} \leftarrow [\mathbf{q}^T_{n/b}\ \mathbf{q}^T_n]^T$
  end for
  $\mathbf{Q} = [\mathbf{q}_1\ \mathbf{q}_2 \dots \mathbf{q}_S]$
  $\mathbf{v}' \leftarrow \Delta\mathbf{v}'$
  $\mathbf{Q} \leftarrow \Delta\mathbf{Q}$
  $\begin{bmatrix}Q\\ V\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\bar{\mathbf{q}}\\ \bar{v}'\end{bmatrix}, \begin{bmatrix}\boldsymbol{\Sigma}_{\mathbf{qq}} & \boldsymbol{\Sigma}_{\mathbf{q}v'}\\ \boldsymbol{\Sigma}_{v'\mathbf{q}} & \Sigma_{v'v'}\end{bmatrix}\right)$
  $\iota_m = I(V;Q)$
end for
return m such that $\iota_m = \max(\boldsymbol{\iota})$

III. Results & Discussion

A. Collected Walking Trial Data

Trial data from three uninjured and two injured users was collected as part of a study approved by the IRB of the University of Notre Dame (Protocol 18-04-4650) [33]. One of the injured users, IU-1, had a complete SCI at the middle of the spine (T5) and the second user, IU-2, had an incomplete SCI from the middle to the lower spine (T8 to L2). All users were

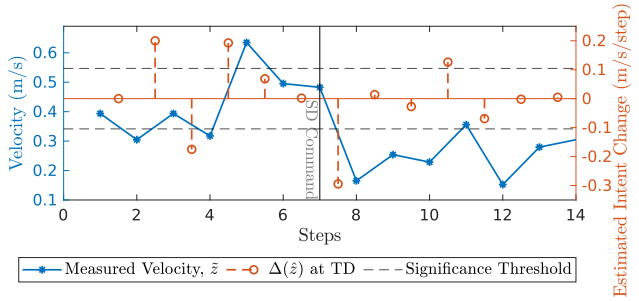

Fig. 4. Output of a single estimator trial for IU-2 using crutches.

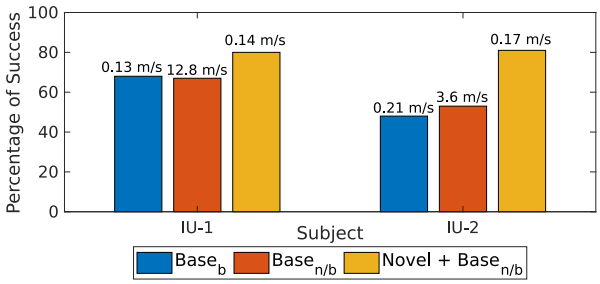

Fig. 5. Percentage of success with and without novel data for both subjects, labeled with RMS error between predicted and measured gait speed changes.

highly experienced in the use of the EksoGT. The subjects used the exoskeleton at a self-selected speed with the assistance of a walker and were at a steady-state gait before being issued a verbal command to either speed up or slow down. The trial sequence was pseudo-random and each subject underwent three SU and SD trials for a total of six trials.

One step before and three steps after the speed change command, for a total of four steps, were chosen as training data from all base and novel trials. For each ambulatory device, i.e., crutches or walker, the base data consisted of 6 trials from each uninjured user in each mode, adaptive and free, for a total of 72 trials. While similar user responses to desired speed change exist for free and adaptive mode, they are evident in different sensor measurements [27]. Despite these differences, more base data was found to result in improved accuracy, even when that data resulted in a mode mismatch between base and novel data. For example, for IU-1, excluding base data from trials in free mode resulted in an estimator accuracy of 69% compared to 80% when both free and adaptive mode data were included in the base set. Up to three trials (i.e., 12 steps' worth of data) were selected as novel data for each injured user using Algo. 1 and used to transform the base data. The estimator was then run on all available trial data for each subject (38-85 steps) out of which at most 12 steps were seen in training. The estimated change in desired speed was compared to the measured change, and if the speed change sign was correctly anticipated, it was considered a successful estimate.

Figure 4 visualizes the output of the estimator for an SD trial with IU-1. The stem plot represents the SU/SD intent signal; a positive value indicates SU and a negative value indicates SD anticipated for the subsequent MS. A significant speed change is expected after the vertical line as it represents the MS closest to when the speed-change command was issued. For an accurate estimate, the value of the signal in the stem plot should be positive for SU and negative for SD after the command is issued. Another metric considered while evaluating estimator performance was the root mean square (RMS) error between the predicted gait speed change at TD and the value measured at the subsequent MS.

The estimator was run in three configurations for both subjects to highlight the benefits of using both novel and base data, as illustrated in Fig. 5. The base data was from walker trials and the novel data was from walker and crutch trials for IU-1 and IU-2, respectively. The first configuration used untransformed base data (Base_b), the second used only the transformed base data ($\text{Base}_\text{n/b}$), and the third used both novel and transformed base data ($\text{Novel+Base}_\text{n/b}$). Minor increases in SU/SD identification accuracy were observed when using only the transformed base data, as the accuracy depends on identifying only the speed changes, and not their magnitude. Despite increases in accuracy, the RMS errors deteriorated and were unacceptable at 12.8 m/s and 3.6 m/s for IU-1 and IU-2 respectively. Adding novel data to the transformed base data increased the speed change estimation accuracy and decreased the RMS errors.

The trials of IU-1 using only base data had an overall success rate of 68%. Upon using novel data, the success rate increased to 80% with a p-value $p = 0.049$ where the null hypothesis was that the success rate would stay the same. Similarly, the success rate for IU-2 increased from 48% to 80% with a p-value $p = 7 \times 10^{-5}$. Therefore, using novel data resulted in statistically significant increases in accuracy.

B. *Efficacy of the Novel Data Selection Algorithm*

The number of possible combinations of novel data in Algo. 1 for training was 35 for walker trials and 12 for trials with crutches for IU-1. There were fewer trials with walking using crutches, as there was data loss from sensors that hindered the identification of gait events. All combinations of these trials were used as novel data with base data from uninjured subjects using a walker to train conditional models that were used in the estimator. The percentages of success of those estimator trials are shown in Fig. 6. The whiskers denote the most extreme points, and the central line denotes the median. The green markers illustrate the success rates observed when the novel data chosen using Algo. 1 was used. If novel data is chosen at random, accuracy may be as low as 59%, however, using the novel data selection algorithm outlined previously ensures a high likelihood of increased success despite not guaranteeing it. Physically, the difference between the results shown in green and the maximum success rate shown by the top whisker would be of at most four misclassified steps. The increase in accuracy seen in Figure 5 after pooling novel and base data further highlights the importance of including user-specific data.

Overall, our analysis showed different outcomes based on the assistive device used in the novel data. This observation motivates the remainder of the analysis herein, which considers the effects of the assistive device on the efficacy of our methods.

C. *Trials Using a Walker*

Estimation for both IU-1 and IU-2 was first performed using base data exclusively from walking trials of uninjured users

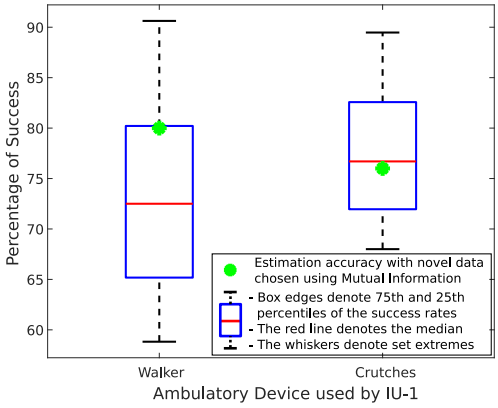

Fig. 6. Percentage of success for IU-1 with the base data from walker trials and the novel dataset chosen randomly vs. using Algo. 1.

TABLE II
CONFUSION MATRIX FOR
NOVEL DATA - WALKER/BASE DATA - WALKER

	Predicted SD		Predicted SU	
	IU-1	IU-2	IU-1	IU-2
Actual SD	80%	58%	19%	78%
Actual SU	20%	42%	81%	22%

using walkers. The performance for these users is summarized by the confusion matrix in Table II. The color of each cell ranges from green to red as the accuracy ranges from 100% to 0% therefore, the higher the accuracy, the greener the cell. The estimator for IU-1 was personalized using one SD and two SU trials. The speed change threshold for this estimator configuration was 0.072 m/s, i.e., if the change in predicted speed at TD compared to the measured speed at the previous MS was smaller than this threshold, the prediction was rejected. Walking trials for IU-1 had 90 steps, 46 having significant speed changes. 80% of SU and 81% of SD changes were accurately detected at TD, for an overall accuracy of 80% with an RMS speed error of 0.14 m/s.

Similarly for IU-2, the change threshold was 0.13 m/s. Out of the 73 total steps, 21 had significant speed changes, and 58% SD and 22% SU changes were accurately detected at TD for an overall accuracy of 43%. The confusion matrix is given in Table II. These trials had higher gait variability as seen from the standard deviation of the steady-state velocity of 0.13 m/s compared to 0.072 m/s for IU-1, which may explain the difficulty in estimation and drop in success rate. Preliminary work suggests other methods to address this low success rate, which are discussed at the end of this section.

D. Trials Using Crutches

The choice of the assistive device affects gait patterns, so estimation was performed for both injured users for walking trials using crutches with the performance summarized in Table III. The base data contained walking trials of uninjured users exclusively using crutches. The speed change threshold was 0.11 m/s. There were 4 significant speed changes in the trials for IU-1 out of which 3 were detected successfully for an overall success rate of 75% with an RMS error of 1.52 m/s. The success rates of SU and SD are shown in Table III.

TABLE III
CONFUSION MATRIX FOR
NOVEL DATA - CRUTCHES/BASE DATA - CRUTCHES

	Predicted SD		Predicted SU	
	IU-1	IU-2	IU-1	IU-2
Actual SD	100%	85%	33%	22%
Actual SU	0%	15%	67%	78%

TABLE IV
CONFUSION MATRIX FOR
NOVEL DATA - CRUTCHES/BASE DATA - WALKER

	Predicted SD		Predicted SU	
	IU-1	IU-2	IU-1	IU-2
Actual SD	78%	78%	25%	17%
Actual SU	22%	22%	75%	83%

The trials for IU-2 had 75 steps, out of which 36 had significant speed changes with a threshold speed of 0.1 m/s. The percentage of success was 81% with 29 speed changes correctly identified and the rates for SU and SD changes are shown in Table III. The RMS error of the speed estimates was 0.2 m/s. A possible explanation for the lower estimator performance for IU-1 is that there may not be enough information about the desired gait speed in the user's gait patterns as evidenced by the amount of MI in the data. The values of MI for the selected novel data for IU-1 and IU-2 were 0.3 and 0.39 respectively. The difference in the ι of the two pairings indicates that the gait feature measurements for IU-2 carried roughly 30% more information about the intended speed than those for IU-1 in this case.

E. Exploring the Interchangeability of Base Data

Interchangeability of base data was studied to explore the effect of ambulatory devices on estimator accuracy by using different devices for the novel and base data. The first pairing was for IU-1 where the novel data was from trials using crutches and base data was from trials using a walker. The overall success rate for this trial, with 17 significant speed changes, was 76% and the RMS error was 0.09 m/s. The confusion matrix for this trial is listed in Table IV. Estimation performed on trial data of IU-2 walking using crutches with base data from trials using a walker yielded a percentage of success of 80% and an RMS error of 0.17 m/s with the confusion matrix also given in Table IV.

These results were compared to estimator trials with IU-1 in which the base data and novel data were both from walking trials with crutches. In this case, the success rate and RMS error were 75% and 1.52 m/s respectively. Surprisingly, this overall success rate was similar and the RMS error was higher than when the base data was from trials with a walker as in the previous paragraph. Further analysis revealed that the task of gait speed estimation was particularly difficult for IU-1 with crutches since 32 of the 38 steps were below the MDC threshold of 0.17 m/s found in literature. In both of the previous cases, the novel data was the same, only the base data was changed. However, pairing the novel data with walker and crutch base data results in MI values of 0.4048 and 0.3007, respectively, so gait features are more informative of gait speed when uninjured users use a walker. As a result, the model

TABLE V
CONFUSION MATRIX FOR
NOVEL DATA - WALKER/BASE DATA - CRUTCHES

	Predicted SD		Predicted SU	
	IU-1	IU-2	IU-1	IU-2
Actual SD	67%	50%	45%	75%
Actual SU	33%	50%	55%	25%

generated using base data from walker trials was able to handle the increased estimation difficulty and increase the estimator accuracy and lower RMS error. This result highlights the flexibility of the estimator to incorporate the most informative base data, even under potentially mismatched conditions.

However, this interchangeability did not hold for every pairing. Estimation was performed for IU-1 with novel data from walking trials with a walker and base data from trials with crutches. Compared to estimation using base data from trials with a walker, the percentage of success dropped to 63% with an RMS error of 0.12 m/s, and the confusion matrix for this trial is given in Table V. There was a less severe deterioration in performance for IU-2. The success rate and RMS error were 38% and 0.26 m/s, respectively; a difference of 5% and 0.046 m/s when compared to the performance of the same novel data paired with data of trials with a walker (see Section III-C).

In general, estimators had higher percentages of success across all tests when using base data of walker trials. This may be as the crutches offer more freedom to move during use than a walker, resulting in more individualized effects on gait patterns across trials. Again, the lower correlation between gait speed and features is evidenced by the MI in the untransformed walker base data (0.3675) being higher than the crutch base data (0.2813), so gait features are more informative of gait speed when uninjured users use a walker. These inconsistencies could possibly be overcome by adding instrumentation such as IMUs [42] to crutches to capture their role in gait dynamics. Additionally, upon expanding the base dataset to include both crutch and walker trials, estimator performance either stayed the same or deteriorated. These results suggest that choosing subsets of the base data along with novel data (e.g., using extensions of Algo. 1) may allow further improvement in estimator performance.

Overall, pooling base and novel data improved estimator accuracy in every case except for trials of IU-2 using a walker. This loss of accuracy was due to noisy measurements during those particular trials. The traces of the model covariance matrices when using novel walker and crutch data were 3.425 and 2.015 respectively (for the same base data). Preliminary work shows that this noise in the IU-2 walker data can be attributed to certain noisy features for this user. Multiple gait features often provide the same information about gait speed while introducing unique noise. In such cases, using a reduced, minimally-redundant set of gait features offers the potential to further improve accuracy. Preliminary results addressing this aspect show increases of over 20% in overall estimation accuracy of the IU-2 trials shown in Tables II and V.

F. Limitations

The training process assumed that the velocity at the next step was a reliable proxy for the user's desired speed at the current step. This approach is likely a worse approximation in adaptive mode than in free mode due to the effects of human-robot coupling. However, it may still accurately capture the direction (SU/SD) of the desired speed change, which is the goal of this work.

The number of possible novel/base data pairings makes it difficult to manually choose a pairing to customize the estimator. The presented method automates that choice but it does not provide any guarantee that the selected pairing is the best possible option. This observation is further supported by the study by Moolchandani et al., [16] where estimation performed with an unoptimized feature set had marginally lower error than when an optimized set was used.

Data from walking trials of only two subjects with iSCIs were used to evaluate this data selection method. These were experienced users and that affects the pHRI during exoskeleton use, as their familiarity with the device may allow them to better predict device behavior and convey their intent more reliably than a novice user. Therefore, it would be beneficial in the future to acquire data from additional trials of both injured and uninjured users with varying degrees of usage proficiency to expand both the novel and base datasets.

Finally, it is noted that the presented results were all obtained in offline evaluation. Assessing methods for integrating the estimated intent into control is an interesting next step, the details of which will have coupling with the intent signals present for real-time estimation.

IV. CONCLUSION & FUTURE WORK

This letter presented a method for personalized estimation of the intended gait speed of a novel exoskeleton user. Data scarcity is addressed by using a small amount of user-specific data to transform an easily accessible base dataset with data from walking trials of uninjured users. This method relies on commonalities in gait patterns observed across subjects and considers 18 gait features to estimate the user's desire to change speeds. Conditional Gaussians were used to construct an estimate of the desired speed, which was then used to infer SU/SD changes. In the future, this method may be extended to use the estimated magnitude of the speed change, though additional work may be required to improve the metric quality of the speed estimates. Human limitations on speed change perceptions would set a lower bound of roughly 0.2 m/s for the accuracy that would be practically noticeable [43].

It was observed that appropriate novel data must be used in training the conditional models for the estimator due to gait variability present across trials. This selection served to capture gait patterns accurately while avoiding forming misleading relationships resulting from noisy trials. However, it is difficult to guarantee that the chosen novel data will result in the best possible estimator performance. It was also found that walker base data led to uniformly better estimator performance over the use of crutch base data, even when the novel data was for user trials with crutches. This result shows that the selection of appropriate base data needs to be considered to ensure optimal estimator performance.

The new data selection algorithm suggests future work to address other configuration options of the data pooling and training process. However, leaving the choice of both base and novel datasets free for selection results in a combinatorial challenge. New methods to address this scaling challenge would be needed. Selecting subsets of the gait features could also be assessed to further improve estimator performance.

REFERENCES

[1] "Spinal cord injury facts and figures at a glance." [Online]. Available: https://www.nscisc.uab.edu/PublicDocuments/fact_figures_docs/Facts%202012%20Feb%20Final.pdf

[2] M. DeVivo, Y. Chen, S. Mennemeyer, and A. Deutsch, "Costs of care following spinal cord injury," *Topics in spinal cord injury rehabilitation*, vol. 16, no. 4, pp. 1–9, 2011.

[3] A. Curt, H. J. Van Hedel, D. Klaus, V. Dietz, and E.-S. study group, "Recovery from a spinal cord injury: significance of compensation, neural plasticity, and repair," *Journal of neurotrauma*, vol. 25, no. 6, pp. 677–685, 2008.

[4] J. Hidler and R. Sainburg, "Role of robotics in neurorehabilitation," *Topics in Spinal Cord Injury Rehab.*, vol. 17, no. 1, pp. 42–49, 2011.

[5] L. Brenner, "Exploring the psychosocial impact of ekso bionics technology," *Arch. of Phys. Med. and Rehab.*, vol. 97, no. 10, p. e113, 2016.

[6] F. Sup, A. Bohara, and M. Goldfarb, "Design and control of a powered transfemoral prosthesis," *The International Journal of Robotics Research*, vol. 27, no. 2, pp. 263–273, 2008.

[7] "Rewalk personal exoskeleton system cleared by FDA for home use," https://ir.rewalk.com/news-releases/news-release-details/rewalktm-personal-exoskeleton-system-cleared-fda-home-use.

[8] G. Hoffman and C. Breazeal, "Cost-based anticipatory action selection for human–robot fluency," *IEEE Transactions on Robotics*, vol. 23, no. 5, pp. 952–961, 2007.

[9] G. Hoffman, "Evaluating fluency in human–robot collaboration," *IEEE Trans. on Human-Machine Systems*, vol. 49, no. 3, pp. 209–218, 2019.

[10] K. Suzuki, G. Mito, H. Kawamoto, Y. Hasegawa, and Y. Sankai, "Intention-based walking support for paraplegia patients with robot suit HAL," *Advanced Robotics*, vol. 21, no. 12, pp. 1441–1469, 2007.

[11] A. Goffer and C. Zilberstein, "Locomotion assisting device and method," Jan. 17 2012, uS Patent 8,096,965.

[12] H. Geyer, A. Seyfarth, and R. Blickhan, "Compliant leg behaviour explains basic dynamics of walking and running," *Proc. Royal Society B: Biological Sciences*, vol. 273, no. 1603, pp. 2861–2867, 2006.

[13] Y. Liu, P. M. Wensing, D. E. Orin, and Y. F. Zheng, "Dynamic walking in a humanoid robot based on a 3D actuated Dual-SLIP model," in *IEEE International Conference on Robotics and Automation*, 2015.

[14] U. H. Lee, J. Bi, R. Patel, D. Fouhey, and E. Rouse, "Image transformation and cnns: A strategy for encoding human locomotor intent for autonomous wearable robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5440–5447, 2020.

[15] N. Thatte, T. Shah, and H. Geyer, "Robust and adaptive lower limb prosthesis stance control via extended kalman filter-based gait phase estimation," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3129–3136, 2019.

[16] P. R. Moolchandani, A. Mazumdar, and A. J. Young, "Design of an intent recognition system for dynamic, rapid motions in unstructured environments," *ASME Letters in Dynamic Systems and Control*, vol. 2, no. 1, p. 011004, 2021.

[17] T. Gambon, J. P. Schmiedeler, and P. M. Wensing, "User intent identification in a lower-extremity exoskeleton via the mahalanobis distance," in *IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, 2020, pp. 1115–1121.

[18] W. J. Sohn, A. Q. Tan, H. B. Hayes, S. Pochiraju, J. Deffeyes, and R. D. Trumbower, "Variability of leg kinematics during overground walking in persons with chronic incomplete spinal cord injury," *Journal of neurotrauma*, vol. 35, no. 21, pp. 2519–2529, 2018.

[19] P. Krawetz and P. Nance, "Gait analysis of spinal cord injured subjects: effects of injury level and spasticity," *Archives of Physical Medicine and Rehabilitation*, vol. 77, no. 7, pp. 635–638, 1996.

[20] M. Tucker, E. Novoseller, C. Kann, Y. Sui, Y. Yue, J. W. Burdick, and A. D. Ames, "Preference-based learning for exoskeleton gait optimization," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 2351–2357.

[21] L. Kilmartin, R. K. Ibrahim, E. Ambikairajah, and B. Celler, "Optimising recognition rates for subject independent gait pattern classification," in *ISSC*. IET, 2009.

[22] R. K. Ibrahim, E. Ambikairajah, B. Celler, N. H. Lovell, and L. Kilmartin, "Gait patterns classification using spectral features," in *IET Irish Signals and Systems Conference*, 2008, pp. 98–102.

[23] A. J. Young, A. M. Simon, N. P. Fey, and L. J. Hargrove, "Classifying the intent of novel users during human locomotion using powered lower limb prostheses," in *International IEEE/EMBS Conference on Neural Engineering*, 2013, pp. 311–314.

[24] A. J. Young and L. J. Hargrove, "A classification method for user-independent intent recognition for transfemoral amputees using powered lower limb prostheses," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 2, pp. 217–225, 2015.

[25] K. Bhakta, J. Camargo, L. Donovan, K. Herrin, and A. Young, "Machine learning model comparisons of user independent & dependent intent recognition systems for powered prostheses," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5393–5400, 2020.

[26] T. K. Best, K. Embry, E. J. Rouse, and R. D. Gregg, "Phase-variable control of a powered knee-ankle prosthesis over continuously varying speeds and inclines," in *IEEE/RSJ Int. Conf. Intell. Robots Syst*, 2021.

[27] R. M. Karulkar and P. M. Wensing, "Using footsteps to estimate changes in the desired gait speed of an exoskeleton user," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 6781–6788, 2021.

[28] B. Hu, E. Rouse, and L. Hargrove, "Benchmark datasets for bilateral lower-limb neuromechanical signals from wearable sensors during unassisted locomotion in able-bodied individuals," *Frontiers in Robotics and AI*, vol. 5, p. 14, 2018.

[29] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones." in *Esann*, vol. 3, 2013, p. 3.

[30] C. A. Fukuchi, R. K. Fukuchi, and M. Duarte, "A public dataset of overground and treadmill walking kinematics and kinetics in healthy individuals," *PeerJ*, vol. 6, p. e4640, 2018.

[31] N. Sylla, V. Bonnet, G. Venture, N. Armande, and P. Fraisse, "Assessing neuromuscular mechanisms in human-exoskeleton interaction," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2014, pp. 1210–1213.

[32] T. M. Gambon, J. P. Schmiedeler, and P. M. Wensing, "Characterizing intent changes in exoskeleton-assisted walking through onboard sensors," in *ICORR*, 2019, pp. 471–476.

[33] ——, "Effects of user intent changes on onboard sensor measurements during exoskeleton-assisted walking," *IEEE Access*, 2020.

[34] V. Rota, L. Perucca, A. Simone, and L. Tesio, "Walk ratio (step length/cadence) as a summary index of neuromotor control of gait: application to multiple sclerosis," *International journal of rehabilitation research*, vol. 34, no. 3, pp. 265–269, 2011.

[35] D. Tkach, H. Huang, and T. A. Kuiken, "Study of stability of time-domain features for electromyographic pattern recognition," *Journal of neuroengineering and rehabilitation*, vol. 7, no. 1, pp. 1–13, 2010.

[36] M. Ison and P. Artemiadis, "The role of muscle synergies in myoelectric control: trends and challenges for simultaneous multifunction control," *Journal of Neural Eng.*, vol. 11, no. 5, p. 051001, 2014.

[37] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[38] D. Austin, T. L. Hayes, J. Kaye, N. Mattek, and M. Pavela, "On the disambiguation of passively measured in-home gait velocities from multi-person smart homes," *Journal of ambient intelligence and smart environments*, vol. 3, no. 2, pp. 165–174, 2011.

[39] M. J. Socie, R. W. Motl, J. H. Pula, B. M. Sandroff, and J. J. Sosnoff, "Gait variability and disability in multiple sclerosis," *Gait & posture*, vol. 38, no. 1, pp. 51–55, 2013.

[40] T. M. Cover, *Elements of information theory*. John Wiley & Sons, 1999.

[41] P. Mohandas Nair, T. George Hornby, and A. Louis Behrman, "Minimal detectable change for spatial and temporal measurements of gait after incomplete spinal cord injury," *Topics in spinal cord injury rehabilitation*, vol. 18, no. 3, pp. 273–281, 2012.

[42] D. Brescianini, J.-Y. Jung, I.-H. Jang, H. S. Park, and R. Riener, "INS/EKF-based stride length, height and direction intent detection for walking assistance robots," in *IEEE International Conference on Rehabilitation Robotics*, 2011, pp. 1–5.

[43] F. Zhang, M. Liu, and H. Huang, "Investigation of timing to switch control mode in powered knee prostheses during task transitions," *PLOS one*, vol. 10, no. 7, p. e0133965, 2015.