A Deep Learning Framework For Clinical Trial Enrichment in Alzheimer's Disease

Luis R. Peraza[1], Richard Manber[1], Colm McGinnity[1], Richard Joules[1], Robin Wolz[1,2]
1 IXICO, London UK | 2 Imperial College London UK.

Introduction

Selection of participants at risk of cognitive decline in clinical trials, known as trial enrichment, increases the probability of trial success. It is estimated that by 2050, 153 million people worldwide will be living with a type of dementia. Hence, innovative trial recruitment strategies are necessary to accelerate treatment development.

Here we present a deep-learning framework for trial enrichment for Alzheimer's Disease (AD) that uses a combination of neuroimaging and clinical/demographic variables as inputs. The framework is designed with built-in redundancy allowing the system to work effectively with missing inputs.

Objectives

- Design an DL-based framework for participant recruitment in clinical trials.
- Select and score participants that are at early disease stage but are predicted to experience an accelerated cognitive decline.
- Save participant screening costs by using structural MRI and subsequent confirmation with amyloid PET imaging.

Methods

We employed T1-MR and amyloid-PET images from ADNI, OASIS-3 and AMYPAD repositories (Table I). Images were pre-processed by validated pipelines for volumetric and amyloid PET SUVR estimation [1-4]. SUVR was estimated with cerebellar grey matter as reference region and the analysed PET databases comprised multiple radiotracers: 18F-florbetapir, 18F-florbetaben and 11C-PIB. Training-set participants were labelled as stable or decline through hierarchical clustering, where a stable participant experienced no significant decline within a 36-month period.

Our framework comprises two phases: Firstly, Siamese convolutional neural network (CNN) encoders were trained with decline/stable targets and neuroimages as inputs. Secondly, a battery of random forest classifiers were trained with CNN outputs, neuroimaging derivatives, and demographics/clinical data inputs and decline/stable targets.

Participant decline/stable prediction is achieved in a sequential approach, initially using structural MRI data and clinical variables. If a participant scores within an uncertainty range in the first prediction stage, a second confirmatory prediction is obtained by adding amyloid-PET data to a second classifier.

Virtual clinical trials (VCT) were implemented by simulating a 50% cognitive decline rate reduction as drug effect. For the VCTs, whole/enriched cohorts were bootstrapped 500 times. Methodology steps are shown in Figure 2.

Figure 1. Clinical Trial Enrichment Framework.

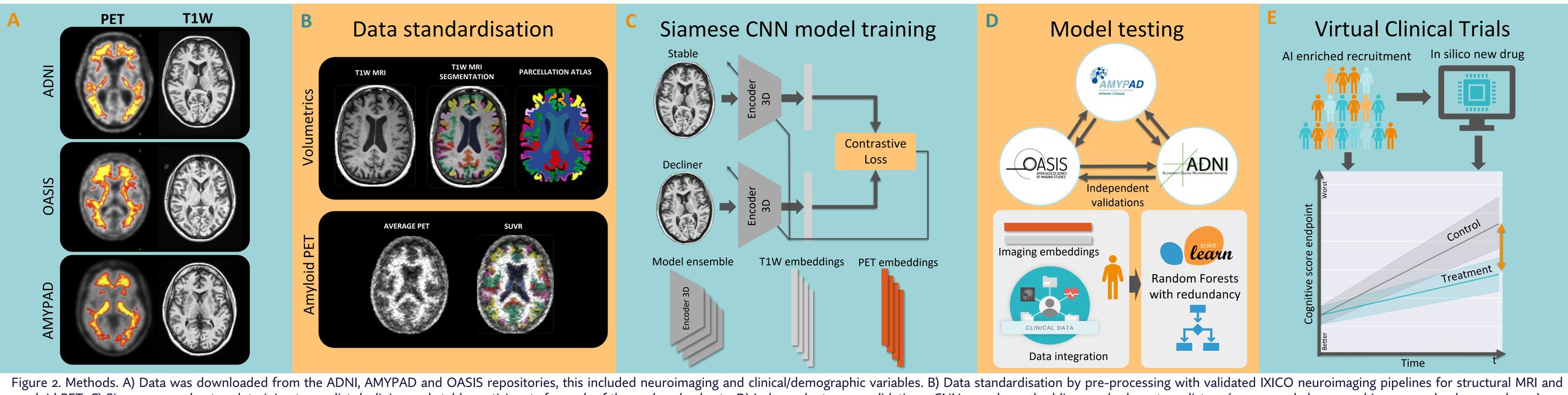

Figure 2. Methods. A) Data was downloaded from the ADNI, AMYPAD and OASIS repositories, this included neuroimaging and clinical/demographic variables. B) Data standardisation by pre-processing with validated IXICO neuroimaging pipelines for structural MRI and amyloid PET. C) Siamese neural network training to predict declining and stable participants for each of the analysed cohorts. D) Independent cross-validations, CNN encoder embeddings and relevant predictors (age, sex, cdrsb, apoe4, hippocampal volume and suvr) are integrated with redundant random forests. E) Virtual clinical trials (VCT) are implemented to test trial enrichment effect using bootstrapping.

Results

Demographics

The AMYPAD database has the largest PET-to-MRI imaging availability, with participants typically in an earlier disease stage compared to the other cohorts. OASIS and ADNI comprised a more heterogenous participant recruitment (Table I).

Table I. Demographics and clinical variables

N >24 months follow-up	Diagnosis (baseline)	Age	Sex (F/M)	CDRSB	MMSE	Education	SUVR	APOE4 (0/1/2)
AMYPAD N_{MRI}=447 N_{PET}=390	CN=443 MCI=4	64.7(6.7)	252/195	0.03(.12)	29.2(1.0)	14.6(3.9)	1.31(0.32)	230/182/35
OASIS N_{MRI}=529 N_{PET}=97	CN=414 MCI=26 AD=89	67.6(9.6)	300/229	0.61(1.37)	28.3(2.4)	15.8(9.6)	1.10(0.4)	311/185/33
ADNI N_{MRI}=924 N_{PET}=270	CN=274 MCI=593 AD=57	72.4(7.2)	451/473	1.04(1.3)	28.1(2.12)	16.3(2.6)	1.48(0.33)	544/300/47

Performance

In independent cross-validation tests, our framework reached a mean balanced accuracy (BAcc) of 83% [5], with a mean sensitivity (recall) of 73% across all studied datasets (Table II).

We also estimated BAcc for optimal hippocampal volume thresholds, which resulted in a mean BAcc of 62%, 57%, 60% for ADNI-, OASIS-, AMYPAD trained thresholds, respectively. Mean BAcc for a majority rule classifier was of 42% (±1).

Table II. Independent validation performance.

Trained with	Tested on	Mean performance	
ADNI	OASIS, ADNI	Balanced Accuracy	94%
		Precision	35%
		Sensitivity	97%
OASIS	ADNI, AMYPAD	Balanced Accuracy	86%
		Precision	67%
		Sensitivity	77%
AMYPAD	ADNI, OASIS	Balanced Accuracy	70%
		Precision	80%
		Sensitivity	43%

Using sequential prediction, 72% of the model's decisions were made with structural MRI data and clinical variables (Figure 3).

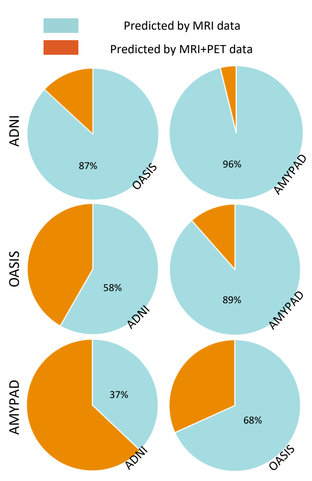

Figure 3. Percentage of model decisions by MRI data and subsequent confirmation by adding PET imaging. Results shown for independent validation.

Highlights

- We independently validated our framework using ADNI, OASIS and AMYPAD datasets.
- Across the analysed datasets, the percentage of participants who cognitively declined was of 15%, 30% in MCI participants, highlighting the importance of this work.
- The framework reached an 83% mean balanced accuracy across all independent datasets.
- Our framework proved it can potentially save 72% of PET imaging costs at recruitment time by using MRI-based prediction followed by PET imaging confirmation.
- The VCT experiments showed that our framework selects participants who experience accelerated decline and who are at early disease stage.

Explainable AI

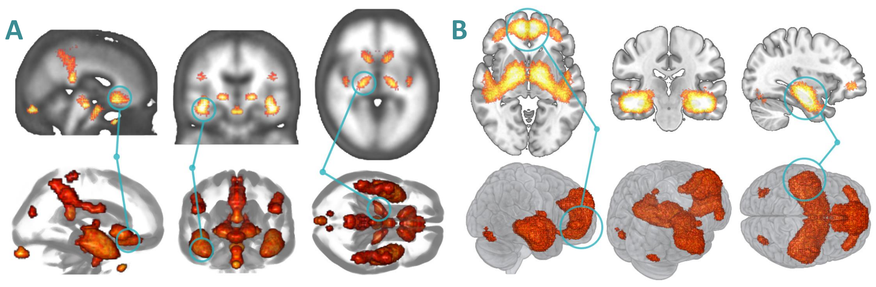

Figure 4. Activation maps from the trained Siamese 3D CNNs. A: Activation maps for the PET images. B: Activation maps for the structural MRIs.

The saliency maps from the trained 3D CNN encoders showed high gradients within brain regions associated with AD. For the PET images these regions were the Temporal, Parietal, Precuneal and Lower-Frontal cortices. We also found important contributions of subcortical regions in PET. For the MRI images, CNN activation regions comprised the Temporal, Frontal cortices as well as Thalamic subcortices (Figure 4).

Virtual Clinical Trials

VCT results are shown for the ADNI cohort as test set using the mean prediction scores from models trained with OASIS and AMYPAD (Figure 5). Simulated recruitment was for participants with a CDRSB between 0 and 3.5 (i.e. CN and MCI). The whole recruited ADNI cohort showed a drug effect size of 0.25 (Cohen's D), while both enriched cohorts (by CDRSB and our framework) at 30% inclusion rate, showed an increased drug effect size of 0.36. However, the mean CDSRB at baseline is much higher for the CDRSB-enriched group, indicating that these participants are in a more advanced disease stage.

Figure 5. Virtual clinical trial for the ADNI cohort. Cohorts enriched by CDRSB and our AI framework. Error shades show the 95% confidence interval for the mean.

Conclusions

We designed an DL-based framework for clinical trial enrichment in AD. Our independent validations showed that our framework will perform well on new clinical trials aiming to evaluate disease modifying therapies [6].

References

1. Lopes-Alves I et al. "Quantitative amyloid PET in Alzheimer's disease: the AMYPAD prognostic and natural history study", 2020.
2. Palombit A., et al. "Regional differences of amyloid PET SUVR induced by spatial smoothing and the role of reference region", 2020.
3. Weatheritt J, Rueckert D, Wolz R. "Transfer learning for brain segmentation: pre-task selection and data limitations", 2020.
4. Wolz R et al. "LEAP: Learning embeddings for atlas propagation", 2010.
5. Brodersen, K.H., et al. "The balanced accuracy and its posterior distribution", 2010.
6. Peraza LR, et al. "Clinical Trial Enrichment in Alzheimer's Disease", on preparation.