DeepPET-3D: A Deep Learning based 3D-CNN model for diagnosis of Alzheimer's Disease using 18-FDG-PET

Dishant Padalia[a,1], Darshil Mehta[a,1], Kaushik Metha[a], Anoushka Bhat[a], Pragya Gupta[a] and Ninad Mehendale[a,*]

[a]K.J. Somaiya College of Engineering, Mumbai, Maharashtra, India, 400077

ARTICLE INFO

Keywords:
Alzheimer's Disease
Deep Learning
Positron Emission Tomography
3D-CNN
Neuroimaging

ABSTRACT

Diagnosis of Alzheimer's disease is a challenging task, and detection can potentially help prevent its progression toward Dementia. Computer-aided techniques incorporating artificial intelligence have proven effective for diagnosing medical images. Deep learning tools like convolutional neural networks assist in extracting relevant visual information and thus avoiding manual feature extraction and interpretation. In this research, we propose a novel 3D-CNN model for the early detection of AD using positron emission tomography (PET). The dataset acquired from ADNI consists of 3D images of the brain segregated into three classes, namely Alzheimer's Disease (AD), Mild-Cognitive Impairment (MCI), and Cognitive Normal (CN). After data acquisition, we performed various pre-processing methods like thresholding, normalization, volume-reduction, and image augmentation. This study performed two types of experiments: multi-class and binary classification. The multi-class classification achieved an accuracy of 92.31%. Furthermore, the accuracy for AD vs CN, CN vs MCI, and MCI vs AD was 94.79%, 93.28%, and 96.91%, respectively.

1. Introduction

Alzheimer's disease is a brain disorder that progressively impairs the brain cells, resulting in memory deterioration. The key symptoms of Alzheimer's disease are disorientation, mood, behavior changes, and difficulty speaking, swallowing, and walking. According to the statistics provided by WHO, approximately 55 million people have Alzheimer's disease in the entire world, and it is expected to grow beyond 139 million by 2050 WHO (2021). The brain typically shrinks to some degree in healthy ageing but does not lose neurons in large numbers. On the contrary, for a person suffering from AD, extensive loss of neurons results in the significant shrinking of the brain Alzheimer'sAssociation (2022).

One of the standard methods to detect AD is screening the brain using Positron Emission Tomography (PET) Scan. A PET scan is an imaging test that can help reveal your tissues and organs' metabolic or biochemical function. The PET scan uses a radioactive drug (tracer) to indicate peculiar metabolic activity. Compared to computerized tomography (CT) and magnetic resonance imaging (MRI), a PET scan can discover abnormality of the tracer in a disease sooner. Before the screening, a small amount of radiotracer or radiopharmaceutical is injected into the peripheral vein. A gamma camera system is used to examine the infected organ and tissues during the scan RadiologyInfo (2021).

Manual diagnosis of AD in a PET Scan is a challenging and time-consuming task, and the probability of getting an incorrect diagnosis when manually examining the scan is high. Artificial Intelligence (AI) can be utilized to obtain results instantaneously and accurately to subdue this challenge. Various AI techniques can be employed to extract relevant features from the scan, and one of the most widely used AI techniques is the Deep Learning-based approach.

Convolutional Neural Networks(CNN) are state-of-the-art applications that extract visual information from a given set of input to carry out generative and descriptive tasks. CNN provides an amalgamation of various techniques that can be exploited for learning image representation and feature classification. The scarcity of experts, expensive consultation charges, and the complexities and anomalies in the medical images are complicated to identify. Therefore, computer-aided deep learning tools can significantly impact diagnosis by providing more accurate results than humans MayoClinic (2021); Shen, Wu and Suk (2017).

A PET scan is a collection of individual slices combined to form a 3D image of the brain. However, the methodology used in most of the research done for this problem has considered the slices of the brain as individual images instead of taking into account the collection of slices as a particular 3D entity. The majority of the previous research has used transfer learning and ensemble-based approaches. Alternatively, the uniqueness of our proposed solution takes into account the 3D scans, which are used to train a novel 3D CNN model to categorize a given PET scan into Alzheimer's disease (AD), Mild Cognitive Impairment (MCI), and Cognitive Normal (CN). A graphical representation of our proposed model is illustrated in Fig. 1.

2. Related Work

Several studies and research work has been carried out in diagnosis from medical images such as Positron Emission Tomography (PET) scans. A. Nordbert et al. Nordberg, Rinne, Kadir and Löm (2010) explored the use of PET in diagnosing AD. In this paper, various tracers were reviewed

*Corresponding author (ninad@somaiya.edu)
ORCID(s):
[1]These co-first authors contributed equally to this work.

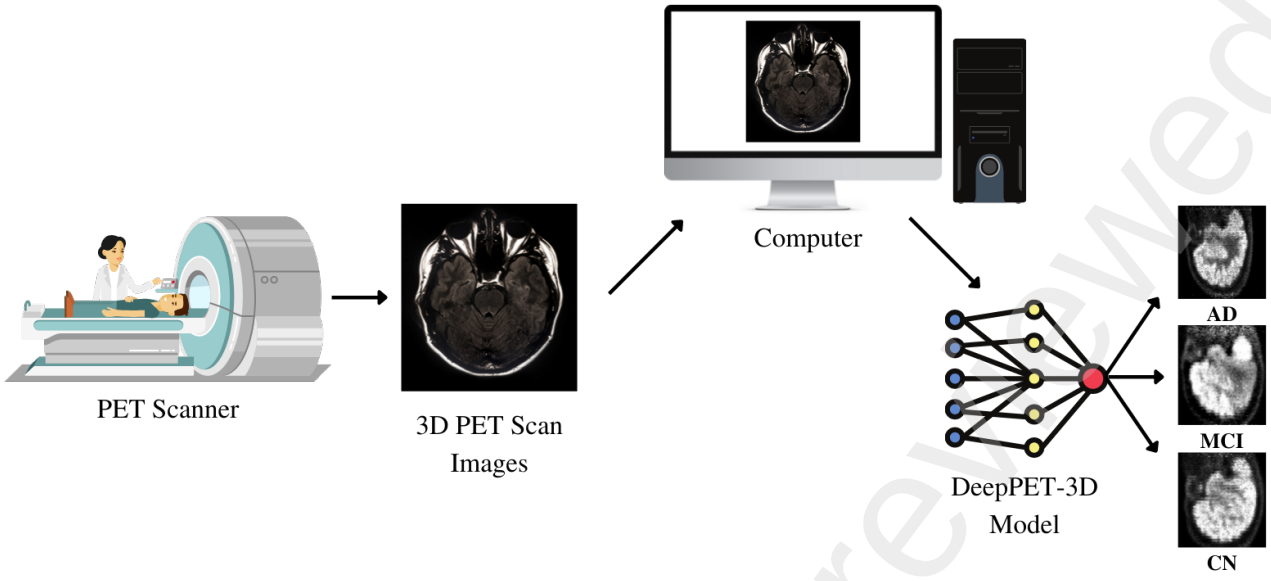

Figure 1: Schematic Flow Diagram. The PET scanner provides the 3D PET scan images of a patient to screen the Alzheimer's Disease. This input scan is passed through the proposed DeepPET-3D model that categorizes the patient's scan as Cognitively Normal, Mild Congentively Impairment, or Alzheimer's Disease.

that could be used to detect AD. The most effective and widely used tracer for early diagnosis of AD is glucose analog 2-[18F]-fluoro-2-deoxy-d-glucose (18F-FDG). It is used in PET imaging because the correlation between the degree of reduction in brain glucose metabolism and the severity of cognitive impairment is high in patients with early AD symptoms. Also, this tracer achieved a sensitivity of up to 90% to detect AD.

In recent years, image understanding systems that exploit machine learning (ML) techniques have been rapidly evolving. Convolutional neural networks are becoming a mainstream solution for analyzing medical images Yu, Yang, Zhang, Armstrong and Deen (2021); Tajbakhsh, Shin, Gurudu, Hurst, Kendall, Gotway and Liang (2016). For example, H. Suk et al. Suk and Shen (2013) proposed a deep learning-based feature representation model for AD/MCI/HC classification using a stacked autoencoder (SAE). The SAE was used to discover latent representation from neuroimaging and low-level biological features. This research obtained images of 51 AD patients, 99 MCI patients, and 52 HC patients from the ADNI dataset. Then, the selected feature information from the SAE was given as input to a multi-kernel support vector machine (SVM) for classification. Three binary classification problems were considered: 1) AD vs HC, which achieved an accuracy of 95.9%, 2) MCI vs HC with an accuracy of 85%, and 3) MCI-C vs MCI-NC with an accuracy of 75.8%.

In parallel with Suk's work, S. Liu et al. Liu, Liu, Cai, Pujol, Kikinis and Feng (2014b) developed an early diagnosis method for AD based on stacked sparse autoencoders with output as a softmax activation layer. The dataset used for the proposed methodology was acquired from ADNI that consisted of MRI images of 311 subjects that included 65 AD, 67 cMCI, 102 ncMCI, and 77 normal control patients. The proposed methodology was evaluated on both binary and multi-class classification. For binary, the accuracy for AD vs NC was 87.76%, whereas for MCI vs NC, it was 76.92%. For multi-class, the labels included were NC, cMCI, ncMCI, and AD, which yielded an accuracy of 47.42%.

S. Liu et al. Liu, Liu, Cai, Che, Pujol, Kikinis, Feng and Fulham (2014a) improved the model proposed in the previous research Liu et al. (2014b) by implementing a zero masking strategy on stacked auto-encoders (SAE). The dataset consisted of both PET and MR images of 331 and 758 subjects, respectively, unlike the previous research. The accuracy for both binary and multi-class classification improved with the help of this technique. Using SAE with zero masking, the accuracy for AD vs NC, MCI vs NC, and multi-class improved to 91.4%, 82.1%, and 53.79%, respectively.

In addition to extracted features from the scans, patients' demographics can be prominent in determining Alzheimer's disease. S. Singh et al. Singh, Srivastava, Mi, Caselli, Chen, Goradia, Reiman and Wang (2017) performed multiple binary classification experiments using different combinations of the classes, both with and without patient demographics. The proposed method involved applying dimensionality reduction on max-pooled and mean-pooled data. Then this data is fed to a neural network model, which performs binary classification. The images acquired from the ADNI2 dataset consisted of four classes: AD, CU, L-MCI, and E-MCI.

A promising alternative to devise a neural network architecture from scratch is to use pre-trained weights that have been trained on a large-scale collection of annotated images,

such as Imagenet Deng, Dong, Socher, Li, Li and Fei-Fei (2009). C. Zheng et al. Zheng, Xia, Chen, Yin and Zhang (2018) proposed an ensemble of pre-trained AlexNet models for the early diagnosis of AD. The solution proposed for this problem incorporated automated anatomical labelling (AAL) cortical parcellation map to detect anatomical brain volumes. These volumes were later used to fine-tune each AlexNet model. The label that got maximum votes from the ensemble network was classified as output. The dataset was obtained from ADNI for this research, including 241 AD, 306 mMCI, 127 sMCI, and 288 NC cases. Data augmentation techniques like shifting and rotation were used to increase the size of the dataset. The proposed approach achieved an accuracy of 91% for AD vs NC, whereas an accuracy of 85% was obtained for mMCI vs sMCI.

The slices in the PET scan are sequential; Recurrent Neural Networks Sherstinsky (2020) (RNN) can be used since they are memory-based neural networks storing previous cell state and hidden state information. M. Liu et al. Liu, Cheng, Yan and Initiative (2018) proposed a novel architecture that incorporated a combination of both CNN and RNN frameworks to create an ensemble-based classification model. The 2D CNN model captures the intra-slice features while the Bidirectional Gated Recurrent Unit network learns inter-slice characteristics. This research used image slices with sagittal, axial, and coronal orientations as input to each model. The weighted sum of the outputs obtained from each model was then used as an input for the final classification. Two binary classification models were trained that achieved an accuracy of 91.2% and 78.9% for AD vs NC and MCI vs NC, respectively.

A PET scan is a collection of 2D slices in sequential order that forms a three-dimensional structure. Most of the methodologies employ 2D-CNN architecture by considering each slice a separate image for classification. However, these individual slices fail to capture information from the entire brain. Hence, 3D-CNN architecture can be used to classify the whole three-dimensional brain scan as a whole. For instance, H. Choi et al. Choi, Jin and Initiative (2018) developed an automatic image interpretation system using Convolutional Neural Networks (CNN) to predict Alzheimer's disease. The data collected from ADNI was used to train a 3D-CNN model that performed binary classification between AD and NC that obtained an accuracy of 96%. The model parameters were transferred as the imaging features to categorize MCI converter and non-converter, achieving an accuracy of 84.2%.

In addition to Choi's work, T. Jo et al. Jo, Nho, Risacher and Saykin (2020) initially trained a 3D-CNN model to perform binary classification on Tau-PET scans of AD and CN. This model was applied to Tau-PET scans of MCI patients to generate AD probability scores between 0 and 1. Scores closer to 1 indicate AD symptoms, whereas scores nearer to 0 indicate typical brain characteristics. The deep learning model for the prediction of AD and CN yielded an accuracy of 90.8%. The probability scores computed substantiated that there exists a positive correlation between AD and L-MCI.

K. Etminan et al. Etminani, Soliman, Davidsson, Chang, Martínez-Sanchis, Byttner, Camacho, Bauckneht, Stegeran and Ressner (2022) proposed a solution to develop and validate a 3D deep learning model to predict a final clinical diagnosis of Alzheimer's disease. A 3D image classification model was developed with reference to the VGG16 model, and it was trained using data acquired from ADNI and EDLB. The proposed model achieved an area under the ROC curve of 96.2% for DLB, 96.4% for AD, 71.4% for MCI, and 94.7% for CN. Based on the results observed, the proposed model achieved competitive performance compared to human readers in predicting the final diagnosis of common neurodegenerative disorders.

3. Materials and Methods

3.1. Dataset

The Dataset used in our research was acquired from the Alzheimer's Disease Neuroimaging Initiative Mueller, Weiner, Thal, Petersen, Jack, Jagust, Trojanowski, Toga and Beckett (2005) database. ADNI is a global research study that devises and evaluates the treatments used for the early detection of Alzheimer's Disease. The main objective of ADNI is to prevent the progression of Alzheimer's Disease and identify potential dead brain cells due to atrophy or brain shrinkage. The ADNI Image Collections provides repositories that scientific fraternities can research to develop optimum medical image analysis methodologies.

The ADNI database comprises PET Scans of multiple subjects, of which we selected 624 studies containing three classes, of which 160, 231, and 233 are image sets of AD, MCI, and CN, respectively. Sample demographics of patients' are given in Table 1. Each 3D PET Scan consists of 210 or 282 slices of the brain. The radiopharmaceutical used in these 3D scans is 18F-FDG (18 Fluorine-fluorodeoxyglucose) and the radioisotopes used are C-11 (Carbon-11) and F-18 (Fluorine-18). The image sets consisted of 3 orientations: axial, coronal, and sagittal. We selected axial orientation, which is the top view of the brain, as it provides better clarity of the dead brain cells than sagittal and coronal orientation.

3.2. Data Preparation

The slices of the image set were in DI-COM (.dcm) format. DI-COM stands for Digital Imaging and Communications in Medicine. It is a conventionally and universally accepted format for managing medical images. The intensity value of a DI-COM image ranges from -2^{15} to 2^{15}. The shape of each slice in the image set is 128 x 128. We performed voxel normalization to standardize the intensity values. Then we used min-max normalization to scale the pixel values between 0 and 1. After testing various minimum and maximum values, we concluded that the optimal threshold range is 0 to 20000. The volume was resized to 128 x 128 x 128 (height x width x depth) to maintain the uniformity of the number of slices in the image set. Subsequently, we

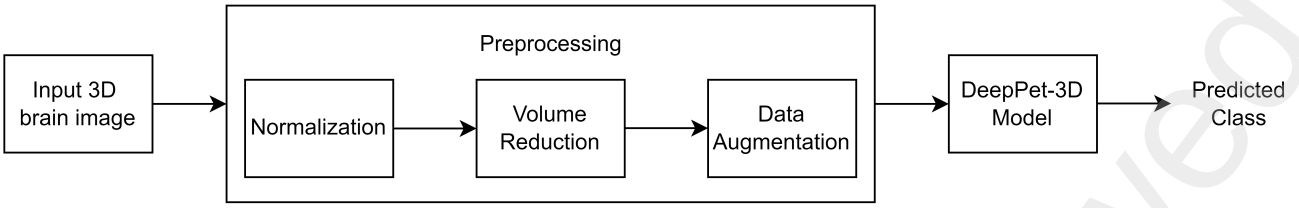

Figure 2: The proposed method is based on a 3D-CNN Framework. The input 3D image of the brain is first preprocessed before passing it to the model for training. Preprocessing has three significant steps: Normalization, in which voxel normalization is used to standardize the pixel values; volume Reduction in which redundant top and bottom image slices are removed; and lastly, Data Augmentation to increase the size of the dataset. The preprocessed dataset is passed to the 3D CNN model for feature extraction. Finally, the Alzheimer's Disease stage from the input image is detected.

Table 1
Patients' Demographics

Label	Gender Count		Average Age	
	Male	Female	Male	Female
AD	90	39	76.67 ± 8.15 (57 - 91)	76.23 ± 5.66 (63 - 86)
MCI	141	59	78.35 ± 6.66 (58 - 91)	74.92 ± 5.79 (59 - 85)
CN	97	102	77.90 ± 6.42 (62 - 90)	74.92 ± 5.79 (59 - 85)
ALL	328	200	77.90 ± 6.42 (62 - 90)	75.72 ± 6.13 (56 - 87)

inspected the slices and concluded that the top and bottom 32 slices were redundant; hence they were eliminated from the volume. Slices from each class are visualized in Fig. 3.

3.3. Augmentation

Image augmentation generates training images using image processing techniques such as random rotation, shifting, flipping, and colour transformation. The 624 image sets were augmented by rotating the image sets by 13 degrees to the left and right sides. Consequently, the number of image sets increased from 624 to 1500 training image sets, 62 validation image sets and 62 testing image sets. Original and augmented slices can be observed in Fig. 4.

3.4. Training Data Preparation

The labels (AD, MCI, and CN) were one-hot encoded for training the Convolutional Neural Network (CNN) model. The model is initially fit on a training dataset. Initially, the model learns the relevant features of images from the training dataset. Then the model is analyzed on a validation dataset to estimate how well the model is performing on images that are not being used in training. Lastly, the final model is assessed on the test dataset to get an unbiased evaluation. Out of the 624 image sets, we used 62 image sets for validation, and 62 image sets for testing and the remaining 500 image sets were augmented using rotation to 1500 image sets.

3.5. Model Architecture

Fig. 5 represents the proposed model architecture to predict Alzheimer's disease (AD). The architecture comprises convolutional and max-pool layers for image feature extraction, dropout, and batch normalisation to prevent over-fitting and fully-connected dense layers for predicting the outcome.

The first layer in our model is the input layer. It takes an image of dimension 128 x 128 x 64 x 1, where 128 x 128 is the size of the image slice, 64 represents the number of slices, and 1 indicates the number of channels (1 channel means a black and white image) of the image set.

The model consists of 4 blocks; each block of our architecture consists of a 3D convolutional layer, a 3D max-pooling layer followed by a batch-normalisation layer. The 3D convolutional layer performs convolution operation using the specified number of filters on the input data. 3D convolutional layers will be more expensive in the required computational power but allow the model to retrieve many relevant insights. The convolutional layer is followed by a 3D Max Pooling layer that downsamples the input by computing the maximum value over a window of specified pool size for each channel. Furthermore, it reduces the feature map's size, making it convenient for successive convolutional layers to extract new features. To prevent overfitting, we used

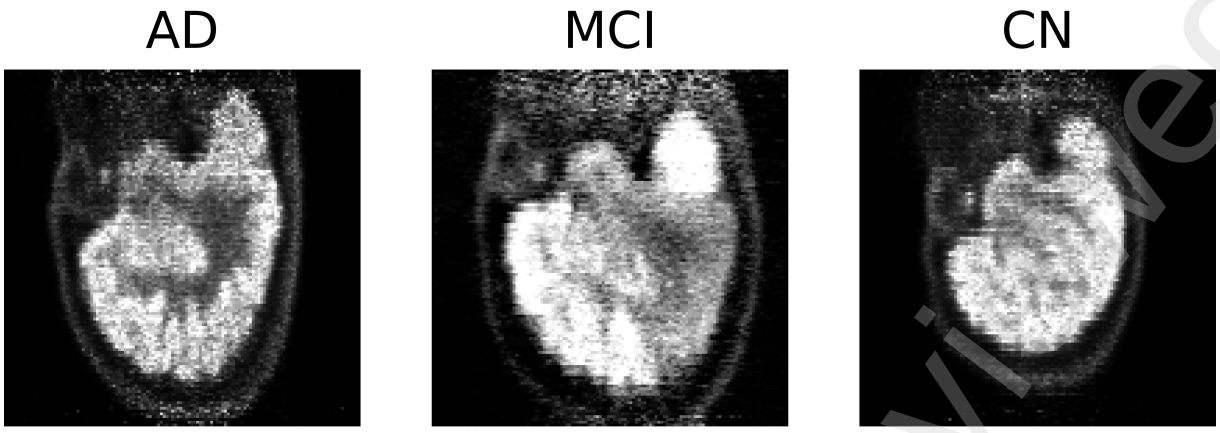

Figure 3: Visualization of random slices of each class in the Dataset

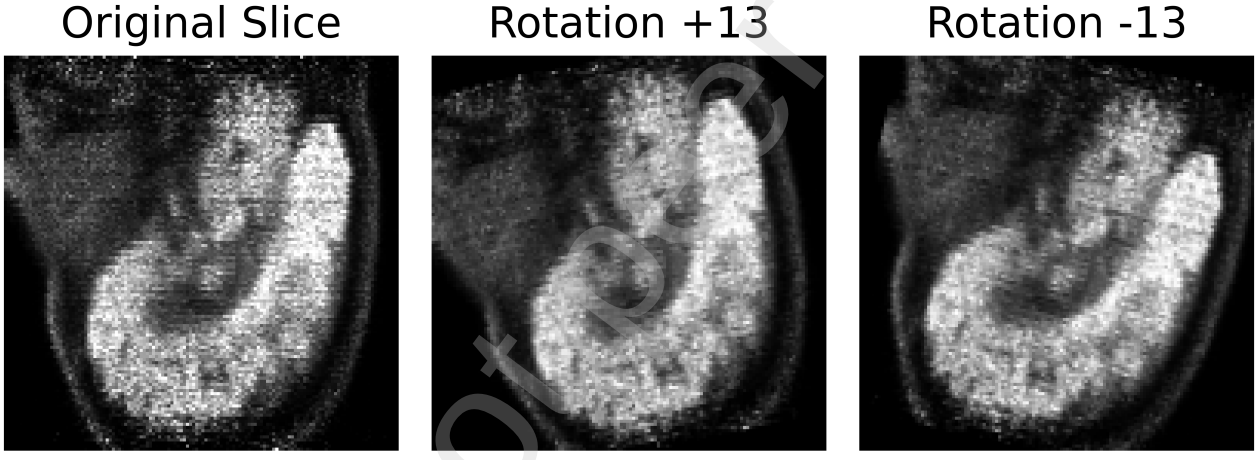

Figure 4: Visualization of random augmented slice from a random image

batch-normalisation at the end of each block. Batch Normalisation uses a transformation technique that maintains the mean output close to 0 and the standard deviation close to 1, which helps the model reduce overfitting by normalising the input data. In addition to batch-normalisation, Dropout Layer was used to generalise the model. It randomly sets the output of a certain percentage of neurons to zero while training. For classification, the extracted features were fed to fully-connected Dense layers.

The number of filters in the convolutional layers of the four blocks are 32, 64, 128, and 256, respectively. Each convolutional layer uses a kernel of size 3 x 3 x 3 and an activation function as 'ReLU.' The pool size in every max-pool layer was set to 3 x 3 x 3. The convolutional and max-pool layers padding was selected as 'same.'

The output of the final block is converted into a one-dimensional array using the Flatten layer. This array is passed to the fully-connected dense layer with 512 units with ReLU activation. Finally, a Dense layer with softmax activation predicted the probability of three labels, namely AD, MCI, and CN. A Dropout layer with a rate of 0.3 was used to avoid overfitting.

3.6. Model Training

We used the TensorFlow framework Abadi, Barham, Chen, Chen, Davis, Dean, Devin, Ghemawat, Irving and Isard (2016) to train and evaluate the performance of our model. After preprocessing and augmentation, the images were loaded onto a Windows 10 operating system machine. The machine has a six-core Intel Xeon Silver 3.2-GHz processor, 96 GB of DDR4 RAM, and a Quadro RTX 5000 graphical processing unit with CUDA 11.2 and CuDNN 8.1 (Nvidia).

For training the model, we used Adam, a first-order gradient-based stochastic optimization function, with a learning rate of 0.001, categorical cross-entropy as the loss

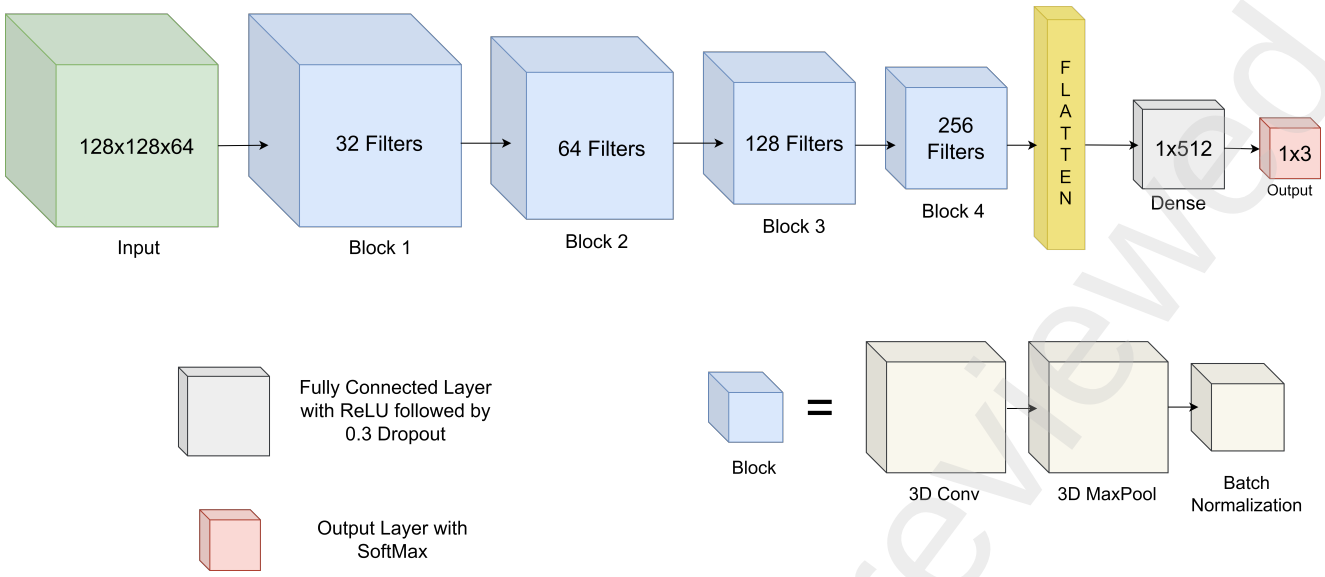

Figure 5: Model Architecture

function, batch size of 8, and softmax activation function in the output layer.

An optimizer is a function that alters the model's parameters, which helps reduce the overall loss and improve the accuracy. Adam (Adaptive Moment Estimation) optimizer Kingma and Ba (2014) is used in our model as it is computationally efficient and requires less memory space. Adam algorithms combine the heuristics of both Momentum and RMSProp. Here are the updated equations.

$$For\ each\ Parameter\ W^j$$
$$V_t = \beta_1 * V_{t-1} - (1-\beta_1) * G_t$$
$$S_t = \beta_s * S_{t-1} - (1-\beta_s) * G_t^2$$
$$\Delta W_t = -\eta \frac{V_t}{\sqrt{S_t + \epsilon}} * G_t$$
$$W_{t+1} = W_t + \Delta W_t$$

where :
η : *Initial learning rate*
G_t : *Gradient at time t along* W^j
V_t : *Exponential Average of gradients*
S_t : *Exponential Average of squares of gradients*
$\beta_1\ \beta_2$: *Hyperparameters*

There are two loss functions for multiclass classification: Categorical Cross-entropy and Sparse Categorical Cross-entropy. Since our labels are one-hot encoded, we used the Categorical Cross entropy loss function. The loss function measures the model's prediction error, and then it calculates the gradients and updates the weights of the model. Similarly, we used Binary Cross-entropy for binary classification to compare the predicted probabilities to the ground truth, which can be either 0 or 1.

Binary Cross Entropy :

$$Loss = \frac{-1}{N} \sum_{i=1}^{N} y_i . \log \hat{y}_i + (1 + y_i) . \log(1 - \hat{y}_i)$$

Categorical Cross Entropy :

$$Loss = -\sum_{i=1}^{N} y_i . \log \hat{y}_i$$

where :
N : *Number of classes*
y_i : *Actual label*
$\hat{y}_i$: *Predicted label*

The multiclass classification model uses the softmax activation function to predict the class of the image set in the output layer. Softmax computes the probabilities of each class such that the resultant probability adds up to 1. The label with the maximum probability is classified as the output. On the other hand, the sigmoid activation function is used for binary classification. The output value ranges between 0 and 1. If the computed value is above 0.5, it is classified as 1; otherwise, 0.

Softmax Activation Function :

$$f(s)_i = \frac{e^s_i}{\sum_j^C e^{s_j}}$$

Sigmoid Activation Function :

$$S(x) = \frac{1}{1 + e^{-x}}$$

Table 2
Results obtained from different experiments for Multiclass Classification

No. of parameters	Train Accuracy	Train Loss	Validation Accuracy	Validation Loss	Test Accuracy	Test Loss
206,515	98.74	0.0016	90.5	0.3297	90.56	0.3208
556,131	99.62	0.0004	94.5	0.158	91.33	0.1487
1,690,819	**95.78**	**0.0003**	**94.44**	**0.3202**	**92.31**	0.3531
5,702,019	99.18	0.0431	86.54	1.6435	81.13	1.281
16,842,511 (AlexNet)	97.32	0.1012	93.06	0.3435	84	0.6315
20,691,715	100	0.0066	91	0.2935	89.93	0.2595
50,357,8833 (ResNet50)	58.48	0.9087	57.73	0.8891	50	0.9968

4. Results and Analysis

4.1. Evaluation of Proposed Architecture

The dataset used for research was obtained from ADNI, which consisted of 624 18-FDG PET scans of the brain. Out of these 624 image sets, 62 image sets were used for validation, and 62 image sets were used for testing. The remaining 500 image sets were augmented using rotation at 13 degrees clockwise and anti-clockwise, yielding 1500 training images.

To evaluate the model's performance, we used metrics such as accuracy, precision, recall and F1-score.

4.2. Multi-class Classification

We trained various models, that included custom model and a few state-of-the-art architectures such as AlexNet Krizhevsky, Sutskever and Hinton (2012) and ResNet50 He, Zhang, Ren and Sun (2016). From Table 2, we can observe that increase in the number of parameters resulted in poor performance of models. The optimal performance on the dataset can be achieved by the model with the proposed architecture. We tuned the proposed model for further verification by using different permutations and combinations of parameters such as kernel size, pool size, and the number of layers.

From Fig. 6-8, it can be observed that the model trained with a kernel size of 3, pool size of 3, and 4 layers that is the proposed architecture attained the best performance compared to other combinations. This architecture achieved 95.78%, 94.44% and 92.31% for training, validation and testing, respectively. Other metrics used for evaluation included recall, precision and F1-score. For training, the model achieved 95.69%, 95.78% and 95.73% for recall, precision and F1-score, respectively. Whereas for validation, the model achieved 92.31% for recall, precision and F1-score.

4.3. Binary Classification

For binary classification, we replaced the categorical cross-entropy loss function with the binary cross-entropy loss function and used the sigmoid activation function instead of the softmax activation function. The proposed architecture was reused for classification between AD vs CN, CN vs MCI and MCI vs AD, and it achieved an accuracy of 98.5%, 97.18% and 98.04% for training and 94.79%, 93.28% and 96.91% for testing respectively.

4.4. Comparison with other research works

From Table 4, it can be observed that the proposed architecture achieved the optimal accuracy of 92.31% for multiclass classification compared with the models presented in other studies. In the case of binary classification, the model developed by H. Suk et al. Suk and Shen (2013) achieved the highest accuracy of 95.9% for AD vs CN, whereas our model attained a competitive accuracy of 94.79%, which proves the claim made by the researchers regarding the discovery of latent features that are helpful for the improvement of the accuracy. For CN vs MCI, our model obtained an accuracy of 93.28%, which was the highest compared with other reviewed models. Furthermore, there was barely any study regarding AD vs MCI. Thus, this encouraged us to test our custom model for binary classification between AD and MCI, where we achieved an accuracy of 96.91%.

5. Discussion

The dataset used for this study consists of 3 stages of Alzheimer's disease: Cognitive Normal (CN), Mild Cognitive Impairment (MCI) and Alzheimer's disease (AD). However, the MCI stage can be further split into Early MCI (E-MCI) and Late MCI (L-MCI), which further insights into preventing progression to AD and the degree of treatment required. By including these classes, the diagnosis can be more accurate and precise. Along with the classification of the stages, a model can be developed to identify and segment areas in the 3D scan that show signs of brain shrinkage or dead neurons.

6. Conclusion

This study explores a Deep Learning approach for early detection of Alzheimer's Disease that helps the patients avoid its adverse effects that might occur due to progression in future. The Convolutional Neural Network is a robust tool that can be used to classify medical diseases since it can process an input image/scan and classify it into its

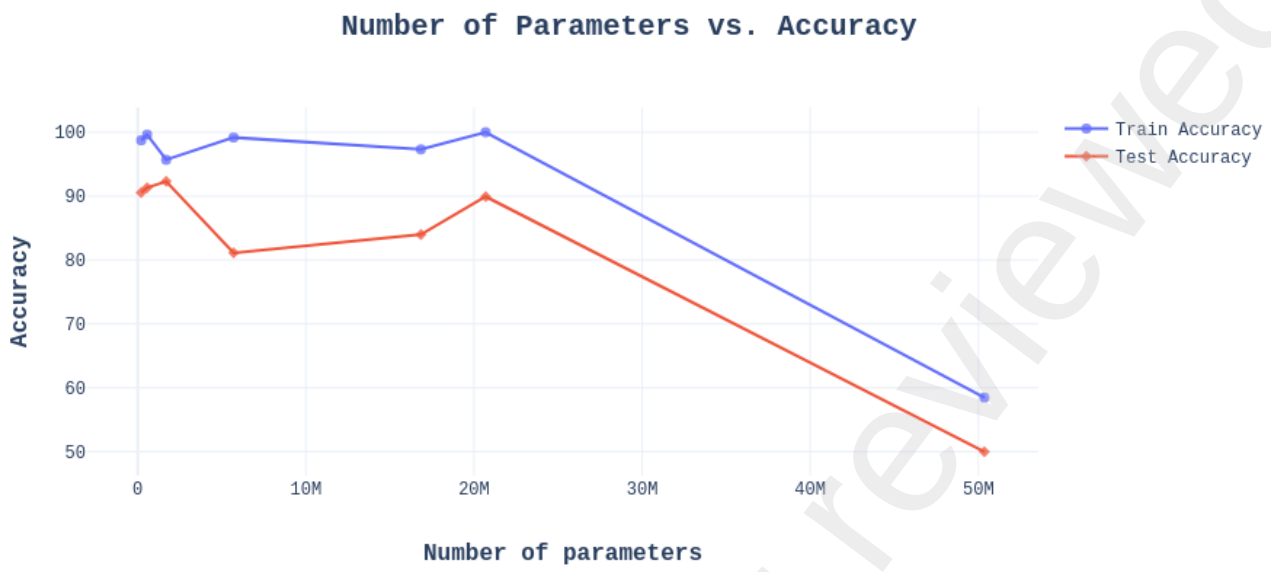

Figure 6: Number of Parameters vs. Accuracy

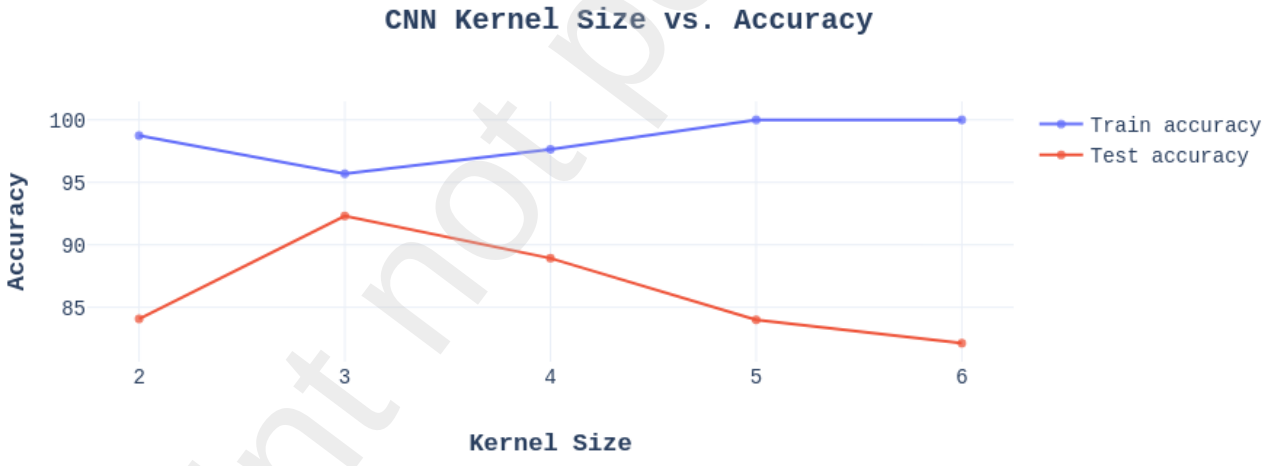

Figure 7: Kernel Size vs. Accuracy

Table 3
Results obtained from different experiments for Binary Classification

Experiment	Train Accuracy	Validation Accuracy
AD vs. CN	98.5	94.79
CN vs. MCI	97.18	93.28
MCI vs. AD	98.04	96.91

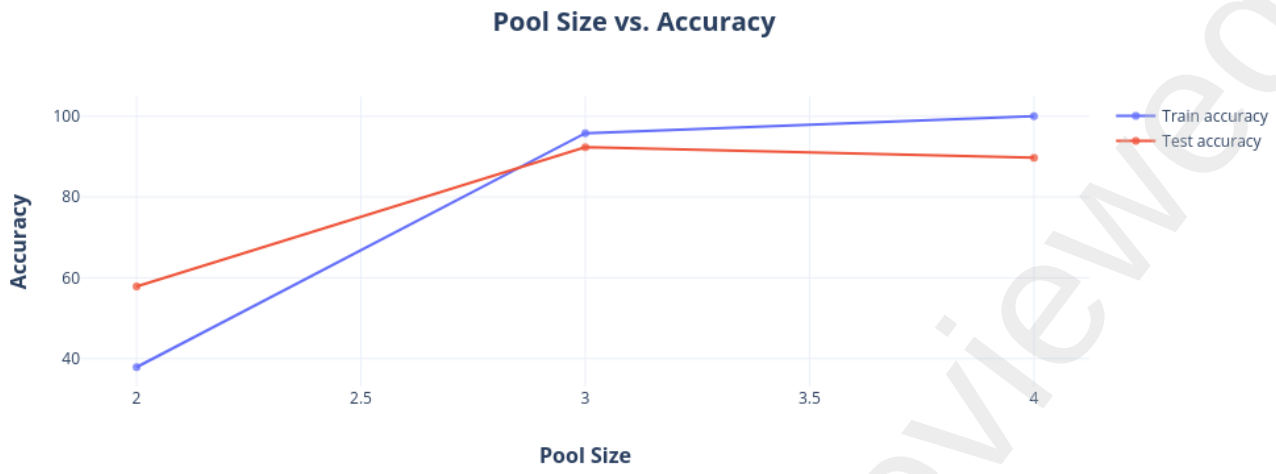

Figure 8: Pool size vs. Accuracy

Table 4
Comparison table with methods reported in the literature

Research Paper	Model Architecture	Mutliclass Accuracy	Binary Accuracy AD vs. CN	Binary Accuracy CN vs. MCI
Suk and Shen (2013)	SAE + SVM	-	**95.9**	85
Liu et al. (2014b)	SAE	47.42	87.76	76.92
Liu et al. (2014a)	Improved SAE	53.79	91.4	82.1
Zheng et al. (2018)	EnAlexNet	-	91	-
Liu et al. (2018)	CNN + RNN	-	91.2	78.9
Choi et al. (2018)	3D-CNN	-	96	-
Jo et al. (2020)	3D-CNN	-	90.8	-
Proposed Architecture	3D-CNN	**92.31**	94.79	**93.28**

corresponding class. Furthermore, the DeepPET-3D model provides better accuracy than the previous research for multi-class classification of 3D brain image slices. This method can be used by doctors for mass screening of patients, reducing the workload of radiologists. The proposed method can efficiently classify the patient's PET scan as AD, MCI or CN.

Authors' contributions:
Conceptualization was done by Dr. Ninad Mehendale, Pragya Gupta, Dishant Padalia and Darshil Mehta. All the experiments/code executions were performed by Dishant Padalia, Darshil Mehta and Kaushik Metha. The formal analysis was performed by Dishant Padalia and Darshil Mehta. Manuscript writing- original draft preparation was done by Dishant Padalia, Darshil Mehta, and Kaushik Metha. Review and editing was done by Dr. Ninad Mehendale, Dishant Padalia, and Darshil Mehta. Visualization work was carried out by Anoushka Bhat.

Acknowledgments:
The authors would like to thank the colleagues.

Funding information:
No funding was involved in the present work.

Consent for publication:
Authors have taken all the necessary consents for publication from participants wherever required.

Code availability:
All the codes used in this study are provided in the supplementary material.

Compliance with ethical standards

Conflict of interest:
The authors declare that they have no conflict of interest.

Consent to participate:
This article does not contain any studies with animals or humans performed by any of the authors. Informed consent was not required as there were no human participants. All the necessary permissions were obtained from Institute

Ethical committee and concerned authorities.

Ethics approval:

All authors consciously assure that the manuscript fulfills the following statements:

1. This material is the authors' own original work, which has not been previously published elsewhere.
2. The paper is not currently being considered for publication elsewhere.
3. The paper reflects the authors' own research and analysis in a truthful and complete manner.
4. The paper properly credits the meaningful contributions of co-authors and co-researchers.
5. The results are appropriately placed in the context of prior and existing research.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., 2016. $$TensorFlow$$: A System for $$Large-Scale$$ Machine Learning, in: 12th USENIX symposium on operating systems design and implementation (OSDI 16), pp. 265–283.

Alzheimer'sAssociation, 2022. Alzheimer's Effects on Brain.

Choi, H., Jin, K.H., Initiative, A.D.N., 2018. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. Behavioural brain research 344, 103–109.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, pp. 248–255.

Etminani, K., Soliman, A., Davidsson, A., Chang, J.R., Martínez-Sanchis, B., Byttner, S., Camacho, V., Bauckneht, M., Stegeran, R., Ressner, M., 2022. A 3D deep learning model to predict the diagnosis of dementia with Lewy bodies, Alzheimer's disease, and mild cognitive impairment using brain 18F-FDG PET. European journal of nuclear medicine and molecular imaging 49, 563–584.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

Jo, T., Nho, K., Risacher, S.L., Saykin, A.J., 2020. Deep learning detection of informative features in tau PET for Alzheimer's disease classification. BMC bioinformatics 21, 1–13.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 .

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25.

Liu, M., Cheng, D., Yan, W., Initiative, A.D.N., 2018. Classification of Alzheimer's disease by combination of convolutional and recurrent neural networks using FDG-PET images. Frontiers in neuroinformatics 12, 35.

Liu, S., Liu, S., Cai, W., Che, H., Pujol, S., Kikinis, R., Feng, D., Fulham, M.J., 2014a. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Transactions on Biomedical Engineering 62, 1132–1140.

Liu, S., Liu, S., Cai, W., Pujol, S., Kikinis, R., Feng, D., 2014b. Early diagnosis of Alzheimer's disease with deep learning, in: 2014 IEEE 11th international symposium on biomedical imaging (ISBI), pp. 1015–1018.

MayoClinic, 2021. Positron emission tomography scan.

Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L., 2005. The Alzheimer's disease neuroimaging initiative. Neuroimaging Clinics 15, 869–877.

Nordberg, A., Rinne, J.O., Kadir, A., Löm, B., 2010. The use of PET in Alzheimer disease. Nature Reviews Neurology 6, 78–87.

RadiologyInfo, 2021. Positron emission tomography/CT.

Shen, D., Wu, G., Suk, H.I., 2017. Deep learning in medical image analysis. Annual review of biomedical engineering 19, 221–248.

Sherstinsky, A., 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena 404, 132306.

Singh, S., Srivastava, A., Mi, L., Caselli, R.J., Chen, K., Goradia, D., Reiman, E.M., Wang, Y., 2017. Deep-learning-based classification of FDG-PET data for Alzheimer's disease categories, in: 13th International Conference on Medical Information Processing and Analysis, p. 105720.

Suk, H.I., Shen, D., 2013. Deep learning-based feature representation for AD/MCI classification, in: International conference on medical image computing and computer-assisted intervention, pp. 583–590.

Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B., Liang, J., 2016. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging 35, 1299–1312.

WHO, 2021. Dementia.

Yu, H., Yang, L.T., Zhang, Q., Armstrong, D., Deen, M.J., 2021. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. Neurocomputing 444, 92–110.

Zheng, C., Xia, Y., Chen, Y., Yin, X., Zhang, Y., 2018. Early diagnosis of Alzheimer's disease by ensemble deep learning using FDG-PET, in: International Conference on Intelligent Science and Big Data Engineering, pp. 614–622.