



Research paper

Multi-instance learning attention model for amyloid quantification of brain sub regions in longitudinal cognitive decline

R. Divya^{*,1}, R. Shantha Selva Kumari², for the Alzheimer's Disease Neuroimaging Initiative³

Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi 626 005, Tamil Nadu, India



ARTICLE INFO

Keywords:

Alzheimer's disease
Attention
Standardized uptake value ratio
Multi-instance learning

ABSTRACT

Amyloid PET scans help in identifying the beta-amyloid deposition in different brain regions. The purpose of this study is to develop a deep learning model that can automate the task of finding amyloid deposition in different regions of the brain only by using PET scan and without the corresponding MRI scan. 2647 ¹⁸F-Florbetapir PET scans are collected from Alzheimer's Disease Neuroimaging Initiative (ADNI) from multiple centres taken over a period. A deep learning model based on multi-instance learning and attention is proposed which is trained and validated using 80% of the scans and the remaining 20% of the scans are used for testing the model. The performance of the model is validated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The proposed model is further tested upon an external dataset consisting of 1413 ¹⁸F-Florbetapir PET scans from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study. The proposed model achieves MAE of 0.0243 and RMSE of 0.0320 for summary Standardized Uptake Value Ratio (SUVR) based on composite reference region for ADNI test set. When tested on the A4-study dataset, the proposed model achieves MAE of 0.038 and RMSE of 0.0495 for summary SUVR based on the composite region. The results show that the proposed model provides less MAE and RMSE when compared with existing models. A graphical user interface is developed based on the proposed model where the predictions are made by selecting the files of ¹⁸F-Florbetapir PET scans.

1. Introduction

Alzheimer's disease is a chronic neurodegenerative disease that affects older people across the world. Dementia progressively worsens over the years making the persons affected to lose their ability to respond to their surroundings. The main biomarkers of Alzheimer's disease are the abnormal accumulation of beta-amyloid proteins in between the neurons and neurofibrillary tau tangles inside the nerve cells (Hampel et al., 2009; Verde, 2022). During the autopsy of the Alzheimer's disease-affected brains, amyloid plaques and tau tangles are found to be high and in predictable patterns in the brain regions accountable for memory when compared to healthy adults' brains. The metabolism changes occur in the person affected by this disease even before the structural changes take place (Jack et al., 2013). Thus, functional brain scanning will be able to better identify these abnormal

deposits even at the earlier stage of the disease. The scans visualize amyloid plaques present in the brain, which are the prime suspects in damaging nerve cells in Alzheimer's disease-affected brain. Before the usage of amyloid PET scans, these plaques could only be detected during the autopsy. Amyloid PET scanning highlights the amyloid plaques found in the brain of living people. The abnormal beta-amyloid deposits are identified with the help of the radiotracers like Florbetapir, Florbetaben, and Flutemetamol used in Positron Emission Tomography scans. Of these, Florbetapir PET scans are more commonly used by neurologists and radiologists in Alzheimer Disease Neuroimaging Initiative (ADNI) as a research biomarker to determine whether a scan is amyloid positive or amyloid negative. These results can further help in the trials involving treatments for Alzheimer's disease by having anti-amyloid medications (Sperling et al., 2020). In particular, for amyloid-targeted therapies, amyloid PET imaging plays a strategic role in therapeutics and

* Corresponding author.

E-mail addresses: rdivya@mepcoeng.ac.in (R. Divya), rshantha@mepcoeng.ac.in (R. Shantha Selva Kumari).

¹ 0000-0001-8469-3534.

² 0000-0003-4123-7744.

³ Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

<https://doi.org/10.1016/j.brainres.2024.149103>

Received 9 November 2023; Received in revised form 21 May 2024; Accepted 26 June 2024

Available online 30 June 2024

0006-8993/© 2024 Elsevier B.V. All rights reserved, including those for text and data mining, AI training, and similar technologies.

diagnostics in the early treatment of the disease and also for amyloid-positive cognitively normal adults. Usually, a trained reader performs the visual assessment of the scan, determining whether the presence of amyloid depositions from the uptake of the radiotracer could be considered amyloid positive or negative. The experience of the trained reader is instrumental in the visual assessment especially when the scans have lower levels of beta-amyloid depositions. It will be beneficial to have an automated system to perform this task.

In recent years, artificial intelligence is influencing the medical field specifically in automatic diagnostics. Artificial intelligence is becoming a common practice in many fields (Amrutha et al., 2022; Ahila Priyadarshini et al., 2021; Vijayakumari and Rashmita, 2022; Kim and Lee, 2022; Lyu and Liu, 2021). In the medicinal field, artificial intelligence is used in the identification of pathogens, DNA sequencing, drug discovery, medical imaging, and informatics. Extensively, deep learning is applied in medical imaging (Sathananthavathi and Indumathi, 2022; López-Labraca et al., 2022; Pham et al., 2021; Jiang et al., 2021; Hwang et al., 2021). In the case of Alzheimer's disease, numerous machine learning and deep learning algorithms are used for the classification of disease stages based on MRI images (Feng et al., 2020; Divya and Shantha Selva Kumari, 2021; Divya and Kumari, 2023; Bae et al., 2021; Prasad et al., 2015). Recently, convolutional neural networks (CNN) dominate with better results in various disease classification tasks. Furthermore, using spatial and channel attention mechanisms to focus attention toward certain regions of the images has led to better performance results with deep learning classification algorithms (del Amor et al., 2021; Cheng et al., 2022; Divya and Shantha Selva Kumari, 2023). Another popular algorithm in medical image classification uses a weakly supervised technique called Multi-Instance Learning (MIL) algorithm (Hu et al., 2021; Chikontwe et al., 2021; Liu et al., 2022) where images are split into bags and are positive if at least one bag is positive otherwise negative.

Previously, machine learning algorithms were used in determining the amyloid positivity of the scans (Zukotynski et al., 2019; Kim et al., 2020). Though the recent works achieved good results, there is a need for a system that has better accuracy across different datasets so that it would be a generalized model.

In the earlier days, many machine learning and deep learning models were developed with the purpose to classify Alzheimer's disease into different stages. Abnormal amyloid deposition is one of the characteristics of Alzheimer's disease. Pfeil et al. claim that although beta-amyloid positive scans are linked to a greater possibility of progression to Alzheimer's Disease (AD) from Mild Cognitive Impairment (MCI), this binary categorization does not provide necessary information about the time for conversion to dementia (Pfeil et al., 2021). There is a greater amyloid load in precuneus, subcortical, and parietal areas in CN-to-MCI/AD progressors, and cingulate, temporal, and frontal regions in MCI-to-AD progressors. Importantly, these localized patterns indicated the progression of AD in both the short and long term, which is independent of cerebrospinal fluid biomarkers. These findings considered unique regional patterns of beta-amyloid load as a better biomarker for disease progression risk in CN and MCI patients. Hence, a model that is proposed to find the amyloid depositions in the four different brain regions could be used to analyze the scans taken for the patient over a period of time to check whether the disease has progressed or not.

While identifying amyloid positive PET scans from amyloid negative PET scans is essential, it would provide more information if the Standardized Uptake Value Ratio (SUVR) in different parts of the brain could be obtained. This will help in the targeted drug therapy for the patients. Automating this process will be an additional aid to neurologists and clinicians. When SUVR is calculated manually, initially, the PET scan must be properly aligned and registered to a MRI scan. Then, the PET scan has to be segmented based on the different regions based on MRI scan. Errors may occur throughout this process. The primary objective of this work is to identify the standardized uptake value ratio in various regions of the brain and to check whether the Florbetapir PET scan is

Table 1
Details of the scans considered in ADNI trial.

Category	CN	MCI	AD
Image Count	1052	1195	400
Age	75.58 ± 7.06	74.36 ± 8.00	76.37 ± 7.73
Gender (F/M)	565 / 487	512 / 683	174 / 226
Amyloid Positive Image Count	322	603	345
Amyloid Negative Image Count	730	592	55

positive or negative using deep learning techniques. Hence in this work, MIL attention deep learning architecture is proposed to perform a multiple regression task of finding SUVR in various regions of the brain with the usage of Florbetapir PET scans alone without the necessity for MRI scans.

The main contributions of the study are summarized as:

1) Multiple regression

Multiple Regression is proposed for predicting the frontal SUVR, lateral temporal SUVR, anterior/posterior cingulate SUVR, lateral parietal SUVR, and summary SUVR of the brain from the Florbetapir PET scans. Summary SUVR is used for determining amyloid positivity.

2) MIL attention deep learning model

Usually, MIL-based weak supervised techniques are used for classification tasks. In this proposed work MIL attention deep learning technique based on the weakly supervised technique is used for multiple regression tasks where it is used for finding SUVR in different sub-regions of the brain and tells whether the scan is amyloid positive or negative. It uses the patch-nets to extract the important features and has patch scores for those corresponding features which helps in making the target prediction.

3) Extensive experiments for validating the proposed model

The proposed model is trained, validated, and tested on Florbetapir PET scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Numerous experiments are conducted to arrive at the proposed model. The proposed model is further tested upon an external dataset of Florbetapir scans from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study.

4) Graphical User Interface

A Graphical User Interface (GUI) is developed which will aid in the usage of the proposed model for amyloid analysis of Florbetapir PET scans.

The rest of the paper is organized as follows: In Section 2, the data used for training the deep learning model, the proposed workflow, and the model design are presented. Experimental results are discussed in Section 3 and Section 4 concludes the paper.

2. Methods

2.1. Subjects

Datasets used in this study are acquired from Alzheimer Disease Neuroimaging Initiative (ADNI) (<https://adni.loni.usc.edu>) and Anti-Amyloid Treatment in Asymptomatic Alzheimer's (A4) study (Sperling et al., 2020).

The ADNI dataset consists of study participants from ADNI1, ADNI2, ADNIGO, and ADNI3 who had undergone longitudinal ¹⁸F-Florbetapir imaging over a period of years at 57 sites. A total of 2647 Florbetapir ¹⁸F [AV-45] scans of cognitive normal (CN), Mild Cognitive Impairment

Table 2
Details of the scans considered in the A4-study trial.

Category	Amyloid Positive Scans	Amyloid Negative Scans
Image Count	489	924
Age	72.23 ± 5.16	71.22 ± 4.59
Gender (F/M)	300 / 189	545 / 379

(MCI), and Alzheimer’s Disease (AD) subjects are used. The details of the scans from ADNI that are considered for developing the proposed model are shown in Table 1.

A4 study dataset consists of Florbetapir PET scans taken from participants belonging to three research groups – elevated amyloid level group, not elevated amyloid level group belonging to Longitudinal Evaluation of Amyloid Risk and Neurodegeneration (LEARN) observation, and not elevated amyloid level group not belonging to LEARN observation. A total of 1413 Florbetapir scans from all three research categories are considered in this study. The details of the scans from the A4 study that are considered for developing the proposed model are shown in Table 2.

2.2. Proposed workflow

The proposed workflow is depicted in the Fig. 1. It details the steps involved in the model development for automatic amyloid quantification in different brain regions.

Initially scans from the training dataset of ADNI are Anterior Commissure-Posterior Commissure (AC-PC) corrected and normalized to the same PET template. For this process, the origin point is set manually to the Anterior Commissure point so that the next step normalization to the PET template will occur successfully without failures. Then a deep learning model involving patch-nets, multi-instance learning, and an attention module is trained to predict standardized uptake value ratios from different regions of the brain based on the training dataset. The developed model is then tested on the normalized scans of the test dataset of ADNI and A4-study.

2.3. Florbetapir PET pre-processing

¹⁸F-Florbetapir scans which are acquired for 20 min are co-registered and resized to have uniform size of 160x160x96 are downloaded from the ADNI website (adni.loni.usc.edu). The scans are then corrected according to the AC-PC line and they are normalized to the template created from 100 Florbetapir scans of ADNI (Iaccarino et al., 2022) using MATLAB R2020b and SPM12-Statistical Parametric Mapping (https://www.fil.ion.ucl.ac.uk/spm/). After all these processes the resulting processed volumes have the size of 101x116x96 with a voxel size of 2 × 2 × 2 mm³.

Florbetapir PET scans that are downloaded from the A4 study were collected from 50 to 70 min post-injection and generally reconstructed in 4x5-minute frames available as NIfTI (Neuroimaging Informatics Technology Initiative) files. These Florbetapir PET scans from the A4 study which have different volume sizes are adjusted for the AC-PC line and normalized to the PET template created from ADNI resulting in processed volumes of size 101x116x96 with a voxel size of 2 × 2 × 2 mm³.

From the UC Berkley ¹⁸F-Florbetapir analysis data taken from the ADNI website, Standardized uptake value ratios (SUVR) for the amyloid

PET scans are downloaded. In addition to SUVR, the uptakes in different grey matter regions of the brain – frontal, cingulate, parietal, and temporal regions are also downloaded.

The brain regions considered for amyloid deposition estimation in frontal, cingulate, parietal, and temporal regions are shown in Table 3.

For performing this analysis, registration to the closest in time MRI of the patient is done which is then followed by skull stripping and then cortical and subcortical regions segmentations are done by FreeSurfer software (https://surfer.nmr.mgh.harvard.edu), and then SUVR is calculated for a particular region with reference to a composite region consisting of the white matter of the brain, cerebellum and some small brain regions. This involves many calculations and the segmentations must be checked to get the correct results. From the A4 study, the SUVR is listed for different regions than specified by the ADNI. Hence only composite reference region-based SUVR and the amyloid positive or negative result is considered for the testing of the model.

For longitudinal analysis of amyloid PET scans, amyloid positivity is determined if the standardized uptake value ratio according to the composite reference region is greater than the cut-off value of 0.78 (Landau et al., 2015). The composite reference region includes the entire cerebellum, brainstem/pons, and eroded subcortical white matter. For the ADNI dataset, SUVRs based on composite reference region are accessible. For the A4 study, SUVRs based on whole cerebellum cut-off are only accessible. When the whole cerebellum is considered, the cut-off used is 1.11 to determine amyloid positivity. The linear regression equation for converting SUVR based on the whole cerebellum to SUVR based on the composite reference region is given by eq.1 as mentioned in Florbetapir (AV45) processing methods (https://adni.loni.usc.edu)

$$y = 0.630x + 0.080 \tag{1}$$

Since longitudinal analysis is performed, the whole cerebellum cut-off is transformed to 0.78 cut-off for composite region reference using linear regression results. Similarly, SUVRs based on the composite reference region are determined using the eq.1 for scans from the A4 study.

2.4. Multi-Instance learning attention deep learning model

Multi Instance Learning (MIL) is a weakly supervised learning technique where training data is assumed to be consisting of bags of

Table 3
Brain regions considered for determining SUVR in frontal, cingulate, parietal, and temporal regions.

Frontal Region	Anterior/Posterior Cingulate Region
1. caudalmiddlefrontal	1. caudalanteriorcingulate
2. lateralorbitofrontal	2. isthmuscingulate
3. medialorbitofrontal	3. posteriorcingulate4.
4. parsopercularis	rostralanteriorcingulate
5. parsorbitalis	
6. parstriangularis	
7. rostralmiddlefrontal	
8. superiorfrontal9. frontalpole	
Lateral Parietal Region	Lateral Temporal Region
1. inferiorparietal	1. inferior temporal
2. precuneus	2. middletemporal3. superior temporal
3. superiorparietal4. supramarginal	

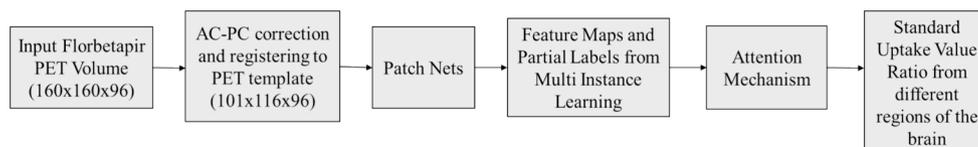


Fig. 1. Design Workflow. The steps involved in developing the proposed model to predict SUVR in different regions of the brain is explained.

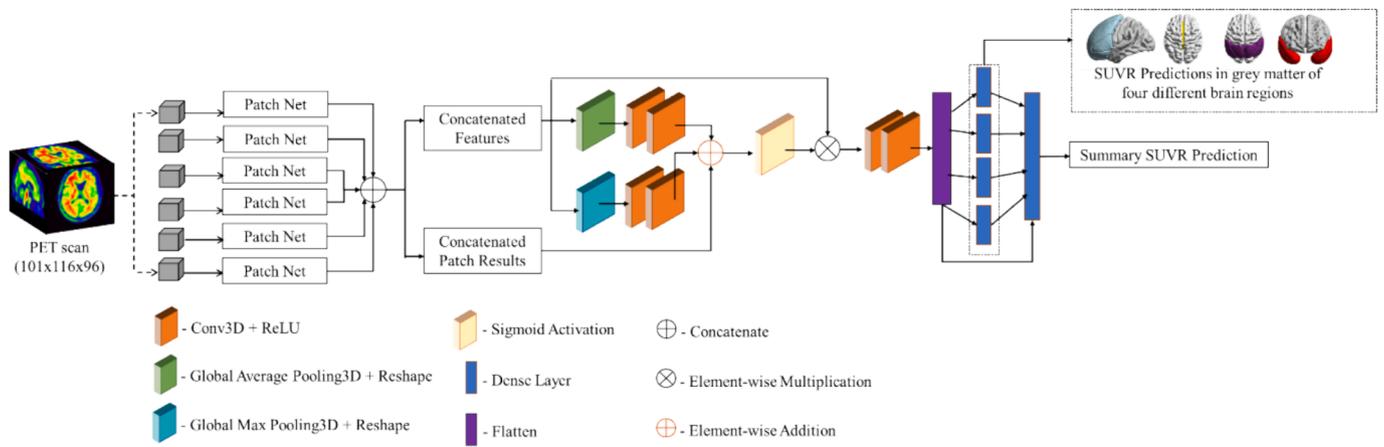


Fig. 2. Illustration of Multi-Instance Learning Attention Deep Learning Model. The different modules – patch net, features obtained from patch nets, patch scores, attention block, multi-regression are present in the proposed MIL attention deep learning model.

instances where each bag consists of a single label Y for the set of instances $X = \{x_1, x_2, \dots, x_n\}$. The individual labels for the instances (y_1, y_2, \dots, y_n) within a bag are unknown during the training process. Usually, a bag is positive if it has at least one instance as positive otherwise the bag is treated as negative. MIL formulation fits the task of amyloid analysis of PET scans. In the proposed model, the entire input scan is a bag and the instances of the bag are the non-overlapping patches. Here, the bag is assumed to be positive if it has features corresponding to the amyloid positive scan otherwise the bag is assumed to be negative.

The architecture of the deep learning model based on MIL and attention mechanism is depicted in Fig. 2.

2.4.1. Patch nets

The Patch Nets are used to extract discriminative features from each non-overlapping patch using spatial attention module. This helps in identifying the micro-functional change caused due to difference in the amyloid deposition of the brain due to Alzheimer’s disease. All the Patch Nets in the proposed model has the same architecture. The Patch Nets consist of five 3D convolutional layers and a 3D max-pooling layer. The

first two convolutional layers have a filter size to be of 5x5x5. The next two convolutional layers have a filter size of 3x3x3 and the last convolutional layer has a filter size to be 1x1x1. The number of filters for the convolutional layers 1 to 5 are 8,16,16,32, and 64 respectively. All the convolutional layers have a unit stride except convolutional layers 3 and 4 which have a stride of 2. Rectified Linear Unit (ReLU) activation is performed after every convolutional operation. The max-pooling operation has a filter size of 2 x 2 x 2 with unit stride. Two modules are created from the output of the fifth convolutional layer. One of the branches is the global average pooling followed by a dense layer with sigmoid activation. It provides a score which is used to identify the discriminative pathological locations. The other branch consists of spatial attention module. It is designed based on the proposed spatial attention in (Woo et al., 2018). Two feature maps are created along the channel axis (channel max pooling and channel average pooling) to compute spatial attention. These feature maps are concatenated and given as the input to the convolutional layer with the same padding, unit stride and a single filter of size 3 x 3 x 3. The spatial attention map is obtained by passing the output from the convolutional layer to the

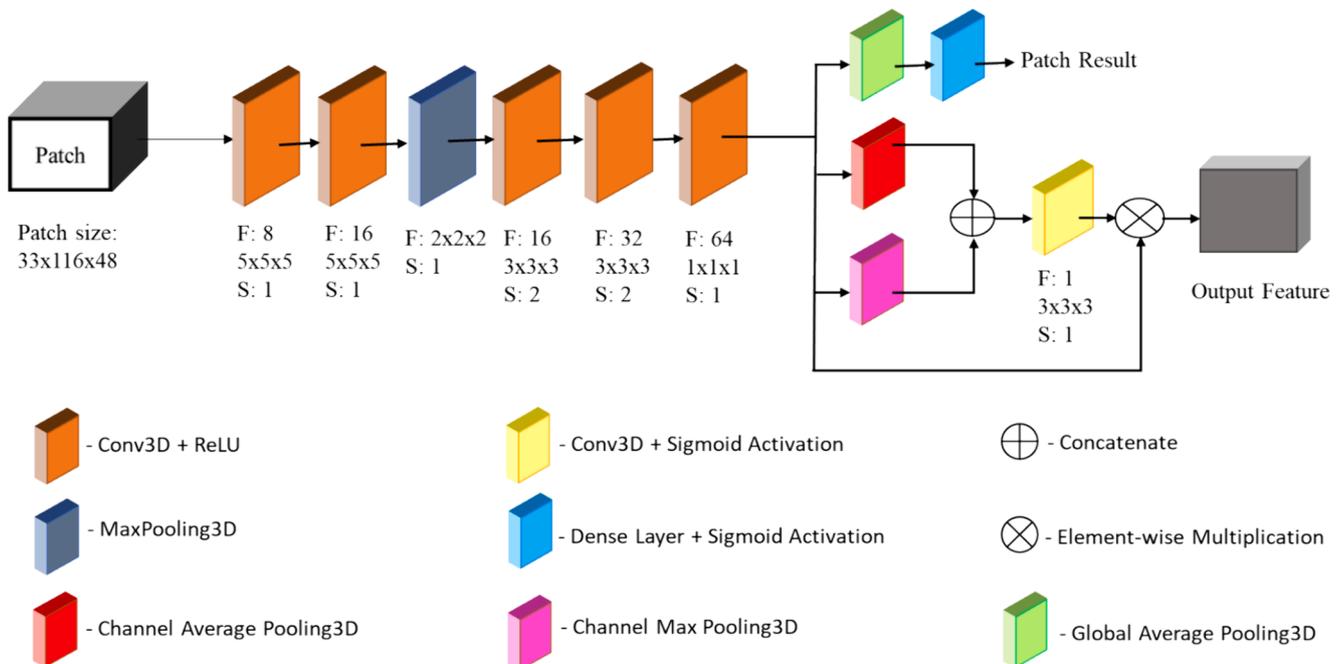


Fig. 3. Illustration of Patch Net using Multi Instance Learning. The Patch net modules are 3D CNN backbone, spatial attention and generation of the patch result.

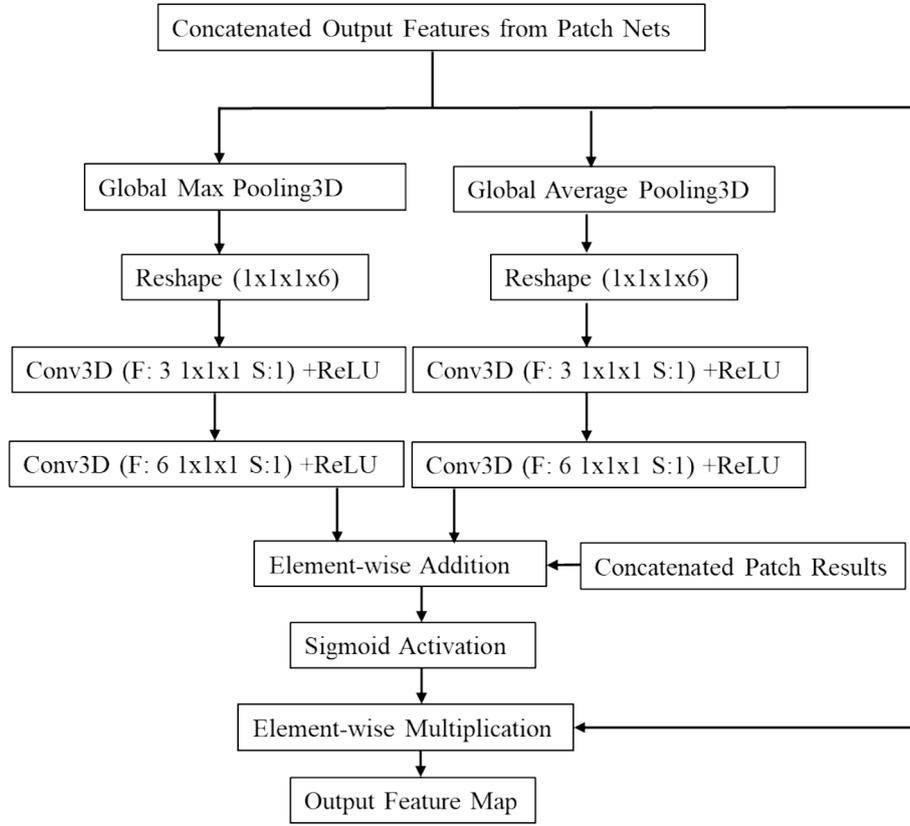


Fig. 4. Attention MIL Pooling. Global Maxpooling and Global Average Pooling is performed followed by the 3D Convolutional layers and they are added together with the patch results followed by sigmoid activation and multiplied with the input feature map.

sigmoid activation. It details the spatially varying roles of different parts of the input patches. The spatial attention map is then calculated by eq.2

$$Spatialattention(SA) = \sigma(f[Concatenate(CMP; CAP)]) \quad (2)$$

where CMP represents channel max pooling, CAP represents channel average pooling, $[\cdot]$ represents concatenate operation, f represents the convolution operation and σ denotes the sigmoid activation function. The output feature descriptors are obtained by

$$Output\ Feature\ Maps = SA \otimes Feature\ Maps\ of\ fifth\ convolution\ layer \quad (3)$$

where \otimes denotes element-wise multiplication.

The design of a single patch net is depicted in Fig. 3.

2.4.2. MIL attention module

To know how each patch influences the final prediction, an attention MIL pooling operation is performed to learn the patch-attention map. The output feature maps from the spatial attention layer in the patch-net are average pooled along the channel axis. Let this compressed feature along the channel axis be called F_i for the i^{th} patch-net. The global feature map is created by concatenating all the patch features that have been average pooled as in eq.4.

$$Global\ Feature\ Map = [F_1; F_2; F_3 \dots F_n] \quad (4)$$

where $[\cdot]$ indicates concatenate operation, n represents the number of patches. Similarly, the patch scores from the patch-nets are concatenated as in eq.5.

$$Patch\ scores = [P_1; P_2; P_3 \dots P_n] \quad (5)$$

where $[\cdot]$ indicates concatenate operation, n represents the number of patches. The global feature map is then passed to two branches simultaneously. One branch consists of global max pooling followed by two

convolution layers with ReLU activation. The first convolutional layer consists of $n/2$ filters of size $1 \times 1 \times 1$ and the second convolutional layer consists of n filters of size $1 \times 1 \times 1$. The other branch consists of global average pooling followed by two convolution layers with ReLU activation. These two convolutional layers are designed the same as in the first branch. The operation is explained in eq.6 and eq.7.

$$F_{Max} = ReLU(Conv3D(ReLU(Conv3D(GMP(Global\ Feature\ Map)))))) \quad (6)$$

$$F_{Avg} = ReLU(Conv3D(ReLU(Conv3D(GAP(Global\ Feature\ Map)))))) \quad (7)$$

The patch scores are extended to the same size as the two feature maps F_{Max} and F_{Avg} . Elementwise summation operation is performed on these three feature maps and then passed onto the sigmoid activation function which is then multiplied with the global feature map. The operation is detailed in eq.8.

$$F_{output} = Global\ Feature\ Map \otimes (\sigma(F_{Max} + F_{Avg} + Patch\ scores)) \quad (8)$$

where \otimes represents multiplication.

Instead of using the max MIL pooling and mean MIL pooling, attention MIL pooling is used to improve the performance of the model to produce better prediction results. The attention MIL pooling is depicted in Fig. 4.

2.4.3. Regression

Instead of passing the F_{output} features directly to the dense layers to know the relationship among the features, F_{output} features are given as input to two convolutional layers with $1 \times 1 \times 1$ filter size and unit stride. The number of filters used in the convolutional layers are 16 and 32 respectively. These convolutional layers help to better learn the attention-based features from MIL. ReLU activation is used after every

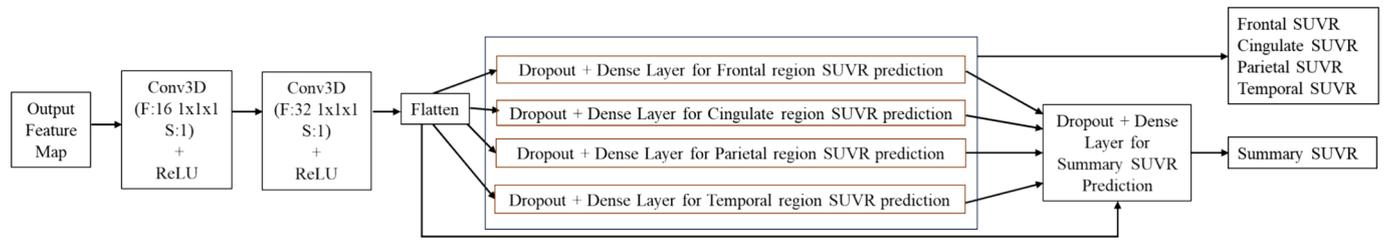


Fig. 5. Regression to determine SUVR in different brain regions. This block diagram represents how the SUVR is calculated in different brain regions and how the summary SUVR is predicted.

convolutional layer. The features are then flattened and then passed onto four dense layers in parallel. The first four dense layers have single neuron each to determine the SUVR in four different brain subregions. The results from these four dense layers and the flattened features are then concatenated and passed on to the final dense layer with single neuron to determine the summary SUVR based on composite reference region. Dropout rate of 0.3 is used before each dense layer. MIL attention-based regressor is designed to estimate the summary SUVR from a Florbetapir scan. Multiple regression is performed to calculate the SUVR in four different brain regions and these values are also used in the calculation of summary SUVR based on composite region which is depicted in Fig. 5.

The summary SUVR value is not only based on the SUVR values considered from Frontal region, Cingulate region, Parietal region, and temporal region, whereas it is also based on the other brain regions. Hence, in addition to those SUVR values, all the flattened features are considered for the prediction of Summary SUVR.

2.4.4. Performance metrics

Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to analyse the performance of the proposed model. MAE measures the average magnitude of the errors in a set of predictions, without considering the direction. RMSE gives the square root of the average of squared differences between prediction and actual observation. Their formulae are expressed in eq.9 and eq.10.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{10}$$

3. Results and discussions

Two datasets – ADNI and A4 study are used to evaluate the proposed multi-instance learning and attention-based model. ADNI dataset is separated as 80 % for training and validation and 20 % for hold-out testing. Training and validation data consists of 2116 florbetapir PET scans from ADNI. A separate 531 florbetapir PET scans from ADNI are used as a hold-out test set to evaluate the model. Since longitudinal data is considered, PET scans are split patient-wise to ensure no data leak among training, validation, and test sets. The developed model is further tested on another external dataset A4 study from which 1413 scans are used. Instead of loading all the data into RAM at the same time, TensorFlow records are created for easy batch processing of the deep learning system. MAE loss and RMSE metric is used for training the model. Early stopping is used during training to monitor the validation loss with the patience of 30 epochs. The learning rate starts from 0.01 and is reduced by a factor of 0.1 till it reaches 1E-4 when the validation loss does not decrease for 10 epochs. Adam optimiser is used during the training of the deep learning model. The proposed model is implemented with TensorFlow and Keras framework using Intel Xeon Silver 4210R processor with 10 cores. Weights are shared by all the Patch-Nets,

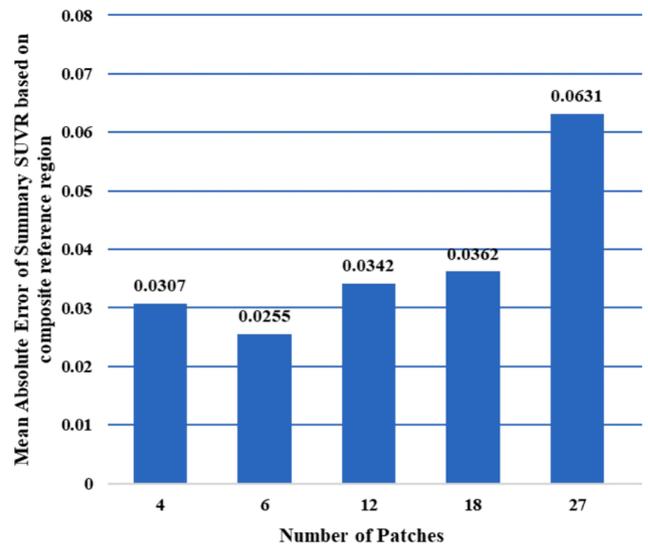


Fig. 6. Mean Absolute Error of Summary SUVR based on composite reference region according to the number of input patches. This bar chart represents the MAE measures found when different patches are considered. These MAE values are for the validation data.

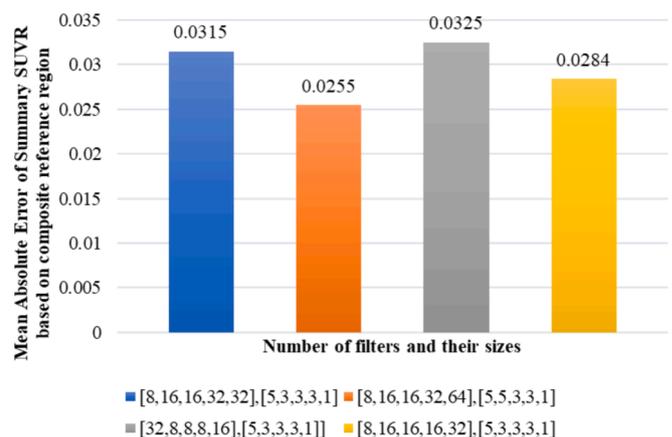


Fig. 7. Mean Absolute Error of Summary SUVR based on composite reference region according to the number of filters and their sizes in the conv layers of patch-nets. This bar chart represents the MAE measures found when different number of filters and their sizes are considered. These MAE values are for the validation data.

which in turn reduces the number of trainable parameters. Various experiments are performed to choose the hyperparameters for the model by training on the ADNI training dataset. Finally, the model with chosen hyperparameters is used to test on the hold out ADNI dataset. The final model is also used to test the A4-study dataset to ensure generalizability

Table 4
MAEs and RMSEs estimated on the amyloid positive and negative scans.

Category	Mean Absolute Error		Root Mean Square Error	
	Amyloid Positive Scans	Amyloid Negative Scans	Amyloid Positive Scans	Amyloid Negative Scans
	ADNI test data			
Frontal SUVR	0.0727	0.0525	0.0938	0.0667
Anterior/Posterior Cingulate SUVR	0.0784	0.0694	0.1035	0.0869
Lateral Parietal SUVR	0.0703	0.0518	0.0894	0.0676
Lateral Temporal SUVR	0.0705	0.0485	0.0901	0.0631
Summary SUVR	0.0274	0.0215	0.0370	0.0268
	A4 study test data			
Summary SUVR	0.0432	0.0354	0.0573	0.0448

and robustness of the model.

Different validation runs are performed by varying the number of input non-overlapping patches of the Florbetapir PET scan. The number of patches that are tested for are 4,6,12,18, and 27. The resulting MAEs are shown in Fig. 6. When the number of patches is increased, the MAE value also increases for the summary SUVR. An optimum value is reached when the number of input patches is 6.

Different validation runs are performed by varying the number of filters and their sizes for the convolution layers of the patch-nets as [(8,16,16,32,32), (5,3,3,3,1)], [(32,8,8,8,16), (5,3,3,3,1)], [(8,16,16,32,64), (5,5,3,3,1)], and [(8,16,16,16,32), (5,3,3,3,1)]. The results are depicted in Fig. 7.

After the validation processes, the number of filters for the final model are chosen as (8,16,16,32,64) and the corresponding filter sizes are chosen as $5 \times 5 \times 5$, $5 \times 5 \times 5$, $3 \times 3 \times 3$, $3 \times 3 \times 3$, and $1 \times 1 \times 1$ respectively based on the lowest MAE obtained for the summary SUVR. The number of input patches is selected to be 6. This model is used for five-fold cross-validation with training and validation data considered together, utilising MAE as the performance metric. To prevent data leaking, the split is designed such that scans of a certain individual belong entirely within each fold and are not split between folds. Upon five-fold cross-validation, the MAE of 0.0258 ± 0.0002 and RMSE of 0.0340 ± 0.0002 is achieved for summary SUVR. The training time of the MIL attention-based model for each fold took nearly three hours. Since the model performance is roughly the same for each fold during five-fold cross validation the same number of filters with their corresponding filter sizes and 6 patches are chosen for the final model. Testing was completed within a few seconds upon running this developed model on the Intel Xeon Silver 4210R processor. The hold out ADNI test dataset and A4 study test dataset are used only after the finalisation of the model selection and completion of hyperparameter tuning.

MAEs and RMSEs for different brain subregions SUVR are listed in Table 4 with respect to the amyloid positive scans and amyloid negative scans from both ADNI test data and A4 study test data.

The proposed model achieves MAE of 0.0621 for frontal SUVR, 0.0737 for anterior/posterior cingulate SUVR, 0.0605 for lateral parietal SUVR, 0.0589 for lateral temporal SUVR, and 0.0243 for cortical composite SUVR based on composite reference region for the entire ADNI test set. RMSE values are found to be 0.0807 for frontal SUVR, 0.0951 for anterior/posterior cingulate SUVR, 0.0787 for lateral parietal SUVR, 0.0771 for lateral temporal SUVR, and 0.0320 for cortical composite SUVR based on composite reference region for the entire ADNI test set. When tested on the A4-study dataset the proposed model achieves MAE of 0.038 and RMSE of 0.0495 for cortical composite SUVR based on the composite region. Amyloid positivity is decided if the cortical composite

Table 5
Performance comparison of different works on ADNI.

Model	RMSE on ADNI Test data	MAE on ADNI Test data
ResNet50	0.0794	0.0595
ResNet101	0.0722	0.0559
ResNet152	0.0666	0.0504
VGG19	0.1488	0.1293
Reith F et al. (2020) (Reith et al., 2020)	0.059 ± 0.005	–
Reith F et al. (2021) (Reith et al., 2021)	0.0339 ± 0.0003	–
Maddury S et al. (2023) (Maddury and Desai, 2023)	–	0.0441
Proposed MIL attention model	0.0320	0.0243

SUVR based on the composite region is greater than 0.78. Based on this calculation, accuracy of 97.36 %, sensitivity of 97.85 % and specificity of 96.82 is achieved on ADNI test set while determining amyloid positive or amyloid negative scans from predicted summary SUVR. When tested upon the A4 study test set, the proposed model achieves 94.55 % accuracy, 94.27 % sensitivity and 94.70 % specificity while using predicted summary SUVR for determining amyloid positivity.

The proposed multi-instance learning attention-based model for SUVR regression is compared with the Residual Network (ResNet) and Visual Geometry Group (VGG)-19 models constructed using transfer learning by fine-tuning ResNet and VGG-19 weights pre-trained on the ImageNet dataset of natural pictures in Table 5. Table 5 shows that the proposed multi-instance learning attention model has the lowest RMSE and MAE when compared to the other models. The slices 40, 50, and 60 of the 96 scan slices are utilised as input to the ResNet and VGG19 models, respectively, while the complete volume $101 \times 116 \times 96$ is given as input to the proposed model.

The proposed model has 59,863 parameters. When compared to existing models, the proposed model has fewer parameters but provides superior results with less MAE and RMSE. This is because multiple brain regions are considered for the computation of SUVR in PET scans, which may not be the case when only three slices in the brain region are considered.

The performance of the proposed multi-instance learning attention model is also analysed by comparing it with the other recent works. In (Kim et al., 2019), a 3D convolutional neural network was proposed to estimate the composite SUVR on ADNI Florbetapir PET scans test data with the mean absolute error of 0.060. In (Reith et al., 2020), ResNet-50 and ResNet-152 were used to predict SUVR in Florbetapir PET scans with a mean RMSE score of 0.059 upon cross-validation on the ADNI dataset. In (Reith et al., 2021), ResNet-50 and Gradient boosting decision tree algorithm was used to predict the SUVR in Florbetapir PET scans with a root mean squared error of 0.0339 ± 0.0027 in the ADNI dataset. In (Maddury and Desai, 2023), RegNet X064 and gradient boosting tree were utilized with three input slices to achieve MAE of 0.0441. The proposed model achieved MAE of 0.0258 ± 0.0002 and RMSE of 0.0340 ± 0.0002 for summary SUVR upon five-fold cross-validation. The proposed multi-instance learning attention-based model works well when compared with the recent works by predicting summary SUVR with MAE of 0.0243 and RMSE of 0.0320 on ADNI test set and MAE of 0.038 and RMSE of 0.0495 on A4 study test set. From Table 5, RMSE and MAE are the lowest for the proposed MIL attention model when compared with the other recent works. The proposed method provides better accuracy in detecting amyloid positive scans across different datasets.

The longitudinal Florbetapir PET scans from the ADNI dataset contains participants belonging to CN, MCI, and AD disease category having amyloid positive and amyloid negative scans. The performance of the proposed multi-instance learning attention model for the entire test data, test data containing only amyloid positive scans, and test data containing only amyloid negative scans are shown in Fig. 8 and Fig. 9 for

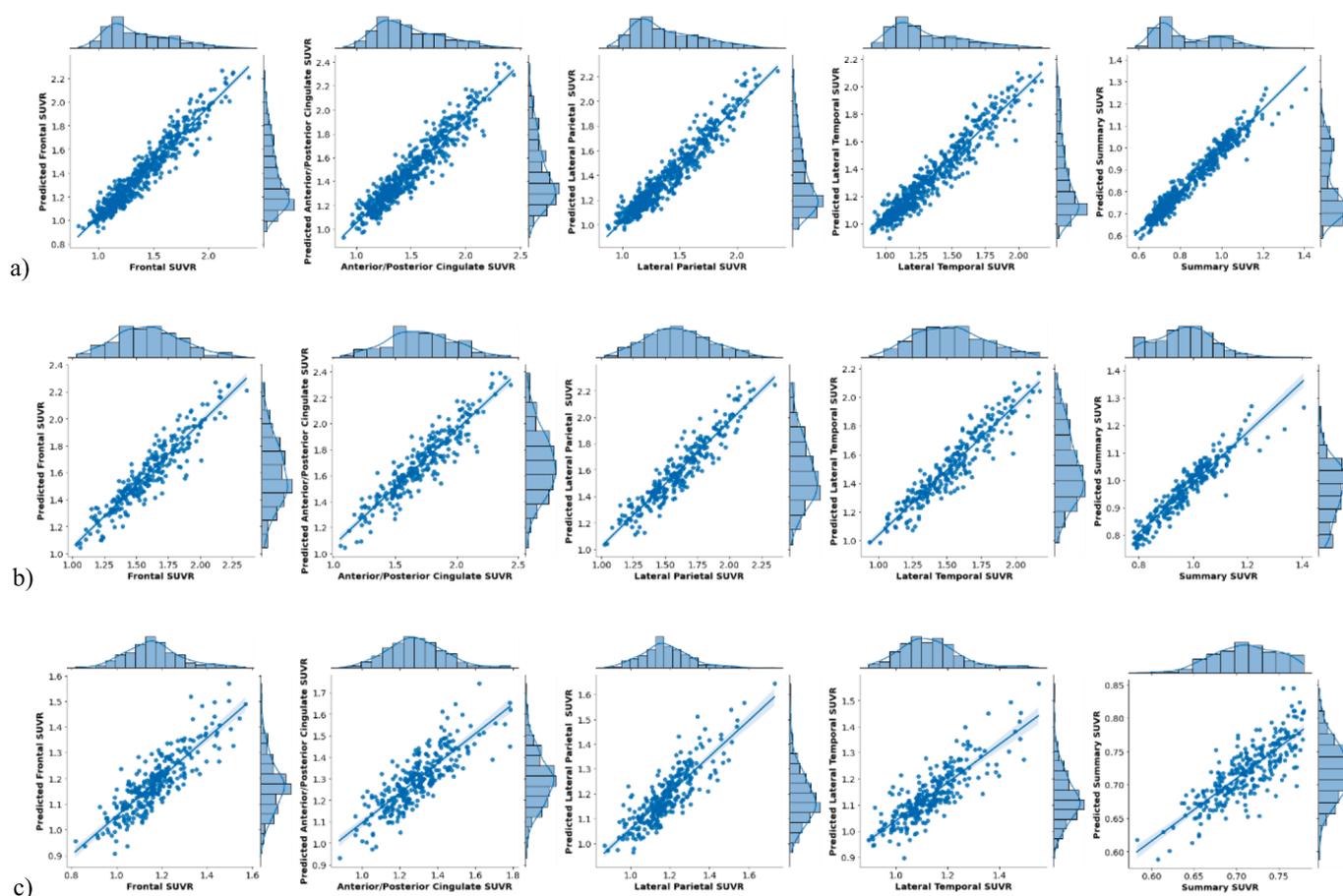


Fig. 8. Scatter Plot of SUVR and predicted SUVR in different regions of the brain with their corresponding histograms on the a) ADNI test data b) Amyloid positive ADNI test data C) Amyloid negative ADNI test data. It represents the scatter of the predicted SUVR values of ADNI test data from the target values.

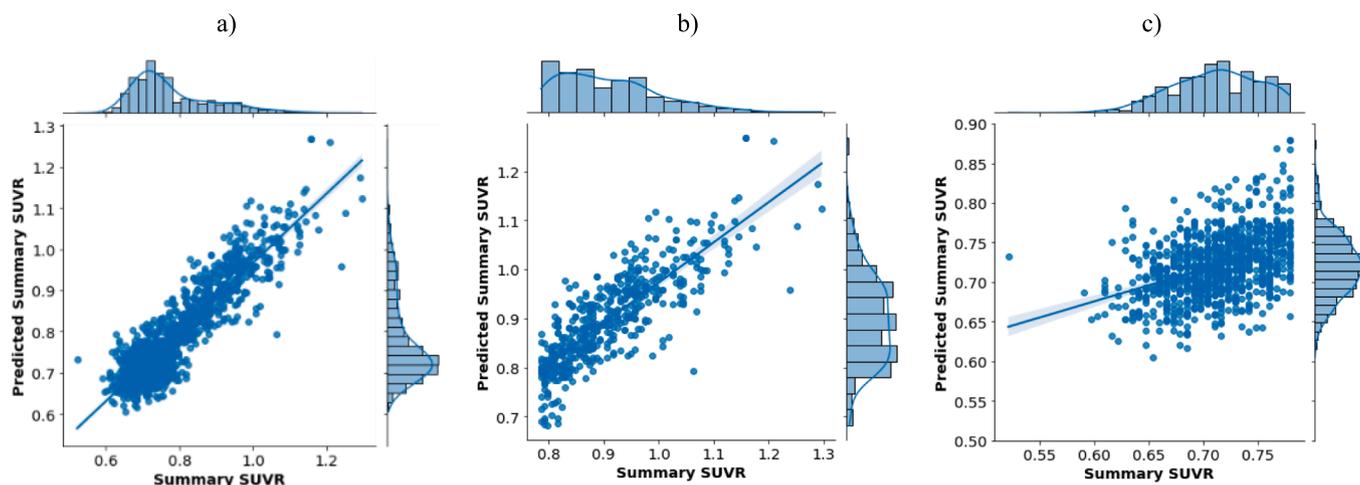


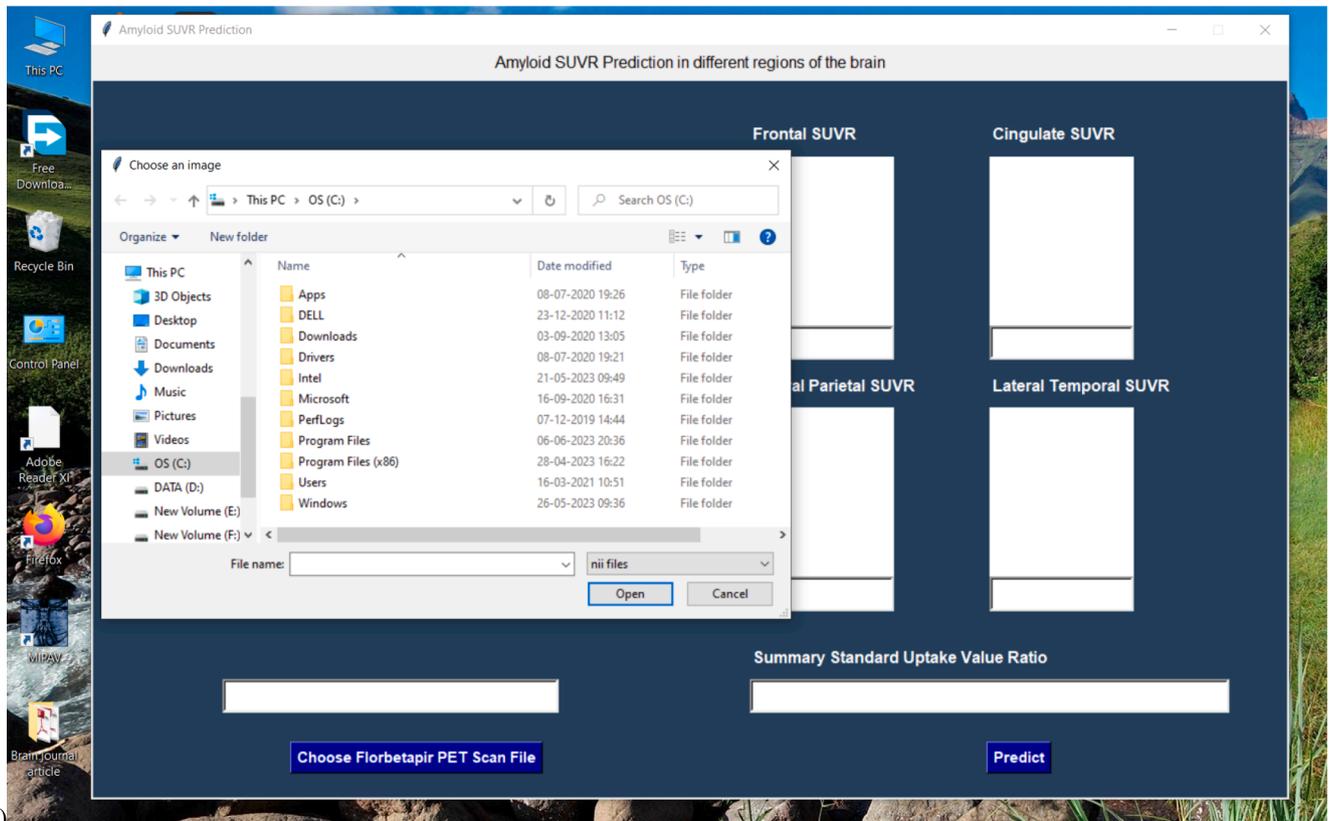
Fig. 9. Scatter Plot of Summary SUVR and predicted summary SUVR with their corresponding histograms on the a) A4-study test data b) Amyloid positive A4-study test data c) Amyloid negative A4-study test data. It represents the scatter of the predicted SUVR values of A4 study test data from the target values.

both ADNI and A4 study datasets.

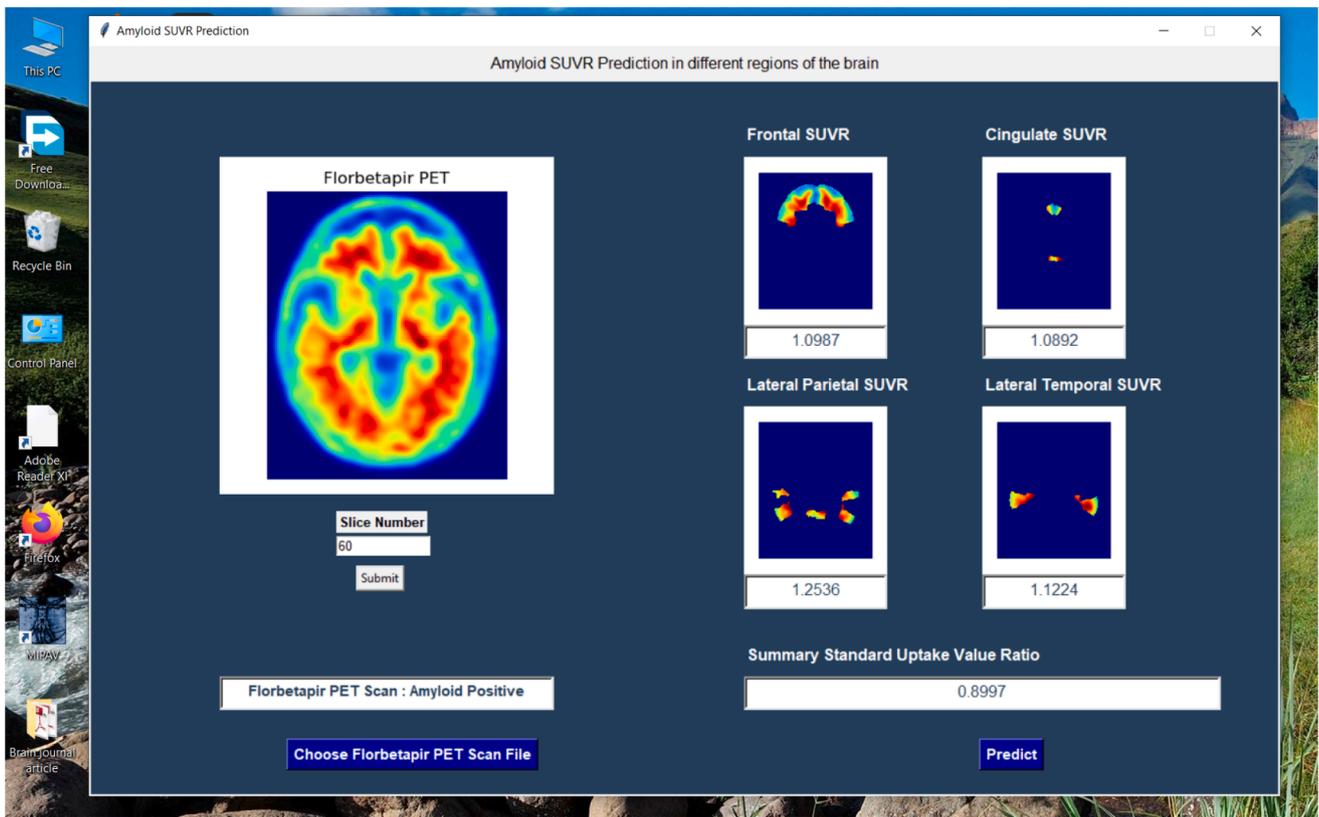
Histograms are incorporated into the scatter plots of the test dataset to visualize the distribution of predicted SUVRs and actual SUVRs along different brain regions. From Fig. 8 and Fig. 9, the scatter is less in amyloid positive scans when compared to the amyloid negative scans. This shows that SUVR for amyloid positive scans are predicted with less errors leading to better diagnosis.

In this work, the amyloid SUVR quantification of Florbetapir PET

scans across different brain regions is carried out which is then used for determining amyloid positivity. Targeted drugs for amyloid pathology in neurodegenerative diseases is a possibility when SUVRs can be detected over different cortical regions of the brain. Future investigations can be carried out to include MMSE scores, apolipoprotein (APOE) allele information, and other modalities of imaging to explore further in this field of amyloid quantification and Alzheimer’s disease diagnosis.



a)



b)

Fig. 10. GUI Interface for the proposed MIL-attention model a) Input Screen b) Sample Prediction. It displays how the file is selected using 'Choose Florbetapir PET Scan File' button and how the predictions appear on the screen when the 'Predict' button is clicked on.

Based on the proposed multi-instance learning attention model, a graphical user interface (GUI) is created using tkinter library. The GUI is designed such that the users who like to know the value of SUVR in different regions of the brain of a patient's amyloid PET scan can simply upload the file into the system. By using the GUI, the scan file can be easily selected by clicking on the button that says 'Choose Florbetapir PET Scan File'. Only .nii files can be selected using this GUI. This graphical user interface is shown in Fig. 10.

The screen of the graphical user interface is split into two frames such that the left frame displays a slice of the Florbetapir PET scan and tells whether the scan is amyloid positive or not. The right frame displays the SUVR predictions across different brain. This GUI is an easy tool to operate with little computer knowledge without knowing the complexity behind the model to predict the results.

4. Conclusion

In this work, multi-instance learning, spatial attention, and channel attention are exploited to create the deep learning model. The proposed model has its unique ability to quantify the standardized uptake value ratio in the different regions of the brain based on Florbetapir PET scans and determine whether the scan is amyloid positive or negative. The proposed model having been tested on a large number of scans from ADNI and A4-study proves that the proposed model is robust in detecting summary SUVR with 0.0243 MAE for ADNI test scans and 0.038 MAE for A4-study scans. The GUI developed is easier to use for finding the predictions without knowing the complexity behind the model to perform the predictions. The automated amyloid analysis of the scans will be valuable in the Alzheimer's prognosis.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRedit authorship contribution statement

R. Divya: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **R. Shantha Selva Kumari:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

ADNI Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). Data collection and sharing for this project were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

A4 study data used in the preparation of this article were obtained from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's study.

The Anti-Amyloid Treatment in Asymptomatic Alzheimer's study (the A4 Study for short) is a landmark public-private partnership, funded by the National Institute on Aging/NIH, Eli Lilly and Company, and several philanthropic organizations. The A4 trial is coordinated by the University of Southern California's Alzheimer's Therapeutic Research Institute, with study sites in multiple locations.

Ethical Approval

Approval for the ADNI protocol and A4 study protocol has been granted by the institutional review board at the study sites.

Consent to participate

Not applicable.

Consent for publication

All the authors agreed to the publication of the article.

Authors' contributions

Conception and study design (all authors); statistical analyses (all authors); interpretation of results (all authors); drafting the manuscript or revising it critically for important intellectual content (all authors); approval of the final version to be published and agreement to be accountable for integrity and accuracy of all aspects of the work (all authors).

References

- Ahila Priyadarshini, R., Arivazhagan, S., Arun, M., 2021. A deep learning approach for person identification using ear biometrics. *Appl. Intell.* 51, 2161–2172. <https://doi.org/10.1007/s10489-020-01995-8>.
- Amrutha, E., Arivazhagan, S., Sylvia Lilly Jebarani, W., 2022. MixNet: A Robust Mixture of Convolutional Neural Networks as Feature Extractors to Detect Stego Images Created by Content-Adaptive Steganography. *Neural Process. Lett.* 54, 853–870. <https://doi.org/10.1007/s11063-021-10661-0>.
- Bae, J., Stocks, J., Heywood, A., et al., 2021. Transfer learning for predicting conversion from mild cognitive impairment to dementia of Alzheimer's type based on a three-dimensional convolutional neural network. *Neurobiol. Aging* 99, 53–64. <https://doi.org/10.1016/j.neurobiolaging.2020.12.005>.
- Cheng, Z., Qu, A., He, X., 2022. Contour-aware semantic segmentation network with spatial attention mechanism for medical image. *Vis. Comput.* 38, 749–762. <https://doi.org/10.1007/s00371-021-02075-9>.
- Chikontwe, P., Luna, M., Kang, M., et al., 2021. Dual attention multiple instance learning with unsupervised complementary loss for COVID-19 screening. *Med. Image Anal.* 72, 102105 <https://doi.org/10.1016/j.media.2021.102105>.
- del Amor, R., Launet, L., Colomer, A., et al., 2021. An attention-based weakly supervised framework for spitzoid melanocytic lesion diagnosis in whole slide images. *Artif. Intell. Med.* 121, 102197 <https://doi.org/10.1016/j.artmed.2021.102197>.
- Divya R, Kumari RSS, Initiative for the ADN (2023) Detection of Alzheimer's disease from temporal lobe grey matter slices using 3D CNN. Doi: 101080/1368219920232173548 70:578–587.
- Divya R, Shantha Selva Kumari R, Initiative the ADN (2021) Genetic algorithm with logistic regression feature selection for Alzheimer's disease classification. *Neural Comput. Appl.* 33:8435–8444. Doi: 10.1007/s00521-020-05596-x.
- Divya, R., Shantha Selva Kumari, R., 2023. SUVR quantification using attention-based 3D CNN with longitudinal Florbetapir PET images in Alzheimer's disease. *Biomed. Signal Process. Control* 86, 105254. <https://doi.org/10.1016/j.bspc.2023.105254>.
- Feng, J., Zhang, S.-W., Chen, L., 2020. Identification of Alzheimer's disease based on wavelet transformation energy feature of the structural MRI image and NN classifier. *Artif. Intell. Med.* 108, 101940 <https://doi.org/10.1016/j.artmed.2020.101940>.
- Hampel, H., Broich, K., Hoessler, Y., Pantel, J., 2009. Biological markers for early detection and pharmacological treatment of Alzheimer's disease. *Dialogues Clin. Neurosci.* 11, 141–157. <https://doi.org/10.31887/DCNS.2009.11.2/hhampel>.
- Hu, T., Zhang, L., Xie, L., Yi, Z., 2021. A multi-instance networks with multiple views for classification of mammograms. *Neurocomputing* 443, 320–328. <https://doi.org/10.1016/j.neucom.2021.02.070>.
- Hwang, I., Yeon, E.K., Lee, J.Y., et al., 2021. Prediction of brain age from routine T2-weighted spin-echo brain magnetic resonance images with a deep convolutional neural network. *Neurobiol. Aging* 105, 78–85. <https://doi.org/10.1016/j.neurobiolaging.2021.04.015>.
- Iaccarino, L., la Joie, R., Koeppe, R., et al., 2022. rPOP: Robust PET-only processing of community acquired heterogeneous amyloid-PET data. *Neuroimage* 246, 118775. <https://doi.org/10.1016/j.neuroimage.2021.118775>.
- Jack, C.R., Knopman, D.S., Jagust, W.J., et al., 2013. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* 12, 207–216. [https://doi.org/10.1016/S1474-4422\(12\)70291-0](https://doi.org/10.1016/S1474-4422(12)70291-0).
- Jiang, M., Zhai, F., Kong, J., 2021. A novel deep learning model DDU-net using edge features to enhance brain tumor segmentation on MR images. *Artif. Intell. Med.* 121, 102180 <https://doi.org/10.1016/j.artmed.2021.102180>.
- Kim, J.P., Kim, J., Kim, Y., et al., 2020. Staging and quantification of florbetapir PET images using machine learning: impact of predicted regional cortical tracer uptake and amyloid stage on clinical outcomes. *Eur. J. Nucl. Med. Mol. Imaging* 47, 1971–1983. <https://doi.org/10.1007/s00259-019-04663-3>.

- Kim, D., Lee, J., 2022. Predictive evaluation of spectrogram-based vehicle sound quality via data augmentation and explainable artificial intelligence: Image color adjustment with brightness and contrast. *Mech. Syst. Sig. Process.* 179, 109363 <https://doi.org/10.1016/j.ymssp.2022.109363>.
- Kim, J.-Y., Suh, H.Y., Ryoo, H.G., et al., 2019. Amyloid PET Quantification Via End-to-End Training of a Deep Learning. *Nucl. Med. Mol. Imaging* 53, 340–348. <https://doi.org/10.1007/s13139-019-00610-0>.
- Landau, S.M., Fero, A., Baker, S.L., et al., 2015. Measurement of longitudinal β -amyloid change with 18F-florbetapir PET and standardized uptake value ratios. *J. Nucl. Med.* 56, 567–574. <https://doi.org/10.2967/jnumed.114.148981>.
- Liu, S., Yang, E., Liu, Y., Zhao, S., 2022. DMA-Net: Dual multi-instance attention network for X-ray image classification. *IET Image Process* N/a. <https://doi.org/10.1049/ipr2.12560>.
- López-Labraca, J., González-Díaz, I., Díaz-de-María, F., Casado, A.F., 2022. An interpretable CNN-based CAD system for skin lesion diagnosis. *Artif. Intell. Med.*, 102370 <https://doi.org/10.1016/j.artmed.2022.102370>.
- Lyu, W., Liu, J., 2021. Artificial Intelligence and emerging digital technologies in the energy sector. *Appl. Energy* 303, 117615. <https://doi.org/10.1016/j.apenergy.2021.117615>.
- Maddury, S., Desai, K., 2023. DeepAD: A deep learning application for predicting amyloid standardized uptake value ratio through PET for Alzheimer's prognosis. *Front. Artif. Intell.* 6, 4. <https://doi.org/10.3389/FRAI.2023.1091506/BIBTEX>.
- Pfeil, J., Hoenig, M.C., Doering, E., et al., 2021. Unique regional patterns of amyloid burden predict progression to prodromal and clinical stages of Alzheimer's disease. *Neurobiol. Aging* 106, 119–129. <https://doi.org/10.1016/J.NEUROBIOLAGING.2021.06.014>.
- Pham, V.-T., Tran, T.-T., Wang, P.-C., et al., 2021. EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images. *Artif. Intell. Med.* 115, 102065 <https://doi.org/10.1016/j.artmed.2021.102065>.
- Prasad, G., Joshi, S.H., Nir, T.M., et al., 2015. Brain connectivity and novel network measures for Alzheimer's disease classification. *Neurobiol. Aging* 36, S121–S131. <https://doi.org/10.1016/J.NEUROBIOLAGING.2014.04.037>.
- Reith, F., Koran, M.E., Davidzon, G., Zaharchuk, G., 2020. Application of deep learning to predict standardized uptake value ratio and amyloid status on 18F-florbetapir PET using ADNI data. *Am. J. Neuroradiol.* 41, 980–986. <https://doi.org/10.3174/ajnr.A6573>.
- Reith, F.H., Mormino, E.C., Zaharchuk, G., 2021. Predicting future amyloid biomarkers in dementia patients with machine learning to improve clinical trial patient selection. *Alzheim. Dement. (new York, N y)* 7, e12212.
- Sathananthavathi, V., Indumathi, G., 2022. Atrous Fully Convolved Depth Concatenated Neural Network with Enriched Encoder for Retinal Artery-Vein Classification. *IETE J. Res.* 1–10 <https://doi.org/10.1080/03772063.2022.2098181>.
- Sperling, R.A., Donohue, M.C., Raman, R., et al., 2020. Association of Factors With Elevated Amyloid Burden in Clinically Normal Older Individuals. *JAMA Neurol.* 77, 735–745. <https://doi.org/10.1001/jamaneurol.2020.0387>.
- Verde, F., 2022. Tau proteins in blood as biomarkers of Alzheimer's disease and other proteinopathies. *J. Neural Transm.* 129, 239–259. <https://doi.org/10.1007/s00702-022-02471-y>.
- Vijayakumari, B., Rashmita, S., 2022. Abnormality Detection in Kidney Ultrasound Images by Various Classifiers with FPGA. *IETE J. Res.* 1–10 <https://doi.org/10.1080/03772063.2022.2089245>.
- Woo S, Park J, Lee J-Y, Kweon IS (2018) CBAM: Convolutional Block Attention Module. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y (eds) *Computer Vision – ECCV 2018*. Springer International Publishing, Cham, pp. 3–19.
- Zukotynski, K., Gaudet, V., Kuo, P.H., et al., 2019. The Use of Random Forests to Classify Amyloid Brain PET. *Clin. Nucl. Med.* 44.