Comparing the Outcomes of Population-averaged and Personalised Input Feature Selection for Transhumeral Prosthetic Interfaces

Tianshi Yu, Ricardo Garcia-Rosas, Alireza Mohammadi, Ying Tan, Peter Choong and Denny Oetomo

***Abstract*— An interface for a powered upper limb prosthesis is designed to achieve the target poses intended by the human user as interpreted from input features measured from a variety of sensors typically worn on the residual limb. Selection of which input features to use in the interface is usually done based on which features provide the best likelihood to identify the intended target poses. This work investigates the outcomes of the prosthetic interfaces constructed with input features selected based on the performance of all subjects (population-averaged) vs. the performance of individual subjects (personalised). Specifically, two outcomes are evaluated. First, the diversity in the resulting input feature sets across different subjects is evaluated. Second, the accuracy in identifying the intended target pose using the input features selected through personalised and population-averaged data are analysed and compared. The experiment was conducted with 10 able-bodied subjects in a virtual reality (VR) platform. The corresponding kinematic information of their arm movements and muscle activities were recorded and utilised to construct the input features in the scenario of transhumeral prostheses. The outcomes show a significant variation in the resulting input feature sets from subject to subject. The accuracy of the personalised feature sets is found to be higher in classifying the intended target poses. For the case of selecting 2 out of the 41 input features constructed from the recorded sensor data, none of the 10 subjects shares any common set of input features and the accuracy is improved by $10.4 \pm 6.8\%$ with the personalised selection. The outcomes demonstrate the extent of person-to-person variation that needs to be taken into account when designing prosthetic interfaces and the quantified potential gain (albeit in a specific application) from personalising interfaces.**

I. Introduction

Recent studies and development in upper limb powered prostheses seek to provide coordinated movements in the execution of tasks and activities of daily living [1]–[4]. The designed prosthetic interface takes as its input the characteristics of sensor signals measured from the residual limb (termed input features), which reflect the human user intention. The interface then regulates the motors in the prosthesis to realise a coordinated motion across the multiple joints of the prosthesis, which improves the efficiency of completing tasks [2], [4], [5].

To achieve the desired functionality of the prostheses, the input features need to be constructed such that the interface is able to accurately identify the prosthetic target poses intended by the human user [6]. Some input features naturally perform worse than the others for a given set of intended target pose due to the information content they contain. For example, in the case of a transhumeral prosthesis, the shoulder flexion/extension movements generally cannot tell us what a person intends to do with the elbow [7], [8]. It is therefore important to evaluate the information contained by the input features and select those with the most useful information to identify the intended target poses. Conventionally, input features based on surface electromyography (sEMG) were used heavily in the prosthetic interfaces. More recently, the observations of upper extremity inter-joint coordination [9], [10] prompted more kinematic features from the residual limb movement to be chosen [2], [4], [11]–[13], such as shoulder and upper body movements. The most recent studies combined both modalities to obtain better performance [14]–[16]. In all these studies, the input features for the interfaces were designed based on the overall subject data and movement behaviour. From the practical point of view, such population-averaged input features can be preferred for wider implementation. However, how much difference actually exists from person to person and how much effect such variation has on the performance of the interface is yet to be investigated. Some recent work searched for the optimum sensor set among the population (mostly sEMG sensors) [17], [18] or for individual [19], [20]. Subject-to-subject variations in the resulting IMU-EMG electrode positions were observed and reported in [20], which further highlights the need to understand the extent and the effect of such variations.

In this paper, the person-to-person variations of the input feature sets that best identify the required set of intended target poses are investigated. This is especially important if we seek to utilise a low number of input features (corresponding to a low number of sensors to be mounted on the prosthesis to obtain the human intention), while still achieving an acceptable accuracy in identifying the intended target poses. Specifically, two outcomes are sought. First, the extent of the diversity in the resulting input features selected when optimised for each subject (personalised) is evaluated. A comparison is made to the resulting input features if the interface had been designed based on the data from all the subjects (population-averaged). Second, the accuracy of the prosthetic interface in identifying the intended target pose is evaluated for the personalised input features for each subject, and compared to the accuracy achieved when the population-averaged input feature set is used. In this paper, no constraint was made to the modality of the input features, as long as the information obtained constitute an input feature. As such, both kinematics and sEMG input features were used in the example. The study in this paper was conducted on a

This project is funded by the Valma Angliss Trust and The University of Melbourne. T. Yu, R. Garcia-Rosas, D. Oetomo, and Y. Tan are with the School of Electrical, Mechanical and Infrastructure Engineering, and P. Choong with the Department of Surgery, The University of Melbourne, VIC 3010, Australia. {tianshiy}@student.unimelb.edu.au; {garcia.r,alireza.mohammadi,yingt,pchoong,doetomo}@unimelb.edu.au.

transhumeral prosthetic interface scenario.

The outcomes of this paper not only inform the extent of the person-to-person variations in the design of a prosthetic interface, but also quantify the potential performance improvement that can be obtained by using a personalised interface compared to that designed using population average data/behaviours.

II. Input Feature Selection

The input feature selection algorithm is constructed based on the concept of class separability which selects n_d features from a feature pool of total D candidate features. To provide context, note that 41 input features were constructed out of the sensors used in this study, listed in Tab. I. They consisted of 6 kinematics-based input features (A1 to C2) and 35 surface electromyography (sEMG) based input features (D1 to J5). The intended target poses are shown in Fig. 1, showing 9 possible target poses, made up of 3 discrete shoulder flexion angles and 3 elbow flexion angles. However, as a transhumeral scenario was being emulated, the target poses of interest were those of the elbow, thus there were only 3 target poses (T1 - T3). The subsection below explains how to evaluate the separability of the target poses from observing the available input features.

A. Target pose separability

Scatter-matrix based class separability measure [21], [22] is used in this work for its simplicity and generality. Denote F as the set containing the overall D candidate features, $F = \{f_1, ..., f_D\}$. The overall intended target pose set is denoted as P with C target classes with $P = \{T_1, ..., T_C\}$.

With the scatter matrices, the target pose separability provided by any feature set $F_s \subseteq F$ for any class set of $P_s \subseteq P$ can be calculated. Assume the sampled data with n samples are arranged in a matrix $\boldsymbol{X} = [\boldsymbol{x}_1, ..., \boldsymbol{x}_n] \in \mathbb{R}^{d\times n}$, where d is the number of the features that are of interest to evaluate and $d \leq D$. Each sample $\boldsymbol{x}_i$ $(i = 1, ..., n)$ is assigned to a label with c classes in an intended target subset P_s. Denote n_j as the number of samples in the j^{th} class, i.e., $\sum_{j=1}^{c} n_j = n$. By the relabeling of n samples, denote $\boldsymbol{x}_{j,k}$ $(j = 1, \ldots, c,\ k = 1, \ldots, n_j)$ to represent the k^{th} samples of the j^{th} class. The inter-class and intra-class scatter matrix $\boldsymbol{S}_B$ and $\boldsymbol{S}_W$ are

$$\boldsymbol{S}_B = \sum_{j=1}^{c} n_j(\boldsymbol{m}_j - \boldsymbol{m})(\boldsymbol{m}_j - \boldsymbol{m})^T, \tag{1}$$

$$\boldsymbol{S}_W = \sum_{j=1}^{c}\sum_{k=1}^{n_j}(\boldsymbol{x}_{j,k} - \boldsymbol{m}_j)(\boldsymbol{x}_{j,k} - \boldsymbol{m}_j)^T, \tag{2}$$

where $\boldsymbol{m}_j \in \mathbb{R}^d$ is the mean of the j^{th} class, computed from the sampled data in the j^{th} class while $\boldsymbol{m} \in \mathbb{R}^d$ is the mean of all samples. Finally, the target pose separability s is evaluated by

$$s = tr\left(\boldsymbol{S}_W^{-1}\boldsymbol{S}_B\right) \text{ if } \ det(\boldsymbol{S}_W) \neq 0, \tag{3}$$

where $det(\cdot)$ denotes the determinant of a square matrix. Note that the measure s is always positive and the lower the s, the worse the separability and vice versa. The determinant $det(\boldsymbol{S}_W)$ is zero only if at least one of the features is a linear combination of the other features, which is unlikely due to human movement variability.

B. Input features ranking and selection

A filter method is used to rank the total D features based on an objective function J which is decoupled from the interface scheme. The feature with a higher rank is given a higher priority to be selected. The forward sequential search (SFS) is applied to rank the features, which incrementally searches the next feature with the highest objective function value when used in combination with the previously ranked ones. The process is repeated until the number of ranked features reaches the overall number D. After the ranking, the first n_d features will be selected.

The objective function J used for feature ranking is formulated as

$$J = \beta s, \tag{4}$$

where scaling factor $\beta \in [0, 1]$ is used to modify the value s obtained from (3). This is required to mitigate the scenario of highly uneven separability, where a high value of s can result from some very highly separable class pairs while the others in the set perform far below the expectation. It is noted that the (3) is an averaged behaviour without showing the variance with respect to the average. In our application, it is important to be able to accurately distinguish each target pose from the others. Therefore, scaling factor β is introduced to scale down the input feature sets where the worst separability between two classes in the target pose set is below the lower bound. Denote s_p as the minimum separability between all the possible combinations of two classes and s_{p^-} as the lower bound of s_p which is a tuning variable. The β takes the form

$$\beta = \min\left\{\frac{s_p}{s_{p^-}}, 1\right\}. \tag{5}$$

III. Experimental Methodology

In this section, the experimental protocol is presented, followed by the methodology to perform the comparison between the outcomes from the population-averaged and personalised input feature sets.

A. Experimental protocol

Ten able-bodied (7 male, 3 female; all were right-handed;) were recruited in this study. The age range was $[24, 32]$ with a median at 27. The experimental protocol was approved by the University of Melbourne Human Research Ethics Committee, project ID 11878. Informed consents were received from all subjects. Subjects were asked to perform forward-reaching tasks towards targets placed along the parasagittal plane of extending the upper limb forward, for ten iterations each. The targets, shown as a small sphere that the subject should reach with their hand, were displayed in the head-mounted display based virtual reality (VR) environment, see Fig. 1. The position of the target was adjusted for each subject to reflect the intended shoulder and elbow flexion

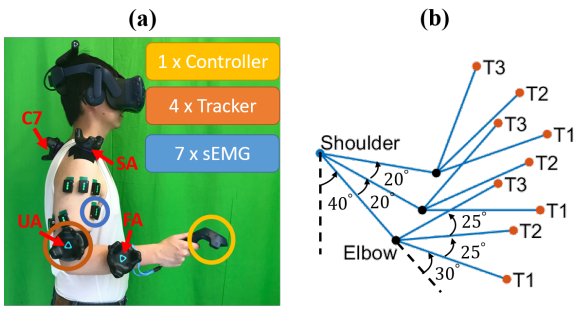

Fig. 1. (a) Experiment setup and sensor deployment, (b) Target set within the parasagittal plane, T1-T3 denote the intended target poses.

angles. The subjects were required to hold their arm position upon reaching the target for one second to allow the quasi-static readings upon reaching to be collected. A set of kinematic and sEMG signals associated with the human subject upper limb and upper body movements upon reaching the target poses were recorded.

Nine reaching targets set in the parasagittal plane were set based on the joint space human arm displacements, as illustrated in Fig. 1(b). To set the locations of the targets, subjects were required to reach the target joint poses as shown in Fig. 1(b) during the initialisation of the experiment. In this study, only the prosthetic poses are of interest, thus $C = 3$ intended target pose classes, denoted as T_i, $(i = 1, 2, 3)$, are extracted from the 9 targets as labelled in Fig. 1(b). It is worth noting that alternatively, one can consider the pose of the entire arm as the intended target poses, if the application is extended beyond the prosthetic application, in cases where the upper extremity poses are of interest (e.g. for a general human-machine interface problem).

B. Data collection and feature extraction

The body and arm kinematic signals and upper arm sEMG signals were collected using wearable sensors with a sampling rate of 90 Hz and 1,111 Hz, respectively. Fig. 1(a) shows the sensor deployment. Upper body and upper arm postural data were acquired through three HTC VIVE Trackers (with motion capture sensors and an embedded Inertial Measurement Unit (IMU)) attached to the subjects' upper arm (UA), shoulder acromion (SA) and C7 vertebrae (C7). Another tracker on the forearm (FA) and the controller in the hand were utilised to control the arm in the VR environment only. Seven Delsys Trigno sEMG electrodes were attached to the dominant upper arm of the subjects: two on the biceps long/short heads, two on the triceps lateral/long heads, three on the anterior, middle and posterior of the deltoid. The raw sEMG signals were filtered by a 4^{th} order Butterworth band-pass filter with 10-500 Hz passband. Outliers of more than three standard deviations from the mean were removed.

The candidate features were extracted from the quasi-static data during the last-second holding period upon reaching the targets in a coordinated fashion, and were summarised in Tab. I. For sEMG features, time-domain features were selected as they were reported to be superior to frequency-domain features in classifying the intended target poses [23] and are of low computational cost. To extract the time-domain features for each of the 7 signal sites of sEMG, a sliding window of 200ms was applied, with overlapping 100ms (resulting 10Hz sampling rate). Assume that the sEMG signal $\boldsymbol{y}$ has n_s sampling points in a sliding window and $\boldsymbol{y}_i$ be the i^{th} sample. The following time-domain sEMG features were extracted, consistent with [23].

Mean absolute value (MAV) is usually used to detect the onset of the sEMG signal for prostheses control.

$$\text{MAV} = \frac{1}{n_s}\sum_{i=1}^{n_s}|\boldsymbol{y}_i|. \tag{6}$$

Root mean square (RMS) captures the envelope of the sEMG signals and reflects the strength of muscle contraction.

$$\text{RMS} = \frac{1}{n_s}\sqrt{\sum_{i=1}^{n_s}\boldsymbol{y}_i^2}. \tag{7}$$

Wave length (WL) calculates the cumulative length of sEMG waveform within the sliding window.

$$\text{WL} = \sum_{i=1}^{n_s-1}|\boldsymbol{y}_{i+1} - \boldsymbol{y}_i|. \tag{8}$$

Before listing the following sEMG features, first define the $\text{sgn}(\cdot)$ function as

$$\text{sgn}(x) = \begin{cases} 1, & \text{if } x > \text{threshold} \\ 0, & \text{otherwise} \end{cases}. \tag{9}$$

Zero crossing (ZC) is a rough estimation of the frequency characteristics of the sEMG sequence.

$$\text{ZC} = \sum_{i=1}^{n_s-1}\text{sgn}(-\boldsymbol{y}_i\boldsymbol{y}_{i+1}). \tag{10}$$

Slope sign change (SSC) measures the times that the sEMG sequence changes the slope sign, which also shows

TABLE I

KINEMATIC & SEMG FEATURES AND SIMPLIFIED LABELS

Label	Features
A1	Shoulder flexion/extension
A2	Shoulder adduction/abduction
B1	Scapular protraction/retraction
B2	Scapular depression/elevation
C1	Trunk flexion/extension
C2	Trunk left/right bending

(a) Kinematic features and the simplified labels for each feature.

Features	Biceps		Triceps		Deltoid		
	Short head	Long Head	Lateral Head	Long Head	Anterior	Middle	Posterior
MAV	D1	E1	F1	G1	H1	I1	J1
RMS	D2	E2	F2	G2	H2	I2	J2
WL	D3	E3	F3	G3	H3	I3	J3
ZC	D4	E4	F4	G4	H4	I4	J4
SSC	D5	E5	F5	G5	H5	I5	J5

(b) sEMG features and the simplified labels arranged in matrix form. For example, the mean absolute value (MAV) of biceps short head sEMG is simplified as D1.

the frequency characteristics of the sEMG signal.

$$\text{SSC} = \sum_{i=1}^{n_s-1} \text{sgn}\left((\boldsymbol{y}_i - \boldsymbol{y}_{i-1})(\boldsymbol{y}_{i+1} - \boldsymbol{y}_i)\right). \quad (11)$$

Kinematic postural data were downsampled to 10Hz. The kinematic features were: shoulder flexion/extension (f/e) and adduction/abduction (abd/add); scapular protraction/retraction (p/r) and depression/elevation (d/e); and trunk flexion/extension (f/e) and left/right bending (lb/rb). Since the targets were in the parasagittal plane, shoulder internal/external rotation and trunk rotation were not considered. The motion trajectory of position, velocity and acceleration in the motion were not considered.

In total, $D = 41$ candidate features were extracted. The details are summarised in Tab. I. For the convenience, each feature is coded with a label representing its category and the place in the category. For each iteration of reaching, the recorded data matrix with D features was of size 41×10 and was concatenated one after another. The overall samples of each feature were normalised to zero mean and unit variance.

C. Comparison of the outcomes: population-averaged and personalised input feature sets

1) Feature selection: The population-averaged and personalised cases differ in terms of the data set used to determine the input feature set. The personalised feature sets, denoted as F_{per} are selected based on the data from each subject, while the population-averaged feature sets denoted as F_{pop} are based on the overall dataset. The number of selected features is denoted as n_d whose maximal value N_d in this study is the number when the classification accuracy saturates with the increasing number of selected features. The stop criterion is defined as the step-wise improvement of classification accuracy of less than 1% for the consecutive possible values of n_d.

2) Feature evaluation: The selected feature set is evaluated on the dataset from each subject by its classification performance based on the linear discriminant analysis (LDA) classifier. The average classification accuracy from 10-fold cross-validation is selected as the performance metric. The classification evaluations are applied on both the population-averaged input feature sets F^i_{pop} $(i = 1, 2, \ldots, N_d)$ and the personalised input feature sets F^j_{per} $(j = 1, 2, \ldots, N_d)$. Here, the superscript i and j denote the number of selected input features. Thus, each input set F^i_{pop} and F^j_{per} would result in a classification accuracy value (performance) for each subject. It should be noted that for each subject, the selected features of F^j_{per} can be different. The cases where $n_d < N_d$ are evaluated to see how much the number of input features can be reduced while maintaining functional performance.

3) Statistical analysis: The Non-parametric method is selected because the obtained classification performance does not fit the normality assumption. First, a Friedman test is performed to detect if there is any significant difference between the performance of all possible F^i_{pop} and F^j_{per}. If the null hypothesis is rejected, then a post hoc pairwise comparison is conducted. The performance of F^i_{pop} and F^j_{per} with $i = j = 1, 2, \ldots, N_d$ are compared using the two-sided Wilcoxon signed-rank test followed by *Bonferroni* $p-$value correction to control the family-wise error as suggested in [24]. However, unlike [24], we only compare the n_d interested pairs with $i = j$ and correct the corresponding $p-$values. The significance level is set to be $\alpha = 0.05$.

IV. Results and Discussion

The first outcome: the resulting input feature sets when constructed for each subject and for the overall population are presented in the first subsection. The second outcome: the resulting accuracy of the personalised and population-averaged constructed input features are presented and discussed in the second subsection.

A. Input feature selection

Tab. II and III list the input feature selection results for $n_d = 2$ and $n_d = N_d = 8$, respectively. This means that out of the 41 available input features, only n_d input features were selected. The effect is similar to the dimensionality reduction process in Principal Component Analysis (PCA) techniques, but with a significant difference in that, the PCA does not select features but projects the original feature space to a lower dimension space, which will contain components from all the original features. In contrast, input feature selection techniques directly select the input features and do not form new features out of the combinations of the original features.

Tab. II shows the resulting input features for each subject (S1 - S10) and that selected based on the population average (labelled Pop) when n_d is small ($n_d = 2$). A high degree of diversity in the selected input features can be observed. No two subjects ended up with the same input feature set. Only one subject (S7) ended up with the same input feature set as the population-averaged set. Subjects S9 and S10 were observed to "utilise" sEMG-only features, whereas the combination of kinematic and sEMG input features scored higher for most other subjects.

Tab. III shows the resulting input features when n_d is the maximal value N_d. In this case, N_d was determined based on Fig. 2 which demonstrates the average classification accuracy among the subject datasets versus the number of selected input features n_d. The dotted line illustrates the determined value at $n_d = N_d = 8$ where the classification accuracy meets the stop criterion as described in Section III-C, where adding another input feature will not improve the accuracy by more than 1%.

From the Tab. III, it can be observed that the resulting input feature sets still demonstrated a high degree of diversity. However, several input features were prominent, such as C1 (appears 9 times in 10 subjects), B1 and B2 (8 times in 10 subjects) and C2 (7 out of 10 subjects). Despite their prevalence, it ought to be noted that C1, C2, B1 and B2 are kinematics-based features and that alone they cannot be used to determine uniquely the intended target pose, because different target elbow poses should share the same shoulder and trunk pose, see Fig. 1. Referring back to Tab. II, this observation was confirmed (when only 2 input features can be selected, separability of the target poses were not achieved in any subjects (except S7) using exclusively kinematics-based input features).

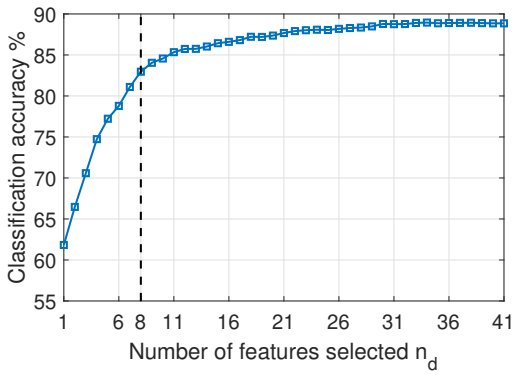

Fig. 2. Classification accuracy vs. the number of selected features. The dotted line shows the N_d determined in this study where the consecutive n_d improve the performance by less than 1%.

B. Resulting target pose classification accuracy

Up to this point, we have discussed the variation of the individual resulting input feature sets between the subjects. Here, we investigate whether the resulting variation would result in a statistically significant difference in performance (the intended target pose classification accuracy). The performance of the population-averaged input feature sets F^i_{pop}, $(i = 1, 2, \ldots, 8)$ and personalised input feature sets F^j_{per}, $(j = 1, 2, \ldots, 8)$ were evaluated through 10-fold cross validation classification accuracy, thus each input feature set resulted in 10 values. As a result, a table with 10 rows of classification evaluations and 16 columns for different input feature sets can be sorted for statistical analysis.

Fig. 3 illustrates the box-and-whisker plot for the performance evaluation. It can be seen that the F^j_{per} performed better than F^i_{pop} when $i = j$. The Friedman test rejected the null hypothesis ($p < 0.01$), indicating there exists a significant difference among the groups in the resulting classification accuracy. The pairwise Wilcoxon tests were then performed and results are annotated in Fig. 3 as well. The F^j_{per} performed significantly better than the F^i_{pop} for $i = j = 2, \ldots, 8$ with $p < 0.05$ except for $i = j = 1$ whose p value level was $0.05 < p < 0.1$. The conclusion can be made that the F^j_{per} significantly outperform the F^i_{pop} in terms of predicting the prosthetic elbow intended target poses.

Furthermore, the detailed improvement of applying the F_{per} can be obtained from Fig. 4. It shows the classification accuracy difference when applying the F_{per} compared to the F_{pop}. The mean values are plotted as the solid line covered by the shaded area representing the standard deviation. The improvement at $n_d = 2$ and 8 were $10.4\% \pm 6.8\%$ and $5.4\% \pm 5.4\%$, respectively. The highest improvement was $12.4\% \pm 6.5\%$ which occurred at $n_d = 3$. In order to achieve the equivalent accuracy as F_{per} with only 3 features, 7 features are required for the F_{pop} as illustrated in Fig. 3. It should be noted that multiple sEMG input features were captured by a single sEMG electrode (e.g. input features D1-D5 were all obtained from the electrode on the biceps short head). Here, the F^3_{per} required fewer or equal number of electrodes (2.3 electrodes on average) than F^7_{pop} (3 electrodes). Further optimisation in personalised input feature design can be done to consider this factor and reduce the number of electrodes required.

TABLE II

n_d = 2 FEATURE SELECTION RESULTS

Label	Pop*	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
A2				●							
B1	●	●	●		●			●	●		
C1	●							●			
C2						●					
D1						●					
D2											●
D5										●	●
E1										●	
E2			●						●		
E3		●									
E5							●				
G2							●				
H4				●							
I3					●						

●: the feature is selected; *: Pop represents the population-averaged selection.

TABLE III

$n_d = N_d$ = 8 FEATURE SELECTION RESULTS

Label	Pop*	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
A1		●	●		●						
A2			●	●				●			
B1	●	●	●		●		●	●	●	●	●
B2	●	●	●	●	●	●			●	●	●
C1	●	●	●		●	●	●	●	●	●	●
C2	●		●	●		●	●	●	●	●	
D1						●	●				
D2											●
D3	●				●			●			
D4		●									
D5	●								●	●	●
E1	●				●		●			●	●
E2			●	●					●		
E3		●								●	
E5					●		●		●		
F2						●					
F3				●							
G2		●					●	●			●
H2				●							
H4				●							
H5						●					
I1								●			
I3			●		●				●		
I4		●								●	
I5				●		●					
J1						●					
J2											●
J3	●						●				
J4								●			

●: the feature is selected; *: Pop represents the population-averaged selection.

V. Conclusion

The outcomes of this study demonstrated a high degree of variability in the selected input feature sets among the 10 subjects, in carrying out the same task in commanding the elbow prosthesis to the same intended target poses. When the overall population data was used to determine the input features, the resulting input features for the population were only shared with one out of 10 subjects when only two input features were selected, and none out of 10 subjects when 8 input features were selected. This difference in the input feature selections translated into a significant performance difference in terms of the accuracy in identifying the intended target pose through the prosthetic interface. When tested against the interface constructed using population-averaged data, the personalised interface demonstrated improvements in the performance of up to $12.4\% \pm 6.5\%$. Note that this result was obtained from able-bodied subjects. In the

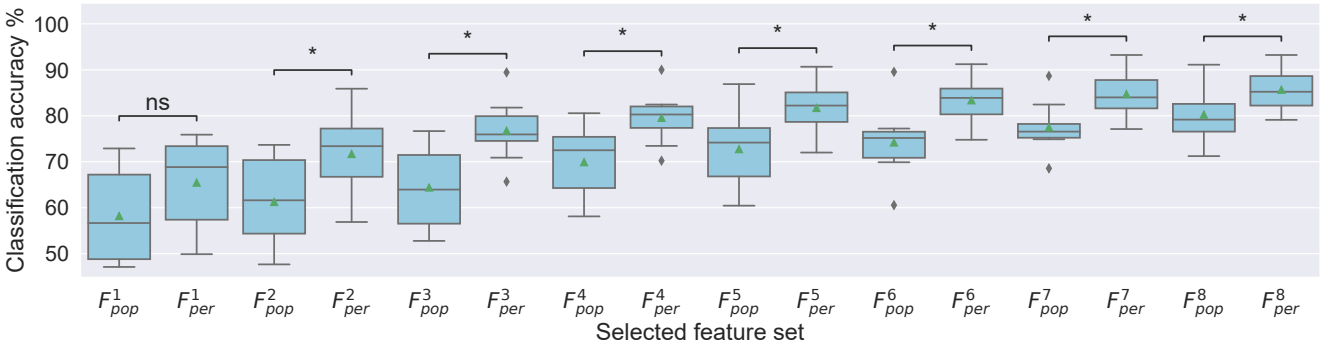

Fig. 3. The box-and-whisker plots show classification performance of each interested feature set. F^i_{pop} and F^j_{per} represent the population-average and personalised feature set with i and j features. The results of statistical analysis are demonstrated through the p-value annotations where ns represents $0.05 < p < 0.1$; * represents $p < 0.05$.

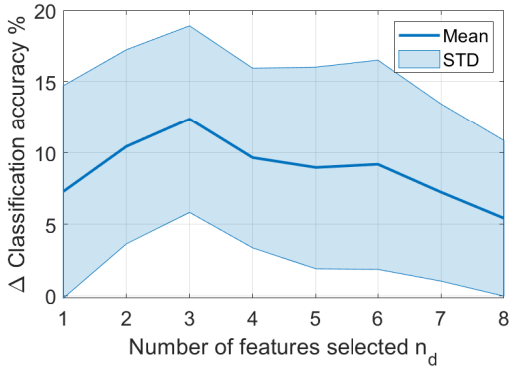

Fig. 4. The plot shows the classification accuracy difference in mean and standard deviation (the shaded area) between the personalised and the population-averaged by subtracting the accuracy of F_{pop} from F_{per}. The mean accuracy improvement reaches the peak when $n_d = 3$.

context of the prosthetic application, it is expected that the anatomical differences of persons living with limb loss would introduce an even more significant diversity, thus further highlighting the need for input features used in the prosthetic interfaces to be personalised to individual users.

REFERENCES

[1] F. Cordella *et al.*, "Literature review on needs of upper limb prosthesis users," *Front. Neurosci.*, vol. 10, p. 209, 2016.

[2] M. Merad *et al.*, "Assessment of an automatic prosthetic elbow control strategy using residual limb motion for transhumeral amputated individuals with socket or osseointegrated prostheses," *IEEE Trans. Med. Robot. Bionics.*, vol. 2, no. 1, pp. 1–1, 2020.

[3] A. Mohammadi *et al.*, "A practical 3d-printed soft robotic prosthetic hand with multi-articulating capabilities," *PloS One*, vol. 15, no. 5, p. e0232766, 2020.

[4] S. Mick *et al.*, "Shoulder kinematics plus contextual target information enable control of multiple distal joints of a simulated prosthetic arm and hand," *J. Neuroeng. Rehabil.*, vol. 18, no. 1, pp. 1–17, 2021.

[5] R. Garcia-Rosas, Y. Tan, D. Oetomo, C. Manzie, and P. Choong, "Personalized online adaptation of kinematic synergies for human-prosthesis interfaces," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 1070–1084, 2021.

[6] J. Huang, G. Li, H. Su, and Z. Li, "Development and continuous control of an intelligent upper-limb neuroprosthesis for reach and grasp motions using biological signals," *IEEE Trans. Syst., Man, Cybern., Syst.*, pp. 1–11, 2021.

[7] T. Yu, R. Garcia-rosas, A. Mohammadi, Y. Tan, P. Choong, and D. Oetomo, "Separability of input features and the resulting accuracy in classifying target poses for active transhumeral prosthetic interfaces," in *Procs Int'l Conf IEEE Eng Med & Biol Soc. (EMBC)*, 2021.

[8] R. Garcia-Rosas, D. Oetomo, C. Manzie, Y. Tan, and P. Choong, "Task-space synergies for reaching using upper-limb prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2966–2977, 2020.

[9] T. R. Kaminski, C. Bock, and A. M. Gentile, "The coordination between trunk and arm motion during pointing movements," *Exp. Brain Res.*, vol. 106, no. 3, pp. 457–466, 1995.

[10] G. Averta *et al.*, "On the time-invariance properties of upper limb synergies," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 7, pp. 1397–1406, 2019.

[11] R. R. Kaliki, R. Davoodi, and G. E. Loeb, "Evaluation of a noninvasive command scheme for upper-limb prostheses in a virtual reality reach and grasp task," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 3, pp. 792–802, 2013.

[12] M. Legrand, N. Jarrassé, F. Richer, and G. Morel, "A closed-loop and ergonomic control for prosthetic wrist rotation," in *IEEE Int'l Conf Rob & Autom*, pp. 2763–2769, 2020.

[13] R. Garcia-Rosas, T. Yu, D. Oetomo, C. Manzie, Y. Tan, and P. Choong, "Exploiting inherent human motor behaviour in the online personalisation of human-prosthetic interfaces," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 1973–1980, 2021.

[14] A. Akhtar, N. Aghasadeghi, L. Hargrove, and T. Bretl, "Estimation of distal arm joint angles from EMG and shoulder orientation for transhumeral prostheses," *J. Electromyogr. Kinesiol.*, vol. 35, pp. 86–94, 2017.

[15] D. Blana, T. Kyriacou, J. M. Lambrecht, and E. K. Chadwick, "Feasibility of using combined EMG and kinematic signals for prosthesis control: A simulation study using a virtual reality environment," *J. Electromyogr. Kinesiol.*, vol. 29, pp. 21–27, 2016.

[16] N. A. Alshammary, D. A. Bennett, and M. Goldfarb, "Synergistic elbow control for a myoelectric transhumeral prosthesis," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 468–476, 2018.

[17] H. J. Hwang, J. Mathias Hahne, and K. R. Müller, "Channel selection for simultaneous and proportional myoelectric prosthesis control of multiple degrees-of-freedom," *J. Neural Eng.*, vol. 11, no. 5, 2014.

[18] G. R. Naik, A. H. Al-Timemy, and H. T. Nguyen, "Transradial amputee gesture classification using an optimal number of semg sensors: An approach using ICA clustering," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 8, pp. 837–846, 2016.

[19] H. Huang, P. Zhou, G. Li, and T. A. Kuiken, "An analysis of EMG electrode configuration for targeted muscle reinnervation based neural machine interface," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 1, pp. 37–45, 2008.

[20] A. Krasoulis, S. Vijayakumar, and K. Nazarpour, "Multi-grip classification-based prosthesis control with two EMG-IMU sensors," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 2, pp. 508–518, 2019.

[21] K. Fukunaga, *Introduction to statistical pattern recognition.* Computer science and scientific computing, Academic Press, 1990.

[22] L. Zhou, L. Wang, and C. Shen, "Feature selection with redundancy-constrained class separability," *IEEE Trans. Neural Netw.*, vol. 21, no. 5, pp. 853–858, 2010.

[23] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7420–7431, 2012.

[24] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *J. Mach. Learn. Res.*, vol. 7, pp. 1–30, 2006.