Indian Sign Language Recognition Using Optimized Neural Networks

Sirshendu Hore, Sankhadeep Chatterjee, V. Santhi, Nilanjan Dey, Amira S. Ashour, Valentina Emilia Balas and Fuqian Shi

Abstract Recognition of sign languages has gained reasonable interest by the researchers in the last decade. An accurate sign language recognition system can facilitate more accurate communication of deaf and dumb people. The wide variety of Indian Sign Language (ISL) led to more challenging learning process. In the current work, three novel methods was reported to solve the problem of recogni-

S. Hore (✉)
Department of Computer Science and Engineering,
Hooghly Engineering and Technology College, Chinsurah, India
e-mail: sirshendu.hore@hetc.ac.in

S. Chatterjee
Department of Computer Science and Engineering, University of Calcutta,
Kolkata, India
e-mail: sankha3521@gmail.com

V. Santhi
VIT University, Vellore, India
e-mail: vsanthinathan@gmail.com

N. Dey
Department of Information Technology, Techno India College of Technology,
Kolkata, India
e-mail: neelanjandey@gmail.com

A.S. Ashour
Faculty of Engineering, Department of Electronics
and Electrical Communications Engineering, Tanta University, Tanta, Egypt
e-mail: amirasashour@yahoo.com

A.S. Ashour
College of CIT, Taif University, Taif, Kingdom of Saudi Arabia

V. Balas
Faculty of Engineering, Aurel Vlaicu University of Arad, Arad, Romania
e-mail: valentina.balas@uav.ro

F. Shi
College of Information and Engineering, Wenzhou Medical University,
Wenzhou 325035, People's Republic of China
e-mail: sfq@wmu.edu.cn

V.E. Balas et al. (eds.), *Information Technology and Intelligent Transportation Systems*, Advances in Intelligent Systems and Computing 455,
DOI 10.1007/978-3-319-38771-0_54

tion of ISL gestures effectively by combining Neural Network (NN) with Genetic Algorithm (GA), Evolutionary algorithm (EA) and Particle Swarm Optimization (PSO) separately to attain novel NN-GA, NN-EA and NN-PSO methods; respectively. The input weight vector to the NN has been optimized gradually to achieve minimum error. The proposed methods performance was compared to NN and the Multilayer Perceptron Feed-Forward Network (MLP-FFN) classifiers. Several performance metrics such as the accuracy, precision, recall, F-measure and kappa statistic were calculated. The experimental results established that the proposed algorithm achieved considerable improvement over the performance of existing works in order to recognize ISL gestures. The NN-PSO outperformed the other approaches with 99.96 accuracy, 99.98 precision, 98.29 recall, 99.63 F-Measure and 0.9956 Kappa Statistic.

Keywords Indian sign language · Neural network · Genetic algorithm · Evolutionary algorithm · Particle swarm optimization

1 Introduction

Sign language is a widely used medium of communication for deaf and dumb people. Developing systems to facilitate sign language handling is a significant subject of researchers' interest. Thus, recognition of sign languages is a challenging process due to the involvement of several technological aspects. Automated systems are urgent to serve the sign language interpretation purpose easily. Such systems facilitate the effective communication with deaf and dumb more reliable, accurate and robust processes. Sign languages have a wide variety all over the world, such as the American Sign Language, Australian Sign Language, German Sign Language, British Sign Language and Indian Sign Language (ISL). They differ in terms of morphology, grammar and syntax.

Large variation of ISL can be found in Indian sub-continent. However, the grammatical structure is same for all dialects [1]. In Indian sub-continent, there are approximately more than six million deaf and dumb people [2].

Extensive studies have been conducted using neural network in various applications [3–7]. Sign language recognition of ISL attracted researchers to improve recognition techniques in the last decade. A four step method to recognize alphabets of ISL has been proposed in [8]. In [9], a statistical approach to classify and recognize dynamic gestures of ISL in real time was proposed. Orientation of histogram has been used as a key feature for classification. Edge orientation of the sequence of dynamic ISL gestures has been calculated by employing a simple algorithm. K-nearest neighbor and Euclidean distance have been used for classification. The authors have reported accuracy range of 51.35 to 100 % and about 64.42 to 100 % while using Euclidean distance for 18 and 36 beans; respectively. Oszust and Wysocki (2010) [10] proposed an approach for automatic signed expressions recognition based on modeling gestures with subunits. The authors applied the evolutionary optimization technique to determine the cut points required for the proposed method. Krishnaveni and Radha (2011) [11] suggested a method for potential multilevel thresholds image segmentation method based on maximum entropy and PSO to estimate the human gestures. Results proved that computing using PSO can promptly convergence the results with high computational efficiency.

Recognition of South Indian Sign Language gestures have been proposed by Rajam et al. [12]. A set of 32 gestures, each depicting different postures of five fingers, has been used as the sign language. The authors have reported an accuracy of 98.125 % with 320 data instances for training and 160 for testing. Deora et al. [13] proposed a human computer interface capable of recognizing ISL gestures. The authors reported the recognition complexity due elaborating both hands and overlapping hands cases. A recent research has reported a Support Vector Machine based recognition system of ISL [14]. A novel method has been proposed in [15] to recognize ISL gestures for Humanoid Robot Interaction. The authors have proposed an efficient approach to communicate between a HOAP-2 robot and a human being. Feature extraction is carried out combining Genetic algorithm and Euclidean distance.

In the current work, three novel techniques have been proposed to recognize the ISL gestures. The NN has been trained using Genetic Algorithm, Evolutionary Algorithm and Particle Swarm Algorithm to enhance its performance. A set of 22 ISL gestures has been used to test the performance of the proposed work with ten images for each of the gestures. For the experimental results, 70 and 30 % of the data set has been used to train and test the NN, respectively.

The structure of the remaining sections is as follows. The background of the NN, GA and PSO techniques is introduced. The methodology followed by the results and discussion in Sects. 3 and 4; respectively. Finally, the conclusion is presented in Sect. 5.

2 Background

In this section the basics of the NN, GA and PSO are conducted as follows.

2.1 Artificial Neural Networks

Predominantly, artificial neural network (ANN) is used in numerous applications. It can be considered non-linear, highly parallel, robust, fault tolerant network. ANN has the ability to handle imprecise and fuzzy information easily [16].

In the ANN, neurons receive inputs (x_j) as stimuli from the environment. Then, these inputs are combined using their corresponding weights (w_j) to form 'net' input (net_j). The 'net' input is passed through a non-linear threshold filter to get the output (y) that can pass to another neuron in the network. The neuron can be activated if exceeds the threshold or bias 'θ_j' value of the concerned neuron. The net input is calculated using the Eq. (1), which computes the input for 'n' input signals by adding the dot products of weight (w) and strength (x) of each signal.

$$net_j = \sum_{i=1}^{n} w_{ij} x_i \tag{1}$$

In order to calculate the output signal, the calculated net input is compared against a threshold value as shown in Eq. 2.

$$\text{y} = \begin{cases} 1, & \text{if } net_j \geq \theta_j \\ 0, & \text{if } net_j < \theta_j \end{cases} \tag{2}$$

A perceptron learning rule is to be used in order to find the optimal weight vector. Several network architectures have been used to enhance the NN performance, such as the two-layer perceptron feed-forward network that can be used for the MLP-FFN experiments.

2.2 Genetic Algorithm

Genetic algorithm has been implemented to solve various optimization problems [17, 18]. The GA starts any problem solving using a set of initial solutions. It continuously applies crossover and mutations on the solutions to produce better offspring. The survival of any offspring depends on the fitness that based on the problem under concern. In each generation the best offspring survives and gradually produces better solutions in next generations until the required accuracy is achieved.

In the current work, determination of the NN initial weight vector is considered an optimization problem. Thus, the GA is applied to optimize the initial weight vector of the NN using the following algorithm, using the Root Means Square Error (RMSE) as a fitness function.

Algorithm: Genetic Algorithm

Start

Input: initial population represented by N numbers of chromosomes, which is randomly generated and takes value between ‘0’ to ‘1’.

Calculating fitness values: a fitness function need to be defined for evaluating individual solution or chromosomes.

Selection

- The $RMSE_i$ value is calculated for each solution in population
- Calculate the average to find the $RMSE_{ave}$
- Select randomly the $RMSE_r$ from predefined closed interval [0, $RMSE_{ave}$]
- Calculate the difference ($RMSE_r - RMSE_i$) for every solution.

If the obtained result is $\leq$ '0'
Then i^{th} individual is selected
End If

- Repeat the previous step until the number of solutions selected for next generation is equal to the number of solutions in the population.

Cross-over using the selected chromosomes
Mutation: Genes from randomly selected position are swapped to create new individual solution.
Termination: check the termination condition based on the selected condition
End

2.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) considers a population of particles (candidate solutions) in D-dimensional search space [19]. To estimate the particle's ability to realize the objective, each particle is associated with a fitness value. Primarily, the particles are located randomly in the search space. The swarm moves inside the search space to accomplish the optimal fitness. Particles are related to account for the best solution in the hyperspace, which denoted by '*pbest*'. The best value realized by any particle of the swarm is known as the Global best '*gbest*'. Each particle's position and velocity are initialized randomly. The fitness values of the particles are calculated after every iteration. The necessary adjustments are performed to the position/ velocity of each particle to move them to optimal fitness. The general algorithm of PSO to be used with the NN is as follows.

Algorithm: Particle Swarm Optimization

Start
The particles of the Swarm is placed at random positions with zero velocity
for n : 1 : Swarm size ***do***
Compute fitness
end for
for i : 1 : number of iterations ***do***
for j : 1 : Swarm size ***do***
Update *pbest*
Update *gbest*
Adjust position and velocity
Calculate fitness for the new population
end for
end for
End

Dixit et al. (2015) [20] introduced an exhaustive survey on nature inspired optimization algorithms including the evolutionary optimization strategy. Accordingly, the GA, EA and the PSO optimization algorithms were employed to support the NN for the Sign language recognition approach.

3 Experimental Methodology

In the current work, various models using Neural Network (NN), MLP-FFN, NN-GA, NN-Evolutionary, and NN-PSO have been utilized. The conjugate gradient algorithm has been used with the NN and MLP-FFN modules for learning [21]. The basic steps elaborated in the proposed methodology for sign language recognition are preprocessing, data cleaning and data normalization. Significant features/ attributes for classifying the dataset are extracted from the datasets accurately using effective features. If optimum feature set is combined with the Bayes [22] classifier it would result in minimum error for the given distributions. Therefore, theoretically the Bayes error would be considered as the optimum measure of the feature effectiveness. After the preprocessing step, data cleaning is required to resolve the missing values and noise issues. It is required to remove noise and fill up empty entries by suitable values through statistical analysis. Additionally, in order to reduce the distance between attribute values, data normalization is carried out by keeping the value of range between -1 to $+1$.

Afterward, the classification is conducted by dividing the datasets into two parts, namely (i) 70 % of the data as training dataset and (ii) 30 % as testing dataset. In training phase, the training dataset is supplied to different algorithms to build the required classification model. In testing phase, the classification models obtained from the training phase is employed to test the accuracy of the proposed models.

To measure the proposed performance, several metrics such as the correlation coefficient, Kappa statistic, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), True Positive rate (TP rate), and F-measure were calculated [23]. These metrics are defined as follows. The RMSE is used to measure the difference between the values anticipated by a classifier and the values actually discovered from the surroundings of the system being modeled. The RMSE of a classifier prediction with respect to the computed variable v_{c_k} is evaluated by: $RMSE = \sqrt{\frac{\Sigma_{k=1}^{n}(v_{d_k} - v_{c_k})^2}{n}}$. Where v_{d_k}, denotes the originally observed value of k^{th} object and v_{c_k} denotes the predicted value by the classifier.

The confusion matrix is a tabular representation of the classification performance. As illustrated in Table 1, each column of the matrix denotes the examples in a predicted class, while each row indicates the examples in an actual class, where (i) True positive (tp) is the number of positive instances categorized as positive, (ii) False positive (fp) is the number of negative instances categorized as positive, (iii) False negative (fn) is the number of positive instances categorized as negative and (iv) True negative (tn) is the number of negative instances categorized as negative.

Table 1 Typical example of confusion matrix of a binary classification problem

Actual class	Predicted class	
	Positive	Negative
Positive	tp	fp
Negative	fn	tn

Other metrics are derived from confusion matrix, namely: (i) the accuracy which defined as the ratio of the correctly classified instances sum to the total number of instances, which expressed as ($Accuracy = \frac{tp+tn}{tp+fp+fn+tn}$), (ii) the precision is defined as the ratio of correctly classified data in positive class to the total number of data classified to be in the positive class, which given by ($Precision = \frac{tp}{tp+fp}$), (iii) the recall (TP rate) is given by ($Recall = \frac{tp}{tp+fn}$), iv) F-measure which is a combined representation of the precision and the recall and represented by: ($F - measure = 2 * \frac{Precision * Recall}{Precision + Recall}$).

Another performance metric is the Kappa Statistic, which is a statistical measure denoted by k [24]. The value of k is estimated using: ($k = \frac{prob\,(0) - prob\,(E)}{1 - prob\,(E)}$). Where, $prob(0)$ is the probability of the observed agreements and $prob(E)$ is the same for agreements expected by chance.

4 Results and Discussion

Recently, various studies are concerned with neural network and optimization approaches in several applications [25, 26]. Meanwhile, the current study suggested a novel approach to optimize the input weight vector to the NN to achieve minimum error using optimization algorithms. The proposed work is tested with a dataset consists of 22 ISL gestures with 10 images for each gesture. Table 2 depicts the performance metrics values of the proposed algorithms along with NN and MLP-FFN classifiers.

Table 2 illustrated that with respect to the recall values, the NN is has attained 92.86 %, while using the NN-PSO is 98.29 %, which outperformed all other approaches. In addition, in terms of the F-measure values, the NN-PSO approach established its superiority compared to the other tested approaches. The obtained accuracy and precision for proposed the different compared approaches are demonstrated in Figs. 1 and 2; respectively.

Table 2 along with Figs. 1 and 2 illustrated that the accuracy of the NN approach is 93.64 %, which improved to be 96.7 % using the MLP-FFN approach. In terms of the accuracy, the results established that the NN-PSO based approach achieved 99.96 % accuracy, which outperformed both the NN-GA and the NN-EA. In terms of precision, the NN approach has attained the least precision value of 94.55 %, while the NN-PSO has achieved the superior precision value of 99.98 %. Finally, the kappa

Table 2 Calculated accuracy of proposed methodology

Algorithm / Performance measure	NN	MLP-FFN	NN-GA	NN-Evolutionary (NN-EA)	NN-PSO
Accuracy	93.64	96.7	97.83	98.86	**99.96**
Precision	94.55	97.25	98.43	99.83	**99.98**
Recall	92.86	94.67	97.71	97.75	**98.29**
F-measure	93.69	95.94	98.07	98.78	**99.63**
Kappa statistic	0.9127	0.9359	0.9772	0.9822	**0.9956**

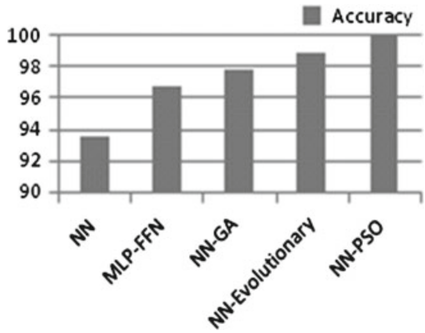

Fig. 1 Accuracy analysis for different proposed methodologies

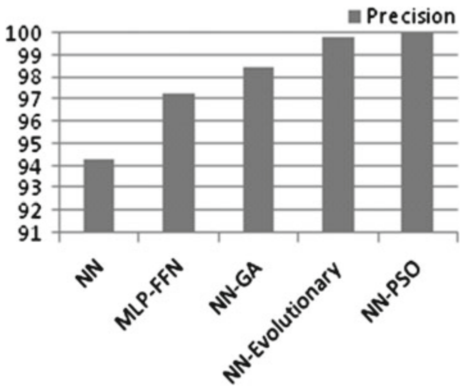

Fig. 2 Analysis of precision for different proposed methodologies

statistic profoundly establishes the improvement claims achieved by the NN-PSO as indicted in Fig. 3.

Figure 3 along with Table 2 proved that the NN-PSO gained the superior Kappa Statistic value of 0.9956 compared to the other methods. The confusion matrix of the testing phase for the NN (as an example) is illustrated in Table 3.

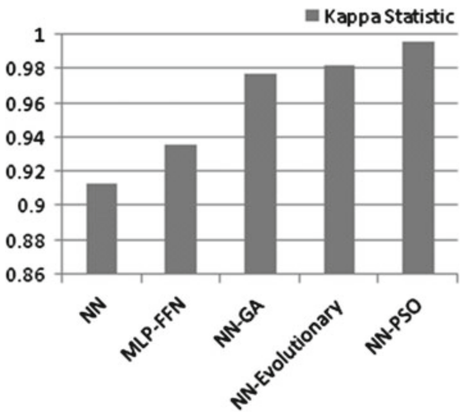

Fig. 3 Kappa statistic measure for different proposed methodologies

Table 3 Confusion matrix of testing phase for neural network

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20	T21	T22
P1	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P2	0	9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
P3	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P4	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P5	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P6	0	0	0	0	0	9	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
P7	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P8	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	2	0	0	0
P9	0	1	0	1	0	0	0	0	8	0	0	0	0	0	0	1	1	0	0	0	0	0
P10	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0	0
P12	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0	0	0
P13	0	0	0	0	0	1	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0
P14	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0	0
P15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0	0	0
P16	0	0	2	0	0	0	0	0	0	1	0	0	0	0	0	7	0	0	0	0	0	0
P17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0	0
P18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0	0	0
P19	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0
P20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	0
P21	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0
P22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10

From the preceding results, it is established that the NN-PSO outperformed all the other tested methods for the recognition of the ISL sign language.

5 Conclusion

The present work proposes three novel strategies to recognize Indian Sign Language using a dataset of 22 ISL gestures with 10 images for each. Based on the literature review, it is observed that the traditional learning algorithms significant problems as they have a chance of getting trapped into local optima while optimizing an objective. Thus the proposed optimization algorithms were elaborated to solve this problem. Moreover, the previous existing studies have used only the accuracy for the performance analysis, while it is observed that the accuracy of an algorithm varies greatly with the variation in the number of instances in different classes. Thus, precision, recall, F-Measure and kappa statistic have been considered as metrics to evaluate the proposed algorithms. Experimental results have revealed that among the three proposed methods, the NN-PSO approach has performed well compared to other algorithms in terms of all metrics. It achieved 99.96 accuracy, 99.98 precision, 98.29 recall, 99.63 F-Measure and 0.9956 Kappa Statistic.

References

1. Zeshan U (2003) Indo-Pakistani sign language grammar: a typologicaloutline. Sign Lang Stud 3(3):157–212
2. Naik SM, Naik MS, Sharma A (2013) Rehabilitation of hearing impaired children in India-an update. Otolaryngol Online J 3(1):20–31
3. Bharathi S, Sudhakar R, Balas VE (2014) Biometric recognition using fuzzy score level fusion. Int J Adv Intell Paradigms 6(2):81–94
4. Samanta S, Ahmed SKS, Salem MA, Nath SS, Dey N, Chowdhury SS (2014) "Haralick features based automated glaucoma classification using back propagation neural network." In: 2014 international conference on frontiers of intelligent computing: theory and applications (FICTA), special session: advanced research in computer vision, image and video processing, pp 14–15
5. Senthilkumar S, Piah ARM (2012) An improved fuzzy cellular neural network (IFCNN) for an edge detection based on parallel RK (5, 6) approach. Int J Comput Syst Eng 1(1):70–78
6. Maji P, Chatterjee S, Chakraborty S, Kausar N, Dey N, Samanta S (2015) Effect of Euler number as a feature in gender recognition system from offline handwritten signature using neural networks. In: 2nd international conference on computing for sustainable global development. INDIACom-2015, Bharati Vidyapeeth, New Delhi (INDIA), 11–13th March 2015, pp 1869–1873
7. Kotyk T, Ashour AS, Chakraborty S, Dey N, Balas VE (2015) Apoptosis analysis in classification paradigm: a neural network based approach. In: Healthy World conference 2015—a healthy World for a happy life, Kakinada (AP), India, pp 17–22
8. Ghotkar AS, Khatal R, Khupase S, Asati S, Hadap M (2012) Hand gesture recognition for indian sign language. In: 2012 international conference on computer communication and informatics (ICCCI), pp 1–4
9. Nandy A, Prasad JS, Chakraborty P, Nandi GC, Mondal S (2010) Classification of Indian sign language in real time, pp 52–57
10. Oszust M, Wysocki M (2010) "Determining subunits for sign language recognition by evolutionary cluster-based segmentation of time series", artificial intelligence and soft computing. Springer, Berlin
11. Krishnaveni M, Radha V (2011) Improved histogram based thresholding segmentation using PSO for sign language recognition. Int J Eng Sci Technol 3(2):1014–1020
12. Rajam PS, Balakrishnan G (2011) Real time Indian sign language recognition system to aid deaf-dumb people. In: 2011 IEEE 13th international conference on communication technology (ICCT), pp 737–742. IEEE
13. Deora D, Bajaj N (2012) Indian sign language recognition. In: 1st international conference on emerging technology trends in electronics, communication and networking (ET2ECN), pp. 1–5. IEEE
14. Agrawal SC, Jalal AS, Bhatnagar C (2014) Redundancy removal for isolated gesture in Indian sign language and recognition using multi-class support vector machine. Int J Comput Vis Robot 4(1):23–38
15. Nandy A, Mondal S, Prasad JS, Chakraborty P, Nandi GC (2010) Recognizing and interpreting Indian sign language gesture for human robot interaction. In: 2010 international conference on computer and communication technology (ICCCT), pp 712–717
16. Jain AK, Mao J, Mohiuddin KM (1996) Artificial neural networks: a tutorial. Comput, 1–44
17. Kumar BV, Karpagam M (2015) Differential evolution versus genetic algorithm in optimising the quantisation table for JPEG baseline algorithm. Int J Adv Intell Paradigms 7(2):111–135

18. Dey N, Ashour AS, Beagum S, Pistola DS, Gospodinov M, Gospodinova EP, Tavares JMRS (2015) Parameter optimization for local polynomial approximation based intersection confidence interval filter using genetic algorithm: an application for brain MRI image de-noising. J Imaging 1:60–84
19. Chakraborty S, Samanta S, Mukherjee A, Dey N, Chaudhuri SS (2013) Particle swarm optimization based parameter optimization technique in medical information hiding. In: 2013 IEEE international conference on computational intelligence and computing research (ICCIC), Madurai, 26–28 Dec 2013
20. Dixit M, Upadhyay N, Silakari S (2015) An exhaustive survey on nature inspired optimization algorithms. Int J Softw Eng Appl 9(4):91–104
21. Hornik K (1991) Approximation capabilities of multilayer feed forward networks. Neural Netw 4:251–257
22. Richard MD, Lippmann R (1991) Neural network classifiers estimate Bayesian a posteriori probabilities. Neural Comput 3:461–483
23. Stehman SV (1997) Selecting and interpreting measures of thematic classification accuracy. Remote Sens Environ 62(1):77–89
24. Carletta J (1996) Assessing agreement on classification tasks: the kappa statistic. Comput Linguist 22(2):249–254
25. Ashour AS, Samanta S, Dey N, Kausar N, karaa WBA, Hassanien AE (2015) Computed tomography image enhancement using Cuckoo search: a log transform based approach. J Signal Inf Process 6(3):244–257
26. Karaa WBA, Ashour AS, Sassi DB, Roy P, Kausar N, Dey N (2016) MEDLINE text mining: an enhancement genetic algorithm based approach for document clustering. Applications of intelligent optimization in biology and medicine. Springer International Publishing, Switzerland, pp 267–287