Mild Cognitive Impairment Stratification by Artificial Intelligence on FDG PET

Joan Prats-Climent

Instituto de Instrumentacion para Imagen Molecular

Maria Teresa Gandia-Ferrero (✉ maite.21.xu@gmail.com)

Health Research Institute La Fe: Instituto de Investigacion Sanitaria La Fe <https://orcid.org/0000-0003-0072-057X>

Irene Torres-Espallardo

La Fe University and Polytechnic Hospital Medical Imaging Clinical Area: Hospital Universitario La Fe
Area Clinica de Imagen Medica

Lourdes Álvarez-Sánchez

La Fe University and Polytechnic Hospital: Hospital Universitari i Politecnic La Fe

Begoña Martínez-Sanchis

La Fe University and Polytechnic Hospital Medical Imaging Clinical Area: Hospital Universitario La Fe
Area Clinica de Imagen Medica

Consuelo Cháfer-Pericás

La Fe University and Polytechnic Hospital: Hospital Universitari i Politecnic La Fe

Ignacio Gómez-Rico-Junquero

Health Research Institute La Fe: Instituto de Investigacion Sanitaria La Fe

Leonor Cerdá-Alberich

Health Research Institute La Fe: Instituto de Investigacion Sanitaria La Fe

Fernando Aparici-Robles

La Fe University and Polytechnic Hospital Medical Imaging Clinical Area: Hospital Universitario La Fe
Area Clinica de Imagen Medica

Miquel Baquero-Toledo

La Fe University and Polytechnic Hospital: Hospital Universitari i Politecnic La Fe

María José Rodríguez-Álvarez

Instituto de Instrumentacion para Imagen Molecular

Luis Martí-Bonmatí

La Fe University and Polytechnic Hospital Medical Imaging Clinical Area: Hospital Universitario La Fe
Area Clinica de Imagen Medica

Keywords: PET, Artificial Intelligence, Deep Learning, Alzheimer, Mild Cognitive Impairment, Neurodegenerative diseases

Posted Date: December 20th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1166745/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Purpose

The purpose of this project is to develop and externally validate a Deep Learning (DL) FDG PET imaging algorithm able to identify patients with Alzheimer's Disease (AD), Frontotemporal Degeneration (FTD) and Dementia with Lewy Bodies (DLB) among a group of patients with Mild Cognitive Impairment (MCI).

Methods

A 3D Convolutional neural network, trained using images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, was implemented. The ADNI dataset used for training and testing the model consisted of 822 subjects (472 AD and 350 MCI). The external validation was performed on an independent dataset from our hospital. The hospital real world dataset contains 90 subjects with MCI: 71 patients that developed a neurodegenerative disease (64 AD, 4 FTD and 3 DLB) and 19 subjects without associated neurodegenerative disease.

Results

The ADNI model had 79% accuracy, 88% sensitivity and 71% specificity in the identification of patients with neurodegenerative diseases tested on the 10% ADNI dataset, achieving an area under the receiver operating characteristic curve (AUC) of 0.897. When used for the external validation, the model preserved 77% accuracy, 75% sensitivity, 84% specificity and 0.860 area under the ROC curve.

Conclusion

This model based on FDG PET images can help the early non-invasive prediction of neurodegenerative diseases in MCI patients in standard clinical settings with an overall 77% classification accuracy.

Introduction

Mild cognitive impairment (MCI) can stay stable or progress to dementia. Dementia is a clinical syndrome defined by a persistent deterioration of higher brain functions, such as memory, orientation, calculation, language, and spatial perception. This deterioration entails a loss of the patient's autonomy, and a detriment to their social, work and leisure activity.

Dementia cases increase exponentially with age, with a true worldwide epidemic due to the progressive aging of the population. Dementia is the main cause of dependency and disability in the elderly, and entails a very significant morbidity, which involves a major economic, social and health cost, that principally falls on family members [1].

Alzheimer's disease (AD) is the most common form of dementia. According to the 2016 report by Alzheimer's Disease International, it is estimated that around 46 million people suffer from dementia worldwide, with projection of 131 million people in 2050 [2]. An early recognition of patients with dementia presenting with MCI is clinically challenging, specific diagnosis might be delay before deterioration is evident. Due to the high number of patients with MCI, the ability to diagnose and classify early stage dementia will have an impact on reducing the cost of long-time care by making more knowledgeable decisions regarding clinical interventions and treatment planning [3].

In the last years, developments of diagnostic tests improved the detection of dementia patients at MCI stage. Cerebrospinal fluid (CSF) amyloid-beta 42 ($A\beta_{42}$), $A\beta_{40}$, $A\beta_{42} / A\beta_{40}$ ratio, threonine-181-phosphorylated-tau (p-tau), and total-tau (t-tau) are reliable biomarkers for amyloidosis, tauopathy, and neurodegeneration. These biomarkers allow to distinguish AD at MCI stages. CSF, obtained by lumbar puncture, is an invasive procedure that should not be used in subjects with MCI but a low probability to have a neurodegenerative dementia.

Positron emission tomography (PET) provides a non-invasive quantification of brain metabolism, having been in the detection and diagnosis of neurodegenerative diseases even in the initial stages when the patient only presents with MCI. Although Amyloid PET is recognised as the tracer of choice, availability and price might be a limiting factor. The most common and easily accessed PET radiopharmaceutical [^{18}F]FDG [4].

FDG PET has demonstrated diagnostic and prognostic utility in the evaluation of patients with cognitive impairment and in the distinction between primary neurodegenerative disorders and other cognitive impairment [5, 6, 7]. However, sometimes it can be challenging to subjectively differentiate between MCI and neurodegenerative pathologies. To improve the precision and predictivity of FDG PET in this setting, a Deep Neural Network model was trained and externally validate to identify patients with an associated neurodegenerative disease among a group of MCI patients.

Materials And Methods

Participants

Patients with MCI in the neuropsychological evaluation in absence of functional impairment in daily living activities were recruited. MCI was defined by a regulated neuropsychological evaluation with a global Clinical Dementia Rating (CDR) scale equal to 0.5 (questionable dementia). All cases were retrospectively recruited from the Cognitive Disorders Consultation of the Neurology Service. Patient's data was obtained from the hospital electronic clinical records and images from the Picture Archiving and Communication System (PACS), from 2013 through 2020. Patients who presented with cognitive functions deterioration, mainly memory, in the absence of functional impairment for activities of daily living, were included. Those patients without an FDG PET imaging were excluded. Approval from the

Ethics Committee was obtained for the study and the patients information consent was waived due to the retrospective non-interventional observational research on data design of the study (**Fig. 1**).

Finally, 90 subjects who met the eligibility criteria were included. The final diagnosis was established after following up for at least one year, considering physical and neuropsychological examination, structural MR imaging, and CSF molecular analysis. Furthermore, once the patients were chosen meeting the inclusion criteria, the database was revised for greater precision in the clinical diagnosis, considering the follow-up of the patient. Patients were stratified into 71 MCI cases with associated neurodegenerative disease (64 with AD, 4 Frontotemporal Degeneration (FTD) and 3 Dementia with Lewy Bodies (DLB) patients), and 19 patients with MCI but without an associated neurodegenerative disease (Figure 1). Overall, there were 52 women and 38 men, aged between 46 and 78 years, 14 having university studies, 21 secondary studies, 54 primary studies and 1 with no studies. The demographic and neuropsychological information is collected in **Table 1**.

FDG-PET

The FDG PET images were retrieved from the PACS. All patients were acquired in a Philips Gemini TF 64 PET/CT scanner. Acquisition protocol followed the European Association of Nuclear Medicine procedure guidelines [8]. All patients were positioned comfortably in a quiet, dimly lit room about 20-30 minutes before [^{18}F]FDG administration and during the uptake phase of [^{18}F]FDG. Images were acquired between 20- and 35-minutes post-injection, for a total scan length between 10 and 20 minutes. All images were validated by a nuclear medicine physician for quality.

Images were downloaded in the DICOM format with 128 x 128 x 90 matrix resolution, and 2 x 2 x 2 mm voxel size. Images were co-registered to a Dementia-Specific FDG PET template from the Montreal Neurological Institute [9] and smoothing with a 4mm FWHM filter. The software Statistical Parametric Mapping 12 (SPM12) in MATLAB R2014b was used for this task [10]. All scans were resized to a 110 x 92 x 92 resolution with 2 x 2 x 2 mm voxel size.

Neural network

A 3D CNN was implemented to separate MCI cases with from without associated neurodegenerative disease. The CNN consisting of 3 convolutional and pooling blocks attached to a fully connected layer with rectified linear units as activation function. Adam optimizer [11] with $1.5 \cdot 10^{-3}$ learning rate and binary cross-entropy loss function where chosen, with a batch size set to 8. Dropout layers were attached after the convolution blocks and the fully connected layer [12]. This reduced number of convolutional blocks and filters, together with the addition of dropout layers, was ensured to control overfitting in the training process. A depiction of the network architecture and parameters is shown in **Table 2**, and the architecture of the neural network is illustrated in **Fig. 2**.

Python's framework for DL Keras [13] was used for building the model. The 3D CNN was launched on a Linux operating system (Ubuntu 18.04) with 2 Nvidia GeForce RTX 2080 GPUs and Intel i7-7800x CPU.

The 3D CNN was trained and validated on a dataset downloaded from the Alzheimer's Disease Neuroimaging Initiative (ADNI) open database [14]. This dataset consists of 822 subjects (472 AD, 350 MCI). FDG PET brain images acquired at the baseline visit of each subject were used as input for the network. The ADNI diagnosis (baseline diagnosis) considered neuropsychological testing, imaging, and fluid sample. 90% of ADNI dataset was used for training and internal validation of the algorithm and the remaining 10% is used for testing the effectiveness of the algorithm. ADNI image pre-processing included dynamic frame co-registration and averaging. Further pre-processing steps were carried out including co-registration to a Dementia-Specific FDG PET template [9] and smoothing with a 4mm FWHM filter. The software SPM12 [10], was used for this task. Once the network was trained and tested, the model was applied to our hospital independent set of images to perform an external validation of the CNN performance.

Results

To evaluate the performance of the developed ADNI model, several classification metrics such as Sensitivity, Specificity, Accuracy, and area under the receiver operating characteristic curve (AUC) were considered. The CNN tested on the 10% ADNI dataset (41 MCI and 41 AD subjects) yielded a sensitivity and specificity of 88% (36 out of 41 AD subjects correctly identified) and 71% (29 out of 41 MCI) respectively, with an accuracy of 79% in classifying AD from MCI patients. An AUC of 0.897 was obtained.

This same model was applied to our institution dataset of 90 patients, 19 MCI without an associated neurodegenerative disease and 71 MCI with an associated neurodegenerative disease patient. The sensitivity and specificity of 75% (53 out of 71 MCI with an associated neurodegenerative disease identified) and 84% (16 out of 19 MCI without an associated neurodegenerative disease identified) values were respectively obtained, with an accuracy of 77% and an AUC of 0.860. The following table and figures provide an overview of this information: Table 3 includes the comparison between ADNI and La Fe classification metrics. For the hospital dataset, the confusion matrix is depicted in Fig. 3 and the receiver operating characteristic curve in Fig. 4.

In order to check for the interpretability of our model, we investigated several visualization tools for generating the network attention maps, highlighting relevant features of the input image which the learned model uses in the final prediction. Saliency maps [15], SmoothGrad [16] and Grad-CAM [17] visualization algorithms were computed and averaged to both ADNI validation and our institution dataset. Results from the validation set are depicted in Fig. 5, Fig. 6, and Fig. 7.

Discussion

In this study, we proposed a Deep Learning algorithm that classifies patients who present mild cognitive impairment into two diagnostic groups by using FDG PET imaging modality. The classification allows us to distinguish between those who present a neurodegenerative disease and those who doesn't.

The DL neural network was trained and validated first on the ADNI dataset. Then, we validated the network on a dataset composed of 90 patients whose images were acquired in the hospital La Fe in order to compare results. 71 out of 90 patients were diagnosed as MCI without an associated neurodegenerative disease. The obtained results show that the model trained on the ADNI dataset can be applied on the hospital images preserving 77% of accuracy in the diagnosis classification. The sensitivity and specificity of the validation were 75% and 84% respectively. This validation has demonstrated the reproducibility of the ADNI based trained network in an external dataset. The network exhibits acceptable performance when evaluated on both datasets, enabling to learn the non-complex features in the brain image when performing classification.

A possible source of error can be, on the one hand, the selection bias, because all ADNI subjects had not a CDR = 0.5, which means that the cognitive impairment might not be in the beginnings, while all patients from the validation dataset had CDR = 0.5. Furthermore, the ADNI dataset also included only AD neurodegenerative disease, while in our validation dataset DLB and DFT patients were included in the MCI with an associated neurodegenerative disease. On the other hand, unlike magnetic resonance images, there is not a standardized method to process or prepare the FDG PET images, so our considerations may have not be the optimal. Finally, an appropriate partial volume correction (PVC) on FDG PET images might have enhanced the brain network structure analysis and improved classification performance [18]. No PVC was considered due to the lack of MR in part of the 90 patients.

Apart from that, as shown in the attention maps of our network, Saliency maps and SmoothGrad appear to be highlighting posterior cingulate and superior parietal areas. This brain sections agree with the FDG endophenotype of AD [19]. However, Grad-CAM method show right prefrontal brain area being focused, which is not typical of AD. This leads us to speculate that SmoothGrad is the most reliable method because it is consistent with the FDG endophenotype of AD. Still, this visualization methods for debugging the neural network need to be treated carefully, since they are very fragile to small perturbations in the input image [20]. Further research in this topic is needed in order to achieve more robustness in the interpretation of neural networks.

There are several AI studies which their aim is to classify neurodegenerative diseases [21, 22, 3, 23]. Although some of them achieve better results for the accuracy, specificity, or sensitivity metrics, none of them has done an external validation to verify its reproducibility. Moreover, there are few studies which only use PET imaging for the classification, most of them also used MR and/or other biomarkers.

This tool can help for the early non-invasive prediction of neurodegenerative diseases in patients with MCI. Furthermore, we are developing a new neural network architectures to improve results by applying transfer learning from the DL neural network trained on the ADNI images to the hospital dataset as well as to include other variables with the purpose of getting better classification results.

Declarations

Funding

This work was financially supported by INBIO 2019 (DEEPBRAIN), INNVA1/2020/83 (Generalitat Valenciana, DEEPPET) and PID2019-107790RB-C22, (AGENCIA ESTATAL DE INVESTIGACION).

Competing interests

The authors declare that they have no competing interests.

Ethics approval

Approval was granted by the Ethics Committee for the study and the patients information consent was waived due to the retrospective non-interventional observational research on data design of the study.

Acknowledgements

Data used in the training of the deep learning network were obtained from the ADNI database (<http://www.loni.ucla.edu/ADNI>). In consideration of that, the investigators within the ADNI provided data but did not participate in analysis or writing of this report.

References

1. A. V. G. a. M. E. O. a. S. L. V. a. M. L. L. a. C. L. d. S. d. M. a. C. P. Jurczynska, "Report by the Spanish Foundation of the Brain on the social impact of Alzheimer disease and other types of dementia," *Neurologia*, vol. 36, pp. 39-49, 1 2021.
2. M. P. a. A. C.-H. a. M. K. a. M. G. a. & M. Karagiannidou, "World Alzheimer Report 2016 Improving healthcare for people living with dementia. Coverage, Quality and costs now and in the future. Alzheimer's Disease International (ADI).," 2016 .
3. Y. G. a. R. K. L. a. G. R. Kwon, "Prediction and Classification of Alzheimer's Disease Based on Combined Features From Apolipoprotein-E Genotype, Cerebrospinal Fluid, MR, and FDG-PET Imaging Biomarkers," *Frontiers in Computational Neuroscience*, vol. 13, 10 2019.
4. J. R.-B. L. R.-B. J. G.-B. M. C.-R. A. S.-L. L. M. G.-S. I. R.-Z. M. R. J. V.-C. C. M.-T. C. S.-M. C. A. g. I. R.-P. R. R.-R. C Gamez-Cenzano, "Impact of PET brain imaging using F18-FDG and F18-FLORBETAPIR in patients with cognitive impairment," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 42, no. 1, 2015.
5. D. H. S. S. a. L. M. a. L. E. a. W. C. a. G. W. Small, "PET Scans Obtained for Evaluation of Cognitive Dysfunction.," *Seminars in nuclear medicine*, vol. 38, 2008.
6. C. M. a. E. M. a. R. M. Subramaniam, "Brain PET in the diagnosis of Alzheimer's disease," *Clinical Nuclear Medicine*, vol. 39, 2014.
7. R.-B. L. G.-B. J. R.-R. R. C.-F. J. T.-S. J. S.-M. C. V.-C. J. G.-S. L. L.-T. E. P.-E. A. M.-T. C. Gámez-Cenzano C, "Role of 18F-FDG-PET and amyloid-PET imaging on patient management in mild cognitive impairment or dementia," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 43, no. 1, 2016.

8. S. A. T. V. B. J. B. F. N. K. N. J. D. Ö. L. K. K. T. P. B. K. V. Laere Andrea Varrone, "EANM procedure guidelines for PET brain imaging using [18F]FDG, version 2," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, no. 12, 2009.
9. P. A. a. C. C. a. G. F. a. P. A. a. C. A. a. C. I. a. G. M. C. a. F. G. a. F. K. a. A. J. a. P. D. a. { Della Rosa, "A standardized [18F]-FDG-PET template for spatial normalization in statistical parametric mapping of dementia," *Neuroinformatics*, vol. 12, no. 4, 10 2014.
10. K. F. a. J. A. a. S. K. a. T. N. a. W. Penny, *Statistical Parametric Mapping*, London: Academic Press, 2007.
11. D. P. a. B. J. Kingma, "Adam: A Method for Stochastic Optimization," 2014.
12. N. a. H. G. E. a. K. A. a. S. I. a. S. R. Srivastava, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 1, 2014.
13. F. a. o. Chollet, "Keras," 2005. [Online]. Available: <https://github.com/fchollet/keras>.
14. "Alzheimer's Disease Neuroimaging Initiative," 2003. [Online]. Available: <http://adni.loni.usc.edu/>.
15. K. S. a. A. V. a. A. Zisserman, "Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps," *CoRR*, vol. abs/1312.6034, 2014.
16. N. T. B. K. F. V. M. W. Daniel Smilkov, "Smoothgrad: removing noise by adding noise," *Workshop on Visualization for Deep Learning, ICML*, 2017.
17. R. R. a. C. M. a. D. A. a. V. R. a. P. D. a. B. D. Selvaraju, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," pp. 618-626, 2017.
18. C. H. N. G. J. D. L. M. V. K. A. J. J. S. G. E. F. & Q. L. Jiarui Yang, "Partial volume correction for PET quantification and its impact on brain network in Alzheimer's disease," *Scientific Reports*, vol. 7, 2017.
19. K. A. F. N. C. S. R. A. & K. W. E. Johnson, "Brain imaging in Alzheimer disease," *Cold Spring Harbor*, vol. 2, no. 4, 2012.
20. A. A. A. & Z. J. Ghorbani, "Interpretation of Neural Networks Is Fragile," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 3681-3688, 2019.
21. S. J. K. M. T. H. H. R. J. N. L. D. C. T. A. M. M. A. C. B. S. F. R. H. S. Z. K. N. L. S. Y. H. R. H. P. M. H. D. F. B. Ding Y, "A deep learning model to predict a diagnosis of Alzheimer disease by using 18 F-FDG PET of the brain," *Radiology*, vol. 290, no. 3, pp. 456-464, 2019.
22. C. E. A. Y. P. W. T. Z. F. H. H. X. X. a. L. B. Feng, "Deep Learning Framework for Alzheimer's Disease Diagnosis via 3D-CNN and FSBi-LSTM," *IEEE Access*, vol. 7, 2019.
23. R. K. L. G.-R. K. a. t. A. D. N. I. Yubraj Gupta, "Prediction and Classification of Alzheimer's Disease Based on Combined Features From Apolipoprotein-E Genotype, Cerebrospinal Fluid, MR, and FDG-PET Imaging Biomarkers," *Frontiers in Computational Neuroscience*, vol. 13, 16 10 2019.

Tables

Table 1 Demographical neuropsychological and cerebrospinal fluid characteristics of the studied sample

Groups	MCI with neurodegenerative disease	MCI without neurodegenerative disease
N of subjects	71	19
Female/Male	41/30	11/8
Age*	69 ± 5	66 ± 7
US/SS/PS/NS	12/16/42/1	2/5/12/0
MMSE*	24 ± 5	25 ± 2
FAQ*	6.94 ± 6.38	5.68 ± 6.16
RBANs*	338 ± 61	357 ± 65
Aβ42 CSF*	559 ± 267	1003 ± 384
p-tau CSF*	110 ± 65	43 ± 17
t-tau CSF*	735 ± 506	286 ± 113

Note: * Values are presented as mean ± standard deviation (SD).

MCI = Mild Cognitive Impairment, US = university studies, SS = secondary studies, PS = primary studies, NS = no studies, MMSE = Mini-Mental State Examination, FAQ = Functional Activities Questionnaire, RBANs = Repeatable Battery for the Assessment of Neuropsychological Status, CSF = cerebrospinal fluid.

Table 2 3D-CNN architecture and hyperparameters

Layer ID	Layer	Kernel number	Kernel size
0	input		
1	Conv3D-1	2	(3,3,3)
2	Max Pooling3D-1		(2,2,2)
3	Dropout (0.4)		
4	Conv3D-2	8	(3,3,3)
5	Max Pooling3D-2		(2,2,2)
6	Dropout (0.4)		
7	Conv3D-3	16	(3,3,3)
8	Max Pooling3D-3		(2,2,2)
9	Dropout (0.4)		
10	Fully Connected	128	
11	Dropout (0.5)		
12	Sigmoid	2	

Table 3 3D-CNN classification metrics

Test Set	Accuracy	Sensitivity	Specificity	AUC
ADNI 10%	0.79	0.88	0.71	0.897
La Fe	0.77	0.75	0.84	0.860

Figures

Figure 1

Eligibility criteria for the participants

Figure 2

Neural network architecture

Figure 3

Confusion matrix for La Fe dataset

Figure 4

AVERAGE SALIENCY MAP

Figure 5

Saliency maps visualization algorithm computed and averaged on La Fe dataset

AVERAGE SMOOTHGRAD

Figure 6

SmoothGrad visualization algorithm computed and averaged on La Fe dataset

AVERAGE GRADCAM

Figure 7

Grad-CAM visualization algorithm computed and averaged on La Fe dataset