Computers in Biology and Medicine 177 (2024) 108622

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

CCADD: An online webserver for Alzheimer's disease detection from brain MRI

Priyanka Panigrahi [a,b], Subhrangshu Das [a,**], Saikat Chakrabarti [a,b,*]

[a] Structural Biology and Bioinformatics Division, Council for Scientific and Industrial Research (CSIR) - Indian Institute of Chemical Biology (IICB), TRUE Campus, Kolkata, 700091, West Bengal, India
[b] Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, Uttar Pradesh, India

A B S T R A C T

Alzheimer's disease (AD) imposes a growing burden on public health due to its impact on memory, cognition, behavior, and social skills. Early detection using non-invasive brain magnetic resonance images (MRI) is vital for disease management. We introduce CCADD (Corpus Callosum-based Alzheimer's Disease Detection), a user-friendly webserver that automatically identifies and segments the corpus callosum (CC) region from brain MRI slices. Extracted shape and size-based features of CC are fed into Support Vector Machines (SVM), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN) classifiers to predict AD or Mild Cognitive Impairment (MCI). Exhaustive benchmarking on ADNI data reveals high prediction accuracies for different AD severity levels. CCADD empowers clinicians and researchers for AD detection. This server is available at: http://www.hpppi.iicb.res.in/add.

1. Introduction

The intricate complexities of the human brain encompass its multi-faceted functions, from movement and perception to thought, speech, learning, and emotion. Within this realm, dementia, including Alzheimer's disease (AD), spans a range of conditions disrupting cognitive and behavioral aspects. These disturbances encompass memory, attention, language, and higher cortical functions, eroding patients' ability to engage with their surroundings and perform daily tasks. Despite its identification over a century ago, substantial headway in effectively treating this progressive disorder has remained elusive [1,2].

Mild cognitive impairment (MCI) represents a transitional phase between normal aging and dementia, elevating the risk of AD and other dementias [3–7]. Identifying MCI as an AD precursor has spurred hopes of preventing or altering neurodegeneration in its early stages [8]. Considering cognitive impairment as a consequence of neurodegeneration, tracking structural brain changes, which manifest ahead of molecular and behavioral symptoms, could offer insights into MCI's pathophysiology [9]. While initial studies featured a cross-sectional approach, they suggested that although structural distinctions in specific brain regions aren't exclusive to MCI or AD (also present in "normal" aging), the pattern of regional atrophy rates and the progression of atrophy exhibit distinct characteristics, especially in AD [10–19]. Furthermore, these studies unveiled differing regional atrophy rates in MCI and AD.

Magnetic resonance imaging (MRI) for assessing cerebral structure has become pervasive in recent decades. Structural magnetic resonance imaging (sMRI) unveils anatomical brain morphology anomalies within grey and white matter. Notably, the spatial pattern of grey matter atrophy, revealed by T1-weighted MRI, closely mirrors the distribution of specific pathology tied to diverse dementia syndromes [20]. For instance, in AD, atrophy corresponds to the distribution of neurofibrillary tangles [21]. Moreover, the extents and pace of volume reduction, particularly in white matter, correlate strongly with cognitive deficits [22].

In this context, identifying region-specific atrophy rates in MCI holds promise for detecting early AD development and gauging expected structural changes in clinical trials. Longitudinal studies have pinpointed specific brain areas linked to MCI pathology, such as cingulate gyri, caudate nucleus, hippocampus [23,24], entorhinal cortex, and frontotemporal lobe, often serving as markers for AD [16–19,25–28]. A study addressed the challenging classification of the early mild cognitive impairment group from the cognitively normal group by considering various essential features such as left hemisphere (LH) lateral ventricle volume, LH precuneus volume, LH superior parietal volume, LH precentral thickness, right hemisphere (RH) lingual volume, RH lateral ventricle volume, LH lingual volume by using a Gaussian-based model [29]. Volumetric changes at different brain regions can be analyzed with

* Corresponding author. CSIR-IICB, TRUE Campus, CN-6, Sector V, Salt Lake, Kolkata, WB, 700091, India.
** Corresponding author.
E-mail addresses: subhrangshu73das@gmail.com (S. Das), saikat@iicb.res.in, saikat273@gmail.com (S. Chakrabarti).

https://doi.org/10.1016/j.compbiomed.2024.108622
Received 11 October 2023; Received in revised form 26 February 2024; Accepted 15 May 2024
Available online 15 May 2024

the progression stages of MCI. It is seen that the volumes of some regions decrease, and some of them increase with the progression of the disease [30]. Texture-based features of brain regions and voxel-based parameters also play a role in AD detection [31,32]. supervised Gaussian discriminative component analysis (GDCA) algorithmwas also implemented to delineate subtle changes in early mild cognitive impairment (EMCI) patients [33].

Corpus callosum (CC) region has also gained attention for AD detection [34–39]. Corpus callosum is the mediator between two cerebral hemispheres which are very much involved in cognitive function and movement control. Researchers have found that thick bundle of nerve fibers of CC are mostly made of large pyramidal neurons in layers III and V of association neocortex, which form a subset of the intracortical projecting pyramidal neurons that might be early and specifically affected by Alzheimer's disease (AD) pathological characteristics [18]. Recent work by our group indicated that CC atrophy significantly aids in detecting AD, particularly its milder forms [40]. Employing MATLAB-based image processing and support vector machine (SVM) classification models, CC atrophy-based features achieved approximately 90 % sensitivity and 80 % specificity in distinguishing healthy individuals from those with severe/mild/moderate AD.

Various studies have used multimodal images such as magnetic resonance imaging (MRI), positron emission tomography (PET), resting-state functional MRI (rs-fMRI), diffusion tensor imaging (DTI) data and demographic information to improve early Alzheimer's disease [41–43]. In a recent study, authors built a multidirectional perception generative adversarial network (MP-GAN) to visualize morphological features for whole-brain MR images in AD. It can generate realistic data without explicitly modeling the probability density function. It can identify the complex lesions that may not be found within one brain region [41]. In another study, authors designed a generator with an interactive hyperedge neurons module to better capture complex relationships between modalities. An Optimal Hypergraph Homomorphism algorithm was also used to construct robust hypergraph structures. Experimental results on ADNI data show the proposed method can effectively extract discriminative features and improve classification performance for AD diagnosis compared to single modalities [42]. Another group built a framework called Multimodal Representation Learning and Adversarial Hypergraph Fusion (MRL-AHF) to improve early Alzheimer's disease prediction using multimodal neuroimaging data including fMRI, DTI and MRI and provided a possible way to understand the underlying mechanisms of disorder's progression by analyzing the abnormal brain connections [43].

Numerous studies have showcased the utility of MRI data and machine learning for effective AD detection. Despite their availability in software repositories, many of these methods demand advanced computational expertise to navigate intricate dependencies. Notably, no user-friendly online platform exists for catering both general and expert users who seek to assess AD probability for a given brain MRI. Addressing this gap, we developed CCADD (Corpus Callosum-based Alzheimer's Disease Detection), an all-encompassing platform that integrates in-house Python-based image analysis algorithms and machine learning protocols for AD detection from uploaded brain MRI slices.

CCADD server accepts raw and processed MRI data as input and facilitates automatic pre-processing and segmentation of sagittal slices to identify CC regions, followed by feature extraction for classification against pre-optimized trained models derived from large number of AD patients and healthy samples. The package includes multiple state-of-the-art machine learning protocols, such as SVM, Random Forest, and eXtreme Gradient Boosting, KNN, and ANN, respectively. Rigorous benchmarking analyses involving substantial patient and control data affirm the models' remarkable accuracy in predicting varying AD categories. This freely accessible software holds considerable value for the scientific, medical, and general communities, enabling validation of radiologist-based detection and identification of structural changes for expert validation and clinical insights.

2. Materials and methods

2.1. Data collection

Structural MR images from sagittal view were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://www.loni.ucla.edu/ADNI). 193 samples from the healthy cohort and 740samples from the AD cohort, including mild, moderate, and severe patients categorized according to Clinical Dementia Rating (CDR) scoring function definition, were collected, while 207, 447,and 279 samples for healthy, AD, and MCI cohorts according to the ADNI clinical scoring scheme were also collected. CDR score is calculated based on testing six different cognitive and behavioral domains such as memory, orientation, judgment and problem-solving, community affairs, home and hobbies performance, and personal care. ADNI grouping is done based on a defined clinical protocol under the supervision of Ronald Petersen.

2.2. MRI pre-processing and processing

Raw MRI data must be processed before it can be used in the CCADD server. Hence, a separate module is provided within the CCADD server where pre-processing of the raw 3D MRI data can be performed to extract sagittal view slices and subsequently passed onto the next module, where the corpus callosum (CC) region is detected automatically. A 3D MRI raw data can be a stack of DICOM (.dcm) image files or a single NIFTI (.nii) file.3D MR images (NIFTI format) are converted into 2D slices, and 2D images (DICOM format) are converted into JPEG format. Images from the sagittal view were used for further processing to keep all the images in a similar orientation.

2.2.1. Automatic detection of corpus callosum region from sagittal slices

An algorithm was developed for automatic detection and subsequent extraction of the corpus callosum region from the sagittal view images. The algorithm uses a set of morphological features to determine whether a segmented region is an accurate segmentation of the CC. The primary concept of the algorithm is to use pattern of the cumulative change in angle across the periphery of the segmented region as the main decision making feature. As shown in Fig. 1A, the algorithm starts at a point on the periphery of the segmented region and traverses across it to complete a closed walk, keeping track of the change in angle at each step. Now, a step from the current pixel to its next one can be possible in 8 different ways (Fig. 1B). For example, in Fig. 1A–if the pivot variable traverses from pixel 1 to pixel 9, the sequence of the change in angle will be 0°,-90°,0°,-45°,-90°,-135°,180°,and 90°; hence the sequence of the cumulative change in angle will be 0°,-90°,-90°,-135°,-225°,-360°,-180°,and −90°. During the traversal across the periphery, the pivot variable may encounters two types of path, one through horizontal/vertical steps (four-neighborhood points) and another through diagonal steps (eight-neighborhood points). First priority is given to top, bottom, left and right neighbors using horizontal/vertical steps while diagonal steps to traverse the bottom-left and bottom-right neighbors is given second priority. As it is closed walk across the periphery, no two untraversed pixel in the neighborhood will come from same priority category. For example, while the pivot variable holds its position at pixel 2 it will prioritize pixel 3 (bottom) over pixel 4 (bottom-right) as its next step. If the curve representing the cumulative change in angle across the periphery has only one local maxima point and exactly two local minima points, with the maxima point occurring between the two minima points, then the segmented region is considered to be a likely segmentation of the CC. Hence the no. of minima to no. of maxima ratio will always be 2 for CC like structure where the above ratio is mostly less than 1 (Fig. 1D) for non-CC structures, as no. of minima is either 0 or more than 6for those structures. This feature is based on the observation that the CC has a characteristic shape that is characterized by a single protrusion (the genu) and two indentations (the rostrum and splenium).

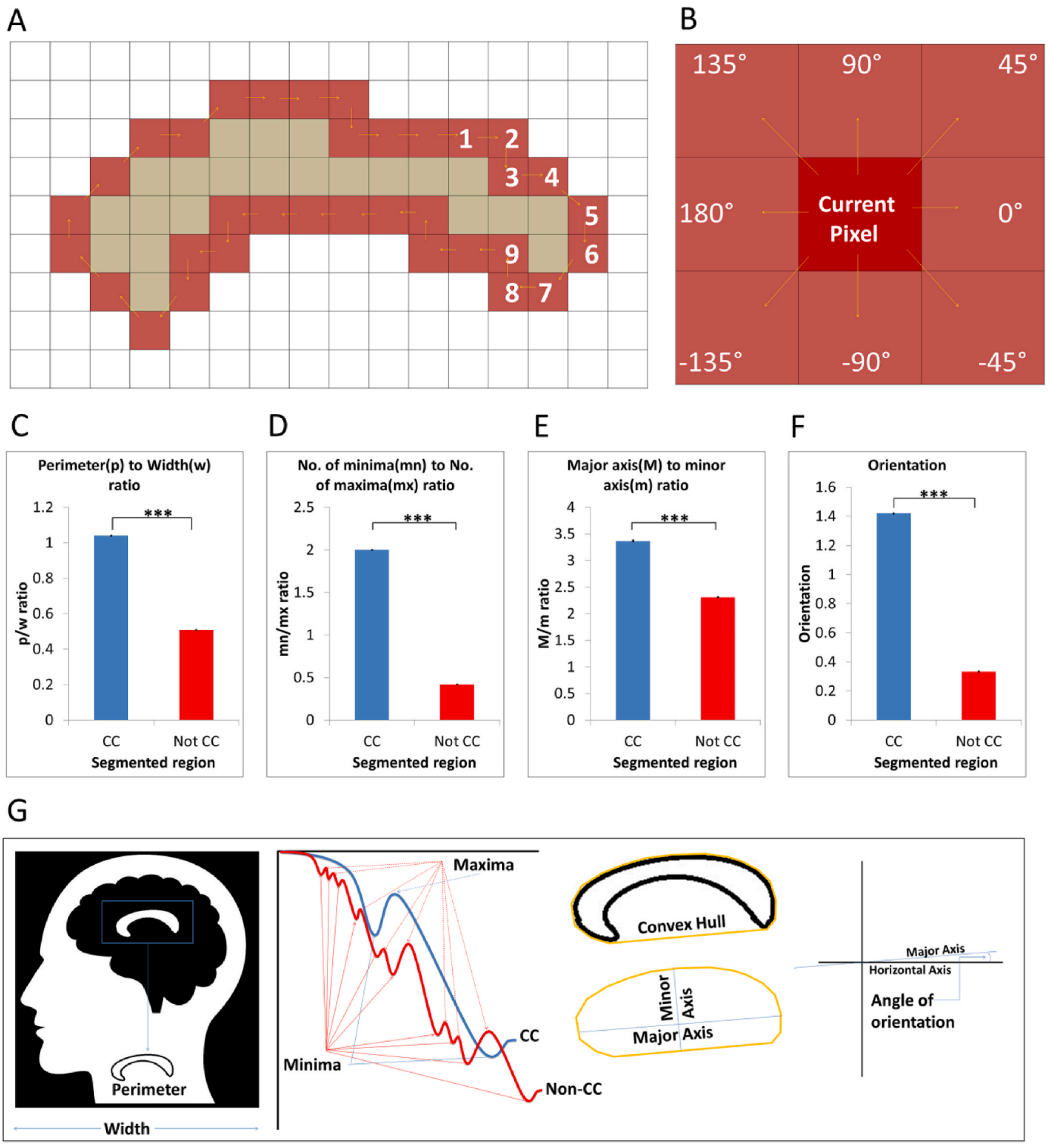

Fig. 1. Concept of the CC detection algorithm. Panel A shows the closed walk across a segmented region's periphery (dark pixels). Panel B shows all possible directions with the current pixel's assigned angle (in degrees). Panel C shows the mean ± standard deviation of perimeter to width ratio between segmented CC and non-CC structures. Panel D shows the mean ± standard deviation of no. of minima to no. of maxima ratio between segmented CC and non-CC structures. Panel E shows the mean ± standard deviation of the major axis to the minor axis ratio between segmented CC and non-CC structures. Panel F shows the mean ± standard deviation of orientation of the encapsulating convex hull between segmented CC and non-CC structures. Panel G shows the cartoon representation of the whole MR image, image width, segmented CC region, and its perimeter. It also shows the minima and maxima in case of correctly predicted CC and non-cc structure, convex hull, major axis, minor axis, and angle of orientation for better understanding.

The cumulative change in angle across the periphery of the CC will, therefore, have one local maxima peak and two local minima through, corresponding to the genu, rostrum, and splenium, respectively.

In addition to the cumulative change in the angle feature, the algorithm also checks that the segmented region satisfying the following morphological criteria.

- The number of pixels in the periphery of the smoothed CC must be between 50 pixels more and less than the width of the image. As the input images are of 256x256 or 192x192 dimension, it has been observed that length of the periphery occurs between 256 ± 50 and 192 ± 50 respectively for an accurately segmented CC. Hence the perimeter to with ratio will fall in the range of $\frac{256-50}{256}$ and $\frac{256+50}{256}$ or $\frac{192-50}{192}$ and $\frac{192-50}{192}$, i.e., 0.8 and 1.2 or 0.74 and 1.26 (Fig. 1C).
- The length of the major axis of the convex hull of the CC is mainly more than two times the length of the minor axis, whereas the ratio of major to minor axis is less than 2 in most of the cases where the segmented region is not a corpus callosum (Fig. 1E).
- The orientation of the convex hull of the CC is mainly around 1.25 whereas that of Non-CC structures varies from −0.5 to 1.25 with a mean of 0.5 (Fig. 1F).

The convex hull of a set of points is the smallest convex polygon containing all the points. The major axis of the convex hull is the longest axis of the polygon, and the minor axis is the shortest axis. The orientation of the convex hull is the angle between the major axis and the horizontal axis (Fig. 1G).

If the segmented region satisfies all of these criteria, then it is considered to be a correct segmentation of the CC.

2.3. Feature extraction

Segmented and extracted CC regions from the AD and healthy cohorts were used for extracting features representing the shape and size of the CC for each sagittal slice. Several features describing the area, length, coordinates, geometric shape, and intensity of the CC regions were calculated using in-house Python codes where *regionprops* function of *scikit-image* module extracts different structural properties of a segmented region. The details of the feature extraction method were thoroughly explained in one of our previous reports [40] on CC based AD diagnosis.

2.4. Machine learning protocols

Various supervised machine learning (e.g., SVM, RF, XGBoost, and KNN) and deep learning (ANN) techniques are implemented to build and test the training models in distinguishing AD patients' CC from healthy ones. The following section briefly describes the four machine learning (ML) classification protocols and one deep learning (DL) learning based approach implemented within the CCADD server.

2.4.1. Support vector machine (SVM)

Support Vector Machine (SVM) is a supervised binary classification algorithm that takes labeled data as input from two different classes and outputs a model/classifier, which can distinguish/classify between the classes when given new labeled data. *Scikit-learn* Python package was used, and the regularization parameter c value was set at 1000, while the radial basis function (RBF) kernel was used for classification. The regularization parameter c was determined through a grid search, ranges from 1000 to 50000, by a parameter selection tool named *grid. py*, provided by *LIBSVM*, a well-known library for SVM. It generates a contour of cross-validation accuracy using *Gnuplot* to understand the optimal value of the regularization parameter c for the given dataset.

2.4.2. Random Forest (RF)

Random Forest is a supervised machine learning-based algorithm that unites the output of multiple decision trees and results into a single decision and improves the model's performance. To avoid overfitting, RF utilizes an ensemble of diverse decision trees. It applies the majority voting algorithm, which makes predictions based on each decision tree model. We have used the Random Forest Classifier module of *scikit-learn* Python package to classify between demented and healthy cohorts. *Random Forest Classifier* with the default parameter of *n-estimators* and default number of decision trees were used to build our models. For the RF classification, we have used *scikit-learn 0.24.2* where the default tree quantity parameter was set at 100.

2.4.3. XGBoost

Extreme Gradient Boosting is a supervised machine learning algorithm. It implements the gradient-boosted decision trees for maximum speed and performance. It provides a parallel tree boosting that solves problems in a fast and accurate way. XGBoost provides a larger number of hyperparameters which determines the learning process of the algorithm. We tuned few hyperparameters (*learning rate, gamma, n_estimators, max_depth, min_child_weight, subsample, colsample_bytree, reg_lambda, reg_alpha*) using *RandomizedSearchCV* of *scikit-learn* module. The best of 50 randomly taken hyperparameter combinations were chosen based on accuracy in different categories. However, the performance of the models with default parameters (without tuning) using the exact greedy algorithm to construct the trees was notably better than best tuned hyperparameter combinations [Supplementary Table S1]. Hence, we set default values of various parameters to build our models. We have used *gbtree* booster, which uses tree-based models. We took the default *gamma* value of 0. The tree booster parameters like *gamma, regalpha, reglambda, colsample_bytree, colsample_bylevel,* and*colsample_bynode* were set at default values. We have used 100 trees for our model building. The maximum depth of a tree was set to 6.

2.4.4. K-Nearest Neighbors (KNN)

k-Nearest Neighbors classifier is a learning algorithm. It is a non-parametric algorithm used for classification and regression tasks. We have used the KNN classifier module of *scikit-learn* Python package to classify between demented and healthy cohorts. For this classification we set the *n neighbors* to 3.

2.4.5. Artificial Neural Network (ANN)

Artificial neural networks are brain-inspired systems that are intended to replicate the way that humans learn. Neural networks consist of input and output layers and a hidden layer consisting of units that transform the input into something that the output layer can use. ANNs have three interconnected layers. We built a sequential model for our classification task. We have used *relu* and *sigmoid* activation functions and *adam* optimizer. Here, we considered 500 epochs with batch size 10.

Features related to the shape and sizes of the CC were utilized for classification via a 100-fold cross-validation protocol. We have used a 100-fold cross-validation method where 100 times randomly selected, 80 % of the data has been used to train the model, and the rest 20 % has been used as validation data. In both the train and test datasets, the ratio of healthy and demented samples ratio was the same, and 100 models were generated. Further, we validate the 100 models on the validation dataset. These training models can classify any given query CC features and predict the likelihood of it belonging to AD or a healthy population. Separate training models considering the severe, mild, moderate, and all demented patients to normal/healthy samples were constructed using both CDR and ADNI-based definitions of the disease severity. Users of the CCADD server can test the AD likelihood of an MRI sample/data by using these models individually as well as all together. The best-performing 25 training models from each AD category are kept for prediction.

We estimated the learning rate of test and train data using model

accuracy with varying train data sizes. The learning rates increase with train data size. We have plotted the mean with the standard deviation of the model accuracy for train and test data in Supplementary Fig. S1.

CCADD server accepts query MRI data as input and, at the backend, extracts the CC shape and size-based features from this input file. Further, the features of each query CC region were classified against the previously described trained models for the prediction of whether the particular CC image is AD or not.

2.5. Webserver options

CCADD (Corpus Callosum-based Alzheimer's Disease Detection) webserver has been developed on PHP and CGI-PERL frontend platforms. At the backend, the classification and prediction schemes are encoded using Python codes. Given brain MRI data, the CCADD server attempts to predict whether the corpus callosum (CC) atrophy (if there is any) estimated from the query patient's MRI significantly resembles the atrophy derived from the brain MRI of large numbers of clinically determined AD patients. Users can test the resemblance of the CC features derived from mild, moderate, and severe AD patients. The server is available at http://www.hpppi.iicb.res.in/add.

2.5.1. Input option

As mentioned earlier, raw input data can be pre-processed using a separate module under the 'PRE-PROCESSING' tab, where CC slices can be extracted from the 3D MRI data that comes as a standard output of clinically used MRI machines. The output option is given under the 'SERVER' tab of the CCADD website.

Similarly, processed and extracted sagittal slices of brain MRI can be used directly for AD prediction. Classification methods, reference prediction models, and clinical scoring methods need to be selected for the prediction of the AD status of the query slice. Users can select single or multiple classification and combined methods simultaneously to predict the AD status of their query brain MRI slices.

2.5.2. Output option

CCADD server provides an output page where average probabilities of both AD/Dementia and healthy for each query sagittal slice for which the CC region was detected are given. CC shape and size-based feature matrix for each query slice are compared against the best 25 training models generated and evaluated beforehand, and the probabilities (likelihood) of the features to be AD-like and healthy-like are given accordingly. A voting column is also provided where the number of models out of 25 possessing higher AD probability for the particular query slice is mentioned. Based on the average AD/healthy probability and the voting count, an overall prediction outcome (demented/healthy) is mentioned for each query slice. Final AD status prediction for the patient/sample is provided based on the voting outcome of the number of slices for which CC was detected, and CC-based features were matched against trained models.

2.6. Benchmarking

Train-Test: Benchmarking of the CCADD server was done using MRI data downloaded from the ADNI database. According to CDR classification, we have built our training models based on different categories, i.e., mild vs. healthy, moderate vs. healthy, severe vs. healthy, and all demented vs. healthy. For mild vs. healthy and moderate vs. healthy category, 540 sagittal slices were taken from the demented cohort, and 540 sagittal slices were taken from the healthy cohort. For all demented vs. healthy category, we took 540 slices each from healthy and demented cohorts. Moreover, we have maintained the 1:1:1 image slice ratio among all the three demented groups i.e., 180 image slices from mild, moderate, and severe groups. We took 201 slices from each cohort (healthy and demented) for the severe vs. healthy category. In all the categories, we have maintained the ratio to 1:1 for healthy and demented cohorts in our train-test exercise. At patient-level benchmarking, if 50 % or more of the slices are predicted as demented, the patient is considered to have dementia. For each category of AD (mild/moderate/severe), separate train-test datasets were created, keeping an equal ratio (1:1) of demented and healthy samples/slices. According to ADNI-based classification, we have built models based on two categories, i.e., AD vs. healthy and MCI vs. healthy, for both categories, 605 image slices were used from each healthy and demented cohort to perform the train-test exercise.

Validation: For CDR-based classification, validation datasets were also created for each category of AD (mild/moderate/severe) consisting of 39 mild, 39 moderate, and 13 severe samples and 106 healthy samples, respectively. Similarly, for ADNI-based classification, 74 healthy, 41 AD, and 41 MCI samples were taken. We maintained the healthy and demented sample ratio at 1:1 for each category while validating our classifier's performance. Separate benchmarking was performed for five classification algorithms (e.g., SVM, RF, XGBoost, KNN, ANN). Benchmarking was done to identify AD/healthy slices and/or AD/healthy patients, respectively. Accuracy, sensitivity, specificity, precision, and F1 scores of the models using the separate validation dataset were calculated to evaluate the server's performance. The top 25 models for each category from the validation experiments are kept at the backend of the server for live prediction. Tables 1 and 2 provide the data distribution within the train, test, and validation sets for CDR and ADNI-based disease/healthy categorization.

Performance of the CCADD server was also testified using one MRI sample at a time. Each sample from the validation dataset was submitted to the CCADD server for prediction of the AD or healthy status at the patient level. Same data was used for AD prediction via a software package named BAAD (Brain Anatomical Analysis using Diffeomorphic deformation) [44]. BAAD is based on a VBM (voxel-based morphometry) algorithm to extract features of the brain shape. It provides AD detection based on volumetric atrophy estimation of individual ROIs including corpus callosum as well as all combined ROIs (139 brain regions). To the best of our knowledge, this is the only online resource which could easily be installed locally without any prior technical knowledge and accepts user provided MRI in NIFTI/DICOM format. Receiver operating curves (ROC) were generated by plotting the true positive rate (TPR) and false positive rate calculated at each CCADD and BAAD score threshold. Performances of the CCADD models were compared with that of the BAAD modules using CC and hippocampus based prediction along with all ROI.

3. Results

3.1. Dataset and training models

Notable mismatches were found among clinical definition and scoring patterns for mild and moderate categorization suggested by CDR scoring and MCI grouping by ADNI experts (Supplementary Fig. S2). These observations prompted us to prepare separate training models using CDR and ADNI clinical scoring-based AD categorizations. We tested with different machine learning (ML) and deep learning (DL) techniques, out of which, logistic regression did not perform well [Supplementary Tables S2–S5]. Thus, we have considered five different classification protocols (SVM, RF, XGBoost, KNN and ANN) in the CCADD server, four models from the CDR category (mild vs. healthy, moderate vs. healthy, severe vs. healthy, all demented vs. healthy) while two models (MCI vs. healthy, AD vs. healthy) were generated for prediction of AD status of a given MRI sample for each of the classification protocols. 100 models were generated by randomly mixing 80 percent of the AD and healthy samples from each category to build the training models, while the remaining 20 percent were used for testing the accuracy of the corresponding training model. Query CC features are matched against the 100 train models, and average AD/healthy probabilities are calculated and provided on the results page of the CCADD

Table 1
Data size and distribution (CDR score-based AD grouping).

Category	No. of Slices			Category	Training Models									
					Mild Vs. Healthy			Moderate Vs. Healthy			Severe Vs. Healthy		All Demented Vs. Healthy		
					Available	Randomly selected Train	Randomly selected Test	Available	Randomly selected Train	Randomly selected Test	Available	Randomly selected Train	Randomly selected Test	Available	Randomly selected Train	Randomly selected Test
Healthy	690	**Data**	**Train-Test**	**Healthy**	540	432	108	540	432	108	640	161	40	540	432	108
Mild	1523	**Splitting**		**Mild**	1373	432	108							1473	432	108
				Moderate				852	432	108				952		
				Severe							201	161	40	201		
Moderate	1002		**Validation**	**Healthy**		150			150			50			150	
				Mild		150									50	
Severe	251			**Moderate**					150						50	
				Severe								50			50	

server. The best 25 models were selected from this exercise to be incorporated within the CCADD server.

3.2. Training-testing and model validation

Performance of the models tested both at slice and patient sample level. Average test accuracy, sensitivity, specificity, and F1 score (mean of top 25 model results) were calculated for the respective test sets. Performance of the train models was further testified using the separate validation cohort consisting of equal ratio (1:1) of healthy and demented samples (Tables 1 and 2). Tables 3 and 4 provide the performance results of the CCADD models obtained for the training-test and validation cohorts using the CDR and ADNI-based disease categorizations, respectively. For the validation cohort, we have also tested the performance using a separate train model where 80 percent train and 20 percent test data were combined into one model. Moderate performance improvement was observed using these 100 percent models (Supplementary Tables S6–S7).

In the CDR grouping, for all demented and severe categories from the validation cohort, SVM performs best (in terms of F1 score) in differentiating severe and all demented samples from the healthy ones; however, for mild and moderate groups RF and XGBoost outperform others (Table 3). Similar trends were observed for experiments done against 100 percent models (Supplementary Table S6). Models built based on ADNI clinical scoring suggest better classification performances of RF for the AD group and XGBoost for the MCI group (Table 4 and Supplementary Table S8), respectively.

Overall, our CDR models perform pretty well in distinguishing all demented, severe, moderate, and mild AD samples from healthy ones with average F1 scores of 0.81, 0.88, 0.83, and 0.79, respectively. In addition to the F1 score, other metrics like accuracy, sensitivity, and specificity also indicate a similar trend for our models. ADNI models produce average F1 scores of 0.86 and 0.79 for AD and MCI groupings, respectively.

Further, we have also tested the performance of the machine learning and deep learning-based CCADD models, considering one patient sample at a time. Each sample (AD and healthy) from the CDR and ADNI based validation datasets was submitted to the CCADD server individually using one model at a time. Percentage of correctly predicted AD and healthy samples are provided in Figs. 2 and 3, respectively. RF classifiers using 'All demented' and 'Moderate' models performed better in correctly identifying all types of AD samples (categorized by CDR scores) compared to that achieved by other classifiers. However, SVM 'Mild' and 'Severe' models performed better than others in identifying all AD samples (Fig. 2A and Supplementary Table S8). Similarly, RF AD and MCI models also performed relatively better in identifying all demented samples (AD and MCI categorization by ADNI grouping) (Fig. 3A and Supplementary Table S10). XGBoost models are clearly more accurate in correctly identifying healthy samples categorized by both CDR and ADNI groupings (Figs. 2B and 3B, Supplementary Tables S9 and S11).

Each classifier in principle could be attuned to predict certain types of sample and may fail to identify certain other AD types. Hence, we investigated whether combining the results from multiple classifiers in term of union and/or intersection of their predictions could improve the overall accuracy. The union of the prediction was done by pooling the correctly predicted instances whereas the intersection of prediction was counted if a sample (AD or healthy) is correctly predicted by at least three of the five classifiers. Figs. 2 and 3, Supplementary Tables S8–S11 show the results of union and intersection prediction of the validation datasets. It is not quite surprising that 'union prediction' improves the accuracy coverage significantly. However, the 'intersection prediction' also provides very high prediction accuracy for both healthy and AD samples. 'Intersection prediction' adds reliability to the prediction as the predicted class (AD or healthy) is suggested by multiple (at least three) classifiers. We believe, this combined end-to-end prediction mode would

Table 2
Data size and distribution (ADNI-based AD grouping).

Category	No.of Slices			Category	Training Models					
					AD Vs. Healthy			MCI Vs. Healthy		
					Available	Randomly selected Train	Randomly selected Test	Available	Randomly selected Train	Randomly selected Test
Healthy	756	**Data Splitting**	**Train-Test**	**Healthy**	605	484	121	605	484	121
				AD	1623	484	121			
AD	1774			**MCI**				785	484	121
			Validation	**Healthy**		151			151	
MCI	936			**AD**		151				
				MCI					151	

Table 3
Benchmarking results using CDR score-based AD grouping.

Classifiers	TEST				VALIDATION			
	Accuracy	Sensitivity	Specificity	F1Score	Accuracy	Sensitivity	Specificity	F1Score
	Mild Vs. Healthy				**Mild Vs. Healthy**			
RF	0.81 ± 0.009	0.80 ± 0.064	0.82 ± 0.061	0.80 ± 0.017	**0.82 ± 0.010**	**0.80 ± 0.039**	**0.84 ± 0.045**	**0.82 ± 0.012**
SVM	0.81 ± 0.010	0.78 ± 0.046	0.84 ± 0.042	0.80 ± 0.015	0.80 ± 0.008	0.77 ± 0.046	0.84 ± 0.043	0.80 ± 0.014
XGB	0.81 ± 0.009	0.80 ± 0.059	0.83 ± 0.053	0.81 ± 0.016	0.82 ± 0.012	0.83 ± 0.033	0.81 ± 0.030	0.82 ± 0.014
KNN	0.79 ± 0.008	0.73 ± 0.048	0.85 ± 0.050	0.78 ± 0.013	0.80 ± 0.008	0.74 ± 0.022	0.85 ± 0.022	0.78 ± 0.009
ANN	0.76 ± 0.018	0.73 ± 0.063	0.79 ± 0.054	0.75 ± 0.026	0.75 ± 0.015	0.76 ± 0.038	0.74 ± 0.042	0.75 ± 0.017
	Moderate Vs. Healthy				**Moderate Vs. Healthy**			
RF	0.86 ± 0.010	0.84 ± 0.036	0.89 ± 0.037	0.86 ± 0.012	0.84 ± 0.006	0.82 ± 0.042	0.86 ± 0.042	0.84 ± 0.009
SVM	0.87 ± 0.006	0.86 ± 0.034	0.89 ± 0.035	0.87 ± 0.007	0.85 ± 0.007	0.79 ± 0.040	0.90 ± 0.038	0.84 ± 0.011
XGB	0.87 ± 0.010	0.83 ± 0.050	0.91 ± 0.050	0.86 ± 0.013	**0.86 ± 0.006**	**0.81 ± 0.038**	**0.90 ± 0.033**	**0.85 ± 0.01**
KNN	0.86 ± 0.014	0.82 ± 0.023	0.90 ± 0.022	0.86 ± 0.015	0.84 ± 0.009	0.79 ± 0.020	0.90 ± 0.019	0.83 ± 0.010
ANN	0.84 ± 0.012	0.82 ± 0.052	0.87 ± 0.047	0.84 ± 0.016	0.83 ± 0.009	0.77 ± 0.044	0.88 ± 0.040	0.81 ± 0.014
	Severe Vs. Healthy				**Severe Vs. Healthy**			
RF	0.90 ± 0.015	0.89 ± 0.053	0.91 ± 0.046	0.90 ± 0.017	0.87 ± 0.011	0.86 ± 0.074	0.89 ± 0.069	0.87 ± 0.016
SVM	0.90 ± 0.013	0.89 ± 0.051	0.92 ± 0.047	0.90 ± 0.015	0.90 ± 0.013	0.89 ± 0.025	0.92 ± 0.034	0.90 ± 0.012
XGB	0.90 ± 0.013	0.90 ± 0.042	0.91 ± 0.039	0.90 ± 0.015	**0.88 ± 0.017**	**0.85 ± 0.054**	**0.92 ± 0.049**	**0.88 ± 0.019**
KNN	0.88 ± 0.016	0.86 ± 0.049	0.91 ± 0.047	0.88 ± 0.017	0.88 ± 0.009	0.86 ± 0.030	0.89 ± 0.026	0.87 ± 0.011
ANN	0.87 ± 0.016	0.88 ± 0.047	0.86 ± 0.054	0.87 ± 0.015	0.88 ± 0.013	0.88 ± 0.033	0.88 ± 0.030	0.88 ± 0.014
	All Demented Vs. Healthy				**All Demented Vs. Healthy**			
RF	0.83 ± 0.009	0.80 ± 0.045	0.86 ± 0.043	0.83 ± 0.013	0.82 ± 0.008	0.78 ± 0.034	0.86 ± 0.033	0.81 ± 0.010
SVM	0.84 ± 0.012	0.79 ± 0.051	0.89 ± 0.048	0.83 ± 0.017	**0.84 ± 0.006**	**0.82 ± 0.031**	**0.86 ± 0.029**	**0.84 ± 0.009**
XGB	0.84 ± 0.010	0.82 ± 0.042	0.86 ± 0.042	0.84 ± 0.012	0.82 ± 0.005	0.79 ± 0.038	0.86 ± 0.039	0.82 ± 0.009
KNN	0.82 ± 0.012	0.74 ± 0.045	0.90 ± 0.047	0.80 ± 0.015	0.81 ± 0.008	0.73 ± 0.017	0.88 ± 0.019	0.79 ± 0.009
ANN	0.80 ± 0.008	0.75 ± 0.056	0.84 ± 0.052	0.79 ± 0.016	0.81 ± 0.012	0.77 ± 0.040	0.86 ± 0.039	0.80 ± 0.015

Table 4
Benchmarking results using ADNI-based AD grouping.

Classifiers	TEST				VALIDATION			
	Accuracy	Sensitivity	Specificity	F1Score	Accuracy	Sensitivity	Specificity	F1Score
	AD Vs. Healthy				**AD Vs. Healthy**			
RF	0.88 ± 0.008	0.86 ± 0.037	0.90 ± 0.032	0.88 ± 0.010	**0.88 ± 0.005**	**0.86 ± 0.027**	**0.90 ± 0.028**	**0.88 ± 0.006**
SVM	0.88 ± 0.012	0.85 ± 0.033	0.92 ± 0.037	0.88 ± 0.012	0.87 ± 0.006	0.85 ± 0.029	0.89 ± 0.024	0.87 ± 0.008
XGB	0.88 ± 0.006	0.83 ± 0.035	0.92 ± 0.034	0.87 ± 0.008	0.87 ± 0.006	0.82 ± 0.031	0.92 ± 0.027	0.86 ± 0.009
KNN	0.86 ± 0.012	0.82 ± 0.022	0.90 ± 0.021	0.85 ± 0.013	0.87 ± 0.005	0.81 ± 0.016	0.93 ± 0.016	0.86 ± 0.006
ANN	0.85 ± 0.007	0.83 ± 0.041	0.88 ± 0.039	0.85 ± 0.011	0.86 ± 0.009	0.82 ± 0.031	0.89 ± 0.030	0.85 ± 0.011
	MCI Vs. Healthy				**MCI Vs. Healthy**			
RF	0.81 ± 0.009	0.84 ± 0.041	0.79 ± 0.043	0.82 ± 0.011	0.81 ± 0.005	0.80 ± 0.045	0.83 ± 0.047	0.81 ± 0.009
SVM	0.82 ± 0.009	0.82 ± 0.031	0.82 ± 0.024	0.82 ± 0.012	0.78 ± 0.009	0.81 ± 0.039	0.76 ± 0.037	0.79 ± 0.012
XGB	0.83 ± 0.012	0.84 ± 0.058	0.82 ± 0.063	0.83 ± 0.014	**0.82 ± 0.009**	**0.81 ± 0.037**	**0.84 ± 0.034**	**0.82 ± 0.012**
KNN	0.80 ± 0.011	0.76 ± 0.026	0.83 ± 0.023	0.79 ± 0.013	0.78 ± 0.007	0.75 ± 0.021	0.81 ± 0.015	0.77 ± 0.010
ANN	0.76 ± 0.018	0.76 ± 0.045	0.76 ± 0.049	0.76 ± 0.020	0.75 ± 0.015	0.76 ± 0.038	0.74 ± 0.042	0.75 ± 0.017

be quite useful for the users of CCADD webserver.

3.3. Comparison study

To compare the performance of the CCADD server, we ran the validation dataset using the BAAD software package [44], which supports the clinical diagnosis of Alzheimer's disease (AD) from brain MRI using multiple region of interest (ROI) based atrophy analysis. It presents a possibility of AD for the given MRI sample by a value from 0 to 1 as Alzheimer's score (ADS). An ADS >0.5 suggests the possibility of AD spectrum. Our validation datasets from different AD groupings were supplied to the CCADD server and BAAD package, and their performances were measured in terms of receiver operating curve (ROC) analysis (Figs. 4–5). BAAD software also provides a score (z-score) for individual brain regions based atrophy. We extracted the scores for four sub-regions of the hippocampus and two properties of the corpus

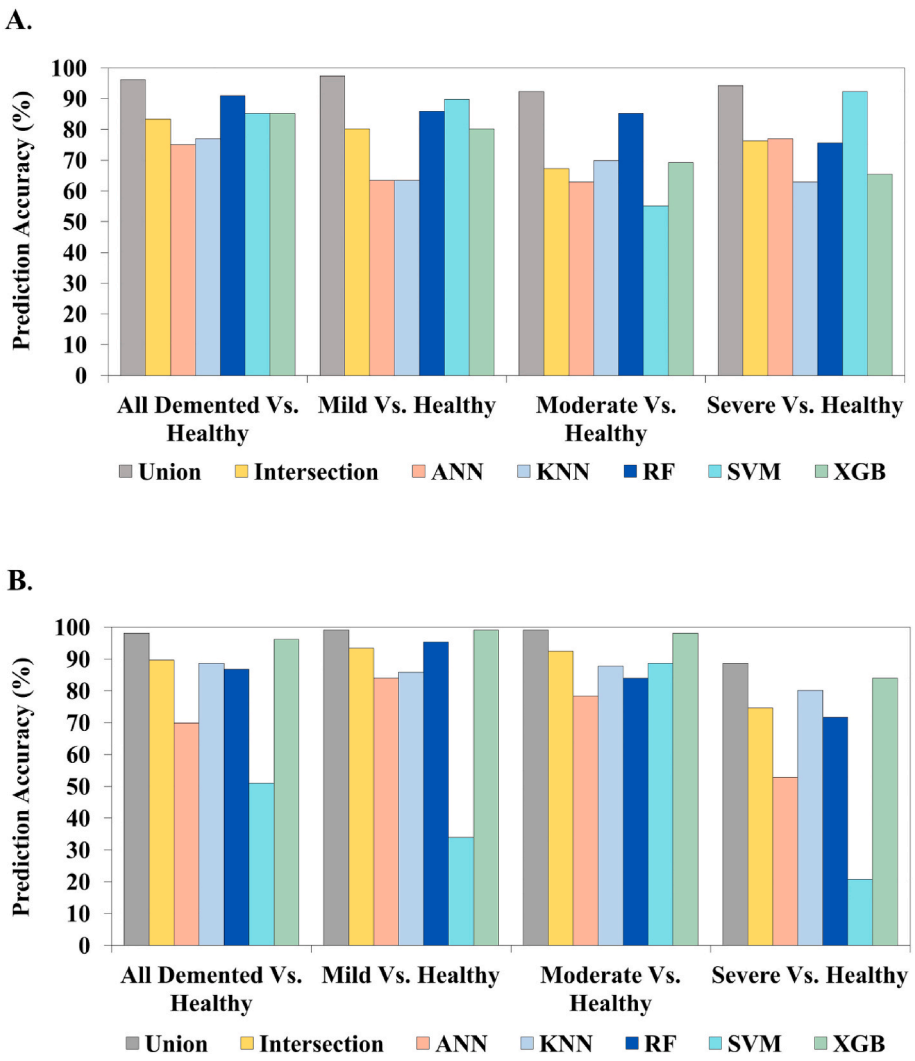

Fig. 2. Performance of the CCADD training models from various classifiers in the identification of healthy and AD samples. Panel A shows the average percentage of accurate (y-axis) identification of all demented samples across various AD categories while panel B plots the average accuracy of identification of all healthy samples. Different categories (mild, moderate, severe, and all demented) of AD samples from the validation set were run against all the models generated by RF, XGBoost, SVM, KNN, and ANN, respectively. Accuracies of the union and intersection approaches were calculated as described before. CDR-based models and validation dataset were used here.

callosum (AHL: Anterior Hippocampus Left; PHL: Posterior Hippocampus Left; AHR: Anterior Hippocampus Right; PHR: Posterior Hippocampus Right; Corpus Callosum Area; and Corpus callosum Volume) and further utilized these scores to generate corresponding ROC plots. Here, we compared the performances of CCADD ML and Dl models with the BADD module that provides AD detection based on volumetric atrophy estimation of various ROIs including 139 brain regions. Similarly, as hippocampal atrophy is one of the widely accepted ROI markers for AD detection, we also compared the CCADD performance with that derived from the atrophy scores of four regions of hippocampus (AHL, PHL, AHR, and PHR, respectively) supplied by the BAAD package. Figs. 4–5, show CCADD server algorithms especially RF, XGBoost, and KNN models clearly outperform the BAAD package for identification of mild/moderate/severe/all demented samples with respect to their corresponding healthy samples. Similar trend is observed for both CDR and ADNI based grouping of the demented and healthy samples. Further, we also compared the performances of our CC shape and size based features and models with that estimated from the CC atrophy features (CC Area; and CC Volume) based scores of the BADD package (Supplementary Figs. S3–S4). Interestingly, we found CCADD outperforms BAAD CC based scores quite remarkably, indicating the strength of our CC features in distinguishing AD and healthy samples.

3.4. Webserver output

CCADD server provides a front end to upload MRI data (sagittal view slices) and subsequently to run the back-endmachine learning based algorithms for prediction of the AD status of the query sample. Users need to select the specific clinical scoring protocol used for generating the training models (e.g., CDR or ADNI) as well as the specific prediction category (e.g., mild/moderate/severe/all demented) for the prediction. Additionally, specific classification algorithm (e.g., SVM/RF/XGBoost/KNN/ANN) and the training model size (e.g., 80 % or 100 %) need to be selected before the run. In classification category we also included a combined algorithm selection option (end-to-end model/ALL).

At the back end raw images are processed and segmented to identify the CC regions automatically. Feature extraction module calculates the CC shape and size-based features. Once the feature matrix for the query sample is created it is submitted to the ML/DL based prediction module using selected training model. The output of the prediction is displayed per slice basis and based on the average AD and/or healthy probability score, each slice is predicted as demented or healthy. Query sample is

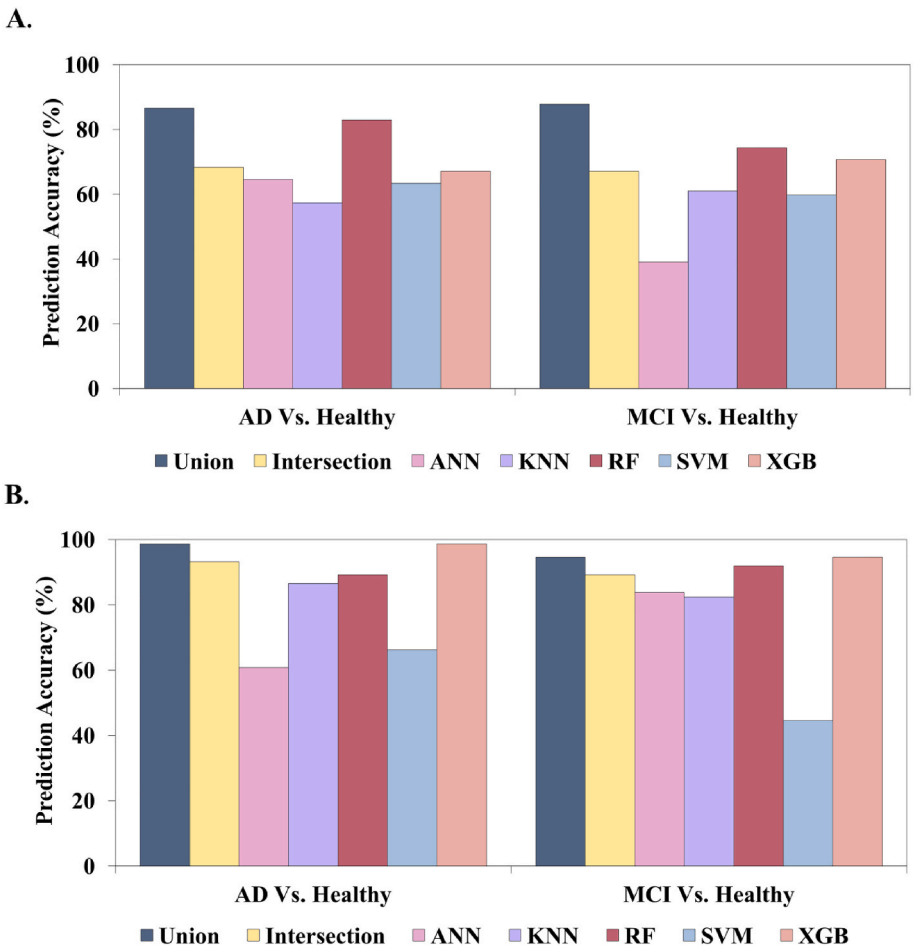

Fig. 3. Performance of the CCADD training models from various classifiers in the identification of healthy and AD samples. Panel A shows the average percentage of accurate (y-axis) identification of all demented samples across various AD categories while panel B plots the average accuracy of identification of all healthy samples. Different categories (mild, moderate, severe, and all demented) of AD samples from the validation set were run against all the models generated by RF, XGBoost, SVM, KNN, and ANN, respectively. Accuracies of the union and intersection approaches were calculated as described before. ADNI-based models and validation datasets were used here.

run against best 25 pre-generated models and average probabilities for AD and healthy are provided in the server output display. For the sample/patient based prediction a voting protocol is implemented based on the majority of CC slices being predicted as demented or healthy. Additionally, individual CC shape and size-based feature values are tabulated for the query sample along with subsequent reference values calculated from all the healthy cohorts of our dataset. Statistical significance of the difference of each feature between healthy average and the sample was calculated and resultant –values are provided in the output display. Finally, 3D-reconstruction of the identified CC regions is performed where each pixel from the CC regions is converted into Cartesian co-ordinate and resultant 3D image of the CC and its volume are displayed using a web-based display window [45]. A representative 3D image and volume of the CC from the healthy population are also displayed for qualitative comparison. Fig. 6 provides a snapshot of the CCADD homepage and excerpts of its output options.

4. Discussion

We present an integrated platform that merges image analysis algorithms and machine learning protocols for Alzheimer's disease (AD) detection from brain MRI slices. Our online package, CCADD, incorporates advanced classification methods, including Support Vector Machines (SVM), Random Forest (RF), eXtreme Gradient Boosting (XGBoost), k-Nearest Neighbor (KNN) and Artificial Neural Network (ANN) catering to both raw and processed MRI data. CCADD automatically preprocesses and segments sagittal slices to identify corpus callosum (CC) regions, obviating manual involvement. Subsequently, the feature extraction module calculates CC shape and size-based features for classification. These features are matched against pre-optimized models derived from AD patient and healthy sample data, yielding predictions regarding the AD status of a given CC image. Rigorous benchmarking, utilizing an extensive dataset from the ADNI database, highlights the platform's efficiency and robustness. Our models achieve remarkable prediction accuracies across varying AD severity categories. CCADD's rapid and precise predictions are invaluable for the general population, scientific researchers, and particularly the medical community. It complements manual radiology observations and reports, bolstering pre- and post-validation efforts.

The combination of structural MRI (sMRI) data, image processing algorithms, and machine learning classification protocols is quite successful in the detection and differentiation of AD samples from healthy ones. Numerous studies over the last couple of decades reported reasonable to very high accuracies for the detection of AD using either ROI-based or whole-brain image analysis approaches. ROI-based approaches aided by traditional machine protocols implementing atrophy and volumetric parameters of specific brain regions are widely used in differentiating AD and MCI cohorts from healthy controls [46–48]. sMRI is very efficient in delineating variations in anatomical structures and morphology of brain tissues. It is a relatively cheaper and safer

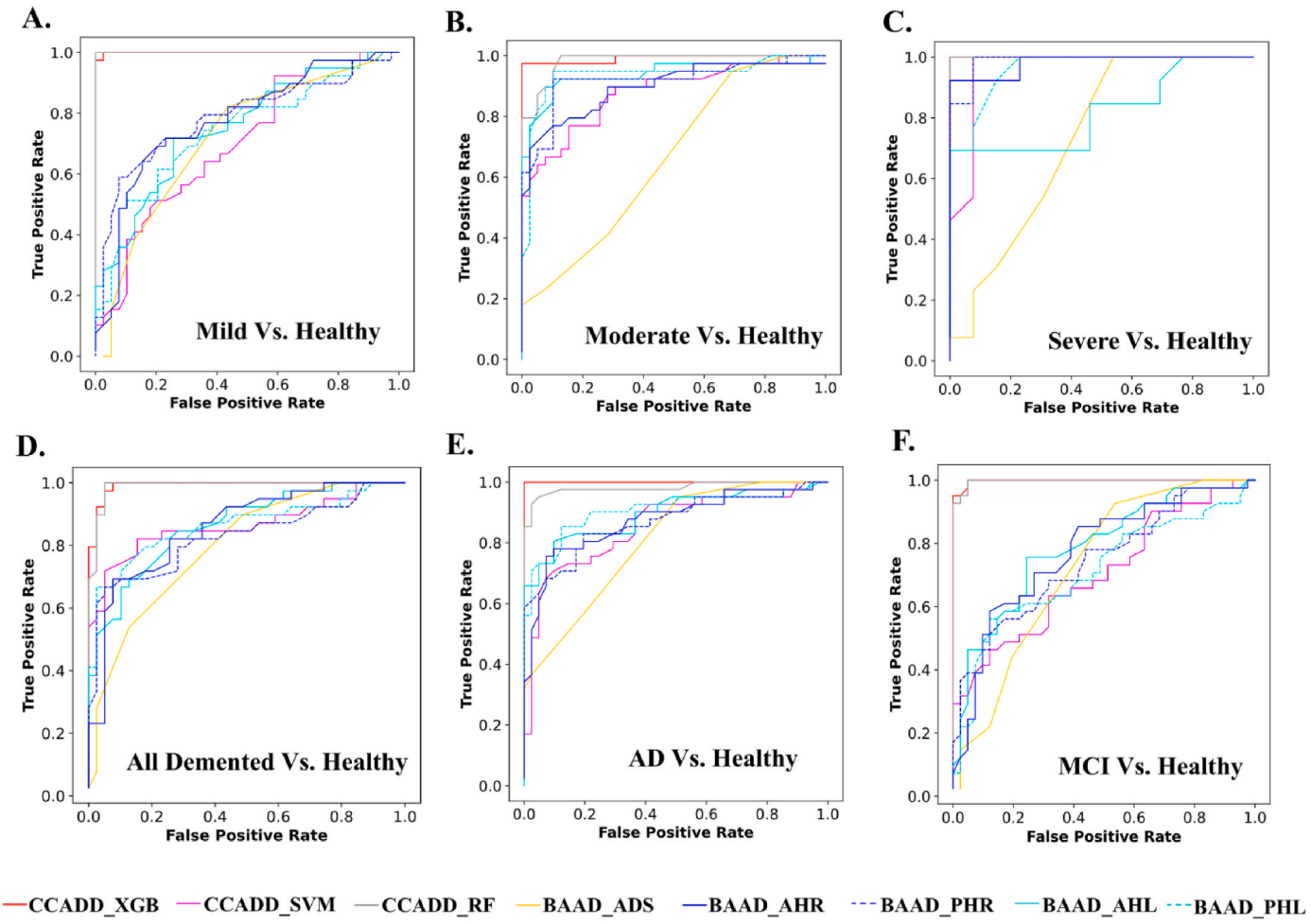

Fig. 4. Comparison of performance of the CCADD server. Receiver operating characteristics (ROC) plots were generated using the validation datasets from various CDR-based (mild, moderate, severe, and all demented) (A–D) and ADNI-based groupings (E–F) against the CDR and ADNI-based training models (RF, SVM and XGBoost). Same datasets were used in BAAD package, and the corresponding ROC plots were generated using overall Alzheimer's score (ADS) as well as individual region-specific atrophy scores for the hippocampus (AHL: Anterior Hippocampus Left; PHL: Posterior Hippocampus Left; AHR: Anterior Hippocampus Right; PHR: Posterior Hippocampus Right), respectively.

diagnostic method and highlights AD-specific tissue atrophy efficiently [49,50].

Similarly, machine learning (ML) and deep learning [DL] methods are widely used for AD identification from the sMRI data. Although DL methods are proven to be more accurate and robust in detecting AD and its classes, they usually perform best when the data and feature sizes are much larger and higher, respectively. In addition, DL also requires superior computational resources and expertise. DL methods such as convolutional neural networks (CNNs) are capable of extracting features directly from the image data compared to traditional machine learning methods that require predetermined expert-driven feature calculation *a-priori*. However, the newer generation of ML methods such as random forest (RF), XGBoost and KNN aided by improved and automatic segmentation of MRI images along with automatic identification features of known and novel AD biomarker ROIs, can be of great value to make the applications much more efficient and useful especially for rapid identification of AD/MCI for the query MRI data.

The corpus callosum (CC) emerges as a promising AD biomarker, underscored by numerous studies [34–40]. These studies have established that in AD, morphological changes can appear not only in the grey matter cortical area but also in the white matter tract. Changes in the shape, size, and circularity of CC were observed in MCI and AD, and subsequently, these altered features could distinguish AD/MCI cohorts from healthy samples quite successfully when implemented via machine learning-based classification [40]. Detection of ROI (e.g., CC) from each slice of the MR images is a critical step and requires manual interventions. To overcome the manual involvement and to reduce the error and bias associated with manual selection, we developed an algorithm to detect CC-like sub-structures automatically from the brain slices. We performed automatic detection and segmentation of corpus callosum (CC) and further extracted shape and size-based features of CC visible from the sagittal slices of the brain MRI scans to feed into multivariate pattern analysis using the most widely used machine learning and deep learning techniques like SVM, RF, XGBoost, KNN, and ANN, respectively. Automatic detection and segmentation of CC followed by feature extraction are relatively less error-prone and much faster than other known ROIs such as the hippocampus and therefore making this approach suitable for web-based prediction software. Our rigorous benchmarking analyses using large AD and healthy cohorts also support its utility and efficacy. Performances of the test datasets and, more importantly for the validation datasets are very encouraging given the larger data size in various categories of AD. It has been observed that some classification algorithms are particularly more successful in identifying healthy samples correctly, while AD samples, especially mild and moderate ADs are best predicted/classified by RF/XGBoost. 100 % models, which are a combination of 80 % training and corresponding test (20 %) datasets, perform relatively better than the only training (80 %) models. In addition, a more realistic validation of CCADD server was

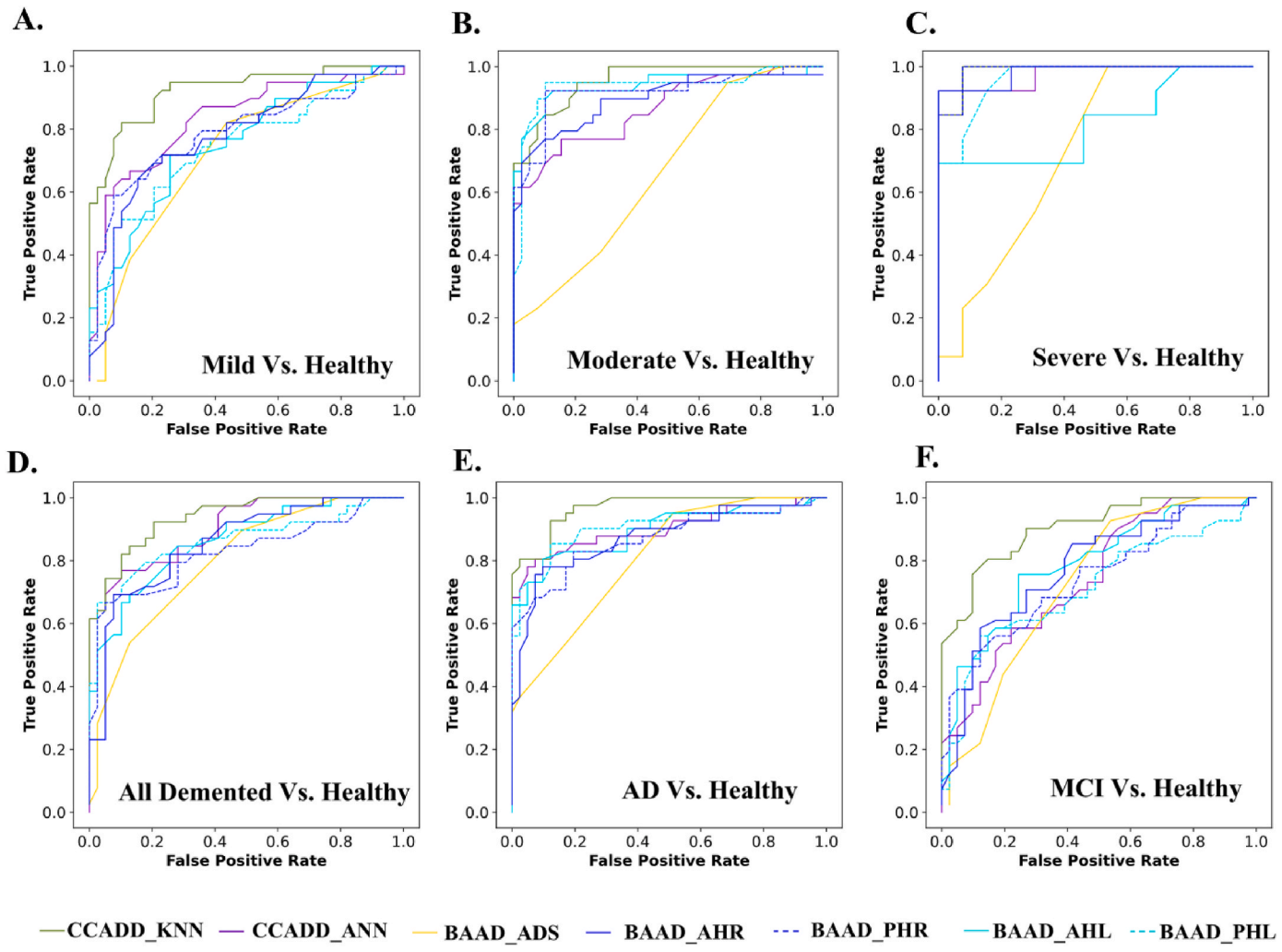

Fig. 5. Comparison of performance of the CCADD server. Receiver operating characteristics (ROC) plots were generated using the validation datasets from various CDR-based (mild, moderate, severe, and all demented) (A–D) and ADNI-based groupings (E–F) against the CDR and ADNI-based training models (KNN and ANN). Same datasets were used in BAAD package, and the corresponding ROC plots were generated using overall Alzheimer's score (ADS) as well as individual region-specific atrophy scores for the hippocampus (AHL: Anterior Hippocampus Left; PHL: Posterior Hippocampus Left; AHR: Anterior Hippocampus Right; PHR: Posterior Hippocampus Right), respectively.

also performed, where each patient and healthy samples were tested independently. Some of the classifiers perform very well for all the categories. Further, a comprehensive comparison of performance of the CCADD server was done with respect to an available brain image processing and AI based package named BAAD. ROC plots (Figs. 4–5 and Supplementary Figs. S3–S4) clearly indicate much better performances of XGBoost, RF and KNN models of the CCADD server with respect to that achieved by the BAAD package using all the ROIs, hippocampus, and CC atrophies, respectively.

For ROI-based AD prediction, the most critical steps are segmentation, identification, and, subsequently, feature extraction of the ROIs to differentiate between AD/MCI and healthy samples. Researchers commonly use Free Surfer [51], SPM [52], FSL [53], DIPY [54], and MATLAB [55] software packages to assess different neuro regions of the brain. Through these methods, various areas can be identified by means of handcrafted feature extraction, which has significantly enhanced AD detection and treatment [56]. However, these methods require a large amount of computing power, expertise computation skills, and hefty subscription fees. Hence, the development of a simple-to-use, fast, accurate, and freely available platform for the detection of AD using MRI data is warranted. CCADD server embeds automatic selection and subsequent segmentation of the corpus callosum region from each sagittal brain slice MR image with high accuracy [40], alleviating the requirement of expert manual intervention. It generates results within 2–5 min for most of the query MRI data with reasonable size. Benchmarking results are encouraging for its efficacy in the prediction/detection of AD probability. Hence, this online software package will be valuable for the general, scientific, and medical communities for further revalidation/support of the MRI-based observation by domain experts and clinicians. However, it is important to note that it is an academic software package enabled by atrophy-based features of only one ROI, and therefore, it is limited to detect other abnormalities and finer anomalies that experienced radiologists and/or neurologists usually detect. No clinical decisions and diagnoses should be taken on the basis of CCADD prediction without clinical guidance. In the future, we aim to add morphological alterations of other critical regions of interest in combination with CC atrophy to further improve the accuracy of the CCADD server.

CRediT authorship contribution statement

Priyanka Panigrahi: Data curation, Formal analysis, Investigation, Methodology, Software. **Subhrangshu Das:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software,

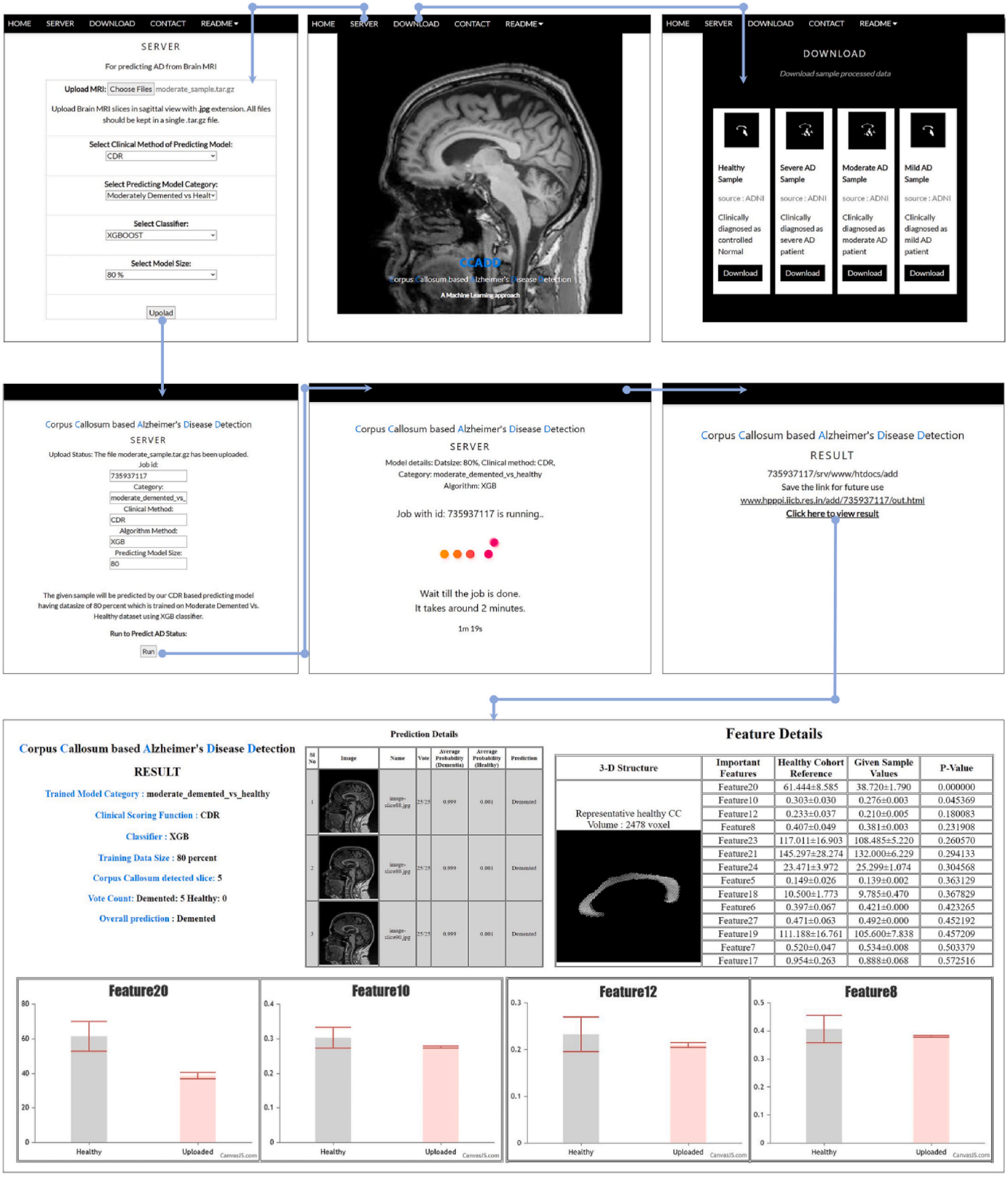

Fig. 6. Snapshot of the CCADD server and excerpts of its sample result pages.

Validation, Writing – review & editing. **Saikat Chakrabarti:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Subhrangshu Das reports financial support was provided by India

Ministry of Science & Technology Department of Biotechnology.

Acknowledgements

The authors acknowledge CSIR-Indian Institute of Chemical Biology for infrastructural support. SC and SD thank DBT for research grant (BT/PR40137/BTIS/137/35/2022) and post-doctoral fellowship. PP acknowledges DST-INSPIRE fellowship from Department of Science and Technology (DST), Government of India.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108622.

References

[1] D. Mehta, R. Jackson, G. Paul, J. Shi, M. Sabbagh, Why do trials for Alzheimer's disease drugs keep failing? A discontinued drug perspective for 2010-2015, Expert OpinInvestig Drugs 26 (2017) 735–739.
[2] C.A. Lane, J. Hardy, J.M. Schott, Alzheimer's disease, Eur. J. Neurol. 25 (2018) 59–70.
[3] C. Flicker, S.H. Ferris, B. Reisberg, Mild cognitive impairment in the elderly: predictors of dementia, Neurology 41 (1991) 1006–1009.
[4] R.C. Petersen, G.E. Smith, S.C. Waring, R.J. Ivnik, E.G. Tangalos, E. Kokmen, Mild cognitive impairment: clinical characterization and outcome, Arch. Neurol. 56 (1999) 303–308.
[5] B. Winblad, K. Palmer, M. Kivipelto, V. Jelic, L. Fratiglioni, L.O. Wahlund, A. Nordberg, L. Backman, M. Albert, O. Almkvist, et al., Mild cognitive impairment–beyond controversies, towards a consensus: report of the International working group on mild cognitive impairment, J. Intern. Med. 256 (2004) 240–246.
[6] S. Gauthier, B. Reisberg, M. Zaudig, R.C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, et al., Mild cognitive impairment, Lancet 367 (2006) 1262–1270.
[7] F.E. Taragano, R.F. Allegri, C. Lyketsos, Mild behavioral impairment: a prodromal stage of dementia, Dement Neuropsychol 2 (2008) 256–260.
[8] J. Fessel, Prevention of Alzheimer's disease by treating mild cognitive impairment with combinations chosen from eight available drugs, Alzheimers Dement (N Y) 5 (2019) 780–788.
[9] Y.T. Quiroz, R.A. Sperling, D.J. Norton, A. Baena, J.F. Arboleda-Velasquez, D. Cosio, A. Schultz, M. Lapoint, E. Guzman-Velez, J.B. Miller, et al., Association between amyloid and tau accumulation in young adults with autosomal dominant alzheimer disease, JAMA Neurol. 75 (2018) 548–556.
[10] R.J. Killiany, T. Gomez-Isla, M. Moss, R. Kikinis, T. Sandor, F. Jolesz, R. Tanzi, K. Jones, B.T. Hyman, M.S. Albert, Use of structural magnetic resonance imaging to predict who will get Alzheimer's disease, Ann. Neurol. 47 (2000) 430–439.
[11] N.C. Fox, S. Cousens, R. Scahill, R.J. Harvey, M.N. Rossor, Using serial registered brain magnetic resonance imaging to measure disease progression in Alzheimer disease: power calculations and estimates of sample size to detect treatment effects, Arch. Neurol. 57 (2000) 339–344.
[12] D. Chan, N.C. Fox, R. Jenkins, R.I. Scahill, W.R. Crum, M.N. Rossor, Rates of global and regional cerebral atrophy in AD and frontotemporal dementia, Neurology 57 (2001) 1756–1763.
[13] R.I. Scahill, J.M. Schott, J.M. Stevens, M.N. Rossor, N.C. Fox, Mapping the evolution of regional atrophy in Alzheimer's disease: unbiased analysis of fluid-registered serial MRI, Proc Natl AcadSci U S A 99 (2002) 4703–4707.
[14] N.C. Fox, W.R. Crum, R.I. Scahill, J.M. Stevens, J.C. Janssen, M.N. Rossor, Imaging of onset and progression of Alzheimer's disease with voxel-compression mapping of serial magnetic resonance images, Lancet 358 (2001) 201–205.
[15] D. Chan, N.C. Fox, R. Jenkins, R.I. Scahill, W.R. Crum, M.N. Rossor, Rates of global and regional cerebral atrophy in AD and frontotemporal dementia, Neurology 57 (2001) 1756–1763.
[16] C.R. Jack Jr., R.C. Petersen, Y. Xu, P.C. O'Brien, G.E. Smith, R.J. Ivnik, E. G. Tangalos, E. Kokmen, Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease, Neurology 51 (1998) 993–999.
[17] C.R. Jack Jr., R.C. Petersen, Y. Xu, P.C. O'Brien, G.E. Smith, R.J. Ivnik, B.F. Boeve, E.G. Tangalos, E. Kokmen, Rates of hippocampal atrophy correlate with change in clinical status in aging and AD, Neurology 55 (2000) 484–489.
[18] S.J. Teipel, W. Bayer, G.E. Alexander, Y. Zebuhr, D. Teichberg, L. Kulic, M. B. Schapiro, H.J. Moller, S.I. Rapoport, H. Hampel, Progression of corpus callosum atrophy in Alzheimer disease, Arch. Neurol. 59 (2002) 243–248.
[19] A.T. Du, N. Schuff, J.H. Kramer, S. Ganzer, X.P. Zhu, W.J. Jagust, B.L. Miller, B. R. Reed, D. Mungas, K. Yaffe, et al., Higher atrophy rate of entorhinal cortex than hippocampus in AD, Neurology 62 (2004) 422–427.
[20] P.M. Thompson, K.M. Hayashi, G. de Zubicaray, A.L. Janke, S.E. Rose, J. Semple, D. Herman, M.S. Hong, S.S. Dittmer, D.M. Doddrell, et al., Dynamics of gray matter loss in Alzheimer's disease, J. Neurosci. 23 (2003) 994–1005.
[21] M.A. DeTure, D.W. Dickson, The neuropathological diagnosis of Alzheimer's disease, Mol. Neurodegener. 14 (2019) 32.
[22] O. Piguet, K.L. Double, J.J. Kril, J. Harasty, V. Macdonald, D.A. McRitchie, G. M. Halliday, White matter loss in healthy ageing: a postmortem analysis, Neurobiol. Aging 30 (2009) 1288–1295.
[23] G. Uysal, M. Ozturk, Hippocampal atrophy based Alzheimer's disease diagnosis via machine learning methods, J. Neurosci. Methods 337 (2020) 108669.
[24] H.C. Achterberg, L. Sørensen, F.J. Wolters, W.J. Niessen, M.W. Vernooij, M. A. Ikram, M. Nielsen, M. de Bruijne, The value of hippocampal volume, shape, and texture for 11-year prediction of dementia: a population-based study, Neurobiol. Aging 81 (2019) 58–66.
[25] D. Holilah, A. Bustamam, D. Sarwinda, Detection of Alzheimer's disease with segmentation approach using K-Means Clustering and Watershed Method of MRI image, J. Phys. Conf. 1725 (2021) 1.
[26] R. Wolz, V. Julkunen, J. Koikkalainen, E. Niskanen, D.P. Zhang, D. Rueckert, H. Soininen, J. Lotjonen, I. Alzheimer's Disease Neuroimaging, Multi-method analysis of MRI images in early diagnostics of Alzheimer's disease, PLoS One 6 (2011) e25446.
[27] C. Fennema-Notestine, D.J. Hagler Jr., L.K. McEvoy, A.S. Fleisher, E.H. Wu, D. S. Karow, A.M. Dale, I. Alzheimer's Disease Neuroimaging, Structural MRI biomarkers for preclinical and mild Alzheimer's disease, Hum. Brain Mapp. 30 (2009) 3238–3253.
[28] J.H. Morra, Z. Tu, L.G. Apostolova, A.E. Green, C. Avedissian, S.K. Madsen, N. Parikshak, X. Hua, A.W. Toga, C.R. Jack Jr., et al., Automated 3D mapping of hippocampal atrophy and its clinical correlates in 400 subjects with Alzheimer's disease, mild cognitive impairment, and elderly controls, Hum. Brain Mapp. 30 (2009) 2766–2788.
[29] P. Forouzannezhad, A. Abbaspour, C. Li, C. Fang, U. Williams, M. Cabrerizo, A. Barreto, J. Andrian, N. Rishe, R.E. Curiel, D. Loewenstein, R. Duara, M. Adjouadi, A Gaussian-based model for early detection of mild cognitive impairment using multimodal neuroimaging, J. Neurosci. Methods 333 (2020) 108544.
[30] G. Uysal, M. Ozturk, Comparative analysis of different brain regions using machine learning for prediction of EMCI and LMCI stages of Alzheimer's disease, Multimed. Tool. Appl. 83 (2023) 1–16.
[31] K. Vaithinathan, L. Parthiban, Alzheimer's Disease Neuroimaging Initiative, A novel texture extraction technique with T1 weighted MRI for the classification of Alzheimer's disease, J. Neurosci. Methods 318 (2019) 84–99.
[32] F. Zhang, S. Tian, S. Chen, Y. Ma, X. Li, X. Guo, Voxel-based morphometry: improving the diagnosis of Alzheimer's disease based on an extreme learning machine method from the ADNI cohort, Neuroscience 414 (2019) 273–279.
[33] C. Fang, C. Li, P. Forouzannezhad, M. Cabrerizo, R.E. Curiel, D. Loewenstein, R. Duara, M. Adjouadi, Alzheimer's Disease Neuroimaging Initiative, Gaussian discriminative component analysis for early detection of Alzheimer's disease: a supervised dimensionality reduction algorithm, J. Neurosci. Methods 344 (2020) 108856.
[34] S. Elahi, A.H. Bachman, S.H. Lee, J.J. Sidtis, B.A. Ardekani, I. Alzheimer's Disease Neuroimaging, Corpus callosum atrophy rate in mild cognitive impairment and prodromal Alzheimer's disease, J Alzheimers Dis 45 (2015) 921–931.
[35] S. Kamal, I. Park, Y.J. Kim, Y.J. Kim, U. Lee, Alteration of the corpus callosum in patients with Alzheimer's disease: deep learning-based assessment, PLoS One 16 (2021) e0259051.
[36] A.H. Bachman, S.H. Lee, J.J. Sidtis, B.A. Ardekani, Corpus callosum shape and size changes in early Alzheimer's disease: a longitudinal MRI study using the OASIS brain database, J Alzheimers Dis 39 (2014) 71–78.
[37] M. Di Paola, G. Spalletta, C. Caltagirone, In vivo structural neuroanatomy of corpus callosum in Alzheimer's disease and mild cognitive impairment using different MRI techniques: a review, J Alzheimers Dis 20 (2010) 67–95.
[38] M. Di Paola, E. Luders, F. Di Iulio, A. Cherubini, D. Passafiume, P.M. Thompson, C. Caltagirone, A.W. Toga, G. Spalletta, Callosal atrophy in mild cognitive impairment and Alzheimer's disease: different effects in different stages, Neuroimage 49 (2010) 141–149.
[39] M. Zhu, X. Wang, W. Gao, C. Shi, H. Ge, H. Shen, Z. Lin, Corpus callosum atrophy and cognitive decline in early Alzheimer's disease: longitudinal MRI study, Dement. Geriatr. Cogn. Disord 37 (2014) 214–222.
[40] S. Das, P. Panigrahi, S. Chakrabarti, Corpus callosum atrophy in detection of mild and moderate Alzheimer's disease using brain magnetic resonance image processing and machine learning techniques, J Alzheimers Dis Rep 5 (2021) 771–788.
[41] W. Yu, B. Lei, S. Wang, Y. Liu, Z. Feng, Y. Hu, Y. Shen, M.K. Ng, Morphological feature visualization of Alzheimer's disease via multidirectional perception gan, IEEE Transact. Neural Networks Learn. Syst. 34 (2023) 4401–4415.
[42] J. Pan, B. Lei, Y. Shen, Y. Liu, Z. Feng, S. Wang, Characterization multimodal connectivity of brain network by hypergraph gan for Alzheimer's disease analysis, Pattern Recognition and Computer Vision (2021) 13021.
[43] Q. Zuo, B. Lei, Y. Shen, Y. Liu, Z. Feng, S. Wang, Multimodal representations learning and adversarial hypergraph fusion for early Alzheimer's disease prediction. Chinese Conference on Pattern Recognition and Computer Vision, 2021.
[44] A.H. Syaifullah, A. Shiino, H. Kitahara, R. Ito, M. Ishida, K. Tanigaki, Machine learning for diagnosis of AD and prediction of MCI progression from brain MRI using brain anatomical analysis using diffeomorphic deformation, Front. Neurol. 11 (2020) 576029.
[45] R.M. Hanson, J. Prilusky, R. Zhou, T. Nakane, J.L. Sussman, JSmol and the next-generation web-based representation of 3D molecular structure as applied to proteopedia, Isr. J. Chem. 53 (2013) 207–216.
[46] A. Shukla, R. Tiwari, S. Tiwari, Review on alzheimer disease detection methods: automatic pipelines and machine learning techniques, Sci 5 (2023) 13.

[47] S. Afzal, M. Maqsood, U. Khan, I. Mehmood, H. Nawaz, F. Aadil, O.Y. Song, Y. Nam, Alzheimer disease detection techniques and methods: a review, International Journal of Interactive Multimedia and Artificial Intelligence 6 (7) (2021) 26–38.
[48] T. Goel, R. Sharma, M. Tanveer, P.N. Suganthan, K. Maji, R. Pilli, Multimodal neuroimaging based Alzheimer's disease diagnosis using evolutionary RVFL classifier, IEEE J Biomed Health Inform (2023).
[49] M.A. Ebrahimighahnavieh, S. Luo, R. Chiong, Deep learning to detect Alzheimer's disease from neuroimaging: a systematic literature review, Comput. Methods Progr. Biomed. 187 (2020) 105242.
[50] H. Hampel, K. Broich, Y. Hoessler, J. Pantel, Biological markers for early detection and pharmacological treatment of Alzheimer's disease, Dialogues Clin. Neurosci. 11 (2009) 141–157.
[51] M. Reuter, N.J. Schmansky, H.D. Rosas, B. Fischl, Within-subject template estimation for unbiased longitudinal image analysis, Neuroimage 61 (2012) 1402–1418.
[52] J. Ashburner, Computational anatomy with the SPM software, Magn. Reson. Imaging 27 (2009) 1163–1174.
[53] M. Jenkinson, C.F. Beckmann, T.E. Behrens, M.W. Woolrich, S.M. Smith, Fsl, Neuroimage 62 (2012) 782–790.
[54] E. Garyfallidis, M. Brett, B. Amirbekian, A. Rokem, S. van der Walt, M. Descoteaux, I. Nimmo-Smith, C. Dipy, Dipy, a library for the analysis of diffusion MRI data, Front. Neuroinf. 8 (2014) 8.
[55] The MathWorks Inc., MATLAB Version: 9.13.0 (R2022b), The MathWorks Inc, Natick, Massachusetts, 2022.
[56] A. Bernstein, R. Akzhigitov, E. Kondrateva, S. Sushchinskaya, I. Samotaeva, V. Gaskin, MRI brain imagery processing software in data analysis, Proceedings of the 13th International Conference 9 (2018) 61–74.