A Cascaded Mutliresolution Ensemble Deep Learning Framework for Large Scale Alzheimer's Disease Detection Using Brain MRIs

Imran Razzak, Saeeda Naz, Hamid Alinejad-Rokny, Tu N. Nguyen, and Fahmi Khalifa

Abstract—Alzheimer's is progressive and irreversible type of dementia, which causes degeneration and death of cells and their connections in the brain. AD worsens over time and greatly impacts patients' life and affects their important mental functions, including thinking, the ability to carry on a conversation, and judgment and response to environment. Clinically, there is no single test to effectively diagnose Alzheimer disease. However, computed tomography (CT) and magnetic resonance imaging (MRI) scans can be used to help in AD diagnosis by observing critical changes in the size of different brain areas, typically parietal and temporal lobes areas. In this work, an integrative mulitresolutional ensemble deep learning-based framework is proposed to achieve better predictive performance for the diagnosis of Alzheimer disease. Unlike ResNet, DenseNet and their variants proposed pipeline utilizes PartialNet in a hierarchical design tailored to AD detection using brain MRIs. The advantage of the proposed analysis system is that PartialNet diversified the depth and deep supervision. Additionally, it also incorporates the properties of identity mappings which makes it powerful in better learning due to feature reuse. Besides, the proposed ensemble PartialNet is better in vanishing gradient, diminishing forward-flow with low number of parameters and better training time in comparison to its counter network. The proposed analysis pipeline has been tested and evaluated on benchmark ADNI dataset collected from 379 subjects patients. Quantitative validation of the obtained results documented our framework's capability, outperforming state-of-the-art learning approaches for both multi-and binary-class AD detection.

Index Terms—Ensemble PartialNet, alzheimer disorder, dementia, early diagnosis

1 Introduction

Alzheimer's is an irreversible type of dementia, or a brain disorder, that generally affects elder people. The majority of people with Alzheimer's disease (AD) are 65 and older, which greatly impact the lives of the elderly. Recent statistics showed that AD is considered as the 4th leading cause of death followed by cardiovascular disease, cancer, and stroke worldwide [1], [2]. Clinically, AD is a chronic neuro-degenerative disorder caused by abnormal accumulation of neurofibrillary tangles and amyloid plaques in the brains of Alzheimer's paitent. This gradually results in decline and deterioration of patient's cognition, mental and memory, often referred to as senile dementia, which eventually lead to severe personality changes (e.g., writing, speaking, reading, the ability to carry out the simplest tasks, and other abnormal behaviour symptoms. This not only imposes bad impacts on the daily-life tasks of patients, but also demonstrates difficulties in the identification of their family members.

AD progress through three stages: early, moderate cognitive stage and late (see Fig. 1). Clinically those stages are identified using a set of biomarkers, see Fig. 2. Early diagnosis of AD is referred as early AD onset. In 2022, about 6.5 million Americans age 65 and older are living with AD and 55 worldwide with approximately 5% have early onset AD or the so-called "younger onset" [4]. At this stage, mild memory, difficulties in planning, problem solving or finishing daily tasks are the most common symptoms. Patients in the moderate cognitive stage act very aggressively, but patients in the late stage suffer from heart failure and death-causing respiratory system dysfunction [5], [6]. Like other diseases (e.g., Parkinson's) AD has no cure. Nonetheless, there are some treatments that can effectively change disease progression by temporarily slowing the worsening of the symptoms and thus improve patients' quality of life (See Table 1). Thus, over the coming decades, AD will be a global challenge due to the growing life expectancy of people.

Generally speaking, Alzheimer is a progressive disease process from the onset of pathological changes to the appearance of clinical symptoms. The AD diagnosis typically include various testing methods, such as biomarker testing, neuropsychological testing, and neuroimaging testing. The definitive approach for AD diagnose is the postmortem examination of the brain tissue. Nonetheless, various clinical procedures and/or tests can be utilized to evaluate patients' mental abilities, diagnose dementia, and to exclude other conditions. Furthermore, specialized brain-scans, such as

- *Imran Razzak and Hamid Alinejad-Rokny are with the University of New South Wales, Sydney, NSW 2052, Australia. E-mail: imran.razzak@ieee.org, h.alinejad@unsw.edu.au.*
- *Saeeda Naz is with GPGC No. 1, Higher Education Department, Abbottabad 22010, Pakistan. E-mail: saeedanaz292@gmail.com.*
- *Tu N. Nguyen is with the Department of Computer Science, Kennesaw State University, Marietta, GA 30060 USA. E-mail: tu.nguyen@kennesaw.edu.*
- *Fahmi Khalifa is with the Department of Electrical and Computer Engineering, Morgan State University, Baltimore, MD 21251 USA. E-mail: fahmi.khalifa@morgan.edu.*

Manuscript received 26 December 2021; revised 29 July 2022; accepted 28 October 2022. Date of publication 2 November 2022; date of current version 8 August 2024.
(Corresponding author: Imran Razzak.)
Digital Object Identifier no. 10.1109/TCBB.2022.3219032

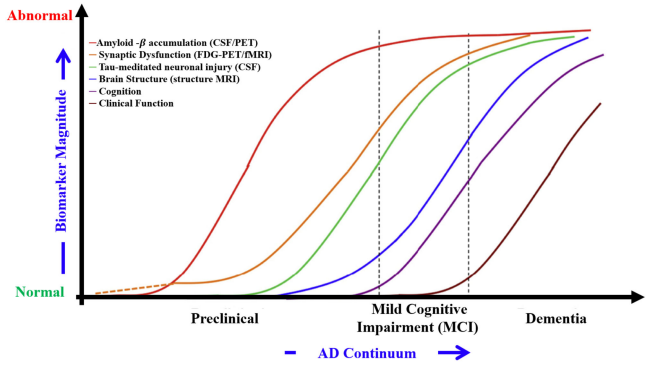

Fig. 1. Graphical depiction of bio-markers as indicators of AD [3].

resting-state functional MRI (rs-fMRI), structural magnetic resonance imaging (sMRI), computed tomography (CT) and positron emission tomography (PET) are widely explored for AD diagnosis as well as for follow-up of AD progression [8], [9], [10], [11], [12]. Those MR brain images can can depict/capture brain structural abnormalities and abnormal functions, see example in Fig. 3. Noninvasive neurodegenerative disease identification using brain images have shown an exponential growth in recent years. Namely, computer-aided diagnosis (CAD) systems utilizing brain images have been extensively used as a supplement to improve brain disorder diagnosis [13], [14], [15]. This is coupled with recent advances in machine learning (ML) and artificial intelligence (AI) tools and data fusion to develop streamlined workflows for improved diagnosis of Alzheimer and its prodromal dementia stage, mild cognitive impairment (MCI), from the normal controls (NC). In literature, various ML techniques and methods have been developed to identify neurodegenerative diseases, AD, based on brain imaging data diseases. Particularly, deep learning (DL) methods, i.e., convolutional neural networks (CNN), have been widely adopted in the field of neurodegenerative diseases diagnosis [16], [17]. CNN has been found to be the most widely used and represented technique in most research work when it comes to brain image processing and analysis. Practically, CNN uses 2D/3D brain images as the input and automatically learn meaningful higher-level local and global features, thus eliminating multiple measurement errors induced by the conventional hand-crafted feature.

In general, the similarity of MR brain images between AD and normal controls complicates the detection of AD. Although several research work in MRI-based AD diagnosing exist, they focused on improving and altering multiple network architectures or ensembled CNN models in order to provide high-accuracy AD diagnosis. Accordingly, In this paper,

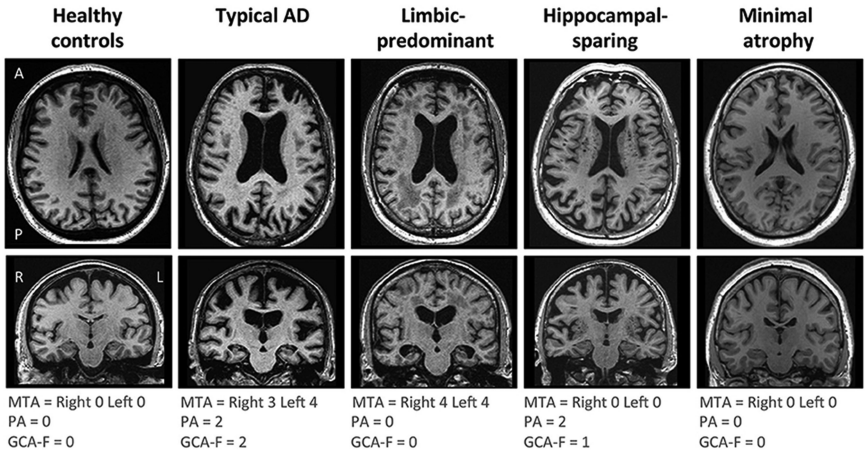

Fig. 2. Visual examples: Pattern of brain atrophy in different Alzheimer subtypes [7].

TABLE 1
Determining Alzheimer's Symptoms

Early	Mild	Late
Alzheimer's symptoms are mild and characterized by general forgetfulness Symptoms	Alzheimer's symptoms are more disabling and additional care my be required	Alzheimer's symptoms are significant and apparent
Forgets recently read material	Delusions, compulsions or repetitive behaviour	Significant personality and behaviour change
Trouble organizing/planning	Agitation, restlessness and anxiety	Loss of ability to hold conservatio0n
Forgets where valueables have been placed	Need assistance with getting dressed	Difficulty in moving and eating
Trouble in managing money	Bladder and bowel issues	Loss of Bladder and bowl control
Forgets recent events, names and about own identity	Trouble learning new things	Lack of awarness of recent activities
Trouble with challenging task at work	Problem with reading and writing	Highly suscepitible to infections
Wanders and becomes lost in familiar places	Sleep pattern disturbance	Unable to do daily activities

we present an integrative deep framework that possess ability to provide robust and accurate AD diagnosis by highlighting the structural similarity of AD image classes while maximizing the variance between classes. Particularly, the proposed framework is based on multi-level structure that incorporates a deep multi-resolutional ensemble PartialNet. Identity mappings and diversified depth properties are maintained at each of the individual levels. For better feature representation, the proposed architecture, integrates additional blocks, namely deep supervision and transition blocks. Further improvement of the gradient propagation as well as information flow, is achieved by utilizing partial connectivity pattern by connecting some (but not all) of its subsequent layers.

The *major contributions of our work is five- fold*: (i) an efficient feature reuse for better learning by employing a PartialNet architecture that limits the connection to its preceding layers, compared with DenseNet that utilize each layer in a CNN to all preceding layers, besides, we have created ensemble bottleneck to force the network to learn different level features (lower, intermediate and high level); (ii) A multi-resolution ensemble is employed with block concept to limit the skip connection within a block. This makes each block behaves like a unique structure which makes it powerful to contribute more to gradient magnitude compared to its counter deep networks; (iii) To help in intermediate features learning, the proposed architecture incorporates both supervision as well as transition layers in each PartialNet block; (iv) An enhancement of the gradient flow gradient flow (in comparison to partially-dense path) by bounding the paths in each block as well as creating bottleneck; and (iv) Considerable gain in performance is achieved when compared with benchmark AD diagnosis methods for both binary, or two-level (i.e., AD-vs-MCI, MCI-vs-CN, AD-vs-CN) and multi-class (CN, MCI, and AD) classifications using AD neuroimaging initiative (ADNI) dataset.

This paper is organized as follows. The literature's related work is provided and discussed in Section 2. Section 3.2 provides the detailed description of the proposed analysis pipeline. This is followed by the experimental results, system evaluation, and the comparison with state-of-the-art methods in Section 4. In Section 5, a detailed discussion of the presented results, observations and limitations are provided. Finally, Section 6 provides concluding remarks and suggested future directions.

2 Related Work

The diagnosis of AD at early stage is extremely important for possible treatment to delay disease progression to advanced stages. In literature, many research work has been developed and introduced in recent years in an effort to provide early-stage diagnostic tools of AD This has been supported by the exponential advances in ML techniques, specially those based on deep learning (DL) algorithms. In this section, we will overview different ML and DL-based algorithms and methods introduced by the research community for AD diagnosis and classification employed in this direction.

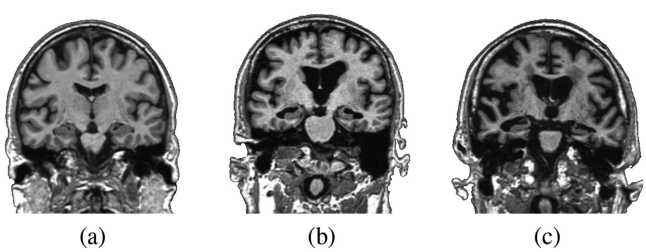

Fig. 3. Examples of sMRI showing brain atrophy patterns for (a) normal cognitive (CN); (b) mild cognitive impairment (MCI), and (c) an Alzheimer's disease (AD) subject. Adopted from [18].

In the literature, various deep learning based algorithms have been proposed recently to recognize and detect AD. Particularly, A hybrid multi-class framework by Bhatkoti and Paul [19] was introduced for early diagnosis of AD using a modified k-sparse autoencoder (KSA) classification to locate neutrally degenerated areas of the brain. Experiments were conducted using 150 images for MRI scans as well as CSF and PET images from ADNI study. The reported results showed that modified KSA outperformed zero-masking strategy and traditional KSA in terms of overall accuracy 83.143% accuracy for boosted pair predictors with 100 classifier combinations compared to 71.327% for 50 classifier combinations. Saman et al. [20] introduced a DL framework for AD diagnosis using 43 patient cohort (28 AD and 15 NC) collected from ADNI dataset. Extensive analysis steps including skull stripping, motion correction, spatial smoothing, image registration and noise removal are applied as preprocessing. Then, a DL model, called Le-Net, is fed with the processed data and achieved a 96.85% classification accuracy. A similar framework by Mathew et al. [21] processed a 151 subset of ADNI MR images (71 NC and 87 AD) for AD diagnosis. They applied several pre-processing stages, including data cropping and resize, intensity normalization, and reorientation. Feature extraction is conducted using principal component analysis (PCA) and discrete Wavelet transform (DWT) and a support vector machine (SVM) is used classification. Their pipeline achieved 91% (84%) accuracy for MCI versus CN (AD versus CN). Another approach by Iftikhar et al. [22] introduced an ensemble-based pipeline to classify MCI and AD patient cohort. The ensemble classifier was fed by both volumetric cortex and cortical thickness features. The framework analyzed a total of 180 patients (60 NC, 60 AD, and 60 MCI) and achieved specificity, sensitivity, and classification accuracy of 91.66%, 92% and 89% respectively. A deep 3D CNN for AD diagnosis by Asl et al. [23] employed convolutional auto encoder (CAE) for local feature extraction from 3D skull-stripped and spatially normalized brain data. Then, fine-tuned transfer learning is utilized to identify AD patient on ANDI from CAD Dementia data set. The pipeline evaluation were carried out on CAD Dementia data set using a 210 patient cohort: 70 for each CN, MCI and AD class. The study documented that the fine tune technique demonstrated significantly better performance and achieved a CN versus AD classification with a 97.6% accuracy. Farooq et al. [24] examined different DL methods for multiclass AD classification on ADNI data sets. Particularly, ResNet-152 and ResNet-18, and GoogLeNet architectures are used to analyze a four-cohort patient group containing 22, 33, 45, and 449 cases for LMCI, AD, CN, and MCI, respectively. Experimental evaluation showed that GoogLeNet showed marginal performance (98.8% accuracy)

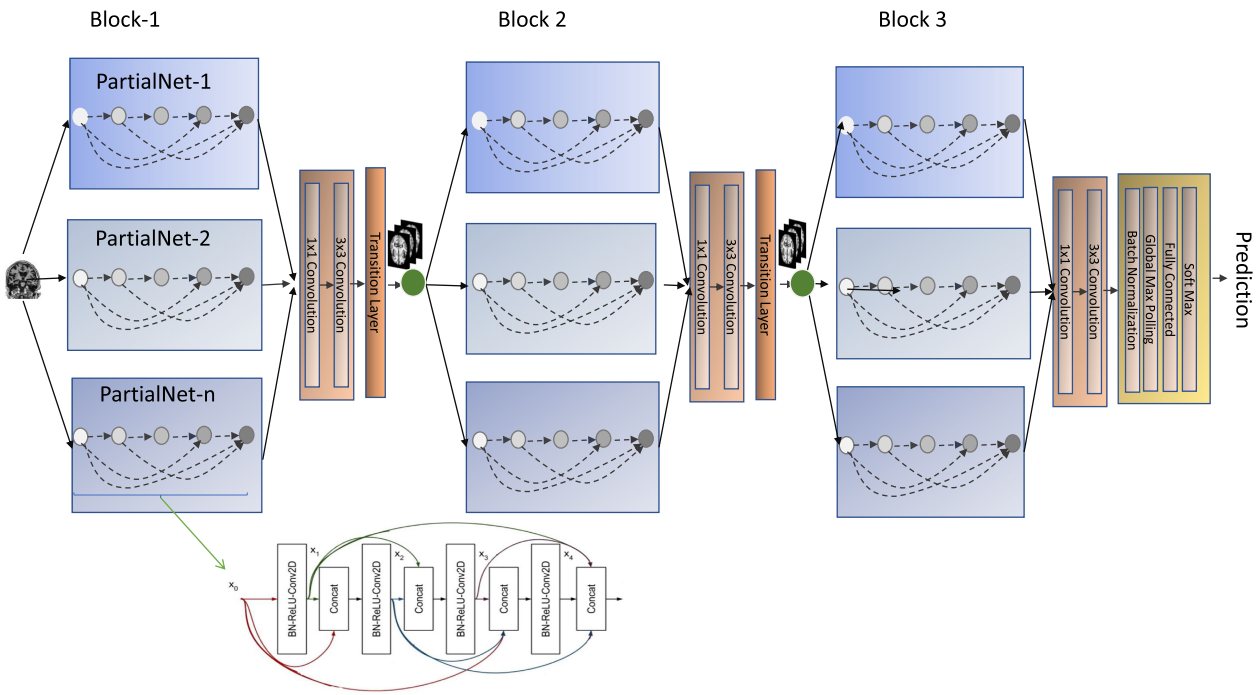

Fig. 4. Schematic diagram of the proposed ensemble pipeline, which consists of three blocks network of and n number of PartialNet.

compared with ResNet-18 and ResNet-152 (98.01%, and 98.14% respectively). Using brain MRIs, a simple and effective method by Backstorm et al. [25] was proposed to detect AD, which is called 3D ConvNet architecture. Their analysis contained a series of preprocessing steps, such as cortical reconstruction, edge trimming, data resizing, and intensity normalization. Followed by an automated features extraction from the prepossessed images using the 3D ConvNet. A total of 340 experimental data was gathered from ADNI dataset, which included 1190 MRI scans of 199 AD patients (96 female and 103 male) and 141 NC (66 female and 75 male) and achieved 98.78% AD accuracy. An fMRI-based framework was introduced by Kazemi and Houghten [26] to differentiate between AD patients at different disease stages. The analysis started with preprocessing of the images including brain extraction, spatial normalization, slice timing correction, spatial smoothing, and high-pass filtering. Finally, AlexNet architecture is utilized for classification using a total of 197 subjects to differentiate between five-classes patients: CN, AD, late MCI (LMCI), early MCI (EMCI) and subjective memory complaints (SMC). A similar data split scenario to [25] was employed during experiments and the system achieved an overall accuracy of 97.63%. Moreover, the perclass accuracies for EMCI, LMCI, AD, CN and SMC were 95.64%, 95.89%, 94.97%, 98.34% and 94.55%, respectively. A hybrid transfer leaning-based technique by Ghahnavieh et al. [27] was proposed to detect AD using MRIs from collected from ADNI dataset. Their method integrated both CNN and recurrent neural network (RNN) to better consider the association between the MRI slices. The CNN is used as a feature extractor while the RNN is utilized to infer the relationship between slices, which showed an overall improvement of the diagnostic performance. Recently, Zhang et al. [1] developed a DL-based classification framework using MR and PET brain images. Their strategy included stacked autoencoder (SAE), a regression layer of softmax, and minimal labeled learning samples. The framework is tested on a cohort of 311 ADNI neuroimaging subjects (65 AD, 67 converter MCI, 102 non-converters MCI (ncMCI), and 77 NC). The results documented a 88.58% and 47.42% accuracies for binary for 4-class classification, respectively. A hybrid multimodal approach by Ju et al. [28] for diagnosing Alzheimer patient integrated both textual and imaging (MRI) data. Textual information include age, gender, and genomics data and are used to find prevalence between MCI and age, gender, and ApoE. The analysis is conducted using a patient cohort from ADNI-2 dataset (91 MCI and 79 NC). Preprocessing of the data is performed using the method in [29]. A similar study using fMRI data and correlation coefficient (CC) using various ML algorithms (e.g., SVM, KR, LDR, and AE) demonstrated a performance gain for accuracy, sensitivity, andspecificity. Particularly, 86.47%, /92%, /81% using SVM. The results showed that CC could be used to enhance the diagnostic performance. More recently, Ashraf et al. [30] investigated multiple pre-trained CNNs backbones for the classification of CN, MCI and AD. The performance of thirteen well-known architectures are assessed and compared using augmented data of ADNI dataset. Experimental results reported that the DensNet architectures demonstrated the highest accuracy (99.05%) using 5-fold cross validation scenario. In this study, the the authors investigated the freeze features of multiple CNN architectures [31].

3 Proposed Alzheimer Diagnosis Framework

A schematized diagram in Fig. 4 shows the proposed framework for AD detection using MRI. The pipeline is an integrative DL architecture that integrates multi-resolutional ensemble PartialNet, which is incorporate to enhance the predictive performance for AD diagnosis. The input or DL-based ensemble architecture is a subset of ADNI MRI data

sets. Our framework include preprocessing, multi-resolutional features extraction, ensemble learning, and classification. Section 3.2 briefly describes the components of deep partial Net architecture. The developed system has the capability to provides both binary and multi-class classifications. Before any data analysis is performed, the input images are preprocessed to help in targeting the region-of-interest (ROI). This is followed by the actual stage of multi-resolution features extraction and classification. Next, a full description of all the stages of preprocessing sand deep multi-resolutional ensemble learning are detailed.

3.1 Prepossessing

Before any data analysis, we applied different preprocessing techniques to further enhance the input data and improve the performance. First, the raw MR images are converted into 1-channel image of different sizes followed by resizing and and cropping of an image in order to discard the white regions and thus enhance the quality of extracted image. Furthermore, ROI extraction is performed which determines the extreme points in the contours along with x, y-coordinates. Beside, several data augmentations are performed to reduce the possibility of overfitting during training. Popular augmentation including, illumination, zoom-in and zoom-out, horizontal and vertical flipping, rotation (90^0, 180^0 and 270^0) are employed. In total, the reprocessing stage results in significant increase in dataset size to 37,590.

3.2 Cascaded Partial Deep Network

In general, the deeper the DenseNet, i.e., the increased dense block, the more exponential increase in computational. This also results in large number of network parameters that need complex optimization. Compared to the DenseNet architecture that utilize each layer in CNN to all preceding layers, a deep multi-resolutional PartialNet is proposed. The developed architecture incorporates identity mapping in a way that it does not impact the performance. Besides, it also diversify the depth and deep supervision as well; thus, it implement reuse of features, which in turns produces better representation for the input. To further improve the performance and converge to the expected output, we limit the structure of proposed framework to blocks assisted with supervision block and transition block followed by ensemble of PartialNet unit in each block, which helps the PartialNet to not only learn high level or intermediate level features but also low level features.

Furthermore, depth (layers) does not contribute to the propagation of gradient and act like ensembling of the same network. Beside to DenseNet and ResNet, PartialNet consist of path with different length, as a result PartialNet contributes more towards the gradient. To further deal with challenges mentioned earlier, in this work, the concept of block level ensembling is added that limit the skip connection within PartialNet by forcing it to block level i.e., we have utilized mulit-resolution policy. Thus, every block in our network behaves like a unique structure and contribute more to the gradient magnitude than deeper networks. Moreover, length of residual path plays important role to magnitude of gradient. As we have limited the skip path in each block, thus, we have better gradient flow in comparison to partially DenseNet and ResNet. PartialNet is based on DenseNet; thus PartialNet is aided with reuse of features, which at the end able to learn better feature representation. Fig. 4 schematizes the proposed analysis pipeline. As readily seen, partially connectivity within each block and bottleneck at each block improve the flow of information. Moreover, such structure alleviates the vanishing gradient problem.

Our multi-resolutional ensemble architecture consists of three networks (Fig. 4) each of those networks consists of four blocks. Individual blocks consist of three consecutive layers: a partially connected dense layer, a supervision layer, and a transition layer. Finally, we concatenated the features from each block to create the bottleneck, which enforces the network to learn better features. Compared with densely connected layer, partial connectivity is introduced instead. As a result, the proposed architecture tangible impact will be four fold

- network parameter reduction
- better feature learning
- reducing over fitting
- improving information flow.

The latter is achieved by directly connecting some of subsequent layers in the block and finally concatenating the feature map after each block by creating bottleneck, i.e., feature map received at l^{th} layer from block B_i simply receive the information from the earlier layers $l - 1/2$ from same block (Fig. 4). Finally, features form each Partialnet are concatenated as a output of block for input to next block.

$$P_l^j = H_l^j([\mathbf{P_0^j}, \mathbf{P_2^j}, \ldots, \mathbf{P_{1-2}^j}]). \tag{1}$$

Here "++" denotes the concatenation of different PartialeNet (P^j), H_l^j represents composite non-liner transformation which consist of batch normalization, ReLU (leaky rectified linear unit) and convolution $3 \times 3 \times 3$. As discussed earlier, we have created block of smaller size, where in each block. each layer (l_i) receives the information (features) from its preceding layers (l_n). $[x_0, x_1, \hat{a}, x_{l-1}]$ represents the connectivity of all previous layers ($l - 1$) to single tensor.

Unlike DenseNet and its varaints, the PartialNet block consist of partial connectivity of earlier layer. The layers in PartialNet consist of set of $1 \times 1 \times 1$ and $3 \times 3 \times 3$ convolution followed by bottleneck of different PartialNet though supervision and transition layers that consist of 1×1 and 3×3 convolution for supervision and transition block of BN, 1×1 convolution and 2×2 pooling which filters out the information as well as forces the PartialNet to learn intermediate features as well as low level features. Every PartialNet component of multi-path block J process the information with different depth and scale. We have added the bottleneck and concatenated the features at block level.

$$P_i = \mathbf{P}_1^j + +\mathbf{x}_2^j + + \cdots \ldots + +\mathbf{P}_n^j \tag{2}$$

$$P = P_0 + +P_1 + +P_2 + + \ldots + +P_n. \tag{3}$$

Where P is the concatenated features from each block, P_i^j is the output of jth PartialNet, and n is the total number of ensemble PartialNets in each block.

In tradition DenseNet, a layer takes input from all its earlier layers, which results in network multiplicity to $2nk$,

TABLE 2
Study Dataset Details

Class	Number of Cases/M/F)	# Scans
Alzheimer's disease (AD)	95/44 m/51f	1,425
Mild cognitive impairment (MCI)	138/66 m/72f	1,021
Cognitive normal (CN)	146/87 m/59f	1,479

		Multiresolution Architecture		
	Class	DenseNet	PartialNet	Ensemble PartialNet
MCI versus AD versus CN	AD	96.77	97.87	98.06
	MCI	97.43	98.74	99.13
	CN	94.43	97.31	98.01
	Overall	96.72	98.23	98.65
CV versus AD	AD	76.43	85.11	85.44
	CN	83.22	88.65	88.57
	Overall	80.11	87.11	87.92
MCI versus AD	AD	95.98	98.43	99.29
	MCI	97.23	99.87	99.86
	Overall	96.65	99.26	99.47
MCI versus CN	CN	95.23	100	100
	MCI	97.43	100	100
	Overall	96.77	100	100

"M" and "F" stand for male and female, respectively
MCI: mild cognitive impairment; AD: Alzheimer's disease, and CN: cognitive normal

however, multi-residual network has far less connection. In comparison to both ResNet and DenseNet, PartialNet block has average skip connection ($Ens\_ResNet < Ens\_PartialNet < Ens_DenseNet$).

Growth rate, k, of the PartialNet architecture is an important hyperparameter that controls the the rate at which layer's size in individual blocks grows. It particularly acts as a regulator to control information flow between successive layers. For instance, when $k = 16$ a filter size of 16 is used at each layer in individual block. During our experiments, smaller k showed enhanced performance and thus more efficient information transfer between layers. Additionally, the bottleneck layer (1×1 convolution) is introduced before each 3×3 convolution in each block to reduce the input feature-maps as in Lodhi et al. [32]. The ensemble of mulitresolutional PartialNet is performed after each block.

4 Experimental Results

The proposed architecture has been evaluated on MRI brain data obtained from the ADNI database. [33] ADNI is multisite longitudinal study with the main purpose to support active research for the diagnosis of AD at the early stage of MCI through the integration of various biomarkers. Particularly, clinical, imaging (e.g., MRI and PET), genetic, and bio specimen biomarkers are collected for that purpose. The datasets (ADNI, ADNI 2, ADNI 3, and ADNI GO) includes 1800 females and males subjects. In our work, we collected T1-weighted structural MR images for a cohort of 379 subjects (197 male and 182 female). The collected subjects contained 146 NC, 138 MCI, and 95 AD. Since each subject in the ADNI database has multiple MR scans collected at different times, in this study the minimum and maximum scan number for the selected subjects are 3 and 15, respectively. Summary of study cohort statistics is given in Table 2.

The proposed pipeline is implemented using MATLAB programming environment. All experiments of the proposed PartialNet architecture were conducted on a machine with NVIDIA RTX5000 GPU. Various experiments were performed using different system parameter, i.e., block sizes, multiresolution networks, and growth rates. The network is trained from input data for 70 epochs with each epoch consisting of 32 batches. We employed the stochastic gradient descent (SGD) optimization method for training and a 10-fold cross strategy has been utilized in our experiments.

In this work, several known classification metrics are used for the performance evaluation. The first is the accuracy, which is a basic evaluation metric, which calculates how much an instance class is correctly predicted by a given model in the validation set. Additional metrics are used, such as sensitivity and specificity, and are defined respectively as

$$\begin{aligned} \mathrm{ACC} &= \frac{\mathrm{TN} + \mathrm{TP}}{\mathrm{TN+TP+FN+FP}} \\ \mathrm{SEN} &= \frac{\mathrm{TP}}{\mathrm{TP+FP}} \\ \mathrm{SPE} &= \frac{\mathrm{TN}}{\mathrm{TN+FP}}, \end{aligned} \tag{4}$$

where TN, TP, FP, and FN represent true positive, true negative, false negative, and false-positive, respectively. In addition to quantitative metrics, network performance is qualitatively evaluated using accuracy/loss graphs against the number of epochs. In the training phase, loss of training and testing and also the loss of validation data.

Another quantitative evaluation is performed using the confusion matrix. The confusion matrix is very helpful to examine which classes, if any, were being misclassified more. Confusion matrices are not just useful in model evaluation but also model monitoring and model management. Additional indices can be measure from the generated confusion matrix that includes the precision, F1-score, and recall. Generally, both confusion matrix and related metrics are used together for the assessment/evaluation of classification models. Furthermore, the ablation study of proposed pipeline is also performed using similar multi-DenseNet fusion with partially and fully connected blocks for fair comparison.

Evaluation was performed on the publicly available ADNI dataset, using the multi-resolutional ensemble architecture. The advantage of the proposed multi-resolutional ensemble architecture is highlighted by comparing it against other networks for AD diagnosis. For all networks' training, the SGD method is used and other network hyperparameters, including the learning rate, weight decay, and Nesterov moment are set to 0.1, 10^{-4} and 0.9, respectively. Additionally, the learning rate is reduced by 10% factor at both 50% and 75% of the training epochs. Finally, we set the number of ensemble PartialNet are set to 3 (3x3, 5,5, 7x7) and the number of blocks to 4.

Framework evaluation is performed for both both binary classification and multiclass classification. First, binary

TABLE 3
Comparison With SOTA DNN-Based Literature Systems for the Identification of AD Using ADNI Dataset

Study	Tehnique	Modality	Classification	Accuracy (%)
Backstrom et al. [25]	3D-ConvNet	MRI	2 way	98.41
DenseNet	Multiresolution Ensemble	MRI	2 way	97.45
Gorji et al. [34]	pseudo Zernike moments	MRI	2 way	95.69
Hon et al. [35]	Transfer Learning	MRI	2 way	96
PartialNet [36]	Multiresolution Ensemble	MRI	2 way	98.82
Ensemble PartialNet	Multiresolution PartialNet Ensemble	MRI	2 way	99.05
Tong et al. [37]	NGF	PET+MRI+CSF	3 way (NC, MCI, AD)	60.2
Suk et al. [38]	DW-S MTL	PET+MRI+CSF	3 way (NC, MCI, AD)	62.93
Gupta et al. [39]	SAE	MRI	3 way (NC, MCI, AD)	85
Hosseni et al. [23], [40]	DSA-3DCNN	MRI	3way (NC, MCI, AD)	94.8
Lui et al. [41]	SAE-Zeromask	MRI+PET	4 way (NC, AD, cMCI, and neMCI)	53.8
Farooq et al. [24]	GoogleNet	MRI	4 way (NC, AD,MCI, and LMCI)	98.88
Farooq et al. [24]	ResNet-152	MRI	4 way (ANC, D,MCI, and LMCI)	98.14
Kazemi et al. [26]	AlexNet	MRI	4 way (CN, AD,SMC, EMCI, and LMCI)	97.63
PartialNet [36]	Multiresolution Ensemble	MRI	3 way (MCI,AD and CN)	98.23
Ensemble PartialNet	Multiresolution PartialNet Ensemble	MRI	3 way (MCI, AD, and CN)	98.46

classification is considered to differentiate subjects among different categories (i.e., MCI versus CN, AD versus CN, and MC versus AD). The summary of obtained results are given in Tables 3 and 4. As readily seen in the Tables 5 and 4, our proposed framework showed significantly higher performance in comparison to its counter network. Beside the accuracy, we can also observe that proposed framework has significantly less computational and space complexity.

To obtain better network and parameters, several experiments have been performed. Table 5 describes the ablation studies for both binary and multi-class problem of the proposed architecture and compare with DensNet and other state of the art network. We can observe that performance of proposed framework is degraded if we add the supervision and transition block within PartialNet, thus we have added during bottleneck. We can observe that our framework showed better performance in comparision to DenseNet, ResNet and our earlier network [36]. It may be due to the bottleneck after each block and ensemble of different PartialNet, besides, DenseNet and ResNet have path of different length, while we enforces the limitation of skip paths i.e., we not only enforces the skip connections in each block but also created bottleneck which results in better gradient flow.

5 Discussion

Alzheimer's disease (AD) is the most common cause of dementia and it is a slowly progressive brain disease, which starts many years before the emergence of the associated symptoms. Recent statistical data showed that an estimated 6.2 million Americans age 65+ are living with AD today [45]. Like other fatal diseases, early detection of the onset of AD can be help in the diagnosis, evaluation, and the development of treatments that slow or stop its the progression. Thus, automated systems that utilize patient data (including imaging, clinical, genetic, etc) could be valuable decision support tools for AD detection and could give significant downstream follow-up/treatment implications.

TABLE 4
Comparative Analysis (Accuracy) of Various SOTA on ADNI (Baseline) With 813 (187 AD, 228 NC, and 398 MCI) Subjects

–	CN versus AD	MCI versus AD
Hon and Khan [35]	84.14	82.26
Wang et al. [42]	79.93	76.86
Snapshot Ensemble [43]	54.93	68.02
Ensemble Transfer [44]	85.27	83.11
PartialNet [36]	87.11	99.26
Ensemble PartialNet	**87.92**	**99.38**

For that reason, we have developed a deep learning (DL) system for the detection of AD. The proposed DL has been validated experimentally on benchmark dataset of brain MR images, which contains three classes: NC, MCI, and AD. The proposed analysis system utilized a novel multi-resolutional ensemble PartialNet model, which possesses multiple advantages over dense networks, including (1) efficient feature reuse for better learning, (2) utilized the block concept to limit the skip connection within a block, which makes each block not only behaves like a unique structure but also contributes more to gradient magnitude than deeper architectures; and (3) incorporate supervision and transition blocks in each PartialNet block.

Developing an architecture for AD detection with high detection accuracy is our ultimate goal. Thus, in this study we have conducted several experiments in order to thoroughly evaluate our system's performance. This has been performed using various widely-used classification metrics, such as precision, recall, accuracy and F1-Score, area under the ROC. Additionally, data augmentation (i.e., rotation, zoomming, flipping) has been applied before being fed into the proposed architecture to prevent overfitting during network training. This results in an increase of the overall number of database images to 37,590.

As the results demonstrate, the proposed multi-resolution ensemble PartialNet structure documented a considerable gain in classification performance in comparison with benchmark approaches for both binary, or the two-level (e.g., AD-MCI) and multitclass (AD versus MCI versus CN) classifications using Alzheimer's disease neuroimaging initiative (ADNI) dataset. Further, our results showed higher detection performance in comparison to state-of-the-art (SOTA) DL-based methods. Particularly, results' evaluation

of PartialNet and DenseNet are summarized in Table 5. As shown in the Table, our system compared with DenseNet achieved a binary-classification accuracy of 100%, 99.38%, and 85.32% for MCI versus CN, MCI versus AD, and NC versus AD, respectively. While DenseNet showed 96.72 %, 80.11 %, and 96.65/97.76%, respectively, for the same binary tasks. We can also notice that Ensemble PartialNet multi-classification results (NC versus MCI versus AD) is 98.46% accuracy (both system sensitivity and specificity are 98+) in comparison with a 96.72% obtained using the DenseNet ensemble. It is worth mentioning that, the comparison of the proposed system with its counterpart DenseNet used same set of parameter and ensemble framework.

Furthermore, comparison with SOTA benchmark approaches on ADNI data sets are given in Tables 3 and 4. As evident from the presented results, the proposed ensemble PartialNet considerably outperformed all compared approaches. As readily seen, our system using ensemble PartialNet showed higher performance for both binary and multi-class classification for AD diagnosis. We can observe the gain in performance in terms of all parameters (accuracy, specificity, and sensitivity). Our proposed framework showed 99.11/99.23/97.54 sensitivity in comparison to Arfa et al. [24], Gupta et al. [46], and Hossein et al. [40] 98/97/97, 96/74/88, 100/80/47 sensitivity, for (AD/MC/NC) respectively. We can observe from the results that proposed framework has shown promising and consistent behaviour by achieving the sensitivity, specificity and accuracy above 98.24%, 98.02% and 98.46%, respectively. The higher performance can be explained in part by the fact that our ensemble PartialNet architecture follows a simple connectivity and forces the network to extract low, middle and high level features of features.

To this end, we can summarize the advantage of ensemble PartialNet framwork as: it showed considerably better generalization than its counter/parent networks; it incorporate the properties of identity mappings and diversified the depth through bottleneck after each block that forcest the network to extract low, mid and high level features, besides, it help the network to considers the feature reuse; the partial connectivity in each block help to deal with vanishing gradient and diminishing the forward flow with better training time as well as much lower number of parameters. Experiment on ADNI dataset showed considerable improvement in diagnostic performance both multiclass and binary class problem. Based on our experiments, we have the following *key observations:*

- the concept of block followed by bottleneck forces the network to extract low, mid and high level features and contributes more to gradient magnitude.
- The size of PartialNet and number of blocks depends upon the problem. For large datasets consisting of complex image may requires larger PartialNet blocks and more number of blocks.

6 Conclusion

In sum, we have presented a mulit-resolutional ensemble architecture tailored to AD detection using brain MRIs. By observing the reported experimental results, we can conclude that an the proposed multi-resolution ensembel PartialNet network showed considerably better generalization than DenseNet, the parent network. Our future pan is to explore group equivariant PartialNet to aid translations, reflections and rotations capability. Beside, we also plan to explore diver prediction to understand biological relevance that will help to observe the disease progression. Further, multi-input data fusion including other imaging types (i.e., PET) and clinical markers is another venue for our future work. This multi-input data system will help in providing more comprehensive markers (images as well as clinical) that can help physicians for more appropriate personalized treatment. In addition, to highlight our model, we will try to test our model with different and more diverse data sets.

References

[1] F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, and X. Zhang, "Multi-modal deep learning model for auxiliary diagnosis of alzheimer's disease," *Neurocomputing*, vol. 361, pp. 185–195, 2019.
[2] A. Bernasconi, A. Canakoglu, M. Masseroli, and S. Ceri, "META-BASE: A novel architecture for large-scale genomic metadata integration," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 19, no. 1, pp. 543–557, Jan./Feb. 2022.
[3] C. R. Jack Jr et al., "Tracking pathophysiological processes in alzheimer's disease: An updated hypothetical model of dynamic biomarkers," *Lancet Neurol.*, vol. 12, no. 2, pp. 207–216, 2013.
[4] M. F. Mendez, "Early-onset Alzheimer's disease," *Neurologic Clin.*, vol. 35, no. 2, pp. 263–281, 2018.
[5] A. Mehmood, M. Maqsood, M. Bashir, and Y. Shuyuan, "A deep siamese convolution neural network for multi-class classification of alzheimer disease," *Brain Sci.*, vol. 10, no. 2, 2020, Art. no. 84.
[6] X. Zhou, Y. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 912–921, May/Jun. 2021.
[7] D. Ferreira, J. B. Pereira, G. Volpe, and E. Westman, "Subtypes of alzheimer's disease display distinct network abnormalities extending beyond their pattern of brain atrophy," *Front. Neurol.*, vol. 10, 2019, Art. no. 524.
[8] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Classification in BioApps*, Berlin, Germany: Springer, 2018, pp. 323–350.
[9] A. Rehman, S. Naz, and I. Razzak, "Leveraging Big Data analytics in healthcare enhancement: Trends, challenges and opportunities," *Multimedia Syst.*, vol. 28, pp. 1339–1371, 2022.
[10] X. Zhou, W. Liang, I. Kevin, K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous Big Data recommendations," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 171–178, Feb. 2021.
[11] L. Tan et al., "Toward real-time and efficient cardiovascular monitoring for COVID-19 patients by 5G-enabled wearable medical devices: A deep learning approach," *Neural Comput. Appl.*, vol. 4, pp. 1–14, 2021.
[12] J. Huang, L. Ren, X. Zhou, and K. Yan, "An improved neural network based on SENet for sleep stage classification," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 10, pp. 4948–4956, Oct. 2022.
[13] A. Naseer, M. Rani, S. Naz, M. I. Razzak, M. Imran, and G. Xu, "Refining parkinson's neurological disorder identification through deep transfer learning," *Neural Comput. Appl.*, vol. 32, no. 3, pp. 839–854, 2020.
[14] M. I. Razzak, M. Imran, and G. Xu, "Efficient brain tumor segmentation with multiscale two-pathway-group conventional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 5, pp. 1911–1919, Sep. 2019.
[15] A. Rehman, S. Naz, M. I. Razzak, F. Akram, and M. Imran, "A deep learning-based framework for automatic brain tumors classification using transfer learning," *Circuits Syst. Signal Process.*, vol. 39, no. 2, pp. 757–775, 2020.
[16] M. A. Ebrahimighahnavieh, S. Luo, and R. Chiong, "Deep learning to detect alzheimer's disease from neuroimaging: A systematic literature review," *Comput. Methods Programs Biomed.*, vol. 187, 2020, Art. no. 105242.

[17] X. Zhou, W. Liang, I. Kevin, K. Wang, and S. Shimizu, "Multi-modality behavioral influence analysis for personalized recommendations in health social media environment," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 5, pp. 888–897, Oct. 2019.
[18] P. Vemuri and C. R. Jack, "Role of structural MRI in alzheimer's disease," *Alzheimer's Res. Ther.*, vol. 2, no. 4, pp. 1–10, 2010.
[19] P. Bhatkoti and M. Paul, "Early diagnosis of alzheimer's disease: A multi-class deep learning framework with modified k-sparse autoencoder classification," in *Proc. Int. Conf. Image Vis. Comput. New Zealand*, 2016, pp. 1–5.
[20] S. Sarraf et al., "DeepAD: Alzheimer's disease classification via deep convolutional neural networks using MRI and fMRI," *BioRxiv*, 2016, Art. no. 070441.
[21] J. Mathew, L. Mekkayil, H. Ramasangu, B. R. Karthikeyan, and A. G. Manjunath, "Robust algorithm for early detection of alzheimer's disease using multiple feature extractions," in *Proc. IEEE Annu. India Conf.*, 2016, pp. 1–6.
[22] M. A. Iftikhar and A. Idris, "An ensemble classification approach for automated diagnosis of alzheimer's disease and mild cognitive impairment," in *Proc. Int. Conf. Open Source Syst. Technol.*, 2016, pp. 78–83.
[23] E. Hosseini-Asl, R. Keynton, and A. El-Baz, "Alzheimer's disease diagnostics by adaptation of 3D convolutional network," in *Proc. IEEE Int. Conf. Image Process.*, 2016, pp. 126–130.
[24] A. Farooq, S. Anwar, M. Awais, and S. Rehman, "A deep CNN based multi-class classification of alzheimer's disease using MRI," in *Proc. IEEE Int. Conf. Imag. Syst. Techn.*, 2017, pp. 1–6.
[25] K. Bäckström, M. Nazari, I. Y.-H. Gu, and A. S. Jakola, "An efficient 3D deep convolutional network for alzheimer's disease diagnosis using MR images," in *Proc. IEEE 15th Int. Symp. Biomed. Imag.*, 2018, pp. 149–153.
[26] Y. Kazemi and S. Houghten, "A deep learning pipeline to classify different stages of alzheimer's disease from fMRI data," in *Proc. IEEE Conf. Comput. Intell. Bioinf. Comput. Biol.*, 2018, pp. 1–8.
[27] A. Ebrahimi-Ghahnavieh, S. Luo, and R. Chiong, "Transfer learning for alzheimer's disease detection on MRI images," in *Proc. IEEE Int. Conf. Ind. 4.0, Artif. Intell. Commun. Technol.*, 2019, pp. 133–138.
[28] R. Ju, C. Hu, P. Zhou, and Q. Li, "Early diagnosis of alzheimer's disease based on resting-state brain networks and deep learning," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 16, no. 1, pp. 244–257, Jan./Feb. 2019.
[29] C.-G. Yan, X.-D. Wang, X.-N. Zuo, and Y.-F. Zang, "DPABI: Data processing & analysis for (resting-state) brain imaging," *Neuroinformatics*, vol. 14, no. 3, pp. 339–351, 2016.
[30] A. Ashraf, S. Naz, S. H. Shirazi, I. Razzak, and M. Parsad, "Deep transfer learning for alzheimer neurological disorder detection," *Multimedia Tools Appl.*, vol. 80, pp. 30 117–30 142, 2021.
[31] S. Naz, A. Ashraf, and A. Zaib, "Transfer learning using freeze features for alzheimer neurological disorder detection using adni dataset," *Multimedia Syst.*, vol. 28, pp. 85–94, 2022.
[32] B. Lodhi and J. Kang, "Multipath-DenseNet: A supervised ensemble architecture of densely connected convolutional networks," *Inf. Sci.*, vol. 482, pp. 63–72, 2019.
[33] C. Sandeep, A. S. Kumar, and M. Susanth, "The online datasets used to classify the different stages for the early diagnosis of alzheimer's disease (AD)," *Int. J. Eng. Adv. Technol.*, vol. 6, no. 4, pp. 38–45, 2017.
[34] H. Gorji and J. Haddadnia, "A novel method for early diagnosis of alzheimer's disease based on pseudo zernike moment from structural MRI," *Neuroscience*, vol. 305, pp. 361–371, 2015.
[35] M. Hon and N. M. Khan, "Towards alzheimer's disease classification through transfer learning," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, 2017, pp. 1166–1169.
[36] I. Razzak, S. Naz, A. Ashraf, F. Khalifa, M. R. Bouadjenek, and S. Mumtaz, "Mutliresolutional ensemble PartialNet for alzheimer detection using magnetic resonance imaging data," *Int. J. Intell. Syst.*, vol. 37, pp. 6613–6630, 2022.
[37] W. Lin et al., "Convolutional neural networks-based MRI image analysis for the alzheimer's disease prediction from mild cognitive impairment," *Front. Neurosci.*, vol. 12, 2018, Art. no. 777.
[38] H.-I. Suk and D. Shen, "Deep learning-based feature representation for AD/MCI classification," in *Proc. Int. Conf. Med. Image Comput. Comput.- Assist. Intervention*, 2013, pp. 583–590.
[39] Y. Gupta et al., "Early diagnosis of alzheimer's disease using combined features from voxel-based morphometry and cortical, subcortical, and hippocampus regions of MRI T1 brain images," *PLoS One*, vol. 14, no. 10, 2019, Art. no. e0222446.
[40] E. Hosseini-Asl et al. "Alzheimer's disease diagnostics by a deeply supervised adaptable 3 D convolutional network," *Front. Biosci.*, vol. 23, no. 3, pp. 584–596, Jan. 2018.
[41] M. Liu et al., "A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in alzheimer's disease," *NeuroImage*, vol. 208, 2020, Art. no. 116459. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S105381191931050X
[42] S.-H. Wang, P. Phillips, Y. Sui, B. Liu, M. Yang, and H. Cheng, "Classification of alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling," *J. Med. Syst.*, vol. 42, no. 5, pp. 1–11, 2018.
[43] C. Zheng, Y. Xia, Y. Pan, and J. Chen, "Automated identification of dementia using medical imaging: A survey from a pattern classification perspective," *Brain Informat.*, vol. 3, no. 1, pp. 17–27, 2016.
[44] M. Tanveer, A. H. Rashid, M. Ganaie, M. Reza, I. Razzak, and K.-L. Hua, "Classification of alzheimer's disease using ensemble of deep neural networks trained through transfer learning," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 4, pp. 1453–1463, Apr. 2022.
[45] J. Wiley, "Alzheimer's disease facts and figures," *Alzheimers Dement*, vol. 17, pp. 327–406, 2021.
[46] A. Gupta, M. S. Ayhan, and A. S. Maida, "Natural image bases to represent neuroimaging data," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013, pp. III-987–III-994. [Online]. Available: http://dl.acm.org/citation.cfm?id=3042817.3043047