Feature Decoupling for Multimodal Locomotion and Estimation of Knee and Ankle Angles Implemented by Multi-Model Fusion

Xisheng Yu and Zeguang Pei

***Abstract*— Many challenges exist in the study of using orthotics, exoskeletons or exosuits as tools for rehabilitation and assistance of healthy people in daily activities due to the requirements of portability and safe interaction with the user and the environment. One approach to dealing with these challenges is to design a control system that can be deployed in a portable device to identify the relationships that exist between the gait variables and gait cycle for different locomotion modes. In order to estimate the knee and ankle angles in the sagittal plane for different locomotion modes, a novel multimodal feature-decoupled kinematic estimation system consisting of a multimodal locomotion classifier and an optimal joint angle estimator is proposed in this paper. The multi-source information output from different conventional primary models are fused by assigning the non-fixed weight. To improve the performance of the primary models, a data augmentation module based on the time-frequency domain analysis method is designed. The results show that the inclusion of the data augmentation module and multi-source information fusion modules has improved the classification accuracy to 98.56% and kinematic estimation performance (PCC) to 0.904 (walking), 0.956 (running), 0.899 (stair ascent), 0.851 (stair descent), respectively. The kinematic estimation quality is generally higher for faster speed (running) or proximal joint (knee) compared to other modes and ankle. The limitations and advantages of the proposed approach are discussed. Based on our findings, the multimodal kinematic estimation system has potential in facilitating the deployment for human-in-loop control of lower-limb intelligent assistive devices.**

***Index Terms*— Intelligent assistive devices, motion intent recognition, multimodal locomotion decoupling, joint angle estimation.**

I. Introduction

DISEASES such as stroke and osteoarthritis can affect people's ability in daily activities, including walking, running, ascending and descending stairs [1], [2]. Wearable assistive robots (such as exosuit) can assist wearers and reduce their metabolism, thus helping patients to complete daily rehabilitation or enhancing the motion capacity of healthy people [3], [4]. During the assisting process, the wearable robot interacts with the wearer and the environment through the sensing system to accomplish motion intent recognition and implement assistance planner [5], [6]. Human lower limbs and joints do not have the same biomechanics in the sagittal plane due to functional differences in neuromuscular activity during different locomotion, which makes developing intelligent assistive devices that assist multiple gaits challenging [7], [8].

To obtain better human-machine interaction (HMI) performance in multimodal locomotion, the gradually accepted way is to use the locomotion classifier as a high-level controller to switch the mid-level motion planner and then control the low-level actuator of the exoskeletons [4], [6]. A wide range of signals can be used as inputs for high-level controllers to recognize current locomotion mode [9]. In the motion planner, relationships between gait variables are created such as the kinematics and kinetics of the joints. In particular, joint angles play an important role in providing insights and fundamental information for human gait analysis and development of exoskeletons. For example, the range of motion of the joints of the lower extremity can be used to understand the impact of abnormal gait on the energy efficiency and joint compensation during walking [10], [11] and to provide feedback for low-level actuator position control in the soft exosuit to minimize muscle efforts [12], [13].

Traditionally, marker-points data are collected by an optical motion capture system and implemented into computational musculoskeletal modeling software such as Open-Sim [14] to calculate the joint angle. Although reliable measurement results of joint angles can be obtained in this method, it has some limitations. The major limitation is that this method can only work indoors where space is restricted due to the need for a large number of fixed motion capture components. However, it has been suggested that the use of wearable inertial measurement unit (IMU) sensors and forward kinematics model can replace traditional motion capture cameras to estimate human movements and calculate joint angles outside the laboratory environment [15], [16]. In order to collect kinematic information of lower extremity multi-joints, a full set of IMU sensors (minimum of 7) must be worn by the subjects. However, the

Manuscript received 21 February 2024; revised 14 May 2024; accepted 14 June 2024. Date of publication 19 June 2024; date of current version 25 June 2024. This work was supported by the Fundamental Research Funds for the Central Universities under Grant 2232023Y-01. *(Corresponding author: Zeguang Pei.)*

The authors are with the College of Mechanical Engineering, Donghua University, Shanghai 201620, China (e-mail: zgpei@dhu.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2024.3416530, provided by the authors.
Digital Object Identifier 10.1109/TNSRE.2024.3416530

use of each IMU sensor requires calibration setup to determine the position and orientation of the sensors before collecting motion data. It is very challenging to maintain the relative stability of all sensors and limbs during movement [17]. Moreover, the large number of sensors also places higher demands on the performance of the equipment used for data processing, which is not favorable for implementation in controllers of portable assistance devices.

It has been hypothesized that significant redundancy can occur in multi-sensor human motion capture systems [18]. The use of a large number of sensors does not provide a large benefit in improving motion accuracy. To make the system simple and increase the reliability, some approaches based on statistical analysis models or mathematical models have been proposed to predict target joint angular trajectories using fewer wearable sensors during human locomotion [19], [20], [21], [22], [23], [24]. However, the accuracy of the regression statistical model is affected by the distributional properties of the sample statistics. Then the inputs of mathematical models have to be discretely mapped to the estimated outputs, and intermediate parameters such as switching rules, velocity estimation, gait percentage identification, and look-up table design are required in the computational process. These methods are too resource-intensive to be used for real-time gait tracking. The limitations imposed by these model-based kinematics estimation methods create difficulties in estimating gait variables quickly and accurately.

Different types of wearable sensors generate a large amount of gait data in multiple locomotion. The data-driven method is considered to be a promising approach in areas such as gait analysis and intelligent assistive devices development. This method can effectively extract features and implicit information from the large amount of data. Recently, researchers have used deep neural networks to map input signals to the target object space for locomotion mode recognition [25], [26], [27] and continuous kinematic estimation [19], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39]. A wide range of inputs are used in these works, such as the partly limb or joint kinematics from IMUs, electromyography (EMG) signals, force sensitive resistors sensors (FSRs) and human physiological characteristics (e.g., age, gender, height, mass and BMI). However, the sensors for collecting raw data may be affected by multiple noises which originate from multi-frequency vibrations of the limbs, motion artifacts and electric fields in the environment. The gait data obtained from sensor acquisition systems may not provide reliable kinematic estimation in practical applications. All these limitations are further aggravated by utilizing conventional deep learning models such as Long Short-Term Memory (LSTM), Convolutional Neural Networks (CNN) and Gate Recurrent Unit (GRU) for kinematic estimation.

Our contributions: This work aims to estimate knee and ankle angles in sagittal plane using fewer input signals (e.g. hip kinematics and individual characteristic parameters) under multimodal locomotion including walking, running, stair ascent and stair descent. To do so, a novel multimodal feature-decoupled kinematic estimation system is built by fusing the outputs of conventional deep learning primary models with high performance. Four deep learning models consisting of CNN, LSTM, and GRU for time-series data predictions are used as primary models. To address the challenges posed by the fewer input signals [39], the data augmentation module is designed to improve estimation performance of conventional primary models and investigate the effect of input sets with different sizes on the prediction performance. We further develop two novel composite models to improve the locomotion mode recognition accuracy (LMR-FM-Net) and estimation performance of joint angles (Kinematic-FM-Net) by combining output from multiple primary models through the fusion modules in composite models. By conducting extensive evaluation with model ablation studies, the results suggest that the proposed composite models outperform each of the primary models individually or the other combinations. Furthermore, the complexity of the system is minimized so that it consumes less computational resources. This system has the potential to be deployed in real-time HMI controllers for portable intelligent assistive devices, such as exoskeletons or exosuit.

II. Related Work

In this section, related works of motion intent recognition and kinematic estimation methods based on wearable sensors will be discussed. First, the locomotion classification methods for high-level controller are discussed. Then, kinematic estimation methods based on fewer IMUs and mathematical models are discussed. After that, the data-driven methods for kinematic estimation and their limitations are discussed.

Accurate locomotion classification by high-level controller of exoskeleton for multimodal motion is essential for fast switching of motion planning and the provision of effective assistance. In [4], a binary classification algorithm based on the threshold of the potential energy fluctuations is proposed to recognize human motion states (walking or running). The recognition result is used as a switching signal for the actuation profiles of the portable exosuit. Since different features are used to detect multiple gait modes in different periods of a reference-point transformation period (RPTP), it is difficult to use a threshold-based method to detect the gait modes. A fast gait pattern detection method based on a fuzzy logic algorithm with human sensor system is proposed in [6]. The results consist of five modes including level walking (LW), stairs ascent (SA), stairs descent (SD), ramp ascent (RA) and ramp descent (RD), which can be used to select appropriate kinematic and kinetic models for the motion planner. The model complexity in these methods is related to the kind of motion patterns, which is challenging for real-time gait detection applications for multimodal motion.

The data-driven methods are being increasingly used as a substitution for the locomotion classification. In [26], researchers proposed an ensemble learning-based hybrid deep learning framework to recognize the multitask human walking activities using human gait patterns. The IMU sensors is placed on the chest, left thigh, and right thigh to collect the data for 7 different activities. The results show that the hybrid deep learning framework has provided a promising classification accuracy of 99.34% over other models. Similarly, to detect more movement patterns, a method for recognizing human activity from wearable sensors based on SensCapsNet, which is able to recognize 19 movement patterns (including upper and lower), is proposed in [27]. These methods demonstrate the efficiency and potential of using wearable sensors and neural network models for human activity recognition.

To make the system simple and increase the reliability, some approaches based on mathematical models have been proposed to predict the trajectories of target joint. In [19] and [20], discrete polynomial models were developed for seven different subphases of the gait cycle to create joint trajectories as a function of the gait percentage. In addition, thigh angles and thigh angle integrals have also been used to create quasi-circular curves and predict knee and ankle angles as a function of the estimated gait percentage based on the discrete Fourier transform (DFT) [21]. The work is then further extended in [22] and [23] to take the effects of different speeds and slopes into account. Different sets of phase variables (virtual constraints) are generated for each speed and slope condition. Individual characterization parameters such as gender, age and BMI have also been used to build multiple regression models to predict lower extremity joint angles in the sagittal plane [24]. Each of these methods utilize simulation or experimental data to create a functional relationship between joint angles and gait percentage or other parameters. Different sets of phase variables are generated for different joints, velocities and motion conditions in all function models, while switching rules are designed separately. All of these models are too resource-intensive to provide joint angles estimation in real time.

Due to the limitations of model-based methods and wearable sensing systems, researchers have focused on estimating joint angles by deep learning models and multiple types of wearable sensors (e.g., IMUs in [30], [31], [35], [37], [39], FSRs in [36] and EMGs in [29], [32], [33], [38]). In [31], an estimation approach based on a nonlinear auto-regressive model with wavelets theory was developed that continuously mapped the inputs (thigh and shank angle from IMUs) to the outputs (knee and ankle trajectories), which did not require intermediate parameters. In [33], a framework of online prediction method of joint angles by long short-term memory (LSTM) neural network based on surface EMG signals from eight muscles was proposed. The results show that the proposed method can realize accurate online joint angle prediction. In [36], a method utilizing force myography (FMG) was employed to model and estimate knee joint angles during walking and running. An eight-channel FSRs data acquisition system was created to gather data wirelessly. This data was then utilized to train an artificial neural network (ANN) to predict knee joint angles. However, the experiment was performed for one subject. In addition to wearable sensors, speed, anthropometrics and demographics are also used as inputs of different deep learning algorithms to estimate joint angles [28], [34].

The traditional network models are used in most of the current studies. Therefore, the advanced models with more complex structures or multi-model fusion have been proposed to improve estimation performance [30], [38], [39]. In [39], as a novel framework, DeepBBWAE-Net was proposed that

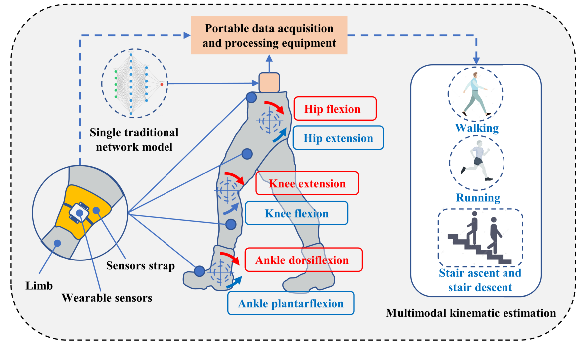

Fig. 1. The graphical description of the problems to be addressed.

implements ensemble techniques such as bagging, boosting, and weighted averaging to improve kinematic predictions. In the framework, five conventional deep learning networks are used as the base learners. Their developed algorithm estimates unilateral joint angles using two shoe-mounted IMU sensors in different walking conditions. This study provides a practical and accurate estimation of joint kinematics which can reduce the number of IMU sensors. The estimation performance is improved compared to other deep learning models. However, it will consume more digital resources due to more complicated networks and learning parameters. In addition, the IMU sensor fixed to the foot is not suitable for the patients wearing prostheses due to transfemoral amputation [31]. Another limitation of this work is the limited dataset.

Most of the current studies focus on the kinematic estimation under single locomotion mode [33], [34], [36] or discrete multi-mode motion [29], [32]. The limitations of these studies are compounded further with the use of conventional deep learning models and more wearable sensors, which limit more accurate multi-variable human kinematics estimations. The problems to be addressed are shown in Fig. 1. To address the limitations imposed by these discussed works, a novel deep learning model fusion method is proposed to continuously and smoothly estimate lower limb knee and ankle joint angles using proximal joint (hip) kinematics that can be obtained from fewer IMUs under multiple outdoor locomotion. This estimation method is not only applicable to assist in healthy people, but can also be used for patients with transfemoral amputations.

III. Methods

A. System Architecture

Fig.2 illustrates the architecture of a multimodal kinematic estimation system for knee and ankle angles. At the start, raw data including age, sex, height, weight, body mass index (BMI) and sagittal plane multi-joint kinematics in different locomotion modes are collected from publicly available datasets. For this purpose, three datasets involving 115 healthy subjects and four modes are used, including 42 in walking [40], 28 in running [41], 43 in stair ascent, and 43 in stair descent (41 of the subjects are included both in stair ascent and stair descent) [42]. The kinematic data from different subjects are preprocessed for classification and normalization to create initial datasets and locomotion labels. Then, the data augmentation module extends the hip kinematic data to get different sizes of input sets for the training and testing of deep neural networks.

After the data preprocessing, the system is divided into two parts. The upper layer is a locomotion classifier which performs feature decoupling for multimodal tasks. This can address the problem posed by heterogeneous data to deep neural networks for kinematic modeling. It enables the primary models to be more focused on the estimation for a particular locomotion. LMR-FM-Net, the multimodal locomotion classifier, is trained and tested to take locomotion modes recognition. It consists of LMR-Nets and a fusion module, with the LMR-Nets consisting of four primary models including LSTM-net, GRU-net, CNN-LSTM-net, and CNN-GRU-net. Taking the augmented hip kinematic data and individual characteristic parameters as input, LMR-FM-Net outputs one of the four modes as the result of the mode recognition.

The lower layer is the joint angle estimator which relies on multi-source information fusion to accomplish optimal estimation of the gait variables in the identified mode. Kinematic-FM-Net, the optimal joint angle estimator, consists of four Kinematic-Nets and a fusion module based on the Kalman filter. The type of primary models used in Kinematic-Nets is the same as in LMR-Nets, with the difference that the last layer is a regression layer instead of a classification layer. The input of the angle estimator includes the results of locomotion mode recognition from the LMR-FM-Net, the augmented hip kinematics and the individual characteristic parameters. The optimal parameters of the primary models are different for each locomotion mode. After matching the mode according to the result from LMR-FM-Net, Kinematic-Net is used to estimate other joint angles of the lower limb in the sagittal plane. The fusion module based on the Kalman filter integrates the output of each primary model for information fusion and iterative updating. The ultimate output is expected to approximate the real joint angles.

B. Dataset Description

After removing the abnormal data from the public datasets, the joint kinematics data and individual characterization parameters are used to train and test the deep network models for locomotion modes classification and joint angles estimation. The trained models can adapt to the biomechanical properties of most populations. Data from the hip, knee, and ankle joints in the sagittal plane are used to construct the deep learning model, while pelvic and foot data are discarded. We focus on predicting joint kinematics rather than limb kinematics.

The first dataset (publicly available), in which 42 subjects (24 young adults (27.6 $\pm$ 4.4 years, 171.1 $\pm$ 10.5 cm, 68.4 $\pm$ 12.2 kg, BMI: 23.2 $\pm$ 4.0) and 18 older adults (62.7 $\pm$ 8.0 years,161.8 $\pm$ 9.5 cm, 66.9 $\pm$ 10.1 kg, BMI: 25.6 $\pm$ 3.6)) walked at different speeds (subject-specific, 0.36 $\sim$ 2.23 m/s), is obtained from [40]. Each participant was asked to perform walking trials in the over-ground walking and treadmill walking conditions. Under the over-ground walking, the participant first walked at a self-selected comfortable speed (level OC), and then at speeds 30% faster (level OF) and 30% slower (level OS) than the comfortable speed. In addition, the participants walked on the treadmill at eight different controlled speeds corresponding to levels T01, T02, ..., and T08. Within these 11 speed levels, walking speed was not the same across participants. The second dataset (publicly available), in which 28 subjects (34.75 $\pm$ 6.69 years, 176.0 $\pm$ 6.8 cm, 69.6 $\pm$ 7.7 kg, BMI: 22.4 $\pm$ 1.6) were asked to run on a treadmill at fixed speeds of 2.5, 3.5, and 4.5 m/s, respectively, is obtained from [41]. Observation of the joint angle curves in different planes for walking and running illustrates that the speed has a significant effect on each joint kinematics. This may be due to a change in the proportion of the swing phase and support phase in the gait cycle [4]. The third dataset (publicly available), in which 45 subjects (6-72 years, 116.6-187.5 cm,18.2-110 kg, BMI:11.7-33.6) ascended or descended stairs at several speeds (stair ascending:0.49 $\pm$ 0.06 m/s, stair descending: 0.52 $\pm$ 0.15 m/s), is obtained from [42]. In the first and third databases, each subject walked at self-selected speeds which were not necessarily similar to another participant's speeds. Unlike the second database, the locomotion of participants was not required to be set at fixed speeds. Other detailed information on the protocol of the experiment is described in [40], [41], and [42].

C. Data Augmentation Module

Data augmentation can effectively expand the data samples and prevent overfitting of the training model, which is critical for the successful use of deep learning models as it is a useful tool for increasing the quality and dimension of the input features [43]. It has been shown to be effective in many applications such as time series forecasting [44], [45]. In addition, Data augmentation can minimize sensor inputs to reduce the requirements of marker data and sensor for the estimation of gait variables [46], thus have potential for gait analysis as well as assistive device design. Time-domain, frequency-domain and time-frequency domain transforms are typical data augmentation methods. Among them, time-frequency analysis is widely used for time series prediction [47]. In this paper, the data augmentation module is designed based on Hilbert-Huang transformation (HHT) and Hilbert transformation (HT) of multiresolution analysis based on maximal overlap discrete wavelet transform (MODWT-MRA). It enriches the feature types and increases the dimension of the data series. Raw signals, augmented time-frequency domain signals and individual parameters are combined as inputs for the primary deep neural networks.

In [43], [48], and [49], the HHT, which contains empirical mode decomposition (EMD) and the Hilbert transform, is proposed to analyze the gait signals. The EMD method allows the decomposition of any signal into a finite small number of intrinsic mode functions (IMFs). That offers a possibility to exploit the information hidden in the gait signals. The shortcoming of EMD is that the number of IMFs decomposed from the same type of signal is uncertain. The IMF of the hip signal for one gait cycle is less than or equal to 2 in this study. Input feature size of deep neural network is required

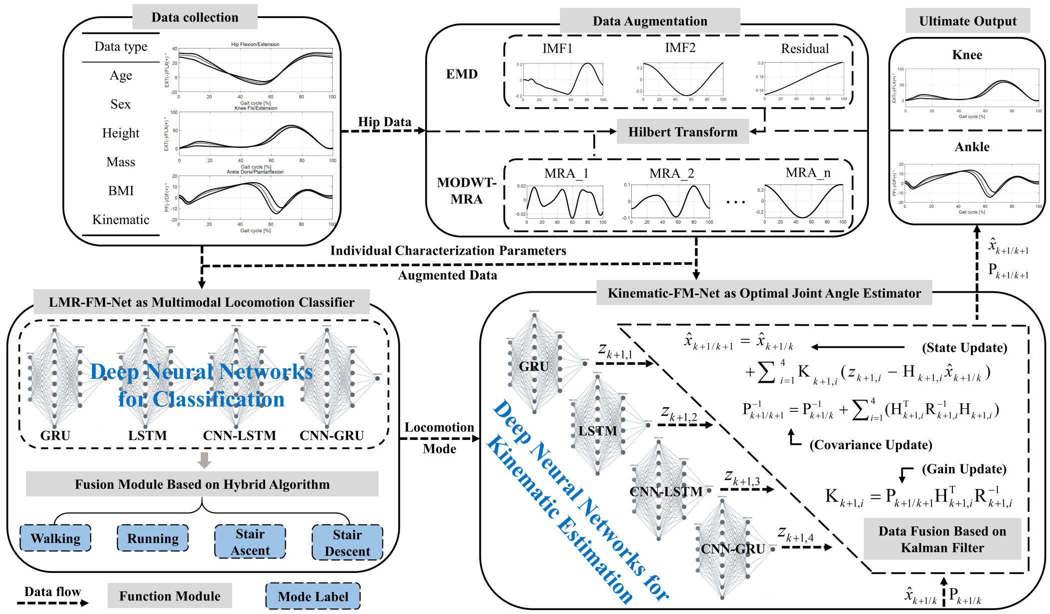

Fig. 2. Architecture of the multimodal kinematic modeling system. The EMD (empirical mode decomposition) and MODWT-MRA (multi-resolution analysis based on maximum overlap discrete wavelet transform) blocks in the data augmentation module decompose the hip kinematic signals into many IMFs and MRAs in the time-frequency domain. The Hibbert transform of the sub-signals expands the input feature dimension of the deep network model. The GRU, LSTM, CNN-LSTM, and CNN-GRU are the primary models used in this system. Both LMR-FM-Net and Kinematic-FM-Net are generated by fusion of primary models.

to be a fixed constant. Therefore, the first component of the IMFs is selected to apply the HT during the preprocessing of the dataset. The analytic signal $z(t)$ is defined by:

$$z(t) = u_1(t) + \frac{1}{\pi} PV \int_{-\infty}^{\infty} \frac{u_1(\tau)}{t-\tau} \mathrm{d}\tau \cdot j \tag{1}$$

where $u_1(t)$ is the first IMF of gait cycle signal, PV denotes the Cauchy principal value of the integral. The calculation of the instantaneous frequency and instantaneous energy of $z(t)$ is described in [43].

Among many variants of wavelet transform, maximum overlap discrete wavelet transforms (MODWT) is especially attractive for time series analysis [50], [51]. It can decompose the signal into multiple wavelet components with different frequency domain scales to realize the sparse representation of the signal. One of the advantages is that no mode mixing occurs with MODWT. Assuming that the hip kinematic signals are samples of a function $f(x)$ evaluated at N time points, the function can be expressed as a linear combination of the scaling function $\phi(x)$ and wavelet $\psi(x)$ at varying scales and translations as follows:

$$\begin{aligned} f(x) = &\sum_{k=0}^{N-1} c_k 2^{\frac{-J_0}{2}} \phi(2^{-J_0}x - k) \\ &+ \sum_{j=1}^{J_0} \sum_{k=0}^{N-1} d_{j,k} 2^{\frac{-j}{2}} \psi(2^{-j}x - k) \end{aligned} \tag{2}$$

where J_0 is the number of levels of wavelet decomposition, the first sum is the coarse scale approximation of the signal, c_k are the N scaling coefficients, $d_{j,k}$ are the $(J_0 * N)$ detail coefficients.

TABLE I
INPUT SETS OF THE PRIMARY DEEP NEURAL NETWORKS

features	θ	ω	α	ICPs	HHT	MODWT-MRA-HT	D_i
Set 1	$*$			$*$			6
Set 2	$*$	$*$	$*$	$*$			8
Set 3	$*$	$*$	$*$	$*$	$*$ with θ		11
Set 4	$*$	$*$	$*$	$*$		$*$ with θ	17
Set 5	$*$	$*$	$*$	$*$	$*$ with θ,ω,α		17
Set 6	$*$	$*$	$*$	$*$		$*$ with θ,ω,α	35
Num. of signals	1	1	1	5	3 /feature	9 /feature	

ICPs is the individual characterization parameters, including age, sex, height, weight and BMI. $*$ denotes that the feature is included.

Like EMD, MODWT-MRA is also a signal decomposition method based on the wavelet transform, that can return the projections of the function $f(x)$ onto the various wavelet subspaces and final scaling space. Each row in projections is an MRA component of $f(x)$ onto a different subspace. This means the original signal can be recovered by adding all the components. By default, the maximum decomposition level is rounded to negative infinity of $\log_2(N)$. In this paper, the number of levels (J_0) is set to 3. Similar to the HHT, the HT of the MRA components can obtain time-frequency domain information of the signal in the space of different frequency scales. The decomposed components and results of HT are used as input features for the primary models. Table I shows the different input sets used to train the different networks.

D. LMR-FM-Net and Kinematic-FM-Net

In this paper, LMR-FM-Net and Kinematic-FM-Net are built with four primary models and a fusion module,

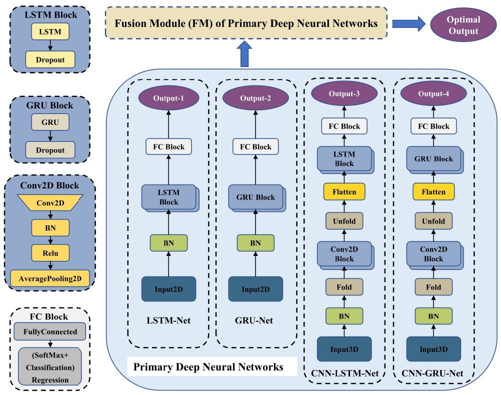

Fig. 3. Entire structure of LMR-FM-Net and Kinematic-FM-Net.

respectively, as shown in Fig. 3. The primary models are conducted by different deep learning layers including LSTM, GRU, Conv2D and fully-connected dense layers. There are differences in the predictive performance of these primary models due to their different structures. The loss function of each primary model is minimized so that the final fusion model can have better prediction performance. The results of multiple primary models are fused using a hybrid algorithm in the sequence-to-last model for locomotion mode classification. For the kinematic estimation model, a fusion method based on Kalman filtering is used with the number of iterative steps equal to the sequence length. All the components of LMR-FM-Net and Kinematic-FM-Net will be described in this section.

1) Primary Models: LSTM-Net: The LSTM-Net solves the problems of gradient vanishing and gradient explosion in RNN networks. It consists of an input layer followed by a Batch Normalization (BN) layer, a LSTM block and a fully connected (FC) Block. Due to the heterogeneity of input data types, the BN layer is applied after the input signal to perform an operation similar to the standard normalization of the input data. In the LSTM block, a bidirectional LSTM (Bi-LSTM) layer and a dropout layer are used. A dropout layer is added after the Bi-LSTM layer to avoid overfitting during the training of the model.

GRU-Net: Like LSTM-Net, GRU-Net is a type of RNN. It realizes the dependence on past information in time-series neural networks through gate control units. The difference is that the cell structure is simplified into an update gate and reset gate in the GRU-net. The network structure is more concise and has fewer computational parameters. In GRU-net, the same architecture as LSTM-net is used with only the LSTM block being replaced by the GRU block which has a GRU layer and a dropout layer.

CNN-LSTM/GRU-Net: There is the CNN-LSTM-Net from LSTM-Net by adding the series-connectable Conv2D block. To perform the convolutional operations on each time step independently, a sequence folding layer is included before the Conv2D block. In the Conv2D block, a conv2D layer is used followed by a BN layer, a Relu layer and an Averagepooling2D layer. The Conv2D layers are capable of extracting local features of a one-dimensional time series and maintaining continuity. The BN layer helps to reduce internal co-variance shift, while the Relu layer is applied to enhance the nonlinear fitting ability of the model. Then, an Averagepooling2D layer is applied to reduce the feature space, which helps reduce the model's complexity and select dominant features. Since LSTM layers expect vector sequence input, to restore the sequence structure and reshape the output of the convolutional layers to sequences of feature vectors, a sequence unfolding layer and a flatten layer are inserted between the Conv2D block and the LSTM block. Following the same architecture as CNN-LSTM-net, CNN-GRU-net is built by replacing the LSTM block with the GRU block.

2) Fusion Module of LMR-FM-Net: For LMR-Nets, hard voting is initially performed based on the outputs of the four main models. However, this may not ensure the accuracy of the results. Since the number of neural network models is even, the voting results of the four primary models may be 2:2 and 1:1:1:1. In order to avoid the impact of these situations on the recognition accuracy, a fusion module is designed to create the LMR-FM-Net using a hybrid algorithm based on hard voting and Markov chain. Firstly, in the fusion module, hard voting is used. The prediction of the output (categories) depends on the category with the highest frequency of output from the primary model. Then, to solve the misclassification caused by the voting result of 50% vs 50%, a state transfer process based on the Markov chain is designed. The specific algorithmic process is presented as algorithm S1 in Supplement materials.

3) Fusion Module of Kinematic-FM-Net: In Kinematic-Nets, initially the simple average of the output is taken from four primary models. However, this may not ensure optimal performance from four models as equal weights are assigned to all the models. Since the performance of each model will be different, proper gain weight needs to be assigned to ensure the best performance gain from these four models. Therefore, a fusion module based on Kalman filter is designed to create the Kinematic-FM-Net. By treating the output of each neural network as a joint angle sensor observation, data fusion of multi-source sensors based on the Kalman filter is implemented in this paper. The output of the fusion module represents the optimal estimation of the knee and ankle angles. For the time series with length $= 101$, optimal estimation of joint angles is taken by iterative updating of gain and covariance matrices at each step. The specific algorithmic process is presented as algorithm S2 in Supplement materials.

E. Bayesian Optimization and Evaluation Procedures

Tuning of model hyperparameters is required when using deep learning models. Bayesian optimization seems to be a good choice. For an overview of Bayesian optimization, see Snoek et al. [52]. The objective function of Bayesian optimization is to minimize the RMSE of the neural network model. For different modes of network model training, Adam [53] is used as the optimizer. All the models are optimized and trained with the maximum Epochs of 200 and minibatch size of 16 in MATLAB (MathWorks, USA) with RTX3060Ti. GPU (NVIDIA, CA). The optimized hyperparameters of the primary deep neural network are listed as Table S.1 in Supplement materials.

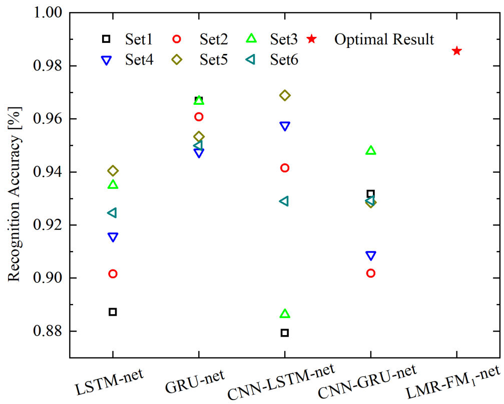

Fig. 4. Classification accuracy of different network models.

In this work, five-fold cross-validation is implemented to evaluate the performance of all the deep learning models. This procedure is carried out to recognize the locomotion modes and estimate knee and ankle angles. For locomotion mode recognition, the proportion of correct labels recognized in the test set is calculated. The root means square error (RMSE), Normalized RMSE (NRMSE) and the Pearson Correlation Coefficient (PCC) are used to evaluate the estimation quality of the joint angles for the above four locomotion modes. The angle range of different joints varies considerably, e.g., the knee has the largest angle during walking which is much larger than the ankle. The value of RMSE cannot be used to compare the kinematic estimation performance of different motion patterns and joints. Therefore, PCC and NRMSE are used as the major evaluation metric. To obtain the NRMSE, the RMSE is normalized by the range of joint angles (the difference between the maximum and minimum values) measured by the corresponding experiment. Repeated-Measures Analysis of Variance (ANOVA) with a least significant difference correction post hoc test is conducted for RMSE, NRMSE and PCC separately with a *p-value* less than 0.05 to determine if our proposed method significantly improves the kinematics estimation compared to other models.

IV. Results

A. Locomotion Mode Classification

Fig. 4 shows the variation in the recognition accuracy of the multiple primary deep neural networks with respect to different sizes of input sets. For LSTM-Net and CNN-LSTM-Net, the performance is best when the input is Set5. The accuracies are 94.06 $\pm$ 1.54 and 96.89 $\pm$ 2.95 [%] (mean $\pm$ std), respectively. For GRU-Net, the most appropriate input is Set1. The accuracy is 96.68 $\pm$ 1.79 [%]. However, when the input is Set3, the result is close to Set1 and better than other sets with an accuracy of 96.67 $\pm$ 2.54 [%]. For CNN-GRU-Net, the most appropriate input is Set3. The accuracy is 94.79 $\pm$ 2.82 [%]. According to Fig. 4, it can be seen that the accuracy of the LMR-FM-Net fused by the primary models with optimal performance is much higher than that of each component model. The recognition accuracy is 98.56 $\pm$ 1.67 [%].

B. The Impact of Input Size on Estimation for Different Primary Models

Fig. 5 shows the variation of the mean $\pm$ std of RMSE of the primary models with respect to different sizes of input set in four locomotion modes. It is the average RMSE of the knee and ankle angle estimations, which are used to evaluate the performance and determine the optimal input sets of the primary models. In Fig. 5A, it has a higher average RMSE for LSTM-Net in the stair descent. For all sets, the values significantly decrease in the order of stair descent > stair ascent > running > walking. This trend can also be seen in Fig. 5B. For Fig. 5C and 5D, the RMSE is largest in the stair descent mode and smallest in the walking mode, which is consistent with Fig. 5A and 5B. Unlike Fig. 5A and 5B, the RMSE of running is sometimes greater than the RMSE of stair ascent in Fig. 5C and 5D. Fig. 5A also shows that increasing the dimension of the input set does not necessarily lead to a continuous decline in the mean RMSE. This is also found in Fig. 5B to 5D.

The results of the selected input set for the primary models in different locomotion modes are shown as Table S.2 in Supplement materials. It should be noted that, CNN-LSTM-Net and CNN-GRU-Net all have the best performance at Set2 in stair descent and stair ascent, respectively. Compared with the determined input set, the difference is very small. Therefore, a compromise between the network training complexity and estimation performance can be reached.

C. The Impact of the Number of Primary Models in the Fusion Module on the Angle Estimation Performance

The fusion model improves the performance of joint angle estimation for all locomotion modes by assigning iteratively non-fixed weights at every step to the prediction of primary models (GRU-Net, LSTM-Net, CNN- GRU-Net, and CNN-LSTM-Net). Table II demonstrates that integrating the fusion module into Kinematic-Nets improves the performance of the primary models in all locomotion modes, i.e., significantly reducing the average RMSE and increasing the PCC. Among them, the fusion module based on four primary models has the best performance.

D. RMSE, PCC, and NRMSE for Knee and Ankle Estimation in Different Locomotion Modes

Fig. 6 shows the RMSE (A), PCC (B) and NRMSE(C) values for knee and ankle estimations of all subjects in different modes. For the estimation results of Kinematic-FM-Net in different locomotion modes, blue circles represent the estimation of the knee angle, red triangles represent the estimation results of the ankle angle and cyan squares represent the average value of each joint angle. For better comparison, the results are also presented as Table S.3 in Supplement materials.

E. Knee and Ankle Angles: Estimated Vs. Gait Data

As an example, the gait cycles of several subjects in different locomotion modes are plotted in Fig. 7 to demonstrate a sample qualitative comparison of actual value and estimation from the model. Then the estimations of knee

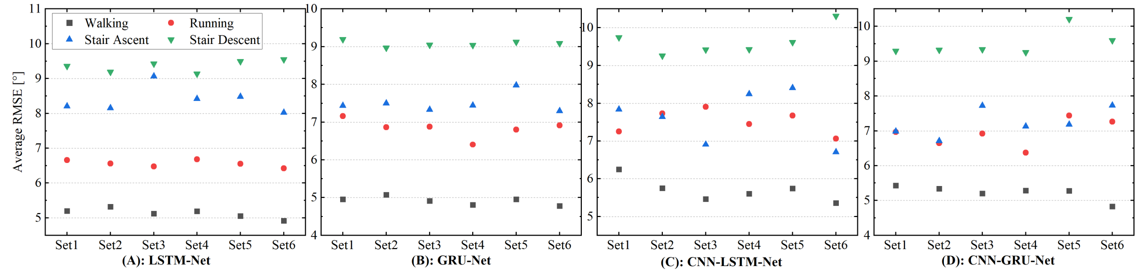

Fig. 5. The change of mean RMSE for the primary models with respect to different input sets under multimodal locomotion.

TABLE II
EFFECT OF FUSION OF DIFFERENT MODELS ON RMSE AND PCC FOR MULTIMODAL KINEMATIC MODELING

Models	Walking		Running		Stair ascent		Stair descent	
	RMSE (°)	PCC	RMSE (°)	PCC	RMSE (°)	PCC	RMSE (°)	PCC
L	4.91 ± 0.18	0.891 ± 0.015	6.42 ± 1.14	0.954 ± 0.008	8.02 ± 0.69	0.858 ± 0.019	9.14 ± 1.13	0.828 ± 0.039
G	4.78 ± 0.24	0.899 ± 0.021	6.41 ± 0.73	0.952 ± 0.009	7.29 ± 0.62	0.869 ± 0.022	9.04 ± 1.30	0.844 ± 0.027
CL	5.36 ± 0.71	0.859 ± 0.056	7.07 ± 1.09	0.945 ± 0.009	6.70 ± 0.45	0.890 ± 0.020	9.43 ± 1.21	0.833 ± 0.041
CG	4.83 ± 0.41	0.902 ± 0.015	6.36 ± 0.70	0.954 ± 0.009	7.13 ± 0.56	0.874 ± 0.021	9.25± 1.47	0.842 ± 0.035
L+G	4.78 ± 0.24	0.899 ± 0.019	6.44 ± 0.92	0.952 ± 0.010	7.29 ± 0.60	0.870 ± 0.020	8.96 ± 1.37	0.845 ± 0.031
L+CL	4.86 ± 0.16	0.894 ± 0.015	6.54 ± 1.18	0.953 ± 0.009	6.69 ± 0.45	0.891 ± 0.020	9.01 ± 1.15	0.832 ± 0.039
L+CG	4.83 ± 0.41	0.901 ± 0.015	6.34 ± 0.85	0.954 ± 0.009	6.98 ± 0.35	0.880 ± 0.014	8.92 ± 1.45	0.839 ± 0.040
G+CL	4.80 ± 0.26	0.897 ± 0.020	6.50 ± 0.90	0.951 ± 0.009	6.67 ± 0.47	0.891 ± 0.020	8.86 ± 1.31	0.848 ± 0.029
G+CG	4.79 ± 0.26	0.899 ± 0.020	6.39 ± 0.80	0.953 ± 0.009	6.99 ± 0.36	0.879 ± 0.015	8.94 ± 1.40	0.847 ± 0.030
CL+CG	4.81 ± 0.43	0.901 ± 0.015	6.42 ± 0.86	0.953 ± 0.009	6.65 ± 0.40	0.892 ± 0.019	9.08 ± 1.44	0.847 ± 0.034
L+G+CL	4.71 ± 0.24	0.901 ± 0.018	6.38 ± 0.91	0.953 ± 0.009	6.56 ± 0.51	0.895 ± 0.020	8.72 ± 1.24	0.849 ± 0.030
L+G+CG	4.74 ± 0.25	0.901 ± 0.019	6.32 ± 0.81	0.954 ± 0.009	6.83 ± 0.27	0.885 ± 0.013	8.78 ± 1.32	0.848 ± 0.030
L+CL+CG	4.70 ± 0.37	0.902 ± 0.018	6.29 ± 0.86	0.955 ± 0.009	6.52 ± 0.40	0.896 ± 0.019	8.74 ± 1.11	0.841 ± 0.035
G+CL+CG	4.72 ± 0.26	0.900 ± 0.019	6.32 ± 0.81	0.955 ± 0.009	6.51 ± 0.42	0.896 ± 0.019	8.70 ± 1.32	0.852 ± 0.027
Kinematic-FM-Net	**4.68 ± 0.25***	**0.904 ± 0.013***	**6.23 ± 0.83***	**0.956 ± 0.009***	**6.45 ± 0.45***	**0.898 ± 0.019***	**8.65 ± 1.25***	**0.851 ± 0.028***

In the table, L stands for LSTM-Net, G stands for GRU-Net, CL stands for CNN-LSTM-Net and CG stands for CNN-GRU-Net. The bold numbers represent the result of Kinematic-FM-Net which has the best performance. * denotes significant differences with other models.

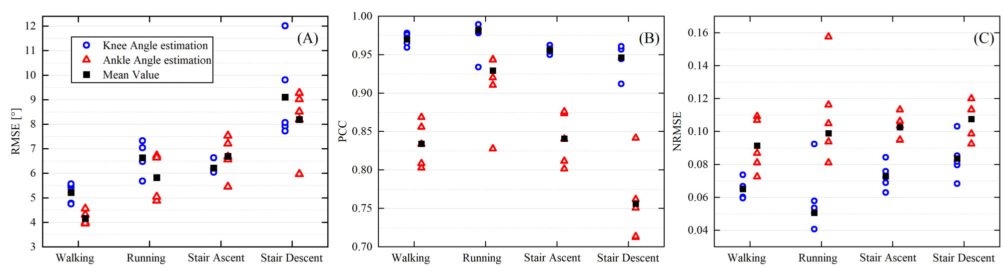

Fig. 6. (A) RMSE, (B) PCC and (C) NRMSE values for knee and ankle angle estimations in different locomotion modes.

and ankle angles are compared with gait data obtained from thie optical motion capture system according to multimodal kinematic modeling explained in the Section II. To validate the trajectories, we applied the estimated knee and ankle trajectories of one subject under walking to the human walking model built in Automatic Dynamic Analysis of Mechanical Systems (ADAMS). The walking simulation is performed in ADAMS. The simulation results are shown as Fig. S1 in the Supplementary material. From the results, the robot model is able to walk stably following the desired joint trajectories.

F. Comparison With State-of-the-Art

A number of studies have suggested different methods to estimate knee or ankle angles. In those works, different input sources, algorithms and data-driven models are proposed and used, and the approaches are verified on different numbers of participants. Table S.4, S.5, and S.6 in Supplement materials present a summary of the related results obtained from several studies and make a comparison with the results obtained in this study. The tables show that both mechanical and EMG signals are used to estimate the joint angles depending on the algorithm. Furthermore, biometric data such as age and height are used in a number of studies as well, e.g., [20] and [34]. The results of this current study are in the range reported by other studies.

V. DISCUSSION

In this study, hip kinematic information and individual biometric data of the subject are used to estimate knee and ankle angles for multiple motion conditions in daily life.

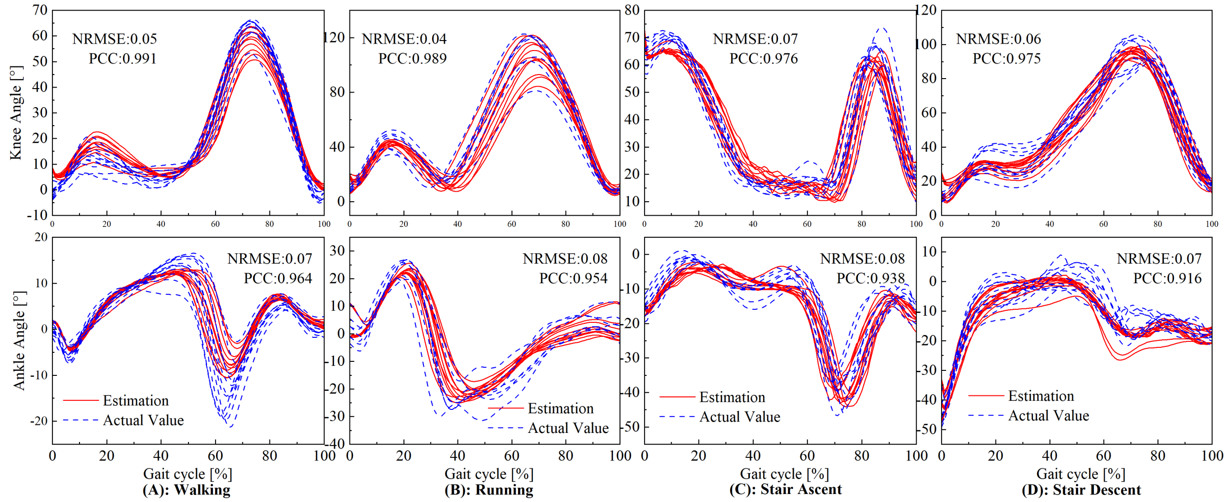

Fig. 7. The estimated knee and ankle angles vs. gait data for various locomotion modes are explained in the methods section.

To this end, a multimodal feature decoupled kinematic modeling system fusing primary deep neural networks is constructed and the results are shown. Our system model is applied to the public dataset and five-fold cross-validation is performed. A wide range of speeds and population characteristics are included in the datasets, which makes our model have a high generalization capacity. In addition, as our model is established and evaluated on the public datasets, other researchers can validate our model with their algorithms and also have the potential to improve our contribution in this field further.

In this paper, a data augmentation module is designed based on HHT and MODWT-MRA-HT. This module is capable of signal decomposition and time-frequency domain feature expansion for hip kinematic signals. Firstly, EMD and MODWT-MRA are utilized for signal decomposition, while the time-frequency domain information (instantaneous frequency and energy) of the analytic signal can be obtained after the Hibbert transform. Then, raw and augmented analytic signals are combined to produce six different input sets for the primary models. Fig. 4 and 5 demonstrate the locomotion mode recognition and kinematic estimation performance of the four primary models for different input size. The classification performance of LSTM-Net, GRU-Net, CNN-LSTM-Net and CNN-GRU-Net is improved when the HHT-based data augmentation is used. It is worth noting that when the input is Set3, the accuracy of GRU-Net is 0.01 less than the optimal (Set1). For lower extremity joint kinematics estimation of the primary models, the inclusion of the data augmentation module is beneficial. In particular, from the performance of the primary models in Fig. 5 and Table S.2, the data augmentation of MODWT-MRA-HT is better. From the results, HHT is suitable for time series classification, while MODWT-MRA-HT is more suitable for kinematic estimation. This may be due to the fact that continuously kinematic estimation is harder and requires more input features than mode classification. Overall, the inclusion of the data augmentation module can result in improved classification and regression performance of the conventional primary models.

The LMR-FM-Net and Kinematic-FM-Net proposed in this paper consist of four primary models constructed by different traditional deep learning layers and multi-source information fusion algorithms. Multimodal feature decoupling is achieved using deep learning networks for locomotion mode recognition. This enables the wearable assistive device to quickly and accurately update the control strategy in multi-tasking scenarios. LMR-FM-Net is built by integrating the four primary models together using the hybrid algorithm based on ensemble voting and Markov chain. From the results in Fig. 4, the classification accuracy of the four optimal primary models is greater than 94% with some differences in their network structures, thus can be used as base classifiers for ensemble learning. Compared with the base model, LMR-FM-Net with fusion module can effectively improve the accuracy of locomotion mode recognition.

The primary models for kinematic estimation with different network parameters are developed for multiple motion tasks. The kinematic estimation performance of each base model is improved so that the estimation accuracy of the Kinematic-FM-Net can maximized based on the Kalman filter. From the results in Table II, the fusion model integrated with four primary models outperforms other combination approaches or single models. This validates that our approach of fusing different types of primary models can improve locomotion mode classification accuracy and kinematic estimation performance. Moreover, the direction of further reducing the RMSE of joint angle estimation can be seen through the fusion of more primary models.

According to the results shown in subsections III-C, D and E, the estimation quality is generally higher for faster movement or proximal joints. Table II shows that all primary models and their combinations have the highest average PCC and the smallest standard deviation in running mode. The minimum speed in this mode is 2.5 m/s which is much greater than the speeds in the other modes. This might potentially have some connections to the fact that a higher speed of motion is energetically more efficient [54]. The estimation performance of different speeds in the first and second datasets is also investigated and comparatively analyzed. For walking, the PCC of the three speeds under over-ground walking condition is 0.906 ± 0.032, 0.911 ± 0.020, and 0.919 ± 0.016 (OS, OC, and OF, respectively). For walking on a treadmill, the estimated performance is also positively correlated with speed (T01: 0.855 ± 0.029, T03: 0.893 ± 0.020, T05: 0.909 ± 0.023 and T07: 0.915 ± 0.017, respectively). The PCC of the three speeds under running condition is 0.939 ± 0.013, 0.960 ± 0.013, and 0.961 ± 0.015 (2.5 m/s, 3.5 m/s and 4.5 m/s, respectively). The results show that the same discipline exists when comparing the estimated performance of different speeds. Nevertheless, whether and how kinematics and gait energetics could be correlated and impact the estimation performance requires further investigations. On the other hand, the results of stair ascents and descents do not support the above statement. Though the mean velocities of the two modes are similar (0.49 and 0.52), the values of NRMSE and PCC differ considerably. This may be due to the different biomechanical properties of lower limb joints in these two modes. Unlike walking and running, limb strength of the subjects, such as the power of the rectus femoris, had a greater influence on the gait cycle during stair ascent and descent. There is no significant correlation between gait behavior and speed in these two modes.

From the results in Fig. 6 and Table S.3, the performance of the ankle estimation is inferior to the knee. There is a greater correlation between the hip and knee for each locomotion mode compared to the ankle. Furthermore, the standard deviation at the ankle is higher in comparison to the knee. This may be due to the fact that the hip and knee are proximal joints that are more kinematically related to each other. The same result has been seen in the [31] where the estimation performance of the ankle angle using the shank angle is better than the estimation using the thigh angle. The estimation of the ankle angle in stair descending has the poorest performance compared to the other models. This may be due to the low contribution of this joint during stair descent. The complex kinematic properties of this joint create difficulties in estimating the angles of the primary models. Another explanation could be that the number of input sources or type of inputs, or even the processing algorithm is not suitable for that locomotion mode. The results might be improved if, e.g., the inputs are combined with EMG signals or if the number of inputs is increased [29], [32], [33]. As can be seen in Fig. 7, the gait cycle profile of the knee joint is simpler compared to the ankle. This enables the deep learning model to have a higher kinematic modeling performance of the knee. These characteristics can be considered when designing the control strategies for wearable assistive robots. The above discussions can form starting points for further investigations.

Many studies have applied deep learning algorithms to estimate the kinematics of different locomotion. In those works, different algorithms and inputs are proposed and used, while the approaches are verified on different numbers of participants. Table S.4, S.5 and S.6 present the performance of multiple deep learning algorithms for kinematics in other works. However, it is not valid to take a direct comparison of their results with those of this study due to the different sensor modalities, number of sensors, motion conditions and number of subjects. Moreover, those datasets are not publicly available, which makes it challenging to apply our system models to make comparison. Therefore, the differences between their findings and ours can only be discussed qualitatively. Interestingly, a part of our findings is similar to what other researchers have reported in different studies. For example, some studies report better estimation performance (PCC and NRMSE) for the knee in comparison to the ankle joint in all modes. Lu et al. [29] implemented kinematic estimation of lower limb joints from sEMG signals based on the Conv-LSTM in multiple locomotion modes. Unlike our model, static and anthropometric characteristics (e.g., age, sex, height, weight, BMI, etc.) were not used in their study. These parameters can provide necessary information for human kinematics estimation. Fusing this information into the model has the potential to improve prediction performance. In addition, results reported by [29], [32], and [33] show that our estimation performance for the ankle joint is a little worse than theirs. Their algorithms may benefit from the EMG signal as an input source. However, the acquisition and processing procedures of EMG signals are complicated, while the signal accuracy is affected by the external environment and limb movement. Therefore, it is not convenient for the development of portable wearable assistive devices. More importantly, the generalization capability of our model is better than others due to the number of subjects, while the range of test speeds in our work is much larger than others. Future research on motion planners of wearable assistive devices can be applied to healthy people.

This work aims to recognize locomotion mode and create a kinematics connection between the hip and the joint beneath it, i.e., the knee (ankle) joint. Due to the estimated knee and ankle angles, the system model presented in this study can be applied to the real-time control of wearable intelligent assistive devices, such as the case in [21], [55], and [56]. The embedded platform on which the model is deployed can be used as a high-level controller to convert all inputs into the desired angle for the target joints. Finally, the command of the actuator will be a function of hip kinematics and individual characterization parameters. Hip kinematics can be obtained from IMUs mounted on the thigh and upper limb. In these cases, the actuator is expected to have no limitations on the range of joint kinematics and to output assistive torque.

Although this paper provides an accurate prediction of joint angles with extensive locomotion conditions, speeds, good number of subjects and three independent public datasets, there are several limitations to this study. Firstly, a relatively few number of primary models (four) is used to estimate single-limb kinetics. This may limit the performance of the final fusion model. More specifically, by meeting the configuration parameters of the portable device, the estimation performance of the fusion model can be further improved by using more high-performance primary models. However, a trade-off between accuracy and computation complexity need to be made. Secondly, there are interactions and effects between the bilateral limbs during movement. IMU sensors mounted on bilateral lower limbs will collect more meaningful kinematic information, which may be able to further enhance the estimation. However, the number of IMU sensors needs to be minimized for system robustness and wearer comfort. Therefore, the use of bilateral hip kinematics as an input source can be considered, which can utilize the common information from the IMU sensors mounted at the upper limb. Finally, compared to the other modalities, the estimation of stair descent has the highest error values. In particular, the ankle has a lower PCC and the highest NRMSE. Therefore, a direction for future improvement towards improving the ankle estimation performance may be to add other relevant input sources (e.g., EMG signal) and make it roughly equal to other locomotion.

In this work, we have decent accuracy for healthy individuals from public datasets. However, patients with gait abnormalities who have stroke and musculoskeletal issues may not get an accurate estimation from our system model due to the lack of training data. Data-driven methods based on deep learning need to be supported by a large amount of high-quality data. The models we constructed using publicly available datasets can meet the basic training requirements. However, during the development of controllers for portable intelligent assistive devices, the predicted results are often not accurate enough due to the discrepancy between constructed and actual data. Thus, transfer learning may be a probable future direction to address the limitation [57]. Based on the rational use of public datasets to construct the model, it is also possible to obtain good results in the actual testing process. This can accelerate the development of prototypes.

As we use extensive datasets in the system model, our outcome provides reliable estimates of joint angles similar to state-of-the-art studies. Thus, our algorithm can be used to measure the joint angles in the clinic or other research labs where they cannot accommodate prohibitive measurement modalities such as motion capture cameras. Our model also has the potential to serve as a platform that enables updates or modifications via retraining of the model to accommodate kinematics estimation of different subject populations who have musculoskeletal issues or aging in the future.

VI. Conclusion

In this study, a novel multimodal feature decoupled kinematic estimation system is proposed for estimating sagittal knee and ankle joint angles in multiple locomotion conditions and speeds. From extensive evaluation of the developed model, the design choices are justified. This accurate estimation method will enable tracking of kinematics parameters outside the lab and removing the limitation of traditional optical motion capture systems. There are still some limitations in this paper that could be further improved, such as the fewer number of primary models, interference between bilateral limbs, and low accuracy of the stair descent. To verify the feasibility of the proposed method, one direction for future work is to design a portable smart assistive device based on wearable sensors. In addition, to obtain a more comprehensive algorithm, the work can be continued to investigate the functionality of the proposed estimation method for other locomotion, such as swimming.

Acknowledgment

The gait data from three different publications [40], [41], and [42] are used in this study. Hereby, the authors would like to express their gratitude to the authors of these publications for sharing their data.

References

[1] E. Andrenelli et al., "Features and predictors of activity limitations and participation restriction 2 years after intensive rehabilitation following first-ever stroke," *Eur. J. Phys. Rehabil. Med.*, vol. 51, no. 5, pp. 575–585, Oct. 2015.

[2] V. D. O. Fernandes, B. D. S. Moreira, G. A. S. C. de Melo, N. C. P. de Avelar, H. S. Costa, and A. D. C. Bastone, "Factors associated with fear of falling in older women with knee osteoarthritis: A cross-sectional study," *Geriatric Nursing*, vol. 55, pp. 333–338, Jan. 2024.

[3] L. N. Awad et al., "A soft robotic exosuit improves walking in patients after stroke," *Sci. Transl. Med.*, vol. 9, no. 400, Jul. 2017, Art. no. eaai9084.

[4] J. Kim et al., "Reducing the metabolic rate of walking and running with a versatile, portable exosuit," *Science*, vol. 365, no. 6454, pp. 668–672, Aug. 2019.

[5] Y. David Li and E. T. Hsiao-Wecksler, "Gait mode recognition and control for a portable-powered ankle-foot orthosis," in *Proc. IEEE 13th Int. Conf. Rehabil. Robot. (ICORR)*, Jun. 2013, pp. 1–8.

[6] W. Huo, S. Mohammed, Y. Amirat, and K. Kong, "Fast gait mode detection and assistive torque control of an exoskeletal robotic orthosis for walking assistance," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1035–1052, Aug. 2018.

[7] D. J. Farris and G. S. Sawicki, "The mechanics and energetics of human walking and running: A joint level perspective," *J. Roy. Soc. Interface*, vol. 9, no. 66, pp. 110–118, Jan. 2012.

[8] D. Samuel, P. Rowe, V. Hood, and A. Nicol, "The biomechanical functional demand placed on knee and hip muscles of older adults during stair ascent and descent," *Gait Posture*, vol. 34, no. 2, pp. 239–244, Jun. 2011.

[9] R. Baud, A. R. Manzoori, A. Ijspeert, and M. Bouri, "Review of control strategies for lower-limb exoskeletons to assist gait," *J. NeuroEng. Rehabil.*, vol. 18, no. 1, p. 119, Jul. 2021, doi: 10.1186/S12984-021-00906-3.

[10] L. Ballaz, S. Plamondon, and M. Lemay, "Ankle range of motion is key to gait efficiency in adolescents with cerebral palsy," *Clin. Biomech.*, vol. 25, no. 9, pp. 944–948, Nov. 2010.

[11] W. R. Stevens, D. A. Podeszwa, and K. Tulchin-Francis, "Compensatory sagittal plane ankle gait mechanics: Are they present in patients with a weak or stiff hip?" *Gait Posture*, vol. 74, pp. 250–254, Oct. 2019.

[12] L. Chen et al., "A portable waist-loaded soft exosuit for hip flexion assistance with running," *Micromachines*, vol. 13, no. 2, p. 157, Jan. 2022.

[13] H. D. Lee, H. Park, B. Seongho, and T. H. Kang, "Development of a soft exosuit system for walking assistance during stair ascent and descent," *Int. J. Control, Autom. Syst.*, vol. 18, no. 10, pp. 2678–2686, Oct. 2020.

[14] S. L. Delp et al., "OpenSim: Open-source software to create and analyze dynamic simulations of movement," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 11, pp. 1940–1950, Nov. 2007.

[15] I. Weygers, M. Kok, M. Konings, H. Hallez, H. De Vroey, and K. Claeys, "Inertial sensor-based lower limb joint kinematics: A methodological systematic review," *Sensors*, vol. 20, no. 3, p. 673, Jan. 2020.

[16] T. Seel, J. Raisch, and T. Schauer, "IMU-based joint angle measurement for gait analysis," *Sensors*, vol. 14, no. 4, pp. 6891–6909, Apr. 2014.

[17] T. Zimmermann, B. Taetz, and G. Bleser, "IMU-to-segment assignment and orientation alignment for the lower body using deep learning," *Sensors*, vol. 18, no. 1, p. 302, Jan. 2018.

[18] M. Mundt, A. Koeppe, F. Bamer, S. David, and B. Markert, "Artificial neural networks in motion analysis—Applications of unsupervised and heuristic feature selection techniques," *Sensors*, vol. 20, no. 16, p. 4581, Aug. 2020.

[19] V. B. Semwal et al., "Development of the LSTM model and universal polynomial equation for all the sub-phases of human gait," *IEEE Sensors J.*, vol. 23, no. 14, pp. 15892–15900, Jul. 2023.

[20] V. B. Semwal, C. Kumar, P. K. Mishra, and G. C. Nandi, "Design of vector field for different subphases of gait and regeneration of gait pattern," *IEEE Trans. Autom. Sci. Eng.*, vol. 15, no. 1, pp. 104–110, Jan. 2018.

[21] D. Quintero, D. J. Villarreal, D. J. Lambert, S. Kapp, and R. D. Gregg, "Continuous-phase control of a powered knee-ankle prosthesis: Amputee experiments across speeds and inclines," *IEEE Trans. Robot.*, vol. 34, no. 3, pp. 686–701, Jun. 2018.

[22] K. R. Embry and R. D. Gregg, "Analysis of continuously varying kinematics for prosthetic leg control applications," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 262–272, 2021.

[23] E. Reznick, K. Embry, and R. D. Gregg, "Predicting individualized joint kinematics over a continuous range of slopes and speeds," in *Proc. 8th IEEE RAS/EMBS Int. Conf. Biomed. Robot. Biomechatronics (BioRob)*, Nov. 2020, pp. 666–672.

[24] F. Moissenet, F. Leboeuf, and S. Armand, "Lower limb sagittal gait kinematics can be predicted based on walking speed, gender, age and BMI," *Sci. Rep.*, vol. 9, no. 1, 2019, Art. no. 9510, doi: 10.1038/s41598-019-45397-4.

[25] V. B. Semwal, N. Gaud, P. Lalwani, V. Bijalwan, and A. K. Alok, "Pattern identification of different human joints for different human walking styles using inertial measurement unit (IMU) sensor," *Artif. Intell. Rev.*, vol. 55, no. 2, pp. 1149–1169, Feb. 2022.

[26] V. B. Semwal, A. Gupta, and P. Lalwani, "An optimized hybrid deep learning model using ensemble learning approach for human walking activities recognition," *J. Supercomput.*, vol. 77, no. 11, pp. 12256–12279, Nov. 2021.

[27] C. Pham et al., "SensCapsNet: Deep neural network for non-obtrusive sensing based human activity recognition," *IEEE Access*, vol. 8, pp. 86934–86946, Jan. 2020, doi: 10.1109/ACCESS.2020.2991731.

[28] S. K. Challa, A. Kumar, V. B. Semwal, and N. Dua, "An optimized-LSTM and RGB-D sensor-based human gait trajectory generator for bipedal robot walking," *IEEE Sensors J.*, vol. 22, no. 24, pp. 24352–24363, Dec. 2022.

[29] Y. Lu, H. Wang, B. Zhou, C. Wei, and S. Xu, "Continuous and simultaneous estimation of lower limb multi-joint angles from sEMG signals based on stacked convolutional and LSTM models," *Expert Syst. Appl.*, vol. 203, Oct. 2022, Art. no. 117340.

[30] M. Sharifi-Renani, M. H. Mahoor, and C. W. Clary, "BioMAT: An open-source biomechanics multi-activity transformer for joint kinematic predictions using wearable sensors," *Sensors*, vol. 23, no. 13, p. 5778, Jun. 2023.

[31] M. Eslamy and A. F. Schilling, "Estimation of knee and ankle angles during walking using thigh and shank angles," *Bioinspiration Biomimetics*, vol. 16, no. 6, Nov. 2021, Art. no. 066012.

[32] E. V. Zabre-Gonzalez, L. Riem, P. A. Voglewede, B. Silver-Thorn, S. R. Koehler-McNicholas, and S. A. Beardsley, "Continuous myoelectric prediction of future ankle angle and moment across ambulation conditions and their transitions," *Front Neurosci-Switz*, vol. 15, 2021, Art. no. 709422, doi: 10.3389/fnins.2021.709422.

[33] Q. Song, X. Ma, and Y. Liu, "Continuous online prediction of lower limb joints angles based on sEMG signals by deep learning approach," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107124.

[34] V. B. Semwal, R. Jain, P. Maheshwari, and S. Khatwani, "Gait reference trajectory generation at different walking speeds using LSTM and CNN," *Multimedia Tools Appl.*, vol. 82, no. 21, pp. 33401–33419, Sep. 2023.

[35] D. Chow, Z. Iqbal, L. Tremblay, C.-Y. Lam, and R.-B. Zhao, "Cross-leg prediction of running kinematics across various running conditions and drawing from a minimal data set using a single wearable sensor," *Symmetry*, vol. 14, no. 6, p. 1092, May 2022.

[36] A. Kumar, A. K. Godiyal, and D. Joshi, "Force myography based continuous estimation of knee joint angle using artificial neural network," in *Proc. IEEE 5th Int. Conf. Converg. Technol. (I2CT)*, Mar. 2019, pp. 1–3.

[37] M. Gholami, C. Napier, and C. Menon, "Estimating lower extremity running gait kinematics with a single accelerometer: A deep learning approach," *Sensors*, vol. 20, no. 10, p. 2939, May 2020.

[38] A. Rezaie Zangene et al., "An efficient attention-driven deep neural network approach for continuous estimation of knee joint kinematics via sEMG signals during running," *Biomed. Signal Process. Control*, vol. 86, Sep. 2023, Art. no. 105103.

[39] M. S. B. Hossain, J. Dranetz, H. Choi, and Z. Guo, "DeepBBWAE-Net: A CNN-RNN based deep superlearner for estimating lower extremity sagittal plane joint kinematics using shoe-mounted IMU sensors in daily living," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 8, pp. 3906–3917, Aug. 2022, doi: 10.1109/JBHI.2022.3165383.

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

[41] R. K. Fukuchi, C. A. Fukuchi, and M. Duarte, "A public dataset of running biomechanics and the effects of running speed on lower extremity kinematics and kinetics," *PeerJ*, vol. 5, p. e3298, May 2017.

[42] T. Lencioni, I. Carpinella, M. Rabuffetti, A. Marzegan, and M. Ferrarin, "Human kinematic, kinetic and EMG data during different walking and stair ascending and descending tasks," *Sci. Data*, vol. 6, no. 1, 2019, Art. no. 309, doi: 10.1038/s41597-019-0323-z.

[43] 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, p. 10, Feb. 2021.

[44] K. Bandara, H. Hewamalage, Y.-H. Liu, Y. Kang, and C. Bergmeir, "Improving the accuracy of global forecasting models using time series data augmentation," *Pattern Recognit.*, vol. 120, Dec. 2021, Art. no. 108148.

[45] Q. Wen et al., "Time series data augmentation for deep learning: A survey," in *Proc. 13th Int. Joint Conf. Artif. Intell.*, Aug. 2021, pp. 4653–4660.

[46] G. Khademi and D. Simon, "Toward minimal-sensing locomotion mode recognition for a powered knee-ankle prosthesis," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 3, pp. 967–979, Mar. 2021.

[47] N. E. Huang et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. Lond. Ser. Math. Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[48] A. Goshvarpour and A. Goshvarpour, "Nonlinear analysis of human gait signals," *Int. J. Inf. Eng. Electron. Bus.*, vol. 4, no. 2, pp. 15–21, Apr. 2012.

[49] G.-S. Huang, C.-C. Wu, and J. Lin, "Gait analysis by using tri-axial accelerometer of smart phones," in *Proc. Int. Conf. Computerized Healthcare (ICCH)*, Dec. 2012, pp. 29–34.

[50] D. B. Percival and A. T. Walden, *Wavelet Methods for Time Series Analysis* (Cambridge Series in Statistical and Probabilistic Mathematics). Cambridge, U.K.: Cambridge Univ. Press, 2000.

[51] Q. Wen, K. He, L. Sun, Y. Zhang, M. Ke, and H. Xu, "RobustPeriod: Robust time-frequency mining for multiple periodicity detection," in *Proc. Int. Conf. Manage. Data*, Jun. 2021, pp. 2328–2337.

[52] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," presented at the 25th Int. Conf. Neural Inf. Process. Syst., vol. 2. Lake Tahoe, NV, USA, 2012.

[53] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimisation," 2014, *arXiv:1412.6980*.

[54] M. Eslamy, "Emulation of ankle function for different gaits through active foot prosthesis: Actuation concepts, control and experiments," Ph.D. dissertation, Dept. Mech. Eng., Technische Universität, Darmstadt, Germany, 2014.

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

[56] T. Elery, S. Rezazadeh, C. Nesler, and R. D. Gregg, "Design and validation of a powered knee-ankle prosthesis with high-torque, low-impedance actuators," *IEEE Trans. Robot.*, vol. 36, no. 6, pp. 1649–1668, Dec. 2020.

[57] J. Zou, X. Zhang, Y. Zhang, J. Li, and Z. Jin, "Prediction on the medial knee contact force in patients with knee valgus using transfer learning approaches: Application to rehabilitation gaits," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106099.