International Journal of Electrical and Computer Engineering (IJECE)
Vol. 13, No. 5, October 2023, pp. 5126~5134
ISSN: 2088-8708, DOI: 10.11591/ijece.v13i5.pp5126-5134

An effective feature selection using improved marine predators algorithm for Alzheimer's disease classification

Preeti Sadanand Topannavar[1], Dinkar M. Yadav[2]
[1]Department of Electronics and Telecommunication, G. H. Raisoni College of Engineering and Management, Pune, India
[2]Department of Electronics and Telecommunication, S. N. D. College of Engineering and Research Center, Yeola, India

Article Info

Article history:

Received Dec 19, 2022
Revised Mar 27, 2023
Accepted Apr 7, 2023

Keywords:

Alzheimer's disease
Feature selection
Improved marine predators algorithm
Multiclass support vector Machine
ResNet-18

ABSTRACT

Alzheimer's disease (AD) is an irremediable neurodegenerative illness developed by the fast deterioration of brain cells. AD is mostly common in elder people and it extremely disturbs the physical and mental health of patients, therefore early detection is essential to prevent AD development. However, the precise detection of AD and mild cognitive impairment (MCI) is difficult during classification. In this paper, the Residual network i.e., ResNet-18 is used for extracting the features, and the proposed improved marine predators algorithm (IMPA) is developed for choosing the optimum features to perform an effective classification of AD. The multi-verse optimizer (MVO) used in the IMPA helps to balance exploration and exploitation, which leads to the selection of optimal relevant features. Further, the classification of AD is accomplished using the multiclass support vector machine (MSVM). Open access series of imaging studies-1 (OASIS-1) and Alzheimer disease neuroimaging initiative (ADNI) datasets are used to evaluate the IMPA-MSVM method. The performance of the IMPA-MSVM method is analyzed using accuracy, sensitivity, specificity, positive predictive value (PPV) and matthews correlation coefficient (MCC). The existing methods such as the deep learning-based segmenting method using SegNet (DLSS), mish activation function (MAF) with spatial transformer network (STN) and BrainNet2D are used to evaluate the IMPA-MSVM method. The accuracy of IMPA-MSVM for the ADNI dataset is 98.43% which is more when compared to the DLSS and MAF-STN.

Corresponding Author:

Preeti Sadanand Topannavar
Department of Electronics and Telecommunication, G. H. Raisoni College of Engineering and Management
Wagholi, Pune, Maharashtra, India
Email: topannavarp@gmail.com

1. INTRODUCTION

Alzheimer's disease (AD) is a common neurodegenerative disease characterized by hidden onset and continuous expansion. AD causes changes in memory, behavior, and cognitive functioning, as well as the patient's risk of dying after about 3 to 10 years [1]–[3]. AD slowly develops from mild to moderate, to serious stages of dementia. This happens because of the irregular development of proteins like tau tangles and amyloid plaques in the brain [4]. The World Alzheimer's Report of 2020 stated that there were 5.8 million patients over 65 years of age affected by Alzheimer's in the United States, with AD being the sixth main reason leading to deaths in 2018. Further, the number of dementia patients may probably reach up to 150 million in the year 2050 [5]. The people identified with mild cognitive impairment (MCI) noticeably have a higher risk of developing AD, as MCI is observed as a transitional phase between healthy cognitive aging and dementia. 60%

of patients having MCI develop dementia within ten years of diagnosis, others however endure cognitive stablility or recover their normal cognitive (NC) functions [6]–[8].

Still, the efficiency of drugs for AD treatment is restricted to data and no treatment is stated to reverse or avoid AD development [9], [10]. Conventionally, the precise identification of AD, MCI and NC mainly depend on neuropsychological tests. However, neuropsychological test is subjective and appropriate for a person having certain clinical symptoms. Further, magnetic resonance imaging (MRI), essential role in brain disease identification along with the growth of medical imaging modalities [11], positron emission tomography, and single-photon emission computed tomography perform an [12]. MRI is an effective approach for nonaggressive in vivo imaging of the human brain. It is preferred over other technologies because of its ability to detect the structural variations that occur in neurodegenerative diseases, and hence has more importance in AD analysis and prediction [13]. The dimension of extracted features is comparatively higher than the sample size in neuroimaging. Subsequently, the classification performance is affected due to irrelevant features. Hence, feature selection that eliminates irrelevant features is considered a primary step while diagnosing AD [14], [15].

The related works about the classification of AD are given as follows: Buvaneswari and Gayathri [16] developed deep learning-based segmenting namely SegNet to identify the features related to AD. Further, the classification of accurate AD was obtained using ResNet-101. The strong features of the labels which included the graded features were established by the properties of every segmented image. Therefore, these graded features were used to enhance the classification. The developed SegNet did not consider entire brain regions as the region of interest (ROI) while performing the feature extraction. Sun *et al.* [17] presented a model for early identification of AD using a deep learning model depending on ResNet-50. In ResNet, the mish activation function (MAF) was chosen instead of the Relu function, followed by spatial transformer network (STN) incorporated into the input layer and the improved ResNet-50. The incorporation of STN was used to improve the spatial invariance that was used to enhance the feature-extracting capacity. An appropriate feature selection was required to choose the optimum features from the feature vector. Saratxaga *et al.* [18] developed 2D and 3D networks in either a slice-level approach or a subject-level method for predicting AD. Custom networks i.e., BrainNet2D and BrainNet3D were developed together with a popular architecture and transfer learning methods such as fine-tuning along with ImageNet weights. A huge set of examples were used in the 2D network which was then used to decrease the overfitting issue. The accuracy of the BrainNet3D was less, even though it was processed with all slices.

Cui *et al.* [19] presented adaptive least absolute shrinkage and selection operator (LASSO) logistic regression according to the particle swarm optimization (PSO) for detecting AD. Initially, the PSO was utilized for a global search for removing the irrelevant features and decrease the operational time. Later, the adaptive LASSO was performed in the local search for selecting the optimal features to classify AD. Here, the classification accuracy was sensitive to the tuning parameter. Suresha and Parthasarathy [20] developed the detection of AD using grey wolf optimization based clustering algorithm (GWOCA) and deep neural network (DNN). The adaptive histogram equalization and GWOCA were developed to denoise and segment the brain tissues. Next, the local ternary pattern, dual-tree complex wavelet transform and Tamura feature extraction were used to obtain the features. Subsequently, the ReliefF was used to select the optimal features for a precise classification of AD's abnormality and normality. The ReliefF used in this AD detection was not operated based on a learning basis, hence there was no assurance of acquiring an efficient feature subset.

The contributions are summarized as follows: i) for improving the intensity of pixels, normalization is performed over the input images. The ResNet-18 is used for feature extraction because each layer learns from the residual functions by taking reference to its input layer and ii) improved marine predators algorithm (IMPA) is proposed for feature selection because of its effective balance among the exploration and exploitation processes developed using the multi-verse optimizer (MVO). Further, the classification of AD is done by using multiclass support vector machine (MSVM), because it can handle high-dimension spaces and control nonlinear problems.

The remaining paper is arranged as follows: section 2 delivers a detailed explanation of the IMPA-MSVM. The outcomes of IMPA-MSVM are provided in section 3. Further, the conclusion is presented in section 4.

2. IMPA-MSVM METHOD

In this research, the classification of brain images is enhanced to take early action for patients who are suffering from AD and MCI. The classification accuracy of different stages of AD is enhanced using IMPA based feature selection. The important process of IMPA-MSVM method is data acquisition, preprocessing, feature extraction, feature selection, and classification. The block diagram of the overall IMPA-MSVM method is shown in Figure 1.

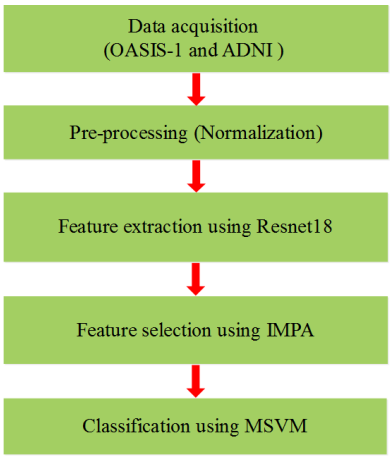

Figure 1. Block diagram of overall IMPA-MSVM method

2.1. Data acquisition

The classification of Alzheimer's disease using the proposed method is analyzed with two different datasets namely Open access series of imaging studies-1 (OASIS-1) [21] and Alzheimer disease neuroimaging initiative (ADNI) [22]. OASIS-1 is a neuroimaging dataset that has a cross-sectional collection of 416 subjects. Among 416 subjects, 100 subjects are over 60 years of age and identified with very light or moderate Alzheimer's using clinical dementia rating (CDR) value. OASIS-1 has a T1-weighted MRI scan for each subject along with 20 subjects without dementia. These 20 subjects' information is scanned twice on continuous visits, and used for a baseline evaluation with a list of explanations. On the other hand, the ADNI was launched in 2003 which is a multisite, longitudinal study intended to create genetic, clinical, imaging and biospecimen biomarkers to achieve an early prediction of Alzheimer's. ADNI consists of T1 weighted structural MRI data for three different classes such as: 1,425 scans for AD, 1,021 scans for cognitive normal (CN) and 1,479 scans for mild cognitive impairment (MCI).

2.2. Preprocessing using normalization

The normalization is applied over the input images from the OASIS-1 and ADNI datasets. The input image's pixel intensity is enhanced by changing the pixel range through normalization as expressed in (1).

$$D' = (D - min)\frac{newmax - newmin}{max - min} + newmin \quad (1)$$

where, D denotes the input image; the minimum and maximum intensity values of input are denoted as min and max; the preprocessed image is denoted as D' and it has intensity values of $newmin$ and $newmax$.

2.3. Feature extraction using ResNet-18

In feature extraction, the pre-trained network of ResNet-18 [23] is used to extract the features. The architecture of ResNet18 is shown in Figure 2. The preprocessed image of D' is given as input to the ResNet-18 architecture. The developed ResNet-18 has 16 convolution layers, 2 downsampling layers and few fully connected layers. The eigenvector i.e., the feature map of the last convolution layer is obtained, once the average pooling is done in ResNet-18. The obtained eigenvector has numerous probabilities. From the convolutional layers, the extracted features are obtained and it is further processed under the feature selection process.

2.4. Feature selection using IMPA

In the proposed method, the IMPA is utilized to extract the optimum features out of the the overall features extracted from ResNet-18. This IMPA based feature selection removes the irrelevant information from the extracted features of ResNet-18 to enhance the accuracy. Generally, the MPA [24] imitates the foraging strategy namely Levy and Brownian movements in ocean predators, and also the optimal searching in the biological communication between Predator and Prey. The IMPA-based feature selection has four main phases according to the different velocity ratios and it replicates the normal behavior of predators and prey.

Phase 1: The first phase is initialized when there is a high-velocity ratio (velocity $v \geq 10$) in the event of the prey traveling faster, as the appropriate policy of the predator is to be stationary. The exploration is essential during the initial iteration of optimization. The mathematical expression for this 1st phase is expressed in (2).

$$\begin{aligned} &While\ Iter < \tfrac{1}{3}\max\ \_Iter \\ &\overrightarrow{stepsize_i} = \vec{R}_B \otimes (Elite_i - \vec{R}_B \otimes \overrightarrow{Prey_i}) \quad i = 1,2,\dots,n \\ &\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P.\vec{R} \otimes \overrightarrow{stepsize_i} \end{aligned} \tag{2}$$

where, $\overrightarrow{Prey_i}$ is considered as the feature vectors extracted using ResNet-18.

According to the Brownian motion's normal distribution, the $\vec{R}_B$ contains the random values. The motion of prey is replicated by multiplying the R_B with prey; P is a constant number that is equal to 0.5 and the uniform random number in the range of $[0, 1]$ is denoted as R. This phase is accomplished in 1/3 of iterations when the motion velocity is higher for allowing higher exploration levels.

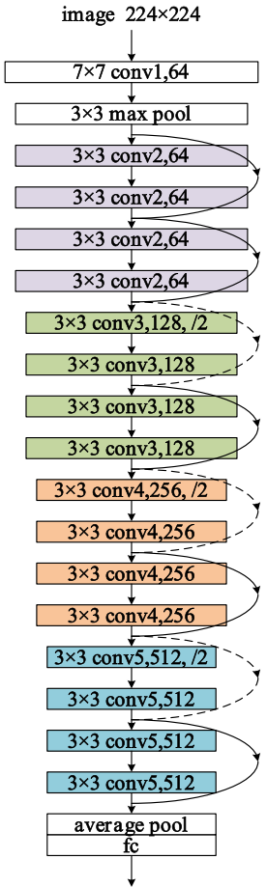

Figure 2. Architecture of ResNet18 for feature extraction

Phase 2: The second phase is accomplished when both the predator and prey are concurrently traveling in the search space. This kind of motion is accomplished when the exploration tends to get transformed into exploitation. This phase comprises both the exploration and exploitation features, while the predator is accountable for exploration and the prey is accountable for exploitation.

An appropriate approach for predator is Brownian and for prey is levy, when there is a unit velocity ratio ($v \approx 1$) which is expressed in (3).

$$\begin{aligned} &While\ \tfrac{1}{3}max\ \_Iter < Iter < \tfrac{2}{3}max\ \_Iter \\ &\overrightarrow{stepsize_i} = \vec{R}_L \otimes (\overrightarrow{Elite_i} - \vec{R}_L \otimes \overrightarrow{Prey_i}) \quad i = 1,2,\dots,n/2 \\ &\overrightarrow{Prey_i} = \overrightarrow{Prey_i} + P.\vec{R} \otimes \overrightarrow{stepsize_i} \end{aligned} \tag{3}$$

where, the random number vectors according to the levy are denoted as $\vec{R}_L$ that used for 1st half of the population. The multiplication of $\vec{R}_L$ is used to replicate the prey's motion in levy manner, whereas the incorporation of step size to the prey's location replicates the prey's movement. A huge amount of step sizes in levy are small. The mathematical expression for 2nd half of the population is shown in (4).

$$\overrightarrow{stepsize_i} = \vec{R}_B \otimes (\vec{R}_B \otimes \overrightarrow{Elite_i} - \overrightarrow{Prey_i}) \quad i = \frac{n}{2}, \dots, n$$
$$\overrightarrow{Prey_i} = \overrightarrow{Elite_i} + P.CF \otimes \overrightarrow{stepsize_i} \tag{4}$$

where, $CF = \left(1 - \frac{Iter}{\max\_Iter}\right)^{2\frac{Iter}{\max\_Iter}}$ is the adaptive parameter used to handle the step size for predator motion. The multiplication of $\vec{R}_B$ is used to stimulate the predator motion in a Brownian manner. According to the predator motion, the prey motion is simulated using Elite for updating its location.

Phase 3: If the predator's movement is faster than the prey's low-velocity ratio, the third phase is initialized and connected with higher exploitation ability. If the low-velocity ratio $v = 0.1$, then the appropriate approach for Predator is a levy. The updation of step size and prey for third phase is expressed in (5).

$$While\ Iter < max\_Iter$$
$$\overrightarrow{stepsize_i} = \vec{R}_L \otimes (\vec{R}_L \otimes \overrightarrow{Elite_i} - \overrightarrow{Prey_i}) \quad i = 1,2, \dots, n/2$$
$$\overrightarrow{Prey_i} = \overrightarrow{Elite_i} + P.CF \otimes \overrightarrow{stepsize_i} \tag{5}$$

Phase 4: The operators of MVO are included in IMPA for obtaining the better solution (i.e., optimal features). The prey is updated based on the first half of phase 2, when $r3 \geq WAP$; Otherwise, the second half is executed, when $r3 < WAP$ and the WAP is defined in (6).

$$WAP = min + Iter \times \left(\frac{min - max}{Iter}\right) \tag{6}$$

The location update of prey is expressed in (7).

$$\overrightarrow{Prey_i} = \begin{cases} \overrightarrow{Prey_i} + TDR \times \left((ub_i - lb_i) \times r_5 + lb_i\right) & r4 < 0.5 \\ \overrightarrow{Prey_i} - TDR \times \left((ub_i - lb_i) \times r_5 + lb_i\right) & r4 \geq 0.5 \\ \overrightarrow{Prey_i} & \end{cases} \tag{7}$$

where, $r3, r4$ and $r5$ denote the random numbers and TDR is expressed in (8).

$$TDR = 1 - \frac{Iter^{1/P1}}{max\_Iter^{1/P1}} \tag{8}$$

where, $P1$ is the constant value that is set as 6 for controlling the exploitation accuracy. Further, the remaining point which causes the variation in marine predator's behavior is eddy formation or fish aggregating devices (FADs) effects that are expressed in (9).

$$\overrightarrow{Prey_i} = \begin{cases} \overrightarrow{Prey_i} + CF\left[X_{min} + \vec{R} \otimes (\bar{X}_{max} - \bar{X}_{min})\right] \otimes \vec{U} & if\ r \leq FADs \\ \overrightarrow{Prey_i} + [FADs(1-r) + r]\left(\overrightarrow{Prey_{r1}} - \overrightarrow{Prey_{r2}}\right) & if\ r > FADs \end{cases} \tag{9}$$

where, r is the uniform random number generated in the range of [0,1]; $\bar{X}_{min}$ and $\bar{X}_{max}$ are the lower and upper boundary of the dimensions respectively; $FADs$ is equal to 0.2 which creates the impact in the optimization process; the binary vector which has 0 and 1 is denoted as U and it is defined by creating a random vector in [0, 1]. If the array is less than 0.2, the array is changed to 0; otherwise, the array is changed to 1 when the array is greater than 0.2.

The IMPA selects the optimal features according to two distinct fitness functions namely the classification accuracy and the number of features, as expressed in (10).

$$Fitness\left(\overrightarrow{Prey_j}\right) = \vartheta Acc\left(\overrightarrow{Prey_i}\right) + \delta\left(1/\left|\overrightarrow{Prey_i}\right|\right) \tag{10}$$

where, the $\overrightarrow{Prey_i}$ defines the chosen random feature subset; $Acc\left(\overrightarrow{Prey_i}\right)$ is the classification accuracy of $\overrightarrow{Prey_i}$; $\left|\overrightarrow{Prey_i}\right|$ defines the number of features; ϑ and δ is the random number generated in the range of $[0, 1]$ that is

used to define the relation among the accuracy and number of the selected subset. The fitness of each $\overrightarrow{Prey_j}$ i.e., feature subset is computed for each iteration and the $\overrightarrow{Prey_j}$ with best fitness is saved as $\overrightarrow{Prey_{best}}$. Further, this $\overrightarrow{Prey_{best}}$ is used in the classification for detecting AD.

2.5. MSVM based classification

MSVM is used to classify the AD, NC, and MCI according to the features ($\overrightarrow{Prey_{best}}$) selected from IMPA. Generally, the conventional support vector machine (SVM) is generated for binary classification, hence the SVM is transformed into MSVM [25] to classify different types of Alzheimer's disease. The one-against-one approach is used in MSVM for multi class classification. Further, the radial basis function is used as the kernel in MSVM for nonlinear issues.

3. RESULTS AND DISCUSSION

The implementation and simulation of the IMPA-MSVM method is performed using MATLAB R2020a software where the system is configured with an i5 processor with 8GB of RAM. The datasets used to analyze the IMPA-MSVM method is OASIS-1 and ADNI, where training takes 80% and testing takes 20% of data from each dataset. Here, the data is randomly taken for training and testing based on the iterations. The IMPA-MSVM method is examined using accuracy, sensitivity, specificity, PPV and MCC which are expressed in (11) to (15).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100 \quad (11)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100 \quad (12)$$

$$Specificity = \frac{TN}{TN+FP} \times 100 \quad (13)$$

$$PPV = \frac{TP}{TP+FP} \times 100 \quad (14)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TN+FN)(TP+FP)(TN+FP)(TP+FN)}} \times 100 \quad (15)$$

where, TP, TN, FP and FN denote true positive, true negative, false positive and false negative.

The OASIS-1 dataset has different classes such as with different CDR values. The classes of OASIS-1 are cognitive normal with 336 samples and $CDR = 0$, very-mild dementia with 70 samples and $CDR = 0.5$, mild dementia with 28 samples and $CDR = 1$ and moderate dementia with 2 samples and$CDR = 2$. Subsequently, this OASIS-1 is labeled as three class problem i.e., $CDR = 0$ is first class, $CDR = 0.5$ is second class and $CDR = 1$ and $CDR = 2$ is third class. Because, the class of $CDR = 2$ has only 2 samples and $CDR = 1$ has 28 samples, it creates inaccuracy in classification because of under-representation. On the other hand, randomly 1,000 scans are taken from AD, NC and MCI to avoid the data imbalance issue during classification.

3.1. Performance analysis of the IMPA-MSVM method

In this section, the IMPA-MSVM method is analyzed with various classifiers and with various feature selection approaches. The different classifiers used to evaluate the IMPA-MSVM method are random forest classifier (RFC), K-nearest neighbour (KNN) and decision tree (DE). The performance analyses of IMPA-MSVM & different classifiers with feature selection (WFS) and without feature selection (WOFS) are shown in Table 1. The graphically illustrated results of IMPA-MSVM with ADNI dataset for WOFS and WFS are shown in Figures 3 and 4 respectively. From the examination, it is concluded that the MSVM achieves better classification performances in WOFS as well as WFS. For example, the accuracy of IMPA-MSVM with WFS for the ADNI dataset is 98.43%, which is higher when compared to the RFC, KNN, and DE classifiers. Besides, the accuracy of IMPA-MSVM with WFS for OASIS-1 dataset is 94.65%, which is higher in relation with RFC, KNN and DE. The MSVM with IMPA provides better classification among the multiple classifiers taken for evaluation, because it can handle high-dimension spaces and control the non linear problems. Further, the optimal selection of features eliminates the irrelevant features that help to improve the accuracy.

Additionally, the different feature selection approaches such as bat optimization algorithm (BOA), grey wolf optimization (GWO) and conventional MPA are used to evaluate the performance of IMPA-MSVM. Table 2 shows the performance analyses of IMPA-MSVM and different feature selection approaches where the performance is evaluated using ADNI and OASIS-1 datasets. An example of a graphical illustration of

IMPA-MSVM for the ADNI dataset is shown in Figure 5. This analysis shows that the IMPA provides better classification accuracy for both the OASIS-1 and ADNI datasets. For example, the accuracy of IMPA-MSVM method for ADNI dataset is 98.43% which is higher compared to the BOA, GWO and MPA. The reason for IMPA with better performance is the incorporation of MVO parameters, which creates an effective balance between the exploration and exploitation processes that help to search for an optimal feature subset.

Table 1. Performance analysis of IMPA-MSVM for different classifiers

Data set	Feature selection	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	MCC (%)
ADNI	WOFS	RFC	87.4	84.69	84.53	80.51	87.36
		KNN	92.71	91.99	90.32	93.97	90.52
		DE	90.25	89.26	88.51	90.45	87.28
		MSVM	95.7	94.51	92.36	94.93	93.17
	WFS	RFC	90.08	92.98	89.86	89.43	90.61
		KNN	96.67	95.96	95.07	92.82	91.87
		DE	93.33	95.92	94.41	92.65	92.35
		MSVM	98.43	98.86	98.02	97.04	96.07
OASIS-1	WOFS	RFC	82.7	84.65	82.31	79.87	79.32
		KNN	82.6	84.96	85.8	80.33	79.55
		DE	85.47	87.38	83.9	80.89	81.74
		MSVM	85.73	89.34	89.91	83.14	84.04
	WFS	RFC	88.89	86.78	86.86	87.45	82.33
		KNN	90.08	93.03	92.6	92.4	91.42
		DE	92.76	89.07	90.12	90.26	88.04
		MSVM	94.65	94.32	95.81	95.69	94.06

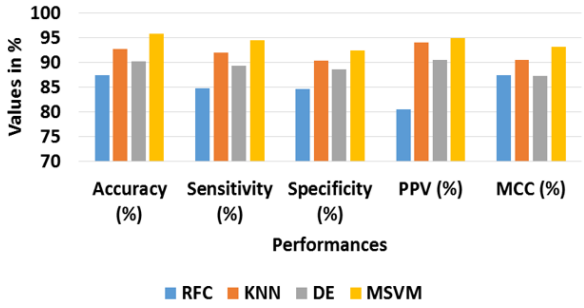

Figure 3. Graphical comparison of classifiers WOFS for ADNI dataset

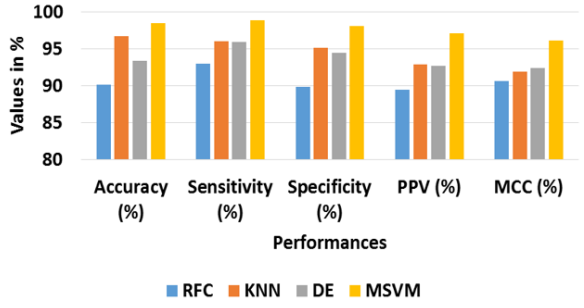

Figure 4. Graphical comparison of classifiers WFS for ADNI dataset

Table 2. Performance analysis of IMPA-MSVM for different feature selection approaches

Data set	Feature selection	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	MCC (%)
ADNI	BOA	89.62	90.87	89.08	89.7	91.03
	GWO	90.6	91.65	91.02	89.49	89.91
	MPA	95.63	92.56	94.06	94.26	92.02
	IMPA	98.43	98.86	98.02	97.04	96.07
OASIS-1	BOA	77.49	84.79	83.09	75.6	77.97
	GWO	75.11	82.58	82.41	73.18	73.39
	MPA	79.72	81.07	65.98	73.24	76.07
	IMPA	94.65	94.32	95.81	95.69	94.06

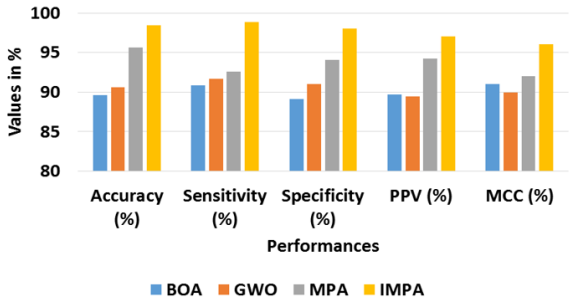

Figure 5. Graphical comparison of feature selection approaches for ADNI dataset

3.2. Comparative analysis

Existing researches such as deep learning-based segmenting method using SegNet (DLSS) [16], MAF-STN [17] and BrainNet2D [18] are used to evaluate the efficiency of the IMPA-MSVM method. Table 3 shows the comparative analysis of IMPA-MSVM with DLSS [16], MAF-STN [17] and BrainNet2D [18]. From Table 3, it is concluded that the IMPA-MSVM achieves improved performance than the existing researches. The accuracy of IMPA-MSVM for the OASIS-1 dataset is 94.65% which is higher than the BrainNet2D [18]. The optimal feature selection using IMPA helps to improve the multi-class classification of AD.

Table 3. Comparative analysis of IMPA-MSVM

Dataset	Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)
ADNI	DLSS [16]	96.3	96.7	93.9	NA
	MAF-STN [17]	97.1	95.3	NA	95.5
	IMPA-MSVM	98.43	98.86	98.02	97.04
OASIS-1	BrainNet2D [18]	84	NA	NA	NA
	IMPA-MSVM	94.65	94.32	95.81	95.69

4. CONCLUSION

The precise diagnosis of AD in its initial stage such as the condition of MCI, is necessary for timely treatment and to slow down the AD development. In this paper, the ResNet-18-based feature extraction and IMPA are used to eliminate the inappropriate features during the classification. The MVO operators used in the IMPA provide balance among exploration and exploitation processes. Further, the MSVM is used to classify AD according to the features selected from the IMPA. The performance evaluation shows that the IMPA-MSVM provides better classification results in both the OASIS-1 and ADNI datasets when compared to DLSS, MAF-STN, and BrainNet2D. The accuracy of IMPA-MSVM for the ADNI dataset is 98.43% which is relatively higher compared to DLSS and MAF-STN. In the future, deep learning classifiers can be used to improve the detection of AD classes.

REFERENCES

[1] Z. Fan, F. Xu, X. Qi, C. Li, and L. Yao, "Classification of Alzheimer's disease based on brain MRI and machine learning," *Neural Computing and Applications*, vol. 32, no. 7, pp. 1927–1936, Apr. 2020, doi: 10.1007/s00521-019-04495-0.

[2] Z. Xu, H. Deng, J. Liu, and Y. Yang, "Diagnosis of Alzheimer's disease based on the modified tresnet," *Electronics*, vol. 10, no. 16, Aug. 2021, doi: 10.3390/electronics10161908.

[3] Y. Zhang, S. Wang, K. Xia, Y. Jiang, and P. Qian, "Alzheimer's disease multiclass diagnosis via multimodal neuroimaging embedding feature selection and fusion," *Information Fusion*, vol. 66, pp. 170–183, Feb. 2021, doi: 10.1016/j.inffus.2020.09.002.

[4] R. Divya and R. Shantha Selva Kumari, "Genetic algorithm with logistic regression feature selection for Alzheimer's disease classification," *Neural Computing and Applications*, vol. 33, no. 14, pp. 8435–8444, Jul. 2021, doi: 10.1007/s00521-020-05596-x.

[5] C. Liang, H. Lao, T. Wei, and X. Zhang, "Alzheimer's disease classification from hippocampal atrophy based on PCANet-BLS," *Multimedia Tools and Applications*, vol. 81, no. 8, pp. 11187–11203, Mar. 2022, doi: 10.1007/s11042-022-12228-0.

[6] R. E. Turkson, H. Qu, C. B. Mawuli, and M. J. Eghan, "Classification of Alzheimer's disease using deep convolutional spiking neural network," *Neural Processing Letters*, vol. 53, no. 4, pp. 2649–2663, Aug. 2021, doi: 10.1007/s11063-021-10514-w.

[7] A. Giovannetti *et al.*, "Deep-MEG: spatiotemporal CNN features and multiband ensemble classification for predicting the early signs of Alzheimer's disease with magnetoencephalography," *Neural Computing and Applications*, vol. 33, no. 21, pp. 14651–14667, Nov. 2021, doi: 10.1007/s00521-021-06105-4.

[8] J. E. Santos Toural, A. Montoya Pedrón, and E. J. Marañón Reyes, "Classification among healthy, mild cognitive impairment and Alzheimer's disease subjects based on wavelet entropy and relative beta and theta power," *Pattern Analysis and Applications*, vol. 24, no. 2, pp. 413–422, May 2021, doi: 10.1007/s10044-020-00910-8.

[9] X. Hao *et al.*, "Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of Alzheimer's disease,"

Medical Image Analysis, vol. 60, Feb. 2020, doi: 10.1016/j.media.2019.101625.
[10] Z. Fan, F. Xu, C. Li, and L. Yao, "Application of KPCA and AdaBoost algorithm in classification of functional magnetic resonance imaging of Alzheimer's disease," *Neural Computing and Applications*, vol. 32, no. 10, pp. 5329–5338, May 2020, doi: 10.1007/s00521-020-04707-y.
[11] H. Lao, X. Zhang, Y. Tang, and C. Liang, "Alzheimer's disease diagnosis based on the visual attention model and equal-distance ring shape context features," *IET Image Processing*, vol. 15, no. 10, pp. 2351–2362, Aug. 2021, doi: 10.1049/ipr2.12218.
[12] X. Zhou *et al.*, "Enhancing magnetic resonance imaging-driven Alzheimer's disease classification performance using generative adversarial learning," *Alzheimer's Research & Therapy*, vol. 13, no. 1, Dec. 2021, doi: 10.1186/s13195-021-00797-5.
[13] S. Salunkhe *et al.*, "Classification of Alzheimer's disease patients using texture analysis and machine learning," *Applied System Innovation*, vol. 4, no. 3, Aug. 2021, doi: 10.3390/asi4030049.
[14] W. Shao, Y. Peng, C. Zu, M. Wang, and D. Zhang, "Hypergraph based multi-task feature selection for multimodal classification of Alzheimer's disease," *Computerized Medical Imaging and Graphics*, vol. 80, Mar. 2020, doi: 10.1016/j.compmedimag.2019.101663.
[15] Z. Jiao, S. Chen, H. Shi, and J. Xu, "Multi-modal feature selection with feature correlation and feature structure fusion for MCI and AD classification," *Brain Sciences*, vol. 12, no. 1, Jan. 2022, doi: 10.3390/brainsci12010080.
[16] P. R. Buvaneswari and R. Gayathri, "Deep learning-based segmentation in classification of Alzheimer's disease," *Arabian Journal for Science and Engineering*, vol. 46, no. 6, pp. 5373–5383, Jun. 2021, doi: 10.1007/s13369-020-05193-z.
[17] H. Sun, A. Wang, W. Wang, and C. Liu, "An improved deep residual network prediction model for the early diagnosis of Alzheimer's disease," *Sensors*, vol. 21, no. 12, Jun. 2021, doi: 10.3390/s21124182.
[18] C. L. Saratxaga *et al.*, "MRI deep learning-based solution for Alzheimer's disease prediction," *Journal of Personalized Medicine*, vol. 11, no. 9, Sep. 2021, doi: 10.3390/jpm11090902.
[19] X. Cui *et al.*, "Adaptive LASSO logistic regression based on particle swarm optimization for Alzheimer's disease early diagnosis," *Chemometrics and Intelligent Laboratory Systems*, vol. 215, Aug. 2021, doi: 10.1016/j.chemolab.2021.104316.
[20] H. S. Suresha and S. S. Parthasarathy, "Detection of Alzheimer's disease using grey wolf optimization based clustering algorithm and deep neural network from magnetic resonance images," *Distributed and Parallel Databases*, vol. 40, no. 4, pp. 627–655, Dec. 2022, doi: 10.1007/s10619-021-07345-y.
[21] D. S. Marcus, T. H. Wang, J. Parker, J. G. Csernansky, J. C. Morris, and R. L. Buckner, "Open access series of imaging studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults," *Journal of Cognitive Neuroscience*, vol. 19, no. 9, pp. 1498–1507, Sep. 2007, doi: 10.1162/jocn.2007.19.9.1498.
[22] "ADNI|Alzheimer's disease neuroimaging initiative." 2017, Accessed: Jan. 09, 2023. [Online]. Available: http://adni.loni.usc.edu/.
[23] X. Ou *et al.*, "Moving object detection method via ResNet-18 with encoder–decoder structure in complex scenes," *IEEE Access*, vol. 7, pp. 108152–108160, 2019, doi: 10.1109/ACCESS.2019.2931922.
[24] D. S. A. Elminaam, A. Nabil, S. A. Ibraheem, and E. H. Houssein, "An efficient marinepredators algorithm for feature selection," *IEEE Access*, vol. 9, pp. 60136–60153, 2021, doi: 10.1109/ACCESS.2021.3073261.
[25] M. Zhang, Y. Yuan, R. Wang, and W. Cheng, "Recognition of mixture control chart patterns based on fusion feature reduction and fireworks algorithm-optimized MSVM," *Pattern Analysis and Applications*, vol. 23, no. 1, pp. 15–26, Feb. 2020, doi: 10.1007/s10044-018-0748-6.

BIOGRAPHIES OF AUTHORS

Preeti Sadanand Topannavar is currently a research scholar in the electronics and telecommunication department at G. H. Raisoni College of Engineering and Management, Pune, India. She is having total teaching experience of 17 years and currently working as Assistant professor at JSPM's Imperial College of Engineering and Research, Pune India. She has received a Bachelor's degree in electronics and telecommunication engineering from Amaravati University and M.E. degree from Sinhgad College of Engineering Pune in 2002 and 2010 respectively. Her current research interests include biomedical image processing, human-computer interactions, machine learning. She has published various research papers in different international and national journals and conferences. She can be contacted at topannavarp@gmail.com.

Dinkar M. Yadav received Bachelor's degree, Master's Degree and Ph.D. in electronics and telecommunication engineering. He is having total teaching experience of over 28 years and research Experience of 8 years. He is currently working as professor and principal at S. N. D. college of engineering and research center, Yeola, Maharashtra, India. His research interests include biomedical image processing, communication engineering, digital signal processing, neural networks. He has published several research papers in various SCI/Scopus listed journals and more than 20 research papers in peer reviewed international conferences. He is recognised Ph.D. guide for various repute universities such as Savitribai Phule Pune University, Dr. Babasaheb Ambedkar Marathwada University Aurangabad JJTU Rajasthan, Modi University, Jaipur He has guided more than 15 students for Ph.D. He can be contacted at dineshyadav800@gmail.com.