ARTICLE OPEN ACCESS

OPTIMIZING SUPPORT VECTOR MACHINE USING MODIFIED CLONAL SELECTION FOR BRAIN COMPUTER INTERFACE

Padmavath[1*] and Ranganathan [2]

[1]*Dept. of Electronics and Communication Engineering, Dhanalakshmi Srinivasan College of Engineering and echnology, Mamallapuram, Chennai., TN, INDIA*

[2]*Dept. of Electrical and Instrumentation, Dr. Mahalingam College of Engineering & Technology, Pollachi, TN, INDIA*

ABSTRACT

Brain Computer Interface's (BCI) central element, is a translation algorithm converting electrophysiological input from user into output capable of controlling external devices.Many studies over the past two decades have shown that people and animals can use brain signals to convey their intent to a computer using BCIs. This is possible through use of sensors that capture signals in the brain, corresponding to certain thought forms. The kernel parameters setting for SVM in training process impacts on the classification accuracy. The Modified Clonal Selection Algorithm (CLONALG) is one such system inspired by the clonal selection theory of acquired immunity, which has shown success on broad range of engineering problem domains.

Published on: 10th– August-2016

KEY WORDS

Brain Computer Interfaces (BCIs), ElectroCorticoGraphy (ECoG), Support Vector Machine (SVM), Clonal Selection Algorithm (CLONALG).

***Corresponding author: Email:** ssmadhu80@gmail.com, profsks@rediffmail.com

INTRODUCTION

A BCI is a hardware that allows humans to interact with a computer through brainwaves. Presently, BCI is used for healthcare, and education based on neural-feedback which is a type of brainwave using bio-feedback. BCI controls computers using human brain waves. BCIs convert brain signals into outputs communicating user's intent [1]. As this new communication channel is independent of peripheral nerves and muscles, it is resorted to by those with severe motor disabilities.

To achieve this, a BCI system consists of four sequential components are signal acquisition, feature extraction, feature translation, and device output. These four components are controlled by an operating protocol that defines the onset and timing of operation, the details of signal processing, the nature of the device commands, and the oversight of performance. An effective operating protocol allows a BCI system to be flexible and to serve the specific needs of each user.

BCIs use invasive and non-invasive methods. ElectroEncephaloGraphic activity (EEG) [2] from the scalp is used by non-invasive BCIs. Though convenient, safe and inexpensive, they are susceptible to artifacts like electromyography (EMG) signals, which have low spatial resolution and so need much user training. Single-neuron activity recorded in the brain is used by invasive BCIs. Though having higher spatial resolution and providing control signals with much freedom, BCIs still are dependent on electrodes in the cortex and so have problems ensuring stable long-term recordings.

ECoG is acute recording of electrical activity directly from cortical surface during exposure in surgical treatment of epilepsy [3, 4]. Recent studies emphasized the intraoperative ECoG importance for precise epileptic focus localization and good surgical outcome. ECoG is not invasive, as neuronal recordings as the brain is not entered into. It has a higher Signal-to-Noise Ratio (SNR) than EEG and also higher spectral/spatial resolution [5] which necessitates re-engineering of signal processing and classification techniques found in conventional EEG-based

THE IIOAB JOURNAL ISSN: 0976-3104

BCIs. Extreme data scarcity due to limited time available for volunteering patients is an obstacle to characterize information in ECoG signals.

A huge challenge in designing BCI is the selection of relevant features from a huge set of potential features. High dimensional features vectors are not good because of the curse of dimensionality in trainingclassification protocols. Features selection may be carried out studying all potential subsets of features. But the quantity of possibilities increases in an exponential fashion, making extensive searches impractical for even moderate quantities of features. Certain more effective optimization protocols may be employed with the objective of decreasing quantity of features and at the same time improving classification performance [6].

The typical CLONALG model implies the choosing of antibodies (candidate solutions) on the basis of affinity either by matching antigen patterns or through evaluation of patterns by cost functions. Chosen antibodies are vulnerable to cloning proportional to affinity, and hyper-mutation of clones inversely proportional to clone affinity. Resulting clonal-set competes with antibody populations for membership in subsequent generations and finally, low-affinity population members are substituted by arbitrarily created antibodies [7]. CLONALG is the abbreviation of the clonal algorithm and has been inspired by the following elements of the clonal selection theory [8]

Maintenance of a specific memory set

- Selection and cloning of most stimulated antibodies
- Death of non-stimulated antibodies
- Affinity maturation (mutation)
- Re-selection of clones proportional to affinity with antigen
- Generation and maintenance of diversity

A mixture kernel function based on radial based and polynomial kernel was introduced and the parameters of this new kernel function were optimized. Their algorithm gives the better results than normal SVM in fault diagnosis. But it has some disadvantages. Firstly, their immune optimization method refers to crossover parameter. But original immune optimization algorithm (CLONALG) has not crossover operator. A clonal selection algorithm whose name is determined as CLONALG by them. The objective function of the immune optimization method is calculated based on training of SVM [9]. Because of these properties, clonal selection converges faster than genetic algorithm and does not catch local minimum.

Hybrid algorithms [10] are the combination of exact algorithms and Meta-heuristics. In the scientific community, the term “meta-heuristic” refer to general purpose approximated optimization methods, such as Tabu search, evolutionary computation, and simulated annealing, among others. A general classification of meta-heuristic algorithms grouped into two categories: Collaborative combinations: In an environment of collaboration, the algorithms exchange information, but are independent. The exact and meta-heuristic algorithms may be executed sequentially, in parallel or intertwined. Integrated combinations: In integrated methods, an algorithm is a subordinated component of another algorithm. In the integrated combinations category, there are two subcategories: the meta-heuristic algorithm is the master and controls the calls to the exact algorithm and the exact algorithm is the master and calls the meta-heuristic algorithm.

In this paper, proposed the optimized SVM using modified CLONALG. Section 2 deals with literature related to this work, section 3 describes the methods used in the work, section 4 deals with results and discusses obtained results and finally section 5 concludes the work.

RELATED WORKS

BCI using Electroencephalogram signal was discussed by Shende&Jabade [11]. If BCI aims to control robotics machinery with better accuracy. The use of BCI is to control the wheelchair/robotic limb movement for a disable person and to access security systems. BCI can discover emotions which is can control the surrounding environment like controlling the interior of a car/house.

A low-cost non-invasive BCI hybridized with eye tracking described by Kim et al., [12] also discussed its feasibility through a Fitts' law-based quantitative evaluation method. Non-invasive BCI received a lot of attention recently. BCI applications in real life need to be user-friendly and easily portable. In the new work, an approach to

THE IIOAB JOURNAL
ISSN: 0976-3104

realize a real-world BCI, EEG-based BCI combined with eye tracking was investigated. The new hybrid BCI system was discussed regarding a practical interface scheme. Though further advancement was required, the new hybrid BCI system had the potential to be useful in a natural/intuitive manner.

A novel driver-vehicle interface for the individuals with severe neuromuscular disabilities to use intelligent vehicles through P300 and steady-state visual evoked potential (SSVEP) BCIs to select destination and test its performance in the lab and in real driving conditions was proposed by Fan et al., [13]. The new interface has 2 components: a selection component based on a P300 BCI and a confirmation component based on an SSVEPBCI. The proposed system improved destination selection accuracy compared to a single P300 BCI-based selection system, specifically for participants with relatively low accuracy in using P300 BCI.

Krusienski et al., [14] presented a preliminary analysis of the relationship between EEG and ECoG event-related potentials (ERPs) recorded from a single patient using a BCI speller. The patient carried out one experimental session through usage of BCI spelling paradigms controlled by scalp-recorded EEG before ECoG grid implantations as well as one identical session controlled by ECoG post grid implantations. The patient was capable of achieving near perfect spelling precision through EEG as well as ECoG. An offline analysis of the average ERPs was performed to assess how accurately the average EEG ERPs could be predicted from the ECoG data. The preliminary results indicated that EEG ERPs can be accurately estimated from proximal asynchronous ECoG data using simple linear spatial models.

The robust nature of Least-Square Support Vector Machines (LS-SVMs) for classifying multi-class self-paced MI temporal features while tuning hyper parameters automatically was investigated by Hamedi et al., [15]. MI EEG signals were pre-processed/segmented into non-overlapped distinctive time slots. The new method was evaluated/compared to three classifiers. Results indicated LS-SVM's high potential to classify different MI's by getting average 89.88±8.00 classification accuracy when using Sign Slop Changes (SSC) features.

A multi-ganglion Artificial Neural Network (ANN) based Feature Learning (ANNFL) method to extract the deep feature structure of a single-trial multi-channel ERP signals and improve classification accuracy proposed by Gao et al., [16] extracted feature vectors and classified them using SVM. The method outperformed a PCA and conventional three-layer auto-encoder leading to higher classification accuracies in five subjects' BCI signals than with using single-channel temporal features. ANNFL is an unsupervised feature learning method, which automatically learns feature vector from EEG data providing more effective feature representation than the PCA and single-channel temporal feature extraction methods.

Parallel multi-objective optimization methods for coping with high-dimensional features selection problems were proposed by Kimovski et al., [17]. Many parallel multi-objective evolutionary alternatives were proposed and evaluated using synthetic/BCI benchmarks. Results showed that cooperation of parallel evolving subpopulations improved solution quality and computing time speedups based on a parallel alternative and data profile.

A neural classifier optimized using the Backtracking Search optimization Algorithm (BSANN) to classify 3 mental tasks consisting of a right or left hand movement imagination and word generation was presented by Agarwal et al., [18]. The new method BSANN was tested on a publicly available BCI Competition 3-5 datasets. Result showed that BSANN exhibited better results than 21 other algorithms for mental tasks classification regarding classification accuracy.

Ding [19] proposed a new strategy combining with the SVM classifier for features selection that retains sufficient information for classification purpose. For improving classification precision, the variables optimization of the penalty constant C as well as the bandwidth of Radial Basis Function (RBF) kernels is a significant step in the establishment of effective as well as high-performance SVM model. Aiming at optimizing the parameters of SVM, also presented a grid based Ant Colony Optimization (ACO) algorithm to choose parameters C and Â automatically for SVM instead of selecting parameters randomly by human's experience and traditional grid searching algorithm, so that the classification feature numbers can be reduced and the classification performance may be enhanced concurrently. Experiments prove the feasibility as well as efficacy of the method.

Gonzalez et al., [20] proposed a method for classifying single-trial ERPs using a combination of the Lifting Wavelet Transform (LWT), SVM and Particle Swarm Optimization (PSO). In particular, the LWT filters, the set of EEG channels and SVM parameters that maximize the classification accuracy are searched using PSO. The

THE IIOAB JOURNAL
ISSN: 0976-3104

authors evaluated the method's performance through offline analysed on the datasets from the BCI Competitions II and III. The proposed method achieved in most cases a similar or higher classification accuracy than that achieved by other methods, and adapted wavelet basis functions and channel sets that match the time-frequency and spatial properties of the P300 ERP.

Wang et al., [21] used GA-SVM hybrid algorithm with two purposes: Selecting of the optimal feature subset and deciding the parameters for SVM classifier after the features extracted though the algorithm called Sample Entropy. Compared with GA-based feature selection and GA-based parameters optimization for SVM, the GA-SVM hybrid algorithm has fewer input features and gain much higher classification accuracy.

Rathipriya et al., [22] suggested a hybrid algorithm to advance the classification achievement rate of MI-based ECoG in BCIs. To verify the effectiveness of the suggested classifier, the authors restored the SVM classifier with the identical features extracted from the cross-correlation method for the classification. The performances of those procedures are assessed with classification correctness through a 10-fold cross-validation procedure. The authors furthermore consider the performance of the suggested procedure by comparing it with existing system.

METHODOLOGY

In this section, GA and Modified CLONALG are described. Genetic algorithm is a heuristic approach for resolving optimization issues. Currently, the approach is utilized in several research areas, but it initiated in the genetic sciences. Therefore, most terms which are utilized for describing the optimization problems are inherited directly from biological terms. To run the optimization process with a GA, first, the problem environment has to be defined, that is, a way of encoding problem solutions to the form of genetic algorithm individuals, fitness functions which are utilized for evaluation of individuals in all generations, genetic operations which are utilized for mixing as well as modifying individuals, an approach for choosing individuals as well as other extra genetic algorithm variables. Once problem environment is defined, the selected GA is applied to process individuals by a given number of generations.

Genetic Algorithm (GA)

The general scheme of the classic GA, i.e. Holland algorithm from 1975, may be delineated as given below. In the initial stages of the protocol, a set of arbitrarily selected individuals, each coding one solution of the problem at hand, is created. The individuals are ranked as per selected criteria, in the form of fitness functions. Later, solutions of small values of fitness functions are discarded from the set of solutions. They are replaced by new solutions, which are created by combining together parts of solutions of high fitness (crossover stage). From time to time random alterations are made in the existing solutions. The alterations permit exploration of completely new regions of the problem space. The whole procedure is iterated till adequate solutions are discovered.

Different GAs can be used for feature selection. From all the protocols, the typically utilized one is the protocol which necessitates coding of all extricated features. As per this method, all genes of an individual relate to a single feature and contains the data as to whether the feature exists in the specified individual or not. The protocol is compatible with the traditional Holland algorithm which implies that it begins from arbitrary population and it employs one-gene mutation as well as one-point crossover. The quality of individuals generated in consequent iterations is appraised as per the accuracy of classifiers created separately for all individuals [23].

Because of arbitrary selection of genes to individuals of the initial population, all individuals contain around half of all potential features (assuming uniform distribution). In the case of spaces comprised of features extricated from raw EEG signals (comprised generally of a minimum of 102-103 features), initiating a procedure of looking for the optimum set of features from the middle of the set is not a profitable solution as it can disable considerable decrease of features from the set. This is because of direction of optimization procedure to increase classification precision that favors individual of greater accuracy, i.e. individuals that code solutions generating classifiers of higher number of free parameters (and so, higher number of features).

Theoretically, the optimization process has not to be guided purely by the classifier results. It is possible, for instance, to equip genetic algorithm fitness function with penalty term that penalizes individuals coding too huge quantity of features. It is possible to develop certain specialized genetic operators converting these unwelcome individuals to individuals carrying smaller number of features. In practice, however, the scale of the required reduction of the feature set is so large that it is extremely difficult to develop the stable function penalizing individuals which carry several features or functions to convert the individuals. Improved solution is to run genetic algorithm with individuals of restricted quantity of features, coding merely small subsets of the entire features set.

For applying the solution, certain alterations are to be made in the genotype. Firstly, it is not necessary to stick to binary coding; a better solution is to utilize integer genes. Second, all genes ought to encode index of a single feature from an entire set of features. With this method, a single individual comprises indexes of features which are to be delivered to classifier inputs. The adequate quantity of features (that is, the quantity of genes contained in a single individual) is set by user prior to launch of the protocol, with respect to the quantity of recorded observations as well as applied classifiers.

THE IIOAB JOURNAL
ISSN: 0976-3104

When these coding methods are employed, individuals with two or more equal genes may occur because of genetic operations or because of arbitrary selection of initial population. In some applications, e.g. in the travelling salesman problem, such an individual indicating a double visit in one city would be eliminated as an incorrect one. But in the case of features selection issue, guided by the classification accuracy, such an individual is not regarded as defective – rather, it may even be required. Repair is necessary as several usages of the same feature in the classifier do not make sense, but the repair involves discarding all but one of the genes coding the feature. Such individuals are desirable because if the precision of the classifier utilizing features coded in the individual was considerably high to permit the individual surviving the selection procedure, it would denote that further decrease in the quantity of features is possible.

The genetic algorithm controls a population of potential solutions to problems. The solutions are coded as binary chains. The group of chains represent the genetic material of a set of individuals. Artificial operators of selection, crossovers as well as mutations are employed in a stochastic search procedure for finding best individual through simulation of natural evolutionary procedure. All candidate solutions are linked to fitness values that measure the excellence of solutions. Hence, the fitness simulates environmental pressure of Darwin's natural evolution. A simplified GA pseudo-code structure [24]:

1. *Initialization of population*
2. *Evaluation of population*
3. *While Better fitness < Fitness Required do*

Selection of parents

Crosses and mutations

Evaluation of population

End While

GA is adaptive heuristic search protocol on the basis of evolutionary concepts of natural selection as well as genetics. The fundamental notion of GA is that it is formulated to mimic procedures in natural systems necessary for evolution. The main operator of GA to search in pool of possible solutions is Crossover, Mutation and selection.

The genetic search process is iterative: evaluating, selection and recombining string in the population during each one of iterations (generation) until reaching some termination condition. Evaluating all strings is based on fitness functions that are problem dependent. It defines which of the potential solutions are better. It corresponds to environmental determination of survivability in natural selection.

Selection of a string, which represents a point in the search space, depends on the string's fitness relative to those of other strings in the population, those points that have relatively low fitness [25].

Mutation, like in natural environments, is a low probability operator and merely flips bit. The objective of mutation is the introduction of new genetic material into an existing individual; that is, to add diversity to the genetic characteristics of the population. Mutation is used in support of crossover to ensure that the full range of allele is accessible for each gene.

Crossover, whereas, is employed with great probability. It is a randomized though structured operator that permits information exchange between points. The goal is the preservation of fittest individual without introducing any new value.

The suggested method to the utilization of genetic algorithms for feature selection involves encoding a set of d, Feature s as a binary string of d elements, in which a 0 in the string indicates that the corresponding Feature has to be omitted, and 1 that it has to be included. The coding strategy denotes presence or absence of a certain Feature from the Feature space **[Figure 1]**. The length of chromosome equal to Feature space dimensions.

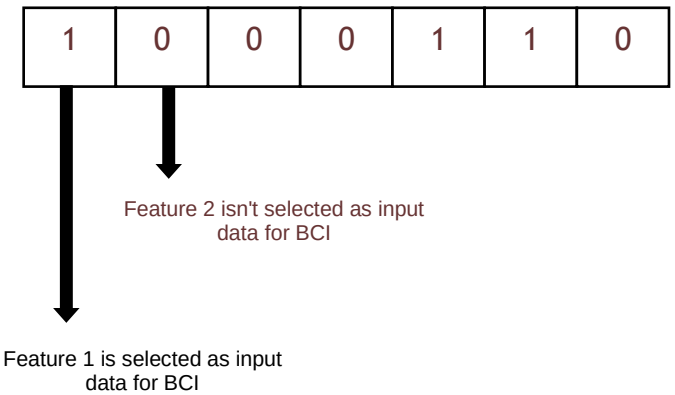

Fig: 1. Schema of the proposed GA-based feature selection approach

THE IIOAB JOURNAL
ISSN: 0976-3104

GAs are computational models inspired by evolution. These algorithms encode a potential solution to a specific problem on a simple chromosome-like data structure, and apply recombination operators to these structures in such a way as to preserve critical information. Genetic algorithms are typically seen as function optimizers, though the range of issues to which genetic algorithms are employed is vast [26].

The present study uses a GA for feature selection. Thus, each member of the population was encoded with a binary string of length equal to the feature set size. Each bit of these strings represented one specific feature. If the bit value was '1', this feature was used for classification, if the value was '0' it was not used for classification. Therefore, each member of the population represented a feature subset.

The fitness value of each member of the population was calculated as the kappa coefficient achieved using the corresponding feature subset for classifying. This criterion of accuracy has also into account the distribution of wrong classifications and it was chosen because it was used to evaluate the submissions to the dataset. Population size was equal to the feature set size. The elitist selection was set to 2 and the roulette selection method was used. Single-point crossover with probability of 0.8 and uniform mutation with probability of 0.1 were applied to every generation in order to create the next population. The number of generations was set to 50. The GA searches for all feature subset sizes smaller than 15 features.

A GA with aggressive mutation has been proposed, where a detailed description of it can be found. In short, the most important algorithm features are as follows:

Step 1: An individual is composed of integer genes. The number of genes in an individual (N) is fixed for the whole algorithm and is set either by the user or automatically based on the number of features, observations, and classifier parameters. Each gene either contains an index of one feature from the feature space or is equal to zero

Step 2: An initial population of M individuals is created randomly by choosing values from the interval f0; 1; 2; . . . ; Pg, where the values from 1 to P correspond to the feature indexes and the value zero corresponds to \none feature" state.

Step 3: The order of two main GA steps is reversed comparing to the classic scheme first the reproduction takes place and then the selection is performed.

Step 4: The basic genetic operation used in the algorithm is a mutation. This is a very aggressive form of mutation, as not only is each individual from the mother population mutated, but also each gene of that individual. The mutation scheme is as follows [27]:

$$\begin{aligned}
&for\ i\ =\ 1\ to\ M\\
&take\ an\ individual\ i\\
&for\ g = 1\ to\ N\\
&take\ a\ gene\ g\\
&assign\ a\ random\ value\ from\ the\\
&\text{int}erval\ \{0,1,2,...,P\}\ to\ the\ gene\ g\\
&save\ the\ individual\ i\ as\ a\ new\ individual
\end{aligned}$$

Step 5: The second genetic operation used in the algorithm is the classic Holland crossover performed on the mother population.

Step 6: After reproduction, the population is composed of: M mother individuals, NM of new individuals created during mutation and M new individuals created during crossover. All of these individuals are then evaluated according to their classification capabilities (i.e. a classifier is implemented and validated for each individual).

Step 7: The selection step is based on the discarding strategy in which only M individuals providing the highest classification accuracy remain in the population. Since all the best individuals from the last algorithm step (individuals from the mother population) take part in the selection process, this strategy guarantees that the best individual from the next population has at least the same fitness value as the best individual from the previous population. The population created in the selection process is a mother population for the next algorithm step.

Step 8: The reproduction and selection steps are repeated by a predetermined number of iterations.

The flowchart of GA as shown in **[Figure- 2]** [28]:

THE IIOAB JOURNAL
ISSN: 0976-3104

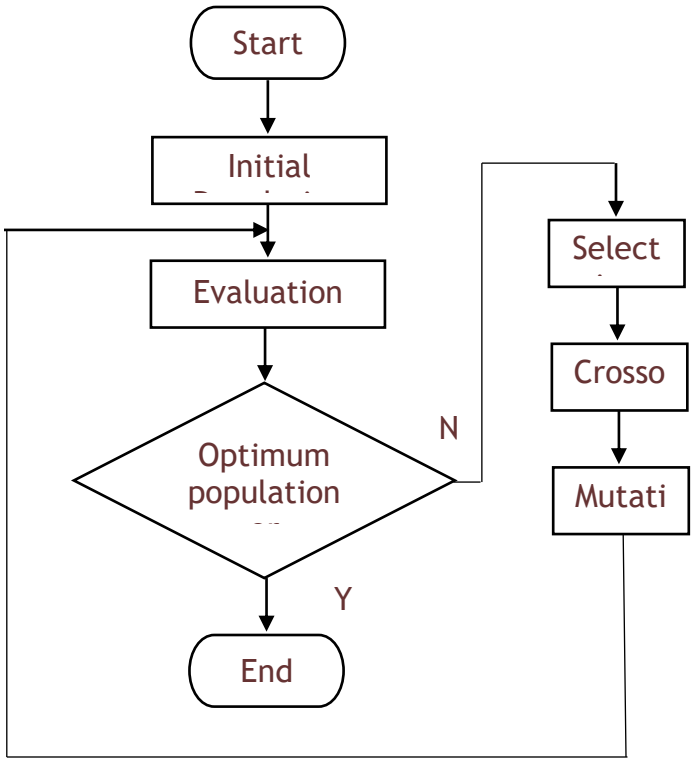

Fig: 2. Flowchart of Genetic Algorithm

PROPOSED MODIFIED CLONALG

The proposed Modified CLONALG consists in an outer (GA) search loop where the current population is checked for constraint violation and then divided into feasible (antigens) and infeasible individuals (antibodies). If there are no feasible individuals, the best infeasible one (that with the lowest constraint violation) is moved to the antigen population. Here, AIS is given as inner loop wherein antibodies are first cloned and then mutated. Next, the distances (affinities) between antibodies and antigens are computed. Those with higher affinity (smaller sum of distances) are selected thus defining the new antibodies (closer to the feasible region). This (CLONALG) cycle is repeated a number of times. The resulting antibody population is then passed to the GA where constraint violations are computed as well as fitness function values for the feasible individuals. The selection operation is then performed in order to apply recombination and mutation operators to the selected parents producing a new population and finishing the external genetic algorithm loop. The selection process in the genetic algorithm comprises in binary tournaments wherein all individuals are chosen once and the opponent is arbitrarily drawn, with substitutes from the population. The tournament rules are as follows:

- Feasible individuals are preferred to infeasible ones
- Between two feasible individuals, that with the greater fitness value is chosen, and
- Between two infeasible individuals, that with the lesser constraint violation is selected. It is to be observed here the affinity is computed from the sum of phenotypical distances between individuals, employing a standard Euclidean vector norm.

Pseudo-code for the suggested hybrid is as follows:

THE IIOAB JOURNAL
ISSN: 0976-3104

Begin
for $i = 0$ *to number Of Generations GA do*
computeViolation();
dividePopulation();
antibodies $<-$ *infeasiblePop*();
antigens $<-$ *TopFeasible*();
for $j = 0$ *to number Of Iterations CLONALG do*
cloneAntibodies();
mutateAntibodies(); *computeDistanceAntibodiesAntigens*();
antibodies $<-$ *selectBetterAntibodies*();
end $-$ *do*;
computeViolationAntibodies();
computeFitnessFeasiblePop();
tournamentSelection();
crossover();
mutation();
end $-$ *do*;
End

The main goal of the research reported in this paper was to explore alternative constraint-handling schemes for GAs used in global optimization. In the past, it have explored the use of penalty functions in engineering optimization. However, this previous work indicated a high correlation between performance of the GA and the fine-tuning of its penalty factors. Additionally, an important issue for us was not to increase in an important way, the number of fitness function evaluations (as when using, for example, the coevolutionary penalties, which do not require a manual fine-tuning of the penalty factors, but whose computational cost is extremely high).

The previous requirements led us to the development of the approach proposed in this paper. In the proposed approach, it use the search engine of the GA to conduct the search towards the global optimum. However, the GA is hybridized with a scheme inspired on an artificial immune model, which acts as a local search mechanism that helps the GA to reach the feasible region in a more efficient way. Since this local search mechanism is based only on similarities between chromosomic strings, no additional evaluations of the fitness function are required. Thus, it keep a low computational cost for the approach, which was one of its main design goals.

The proposed algorithm emulates the invaders recognition process by combining antibodies' libraries in order to attain antigen specificity. Furthermore, the purpose is to learn to identify the proper antibodies. The search process of the approach is led by a GA. Thus, what it propose is a scheme in which a simple emulation of an artificial immune system is embedded into a GA. Note however that the computational complexity of the approach is not really O (N^2), because the internal scheme (i.e., the artificial immune system) does not evaluate the original fitness function of the problem as it will see later on **[Figure 1]**. This internal scheme (which is indeed another GA) guides its search based on string similarities and not on objective function values [29].

A Serial Version of the Proposed Algorithm: The serial algorithm version to handle constraints using an immune system is described below **[Figure- 1]**:

Step 1: Generate randomly an initial population for the GA.
BEGIN inner GA
Step 2: If the initial population contains a mixture of feasible and infeasible individuals, then it divide the population in two groups. The first group contains the infeasible individuals, which are denominated "antibodies", and the second contains the feasible individuals, which are called "antigens".
Step 3: If none of the individuals in the initial population is feasible, then it use the magnitude of constraint violation of each individual as its fitness. Then, it use the best individual in the population as the "antigen", where "best" refers to the individual with the lowest amount of constraint violation.
Step 4: Select randomly a sample of antibodies of size σ .
Step 5: The fitness of the sample of antibodies is computed according to their similarity with a set of antigens in the following way:

THE IIOAB JOURNAL ISSN: 0976-3104

- An antigen is randomly selected from the antigens population.
- Each antibody in the sample is compared against the antigen selected, and it compute the result of the comparison, to which it will call Z (matching magnitude). Z represents a distance (normally but not necessarily Euclidean) measured at the genotype level (i.e., at the level of the chromosomic encoding). Z is computed using:

$$Z = \sum_{i=1}^{L} t_i$$

Where t_i = 1 if there is a matching at position i=1.....L (L is the length of the chromosome), or zero if there is no match. A large Z value means a high matching between the two strings compared and, therefore, a high fitness value.

Step 6: Based on the fitness computed in the previous step, the population of antibodies is reproduced in a traditional GA (using crossover and mutation).

Step 7: The process is repeated from the fourth step until convergence (e.g., when the mean and the maximum fitness in the population are practically the same) or until it reach a maximum number of iterations.

Step 8: Individuals are returned to the external GA and it proceed in the conventional way.

END inner GA

Step 9: Apply binary tournament selection (with the objective function of the problem) using special rules (as described below).

Step 10: Apply crossover and mutation in a conventional way.

Step 11: The process is repeated from step 2 until reaching stopping condition.

The binary tournament used in step 9 is defined in the following way (two individuals are compared each time):

- If one individual is infeasible and the other one is feasible, then the feasible individual wins.
- If both individuals are feasible, then the one with the highest fitness value is the winner.
- If both individuals are infeasible, then the winner is the one with the lowest constraint violation value.

Few problems are required to be mentioned. First, as it indicated before, the approach is really using a GA embedded inside another GA used to optimize a certain function. However, the GA that is run with the emulation of the immune system does not use the fitness function directly; it only computes Hamming distances, which are very inexpensive with respect to evaluating the objective function of the problem. Also, the implicit premise of the technique is that, under certain conditions, the reduction of genotypic differences between two individuals will produce, as a consequence, a phenotypic similarity, which, in the case, will make that an infeasible individual approaches the feasible region. The algorithm is an extension of the proposal of Hajela& Lee.

To clarify the way in which the approach works, it provide next both, the internal and the external GA's adopted:

Internal GA

Step 1: Initialize the fitness of all antibodies to zero.

Step 2: Compute the fitness of the antibody pool based on similarity to the antingens (or based on complementarity); this requires the following specific steps:

- An antigen is selected at random.
- A sample of antibodies of size μ is selected from the antibody pool without replacement.
- The match score of each antibody is computed by comparing against the selected antigen, and the antibody with the highest score has the match score added to its fitness value; the fitness of the other antibodies is unchanged.
- The antibodies are then returned to the antibody population, and the process is repeated a number of times (typically two or three times the antibody population size).

Step 3: Based on the fitness computed in Step 2, a GA simulation is conducted with prescribed probabilities of crossover and mutation to evolve the antibody population through one generation of evolution.

Step 4: The process is then repeated from Step 1 until convergence in the antibody population is attained.

External GA

Step 1: A population of designs is randomly generated.

Step 2: The fitness function, a composite of the objective function and a penalty associated with constraint violation, is obtained for the entire population.

Step 3: Members within the top 3% of the population obtained at the end of Step 2 are designated as antigens, and the entire population (including the antigens) is defined as the starting population of antibodies.

Step 4: Using and antibody sample size μ smaller than the number of antigens, the degree of match Z is obtained for each member of the population according to the steps described before.

Step 5: The match score of each design is used as a fitness measure in a traditional selection or reproduction operation. During this reproduction operation, the size of the population is unchanged.

Step 6: The crossover and mutation operations are performed on the new population of antibodies formed in Step 5.

Step 7: The process is then repeated from Step 2 with an intent of evolving the population to maximize the Z function and cycled to convergence.

Note that the proposed approach is designed to operate only on binary strings. Although it know that the binary alphabet can be used to encode any type of decision variables, it may be useful in some cases to use alternative encodings. Should that be the case, the proposed approach is not directly applicable, and its generalization to alphabets of higher cardinality (e.g., real-numbers encoding) remains as an open research area.

THE IIOAB JOURNAL
ISSN: 0976-3104

At this point, it is important to clarify the main weaknesses that it identified in Hajela's algorithm and that led us to develop the algorithm proposed herein:

Hajela's approach requires a penalty function in order to sort the population and assign the antigens. This makes necessary to evaluate twice the objective function of the problem (per individual) at each generation of the external GA. This may become considerably expensive (computationally speaking) when dealing with real-world applications. In contrast, the approach only evaluates once the objective function (for each individual) per generation, since it do not use a penalty function. It relate the values of the antigens to the constraint violation of each solution. This keeps us from evaluating the fitness function more than once and makes unnecessary to sort the population.

In Hajela's approach, the computation of the fitness Z in the internal GA is performed through a cycle in which the population of antibodies must be traversed several times. In the case, it compute Z in the internal GA by performing a single traversal of the antibodies.

The approach of Hajela& Lee is only validated with a few engineering optimization problems, and no information about its computational cost is provided. In the case, it have used some benchmarks reported in the evolutionary computation literature and it have compared the results against a highly competitive constraint-handling technique which is representative of the state-of-the-art in the area (the homomorphous maps).

A schematic of the serial version of the algorithm based on the artificial immune system as shown in **[Figure- 3]**

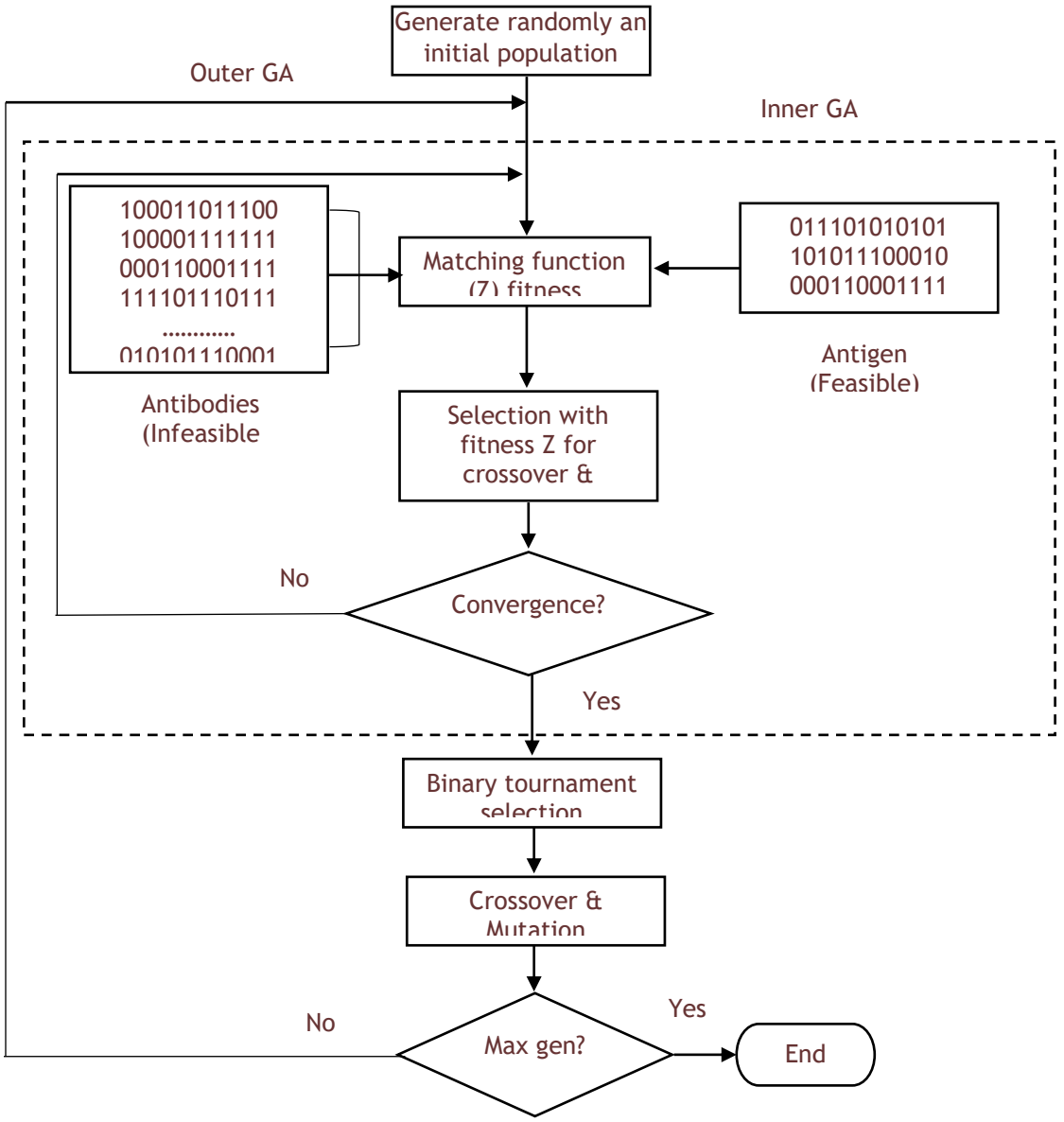

Fig: 3 (A). schematic of the serial version of the algorithm based on the artificial immune system

THE IIOAB JOURNAL
ISSN: 0976-3104

RESULTS

In this section, the Autoregressive - Wavelet fused features, Modified CLONALG Feature Selection, SVM classifier (ARWAMCSVM), Autoregressive - Wavelet fused features, Hybrid CLONALG Feature Selection, SVM classifier (ARWAHCSVM), Autoregressive - Wavelet fused features, Modified CLONALG Feature Selection, Optimized SVM classifier (ARWAMCSVM-Opt) and Autoregressive - Wavelet fused features, Hybrid CLONALG Feature Selection, Hybrid Optimized SVM classifier (ARWAHCSVM-HOpt) are evaluated. **[Table- 1]** shows the summary of results obtained. **[Figure- 4], [Figure- 5], [Figure- 6]** shows the classification accuracy, precision and recall respectively.

Table: 1. Summary of Results

	Classification Accuracy	Precision	Recall
ARWAMCSVM	84.17	0.84205	0.8417
ARWAHCSVM	92.81	0.92815	0.9281
ARWAMCSVM-Opt	95.68	0.95695	0.9568
ARWAHCSVM-HOpt	96.76	0.9678	0.9676

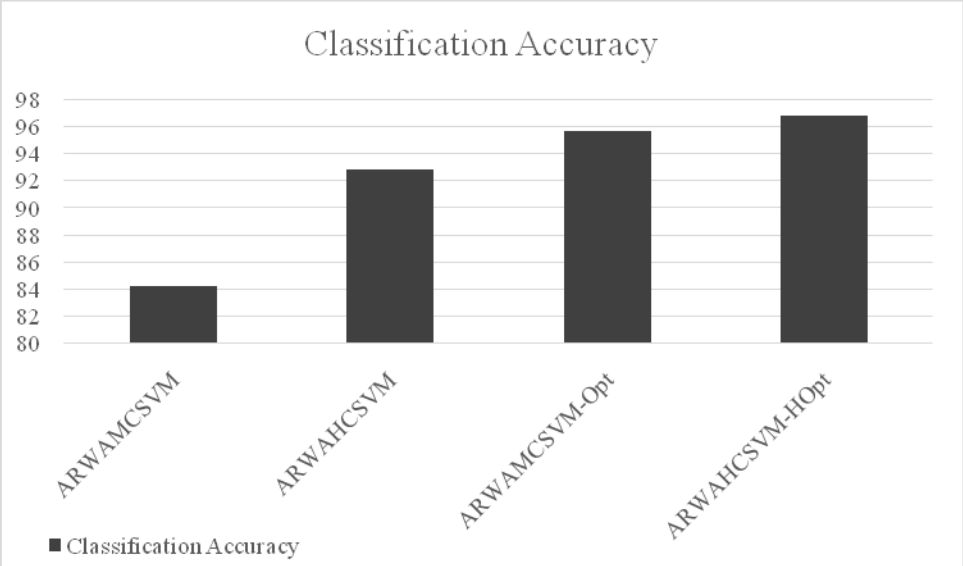

Fig: 4 .Classification Accuracy

From **Table- 1** and **[Figure -4]** it is observed that the classification accuracy of ARWAHCSVM-HOpt performs better by 13.9% than ARWAMCSVM, by 4.17% than ARWAHCSVM and by 1.12% than ARWAMCSVM-Opt.

THE IIOAB JOURNAL
ISSN: 0976-3104

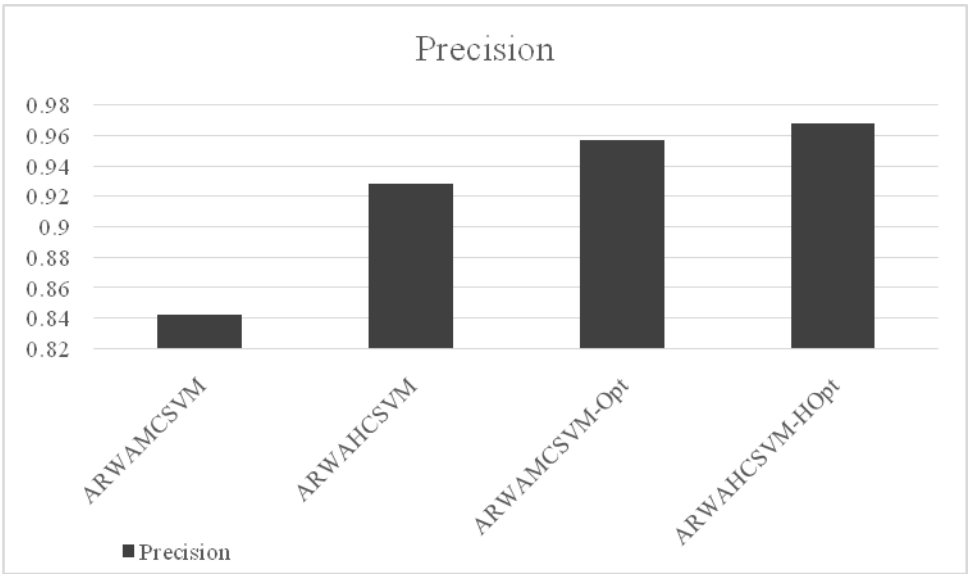

Fig: 5. Precision

From **Table- 1** and **[Figure- -5]** it is observed that the precision of ARWAHCSVM-HOpt performs better by 13.9% than ARWAMCSVM, by 4.18% than ARWAHCSVM and by 1.13% than ARWAMCSVM-Opt.

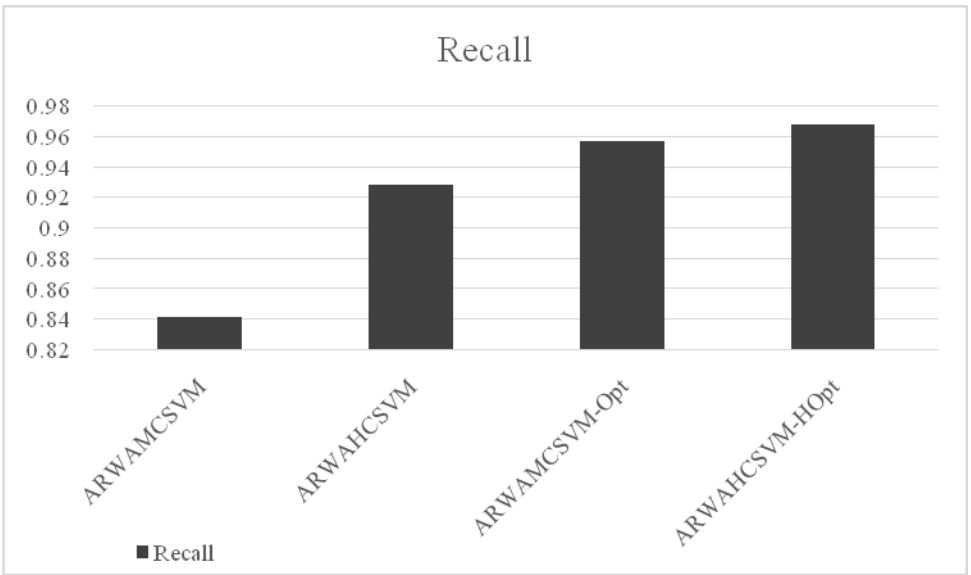

Fig: 6.Recall

From **Table- 1** and **[Figure- 6]** it is observed that the recall of ARWAHCSVM-HOpt performs better by 13.9% than ARWAMCSVM, by 4.17% than ARWAHCSVM and by 1.12% than ARWAMCSVM-Opt.

CONCLUSION

ECG includes a spatial scale between EEG and intra-cortical microelectrode recording, and ECoG offers a balance between invasiveness, spatiotemporal resolution, and signal stability for BCI applications. BCI has progressed, but it is slowed by many factors including noise in brain signals, muscular artefacts and inconsistency and variability of user attention/intentions. In this paper proposed modified CLONALG optimizes SVM. For

THE IIOAB JOURNAL
ISSN: 0976-3104

intensification, the strategy works with many clones to improve. Experiments were undertaken through tenfold cross validation and accuracy achieved is satisfactory but further work is needed for classification accuracy improvement.

CONFLICT OF INTEREST

The authors declare no conflict of interests.

ACKNOWLEDGEMENT

None

FINANCIAL DISCLOSURE

The authors report no financial interests or potential conflicts of interest.

REFERENCES

[1] Ko M, Bae K, Oh G, Ryu T. [2009] A study on new gameplay based on brain-computer interface. In Breaking New Ground: Innovation in Games, Play, Practice and Theory: Proceedings of the 2009 Digital Games Research Association Conference, Brunel University.

[2] Leuthardt EC, Schalk G, Wolpaw JR, Ojemann J G, Moran DW. [2004] A brain–computer interface using electrocorticographic signals in humans, *Journal of neural engineering*, 1(2):63.

[3] Miller KJ, Abel TJ, Hebb AO, Ojemann JG. [2011] Rapid online language mapping with electrocorticography. *Journal of neurosurgery Pediatrics*, 7(5):482.

[4] Al-Qudah AA, Tamimi AF, Ghanem S. [2000] Electrocorticography in the management of surgically treated epileptic patients. *Neurosciences*, 5(1):22-25.

[5] Shenoy P, Miller KJ, Ojemann JG, Rao RP. [2008] Generalized features for electrocorticographic BCIs. Biomedical Engineering, *IEEE Transactions on*, 55(1):273-280.

[6] Nicolas-Alonso LF, Gomez-Gil J. [2012] Brain computer interfaces, a review, *Sensors*, 12(2):1211-1279.

[7] Brownlee J. [2007] Clonal selection algorithms. Complex Intelligent Systems Laboratory, Swinburne University of Technology, Australia.

[8] Koklu M, Kahramanli H, Allahverdi N. [2012] A New Approach To Classification Rule Extraction Problem by the Real Value Coding.nternational *Journal of Innovative Computing, Information and Control*, 8(9).

[9] Aydin I, Karakose M. Akin E. [2011] A multi-objective artificial immune algorithm for parameter optimization in support vector machine.*Applied Soft Computing*, 11(1): 120-129.

[10] Buzón-Cantera IE, Mora-Vargas J, Ruiz, A, Soriano P. [2015]A Hybrid Optimization Model: An Approach for the Humanitarian Aid Distribution Problem, *Applied Mathematical Sciences*, 9(127): 6329-6346.

[11] Shende PM, Jabade VS. [2015, January] Literature review of brain computer interface (BCI) using Electroencephalogram signal. In Pervasive Computing (ICPC), 2015 International Conference on (pp. 1-5). IEEE.

[12] Kim M, Kim BH, Jo S. [2015] Quantitative Evaluation of a Low-Cost Noninvasive Hybrid Interface Based on EEG and Eye Movement. Neural Systems and Rehabilitation Engineering, *IEEE Transactions on*, 23(2): 159-168.

[13] Fan XA, Bi L, Teng T, Ding H, Liu Y. [2015] A brain–computer interface-based vehicle destination selection system using P300 and SSVEP signals. Intelligent Transportation Systems, *IEEE Transactions on*, 16(1): 274-283.

[14] Krusienski DJ, Shih JJ. [2010, August] A case study on the relation between electroencephalographic and electrocorticographic event-related potentials. In Engineering in Medicine and Biology Society (EMBC), 2010 *Annual International Conference of the IEEE*, 6019-6022.

[15] Hamedi M, Salleh S, Ting CM. [2014, September] Multiclass self-paced motor imagery temporal features classification using least-square support vector machine. In Functional Electrical Stimulation Society Annual Conference (IFESS), 2014 IEEE 19th International (pp. 1-5). *IEEE*.

[16] Gao W, Guan JA, Gao J, Zhou D. [2015] Multi-ganglion ANN based feature learning with application to P300-BCI signal classification. *Biomedical Signal Processing and Control*, 18: 127-137.

[17] Kimovski D, Ortega J, Ortiz A, Baños, R. [2015]Parallel alternatives for evolutionary multi-objective optimization in unsupervised feature selection. *Expert Systems with Applications*, 42(9): 4239-4252.

[18] Agarwal SK, Shah S, Kumar R. [2015] Classification of mental tasks from EEG data using backtracking search optimization based neural classifier. *Neurocomputing*.

[19] Ding S. [2009, November] Feature selection based F-score and ACO algorithm in support vector machine. In Knowledge Acquisition and Modeling, 2009. KAM'09. Second International Symposium on *IEEE*, 1: 19-23).

[20] Gonzalez A, Nambu I, Hokari H, Iwahashi M, Wada Y. [2013, October] Towards the Classification of Single-Trial Event-Related Potentials Using Adapted Wavelets and Particle Swarm Optimization. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 3089-3094*). IEEE*.

[21] Wang L, Xu G, Wang J, Yang S, Guo L, Yan W. [2011, July] GA-SVM based feature selection and parameters optimization for BCI research. In Natural Computation (ICNC), 2011 Seventh International Conference on, 1: 580-583). *IEEE*.

[22] Rathipriya N, Deepajothi S, Rajendran T. [2013] Classification of motor imagery ecog signals using support vector machine for brain computer interface. In Advanced Computing (ICoAC), 2013 Fifth International Conference on (pp. 63-66). *IEEE*.

[23] Rejer I, Lorenz K. [2013] Genetic algorithm and forward method for feature selection in EEG feature space.

THE IIOAB JOURNAL
ISSN: 0976-3104

Journal of Theoretical and Applied Computer Science, 7(2): 72-82.

[24] Atum Y, Gareis I, Gentiletti G, Acevedo R, Rufiner L. [2010] Genetic feature selection to optimally detect P300 in brain computer interfaces. In Engineering in Medicine and Biology Society (EMBC), 2010 *Annual International Conference of the IEEE ,IEEE*, 3289-3292.

[25] Ghanbari AA, Broum A, Navidi H, Ahmadi A. [2012] Brain computer interface with genetic algorithm. *International Journal of Information and Communication Technology Research*, 2.

[26] Corralejo R, Hornero R, Alvarez D. [2011, August] Feature selection using a genetic algorithm in a motor imagery-based Brain Computer Interface. In Engineering in Medicine and Biology Society, EMBC, 2011 *Annual International Conference of the IEEE* (pp. 7703-7706). *IEEE*.

[27] Rejer I. [2015] Genetic Algorithms for Feature Selection for Brain–Computer Interface. *International Journal of Pattern Recognition and Artificial Intelligence*, 29(05):1559008.

[28] Moon SN, Bawane N. [2015] OPTIMAL FEATURE SELECTION BY GENETIC ALGORITHM FOR CLASSIFICATION USING NEURAL NETWORK.

[29] Coello CAC, Cortés NC. [2004] Hybridizing a genetic algorithm with an artificial immune system for global optimization. Engineering Optimization. Doi: 10.1080/03052150410001704845