Boosting Personalized Musculoskeletal Modeling With Physics-Informed Knowledge Transfer

Jie Zhang, *Member, IEEE*, Yihui Zhao, Tianzhe Bao, Zhenhong Li, *Member, IEEE*, Kun Qian, Alejandro F. Frangi, *Fellow, IEEE*, Sheng Quan Xie, *Senior Member, IEEE*, and Zhi-Qiang Zhang, *Member, IEEE*

***Abstract*— Data-driven methods have become increasingly more prominent for musculoskeletal modeling due to their conceptually intuitive simple and fast implementation. However, the performance of a pretrained data-driven model using the data from specific subject(s) may be seriously degraded when validated using the data from a new subject, hindering the utility of the personalized musculoskeletal model in clinical applications. This article develops an active physics-informed deep transfer learning framework to enhance the dynamic tracking capability of the musculoskeletal model on the unseen data. The salient advantages of the proposed framework are twofold. First, for the generic model, physics-based domain knowledge is embedded into the loss function of the data-driven model as soft constraints to penalize/regularize the data-driven model. Second, for the personalized model, the parameters relating to the feature extraction will be directly inherited from the generic model, and only the parameters relating to the subject-specific inference will be fine-tuned by jointly minimizing the conventional data prediction loss and the modified physics-based loss. In this article, we use the synchronous muscle forces and joint kinematics prediction from surface electromyogram (sEMG) as the exemplar to illustrate the proposed framework. Moreover, convolutional neural network (CNN) is employed as the deep neural network to implement the proposed framework, and the physics law between muscle forces and joint kinematics is utilized as the soft constraints. Results of comprehensive experiments on a self-collected dataset from eight healthy subjects indicate the effectiveness and great generalization of the proposed framework.**

***Index Terms*— Personalized musculoskeletal model, physics-informed deep transfer learning, surface electromyogram (sEMG), wrist muscle forces and joint kinematics estimation.**

Manuscript received 26 July 2022; revised 25 October 2022; accepted 14 November 2022. Date of publication 8 December 2022; date of current version 11 January 2023. This work was supported in part by the U.K. Engineering and Physical Sciences Research Council (EPSRC) under Grant EP/S019219/1 and Grant EP/V057782/1 and in part by the European Union (EU) Marie Curie Individual Fellowship under Grant 101023097. The Associate Editor coordinating the review process was Dr. Adam G. Polak. *(Corresponding authors: Zhi-Qiang Zhang; Alejandro F. Frangi.)*

Jie Zhang is with the School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT Leeds, U.K., and also with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: eenjz@leeds.ac.uk).

Yihui Zhao, Zhenhong Li, Kun Qian, Sheng Quan Xie, and Zhi-Qiang Zhang are with the School of Electronic and Electrical Engineering, University of Leeds, LS2 9JT Leeds, U.K. (e-mail: eenyzhao@leeds.ac.uk; z.h.li@leeds.ac.uk; fbskq@leeds.ac.uk; s.q.xie@leeds.ac.uk; z.zhang3@leeds.ac.uk).

Tianzhe Bao is with the School of Rehabilitation Sciences and Engineering, University of Health and Rehabilitation Sciences, Qingdao 261000, China (e-mail: tianzhe.bao@uor.edu.cn).

Alejandro F. Frangi is with the School of Computing, University of Leeds, LS2 9JT Leeds, U.K., also with the Alan Turing Institute, NW1 2DB London, U.K., and also with the Department of Electrical Engineering, KU Leuven, 3000 Leuven, Belgium (e-mail: a.frangi@leeds.ac.uk).

Digital Object Identifier 10.1109/TIM.2022.3227604

I. Introduction

COMPUTATIONAL musculoskeletal modeling aims to understand the mechanism of the nervous system to learn and adapts to physiological modifications, which is critical to various clinical applications, such as planning rehabilitative treatments, prostheses and robotics control, and designing assistive devices [1], [2], [3]. Physics-based musculoskeletal modeling methods could interpret transformation among neural excitation, muscle dynamics, and joint kinematics and kinetics using experimental measurements from markers and sensors [4], [5]. Physiological quantities, e.g., muscle forces and joint moment, could be successively estimated with electromyograms (EMGs), foot-ground reaction forces (GRFs), segmental body kinematics, and so on. However, such set of methods is time-consuming and slow, especially for complex models in high-dimensional spaces, hindering the large-scale implementation in real-time application scenarios [6].

To tackle the slowness of physics-based musculoskeletal modeling methods, machine-/deep-learning-based data-driven methods are recently used to build mappings between EMGs, and muscle forces/joint kinematics [7], [8], [9], [10], [11] to reduce the time consumption in musculoskeletal model building [12]. For instance, Tang et al. [13] developed a modified framework to estimate the muscle force using surface EMG (sEMG) based on an encoder–decoder network. Wimalasena et al. [14] proposed a large-scale unsupervised deep learning method to achieve the spatial and temporal representations of the multimuscle activation from EMG measurements. Burton et al. [15] employed four machine/deep learning methods, including convolutional neural network (CNN), recurrent neural network (RNN), fully connected neural network, and principal component regression, to estimate the lower extremity muscle and joint contact forces from joint kinematics, GRFs, and anthropometrics. However, all these data-driven methods would require a large number of training data to make the model personalized. The performance of a pretrained data-driven model using the data from specific subject(s) may be seriously degraded when validated using the data from a new subject.

The targeted outputs of the data-driven model are normally derived from the physics-based model using static optimization, and it is very time-consuming to get a large number of training data. To overcome this issue, transfer learning has been explored recently to investigate how to make a pretrained model work for a new subject with only limited training data

from him/her. Dao [16] developed a deep transfer learning strategy based on long short-term memory (LSTM) to predict skeletal muscle forces from kinematic measurements during a gait cycle. Bao et al. [17] proposed a novel two-stream CNN for supervised domain adaptation to reduce domain shift effects on wrist kinematics estimation in the intersubject circumstance. Wang et al. [18] utilized EMG signals and acceleration data as the multimodal input of the deep learning model to enhance the adaptability to the effects of arm movements, and transfer learning was considered to accelerate the convergence of deep learning model and avoid overfitting problem. Kim et al. [19] proposed a modified subject-transfer framework, in which supportive CNN classifiers were pretrained using EMG signals from several subjects, and then transfer learning was utilized to fine-tune these classifiers to enhance the robustness of the proposed method in terms of intrauser variability. However, state-of-the-art methods are actually "black-box," and its narrow focus on the input–output mappings may be inherently deterministic without considering the explicit physics modeling of the underlying neuromechanical processes [20], [21], [22], [23], [24], [25].

To tackle the aforementioned issues, a physics-informed deep transfer learning framework for personalized musculoskeletal modeling is proposed to learn the mappings from EMGs to muscle forces and joint kinematics in this article. The proposed framework consists of a generic network and a personalized network. Specifically, a generic network is first trained with sEMG measurements from several subjects by minimizing the loss function, which jointly considers the minimization of the conventional data prediction loss and a physics-based loss. The parameters relating to the generic information of subjects are then directly shared to the personalized network as the backbone. After that, the parameters relating to the subject-specific information in the personalized network could be achieved with limited available subject-specific data based on the modified loss function. In this manner, the derived personalized network will contain mappings between kinematic measurements, and internal forces and muscle activations with the shared parameters from the generic network, which could help the personalized network hold the generalization of the generic network achieved from subjects' data. More importantly, with many parameters frozen, over-fitting problem could be avoided during the knowledge transfer phase. sEMG-based wrist muscle forces and joint kinematics estimation is considered as the exemplar to verify the feasibility of the proposed framework. Additionally, CNN is employed as the deep neural network to implement the proposed framework. The main contributions can be summarized as follows.

1) A deep learning and physical knowledge integration-based knowledge transfer framework for personalized musculoskeletal modeling is developed. The twin neural networks architecture, i.e., the generic network and the personalized network, could significantly reduce the training data required for any new subject.
2) A modified loss function is designed by embedding physics law, in which the physics-based domain knowledge is utilized as soft constraints to penalize/regularize the loss function to enhance the robustness and prediction performance of deep neural networks, and make the intermediate functional relationships of the deep learning-based "black-box" modeling be reflected and controlled by the underlying physical mechanisms.
3) Comprehensive experiments on a self-collected dataset from eight healthy subjects are performed, and the results demonstrate that the proposed framework could achieve better prediction performance than baseline methods with less training data and significantly reduce the time-consumption on model retraining.

The remaining of the article is organized as follows: Methodology is detailed in Section II, including the problem formulation, main framework of the proposed physics-informed deep transfer learning method, design of the generic network, and the personalized network. Materials and experimental methods are presented in Section III. Experimental results are reported in Section IV, followed by discussions in Section V. Finally, conclusions are given in Section VI.

II. Methodology

A. Problem Formulation

In this article, we will use sEMG measurements to predict human muscle force and joint kinematics as the exemplar. As shown in Fig. 1, after building a generic model from EMG_t^n and the corresponding time steps t of m subjects to predict muscle forces F_t^n and joint angles $\theta_t (n = 1, \ldots, N, t = 1, \ldots, T)$, how can we apply this model to a new subject $m + 1$. Here, N is the number of muscles at the joint of interest and T denotes the total sample number. Since achieving the ground truths of these subject-specific data is time-consuming and labor-intensive, it would be unwise to ignore the training data from the first m subjects and extract the large amount of ground-truth data for the $(m+1)$th subject. In Sections II-B–II-D, we will introduce a twin physics-informed deep neural networks strategy to make the generic model personalized for the $(m + 1)$th subject with limited training data from him/her.

B. Main Framework of Twin Physics-Informed Deep Neural Networks for Personalized Musculoskeletal Modeling

Fig. 1 depicts the main framework of the proposed physics-informed deep transfer learning method for musculoskeletal modeling, in the context of muscle forces and joint kinematics estimation from sEMG. As shown in Fig. 1, it consists of two deep neural networks, including a physics-informed generic network and a physics-informed personalized network.

In the generic network, CNN is employed to automatically extract more discriminative features to build mappings between sEMGs and joint motions/muscle forces. Specifically, the measured sEMGs from several subjects and the corresponding time steps are first fed into CNN, and then the predicted joint angles and muscle forces are achieved with the extracted features. Such predictions should also satisfy the physical equation of motion, which is then taken as the soft constraint to penalize/regularize the loss function

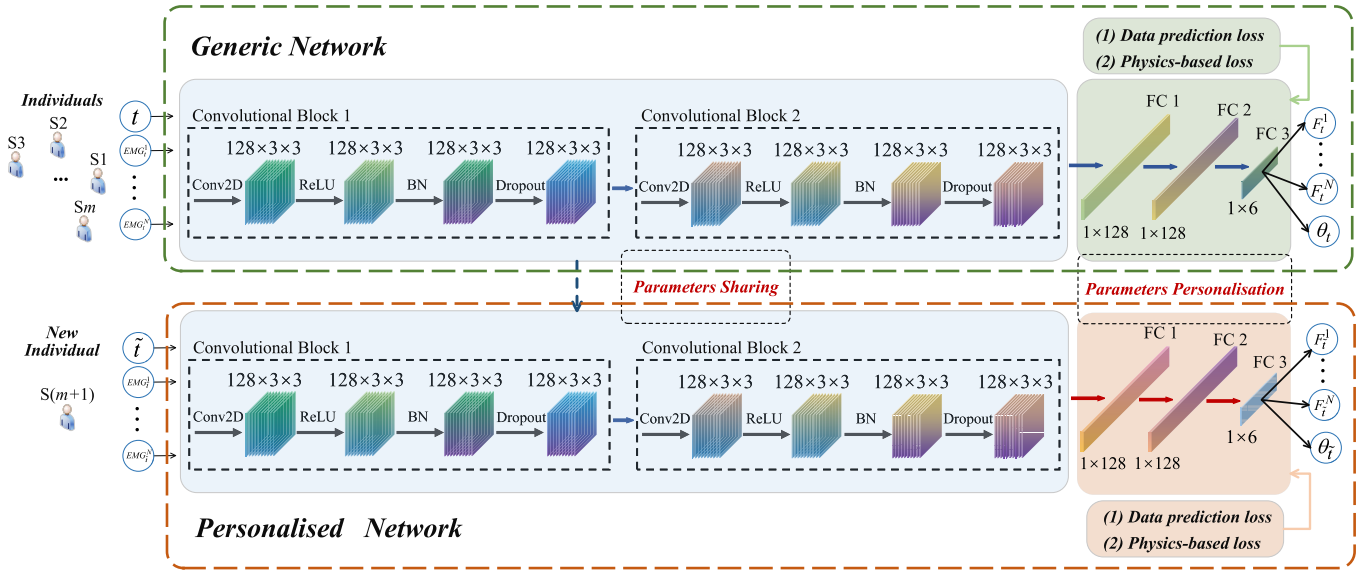


Fig. 1. Main framework of the proposed physics-informed deep transfer learning method. In this study, inputs of the generic network are sEMG measurements and time steps of m subjects, while its outputs are muscle forces F_t^n and joint angles $\theta_t (n = 1, \ldots, N, t = 1, \ldots, T)$. Inputs of the personalized network are sEMG measurements and the corresponding time steps of the $(m+1)$th subject, and its outputs are muscle forces $F_{\tilde{t}}^n$ and joint angles $\theta_{\tilde{t}} (\tilde{t} = 1, \ldots, \tilde{T})$ of the subject. n denotes the number of muscles at the joint of interest; t and $\tilde{t}$ are the time steps of the m subjects and the $(m+1)$th subject, respectively. We use the difference of background color of each component to demonstrate the parameter sharing and parameter personalisation phases.

of CNN. In this manner, the modified loss function of the physics-informed generic network jointly minimizes two components, i.e., the conventional mean square error (MSE) loss and the physics-based loss. When the new measured sEMGs of another subject are available, the parameters relating to the generic information of subjects, i.e., the parameters in the feature extraction phase, in the generic network will be directly shared to the personalized network as the backbone, which could help the personalized network hold mappings between kinematic measurements, and internal forces and muscle activations, and also guarantee the generalization. After that, the parameters relating to the subject-specific information, e.g., the parameters in fully connected layers, in the personalized network could be achieved through fine-tuning the corresponding parameters in the generic network with the limited subject-specific data by jointly considering the data prediction error and the physical constraint.

In this manner, the derived personalized network will contain mappings between kinematic measurements, and internal forces and muscle activations with the shared parameters from the generic network. More importantly, with many parameters frozen, over-fitting problem can be avoided during the knowledge transfer phase, and it can also significantly reduce the number of required training data and time consumption.

C. Design of Generic Network

We detail the network architecture and the modified loss function of the generic network, respectively.

1) Architecture of Generic Network: The generic network is a CNN with two convolutional blocks, two fully connected blocks, and one regression block, i.e., FC 3 in Fig. 1. To be specific, each convolutional block consists of one convolutional layer, one ReLU layer, one batch normalization layer, and one dropout layer. In the convolutional block, the kernel size is 3, the boundary padding is 3, and the stride is 1. There are 128 kernels in the convolutional layer and a ReLU layer is added subsequently to the convolutional layer. The batch normalization layer is employed for mitigating alternation made by the convolutonal layer. Each fully connected block consists of one ReLU layer, one normalization layer, and one dropout layer. The number of hidden nodes is 128. Outputs of the second fully connected block are then fed into the regression block for the muscle forces and joint kinematics estimation.

2) Loss Function of Generic Network: The loss function of the generic network consists of the MSE loss and the physics-based loss. The MSE loss is to minimize the MSE of the ground truth and prediction, while the physics-based loss preserves the physical constraints during movements

$$\mathcal{L} = \mathcal{L}_d + \mathcal{L}_p \tag{1}$$

$$\mathcal{L}_d = \mathrm{MSE}_{\mathrm{gen}}(F) + \mathrm{MSE}_{\mathrm{gen}}(\theta) \tag{2}$$

$$\mathcal{L}_p = \Theta(F, \theta) \tag{3}$$

where $\mathcal{L}_d$ represents the data prediction loss of the muscle force and the joint angle, $\mathcal{L}_d$ is the loss function imposed by the physics law, which could be utilized to penalize/regularize the generic network for performance enhancement. $\Theta(F, \theta)$ denotes the function of predicted variables.

The MSE loss can be represented as

$$\mathrm{MSE}_{\mathrm{gen}}(F) = \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{N} (F_t^n - \hat{F}_t^n)^2 \tag{4}$$

$$\mathrm{MSE}_{\mathrm{gen}}(\theta) = \frac{1}{T} \sum_{t=1}^{T} (\theta_t - \hat{\theta}_t)^2 \tag{5}$$

where F_t^n and θ_t are the force of muscle n and the joint angle at time step t, and $\hat{F}_t^n$ and $\hat{\theta}_t$ denote the corresponding predictions, respectively. Furthermore, T is the total sample number of subjects, and N is the number of muscles at the joint of interest.

The equation of motion, which could reflect underlying relationships among the muscle force and kinematics, is utilized to design the physics-based loss

$$\Theta(F,\theta)=\frac{1}{T}\sum_{t=1}^{T}(M(\theta_t)\ddot{\theta}_t+C(\theta_t,\dot{\theta}_t)+G(\theta_t)-\tau_t)^2 \quad (6)$$

where $M(\theta_t)$, $C(\theta_t,\dot{\theta}_t)$, $G(\theta_t)$, and θ_t are the mass matrix, the Centrifugal and Coriolis force, the gravity, and the joint angle, respectively. Additionally, τ_t is the joint torque, which can be calculated through the summation of the product of the moment arm and muscle force

$$\tau_t=\sum_{n=1}^{N} r_n F_t^n \quad (7)$$

where r_n is the moment arm of the muscle n, which is exported from OpenSim.

D. Design of Personalized Network

The personalized network has the similar architecture with the generic network, thus we only demonstrate its modified loss function and the knowledge transfer process for musculoskeletal model personalisation.

1) Loss Function of Personalized Network: The loss function of the personalized network also consists of the MSE loss and the physics-based loss

$$\tilde{\mathcal{L}}=\tilde{\mathcal{L}}_d+\tilde{\mathcal{L}}_p \quad (8)$$

$$\tilde{\mathcal{L}}_d=\mathrm{MSE}_{\mathrm{per}}(F)+\mathrm{MSE}_{\mathrm{per}}(\theta) \quad (9)$$

$$\tilde{\mathcal{L}}_p=\Xi(F,\theta) \quad (10)$$

where $\tilde{\mathcal{L}}_d$ represents the prediction loss of the muscle force and the joint angle, and $\tilde{\mathcal{L}}_p$ is the loss function imposed by the physics law.

Specifically, the MSE loss is

$$\mathrm{MSE}_{\mathrm{per}}(F)=\frac{1}{\tilde{T}}\sum_{\tilde{t}=1}^{\tilde{T}}\sum_{n=1}^{N}(F_{\tilde{t}}^n-\hat{F}_{\tilde{t}}^n)^2 \quad (11)$$

$$\mathrm{MSE}_{\mathrm{per}}(\theta)=\frac{1}{\tilde{T}}\sum_{\tilde{t}=1}^{\tilde{T}}(\theta_{\tilde{t}}-\hat{\theta}_{\tilde{t}})^2 \quad (12)$$

where $F_{\tilde{t}}^n$ and $\theta_{\tilde{t}}$ are the force of the muscle n and the joint angle at time step $\tilde{t}$, and $\hat{F}_{\tilde{t}}^n$ and $\hat{\theta}_{\tilde{t}}$ denote the corresponding predictions, respectively. Furthermore, $\tilde{T}$ is the total sample number of the new subject.

Similar to the loss function of the generic network, the equation of motion is also employed as the physics-based loss of the personalized network

$$\Xi(F,\theta)=\frac{1}{\tilde{T}}\sum_{\tilde{t}=1}^{\tilde{T}}(M(\theta_{\tilde{t}})\ddot{\theta}_{\tilde{t}}+C(\theta_{\tilde{t}},\dot{\theta}_{\tilde{t}})+G(\theta_{\tilde{t}})-\tau_{\tilde{t}})^2 \quad (13)$$

where $M(\theta_{\tilde{t}})$, $C(\theta_{\tilde{t}},\dot{\theta}_{\tilde{t}})$, $G(\theta_{\tilde{t}})$, and $\theta_{\tilde{t}}$ are the mass matrix, the Centrifugal and Coriolis force, the gravity, and the joint angle, respectively. Similarly, $\tau_{\tilde{t}}$ could be calculated as

$$\tau_{\tilde{t}}=\sum_{n=1}^{N} r_n F_{\tilde{t}}^n. \quad (14)$$

2) Musculoskeletal Model Personalisation: In order to transfer the useful knowledge to the personalized network, the parameters relating to the generic information of subjects in the generic network are first fixed and then directly shared to the personalized network, enabling the updated personalized network to hold mapping relationships between sEMG measurements, and muscle forces and joint angles of subjects. These inherited parameters could help enhance the generalization of the personalized network. In the parameter penalisation phase, optimal parameters in fully connected layers, which actually relate to the subject-specific information, could be achieved by minimizing the modified loss function only with the limited available subject-specific data as inputs of the personalized network.

III. Materials and Experimental Settings

In this section, data collection, data preprocessing, baseline methods, and evaluation criteria are detailed, respectively.

A. Data Collection

Approved by the MaPS and Engineering Joint Faculty Research Ethics Committee of University of Leeds (MEEC 18-002), eight able-bodied subjects, including four males and four females, aged 20–30, were recruited to record data for the experiment. All subjects gave signed consent. Specifically, subjects were informed to maintain a fully straight torso with the 90° abducted shoulder and the 90° flexed elbow joint. The continuous wrist flexion/extension motion was recorded by the VICON motion capture system. The joint motions were computed by the upper limb model using 16 reflective markers (sampled at 250 Hz). In addition, sEMGs were measured by Avanti Sensors (sampled at 2000 Hz) from the main wrist muscles, i.e., the flexor carpi radialis (FCR), the flexor carpi ulnaris (FCU), the extensor carpi radialis longus (ECRL), the extensor carpi radialis brevis (ECRB), and the extensor carpi ulnaris (ECU), respectively. The electrodes were allocated by palpation and evaluated by performing contraction while looking at the signal before the experiment. Moreover, sEMGs and motion data were synchronized and resampled at 1000 Hz. Five repetitive trials were done for each subject, and a 3-min break was given between trials to prevent the muscle fatigue. We collected 80 000 samples in total (10 000 samples from each subjects).

B. Data Preprocessing

The measured sEMGs were then bandpass filtered (20 and 450 Hz), fully rectified, and low-pass filtered (6 Hz), respectively. After that, they were normalized concerning the maximum voluntary contraction recorded before the experiment, resulting in the enveloped EMG. The markers' data were

used to compute the wrist kinematics via the inverse kinematic tool, and the joint torque and wrist muscle forces were then achieved from the inverse dynamic and the computed muscle control tools, which ensured the computed motion was consistent with the measured joint motion. Each wrist motion trial consists of time steps, filtered sEMGs, wrist muscle forces, and wrist joint angles, respectively.

C. Baseline Methods

To verify the effectiveness of the proposed physics-informed deep transfer learning framework, several methods are chosen as baseline methods for the comparison. In addition, we define the generic domain as sufficient labeled sEMG measurements collected from several subjects could be provided for the generic network training, denoted by D_G, while personalized domain as only limited labeled sEMG measurements collected from another specific subject are available for musculoskeletal model training or personalized network training, denoted by D_P.

1) CNN With Training Data From D_G (CNN-1): CNN-1 only involves conventional supervised learning with two convolutional blocks, two fully connected blocks, and one regression block. The measured sEMGs from D_G are for CNN-1 training, while sEMGs from D_P for testing. Stochastic gradient descent with momentum optimizer is employed for CNN-1 training, the batch size is set as 1, the maximum iteration is set as 2000, and the initial learning rate is 0.001.

2) CNN With Training Data From D_P (CNN-2): CNN-2 also not involves knowledge transfer and is with the same architecture as CNN-1, but its training data are the sEMGs measured from the new subject in D_P. Its training strategy and parameter settings are the same as CNN-1.

3) CNN With Knowledge Transfer (CNN-KT): CNN-KT shares the same architecture, training strategy, and parameter settings as CNN-1 and CNN-2. Differently, CNN-KT is first pretrained using the data from D_G, and then its parameters relating to the generic information of subjects, i.e., the parameters in the feature extraction phase, are fixed. After that, the data from the new subject are utilized for network retraining, and the parameters relating to the subject-specific information, i.e., the parameters in fully connected layers, are achieved by fine-tuning strategy.

4) Physics-Informed CNN With Training Data From D_G (Pi-CNN-1): Pi-CNN-1 has the same architecture and training data as CNN-1. However, its loss function is to jointly minimize the MSE loss and the physics-based loss, i.e., the loss function illustrated in (1).

5) Physics-Informed CNN With Training Data From D_P (Pi-CNN-2): Pi-CNN-2 has the same architecture and training data as CNN-2, and its loss function is the same as the loss function of Pi-CNN-1.

D. Evaluation Criteria

In this article, two commonly used evaluation criteria, including root mean square error (RMSE) and Pearson's correlation coefficient (CC), are considered to quantify the performance of the proposed framework. To be specific, RMSE is

$$\mathrm{RMSE} = \sqrt{\frac{1}{\tilde{T}}\sum_{\tilde{t}=1}^{\tilde{T}}\left(y_{\tilde{t}} - \hat{y}_{\tilde{t}}\right)^2} \tag{15}$$

where $y_{\tilde{t}}$ and $\hat{y}_{\tilde{t}}$ indicate the ground truth and the corresponding predicted value, respectively.

CC is defined as

$$\mathrm{CC} = \frac{\sum_{\tilde{t}=1}^{\tilde{T}}\left(y_{\tilde{t}} - \overline{y_{\tilde{t}}}\right)\left(y_{\tilde{t}} - \overline{\hat{y}_{\tilde{t}}}\right)}{\sqrt{\sum_{\tilde{t}=1}^{\tilde{T}}\left(y_{\tilde{t}} - \overline{y_{\tilde{t}}}\right)^2}\sqrt{\sum_{\tilde{t}=1}^{\tilde{T}}\left(y_{\tilde{t}} - \overline{\hat{y}_{\tilde{t}}}\right)^2}} \tag{16}$$

where $\overline{y_{\tilde{t}}}$ and $\overline{\hat{y}_{\tilde{t}}}$ are the mean of the ground truth and predicted value, respectively.

IV. Results

In this section, the effectiveness of the proposed physics-informed deep transfer learning framework for musculoskeletal modeling penalisation is verified. Specifically, the training process of the proposed framework is first shown to demonstrate its convergence. Performance evaluation for the single-to-single and the multiple-to-single scenarios are then performed to depict the predicted performance of the proposed framework. It should be noted that all the training of the proposed framework and baseline methods is carried out on a workstation with GeForce RTX 2080 Ti graphic cards and 128GB RAM.

A. Training Process of the Proposed Framework

In the experiments, we use 60% of the data to train the proposed framework and all the selected comparison methods, 20% as the validation dataset, and the rest 20% as the testing dataset. During the training phase, we set the batch size as 1, and the embedded CNN is trained by stochastic gradient descent with momentum. The maximum iteration is set as 2000, the initial learning rate is 0.001, and the dropout rate is 0.5. The separate losses of wrist muscle forces and wrist angle estimation of Pi-CNN-1 and the proposed framework (also denoted by Pi-CNN-KT in the experiments in the following) during the training are depicted in Fig. 2. Observed from Fig. 2, the proposed framework has faster convergence speed than Pi-CNN-1, and its losses are more stable with less local oscillations than the ones of Pi-CNN-1. Compared with Pi-CNN-1, the twin neural networks mechanism in the proposed framework could accelerate the convergence of networks: for example, the loss of ECU in Fig. 2(b) drops significantly at the beginning of the training, and then converges.

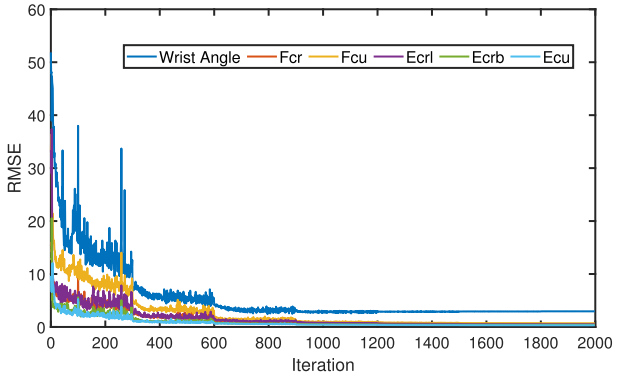
B. Performance Evaluation in Single-to-Single Scenario

We first evaluate the performance of the proposed framework in the single-to-single scenario. The generic network is first trained using the data from one specific subject, i.e., the mth subject ($m = 1, 2, \ldots, 7$), and then the personalized network is fine-tuned using the data from the eighth subject. Fig. 3 illustrates the representative predicted results of the proposed framework and baseline methods, including the wrist flexion

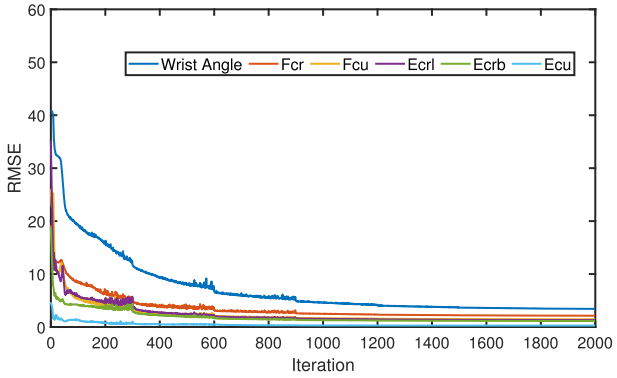
TABLE I

DETAILED RMSES AND CCS OF THE PROPOSED FRAMEWORK AND BASELINE METHODS IN THE SINGLE-TO-SINGLE SCENARIO

Outputs	Methods	S1	S2	S3	S4	S5	S6	S7	Outputs	Methods	S1	S2	S3	S4	S5	S6	S7
Angle	**Pi-CNN-KT**	**10.67/0.97**	**9.25/0.97**	**11.95/0.96**	**9.81/0.97**	**9.94/0.97**	**10.35/0.97**	**11.33/0.96**	FCR	**Pi-CNN-KT**	**4.45/0.95**	**4.02/0.96**	**4.17/0.96**	**3.99/0.97**	**4.21/0.96**	**4.32/0.96**	**4.02/0.96**
	CNN-KT	12.15/0.95	13.96/0.93	15.03/0.94	13.97/0.94	11.27/0.95	12.27/0.95	12.41/0.94		CNN-KT	4.53/0.95	4.41/0.95	4.95/0.95	4.27/0.96	4.79/0.94	4.61/0.95	4.79/0.95
	Pi-CNN-1	16.55/0.93	17.80/0.93	15.39/0.93	16.53/0.93	14.91/0.94	16.27/0.93	15.99/0.93		Pi-CNN-1	5.89/0.92	4.72/0.95	4.39/0.96	5.77/0.93	5.72/0.92	5.66/0.93	5.70/0.93
	CNN-1	67.21/0.31	59.26/0.35	71.22/0.29	70.09/0.31	62.37/0.32	76.29/0.28	65.52/0.32		CNN-1	14.82/0.57	12.98/0.62	17.66/0.52	15.93/0.57	13.98/0.61	14.56/0.59	13.87/0.61
	Pi-CNN-2	14.95/0.94	16.21/0.93	17.98/0.93	15.63/0.93	12.96/0.95	14.56/0.94	13.29/0.94		Pi-CNN-2	4.82/0.95	4.69/0.95	5.63/0.93	4.81/0.94	4.98/0.94	5.03/0.95	4.91/0.95
	CNN-2	17.24/0.93	19.22/0.92	18.31/0.93	18.09/0.92	14.15/0.94	16.98/0.98	16.90/0.93		CNN-2	5.79/0.93	5.66/0.93	6.07/0.93	5.75/0.93	5.95/0.93	5.62/0.93	5.53/0.94
FCU	**Pi-CNN-KT**	**2.82/0.99**	**3.09/0.98**	**4.51/0.98**	**2.99/0.98**	**3.29/0.98**	**3.02/0.98**	**2.77/0.99**	ECRL	**Pi-CNN-KT**	**3.13/0.99**	**4.19/0.97**	**3.83/0.99**	**3.29/0.99**	**3.02/0.99**	**3.22/0.99**	**3.09/0.99**
	CNN-KT	3.05/0.98	3.82/0.98	5.23/0.97	3.93/0.98	3.73/0.98	3.72/0.98	3.19/0.98		CNN-KT	3.61/0.99	4.92/0.95	4.20/0.97	3.55/0.99	3.94/0.98	3.69/0.99	3.67/0.99
	Pi-CNN-1	4.21/0.98	4.53/0.98	5.89/0.96	4.10/0.98	3.96/0.98	4.69/0.97	4.11/0.98		Pi-CNN-1	4.68/0.98	5.87/0.95	5.31/0.93	4.77/0.97	4.31/0.97	4.06/0.98	3.98/0.97
	CNN-1	21.41/0.83	23.60/0.82	31.99/0.80	19.57/0.83	18.28/0.86	20.33/0.83	22.30/0.82		CNN-1	26.81/0.85	30.95/0.82	27.61/0.86	29.30/0.87	27.90/0.86	24.91/0.85	22.87/0.83
	Pi-CNN-2	3.34/0.98	4.27/0.98	5.78/0.97	3.91/0.98	4.20/0.98	3.99/0.98	3.51/0.98		Pi-CNN-2	4.21/0.97	5.51/0.93	4.72/0.97	4.56/0.98	4.07/0.98	3.98/0.98	3.92/0.98
	CNN-2	4.36/0.98	4.99/0.97	6.31/0.96	4.25/0.98	4.07/0.98	4.51/0.97	4.27/0.98		CNN-2	4.59/0.98	5.72/0.94	5.17/0.95	4.72/0.98	4.19/0.97	4.26/0.97	4.05/0.97
ECRB	**Pi-CNN-KT**	**2.45/0.98**	**2.65/0.98**	**2.32/0.98**	**2.29/0.98**	**2.02/0.98**	**2.36/0.97**	**2.22/0.98**	ECU	**Pi-CNN-KT**	**0.61/0.98**	**0.58/0.97**	**0.62/0.98**	**0.55/0.98**	**0.63/0.98**	**0.66/0.98**	**0.57/0.98**
	CNN-KT	2.72/0.97	2.89/0.97	2.66/0.98	2.98/0.97	2.58/0.98	2.77/0.98	2.57/0.98		CNN-KT	0.71/0.98	0.82/0.97	0.76/0.98	0.81/0.97	0.75/0.97	0.76/0.98	0.69/0.98
	Pi-CNN-1	3.65/0.96	3.89/0.96	3.26/0.97	3.98/0.94	3.55/0.96	3.27/0.97	3.39/0.96		Pi-CNN-1	0.86/0.96	0.99/0.97	0.81/0.97	0.92/0.97	0.85/0.97	0.79/0.97	0.82/0.97
	CNN-1	16.06/0.75	15.97/0.75	12.50/0.77	16.28/0.73	13.33/0.78	15.99/0.75	14.27/0.76		CNN-1	4.21/0.69	6.30/0.59	4.39/0.69	4.37/0.70	4.14/0.71	4.06/0.75	4.13/0.70
	Pi-CNN-2	3.12/0.97	3.57/0.96	3.19/0.97	3.55/0.96	3.05/0.98	3.11/0.98	2.98/0.98		Pi-CNN-2	0.62/0.98	0.80/0.97	0.71/0.98	0.85/0.97	0.76/0.98	0.72/0.98	0.75/0.98
	CNN-2	3.29/0.97	4.01/0.95	3.21/0.97	4.26/0.95	3.79/0.97	3.50/0.96	3.20/0.97		CNN-2	0.93/0.96	1.16/0.93	0.99/0.97	0.91/0.96	0.82/0.97	0.80/0.98	0.86/0.98



(a)



(b)

Fig. 2. Convergence illustration of the proposed framework. The proposed framework has faster convergence speed with less local oscillations. (a) Separate losses of Pi-CNN-1. (b) Separate losses of Pi-CNN-KT.

angle, muscle force of FCR, muscle force of FCU, muscle force of ECRL, muscle force of ECRB, and muscle force of ECU, respectively. According to Fig. 3, CNN-1 achieves the worst predicted results, because it does not involve the knowledge transfer of the new subject, leading to its poor generalization on the unseen data. The predicted results of CNN-KT are better than CNN-1, indicating the effectiveness of the twin neural networks mechanism employed in the musculoskeletal model personalisation. In addition, Pi-CNN-1 achieves better predicted performance than CNN-1; it means that the proposed physics-based loss function could help enhance the performance. The proposed framework achieves the best predicted performance, which indicates that the combination of the physics-based domain knowledge embedding and the twin neural networks mechanism could significantly strengthen the performance of the data-driven musculoskeletal model, and enhance its robustness and generalization.

Table I lists the detailed RMSEs and CCs of the proposed framework and baseline methods in the single-to-single scenario. Observed from Table I, the proposed framework could achieve smaller RMSEs and higher CCs than baseline methods. To be specific, among the methods without physics-based domain knowledge embedding, i.e., CNN-1, CNN-2, and CNN-KT, CNN-KT has the smallest RMSEs and the highest CCs, because the transferred knowledge of the new subject makes the personalized network contain subject-specific information. Moreover, the performance of CNN-2 is better than CNN-1, the main reason is that CNN-1 does not learn any statistical characteristics of the new subject; it also indicates the poor generalization of the conventional data-driven musculoskeletal model on the unseen data. Finally, we can find that the proposed framework outperforms CNN-KT; it means that the physics-based domain knowledge in the modified loss function could provide informative constraints to penalize/regularize CNN utilized in the proposed framework for performance improvement.

To further verify the feasibility of the proposed framework, a pairwise analysis between the proposed method and each comparison method is considered. One-way analysis of variance (ANOVA) is conducted for statistical analysis of the proposed framework and baseline methods, where RMSE is the response variable. A post-hoc analysis using Tukey's Honest Significant Difference test is applied. The significance level is set at $p < 0.05$. Figs. 4 and 5 depict the RMSEs and CCs of the proposed framework and baseline methods; we can find that the proposed framework achieves better predicted performance no matter which subject's data are employed to train the generic network. Table II details the multiple

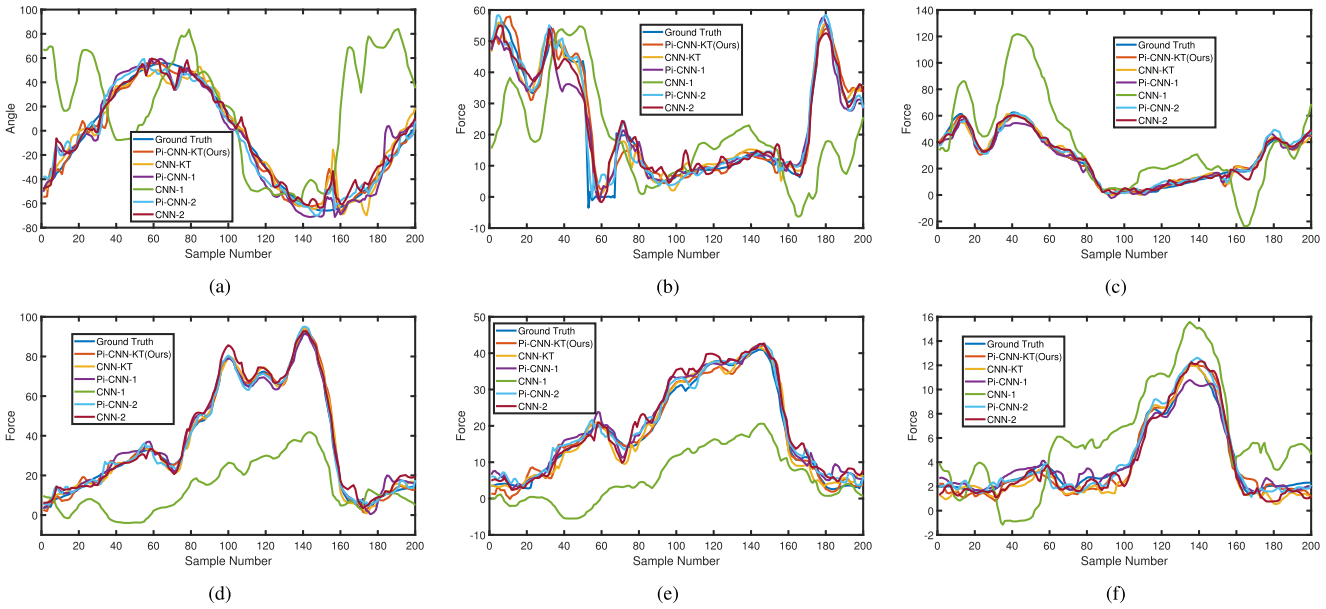


Fig. 3. Representative predicted results of the proposed framework and baseline methods in the single-to-single scenario. The predicted outputs are the wrist angle, muscle force of FCR, muscle force of FCU, muscle force of ECRL, muscle force of ECRB, and muscle force of ECU. Predicted results of (a) wrist angle, (b) FCR, (c) FCU, (d) ECRL, (e) ECRB, and (f) ECU.

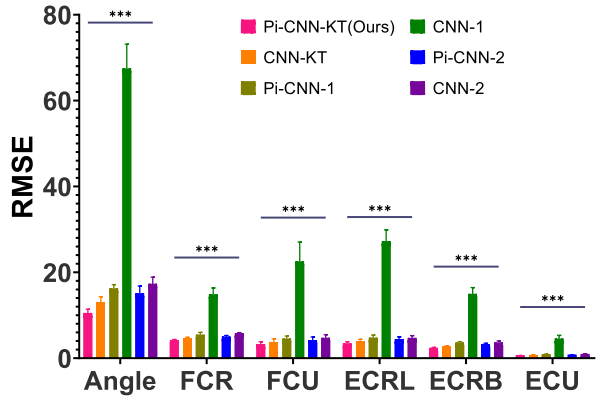


Fig. 4. RMSEs of the proposed framework and baseline methods with different subjects' data as the training samples for the generic network in the single-to-single scenario. The proposed framework achieves smaller RMSEs than baseline methods. The significance level is set as 0.05 ($^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$).

TABLE II
MULTIPLE COMPARISON ON RMSE BETWEEN THE PROPOSED FRAMEWORK AND BASELINE METHODS IN THE SINGLE-TO-SINGLE SCENARIO

RMSE							
Methods		**Angle**	**FCR**	**FCU**	**ECRL**	**ECRB**	**ECU**
CNN-1	**Pi-CNN-KT**	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Pi-CNN-1	**Pi-CNN-KT**	< 0.01	< 0.05	0.620	0.176	< 0.01	0.458
CNN-KT	**Pi-CNN-KT**	0.270	0.633	0.969	0.870	0.686	0.854
Pi-CNN-2	**Pi-CNN-KT**	< 0.05	0.137	0.844	0.390	0.071	0.891
CNN-2	**Pi-CNN-KT**	< 0.001	< 0.01	0.5	0.198	< 0.01	0.270
CC							
Methods		**Angle**	**FCR**	**FCU**	**ECRL**	**ECRB**	**ECU**
CNN-1	**Pi-CNN-KT**	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Pi-CNN-1	**Pi-CNN-KT**	< 0.001	< 0.05	0.457	< 0.05	< 0.05	0.852
CNN-KT	**Pi-CNN-KT**	< 0.01	0.667	0.848	0.865	0.999	0.999
Pi-CNN-2	**Pi-CNN-KT**	< 0.01	0.261	0.848	0.177	0.720	1
CNN-2	**Pi-CNN-KT**	< 0.001	< 0.01	0.291	0.062	< 0.05	0.778

comparison results in terms of RMSEs and CCs. The multiple comparison correction method is Dunnett's test. The proposed framework is regarded as the control group. The comparison results indicate that the proposed framework is statistically superior to baseline methods.

C. Performance Evaluation in Multiple-to-Single Scenario

Aside from the single-to-single scenario, we further evaluate the performance of the proposed framework in the multiple-to-single scenario. Specifically, we first randomly choose seven subjects from the eight subjects, and the generic network is then trained using the data from the chosen seven subjects, and the data from the rest one subject is utilized to fine-tune the personalized network.

Fig. 6 illustrates the representative predicted results of the proposed framework and baseline methods in the multiple-to-single scenario. Similar to the single-to-single scenario, comparing with baseline methods, the proposed framework could achieve more satisfactory predicted results, and fit the ground truth curves well. Table III details subject's RMSEs and average CCs of the proposed framework and baseline methods, where the six predicted outputs are normalized. Table IV details output's RMSEs and average CCs of the proposed framework and baseline methods, in which average values of the eight subjects' predicted outputs are calculated. According to Tables III and IV, the proposed framework still could achieve smaller RMSEs and higher CCs under different subjects' data for training the generic network and the personalized network, indicating its great tracking capability.

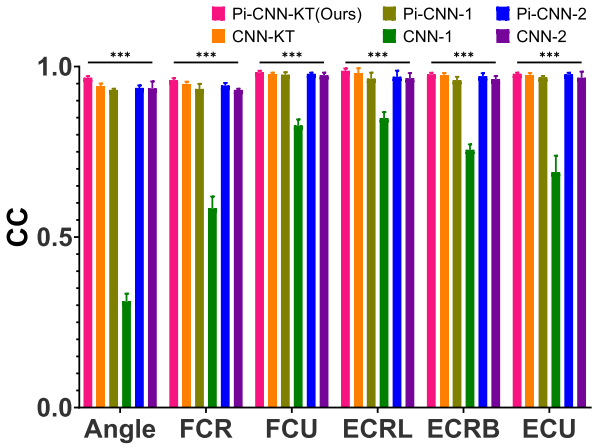


Fig. 5. CCs of the proposed framework and baseline methods with different subjects' data as the training samples for the generic network in the single-to-single scenario. The proposed framework achieves higher CCs than baseline methods. The significance level is set as 0.05 ($^{***}p < 0.001$, $^{**}p < 0.01$, and $^{*}p < 0.05$).

TABLE III

SUBJECT'S NORMALIZED RMSES AND AVERAGE CCS OF THE PROPOSED FRAMEWORK AND BASELINE METHODS IN THE MULTIPLE-TO-SINGLE SCENARIO

Metrics	Methods	S1	S2	S3	S4	S5	S6	S7	S8
Normalized RMSE	**Pi-CNN-KT**	**0.10**	**0.10**	**0.09**	**0.10**	**0.08**	**0.10**	**0.11**	**0.09**
	CNN-KT	0.13	0.11	0.09	0.13	0.12	0.12	0.14	0.12
	Pi-CNN-1	0.52	0.55	0.50	0.54	0.54	0.51	0.52	0.51
	CNN-1	0.43	0.41	0.39	0.42	0.43	0.46	0.45	0.45
	Pi-CNN-2	0.23	0.24	0.20	0.27	0.22	0.22	0.20	0.19
	CNN-2	0.19	0.17	0.20	0.15	0.17	0.16	0.17	0.16
CC	**Pi-CNN-KT**	**0.97**	**0.96**	**0.97**	**0.98**	**0.98**	**0.97**	**0.96**	**0.98**
	CNN-KT	0.95	0.96	0.95	0.94	0.97	0.94	0.93	0.96
	Pi-CNN-1	0.55	0.57	0.56	0.53	0.54	0.55	0.53	0.53
	CNN-1	0.57	0.55	0.51	0.55	0.53	0.56	0.57	0.54
	Pi-CNN-2	0.90	0.93	0.91	0.91	0.90	0.92	0.92	0.92
	CNN-2	0.91	0.91	0.90	0.92	0.91	0.88	0.89	0.93

TABLE IV

OUTPUT'S AVERAGE RMSES AND AVERAGE CCS OF THE PROPOSED FRAMEWORK AND BASELINE METHODS IN THE MULTIPLE-TO-SINGLE SCENARIO

Metrics	Methods	Angle	FCR	FCU	ECRL	ECRB	ECU
Average RMSE	**Pi-CNN-KT**	**10.27**	**5.19**	**4.61**	**1.12**	**1.27**	**0.42**
	CNN-KT	14.40	5.71	6.17	1.51	1.35	0.71
	Pi-CNN-1	41.21	15.41	31.33	23.03	11.12	2.76
	CNN-1	38.39	15.59	29.02	8.87	13.45	5.16
	Pi-CNN-2	26.41	7.38	8.41	1.75	2.18	1.02
	CNN-2	20.23	7.49	7.55	1.53	1.63	1.49
CC	**Pi-CNN-KT**	**0.95**	**0.94**	**0.97**	**0.99**	**0.98**	**0.99**
	CNN-KT	0.93	0.92	0.95	0.99	0.97	0.98
	Pi-CNN-1	0.78	0.14	-0.17	0.85	0.83	0.76
	CNN-1	0.80	0.13	0.51	0.87	0.05	0.85
	Pi-CNN-2	0.86	0.86	0.92	0.99	0.96	0.97
	CNN-2	0.89	0.86	0.94	0.99	0.95	0.95

V. DISCUSSIONS

In this section, effects of dataset sizes and time consumption in physics-informed knowledge transfer phase are first analyzed, and then essential advantages of physics-informed deep transfer learning in facilitating the musculoskeletal modeling personalisation and flexibility of the proposed physics-informed deep learning framework are discussed.

TABLE V

RMSES OF THE PROPOSED FRAMEWORK UNDER DIFFERENT DATASET SIZES IN THE KNOWLEDGE TRANSFER PHASE

Scenarios	Outputs	20%	40%	60%	80%
Single	Angle	10.26	11.18	10.67	10.59
	FCR	4.75	4.51	4.45	4.63
	FCU	3.22	3.03	2.82	3.67
	ECRL	2.16	3.27	3.13	4.11
	ECRB	2.56	2.58	2.45	2.15
	ECU	0.72	0.59	0.61	0.79
Multiple	Angle	10.25	9.52	8.81	9.22
	FCR	5.71	5.23	4.97	4.21
	FCU	6.17	5.51	5.43	5.26
	ECRL	0.98	0.92	1.03	1.05
	ECRB	1.27	1.23	1.25	1.36
	ECU	0.61	0.47	0.60	0.63

A. Effects of Dataset Sizes on Physics-Informed Knowledge Transfer

Table V depicts the RMSEs of the proposed framework under different training dataset sizes in the knowledge transfer phase, from 20% to 80% of the data from the new subject, used for fine-tuning the personalized network both in the single-to-single and the multiple-to-single scenarios. Observed from Table V, the predicted performance of the proposed framework is relatively stable with the increase of the number of the data used for knowledge transfer, which indicates that the proposed framework is not sensitive to the training dataset sizes during the knowledge transfer phase.

B. Time Consumption of Physics-Informed Knowledge Transfer

Table VI lists the time consumption of the proposed framework, Pi-CNN-1, CNN-1, and CNN-KT both in the single-to-single and the multiple-to-single scenarios. Accordingly, the time consumption of the proposed framework is much less than Pi-CNN-1 and CNN-1, especially in the multiple-to-single scenario. Because both Pi-CNN-1 and CNN-1 need to train the data-driven model using the whole data, while the proposed framework only utilizes part of the data from the new subject to optimize the parameters relating to the subject-specific information in the personalized network. Additionally, the time consumption of CNN-KT is less than the proposed framework, because the physics-based domain knowledge is embedded into the loss function to penalize/regularize the neural network during the knowledge transfer phase, but CNN-KT only needs to minimize the MSE loss.

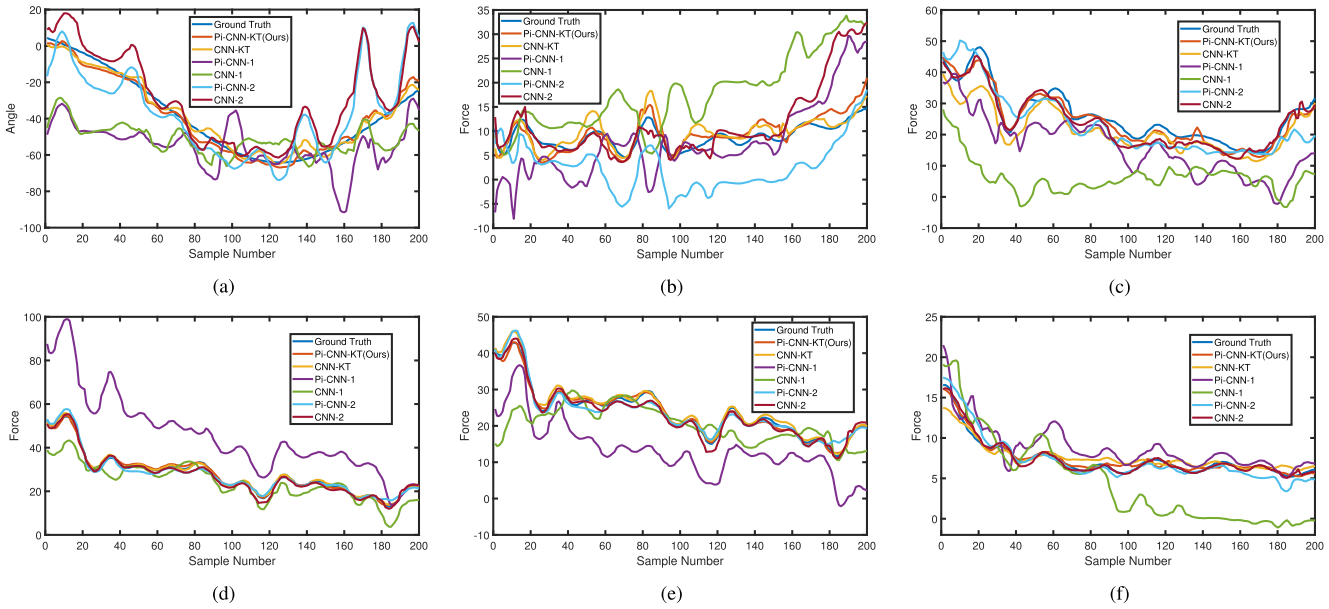


Fig. 6. Representative predicted results of the proposed framework and baseline methods in the multiple-to-single scenario. The personalized network is trained by the data from one specific subject, while the rest seven subjects' data are utilized for training the generic network. Predicted results of (a) wrist angle, (b) FCR, (c) FCU, (d) ECRL, (e) ECRB, and (f) ECU.

TABLE VI
COMPARISONS OF TIME CONSUMPTION (MINUTE)

Scenarios	Pi-CNN-1	CNN-1	Pi-CNN-KT	CNN-KT
Single	80	64	**33**	27
Multiple	361	305	**37**	29

C. Physics-Informed Deep Transfer Learning to Facilitate Musculoskeletal Modeling Personalisation

As mentioned earlier, the generic musculoskeletal model without particular personalisation may be enough when we want to investigate the musculoskeletal phenomena decoupled from individuals or groups, such as how muscles are neurally recruited, how muscles transfer force around multiple joints, or motor control principles, and so on. However, when we study the musculoskeletal function of the specific subject, musculoskeletal model with the unique anatomy and neurophysiology of the subject is necessary [26], [27]. Additionally, understanding the mechanism underlying the specific individual's musculoskeletal function is important for rehabilitation, such as design of the personalized assistive device and human-machine interface, and formulating personalized rehabilitation intervention strategy for the specific subject based on his/her anatomy and impairment [28], [29], [30].

One of the challenging issues for musculoskeletal modeling personalisation is the shifts of statistical characteristics of the collected EMG signals from different individuals mainly caused by the diversity of the anatomical, physiological, and biochemical characteristics between individuals, which may seriously degrade the performance of the conventional data-driven model [19], [31]. Therefore, conventional data-driven methods usually require a large number of subject-specific data to retrain the existing model, which is time-consuming and labor-intensive [32], [33]. In addition, data-driven methods highly rely on the quality of the collected EMG signals, and the ill-conditioned and noisy training data may impose negative and unpredictable results. Although some existing transfer learning frameworks, such as MetaSleepLearner [34], EEGWaveNet [35], and MIN2Net [36], could achieve satisfactory performance in subject-independent scenarios, they were developed without considering the physics-based domain knowledge, and such kind of "black-box" solutions cannot reflect the underlying physical mechanisms during the modeling process. Differently, the proposed framework is more robust and with better generalization by integrating physics-based domain knowledge that from kinematic measurements and the physical understanding of the neuromusculoskeletal coupling into the deep neural networks to reflect physical or physiological mechanisms [37], [38], [39], [40]. Furthermore, less subject-specific data are required in the knowledge transfer phase due to the twin neural networks mechanism.

D. Flexibility of Physics-Informed Deep Learning Framework

The aim of our work is to develop the next-generation physics-informed data-driven musculoskeletal models, which could seamlessly integrate the existing physics-based domain knowledge into the deep learning-based data-driven models. The proposed method is only an attempt to bring physics information into data-driven models to overcome its limitations by creating data reflecting the underlying physical mechanisms, and also bring the powerful deep learning techniques into physics-based musculoskeletal models to reduce computational demands in data processing and improve execution speed for real-time applications. Moreover, the embedded

TABLE VII
COMPARISONS BETWEEN PI-CNN AND PI-LSTM

	Angle	FCR	FCU	ECRL	ECRB	ECU
Pi-CNN	9.35	5.17	3.98	2.77	2.61	0.81
Pi-LSTM	9.92	5.26	3.76	2.99	2.52	0.75

physics law illustrates the relationship between sEMG, and muscle forces and joint kinematics. We choose the wrist joint as the exemplar, by considering that wrist primary muscles are superficial muscle and can be easily measured, and wrist muscle forces and joint kinematics estimation is a kind of application widely considered in the existing works. It could also be generalized to other joints and more scenarios. In this article, we use CNN to implement the proposed physics-informed deep learning framework. The proposed framework is actually flexible; the deep neural network can be replaced depending on the specific requirements, such as LSTM and ResNet.

Table VII lists the detailed comparison results between Pi-CNN (CNN as the deep neural network in the proposed physics-informed deep learning framework) and Pi-LSTM (LSTM as the deep neural network in the proposed physics-informed deep learning framework). According to Table VII, we can find that Pi-CNN and Pi-LSTM achieve comparable performance, indicating that different types of deep neural networks can be embedded into the proposed physics-informed deep learning framework.

VI. CONCLUSION

In this article, a physics-informed deep transfer learning framework is designed for musculoskeletal modeling personalisation. Different from conventional transfer learning methods, physics-based domain knowledge is integrated into deep neural networks as soft constraints to penalize/regularize the loss function of CNNs utilized in the proposed framework. The embedded physics-based domain knowledge could help strengthen robustness and generalization, and the number of required subject-specific data is reduced during the knowledge transfer phase. Moreover, the utility of the twin neural network mechanism significantly reduces computational burdens in personalized network fine-tuning. Comprehensive experiments on wrist muscle forces and joint angle estimation demonstrate the effectiveness of the proposed framework. It should be noted that, in this study, we assume that all the data used in the experiments are labeled, but the available data of the new subject may be unlabeled in clinical applications. Therefore, we will focus on the future physics-informed unsupervised transfer learning and physics-informed semi-supervised transfer learning strategies.

REFERENCES

[1] Q. Li, Z. Luo, and J. Zheng, "A new deep anomaly detection-based method for user authentication using multichannel surface EMG signals of hand gestures," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.

[2] K. J. Bennett et al., "EMG-informed neuromusculoskeletal models accurately predict knee loading measured using instrumented implants," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 7, pp. 2268–2275, Jan. 2022.

[3] J. Weng, E. Hashemi, and A. Arami, "Adaptive reference inverse optimal control for natural walking with musculoskeletal models," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 1567–1575, 2022.

[4] M. K. Jung et al., "Intramuscular EMG-driven musculoskeletal modelling: Towards implanted muscle interfacing in spinal cord injury patients," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 1, pp. 63–74, Jun. 2021.

[5] Y. Zhao, Z. Zhang, Z. Li, Z. Yang, A. A. Dehghani-Sanij, and S. Xie, "An EMG-driven musculoskeletal model for estimating continuous wrist motion," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 3113–3120, Dec. 2020.

[6] D. Xiong, D. Zhang, X. Zhao, and Y. Zhao, "Deep learning for EMG-based human-machine interaction: A review," *IEEE/CAA J. Automat. Sinica*, vol. 8, no. 3, pp. 512–533, Mar. 2021.

[7] L. Rane, Z. Ding, A. H. McGregor, and A. M. J. Bull, "Deep learning for musculoskeletal force prediction," *Ann. Biomed. Eng.*, vol. 47, no. 3, pp. 778–789, Mar. 2019.

[8] A. De Brabandere, J. Emmerzaal, A. Timmermans, I. Jonkers, B. Vanwanseele, and J. Davis, "A machine learning approach to estimate hip and knee joint loading using a mobile phone-embedded IMU," *Frontiers Bioeng. Biotechnol.*, vol. 8, p. 320, Apr. 2020.

[9] N. Makaram, P. A. Karthick, and R. Swaminathan, "Analysis of dynamics of EMG signal variations in fatiguing contractions of muscles using transition network approach," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–8, 2021.

[10] J. Zhang, Y. Li, W. Xiao, and Z. Zhang, "Robust extreme learning machine for modeling with unknown noise," *J. Franklin Inst.*, vol. 357, no. 14, pp. 9885–9908, Sep. 2020.

[11] T. Bao, S. A. R. Zaidi, S. Xie, P. Yang, and Z.-Q. Zhang, "A CNN-LSTM hybrid model for wrist kinematics estimation using surface electromyography," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–9, 2020.

[12] T. Bao, S. Q. Xie, P. Yang, P. Zhou, and Z.-Q. Zhang, "Toward robust, adaptiveand reliable upper-limb motion estimation using machine learning and deep learning—A survey in myoelectric control," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 8, pp. 3822–3835, Aug. 2022.

[13] X. Tang, X. Zhang, M. Chen, X. Chen, and X. Chen, "Decoding muscle force from motor unit firings using encoder–decoder networks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 2484–2495, 2021.

[14] L. N. Wimalasena et al., "Estimating muscle activation from EMG using deep learning-based dynamical systems models," *J. Neural Eng.*, vol. 19, no. 3, May 2022, Art. no. 036013.

[15] W. S. Burton, C. A. Myers, and P. J. Rullkoetter, "Machine learning for rapid estimation of lower extremity muscle and joint loading during activities of daily living," *J. Biomech.*, vol. 123, Jun. 2021, Art. no. 110439.

[16] T. T. Dao, "From deep learning to transfer learning for the prediction of skeletal muscle forces," *Med. Biol. Eng. Comput.*, vol. 57, no. 5, pp. 1049–1058, 2019.

[17] T. Bao, S. A. R. Zaidi, S. Xie, P. Yang, and Z.-Q. Zhang, "Inter-subject domain adaptation for CNN-based wrist kinematics estimation using sEMG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1068–1078, 2021.

[18] W. Wang, B. Chen, P. Xia, J. Hu, and Y. Peng, "Sensor fusion for myoelectric control based on deep learning with recurrent convolutional neural networks," *Artif. Organs*, vol. 42, no. 9, pp. E272–E282, Sep. 2018.

[19] K.-T. Kim, C. Guan, and S.-W. Lee, "A subject-transfer framework based on single-trial EMG analysis using convolutional neural networks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 1, pp. 94–103, Jan. 2020.

[20] M. Sartori, D. G. Llyod, and D. Farina, "Neural data-driven musculoskeletal modeling for personalized neurorehabilitation technologies," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 5, pp. 879–893, May 2016.

[21] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, "Physics-informed machine learning," *Nat. Rev. Phys.*, vol. 3, no. 6, pp. 422–440, 2021.

[22] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, Mar. 2019.

[23] J. Willard, X. Jia, S. Xu, M. Steinbach, and V. Kumar, "Integrating scientific knowledge with machine learning for engineering and environmental systems," *ACM Comput. Surv.*, vol. 55, no. 4, pp. 1–37, May 2023.

[24] Y. Zhao et al., "Adaptive cooperative control strategy for a wrist exoskeleton using model-based joint impedance estimation," *IEEE/ASME Trans. Mechatronics*, early access, Oct. 19, 2022, doi: 10.1109/TMECH.2022.3211671.

[25] J. Zhang, Y. Li, W. Xiao, and Z. Zhang, "Non-iterative and fast deep learning: Multilayer extreme learning machines," *J. Franklin Inst.*, vol. 357, no. 13, pp. 8925–8955, 2020.

[26] Y. Hiasa, Y. Otake, M. Takao, T. Ogawa, N. Sugano, and Y. Sato, "Automated muscle segmentation from clinical CT using Bayesian U-Net for personalized musculoskeletal modeling," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1030–1040, Sep. 2019.

[27] M. Sartori, J. Rubenson, D. G. Lloyd, D. Farina, and F. A. Panizzolo, "Subject-specificity via 3D ultrasound and personalized musculoskeletal modeling," in *Converging Clinical and Engineering Research on Neurorehabilitation II*. Springer, 2017, pp. 639–642.

[28] G. Valente et al., "Are subject-specific musculoskeletal models robust to the uncertainties in parameter identification?" *PLoS ONE*, vol. 9, no. 11, Nov. 2014, Art. no. e112625.

[29] M. M. Arones, M. S. Shourijeh, C. Patten, and B. J. Fregly, "Musculoskeletal model personalization affects metabolic cost estimates for walking," *Frontiers Bioeng. Biotechnol.*, vol. 8, Nov. 2020, Art. no. 588925.

[30] C. Pizzolato et al., "Targeted Achilles tendon training and rehabilitation using personalized and real-time multiscale models of the neuromusculoskeletal system," *Frontiers Bioeng. Biotechnol.*, vol. 8, 2020, Art. no. 878.

[31] X. Chen, Y. Li, R. Hu, X. Zhang, and X. Chen, "Hand gesture recognition based on surface electromyography using convolutional neural network with transfer learning method," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 4, pp. 1292–1304, Apr. 2020.

[32] S. Goswami, C. Anitescu, S. Chakraborty, and T. Rabczuk, "Transfer learning enhanced physics informed neural network for phase-field modeling of fracture," *Theor. Appl. Fract. Mech.*, vol. 106, Apr. 2020, Art. no. 102447.

[33] S. Chakraborty, "Transfer learning based multi-fidelity physics informed deep neural network," *J. Comput. Phys.*, vol. 426, Feb. 2021, Art. no. 109942.

[34] N. Banluesombatkul et al., "MetaSleepLearner: A pilot study on fast adaptation of bio-signals-based sleep stage classifier to new individual subject using meta-learning," *IEEE J. Biomed. Health Inform.*, vol. 25, no. 6, pp. 1949–1963, 2020.

[35] P. Thuwajit et al., "EEGWaveNet: Multiscale CNN-based spatiotemporal feature extraction for EEG seizure detection," *IEEE Trans. Ind. Informat.*, vol. 18, no. 8, pp. 5547–5557, Dec. 2021.

[36] P. Autthasan et al., "MIN2Net: End-to-end multi-task learning for subject-independent motor imagery EEG classification," *IEEE Trans. Biomed. Eng.*, vol. 69, no. 6, pp. 2105–2118, Dec. 2021.

[37] P. A. K. Reinbold, L. M. Kageorge, M. F. Schatz, and R. O. Grigoriev, "Robust learning from noisy, incomplete, high-dimensional experimental data via physically constrained symbolic regression," *Nature Commun.*, vol. 12, no. 1, pp. 1–8, May 2021.

[38] S. Wang, X. Yu, and P. Perdikaris, "When and why PINNs fail to train: A neural tangent kernel perspective," *J. Comput. Phys.*, vol. 449, pp. 1–28, Jan. 2022.

[39] C. Rudin, C. Chen, Z. Chen, H. Huang, L. Semenova, and C. Zhong, "Interpretable machine learning: Fundamental principles and 10 grand challenges," *Statist. Surv.*, vol. 16, pp. 1–85, Jan. 2022.

[40] Y. Yang and P. Perdikaris, "Adversarial uncertainty quantification in physics-informed neural networks," *J. Comput. Phys.*, vol. 394, pp. 136–152, Oct. 2019.