Adapting Subject-Independent Task-Specific EEG Feature Masks using PSO

Adham Atyabi, Martin Luerssen, Sean P. Fitzgibbon, David M. W. Powers

School of Computer Science, Engineering, and Mathematics,
Flinders University, Adelaide, Australia
{Adham.Atyabi, Martin.Luerssen,Sean.Fitzgibbon, David.Powers}@flinders.edu.au

Abstract—**Dimension reduction is an important step toward asynchronous EEG based BCI systems, with EA based Feature/Electrode Reduction (FR/ER) methods showing significant potential for this purpose. A PSO based approach can reduce 99% of the EEG data in this manner while demonstrating generalizability through the use of 3 new subsets of features/electrodes that are selected based on the best performing subset on the validation set, the best performing subset on the testing set, and the most commonly used features/electrodes in the swarm. This study is focused on applying the subsets generated from 4 subjects on a 5th one. Two schemes for this are implemented based on i) extracting separate subsets of feature/electrodes for each subject (out of 4 subjects) and combining the final products together for use with the 5th subject, and ii) concatenating the preprocessed EEG data of 4 subjects together and extracting the desired subset with PSO for use with the 5th subject. The results indicate the feasibility of generating subsets of feature/electrode indexes that are task specific and can be used on new subjects.**

I. Introduction

Particle Swarm Optimization (PSO) is an unsupervised, population based approach inspired by bird flocking and fish schooling. PSO evolves a population of particles that represent possible solutions in the search space. Several studies employed PSO for dimension reduction of EEG signals [1]-[6]. It is common in Evolutionary Algorithm (EA) based dimension reduction studies to report the achieved performance with a subset of features or electrodes that are tuned on a testing set. This provides a degree of contamination between training and testing sets that can be prevented by introducing a third set that has not been involved in the feature/electrode reduction procedure.

We previously investigated the use of PSO for generating a subset of features that are optimal for a subset of electrodes in [7]. This study is focused on further assessing the re-usability of the PSO-selected subsets of features/electrodes by combining subsets of feature/electrode indexes from 4 subjects and applying this to a 5th subject. Section II discuss the used PSO paradigm for feature/electrode reduction. Section III presents the experimental setup. The results are demonstrated in IV.

II. PSO Paradigm for Feature/Electrode Reduction

A. Basic PSO

Particle swarm optimization (PSO) is a population-based method inspired from animals social behaviors. PSO searches for possible solutions through local and global interactions of particles (members of population). Each Particle in the population contains a velocity and a position in the search space denoted as $V_{i,j}$ and $X_{i,j}$. i represents the particle's index and j is the dimension in the search space respectively. Particles have a limited memory including *Local-Best* and *Global-Best* positions. *Local-Best* is the best solution (position in the search space) found by a particle and *Global-Best* is the best solution found by the particles and their neighbors [8]. In basic PSO, particles update their velocity and position in the search space ($V_{i,j}$ and $X_{i,j}$) using Eqs. 1 and 2 respectively.

$$V_{i,j}(t) = w \times V_{i,j}(t-1) + c_1 r_{1,j} \times (p_{i,j}(t-1) - x_{i,j}(t-1)) \\ + c_2 r_{2,j} \times (g_{i,j}(t-1) - x_{i,j}(t-1)) \tag{1}$$

$$x_{i,j}(t) = x_{i,j}(t-1) + V_{i,j}(t) \tag{2}$$

In Eq. 1, $r_{1,j}$ and $r_{2,j}$ are random values in the range of 0 and 1. Acceleration coefficients (c_1 and c_2) control the impact of social and cognitive components and w is the inertia weight. $p_{i,j}$ and $g_{i,j}$ are the local and global best positions (*PBest* and *Gbest*) and can be updated using Eq. 3 and 4 respectively.

$$p_i(t) = \begin{cases} p_i(t-1) & \text{if} & f(x_i(t)) >= f(p_i(t-1)) \\ x_i(t) & \text{otherwise} \end{cases} \tag{3}$$

$$g(t) = argmin \left\{ (p_1(t)), f(p_2(t)), ... f(p_s(t)) \right\} \tag{4}$$

Neighborhood topology controls the knowledge sharing process (sharing p_i) among particles. Local and global neighborhood topologies respectively allow nearby particles and all particles to share information with each other. Consequently, a Local neighborhood topology permits the existence of more than one *Gbest* while a global neighborhood topology only allows one [9].

Inertia weight controls the impact of last step finding in the generation of the new solution. Large and small inertia weights cause exploration (avoiding local-minima) and exploitation (converging toward the optimal solution) respectively. Random, Fixed, and Linearly Decreasing Inertia Weight (*LDIW*)) are common methods for controlling the w [10]. This study employed the *LDIW* represented in equation 5,

$$w = (w_1 - w_2) \times \frac{(maxiter - t)}{maxiter} + w_2 \tag{5}$$

where, w_1 and w_2 are the initial and final inertia weight, respectively. t is the current iteration and *maxiter* is the termination iteration.

Algorithm 1 Basic PSO's Pseudo-code

Initialization: Randomly initialize the population by using random values for parameters of each particle such as starting position (*X*) and starting velocity (*V*).

Initial Evaluation: Evaluate all members of the population (particles) using fitness function *f*.

repeat

    Updating the population: Update the velocity in each particle using equation 1 and update the mask by applying the new velocity to equation 2.

    Evaluation: Evaluate all members of the population (particles) using fitness function *f*.

    Update Bests: Update Personal (*p*) and Global Best (*g*) using equation 3 and 4.

until (*Termination:* the maximum iteration is achieved or the best member of the population (Global Best) is reached to the desired optimum)

B. PSO based feature and electrode reduction

The common approach for EA based dimension reduction is to apply the EA method on either the feature or electrode dimension. The used PSO paradigm in this study allows the election of a subset of feature indexes for a down-selected set of electrodes. The electrode reduction phase is governed by dividing the electrode space (human head) to 16 equally distant regions denoted as *Areas*. It is desired to select the electrodes from areas that nonscientificly represent the performed task by subjects.

Prior to describing the used PSO paradigm, it is essential to discuss some more general aspects of the EEG signal assessment procedure used in this study.

*1) **EEG Signal Analysis:*** EEG signal classification usually include steps such as signal acquisition, preprocessing (optional), artefact removal (optional), feature extraction, dimension reduction (optional), and classification. In this study, dataset IVa from BCI competition III is used [15], which contains the EEG recording of 5 healthy subjects (*aa, al, av, aw, ay*) performing two classes of motor imagery tasks (right hand and left foot imaginary movement).

The task period is set to 3.5s and 280 randomly presented trials are performed. The first and the last 0.5s of the task periods are omitted, since these two periods can be considered as prior and post task transition periods without much contribution toward EEG signal classification [11]. The resultant 2.5s super-epochs are divided to five 0.5s sub-epochs to provide consistency with other studies on the same dataset. The dataset is demeaned and commonly average referenced, no artefact removal method is applied and frequency features (Abs-DFT) are extracted. $n \times k$ fold Cross Validation (CV) (n=10, k=20) is applied. In each fold, three sets of training, validation, and testing representing 0.9, 0.05, and 0.05 ratio

are generated. The CV is applied in a way that sub-epochs (0.5s) originating in a super-epoch (2.5s) only get involved in one of the sets.

Sigmoid ELM is used as the primary classifier. The number of hidden nodes is set to 80 which is twice the optimal number of used hidden nodes for a Single Layer Back Propagation Neural Network (SLBPNN) employed in previous studies with the same dataset [11]. A higher number of hidden nodes (up to 200) showed no significant improvement in classification performance.

Polynoimal SVM and a modified Peceptron with early stopping are used as alternative classifiers. The early stopping in the modified perceptron is implemented in a way to stop the learning procedure i) after certain number of internal cycles during which the learning error is not improved while the achieved accuracy is above chance level, ii) the learning error is zero, and iii) maximum internal cycles are reached. Bookmaker Informedness is used to assess the classification performance due to being more informative method in comparison with recall, precision, and accuracy. It take advantage from contingency table to calculate informedness. A detailed discussion can be found in [12], [13], and [14].

For more clarity, a flowchart of the procedure is illustrated in Fig. 1.

*2) **PSO Paradigm Used:*** Assume a 2 layer swarm notation in which the population is the first layer and each member of the population is a sub-swarm by itself. The sub-swarms (denoted as *Mask*s) have limited memory containing a vector of selected electrode indexes (denoted *ELV*) and a set of particles bundled in a matrix notation in a way that each particle (each row of the matrix) represents a set of indexes from feature-space that are down selected by PSO for the designated electrode in the *ELV*. This matrix is called *Feature Set Matrix* and denoted by *FSM*.

Given an EEG dataset with *N* electrodes and *K* feature points in each electrode, *ELV* vector in each *Mask* of the swarm contain *n* out of *N* electrodes while the *FSM* contains $n \times k$ out of $N \times K$ possible features (indexes).

In addition, each *mask* contains a velocity vector V_i for $i \in [1...n]$ representing the previous velocity for each particle in the feature set matrix. Adopting this notation for Eq. 2, each particle in the sub-swarm can be denoted as $X_{i,j}$ for $i \in [1...n]$ and $j \in [1...k]$.

In the swarm, $Pbest_i$ for $i \in [1...n]$ represent the personal best findings of the *masks*. This is equivalent to a copy of the matrix of the masks that received the highest classification performance.

Therefore, a member of the swarm (sub-swarm) called Sub_i contains a mask with *n* electrodes and $n \times k$ features. In each iterations, the sub-swarms share their best achieved mask (the mask with the highest classification performance) called *Pbest*. *Gbest* is considered as the global best and represents the mask that the best *Pbest* belong to.
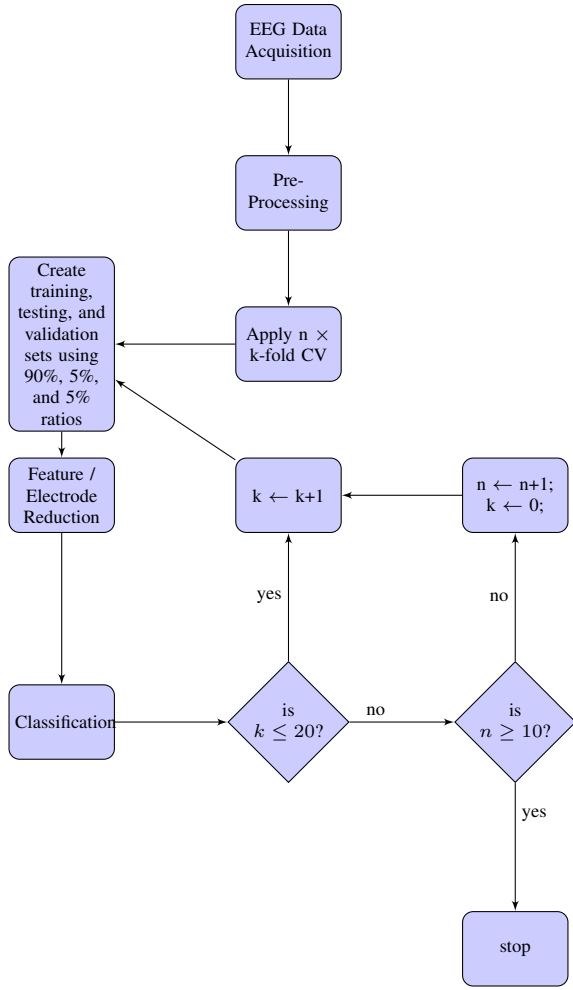
Fig. 1. The applied steps for EEG data processing

Assuming the mentioned 2 layer swarm notation, the following steps are taken for feature and electrode reduction:

1) Initialization:
   - Create a Mask with random index values in *ELV* vector and *FSM* (Sub_1).
   - Create a Mask with random index values in *ELV* vector and *FSM* (*Gbest*).

2) Initial Evaluation: Evaluate the Sub_1 by applying its *FSM* and *ELV* to the EEG training and validation sets and assessing the classification performance using Sigmoid ELM with 80 hidden nodes.

3) Update the population:
   - Isolate the top 10 sub-swarms and eliminate the rest.
   - For each Sub_i for $i \in [1..10]$
     a) Generate a new *mask* with new set of electrodes (*ELV*) that are positioned in nearby *areas*. *FSM* and *Pbest* are inherited from the parent *mask*.
     b) Generate a new *mask* with new set of electrodes (*ELV*) that are positioned in the same *areas*.

FSM and *Pbest* are inherited from the parent *mask*.
   c) Generate a child *mask* with new set of $n \times k$ indexes of features using velocity vector *v* and position matrix *x* and update equations 5, 1 and 2. *ELV* and *Pbest* are inherited from the parent *mask*.
   d) Generate a child *mask* with new set of randomly chosen electrodes (*ELV*) that are positioned in the same areas and a new set of $n \times k$ features using velocity vector *v* and position matrix *x* and update equations 5, 1 and 2. *Pbest* is inherited from the parent *mask*.

4) Re-Evaluation: Evaluate the Sub_i for $i \in [1...t]$ where t is the maximum number of sub-swarms in the swarm by applying their *FSM* and *ELV* to the EEG training and validation sets and assessing the classification performance using Sigmoid ELM with 80 hidden nodes.

5) Update *Pbest* and *Gbest*.

6) go to step 3 if *Gbest* informedness is below 0.9 or maximum iteration is not reached.

7) Final Evaluation: re-evaluate the *Gbest* by applying its *feature set matrix* and *ELV* to the EEG training and testing sets and assessing the classification performance using Sigmoid ELM with 80 hidden nodes, Polynomial SVM and Modified Perceptron with Early Stopping.

8) Stop.

As mentioned in the procedure, Sigmoid ELM's informedness results are used as the fitness criteria for evaluating the potential of the generated masks in the swarm.

The testing set is only introduced in the final evaluation stage in order to provide unbiased results. The masks are generated from separate data in the training and the validation sets. In this study, two variations of PSO is used. The variation is in terms of the parameter adjustment. The used parameter adjustments are presented in table I.

TABLE I
THE INITIAL SETUP OF PSO

Parameters	PSO1 The Value	PSO2 The Value	The Adjustment Method
c_1	0.5	0.5	Fixed Acceleration
c_2	0.5	2.5	Coefficient (FAC)
w_1	0.2	0.2	Linearly Decreasing
w_2	1	1	Inertia Weight (LDIW)

III. EXPERIMENTAL SETUP

As mentioned in previous sections, in [7], a PSO paradigm that allows the reduction of feature and electrode dimension concurrently is used on 2 class dataset containing EEG signal of 5 subjects performing motor imagery tasks. 3 sets of training, validation, and testing are used and the output of the PSO paradigms are meant to be tuned on the validation set and further evaluated by the testing set.

The average primary results achieved by either of the used PSOs from the reduced dimension validation set (using the

masks generated by PSO) reached 0.48 informedness (averaged across subjects) using Sigmoid ELM while the averaged informedness across subjects using the combination of full-set and Polynomial SVM reached 0.478. The achieved informedness (averaged across subjects) from PSO1 and PSO2 on the reduced dimension testing set reached 0.11 with Polynomial SVM in PSO1 and 0.19 with modified Perceptron in PSO2. These results are presented in Table II.

TABLE II
COMPARISON RESULTS OF TWO PSO BASED DIMENSION REDUCTION METHODS USING SIGMOID ELM, POLYNOMIAL SVM AND MODIFIED PERCEPTRON WITH EARLY STOPPING ON THE TESTING SET.

| Subject | Full-Set | | | | | |
	SELM	PSVM	Perc			
AA	0.08	0.38	0.11			
AL	0.31	0.72	0.52			
AV	0.10	0.21	0.15			
AW	0.18	0.58	0.31			
AY	0.09	0.51	0.00			
	PSO1			PSO2		
	SELM	PSVM	Perc	SELM	PSVM	Perc
AA	0.02	–	0.02	0.04	–	0.01
AL	0.21	0.21	0.11	0.22	0.22	0.11
AV	0.03	0.03	0.01	0.01	0.02	0.01
AW	0.07	0.07	0.02	0.07	0.06	0.01
AY	0.18	0.13	0.08	0.18	0.12	0.8

This indicates the lack of generalizability in the solutions (masks) generated by PSO. To address this problem, 3 new masks are considered based on

1) The best performing mask on the validation set.
2) The best performing mask on the testing set.
3) The most commonly used indexes by the masks during the 10×20 CV.

The results indicate improvement of at least 0.1 informedness in most cases using either of the suggested new masks. In addition, in some subjects, the best performing masks on the testing and the validation sets reached to maximum performance of 0.53 using Polynomial SVM. The results achieved with the combination of either of these new masks and Polynomial SVM is illustrated in Table III.

TABLE III
THE HIGHEST BOOKMAKER INFORMEDNESS PERFORMANCE ACHIEVED BY ANY COMBINATION OF PSO BASED DIMENSION REDUCTION METHODS AND THE POLYNOMIAL SVM.

| Subject | PSO1 | | | PSO2 | | |
	Val	Tes	Com	Val	Tes	Com
AA	0	0.04	0.11	0.11	0.05	0.12
AL	0.53	0.44	0.36	0.43	0.43	0.30
AV	0.07	0.03	0.24	0.12	0.11	0.20
AW	0.13	0.12	0.37	0.30	0.24	0.27
AY	0.36	0.30	0.19	0.35	0.34	0.35

Two experiments are designed to further analyze the achievements with the new 3 masks and generate a mask that is task (not subject) specific so that it can be used on a subject whose data was not involved in its creation.

A. Experiment 1: Analysis of Masks at Subject Level

In this experiment, the previously generated masks by either of PSO1 or PSO2 for each subject are used in the following way:

1) Given that 10×20 CV is used for generating the masks, 200 masks each representing the *Gbest* of the swarm for a particular fold is reported in each subject.
2) For each subject, the remaining masks from other subjects performing the same tasks are concatenated together creating a *Meta-set* of $4 \times 10 \times 20$ masks (5 subjects were involved in the EEG data acquisition).
3) For each subject the following masks are extracted from its *meta-set*:
   - The best performing mask on the validation set.
   - The best performing mask on the testing set.
   - The most commonly used indexes by the masks.
4) The potential of the new 3 masks containing selected feature/electrode indexes extracted from the *meta-set* of each subject are assessed by applying them to the subject's EEG data and using Sigmoid ELM, Polynomial SVM, and modified Perceptron with early stopping as classifiers (in a 10×20 CV paradigm).

B. Experiment 2: Analysis of Masks at Task Level

In this experiment, the described PSO paradigm explained in section II-B is used to generate a set of masks that are task specific with the capability of being applied to subjects that were not involved in the mask acquisition procedure. To do so, the following steps are used in each subject:

1) The EEG data of the other 4 subjects are concatenated together. To preserve the underlying characteristics of each one of the 4 datasets used in the concatenation, they are first preprocessed individually (demeaned and commonly average referenced and Frequency features (Abs-DFT) are extracted).
2) A 10×20 CV is used and the dataset is divided to three sets of the training, the validation and the testing with the ration of 0.9 0.05 and 0.05.
3) In each fold, the PSO paradigm explained in section II-B is used to generate a mask (*GBest* of the swarm) containing down-selected feature/electrode indexes.
4) For each subject the following masks are extracted:
   - The best performing mask on the validation set.
   - The best performing mask on the testing set.
   - The most commonly used indexes by the masks.
5) The potential of the new 3 masks is assessed by applying them to the subject's EEG data and using Sigmoid ELM, Polynomial SVM, and modified Perceptron with early stopping as classifiers (in a new 10×20 CV paradigm). The subject's EEG data is introduced for the first time.

IV. RESULTS

A. Experiment 1: Analysis of Masks at Subject Level

For an easier view, the results are divided into groups of two subjects and presented in Figs. 2, 3, and 4. With the exception of subject *aa*, the combination of the common mask and the Polynomial SVM reached a higher classification performance compared to the other two masks (best performing masks on the testing and validation sets). There is still a clear
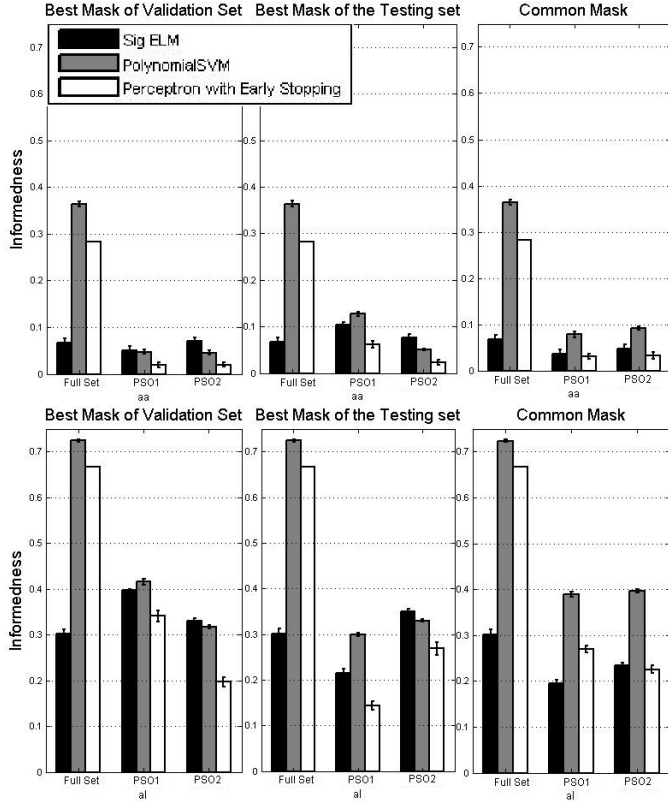
Fig. 2. The comparison results of three new set of masks coined as *Best Mask of Validation set*, *Best Mask of Testing Set*, and *Common Mask* on subject *aa & al* using Sigmoid ELM, Polynomial SVM and modified Perceptron with early stopping (experiment 1)

Fig. 3. New mask comparison results, as in Fig. 2, but for subjects *av & aw* (experiment 1).
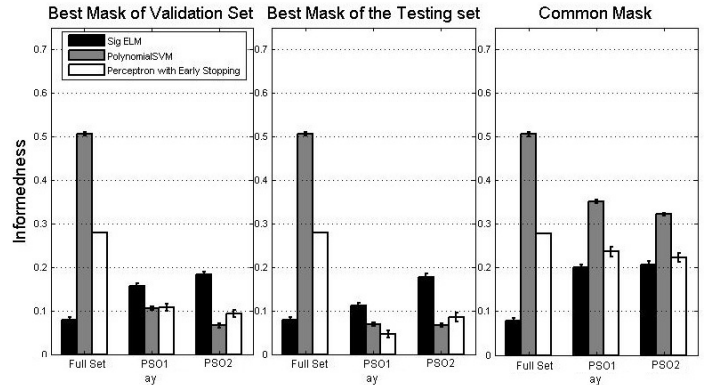
Fig. 4. New mask comparison results, as in Fig. 2, but for subject *ay* (experiment 1).

disadvantage across subjects in terms of the classification performance using the reduced dimension sets with either of the three masks, as in all subjects, the combination of full-set and the Polynomial SVM reached to the highest performance. Despite such a disadvantage, a noticeable Informedness gain is still observed with the set of masks that are tuned on other subjects' signals. The final testing set was not used in either of PSO paradigm steps, which guarantees non-contamination of the testing and training sets.

B. Experiment 2: Analysis of Masks at Task Level

For an easier view of the results, they are divided into groups of two subjects and presented in Figs. 5, 6, and 7. The common mask shows clear superiority in comparison with other two masks. In addition to the achieved informedness, in subjects *ay, av, and aa*, the combinations of the common mask and the Polynomial SVM is marginally performing as well as the combinations of the full-set and the Polynomial SVM. Given that the results indicate the achieved performance with the testing set that were intentionally kept blind to the PSO feature/electrode reduction paradigm and considering the fact that the used paradigm reduce 99% of the data, such a performance is noteworthy.
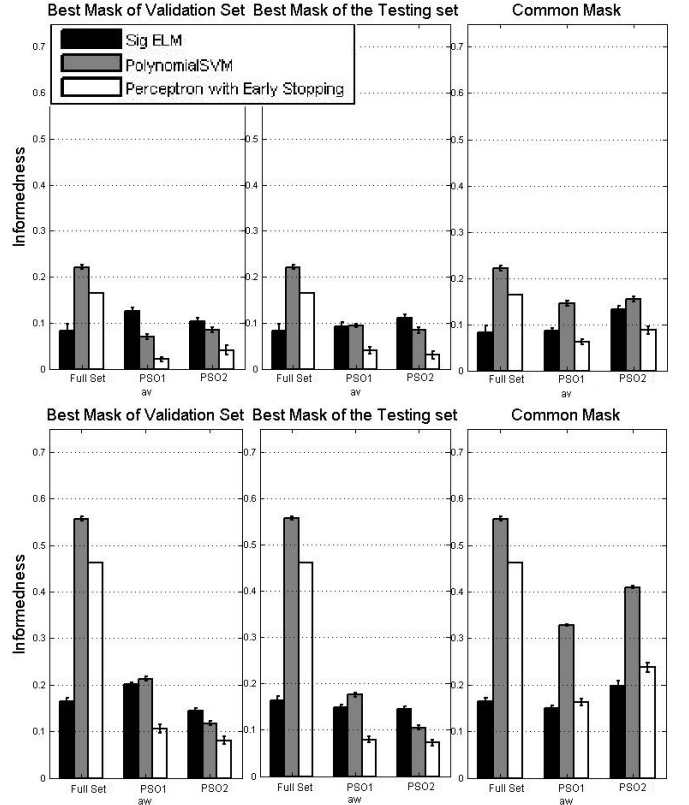
V. CONCLUSION AND DISCUSSION

Dimension reduction is an important step toward asynchronous EEG based BCI systems. A PSO-based approach that reduces Feature and Electrode dimensions concurrently shows potential for use in EEG studies, as it reduced 99% of the EEG data and demonstrated generalizability through the use of 3 new subsets of features/electrodes that are selected based on the best performing subset on the validation set, the best performing subset on the testing set, and the most commonly
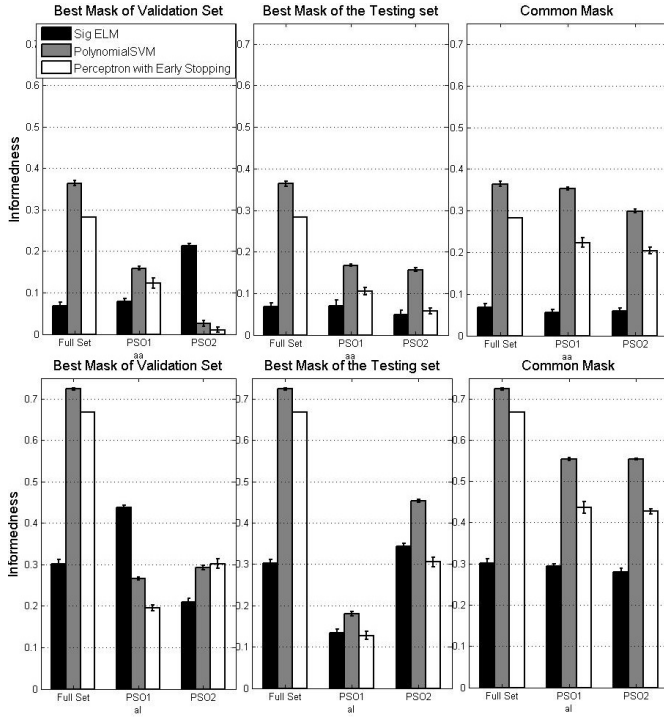
Fig. 5. The comparison results of three new set of masks coined as *Best Mask of Validation set*, *Best Mask of Testing Set*, and *Common Mask* on subject *aa & al* using Sigmoid ELM, Polynomial SVM and modified Perceptron with early stopping (experiment 2)
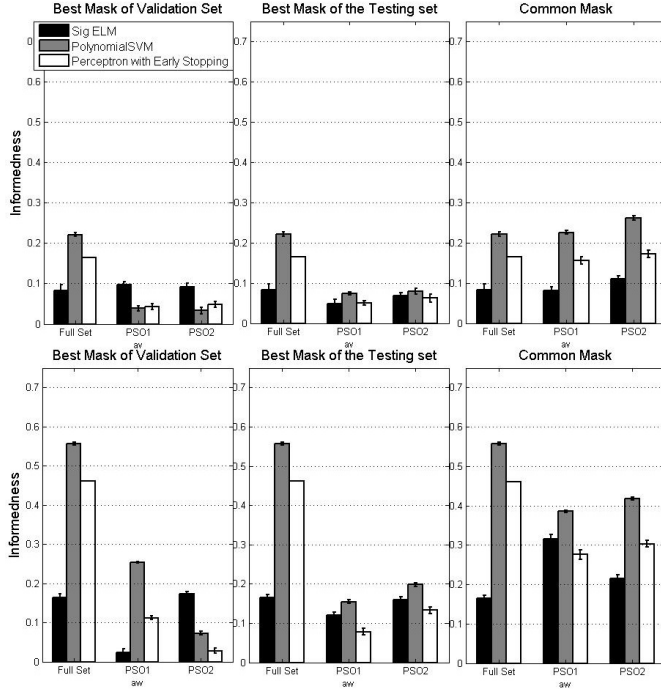

Fig. 6. New mask comparison results, as in Fig. 5, but for subjects *av & aw* (experiment 2).

used features/electrodes. This study provided further analysis on these new subsets by assessing their re-usability. To do so, two different approaches are considered.
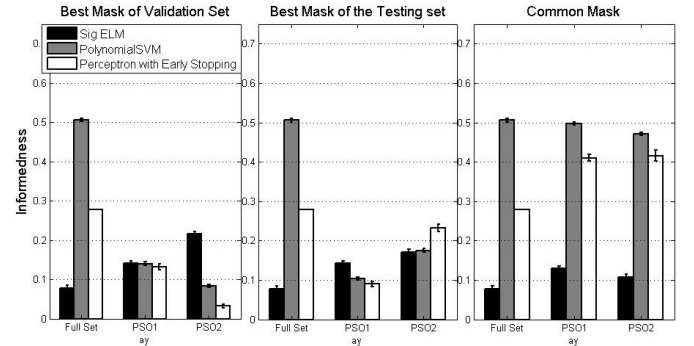

Fig. 7. New mask comparison results, as in Fig. 5, but for subject *ay* (experiment 2).

In the first approach the subsets generated with PSO from 4 subjects are concatenated together and 3 new subsets representing the best performing subsets in the testing, and the validation sets and the most commonly used indexes are extracted and their potentials are measured using the EEG data of the 5th subject. Given that this procedure prevents the contamination between the training and the testing sets and the resulting subsets are tuned on other subjects, the achieved results are noteworthy.

In the second approach, the EEG data of 4 subjects are concatenated together after passing the preprocessing and feature extraction stages. The PSO paradigm is used to generate sets of subsets for the concatenated dataset. Three new subsets representing the best performing subsets on the testing and the validation sets and the most commonly used indexes are extracted. The 5th subject's EEG data is used to asses the potential of the 3 subsets. The results indicate the superiority of the commonly used indexes in comparison with the other two subsets, in addition to reaching a marginally equal or better classification performance in comparison to the full-set (in 3 out of 5 subjects). Given that the subject's EEG data was only introduced to evaluate the potential of the extracted subsets and was not involved in the swarm evaluation stages, no contamination between the training and the testing set can be considered. The results indicate the possibility of shortening the on-line training phase in EEG study through generating subsets of feature/electrode indexes that are originated from other subjects' signals that performed the same task in an off-line mode.

VI. FUTURE WORK

Despite the fact that the results clearly show the possibility of generating task-tuned masks that reduce 99% of the EEG data (especially with the combination of the common mask and the Polynomial SVM as in experiment 2), there are lots of questions that need to be answered. Given that in this dataset - based on the achieved informedness with the full-set - subjects can be categorized as strong (*al*), normal or middle range (*aw* and *ay*), and weak (*aa* and *av*), it is noticeable that in subjects *aa, av*, and *ay*, the achieved informedness with combinations of the common mask and Polynomial SVM are as well as or

close to the best achieved results with full-set. In all of these subjects, the mask is generated from the EEG data that of a strong subject and at least one normal subject. However, the impact of such combination is not clear. In other words, the variations of subjects (in terms of their performance strength) that is required in order to create a fine-tuned task-specific mask that is suitable for all other subjects with different performance strengths is not clear. In case of generating such mask, it is still not clear if it is possible to improve the weak subject's performance to a level as high as the strong or normal subjects' performances. The contribution of weaker subjects are also unclear. Considering the achieved results with subjects *aa* and *av*, it is not clear if the improvement is due to the inclusion of one strong and two middle range strength subjects in the mask generating procedure or due to the inclusion of one weak subject. A more conservative fusion of masks that consider the performance strength of the subjects in generating the task-specific masks might reach to a higher performance compared with the employed paradigm.

REFERENCES

[1] T. Nakamura, S. I. Ito, Y. Mitsukura and H. Setokawa, *A method for Ealuating the Degree of Humans Preference based on EEG analysis*, Intelligent Information Hiding and Multimedia Signal Processing, 2009. IIH-MSP '09. Fifth International Conference on, 732 735, 2009.

[2] J. Jin, X., Wang and J., Zhang, *Optimal Selection of EEG Electrodes via DPSO Algorithm*, Proceeding of the 7th world congress on intelligent control and automation, pp. 5095-5099, 2008.

[3] N. A. Moubayed, B. A. S., Hasan, J. Q., Gan, A., Petrovski and J., McCall, *Binary-SDMOPSO and its application in channel selection for brain computer interfaces*, Computational Intelligence (UKCI), 2010 UK Workshop on, pp. 1-6, 2010.

[4] E. Y. Tov and G. F., Inbar, *feature selection for the classification of movements from single movement-related potentials*, IEEE transactions on neural systems and rehabilitation engineering, 10(3), pp. 170-177, 2002.

[5] B. A. S. Hasan, J. Q., Gan and Q., Zhang, *Multi-Objective Evolutionary Methods for channel selection in brain Computer interface: Some Preliminary Experimental Results*, Evolutionary Computation (CEC), 2010 IEEE Congress on, pp. 1-6, 2010.

[6] B. A. S. Hasan and J. Q., Gan, *Multi-Objective Particle Swarm Optimization for Channel Selection in Brain Computer Interface*, The UK Workshop on Computational Intelligence (UKCI2009), Nottingham, UK, 2009.

[7] A. Atyabi, M. Luerssen, S. P. Fitzgibbon, D. M. W. Powers, *Dimension Reduction in EEG Data using Particle Swarm Optimization*, submitted.

[8] Y. K., Hwang and P. C., Chen, *A Heuristic and Complete Planner for the Classical Movers Problem*, Proceedings of the 1995 IEEE International Conference on Robotics and Automation, IEEE, pp. 729-736, 1995.

[9] Deneb, *TELEGRIP User Manual*, Deneb Robotics Inc, 1994.

[10] P. C., Chen, Y. K., Hwang SANDROS, *A Dynamic Graph Search Algorithm for Motion Planning*, IEEE Transactions on Robotics and Automation, Vol. 14, no. 3, pp. 390-403, 1998.

[11] S., Fitzgibbon: *A Machine Learning Approach to Brain-Computer Interfacing*, PhD Thesis, School of Psychology, Faculty of Social Sciences, Flinders University, 2007.

[12] D. M. W., Powers: *Recall and Precision versus the Bookmaker*, International Conference on Cognitive Science (ICSC-2003), pp. 529–534, 2003.

[13] D. M. W., Powers: *The Problem of Kappa. 13th Conference of the European Chapter of the Association for Computational Linguistics*, Avignon France, April 2012.

[14] D. M. W., Powers: *Evaluation: From Precision, Recall and F-Measure to ROC., Informedness, Markedness & Correlation*, Journal of Machine Learning Technologies, 2(1), pp. 37-63, 2011.

[15] B. Blankertz, K.-R. Muller, D. J. Krusienski, G. Schalk, J.R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. del R. Millan, M. Schroder and N. Birbaumer: *The BCI competition III:Validating alternative approaches to actual BCI problems*. Neural Syst. Rehabil. Eng., vol. 14, no. 2, pp. 153-159, 2006.

[16] A., Atyabi, S. P., Fitzgibbon and D. M. W., Powers: *Multiplying the mileage of your dataset with subwindowing*, BI'11 Proceedings of the 2011 international conference on Brain informatics, Lanzhou, China, pp. 173–184, 2011.