An Improved Attribute Subset Selector for Alzheimer's Disease Prediction

S. Sarumathi[1,*], N. Reshma[1], Sharmila Mathivanan[2], and S. Malarkhodi[1]

[1]K. S. Rangasamy College of Technology, Tiruchengode, Tamil Nādu, India
[2]FPT Greenwich University, Ho Chi Minh City, Vietnam
Email: sarumathi@ksrct.ac.in (S.S.); reshmafarhana17@gmail.com (N.R.); SharmilaM@fe.edu.vn (Sh.M.); malarkhodi@ksrct.ac.in (S.M.)
*Corresponding author
Manuscript received June 25, 2023; revised September 15, 2023; accepted May 14, 2024; published August 28, 2024

***Abstract*—Alzheimer's Disease (AD) is one of the prevalent diseases which is a neurological condition that impairs brain activities like reading, writing, thinking, and remembering. The death rate due to AD would be reduced by providing proper treatment based on the stage of the disease. This can be determined by using data mining techniques. A data mining technique, Binary version of the Artificial Bee Colony (BABC) algorithm was proposed to choose the best features from statistical and volumetric information of Magnetic Resonance Images (MRIs) of the brain. However, the accuracy of BABC is low due to slow convergence. So, in this article, an Improved Artificial Bee Colony (IABC) algorithm is introduced to enhance the AD prediction accuracy. It can be achieved by improving the exploration and exploitation process of BABC. In the employee bee phase of IABC, a novel search equation is used that enhances the probabilities for onlookers' bees to determine the best positions and change the number of bad ones by the fresh ones in the following phase. Furthermore, Particle Swarm Optimization (PSO) is utilized to create a fresh position changing an un-updated location in the scout bee phase. IABC is enhancing the AD prediction efficiency and interpretability by identifying the most relevant predictors, reducing dimensionality, and improving model generalization for the AD prediction. Furthermore, it also improves the exploration-exploitation process of feature selection. From the empirical findings, it is proved that the proposed IABC with Random Forest (IABC-RF) has 10.52%, 8.57%, 7.87%, and 6.8% better accuracy, precision, recall and F-measure than BABC with K-Nearest Neighbor (BABC-KNN) for the AD prediction.**

***Keywords*—Alzheimer Disease (AD), Artificial Bee Colony (ABC) algorithm, feature selection, Improved Artificial Bee Colony (IABC) algorithm**

I. Introduction

There are many people throughout the world who are affected by Alzheimer Disease (AD) [1] which is an incurable type of memory loss. It is described as a gradual neuropsychological disorder with a profound impact on mindset and behavior functions, which significantly lowers the life expectancy for the AD patients and causes significant mood swings. Early diagnosis may facilitate access to the services and support resources, as well as medication for symptoms. They will be able to participate in choices about their welfare, residential circumstances, finances, and legal issues as a result. When given a prompt diagnosis, patients are frequently able to take part in this preparation and choose who will make financial and medical choices on their behalf as the condition progresses [2, 3].

Nowadays, AD identification is still not reliable till the patient reaches the noticeable phase of AD. One of the reasons for this is having a massive amount of high-dimensional data which results in the complexity of analysis. To overcome this problem, approaches such as data mining can be used to effectively predict AD. Data mining [4, 5] is the process of determining patterns and knowledge and patterns from huge volumes of data. It involves the utilization of statistical models and algorithms to identify knowledge and patterns. Here are some commonly used data mining techniques that can be applied in AD disease prediction:

Data pre-processing: Before applying data mining techniques, it is important to pre-process the data [6]. This involves cleaning the data, handling missing values, and transforming the data into a suitable format for analysis. Feature selection: Feature selection [7] is the process of identifying the most relevant features or variables that contribute significantly to the AD prediction. Techniques such as correlation analysis, information gain, and principal component analysis can be used to select the most informative features. Classification algorithms: Once the relevant features are identified, various classification algorithms [8] can be applied to build predictive models. Some commonly used algorithms include decision trees, random forests, Support Vector Machines (SVM) [9], logistic regression, and artificial neural networks [10]. These algorithms learn from the available data and generate a model that can predict the presence or absence of AD based on the selected features.

Cross-validation and model evaluation: It is essential to evaluate the performance of the predictive models to ensure their reliability. Cross-validation techniques, such as k-fold cross-validation, can be used to assess the model's performance on different subsets of the data. Metrics like accuracy, sensitivity, specificity, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) can be used to evaluate the model's performance. Ensemble methods: Ensemble methods [11] combine multiple predictive models to improve prediction accuracy. Techniques such as bagging, boosting, and stacking can be employed to create an ensemble of models that collectively make predictions.

Validation with independent datasets: Once the predictive model is built, it is important to validate its performance on independent datasets. This helps to ensure that the model's predictive ability is not limited to the dataset on which it was trained. Interpretation and feature importance: Data mining techniques can also provide insights into the most important features contributing to the AD prediction. This can help researchers and clinicians better understand the underlying factors and mechanisms associated with the disease.

Feature selection is one of the processes of data mining which reduces the dimensionality of data for reducing the

DOI: 10.7763/IJCTE.2024.V16.1357

complexity and overfitting of the classification model. Binary version of the Artificial Bee Colony (BABC) algorithm was used as a feature selector to classify brain volumetric data [12]. In BABC, the source of food was considered as a possible fine-tune feature subset for the AD prediction and the source of food was generated as 0's and 1's. Then, the processes were continued to select the most important features in the AD Neuroimaging Initiative (ADNI) database. The selected features were processed in three different classifiers for the prediction of AD. However, the accuracy of BABC is low due to premature convergence and slow convergence.

So, in this article, the Improved Artificial Bee Colony (IABC) algorithm is proposed to enhance the AD prediction accuracy. At first, quantitative and analytical information of brain Magnetic Resonance Imaging (MRI) images are gathered from volBrain. After that, the images are given as input to IABC where a novel search equation is used in the employee bee phase to enhance the probabilities for onlooker bees to determine the best position. Moreover, several bad ones are replaced by good ones in the onlooker bee phase. In the scout bee phase, Particle Swarm Optimization (PSO) is used to create a novel position changing an un-updated position. The selected features by IABC are processed in Random Forest (RF), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) for the prediction of AD effectively.

This paper's related works are presented in the next section. The suggested method is presented in Section III. The results, discussion, and performance of the suggested system are shown in Section IV. Ultimately, we conclude the study with some thoughts in Section V.

II. Literature Survey

A search method [13] was introduced for the AD prediction by finding complex biomarkers. This method uses association rule mining techniques to find out combinatorial biomarkers which include psychiatric test information and clinical laboratory information. Many new relations were identified by using the search method. However, this method is still not reliable in predicting AD.

Feature selection [14] method and classifier were employed for early diagnosis of AD. Eight feature selection algorithms [15, 16] were used in the pre-processing stage for developing classifiers. In addition to this, the selected features were evaluated regarding previous clinical findings for analyzing the efficiency of feature selection methods to choose biologically consistent features. Through the utilization of the feature selection method, the classifier achieves an accuracy of over 90% while also reducing the dimensionality of the data in AD. A large number of individuals would be included to validate the findings.

A feature selection method [17] was proposed for the prediction of AD. Here, feature selection was carried out independently for each modality and coupled group-sparsity regulazier with a joint choice of common characteristics among various modalities. This method compensates for the deficiencies in the conventional subject-based multi-modal feature selection method and thoroughly takes into account the connection between feature nodes and the local geometric structure of the feature space. However, certain additional modalities that were not focused on in this study could include significant data that might substantially boost generalization ability.

A multi-view feature selection method [18] was proposed for the early diagnosis of AD. Initially, the features were clustered, and then unnecessary features were eliminated using lasso learning. Lasso learning is a machine learning technique that performs both feature selection and regularization by adding a penalty to the absolute value of the coefficients, encouraging sparse solutions where irrelevant features are set to zero. The selected features were used in the classification model for the diagnosis of AD. This algorithm not only eliminates data outliers' interference but also diminishes the need for storage space. This algorithm will be improved by using local structure learning.

A combined method [19] was proposed for the volumetric feature-based AD diagnosis. In the combined method, a two-stage ensemble Hough deep learning method was utilized for the automatic localization of the right and left hippocampi. Then, 2D slices were extracted from 3D patches that were trained in deep learning diagnosis of AD. It operates entirely on autopilot and achieves a greater level of accuracy compared to alternative approaches. However, the computational complexity of this method is high.

A novel multi-class framework [20] was introduced for the prediction of AD using fusion based on multimodal neuroimaging and embedding feature selection methods. The optimization process used in the framework was convergence to a global optimum. However, the space complexity of this framework is high when the number of subjects is large. An AD early diagnostic method [21] based on different learning methods for the prediction of AD. This method has the potential to significantly improve the accuracy of AD classification. The precision would be increased marginally by using a Principal Component Analysis (PCA)-based feature range.

Cognitive-based 3-tiered machine learning [22] was developed for predicting Mild Cognitive Impairment (CGI) and AD. Even though this method has better accuracy; the final AD prediction of the total subjects was affected because of the limited number of subjects studied.

An efficient feature selection [23] method was proposed for the detection of the AD patients. This method combined the greedy searching heuristic and the Fisher score ranking method for the prediction of AD. The features that are chosen by the efficient feature selection method were processed in SVM and KNN to detect the AD patients. This method provides better sensitivity and specificity than other methods. However, this method finds a better minimal set of features but not the best set.

A kernel-based approach [24] was suggested for conducting Functional Connectivity Analysis and Channel Selection based on Electroencephalogram (EEG) data in the context of AD. Nevertheless, while this approach maintains consistency, there is room for additional evaluation of its resilience against volume conduction effects in EEG. A method [25] for AD classification was introduced that utilized feature fusion, resulting in the highest accuracy; nonetheless, this method exhibits a high level of computational complexity.

III. Proposed Methodology

In this section, the Particle Swarm Optimization (PSO) method used in the proposed method is described and the proposed IABC method is described in detail to select the most important features for better identification of Alzheimer's disease. Initially, the statistical and volumetric data of brain Magnetic Resonance Images (MRI) are collected from volBrain. Then, the collected images are given as input to the proposed BMABC method which selects the most appropriate features to identify AD. The selected features are processed in KNN, RF, and SVM to determine the AD effectively. Fig. 1 shows the block diagram of IABC-based AD prediction process.

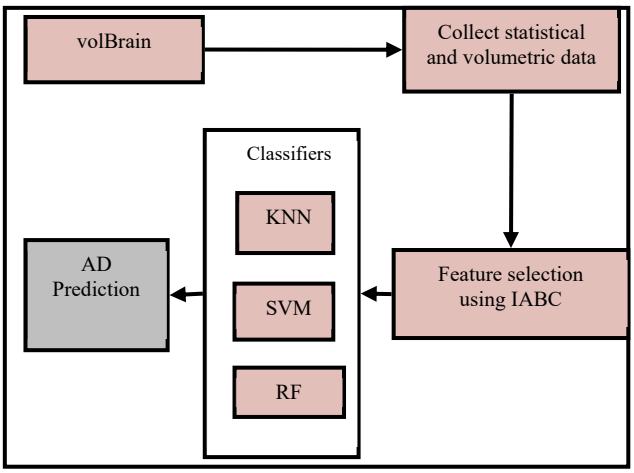

Fig. 1. Block diagram of IABC-based AD prediction process.

A. Feature Selection and Particle Swarm Optimization

The process of choosing a subset of essential features from a broader collection of attributes to employ in the design of a model is known as feature selection. The main intention of the feature selection process is to enhance the accuracy and interpretability of the classification model, as well as to reduce the complexity and overfitting of the model. One of the best optimization methods is Particle Swarm Optimization (PSO) which draws inspiration from the individual interaction and continuous activities and interaction of fish, birds, and insects. This algorithm starts with initializing the population by a user. Each swarm in the population randomly chooses the features from the dataset. The key concept of PSO is dealing with changes in velocity. For ith particle in the d-dimension, it could update its velocity and position by using the following Eqs. (1) and (2).

$$\mathcal{V}_{id} = w_{id}\,\mathcal{V}_{id} + b_1 s_1(p_{id} - \mathcal{X}_{id}) + b_2 s_2(p_{id} - \mathcal{X}_{id}) \quad (1)$$

$$\mathcal{X}_{id} = \mathcal{X}_{id} + \mathcal{V}_{id} \quad (2)$$

In Eq. (1), $\mathcal{V}_{id}$ is the velocity of the ith particle, $w_{id}$is the inertia weight, b_1 , and b_2 are the cognitive learning parameter and s_1 and s_2 are the social collaboration parameter respectively, $\mathcal{X}_{id}$ means the position of the ith particle, $p_i = (p_{i1}, p_{i2}, \ldots, p_{id})$ defines the best previous position (i.e., the position with the highest fitness value). By using Eqs. (1) and (2) an optimized solution is obtained for the given problem.

B. Data Collection

An online system called volBrain is used to collect the brain's volumetric data through MRI images. This system accepts the original MRI images in Neuroimaging Informatics Technology Initiative (NIFTY) format and pre-processes those images and produces an article in Portable Document Format (PDF) and Comma Separated Values (CSV) formats highlighting the volumetric outcome by creating an autonomous segmentation of the brain from the data. The dataset used in this proposed work is generated by integrating the articles generated. The volBrain system's output includes the proportion of each brain component, the proportion of components according to the total brain area, and the asymmetry of the groups depending on the right and left sides.

C. Feature Selection Using Improved Artificial Bee Colony Algorithm

The data from volBrain is given as input to the IABC algorithm where the feature selection process is carried out to fasten the classifier model and enhance the performance of determination of AD. The feature selection process also cut off the unnecessary, noisy, irrelevant, and inessential features in the given data. The IABC algorithm is inspired by honeybee swarm behavior which consists of three types of population bees employed bees, onlooker bees, and scout bees all working together to choose the best features for the classification of AD.

Initially, the population is created based on Search Space Division (SSD). For the employed bee phase, the search equation of BABC is enhanced by steadily utilizing the details of best features to hasten the search. For the onlooker bee phase, twenty-five percent of employed bees are chosen with the same likelihood for extra search moves. After that, five percent of bad positions are interchanged with the new ones built through utilizing the details of the present best features and the count of the best features as the scaling factor for presenting long-distance moves in the situation that more number of the best features are found out for multimodal functions.

For the scout bee phase, PSO is applied to build the new position by passing an un-updated position to a new position based on the distance to a better solution which is the best subset of features for the prediction of AD. The overall process of IABC for feature selection is described as follows:

1) Initialization phase

Here, a search space division is used to provide high-quality initial solutions, i.e., selecting a subset of features and creating the h^{th} food source t_h.

$$t_{h,g} = L_g + \frac{(\emptyset_{h,g} + 2h - 1)(U_g - L_g)}{2M} \quad (3)$$

In Eq. (3), $h = 1,2,\ldots,M$ and $g = 1,2,\ldots,D$, where M is the number of bees, L_g and U_g are the lower and upper bounds for the dimension g, $\emptyset_{h,g}$ is a random number in the range of (0,1) and D denotes the number of features.

2) Employed bee phase

A new food source v_h is obtained by using the best position t_{best} which is described as follows:

$$v_{h,g} = \begin{cases} t_{best,g} + \emptyset_{h,g}(t_{h,g} - v_{k,g}); & g = g^* \\ t_{h,g}; & g \neq g^* \end{cases} \quad (4)$$

In Eq. (4), k is randomly selected from 1 to M such that $k \neq h$, g^* is randomly selected from 1 to D and $\emptyset_{h,g}$ is a random number in [–1,1]. When $fitness(v_h) < fitness(t_h)$, t_h is replaced with v_h. On the other hand, t_h is returned. Here a fitness value is the classification accuracy.

3) Onlooker bee phase

The decisions in onlooker bees are made based on probability values which are assigned as $p(h) = 0.25$, when$(b_1, b_2, s_1, s_2) < p(h)$, a new position v_h is created by using Eq. (4). When $fitness(v_h) < fitness(t_h)$, t_h is replaced with v_h. On the other hand, t_h is returned. The following equation is used to replace the worst population with the new ones.

$$t_{x_l} = b\left[h_{best} + \emptyset_l(t_{x_l} - t_{q_1}) + \delta_l(t_{best} - t_{q_2})\right] \quad (5)$$

In Eq. (5), $x_l, l = 1,2, \ldots, [0.05M]$ are the indexes of 5% worst positions, q_1 and q_2 are indexes that are randomly selected from 1 to M in the case of $q_1 \neq q_2 \neq x_l$ for all $l, \emptyset_l$ and δ_l are the random numbers in the range of [–1,1], and b represents the count of the best positions obtained from the previous generation.

4) Scout bee phase

Algorithm 1. IABC based feature selection

Input: Data from volBrain
Output: Selected features
1. Create a high-quality initial solution (i.e., a subset of features) using $t_{h,g}$
2. Obtain a new food source v_h by using the best position t_{best}
3. Calculate $fitness(t_h)$ and $fitness(v_h)$
4. **if** $fitness(v_h) < fitness(t_h)$
5. **do**
6. Replace t_h with v_h
7. Replace worst population with new ones using t_{x_l}
8. **end if**
9. **else**
10. Keep t_h as best solution
11. Use PSO strategy to generate a new position for an un-updated position t_h
12. Return the best solution (i.e., a subset of features)

In the scout bee phase, a PSO strategy is used to generate a new position for an un-updated position t_h which is given as follows:

$$v_h = w_h v_h + b_1 s_1(p_q - x_h) + b_2 s_2(p_q - x_h) \quad (6)$$

$$t_h = t_h + v_h \quad (7)$$

In the Eqs. (6) and (7), q is the first index such that $fitness(t_q) < fitness(t_h)$.

From the above process of IABC (Algorithm 1), the best subset of features is selected and those features are given as input to different classifiers such as KNN, RF, and SVM to predict AD.

IV. Results and Discussion

In this part, the experimental results of existing BABC and proposed IABC with different classifiers for the AD prediction are illustrated. For the experimental purpose, MRI images from the ADNI database [12] are used. This database consists of two classes Alzheimer Disease (AD) and Healthy Controls (HC). For the experimental purpose, a total of 175 patients with AD and 144 healthy persons are considered. The information consists of 167 women and 152 men in the age range of 55–91 years. BABC and IABC with different classifiers are executed in MATLAB 2018a and run on a Microsoft Windows 7 with an Intel processor running at 2.70 GHz and 4GB memory. The performance of BABC and IABC with different classifiers is tested in terms of accuracy, precision, recall, and F-measure.

A. Accuracy

The percentage of occurrences that are successfully categorized using the chosen features is known as accuracy. It is determined by dividing the total count of AD-affected individuals who were accurately predicted (true positive) by the total count of healthy individuals who were accurately predicted (true negative). It is determined as,

$$Accuracy = \frac{True\ Positive\ (TP) + False\ Negative\ (FN)}{TP + True\ Negative\ (TN) + False\ Positive(FP) + FN} \quad (8)$$

where, if the class label is positive and the AD prediction outcome is positive then it is TP. If the class label is negative and the AD prediction outcome is negative then it is TN. If the class label is negative and the AD prediction outcome is positive then it is FP. If the class label is positive and the AD prediction outcome is negative then it is FN.

Table 1 shows the comparison of BABC and IABC with different classifiers for the AD prediction in terms of accuracy. Fig. 2 shows the testing performance between BABC and IBAC with different classifiers for the AD prediction in terms of accuracy. At the 20th iteration, the accuracy of IABC-RF is 10.52%, 10.12%, 5.66%, and 2.7% greater than BABC-KNN, BABC-SVM, IABC-KNN, and IABC-SVM respectively. From this result, it is proved that the proposed IABC-RF has higher accuracy than other methods for the AD prediction.

Table 1. Comparison of accuracy

Iteration	BABC-KNN	BABC-SVM	BABC-RF	IABC-KNN	IABC-SVM	IABC-RF
20	82.7	83	86.5	87.3	89	91.4
40	83	84	86.52	88.6	89.4	91.43
60	83.2	84	86.54	88.9	89.6	91.44
80	83.5	84	86.57	89.2	89.8	91.45
100	83.6	84	86.59	89.4	89.9	91.46

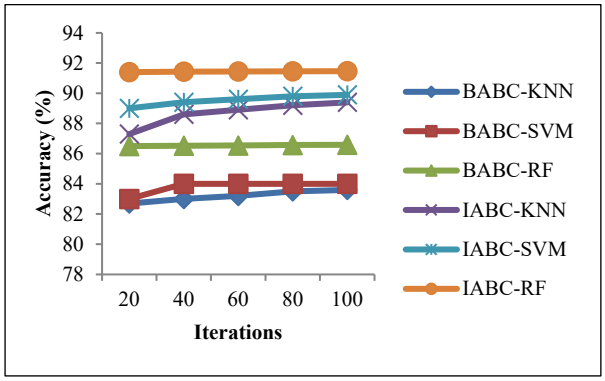

Fig. 2. Comparison of BABC and IBAC with different classifiers in terms of accuracy.

B. Precision

Precision is the measure to find the capacity of an AD classification model to recognize only the relevant instances in the dataset. It is calculated as,

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

Table 2 shows the comparison of BABC and IABC with different classifiers for the AD prediction in terms of precision. The precision for BABC and IABC with different classifiers under different numbers of iterations is shown in Fig. 3.

Table 2. Comparison of precision

Iteration	BABC-KNN	BABC-SVM	BABC-RF	IABC-KNN	IABC-SVM	IABC-RF
20	85.4	86.1	89.2	90.1	92	94.2
40	86	87	89.56	91.3	92.5	94.23
60	86.5	87.02	89.59	91.5	92.7	94.26
80	86.9	87.1	89.63	91.6	92.8	94.34
100	87	87.2	89.71	91.8	92.9	94.46

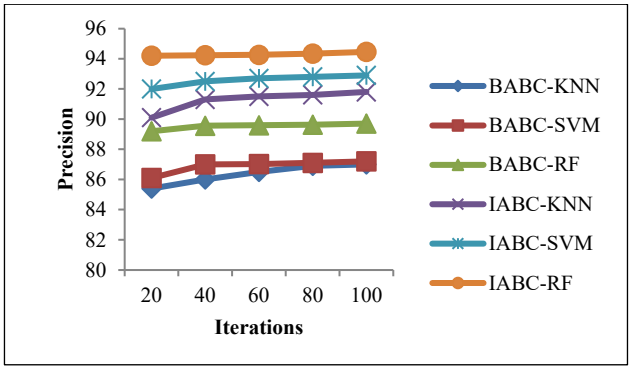

Fig. 3. Comparison of BABC and IBAC with different classifiers in terms of precision.

This analysis indicates that the proposed IABC-RF method achieves higher precision than other methods for the AD prediction. For instance, the precision of IABC-RF is 8.57%, 8.32%, 5.29%, 2.9%, and 1.68% greater than BABC-KNN, BABC-SVM, IABC-KNN, and IABC-SVM. Respectively, at 100th iteration.

C. Recall

Recall can measure the AD classification methods' capacity to identify all the data instances of interest in a dataset. It is calculated as,

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

Table 3 shows the comparison of BABC and IABC with different classifiers for the AD prediction in terms of recall. The recall for BABC and IABC with different classifiers under different numbers of iterations is shown in Fig. 4. This analysis indicates that the proposed IABC-RF method achieves higher recall than other methods for the AD prediction. For instance, the precision of IABC-RF is 7.87%, 7.49%, 5.17%, 3.11%, and 2.1% greater than BABC-KNN, BABC-SVM, IABC-KNN, and IABC-SVM respectively at 100th iteration.

Table 3. Comparison of recall

Iteration	BABC-KNN	BABC-SVM	BABC-RF	IABC-KNN	IABC-SVM	IABC-RF
20	84.3	85.6	88.4	89.1	91	92.9
40	85.1	86	88.01	90.3	91.1	93
60	85.5	86.25	88.13	90.4	91.3	93.1
80	85.6	86.34	88.41	90.41	91.4	93.2
100	86.7	87	88.92	90.7	91.6	93.52

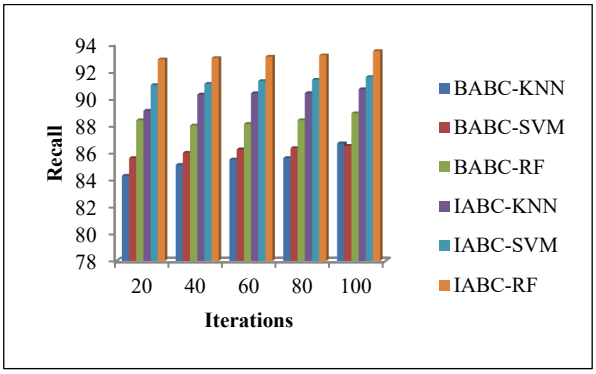

Fig. 4. Comparison of BABC and IBAC with different classifiers in terms of recall.

D. F-Measure

F-measure value is the harmonic mean of precision and recall taking both metrics into account. It is calculated as,

$$F - measure = 2 \cdot \left(\frac{Precision \cdot Recall}{Precision+Recall}\right) \quad (11)$$

Table 4 shows the comparison of BABC and IABC with different classifiers for the AD prediction in terms of F-measure. F-measure for BABC and IABC with different classifiers under different numbers of iterations is shown in Fig. 5. This analysis indicates that the proposed IABC-RF method achieves a higher F-measure than other methods for

the AD prediction. For instance, the F-measure of IABC-RF is 6.8%, 5.61%, 4.43%, 3%, and 2.12% greater than BABC-KNN, BABC-SVM, IABC-KNN and IABC-SVM, respectively, at 100th iteration.

Table 4. Comparison of f-measure

Iteration	BABC-KNN	BABC-SVM	BABC-RF	IABC-KNN	IABC-SVM	IABC-RF
20	84.9	86	89.2	90.3	91.6	92.95
40	86.2	87.3	89	90.64	91.8	93.4
60	87.1	88	89.16	90.95	91.94	93.68
80	87.9	88.6	89.63	91.2	92	93.99
100	88.2	89.2	90.2	91.45	92.24	94.2

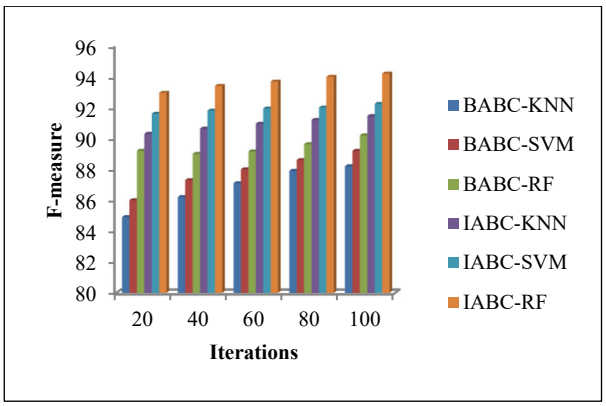

Fig. 5. Comparison of BABC and IBAC with different classifiers in terms of F-measure.

V. Conclusion

This article presents a comprehensive approach utilizing the Improved Artificial Bee Colony (IABC) algorithm for feature selection, enhancing AD identification accuracy. Leveraging brain MRI data collected from VolBrain, the proposed method efficiently selects pertinent features for AD recognition, integrating with classifiers like KNN, RF, and SVM. The IABC algorithm, inspired by honeybee swarm behavior, optimizes feature selection by dynamically adapting to the search space. Experimental results, conducted on ADNI database MRI images, demonstrate superior performance of IABC compared to the existing BABC method across various classifiers in terms of accuracy, precision, recall, and F-measure. Particularly, IABC-RF exhibits notable improvement, outperforming other methods, showcasing its potential in enhancing the AD prediction accuracy.

For future work, further exploration could focus on refining the IABC algorithm parameters to optimize performance and scalability. Additionally, investigating the method's applicability to diverse datasets and its generalization across different neurodegenerative diseases could provide valuable insights into its broader utility in clinical settings. Moreover, integrating advanced imaging techniques and exploring multi-modal data fusion approaches could enhance the model's robustness and predictive capabilities, potentially leading to more effective early the AD diagnosis and personalized treatment strategies.

Data Availability Statement

Publicly available datasets were analyzed in this study. The data is available on http://adni.loni.usc.edu/about/

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

The research was done by Sarumathi S, who also put the algorithm into practice. She assisted with writing the research article in addition to evaluating the findings. Reshma N assisted with data gathering, annotation, implementation, and aggregation of algorithm-derived outcomes. She also helped with the research paper's writing. Sharmila Mathivanan made contributions to the research paper's composition as well as to its implementation and consolidation of the findings. Malarkhodi S assisted in writing the study paper and the work's guidelines. The final draft has been approved by all authors.

References

[1] D. D. Martinelli, "Evolution of Alzheimer's disease research from a health-tech perspective: Insights from text mining," *Int. J. of Inf. Manag. Data Insights*, vol. 2, 100089, 2022.

[2] V. Sathiyamoorthi, A. K. Ilavarasi, K. Murugeswari, S. T. Ahmed, B. A. Devi, and M. Kalipindi, "A deep convolutional neural network based computer aided diagnosis system for the prediction of Alzheimer's disease in MRI images," *Measurement*, vol. 171, 108838, 2021.

[3] L. Brand, K. Nichols, H. Wang, L. Shen, and H. Huang, "Joint multi-modal longitudinal regression and classification for Alzheimer's disease prediction," *IEEE Trans. on Med. Imaging*, vol. 39, pp. 1845–1855, 2019.

[4] M. S. Roobini and M. Lakshmi, "Prediction of Alzheimer Disease using pearson recursive graph convolutional neural network," *Res. Sq.,* 2021.

[5] A. J. Dinu and R. Ganesan, "Early detection of Alzheimer's disease using predictive k-NN instance based approach and t-test method," *Int. J. of Adv. Trends in Comput. Sci. and Eng.*, vol.8, pp. 29–37, 2019.

[6] J. Albright *et al.*, "Forecasting the progression of Alzheimer's disease using neural networks and a novel preprocessing algorithm," *Alzheimer's & Dement. Trans. Res.& Clin. Interv.*, vol. 5, pp. 483–491, 2019.

[7] Y. Shi, C. Zu, M. Hong, L. Zhou, L. Wang, X. Wu, and Y. Wang, "ASMFS: Adaptive-similarity-based multi-modality feature selection for classification of Alzheimer's disease," *Pattern Recognit.*, vol. 126, 108566, 2022.

[8] M. Leela, K. Helenprabha, and L. Sharmila, "Prediction and classification of Alzheimer disease categories using integrated deep transfer learning approach," *Measurement: Sens.*, vol.27, 100749, 2023.

[9] A. Sharma, S. Kaur, N. Memon, A. J. Fathima, S. Ray, and M. W. Bhatt, "Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis," *Neuroscience Inf.*, vol. 1, 100012, 2021.

[10] N. Wang, J. Chen, H. Xiao, L. Wu, H. Jiang, and Y. Zhou, "Application of artificial neural network model in diagnosis of Alzheimer's disease," *BMC Neurol.*, vol. 19, pp. 1–8, 2019.

[11] P. Zhang, S. Lin, J. Qiao, and Y. Tu, "Diagnosis of Alzheimer's disease with ensemble learning classifier and 3D convolutional neural network," *Sens.*, vol.21, 7634, 2021.

[12] M. K. Keleş and Ü. Kiliç, "Classification of brain volumetric data to determine Alzheimer's disease using artificial bee colony algorithm as feature selector," *IEEE Access*, vol. 10, pp. 82989–83001, 2022.

[13] B. Szalkai, V. K. Grolmusz, V. I. Grolmusz, and C. A. M. Diseases, "Identifying combinatorial biomarkers by association rule mining in the CAMD Alzheimer's database," *Arch. of Gerontol. and geriat.*, vol. 73, pp. 300–307, 2017.

[14] L. R. Trambaiolli, N. Spolaôr, A. C. Lorena, R. Anghinah, and J. R. Sato, "Feature selection before EEG classification supports the

diagnosis of Alzheimer's disease," *Clinic. Neurophysiol.*, vol. 128, pp. 2058–2067, 2017.
[15] F. Gu, S. Ma, X. Wang, J. Zhao, Y. Yu, and X. Song, "Evaluation of feature selection for Alzheimer's disease diagnosis," *Front. Aging Neurosci*., vol. 14, 2022.
[16] A. Shukla, R. Tiwari, and S. Tiwari, "Review on Alzheimer disease detection methods: Automatic pipelines and machine learning techniques," *Sci.*, vol. 5, 2023.
[17] W. Shao, Y. Peng, C. Zu, M. Wang, and D. Zhang, "Alzheimer's disease neuroimaging initiative hypergraph based multi-task feature selection for multimodal classification of Alzheimer's disease," *Comput. Med. Imaging Graph*, vol. 80, 101663, 2020.
[18] J. Li, L. Wu, G. Wen, and Z. Li, "Exclusive feature selection and multi-view learning for Alzheimer's disease," *J. of Vis. Commun. and Image Representation*, vol.64, 102605, 2019.
[19] A. Basher, B. C. Kim, K. H. Lee, and H. Y. Jung, "Volumetric feature-based Alzheimer's disease diagnosis from sMRI data using a convolutional neural network and a deep neural network," *IEEE Access*, vol. 9, pp. 29870–29882, 2021.
[20] Y. Zhang, S. Wang, K. Xia, Y. Jiang, and P. Qian, "Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion," *Inf. Fusion*, vol. 66, pp. 170–183, 2021.
[21] M. Sudharsan and G. Thailambal, "Alzheimer's disease prediction using machine learning techniques and Principal Component Analysis (PCA)," *Mater. Today: Proc.,* 2021.
[22] A. Khan and S. Zubair, "Development of a three tiered cognitive hybrid machine learning algorithm for effective diagnosis of Alzheimer's disease," *J. of King Saud University-Comput. and Inf. Sci.*, vol. 34, pp. 8000–8018, 2022.
[23] M. N. Kp and P. Thiyagarajan, "Feature selection using efficient fusion of fisher score and greedy searching for Alzheimer's classification," *J. of King Saud University-Comput. and Inf. Sci.*, vol. 34, pp. 4993–5006, 2022.
[24] R. Gunawardena, P. G. Sarrigiannis, D. J. Blackburn, and F. He, "Kernel-based nonlinear manifold learning for EEG-based functional connectivity analysis and channel selection with application to Alzheimer's disease," *Neurosci*., vol. 523, pp. 140–156, 2023.
[25] A. Asgharzadeh-Bonab, H. Kalbkhani, and S. Azarfardian, "An Alzheimer's disease classification method using fusion of features from brain magnetic resonance image transforms and deep convolutional networks," *Healthc. Anal.*, vol. 4, 100223, 2023.