A Data Driven Approach for Predicting Preferred Ankle Stiffness of a Quasi-Passive Prosthesis

Varun S. Shetty, Ung Hee Lee, *Student Member, IEEE*, Kimberly A. Ingraham, *Student Member, IEEE*, and Elliott J. Rouse, *Senior Member, IEEE*

***Abstract*—Emerging variable-stiffness ankle prostheses can modulate their stiffness to meet differing biomechanical demands. To this end, knowledge of the optimal ankle stiffness is required for each user and activity. One approach is to match the stiffness of prosthesis to the user's preference, but this requires a tuning process to determine each user's preferences. In this work, we seek to estimate user-preferred ankle stiffness using biomechanical data collected from seven subjects during walking at stiffness settings around their preferred stiffness; our hope is an automated method may reduce the time and experimental burden of determining user preferences. We investigated different machine learning algorithms, sensor subsets, and the impact of user-specific training data on estimation accuracy. We found that a long short term memory (LSTM) algorithm trained on user-specific data from only the affected side, were able to predict user preferred ankle stiffness with an RMSE of 5.2% ± 0.3%. The prediction error was less than prosthesis users' ability to reliably sense stiffness changes (7.7%), which highlights the significance of the performance of our proposed method. This study provides a foundation for an automated approach for predicting user-preferred prosthesis mechanics that would ease the burden of tuning these systems in a clinical setting.**

***Index Terms*—Control, deep learning method, prosthetics and exoskeletons, user preference, wearable robotics.**

I. Introduction

ROBOTIC assistive technologies—such as exoskeletons and robotic prostheses—hold the potential to transform the mobility of people with disabilities [1]–[3]. These technologies provide mechanical assistance to their wearer that is able to help offset the biomechanical deficits associated with neuromotor pathologies (*e.g.* limb loss). For example, when using conventional, passive ankle-foot prostheses, people with below-knee amputations walk slower [4], are less stable [5], and fatigue more quickly [6]. Over the past decades, researchers have developed advanced prostheses that seek to address these deficits [1], [2], [7]. Recently developed powered prostheses can have challenges with cost, size, weight, and robustness. Consequently, many researchers have pursued a different approach—termed *quasi-passive* prostheses—that use a small motor and battery to adjust the system's passive mechanical properties [8], [9]. These systems are often smaller, lighter, and lower cost when compared to fully powered versions. Thus, quasi-passive prostheses attempt to strike a balance between the function provided by more sophisticated mechanisms and the challenges associated with the increased mass and complexity.

A promising and emerging type of quasi-passive prostheses can continuously vary their stiffness from step to step. Different tasks, such as walking or stair ascent, have different kinetic and kinematic demands that cannot be achieved with conventional, static prostheses. Variable-stiffness prostheses are able to more closely match the demands of different activities while being able to store and return energy to the wearer. Recent approaches to develop variable-stiffness prostheses have focused on the implementation of variable-length leaf springs [8], [9] and shear thickening fluids [10]. These systems often have a single variable (*e.g.* stiffness) that must be properly "tuned" or adjusted to match the desired kinetic and kinematic patterns.

An open and important research question pertains to how to effectively set the prosthesis stiffness across users and activities. Researchers have proposed to tune the mechanics of assistive technologies to meet specific physiological or biomechanical objectives. A common physiological objective is to reduce the metabolic rate of the wearer. This approach is often combined with an automatic tuning process, a strategy known as "human-in-the-loop" optimization [11]. However, in actuality, prosthesis users likely have shifting objectives depending on the activity or task of the user. These objectives may be influenced by quantitative and qualitative factors (*e.g.* stability, comfort, *etc.*), which can be difficult or impossible to measure. Instead, we propose to leverage user preference as a holistic objective that is able to capture the quantitative and qualitative attributes of user experience. That is, by including the user's perception of their experience—their preferences—we are able to use the user as both the sensor and optimizer to determine the stiffness that maximizes their experience.

Researchers have begun to investigate the role of user preference in the control of assistive technologies [12]–[15]. Tucker *et al.* developed an online tuning process to optimize user preference based on Bayesian methods [16]. Their approach used forced choice experiments with co-active feedback to tune the step trajectory parameters of a lower-limb exoskeleton. Some groups have used a self-tuning process to identify user

Manuscript received September 9, 2021; accepted January 2, 2022. Date of publication January 25, 2022; date of current version February 10, 2022. This letter was recommended for publication by Associate Editor L. D. Michieli and Editor P. Valdastri upon evaluation of the reviewers' comments. This work was supported by the U.S. Army CDMRP PRORP under Grant W81XWH-17-1-0704. *(Corresponding author: Elliott J. Rouse.)*

The authors are with the Department of Mechanical Engineering and Robotics Institute, University of Michigan, Ann Arbor, MI 48109 USA, and also with Neurobionics Lab, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: vashetty@umich.edu; unghee@umich.edu; kaingr@umich.edu; ejrouse@umich.edu).

Digital Object Identifier 10.1109/LRA.2022.3144790

preference, wherein instead of using a forced choice method, users adjust the control parameters to obtain their preferred setting [12], [17]. Our group has also recently quantified the stiffness preferences of both people with amputations and the clinicians that prescribe prostheses [18]. Our work showed that prosthesis users have diverse yet consistent preferences. In addition, we recently demonstrated that the preferred stiffness of prosthesis users maximized their kinematic symmetry at the ankle joint [17]. However, identifying user preference typically involves a self-tuning experimental session that can require specialized engineers to assist, in addition to a lengthy tuning session. In addition, manual preference-based tuning will not likely scale to strategies with a greater number of parameters that must be tuned.

To reduce the experimental demands of determining user preference, we propose to predict user preferences rather than obtaining these data experimentally. Our overarching hypothesis is that there are consistent features that underlie user preference, which can be used to develop a machine learning model able to predict user preferences from biomechanical data. Our hypothesis is supported by common biomechanical trends that emerge across users when walking on a prosthesis with their preferred ankle stiffness. In this study, 1) we implemented and compared the predictions of user preferred prosthesis ankle stiffness using two classical machine learning (ML) approaches and three deep learning (DL) approaches, 2) we analysed and compared the affects of including subject specific data on the predictions of user's preferred prosthesis ankle stiffness, and 3) we analysed and compared three groups of biomechanical signals to understand the effect of data amount and type on estimation accuracy. We trained the models on a dataset of seven subjects who walked on our variable stiffness ankle prosthesis at stiffness settings around their preferred stiffness [17]. Our work demonstrates that biomechanical data can be used with deep learning models to accurately predict user preferences, but some subject-specific data is likely needed. The intent of our work is to introduce a new method to identify user's preferred stiffness and to take the first step toward developing a preference based tuning approach that is more feasible and convenient to incorporate in a clinical setting.

II. Methodology

A. VSPA Foot

The VSPA Foot is a quasi-passive ankle-foot prosthesis that is able to change the stiffness of its ankle joint from step to step [9]. The mechanism uses a variable length leaf spring coupled to a cam-based transmission (Fig. 1). The stiffness is adjusted using a small motor that re-positions the support condition of the leaf spring, changing its effective length. The stiffness can be adjusted during the swing phase of gait, when the ankle is unloaded. The mass of the prototype is approximately 1900 g, and it is able to vary stiffness by nearly an order of magnitude. The mechanism is mechanically passive, but uses a small battery to power an electric motor and embedded system.

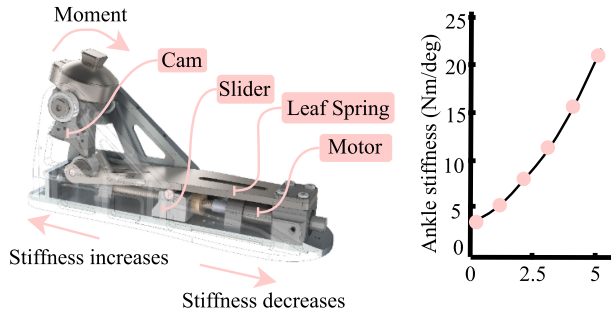


Fig. 1. A schematic of the VSPA Foot, shown with relevant components highlighted (left), and the range of available dorsiflexion stiffness values possible by adjusting the sliding support (right). For more details on the design, please see [9].

B. Experimental Data Collection

The dataset used in this study was collected from seven individuals with unilateral transtibial amputation walking with the VSPA Foot prototype. The experimental protocol was divided into two parts: 1) preferred ankle stiffness identification, and 2) lower extremity biomechanical data collection. During the preferred ankle stiffness identification procedure, participants donned the VSPA Foot before they walked on the treadmill. They were taught to use a mechanical dial to personally customize the stiffness of the VSPA Foot to identify their *preferred stiffness*. Within a single preference-identification trial, participants were encouraged to explore the full range of stiffness values from "uncomfortably stiff" to "uncomfortably soft" before determining the stiffness they preferred. In the experimental protocol, three speeds were investigated :subject's preferred speed (mean 1.27 m/s, standard deviation 0.1 m/s) and $\pm 30\%$ subject's preferred speed; however, they were grouped for the analysis in this work. Thus, participants performed three preference-identification trials at each of three treadmill speeds, for a total of nine preference-identification trials per subject. For a given treadmill speed, the participant's preferred stiffness was calculated as the average preferred stiffness across the three trials. Participants were consistent in selecting their preferred stiffness across walking speeds with the mean coefficient of variance of $9.51 \pm 5.67\%$ across the three speeds and seven subjects.

After the preferred ankle stiffness identification procedure, the lower extremity biomechanical data collection was conducted. Participants walked on the treadmill with the VSPA Foot stiffness set to five different values ($70\%, 85\%, 100\%, 115\%, 130\%$ of their preferred stiffness) at each of three treadmill speeds (preferred speed, $\pm 30\%$ preferred speed). These data were originally collected as part of [17], which includes more details on the experimental protocol. For each treadmill speed, all five stiffness values were tested in a single, consecutive block, with participants continuously walking for 3-4 minutes with a particular stiffness value. While subjects walked on the treadmill, lower-limb biomechanical data were collected using optical motion capture (Vicon Motion Systems Ltd, Oxford, U.K.). 36 reflective spherical markers were attached to the subjects pelvis, legs and feet according to the modified Helen Hayes

model [19]. Kinematics were recorded at a sampling rate of 100 Hz using a 17-camera motion capture system. Markers were automatically labelled using the Nexus software. Ground reaction force data were collected using the force plates embedded in the instrumented treadmill (Bertec, Columbus, OH). During the data collection, approved consent was obtained from all subjects prior to the study with approval from the U-M Medical School Institutional Review Board.

C. Dataset

All biomechanical data were post-processed using OpenSim (Stanford University, Palo Alto, CA) and were scaled on the subject's anthropometry. Inverse kinematic and dynamic analyses were performed using the "*gait10dof18musc*" model in OpenSim v4.0. In total, our dataset comprised 93 biomechanical signals. These signals were grouped into four categories: inverse kinematics (IK), inverse dynamics (ID), ground reaction forces/center of pressure (GRF/COP), and segment kinematics (SK). The IK signals (10 signals) were bilateral 3-DOF hip angle, sagittal knee angle, and sagittal ankle angle. The ID signals (10 signals) were bilateral 3-DOF hip moment, sagittal knee moment, and sagittal ankle moment. The GRF/COP signals (10 signals) were bilateral 3-DOF GRF, and anterior/posterior and mediolateral COP. The SK signals (63 signals) were 3-DOF position, 3-DOF velocity, and 3-DOF acceleration of the pelvis and bilateral foot, shank, and thigh body segments.

We segmented the data into strides, defined from left heel contact to subsequent left heel contact; one stride contained data from all the signals, each interpolated from 0 to 100% gait cycle (101 data points). Pooled across all subjects and speeds, the final dataset comprises 12080 strides (*i.e* samples). Each subject had different number of strides taken throughout the experiment, with an average of 1725 $\pm$ 306 strides.

To evaluate how training with different biomechanical signals affects prediction performance, we divided the signals into three subsets (Table I). The three signal subsets were: 1) All Data, which included all collected signals (93 signals), 2) Affected Side Data, which only included the signals from the user's affected (*i.e.* prosthetic) side, excluding pelvis (42 signals), 3) Affected Side Foot/Shank Data, which only included the signals from the user's affected side foot and shank (25 signals).

TABLE I
BIOMECHANICAL SIGNAL SUBSETS

Subset	Signals	Body Part	# of Signals
All	IK	Bilateral Hip, Knee, Ankle	10
	ID	Bilateral Hip, Knee, Ankle	10
	GRF/COP	Bilateral	10
	SK	Bilateral Foot, Shank, Thigh & Pelvis	63
Total			$S = 93$
Affected Side	IK	Affected Hip, Knee, Ankle	5
	ID	Affected Hip, Knee, Ankle	5
	GRF/COP	Affected Side	5
	SK	Affected Foot, Shank, Thigh	27
Total			$S = 42$
Affected Side Foot/Shank	IK	Affected Ankle	1
	ID	Affected Ankle	1
	GRF/COP	Affected Side	5
	SK	Affected Foot, Shank	18
Total			$S = 25$

We evaluated three signal subsets: All Data, Affected Side Data, and Affected Side Foot/Shank Data. Each signal subset included signals from inverse kinematics (IK), inverse dynamics (ID), ground reaction force (GRF), center of pressure (COP), and segment kinematics (SK). S denotes the total number of signals in each signal subset.

D. Data Prepraration

1) Data Preprocessing: We preprocessed the data to mitigate the effects of noise and irregularities in the dataset and to improve the training efficiency and predictive performance of our algorithms. We performed two data preprocessing steps: a) Min-Max Normalization and b) Feature Extraction.

a) Min-Max normalization: For both classical machine learning (ML) and deep learning (DL) algorithms, we performed Min-Max Normalization to scale the different biomechanical signals in our dataset: $x' = (x - x_{min})/(x_{max} - x_{min})$. For each stride of the biomechanical signals, we normalized each data point (x) with respect to the minimum (x_{min}) and maximum (x_{max}) observed values to generate the scaled data point (x').

b) Feature extraction: For the classical ML models only, we performed feature extraction to remove noise and reduce the model training time. For all biomechanical signals, we extracted six features from each stride: mean, standard deviation, maximum value, minimum value, first value and last value [20].

2) Data Splitting: For both classical ML and DL algorithms, we divided the dataset into training, validation, and testing sets in two ways a) Subject Independent and b) Subject Dependent.

a) Subject independent: In the Subject Independent case, the machine learning models had no information about the test subject during the training process. In this case, we utilized a leave-one-subject-out cross validation. The training and validation sets included data from six subjects. Approximately 76% (mean 75.77%, standard deviation 2.42%) of all strides were included in the training set, while 10% were withheld as a validation set. The seventh subject's data were used as the testing set (approximately 14% of all strides). This process was repeated seven times, until all subjects had been left out.

b) Subject dependent: In the Subject Dependent case, the machine learning models had some information about all the subjects during the training process. In this case, we performed a 7-fold cross validation. For each fold, we split the dataset into training (76% of all strides), validation (10%) and testing (14%) sets. These percentages were chosen to mirror the distribution of data in the Subject Independent case.

E. Machine Learning Algorithms

We implemented two classical machine learning (ML) and three deep learning (DL) algorithms to predict a participant's preferred prosthetic ankle stiffness from biomechanical signals. For all algorithms, we used the validation set to tune the hyperparameters of each model prior to training. We performed a grid search for the hyperparameter tuning, with the objective function

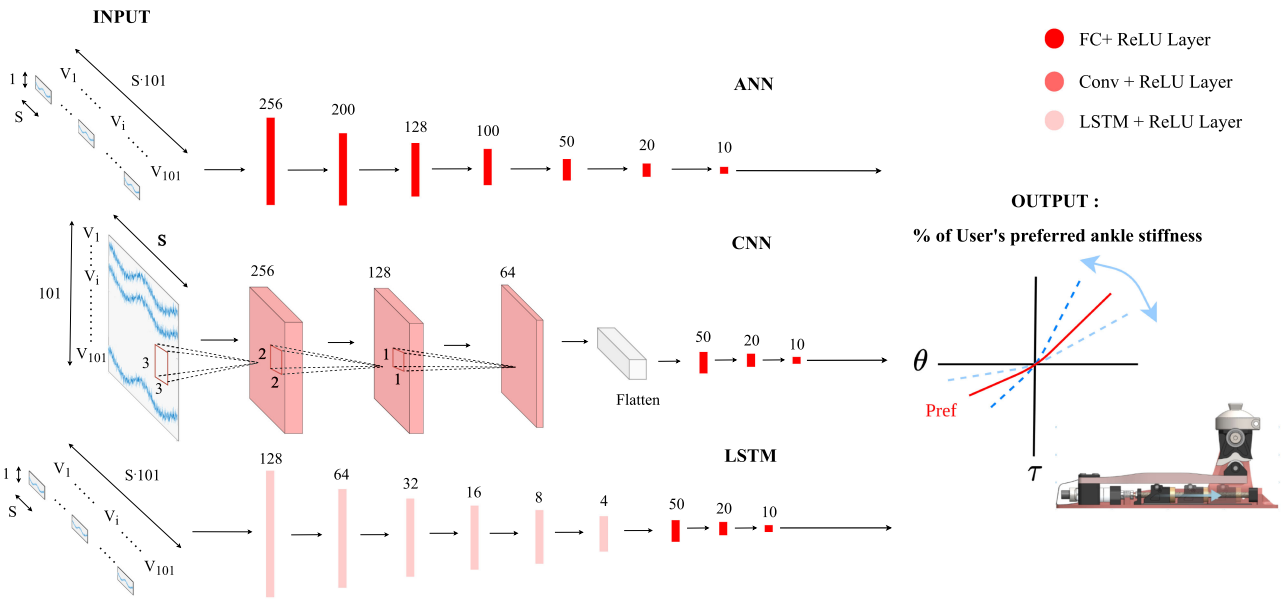


Fig. 2. The depiction of testing our three deep learning algorithms. The input sample is one stride (*i.e.* gait cycle) of data, which includes the S biomechanical signals of the signal subset (see Table I for the value of S for each signal subset). For the ANN and LSTM algorithms, the input sample size is $1 \times (S \cdot 101)$. For the CNN algorithm, the input data is reshaped to the size $101 \times S$. The output for all three networks is the VSPA Foot stiffness, expressed as a percentage of the user's preferred ankle stiffness. V_i (dimensions: $1 \times S$) is a vector that contains the i^{th} $(1 \leq i \leq 101)$ measurement of the gait cycle, concatenated for the S biomechanical signals.

of minimizing the root mean square error (RMSE) between the predicted preferred VSPA Foot stiffness and the ground truth preferred VSPA Foot stiffness. For hyperparameter tuning, we used the entire dataset (All Data) and the Subject Dependent data splitting method.

1) Classical Machine Learning Algorithms: We implemented two classical ML algorithms using Python's *scikit-learn* package [21]: a) K-Nearest Neighbor (KNN) and b) Support Vector Regression (SVR). For the classical ML algorithms, the training data were the six features extracted from each stride of the training set. The training data for the classical ML algorithms had dimensions of $N \times (S \cdot 6)$, where "$\times$" refers to the increase in dimensionality and "$\cdot$" represents scalar multiplication within the same dimension, N is the total number of strides in the training set, and S is the number of signals in the biomechanical signal subset (see Table I). The training labels (dimension: $N \times 1$) were the corresponding ground truth, which is VSPA Foot stiffness expressed as a percentage of the user's preferred stiffness (%PS): 70%, 85%, 100%, 115%, 130%. To test the classical ML algorithms, we used each stride of the testing set (dimension: $1 \times (S \cdot 6)$) as an input to the trained model to predict the corresponding %PS for that stride. This testing method was selected to evaluate each model's performance when using the minimum amount of data collected from one individual (*i.e.,* one stride).

a) K-Nearest neighbor: The KNN algorithm makes predictions based on the similarity of the features in the dataset. The KNN algorithm stores the training data, and the prediction is made based on the weighted average of the distance function for the K nearest points. We selected the Euclidean distance function and a neighbors value of $K = 5$.

b) Support vector regression: The SVR algorithm makes predictions by first mapping the training dataset to a higher dimension based on the kernel, and then finding a hyperplane with optimal margins in the higher dimension [22]. We selected a third degree polynomial kernel, the regularization parameter was 10^3, the error sensitivity parameter (ϵ) was 0.1, and the spread of the kernel (γ) was defined as $1/(number\,of\,features \times training\,data\,variance)$.

2) Deep Learning Algorithms: We implemented three DL models: a) Artificial Neural Network (ANN), b) Convolutional Neural Network (CNN), and c) Long Short Term Memory (LSTM) Network. We used the mean squared error loss function and a stochastic gradient-based optimizer, ADAM [23], to train our DL models. We added L1 and L2 regularization in addition to the loss function to improve the generalizability of the networks. All calculations were performed using TensorFlow [24] and Keras packages [25].

For the DL algorithms, the training data were the time-series data (101 data points) of each stride of the training set. For the ANN and LSTM algorithms, the training data had dimensions of $N \times (S \cdot 101)$, where N is the total number of strides in the training set, and S is the number of signals in the biomechanical signal subset (see Table I). For the CNN algorithm, the training data was transformed to be 3-dimensional, with dimensions of $N \times 101 \times S$. The training labels (dimension: $N \times 1$) were the corresponding ground truth %PS. To test the DL algorithms, we used each stride of the testing set as an input to the trained model to predict the corresponding %PS for that stride. For the ANN and LSTM algorithms, the test data had dimensions of $1 \times (S \cdot 101)$; for the CNN algorithm, the test data had dimensions of $1 \times 101 \times S$. The following describes architecture design of the three algorithms:

a) Artificial neural network: The ANN architecture comprised 7 fully-connected dense layers with 256, 200, 128, 100, 50, 20, and 10 units (Fig. 2). Rectified Linear Units (ReLU) were

TABLE II
MEAN ERROR (% USER'S PREFERRED STIFFNESS)

Machine Learning Algorithms	All data		Affected side data		Affected side Foot and Shank data	
	Sub. Dep.	Sub. Ind.	Sub. Dep.	Sub. Ind.	Sub. Dep.	Sub. Ind.
KNN	-0.3 ± 12.3	2.1 ± 22.0	-0.5 ±11.8	4.1 ± 20.5	0.1 ± 13.1	4.3 ± 20.4
SVR	0.2 ± 6.9	4.5 ± 18.1	0.2 ± 7.9	12.9 ± 19.0	0.2 ± 8.9	13.5 ± 19.7
ANN	0.3 ± 5.5	8.5 ± 16.5	-0.8 ± 5.6	6.5 ± 16.9	-0.9 ± 6.0	-3.8 ± 17.8
CNN	0.3 ± 6.2	8.6 ± 17.1	-0.4 ± 5.8	11.0 ± 14.6	1.2 ± 6.3	8.2 ± 14.9
LSTM	-0.1 ± 5.7	5.3 ± 17.8	0.2 ± 5.2	7.1 ± 16.5	-0.1 ± 5.7	4.1 ± 16.0

The mean and standard deviation (mean ± SD) of the error were calculated across 7 folds of the cross validation. Results are shown for the five machine learning algorithms: K-Nearest Neighbor (KNN), Support Vector Regression (SVR), Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Long Short Term Memory Network (LSTM). Each algorithm was evaluated using three signal subsets: All Data, Affected Side Data, and Affected Side Foot/Shank Data (see Table I), and two data splitting methods: Subject Dependent (Sub. Dep.) and Subject Independent (Sub. Ind.).

TABLE III
ROOT MEAN SQUARE ERROR (RMSE) (% USER'S PREFERRED STIFFNESS)

Machine Learning Algorithms	All Data		Affected Side Data		Affected Side Foot/Shank Data	
	Sub. Dep.	Sub. Ind.	Sub. Dep.	Sub. Ind.	Sub. Dep.	Sub. Ind.
KNN	12.3 ± 0.2	24.5 ± 2.0	11.8 ±0.2	24.3 ± 4.1	13.1 ± 0.1	24.2 ± 5.7
SVR	6.9 ± 0.2	26.0 ± 12.0	7.9 ± 0.5	25.5 ± 10.4	8.9 ± 0.4	28.1 ± 13.6
ANN	5.8 ± 0.5	20.4 ± 6.7	5.7 ± 0.3	21.3 ± 6.7	6.1 ± 0.2	23.0 ± 9.1
CNN	6.3 ± 0.2	23.3 ± 6.4	5.9 ± 0.4	19.7 ± 6.4	6.5 ± 0.2	21.7 ± 8.6
LSTM	5.7 ± 0.4	22.6 ± 6.3	**5.2 ± 0.3**	20.5 ± 6.4	5.7 ± 0.3	20.9 ± 4.7

The mean and standard deviation (mean ± SD) of the root mean square error (RMSE) were calculated across 7 folds of the cross validation. Results are shown for the five machine learning algorithms: K-Nearest Neighbor (KNN), Support Vector Regression (SVR), Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Long Short Term Memory Network (LSTM). Each algorithm was evaluated using three signal subsets: All Data, Affected Side Data, and Affected Side Foot/Shank Data (see Table I), and two data splitting methods: Subject Dependent (Sub. Dep.) and Subject Independent (Sub. Ind.). A Bold number represent the best performing (lowest RMSE) ML algorithm across all signal subsets and data splitting methods.

used as the activation function to account for nonlinearity. We selected a batch size of 128 and used 300 epochs. The learning rate was 10^{-3}, the decay rate was 10^{-5}, the L1 regularization parameter was 10^{-4} and the L2 regularization parameter was 10^{-2}.

b) Convolutional neural network: The CNN architecture consisted of convolutional layers followed by fully connected layers. The CNN architecture had 256, 128, and 64 filters, with kernel sizes of 3x3, 2x2 and 1x1, respectively, with stride of 1 (Fig. 2). ReLU were used as activation functions. We selected a batch size of 128 and used 100 epochs. The learning rate was 10^{-3}, the decay rate was 10^{-5}, the L1 regularization parameter was 10^{-4}, and the L2 regularization parameter was 10^{-3}.

c) Long short term memory network: We implemented the LSTM network to account for the time-dependency of the biomechanical input signals [26]. The LSTM architecture comprised 6 LSTM layers with 128, 64, 32, 16, 8, and 4 units, followed by 3 fully connected layers (Fig. 2). ReLU were used as activation functions. We selected a batch size of 128 and used 400 epochs. The learning rate was 10^{-4}, the decay rate was 10^{-5}, the L1 regularization parameter was 10^{-4}, and the L2 regularization parameter was 10^{-2}.

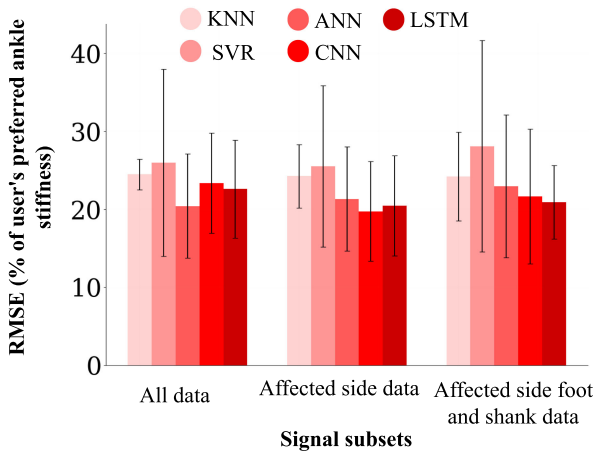
F. Data Analysis and Statistics

We evaluated five machine learning algorithms (KNN, SVR, ANN, CNN, LSTM), three subsets of biomechicanical signals (All Data, Affected Side Data, Affected Side Foot/Shank Data) and two training methods (Subject Dependent, Subject Independent)—in total, we evaluated 30 different models. We evaluated the performance of all models by calculating two types of error: 1) the mean and standard deviation of error, and 2) the root mean square error (RMSE).

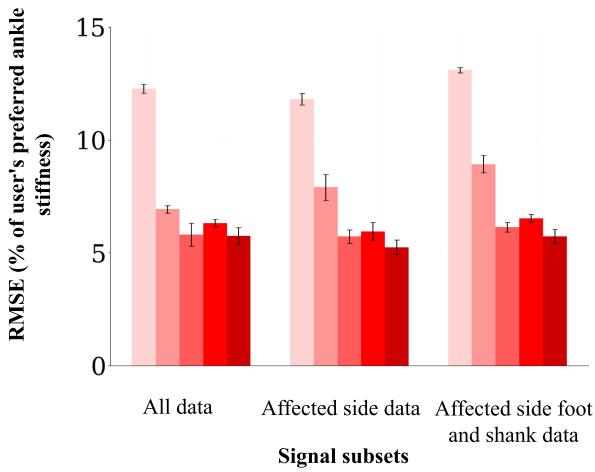
The error was the difference between the predicted %PS and the ground truth %PS. The mean and standard deviation of error were computed across all predicted data points for all folds of the cross validation. We used the mean and standard deviation of the error to examine the distribution of the model errors. The RMSE was calculated for each of the 7 folds of the cross validation, and the mean and standard deviation of RMSE were computed across folds. We calculated RMSE to evaluate the magnitude of the error; a lower RMSE indicates better prediction. Using the mean RMSE values from each of the 30 models, we performed a three-way Analysis of Variance (ANOVA); the outcome was RMSE, and the fixed factors were machine learning algorithm, biomechanical signal subset, and training method (*i.e.* Subject Dependent or Subject Independent). We evaluated the main effects of ANOVA at a significance value of $\alpha = .05$. We calculated average prediction time for our method pooled across all signal subsets, subject dependencies and ML models (total 30 combinations) using NVIDIA Tesla-P100-PCIE GPU.

III. RESULTS

All data are represented as a percentage error from the user's preferred prosthetic ankle stiffness. The mean and standard



(a) Subject Independent



(b) Subject Dependent

Fig. 3. Model performance for (a) Subject Independent and (b) Subject Dependent training methods. Red bars denote the mean RMSE and black error bars denote the standard deviation of RMSE, calculated across seven folds of cross validation. Results are shown for all five ML algorithms and three biomechanical signal subsets.

deviation of error for all 30 models is presented in Table II; the mean and standard deviation of RMSE is presented in Table III and Fig. 3.

There was a significant effect of training method on RMSE (ANOVA, $p < .0001$). Including subject-specific data in the training process (*i.e.,* the Subject Dependent case) resulted in 67% lower RMSE. Pooled across all machine learning algorithms and signal subsets, the RMSE for Subject Dependent training was $7.6\% \pm 2.6\%$, and for Subject Independent training was $23.1\% \pm 2.3\%$ (mean $\pm$ SD).

There was a significant effect of machine learning algorithm on RMSE (ANOVA, $p < .0001$). The three deep learning models (ANN, CNN, LSTM) performed better (*i.e.,* had lower RMSE) than the two classical machine learning models (KNN, SVR). Pooled across training methods and signal subsets, LSTM had the lowest RMSE of $13.4\% \pm 7.9\%$; KNN had the highest RMSE of $18.3\% \pm 6.0\%$.

We did not detect a significant effect of signal subset on RMSE (ANOVA, $p = .35$). Pooled across all training methods and machine learning algorithms, the RMSE was $15.4\% \pm 8.3\%$ for All Data, $14.8\% \pm 7.8\%$ for Affected Side Data, and $15.8\% \pm 8.2\%$ for Affected Side Foot/Shank Data.

For the Subject Dependent case, the LSTM network trained with Affected Side Data had the lowest RMSE of $5.2\% \pm 0.3\%$. For the Subject Independent case, the CNN trained with Affected Side Data had the lowest RMSE of $19.7\% \pm 6.4\%$.

The average prediction time for our method pooled across all signal subsets, subject dependencies and machine learning models was 1.99 ± 2.22 ms.

IV. Discussion

In this letter, we proposed a machine learning based approach for predicting the user's preferred prosthetic ankle stiffness. To this end, we implemented and compared five different ML architectures that estimate the user's preferred stiffness settings from lower extremity biomechanical data. We studied the effects of subject-specific data and biomechanical signal subsets used when training and estimating the user's preferred prosthesis ankle stiffness. The intent of this work is a first step toward reducing the experimental tuning process needed to identify user preferences.

A. Effect of Training With Subject-Specific Data

The inclusion of subject-specific training data significantly improved the algorithms' ability to estimate user preferred ankle stiffness. The mean RMSE error across all ML models trained on all three biomechanical signal subsets decreased from 23.1% to 7.6% when subject specific data were included. Using these distributions, we estimate our best performing ML model (LSTM) would predict a stiffness that would fall within 7.7% of the user's preferred stiffness with a likelihood of 41.5% without subject-specific data, and 99.9% when subject-specific data are included in the training set. This likelihood describes the probability that the LSTM-based approach would predict a stiffness that is imperceptibly close to their preference; we have previously shown that prosthesis users are unable to reliably sense stiffness changes less than 7.7% (*i.e.* JND = 7.7%) [18].

B. Effect of Machine Learning Algorithms

We identified that the three deep learning (DL) models (ANN, CNN, LSTM) performed better than the two classical ML (KNN, SVR) models for all subjects and sensor subsets (biomechanical signal groupings). This may indicate that the DL models are able to extract meaningful features from users' biomechanics that facilitate the estimation of user's preferred ankle stiffness. In addition, the features used with the classical ML methods were selected based on previous work [20]. The hand-crafted features which are used in our study have previously shown promising results in activity recognition and prosthesis control. However, these features may not be optimal and warrant future work [27].

C. Effect of Signal Subset Used in Training the ML Algorithms

The amount and location of biomechanical data included did not improve estimation accuracy when predicting user preferred ankle stiffness. This is especially interesting because the amount of biomechanical data varied greatly between conditions. For example, when all data were included, there were 93 time-series signals compared to 25 signals when only data from the affected foot and shank were included. In addition, previous studies [17][28] showed that kinematic symmetry was correlated with user's prosthesis ankle stiffness. This may indicate that the features important for user's selecting their preferences may not necessarily be crucial for machine learning models. Thus, we hypothesize that most of the biomechanical information required to estimate user-preferred prosthesis stiffness are included in the sensors local to the prosthesis. The relative insensitivity of our estimation error to the amount of data included is helpful for potential clinical applications. It is more convenient to obtain data from sensors within the prosthesis, which do not have the challenges associated with sensors elsewhere (*e.g.* discomfort, obtrusiveness, communication *etc.*).

D. Significance to Control of Wearable Robots

Leveraging user preference is a promising approach for the design and tuning of wearable robotic systems; however, the time required to obtain data on user preference can be limiting. Thus, in this work, we estimated user preferred ankle stiffness from biomechanical data. Our results demonstrate that prediction of user preferred stiffness is possible using biomechanical data, but the approach likely needs to include some user-specific biomechanical data for training. One potential solution could be to acquire training data for a smaller subset of activities, which can be used to predict the preferences and biomechanical trends of other activities. This approach could provide both the benefits of reduced experimental duration while enabling user preference as a method to adjust the stiffness of novel prostheses. In addition, this study showed little drop in performance as the dataset was expanded to include biomechanical information about other joints or legs. This is encouraging as it indicates that sensor data beyond what can be collected on the prosthesis may not be required for accurate prediction of a user's preferred stiffness. Thus, while user-specific data is likely needed, accurate estimation of user preferred stiffness is possible using only data included from the onboard sensors. Using sensors only on the prosthesis will potentially enable the deployment of the proposed system in the real-world.

E. Limitations & Future Work

We observed degraded performance when subject-specific training data was withheld (*i.e.* the Subject Independent case). This demonstrates the potential weak generalizability of our model to a novel user. In future work, we will use meta-learning approaches and a larger and more diverse dataset to improve the performance and generalizability of our models. Larger datasets of people with amputation using experimental hardware can be challenging to obtain. Thus, our work provides initial insight into this approach and lays a foundation for future research.

This investigation is limited to predicting preferred prosthesis stiffness while walking on level ground. User preferred stiffness is expected to vary with biomechanical demands of different activities. However, there is not yet sufficient data on how variable stiffness prostheses affect gait biomechanics across activities. A potential avenue of study would be the extrapolation of models to predict the preferences of new activities, thereby reducing required testing of user preferences. These data could enable the development of new strategies to specify the stiffness of the VSPA Foot or other variable stiffness prostheses.

For this initial investigation, we performed an offline analysis. As a future step, we will assess the model performance in real-time with human subject experiments. An online evaluation would enable us to quantify latency and any biomechanical effects caused by errors in performance.

REFERENCES

[1] A. F. Azocar, L. M. Mooney, J.-F. Duval, A. M. Simon, L. J. Hargrove, and E. J. Rouse, "Design and clinical implementation of an open-source bionic leg," *Nature Biomed. Eng.*, vol. 4, no. 10, pp. 941–953, 2020.

[2] L. Gabert, S. Hood, M. Tran, M. Cempini, and T. Lenzi, "A compact, lightweight robotic ankle-foot prosthesis: Featuring a powered polycentric design," *IEEE Robot. Automat. Mag.*, vol. 27, no. 1, pp. 87–102, Mar. 2020.

[3] A. J. Young and D. P. Ferris, "State of the art and future directions for lower limb robotic exoskeletons," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 2, pp. 171–182, Feb. 2017.

[4] A. Boonstra, V. Fidler, and W. Eisma, "Walking speed of normal subjects and amputees: Aspects of validity of gait analysis," *Prosthetics Orthotics Int.*, vol. 17, no. 2, pp. 78–82, 1993.

[5] N. Vanicek, S. Strike, L. McNaughton, and R. Polman, "Gait patterns in transtibial amputee fallers vs. non-fallers: Biomechanical differences during level walking," *Gait Posture*, vol. 29, no. 3, pp. 415–420, 2009.

[6] R. Waters, J. Perry, D. Antonelli, and H. Hislop, "Energy cost of walking of amputees: The influence of level of amputation," *J. Bone Joint Surg. Amer.*, vol. 58, no. 1, pp. 42–46, 1976.

[7] T. Lenzi, M. Cempini, L. J. Hargrove, and T. A. Kuiken, "Design, development, and validation of a lightweight nonbackdrivable robotic ankle prosthesis," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 2, pp. 471–482, Apr. 2019.

[8] E. M. Glanzer and P. G. Adamczyk, "Design and validation of a semi-active variable stiffness foot prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 12, pp. 2351–2359, Dec. 2018.

[9] M. K. Shepherd and E. J. Rouse, "The VSPA foot: A quasi-passive ankle-foot prosthesis with continuously variable stiffness," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2375–2386, Dec. 2017.

[10] H. Tryggvason, F. Starker, C. Lecomte, and F. Jonsdottir, "Variable stiffness prosthetic foot based on rheology properties of shear thickening fluid," *Smart Mater. Structures*, vol. 29, no. 9, 2020, Art. no. 095008.

[11] J. Zhang *et al.*, "Human-in-the-loop optimization of exoskeleton assistance during walking," *Science*, vol. 356, no. 6344, pp. 1280–1284, 2017.

[12] K. A. Ingraham, C. D. Remy, and E. J. Rouse, "User preference of applied torque characteristics for bilateral powered ankle exoskeletons," in *Proc. 8th IEEE RAS/EMBS Int. Conf. Biomed. Robot. Biomechatronics*, 2020, pp. 839–845.

[13] N. Thatte, H. Duan, and H. Geyer, "A sample-efficient black-box optimizer to train policies for human-in-the-loop systems with user preferences," *IEEE Robot. Automat. Lett.*, vol. 2, no. 2, pp. 993–1000, Apr. 2017.

[14] K. Li *et al.*, "Roial: Region of interest active learning for characterizing exoskeleton gait preference landscapes," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2021, pp. 3212–3218.

[15] S. U. Raschke *et al.*, "Biomechanical characteristics, patient preference and activity level with different prosthetic feet: A randomized double blind trial with laboratory and community testing," *J. Biomech.*, vol. 48, no. 1, pp. 146–152, 2015.

[16] M. Tucker *et al.*, "Preference-based learning for exoskeleton gait optimization," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2020, pp. 2351–2357.
[17] T. R. Clites, M. K. Shepherd, K. A. Ingraham, L. Wontorcik, and E. J. Rouse, "Understanding patient preference in prosthetic ankle stiffness," *J. Neuroengineering Rehabil.*, vol. 18, no. 1, pp. 1–16, 2021.
[18] M. K. Shepherd, A. F. Azocar, M. J. Major, and E. J. Rouse, "Amputee perception of prosthetic ankle stiffness during locomotion," *J. Neuroengineering Rehabil.*, vol. 15, no. 1, pp. 1–10, 2018.
[19] M. P. Kadaba, H. Ramakrishnan, and M. Wootten, "Measurement of lower extremity kinematics during level walking," *J. Orthopaedic Res.*, vol. 8, no. 3, pp. 383–392, 1990.
[20] B. Hu, E. Rouse, and L. Hargrove, "Fusion of bilateral lower-limb neuromechanical signals improves prediction of locomotor activities," *Front. Robot. AI*, vol. 5, no. 78, pp. 1–16, 2018.
[21] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
[22] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, 2004.
[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Representations*, 2015.
[24] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *Proc. 12th USENIX Symp. Operating Syst. Des. Implementation (OSDI 16)*, 2016, pp. 265–283.
[25] F. Chollet *et al.*, "Keras: The python deep learning library," *Astrophys. Source Code Library*, 2018, Art. no. 1806.
[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[27] G. Chandrashekar and F. Sahin, "A survey on feature selection methods," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 16–28, 2014.
[28] M. Major *et al.*, "The effects of prosthetic ankle stiffness on stability of gait in people with trans-tibial amputation," *J. Rehabil. Res. Develop.*, vol. 53, no. 6, pp. 839–852, 2016.