Article

Explainable AI-based Alzheimer's Prediction and Management Using Multimodal Data

Sobhana Jahan [1], Kazi Abu Taher [1], M Shamim Kaiser [2], Mufti Mahmud [3], Md. Sazzadur Rahman[2], A. S. M Sanwar Hosen[4], and In-Ho Ra[5]*

1 Department of Information and Communication Technology, Bangladesh University of Professionals, Dhaka, Bangladesh; sobhanajahanlisa@gmail.com, katahe r@bup.edu.bd
2 Institute of Information Technology, Jahangirnagar University, Dhaka, Bangladesh; msKaiser@juniv.edu, sazzad@juniv.edu
3 Department of Computer Science, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK
Medical Technologies Innovation Facility, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK
Computing and Informatics Research Centre, Nottingham Trent University, Clifton Lane, Nottingham NG11 8NS, UK; muftimahmud@gmail.com, mufti.mahmud@ntu.ac.uk
4 Division of Computer Science and Engineering, Jeonbuk National University, Jeonju 54896, Korea; sanwar@jbnu.ac.kr
5 School of Computer, Information and Communication Engineering, Kunsan National University, Gunsan 54150, Korea; ihra@kunsan.ac.kr
* Correspondence: ihra@kunsan.ac.kr (I.-H.R), M Shamim Kaiser (M.S.K.)

Abstract: According to the World Health Organization (WHO), dementia is the seventh leading reason of death among all illnesses and one of the leading causes of disability among the world's elderly people. Day by day the number of Alzheimer's patients is raising. Considering the increasing rate and the dangers, Alzheimer's disease should be diagnosed carefully. Machine learning is a potential technique for Alzheimer's diagnosis but general users do not trust machine learning models due to the black-box nature. Even, some of those models do not provide the best performance because of using only neuroimaging data. To solve these issues, this paper proposes a novel explainable Alzheimer's disease prediction model using a multimodal dataset. This approach performs a data-level fusion using clinical data, Freesurfer MRI segmentation data, and psychological data. For Alzheimer's disease vs cognitively normal prediction, the random forest classifier provides 100% accuracy. Furthermore, Alzheimer's disease and non-Alzheimer's dementia should be classified properly because their symptoms are similar. To the best of our knowledge, we are the first to present a three-class classification on Alzheimer's disease vs cognitively normal vs non-Alzheimer's dementia and achieved 99.86% accuracy using an ensemble model. Besides, a novel Alzheimer's patient management architecture is also proposed in this work..

Keywords: Machine learning; Dementia; Data-level fusion

check for updates

Citation: Lastname, F.; Lastname, F.; Lastname, F. Title. *Preprints* **2022**, *1*, 0. https://doi.org/

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Alzheimer's disease (AD) is a chronic, progressive neurodegenerative disease that gradually deteriorates memory and cognitive abilities, and it is the most common cause of dementia in older people. Currently, more than 55 million people live with dementia around the globe. Forgetting recently acquired information, important dates or events, difficulty in performing simple daily works, and repeatedly asking the same questions are all classic early symptoms of Alzheimer's disease. In the final stage, patients' behavioral changes are also observed. The disease strikes the majority of people in their mid-60s. Scientists agree that the root cause of this neurological disease is a combination of genetics, long-term environmental conditions, and lifestyle [1]. Though some medications are available, AD is not curable, and the damage it causes is permanent. The most common cause of death of Alzheimer's patients is aspiration pneumonia [2].

Even though Alzheimer's disease is incurable, predicting Alzheimer's can help to identify patients who are at risk of this disorder. As dementia is a common symptom for AD and non-AD patients, proper diagnosis is important. Here, non-AD dementia means

© 2022 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

Vascular Dementia, Dementia with Lewy Bodies Disease (DLBD), Parkinson's disease (PD), and Frontotemporal dementia. Though machine learning (ML) is a very potent technique in Alzheimer's diagnosis, people nowadays are more likely to visit clinics and have their Alzheimer's disease diagnosed. They typically do not believe in ML-based predictions because of the black-box nature of ML models. Even doctors are also willing to rely on clinical diagnosis rather than the ML model's prediction. For these reasons, explaining the ML models' decisions or the features responsible for this decision is the best way to gain the trust of the general people. Explainable Artificial Intelligence (AI) is a magical tool that can annotate a model's decisions as well as decision-making characteristics. Furthermore, the use of only neuroimaging data for AD prediction is very common, and achieving good accuracy is difficult in most cases. Even using single modal data for this type of critical prediction is very risky as sometimes it can produce faulty predictions. To remove these problems, our work proposes an explainable ML model using multimodal data.

The quality of life of AD patients can be improved if there is a constant helping hand with them. But, due to modernization and capitalism, almost all family members are busy with their work. Furthermore, urban people are so busy that they can hardly give constant support and care to Alzheimer's patients. Due to the lack of care and support, the condition of those patients may decrease very rapidly. Besides, the scarcity of Alzheimer's patient care centers is also a problem. A sensor-based and IoT-enabled real-time monitoring and AD patient management system can solve these problems. For this reason, this paper presents a complete model architecture for monitoring and managing Alzheimer's patients.

To the best of our knowledge, we are the first to present this multimodal approach using clinical data, Freesurfer data, and psychological data, which are collected from the OASIS-3 dataset. The main contributions of this research work are:

- We have proposed an AD prediction approach using a multimodal dataset. To ensure multimodality, we have performed data-level fusion using Alzheimer's Disease Research Center (ADRC) clinical data, Freesurfer brain MRI segmentation data, and psychological assessments.
- We have developed two AD prediction models using ML classifiers. Here, Random Forest (RF) has given the best performance for AD vs cognitively normal (CN) prediction. The Voting classifier-based ensemble model provided the best performance for AD vs CN vs non-AD dementia.
- We have made the block box ML decisions to an explainable one where general people can understand the reasons behind any prediction.
- We have proposed a 24/7 AD patient monitoring and management system architecture.

The remainder of this paper is organized as follows: section 2 comes with some recently published similar works. Section 3 contains the methodology of the proposed work. Section 4 shows the performance analysis, and Section 5 contains the concluding remarks.

2. Related Works

Using MRI images and gene data sets, Kamal et al. [3] used SpinalNet and CNN to classify AD from MRI images. The researchers then used microarray gene expression data to classify diseases using KNN, SVC, and Xboost classification techniques. Instead of using gene and image data solely, the authors had combined these two approaches and explained the results using the Local Interpretable Model-agnostic Explanations (LIME) method. For MRI image classification, the accuracy rate of CNN is 97.6%, and for the gene expensive data, SVC outperforms other approaches.

El-Sappagh et al. [4] proposed a two-layered explainable ML model for AD classification. It was a multimodal approach where data from 11 modalities (genetics, medical history, MRI, PET, neuropsychological battery, cognitive scores, etc.) were integrated. RF classifier was used here in the first layer for multiclass classification, and the results were explained using the SHAP framework. In the second layer, binary classification took place, where probable MCI to AD classification took place. For the first layer, the achieved cross-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

validation F1-score was 93.94%, and accuracy was 93.95% (multiclass classifier). For the second layer, accuracy and F1-score were 87.08% and 87.09%, respectively.

Lee et al. [5] proposed a multimodal Recurrent Neural Network (RNN) model to predict AD from the Mild Cognitive Impairment (MCI) stage. In this approach, the authors had integrated subjects' longitudinal Cerebrospinal Spinal Fluid (CSF) and the cognitional performance biomarkers along with cross-sectional neuroimaging data and demographic data. Here, all data were collected from the ADNI website. The proposed model was divided into two layers. Layer one consists of four Gated Recurrent Units (GRUs) where each contains one modality of data. From the 1st layer, a fixed-sized feature vector was produced. Then, the vectors were concatenated to the input for the final layer. The final layer presents the ultimate prediction. From MCI to AD prediction, the proposed model achieved 76% accuracy and 0.83 AUC using data from a single modality, whereas 81% accuracy and 0.86 AUC value had been achieved using multimodal data.

Zhang et al. [6] proposed a multimodal multi-task learning method for predicting multiple features from multimodal data. This method was divided into two parts. First, a multi-task feature selection that selects a common subset of relevant features for multiple variables from each modality. Second, a multi-modal support vector machine that fuses the previously selected features from all modalities to predict multiple variables. Here, all data were collected from the ADNI website. The accuracy of the proposed model was 83.2%1.5% (MCI vs HC) and 93.3%2.2% (AD vs HC).

Buvari et al. [7] proposed a Convolutional Neural Network (CNN) and Neural Network (NN) (dense) based AD prediction multimodal model where MRI and Numerical Freesurfer data were used. Dataset was collected from the OASIS-3 website. The accuracy of numerical, image, and hybrid approaches was 73.593%, 71.429%, and 74.891%. Kumari et al. [8] proposed a Multi-layer CNN-based model for predicting AD using MRI and PET data collected from the OASIS-3 repository. The model achieved an accuracy of 71% using 500 subjects and 74% using 1098 subjects.

3. Proposed Alzheimer's Prediction Model

3.1. Dataset Acquisition and Preparation

The Open Access Series of Imaging Studies (OASIS)-3 [9] is the most recent release in the OASIS, which aims to make neuroimaging datasets freely accessible to the scientific community. The OASIS-3 dataset includes longitudinal neuroimaging, cognitive, clinical, and biomarker data for normal aging and AD. Participants range in age from 42 to 95 years old and include 609 cognitively normal adults and 489 people of cognitive decline. Here, all participants were provided with an identifier, and all dates were deleted and standardized to reflect the days since their enrollment. Many of the Magnetic Resonance (MR) sessions were accompanied by volumetric segmentation files generated by Freesurfer. For this research work, we have used the Alzheimer's Disease Research Center (ADRC) clinical data, Freesurfer volumetric segmentation data, and psychological data. We have used the longitudinal data of 757 unique participants for predicting AD vs CN vs non-AD dementia. For predicting AD vs CN, we have used 744 unique participant data.

3.1.1. ADRC Clinical Data

ADRC clinical data consists of the longitudinal data of unique 1098 participants. There are various features of cognitively normal, AD dementia, uncertain dementia, and some non-AD dementia, for example, Vascular Dementia, DLBD, PD, and Frontotemporal dementia. The important features of these participants were Mini-Mental State Exam (MMSE), age, judgment, memory, APOE (apolipoprotein E gene), Personal Emergency Response System (PERS) care, height, weight, Orient (recent and long-term memory testing), Clinical Dementia Rating (CDR), and Sumbox (clinical dementia rating scale).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

3.1.2. Freesurfer Data

Freesurfer is an open-source software suite that can process and analyze MRI images of the human brain. This dataset gives us the value of volumetric data of different parts of the human brain such as the intracranial, total cortex, left and right hemisphere cortex, subcortical gray, total gray, supratentorial, left and right hemisphere cortical white matter, and cortical white matter.

3.1.3. Psychological Assessment Data

The psychological assessment dataset contains various popular psychological tests such as Boston naming test, Trailmaking A (Trail A), Trailmaking B (Trail B), animals, vegetables, digit symbol, digit span, logical memory, Wechsler Adult Intelligence Scale (WAIS), and so on.

3.2. Multimodal Dataset Preparation

This research work consists of AD prediction using multimodal data. Here, data-level fusion is performed for creating multimodal data. For making it multimodal, we have integrated three different domain datasets which are the clinical data, the Freesurfer data, or more precisely, brain MRI segmentation data, and the psychological assessment data. The data fusion process begins with the integration of three individual datasets. After integration, we have removed all the duplicate data from the dataset. Then, all the NAN values are replaced with the most frequent value of that feature.

3.3. Classification Model

3.3.1. Random Forest

During training, RF builds a large number of individual decision trees. After that, the estimations from all those trees or nodes are combined to form the final prediction, which is in the form of the classes for classification. Here, the reduction in node imperfection weighted by the probability of achieving that node is used to measure the feature's importance. The number of observations that reach the node is divided by the number of samples that yield the chances of a node. The more significant the feature, the greater the value.

For every decision tree, a node's significance can be determined using Gini Importance. Let's assume only two child nodes (binary tree):

$$ni_j = W_j I_j - W_{left(j)} I_{left(j)} - W_{right(j)} I_{right(j)} \tag{1}$$

Here,
ni_j= Importance of node j
W_j= Weighted value of samples that has reached node j
I_j= Impurity value of node j
$left(j)$= Left child of node j
$right(j)$=Right child of node j

The significance of each feature on the decision tree is calculated as follows:

$$Fi_i = \frac{\sum_{j:\, node\ j\ splits\ on\ feature\ i} ni_j}{\sum_{k\ \epsilon\ all\ nodes} ni_k} \tag{2}$$

Here,
Fi_i = Feature i's importance
ni_j = Node j's importance

This can then be normalized to a range between 0 and 1 by dividing the total amount of the values of the feature importance.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

$$norm\ Fi_i = \frac{Fi_i}{\sum_{j\,\epsilon\, all\ features} Fi_j} \tag{3}$$

At the RF level, the resulting value of feature importance is its average across all trees. The amount of significance value of each feature on every tree is measured and divided by the total number or amount of trees.

$$RF\ Fi_i = \frac{\sum_{j\,\epsilon\, all\ tree} norm\ Fi_{ij}}{T} \tag{4}$$

Here,
$RF\ Fi_i$ = The significance of a feature derived from all the trees in the RF model.
$norm\ Fi_{ij}$= in tree j, the normalized feature importance for i.
T = Total number of trees.

3.3.2. Decision Tree

A decision tree is a decision-making aid that employs a tree-like framework of decisions and their potential outcomes. It is indeed one way to show an algorithm made up entirely of conditional control statements.

In general, we know Decision Tree (DT) works in a stepwise fashion and has a tree structure where we divide a node using a feature based on a criterion. There are several splitting criteria used in DTs; we will not go into those conceptual parts in this paper. Gini impurity is one of the powerful techniques for the calculation of DT. Gini impurity is an assessment of how often a randomly selected element from the collection would be classified incorrectly, if it is labeled according to the dispersion of labels in the sample. So, Gini impurity is the assessment of impurity in a node. Its formula is as follow:

$$I_G(n) = 1 - \sum_{i=1}^{k} (Pi)^2 \tag{5}$$

Here,
Pi = Probability value.
k = Total number of samples.

3.3.3. Gradient Boosting

Gradient boosting is an ensemble method in which you produce various weak models (usually decision trees) and integrate them to improve overall performance. Gradient Boosting is made up of three major components:

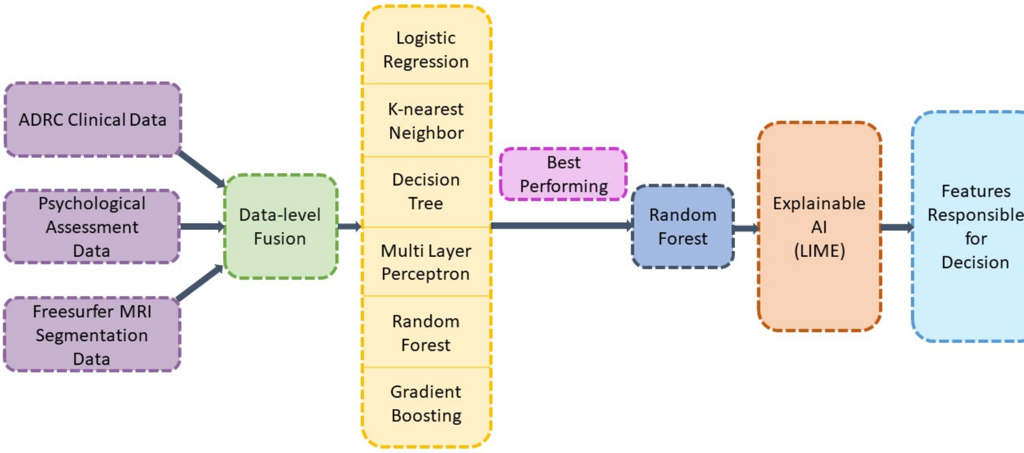
1. Loss function: The loss function's role is to measure how good the model is at predicting things using the data set. This may vary based on the nature of the problem.
2. Weak learner: A weak learner ineffectively classifies data, which is possibly not better than random guessing. In other words, this has a high rate of error.
3. Consecutive approach: The recursive and consecutive approach is used for adding the weak learners. With each iteration, the value of the loss function is reduced.

3.3.4. Majority Voting Ensemble

A majority voting ensemble is an ML model that combines assumptions from several different models. It's a modeling technique that can help one to improve the model's performance. The equation of majority voting is given below:

$$\sum_{p=1}^{P} D_p, j = max_{j=1}^{C} \sum_{p=1}^{P} D_p, j \tag{6}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1



vs CN.jpg

Figure 1. Architecture for AD vs CN classification

Here,
D_p = p^th classifier's decision.
P= Number of classifiers.
C= Number of classes.
$p = 1....P$.
$j = 1....C$.

3.4. Explainable AD Prediction

Explainable AI is a collection of tools and frameworks designed to assist one to understand and interpret the predictions made by the ML models. It allows one to debug and improve model performance and assist to understand the behaviors of the model. In this research work, we have used the LIME [10] method for making the model's decision visible to other general people.

LIME's basic philosophy for result interpretation and visualization is as follows:

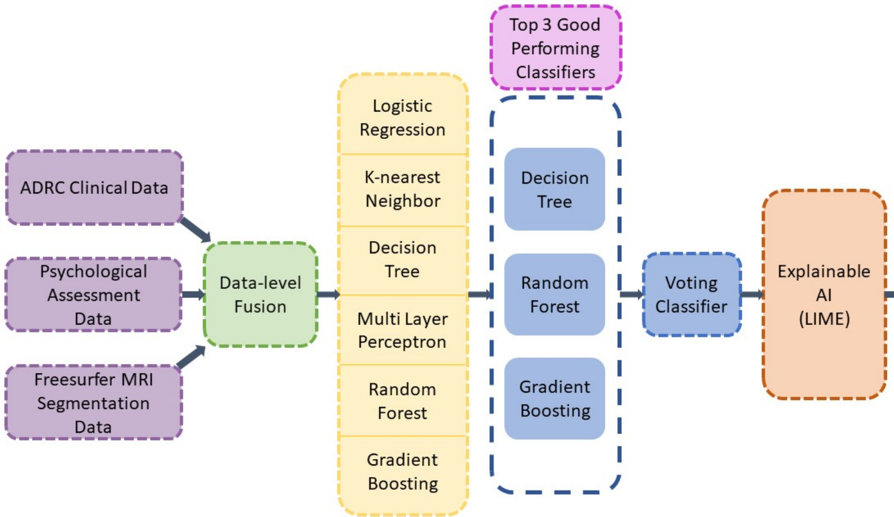
1. Permute the assessment or observation n times for each prediction to describe.
2. Allow the complex model to forecast the outcomes of all permuted observations.
3. Determine the distance between all permutations and the original observation.
4. Convert the range to a score of similarity.
5. From the permuted data, choose m features that best describe the complex model's output.
6. Fit a basic model to the permuted data, describing the complex model's output with the m permuted feature values weighted by their similarity to the original inspection.
7. Take the simple model's feature weights and use them to explain the complex model's local behavior.

$$explanation(x) = argmin_{m \epsilon M} L(f, m, \pi_x) + \omega(m) \tag{7}$$

Here, the explanation model "x" is the model m that minimizes loss L (e.g. mean squared error), and evaluates how close the explanation is to the estimation of the original model f (e.g. the RF model and the ensemble model), while keeping the model complexity $\omega(m)$ low. M is the set of all possible explanations. The similarity measure π_x specifies the size of the neighborhood around the instance x that we recognize for the explanation.

All of the decision-making features and their percentages are examined and viewed in this research work using the LIME explanation. Through this technique, a doctor or a patient can easily understand the reason for the model's decision, as well as the decision-making features and percentage.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1



vs CN vs Non AD.jpg

Figure 2. Architecture for AD vs CN vs non-AD dementia classification

3.5. Proposed AD Prediction Model Architecture

As previously stated, we have performed data-level fusion on three different data domains which are clinical data, Freesurfer MRI segmentation data, and psychological assessment data. So, after completing the data-level fusion, we have trained the model. We have used a variety of ML algorithms to accomplish this and those models are Logistic Regression (LR), K-Nearest Neighbor (KNN), DT, Multi-Layer Perceptron (MLP), RF, and Gradient Boosting. We have discovered that RF is the best performing algorithm for AD vs CN prediction. Besides, we have found that the DT, gradient boosting, and RF are three average performing algorithms for AD vs CN vs non-AD dementia. Because there were three average-performing algorithms, we have developed a customized ensemble technique with the majority voting classifier. This voting classifier outperformed three individual (DT, Gradient Boosting, and RF) algorithms with the help of a soft voting technique. Fig. 1 and 2 show the architecture of the proposed binary and three class classification models. In Fig. 1 we can find that RF is the best performing algorithm and its output is passed to the LIME explainer for achieving the transparency of the model's decision. For three-class classification, our customized ensemble technique is the best performing, which is shown in Fig. 2.

3.6. Proposed AD Patient Management Architecture

If any person is diagnosed as an Alzheimer's patient, we want to keep him under complete monitoring and service. For doing that, a GPS-based wearable sensor band is proposed in this paper, which can do complete monitoring and management. Instead of using only the cloud layer, we have used mist, fog, and cloud to remove data processing latencies and to provide instant responses. Based on the need for data processing resources, mist, fog, or cloud will be instantly selected by the system. Fig. 3 presents a complete view of this proposed patient management architecture. This full management process can be divided into five layers which are described below.

3.6.1. Perception Layer

The perception layer is the lower layer where the sensors of the wearable band capture the raw data from the patient. An optical heart rate sensor, gyroscope, non-invasive blood glucose monitoring sensor, blood pressure monitoring sensor, flex sensor, temperature sensor, and GPS tracker are needed to build this band [11].

- Optical heart rate sensor: Pulse waves, or variations in the volume of a blood vessel caused by the heart pumping blood, are measured by an optical heart rate sensor. A

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

set of optical sensors and a green LED detect pulse waves by measuring the changes in blood volume. This sensor helps to keep track of the AD patients' heat condition [12].

- Gyroscope: A gyroscope sensor is a device that can measure and keep track of a patient's orientation and angular velocity [13]. The concept of conservation of momentum governs the operation of a gyroscope sensor [14]. It functions by maintaining angular momentum. A rotor or spinning wheel is mounted on a pivot in a gyroscope sensor. The pivot allows the rotor to rotate around a single axis, known as a gimbal.
- Blood pressure monitoring sensor: Blood pressure sensors are pushed on the skin to compute the pressure pulse wave and determine the blood pressure [15].
- Non-invasive blood glucose monitoring sensor: The non-invasive blood glucose monitoring sensor system measures blood glucose using the near-infrared (NIR) technique [[16].
- Flex sensor: Patient Monitoring can be done by flex sensor, which helps to recognize a patient's gestures. This sensor, also known as a bend sensor, detects deflection or bending. Typically, the sensor is attached to the patient's body, and the sensor's resistance is varied by bending of the surface [17].
- Temperature sensor: Temperature sensors are used in medical uses to measure the body temperature of patients [18]. The voltage across the diode terminals is the basic working principle of temperature sensors. When the voltage increases, the temperature goes up, resulting in a voltage drop between the emitter and base transistor terminals in a diode [19].
- GPS tracker: Global Positioning System (GPS) tracker is very important for AD patients with dementia [20]. If an AD patient gets missing, this GPS tracker will help to find them out [21].
- LCD: A Liquid Crystal Display (LCD) should be in the wearable band to display necessary, hourly reminders. Exp: Medicines taking reminder, Daily activity reminder [22].
- Mini speaker: A mini speaker will help to provide hourly reminders to the AD patient. This mini speaker will synchronously work with the LCD and serve the AD patient.

3.7. Mist Layer

The mist layer has been added to the framework for processing time-critical data. Mist computing exists directly within the network fabric and operates on the network's edge with the assistance of actuator and sensor controllers. This layer is in charge of the fundamental preprocessing of sensor information such as filtering, data aggregation, and fusion. In this layer, there will also be a comparator and decision-making mechanism. After the sensor data are pre-processed, the required resources for processing those data are calculated. A comparator will check if the necessary resources are less than or equal to the mist resources. If it is, then data are further processed in this layer, and if not, those data are passed to the fog layer [23]. Mist layer's output goes to a rule-based system, and that rule base system will decide the necessary activity like announcing emergency medication and informing family members.

3.8. Fog Layer

The primary driving force behind patient monitoring is the need to process data "on the fly" to detect anomalies, provide near real-time alerts, and instantaneously activate appropriate actions. This necessitates a system with better responsiveness and low latency. Because of their high latency, cloud-based models are ineffective for this purpose. The fog layer brings application services and computing resources closer to the edge, lowering response latency. The fog layer receives those sensor data, which are not processed in the mist layer. It also receives processed data from the mist layer for further calculation if it resembles the patient's critical state. Like mist, the output of the fog layer is routed to a rule-based system, which determines the necessary activity, such as hourly reminders.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

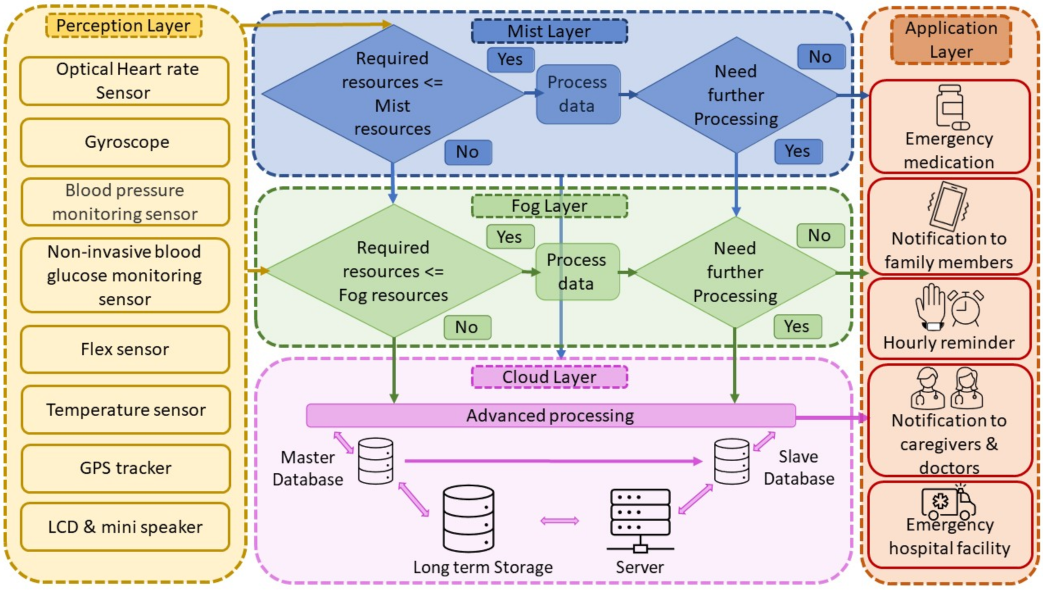


Figure 3. Alzheimer's patient management architecture

3.9. Cloud Layer

The cloud layer can communicate with the mist, fog, and final application layers. AD patients' data from the mist and fog layer are further processed to the cloud layer for long-term storage, as well as big data and advanced analytics. This cloud layer is also connected to a rule-based system.

3.10. Application Layer

The application layer is the final layer that displays necessary information to the patient, relatives, caregivers, doctors, and the hospital authority. Following are some functionalities of the application layer.

- Hourly reminder to patients: With the help of this model, AD patients will get necessary hourly reminders (e.g., Medicines taking reminders, daily activity reminders). From the memory of fog, the information of the hourly reminder comes to the patients' wearable band.
- Emergency medication announcement: The patient will get the announcement of medication in case of an emergency. Suppose, the mist layer processes heart rate data from the sensor and finds the patient is in an emergency state. At that moment, the mist layer will announce the necessary medication information to the patient.
- Notification to family members: Patients' family members will get health updates of the patients. If the patient is in a critical state, the mist layer will send emergency notification to the patient's family member.
- Notification to caregivers and doctors: Caregivers and doctors will get notifications if a patient's condition decreases. Suppose the critical patient's body sensor data needs further processing. In that case, those data come to the fog layer, and the notification goes to the doctor and caregiver in case of an emergency.
- Emergency hospital facility: As we have already stated, the processed data from mist and fog will go to the cloud layer for long-term storage and advanced processing. Suppose the patient is in the critical stage. In that case, the cloud will compute the nearest hospital information with the help of GPS and provide emergency ambulance service and emergency care in the hospital. Even from the cloud database, doctors and caregivers can access the long-term medical history of each particular patient.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

Table 1. Accuracy of ML classifiers using Individual datasets (bold values are best performing results).

Classification Type	ML Classifier	Clinical	Freesurfer	Psychological
AD vs CN	LR	93.23%	**74.08%**	**76.03%**
	KNN	92.45%	67.97%	71.43%
	DT	93.35%	55.01%	59.83%
	MLP	92.84%	72.13%	74.74%
	RF	**94.40%**	60.15%	64.60%
	Gradient Boosting	92.65%	67.97%	74.64%
AD vs CN vs non-AD dementia	LR	91.61%	**65.36%**	**69.94%**
	KNN	88.53%	62.82%	65.90%
	DT	91.68%	51.50%	51.93%
	MLP	91.69%	64.90%	69.17%
	RF	**91.89%**	55.66%	56.94%
	Gradient Boosting	91.52%	62.36%	67.82%

matrix.jpg

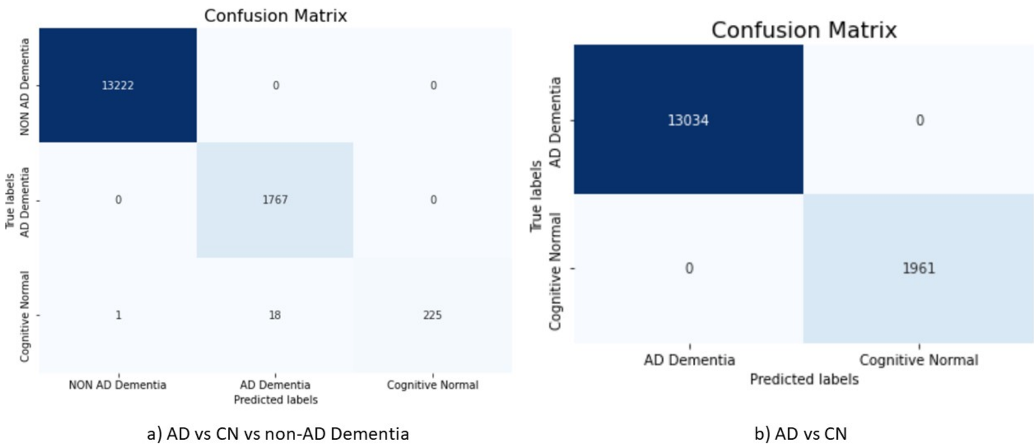


Figure 4. The confusion matrix for a) AD vs CN vs non-AD classification using Voting classifier ensemble technique and b) AD vs CN classification using Random Forest classification technique.

4. Performance Analysis

4.1. Performance Analysis using the Individual Dataset

Clinical data, Freesurfer MRI segmentation data, and psychological data are classified in two ways. One is binary classification (AD vs CN), and another is a three-class classification (AD vs CN vs non-AD dementia). With the help of the six most powerful ML classifiers, this classification has occurred.

From Table 1 we can find that the model performances are not very good using individual datasets. For AD vs CN classification, the accuracy using clinical data, Freesurfer data, and psychological data are 93.40%, 74.08%, and 76.03%, respectively. For AD vs CN vs non-AD dementia classification, the accuracy using clinical data, Freesurfer data, and psychological data are 91.89%, 65.36%, and 69.94%, respectively.

4.2. Performance Analysis using the Multimodal Dataset

After the data level fusion, we get the multimodal dataset which consists of clinical data, Freesurfer data, and psychological data. Table 2 shows the accuracy of different popular classifiers using the multimodal dataset.

Table 2. Accuracy of ML classifiers using multimodal datasets (bold values are best performing results).

Classification Type	ML Classifier	Accuracy	Precision	Recall	F1-score
AD vs CN	LR	98.67%	98.55%	98.82%	99.75%
	KNN	98.02%	98.59%	97.91%	98.74%
	DT	99.89%	99.86%	99.85%	99.90%
	MLP	98.87%	98.92%	98.76%	98.84%
	RF	**100%**	**100%**	**100%**	**100%**
	Gradient Boosting	99.34%	99.25%	99.32%	99.21%
AD vs CN vs non-AD dementia	LR	98.52%	85.96%	71.40%	72.93%
	KNN	98.89%	94.89%	86.04%	89.60%
	DT	99.80%	97.62%	**97.68%**	97.78%
	MLP	99.18%	92.09%	87.55%	89.51%
	RF	99.80%	99.44%	95.77%	97.46%
	Gradient Boosting	99.25%	95.75%	85.85%	89.46%
	Voting classifier (RF, DT, and Gradient Boosting)	**99.86%**	**99.66%**	97.40%	**98.48%**

From Table 2 it is clear that for the binary classification, which means the AD vs CN classification RF is the best performing algorithm that achieves 100% accuracy. Furthermore, in AD vs CN vs non-AD classification, DT, RF, and Gradient Boosting accuracy is 99.80%, 99.80%, and 99.25%. After creating an ensemble model using the Voting classifier, we can get the best accuracy, precision and F1-score that are 99.86%, 99.66%, and 98.48%. The performance of classifiers has increased a lot after using the multimodal dataset. Fig. 4 shows with the confusion matrix of the AD vs CN and AD vs CN vs non-AD classification models.

4.3. Explain Ability of The Model

Using the LIME model explainer, we can easily find out the most important features for any particular decision from the model.

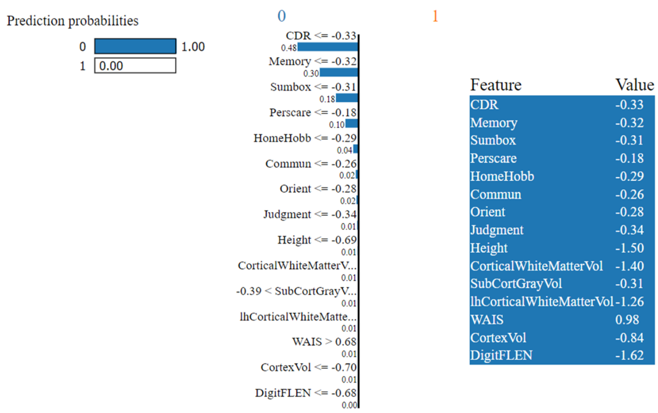
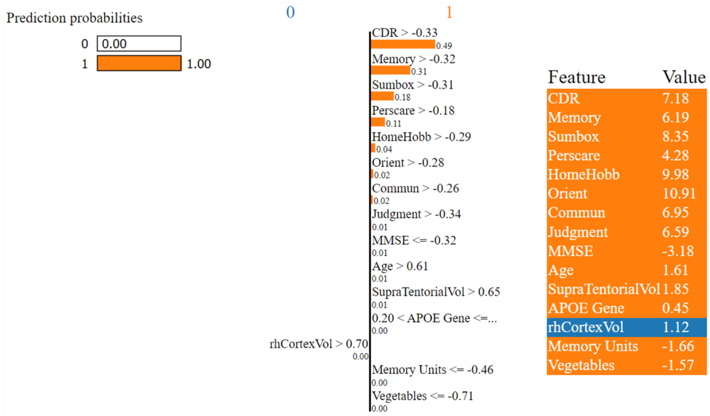


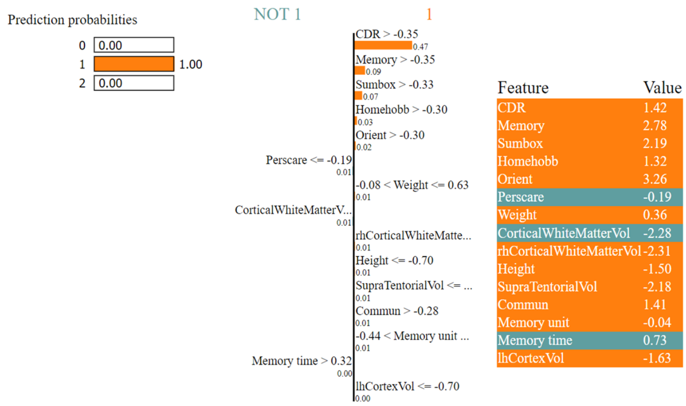
Figure 5. LIME explainer output of a Cognitive Normal (CN) participant. (Class 0 = CN)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1



AD AD vs CN.png

Figure 6. LIME explainer output of an Alzheimer's disease (AD) participant. (Class 1 = AD)



merge AD vs CN vs Non AD.png

Figure 7. LIME explainer output of an Alzheimer's disease (AD) participant for 3 class classification. (Class 1 = AD)

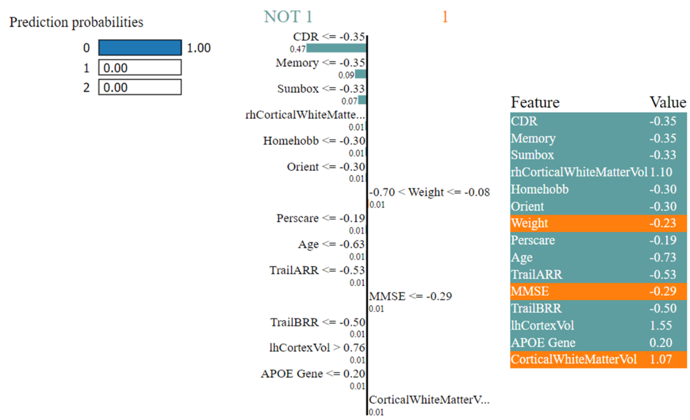
4.3.1. AD vs CN Classification

Fig. 5 and 6 explains the model's prediction with the top 15 features. Here, 0 and 1 labels are CN and AD, respectively. This figure 5 is the explanation of a random CN participant's outcome of the prediction. From this figure, we can see that the influence of CDR, memory, perscare, homehobb, commun, orient, judgement, cortical white matter volume, subcortical white matter volume, left hemisphere cortical white matter volume, WAIS, and cortex volume are 0.48, 0.30, 0.18, 0.10, 0.04, 0.02, 0.02, 0.01, 0.01, 0.01,0.01, 0.01, 0.01, and 0.01. This figure also shows the limit of making any decision. For example: if the CDR value is less than or equal to -0.33, the patient will be diagnosed as CN. We can know the value of the top 15 features for this specific participant from this figure. This explainer also indicates that better decision-making features from all three datasets are important. Similarly, Fig. 6 shows why a participant is diagnosed with AD. It is observed that, for this AD prediction, CDR, memory, sumbox, perscare, homehobb, orient, judgement, MMSE, age, and supratentorial volume are major decision making features. Rather than that, the value of the APOE gene and memory unit also indicates the tendency of AD.

4.3.2. AD vs CN vs Non-AD Dementia Classification

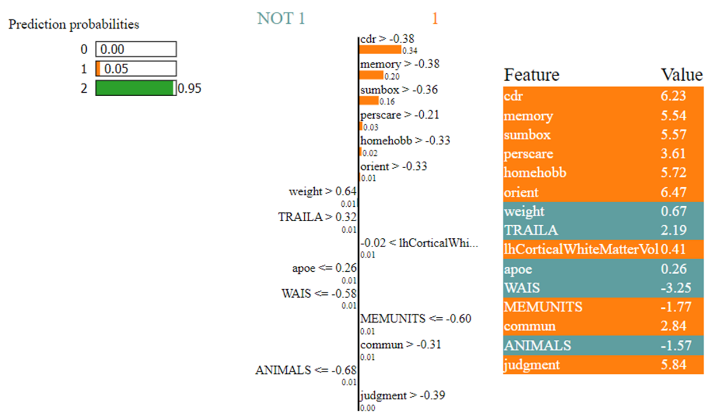
Fig. 7, 8, and 9 explains the reasons behind the model's prediction using the top 15 features, where the labels 0, 1, and 2 represent CN, AD, and non-AD dementia, respectively. Fig. 7 shows that the decision-making features for AD are CDR, memory, Sumbox, homehobb, orient, weight, right hemisphere cortical white matter volume, height, supratentorial volume, commun, memory unit. The values of those features are also mentioned in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1



merge AD vs CN vs Non AD.png

Figure 8. LIME explainer output of a cognitively normal (CN) participant for 3 class classification. (Class 0 = CN)



non ad AD vs CN vs Non AD.png

Figure 9. LIME explainer output of a non-AD dementia participant for 3 class classification. (Class 2 = non-AD dementia)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

Table 3. Performance comparison between RF based AD vs CN prediction works (bold values are best performing results).

Author	Data Modality	Accuracy	Sensitivity	Specificity	AUC
Jain et al. [24]	Oasis-3 (MRI)	87.72%	94.44%	87.07%	-
Kwak et al. [25]	OASIS-3 (MRI)	-	-	-	88.7%
Our proposed work	**OASIS-3 (Clinical, Psychological, and Freesurfer data)**	**100%**	**100%**	**100%**	**100%**

Table 4. Comparison between MRI data and proposed multimodal data (bold values are best performing results).

Classification Type	Model	Data Modality	Accuracy
AD vs CN	RF	MRI data	87.62%
		Proposed Multimodal data	**100%**
AD vs CN vs non-AD dementia	Voting Classifier (RF, DT, and Gradient Boosting)	MRI data	79.31%
		Proposed Multimodal data	**99.86%**

picture. Similarly, Fig. 8 and 9 depict the key features for CN and non-AD dementia participants.

4.4. Performance Comparison

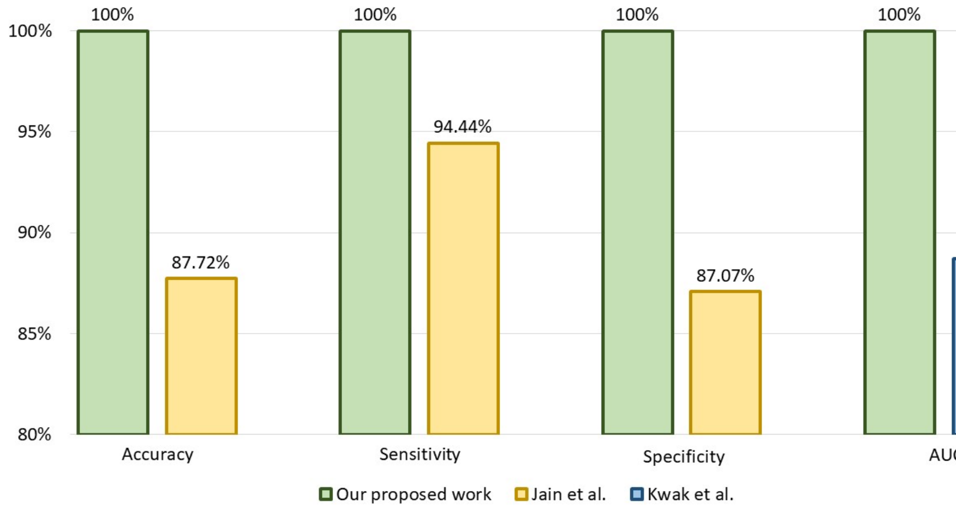
As we have proposed a novel multimodal approach, we could not get exact similar works to compare with. Most of the recent works used MRI data for prediction. For comparison we have used other recently published (2021) works using RF models. Here, the RF model is trained using MRI data collected from the OASIS-3 dataset. From the comparison Table 3, the performance of our work is far better than others. After viewing Fig. 10 we can also state that the accuracy, sensitivity, specificity, and AUC of our work is better than others. So, we can state, instead of using only neuroimaging data (MRI data), we can use multimodal data (Clinical data, psychological assessment data, and Freesurfer data) to increase the model's performance.

Because most recent studies have used MRI datasets, we have also trained our models using MRI datasets from the OASIS-3 dataset to ensure a fair comparison. From the Table 4, for binary classification, the RF classifier's accuracies using MRI and proposed multimodal data are found to be 87.62% and 100%, respectively. Besides, for three-class classification using MRI and proposed multimodal data, the accuracies of the Voting classifier (RF, DT, and Gradient Boosting) model are 78.59% and 99.86%, respectively. As a result, we can conclude that our proposed multimodal approach outperforms single modal data (MRI).

5. Conclusion

Considering the increasing number, dangers, and risks of Alzheimer's, this study proposed a multimodal approach for predicting AD. This multimodal approach consists of data-level fusion on three datasets from three different genres. We have used ADRC clinical data, Freesurfer MRI segmentation numeric data, and psychological data from the OASIS-3 repository. This multimodal approach provided 100% accuracy for AD vs CN classification using RF classifier and 99.86% accuracy for AD vs CN vs non-AD dementia using a voting classifier-based ensemble technique. With the help of RF, DT, and Gradient Boosting algorithm, this custom ensemble model is created. For achieving the trustworthiness of the predictions of this model, the LIME explainer is used here, and all the decision-making features along their values are displayed. From the outcome of the explainer, it is clear that

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1



comparison.jpg

Figure 10. AD vs CN Performance comparison between various Random Forest (RF) models using OASIS-3 dataset

CDR, memory, and Sumbox are the three most important features. Even for perfect decision making, features of all three individual datasets played an important role. This work also provided an entire architecture for AD patient management and 24/7 monitoring. In the future, we will develop an efficient model that can give early AD prediction and predict different AD stages. Even we will implement this proposed management system in real life and evaluate its performance.

Acknowledgments: This work was supported by the National Research Foundation of Korea (NRF) Grant by the Korean Government through the Ministry of Science and ICT (MSIT) under Grant 2021R1A2C2014333.
This work was supported by the research funds of the Information and Communication Technology division of the Government of the People's Republic of Bangladesh in 2020 - 2021.

Data Availability Statement: Data is available at https://www.oasis-brains.org/.

References

1. A Armstrong, R. Risk factors for Alzheimer's disease. *Folia Neuropathologica* **2019**, *57*, 87–105. doi:10.5114/fn.2019.85929.
2. Kukull, W.A.; Brenner, D.E.; Speck, C.E.; Nochlin, D.; Bowen, J.; McCormick, W.; Teri, L.; Pfanschmidt, M.L.; Larson, E.B. Causes of death associated with Alzheimer disease: variation by level of cognitive impairment before death. *Journal of the American Geriatrics Society* **1994**, *42*, 723–726.
3. Kamal, M.S.; Northcote, A.; Chowdhury, L.; Dey, N.; Crespo, R.G.; Herrera-Viedma, E. Alzheimer's Patient Analysis Using Image and Gene Expression Data and Explainable-AI to Present Associated Genes. *IEEE Transactions on Instrumentation and Measurement* **2021**, *70*, 1–7. Conference Name: IEEE Transactions on Instrumentation and Measurement.
4. El-Sappagh, S.; Alonso, J.M.; Islam, S.M.R.; Sultan, A.M.; Kwak, K.S. A multilayer multimodal detection and prediction model based on explainable artificial intelligence for Alzheimer's disease. *Scientific Reports* **2021**, *11*, 2660. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Classification and taxonomy;Computational neuroscience;Data mining;Data processing;Machine learning Subject_term_id: classification-and-taxonomy;computational-neuroscience;data-mining;data-processing;machine-learning.
5. Lee, G.; Nho, K.; Kang, B.; Sohn, K.A.; Kim, D. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Scientific Reports* **2019**, *9*, 1952. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Data integration;Machine learning Subject_term_id: data-integration;machine-learning.
6. Zhang, D.; Shen, D.; Alzheimer's Disease Neuroimaging Initiative. Multi-modal multi-task learning for joint prediction of multiple regression and classification variables in Alzheimer's disease. *NeuroImage* **2012**, *59*, 895–907.
7. Buvari, S.; Pettersson, K. *A Comparison on Image, Numerical and Hybrid based Deep Learning for Computer-aided AD Diagnostics*; 2020.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 March 2022 doi:10.20944/preprints202203.0214.v1

8. Kumari, R.; Nigam, A.; Pushkar, S. MULTI-LAYERED CONVOLUTIONAL NEURAL NETWORK ON MULTI- MODAL IMAGES FOR CLASSIFYING ALZHEIMER DISEASE. p. 6.
9. LaMontagne, P.J.; Benzinger, T.L.; Morris, J.C.; Keefe, S.; Hornbeck, R.; Xiong, C.; Grant, E.; Hassenstab, J.; Moulder, K.; Vlassenko, A.G.; et al. OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease. Technical report, 2019. Company: Cold Spring Harbor Laboratory Press Distributor: Cold Spring Harbor Laboratory Press ISSN: 1901-4902 Label: Cold Spring Harbor Laboratory Press Type: article.
10. Palatnik de Sousa, I.; Vellasco, M.; Silva, E. Local Interpretable Model-Agnostic Explanations for Classification of Lymph Node Metastases. *Sensors* **2019**, *19*, 2969.
11. Niamat Ullah Akhund, T.M.; Mahi, M.J.N.; Hasnat Tanvir, A.N.M.; Mahmud, M.; Kaiser, M.S. ADEPTNESS: Alzheimer's Disease Patient Management System Using Pervasive Sensors - Early Prototype and Preliminary Results. Brain Informatics; Wang, S.; Yamamoto, V.; Su, J.; Yang, Y.; Jones, E.; Iasemidis, L.; Mitchell, T., Eds.; Springer International Publishing: Cham, 2018; Lecture Notes in Computer Science, pp. 413–422.
12. Mohapatra, P.; Preejith, S.P.; Sivaprakasam, M. A novel sensor for wrist based optical heart rate monitor. 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2017, pp. 1–6.
13. Du, J.; Gerdtman, C.; Lindén, M. Signal Quality Improvement Algorithms for MEMS Gyroscope-Based Human Motion Analysis Systems: A Systematic Review. *Sensors* **2018**, *18*, 1123. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
14. Nahian, M.J.A.; Ghosh, T.; Banna, M.H.A.; Aseeri, M.A.; Uddin, M.N.; Ahmed, M.R.; Mahmud, M.; Kaiser, M.S. Towards an Accelerometer-Based Elderly Fall Detection System Using Cross-Disciplinary Time Series Features. *IEEE Access* **2021**, *9*, 39413–39431. Conference Name: IEEE Access.
15. Kim, J.; Chou, E.F.; Le, J.; Wong, S.; Chu, M.; Khine, M. Soft Wearable Pressure Sensors for Beat-to-Beat Blood Pressure Monitoring. *Advanced Healthcare Materials* **2019**, *8*, 1900109. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/adhm.201900109.
16. Yadav, J.; Rani, A.; Singh, V.; Murari, B.M. Near-infrared LED based non-invasive blood glucose sensor. 2014 International Conference on Signal Processing and Integrated Networks (SPIN), 2014, pp. 591–594.
17. Lakshmi, K.J.; Muneshwar, A.; Ratnam, A.V.; Kodali, P. Patient Assistance using Flex Sensor. 2020 International Conference on Communication and Signal Processing (ICCSP), 2020, pp. 00181–00185.
18. Kaiser, M.S.; Al Mamun, S.; Mahmud, M.; Tania, M.H. Healthcare Robots to Combat COVID-19. In *COVID-19: Prediction, Decision-Making, and its Impacts*; Santosh, K.; Joshi, A., Eds.; Lecture Notes on Data Engineering and Communications Technologies, Springer: Singapore, 2021; pp. 83–97.
19. Jesmin, S.; Kaiser, M.S.; Mahmud, M. Artificial and Internet of Healthcare Things Based Alzheimer Care During COVID 19. Brain Informatics; Mahmud, M.; Vassanelli, S.; Kaiser, M.S.; Zhong, N., Eds.; Springer International Publishing: Cham, 2020; Lecture Notes in Computer Science, pp. 263–274.
20. Surendran, D.; Janet, J.; Prabha, D.; Anisha, E. A Study on devices for assisting Alzheimer patients. 2018 2nd International Conference on 2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018, pp. 620–625.
21. Bayat, S.; Babulal, G.M.; Schindler, S.E.; Fagan, A.M.; Morris, J.C.; Mihailidis, A.; Roe, C.M. GPS driving: a digital biomarker for preclinical Alzheimer disease. *Alzheimer's Research & Therapy* **2021**, *13*, 115.
22. Ahmed, Q.A.; Al-Neami, A.Q.H. A Smart Biomedical Assisted System for Alzheimer Patients. *IOP Conference Series: Materials Science and Engineering* **2020**, *881*, 012110. Publisher: IOP Publishing.
23. Asif-Ur-Rahman, M.; Afsana, F.; Mahmud, M.; Kaiser, M.S.; Ahmed, M.R.; Kaiwartya, O.; James-Taylor, A. Toward a Heterogeneous Mist, Fog, and Cloud-Based Framework for the Internet of Healthcare Things. *IEEE Internet of Things Journal* **2019**, *6*, 4049–4062. Conference Name: IEEE Internet of Things Journal.
24. Jain, M.; Rai, C.S.; Jain, J. A Novel Method for Differential Prognosis of Brain Degenerative Diseases Using Radiomics-Based Textural Analysis and Ensemble Learning Classifiers. *Computational and Mathematical Methods in Medicine* **2021**, *2021*, e7965677. Publisher: Hindawi.
25. Kwak, K.; Giovanello, K.S.; Bozoki, A.; Styner, M.; Dayan, E. Subtyping of mild cognitive impairment using a deep learning model based on brain atrophy patterns. *Cell Reports Medicine* **2021**, *2*, 100467.