International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 12, No.1 (Oct-2022)

https://dx.doi.org/10.12785/ijcds/120189

Alzheimer's disease Prediction by Hybrid CNN and SVM Classifier with Metaheuristic Approach

G Stalin Babu[1], S.N.Tirumala Rao[2] and R Rajeswara Rao[3]

[1]*Dept. of CSE, Aditya Institute of Technology and Management, Tekkali, India.*
[1]*Dept. of CSE, JNTUK, Kakinada, India.*
[2]*Dept. of CSE, Narasaraopeta Engineering College, Narasaraopeta, India.*
[3]*Dept. of CSE, JNTUK University College of Engineering, Vizianagaram, India.*

Received 22 Jan. 2021, Revised 15 Jul. 2022, Accepted 23 Jul. 2022, Published 31 Oct. 2022

Abstract: The most common type of dementia is Alzheimer's disease (AD). It is critical to identify the AD at the stage of Mild Cognitive Impairment (MCI) in early. If it is possible to early identification, then it has more chance to cure the disease. This paper implements a novel predictive approach for early detection of AD utilizing Magnetic Resonance Imaging (MRI) images. The developed model involves Feature Extraction, Optimal Feature selection, Classification. At first, the Gray Level Co-Occurrence Matrix (GLCM), Haralick features, and geometric Haralick feature techniques are used to extract the geometric correlation and variances features. This work carries out optimal feature selection using the Combined Grey Wolf -Dragon Updating (CG-DU) hybrid model. This optimization model has been used in Convolutional Neural Network (CNN) for the optimized weights and activation function. Optimally chosen features by CNN are subjected to the Classifier Support Vector Machine (SVM) for AD classification. The final output is obtained from both CG-DU+CNN and SVM outcomes. In the end, the performance of the implemented approach is computed to the existing approaches based on various metrics.

Keywords: Alzheimer's disease Prediction; Feature Extraction; Classification; CNN; SVM; Optimization

1. Introduction

AD is an irreversible and degenerative brain disorder in which subtle brain alterations began before the symptoms appear. Early symptoms of AD include modest cognitive decline that could escalate to severe functional and physical problems [6]. Cortical neurofibrillary degeneration, severe brain shrinkage, indications of extensive limbic, and beta amyloidal deposition are the key indicators. Based on the dynamics of various biomarkers in AD, a computational neurodegenerative disease progression score is presented [7]. Clinical assessments are commonly used to monitor AD development, although biomarkers such as cognitive evaluation, structural MRI, 18-FDG-PET, Electroencephalography (EEG) and Cerebrospinal Fluid (CSF) could be used. Regional brain volume and cortical thickness are common MRI biomarkers for identifying AD progression, but glucose hypo metabolism in neocortical brain areas is the most important biomarker of FDG-PET. A rise in CSF Phospho-tau or t-tau has also been identified as a possible biomarker of illness progression [8].

Genetic information, age, ethnicity, and years of schooling are all risk factors that unconventional measures for AD with Neuroimaging modalities. This additional finding confirmed that age is a key factor in the development of AD. The APOE gene is also widely recognized as the most significant hereditary risk factor. According to reports of WHO, AD is a crucial disease of death in aged people that it is in the third position after cancer and heart diseases. Persons at risk of acquiring AD must be identified for therapeutic treatments tested [9]. A summary of MCI categorization was already given by a number of studies. Early diagnosis might aid in the enrollment of patients and the testing of potential novel Alzheimer's medication therapies. Numerous studies have found that imaging is critical for early identification of AD [10]. Diagnosis of MRI images by doctors could take time and be expensive. So, auto-diagnosis of AD images is effective and efficient with low cost and less time. [14] [15].

EMR data contains heterogeneous structures that make ML techniques difficult to use [11]. Researchers utilized a collection of assorted ambulatory EMR data obtained from major care physician offices across the United States for the study. The study's data came from EMR and IQVIA suppliers, and it was then translated into the OMOP style. ML algorithms used MRI for AD diagnosing. By the Neuroimaging data, an appliance of RNN algorithms was

E-mail address: stalin1227@gmail.com, nagatirumalarao@gmail.com, rrraao.cse@jntukucev.ac.in

created to distinguish from healthy controls of AD patients [12] [13]. Further, the longitudinal EMRs were utilized to examine the course of persistent diseases including AD, and the Long Short-Term Memory (LSTM), Recurrent Neural Network (RNN) was found to be an excellent predictor of AD sequence by utilizing the medical and temporal patterns of patient visits.

The key contribution of the presented approach is listed below:

- Proposes the hybrid classifier that includes CNN and SVM for AD diagnosis, where the activation function and weight of CNN are tuned optimally through a CG-DU Model.

The paper is ordered as: Section 2 depicts the reviews on AD prediction scheme and Section 3 addresses the proposed model for AD prediction. Section 4 portrays the texture feature extraction process. Section 5 describes the proposed CG-DU model for optimal selection of features, weight and activation function. Section 6 depicts the classification using hybrid classifiers: SVM and CNN. Section 7 discussed the results and the conclusion of paper is given in Section 8.

2. LITERATURE REVIEW

In 2020, Branimir et al. [1] have introduced DL techniques for AD prediction. Moreover, the knowledge based on the Medical domain was applied to construct a positive dataset from data relevant to AD. SCRP datasets were fed into an LSTM RNN DL model to predict if a patient is with AD. In an SCRP AD dataset with no. of patients 2,324, risk score prediction utilising drug domain information produced a high out-of-sample score of 98

In 2019, Zhiguang et al. [5] have implemented the DL model for forecasting the risk in AD via brain 18F-FDG PET. From the ADNI database, the 350 MCI participants were chosen as research objects, and the CAFFE was chosen as the DL framework; each participant's FDG PET image features were taken out using a DCNN approach to construct the classification and prediction approaches; thus, the MCI stage features were ranked and classified. The findings revealed that the better specificity and sensitivity were attained in terms of MCI transformation.

In 2019, Solale et al. [4] has offered a new ML algorithm for forecasting the development of AD. Each task was precisely described as regression approaches at a single time point that forecast cognitive scores. As the subjects have different combinations of recording modalities as well as other demographic, genetic, and neuropsychological risk factors, special attention is provided, and each modality might experience a specific pattern of sparsity. Even with an inadequate longitudinal dataset, the results have shown reduced prediction errors.

In 2020, Branimir et al. [1] have introduced DL techniques for AD prediction. Moreover, the knowledge based on the Medical domain was applied to construct a positive dataset from data relavant to AD. SCRP datasets were fed into an LSTM RNN DL model to predict if a patient is with AD. In a SCRP AD dataset, the risk scores AD diagnosis via drugs domain information of 2,324 patients attained larger out-of-sample score with high AUPRC. Even with the naive dataset selection, the model performed considerably better.

In 2020, Baiying et al. [3] have introduced Deep Learning(DL) of longitudinal data for AD diagnosis. Researchers suggest developing a system for predicting clinical ratings based on longitudinal data from various time points. The proposed system is divided into three sections that include feature encoding using deep polynomial networks, feature selection using correntropy regularised joint learning, and ensemble learning for regression using the SVRM. Extensive research on the ADNI public database demonstrated that the adopted model was capable of revealing the link between MRI data and clinical score beats state-of-the-art traditional techniques in score recognition.

The benefits and pitfalls of state-of-the-art models in AD prediction in MRI images are discussed in the sections. Despite the presence of many more deep learning and machine learning techniques, disease prediction at an early stage remains unmanageable. The following are some of the pros and cons of the current works: CNN has superior patient care and assessment, as well as superior early identification of Alzheimer's disease. Nonetheless, appropriate diagnostic tools must be offered, as well as further improvements in the accuracy.

3. PROPOSED MODEL FOR AD PREDICTION

The adopted framework for AD prediction consists of 3 main stages like feature extraction, optimal feature selection and classification. During the feature extraction phase, the texture features such as GLCM (F^{GLMC}), Haralick features (F^{HL}) and geometric Haralick features (F^{GH}) are taken out from the input image . Moreover, from the extracted features, it is planned to select the optimal features via CU-GU model. The optimal selected features are then subjected to the classification phases that include the hybrid classifier like SVM and CNN. In addition, the activation function and weights of the CNN are tuned optimally via a new CG-DU algorithm. Figure. 1 depicts the framework of adopted AD prediction model.

4. TEXTURE FEATURE EXTRACTION PROCESS

At first, the texture features like GLCM, Haralick and geometric Haralick features such as the variance and geometric correlation are taken out during the feature extraction process.

A. GLMC

It is a statistical approach to examine the texture which regards the pixels spatial relationship [22]. The GLCM features are determined as per [16]. Moreover, the GLCM features are indicated as GLCM (F^{GLMC}).

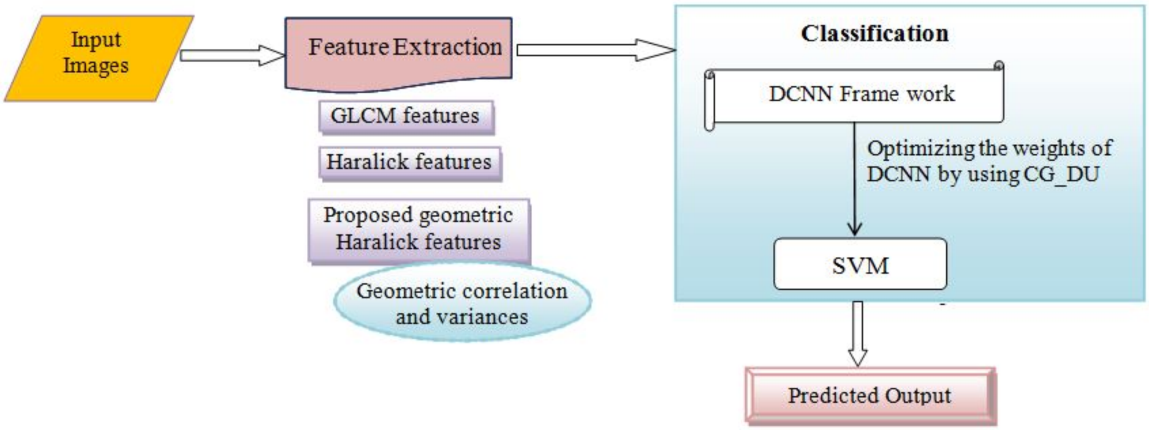

Figure 1. Framework of adopted AD prediction model

B. Haralick texture features

In an image, the Haralick texture features assess both grey scale distribution which considers the pixel's spatial interactions. Normally, these features are obtained by organizing the GLCM approach [23]. These features are signified as (F^{HL}), and it is determined in [16].

C. Geometric Haralick Features

This proposed work developed two haralick features such as proposed variance and geometric correlation. It is determined as per [16]. The features extracted from geometric correlation and variance is specified as . The overall features are known as $FE = (F^{GLMC}) + (F^{HL}) + (F^{GH})$.

5. PROPOSED CG-DU MODEL FOR OPTIMAL SELECTION OF FEATURES, WEIGHT AND ACTIVATION FUNCTION

CU-GU model is used to select the optimal features from the extracted features. Finally, the optimal selected features are indicated as OP.

A. Solution Encoding

The weights and activation function of CNN are optimally tuned using the proposed CG-DU method. Figure. 2 illustrates the input solution of the adopted CG-DU model. Here, denotes the entire number of CNN weights.Further, the objective function obj is determined in Eq. (1), $LOSS$ where is determined using Eq. (7).

$$Obj = Min(Loss) \quad (1)$$

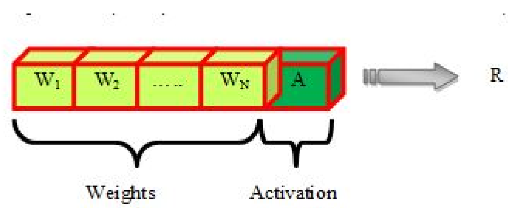

Figure 2. Solution Encoding

B. Proposed CG-DU model

Even though, the traditional DA model [19] provides accurate estimations; however, it suffers from few drawbacks such as slow convergence and minimal internal memory. For overcoming the demerits of traditional DA, the GWO concept [20] is integrated with it for implementing a novel CG-DU model. Hybrid optimization models were report to be capable for definite search troubles [21]. Moreover, the procedure of the adopted CG-DU approach is given in [16].

while the iteration is superior or equal to 2 $(t \geq 2)$, the update is performed by the DA and GWO model. Particularly, when $(f_r^t) < (f_r^{t-1})$ or $t = 2$, the update is performed by DA approach, in which (f_r^{t-1}) specifies the previous fitness and (f_r^t) indicates the current fitness. Else if, $(f_r^t) < (f_r^{t-2})$, the update is done using the GWO model. Else if, the random solution is updated. Further, if the condition $(t \geq 2)$ is not satisfied, the update is done by DA model

6. HYBRID CLASSIFIERS:SVM AND CNN

A. SVM

The optimally selected features OP are provided as the input of SVM. SVM is defined for the operation of simple nonlinear regression [18]. An SVM classifier's main objective is to find a function $K(T)$ that defines the hyperplane or decision boundary. The hyperplane separates two classes of input data points in the most efficient way possible. Where P signifies the margin from the hyperplane distance to the closest data points for both classes. The saddle purpose, which decreases quadratic programming problems to find the separating hyper plane of the SVM model, is used to perform the double issues of the Lagrange task. The maximum margin is used by the SVM classifier

Algorithm for CG-DU Model
Initialization
Calculate the fitness function (f_r^0)
$t \leftarrow 1$
while $(t < t_{max})$
if $(t \geq 2)$
for each solution of S_r
if $(f_r^t) < (f_r^{t-1})$ or $t = 2$
Using the DA model, update the position vector.
else if $(f_r^t) < (f_r^{t-2})$
GWO model is used to update position vector.
else
The random solution will be used for update.
end for
else
The DA model is used to updating the position vector.
end if
$t = t + 1$
end while
return the optimal solution as S_r^*.

to locate the decision boundary among all feasible hyper planes. Further, $||Z||$ reduces the subject for maximizing the P , and it is determined in Eq. (2). Where, p denotes the margin for both classes from the hyper plane distance to the nearest point of data points.

$$min \frac{||Z||^2}{2} subject\ to\ \forall \tilde{o}\ \tilde{x}_{\tilde{O}} \left(Z.T_{\tilde{O}} + \tilde{b} \right) \geq 1 \quad (2)$$

In Eq. (2), $\tilde{b}$ refers to the value of a scalar threshold, $\tilde{O}$ in SVM represents the no. of data inputs,$T_{\tilde{O}}$ the data points in the input, and Z portrays the boundary was specified by a vector. $K(T)$ indicates the attained optimal hyper plane of SVM as given in Eq. (3).

$$K(T) = \sum_{\tilde{O}=1}^{s} \tilde{x}_{\tilde{o}} S_{\tilde{o}} \langle T_{\tilde{o}}.T \rangle + \hat{b} \quad (3)$$

Here,$T_{\tilde{o}}$ is a the support vector with non-zero Lagrange multiplier($S_{\tilde{o}}$). Outside of support vectors, it is not necessary to determine the $K(T)$ by the data points. Furthermore, the SVM classifier, when used with the default configuration, produces excellent information grouping. The perk up area is critical for reducing the precision of the SVM order of wrongly grouped objects. The ordered products were placed near the separating hyper plane by inadvertently. As a consequence, the superiority of the additional equipment is utilised inside the separating strip. The outcome of SVM is indicated as CLS_{SVM} .

B. Optimized CNN

The features that are optimally determined are fed into an optimised CNN[17]. CNN is the known classifier that includes 3 layers: pooling, convolution, and fully connected layers. There are multiple convolution kernels in the convolution layer. Kernel values are used to determine the complete feature map. Here, the total feature map was evaluated through several kernels. Moreover, the feature values in the position (a, b) are represented as $H^v_{a,b,c}$ as given in Eq. (4), and the V^{th} layers matched to w^{th} feature map. Similarly, the w^{th} filter value is determined in the v^{th} layer.

Furthermore, the bias term and the weight vector are symbolized as D^v_c and W^v_c, correspondingly. w^{th} layer at position(a, b) the associated input patches are represented using $U^v_{a,b}$. The activation function, which predicts the nonlinear properties of multi-layer networks, is used to generate non-linearity. Assume the activation value $(V^v_{a,b,c})$ and the nonlinear activation function as $V(\bullet)$ as given by Eq. (5). Further, the shift-variance is evaluated in the pooling layer by minizing the feature maps resolution as represented in Eq. (6). The pooling function is portrayed $pool()$ in ever feature map and for every feature map the local neighbourhood $(V^v_{a,b,c})$ at neighbouring position (a, b) is specified as $J_{a,b}$.

$$H^v_{a,b,c} = W^v_c U^v_{a,b} + D^v_c \quad (4)$$

$$V^v_{a,b,c} = V(H^v_{a,b,c}) \quad (5)$$

$$SP^v_{a,b,c} = pool(V^v_{a,b,c}), \forall (m, n) \in J_{a,b} \quad (6)$$

In CNN, the loss function is determined in Eq. (7). The constraints (Θ) of CNN are linked to the necessary IO input-output relative, and it is specified as $\{(U^p, V^p) ; p \in [1 \cdots \cdot IO]\}$. In addition, the related target values, p^{th} input data, and the output of CNN are represented as U^p,V^p and OUT^p, respectively.

$$Loss = \frac{1}{N}\sum_{p=1}^{x} PL(\Theta; V^p, OUT^p) \tag{7}$$

The classified outcomes of the optimized CNN classifier is denoted as CLS_{CNN} .Thus, the overall classified output CLS is the classified outcomes such as SVM, and CNN, and it is given in Eq. (8).

$$CLS = (CLS_{CNN}, CLS_{SVM)}) \tag{8}$$

7. RESULTS AND DISCUSSIONS

A. Experimental Setup

The presented AD detection with hybrid Classifier + CG-DU model was employed in MATLAB and the outcome was noticed. Furthermore,the data was collected from the ADNI database "adni.loni.usc.edu". The data is spited in to 80% of training and 20% of testing. Moreover, the improvement of the implemented hybrid Classifier + CG-DU approach was computed to other conventional classifiers including SVM [18], CNN [17], RNN [26], DBN [24], LSTM [25], and DCNN+CG-DU [16], respectively. Here, an improvement of the presented approach was computed to optimization algorithms including hybrid Classifier + GOA [28], hybrid Classifier + DHO [30], hybrid Classifier + LA [29], and hybrid Classifier + SMO [27], correspondingly in terms of certain metrics like specificity, accuracy, FPR, precision, FDR, sensitivity, F1-score, MCC, FNR and NPV.

B. Convergence Analysis

The convergence analysis for the presented hybrid Classifier + CG-DU approach over extant schemes like hybrid Classifier + GOA, hybrid Classifier + DHOA, hybrid Classifier + LA, and hybrid Classifier + SMO is shown in Figure. 3 with respect to varied iterations that range from 0, 5, 10, 15 and 20 for data 1 and 2. Further, the adopted scheme has shown lower cost value for all iterations than the extend models and it ensure the superior performance of the proposed hybrid Classifier + CG-DU model.

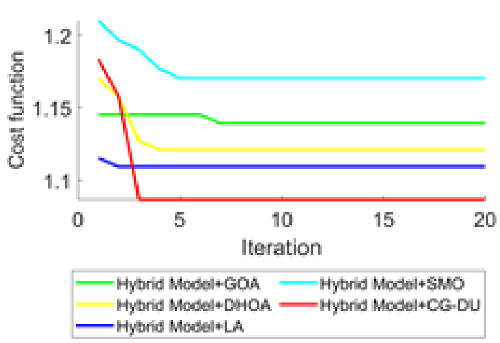

Figure 3. Convergence analysis attained by adopted schemes to the existing appraoches for data 1

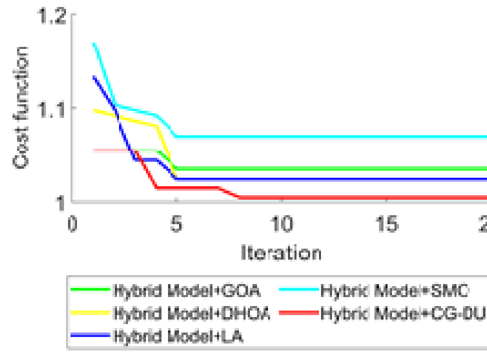

Figure 4. Convergence analysis attained by adopted schemes to the existing appraoches for data 2

In Figure. 4 , the adopted hybrid Classifier + CG-DU scheme has attained minimal cost of ~ 0.003 for data 1 than other traditional hybrid Classifier + GOA (~ 1.14), hybrid Classifier + DHOA (~ 1.125), hybrid Classifier + LA (~ 1.11), and hybrid Classifier + SMO (~ 1.17) models at 20th iteration. Similarly, the cost analysis for the hybrid Classifier + CG-DU model over conventional schemes for data 2 has exhibited lower cost values for all iterations. Therefore, it is shown clearly that the proposed hybrid Classifier + CG-DU model has shown maximum outcomes in cost evaluation than other models.

C. Performance Analysis

The performance of adopted model over the existing schemes based on other metrics is represented in Figure. 5. Moreover, the measures were determined for different learning rates like 60, 70, 80 and 90. On the basis of the graphs, the hybrid Classifier + CG-DU model outperformed the other schemes.

In Figure. 5(a). the hybrid Classifier + CG-DU approach has attained higher accuracy value than other existing SVM, CNN, RNN, DBN, LSTM, and DCNN+CG-DU models when the learning rate is 80. From the Figure. 5. the adopted scheme has shown lower values with better performance than other compared models for negative measures. In addition, the adopted scheme has attained higher MCC value (0.99), which is superior to traditional SVM, CNN, RNN, DBN, LSTM, and DCNN+CG-DU models. Thus, the larger positive measures and minimum negative measures were obtained by the proposed model.

D. Performance Analysis: Traditional Vs. adopted Optimization methods

Table I- II represents the performance analysis of the presented hybrid classifier +CG-DU model to the existing optimization approaches for different learning rate. Moreover, the presented hybrid classifier +CG-DU approach is noted that superior development in accurate AD diagnosis. In addition, the assessment is performed for various learning

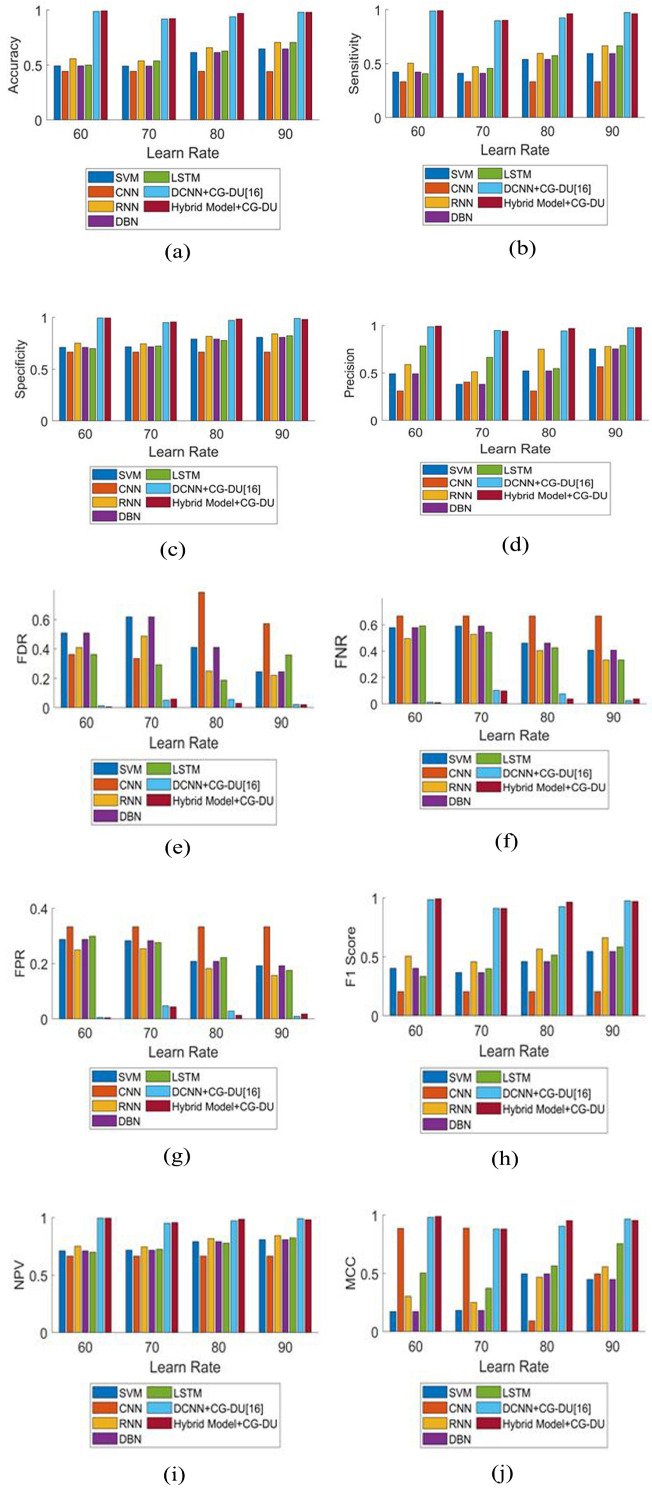

Figure 5. Performance of developed approach over other schemes in terms of (a) Accuracy (b) Sensitivity (c) Specificity (d) Precision (e) FDR (f) FNR (g) FPR (h) F1-Score (i) NPV (j) MCC

TABLE I. PERFORMANCE ANALYSIS OF PROPOSED APPROACH OVER TRADITIONAL OPTIMIZATION MODELS FOR DATASET 1

Measures	Hybrid classifier + GOA [28]	Hybrid classifier + DHO [30]	Hybrid classifier + LA [29]	Hybrid classifier + SMO [27]	Hybrid classifier + CG-DU
FDR	0.11587	0.12302	0.10706	0.11526	0.058664
Sensitivity	0.77778	0.77778	0.80247	0.79012	0.90123
FNR	0.22222	0.22222	0.19753	0.20988	0.098765
Precision	0.88413	0.87698	0.89294	0.88474	0.94134
FPR	0.098725	0.096363	0.08723	0.091797	0.043615
F1-Score	0.76911	0.76546	0.80099	0.78351	0.9112
MCC	0.73246	0.72993	0.76296	0.74651	0.88161
Specificity	0.90128	0.90364	0.91277	0.9082	0.95638
NPV	0.90128	0.90364	0.91277	0.9082	0.95638
Accuracy	0.82692	0.82692	0.84615	0.83654	0.92308

TABLE II. PERFORMANCE ANALYSIS OF PROPOSED APPROACH OVER TRADITIONAL OPTIMIZATION MODELS FOR DATASET 2

Measures	Hybrid classifier + GOA [28]	Hybrid classifier + DHO [30]	Hybrid classifier + LA [29]	Hybrid classifier + SMO [27]	Hybrid classifier + CG-DU
FDR	0.079726	0.08453	0.21624	0.18211	0.04061
Sensitivity	0.85185	0.83951	0.78154	0.80837	0.93827
FNR	0.14815	0.16049	0.21846	0.19163	0.061728
Precision	0.92027	0.91547	0.78376	0.81789	0.95939
FPR	0.066604	0.072351	0.10698	0.092578	0.026374
F1-Score	0.86167	0.84829	0.78006	0.80903	0.94504
MCC	0.82455	0.81019	0.67806	0.7235	0.9247
Specificity	0.9334	0.92765	0.89302	0.90742	0.97363
NPV	0.9334	0.92765	0.89302	0.90742	0.97363
Accuracy	0.88462	0.875	0.79808	0.82692	0.95192

TABLE III. TABLE III. PERFORMANCE ANALYSIS OF PROPOSED APPROACH TO EXTANT CLASSIFIERS FOR DATASET 1

Measures	SVM [18]	CNN [17]	RNN [26]	DBN [24]	LSTM [25]	DCNN+ CG-DU [16]	Hybrid classifier + CG-DU
Specificity	0.7169	0.66667	0.74564	0.7169	0.72414	0.95204	0.95638
Precision	0.38179	0.4053	0.51319	0.38179	0.66658	0.94872	0.94134
Accuracy	0.49038	0.44231	0.53846	0.49038	0.53846	0.91968	0.92308
Sensitivity	0.40973	0.33333	0.47146	0.40973	0.45679	0.89744	0.90123
FPR	0.2831	0.33333	0.25436	0.2831	0.27586	0.047962	0.043615
F1-Score	0.36639	0.20444	0.45869	0.36639	0.39923	0.91162	0.9112
FDR	0.61821	0.33375	0.48681	0.61821	0.29128	0.051282	0.058664
NPV	0.7169	0.66667	0.74564	0.7169	0.72414	0.95204	0.95638
FNR	0.59027	0.66667	0.52854	0.59027	0.54321	0.10256	0.098765
MCC	0.18203	0.88915	0.25048	0.18203	0.37238	0.88044	0.88161

TABLE IV. TABLE III. PERFORMANCE ANALYSIS OF PROPOSED APPROACH TO EXTANT CLASSIFIERS FOR DATASET 2

Measures	SVM [18]	CNN [17]	RNN [26]	DBN [24]	LSTM [25]	DCNN+ CG-DU [16]	Hybrid classifier + CG-DU
Specificity	0.7169	0.66667	0.74564	0.7169	0.72414	0.95204	0.95638
Precision	0.38179	0.4053	0.51319	0.38179	0.66658	0.94872	0.94134
Accuracy	0.49038	0.44231	0.53846	0.49038	0.53846	0.91968	0.92308
Sensitivity	0.40973	0.33333	0.47146	0.40973	0.45679	0.89744	0.90123
FPR	0.2831	0.33333	0.25436	0.2831	0.27586	0.047962	0.043615
F1-Score	0.36639	0.20444	0.45869	0.36639	0.39923	0.91162	0.9112
FDR	0.61821	0.33375	0.48681	0.61821	0.29128	0.051282	0.058664
NPV	0.7169	0.66667	0.74564	0.7169	0.72414	0.95204	0.95638
FNR	0.59027	0.66667	0.52854	0.59027	0.54321	0.10256	0.098765
MCC	0.18203	0.88915	0.25048	0.18203	0.37238	0.88044	0.88161

rate under varied measures. Therefore, from Table I, the adopted scheme is higher accuracy (0.92308) that is higher to the extant optimization approaches. In dataset 2, the precision of the adopted scheme provides maximum value as 0.95939 that is superior to the existing model.

E. Classifier Performance

Table III-Table IV determines the performance of the implemented hybrid classifier +CG-DU model to the convolutional classifiers for various learning percentage in dataset 1 and dataset 2. From the table, the presented hybrid classifier +CG-DU scheme has attained higher accuracy for both dataset. Then, the FNR value of the proposed CG-DU method is lower with better performance over the traditional classifiers including like SVM, CNN, RNN, DBN, LSTM, and DCNN+CG-DU, respectively. The specificity of the presented work for 1st dataset is 0.95638, which is superior to extant classifiers. The negative metrics of the presented hybrid classifier +CG-DU scheme are minimal than the traditional classifiers. Thus, the proposed model has attained the better performance than the extant classifiers.

8. CONCLUSION

This paper has implemented a novel predictive scheme for AD by means of MRI image. Feature Extraction, Optimal Feature Selection, and Classification are all aspects of the developed model. At first, the GLCM, Haralick features and geometric Haralick features such as geometric correlation and variances were taken out. Particularly, this work carried out optimal feature selection using GOA model. Then, the optimally chosen features were subjected for classification by Hybrid Classifier that includes Optimized CNN and SVM, where the activation function and the weight of CNN were optimally selected through a CG-DU scheme. The final output was obtained from the average of both CNN and SVM outcomes. Finally, the implemented scheme's performance compared to an existing approach using quantitative measurements. When looking at the graph, the adopted hybrid Classifier + CG-DU scheme has attained minimal cost of ~ 0.003 for data 1 than other traditional hybrid Classifier + GOA (~ 1.14), hybrid Classifier + DHOA (~ 1.125), hybrid Classifier + LA (~ 1.11), and hybrid Classifier + SMO (~ 1.17) models at 20th iteration. In addition, the adopted scheme has attained higher MCC value (~ 0.99), which was superior to traditional models. The specificity of the presented work for 1st dataset was 0.95638, which is superior to existing classifiers.This research work is utilitarian in the medical field concerning AD disease early and effective detection through MRI images.

Nomenclature	
Abbreviation	**Description**
MRI	Magnetic Resonance Imaging
AD	Alzheimer's Disease
EEG	Electroencephalography
APOE	Apolipoprotein E
PET	Positron Emission Tomography
OMOP	Observational Medical Outcomes Partnership
ML	Machine Learning
CNN	Convolutional Neural Network
EMR	Electronic Medical Records
GLCM	Gray Level Co-Occurrence Matrix
FDG-PET	Fluoro-DeoxyGlucose PET imaging
ADNI	Alzheimer's Disease Neuroimaging Initiative
CG-DU	Grey Wolf and Dragon Updating Model
CAFFE	Convolutional Architecture for Fast Feature Embedding
MCI	Mild Cognitive Impairment
SVM	Support Vector Machine
FNR	False Negative Rate
MCC	Matthews correlation coefficient
DL	Deep Learning
RNN	Recurrent Neural Network
SCRP	Selected Clinically Relevant Positive
FDR	False Detection Rate
SVRM	Support Vector Regression Method
CSF	Cerebrospinal Fluid
FPR	False Positive Rate

References

[1] Branimir LjubicShoumik RoychoudhuryZoran Obradovic, "Influence of medical domain knowledge on deep learning for Alzheimer's disease prediction", Computer Methods and Programs in Biomedicine20 September 2020Volume 197 (Cover date: December 2020)Article 105765

[2] Chihyun ParkJihwan HaSanghyun Park, "Prediction of Alzheimer's disease based on deep neural network by integrating gene expression and DNA methylation dataset", Expert Systems with Applications15 August 2019Volume 140 (Cover date: February 2020)Article 112873

[3] Baiying LeiMengya YangShuqiang Wang, "Deep and joint learning of longitudinal data for Alzheimer's disease prediction", Pattern Recognition28 January 2020Volume 102 (Cover date: June 2020)Article 107247

[4] Solale TabarestaniMaryamossadat AghiliMalek Adjouadi, "A distributed multitask multimodal approach for the prediction of Alzheimer's disease in a longitudinal study", NeuroImage1 November 2019Volume 206 (Cover date: 1 February 2020)Article 116317

[5] Zhiguang YangZhaoyu Liu, "The risk prediction of Alzheimer's disease based on the deep learning model of brain 18F-FDG positron emission tomography", Saudi Journal of Biological Sciences12 December 2019Volume 27, Issue 2 (Cover date: February 2020)Pages 659-665

[6] Calvin, C.M., de Boer, C., Raymont, V. et al. Prediction of Alzheimer's disease biomarker status defined by the 'ATN framework' among cognitively healthy individuals: results from the EPAD longitudinal cohort study. Alz Res Therapy 12, 143 (2020). https://doi.org/10.1186/s13195-020-00711-5

[7] Tepmongkol, S., Hemrungrojn, S., Dupont, P. et al. Early prediction of donepezil cognitive response in Alzheimer's disease by brain perfusion single photon emission tomography. Brain Imaging and Behavior 13, 1665–1673 (2019).

https://doi.org/10.1007/s11682-019-00182-9

[8] De Velasco Oriol, J., Vallejo, E.E., Estrada, K. et al. Benchmarking machine learning models for late-onset alzheimer's disease prediction from genomic data. BMC Bioinformatics 20, 709 (2019). https://doi.org/10.1186/s12859-019-3158-x

[9] Gao, N., Tao, LX., Huang, J. et al. Contourlet-based hippocampal magnetic resonance imaging texture features for multivariant classification and prediction of Alzheimer's disease. Metab Brain Dis 33, 1899–1909 (2018). https://doi.org/10.1007/s11011-018-0296-1

[10] Wang, X., Zhen, X., Li, Q. et al. Cognitive Assessment Prediction in Alzheimer's Disease by Multi-Layer Multi-Target Regression. Neuroinform 16, 285–294 (2018). https://doi.org/10.1007/s12021-018-9381-1

[11] Li, Y., Zhang, L., Bozoki, A. et al. Early prediction of Alzheimer's disease using longitudinal volumetric MRI data from ADNI. Health Serv Outcomes Res Method 20, 13–39 (2020). https://doi.org/10.1007/s10742-019-00206-3

[12] Lei, B., Hou, W., Zou, W. et al. Longitudinal score prediction for Alzheimer's disease based on ensemble correntropy and spatial–temporal constraint. Brain Imaging and Behavior 13, 126–137 (2019). https://doi.org/10.1007/s11682-018-9834-z

[13] Liu, X., Cao, P., Wang, J. et al. Fused Group Lasso Regularized Multi-Task Feature Learning and Its Application to the Cognitive Performance Prediction of Alzheimer's Disease. Neuroinform 17, 271–294 (2019). https://doi.org/10.1007/s12021-018-9398-5

[14] Shigemizu, D., Mori, T., Akiyama, S. et al. Identification of potential blood biomarkers for early diagnosis of Alzheimer's disease through RNA sequencing analysis. Alz Res Therapy 12, 87 (2020). https://doi.org/10.1186/s13195-020-00654-x

[15] P. Wang, Y. Liu and D. Shen, "Flexible Locally Weighted Penalized Regression With Applications on Prediction of Alzheimer's Disease Neuroimaging Initiative's Clinical Scores," in IEEE Transactions on Medical Imaging, vol. 38, no. 6, pp. 1398-1408, June 2019, doi: 10.1109/TMI.2018.2884943.

[16] G Stalin Babu,S N Tirumala Rao and R. Rajeswara Rao, "Automated Assessment for Alzheimer's Disease Diagnosis from MRI images: Meta-Heuristic Assisted Deep Learning Model ",International Journal of Imaging Systems and Technology. doi:10.1002/ima.22650.

[17] Y. LeCun, K. Kavukvuoglu, and C. Farabet, "Convolutional networks and applications in vision", In Circuits and Systems, International Symposium on, pp.253–256, 2010.

[18] E. Avci,"A new intelligent diagnosis system for the heart valve diseases by using genetic-SVM classifier",Expert Systems with Applications, vol.36,no.7,pp.10618-10626,September 2009.

[19] Mohammad Jafari and Mohammad Hossein Bayati Chaleshtari, " Using dragonfly algorithm for optimization of orthotropic infinite plates with a quasi-triangular cut-out", European Journal of Mechanics A/Solids, vol. 66, pp.1-14, 2017.

[20] Seyedali Mirjalili a, Seyed Mohammad Mirjalili, Andrew Lewis, "Grey Wolf Optimizer", Advances in Engineering Software, vol. 69, pp.46–61, 2014.

[21] M. Marsaline Beno, Valarmathi I. R, Swamy S. M and B. R. Rajakumar, "Threshold prediction for segmenting tumour from brain MRI scans", International Journal of Imaging Systems and Technology, Vol. 24, No. 2, pp. 129-137, 2014, DOI: https://doi.org/10.1002/ima.22087.

[22] Dhanashree Gadkari, "Image Quality Analysis using GLCM", 2004.

[23] S. Ohmshankar and C. K. Charlie Paul, "Haralick fetaures based mammogram classification system," Second International Conference on Current Trends In Engineering and Technology - ICCTET 2014, Coimbatore, pp. 409-413, 2014, doi: 10.1109/ICCTET.2014.6966327.

[24] H.Z. Wang, G.B. Wang, G.Q. Li, J.C. Peng, and Y.T. Liu, " Deep belief network based deterministic and probabilistic wind speed forecasting approach", Applied Energy, vol. 182, pp. 80–93, 2016.

[25] Xiaolong Zhou, Jianing Lin, Zhuo Zhang, Zhanpeng Shao, Honghai Liu, "Improved itracker combined with bidirectional long short-term memory for 3D gaze estimation using appearance cues", Neuro computing In press, corrected proof, Available online 20 October 2019.

[26] Ling-Jing Kao, Chih Chou Chiu, "Application of integrated recurrent neural network with multivariate adaptive regression splines on SPC-EPC process",Journal of Manufacturing Systems, vol. 57, pp. 109–118, 2020.

[27] Sharma, Harish, Garima Hazrati, and Jagdish Chand Bansal. "Spider monkey optimization algorithm." In Evolutionary and swarm intelligence algorithms, pp. 43-59. Springer, Cham, 2019.

[28] Shahrzad Saremi, Seyedali Mirjalili and Andrew Lewis, "Grasshopper Optimisation Algorithm: Theory and application", Advances in Engineering Software, vol. 105 , pp. 30–47, 2017.

[29] Rajakumar B R, "Optimization using lion algorithm: a biological inspiration from lion's social behavior", Evolutionary Intelligence, Special Issue on Nature inspired algorithms for high performance computing in computer vision, vol. 11, no. 1-2, pp. 31-52, 2018, DOI: https://doi.org/10.1007/s12065-018-0168-y.

[30] G Brammya, S Praveena, N S Ninu Preetha, R Ramya, B R Rajakumar, D Binu, Deer Hunting Optimization Algorithm: A New Nature-Inspired Meta-heuristic Paradigm, The Computer Journal.2019.

G Stalin Babu is Research Scholar at JNTUK Kakinada in Department of CSE and working as Assistant Professor in Aditya institute of technology and management , His area of research are image processing, Deep learning. Babu completed his M.Tech in CompterScience from JNTU College of enginerring, Hyderabad in year 2010.He published more than 15 journals in International Journals and conferences.

Dr S N Tirumala Rao is working as Professor and Head of the Department of CSE, NarasaraoPeta Engineering College, Narasaraopet, Guntur(AP)-India. His Areas of research are Data Mining, Multi-Core and Parallel Programming and Computer Networks. Rao has valuable professional experience spread across two successful decades. Out of the 21 years, he has spent 13 years in the teaching profession while the other 8 years include experience and understanding of the research industry. He Published more than 30 National and International Journals and organized several National and International Conferences. Books Published: Advanced Unix Programming HI-Tech Publisher.

Dr. R. Rajeswara Rao is presently working as Professor in the Department of Computer Science and Engineering (CSE), JNTU-GV-UCEV, Vizianagaram. Dr. Rao completed his B.Tech in CSE in the year 1999 from V.R.Siddartha Engineering College, Vijayawada. M.Tech in CSE from JNTUH-Hyderabad in the year 2003, PhD in CSE from JNTUH-Hyderabad in the year 2010. He did Post-Doctoral Research from University of Missouri (UOM), Columbia, USA. Dr. Rao is having 19 + Years of Teaching Experience in the field of Computer Science and Engineering. He is a Member of CSI and Sr. Member of IEEE. His Areas of Interest are Artificial Intelligence, Speech Processing, Pattern Recognition, NLP and Cloud Computing.