A Function Approximator Model for Robust Online Foot Angle Trajectory Prediction Using a Single IMU Sensor: Implication for Controlling Active Prosthetic Feet

Sharmita Dey and Arndt F. Schilling

***Abstract*—Lower limb dysfunction hinders the locomotive activity in individuals and compels them to use assistive devices such as prostheses or orthoses to restore the missing locomotive functions. Passive assistive devices have limitations in replacing the lost functionality. On the other hand, active devices could restore more natural locomotion using motorized joints. To make proper use of these embedded motors, a control architecture is required to translate the motion intent of the user into the intended prosthesis joint trajectories. In this article, we propose a temporal convolution-based online foot angle trajectory prediction network (FATP-N) that predicts the foot angular positions during unconstrained walking from the shank angular position measured using a single wearable motion tracker sensor. The data acquisition experiments were designed to reflect natural gait with turns, varying inclines, speeds, and mixed cadences. The trained models were prepared for real-time prediction by compressing them to reduce the storage and time complexity. The validations were performed on different motion conditions, including those that have not been used for model training, to verify the robustness of the model. The proposed FATP-N achieved high accuracy and could generate predictions in real time. The results indicate that the proposed approach offers a possible architecture for real-time control of powered intelligent prostheses.**

Manuscript received 1 December 2021; revised 12 February 2022; accepted 2 March 2022. Date of publication 15 March 2022; date of current version 13 December 2022. This work was supported in part by the German Federal Ministry of Education and Research under Grant INOPRO16SV7657. Paper no. TII-21-5356. *(Corresponding author: Sharmita Dey.)*

Sharmita Dey is with the Applied Rehabilitation Technology Lab (ART-Lab), Department of Trauma Surgery, Orthopedics and Plastic Surgery, University Medical Center Goettingen, 37075 Goettingen, Germany, and also with the Department of Computer Science, University of Goettingen, 37073 Goettingen, Germany (e-mail: sharmita.dey@med.uni-goettingen.de).

Arndt F. Schilling is with the Applied Rehabilitation Technology Lab (ART-Lab), Department of Trauma Surgery, Orthopedics and Plastic Surgery, University Medical Center Goettingen, 37075 Goettingen, Germany (e-mail: arndt.schilling@med.uni-goettingen.de).

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by (Name of Review Board or Committee) (IF PROVIDED under Application No. xx, and performed in line with the (Name of Specific Declaration).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3158935.

Digital Object Identifier 10.1109/TII.2022.3158935

***Index Terms*—Assistive robotics, bionics, convolutional neural networks (CNNs), deep learning, intelligent prostheses, rehabilitation robotics, robotic prostheses, temporal convolutions, trajectory prediction.**

I. Introduction

ANTICIPATING human behavior and human motion prediction have received growing attention in the last decade due to the increasing number of biomimetic human–machine systems. Accurate motion prediction is a challenging task due to the fact that human motion is a result of complex neuromechanical coordination. Precise estimation of future positions and trajectory planning of robotic agents like intelligent upper or lower limb prostheses, power augmentation exoskeletons, and assistive robotic devices (e.g., a smart wheel chair) are essential to impart intelligence and define control strategies for such semiautonomous systems.

Control strategies for such assistive robotic agents need to cofunction with humans and send control commands to the assistive device to perform the desired task in line with the motion of the remaining body [1], thereby, replacing the functioning of the missing or impaired limb. Different methods have been proposed for devising control strategies for such assistive devices [2]. For example, in a state-based control approach, a finite state-action space is typically modeled to develop control schemes for active prostheses/orthoses [3], [4]. A tabular value representation is maintained based on the finite states (i.e., kinematics or kinetics of the user/device) to assign control commands. The different state transitions are governed by switching rules that make decisions based on the sensor input. To support multiple phases of gait (e.g., initial stance, midstance, and early swing) or multiple locomotion tasks (e.g., level-ground walking, walking on inclined ground, and stair ambulation), the number of parameters to be tuned, tabular representations to be maintained, and switching rules to be defined increases [5], leading to an explosion of the parameter space. Pattern recognition algorithms for locomotive intent and gait phase detection could automate some steps [6] (e.g., potentially eliminate switching rules). However, discretization of the gait cycle to finite states and its annotation remains a problem, especially when a finer division of the gait cycle is desired to devise a seamless control

strategy. Moreover, errors and latency in the classification of the intended locomotion tasks and gait phase can lead to the risks of imbalance [7].

More recently, function approximator models have been used to learn a generalized mapping between the sensor input that reflects the motion intent of the prosthetic user and the gait variables to be predicted as output that can be used to control an active prosthesis [8]–[10]. Since such a generalized mapping is learned across different gait phases and locomotion tasks, this method alleviates the need for tuning the parameters, switching between gait phase and tasks, as well as eliminating the need for maintaining a tabular representation discretely. Even though such methods could achieve near-accurate predictions, most of the studies have focused on predicting gait variables at the current time instant using a multitude of sensor variables [8], [10], [11]. These approaches have two potential limitations for predicting gait variables for prosthetic control. First, the prediction delays combined with latency introduced by actuator control and mechanics can lead to delays in actuation [2], [12] and can cause phase differences in the motion of the user and the prosthesis [7]. Second, acquiring inputs from multiple sensors attached to the subjects' body increases the overhead for processing and synchronizing multiple channels in real time.

Therefore, in this article, we aimed at developing a strategy for predicting the foot angular position trajectories in advance while using a "minimal-sensor" setup. To achieve this, we exploited the sequential nature of the gait by which the past temporal patterns of the gait variables could hold information about the current and future joint trajectories. We hypothesized that an appropriate function approximator could learn the temporal relations in the data to predict the desired gait variable trajectory in advance, thus compensating for other delays. Moreover, such a model could potentially minimize the sensor overhead by leveraging history information from a single sensor.

II. Contributions

In this article, we propose a novel temporal convolutional network (TCN)-based framework called foot angle trajectory prediction network (FATP-N), which could use the temporal intralimb synergy during locomotion to predict the future sagittal foot angle trajectory. In contrast to other studies, we used only a single input feature from a single sensor to obtain high prediction accuracy that facilitates minimal-sensor setup and simpler system design. We found that incorporating bouts of history information into the input signal trajectories reduces the need for multiple input features. Moreover, we conducted the experiments without a constrained experimental protocol to replicate walking in the "wild." The subjects were asked to walk freely with turns, variable speeds, inclines, and cadence. We trained individual FATP-N models for each subject from the recorded data. Further, we verified these models using online and offline experiments, which were also conducted on conditions that have not been encountered during training (validated on contralateral turns and higher slopes/inclines while trained on only ipsilateral turns and lower inclines). The trained models were compressed to reduce the storage and time complexity for real-time prediction. The main contributions of our work are summarized as follows.

1) A TCN-based novel framework to predict the sagittal foot positions in advance by exploiting the biological body segment synergies during human gait.
2) A solution to leverage information from a single wearable inertial measurement unit (IMU) ("minimal-sensor" setup).
3) Evaluation experiments testing different walking scenarios encountered during natural, unconstrained locomotion (different inclines, speeds, turns, and cadence).
4) Optimized and compressed models for real-time performance as well as scalability to edge devices.
5) Comparison with different baseline and state-of-the-art data-driven methods for gait trajectory prediction to provide a benchmarking platform for future studies in a similar direction.

III. Related Work

Recently, advanced deep learning algorithms like recurrent neural network (RNN) architectures have gained interest in different fields for time-series prediction. Many authors in biomechanics and rehabilitation have used such algorithms for gait trajectory prediction, which could eventually help design better control strategies for powered assistive devices. Liu *et al.* [13] used a deep spatiotemporal model based on long short-term memory (LSTM) networks for advanced prediction of knee angle trajectories from measurements of other joint angles of both legs for controlling a powered exoskeleton. Liang *et al.* [14] used motion data from wearable sensors to learn a synergy between upper and lower limb trajectories using LSTM networks and predict the reference hip and knee angle trajectories for stroke patients. Mundt *et al.* [15] used feed-forward neural networks and LSTM networks to predict lower limb joint angles and moments from inertial data simulated from optical motion tracking data. Su *et al.* [16] used an LSTM network to predict gait trajectories (lower limb joint angular velocity) and gait phase using multidimensional data from seven IMU sensors attached to the lower limbs and pelvis. He *et al.* [17] used visual signals from a Kinect sensor to predict the lower limb trajectories using an LSTM network. Zaroug *et al.* [18] used LSTM networks for multitimestep forecasting of lower limb trajectories using motion capture data. Fang *et al.* [19] implemented a gait neural network using temporal convolutional networks for gait mode recognition and gait trajectory prediction using data from an array of sensors attached to the lower limbs and pelvis.

The high accuracy of the predicted trajectories obtained in these studies indicates that deep learning algorithms tailored to time-series data could learn the nonlinear relations between gait variables and use sensor history information to forecast gait trajectories. However, some of these studies used either marker-based motion capture data or visual signals, which can be reliably acquired only in laboratory environments. Other studies required acquiring inputs from multiple sensors attached to the subjects' body that may increase overhead for multiple channels processing or synchronization in real time. Alcaraz *et al.* [20]

TABLE I
COMPARISON OF FATP-N APPROACH WITH OTHER RELATED STUDIES

Study	Algorithm	Experiment condition	Data collection method	Input variables	Online validation	Accuracy
Liu et. al. [14]	LSTM	Comfortable gait pattern with varying speeds	Wearable IMUs	6	Yes	–
Liang et. al. [15]	LSTM	Walking on a hallway at two speeds	Wearable IMUs	14	Yes	Error rate = 1.21 – 2.33%
Mundt et. al. [16]	LSTM	Level walking trials at self-selected speeds	Optical motion capture	42	No	nRMSE $>$ 7.3%
Su et. al. [17]	LSTM	Walking on treadmill at five predefined speeds	Wearable IMUs	63 (7 IMU x 9 channels)	No	ρ = 0.98
He et. al. [18]	LSTM	Walking on treadmill at four predefined speeds	Kinect motion sensors	2-4	No	RMSE = 2.4 – 2.7°
Zaroug et. al. [19]	LSTM	Walking on treadmill at a defined speed	3D motion capture	4	No	ρ=0.96 – 0.98
Fang et. al. [20]	TCN	Walking, stairs and slopes	Wearable IMUs	18	No	Error $<$ 7.28%
Ours	TCN	**Walking at varying speeds, cadence, inclines, and turns**	Wearable IMU	**1**	**Yes**	nRMSE = 4.79±0.9% ρ = 0.98±0.01

used data from a single IMU sensor attached to the foot segment to compare different deep learning architectures for joint angle trajectory prediction. However, this method is not feasible for powered prosthesis control for transtibial and transfemoral amputees. Moreover, most of the existing studies (with a few exceptions like [13] and [14]) have not verified the performance of such algorithms for the real-time prediction of gait trajectories. Similarly, many of these studies were performed under constrained locomotion protocols, which do not complement the natural walking conditions of humans. This study addresses these challenges with a unified solution by using compressed temporal convolutional networks that support varying motion conditions while using input from a single wearable sensor to generate real-time predictions on both laptop PC and edge devices. To the best of our knowledge, this is one of the first studies that showcase the potential of a temporal convolutional neural network (CNN) on an edge device for lower limb trajectory prediction.

In Table I, we compare our approach with previous studies that used state-of-the-art sequence prediction models for gait trajectory prediction. It is seen that our approach required only a single input variable to achieve comparable or better performance than the state-of-the-art approaches. However, different studies have quantified their results following different methods, making it difficult to compare the performance of our approach to other studies directly. Moreover, the inputs and predicted variables also varied across studies. To establish a common ground for comparison, we additionally compared our approach to other data-driven models that are potential contenders for our application using our data. This comparison can provide a platform to benchmark future works in a similar direction.

IV. METHODOLOGY

A. Model Selection

Our learning-based model for predicting the foot angular positions was inspired by three main aspects.

1) Relevance to application: Since our application scenario, i.e., gait is an ordered temporal sequence, the history of the temporal kinematic signals could hold information about the current and future trajectories. Therefore, using models that preserve such temporal correlations in data is agreeable. The sequence prediction models such as RNNs and TCNs are tailored to such tasks.
2) Computational efficiency: Computational efficiency is an essential aspect of real-time actuation applications because this decides the reaction time to low-level control commands. Therefore, a model with few parameters is preferred. Deep learning models like CNNs and RNNs use parameter sharing to reduce the number of parameters and facilitate generalization.
3) Performance: To ensure safety in prosthesis control, high prediction accuracy is desired from the model meant to predict the gait variables. Therefore, we choose a model that consistently predicts with high accuracy across different input embeddings and motion conditions.

To select the model that best satisfies these requirements, we compared five different algorithms—linear regression (LR), Gaussian process regression (GPR), gated recurrent units (GRUs), LSTM networks, and TCNs for the predictive accuracy and computational time requirements.

B. Foot Angle Trajectory Prediction Network (FATP-N)

We propose an FATP-N based on a 1-D temporal CNN that captures the history of the shank segment orientation (θ_{shank}) over a sequence of samples to forecast the trajectory of the foot angle (θ_{foot}) at a time point in the future. The FATP-N can be defined as a function $\mathcal{F}$

$$\theta_{\text{foot}}^{t+m} = \mathcal{F}(\theta_{\text{shank}}^{t-T}, \theta_{\text{shank}}^{t-T-1}, \ldots, \theta_{\text{shank}}^{t}) \tag{1}$$

where m is the number of samples to the future at which the θ_{foot} is predicted, and T is the number of samples of history information of θ_{shank} used. In this article, we set $m = 10$ and $T = 15$. The implementation of the FATP-N has been inspired from [21]–[23], which describes a variation of CNNs for sequence modeling tasks by combining 1-D CNNs and causal convolution. We chose a TCN variant over the commonly used RNN architectures like LSTM since TCNs are structurally simpler but computationally more efficient than RNN architectures. Unlike RNNs, which sequentially process the input sequence to compute the hidden states, TCNs can process input sequences in parallel, thus making the computational effort independent of the input sequence lengths. TCNs can preserve temporal information through temporal convolution and pass it through the hidden layers despite the parallel processing.

The FATP-N (see Fig. 1), similar to the TCN implementation in [23] ensures causal behavior by employing convolutions such

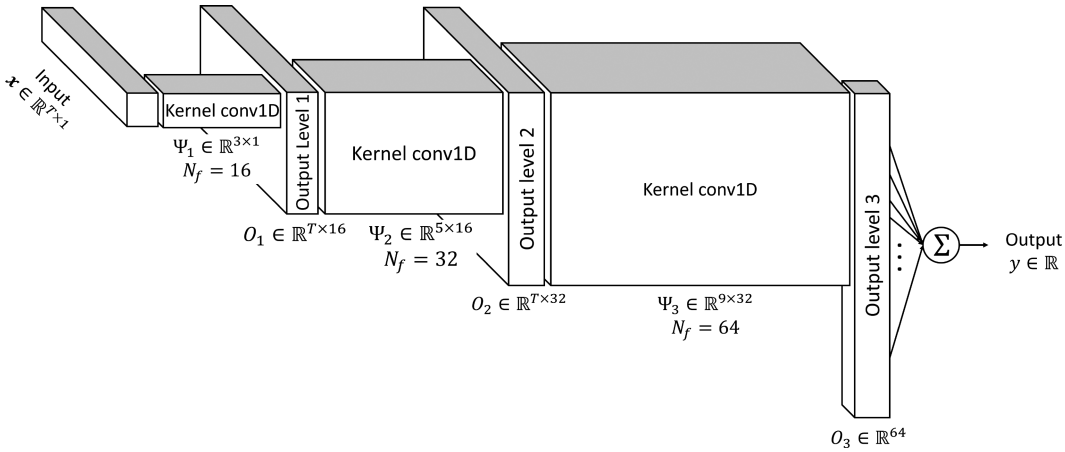

Fig. 1. Structure of the FATP-N with three-layered temporal convolutions followed by a fully connected layer. At layer 1, 16 kernels of size three are used. At level 2, 32 kernels of size 5, and at level 3, 64 kernels of size nine are used.

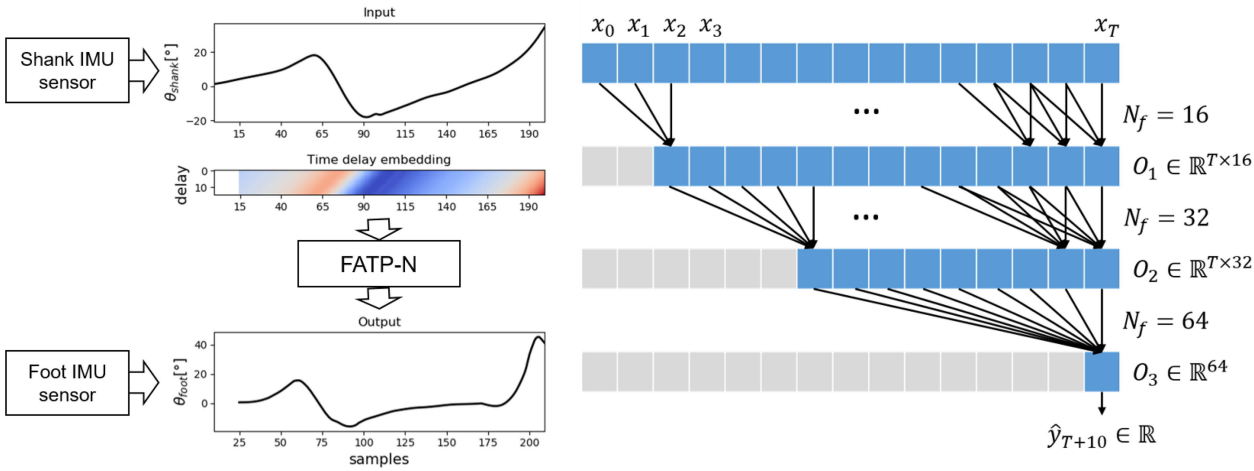

Fig. 2. (Left) Input and output of the FATP-N model: The θ_{shank} measurements obtained from the shank IMU sensor are time-delay embedded with a lookback of 15 samples. The FATP-N model learns to map the time-delay embedded θ_{shank} to the θ_{foot} ten samples to the future. θ_{foot} values for training FATP-N are obtained from an IMU sensor placed on the subject's foot. (Right) The architecture of the FATP-N: A three-level stacked temporal convolutional network with dilation of one at each layer, kernel sizes $K = 3$, 5, and 9 in respective layers. The number of kernel filters N_f in each layer were 16, 32, and 64. At each level, the output feature map has the same length T as the input sequence and a width equal to the number of filters in that layer, N_f.

that the output at the time point, t, is convolved with only the components from time point, t, and earlier in the preceding layer. The causal convolution operation, C, at a time point, t, with a kernel, $\Psi : \mathbb{R}^K \rightarrow \mathbb{R}$ of size K on a 1-D sequence input, $\mathbf{x} \in \mathbb{R}^T$, with a dilation factor of d, can be defined as

$$C(t) = \sum_{i=0}^{K-1} \Psi(i).\mathbf{x}(t - d.i). \tag{2}$$

While typical implementations of the TCN [22], [24] uses the same kernel size across layers with exponentially increasing dilations [23] to allow large receptive fields, we propose a three-level stacked temporal convolutional network with a dilation factor, d, of one and kernel size, K growing with the level of the stack [see Fig. 2 (right)]. Each level of the FATP-N is comprised of causal convolutions and nonlinear activation. At level 1, 16 filters of size 3 and a causal padding were used, leading to feature map, $O_1 \in \mathbb{R}^{15\times 16}$. At the second level, 32 filters of size 5 and a causal padding were used, leading to a feature map $O_2 \in \mathbb{R}^{15\times 32}$. At the final level of the TCN stack, 64 filters of size 9 were used and the last output of the output sequence was considered, leading to a feature map, $O_3 \in \mathbb{R}^{64}$. The final feature map, O_3, is passed through a fully connected layer to obtain the network's final output. This network structure was selected since we observed in our initial analysis that using the same dilation and increasing the kernel size across levels increases accuracy in most cases while keeping the computational time almost similar. Moreover, this practice simplifies the later steps to compress the model.

Other hyperparameters of the model were determined using a grid-search cross validation. We computed the accuracy on a validation set using different hyperparameter combinations from a parameter grid. Specifically, we tested different activation

functions like identity, hyperbolic tangent (tanh), sigmoid, relu, and elu, different optimizers like Adam and SGD, different dropout rates of 0.0, 0.2, and 0.4, and different initializers like He normal, He uniform, Glorot normal, and Glorot uniform. Based on the experimental hyperparameter search and literature review, we used Adam as optimizer, tanh as activation, Glorot uniform as initializer, and a dropout rate of 0.4. Adam [25] is one of the most popular state-of-the-art optimizers generally due to its variable learning rate and momentum for weight updates, which causes faster convergence. Glorot [26] is the recommended initialization corresponding to a tanh activation [27]. The implementations were done using Tensorflow Keras [28].

C. Offline Experiments With IMU Sensor Data

1) Data Acquisition: Shank and foot angular positions were acquired from seven able-bodied subjects (height: 169.6 ± 6.8 cm, mass: 72.3 ± 8.6 kg, age: 42.1 ± 16.6, two females, and five males). Data were obtained at 100 Hz using wireless motion tracker sensors or IMUs (MTw Awinda, Xsens) strapped near the center of mass of the shank and foot of the subject. The walking trials lasted for 35–40 s and were as natural as possible, including varying speeds, cadences, and turns without being strictly constrained to be on a straight line. To assess whether the FATP-N model can generalize its performance to variations in walking patterns, we acquired data from a subject walking on a treadmill at different inclinations (0°, 5°, and 10°) and speeds increasing from 1.5 to 4 m/h. The measurements from the shank sensor were used to provide inputs to the FATP-N model, whereas the foot sensor measurements were used as ground truth for training the FATP-N model and testing the accuracy of its predictions.

2) Input and Output: The time history information of the flexion-extension (sagittal) shank angles (θ_{shank}) were the input to the FATP-N model, whereas the flexion-extension foot angular positions (θ_{foot}) at a future time point were the output to be predicted. A time-delay embedding was performed to ensure causal connections over a look back of 15 samples (150 ms at 100 Hz). The FATP-N mapped each time-delay window of 15 shank angle data samples to the foot angular position, ten samples (100 ms at 100 Hz) into the future (see Fig. 2). The shank and foot angular positions were obtained directly from the wireless motion trackers (IMUs).

3) Model Training and Testing: Three-level stacked FATP-N models were trained and tested for each subject separately using a five-fold cross-validation scheme on the prerecorded data using IMU sensors. During each cross-validation iteration, 80% of data from the subject was used to train the corresponding models, and the remaining 20% was held out for testing the model performance. For a 40-s recording session with 4000 samples (100 Hz), training data consisted of 3200 samples (32 s, 100 Hz), and the test data consisted of 800 samples (8 s, 100 Hz). Accordingly, after the fivefolds of cross-validation, each recorded data point was used for testing once. Furthermore, to assess the generalization capability of the FATP-N, the model was trained on data from a subject walking at varying speeds (1.5 to 4 m/h) and inclines of 0° and 5°. The model performance was then tested on data from a held-out trial with similar walking speed variations and inclines of 0° and 10°. The training inputs and outputs of the model were normalized individually within a range of zero to one. The test inputs and outputs were also normalized using the scale factor obtained from the train data. This was done to simulate a real-time scenario where it is not possible to find the scale factors for input and outputs from real-time data. The coefficient of determination (R^2), correlation coefficient (ρ), and the normalized root mean square error (*NRMSE*) between the actual and predicted trajectories were considered for quantifying the quality of the predictions on the out-of-sample data. Similarly, the inference rates were measured to quantify the computational efficiency.

D. Online Experiments Using IMU Sensors

One of the main challenges of using CNN-based models for real-time applications is the computational complexity of the model due to a large number of parameters. We addressed this challenge by compressing the trained model to a flatbuffer format [29] to make it efficient for real-time predictions both in terms of computational time and storage requirements. We call this compressed model *FATP-N Lite*.

The trained and compressed FATP-N models were used to predict the foot angular positions online as the subjects walked at self-selected speeds. The wireless motion trackers were placed on the shank and foot at the same body locations of the subjects as during the offline data acquisition (in the previous section). The foot sensor measurements were used only to test the prediction accuracy. The subjects were asked to walk as naturally as possible without following any strict experimental protocol. The shank and foot motion data were recorded using IMUs at 100 Hz and were written online to text files. A custom-written program monitored these files continuously, and whenever new data became available, the trained FATP-N model for predicting the foot angular positions was invoked. For each prediction, the previous 15 samples of the shank angular positions were read and normalized using the scaling factor obtained during training and was used as the input to the FATP-N. The FATP-N predictions at each query point were compared with the IMU measured foot angular positions ten samples in the future to quantify the performance of the models. The online trial of each of the subjects lasted 20 s. The online experiment process is illustrated in Fig. 3. An Intel i7 processor (six cores) hardware with 16-GB RAM and a clock speed of 2.59 GHz was used for offline training of the FATP-N models and online predictions.

E. State-of-The-Art and Baseline Comparisons

Finally, we compared the TCN-based FATP-N approach with four other learning-based models; namely, LSTM networks, GRUs, LR, and GPR model with "matern" covariance function. LSTM and GRU are advanced state-of-the-art RNN architectures and potential contenders for our application scenario. LR and GPR models served as representative baselines for linear and nonlinear methods for comparison. Each approach learns a functional relation between history information from the shank segment motion and a future sagittal foot position as given in (1) with a future lookup, m, of ten samples. To ensure that

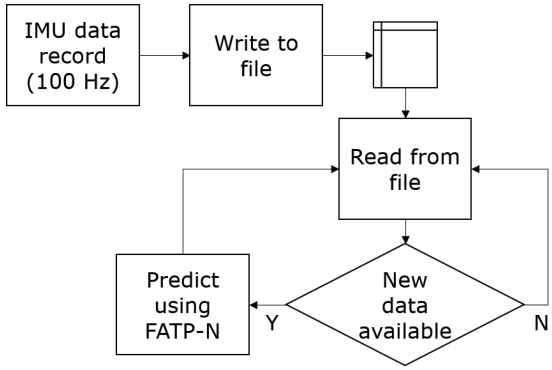

Fig. 3. Online experiment protocol. The shank and foot IMU data are stored into text files, which are polled for new data. When new data becomes available, a compressed FATP-N model is invoked for the prediction.

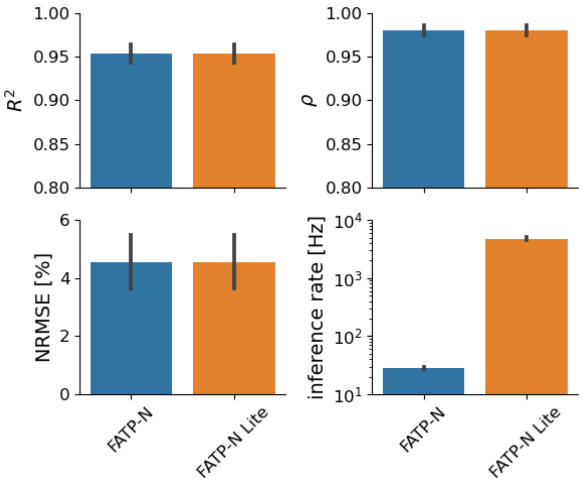

Fig. 4. Mean R^2, correlation coefficient (ρ), NRMSE, and inference rate of FATP-N predictions on prerecorded data across seven subjects. Error bars show one standard deviation from the mean.

this comparison is consistent and sound, we performed this analysis across 17 different sequence lengths, $1 < T < 200$ for all the five different models. Furthermore, for a fair comparison between FATP-N and RNN-based architectures, we selected the network structures such that each model had approximately the same number of tunable parameters. The hyperparameters of LSTM and GRU were determined using grid search cross validation with the same parameter grid used for FATP-N. The algorithms were trained and tested on IMU data following a fivefold cross-validation protocol detailed in Section IV-C3. Finally, we report the R^2 scores, correlation coefficient ρ, and inference rates in samples/s as a function of input sequence length, T.

V. Results

A. Offline Experiments

Fig. 4 shows the performance of the FATP-N models trained and tested using fivefold cross validation on prerecorded data from seven subjects. An R^2 score of 0.95 ± 0.01, *NRMSE* of $4.7 \pm 0.9\%$, and a correlation coefficient, ρ of 0.98 ± 0.01 was reported. Furthermore, the FATP-N model trained on varying speeds and inclines could generalize its performance to similar speeds and new inclines (see Fig. 5), and generate predictions with high accuracy ($R^2 = 0.94$, $\rho = 0.98$, and NRMSE=5%). These results showed that the FATP-N model performance is consistent in adapting to variations in walking patterns.

The offline trained FATP-N models used around 600 kB of memory and generated predictions at a rate of 28.1 ± 1.5 Hz. Since the IMU acquisition rate was 100 Hz, a real-time prediction could not be supported with this prediction rate. To improve the computational efficiency of the FATP-N model for real-time predictions, the models were compressed to a flatbuffer format (FATP-N Lite). The FATP-N Lite models were able to generate predictions at a rate of 4.81 ± 0.25 kHz (150x speed-up in the inference rate) while not affecting the high prediction accuracy (see Fig. 4). Moreover, the FATP-N models required only 275 kB of disk space (compared to 600 kB for the FATP-N model), thus achieving a compression rate greater than two.

TABLE II
Metrics Quantifying Accuracy and Computational Speeds of Online Predictions Using Compressed FATP-N Models

Metric	best	worst	mean
R^2	0.97	0.88	0.93
ρ	0.99	0.97	0.98
NRMSE [%]	4.0	8.0	6.0
Update rate [Hz]	99.8	99.9	99.8
Read rate [Hz]	1082.15	1048.55	1068.9
Prediction rate [Hz]	99.8	98.75	99.4
Loss rate [Hz]	0.1	1.05	0.44

B. Online Experiments

Table II (first three rows) shows the performance metrics of the compressed FATP-N model during the online experiments using IMU data. The online prediction performance was slightly lower than that of the offline predictions. A high correlation coefficient ($\rho > 0.97$) was obtained between all subjects' predicted and sensor-measured foot angular positions. The *NRMSE* stayed below 8% and averaged to 6%. The R^2 scores ranged between 0.88 and 0.97 and averaged to 0.93.

Analogous to the offline predicted trajectories, the online predicted foot angle trajectories closely followed the sensor-measured trajectories (see Fig. 6). Misalignment between the sensor-measured and predicted trajectories occurred mainly during the peak foot angular positions. The highest correlation between the measured and predicted trajectories occurred at a delay of -10 samples except for one subject for whom the delay for the highest correlation was -11 samples. This indicates that the predicted trajectory leads the measured trajectory by $\sim$10 samples, which conforms to how the FATP-N model was trained to predict the foot angle ten samples into the future.

Table II (last four rows) lists the frequencies at which the prediction algorithm monitors (polls) the text file for new data from IMUs, the rate at which the text file was updated with new

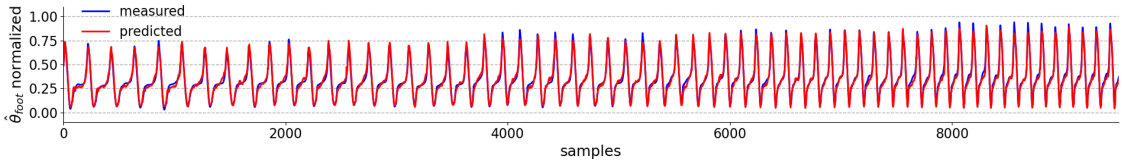

Fig. 5. FATP-N predictions for a subject walking at speeds increasing from 1.5 to 4 m/h and on two different inclines (0° and 10 °). The model was trained on data from another trial during which the subject walked at different speeds (1.5 to 4 m/h) and on inclines of 0° and 5°. Results are shown for 9000 samples acquired at 100 Hz.

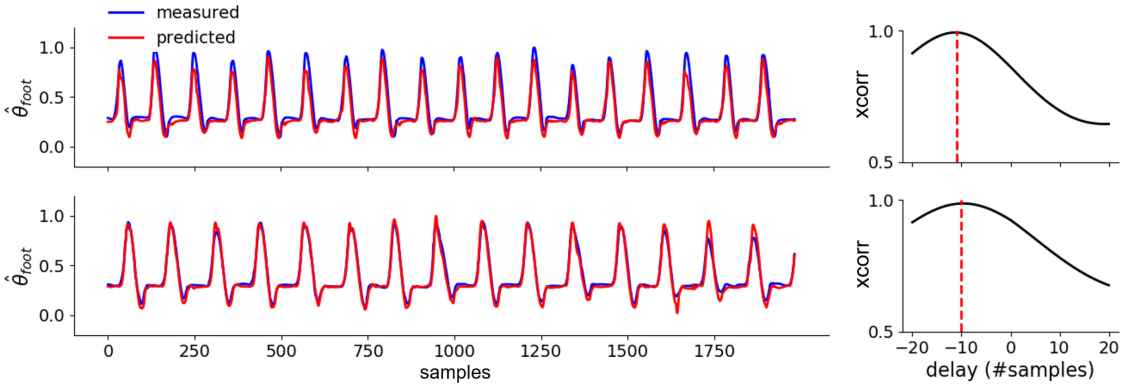

Fig. 6. (Left) Normalized foot angular positions predicted online by the FATP-N models (red) for two subjects (Top: worst R^2 and bottom: best R^2) using shank angular positions acquired at 100 Hz during the 20-s walking trials at self-selected speeds, including turns. The blue curve shows the measured values of normalized foot angle trajectories ten samples (0.1 s at 100 Hz) to the future. (Right) Cross correlation (xcorr) between the measured and predicted normalized θ_{foot} trajectory as a function of the delay between measured and predicted trajectories. A delay less than zero indicates that the predicted trajectory leads the measured trajectory.

data from IMU, the rate at which new predictions are generated, and the rate at which query points are lost (not read or predicted). The polling rate was above 1 kHz, which was more than ten times faster than the file update rate (∼100 Hz). The prediction rates closely matched the file update rate, with loss rates less than one sample per second. It was observed that the file write operation took 120 μs on average. An acquisition rate of 100 Hz ensured that no recorded sample was lost while writing to file. The read-operation took on average 630 μs, and the FATP-N prediction of each query point took on average 320 μs. Since the read and predict operations took less than 1 ms, the proposed algorithm could potentially generate predictions as fast as 1 kHz. Data loss (samples being not read or predicted) occurred when more than one sample was written during the read-predict operation. Since the read-predict operation took less than a millisecond and the acquisition rate was around 100 Hz on average, data loss occurred when the IMU data was being streamed or written in bursts (with less than 1 ms between consecutive samples).

Finally, to assess the model's scalability to edge devices, the compressed FATP-N models (FATP-N Lite) were mounted on a Raspberry Pi 4, and computational requirements were tested. It was found that FATP-N Lite could read new samples and generate predictions at a rate of ∼350 Hz, which is more than three times the data acquisition rate. These results indicate that the proposed FATP-N Lite models are scalable to be used on edge devices, and thus, provide a promising approach for designing onboard controllers for powered prosthesis and exoskeletons.

C. State-of-the-Art and Baseline Comparisons

Fig. 7 shows the prediction accuracies and inference rates of the different algorithms. The prediction accuracy was quantified using the R^2 score and correlation coefficient ρ between the measured and predicted trajectories. The inference rates were measured as the rate at which the test samples were processed. For smaller sequence lengths, all algorithms expect LR gave a good accuracy ($R^2 > 0.9$ and $\rho > 0.95$). With an increase in sequence length up to $T = 100$, the prediction accuracy of GRU, LSTM, and GPR reduce at different rates. On the other hand, LR showed an increase in accuracy with sequence length up to $T = 10$, and then, plateaued at an R^2 of around 0.85 except for a dip at sequence lengths from $T = 75$ to $T = 100$. GPR showed a rapid decrease in accuracy with increasing sequence length, eventually reaching an R^2 of around 0.7 and ρ below 0.9. LSTM network's accuracy also decreased with increasing sequence length, reaching a performance par with the LR model at $T = 50$ before showing a further increase at a large sequence length of $T = 200$. The GRU showed a slight decrease in accuracy with sequence lengths increasing up to $T = 100$ but regained its performance when they were increased further. The TCN gave consistently high accuracy for different sequence lengths with a slight increase in accuracy for larger lengths.

The sequence-based deep learning models like GRU and LSTM regained their initial performance with increasing sequence length. We assume that this is because of the cyclic nature of gait in which larger sequence lengths include

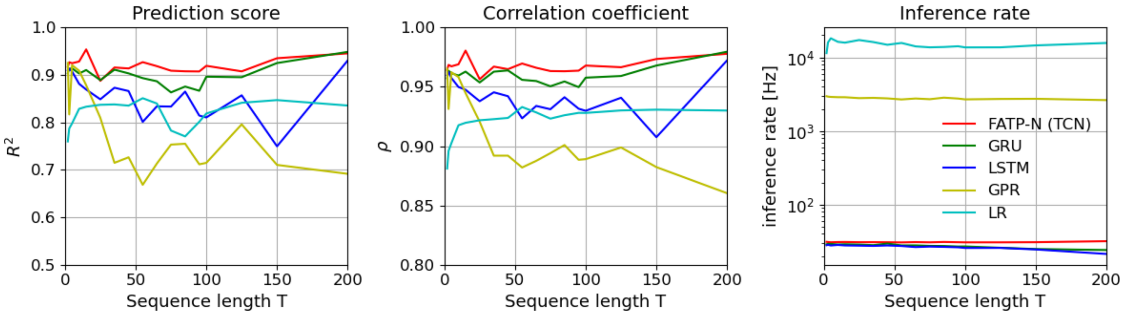

Fig. 7. Prediction score (R^2), correlation coefficient (ρ), and inference speeds of different learning-based models across various input sequence lengths. The algorithms were trained and tested on IMU-acquired kinematic data from seven subjects.

information from the same phase of the previous gait cycle (here, T has a maximum value of 200, which corresponds to 2 s, whereas the time for one gait cycle is slightly above one second on average). The comparisons made it evident that, among the compared algorithms, the TCN offered the best combination of prediction accuracy and computational simplicity. These results inspired us to use the TCN as the base of our FATP-N.

The superior performance of TCNs compared to RNNs across varying sequence lengths may be explained by the fact that TCNs are able to maintain much longer memory than RNNs [23]. TCNs also offer many advantages in sequence prediction tasks compared to RNNs due to their structural and functional differences with RNNs. While RNNs process the input signal sequentially in time, TCNs can process the input sequence as a whole using shared kernels. Furthermore, different kernel sizes allow TCNs to have a flexible receptive field to learn relations at different scales.

VI. Discussion

The main challenges of implementing a real-time control strategy for powered assistive devices like active prostheses are the delays due to data acquisition, control command generation, and motor control. Since human locomotion is a temporal sequence of events involving different limbs, joints, and muscles, this study aimed to predict the θ_{foot} at a future time point using the time history information of θ_{shank}. This was achieved by training a three-level stacked temporal convolutional network (which we call FATP-N) to map the time history information of the shank angular positions to a future foot angular position. The training was done using shank and foot angular position data acquired from wearable motion sensors (IMUs) as subjects performed gait experiments. The trained model was validated on out-of-sample data. Offline validation of the trained models showed high accuracy indicating the potential of the proposed FATP-N models for predicting the foot angle trajectories using only a single input feature (θ_{shank}).

However, deep learning architectures like CNNs are computationally complex to implement for real-time applications. Since real-time prosthesis control is a major focus of the proposed FATP-N architecture, it becomes interesting to make the proposed model suitable for real-time prediction. This was done by compressing the model to a flatbuffer format. The compressed FATP-N models (FATP-N Lite) were less than half the size and more than 150 times faster than the uncompressed models for individual predictions. The online prediction accuracy obtained using the compressed FATP-N models was slightly lower than that of the offline performance. Moreover, we showed that the compressed FATP-N models are scalable to edge devices like Raspberry Pi for real-time prediction, thus presenting a promising approach for onboard controllers for prosthesis/exoskeletons.

In this article, a future look-forward of ten samples (100 ms at 100 Hz) was selected to showcase the ability of FATP-N models to forecast the foot angular position in the future. For all subjects, the predicted θ_{foot} correlated the most with measured θ_{foot}, 10 to 11 samples into the future (see Fig. 6). In a real-world control application, the future predictions could be stored in a buffer and applied by considering the control delays. On the other hand, a large look forward may not be recommended since the user might change his locomotive intent by the time this prediction could be applied for control.

The FATP-N models were tested on various motion conditions ranging from walking with varying speeds, cadences, turns, and inclines. Future work should verify the method with further motion conditions, e.g., walking on different terrains, stair ambulation, running, etc. Furthermore, our strategy remains to be tested on an even larger pool of subjects, for example, amputees. Nevertheless, this study evaluated the feasibility of using the FATP-N for real-time prediction of gait variables using an input feature measured using a single wearable motion sensor. The high accuracy and real-time prediction rates indicate that the predictions from the FATP-N model could be used for online control of powered assistive devices like prostheses and orthoses. These results have implications for developing a real-time position control strategy of the ankle joint or the foot segment of an ankle-foot prosthesis/orthosis, which can assist the user with improved ankle-foot function, especially for ground clearance during the swing phase.

VII. Conclusion

In this article, we proposed a FATP-N, which can predict the foot sagittal angle in advance from the history information of shank sagittal angle. The performance of the FATP-N model was tested in offline and online experiments with wearable

motion trackers. We showed that the input from only a single gait variable obtained from a single IMU sensor was required for near-accurate, robust real-time prediction of the foot angle using the proposed FATP-N network during unconstrained locomotion. We used a compressed version of the trained models to improve the computational efficiency of online predictions. The results indicated that the proposed approach offers a possible architecture for real-time control of powered intelligent prostheses by compensating for the lag due to signal acquisition, processing, and motor control.

VII. Acknowledgment

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

[1] F. Botros, A. Phinyomark, and E. Scheme, "EMG-based gesture recognition: Is it time to change focus from the forearm to the wrist?," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 174–184, Jan. 2020.

[2] M. R. Tucker *et al.*, "Control strategies for active lower extremity prosthetics and orthotics: A review," *J. NeuroEng. Rehabil.*, vol. 12, no. 1, pp. 1–30, Jan. 2015.

[3] B. E. Lawson, "Control methodologies for powered prosthetic interventions in unilateral and bilateral transfemoral amputees," Ph.D. dissertation, Vanderbilt Univ., Nashville, TN, USA, 2014.

[4] G. Chen, Z. Liu, L. Chen, and P. Yang, "Control of powered knee joint prosthesis based on finite-state machine," in *Proc. Chin. Intell. Automat. Conf.*, 2015, pp. 395–403.

[5] B. E. Lawson, A. H. Shultz, and M. Goldfarb, "Evaluation of a coordinated control system for a pair of powered transfemoral prostheses," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2013, pp. 3888–3893.

[6] H. A. Varol, F. Sup, and M. Goldfarb, "Multiclass real-time intent recognition of a powered lower limb prosthesis," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 3, pp. 542–551, Mar. 2010.

[7] H. Huang, F. Zhang, L. J. Hargrove, Z. Dou, D. R. Rogers, and K. B. Englehart, "Continuous locomotion-mode identification for prosthetic legs based on neuromuscular-mechanical fusion," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 10, pp. 2867–2875, Oct. 2011.

[8] M. M. Ardestani *et al.*, "Human lower extremity joint moment prediction: A wavelet neural network approach," *Expert Syst. Appl.*, vol. 41, no. 9, pp. 4422–4433, 2014.

[9] Y. Yun, H.-C. Kim, S. Y. Shin, J. Lee, A. D. Deshpande, and C. Kim, "Statistical method for prediction of gait kinematics with Gaussian process regression," *J. Biomech.*, vol. 47, no. 1, pp. 186–192, 2014.

[10] N. Dhir, H. Dallali, E. M. Ficanha, G. A. Ribeiro, and M. Rastgaar, "Locomotion envelopes for adaptive control of powered ankle prostheses," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 1488–1495.

[11] S. Dey, M. Eslamy, T. Yoshida, M. Ernst, T. Schmalz, and A. Schilling, "A support vector regression approach for continuous prediction of ankle angle and moment during walking: An implication for developing a control strategy for active ankle prostheses," in *Proc. IEEE 16th Int. Conf. Rehabil. Robot.*, 2019, pp. 727–733.

[12] H. F. Maqbool, M. A. B. Husman, M. I. Awad, A. Abouhossein, N. Iqbal, and A. A. Dehghani-Sanij, "A real-time gait event detection for lower limb prosthesis control and evaluation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 9, pp. 1500–1509, Sep. 2017.

[13] D.-X. Liu, X. Wu, W. Du, C. Wang, C. Chen, and T. Xu, "Deep spatial-temporal model for rehabilitation gait: Optimal trajectory generation for knee joint of lower-limb exoskeleton," *Assembly Automat.*, vol. 37, pp. 369–378, 2017.

[14] F.-Y. Liang *et al.*, "Online adaptive and LSTM-based trajectory generation of lower limb exoskeletons for stroke rehabilitation," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2018, pp. 27–32.

[15] M. Mundt *et al.*, "Prediction of lower limb joint angles and moments during gait using artificial neural networks," *Med. Biol. Eng. Comput.*, vol. 58, no. 1, pp. 211–225, 2020.

[16] B. Su and E. M. Gutierrez-Farewik, "Gait trajectory and gait phase prediction based on an LSTM network," *Sensors*, vol. 20, no. 24, 2020, Art. no. 7127.

[17] J. He, Z. Guo, Z. Shao, J. Zhao, and G. Dan, "An LSTM-based prediction method for lower limb intention perception by integrative analysis of kinect visual signal," *J. Healthcare Eng.*, vol. 2020, 2020, Art. no. 8024789.

[18] A. Zaroug, D. T. Lai, K. Mudie, and R. Begg, "Lower limb kinematics trajectory prediction using long short-term memory neural networks," *Front. Bioeng. Biotechnol.*, vol. 8, 2020, Art. no. 362.

[19] B. Fang *et al.*, "Gait neural network for human-exoskeleton interaction," *Front. Neurorobot.*, vol. 14, pp. 58, 2020.

[20] J. C. Alcaraz, S. Moghaddamnia, and J. Peissig, "Efficiency of deep neural networks for joint angle modeling in digital gait assessment," *EURASIP J. Adv. Signal Process.*, vol. 2021, no. 1, pp. 1–20, 2021.

[21] A. V. D. Oord *et al.*, "WaveNet: A generative model for raw audio." SSW, vol. 125, pp. 2, 2016, *arXiv:1609.03499*.

[22] C. Lea, M. D. Flynn, R. Vidal, A. Reiter, and G. D. Hager, "Temporal convolutional networks for action segmentation and detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 156–165.

[23] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.

[24] J. Yan, L. Mu, L. Wang, R. Ranjan, and A. Y. Zomaya, "Temporal convolutional networks for the advance prediction of ENSO," *Sci. Rep.*, vol. 10, no. 1, pp. 1–15, 2020.

[25] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *3rd Int. Conf. Learn. Representations*, May. 2015.

[26] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artif. Intell. Statist.*, 2010, pp. 249–256.

[27] S. K. Kumar, "On weight initialization in deep neural networks," 2017, *arXiv:1704.08863*.

[28] M. Abadi *et al.*, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. [Online]. Available: tensorflow.org

[29] L. Shuangfeng, "Tensorflow lite: On-device machine learning framework," *J. Comput. Res. Develop.*, vol. 57, no. 9, 2020, Art. no. 1839.

Sharmita Dey received the master's degree in computational logic from Institute of Artificial Intelligence, Department of Computer Science, Technical University Dresden, Dresden, Germany. She is currently working toward the Ph.D. degree in learning-based approaches to biomimetic prostheses/orthoses control with the Applied Rehabilitation Technology Lab, Department of Computer Science, University of Goettingen, Goettingen, Germany .

She is currently a Research Affiliate with the NASA Jet Propulsion Laboratory working in intersection of Artificial Intelligence and Robotics. Her research interests include developing machine-learning-based strategies for addressing the gait and locomotion of intelligent context-aware agents, pattern recognition, machine learning, robotics, and human–machine interface control.

Arndt F. Schilling studied medicine and molecular biology from Göttingen and Hamburg, Germany. He received the license to practice medicine in 2000 and the Dr. med. degree in control of bone remodeling from Hamburg University Medical Center, Hamburg, Germany, in 2003. Here he also received the venia legendi in 2010.

He held Professorships with the Technical University Hamburg-Harburg and Technical University Munich, Munich, Germany, from 2008 to 2016 and since then leads the R&D department with the Department of Trauma Surgery Orthopedics and Plastic Surgery, University Medical Center Gttingen, Gttingen, Germany, where he established the Applied Rehabilitation Technology Lab in 2017. He has authored and coauthored more than 100 papers and patents in his research areas. His research interests include physiology and pathology of the locomotor system. Dr. Schilling was the recipient of the numerous scientific awards in his research areas.