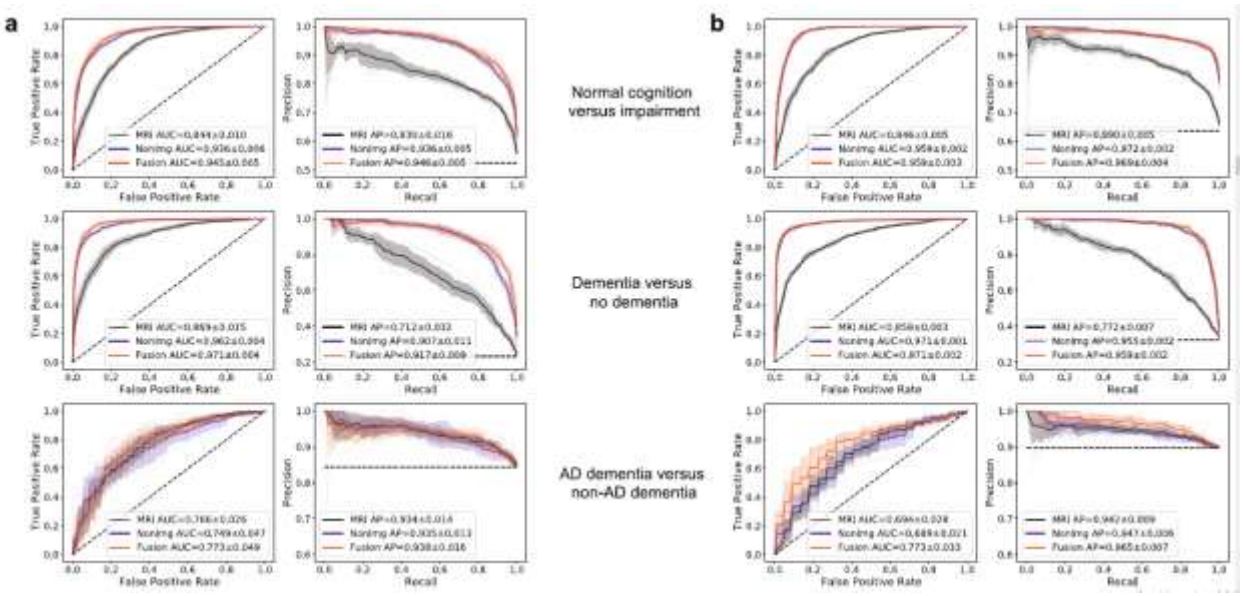
Reference	Objective	DataSet
Multimodal deep learning for Alzheimer's disease dementia assessment. Nat Commun 13, 3404 (2022). https://doi.org/10.1038/s41467-022-31037-5	differential diagnosis of Alzheimer's disease, differentiating between normal cognitive aging, NC, MCI, AD, and nADD, and other dementia etiologies.	(ADNI) dataset (n = 1821), NACC) dataset (n = 4822), (NIFD) dataset (n = 253), (PPMI) dataset (n = 198), (AIBL) dataset (n = 661), (OASIS) dataset (n = 666), (FHS) dataset (n = 313), (LBDSU) (n = 182)

methods	Performance
MRI-only model: (CNN) that uses MRI scans and no other information to complete the COG and the ADD tasks.	The performance metrics, including accuracy, F-1 score, sensitivity, specificity and MCC, were evaluated on various tasks where MRI scans were the sole inputs to the deep learning model. Results for the (a)NACC test set, (b) OASIS (c) combined external datasets. For each table, the columns (COGNC,COGMCI, COGDE) correspond to metrics for a one-versus-rest classification task in which the goal was to individually delineate these three cognitive categories from all others within the overarching COG task. The "COG" column corresponds to the complete COG task of separating each NC/MCI/DE category (i.e., a 3-way classification). The "ADD" columns corresponds to the task of classifying AD and nADD diagnosis given that a DE diagnosis has already been obtained from the COG task. Lastly, the "4-way" column corresponds to the complete classification workflow in which NC, MCI, AD, and nADD cases are delineated in a final four-way classification.

Non-imaging model: A traditional machine learning classifier that uses demographics, past medical history, neuropsychological testing, and functional assessments to complete the COG and the ADD tasks.	After compiling and comparing the performance of numerous machine learning classifiers on non-imaging data, we determined that the CatBoost algorithm yielded the strongest overall performance across all classification tasks. We highlight the performance metrics of this model when trained on non-imaging data only and tested on (a) NACC test set and (b) OASIS. Here, F1 score, sensitivity, specificity, and MCC are reported for various prediction tasks. For each table, the columns (COGNC, COGMCI, COGDE) correspond to metrics for a one-versus-rest classification task in which the goal was to individually delineate these three cognitive categories from all others within the overarching COG task. The "COG" column corresponds to the complete COG task of separating each NC/MCI/DE category (i.e., a 3- way classification). The "ADD" column corresponds to the task of detecting AD diagnosis given that a DE diagnosis has already been obtained from the COG task. Lastly, the "4- way" column corresponds to the complete classification workflow in which NC, MCI, AD, and nADD cases are delineated in a final four-way classification.
• Fusion model: A hybrid model composed of a CNN linked to a CatBoost model. The CNN portion computes the DEMO and the ALZ scores from MRI which are concatenated with non-imaging clinical variables. The CatBoost model then successively completes the COG and the ADD tasks.	*performance of multiple types of fusion models was assessed on the **NACC test** set across all prediction tasks. In addition to the combination of CNNderived features and a CatBoost model, we also combined our CNN with (a) decision tree (b) KNN (c)multilayer perceptron (d)random forest (e) support vector machine (f)XGBoost model. The F-1 score,sensitivity, specificity, andMCC values are reported forvarious predictions. Thecolumn names have identical meaning to those described in prior tables. As previouslymentioned, we judged that thesemodels offered lesser performance than a CatBoost fusion model, and thus these additional classifiers were not used in the final fusion. *The performance of multiple types of fusion models was

	assessed on the **OASIS dataset** across all prediction tasks. In addition to the combination of CNNderived features and a CatBoost model, we also combined our CNN with (a) decision tree (b) K-NN (c) multilayer perceptron (d) random forest (e) support vector machine (f) XGBoost model. The F-1 score, sensitivity, specificity, and MCC values are reported for various predictions. The column names have identical meaning to those described in prior tables.

NACC **OASIS**



WE WILL SEE THAT MODEL MRI (AUC & AP) < NONIMAGING(AUC&AP) < FUSSION MODE(AUC &AP)

Fussion model is the nearest to 1 so is the best.

(a)

MRI

	COG	COG$_{NC}$	COG$_{MCI}$	COG$_{DE}$	ADD	4-way
Accuracy	0.782±0.011 [0.769-0.796]	0.856±0.008 [0.846-0.866]	0.790±0.010 [0.777-0.803]	0.919±0.006 [0.912-0.926]	0.806±0.033 [0.765-0.847]	0.748±0.012 [0.734-0.763]
F-1	0.752±0.013 [0.736-0.768]	0.862±0.009 [0.852-0.873]	0.566±0.022 [0.539-0.593]	0.827±0.012 [0.812-0.842]	0.884±0.019 [0.860-0.908]	0.611±0.027 [0.577-0.645]
Sensitivity	0.752±0.013 [0.736-0.768]	0.863±0.016 [0.843-0.883]	0.563±0.032 [0.524-0.603]	0.831±0.020 [0.806-0.856]	0.878±0.019 [0.854-0.901]	0.612±0.029 [0.575-0.648]
Specificity	0.886±0.006 [0.879-0.893]	0.848±0.014 [0.831-0.865]	0.863±0.013 [0.846-0.880]	0.946±0.007 [0.937-0.955]	0.417±0.140 [0.244-0.590]	0.906±0.004 [0.900-0.911]
MCC	0.638±0.018 [0.616-0.660]	0.711±0.016 [0.691-0.732]	0.428±0.027 [0.395-0.462]	0.775±0.016 [0.755-0.794]	0.283±0.139 [0.110-0.456]	0.517±0.031 [0.478-0.555]

(a)

Non-imaging

	COG	COG$_{NC}$	COG$_{MCI}$	COG$_{DE}$	ADD	4-way
Accuracy	0.659±0.016 [0.639-0.679]	0.760±0.009 [0.749-0.771]	0.718±0.025 [0.687-0.749]	0.840±0.011 [0.826-0.854]	0.849±0.012 [0.834-0.864]	0.645±0.015 [0.626-0.664]
F-1	0.569±0.018 [0.547-0.591]	0.797±0.008 [0.787-0.807]	0.308±0.070 [0.221-0.395]	0.603±0.048 [0.543-0.663]	0.914±0.007 [0.905-0.923]	0.466±0.025 [0.435-0.497]
Sensitivity	0.566±0.014 [0.549-0.583]	0.897±0.042 [0.845-0.949]	0.273±0.109 [0.138-0.408]	0.528±0.076 [0.434-0.622]	0.954±0.011 [0.940-0.968]	0.455±0.022 [0.428-0.482]
Specificity	0.802±0.006 [0.795-0.809]	0.610±0.055 [0.542-0.678]	0.861±0.066 [0.779-0.943]	0.935±0.018 [0.913-0.957]	0.286±0.031 [0.248-0.324]	0.848±0.005 [0.842-0.854]
MCC	0.404±0.022 [0.377-0.431]	0.536±0.016 [0.516-0.556]	0.156±0.033 [0.115-0.197]	0.519±0.043 [0.466-0.572]	0.315±0.047 [0.257-0.373]	0.352±0.036 [0.307-0.397]

(a)

Fusion mode

NACC

	COG	COG$_{NC}$	COG$_{MCI}$	COG$_{DE}$	ADD	4-wa
Accuracy	0.698±0.030 [0.661-0.735]	0.797±0.023 [0.768-0.826]	0.722±0.022 [0.695-0.749]	0.878±0.017 [0.857-0.899]	0.731±0.072 [0.642-0.820]	0.654±0. [0.614-0.
F-1	0.665±0.027 [0.631-0.699]	0.801±0.025 [0.770-0.832]	0.461±0.016 [0.441-0.481]	0.734±0.045 [0.678-0.790]	0.832±0.051 [0.769-0.895]	0.513±0. [0.478-0.
Sensitivity	0.666±0.028 [0.631-0.701]	0.783±0.037 [0.737-0.829]	0.487±0.023 [0.458-0.516]	0.729±0.066 [0.647-0.811]	0.803±0.092 [0.689-0.917]	0.516±0. [0.478-0.
Specificity	0.844±0.014 [0.827-0.861]	0.813±0.014 [0.796-0.830]	0.798±0.034 [0.756-0.840]	0.923±0.008 [0.913-0.933]	0.343±0.105 [0.213-0.473]	0.872±0. [0.858-0.
MCC	0.509±0.040 [0.459-0.559]	0.596±0.044 [0.541-0.651]	0.276±0.029 [0.240-0.312]	0.655±0.054 [0.588-0.722]	0.149±0.136 [-0.020-0.318]	0.388±0. [0.338-0.

(a)

Fusion mode

OASIS

	COG	COG$_{NC}$	COG$_{MCI}$	COG$_{DE}$	ADD	4-way
Accuracy	0.670±0.034 [0.628-0.712]	0.775±0.035 [0.731-0.819]	0.722±0.032 [0.681-0.762]	0.844±0.013 [0.828-0.859]	0.752±0.071 [0.663-0.840]	0.622±0.034 [0.580-0.664]
F-1	0.542±0.014 [0.525-0.559]	0.801±0.038 [0.753-0.848]	0.111±0.035 [0.068-0.155]	0.714±0.029 [0.678-0.751]	0.850±0.052 [0.785-0.914]	0.400±0.018 [0.378-0.423]
Sensitivity	0.587±0.040 [0.538-0.637]	0.716±0.060 [0.642-0.790]	0.437±0.151 [0.250-0.624]	0.608±0.047 [0.550-0.667]	0.799±0.089 [0.688-0.910]	0.437±0.040 [0.387-0.486]
Specificity	0.856±0.011 [0.842-0.869]	0.878±0.018 [0.856-0.900]	0.734±0.036 [0.689-0.778]	0.956±0.020 [0.931-0.980]	0.336±0.121 [0.187-0.486]	0.880±0.009 [0.868-0.891]
MCC	0.427±0.023 [0.398-0.455]	0.573±0.051 [0.510-0.637]	0.075±0.061 [-0.002-0.151]	0.632±0.031 [0.593-0.671]	0.103±0.058 [0.031-0.174]	0.307±0.027 [0.273-0.340]