*International Journal of Engineering & Technology*, 7 (4.28) (2018) 729 -734

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Krill Herd Optimized Feature Selection for Classification of Alzheimer's Disease from MRI Images

**S Sumanth,[1]*, Dr A Suresh, [2],**

[1] *Assistant Professor, Dept. of Computer Science, Smt. V.H.D. Central Institute of Home Science (A)*
[2] *Principal*
*Dept. of Computer Science*
*SIRI PSG Arts and Science College for Women,Sankagiri (tk) - 637 301, Salem District.*
*asuresh1975@yahoo.com*
**Corresponding Author Email:sumanth81s@gmail.com*

**Abstract**

Alzheimer's disease (AD) is a type of dementia that is difficult to detect based on clinical surveillances. AD detection on brain Magnetic Resonance Imaging (MRI) data is major anxiety in the neurosciences. Conventional evaluation of efficient image scans in general relies on manual reorientation, visual reading and semi quantitative exploration in brain sections. The Feature Selection (FS) has been tackled to a greater extent since it has proved itself to be a technique that is able to solve the computational problems that are NP-hard and for finding some optimal feature subsets. The FS works by means of removing the features which are irrelevant or redundant. Here in this work, a Krill Herd Optimized Feature Selection has been proposed for the classification of the MRI images. Using the Krill Herd Algorithm (KHA) happens to be widely accepted recently. This is owing to the fact that it represents a modern optimization that is effective and is a good search process. The segregation of the images from brain MRI into either normal or abnormal is important for analysing a normal patient and considering the ones that have higher chances of abnormalities. This technique of classification known as the fuzzy classifier along with the Neural Network has been proposed for getting a better performance that was accurate.

***Keywords***: *Magnetic Resonance Imaging (MRI;, Alzheimer's Disease (AD)/ Dementia; Feature Selection (FS); Krill Herd Optimized Feature Selection; Fuzzy classifier and Neural Network.;*

## 1. Introduction

Magnetic Resonance Imaging (MRI), an imaging method, developed as a scientific modality above 30 years. Medical image processing is a significant means to categorize and identify different disorders. Medical images data employ imaging methods as Computed Tomography (CT), MRI and mammogram signifying the existence or nonexistence of lesions with patient history. MRI is a scanning device by magnetic fields and computers to confine brain images on film. Dementia is an age-related neurodegenerative disorder that precise cause is yet unfamiliar. Alzheimer's disease (AD), a basic dementia form, is provided by neurons loss and synapses in cerebral cortex and definite subcortical regions. Although the majority dementia patients are older, not all aged people undergo from dementia meaning that dementia is not part of normal ageing. Dementia can occur to other person, however frequent over 65 years. Human beings in 40s and 50s of together sexes can have dementia; however, it is more frequent in men. Dementia is of several varieties and each has its causes. A few general dementia types are: AD, Lewy Body Dementia (LBD) and Fronto Temporal Dementia (FTD).AD is a basic dementia problem. It reports for 50% to 70% of each dementia cases. It is a progressive, degenerative illness offensive the brain causing shrinkage and brain cell desertion. After that, abnormal material constructs up as "tangles" in brain cells center and as "plaques" outside brain cells. Such disrupt messages in the brain demolishing links among brain cells. The dementia diagnose are normally complete by assessment called the scientific assessment of the patient's mind condition that is none other than one's subjective health measures.

The process of Feature selection (FS) is very useful specially whenever there are datasets of a high dimension to be considered for classification process. The FS aims in reducing the size of data in order to make the computation easier since high dimension data burden the process due its complex computation. The Swarm intelligence (SI) is a technique that will be able to solve the computational problems that are NP-hard (the Non-Deterministic Polynomial time). The FS algorithms have exploited the SI based approaches in various applications.

A big problem has been identified in the choice of optimal features that can bring about a distinction among classes. The task of evaluating every available feature subset is painful owing to the very large amount of effort required in terms of computation. The Genetic Algorithm (GA) is an approach for choosing an ideal and best feature subset maintaining the accuracy of classification at the same time.

There are several other algorithms of optimization that have been adapted for enhancing the problem of text clustering that is similar to the Cuckoo Search Algorithm, a Harmony Search Algorithm, the Particle Swarm Optimization, and Genetic Algorithm. The Krill Herd (KH) algorithm has been one of the most recent ones that are based on the swarm which simulates herd behaviour for every individual of the krill. This had been introduced in the year 2012 by Gandom and Alavi for solving the problems of a benchmark function. The KH algorithm works for the purpose of finding a minimum distance of a krill individual from those foods that

have a density that is the highest. This KH algorithm has been applied successfully to several other areas of optimization like the numerical optimization, the graph-based network route, the detection of breast cancer, the neural network, the electrical power system and the data clustering [1].
A Krill Herd Algorithm is the most recent among the heuristic techniques that are applied for deriving solutions that are the best in various tasks of optimization. This algorithm finds its traits from the species called Antarctic krill (the Euphausiasuperba) normally originating from the Southern Ocean. The body of an adult will have a length of about 6 cm, with a weight of around 2 grams. The species normally feed on the phytoplankton. One major feature of the krill along with the inspiration of its current algorithm was the individual krill's capability to mould itself in a large herd which may be many hundreds of meters in terms of length [2].
These Neural Networks (NN) have been employed extensively in pattern classification based on the fact that they have no requirement for details of the classes in terms of their distribution of probability and their a-priori probabilities. The system based on the NN classification has imitated the thinking of humans in some of the scenarios thus providing decisions for the class that demonstrates all possibilities of certain other infections. For the classification of the MR images captured from brain to be either normal or abnormal, the method employs the Artificial Neural Network (ANN) for categorizing the inputs within the collection of their target categories (as either normal or as abnormal) according to the constraints and parameters of feature extraction [3].
Feature Selection-based Krill Herd along with the GA has been proposed for a much better classification of the images of the MRI. The literature related to this proposed work have been explained in Section 2. The details on different techniques approached in this work are explained in Section 3. The results obtained are discussed in Section 4 and the conclusion is duly made in Section 5.

## 2. Related Works

Hecibeyaglou et al., [1] had proposed another novel algorithm based on Particle Swarm Optimization algorithm which was multi-mean for the multilayer feed-forward and training of the artificial neural networks. The solution space can be explored more efficiently in Multi-Mean Particle Swarm Optimization using multiple swarms and can identify some better solutions in place of the Particle Swarm Optimization. Evaluation of the performance of the new multi-mean Particle Swarm Optimization that has been proposed was based on the experiments conducted based on the benchmark datasets. The analysis of these results had demonstrated this algorithm known as multi-mean particle swarm optimization had been performing well and can be rightly adopted to be a novel technique to train the multilayer feedforward Artificial Neural Networks.
Zhang et al., [2] had brought about a proposal of a multilevel thresholding method of colour image segmentation by using another new modified ABC algorithm. The non-availability of information guidance at the time of search has been showing that the ABC algorithm with a poor capacity for local optimization. For fixing this issue, the Krill Herd Algorithm had to be employed. Using the improved local search capability, the KABC will be able to get a faster and better convergence. The results proved that a KABC algorithm enhanced the accuracy of optimization along with speed of convergence and it also proved its effectiveness in robustness.

Bhonsle et al., [4] had proposed another new framework that was used for the de-noising of medical images using an adaptive wavelet thresholding along with a technique of Total Variation in parallel having results fused with one another with a technique of wavelet-based fusion. These were some fused images that were noiseless and consequential that were subject to the restoration of images by means of employing efficiently a new form of hybridized Krill Herd Algorithm along with the Richardson-Lucy approach. The proposed framework performed well, and this was evaluated for both of these images and there was a comparison that was made to the other techniques that have been in existence. The results of the experiment have proved that this proposed framework was able to outperform all the other currently existing techniques in relation to the parameters of performance evaluation.
Zeng et al., [5] presented AD diagnosis that comprise MRI images pre-processing, feature extraction, Principal Component Analysis (PCA), and the Support Vector Machine (SVM) model. In specific, a novel Switching Delayed Particle Swarm Optimization (SDPSO) algorithm had been presented to optimize the SVM parameters. The established framework in accordance with the SDPSO-SVM model is adapted effectively to the AD classification and MCI through MRI scans from ADNI dataset. The proposed algorithm had attained great classification accuracies for six typical cases. moreover, experiment consequences reveal that the presented algorithm performed well on different SVM models and two other state-of-art means with deep learning embedded, thus serving as an efficient AD diagnosis method.
Casillas et al [6] had made a presentation of a feature selection that was genetic which was duly integrated within the genetic algorithm approach algorithm for obtaining a fuzzy rule-based classification system (FRBCS) efficiently. This classification system had a collection of comprehensible fuzzy rules having an ability of high-classification. The process proposed will allocate apriorily the number of chosen features of the candidate fuzzy rules. There was an experimentation that was carried out by making use of a Sonar example base that showed an improvement which was quite significant that was achieved by means of adding the proposed process of feature selection to a multi-stage method of genetic learning or even other methods.

Wang et al., [7] had further introduced another new chaos theory within the process of KH optimization aiming at accelerating the speed of global convergence. There have been several chaotic maps that have been considered in this method of chaotic KH (CKH) for adjusting the primary displacements of a Krill in the course of optimization. There are many problems that have been used for evaluating the behaviour of the CKH. The performance of the proposed method with a suitable chaotic map has been found to be better or even comparable with a KH or certain other robust approaches as portrayed in the results.
In recent times, there is an algorithm of optimization that was based on the Krill Herd (KH) that had been exploited for different applications and this was even more accurate compared to other techniques which were state-of-the-art. An adaptation of KH based on binary representation of the KH technique was proposed by Rodrigues et al [8] and this was validated for the purposes of feature selection in different datasets. These experiments had shown that the technique proposed could outperform all other metaheuristic approaches and was also found to be faster.
The Krill Herd Algorithm (KHA) includes several interesting procedures. A comparison of the KHA algorithm was used for the classification tasks including Artificial Neural Network (ANN) with many different methods which were heuristic were made by Kowalski and Łukasik [9]. This method of the ANN training was verified for the task of classification and for this, there were some benchmark examples that were utilized from the widely available UCI Machine Learning Repository. The evaluation of the training was based on the criterion such as the Classification Error and the Sum of Square Errors. There was a conclusion made that an application of the KHA was able to offer some assuring performance cited with the metrics mentioned above and the time required for the training of the ANN. An enhanced brain image classifier of the MRI that targeted the primary objectives which were the first to get the maximum accuracy of classification and the next was to bring down the features used for classification was proposed by Joans and Sandhiya [10]. Here, the feature selection had been

performed by using the Genetic Algorithm (GA) and the Random Forest Classifier.

## 3. Methodology

The section provides details on the feature selection for classifying dementia based on the Krill Head Optimizer and the Genetic Algorithm.

### 3.1. Feature selection utilizing Genetic Algorithm

In various problem of classification, the features may be large with most of them redundant or irrelevant. The classification is improved by feature reduction and this is done based on a processing goal and a criterion of feature evaluation which was the accuracy of classification. The features of the pattern, from the viewpoint of the type of goal, can be irrelevant (without an effect on the performance of processing) or the relevant ones (with an impact on the performance of processing) or the redundant ones (both dependent and correlated). Thus, for the reduction of the features to one smaller set of the features applying the method of GA-based global search. The GA is the fittest principle of Darwin stating the initial population of the individuals which evolve through the process of natural selection in a way in which the chances of survival is higher for the fittest individuals [11]. The population of the matrices of the competing feature transformation is maintained by the GA.

For the evaluation of every matrix in the population, there are some input patterns that have been multiplied. There is a classifier dividing the patterns into their training set which is used for training classifiers and testing sets used for evaluating the accuracy of classification. The accuracy has been obtained and also returned to a GA to be a measure of the quality of a matrix of transformation for obtaining the transformed patterns. By making use of this information, a GA can search for a new transformation to minimize the dimensionality. The gene is represented as an element in the vector. The vector is in binary representation and hence every bit will be mapped to a feature.

There is a value used for measuring the fitness of chromosomes and then deciding if the chromosome is suitable for the population. The GA also uses three different populations in the process of genetics that has been created randomly. The GA also makes use of three operators which are the reproduction, the crossover, and the mutation. The process in the GA also eradicates the low fitness chromosomes and retains those with high fitness and hence the high fitness chromosomes are passed on to the subsequent generation. The evaluation of the fitness and selection of chromosomes to the next generation takes place until a best individual (or chromosome) has been identified.

The basic pseudocode for Feature selection abstracted from Genetic algorithm has been depicted in Figure 1.

```
Initialize k = 0
Initialize the individuals in the population as P(k)
Evaluate the fitness function for all the individuals in P(k)
While {termination condition is not satisfied}
Do k = k+1 {iterate}
Selection ( )  {select individuals with better fitness}
Crossover ( ) {amalgamation of parent to form new individuals}
Mutation ( ) {Modifying bits in the vector like bit flip}
End while
```

**Fig 1** Pseudocode for the Genetic Feature selection algorithm [12]

The main steps in the algorithm is the initialization process where the population is initialized based on the dataset, the selection process where the individuals are picked based on survival of the fittest, the crossover process where the traits of the individuals are combined, the mutation process where the traits are either deleted or newly added and the termination process. A formula for the calculation of the fitness function has been given below:

The Fitness f(x) = the fitness of that of an individual (i)/sum of the fitness for all the individuals f(i),

### 3.2. Feature selection based on Krill Head optimizer

An attempt has been made by feature selection for identifying one of the most representative among the subsets of the features within the dataset by means of removing all redundant and useless features for the process of classification, thus decreasing the time taken for training and the error of classification for the datasets. But, the process of search being exhaustive for optimal features is quite impracticable owing to the size of the space which is about $2^n$, in which n denotes the actual number of features.

Krill Herd (KH) imitates the behaviour of the Antartic Krill individual based motion. There are two goals to this process which are: (i) increasing the density and (ii) finding food. The Krill uses these goals as the objective for herding through a solution that is the global best. Krill individuals will move to the best solution while searching for the food of a high density. This means, the closer its distance to food of high density, the lower is the objective function. The Krill Herd (KH) is an optimization which is swarm-based and was introduced initially by Gandomi and Alavi [13], duly motivated by the herding behaviour of the Krills adapted to the environmental and biological conditions. A time-dependent position of an individual krill represented as xi in a search space has been administrated by three actions which are: i) the movement induced by that of all the other krill individuals, (ii) a foraging activity, (iii) a random diffusion given as per equation (1):

$$x_i(t+\Delta t) = x_i(t) + \Delta t(N_i + F_i + D_i) \quad (1)$$

Wherein,

$$\Delta t = C_t \sum_{j=1}^{d} (UB_j - LB_j) \quad (2)$$

Where the $N_i$ indicates a movement due to the influence of the another krill individual, $F_i$ denotes the second action due to the foraging activity called as the foraging motion, $D_i$ denotes the third action due to the physical diffusion of that of the ith krill individual, and d denotes the variables needing optimization, wherein the $UB_j$ and the $LB_j$ denote the upper and the lower bounds of that of the jth variable, and $C_t$ denotes a constant number found in [0,2]. The movement of the Krill individuals is due to the mutual effect and at the same time they attempt at maintaining a higher density. The direction of this induced motion that considers an individual i, αi, that are estimated from a local swarm density (the local effect), a target swarm density (a target effect), and finally a repulsive swarm density (the repulsive effect). The movement of the krill individual i is defined as per equation (3) [14]:

$$N_i^{t+1} = N^{\max} * \alpha_i + \omega_n * N_i^t \quad (3)$$

Wherein,

$$\alpha_i = \alpha_i^{local} + \alpha_i^{target} \quad (4)$$

And the $N^{\max}$ denotes a maximum value of the induced speed, the $\omega_n$ denotes an inertia weight for the motion induced given in the range as in [0,1], and the $N_i^{old}$ will be the final motion that is induced. A foraging motion has been formulated based on the location of food and the experience on the previous food location that is expressed for that of the ith krill individual as per (5):

$$F_i^{t+1} = V_f * \beta_i + \omega_f * F_i^t \tag{5}$$

Wherein,

$$\beta_i = \beta_i^{food} + \beta_i^{best}$$

$V_f$ denotes the speed of foraging, $\omega_f$ denotes its inertia weight of a foraging motion as in [0,1], $F_i^{old}$ denotes the final last foraging motion, $\beta_i^{food}$ indicates the food attractive and the $\beta_i^{best}$ denotes the best fitness of that of the ith krill until now. A random process is depicted for the physical diffusion and is expressed as a maximum speed of diffusion with a random vector as per equation (6):

$$D_i = D^{\max} * (1 - \frac{1}{I_{\max}}) * \delta \tag{6}$$

In which the $D^{\max}$ denotes a maximum diffusion speed, the random vector takes the range [−1,1] denoted by δ and the iteration is represented as I.

### 3.3. Classification Algorithm

A classification denotes the method used for identification and the discrimination of the patterns or objects based on attributes done with supervised learning.

#### 3.3.1. Neural Network

The Neural Network (NN) system has a layered architecture with the input layer, output layer, and either one or more hidden layers. Every node is connected from and to a hidden layer. The input layer indicates raw information fed to the network. The network does not change its values and the hidden layer receives its data from its previous layer which is the input layer. The input values from the input layer are modified in the hidden layer using activation function and the weights available between the layers. This is further propagated to the next layer and the same modification process is carried out between the hidden layer and output layer using the connection weights. The output of the network is obtained from the output layer after the application of activation function [15].

Selecting nodes for every layer is dependent on the NN problem that attempts to solve the data network that deals with the quality of this data along with certain other values. The training dataset influence the size of the network based on the nodes in the input and output layer. The manner in which the NN is controlled is by means of adjusting the weights between the weights. The initial weights will be set at the random numbers adjusted at the time of NN training. The processing of the network is done in two passes, the forward and backward pass. The process of evaluating the deviation of the actual output with the target is done in the forward pass and this is the error value in the network. The aim of the training is to bring down this error and hence the process of reducing the error by adjusting the connection weights is done in the backward pass. The network is trained until the desired output is targeted or produces minimum error.

#### 3.3.2. Fuzzy Classifier

On the basis of an object description, the class label to the object may be assigned by the classifier. The classification will be part of the general application field for machine learning and pattern recognition. The input pattern classification will depend on the maximum degree of association measured. For similarities, there is a random pattern used. The fuzzy system objectives like interpretability and accuracy will be based on the degree of maximum association of the pattern classified within a class corresponding to a rule base that is high in dimension. At the time of designing a fuzzy classifier, accuracy and interpretability will be taken to be the primary objectives that are opposite to one another which means one objective may be enhanced at the cost of that of the other.

#### 3.3.3. Fuzzy based Krill Head optimizer

There is a metaheuristic algorithm which reveals a performance which is high and can gain a proper level of balance between exploitation and exploration keeping in mind the progress made to solve the problem. For every step that solves this problem, there may be various conditions happening and the algorithm may also deviate from that of its original to an optimal solution and goes forward directly to the optimal solution. For every metaheuristic algorithm like the KH optimization, there have been certain parameters that determine the amount of contribution by the local and the global search. In the KH algorithm, the inertia weights which are wf and wn have been introduced. The fuzzy rule based on this FKH algorithm has been depicted in Figure 2.

1. If counter is low and Delta is low then Wf is high and Wn is high
2. If counter is low and Delta is middle then Wf is middle and Wn is middle
3. If counter is low and Delta is high then Wf is high and Wn is high
4. If counter is middle and Delta is low then Wf is high and Wn is high
5. If counter is middle and Delta is middle then Wf is middle and Wn is middle
6. If counter is middle and Delta is high then Wf is high and Wn is high
7. If counter is high and Delta is low then Wf is low and Wn is low

**Fig 2**: Fuzzy rule-base of the proposed FKH algorithm

For covering all these probabilistic conditions at the time of the process of search, based on the stochastic nature of these metaheuristic algorithms, one more factor is required for informing the system of fuzzy controller regarding the unwanted and undesirable situations that need correction.

## 4. Results and Discussion

The data that is used for evaluating the algorithms has been got from an Alzheimer's disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). Table 1 to 3 and figure 3 to 5 shows the results evaluated for classification accuracy, sensitivity and specificity respectively for GA feature selection – Fuzzy classifier, GA feature selection -Neural Network, Feature Selection Krill Herd - Fuzzy Classifier and Feature Selection Krill Herd - Neural Network.

**Table 1:** Classification Accuracy for Feature Selection Krill Herd - Neural Network

| Techniques Used | Classification Accuracy |
|---|---|
| GA feature selection - Fuzzy Classifier | 87.3 |
| GA feature selection -Neural Network | 88.27 |
| Feature Selection Krill Herd - Fuzzy Classifier | 89.25 |
| Feature Selection Krill Herd - Neural Network | 90.23 |

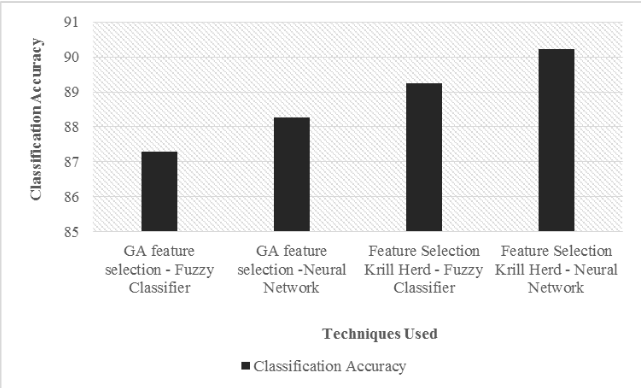


**Fig. 3:** Classification Accuracy for Feature Selection Krill Herd - Neural Network

It is observed from table 1 and figure 3 that the classification accuracy in Feature Selection Krill Herd - Neural Network performs better by 3.3%, by 2.2% and by 1.1% than GA based Feature Selection – Fuzzy Classifier, GA based feature selection – Neural Network and Feature Selection Krill Herd – Fuzzy Classifier respectively.

**Table 2** Sensitivity for Feature Selection Krill Herd - Neural Network

| Techniques Used | Sensitivity for Normal | Sensitivity for AD |
|---|---|---|
| GA feature selection - Fuzzy Classifier | 0.8898 | 0.8065 |
| GA feature selection - Neural Network | 0.898 | 0.8226 |
| Feature Selection Krill Herd - Fuzzy Classifier | 0.9061 | 0.8387 |
| Feature Selection Krill Herd - Neural Network | 0.9143 | 0.8548 |

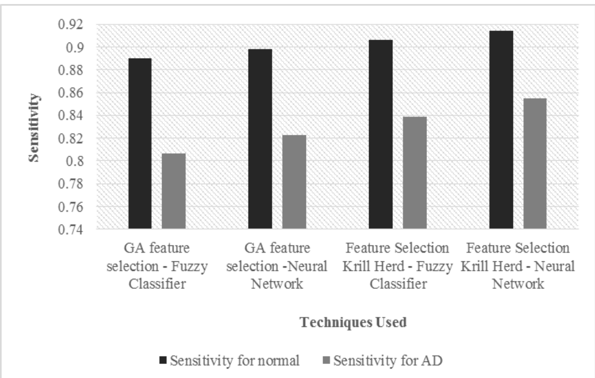


**Fig 4:** Sensitivity for Feature Selection Krill Herd - Neural Network

It is observed from table 2 and figure 4 that the sensitivity for the Normal case in Feature Selection Krill Herd - Neural Network performs better by 2.72%, by 1.79% and by 0.9% than GA based Feature Selection – Fuzzy Classifier, GA based feature selection – Neural Network and Feature Selection Krill Herd - Fuzzy Classifier respectively. Similarly, the sensitivity of AD for Feature Selection Krill Herd - Neural Network performs better by 5.81%, by 3.84% and by 1.9% than GA based Feature Selection – Fuzzy Classifier, GA based feature selection – Neural Network and Feature Selection Krill Herd - Fuzzy Classifier respectively. The overall performance of sensitivity for Normal is better than sensitivity for AD.

**Table 3:** Specificity for Feature Selection Krill Herd - Neural Network

| Techniques Used | Specificity for Normal | Specificity for AD |
|---|---|---|
| GA feature selection - Fuzzy Classifier | 0.8065 | 0.8898 |
| GA feature selection - Neural Network | 0.8226 | 0.898 |
| Feature Selection Krill Herd - Fuzzy Classifier | 0.8387 | 0.9061 |
| Feature Selection Krill Herd - Neural Network | 0.8548 | 0.9143 |

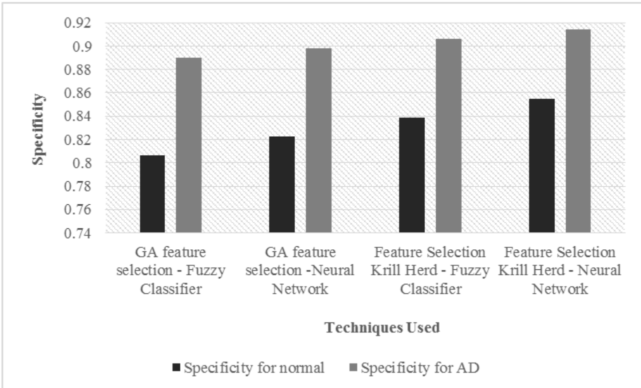


**Fig 4:** Specificity for Feature Selection Krill Herd - Neural Network

It is observed from table 2 and figure 5 that the specificity for the Normal case in Feature Selection Krill Herd - Neural Network performs better by 5.81%, by 3.84% and by 1.9% than GA based Feature Selection – Fuzzy Classifier, GA based feature selection – Neural Network and Feature Selection Krill Herd - Fuzzy Classifier respectively. Similarly, the specificity of AD for Feature Selection Krill Herd - Neural Network performs better by 2.72%, by 1.79% and by 0.9% than GA based Feature Selection – Fuzzy Classifier, GA based feature selection – Neural Network and Feature Selection Krill Herd - Fuzzy Classifier respectively. The overall performance of specificity for Normal is lower than specificity for AD.

## 5. Conclusion

This work presents a framework for classifying MRI images for Dementia. Dementia, an age-related cognitive decline is indicated by degeneration of cortical and sub-cortical structures. A feature selection will attempt to identify a represented feature subset within the dataset by means of removing all the useless and redundant features. On the other hand, the role of GA is to optimise the selection of the best among the features through the fitness function. The work also proposed the Krill Herd algorithm that was based on a feature selection to classify the MRI images better. The results have proved that the accuracy of classification of the Feature Selection Krill Herd - Neural Network has performed better by about 3.3%, 2.2%, and 1.1% compared to the GA based Feature Selection – the Fuzzy Classifier, the GA based feature selection – the Neural Network and the Feature Selection Krill Herd - Fuzzy Classifier.

## References


[1] Hacibeyoglu, Mehmet, and Mohammed H. Ibrahim (2018). "A Novel Multimean Particle Swarm Optimization Algorithm for Non-

linear Continuous Optimization: Application to Feed-Forward Neural Network Training." Scientific Programming.

[2] Zhang, Sipeng, Wei Jiang, and Shin'ichi Satoh (2018). "Multilevel Thresholding Color Image Segmentation Using a Modified Artificial Bee Colony Algorithm." IEICE Transactions on Information and Systems 101, no. 8: 2064-2071.

[3] Kowalski, Piotr A., and SzymonŁukasik (2015). "Experimental study of selected parameters of the krill herd algorithm." In Intelligent Systems' 2014, pp. 473-485. Springer, Cham.

[4] Bhonsle, Devanand, Vivek Kumar Chandra, and G. R. Sinha (2018). "An Optimized Framework Using Adaptive Wavelet Thresholding and Total Variation Technique for De-noising Medical Images." 953-965.

[5] Zeng, N., Qiu, H., Wang, Z., Liu, W., Zhang, H., & Li, Y. (2018). A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. Neurocomputing.

[6] Casillas, J., Cordón, O., Del Jesus, M. J., & Herrera, F. (2001). Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional problems. Information Sciences, 136(1-4), 135-157.

[7] Wang, Gai-Ge, Lihong Guo, Amir H. Gandomi, Guo-Sheng Hao, and Heqi Wang (2014). "Chaotic krill herd algorithm." Information Sciences 274: 17-34.

[8] Rodrigues, D., Pereira, L. A., Papa, J. P., & Weber, S. A. (2014, August). A binary krill herd approach for feature selection. In Pattern Recognition (ICPR), 2014 22nd International Conference on (pp. 1407-1412). IEEE.

[9] Kowalski, P. A., &Łukasik, S. (2016). Training neural networks with krill herd algorithm. Neural Processing Letters, 44(1), 5-17.

[10] Joans, S. M., & Sandhiya, J. A Genetic Algorithm Based Feature Selection for Classification of Brain MRI Scan Images Using Random Forest Classifier. International Journal of Advanced Engineering Research and Science, 4(5).

[11] Kharrat, A., Gasmi, K., Messaoud, M. B., Benamrane, N., &Abid, M. (2010). A hybrid approach for automatic classification of brain MRI using genetic algorithm and support vector machine. Leonardo journal of sciences, 17(1), 71-82.

[12] Vaishali, R., Sasikala, R., Ramasubbareddy, S., Remya, S., &Nalluri, S. (2017, October). Genetic algorithm based feature selection and MOE Fuzzy classification algorithm on Pima Indians Diabetes dataset. In Computing Networking and Informatics (ICCNI), 2017 International Conference on (pp. 1-5). IEEE.

[13] [A. H. Gandomi and A. H. Alavi, (2012) "Krill herd: A new bio-inspired optimization algorithm," Communications in Nonlinear Science and Numerical Simulation, vol. 17, no. 12, pp. 4831–4845.

[14] Abualigah, Laith Mohammad, AhamadTajudinKhader, Mohammed Azmi Al-Betar, and Mohammed A. Awadallah. (2016) "A krill herd algorithm for efficient text documents clustering." In Computer Applications & Industrial Electronics (ISCAIE), 2016 IEEE Symposium on, pp. 67-72. IEEE.

[15] Cilimkovic, M. (2015). Neural networks and back propagation algorithm. Institute of Technology Blanchardstown, Blanchardstown Road North Dublin, 15.