Lower Limb Activity Recognition Based on sEMG Using Stacked Weighted Random Forest

Cheng Shen, Zhongcai Pei, Weihai Chen, *Member, IEEE*, Jianhua Wang, Xingming Wu, and Jianer Chen

***Abstract*— The existing surface electromyography-based pattern recognition system (sEMG-PRS) exhibits limited generalizability in practical applications. In this paper, we propose a stacked weighted random forest (SWRF) algorithm to enhance the long-term usability and user adaptability of sEMG-PRS. First, the weighted random forest (WRF) is proposed to address the issue of imbalanced performance in standard random forests (RF) caused by randomness in sampling and feature selection. Then, the stacking is employed to further enhance the generalizability of WRF. Specifically, RF is utilized as the base learner, while WRF serves as the meta-leaning layer algorithm. The SWRF is evaluated against classical classification algorithms in both online experiments and offline datasets. The offline experiments indicate that the SWRF achieves an average classification accuracy of 89.06%, outperforming RF, WRF, long short-term memory (LSTM), and support vector machine (SVM). The online experiments indicate that SWRF outperforms the aforementioned algorithms regarding long-term usability and user adaptability. We believe that our method has significant potential for practical application in sEMG-PRS.**

***Index Terms*— Biomedical signal, pattern recognition, lower limb activity, electromyography, random forest.**

I. Introduction

With the exacerbation of the global aging trend, there has been a rise in the proportion of individuals experiencing mobility impairments. The lack of mobility not only imposes burden on society and families, but also weakens the confidence in their daily life. Lower limb activity (LLA) is an essential daily behavior in human life, as the lower limb bears the majority of body and serves as a foundation for robust waking ability [1]. To improve the walking ability of elderly and mobility-impaired individuals, a vast array of intelligent lower limb intelligent rehabilitation devices is employed. However, due to the lack of active participation intention of the wearer, traditional human computer interaction (HCI) based on fixed rehabilitation modalities has limited rehabilitation effect [2]. Research indicates that active rehabilitation yields superior outcomes compared to passive rehabilitation. For exoskeleton robots, accurate and efficient LLA intention recognition is a prerequisite for effective rehabilitation training, and it directly affects the overall outcome.

Unlike the inefficient traditional HCI, the bioelectric signal-based control approach is capable of recognizing the wearer's rehabilitation intention. The wearer's intention is appropriately integrated into passive rehabilitation, thereby enhancing the wearer's subjective initiative, and improve the rehabilitation effect. The surface electromyography (sEMG)-based pattern recognition system (sEMG-PRS) shows potential in providing stable and effective intention recognition for rehabilitation devices [3], [4]. The sEMG signal is formed by the superposition of action potentials generated by all muscle fibers on the skin surface [5]. Moreover, the sEMG signal exhibits a low amplitude range of 0-5 mV, a frequency spectrum spanning from 6 to 500 Hz, and a predominant energy distribution within the range of 10-250 Hz [6]. The sEMG-PRS mainly includes four interrelated stages: sEMG data acquisition, pre-processing, feature extraction, and classification. Despite numerous experimental studies with good intention recognition [7], [8], the abandonment rate of lower limb rehabilitation devices remains high. The primary impediment to this phenomenon lies in the gradual decline of recognition performance exhibited by sEMG-PRS during extended periods of clinical application. Over extended periods of use, the performance of sEMG-PRS decreases dramatically due to various factors such as electrode shift, limb position, muscle fatigue, and skin sweating [9]. Furthermore, the generalization across subjects is more complex issue due to factors such as inter-individual differences, variations in muscle strength, and disparities in skin organization [10]. The issue of limited generalization leads to insufficient long-term availability and

Manuscript received 3 July 2023; revised 9 November 2023; accepted 17 December 2023. Date of publication 25 December 2023; date of current version 16 January 2024. This work was supported in part by the National Nature Science Foundation under Grant 62333023, Grant 51975002, Grant U1909215, and Grant 51975029; in part by the Key Research and Development Program of Zhejiang Province under Grant 2021C03050; and in part by the Scientific Research Project of Agriculture and Social Development of Hangzhou under Grant 20212013B11. *(Corresponding author: Weihai Chen.)*

Cheng Shen and Zhongcai Pei are with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China, and also with the Hangzhou Innovation Institute, Beihang University, Hangzhou, Zhejiang 310052, China (e-mail: cshen0322@outlook.com; peizc@buaa.edu.cn).

Weihai Chen is with the School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China (e-mail: whchenbuaa@126.com).

Jianhua Wang and Xingming Wu are with the Hangzhou Innovation Institute, Beihang University, Hangzhou, Zhejiang 310052, China, and also with the School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China (e-mail: jhwangbuaa@126.com; xmwubuaa@163.com).

Jianer Chen is with the Third Clinical Affiliated Hospital, Zhejiang Chinese Medical University, Hangzhou, Zhejiang 310009, China (e-mail: chenje@zcmu.edu.cn).

Digital Object Identifier 10.1109/TNSRE.2023.3346462

user adaptability of sEMG-PRS [11]. Therefore, there is an urgent need to enhance the generalization of sEMG-PRS and bridge the translational gap between laboratory research and clinical applications.

Several approaches have been proposed to enhance the generalizability of sEMG-PRS. The most intuitive approach is to identify the optimal features with strong robustness. Jiang et al. [12] identified the optimal feature combination of sEMG features by searching and selecting among 50 different options, resulting in an impressive cross-day accuracy of 91.9 % for their 11-gesture classifications. Shen et al. [13] proposed a novel feature extraction method to diminish the influence of limb position on sEMG-based pattern recognition. As is often the case, reducting the dimensionality of features (including principal components analysis and linear discriminant analysis) and subsequent processing can generally lead to improved performance [14], [15]. A potential approach for enhancing model generalization is to ensure sufficient training time and diverse training data from a large number of individuals, as this can significantly improve the recognition performance of the sEMG-PRS [16]. By incorporating a substantial amount of disparity data, this approach enhances the comprehensiveness of system recognition. Waris et al. [17] demonstrated that increasing the time span of training data could improve the robustness of the sEMG-PRS over time. However, feature selection is subjective and prone to overestimation the results. Obtaining a substantial amount of training data is time-consuming and unrealistic. In fact, experimental data collection can be tedious and exhausting for individuals, and recruiting a large number of subjects can prove extremely challenging.

With the progress of deep learning, some studies have been introduced to the sEMG-PRS. Transfer learning (TL) is a widely accepted approach for addressing data scarcity while maintaining classification accuracy. In [18], the utilization of TL method was significantly outperformed the non-TL method in addressing the performance degradation caused by the instability of sEMG signals. Depth domain adaptation (DDA) is another widely investigated approach. DDA aligns the data distribution between source and target domains, utilizing limited labeled target data to provide a stable and reliable classification capacity for diverse target domains [19]. Zhang et al. [20] proposed a multi-source synchronize domain adaptation framework that possesses both domain adaptation and domain generalization capability. Zhai et al. [21] proposed a CNN-based self-recalibrating unsupervised classifier to achieve 78.71 % classification ability under intersession scenarios. Shi et al. [22] proposed a new multi-task dual-stream supervised domain adaptation network based on convolutional neural network to realize the long-term reliability and user adaptability simultaneously. Experiments conducted on 12 non-disabled individuals have demonstrate that the proposed method surpasses both convolutional neural network and fine-tuning. Campbell et al. [23] expanded the adaptive domain adversarial neural network (ADANN) method to cross-subject networks. The accuracy of ADANN for intact limbs was shown to reach 86.8 – 96.2% by three dissimilar evaluation. These methods have the potential to enhance the cross-subject and inter-session performance of the sEMG-PRS. However, TL and DDA are implemented from the perspective of global distribution alignment and lack semantic information on sEMG features. Although TL reduces the need for retraining and recalibration, a convolutional neural network still requires a large amount of data to achieve high performance and generalizability. Furthermore, these methods necessitate the utilization of source data and recalibration the trained model, which can cause inconvenience for sEMG-PRS.

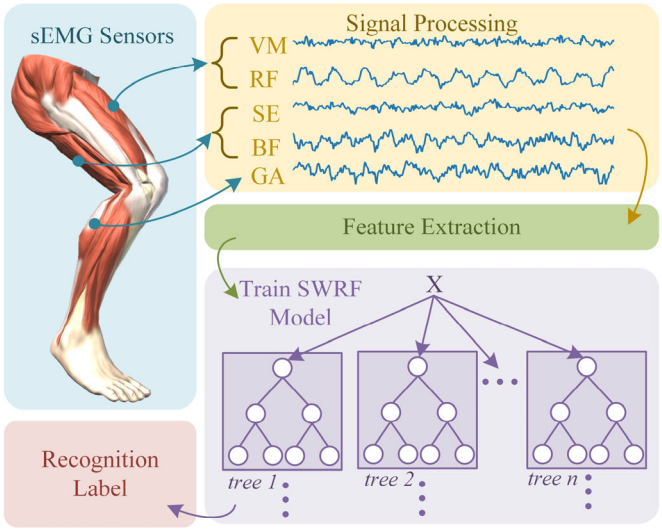

Fig. 1. Entire frame of lower limb activity recognition sEMG-based. First, sEMG signals are acquired and preprocessed from five muscles in the lower limb. Second, the SWRF classifier is trained using the extracted features. Third, the trained SWRF model is employed for action label recognition.

In summary, despite the extensive research conducted to enhance the generalizability of sEMG-PRS for clinical applications. Unfortunately, few studies have investigated a simple model that exhibits both durability in long-term usage and widespread adoption by multiple users [24], [25]. Here, this paper proposes novel classification method, namely stacked weighted random forest (SWRF), which is illustrated in Fig. 1. To validate the efficacy of SWRF, we selected four commonly lower limb activities and recruited four non-disabled subjects to carry out the experiment for five consecutive days. An LLA dataset was utilized to evaluate the robustness and adaptability of SWRF. Afterward, we evaluated the SWRF's generalization performance in three aspects, inter-subject offline dataset, inter-subject single-day, and subject multi-day.

The main contribution of the work is twofold. First, complements the existing sEMG-PRS approach and proposes a classification model with both long-term robustness and multi-subject adaptability, termed SWRF. Second, a multi-subject sEMG signal acquisition experiment is designed to evaluate the performance of SWRF in three aspects, including inter-subject offline dataset, inter-subject single-day, subject multi-day. We noted that a shorter conference version of this paper appeared in [26]. The initial conference paper online experiments were not sufficient, and no offline dataset experiments were performed. This manuscript addresses these issues

TABLE I
DEMOGRAPHIC DATA OF THE SUBJECTS

Subject	Laterality	Gender	Age(years)	Height(cm)	Weight(kg)
1	Right hand	male	30	172	74
2	Right hand	male	27	168	64
3	Left hand	male	25	186	90
4	Right hand	male	24	173	68
	Average		26.50	174.75	74.00

and propose an improved approach based on the weighted random forest presented in the conference paper.

The remainder of the paper is organized as follows. Dataset and sEMG acquisition are presented in Section II. The proposed SWRF is elaborated in Section III. The experiment results and analysis are provided in Section IV. Finally, conclusions are drawn in the last section.

II. DATASET AND SEMG ACQUISITION

A. Dataset Descriptions

The dataset utilized in this work was curated by Oscar Sanchez and Jose Sotelo [27] (http://archive.ics.uci.edu/dataset/278/emg+dataset+in+lower+limb). The dataset contained 11 samples of subjects with some abnormality in the knee previously diagnostic and 11 exhibiting normality. Subjects experienced three movements to analyze behaviors associated with knee muscles, gait, leg extension from a seated position, and flexion of the leg bending while performing sitting, standing and walking tasks. The acquisition process was conducted utilizing four electrodes (vastus medialis, biceps femoris, semitendinosus, and rectus femoris) and knee goniometers. This paper utilized sEMG data from 11 healthy individuals for analysis.

B. Subjects and Lower Limb Movements

Four subjects volunteered to participate in the experiment and the specific details are presented in Table I. None of the subjects claimed to have any motor impairment or disability, and each received comprehensive information regarding the experimental procedures and associated risks. All subjects signed an informed consent.

Rehabilitation exercises such as leg extension and bending are commonly employed for stroke patients. Based on this commonly used movement, we have formulated four leg movements: sitting position (ST), raise thigh (RT), raise calf (RC), straighten leg (SL), as illustrated in Fig. 2. Meanwhile, the sEMG are recorded form a total of five muscles involved in movement, namely: rectus femoris (RF), vastus medialis (VM), biceps femoris (BF), semitendinosus (SE), and gastrocnemius (GA) [28], as illustrated in Fig. 2. Both RF and VM muscles play a crucial role in knee extension movement [29]. SE and BF play an important role in knee flexion movement [30]. GA is related to the stability of standing and walking movement of lower limbs. The position of the sticker sensor corresponds to the muscle position. Due to their location on the opposite side of the lower limb, VM and BF are not depicted in Fig. 2.

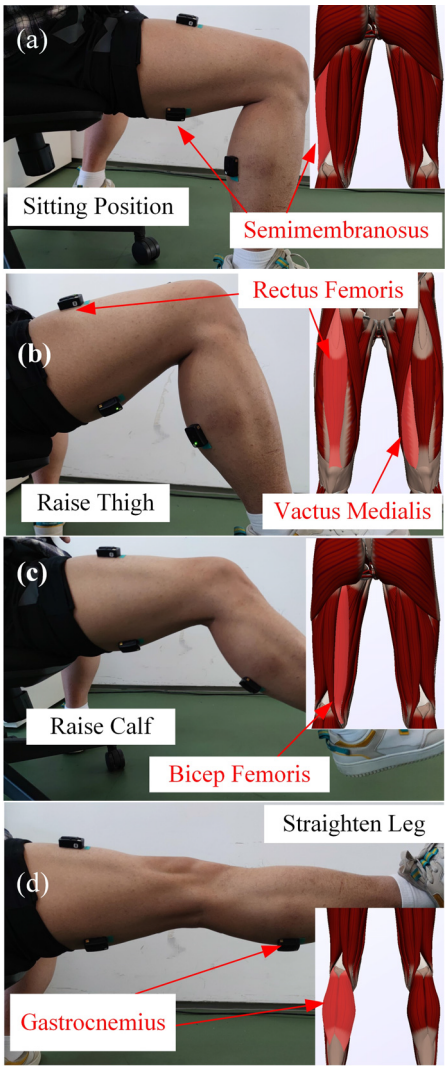

Fig. 2. Four movements of the lower limb and sEMG sensors pastes position and corresponding muscle. (a) Sitting position; (b) Raise thigh; (c) Raise calf; (d) Straighten leg.

C. Acquisition Hardware and Acquisition Protocol

We utilized a Delsys Trigno wireless sEMG System (Delsys Inc, USA) to record sEMG signals. The device has a 16-bit resolution with a bandwidth of 20-4450 HZ and the baseline noise is $<1.25\ \mu V$ (rms). The sampling frequency of this equipment is 2000 Hz. The device utilizes a Bluetooth communication protocol up to a 40 m communication distance. Prior to collection, all skin is wiped with medical alcohol and the sensor is affixed to the skin using double-sided adhesive tape.

The primary objective of this study is to investigate the generalization performance of classifier across subjects and time periods. Hence, we collect sEMG data of lower limb motions from diverse participants and distinct time points. The participants are instructed to perform the experiment thrice daily for five consecutive days. Each experiment consists of five sets of movements. Each set of movement contains four actions: ST, RT, RC, and SL. Each action is performed for a duration of 2 seconds followed by a rest interval of 1 minute to prevent muscle fatigue. Within 2 second of each action

TABLE II
THE CALCULATION OF FIVE FEATURES

Name	Formula
RMS	$RMS_i(k)=\sqrt{\frac{1}{N}\sum_{t=1}^{N}x_i^k(t)^2}$
WL	$WL_i(k)=\sum_{t=1}^{n}\left\|x_i^k(t+1)-x_i^k(t)\right\|$
MAV	$MAV_i(k)=\frac{1}{N}\sum_{t=1}^{N}\left\|x_i^k(t)\right\|$
VAR	$VAR_i(k)=\frac{1}{N-1}\sum_{t=1}^{N}(x_i^k(t)-\frac{1}{N}\sum_{t=1}^{N}x_i^k(t))^2$
MF	$MF_i(k)=\sum_{t=1}^{N}P_{ij}$

Here, x is the sMEG signal, i is the sEMG channel, k is the number of slide window, t is the sEMG points number in the k-th slide window, $P_{i,j}$ is the i-th channel sEMG power spectrum at frequency j.

duration, a total of 10 consecutive samples are collected. Each sample contains 200 sEMG data. Procedures are conducted after the Declaration of Helsinki.

D. sEMG Processing and Features

In the data preprocessing, sEMG are band-pass filtered by butter-worth band pass filter with forth order between 12 Hz and 500 Hz respectively low and high cutoff. Furthermore, a notch filter is used to remove the 50 Hz line interference [13].

sEMG cannot be directly used for training or testing because the data fluctuates seriously and contains abundant noise. After the pre-processing, we extract the sEMG features which are then inputted into the classifier for training or testing. Commonly utilized sEMG features encompass time domain, frequency domain and time-frequency domain features. Time domain features are widely employed due to their superior classification performance and low computational complexity [31]. The widely utilized features are root mean square (RMS), mean absolute value (MAV), zero crossing (ZC), waveform length (WL), slope scope change (SSC), variance (VAR), mean frequency (MF). In previous work [13], we extracted time domain descriptor (TDD) features robust to sensor position changes to overcome the shortage of generalizing the existing sEMG-based pattern recognition. Different combinations of features and classifiers exhibit varying levels of generalization performance. In this paper, we select six features to participate in the experiment and explore the optimal feature combination using the proposed classifier. The five selected features are $xikt P_{i,j}\, j$ shown in Table II, and the other TDD feature is referred to [13].

III. METHODOLOGY

A. Random Forest and Weighted Random Forest

1) Random Forest: The decision tree, as the base learner of RF, plays a pivotal role in RF. The decision tree can be considered as a tree-like model consisting of root nodes, leaf nodes, and intermediate nodes. The process of constructing the decision tree involves the splitting of the root node, traversing through multiple intermediate nodes, and ultimately reaching the leaf node. Among them, nodes are regarded as various features of the sample. The RF consists of multiple unpruned decision trees [32]. The advantages of RF, in comparison to other classifiers, are primarily attributed to the stochasticity of sample selection and feature selection. The RF initially constructs a predetermined number of decision tree classifiers. These classifiers are then utilized to vote on the test samples. Ultimately, the final decision is determined based on majority dominance. The RF has two crucial parameters: the number of trees (*n_estimators*) and the maximum number of selected features (*max_features*). The *n_estimators* parameter should be set sufficiently high and RF exhibits no overfitting tendencies even with an increasing number of trees [31]. Furthermore, the insensitivity of RF to the value of *max_features* allows us to set the *max_features* as the square root of the total number of features. The utilization of out-of-bag (OOB) samples, which are excluded from the construction of individual decision trees, plays a pivotal role in enhancing the performance of RF. The individual decision tree predicts the OOB samples using the following formula [33]:

$$p_k(x_i)=\frac{\sum_{j\in OOB_i} I[y(x_i,tree_j)=k]}{|OOB_i|}, \tag{1}$$

where x_i is the i-*th* observed sample in the OOB; k is the class (binary k = 0, 1); I is an indicator, taking a value of 1 when its parameters are true and 0 otherwise; y is the predicted class of the i-*th* decision tree to x_i, and OOB_i is the sets of decision trees. For a binary classification, the OOB prediction class is determined as 1 when $p(x_i) > 0.5$, and as 0 otherwise. The establishment of RF and the estimating of OOB probabilities can be simultaneously conducted. The main construction process of RF is outlined as follows:

Step 1 Preparing the datasets. The original training sets are randomly sampled with replacement to generate the new training sets.

Step 2 Building decision tree. The building of each decision tree involves the utilization of the new training sets, thereby ensuring randomness. During the process of building the decision tree, each node randomly selects features for splitting without pruning techniques until the specified stopping condition is satisfied.

Step 3 Building random forest. After generating a significant number of decision trees, a majority voting strategy is utilized to predict new samples by aggregating the outcomes of all trees.

2) Weighted Random Forest: The number of decision trees utilized in RF has a significant impact, as the decision trees cannot be modified once created. The decision trees exhibit uneven classification performances. The weight assigned to each decision tree in traditional RF is uniform, disregarding the distinct performance of individual trees. When the performance of individual decision tree classifier in RF is unequal, it is reasonable to assign more weights to the decision tree that ultimately yields accurate predictions. The weighting schemes

employed in the RF are not new [34]. In previous work [26], we employed a weighted voting approach to ascertain the weight of each decision tree. Each decision tree is assumed to be independent with an accuracy of $p_1, p_2, \cdots, p_L$, respectively, in predicting OOB. The formula for obtaining p_i is as follows:

$$p_i = 1 - \frac{1}{m}\sum_{j=1}^{m}(y_{\text{pre},j} - y_{rea,j})^2, \tag{2}$$

where $y_{\text{pre},j}$ denotes the predicted value of the i-*th* decision tree for the j-*th* OOB sample; $y_{rea,j}$ is the true value of the j-*th* OOB sample; m is the number of OOB samples.

Kuncheva and Rodriguez [35] compared the weighted voting method with alternative approaches, including majority vote, recall combiner, naive Bayes combiner. The results indicated that the weighted voting method exhibited superior performance. Inspired by Rodgyruez, we utilized a weighted voting method to determine the weights of each decision tree. The weight ω_i for each decision tree is obtained using the following formula:

$$\omega_i = \frac{p_i}{\sum_{i=1}^{n} p_i}. \tag{3}$$

The final prediction result $Out_{i,j}$ for each decision tree can be obtained by a weighted voting method utilizing the following formula,

$$Out_j = \arg\max_{j \in C} \sum_{i=1}^{L} \omega_i O_{i,j}, \tag{4}$$

where C is the number of classifications; Out_j represents the predicted value of the j-*th* sample; $O_{i,j}$ is the predicted value of the j-*th* sample of the i-*th* decision tree.

B. Stacked Weighted Random Forest

Stacking is an efficient framework that employs a meta-learning algorithm to effectively integrate the predictions of multiple base learners. Stacking generally consists of two layers, namely the base model layer and the meta-learning layer [36]. Specifically, stacking leverages the out-of-bag predictions of the base learner and actual labels from the training data to construct independent and dependent variables for the second level learning task, respectively [37]. The general procedure for stacking is as follows. First, the initial dataset is divided into training and testing datasets. Then, the base learner in the base model layer generates predictions for a subset of the training dataset. The predicted values are utilized as additional training data to construct a new training dataset. Finally, the new training dataset is utilized for training the algorithmic model within the meta-learning layer. Stacking is an ensemble method that aims to enhance accuracy by leveraging the optimal combination of base learners. The stacking method has been widely employed by numerous researchers to effectively integrate multiple models and successfully applied in various scenarios, yielding favorable outcomes [36], [38], [39].

Inspired by the stacking method, we further combine WRF with stacking to propose the SWRF model. The standard RF is utilized as the base learner, while WRF serves as the meta-leaning layer algorithm. The RF has a robust performance on its own. When employing the stacking, WRF assigns more weight to the more powerful RF model, thereby enhancing the overall performance of the model. The implementation of SWRF is as follows:

Step 1. A portion of the raw training data was employed for training and constructing initial RF models.

Step 2. K-fold cross-validation was utilized to obtain out-of-bag predictions for each decision tree.

Step 3. A new dataset was generated by utilizing out-of-bag predictions as input variables and actual labels as dependent variables.

Step 4. The WRF method was utilized to train the new dataset obtained in Step 3, resulting in SWRF.

C. Comparing Classifier and Performance Evaluation

The four classifiers were utilized for comparison with our proposed method, including support vector machine (SVM) [40], standard RF, WRF [26], long short-term memory (LSTM) [41]. The basic parameters of RF, WRF and SWRF are kept consistent to ensure a fair comparison. The basic parameters of RF, WRF and SWRF are *n_trees*=100, *max_depth*=5. The GridSearchCV was utilized to determine the optimal SVM parameters in order to ensure equitable comparisons. The parameters of SVM are $C = 1.0$, *kernel*='rbf'. Default values are selected for other unspecified parameters. The used contrastive LSTM was sourced from [41]. The LSTM approach contains a 2-layer CNN and 2-layer LSTM. The initial CNN contains 32 units and the subsequent layer contains 16 units. The initial LSTM layer contains 64 units and the subsequent layer contains 32 units. Following each CNN layer, subsequent layers of fully connection layer and batch normalization layer are sequentially interconnected. Refer to [41] for other parameters.

The machine learning-based classification system generates four possible outcomes, true negative (TN), true positive (TP), false negative (FN) and false positive (FP). TN and TP are the correct predictions for positive and negative classes. FN and FP are mispredictions for positive and negative classes. In this paper, accuracy (Acc), precision (Pre), recall (Re) and F1-score (F1) are utilized to evaluate the classification model [42]:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \tag{5}$$

$$Precision = \frac{TP}{TP + FP}, \tag{6}$$

$$Recall = \frac{TP}{TP + FN}, \tag{7}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall}. \tag{8}$$

IV. Experimental Results and Analysis

A. Experiments Design

This section details the design of evaluation experiments aimed at comprehensively assessing SWRF performance from

TABLE III
ACCURACY AND PRECISION RESULTS (%) OF DIFFERENT FEATURES COMBINED WITH DIFFERENT CLASSIFIER. **SCORES** IN BOLD INDICATES THE BEST

Multi-Day/ Classifier		Day 1		Day 2		Day 3		Day 4		Day 5		Average	
		Acc	Pre	Acc	Pre	Acc	Pre	Acc	Pre	Acc	Pre	Acc	Pre
SVM	RMS	**88.26**	**92.33**	83.09	89.38	52.14	63.40	50.40	60.50	67.56	54.48	68.29	72.02
	WL	85.11	89.46	88.53	92.84	50.00	57.91	52.01	59.79	66.03	54.15	68.34	70.83
	MAV	86.48	91.85	83.09	90.03	48.60	60.29	51.15	61.34	63.65	51.89	66.59	71.08
	TDD	85.46	91.74	**89.49**	**92.83**	**73.61**	**79.90**	71.70	79.06	**80.95**	**86.12**	80.24	**85.93**
	MF	80.66	82.45	87.36	89.64	59.85	62.55	**98.87**	**98.95**	76.40	77.46	**80.63**	82.21
RF	RMS	98.10	98.28	97.47	97.77	91.18	93.10	97.48	97.77	96.20	96.54	96.09	96.69
	WL	98.08	96.62	**97.98**	**98.17**	**97.60**	**97.78**	94.33	95.47	**97.48**	**97.64**	97.09	97.14
	MAV	98.36	98.48	97.21	97.55	92.92	93.58	96.71	97.02	95.20	95.86	96.08	96.50
	TDD	**98.62**	**98.77**	97.74	98.00	95.95	96.75	**98.22**	**98.37**	96.21	96.67	**97.35**	**97.71**
	MF	90.54	91.53	91.30	92.08	74.48	77.24	93.43	94.27	79.29	82.52	85.81	87.53
WRF	RMS	98.61	98.73	97.98	98.24	93.43	94.90	**98.35**	**98.47**	95.83	96.22	96.84	97.31
	WL	96.46	97.05	97.60	97.89	**97.61**	**97.78**	94.33	94.79	**97.85**	**97.98**	96.77	97.10
	MAV	98.36	98.47	97.60	97.89	92.92	93.82	97.09	97.42	95.83	96.25	96.36	96.77
	TDD	**98.74**	**98.86**	**98.36**	**98.58**	95.08	95.57	98.11	98.34	96.09	96.45	**97.28**	**97.56**
	MF	90.66	91.41	91.54	92.75	75.64	78.65	95.20	95.61	79.67	83.16	86.54	88.32
SWRF	RMS	98.86	98.98	97.86	98.12	95.83	95.15	97.99	98.16	95.59	96.30	97.23	97.34
	WL	96.96	97.34	97.98	98.25	**97.49**	**97.71**	96.20	96.52	96.08	96.83	96.94	97.33
	MAV	98.49	98.64	97.98	97.17	94.83	95.34	97.61	97.73	96.22	96.54	97.03	97.08
	TDD	**98.99**	**99.08**	**98.11**	**98.39**	96.85	97.12	**98.49**	**98.69**	**96.47**	**96.84**	**97.78**	**98.02**
	MF	91.29	92.21	90.66	92.32	76.27	78.28	94.70	95.40	80.17	83.12	86.62	88.27

TABLE IV
RECALL AND F1-SCORE RESULTS (%) OF DIFFERENT FEATURES COMBINED WITH DIFFERENT CLASSIFIER. **SCORES** IN BOLD INDICATES THE BEST

Multi-Day/ Classifier		Day 1		Day 2		Day 3		Day 4		Day 5		Average	
		Re	F1	Re	F1	Re	F1	Re	F1	Re	F1	Re	F1
SVM	RMS	**88.26**	**87.44**	83.09	80.92	52.14	53.03	50.40	49.22	67.56	59.15	68.29	65.95
	WL	85.11	83.75	88.53	87.95	50.00	50.65	52.01	50.49	64.69	58.20	68.07	66.21
	MAV	86.48	85.13	83.09	81.06	48.60	49.47	51.15	51.20	63.65	56.15	66.59	64.60
	TDD	85.46	83.89	**89.49**	**88.84**	**73.61**	**72.38**	71.70	69.76	**80.95**	**76.66**	80.24	78.31
	MF	80.66	80.13	87.36	87.15	59.85	57.13	**98.87**	**98.87**	76.40	75.83	**80.63**	**79.82**
RF	RMS	98.10	98.10	97.47	97.46	91.18	91.31	97.48	97.48	96.21	96.17	96.09	96.10
	WL	96.08	96.06	**97.98**	**97.96**	**97.60**	**97.59**	94.33	94.36	**97.48**	**97.46**	96.69	96.69
	MAV	98.36	98.36	97.21	97.19	92.92	92.93	96.71	97.71	95.20	95.19	96.08	96.28
	TDD	**98.62**	**98.61**	97.74	97.74	95.95	95.93	**98.22**	**98.23**	96.21	96.13	**97.35**	**97.33**
	MF	90.54	90.54	91.30	91.26	74.48	74.87	93.43	93.37	79.29	79.36	85.81	85.88
WRF	RMS	98.61	98.60	97.98	97.99	93.43	93.39	**98.35**	**98.35**	95.83	95.81	96.84	96.83
	WL	94.46	96.47	97.60	97.59	**97.61**	**97.61**	94.33	94.32	**97.85**	**97.84**	96.37	96.77
	MAV	98.36	98.36	97.60	97.58	92.92	92.98	97.09	97.12	95.83	95.84	96.36	96.38
	TDD	**98.74**	**98.75**	**98.36**	**98.36**	95.08	95.06	98.11	98.10	96.09	96.06	**97.28**	**97.27**
	MF	90.66	90.63	91.54	91.60	75.64	75.70	95.20	95.18	79.67	79.80	86.54	86.58
SWRF	RMS	98.86	98.87	97.86	97.83	95.83	95.79	97.99	97.99	95.59	95.57	97.23	97.21
	WL	96.96	96.98	97.98	97.98	**97.49**	**97.49**	96.20	96.22	96.08	96.08	96.94	96.95
	MAV	98.49	98.49	97.98	97.95	94.83	94.87	97.61	97.60	96.22	96.21	97.03	97.02
	TDD	**98.99**	**98.99**	**98.11**	**98.12**	96.85	96.86	**98.49**	**98.47**	**96.47**	**96.44**	**97.78**	**97.78**
	MF	91.29	91.24	90.66	90.80	76.27	76.22	94.70	94.72	80.17	80.26	86.62	86.65

multiple perspectives. The long-term usability and user adaptability of the SWRF are assessed through these experiments. A total of four experiments are designed. The experiment is repeated three times and the mean of the three trials is taken as the final result. Each result is calculated using five-fold cross validation.

1) Feature Generalizability Estimation: In this experiment, the generalizability of different features combinations with various classifier was investigated. We determined the optimal combination by investigating the classification performance of various classifiers combined with different features using multi-day data from a specific subject. The sEMG data of Subject 1 (S1) were randomly selected for this experiment. The model was trained using data obtained from S1 on the first day. The model was validated using data obtained from S1 on other time periods.

2) Offline Estimation on Inter-Subject: In this experiment, the generalizability of the classifier was investigated on the dataset. Data from the first subject were randomly selected to train the model. Data from other subjects was utilized to test the model. Results were shown as the mean of the five test results.

3) Multi-Day Estimation on a Subject: This experiment aimed to evaluate the long-term robustness of the classifier within the same subject. In this case, the training and testing data belonged to the same person at different times. The model was trained using data collected from a subject on the first day. Data from subsequent time points of the same subjects were utilized to validate the model.

4) Single-Day Estimation on Inter-Subject: This experiment aimed to assess the robustness of the classifier across subjects with in the same day. In this case, the training and testing data belonged to different individuals within the same day. The model was trained using data from a subject on a specific day. Data from other subjects in the same day were utilized to test the model.

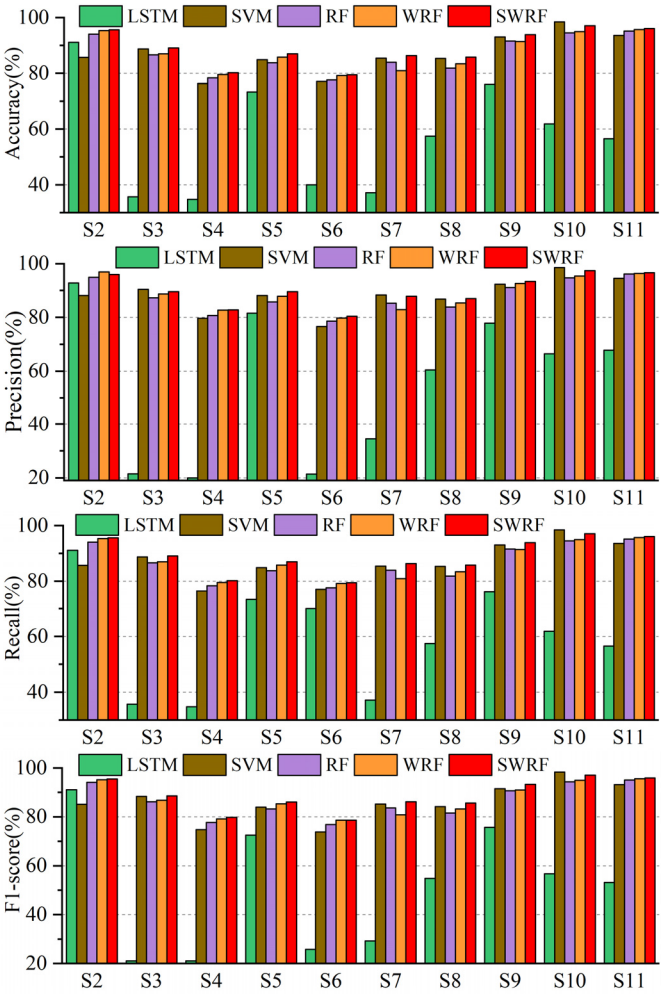

Fig. 3. The quantification calculation results of different classifiers on the dataset.

B. Results of Feature Generalizability Estimation

In this scenario, the sEMG data was obtained from the same individual at different time. Thus, variations in sEMG signals mainly arisen from muscle fatigue, electrode displacement, and peripheral muscle crosstalk within the same subject. Tables III and IV shows the quantitative results for different feature combinations with classifiers. The SVM classification accuracy exhibited a ranged of 48.6 % to 98.87 % over 5 days. The combination of SVM and MAV yielded the lowest accuracy on the third day. The combination of SVM and MF yielded the highest accuracy on the fourth day. The RF classification accuracy exhibited a ranged of 74.48 % to 98.62 % over 5 days. The combination of RF and MF yielded the lowest accuracy on the third day. The combination of RF and TDD yielded the highest accuracy on the first day. The WRF classification accuracy exhibited a ranged of 75.64 % to 98.74 % over 5 days. The combination of WRF and MF yielded the lowest accuracy on the third day. The combination of WRF and TDD yielded the highest accuracy on the first day. The SWRF classification accuracy exhibited a ranged of 76.27 % to 98.99 % over 5 days. The combination of SWRF and MF yielded the lowest accuracy on the third day. The combination of SWRF and TDD yielded the highest accuracy on the fourth day. It was evident that all four classifiers achieved the lowest accuracy on the third day, conversely, the remaining accuracies were significantly higher than those of the third day. The subject's poor adaptation and adjustment to the collection of sEMG data on the third day is a matter of speculation.

Furthermore, the highest mean result achieved over 5 days of SVM was for the combination of SVM and MF (Acc: 80.63 %; Pre: 82.21 %; Re: 80.63 %; F1: 79.82 %). The highest mean result achieved over 5 days of RF was for the combination of RF and TDD (Acc: 97.35 %; Pre: 97.71 %; Re: 97.35 %; F1: 97.33 %). The highest mean result achieved over 5 days of WRF was for the combination of WRF and TDD (Acc: 97.28 %; Pre: 97.56 %; Re: 97.28 %; F1: 97.27 %). The highest mean result achieved over 5 days of SWRF was for the combination of SWRF and TDD (Acc: 97.78 %; Pre: 98.02 %; Re: 97.78 %; F1: 97.78 %). As observed, TDD achieved the highest average classification results in RF, WRF, and SWRF, while ranking second in terms of average classification results in SVM. These results validated that the TDD had significant robustness to time variation. Thereby, the TDD feature was utilized for calculation in subsequent experiments.

Although differences of sEMG signals observed in this experiment cannot include all the influencing factors, the same individual exhibits a high level of representation across different time. Similar results were also obtained through

TABLE V

QUANTIFICATION CALCULATION RESULTS (%) OVER 5 DAYS. **SCORES** IN BOLD INDICATES THE BEST. S *M*-DAY *N* DENOTES THE EXPERIMENT FOR THE *M*-TH SUBJECT ON DAY *N*

Day/ Classifier	S1-Day 1				S1-Day 2				S1-Day 3				S1-Day 4				S1-Day 5			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	68.33	72.58	68.33	67.8	68.83	68.99	69.83	67.06	65.42	78.38	65.42	63.06	57.84	61.45	57.84	53.70
SVM	85.46	91.74	85.46	83.89	91.16	93.34	91.16	90.9	74.75	80.76	74.75	73.19	69.19	74.86	69.19	67.47	79.28	83.80	79.28	73.75
RF	98.62	98.77	98.62	98.61	97.61	97.86	97.61	97.61	92.90	94.26	92.90	92.85	96.98	97.43	96.98	96.93	95.19	95.90	95.19	95.10
WRF	98.74	98.86	98.74	98.75	98.11	98.23	98.11	98.10	93.44	94.11	93.44	93.38	97.85	98.10	97.85	97.83	95.84	96.40	95.84	95.81
SWRF	98.99	99.08	98.99	98.99	**98.48**	**98.61**	**98.48**	**98.48**	**95.33**	**96.05**	**95.33**	**95.32**	**98.86**	**98.95**	**98.86**	**98.86**	**96.22**	**96.57**	**96.22**	**96.20**

Day/ Classifier	S2-Day 1				S2-Day 2				S2-Day 3				S2-Day 4				S2-Day 5			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	98.38	95.68	93.38	93.51	57.04	58.29	57.04	55.43	75.08	82.03	75.08	73.99	55.67	60.13	55.67	53.39	55.35	57.50	55.35	51.60
SVM	97.35	97.74	97.35	97.37	75.02	76.80	75.02	74.21	74.00	71.93	74.00	67.15	70.95	60.65	70.95	62.65	99.12	99.27	99.12	99.13
RF	99.62	99.65	99.62	99.62	95.97	96.28	95.97	95.95	90.53	91.65	90.53	90.58	95.83	95.97	95.83	95.83	99.62	99.64	99.62	99.61
WRF	99.62	99.64	99.62	99.62	96.20	96.69	96.47	96.44	91.04	91.63	91.04	91.12	95.58	95.92	95.58	95.60	99.62	99.65	99.62	99.62
SWRF	99.62	99.65	99.62	99.62	**97.22**	**97.45**	**97.22**	**97.23**	**91.54**	**91.93**	**91.54**	**91.57**	**96.71**	**96.83**	**96.71**	**96.69**	99.62	99.65	99.62	99.62

Day/ Classifier	S3-Day 1				S3-Day 2				S3-Day 3				S3-Day 4				S3-Day 5			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	56.36	50.92	56.36	49.00	32.15	21.87	32.15	22.58	66.12	75.25	66.12	64.97	71.43	72.84	71.43	68.61
SVM	88.63	92.61	88.63	88.36	82.07	84.84	82.07	81.90	88.53	91.24	88.53	88.19	76.55	79.08	76.55	75.70	79.06	84.73	79.06	76.90
RF	98.49	98.70	98.49	98.48	98.48	98.64	98.48	98.49	97.60	97.67	97.60	97.59	97.22	97.47	97.22	97.17	97.85	98.02	97.85	97.85
WRF	98.86	99.04	98.86	98.85	99.12	99.17	99.12	99.12	97.86	97.98	97.86	97.85	98.36	98.49	98.36	98.36	97.98	98.14	97.98	97.99
SWRF	99.12	99.23	99.12	99.12	**99.62**	**99.65**	**99.62**	**99.62**	**98.11**	**98.33**	**98.11**	**98.12**	**98.61**	**98.79**	**98.61**	**98.62**	**98.23**	**98.39**	**98.23**	**98.24**

Day/ Classifier	S4-Day 1				S4-Day 2				S4-Day 3				S4-Day 4				S4-Day 5			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	74.52	75.64	74.52	73.10	43.82	45.25	43.82	38.92	54.18	54.88	54.18	48.73	57.69	66.11	57.69	55.46
SVM	98.24	98.43	98.24	98.24	84.08	89.11	84.08	82.47	84.11	89.95	84.11	84.01	97.86	98.19	97.86	97.88	79.04	85.30	79.04	75.89
RF	98.86	98.98	98.86	98.87	97.60	97.80	97.60	97.60	89.39	90.79	89.39	89.54	97.48	97.77	97.48	97.50	95.71	96.30	95.71	95.67
WRF	98.99	99.08	98.99	98.99	97.90	98.38	98.23	98.21	88.65	90.23	88.65	88.84	97.60	97.80	97.60	97.60	95.32	96.07	95.32	95.28
SWRF	99.37	99.41	99.37	99.37	**98.36**	**98.45**	**98.36**	**98.35**	**90.01**	**91.29**	**90.01**	**90.16**	**98.10**	**98.25**	**98.10**	**98.12**	**96.59**	**97.06**	**96.59**	**96.59**

combined feature and classifier experiments conducted on different individuals within the same day.

C. Results of Offline Estimation on Inter-Subject

In this scenario, the sEMG data was obtained from various individuals within the dataset. Thus, variations in sEMG signals mainly arisen from individual variability. The training data was selected from S1, while the testing data was chosen from other subjects. All calculations utilize the TDD feature. Fig. 3 shows the quantification calculation results of different classifiers on the dataset. It can be observed that the robustness of LSTM is poor and the classification accuracy fluctuates more obviously. The LSTM exhibited the lowest classification accuracy of 34.78 % in S4. The classification accuracy of SVM exhibited substantial fluctuations. The SVM exhibited the highest classification accuracy of 98.46 % among all algorithms at S10 and the lowest classification accuracy of 76.39 % at S4. The RF, WRF and SWRF have relatively stable classification accuracy. The RF achieved the lowest classification accuracy of 77.64 % in S6. The WRF achieved the lowest classification accuracy of 79.22 % in S6. The SWRF achieved the lowest classification accuracy of 79.51 % in S6. Due to inter-individual variability, the application of sEMG-PRS to a large population is challenging, as evidenced by this offline experiment. S4 and S6 had poor classification accuracy, illustrating that S4 and S6 were significantly different from S1.

Furthermore, the average result achieved by LSTM in all subjects were Acc: 56.40 %, Pre: 54.43 %, Re:59.39 %, F1: 50.09 %. The average result achieved by SVM in all subjects were Acc: 86.89 %, Pre: 88.34 %, Re:86.89 %, F1: 85.84 %. The average result achieved by RF in all subjects were Acc: 86.75 %, Pre: 87.82 %, Re:86.75 %, F1: 86.32 %. The average result achieved by WRF in all subjects were Acc: 87.35 %, Pre: 88.84 %, Re:87.35 %, F1: 87.06 %. The average result achieved by SWRF in all subjects were Acc: 89.06 %, Pre: 90.05 %, Re:89.06 %, F1: 88.65 %. As observed, SWRF attains the highest mean classification accuracy among all algorithms, reaching 89.06 %.

D. Results of Multi-Day Estimation on a Subject

Table V shows the quantification calculation results over 5 days for different subjects, respectively. All five subjects utilized the data from the first day to train the model, while data from the subsequent four days were employed for testing. *For Subject 1).* SWRF achieved the highest average result among all algorithms (Acc: 97.58 %; Pre: 97.85 %; Re: 97.58 %; F1: 97.57 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.78 %, 1.32 %, 17.61% and 24.49 %, respectively. *For Subject 2).* SWRF achieved the highest average result among all algorithms (Acc: 96.94 %; Pre: 97.10 %; Re: 96.94 %; F1: 96.95 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.53 %, 0.63 %, 13.65 % and 28.64 %, respectively. *For Subject 3).* SWRF achieved the highest average result among all algorithms (Acc: 98.74 %; Pre: 98.88 %; Re: 98.74 %; F1: 98.74 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.30%, 0.81 %, 15.77 % and 33.53 %, respectively. *For Subject 4).* SWRF achieved the highest average result among all algorithms (Acc: 96.49 %; Pre: 96.89 %; Re: 96.49 %; F1: 96.52 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.79 %, 0.68 %, 7.82 % and 30.44 %, respectively.

From the results, SWRF achieved the highest mean classification accuracy across all subjects over 5 days and exhibited significant long-term usability. Compared with WRF, RF, SVM and LSTM, SWRF improved classification accuracy by an average of 0.60 %, 0.86 %, 13.71 %, 29.28 % across all subjects, respectively.

E. Results of Single-Day Estimation on Inter-Subject

Table VI shows the quantification calculation results on inter-subject for a single day. The S1 data were utilized to train the model and other subject data were utilized for testing. *For Day 1).* SWRF achieved the highest average result among all algorithms (Acc: 99.21 %; Pre: 99.31 %; Re: 99.21 %; F1: 99.21 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.19 %, 0.51 %, 6.22 % and 42.33 %, respectively. *For Day 2).* SWRF achieved the highest average result among all algorithms (Acc: 98.17 %; Pre: 98.34 %; Re:98.17 %; F1: 98.17 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.16%, 0.38 %, 15.89 % and 29.02 %, respectively. *For Day 3).* SWRF achieved the highest average result among all algorithms (Acc: 93.88 %; Pre: 94.47 %; Re:93.88 %; F1: 93.91 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.54 %, 1.69 %, 12.48% and 48.17 %, respectively. *For Day 4).* SWRF achieved the highest average result among all algorithms (Acc: 97.79 %; Pre: 97.99 %; Re:97.79 %; F1: 97.80 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.35 %, 1.02 %, 19.24% and 29.23 %, respectively. *For Day 5).* SWRF achieved the highest average result among all algorithms (Acc: 97.73 %; Pre: 97.89 %; Re:97.73 %; F1: 97.72 %.). Compared to WRF, RF, SVM and LSTM, the average accuracy rate was increased by 0.54 %, 0.91 %, 13.23% and 28.95 %, respectively.

From the results, SWRF achieved the highest average recognition accuracy across all days among subjects and exhibited significant user adaptability. Compared to WRF, RF, SVM and LSTM, SWRF improved classification accuracy by an average of 0.36 %, 0.90 %, 13.41 %, 35.54 % across all days, respectively.

F. Further Analysis on SWRF Algorithm

The analysis of model allows us to gain valuable insights into its areas of expertise and challenges, thereby facilitating future enhancements in its performance. Fig. 10 shows the confusion matrix plots of the five algorithms. The model was trained using the data from the first day of subject 1, and all algorithm were tested with the data from the fifth day of subject1. The average accuracy of all models, as indicated by the confusion matrix result plot, closely aligned with the data presented in Table V. The SWRF achieved

TABLE VI
QUANTITATIVE RESULTS (%) OF INTER-SUBJECT FOR A SINGLE DAY. **SCORES** IN BOLD INDICATES THE BEST.S *M*-DAY *N* DENOTES THE EXPERIMENT FOR THE *M*-TH SUBJECT ON DAY *N*

Multi-Day/ Classifier	S1-Day 1				S2-Day 1				S3-Day 1				S4-Day 1			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	52.55	51.85	52.55	47.06	43.95	47.61	43.95	42.79	31.04	17.38	31.04	19.30
SVM	85.47	91.82	85.47	83.87	97.72	98.02	97.72	97.74	91.30	94.06	91.30	91.07	97.48	97.88	97.48	97.48
RF	98.61	98.66	98.61	98.60	**99.62**	99.64	**99.62**	**99.62**	97.85	98.29	97.85	97.85	98.74	98.95	98.74	98.77
WRF	98.60	98.73	98.60	98.60	**99.62**	**99.65**	**99.62**	**99.62**	98.61	98.83	98.61	98.61	99.24	99.29	99.24	99.24
SWRF	98.98	99.08	98.98	98.98	**99.62**	**99.65**	**99.62**	**99.62**	**98.99**	**99.19**	**98.99**	**98.99**	**99.25**	**99.30**	**99.25**	**99.24**

Multi-Day/ Classifier	S1-Day 2				S2-Day 2				S3-Day 2				S4-Day 2			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	56.95	51.04	56.95	49.92	65.90	64.58	65.90	60.83	53.73	55.22	53.73	50.06
SVM	87.86	92.27	87.86	87.00	75.35	76.70	75.35	74.85	82.21	84.94	82.21	81.96	83.69	88.89	83.69	82.29
RF	97.86	98.10	97.86	97.85	95.46	95.72	95.46	95.44	98.99	99.08	98.99	98.98	**98.85**	98.03	97.85	97.85
WRF	98.36	98.65	98.36	98.39	96.09	96.26	96.09	96.08	**99.37**	99.41	**99.37**	**99.37**	98.22	98.36	98.22	98.22
SWRF	98.36	98.49	98.36	98.36	**96.71**	**97.03**	**96.71**	**96.71**	**99.37**	**99.44**	**99.37**	**99.37**	98.23	**98.41**	**98.23**	**98.23**

Multi-Day/ Classifier	S1-Day 3				S2-Day 3				S3-Day 3				S4-Day 3			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	31.08	19.33	31.08	21.42	26.80	11.94	26.80	15.00	24.98	6.17	24.98	10.99
SVM	74.24	78.85	74.24	72.72	76.51	75.14	76.51	71.96	90.03	91.92	90.03	89.64	84.84	90.51	84.84	84.84
RF	92.94	94.00	92.94	92.96	89.76	91.27	89.76	89.74	97.09	97.24	97.09	97.10	89.00	90.96	89.00	89.28
WRF	95.21	95.69	95.21	95.21	91.16	91.81	91.16	91.16	97.85	**98.07**	97.85	**97.86**	89.14	90.68	89.14	89.30
SWRF	95.70	96.20	95.70	95.72	**92.18**	**92.79**	**92.18**	**92.18**	**97.86**	97.97	**97.86**	97.85	**89.79**	**90.93**	**89.79**	**89.87**

Multi-Day/ Classifier	S1-Day 4				S2-Day 4				S3-Day 4				S4-Day 4			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**100**	**100**	**100**	**100**	65.13	65.88	65.13	60.51	52.15	52.51	52.15	45.46	56.95	51.86	56.95	51.12
SVM	68.79	74.40	68.79	66.94	71.81	60.93	71.83	64.04	75.77	77.43	75.77	74.60	97.85	98.14	97.85	97.87
RF	96.83	97.47	96.83	96.88	95.31	95.76	95.31	95.29	97.46	97.80	97.46	97.48	97.48	97.68	97.48	97.49
WRF	97.73	97.96	97.73	97.73	95.82	96.26	95.82	95.85	98.10	98.24	98.10	98.10	98.10	98.28	98.10	98.11
SWRF	97.86	98.06	97.86	97.87	**96.58**	**96.90**	**96.58**	**96.60**	**98.48**	**98.60**	**98.48**	**98.48**	**98.24**	**98.39**	**98.24**	**98.25**

Multi-Day/ Classifier	S1-Day 5				S2-Day 5				S3-Day 5				S4-Day 5			
	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1	Acc	Pre	Re	F1
LSTM	**99.51**	**99.56**	**99.51**	**99.50**	60.33	69.84	60.33	56.41	64.88	66.18	64.88	60.87	50.38	53.68	50.38	46.15
SVM	79.17	79.66	79.17	73.71	**99.75**	**99.77**	**99.75**	**99.75**	80.31	87.02	80.31	78.20	78.78	82.89	78.78	75.53
RF	95.34	95.91	95.34	95.28	99.62	99.66	99.62	99.62	96.72	97.19	96.72	96.75	95.58	95.99	95.58	95.59
WRF	95.84	96.55	95.84	95.82	99.62	99.65	99.62	99.62	97.60	97.82	97.60	97.59	95.70	96.28	95.70	95.68
SWRF	96.47	96.65	96.47	96.44	99.62	99.64	99.62	99.62	**97.98**	**98.11**	**97.98**	**97.98**	**96.84**	**97.14**	**96.84**	**96.84**

outstanding classification performance on various actions, yet exhibits subpar performance specifically on RC actions. Other algorithms had achieved similar results. It indicated that RC actions were easily confused and difficult to distinguish. The RC was the calf raising action, which was similar to the SL. The action became SL when the angle of calf lift was sufficiently large. We observed that the calf lift angle exhibited considerable variability during the execution of the RC action. The angle may occasionally exceed the optimal range, while at other times it may fall short. Subsequently, the accuracy of the model in recognizing RC action can be enhanced by incorporating additional means to differentiate this action, such as integrating more sensors or detecting muscles that are more representative.

The performance of the SWRF model is also influenced by the training set. Theoretically, the model's generalization can be improved by incorporating a diverse range of training sets and increasing the volume of data. Table VII shows the test comparison results of SVM and SWRF in the case of increasing the training set of days. The training and testing data were both sourced from subject 1. The data obtained from the fifth day was utilized for testing. The training data

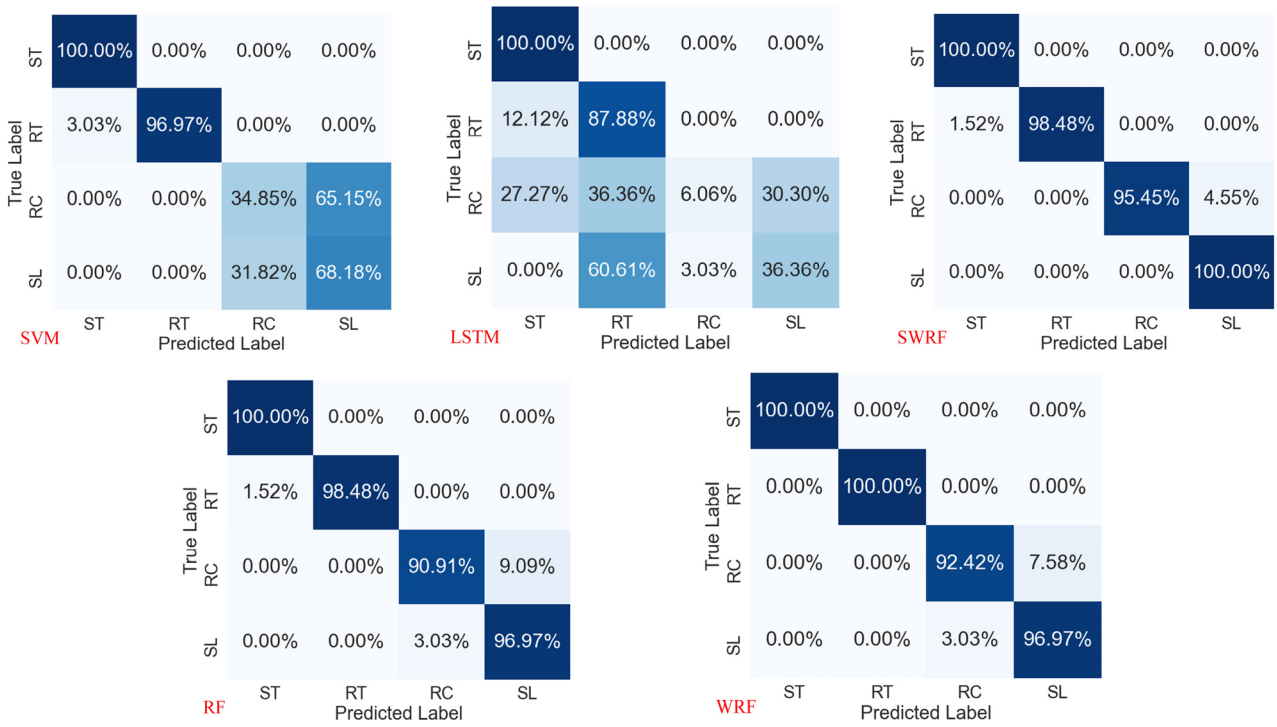

Fig. 4. Confusion matrices for the five algorithms.

TABLE VII
CLASSIFICATION PERFORMANCE WITH VARIED TRAINING SETS (TS) ON THE SAME TEST SET

TS	D1		D1+D2		D3+D4		D1+D2+D3	
Index	Acc	Pre	Acc	Pre	Acc	Pre	Acc	Pre
SVM	79.28	83.80	81.47	82.89	82.01	84.76	84.11	89.26
SWRF	96.22	96.57	96.51	97.02	97.34	97.58	98.10	98.29

TABLE VIII
CLASSIFICATION PERFORMANCE WITH VARIED TRAINING SETS (TS) ON THE SAME TEST SET

TS	S1		S1+S2		S2+S3		S1+S2+S3	
Index	Acc	Pre	Acc	Pre	Acc	Pre	Acc	Pre
SVM	84.84	90.51	85.36	91.69	86.02	91.84	87.81	92.20
SWRF	89.79	90.93	91.19	91.67	90.26	90.89	92.31	92.83

consisted of day 1 (D1), day 1 and day 2 (D1+D2), day 3 and day 4 (D3+D4), as well as the combination of day 1, day 2, and day 3 (D1+D2+D3). The model's performance was observed to improve with an increase in the richness of the training set, as shown in Table VII. Table VIII shows the test comparison results of SVM and SWRF in the case of increasing the training set of subjects. The training and testing data were both sourced from day 3. The data obtained from the subject 4 was utilized for testing. The training data consisted of subject 1 (S1), subject 1 and subject 2 (S1+S2), subject 2 and subject 3 (S2+S3), as well as the combination of subject 1, subject 2, and subject 3 (S1+S2+S3). The model's performance was observed to improve as the training set became more enriched, as shown in Table VIII. Subsequently, the model performance can be improved by enhancing the diversity and quality of the training sets.

V. CONCLUSION

The SWRF algorithm is proposed in this paper to enhance the long-term usability and user adaptability of sEMG-PRS. The standard RF is utilized as the base learner, while WRF serves as the meta-leaning layer algorithm. The offline datasets are utilized to evaluate the SWRF. The offline experimental results demonstrate that the SWRF outperforms classical classification algorithms, exhibiting an average classification accuracy of 89.06 % (LSTM: 56.40%; SVM: 86.89%; RF: 86.75%; WRF: 87.35%). Moreover, two online experiments are conducted to evaluate the performance of SWRF: single-day on inter-subject and multi-day on a subject. The results of single-day on inter-subject online experiment indicate that the SWRF achieves the highest mean classification accuracy across all subjects over 5 days and exhibits significant long-term usability. Compared with WRF, RF, SVM and LSTM, SWRF improves classification accuracy by an average of 0.60 %, 0.86 %, 13.71 %, 29.28 % across all subjects, respectively. The results of multi-day on a subject online experiment indicate that the SWRF achieves the highest average recognition accuracy across all days among subjects and exhibits significant user adaptability. Compared to WRF, RF, SVM and LSTM, SWRF improves classification accuracy by an average of 0.36 %, 0.90 %, 13.41 %, 35.54 % across all days, respectively.

The RF demonstrates strong performance on its own. When employing the stacking, WRF assigns more weight to the more powerful RF model, thereby enhancing the overall performance of the algorithm. By conducting an analysis of the SWRF model, we can implement optimization measures to enhance its weak feature recognition capability, thereby further optimize the SWRF algorithm. Additionally, the performance

of the SWRF can be further improved by enhancing the diversity and quality of the training sets.

REFERENCES

[1] R. R. Torrealba and E. D. Fonseca-Rojas, "Toward the development of knee prostheses: Review of current active devices," *Appl. Mech. Rev.*, vol. 71, no. 3, May 2019, Art. no. 030801.

[2] S. Jezernik, G. Colombo, and M. Morari, "Automatic gait-pattern adaptation algorithms for rehabilitation with a 4-DOF robotic orthosis," *IEEE Trans. Robot. Autom.*, vol. 20, no. 3, pp. 574–582, Jun. 2004.

[3] Y. Du, H. Wang, S. Qiu, W. Yao, P. Xie, and X. Chen, "An advanced adaptive control of lower limb rehabilitation robot," *Frontiers Robot. AI*, vol. 5, no. 5, pp. 1510–1520, Oct. 2018.

[4] C. Shen et al., "STMI: Stiffness estimation method based on sEMG-driven model for elbow joint," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, 2023.

[5] C. Mokri, M. Bamdad, and V. Abolghasemi, "Muscle force estimation from lower limb EMG signals using novel optimised machine learning techniques," *Med. Biol. Eng. Comput.*, vol. 60, no. 3, pp. 683–699, Mar. 2022.

[6] D. Farina et al., "The extraction of neural information from the surface EMG for the control of upper-limb prostheses: Emerging avenues and challenges," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 797–809, Jul. 2014.

[7] A. Vijayvargiya, Khimraj, R. Kumar, and N. Dey, "Voting-based 1D CNN model for human lower limb activity recognition using sEMG signal," *Phys. Eng. Sci. Med.*, vol. 44, no. 4, pp. 1297–1309, Dec. 2021.

[8] M. S. AL-Quraishi et al., "Classification of ankle joint movements based on surface electromyography signals for rehabilitation robot applications," *Med. Biol. Eng. Comput.*, vol. 55, no. 5, pp. 747–758, May 2017.

[9] J. Liu, X. Sheng, D. Zhang, J. He, and X. Zhu, "Reduced daily recalibration of myoelectric prosthesis classifiers based on domain adaptation," *IEEE J. Biomed. Health Informat.*, vol. 20, no. 1, pp. 166–176, Jan. 2016.

[10] J. Liu, X. Sheng, D. Zhang, N. Jiang, and X. Zhu, "Towards zero retraining for myoelectric control based on common model component analysis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 4, pp. 444–454, Apr. 2016.

[11] J. He, D. Zhang, N. Jiang, X. Sheng, D. Farina, and X. Zhu, "User adaptation in long-term, open-loop myoelectric training: Implications for EMG pattern recognition in prosthesis control," *J. Neural Eng.*, vol. 12, no. 4, Aug. 2015, Art. no. 046005.

[12] X. Jiang et al., "Optimization of HD-sEMG-based cross-day hand gesture classification by optimal feature extraction and data augmentation," *IEEE Trans. Human-Mach. Syst.*, vol. 52, no. 6, pp. 1281–1291, Dec. 2022.

[13] C. Shen, Z. Pei, W. Chen, J. Wang, J. Zhang, and Z. Chen, "Toward generalization of sEMG-based pattern recognition: A novel feature extraction for gesture recognition," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022.

[14] U. Côté-Allard et al., "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.

[15] D. Zhang, X. Zhao, J. Han, and Y. Zhao, "A comparative study on PCA and LDA based EMG pattern recognition for anthropomorphic robotic hand," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Hong Kong, May 2014, pp. 4850–4855.

[16] A. Tabor, S. Bateman, and E. Scheme, "Evaluation of myoelectric control learning using multi-session game-based training," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 9, pp. 1680–1689, Sep. 2018.

[17] A. Waris, I. Mendez, K. Englehart, W. Jensen, and E. N. Kamavuako, "On the robustness of real-time myoelectric control investigations: A multiday Fitts' law approach," *J. Neural Eng.*, vol. 16, no. 2, Apr. 2019, Art. no. 026003.

[18] A. Ameri, M. A. Akhaee, E. Scheme, and K. Englehart, "A deep transfer learning approach to reducing the effect of electrode shift in EMG pattern recognition-based control," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 370–379, Feb. 2020.

[19] G. Wilson and D. J. Cook, "A survey of unsupervised deep domain adaptation," *ACM Trans. Intell. Syst. Technol.*, vol. 11, no. 5, pp. 1–46, Oct. 2020.

[20] X. Zhang, L. Wu, X. Zhang, X. Chen, C. Li, and X. Chen, "Multi-source domain generalization and adaptation toward cross-subject myoelectric pattern recognition," *J. Neural Eng.*, vol. 20, no. 1, Feb. 2023, Art. no. 016050.

[21] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, Jul. 2017.

[22] P. Shi, X. Zhang, W. Li, and H. Yu, "Improving the robustness and adaptability of sEMG-based pattern recognition using deep domain adaptation," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 11, pp. 5450–5460, Nov. 2022.

[23] E. Campbell, A. Phinyomark, and E. Scheme, "Deep cross-user models reduce the training burden in myoelectric control," *Frontiers Neurosci.*, vol. 15, May 2021, Art. no. 657958.

[24] H. Xu and A. Xiong, "Advances and disturbances in sEMG-based intentions and movements recognition: A review," *IEEE Sensors J.*, vol. 21, no. 12, pp. 13019–13028, Jun. 2021.

[25] Y. Gu, D. Yang, Q. Huang, W. Yang, and H. Liu, "Robust EMG pattern recognition in the presence of confounding factors: Features, classifiers and adaptive learning," *Expert Syst. Appl.*, vol. 96, pp. 208–217, Apr. 2018.

[26] C. Shen et al., "Lower limb activity recognition using sEMG signals via weighted random forest," in *Proc. IEEE 17th Conf. Ind. Electron. Appl. (ICIEA)*, Dec. 2022, pp. 1151–1156.

[27] O. Sanchez and S. Jose, "EMG dataset in lower limb," UCI Mac-hine Learn. Repository, Tech. Rep., 2014, doi: 10.24432/C5ZW3P.

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

[29] G. A. Lichtwark, K. Bougoulias, and A. M. Wilson, "Muscle fascicle and series elastic element length changes along the length of the human gastrocnemius during walking and running," *J. Biomechanics*, vol. 40, no. 1, pp. 157–164, Jan. 2007.

[30] K. Omuro, Y. Shiba, S. Obuchi, and N. Takahira, "Effect of ankle weights on EMG activity of the semitendinosus and knee stability during walking by community-dwelling elderly," *Rigakuryoho Kagaku*, vol. 26, no. 1, pp. 55–59, 2011.

[31] A. Phinyomark, F. Quaine, S. Charbonnier, C. Serviere, F. Tarpin-Bernard, and Y. Laurillau, "EMG feature evaluation for improving myoelectric pattern recognition robustness," *Expert Syst. Appl.*, vol. 40, no. 12, pp. 4832–4840, Sep. 2013.

[32] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[33] H. Deng, G. Runger, and E. Tuv, "System monitoring with real-time contrasts," *J. Quality Technol.*, vol. 44, no. 1, pp. 9–27, Jan. 2012.

[34] L. V. Utkin, A. V. Konstantinov, V. S. Chukanov, M. V. Kots, M. A. Ryabinin, and A. A. Meldo, "A weighted random survival forest," *Knowl.-Based Syst.*, vol. 177, pp. 136–144, Aug. 2019.

[35] L. I. Kuncheva and J. J. Rodríguez, "A weighted voting framework for classifiers ensembles," *Knowl. Inf. Syst.*, vol. 38, no. 2, pp. 259–275, Feb. 2014.

[36] S. K. Kalagotla, S. V. Gangashetty, and K. Giridhar, "A novel stacking technique for prediction of diabetes," *Comput. Biol. Med.*, vol. 135, Aug. 2021, Art. no. 104554.

[37] D. H. Wolpert, "Stacked generalization," *Neural Netw.*, vol. 5, no. 2, pp. 241–259, Jan. 1992.

[38] G. S. Ghalejoogh, H. M. Kordy, and F. Ebrahimi, "A hierarchical structure based on stacking approach for skin lesion classification," *Expert Syst. Appl.*, vol. 145, May 2020, Art. no. 113127.

[39] W. Jiang, Z. Chen, Y. Xiang, D. Shao, L. Ma, and J. Zhang, "SSEM: A novel self-adaptive stacking ensemble model for classification," *IEEE Access*, vol. 7, pp. 120337–120349, 2019.

[40] X. Yun, X. Ling, G. Lei, L. Zhanhao, and S. Bohan, "An improved SVM method for movement recognition of lower limbs by MIMU and sEMG," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 1, pp. 120–124, 2023.

[41] 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. 2021.

[42] J. Tu, Z. Dai, X. Zhao, and Z. Huang, "Lower limb motion recognition based on surface electromyography," *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104443.