

OPTIMAL FEATURES SELECTION FOR ALZHEIMER'S DISEASE PREDICTION

StalinBabu Gedela^{1,2,*}, Nagatirumalarao Sighakolli³ and Rajeswara Rao Ramisetty⁴

Received: April 28, 2021; Revised: July 20, 2021; Accepted: September 19, 2021

Abstract

A Major Challenge for healthcare in the 21st century is dementia. It is the most common form of Alzheimer's disease (AD). Early detection of AD is necessary for preventing the progression of the symptoms. The goal of this study is to initiate a new predictive model for AD employing MRI images. The developed model involves Feature Extraction, Optimal Feature selection, and Classification. Initially, the Gray-Level Co-Occurrence Matrix (GLCM) and Haralick features are extracted from MRI images. Especially, this work carries out optimal feature selection using a Grasshopper Optimization Algorithm (GOA). Then, the optimally chosen features are classified via Deep Convolution Neural Network (DCNN). Finally, the eminence of the adopted scheme is validated in terms of various measures.

Keywords: Alzheimer's disease, GLCM, DCNN, Feature selection, GOA Algorithm

Introduction

There is a rising attention for Computer-Aided Diagnosis (CAD) systems to assist physicians in diagnosing several disorders such as Alzheimer's disease (AD), Parkinson's disease, schizophrenia, epilepsy and brain lesion accurately (Samper-González *et al.*, 2018; Lahmiri and Shmuel, 2019; Moscoso *et al.*, 2019). "AD is a neurodegenerative disorder that affects a large population". It corrupts the behavior, judgment and memory of the patients. Earlier recognition of AD is necessary for premature therapy, which can lessen the development of disease. For improving the earlier recognition of AD, numerous CAD systems were

introduced to accurately categorize Normal Controls (NC) and AD subjects depending on structural Magnetic Resonance Imaging (MRI) analysis.

In fact, earlier recognition of AD (Batzu *et al.*, 2020) by MRI is of large support to patients and physicians as it is a comparatively non-invasive, low-cost process that can aid diagnosis without human misinterpretations. Thereby, different CAD models were established depending on analysis of structural MRI of the human's brain. (Park *et al.*, 2020; Tadayon *et al.*, 2020).

¹ Department of CSE, Aditya institute of technology and management, Tekkali. E-mail: stalin1227@gmail.com

² Department of CSE, JNTUK, Kakinada.

³ Department of CSE, Narasaraopeta Engineering College, Narasaraopeta. E-mail: nagatirumalarao@gmail.com.

⁴ Department of CSE, JNTUK University College of Engineering, Vizianagaram. E-mail: rrrao.cse@jntukucev.ac.in

* Corresponding author

Along with CAD systems that depends on the investigation of structural MRI to identify AD (McKenzie *et al.*, 2020), other works concerned on the analysis on images attained with Positron Emission Tomography (PET), Diffusion Tensor Imaging (DTI) and Single-Photon Emission Tomography (SPECT) images. In an effort to develop the accuracy of CAD systems, current studies have adopted multi-modal techniques, in which the features are derived from diverse imaging modalities.

The major contributions of this research work are:

- Exploits Grasshopper Optimization Algorithm for carrying out optimal feature selection.

The paper is structured as follows: Section II portrays the reviews on Alzheimer disease. Section III depicts about the framework of proposed model for AD prediction. Further, addresses the texture feature extraction: GLCM, Haralick features. Proposed GOA model for optimal selection of features. The resultants are briefed in Section IV and conclusion is elucidated in Section V.

Literature Review

In this section, we review high quality research articles from reputed journals related to our research problem prediction of Alzheimer's Daises (AD).

Faisal *et al.* (2021) applied Machine Learning techniques to diagnose AD. They predict the AD based on the combined feature selection method. They consider the Alzheimer's disease Neuroimaging Initiative (ADNI) dataset. They used a combination of dimension reduction methods to select the features. They select the best features based on the Recursive Feature Elimination method and Principal Component Analysis. They studied the comparative analysis they applied 5 classifiers Softmax Classifier, Support Vector Machine (SVM), K-Nearest Neighbour (KNN), NaiveBayes and Random Forest and they proved that SVM is best suited for the prediction of AD.

Raj (2020) proposes that the Internet of Medical Things (IoMT) plays important role in the classification and prediction of critical diseases. They applied optimal Deep Learning (DL) for Alzheimer's disease classification. To improve the performance of the Deep learning classifier and selection of optimal features they applied the Opposition-based Crow Search (OCS) algorithm. They expressed the results by using MATLAB and proved that an optimal feature always improves the performance of the model and they achieve an accuracy of 95.22%.

Silvia *et al.* (2019) have developed a deep learning approach for predicting Mild Cognitive Impairment (MCI) and AD. Here, the appliance of

Convolutional Neural Network (CNN) was exploited over 3D T1-weighted images. The CNN was utilized for testing, in which the c-MCI and s-MCI of AD was evaluated. In the end, higher classification accuracy was accomplished by means of ADNI datasets. Thus, the CNN was said to be an effective tool for the enhanced individual patient diagnosis.

Vaithinathan *et al.* (2019) have analyzed the model based on texture with regional atrophies for classification of AD. Initially, the centre image slice was divided into minor blocks and the blocks that were similar to Region of interest (ROI) were marked. These whole marked blocks that relied at the similar location were collected and the textures were evaluated for all blocks on whole slices. At last, the development of the adopted work was examined over other models.

Shaikh and Ali (2019) have developed a creative model for automated tissue segmentation and classification based on SVM and PCA. For exploring the atrophy of the segmented tissues, the texture features from Gray-Level Co-Occurrence Matrix (GLCM) and, Band Expansion Process (BEP) features from discrete wavelet transform (DWT) were merged together. Consequently, pre-processing and classification was done, and the SVM classifier was employed for obtaining the two related features. The experimentation has exposed higher classification accuracy for the developed work than extant ones.

Casanova *et al.* (2018) has examined the feasibility for evaluating anatomical index for predicting the risk factor in AD. This was done with ADNI and MRI data. In this, high-dimensional warping schemes were deployed for processing the studies on MRI data. The ADNI MRI data was employed for computing the high-dimensional classifiers. Finally, the enhancement of the suggested model was proved.

In this, many research works worked on different datasets and approaches on AD predictions using machine learning, soft computing techniques and deep learning.

Dataset Description

For studies, the MRI images data is obtained from ANDI Database "Alzheimer's disease neuroimaging initiative database (ADNI) <http://adni.loni.usc.edu/>". The total dataset is consist of 6400 MRI image of different subjects. For analysis purpose the total dataset are split into two datasets with equal number. In Figure 1 exhibits sample images of different subjects of AD. The Figure 1(a) describes Non Dementia under the class Non-AD the remaining figures that are 1(b)

Very Mild Dementia, Figure 1(c) Mild Dementia and Figure 1(d) Moderate Dementia subjects are under class AD.

Proposed Architecture

Figure 2. demonstrates the picture of adopted prediction model. The developed framework for AD prediction includes two main phases such as, “feature extraction, optimal feature selection and classification”. Initially, texture features like GLCM (F^{GLCM}), Haralick features (F^{HL}) are extracted from the input image as the length of features is more, it is planned to carry out optimal feature selection via GOA model. The final feature set $F = F^{GLCM} + F^{HL} + F^{GH}$ is then subjected to optimized DCNN for classification, which ensures accurate diagnosis of the disease.

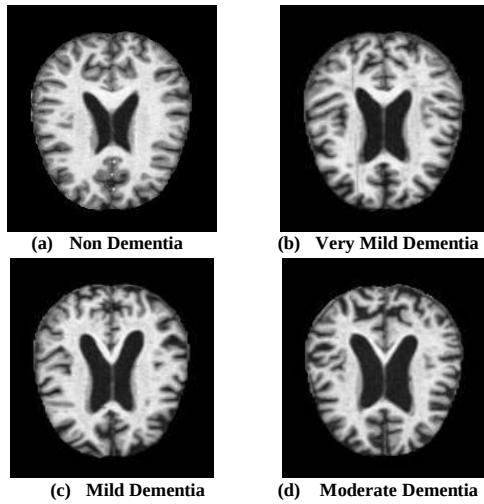


Figure 1. Images of different Demented Subjects

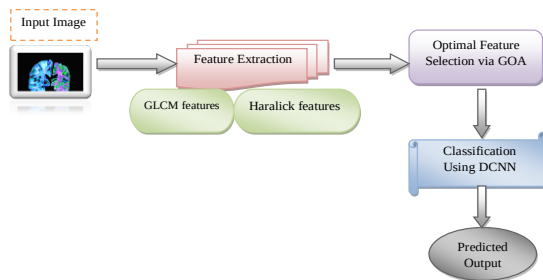


Figure 2. Block diagram of Proposed AD prediction model

Texture Feature Extraction: GLCM, Haralick features

Feature extraction is the first step, where in GLCM, Haralick features are gets extracted from dataset.

GLCM

“GLCM is defined as a statistical method for examining the texture that considers the spatial Relationship of pixels (Gadkari *et al.*, 2004)”. The elucidation on GLCM features is given in Table 1. The extracted features from GLCM are indicated as F^{GLCM} .

Table 1. GLCM features

S.No.	Features	Mathematical Expression
1.	Energy	$E = \sum_u \sum_v g_{du,v}^2$ here $g_{du,v}$ is the $(u,v)^{th}$ entry in GLCM
2.	Entropy	$Ent = -\sum_u \sum_v g_{du,v} \log_2 g_{du,v}$
3.	Contrast	$Con = \sum_u \sum_v (u-v)^2 g_{du,v}$
4.	Variance	$Var = \sum_u \sum_v (u-v)^2 g_{du,v}$, where μ specifies the mean of $g_{du,v}$
5.	Homogeneity	$H = \sum_u \sum_v \frac{1}{1+(u-v)^2} g_{du,v}$
6.	Correlation	$C = \frac{\sum_u \sum_v (uv) g_{du,v} - \mu_x \mu_y}{\sigma_x \sigma_y}$ where $\sigma_x, \sigma_y, \mu_x, \mu_y$, are the std deviations and mean of g_{dx}, g_{dy}
7.	Sum Average	$SA = \sum_{u=2}^{2N_g} u g_{dx+y}(u)$, where N_g indicates the varied gray levels in image.
8.	Sum Entropy	$SE = \sum_{u=2}^{2N_g} g_{dx+y}(u) \log \{g_{dx+y}(u)\}$
9.	Sum Variance	$SV = \sum_{u=2}^{2N_g} (u-SA)^2 g_{dx+y}(u)$
10.	Difference Variance	$DV = \text{variance of } g_{dx-y}$
11.	Difference Entropy	$DE = \sum_{u=0}^{N_g-1} g_{dx-y}(u) \log \{g_{dx-y}(u)\}$
12.	MCC (2 nd higher Eigen value of Q) ⁰⁵	$MCC = \sum_k \frac{g_d(u,k) g_d(v,k)}{g_{dx}(u) g_{dy}(k)}$ where $g_{dx+y}(k) = \sum_{u=1}^{N_g} \sum_{v=1}^{N_g} g_d(u,v)$, $u+v = 2, 3, \dots, 2N_g$
13.	IMC 1	$IMC1 = \frac{HXY - HXY1}{\max\{HX, HY\}}$
		$IMC2 = \sqrt{1 - \exp[-0.2[HXY2 - HXY1]]}$,where
14.	IMC 2	$HXY = -\sum_u \sum_v g_{du,v} \log_2 g_{du,v}$ $HXY1 = -\sum_u \sum_v g_{du,v} \log_2 \{g_{dx}(u) g_{dy}(v)\}$ $HXY2 = -\sum_u \sum_v \left(\frac{g_{dx}(u) g_{dy}(v)}{\log_2 \{g_{dx}(u) g_{dy}(v)\}} \right)$

Haralick texture features

“Haralick texture features evaluate both grey scale distribution in an image that considers the spatial interactions of pixels”. Usually, these types of features are derived by deploying the GLCM model. (Ohmshankar and Paul, 2014). The numerical explanation of GLCM over an $m \times n$ image Im by a displacement $(\Delta y, \Delta x)$ is specified in Equation (1), wherein, j and i signify gray values of Im , q and p point out the spatial position. After computing GLCM, it is normalized by means of the whole count of pixel pairs (Haralick *et al.*, 2019).

$$GLCM_{\Delta y, \Delta x}(j, i) = \sum_{q=1}^m \sum_{p=1}^n \begin{cases} 1, & \text{if } Im(q, p) = i \text{ and } Im(q + \Delta y, p + \Delta x) = j \\ 0, & \text{else} \end{cases} \quad (1)$$

These features together are termed as $F = F^{GLCM} + F^{HL}$. As the length of features is more, it is planned to optimally choose the features using GOA model.

GOA Algorithm

“Grasshopper Optimization Algorithm (GOA)” (Saremi *et al.*, 2017) is inspired by swarming behaviours of grasshoppers. The numerical depiction to expose the swarming behaviour of grasshopper is given in Equation (2), wherein O_i point out i^{th} grasshopper position, V_i point out wind advection, X_i point out gravity force on i^{th} grasshopper and I_i point out social interaction.

$$P_i = I_i + X_i + V_i \quad (2)$$

For creating a random behaviour, Equation (2) is remodelled as revealed in Equation (3), wherein n_1, n_2, n_3 point out random integers among.

$$O_i = n_1 I_i + n_2 X_i + n_3 V_i \quad (3)$$

Equation (4) portrays the social interactions, where a_{ij} indicate the distance amongst i^{th} and j^{th} grasshopper, and it is computed.

As $a_{ij} = |y_j - y_i|$, and $\hat{a}_{ij} = \frac{y_j - y_i}{a_{ij}}$ point out a unit vector from i^{th} to j^{th} grasshopper.

$$I_i = \sum_{j=1, j \neq i}^M r(a_{ij}) \hat{a}_{ij} \quad (4)$$

Equation (5) represents social force's strength r , wherein, g point out the attraction intensity and l point out attractiveness about length scale.

$$r(q) = ge^{\frac{-q}{l}} - e^{-q} \quad (5)$$

The term X and V in Equation (4) is computed as exposed in Equation (6) and Equation (7) respectively, where g indicate gravitational constant, v indicate constant drift and $\hat{e}_u$ indicate unity vector.

$$X_i = -g \hat{e}_u \quad (6)$$

$$V_i = v \hat{e}_w \quad (7)$$

On substituting the values of I , V and X in Equation (10), it is remodelled as per Equation (8), where M point out count of grasshoppers.

$$O_i = \sum_{j=1, j \neq i}^M r(|y_j - y_i|) \frac{y_j - y_i}{a_{ij}} - g \hat{e}_u + v \hat{e}_w \quad (8)$$

Nevertheless, Equation (8) cannot be deployed for resolving the optimization issues and thus it is remodelled as specified in Equation (9), where lb_a and ub_a indicate lower and upper limits. T_a represent value of A^{th} dimension in target and c point out the lessening coefficient.

$$O_i^a = c \left(\sum_{j=1, j \neq i}^M \frac{ub_a - lb_a}{2} r(|y_j^a - y_i^a|) \frac{y_j^a - y_i^a}{a_{ij}} \right) + \hat{T}_a \quad (9)$$

Consequently, in exploration and exploitation phases, the parameter c is minimized based on iteration count and it is specified in Equation (10), where cmx and cmn denotes highest and least values respectively, t indicates current iteration, and L denotes highest iteration count.

$$c = cmx - t \frac{cmx - cmn}{L} \quad (10)$$

Results and Discussions

Experimental Setup

The adopted AD detection approach was implemented in “MATLAB”. The superiority of the GOA model was validated by comparing it with other existing schemes like Firefly Algorithms (FF) (Fister *et al.*, 2013), Genetic algorithm (GA) (McCall *et al.*, 2005). Moreover, the supremacy of adopted model was compared over existing classifiers such as, SVM (Zeng *et al.*, 2018), KNN (Xianlun and Wang, 2019) and NN (Almansour *et al.*, 2019). Here, the analysis was performed regarding varied metrics like accuracy, NPV, MCC and FNR and so on.

Convergence Analysis

The convergence analyses (cost function) for the GOA model over conventional schemes like FF+DCNN, GA+DCNN is shown in Figure 3 with respect to varied iterations that range from 0, 5, 10, 15 and 20 for dataset 1 and 2. According to the results of the study, the proposed approach has the lowest cost value for all iterations as compared to other models, indicating that the developed model is superior. Figure 3(a) shows that the GOA scheme for data 1 has a minimum cost of 1.001, which is 1.89% and 3.39% better than the standard FF+DCNN and GA+DCNN models at the 10th iteration. Similarly, the cost analysis of the GOA+DCNN model for data 2 compared to traditional schemes revealed minimal cost values across all iterations. As a result of the assessment, it is evident that the GOA significantly outperforms the state-of-the-art techniques in terms of cost analysis.

Performance Analysis

In terms of performance, evolved work distinguished from existing models of “positive

measures like accuracy, sensitivity, specificity and precision and negative measures like FPR, FNR and FDR and neutral measures like NPV, F1-score and MCC”. As a result, positive measure resultants are graphically disclosed in Figure.VI, while negative and neutral measure resultants are represented graphically in Figure 4. and Figure 5. Correspondingly. The metrics were examined for varied learning rates such as 60, 70, 80 and 90. The DCNN model outperformed the other approaches when it came to interpreting entire graphs. Minimal negative values, in particular, promise an improved model prediction rate. From Figure. 6(a) the DCNN model for data 1 has achieved better accuracy value, which is 30%, 55% and 60% better than SVM, KNN and NN models when the learning rate is 90. In addition, Table 2. and Table 3. describe the overall analysis of DCNN scheme over traditional schemes for data 1 and 2 respectively on examining the negative measure, In comparison to other models; the developed scheme has exhibited a minimal value. On observing the analysis outcomes, the DCNN model generated

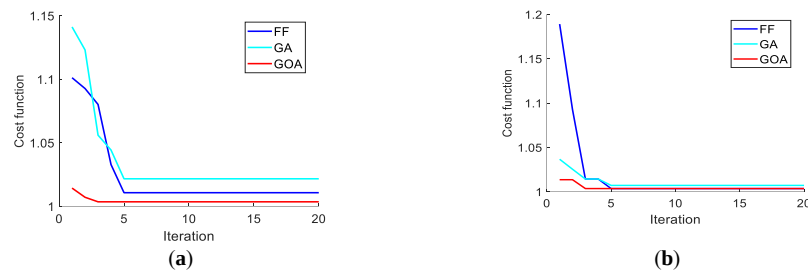


Figure3. Convergence analysis attained by GOA model over other models for (a) data 1 and (b) data 2

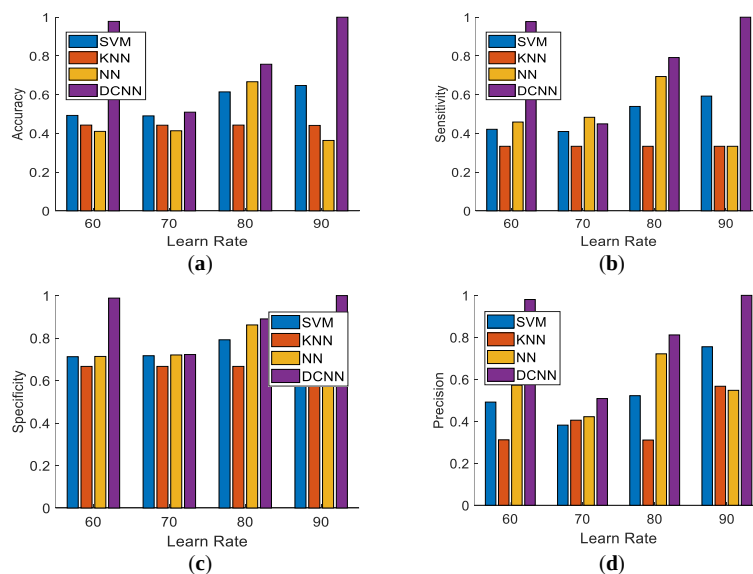


Figure 4. Performance of developed approach over existing approaches in terms of (a) Accuracy (b) Sensitivity (c) Specificity (d) Precision for data 1

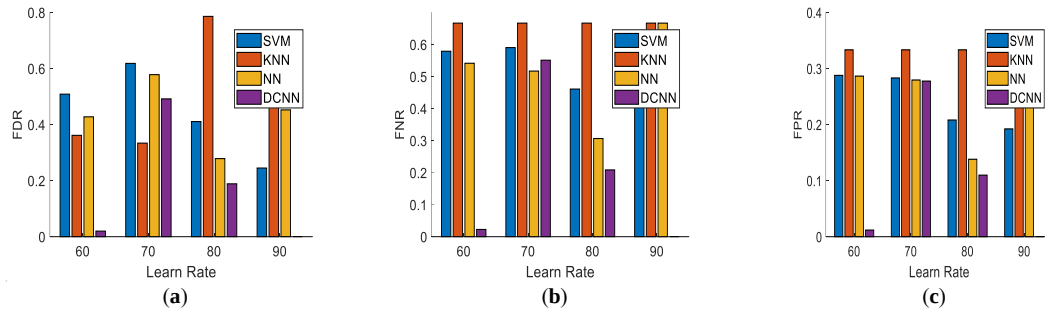


Figure 5. Performance of developed approach over existing approaches in terms of (a) FDR (b) FNR (c) FPR for data 1

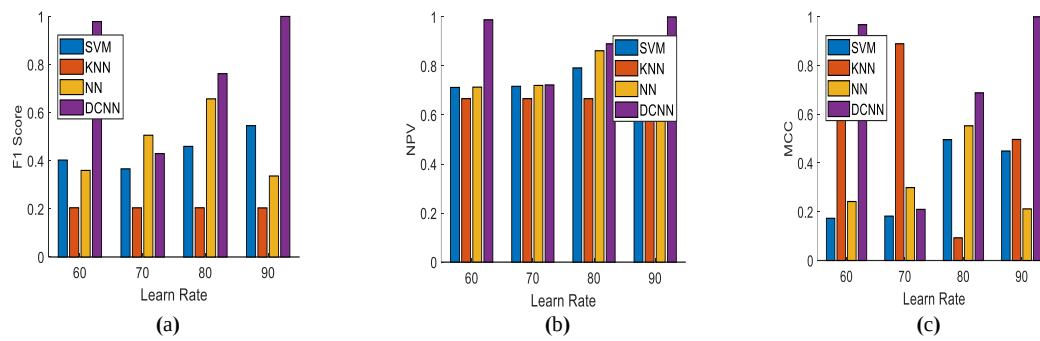


Figure 6. Performance of developed approach over existing approaches in terms of (a) F1-score (b) NPV (c) MCC for data 1

significant precision values for both datasets than the existing schemes. In particular, the adopted scheme for dataset 1 has a better observing accuracy value (0.978), which is superior to traditional, SVM, KNN, and NN models. As a result, the adopted scheme's development has been secured by the high rates of positive measures and lower rates of negative measures.

Analysis on Feature Selection

Since, the optimization model plays a major role in achieving the objective function, it is essential to validate its performance. Hence, in this section, the proposed work with optimal feature selection is compared over the proposed work without feature selection. The corresponding results acquired are tabulated in Table 4. The final feature set is of the size=40. After undergoing the optimal feature selection, the size of the features tends to vary, as only the most relevant features were selected. The size of the features acquire after optimizing with GOA is 23. Moreover, the acquired features before as well as after imply of optimization concept is evaluated in terms of positive performance measures like “accuracy, sensitivity,

specificity and precision”. The accuracy of proposed work before the application of feature selection approach is 0.8. On the other hand, while selecting the optimal feature with GOA, the accuracy of the proposed work boosted up automatically, and has reached the maximal value (0.97857).

The accuracy of the proposed work with feature GOA based optimal feature selection is 17.5% better than the proposed work without GOA based feature section. In addition, the sensitivity of the proposed work with GOA based optimal feature selection is 0.97743, which is the highest value, while compared to the existing optimization techniques like FF=0.84648 and GA=0.88466, In addition, the specificity of the proposed work with GOA based optimal feature selection is 0.98825, which is 8.8% better than the proposed work without GOA feature section. Thereby, from this evaluation, a clear conclusion can be derived that the proposed work generates more precious solutions only after the application of the optimization concept. Moreover, the extracted GOA based features are found to be more superior to the extracted features with FF and GA.

Table 2. Performance Analysis Of Adopted Scheme Over Existing Schemes For Data 1

Learning rate = 60				
Measures	SVM	KNN	NN	DCNN
Sensitivity	0.41564	0.33333	0.40033	0.73101
Accuracy	0.48571	0.44286	0.37313	0.72857
Precision	0.55556	0.38224	0.53228	0.83394
Specificity	0.71169	0.66667	0.68012	0.86988
FPR	0.28831	0.33333	0.31988	0.13012
F1-Score	0.38397	0.20462	0.42185	0.72869
MCC	0.17502	0.047581	0.31205	0.64969
FNR	0.58436	0.66667	0.59967	0.26899
NPV	0.71169	0.66667	0.68012	0.86988
FDR	0.44444	0.40955	0.46772	0.16606
Learning rate = 70				
Sensitivity	0.43841	0.33333	0.41517	0.88889
Accuracy	0.49038	0.44231	0.37234	0.91346
Precision	0.46433	0.81098	0.53224	0.94545
Specificity	0.72305	0.66667	0.70503	0.94828
FPR	0.27695	0.33333	0.29497	0.051724
F1-Score	0.43108	0.20444	0.27738	0.90363
MCC	0.17998	0.23488	0.38889	0.87106
FNR	0.56159	0.66667	0.58483	0.11111
NPV	0.72305	0.66667	0.70503	0.94828
FDR	0.53567	0.31294	0.67291	0.054545
Learning rate = 80				
Sensitivity	0.41893	0.33333	0.48909	0.88889
Accuracy	0.47143	0.44286	0.5	0.91429
Precision	0.38171	0.1971	0.34209	0.92593
Specificity	0.71886	0.66667	0.74762	0.95918
FPR	0.28114	0.33333	0.25238	0.040816
F1-Score	0.3872	0.20462	0.38835	0.89167
MCC	0.16899	0.68416	0.42385	0.86642
FNR	0.58107	0.66667	0.51091	0.11111
NPV	0.71886	0.66667	0.74762	0.95918
FDR	0.61829	0.38635	0.62721	0.074074
Learning rate = 90				
Sensitivity	0.36667	0.33333	0.46667	0.88889
Accuracy	0.41176	0.44118	0.4	0.91176
Precision	0.51494	0.41831	0.44872	0.94444
Specificity	0.69298	0.66667	0.69583	0.94737
FPR	0.30702	0.33333	0.30417	0.052632
F1-Score	0.31096	0.20408	0.67027	0.90303
MCC	0.23596	0.93163	0.30324	0.86974
FNR	0.63333	0.66667	0.53333	0.11111
NPV	0.69298	0.66667	0.69583	0.94737
FDR	0.40158	0.36023	0.55128	0.055556

Table 3. Performance Analysis Of Adopted Scheme Over Existing Schemes For Data 2

Learning rate = 60				
Measures	SVM	KNN	NN	DCNN
Sensitivity	0.4211	0.33333	0.45859	0.97743
Accuracy	0.4928	0.44286	0.41045	0.97857
Precision	0.49157	0.31182	0.57236	0.98016
Specificity	0.7123	0.66667	0.71359	0.98825
FPR	0.2877	0.33333	0.28641	0.011752
F1-Score	0.40294	0.20462	0.35989	0.97872
MCC	0.17311	0.88699	0.242	0.96746
FNR	0.5789	0.66667	0.54141	0.022572
NPV	0.7123	0.66667	0.71359	0.98825
FDR	0.50843	0.36152	0.42764	0.019841
Learning rate = 70				
Sensitivity	0.40973	0.33333	0.48307	0.44908
Accuracy	0.49038	0.44231	0.41346	0.50962
Precision	0.38179	0.4053	0.42188	0.50833
Specificity	0.7169	0.66667	0.72052	0.72249
FPR	0.2831	0.33333	0.27948	0.27751
F1-Score	0.36639	0.20444	0.50621	0.42976
MCC	0.18203	0.88915	0.29908	0.20975
FNR	0.59027	0.66667	0.51693	0.55092
NPV	0.7169	0.66667	0.72052	0.72249
FDR	0.61821	0.33375	0.57812	0.49167
Learning rate = 80				
Sensitivity	0.53917	0.33333	0.69355	0.7916
Accuracy	0.61429	0.44286	0.66667	0.75714
Precision	0.52204	0.31076	0.72145	0.81145
Specificity	0.79191	0.66667	0.8619	0.89024
FPR	0.20809	0.33333	0.1381	0.10976
F1-Score	0.45979	0.20462	0.65738	0.76213
MCC	0.49588	0.092971	0.55264	0.68789
FNR	0.46083	0.66667	0.30645	0.2084
NPV	0.79191	0.66667	0.8619	0.89024
FDR	0.4108	0.78594	0.27855	0.18855
Learning rate = 90				
Sensitivity	0.59259	0.33333	0.33333	1
Accuracy	0.64706	0.44118	0.36364	1
Precision	0.75513	0.56695	0.54762	1
Specificity	0.80775	0.66667	0.64583	1
FPR	0.19225	0.33333	0.35417	0
F1-Score	0.54617	0.20408	0.33669	1
MCC	0.44919	0.49664	0.21168	1
FNR	0.40741	0.66667	0.66667	0
NPV	0.80775	0.66667	0.64583	1
FDR	0.24487	0.57144	0.45238	0

Table 4. Analysis on Feature Extraction: With Optimization

Methods	Feature Size	Feature Size With FS	Accuracy With FS	Sensitivity With FS	Specificity With FS
FF	40	26	0.87857	0.84648	0.92822
GA	40	20	0.9	0.88466	0.94338
GOA	40	23	0.97857	0.97743	0.98825

Conclusion

This paper had developed a novel AD prediction model using GOA model. At first, GLCM, Haralick features were extracted. Subsequently, optimal feature selection takes place, where the optimal features were chosen via a metaheuristic model. Further, classification takes place using DCNN; finally, the superiority of the proposed scheme over

existing models was demonstrated using a variety of metrics. Particularly, the DCNN model for data 1 has achieved better accuracy value, which is 30%, 55% and 60% better than SVM, KNN and NN models when the learning rate is 90. On examining the negative measure, the presented model has revealed a minimal value over other compared models. Therefore, the supremacy of the GOA approach has been confirmed effectively.

References

- Almansour, N.A., Syed, H.F., Khayat, N.R., Altheeb, R.K., Juri, R.E., Alhiyafi, J., and Olatunji, S.O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in biology and medicine.*, 109:101-111.
- Batzu, L., Westman, E., and Pereira, J.B.(2020). Cerebrospinal fluid progranulin is associated with increased cortical thickness in early stages of Alzheimer's disease. *Neurobiology of aging.*, 88:61-70.
- Casanova, R., Barnard, R.T., Gaussoin, S.A., Saldana, S., Hayden, K.M., and Manson, J.E. (2018). Using high-dimensional machine learning methods to estimate an anatomical risk factor for Alzheimer's disease across imaging databases. *Neuroimage.*, 183:401-411.
- Faisal, F.U.R., Khatri, U., and Kwon, G.R. (2021). Diagnosis of Alzheimer's Disease using Combined Feature Selection Method. *J. of Korea Multimedia Society.*, 24(5):667-675.
- Fister, I., Fister Jr., I., Yang, X.S., and Brest, J. (2013). A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation.*, 13:34-46.
- Gadkari, D. (2004). Image quality analysis using GLCM.
- Haralick, R.M., Shanmugam, K., and Dinstein, I.H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics.*, (6):610-621.
- Lahmiri, S. and Shmuel, A. (2019). Performance of machine learning methods applied to structural MRI and ADAS cognitive scores in diagnosing Alzheimer's disease. *Biomedical Signal Processing and Control.*, 52:414-419.
- McCall, J. (2005). Genetic algorithms for modelling and optimisation. *J. of computational and Applied Mathematics.*, 184(1):205-222.
- McKenzie, C., Bucks, R.S., Weinborn, M., Bourgeat, P., Salvado, O., and Gavett, B.E. (2020). Cognitive reserve predicts future executive function decline in older adults with Alzheimer's disease pathology but not age-associated pathology. *Neurobiology of aging.*, 88:119-127.
- Moscato, A., Silva-Rodríguez, J., Aldrey, J.M., Cortés, J., Fernández-Ferreiro, A., and Gómez-Lado, N. (2019). Prediction of Alzheimer's disease dementia with MRI beyond the short-term: Implications for the design of predictive models. *NeuroImage: Clinical.*, 23:101837.
- Ohmshankar, S. and Paul, C.K.C. (2014). Haralick features based mammogram classification system. In: *Second Int. Conference on Current Trends In Engineering and Technology-ICCTET 2014.* p. 409-413
- Park, Y.H., Hodges, A., Risacher, S.L., Lin, K., Jang, J.W., and Ahn, S. (2020). Dysregulated Fc gamma receptor-mediated phagocytosis pathway in Alzheimer's disease: network-based gene expression analysis. *Neurobiology of aging.*, 88:24-32.
- Raj, R. Joshua Samuel. (2020). Optimal feature selection-based medical image classification using deep learning model in internet of medical things. *IEEE Access* 8 (2020):58,006-58,017.
- Samper-González, J., Burgos, N., Bottani, S., Fontanella, S., Lu, P., and Marcoux, A. (2018). Reproducible evaluation of classification methods in Alzheimer's disease: Framework and application to MRI and PET data. *NeuroImage.*, 183:504-521.
- Saremi, S., Mirjalili, S., and Lewis, A. (2017). Grasshopper optimisation algorithm: theory and application. *Advances in Engineering Software.*, 105:30-47.
- Shaikh, T.A. and Ali, R. (2019). Automated atrophy assessment for Alzheimer's disease diagnosis from brain MRI images. *Magnetic resonance imaging.*, 62:167-173.
- Silvia Basaia, Federica Agosta, Luca Wagner, and Elisa Canu, Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deepneural networks. *NeuroImage: Clinical.*, 21(2019).
- Tadayon, E., Pascual-Leone, A., Press, D., Santarnecchi, E., and Alzheimer's Disease Neuroimaging Initiative. (2020). Choroid plexus volume is associated with levels of CSF proteins: relevance for Alzheimer's and Parkinson's disease. *Neurobiology of aging.*, 89:108-117.
- Vaithinathan, K., Parthiban, L., and Alzheimer's Disease Neuroimaging Initiative. (2019). A novel texture extraction technique with T1 weighted MRI for the classification of Alzheimer's disease. *J. of neuroscience methods.*, 318:84-99.
- Xianlun TangTing. and Wang Yuyan Dai. (2019). Motor imagery EEG recognition with KNN-based smooth auto-encoder. *Artificial Intelligence in Medicine.*, (101):101747
- Zeng, N., Qiu, H., Wang, Z., Liu, W., Zhang, H., and Li, Y. (2018). A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease. *Neurocomputing.*, 320:195-202.