# Binary Many-Objective Particle Swarm Optimization of EEG Channel Selection for Emotion Recognition with Recurrent Convolutional Autoencoder

Najwa Kouka, *Student Member, IEEE,* Rahma Fourati, *Member, IEEE,* Raja Fdhila, *Senior Member, IEEE,* Patrick Siarry, *Senior Member, IEEE* and Adel M. Alimi, *Senior Member, IEEE*

***Abstract*—Electroencephalogray devices are popular for both medical and commercial use such as in emotion recognition. One of the remaining challenges in processing EEG signals is the selection of optimal channels. This paper proposes a new approach based on binary multi-swarm optimization handling classification error rate, specificity, sensitivity as objectives. More specifically, we perform an unsupervised feature learning with recurrent convolutional layers based on autoencoder architecture directly from clean EEG signals. In order to provide a scientific evidence, extensive validation on three public affective benchmarks, i. e. DASPS, DEAP, and SEED which are different in channel number, used stimuli, and participant ratings, was carried out with subject independent scheme.**

***Index Terms*—unsupervised EEG feature learning, convolutional LSTM, channel selection, binary particle swarm optimization, local learning strategy.**

## I. Introduction

EMOTION is a psychological response state that involves three distinct components: a subjective experience, a physiological response, and a behavioral or expressive response. The ability to recognize emotion is essential in many fields. For example, human-computer interaction for better interaction with robots or computer. In the medical field, identifying a patient's emotional state can provide an indication for the healthcare professional of the patient's mental, physical state, and progression of the healing process. It may be used also for online games, entertainment and e-learning.

For emotion recognition, many physiological related signals have been used. Among them the electroencephalogram (EEG) that allows to obtain more information about emotional state, by reading scalp electrical activity generated by brain structures.

Acquired EEG signals are generally of multi-channel nature. Indeed, a part of these channels may be irrelevant and redundant, which increase the computational cost and reduce classification accuracy, especially for the high-dimensional data sets. Thus, it is important to select optimal subset of channels that is discriminating from all channels.

In the literature, the methods used for EEG channel selection are derived from feature selection methods [1]. Selecting the optimal feature set is known as feature selection thus using these algorithms to select channel is known as channel selection. The used algorithms for channel selection are mainly classified into three categories e.g., filters, wrapper, embedded and hybrid approaches.

Filtering techniques employs an evaluation method such as distance measure, statistic measure [2] and information measure [3] [4] to evaluate the candidate channel subsets. Filtering techniques have some advantages among which are the high speed, independence from the classifier, but they suffer from the low accuracy, since they do not consider the combinations of different channels. The wrapper methods use classifiers to evaluate the selected channels obtained by selection algorithms. The evaluation of every candidate is obtained by training and testing a classification algorithm. Thus, accuracy of the wrapper methods are better than the filter methods, but they may consume more computing resources. In case of the embedded methods, the channels are selected-based on criteria generated during the training phase of a specific classifier. Recursive channel elimination is adopted to keep only channels with appreciated magnitude. This method is a special cases of the wrapper methods. A hybrid technique is a combination of a filtering technique and a wrapper technique attempting to take advantage of both in avoiding the prespecification of a stopping criterion.

Without loss of generality, in the filtering method-based channel selection, the features extracted from the optimal channel set could produce good results. For wrapper and hybrid method-based channel selection, feature extraction and classification are part of the selection procedure, so they produce the best results.

In fact, finding an optimal subset of channel could regarded as a minimization problem with two objectives: (1) minimizing the error rate of classification (maximizing the performance of classification) and (2) minimizing number of channels. These two objectives are conflicting, and require an optimization algorithm to find the best trade-offs for them. During the past decade, few algorithms have been proposed to solve channel selection as an optimization problem. Among them, the meta-

N. Kouka, R. Fourati, R. Fdhila and A. M. Alimi are with the Research Groups in Intelligent Machines, Department of Computer Engineering and Applied Mathematics, National Engineering School of Sfax, University of Sfax, Sfax 3038, Tunisia.
E-mail: {najwa.kouka.tn, rahma.fourati, raja.fdhila, adel.alimi}@ieee.org
Patrick Siarry is with Laboratoire Image Signaux et Systèmes Intelligents (LISSI), Université de Paris Est, 94400 Créteil, France.
Email: siarry@u-pec.fr

heuristics algorithms have shown a lot of advantages in dealing with optimization problems due to their capability to generate solutions by the strong ability of exploring and exploiting in the full search space.

Examples includes the binary particle swarm optimization (BPSO) algorithm that is a swarm intelligence technique that has been applied and shown their effectiveness in the channel selection. One reason why the BPSO algorithm is widely used is that it can converge more quickly and has less adjustable parameters compared with the other evolutionary method such as genetic algorithms.

In fact, most of studies have been conducted on multi-objective channel selection and not considered two more important objectives which are specificity and sensitivity. By formulating channel selection as a many-objective problem, we can select a subset of optimal channels more relevant than defined as multi-objective problem.

In this paper, a binary new many-objective particle swarm optimization with learning strategy (BMaOPSO) is proposed. The contribution of the proposed BMaOPSO is mainly in maintaining diversity by using a multi swarm and to accelerate convergence by cooperative learning strategy. The main contributions of this paper can be summarized as following:

1) The optimization of channel selection problem is regarded as many objective problem with four conflicting objectives to optimize simultaneously, including enhancing of accuracy, specificity and sensitivity and reducing the channel number.
2) In the BMaOPSO, the updating memory is base on the improvement rate instead of using dominance comparator.
3) In the BMaOPSO, the multiple-swarm based learning strategy is adopted to balance between local exploitation and the global exploration.
4) An elite learning strategy is proposed to promote the information exchange among the sub-swarms. In order to spread the superior information obtained by each sub-swarm, the best particle in each sub-swarm is denoted as elite particle and it will learn the excellent information found by other sub-swarms and bring diversity into its own sub-swarm.
5) In the BMaOPSO, the Unsupervised learning network is used to evaluate a candidate channel subset, leading the algorithm to conduct a new search for the optimal subset based on the evaluation results.
6) Extensive validation on 3 datasets with different characteristics in order to make general evidence, taking in consideration subject independent classification.

## II. Related Work on channel selection

The optimization problem of channel selection mostly rotates around to minimize the channel number and to maximize the performance accuracy that measured by the wrapper methods. This paper mainly reviews the works on channel selection methods based on meta-heuristics methods.

The most commonly used algorithms are genetic algorithms and particle swarm optimization (PSO) algorithms. These algorithms are suitable for the channel selection, since it is a population-based algorithm and it able to find solutions (high accuracy) rapidly through the strong ability of exploring and exploiting in the full search space. Another advantage of using these algorithms is its ability of solving conflicting objectives such as maximize the performance accuracy and minimize channel number.

Table I present the summary of meta-heuristic algorithms applied in selecting channels, illustrating the problem objective, classifier, obtained sub-channels, performance value, subject type and the adopted application.

For instance, Luis et al. solve the problem of EEG channel selection for epileptic-seizure classification with the non-dominated sorting genetic algorithm (NSGA) [5] combined with SVM and KNN classifiers. The channel selection is considered as multi-objective optimization problems (MOPs), including the maximization of classification accuracy and reducing the channel number. Their results show that with 1 and 2 channels, the accuracy attend 1. and 97,5% respectively. To improve the performance of channel selection, Luis et al.included two more objectives: (1) true acceptance rate (TAR) and (2) a true rejection rate (TRR), and applied for subject identification system [6]. For channel evaluation the support vector machines (SVM) is employed. Results conduct that with three channels the system achieves accuracy of 83%, with both TAR and TRR of 100%.

Luis et al. is further propose a channel selection method for a biometric system based on NSGA-III [7] and the local outlier factor (Lof). Here the problem designed with the objectives of minimizing the required number EEG channels and increasing the TAR and TRR. With this method, the most relevant area is around C6, T8, T10 and F5 channels with eyes open and the most relevant area is still around the channels C6, T8, T10 and IZ with eyes closed. A high classification performance is attended with three channels.

In [8], a Binary Particle Swarm Optimization (BPSO) integrated with Extreme Learning Machine (ELM) is proposed to select optimum channel set for seizure detection. In BPSO algorithm, the balance between maximizing accuracy and minimizing number of channel is obtained by using a linear fitness function. The proposed methodology achieved 93.21% detection accuracy by selecting only 6 channels (F3-C3, P3-C3, F4-C4, F8-T8, T7-FT9, and FT10-T8).

To deceit identification system, Annushree et al implement a binary version of the BAT algorithm (binary BAT algorithm) for channel selection [9]. With this method each solution is evaluated based on the cost function that includes the accuracy and the selected channel. The accuracy is obtained by the SVM classifier, achieving average accuracy of 96.8% with 13 channels.

Noor et al [10] use the differential evolution (DE) to solve the problem of channel selection as a single objective problem. The objective is minimize the classification error rate that evaluated by linear discriminant analysis (LDA). The procedure of this method as the sequential forward search methods, which added a channel to the channel set at each step to enhance the performance of the classifier until all channels have been processed. Based on the DE, the accuracy

TABLE I: Existing EEG channel selection methods

| Method | Objectives | Classifier | Optimal channels | Performance measure | Classification scheme* | Application |
|---|---|---|---|---|---|---|
| NSGA-III [5] | Minimize channels<br>Accuracy | SVM and KNN | 1 and 2 channels | Accuracy = 0.975 to 1 | SI | Epileptic-seizure classification |
| NSGA-III [6] | Minimize channels<br>Accuracy<br>TAR and TRR | SVM | 3 channels | Accuracy = 0.83<br>TAR and TRR = 1.00 | SI | subject identification system |
| NSGA-III [7] | Minimize channels<br>TAR and TRR | LOF | Eyes open: C6, T8, T10 and F5<br>Eyes closed: C6, T8, T10 and IZ | Eyes open:<br>TAR = 0.993 ± 0.01<br>TRR= 0.941 ± 0.002<br>Eyes-closed<br>TAR= 0.997 ± 0.02<br>TRR= 0.950 ± 0.05 | SD | Motor movement (EEGMMIDB) |
| BPSO [8] | Minimize channels<br>Accuracy | ELM | F3-C3, P3-C3, F4-C4,<br>F8-T8, T7-FT9, and FT10-T8 | Accuracy = 93.21 % | SI | Seizure detection<br>CHB-MIT Dataset |
| BAT [9] | Minimize channels<br>Accuracy | SVM | 13 Channels | Accuracy = 96.8% | SI | Deceit identification |
| DE [10] | Error Rate | LDA | 6 Channels | Accuracy = 86.85% | SI | Emotion recognition |
| IBGSA [11] | Minimize channels<br>Accuracy | KNN | F3-C3, FP1, FPz, FP2, AF7, AF8, FC5, FC6,T7, TP7, TP8, Cz, PO8 and PO7 | Accuracy = 92.50% | SI | Screening of Alcoholism |

SI: Subject Independent, SD: Subject Dependant.

is enhanced from 80% to 86.85% with 6 channels.

Sandeep et al. propose a Binary Gravitational Search Algorithm (IBGSA) [11] as an optimizer to select the EEG channels for the rapid screening of alcoholism. The evaluation of solution is based on the accuracy and the number of selected channels. The proposed method provides 13 optimum channels with a detection accuracy of 92.5%.

This research emphases on meta-heuristics show high classification accuracy with channel number variate form 1 to 13 channels, improving its efficiency in many real-world application.

A recent method that used a set of swarm-intelligence algorithms, including Grey Wolf Optimizer (GWO), PSO, Cuckoo Search (CS), and Dragonfly Algorithm (DA) to find salient features for emotion recognition [12]. The Dominant channels are identified by analyzing the features selected commonly by all swarm-intelligence algorithms over the different subjects. The proposed method was studied on DEAP dataset and the analysis resulted in 11 channels distributed over all brain regions.

## III. EEG Channel Selection based on BMaOPSO Algorithm

To select the most effective EEG channels, we propose a BMaOPSO algorithm. The pseudo code of BMaOPSO demonstrated in Algorithm 1 and its main components are detailed as follows.

## IV. Problem Formulation

In general, channel selection is an optimization problem considering two major issues. One of them is to get higher classification accuracy and the other is to decrease the number of the selected channels. The emotion recognition required two other important classification measures which are sensitivity and specificity.

The goal of channel selection in this work is to reduce the number of selected channel while maximizing the classification accuracy, specificity, and sensitivity, which can be regarded as a many-objective optimization problem. To consider these four objectives, we design the fitness function based on the linear weighting method. The weighting method used to determine the trade-off between different objectives. Since the classification performance is considered to be more important than the number of selected channels, the weight of channel function assumes a lower value than other weights. The fitness function is defined as follows:

$$fitness(p_i) = \sum_{j=0}^{M} w_j f_j(X_i) \tag{1}$$

Where $j$ denotes the objective function, $M$ denotes the number of objectives, and $w_j$ represents the weight of objective function $j$. In this study, the unsupervised classifier RCAE was employed to evaluate the accuracy, specificity, and sensitivity of selected channels.

### A. Binary Particle Swarm Optimization

The PSO is a population-based algorithm which maintains a swarm of particles. In PSO, each particle represents a solution to a particular problem while the fly process of the population can be regarded as a search process. In each generation, each particle i is defined by a position vector and velocity vector to determine its flying direction. During the evolutionary process, the particle i uses its previous personnel position $p_{Best}$ and the global optimal position of the swarm $g_{Best}$ to adjust its flight trajectory as defined by the following equations:

$$\vec{V}_i(t+1) = w\vec{v}_i(t) + c1r1(\vec{x}_{pBest}(t) - \vec{x}_i(t)) + c2r2(\vec{x}_{gBest}(t) \tag{2}$$

$$\vec{X}_i(t+1) = \vec{X}_i(t) + \vec{V}_i(t+1) \tag{3}$$

Although the standard PSO had gained a great success in solving the problems on continuous space, it was powerless in discrete space. Therefore, a new component is added to PSO which is a necessity to be suitable for discrete problem, and thus a binary version (BPSO) was proposed by Kennedy and Eberhart [13] based on sigmoid activation function. In the basic PSO algorithm, due to the continuous real domain of particle position, particles can move continuously through the search space. However, for the search candidates to be able to move in a binary search space, the equation of updating position is updated as follows: also updated by 2, while the new position of the particle is updated by the following formula:

$$S(\vec{V}_i^d(t+1)) = \frac{1}{1+\exp(\vec{V}_i^d(t+1))} \tag{4}$$

$$\vec{X}_i^d(t) = \begin{cases} 1, & if \quad rand \leq \vec{V}_i^d(t) \\ 0, & otherwise \end{cases} \tag{5}$$

where rand is a random number uniformly distributed between 0 and 1. The BPSO has been widely applied in many fields such as EEG channel selection and feature selection problems.

Since, the PSO/BPSO is originally developed for single optimization problem, the external archive and leader selection are employed, making the algorithm useful in problem with more than two objectives to optimize.

### B. Solutions Encoding

In the channel selection problem, the solution is represented in binary form that can be either bit 0 or 1. Bit "0" means the channel is omitted, while the bit "1" means the channel is selected. The number of "1" is the selected channels number. With this encoding, the algorithm can identify the optimal channel combination instead of focusing on evaluating separate channels.

Hence, each particle's position $p_i$ is defined by a dimensional vector $X_i$ with dimension $D$, which encoded as $X_i = [x_1 x_2 ... x_D]$. The following figure illustrates an example of a solution with 10 dimensions, in which the channel selected are 1, 3 and 10.

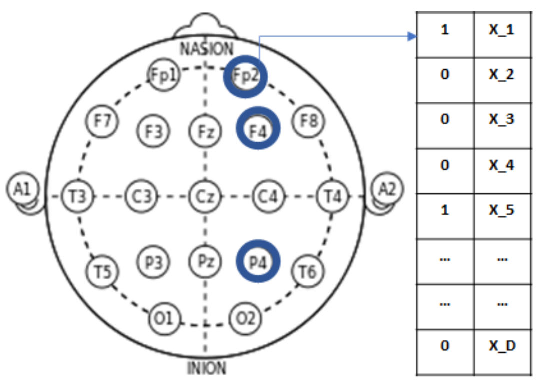



Fig. 1: Solution with D dimension

Individuals are generated by setting only one randomly selected gene to "1" and other channels to "0". This initialization increases the number of solutions with a small size of selected channels and accords with the original intention of finding an appropriate and small channel subset.

### C. Multi-Swarm based Agent Modeling

In the proposed approach, the multi-swarm is used to enhance the exploration of the search space. The multi-swarm is modeled as a Multi Agent System (MAS), by giving the sub-swarm the properties of autonomy and learning. Although, the multi-swarm topology can provide better population exploration, they can not guarantee the exploitation of the population. To deal with this issue, many cooperative interaction strategies have been proposed to balance between the exploitation and exploration. Among them the learning strategies.

The evolutionary process of multiple swarm is presented in Figure 2.

### D. Learning Strategies

In the proposed approach, the learning strategy is dived into global learning strategy and local learning strategy. In the global learning strategy, a new promising solution is generated from different sub-swarm and it used to enhance the exploration ability. While the local learning strategy aim to improve the exploration by learning from the elite particle.

*1) Global Learning Strategy:* In the global learning strategy, we generate a new solutions inherit much more useful knowledge learned from different sub-swarms based on crossover and mutation operators. These two genetic operator have the ability for enhancing the local exploiting and global exploring, respectively [14].

However, the main steps of learning strategy are as follows:

- First, communicating local best solutions within different sub-swarms.
- Second, two parents from these solutions are selected based on the fitness function.
- Third, a potential solutions are generated based on these two parents by the crossover operation.
- Fourth, the generated solutions are mutated for enhancing the diversity.
- Finally, these solutions are then replaced the worst solutions in among sub-swarms.

*2) Local learning strategy:* Personnel learning through memory of solution ($p_{best}$) is an important factor in PSO, as it influences the behaviour of the swarm during the search process. Traditionally, the $p_{best}$ updated only when the new position of the particle is better than the current $p_{best}$. Accordingly, the traditional updating mechanism of $p_{best}$ will not be updated if the classification performance is the same even with less channel number. In order to overcome this limitation, the particle learn from the elite particle that have a greater improvement with a relatively smaller fly distance. The improvement rate [15] is based on both fitness function and the number of channel to reduce. Based on this consideration, the memory update is defined as follows:

$$p_{besti}^t = \begin{cases} X_i, & if \quad I(X_i^t) \geq I(p_{besti}^{t-1}) \\ p_{besti}^{t-1}, & otherwise \end{cases} \tag{6}$$

$$I(X_i) = \frac{f(x_i^t) - f(x_i^{t-1})}{\left\| x_i^t - x_i^{t-1} \right\|} \tag{7}$$

where $\left\| x_i^t - x_i^{t-1} \right\|$ denotes the distance between particle at t and t-1. Since, the position of the particle is defined in binary terms. Therefore, the distance is measured by Hamming distance that represents the number of different bits between two binary positions as defined in 8. However, the greater $I$

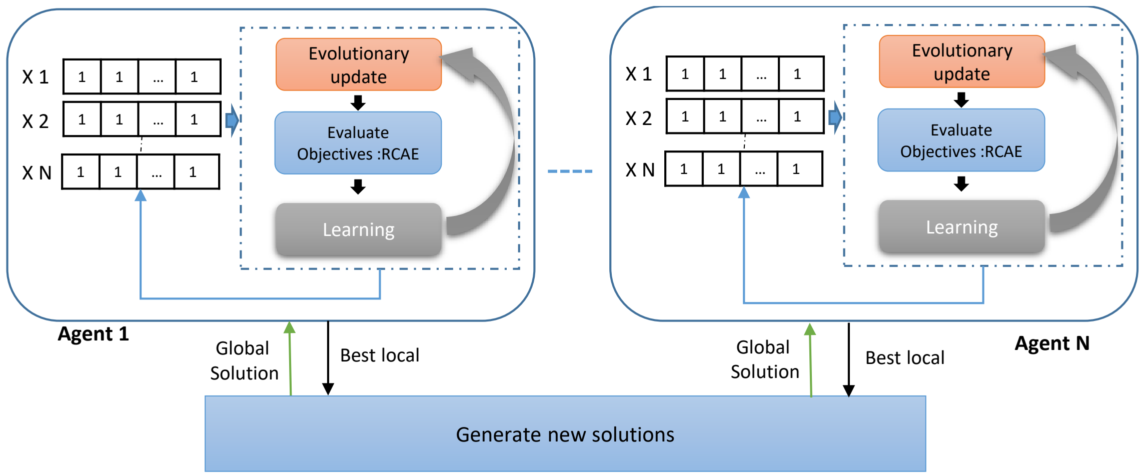



Fig. 2: Concurrent evolution of sub-swarm for channel selection based RCAE

means that a current particle achieve a higher improvement with a relatively smaller fly distance than the previous one.

$$d(X_t, X_{t-1}) = \sum_{k=1}^{D} \begin{cases} 1, & if \quad X_t(k) \neq X_{t-1}(k) \\ 0, & otherwise \end{cases} \tag{8}$$

As $p_{best}$, the selection of leader also influence the exploitation ability of each solution. For that, the solution with best fitness function value and lower channel number is selected to guide the search process.

### *E. Recurrent Convolutional Autoencoder for EEG signals classification*

In this section, the proposed Recurrent Convolutional Autoencoder for EEG signals classification autoencoder is presented.

*1) Basic Concepts:* The autoencoder is a feedforward neural network in which the input is the same as the output. In other words, autoencoders are (unsupervised) learning algorithms that extract features from input data without the need for labeled target datasets. The autoencoder consists of three basic components: the encoder, the code, and the decoder. These function according to their literal meanings. The encoder compresses the input to a 'code,' which is subsequently decoded by the decoder. For this reason, the autoencoder can be used as a dimensionality reduction strategy in time-series forecasting as it can compress the input to a mapped hidden layer. The stacked autoencoder is a hierarchically layered stack of autoencoders and, just like autoencoders, they learn in an unsupervised manner. The model training process involves greedy layer-wise training to minimize the error between the input and output vectors. The subsequent layer of the autoencoder is the hidden layer of the previous one, with each of the layers trained by an optimization algorithm using an optimization function.

2.

### *F. Overall Approach*

At the beginning of the evolutionary process, the swarm with N (number of channel) particles is initialized, where each has a binary components as channels, the initial selected channel is defined according to the index of particle. For example particle $p_i$ has a channel $i$ selected ($X_i = 1$).

In each generation, the evolutionary process of sub-swarms is proceeded in parallel. In each sub-swarm, the leader is selected first based on the fitness function. And each particle then update its position according to its personnel experience and the selected leader. Next, the objectives evaluation based RCAE is performed followed by the evaluation of fitness function. Next, the solution update its memory pbest according to its rate of improvement. Once the particle updating is terminated and mutated, the cooperation among the sub-swarms is achieved by the learning strategy which detailed on above. The above procedure will be repeated until the maximum number of generations is reached. One the optimization is terminated, the archive of BMaOPSO is used as best optimal combination between channels.

## V. Experimental results and discussion

Emotional recognition tasks based on EEG can be divided into subject-dependent and subject-independent ones. In this paper, we focused on subject-independent EEG-based emotion recognition.

This section is related to the evaluation and comparison of BMaOPSO-based EEG channel selection with the two other algorithms. The Binary Multi-Objective PSO (BMOPSO) is a basic binary PSO algorithm, along with a non-dominated sorting genetic algorithm (NSGAII).

In the experiment studies, the parameters of optimization was set as follows:

- **Population size**: 20 solutions
- **Maximum evaluations number**: 200 evaluations
- **Genetic operator**: Binary crossover and binary mutation operators
- **PSO Settings**: The Settings for velocity updating are related to the binary PSOs algorithms in which $c1$ and $c2$= Rand (1.75, 2.0), and $w = 0.7$.

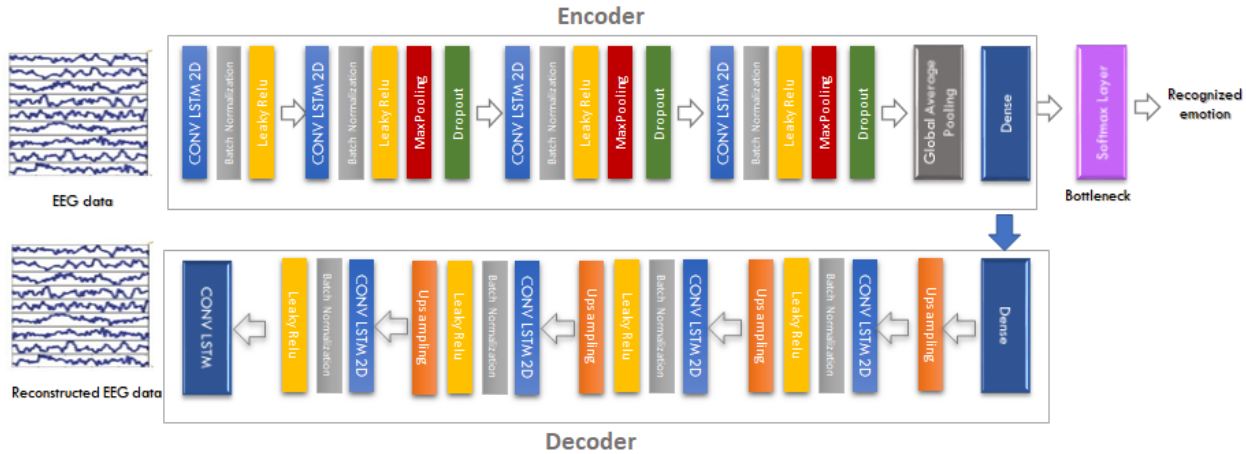



Fig. 3: Model Architecture of Recurrent Convolutional Autoencoder

### A. *Emotional EEG Datasets*

The experiment study was carried out on three public databases widely used in EEG-based emotion recognition: DEAP [16], SEED and DASPS [17]. The details of each dataset is briefly described in below.

*1) DEAP:* item DEAP Dataset [16]: The DEAP dataset was collected from 32 subjects with aged ranged between 19 and 37 years old. They were watching EEG was recorded at a sampling rate of 512 Hz using 32 channels. Figure 4 shows the electrode placements for the EEG. They were watching 40 sets of 1-min music and video clips. The experiment started with a two-minute baseline recording, during which a fixation cross was displayed to the participant (who was asked to relax during this period). Then, the 40 videos were presented in 40 trials. After each trial, the participants were asked to do a self-assessment about their emotional levels, including four different scales, such as valence, arousal, dominance, and liking. Consequently, varied classification problems were yielded such as Low/High valence (LVHV), Low/High arousal.

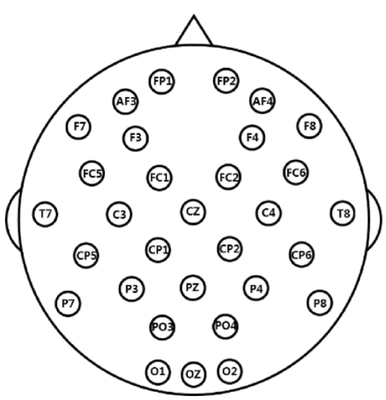



Fig. 4: Placement of EEG electrodes in DEAP

*2) SEED:* item The SEED dataset was collected from 15 subjects (7 males and 8 females) when they were asked to watch 15 film clips. The duration of each film clip was about 4 min. There were 15 trials for each subject and each trial consist of a $5s$ hint before each clip, $45s$ for self-assessment and $15s$ for rest after each clip in one session. EEG data in SEED dataset was collected from 62 electrodes (Figure 6), which includes more information than the DEAP dataset. In this dataset , negative, positive, and neutral are emotion labels that represent the subjects emotion states during each experiment. Label value of negative, positive and neutral is $-1$, $1$, and $0$, respectively.

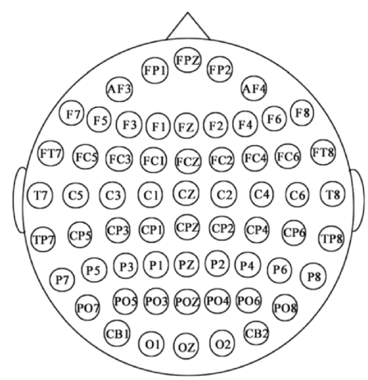



Fig. 5: Placement of EEG electrodes in SEED

*3) DASPS:* This paper uses the publicly available DASPS Dataset for anxiety level detection [17]. The database collected from the 23 participants, where 10 are men and 13 are women with the mean age of 30 years. The EEG signal is acquired using Emotiv EPOC, the wireless EEG headset with 14 channels and 2 mastoids with a 128Hz sampling rate. During the recording of the EEG signal, the study presents six different situations to the subjects with closed eyes and minimized gestures. It divides each situation into two parts: In the first 15 seconds, the psychotherapist recites the situation, and in the next 15 seconds, the subject recalls the situation. For each subject the complete process proceeds with three phases. The first phase includes 5 minutes of Hamilton Pre-Stimuli. In the second phase, the psychotherapist recites the situation for the first 15 seconds, and the subject recalls the situation in the next 15 seconds. The complete process repeats for each

**Algorithm 1** BMaOPSO Algorithm

**Input:** Population $P$ with size $N$, Trained ConvLSTM model
**Output:** External Archive $EA$

**for** $i = 1$ to $N$ **do**
  Initialization of particle $p_i$
  **for** $j = 1$ to $N$ **do**
    **if** $j == i$ **then**
      $X_i[j] = 1$
    **else**
      $X_i[j] = 0$
    **end if**
  **end for**
  Evaluate $f_j, j = 1, ..., k$
  Save non-dominated solutions of $P$ in the archive $EA$
**end for**
$K$ sub-swarm = Ranking $(P)$
**for** $i = 1$ to $K$ **do**
  **while** $g \leq G_{max}$ **do**
    Select leader
    Update velocity
    Update position
    Mutation
    $f_2, f_3, f_4$ =Test ConvLSTM network (selected channels)
    Update Fitness $(f_2, f_3, f_4)$
    **if** $I(X_i^t) \geq I(p_{besti}^{t-1})$ **then**
      $p_{Besti} = p_i$
    **end if**
    $g = g + 1$
    Save solutions in $LA_i$
    Learning strategy
  **end while**
  EA = EA $\cup$ $LA_i$
**end for**

of the six situations for all the subjects. At the end of the six situations, Self Assessment Manikin (SAM) is used to rate how a subject felt during stimulation.

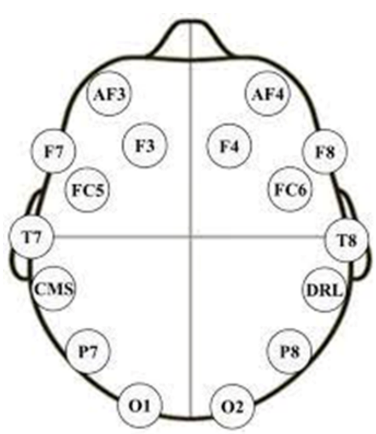



Fig. 6: Placement of EEG electrodes in DASPS

### B. Experiments on the DASPS Dataset

The classification results of DASPS dataset with RCAE and with optimization algorithms, are summarized in Table II. With the help of optimization, the numbers of channels are reduced and hence less number of the channel are required for the analysis. The classifier RCAE with full channel (14) provides 90.70% classification accuracy; whereas, BMOPSO, NSGAII and proposed BMaOPSO provides 89.01%, 89.64% and 90.25% accuracy respectively. Further, the classification accuracy obtained with the help of the proposed algorithm is a very close approximation with the accuracy of non-optimized RCAE. The BMOPSO provides 89.01% with 13 channels and NSGAII provides 89.64% with 11 channels, whereas the BMaOPSO gives 90.25% accuracy with only 9 channels.

TABLE II: DASPS: Accuracy, Sensitivity and Specificity for the selected EEG channels

| Method | #NB | Accuracy (%) | Sensitivity (%) | Specificity (%) |
|---|---|---|---|---|
| **RCAE** | **14** | **90.70** | **89.10** | **96.83** |
| BMOPSO | 13 | 89.01 | 80.06 | 92.80 |
| | 13 | 88.95 | 77.96 | 93.63 |
| | 12 | 88.95 | 81.22 | 92.26 |
| | 12 | 88.71 | 76.89 | 93.50 |
| | 12 | 88.68 | 79.46 | 91.50 |
| NSGAII | 11 | 89.64 | 78.87 | 93.47 |
| | 12 | 89.28 | 83.29 | 91.65 |
| | 10 | 87.80 | 79.48 | 92.45 |
| | 10 | 87.17 | 80.38 | 90.96 |
| | 9 | 86.50 | 76.00 | 91.27 |
| **BMaOPSO** | **9** | **90.25** | **80.39** | **93.47** |
| | 7 | 89.34 | 75.58 | 93.73 |
| | 10 | 89.10 | 79.82 | 91.99 |
| | 11 | 88.74 | 74.99 | 93.81 |
| | 8 | 88.29 | 75.26 | 92.34 |

*1) Interpretation of Dominant Channels in DASPS dataset:* The biological significance of the selected channels for each optimization algorithms are given in Table III.

Figure 7 indicates the selected optimal channels using BMOSPO, NSGAII and the proposed BMaOPSO. In addition, the distribution of solutions (the number of selected channels with the corresponding accuracy) are presented in Figure 8.

The optimization algorithms provide similar channels including 6 channels: AF3, P8, T8, FC6, F4, F8. This is an indication of the fact that the selected channels correspond to the brain region related to emotions. Accordingly, we find that the EEG channels related to emotions are distributed in the front: AF3 F4 and F8; the central: FC6; the temporal: T8; and the parietal: P8.

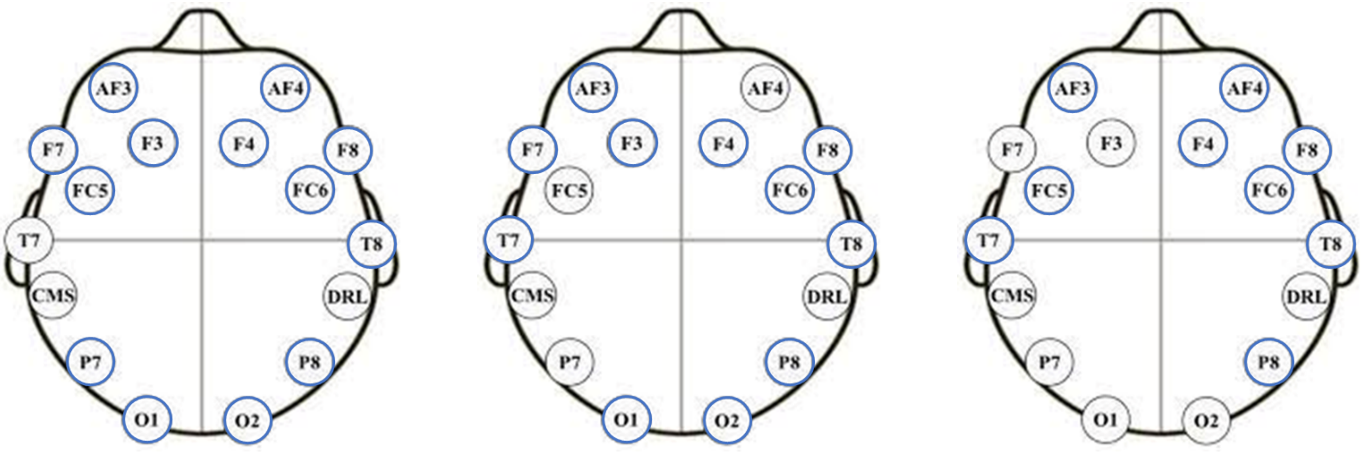
### C. Experiments on the DEAP Dataset

For DEAP database, we have verified the proposed method on valence and arousal labels, respectively. The Performances evaluation of channels selected with different algorithms algorithm are presented in Table IV. For each compared algorithm, we present the EEG channels combination for the first five optimal solution in the Pareto front.

For valence emotion, the RCAE achieves 69.09% classification accuracy with 32 channels. whereas BMOPSO, NSGAII and proposed BMaOPSO provides 67.82%, 68.90% and 69.47% accuracy respectively. It is observed that the best accuracy for valence and arousal is obtained with the proposed BMaOPSO.

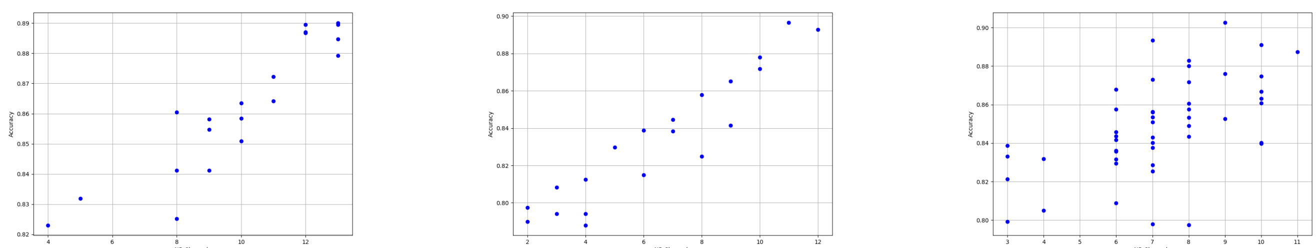
TABLE III: Qualitative comparative analysis for DASPS dataset

| Algorithm | No. | Accuracy | Sensitivity | Specificity | Name of channels | Optimization results | |
|---|---|---|---|---|---|---|---|
| | | | | | | Channels | locations |
| BMOPSO | 13 | 89.01 | 80.06 | 92.80 | AF3, F7, F3, FC5, P7, O1, O2, P8, T8, FC6, F4, F8, AF4 | AF3, P8, T8, FC6, F4, F8 | Frontal: AF3, F4, F8 |
| NSGAII | 11 | 89.64 | 78.87 | 93.47 | AF3, F7, F3, T7, O1, O2, P8, T8, FC6, F4, F8 | | Central: FC6 |
| BMaOPSO | 9 | 90.25 | 80.39 | 93.47 | AF3, FC5, T7, P8,T8, FC6,F4, F8, ’AF4’ | | Temporal: T8 |
| | | | | | | | Parietal: P8 |




(a) BMaOPSO (b) NSGAII (c) BMaOPSO

Fig. 7: DASPS: Channel Selection: (a) using BMOPSO (13 Channels) (b) using NSGAII (11 Channels) (C) using BMaOPSO (9 Channels)



(a) BMaOPSO (b) NSGAII (c) BMaOPSO

Fig. 8: DASPS dataset: Accuracy of the selected channels

*1) Interpretation of Dominant Channels in DEAP dataset:* The location of optimal EEG channels for valence and arousal emotions using the BMOPSO, NSGAII and the proposed method is illustrated in Figure 9 and 10, respectively. In fact, the recognition of valence and arousal emotion involves a different combination of EEG channels as detailed in Table V.

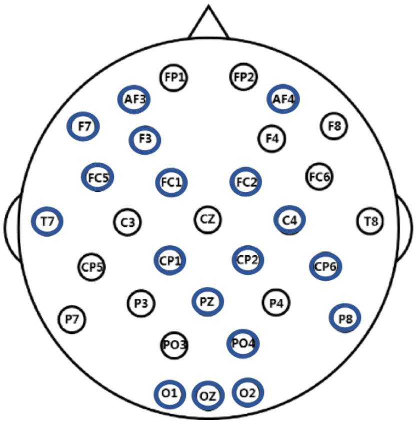
It is observed that the BMaOPSO selects 13 channels for valence: **AF3, F3, FC5, FC1, CP1, P3, PO3, O1, Oz, AF4, FC6, C4, P4** and 15 channels for arousal: **Fp1, AF3, F3, F7, FC5, FC1, P3, P7, PO3, Pz, Fp2, FC2, Cz, T8, P4** to obtain an accuracy which is higher than its competitors. Even that the obtained accuracy of arousal and valence are slightly higher than the classification accuracy of all channels, the channel number falls down to 1/2 of the full channels (32 channels), and it leads to reduce the computational complexity.

## D. Experiments on the SEED Dataset

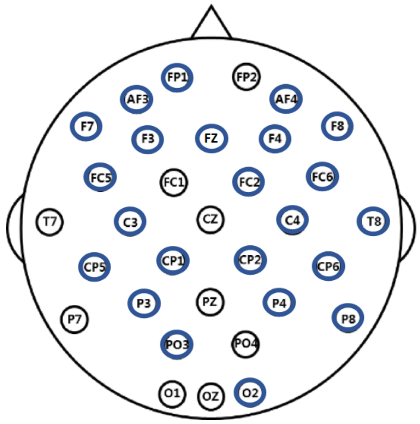
The generated results of the competitive algorithms, including BMOPSO, NSGAII and BMaOPSO over the SEED dataset are presented in Table VII. With this dataset that contain 62 channels, the classifier obtains $87.19\%$ accuracy. With the optimization algorithms, we achieves a higher accuracy with less channel number. The BMOPSO obtains $89.60\%$ with 42 channels; the NSGAII obtains $88.75\%$ with 39 channels; the BMaOPSO obtains $89.04\%$ with 40 channels. By using the proposed method, the accuracy is slightly lower than the BMOPSO algorithm but with lower channel number.

*1) Interpretation of Dominant Channels in SEED dataset:* The significance of selected channels using optimization algorithms is presented in Table VI. In addition, the Figure 13 indicates the distribution of selected optimal channels over the different region of brain.

It is observed that the number of EEG channels can be reduced from 62 to 42, 39 and 42 channel by using BMOPSO,




(a) BMaOPSO




(b) NSGAII

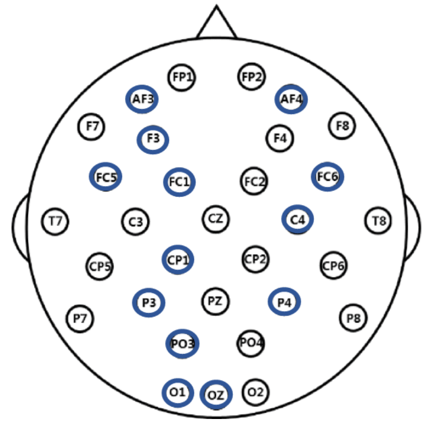
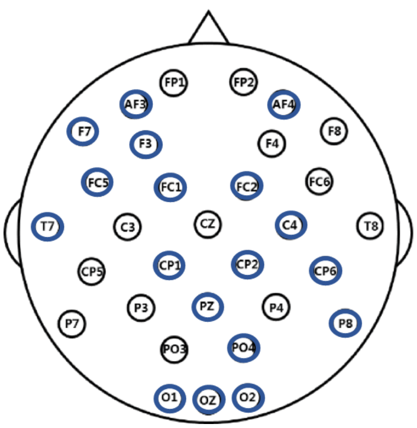
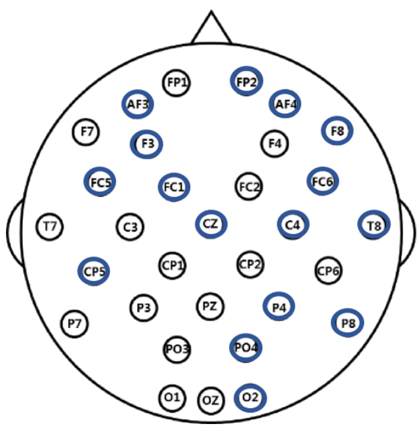


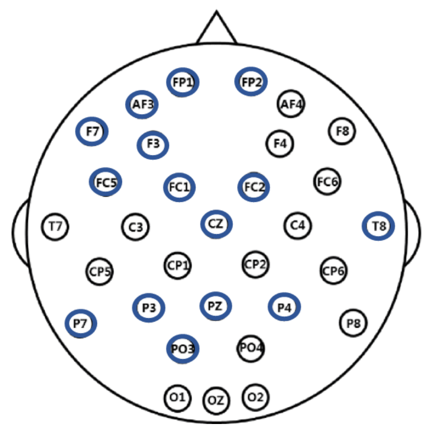

Fig. 9: DEAP Valence: Channel Selection: (a) using BMOPSO (16 Channels) (b) using NSGAII (23 Channels) (C) using BMaOPSO (13 Channels)




(a) BMaOPSO

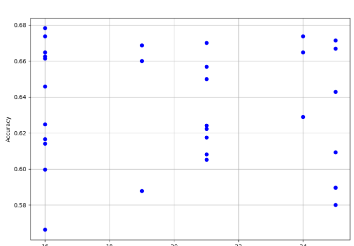



(b) NSGAII

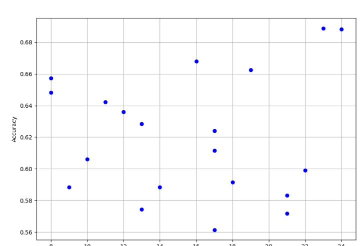



(c) BMaOPSO

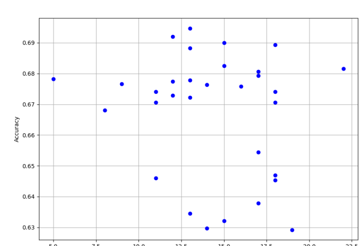
Fig. 10: DEAP Arousal: Channel Selection: (a) using BMOPSO (18 Channels) (b) using NSGAII (16 Channels) (C) using BMaOPSO (15 Channels)




(a) BMaOPSO

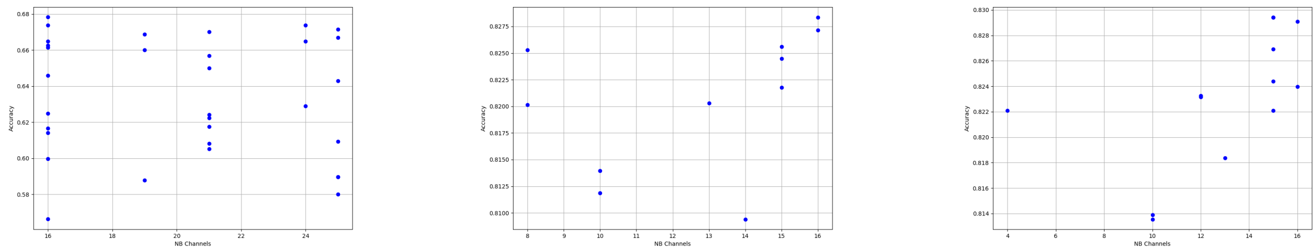



(b) NSGAII



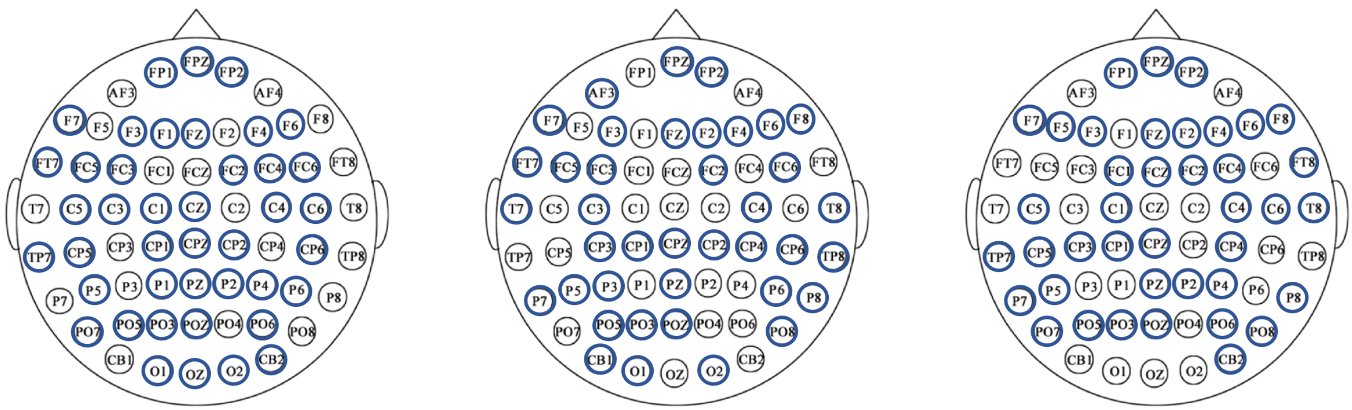

(c) BMaOPSO

Fig. 11: DAEP Valence: Accuracy of the selected channels




(a) BMaOPSO (b) NSGAII (c) BMaOPSO

Fig. 12: DAEP Arousal: Accuracy of the selected channels



(a) BMOPSO (b) NSGAII (c) BMaOPSO

Fig. 13: SEED: Channel Selection: (a) using BMOPSO (42 Channels) (b) using NSGAII (39 Channels) (C) using BMaOPSO (40 Channels)

TABLE IV: DEAP: Accuracy, Sensitivity and Specificity for the selected EEG channels

| Label | Method | #NB | Accuracy (%) | Sensitivity (%) | Specificity(%) |
|---|---|---|---|---|---|
| Valence | **RCAE** | **32** | **69.09** | **75.02** | **61.36** |
| | BMOPSO | 16 | 67.82 | 74.94 | 58.50 |
| | | 24 | 67.38 | 77.72 | 53.95 |
| | | 16 | 67.37 | 75.53 | 56.53 |
| | | 25 | 67.15 | 75.50 | 56.00 |
| | | 21 | 67.02 | 76.58 | 54.46 |
| | NSGAII | 23 | 68.90 | 76.22 | 59.06 |
| | | 24 | 68.82 | 74.50 | 61.30 |
| | | 16 | 66.79 | 71.57 | 60.59 |
| | | 19 | 66.25 | 79.01 | 49.83 |
| | | 8 | 65.73 | 74.59 | 54.10 |
| | **BMaOPSO** | **13** | **69.47** | **74.11** | **63.45** |
| | | 12 | 69.20 | 74.68 | 62.18 |
| | | 15 | 68.99 | 74.34 | 61.98 |
| | | 5 | 67.82 | 72.95 | 61.09 |
| | | 9 | 67.67 | 75.75 | 57.26 |
| Arousal | **RCAE** | **32** | **82.58%** | **80.66%** | **82.58%** |
| | BMOPSO | 18 | 82.44 | 85.79 | 80.57 |
| | | 20 | 82.39 | 88.22 | 79.09 |
| | | 23 | 82.35 | .85.50 | 80.58 |
| | | 16 | 82.32 | 86.94 | 79.77 |
| | | 16 | 82.29 | 87.80 | 79.22 |
| | NSGAII | 16 | 82.84 | 86.38 | 80.81 |
| | | 16 | 82.71 | 83.24 | 82.41 |
| | | 15 | 82.56 | 87.37 | 79.80 |
| | | 8 | 82.53 | 89.17 | 78.73 |
| | | 15 | 82.45 | 82.20 | 82.59 |
| | **BMaOPSO** | **15** | **82.94** | **83.86** | **82.43** |
| | | 16 | 82.91 | 86.55 | 80.93 |
| | | 15 | 82.69 | .83.57 | 82.21 |
| | | 15 | 82.44 | 84.24 | 81.45 |
| | | 16 | 82.40 | 84.62 | 81.20 |

NSGAII and BMaOPSO, respectively.

The results indicate that the compared algorithms selected similar channels: **FPZ, FP2, F7, F3, FZ, F4, F6, FC2 C4, CP1, CPZ, P5, PZ, PO5, PO3, POZ** that distributed over the frontal, central, parietal and occiptal regions. In fact, the selected channels related to the region which is most affected by positive, neutral, and negative emotions.

However, the proposed method ensures a higher accuracy with fewer channels.The fact that each channel has a large amount of data, the reducing of channel number can effectively reduce the computational complexity. After the channel selection, only the channels that related to emotions were retained, leading to reduce the amount of data to analyze.

## VI. Conclusion and Future Work

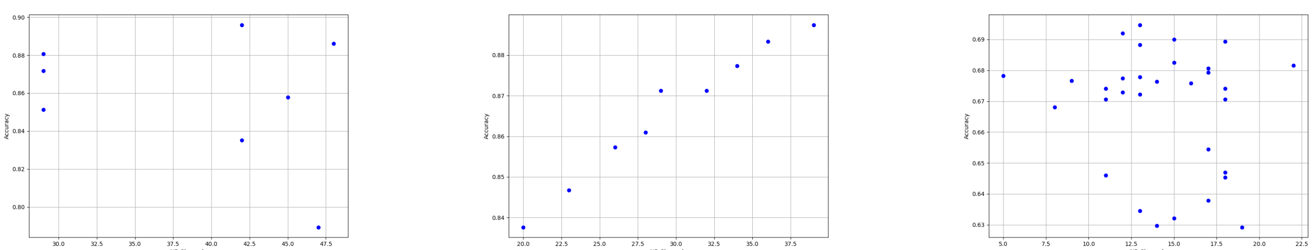
In this paper a BMaOPSO has been proposed for the emotion recognition using a new learning model denoted as RCAE: Recurrent Convolutional Autoencode. The proposed algorithm provides the required channels to detect the emotion. The BMaOPSO is compared with traditional algorithms as BMOPSO and NSGAII algorithms over three well known datasets: DASPS, DEAP and SEED. Experimentally, the proposed algorithm improves its efficiency to detect the emotion with fewer channel number.

TABLE V: Qualitative comparative analysis for DEAP dataset

| | Algorithm | No. | Accuracy | Sensitivity | Specitivity | Name of channels | Optimization results | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | Channels | Locations |
| Valence | BMOPSO | 16 | 67.82 | 74.94 | 58.50 | Fp1, FC5, T7, CP5, CP1, P7, O1, Pz, Fp2, AF4, FC2, C4, T8, CP6, PO4, O2 | AF3, F3 FC5, CP1 AF4, C4 | Frontal: AF3, F3, FC5, AF4 |
| | NSGAII | 23 | 68.90 | 76.22 | 59.06 | Fp1, AF3, F3, F7, FC5, C3, CP5, CP1, P3, PO3, AF4, Fz, F4, F8, FC6, FC2, C4, T8, CP6, CP2, P4, P8, O2 | | Central: CP1, C4 |
| | **BMaOPSO** | **13** | **69.47** | **74.11** | **63.45** | **AF3, F3, FC5, FC1, CP1, P3, PO3, O1, Oz, AF4, FC6, C4, P4** | | |
| Arousal | BMOPSO | 18 | 82.44 | 85.79 | 80.57 | AF3, F3, F7, FC5, FC1, T7, CP1, O1, Oz, Pz, AF4, FC2, C4, CP6, CP2, P8, PO4, O2 | AF3, F3 FC5, FC1 | Frontal: AF3, F3, FC5, FC1 |
| | NSGAII | 16 | 82.84 | 86.38 | 80.81 | AF3, F3, FC5, FC1, CP5, Fp2, AF4, F8, FC6, Cz, C4, T8, P4, P8, PO4, O2 | | |
| | **BMaOPSO** | **15** | **82.94** | **83.86** | **82.43** | **Fp1, AF3, F3, F7, FC5, FC1, P3, P7, PO3, Pz, Fp2, FC2, Cz, T8, P4** | | |

TABLE VI: Qualitative comparative analysis for SEED dataset

| Algorithm | No. | Accuracy | Sensitivity | Specitivity | Name of channels | Optimization results | |
|---|---|---|---|---|---|---|---|
| | | | | | | channels | locations |
| BMOPSO | 42 | 89.60 | 85.34 | 92.82 | FP1, FPZ, FP2, F7, F3, F1, FZ, F4, F6, FT7, FC5, FC3, FC2, FC4, FC6, C5, C3, C1, CZ, C4, C6, TP7, CP5, CP1, CPZ, CP2, CP6, P5, P1, PZ, P2, P4, P6, PO7, PO5, PO3,POZ,PO6, O1, OZ,O2,CB2 | FPZ, FP2, F7, F3, FZ, F4, F6, FC2 C4, CP1, CPZ, P5, PZ, PO5, PO3, POZ | Frontal: FPZ, FP2,F7,F3 FZ, F4, F6 |
| NSGAII | 39 | 88.75 | 84.32 | 92.21 | FPZ, FP2, AF3, F7, F3, FZ, F2, F4, F6, F8, FT7, FC5, FC3, FC2, FC6, T7, C3, C4, T8, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, P7, P5, P3, PZ, P6, P8, PO5, PO3, POZ, PO8, CB1, O1, O2 | | Central: FC2, C4 |
| **BMaOPSO** | **40** | **89.04** | **84.52** | **92.53** | **FP1, FPZ, FP2, F7, F5, F3, FZ, F2, F4, F6, F8, FC1, FCZ, FC2, FC4, FT8, C5, C1, C4, C6, T8, TP7, CP5, CP3, CP1, CPZ, CP4, P7, P5, PZ, P2, P4, P8, PO7, PO5, PO3, POZ, PO6, PO8, CB2** | | Parietal: CP1, CPZ P5, PZ |
| | | | | | | | Occipital: PO5, PO3, POZ |




(a) BMOPSO (b) NSGAII (c) BMaOPSO

Fig. 14: SEED: Accuracy of the selected channels

TABLE VII: SEED: Accuracy, Sensitivity and Specificity for the selected EEG channels

| Method | #NB | Accuracy (%) | Sensitivity(%) | Specificity(%) |
|---|---|---|---|---|
| **RCAE** | **64** | **87.19** | **81.27** | **94.23** |
| **BMOPSO** | **42** | **89.60** | **85.34** | **92.82** |
| | 48 | 88.61 | 84.20 | 92.41 |
| | 29 | 88.06 | 82.94 | 91.84 |
| | 29 | 87.17 | 81.96 | 91.54 |
| | 45 | 85.78 | 80.56 | 90.68 |
| NSGAII | 39 | 88.75 | 84.32 | 92.21 |
| | 36 | 88.34 | 83.51 | 92.08 |
| | 34 | 87.74 | 82.87 | 91.91 |
| | 29 | 87.13 | 81.97 | 91.39 |
| | 32 | 87.12 | 82.04 | 91.45 |
| BMaOPSO | 40 | 89.04 | 84.52 | 92.53 |
| | 31 | 88.93 | 84.34 | 92.46 |
| | 30 | 88.71 | 84.18 | 92.30 |
| | 40 | 88.42 | 83.56 | 92.11 |
| | 30 | 88.42 | 83.56 | 92.11 |

For future work, the proposed algorithm can be investigated to improve the epileptic seizure detection by using a new method to enhance the search process as filtering method.

## References


[1] M. Baig, N. Aslam, and H. Shum, "Filtering techniques for channel selection in motor imagery eeg applications: a survey," *Artificial Intelligence Review*, vol. 53, p. 1207–1232, 2020.

[2] A. Bhattacharyya and R. B. Pachori, "A multivariate approach for patient-specific eeg seizure detection using empirical wavelet transform," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2003–2015, 2017.

[3] F. Qi, W. Wu, Z. L. Yu, Z. Gu, Z. Wen, T. Yu, and Y. Li, "Spatiotemporal-filtering-based channel selection for single-trial eeg classification," *IEEE Transactions on Cybernetics*, pp. 1–10, 2020.

[4] Z. Wang, S. Hu, and H. Song, "Channel selection method for eeg emotion recognition using normalized mutual information," *IEEE Access*, vol. 7, pp. 143 303–143 311, 2019.

[5] L. A. Moctezuma and M. Molinas, "Eeg channel-selection method for epileptic-seizure classification based on multi-objective optimization," *Frontiers in Neuroscience*, vol. 14, p. 593, 2020. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnins.2020.00593

[6] ——, "Multi-objective optimization for eeg channel selection and accurate intruder detection in an eeg-based subject identification system," *Scientific Reports*, vol. 10, 2020.

[7] ——, "Towards a minimal eeg channel array for a biometric system using resting-state and a genetic algorithm for channel selection," *Scientific Reports*, vol. 10, 2020.

[8] G. C. Jana, A. Tripathi, and A. Agrawal, "Eeg channel selection approach for seizure detection based on integrated bpso and elm," in

*2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2020, pp. 856–861.

[9] A. Bablani, D. R. Edla, D. Tripathi, S. Dodia, and S. Chintala, "A synergistic concealed information test with novel approach for eeg channel selection and svm parameter optimization," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 11, pp. 3057–3068, 2019.

[10] N. K. Al-Qazzaz, M. K. Sabir, S. Ali, S. A. Ahmad, and K. Grammer, "Effective eeg channels for emotion identification over the brain regions using differential evolution algorithm," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 4703–4706.

[11] S. Bavkar, B. Iyer, and S. Deosarkar, "Rapid screening of alcoholism: An eeg based optimal channel selection approach," *IEEE Access*, vol. 7, pp. 99 670–99 682, 2019.

[12] E. Yildirim, Y. Kaya, and F. Kiliç, "A channel selection method for emotion recognition from eeg based on swarm-intelligence algorithms," *IEEE Access*, vol. 9, pp. 109 889–109 902, 2021.

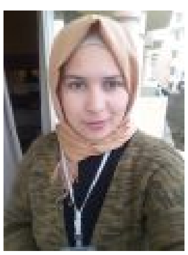
[13] J. Kennedy and R. C. Eberhart, "A discrete binary version of the particle swarm algorithm," in *1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*, vol. 5, 1997, pp. 4104–4108 vol.5.

[14] X.-F. Liu, Z.-H. Zhan, Y. Gao, J. Zhang, S. Kwong, and J. Zhang, "Coevolutionary particle swarm optimization with bottleneck objective learning strategy for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 4, pp. 587–602, 2019.

[15] B. Wei, W. Zhang, X. Xia, Y. Zhang, F. Yu, and Z. Zhu, "Efficient feature selection algorithm based on particle swarm optimization with learning memory," *IEEE Access*, vol. 7, pp. 166 066–166 078, 2019.

[16] S. Koelstra, C. Muhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis ;using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.

[17] A. Baghdadi, Y. Aribi, R. Fourati, N. Halouani, P. Siarry, and A. M. Alimi, "Dasps database," 2021. [Online]. Available: https://dx.doi.org/10.21227/barx-we60



**Najwa Kouka** received the degree in computer science engineering from Engineering Higher Institute of Computer Science and Multimedia of Sfax, Tunisia, in 2015. She is currently pursuing the Ph.D. degree in computer systems engineering with the Research Group in Intelligent Machines, National Engineering School of Sfax, Tunisia. She is currently IEEE student member since October 2015. Her current research interests include multi-objective optimization, many-objective optimization, swarm intelligence, particle swarm optimization, multi-agent system, automated negotiation.

**Rahma Fourati** (IEEE Graduated Student Member'15, Member'15). She was born in Sfax, Tunisia. She received the M.S. degree in 2011 from the Faculty of Economic Sciences and Management of Sfax (FSEGS) and she graduated from the National Engineering School of Sfax (ENIS) with a degree in Computer Engineering. She is currently pursuing the Ph.D. degree in Computer Science Engineering at the ENIS, since October 2015 with the REsearch Group in Intelligent Machines (REGIM). Her research interests include Emotion recognition, Anxiety detection, Affective computing, EEG signals analysis, Recurrent Neural Networks, Support Vector Machines, Simulink Modeling.

**Raja Fdhila** received the engineering degree from the National School of Engineers of Tunisia,in 2005 and the Ph.D degree in computer systems engineering from the National School of Engineers of Sfax,Tunisia, in 2017. She is currently a member of the REGIM-lab: REsearch Group on Intelligent Machines at the National school of engineers of Sfax (ENIS). Her research interests include artificial intelligence, multi-objective optimization, particle swarm optimization and evolutionary algorithms. She has been an IEEE member since 2008.

**Patrick Siarry** (SM'95) is currently a Professor and a Doctoral Supervisor of Automatics and Informatics with the University of Paris-Est Créteil, Créteil, France. He is the Head of one team of the Laboratoire Images, Signaux et Systèmes Intelligents, University of Paris-Est Créteil. He supervises a Research Working Group, in the frame of Centre National de la Recherche Scientifique, META, since 2004. META is interested in theoretical and practical advances in metaheuristics for hard optimization. His current research interests include improvement of existing metaheuristics for hard optimization, adaptation of discrete metaheuristics to continuous optimization, hybridization of metaheuristics with classical optimization techniques, dynamic optimization, and particle swarm optimization.

**Adel M. Alimi** (IEEE Student Member'91, Member'96, Senior Member'00). He was born in Sfax (Tunisia) in 1966. He graduated in Electrical Engineering 1990, obtained a Ph.D. and then an HDR both in Electrical and Computer Engineering in 1995 and 2000, respectively. He is now a Professor in Electrical and Computer Engineering at the University of Sfax. His research interest includes applications of intelligent methods (neural networks, fuzzy logic, evolutionary algorithms) to pattern recognition, robotic systems, vision systems, and industrial processes. He focuses his research on intelligent pattern recognition, learning, analysis and intelligent control of large scale complex systems. He is an Associate Editor and Member of the editorial board of many international scientific journals (e.g. Pattern Recognition Letters, Neurocomputing, Neural Processing Letters, International Journal of Image and Graphics, Neural Computing and Applications, International Journal of Robotics and Automation, International Journal of Systems Science, etc.).