

RESEARCH ARTICLE

Transfer learning-trained convolutional neural networks identify novel MRI biomarkers of Alzheimer's disease progression

Yi Li¹ | Annat Haber¹ | Christoph Preuss² | Cai John¹ | Asli Uyar¹ |
Hongtian Stanley Yang² | Benjamin A. Logsdon³ | Vivek Philip² |
R. Krishna Murthy Karuturi¹ | Gregory W. Carter^{1,2} | The Alzheimer's Disease
Neuroimaging Initiative*

¹ The Jackson Laboratory, Farmington, Connecticut, USA

² The Jackson Laboratory, Bar Harbor, Maine, USA

³ Sage Bionetworks, Seattle, Washington, USA

Correspondence

Gregory W. Carter, The Jackson Laboratory, 600 Main St., Bar Harbor, ME 04609, USA.

E-mail: Gregory.Carter@jax.org

Present address

Cai John, The University of Tennessee, Knoxville, Tennessee, USA

*Data used in 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 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

Funding information

National Institutes of Health, Grant/Award Number: U01 AG024904; Department of Defense, Grant/Award Number: W81XWH-12-2-0012; National Institute on Aging; National Institute of Biomedical Imaging and Bioengineering; NIA, Grant/Award Numbers: P30AG10161, R01AG15819, 363 R01AG17917, R01AG30146, R01AG36836, U01AG32984, U01AG46152

Abstract

Introduction: Genome-wide association studies (GWAS) for late onset Alzheimer's disease (AD) may miss genetic variants relevant for delineating disease stages when using clinically defined case/control as a phenotype due to its loose definition and heterogeneity.

Methods: We use a transfer learning technique to train three-dimensional convolutional neural network (CNN) models based on structural magnetic resonance imaging (MRI) from the screening stage in the Alzheimer's Disease Neuroimaging Initiative consortium to derive image features that reflect AD progression.

Results: CNN-derived image phenotypes are significantly associated with fasting metabolites related to early lipid metabolic changes as well as insulin resistance and with genetic variants mapped to candidate genes enriched for amyloid beta degradation, tau phosphorylation, calcium ion binding-dependent synaptic loss, APP-regulated inflammation response, and insulin resistance.

Discussion: This is the first attempt to show that non-invasive MRI biomarkers are linked to AD progression characteristics, reinforcing their use in early AD diagnosis and monitoring.

KEYWORDS

Alzheimer's disease, convolutional neural networks, deep learning, disease progression, imaging phenotypes, machine learning, magnetic resonance imaging; transfer learning

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, LLC on behalf of Alzheimer's Association

HIGHLIGHTS

- Novel non-invasive magnetic resonance imaging (MRI) biomarkers of Alzheimer's disease progression.
- Transfer learning.
- Deep convolutional neural networks.

1 | BACKGROUND

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that slowly degrades memory and cognitive functions. It is neuropathologically defined by intracellular neurofibrillary tangles and aggregated amyloid beta (A β) plaques,¹ both of which can currently be estimated accurately only *post mortem*. The phenotype in current case/control genome-wide association studies (GWAS) for late onset AD (LOAD)²⁻⁴ are based largely on clinical assessments, in which mild cognitive impairment (MCI) and AD are determined by designed memory and cognitive tests and clinical observations. These criteria fail to reflect early AD hallmark characteristics such as A β plaques and neurofibrillary tangles and highlight advanced AD, leaving the MCI category widely heterogeneous and poorly understood. Consequently, current GWAS for LOAD usually exclude MCI and therefore may miss critical genetic variants associated with early AD characteristics and progression.

Non-invasive brain imaging modalities such as magnetic resonance imaging (MRI) and positron emission tomography (PET) are promising tools for monitoring AD progression and its diagnosis. Imaging provides precise quantitative phenotypes, and numerous methods have been proposed for analyzing neuropathology with MRI.⁵⁻⁹ However, the high dimensionality of these phenotypes makes it challenging to extract concise and interpretable information. Summary measures for pre-defined regions of interest (ROI) are suboptimal for predicting the onset and progression of AD because they are derived independently of AD status.

In this article, we make use of deep convolutional neural networks (CNN)¹⁰⁻¹³ to simultaneously extract relevant features and classify patients using (structural) MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) consortium. Deep CNN have become the state-of-the-art methods for image classification¹⁴ due to their ability to form translation invariant hierarchical image features. To mitigate the scarcity of images in the ADNI dataset, and accommodate the high number of model parameters that need to be learned, we adopt a transfer learning technique.^{15,16} This technique uses an independent data-trained 3D CNN model¹⁷ that is then fine-tuned using our dataset of 1381 images. This greatly augments our image dataset and ensures the learned CNN model is more robust to overfitting. Because our baseline MRI images were taken 3 years prior to clinical labeling and provide holistic snapshots of brain states, our CNN-derived image features reflect earlier and more specific AD characteristics than the memory and cognitive performance features used for assessing AD and

RESEARCH IN CONTEXT

1. **Systematic review:** Case/control genome-wide association studies (GWAS) for late onset Alzheimer's disease (AD) may miss genetic variants relevant for delineating disease stages because the cases highlight advanced AD and widely heterogeneous mild cognitive impairment patients are usually excluded. More precise phenotypes for AD are in demand.
2. **Interpretation:** Convolutional neural networks (CNN) trained on structural magnetic resonance imaging (MRI) and clinical labels integrated AD classification and image feature extraction in one step; transfer learning-trained CNN were more robust to overfitting, yielding more accurate image features that predict AD progression. CNN-derived image phenotypes were significantly associated with metabolites related to early lipid metabolic changes and insulin resistance, and with genetic variants mapped to candidate genes enriched for amyloid beta degradation, tau phosphorylation, calcium ion binding-dependent synaptic loss, APP-regulated inflammation response, and insulin resistance.
3. **Future direction:** Relating the MRI biomarkers to specific regions in original MRI images that drive the AD classification.

MCI. This is supported by the significant associations we find between the CNN-derived phenotypes and early AD-related metabolites and genes (Figure 1). To our knowledge, this is the first attempt to link non-invasive MRI biomarkers with AD progression characteristics.

2 | METHODS**2.1 | MRI and clinically labeled data from the ADNI consortium**

Data for this study were obtained from the ADNI database (adni.loni.usc.edu). ADNI is a longitudinal study in which initial imaging is followed by annual reimaging. MRI images taken at the initial stage are later categorized into four major classes based on their follow-up status: control, AD, stable MCI subjects who maintain the same disease status throughout the follow-up period, and progressive MCI subjects who convert from MCI to AD sometime during the follow-up period. MCI were counted as stable MCI only if they were followed-up for at least 3 years in this study. The conversion and follow-up timelines for the 526 MCI patients are shown in Table S1B in supporting information.

We downloaded 817 screening images from the ADNI-1 cohort, 104 ADNI-GO new participants, and 624 ADNI-2 new participants. Because AD patients rarely convert back, we included 162 Year 1

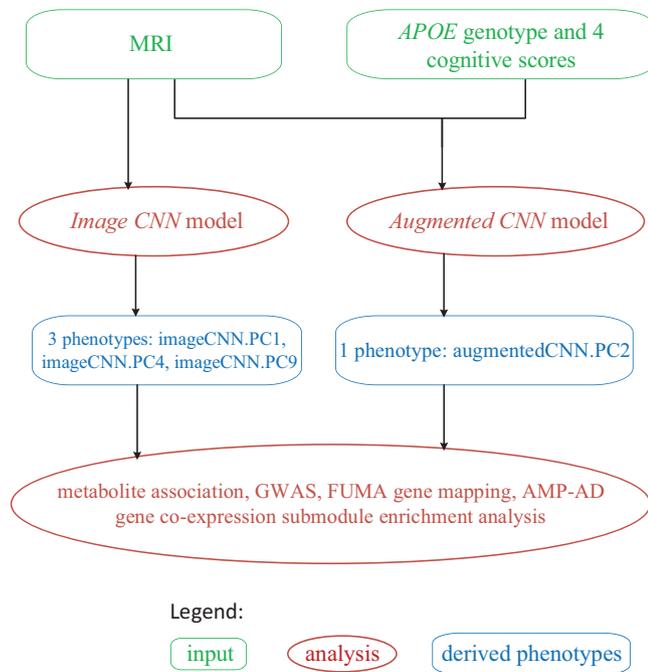


FIGURE 1 Graphic summary of the analytical approach. AMP-AD, Accelerating Medicines Partnership-Alzheimer's Disease; APOE, apolipoprotein E; CNN, convolutional neural network; GWAS, genome-wide association studies; MRI, magnetic resonance imaging

and 95 Year 2 images from ADNI-1 patients who were diagnosed as AD at screening, 155 of which also had MRIs at screening, totaling 1802 images (Table S1A). The adding of Year 1/2 AD images to screening images was expected to help CNN more accurately recognize progression-related image features. However, there were no duplicate subjects in the downstream metabolite and GWAS analysis. Some subjects in ADNI-1 have two MRI scans from the same session; we kept the one in the “Scaled_2” directory as recommended by ADNI MRI core team. We filtered out Year 1 to 2 AD MRIs with the rank of 4 or -1 (based on downloaded MRIMPRANK.csv) and ADNI-GO/2 MRIs with the quality of 4 or none (based on downloaded MAYOADIRL_MRI_IMAGEQC_12_08_15.csv). When there was more than one MRI scan from the same session after filtering, we kept the one with the highest quality or the latest timepoint when qualities were equal.

FreeSurfer^{18,19} software (-autorecon1 option) was applied to correct motions, normalize image intensities, and strip bone tissue, followed by manual checking of sagittal slice 101 of each MRI image, ensuring that the mean intensity of the white matter was around 110 and skull was stripped correctly (FreeSurfer suggestions). Inappropriate skull stripping was rescued by running `mri_watershed` with different watershed thresholds. MRIs with incorrect mean intensity of the white matter or inappropriate skull stripping after rescuing were excluded. To investigate whether co-registration is necessary when applying CNN to brain image analysis, we used the Talairach transformation calculated in FreeSurfer (-autorecon1) to obtain MNI305 atlas-registered MRIs.

Age, four cognitive scores (Mini-Mental State Examination [MMSE], Clinical Dementia Rating-Sum of Boxes [CDR-SOB], Functional Activities Questionnaire [FAQ], Alzheimer's Disease Assessment Scale [ADAS]), and various image summary measures at screening as well as sex, education level, apolipoprotein E (APOE) genotype, and ethnicity were downloaded. The subjects used for training CNN had MRIs and all selected covariates, with only one subject missing cognitive score examination date at screening. Most of the MRIs had examination date within 2 months from that of cognitive scores (Table S1C). The distributions of age, sex, MMSE, and APOE genotype are shown in Table 1.

2.2 | CNN and feature formation

CNN is a type of supervised multiple-layer neural network that adopts learnable convolutional kernels to detect hierarchical image features. Because the same kernel slides over the whole image, the detected image features are translation invariant.²⁰ To reinforce this, input images are often augmented during CNN training via transformations such as multiple scaling and cropping. A loss value at the last layer of a CNN is computed in the forward pass and iteratively minimized by back-propagating the loss to all hidden layers to update their weights based on the stochastic gradient descent rule.²¹

The pre-trained 3D CNN model used for the transfer learning adopted ResNext101 network structure, which consists of 101 layers, and was trained using 300,000 Kinetics video clips.²² Only the parameters in the last few layers of ResNext101 were fine-tuned during the training stage of our dataset. We added nodes to the second-to-last layer in ResNext101 structure to accommodate covariates (Figure S1 in supporting information). Our preliminary classification results showed that progressive MCI was frequently predicted as AD by the CNN model, suggesting that the CNN classification possibly reclassified patients with pending diagnosis. We therefore trained our CNN models with the target classes of controls, stable MCI, and broad AD (AD and progressive MCI).

To maximize the chance of obtaining an accurate CNN model, we generated 10-fold sample splits. The three classes of subjects were evenly divided into 10 folds in a class-wise fashion; for each sample split, one fold was used as an independent test set, the remaining nine folds were randomly split into training and validation sets with the ratio of 9:1. One CNN model was learned on the training and validation sets in each of the 10 sample splits, and the one with the highest classification accuracy on the (independent) test set was selected as the best model for downstream analyses.

The second-to-last layer of our CNN was the only layer that provided input for the class probabilities at the last layer, and therefore contained the features that are the most predictive of the classification. This layer yielded 2048 image features in the adopted ResNext101 structure (Figure S1). Covariates entered CNN at this layer and their effects were passed forward to compute the loss at the last layer, which was propagated backward to all the hidden layers including the second-to-last layer. Hence, when there are covariates in our CNN model, the extracted image features are covariate adjusted. To reduce the

TABLE 1 Demographic assessment and APOE $\epsilon 4$ genotype distribution in ADNI and AIBL data

	ADNI data						AIBL data						
	No. of subjects	Age	Male/female	No. of APOE $\epsilon 4$ copies			No. of subjects	Age	Male/female	No. of APOE $\epsilon 4$ copies			MMSE
				0	1	2				0	1	2	
Control	373	74.3 \pm 6	182/191	274	91	9	107	70.8 \pm 7	51/56	76	30	1	29.1 \pm 1.1
AD	251	74.8 \pm 8	134/117	84	114	53	74	73.2 \pm 8	29/45	23	37	14	20.3 \pm 5.6
sMCI	424	73.1 \pm 8	255/169	246	141	37	10	77.2 \pm 7	8/2	5	4	1	28.0 \pm 1.5
pMCI	230	73.9 \pm 7	134/96	77	114	39	11	74.9 \pm 6	7/4	1	6	4	26.3 \pm 1.7
P-Value ¹		0.047	0.26	5.72 $\times 10^{-30}$				0.28	0.26	4.51 $\times 10^{-10}$			2.93 $\times 10^{-16}$
P-Value ²		0.72	0.64	3.21 $\times 10^{-09}$				0.25	0.43	0.033			0.18

Notes: Age is presented in a mean \pm standard deviation format.

ADNI, sMCI, and pMCI were estimated until 3 years from screening (for CNN training); AIBL, sMCI, and pMCI were estimated until 6 years from baseline (for CNN evaluation).

P-value¹: P value of comparing AD and controls.

P-value²: P value of comparing sMCI and pMCI.

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; AIBL, Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing; APOE, apolipoprotein E; MMSE, Mini-Mental State Examination; pMCI, progressive MCI; sMCI, stable MCI.

number of image features for downstream analyses, we applied principal component analysis (PCA)²³ to the 2048 image features, used the broken stick model²⁴ to estimate the number of PCs needed, followed by a L_1 -norm regularized regression model (Lasso²⁵) to select the most informative PCs for distinguishing stable and progressive MCIs. These PCs are hereafter referred to as CNN-derived image phenotypes. After CNN was trained, covariates were not needed to obtain the image phenotypes, but needed for disease status predictions.

We generated two sets of CNN-derived phenotypes. For the first, we trained a CNN model with age at screening, sex, education level, MRI field strength indicator (1.5T or 3T), and ethnicity as covariates (hereafter *Image CNN*). For the second set, we included APOE genotype as an additional covariate, along with four APOE-correlated cognitive scores at screening (hereafter *Augmented CNN* model).

To evaluate the performance of our CNN-derived image phenotypes, we correlated them with metabolites and genetic variants. We also compared them with conventional image summary measures, cognitive scores, and clinical labels. See supporting information for additional details.

2.3 | AIBL MRIs as a validation dataset

We used MR images from the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL) to evaluate the performance of the trained CNNs. AIBL, designed similarly to ADNI, is a longitudinal study that follows-up participants every 18 months until 6 years from screening. All subjects were assumed to be White, a subset of whose MRIs were provided at the ADNI website. We selected 207 subjects who had one MRI scan at baseline and either remained as controls/MCI/AD for at least 3 years or converted to AD during the 6-year follow-up. There were only 10 stable MCI and 11 progressive MCI among 207 subjects. The MRIs, 60 of which were of 1.5T magnetic strength and 147 of 3T, went through the same pre-processing step

as ADNI MRIs. Participants' characteristics were described in Dang et al.²⁶ and Ellis et al.²⁷ Age, sex, APOE $\epsilon 4$ genotype, and MMSE are summarized in Table 1 and used here as covariates. Education level and cognitive scores of ADAS and FAQ, not available to our access, were assigned values of zero to remove their effects in our CNNs. CDR was available and used here instead of CDR-SOB. Youden's J statistics²⁴ was applied to the class probabilities of CNN to determine the predicted class category.

3 | RESULTS

3.1 | Deep 3D CNN models classify transition from MCI to AD accurately

The confusion matrices for the best CNN models on unregistered MRIs are shown in Table 2. The *Augmented CNN* yielded prediction accuracy of 0.992 for broad AD, 0.986 for controls, and between 0.911 (1-year follow-up) to 0.801 (the final visit) for progressive MCI. In comparison, the *Image CNN* achieved prediction accuracy of 0.913 for broad AD, 0.906 for controls, and between 0.822 (1-year follow-up) to 0.69 (the final visit) for progressive MCI. The *Image CNN* had lower power to differentiate stable MCI from healthy controls. Both models had lower error rate of predicting stable MCI as broad AD with longer follow-up period (0.409 to 0.192 for *Augmented CNN*, 0.376 to 0.2 for *Image CNN*), implying that some of the stable MCI that were predicted by CNN as broad AD converted to AD when tracked for longer than 3 years. We also trained the two CNN models using co-registered MRIs. As shown in Figure 2A and B, CNN performance on non-registered and registered MRIs was not significantly different, verifying that CNN is able to learn translation-invariant image features. However, *Image CNN* with the same structure had a bigger performance difference between training and test samples on registered MRI than on non-registered MRI (Figure 2C), implying that *Image CNN* might be somewhat overfitting for

TABLE 2 Confusion matrix for CNN predictions

Clinical label	<i>Image CNN</i> (trained without APOE genotype and cognitive score as covariates)			<i>Augmented CNN</i> (trained with APOE genotype and cognitive score as covariates)		
	Control	Stable MCI	Broad AD	Control	Stable MCI	Broad AD
3Y_ctrl (373)	0.895 (334)	0.0107 (4)	0.0938 (35)	0.96 (358)	0.0268 (10)	0.0134 (5)
AD (482)	0.0851 (41)	0.00207 (1)	0.913 (440)	0 (0)	0.0083 (4)	0.992 (478)
1Y_sMCI (425)	0.442 (188)	0.181 (77)	0.376 (160)	0.231 (98)	0.36 (153)	0.409 (174)
2Y_sMCI (336)	0.518 (174)	0.226 (76)	0.256 (86)	0.277 (93)	0.435 (146)	0.289 (97)
3Y_sMCI (296)	0.551 (163)	0.24 (71)	0.209 (62)	0.297 (88)	0.476 (141)	0.226 (67)
4Y_sMCI (278)	0.558 (155)	0.241 (67)	0.201 (56)	0.317 (88)	0.478 (133)	0.205 (57)
5Y_sMCI (271)	0.561 (152)	0.232 (63)	0.207 (56)	0.325 (88)	0.48 (130)	0.196 (53)
final_sMCI (255)	0.557 (142)	0.243 (62)	0.2 (51)	0.318 (81)	0.49 (125)	0.192 (49)
1Y_pMCI (101)	0.178 (18)	0 (0)	0.822 (83)	0.0198 (2)	0.0693 (7)	0.911 (92)
2Y_pMCI (190)	0.195 (37)	0.00526 (1)	0.8 (152)	0.0368 (7)	0.0737 (14)	0.889 (169)
3Y_pMCI (230)	0.209 (48)	0.0087 (2)	0.783 (180)	0.0522 (12)	0.0826 (19)	0.865 (199)
4Y_pMCI (248)	0.226 (56)	0.0403 (10)	0.734 (182)	0.0484 (12)	0.109 (27)	0.843 (209)
5Y_pMCI (255)	0.231 (59)	0.0549 (14)	0.714 (182)	0.0471 (12)	0.118 (30)	0.835 (213)
final_pMCI (271)	0.255 (69)	0.0554 (15)	0.69 (187)	0.0701 (19)	0.129 (35)	0.801 (217)

Notes: Number of samples is given in parentheses. The fraction at each entry stands for the ratio of the number of CNN predictions belonging to the column category to the number of samples belonging to the row (clinical) category. That pMCI or AD were predicted as broad AD, and control/sMCI were predicted as non broad AD can be viewed as correct predictions in a broad sense.

Abbreviations: AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; APOE, apolipoprotein E; CNN, convolutional neural network; pMCI, progressive MCI; sMCI, stable MCI.

registered MRI. This is probably because the transferred pre-trained CNN was trained on non-registered images. Hence, all the downstream analyses are based on non-registered MRIs.

3.2 | Trained CNNs applied to AIBL MRIs

CNN analysis was also informative in our AIBL validation dataset, especially for *Image CNN* as the relevant cognitive data were not available for *Augmented CNN*. Area under the curve (AUC) of distinguishing controls from AD was 0.78 for *Image CNN* and 0.76 for *Augmented CNN*. AUC of differentiating stable and progressive MCI were lower—0.61 and 0.6, respectively (Table S3 in supporting information)—probably due to the smaller sample size (21 subjects total). The *Augmented CNN* performed worse than the *Image CNN*, probably because ADAS, FAQ and CDR-SOB were not available. Potential reasons for the reduced AUC compared to ADNI analysis include (1) MRI acquisitions in ADNI and AIBL used slightly different protocols and (2) some overfitting in the trained CNN in spite of the transfer learning, as the trained CNN was fitted to a North American population whereas AIBL data were drawn from Australia.

3.3 | Image phenotypes derived from 3D CNNs

Using Lasso,²⁵ we selected principal components 1, 4, and 9 as informative phenotypes from *Image CNN* model (hereafter imageCNN.PC1, imageCNN.PC4, imageCNN.PC9), explaining 0.257, 0.035, 0.019 of the

variance of the 2048 CNN-derived image features, respectively. Only PC 2 (hereafter augmentedCNN.PC2) was selected from *Augmented CNN* model, explaining 0.064 of the variance of the 2048 image features. Only imageCNN.PC4 and augmentedCNN.PC2 had a high correlation coefficient of 0.72; all other pairwise correlations of PCs were below 0.3.

3.4 | Image phenotypes are associated with early AD-related metabolites

Seven metabolites were found to be significantly ($P < 0.05/55 = 0.0009$) associated with the four CNN-derived phenotypes in the ADNI-1 participants (Table S4 in supporting information), including two phosphatidylcholines (PC) metabolites (PC ae C44:4 associated with imageCNN.PC1, PC aa C32:3 with augmentedCNN.PC2), and three sphingomyelin (SM) metabolites (SM C16:1, SM C18:0, and SM C20:2, with augmentedCNN.PC2). These PC or SM metabolites were previously found to be significantly associated with cerebrospinal fluid (CSF) A β 1-42 and/or CSF tau in ADNI-1 cohort, either directly or indirectly.²⁸ Significant branched-chain amino acids included histidine (with imageCNN.PC9) and isoleucine (with augmentedCNN.PC2), both of which have been previously implicated in insulin resistance.^{28,29} We need to caution the reader that the interpretation may be overstated as the metabolite association was performed only in ADNI-1 participants while image-derived phenotypes were obtained from ADNI-1/GO/2 participants.

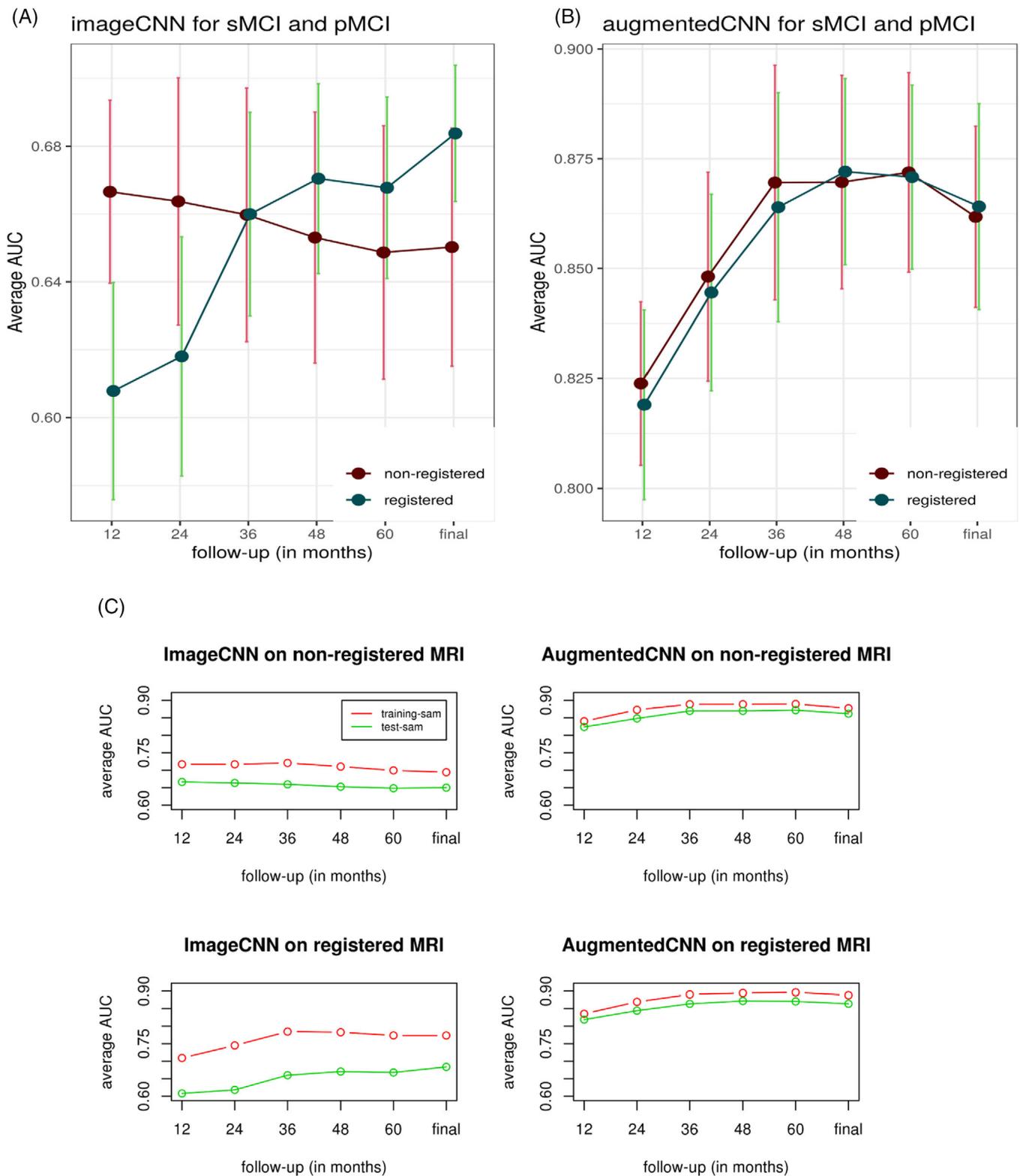


FIGURE 2 Average AUC (area under the receiver operating characteristic curve) of predicting stable and progressive mild cognitive impairment (MCI) among 10 sample splits for six follow-up periods, comparing registered and non-registered images. A, *Image CNN* model on test samples (with error bar). B, *Augmented CNN* model on test samples (with error bar). C, Comparison of average AUC between training and test samples. CNN, convolutional neural network; pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment

3.5 | GWAS using CNN-derived image phenotypes

The results for meta GWAS on imputed SNPs (see [supporting information](#)) using the four CNN-derived image phenotypes are shown in Table 3 and Table S5 in supporting information. QQ plots for all phenotypes show no obvious inflation of large P -values ($> 1 \times 10^{-4}$), with λ_{GC} between 0.993 and 1.01, indicating that our association analyses have accounted for population substructure well (Figure S2 in supporting information).

Using imageCNN.PC1 as a phenotype, we obtained genome-wide significant P -values ($< 5 \times 10^{-8}$) for genetic variants at the *APOE/TOMM40* locus (Figure S2A). However, its QQ plot shows no upward deviation from the diagonal line when single nucleotide polymorphisms (SNPs) at the *APOE/TOMM40* locus are excluded (Figure S2B), suggesting that imageCNN.PC1 is not significantly associated with any genetic variants outside *APOE* given the current sample size. In contrast, using imageCNN.PC4, imageCNN.PC9, and augmentedCNN.PC2 as phenotypes revealed no significant variants at the *APOE/TOMM40* locus ($P < 1 \times 10^{-5}$). Their QQ plots show moderate excess of low P -values even when SNPs at the *APOE/TOMM40* locus are excluded (Figure S2C-2E), suggesting that these phenotypes are significantly associated with variants outside *APOE* given the current sample size.

GWAS based on imageCNN.PC4 and imageCNN.PC9 identified 116 and 41 significant ($P < 1 \times 10^{-5}$) SNPs, respectively (Figure 3A), which were mapped to 17 protein-coding genes (within ± 15 kb) according to FUMA.³⁰ Identified *SLC24A4* (rs12588868, $P = 9.07 \times 10^{-6}$) is a known AD gene;⁴ two genes, *CACNA1C* (rs11062078, $P = 3.14 \times 10^{-6}$) and *DYSF* (rs34707417, $P = 6.38 \times 10^{-6}$), were significantly enriched in the Accelerating Medicines Partnership-Alzheimer's Disease (AMP-AD) gene co-expression submodules (Figure S3 in supporting information). The Gene Ontology (GO) annotations for the enriched submodules include regulation of action potential and calcium-mediated signaling for *CACNA1C* and regulation of endocytosis for *DYSF* (Table S7 in supporting information). *DYSF* has been reported to be significantly associated ($P < 1 \times 10^{-4}$) with AD in an exome array association study.³¹

AugmentedCNN.PC2 identified 130 SNPs with P -values $< 1 \times 10^{-5}$ (Figure 3B), which were mapped to 12 protein-coding genes (within ± 15 kb) according to FUMA. Two genes, *CDH13* (rs67805160, $P = 4.26 \times 10^{-6}$) and *ENSA* (rs112175941, $P = 2.48 \times 10^{-6}$) were significantly enriched in AMP-AD gene co-expression submodules (Figure S3), whose GO annotations include calcium-dependent cell-cell adhesion for *CDH13*, negative regulation of dephosphorylation, and regulation of hormone/insulin secretion for *ENSA* (Table S8 in supporting information).

3.6 | Candidate genes link image phenotypes to AD-related functions

For *Image CNN*, three candidate genes (*SLC24A4*, *CACNA1C*, *DYSF*) are related to the GO term of calcium ion binding, which may in turn play

a role in synaptic plasticity.³² *NCAM2* mediates synaptic adhesion, and $A\beta$ -dependent disruption of *NCAM2* functions in the AD hippocampus contributes to synapse loss.³³ *BRSK1* is the eQTL target gene of significant rs429498 in the AMP-AD RNA-Seq data, and mediates phosphorylation of tau.³⁴ Therefore, we believe that *Image CNN* detected image patterns that are related to calcium ion binding, $A\beta$ -mediated synaptic loss, and tau phosphorylation.

For *Augmented CNN*, *CDH13* negatively regulates axon growth³⁵ and *LMF1* is required for maturation and transport of active lipoprotein lipase (LPL).³⁶ Previous studies have established that LPL is a novel $A\beta$ -binding protein promoting cellular uptake and subsequent degradation of $A\beta$.³⁷ *ENSA* is an inhibitor of protein phosphatase 2A (PP2A)³⁸ that regulates tau phosphorylation directly. *ADCY3* loss-of-function variants increase the risk of obesity and type 2 diabetes.³⁹ *ZC3H12A* (a.k.a. *MCP1P1*), detected by both CNN models, is an *APP*-regulated inflammation respondent in NT2 cells.⁴⁰ All together, we believe that augmentedCNN.PC2 represented both early ($A\beta$ and tau related) and late (insulin resistance/diabetes and inflammation response) AD characteristics.

4 | DISCUSSION

Augmented CNN model achieves higher prediction accuracy than *Image CNN* model in the ADNI cohort. The high accuracy achieved by both models, as well as the four cognitive scores and *APOE* genotype, for distinguishing stable and progressive MCI (Table S2 in supporting information) implies that image, cognitive performance, and genetics have complementary roles in disease status prediction.

imageCNN.PC1 is the only CNN-derived phenotype to identify genome-wide significant SNPs at the *APOE* locus because principal components are uncorrelated with one another by definition and *Augmented CNN*-derived image features were *APOE*-adjusted. This phenotype also had the highest power (AUC = 0.784) to predict the clinical conversion of MCI to AD among all the phenotypes we considered (Table S2), and the highest Pearson correlation with cognitive scores (0.55 with ADAS and 0.52 with CDR-SOB). However, its QQ plot shows no excess of low P -values outside of the *APOE* locus. We therefore believe that imageCNN.PC1 represents image features that are mainly redundant with *APOE* genotype and cognitive performance.

The other three CNN-derived phenotypes show low correlations with cognitive scores (< 0.1) and have relatively low power to predict the clinical conversion of MCI to AD (Table S2). This is probably due to the conversion assessment being largely based on cognitive performance in the first place. However, these phenotypes are associated with early-stage markers of disease. For example, augmentedCNN.PC2 is significantly associated with the largest number of lipid metabolites (three sphingomyelin and one phosphatidylcholines; Table S4), which have been previously found significantly associated with CSF $A\beta$ 1-42 and/or CSF tau.²⁸ Furthermore, imageCNN.PC9 is significantly associated with a metabolite of histidine, which has been implicated in insulin resistance and p-tau.^{29,41} Their most significant SNPs map to

TABLE 3 Genome-wide association results based on the two sets of CNN-derived phenotypes

Lead SNP	Image phenotype	Chr	Position (hg19)	A1	A2	AF ¹ (A2)	GMMAT score (A1)	GMMAT standard error	P-value	SNP type	eQTL genes in AMP-AD	Nearby genes (+15 KB)	AMP-AD logFC ²	Other associated phenotypes [43]
<i>Image CNN model</i>														
rs11558606	PC9	1	230814668	A	G	0.072	5.243	1.153	5.41×10^{-6}	missense		COG2*	0.096	edu
rs6672949 [#]	PC4	1	37985911	C	T	0.361	-12.860	2.673	1.51×10^{-6}	intergenic		ZC3H12A, MEAF6*, SNIP1	-0.222	NA
rs34707417	PC9	2	71708810	T	G	0.236	8.605	1.906	6.38×10^{-6}	intronic		DYSF	NA	NA
rs12361440	PC9	11	74396631	G	A	0.151	-12.19	2.543	1.63×10^{-6}	intergenic		POLD3, CHRDL2	NA	NA
rs11062078	PC4	12	2135487	C	T	0.26	-11.17	2.397	3.14×10^{-6}	intergenic		CACNA1C, DCP1B3	-0.089	edu, high
rs35047	PC9	12	31163186	G	T	0.136	11.516	2.531	5.36×10^{-6}	intergenic	DDX11	TSPAN11*	0.466	NA
rs12588868	PC9	14	92909309	T	C	0.499	16.114	3.631	9.07×10^{-6}	intronic		SLC24A4	NA	edu
rs429498	PC4	19	56455746	G	A	0.441	-12.830	2.664	1.46×10^{-6}	intergenic	BRSK1	NLRP13, NLRP8	NA	NA
rs8115712	PC4	20	4746960	G	A	0.328	11.438	2.535	6.43×10^{-6}	intergenic		RASSF2	NA	NA
rs8116731	PC9	20	16718784	A	G	0.114	-11.820	2.271	1.96×10^{-7}	intronic		SNRPB ² , OTOR	-0.132	NA
rs117100735	PC4	20	18041336	G	T	0.063	4.461	1.005	8.97×10^{-6}	intergenic		OVOL2	NA	NA
rs35278766	PC4	21	22815262	C	T	0.07	6.290	1.374	4.68×10^{-6}	intronic		NCAM2*	0.299	edu, high, cog
<i>Augmented CNN model</i>														
rs112175941	PC2	1	150558293	T	A	0.159	6.594	1.400	2.48×10^{-6}	intergenic	RPRD2, CTSS	ENSA, MCL13	0.407	edu, math, high
rs6698178 [#]	PC2	1	37964765	A	T	0.351	8.954	1.901	2.47×10^{-6}	intronic	EPHA10, GNL2, DNALI, MEAF6	ZC3H12A, MEAF6*	-0.222	NA
rs144916872	PC2	2	25125902	G	T	0.1	5.773	1.171	8.28×10^{-7}	intronic	ADCY3	ADCY3*	0.306	NA
rs8654	PC2	5	96498783	A	G	0.407	-8.740	1.935	6.30×10^{-6}	synonymous	RIOK2	RIOK2	NA	NA
rs7821522	PC2	8	51650500	A	C	0.206	7.136	1.517	2.56×10^{-6}	intronic		SNTG1*	-0.352	NA
rs1426205	PC2	15	49117232	C	A	0.137	6.305	1.333	2.24×10^{-6}	UTR		CEP152, SHC43	-0.327	NA
rs67805160	PC2	16	83734755	G	T	0.167	6.730	1.464	4.26×10^{-6}	intronic		CDH13*	-0.533	edu, math, high, cog
rs4984939	PC2	16	893181	G	A	0.445	-5.378	1.213	9.33×10^{-6}	intergenic		LMF1*	-0.347	edu, math, high, cog

¹Allele frequency.²Log fold change of transcript abundance for AD cases versus controls in AMP-AD RNA-Seq data.³cog, cognitive test performance; edu, educational attainment; high, highest math achievement; math, self-reported math ability; PC, principal component.[#]These two SNPs are in high LD and have P-values $< 1 \times 10^{-5}$ in GWAS from the two CNN models.

*Gene with differential expression in AMP-AD RNA-Seq data.

Abbreviations: AMP-AD, Accelerating Medicines Partnership-Alzheimer's Disease; CNN, convolutional neural network; SNP, single nucleotide polymorphism.

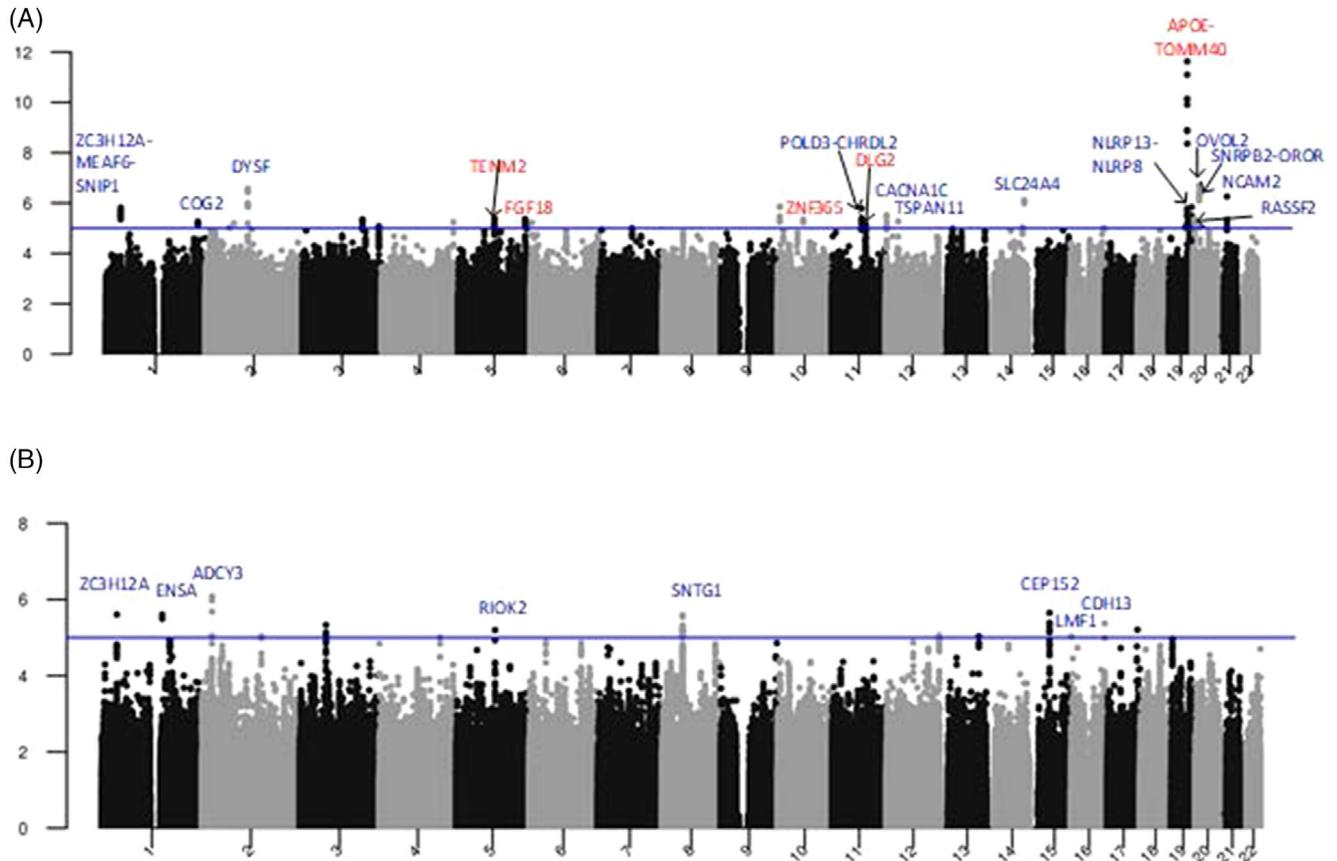


FIGURE 3 Manhattan plots for (A) principal components (PCs) 1, 4, and 9 of the *Image CNN*-derived image features, and (B) PC 2 of the *Augmented CNN*-derived image features. Gene names in red text are for imageCNN.PC1, those in blue text are for imageCNN.PC4, imageCNN.PC9, and augmentedCNN.PC2

protein-coding genes that are enriched for diverse AD stages, ranging from early $A\beta$, tau phosphorylation, and calcium ion binding-related synaptic loss to late energy hypo-utilization and inflammation response. Moreover, our CNN-derived phenotypes compare favorably to other AD-related phenotypes (cognitive scores, image summary measures, and clinical labels) in terms of metabolite association and GWAS findings (Table S4 and S6 in supporting information). Although the CNN-derived image phenotypes could be explained to some degree by a linear combination of image summary measures from ROIs—with the highest explaining R^2 of 0.358 for imageCNN.PC1 (Figure S4A in supporting information), followed by the explaining R^2 of 0.133 for augmentedCNN.PC2 (Figure S4B) the majority of the image phenotypes were unexplained by the ROIs, showing that the CNN-derived image phenotypes provide novel MRI biomarkers.

These findings suggest that our CNN-derived image phenotypes reflect AD progression better than other common phenotypes and refine the genetic associations to key subprocesses for LOAD. Three reasons may explain this. First, the transfer learning technique greatly augments the ADNI image data,¹⁷ making the learned CNN models more robust to overfitting. Second, unlike our image phenotypes, case/control LOAD GWAS often exclude MCI due to their uncertain disease status and could misdiagnose healthy controls that develop AD later, yielding a less precise and less specific phenotype. Third, the cate-

gorical clinical labels are more prone to errors due to the use of thresholds than the continuous CNN-derived image phenotypes.⁴²

Although we have applied FreeSurfer to the downloaded pre-processed MRI images to correct motion and normalize image intensities, we acknowledge that some confounding effects may not have been accounted for, due to the different MRI acquisition parameters adopted at different sites. This could also explain the lower prediction power in the independent AIBL MRIs.

One direction for future studies is to explore CNN training strategies that can better tolerate inaccurate target labels; another one is to identify the regions in the original MRI images that drive the CNN classification.

ACKNOWLEDGMENTS

We would like to thank Jim Peterson for his C++ script of cropping MRIs, and Nikhil Milind for his help in the AMP-AD gene expression submodule analysis. We are grateful to Kwangsik Nho for his advice on MRI pre-processing, and to Andrew J. Saykin for his insightful analysis suggestions. This study was supported by the National Institutes of Health grant U54 AG 054354.

Data used in 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 is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC; Johnson & Johnson Pharmaceutical Research & Development LLC; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The results published here are in part based on data obtained from the AMP-AD Knowledge Portal (<https://doi.org/10.7303/syn2580853>). ROSMAP Study data were provided by the Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago. Data collection was supported through funding by NIA grants P30AG10161, R01AG15819, 363 R01AG17917, R01AG30146, R01AG36836, U01AG32984, U01AG46152; the Illinois Department of Public Health; and the Translational Genomics Research Institute. Mayo RNA-Seq Study data were provided by the following sources: The Mayo Clinic Alzheimer's Disease Genetic Studies, led by Dr. Nilufer Ertekin-Taner and Dr. Steven G. Younkin, Mayo Clinic, Jacksonville, FL using samples from the Mayo Clinic Study of Aging, the Mayo Clinic Alzheimer's Disease Research Center, and the Mayo Clinic Brain Bank. Data collection was supported through funding by NIA grants P50 AG016574, R01 AG032990, U01 AG046139, R01 AG018023, U01 AG006576, U01 AG006786, R01 AG025711, R01 AG017216, R01 AG003949, NINDS grant R01 NS080820, CurePSP Foundation, and support from Mayo Foundation. Study data include samples collected through the Sun Health Research Institute Brain and Body Donation Program of Sun City, Arizona. The Brain and Body Donation Program is supported by the National Institute of Neurological Disorders and Stroke (U24 NS072026 National Brain and Tissue Resource for Parkinson's Disease and Related Disorders), the National Institute on Aging (P30 AG19610 Arizona Alzheimer's Disease Core Center), the Arizona Department of Health Services (contract 211002, Arizona Alzheimer's Research Center), the Arizona Biomedical Research Commission (contracts 4001, 0011, 05-901 and 1001 to the Arizona Parkinson's Disease Consortium), and the Michael J. Fox Foundation for Parkinson's Research.

MSBB data were generated from *post mortem* brain tissue collected through the Mount Sinai VA Medical Center Brain Bank and were

provided by Dr. Eric Schadt from Mount Sinai School of Medicine. Furthermore, Emory study data were supported through funding by NIA grants P50 AG025688, U01 AG046161, and U01 AG061357. Data used to validate trained CNNs are from the Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing (AIBL). AIBL is a study to discover which biomarkers, cognitive characteristics, and health and lifestyle factors determine subsequent development of symptomatic Alzheimer's disease (<https://aibl.csiro.au/>) and supplies a subset of data to ADNI researchers under a collaboration agreement.

AUTHOR CONTRIBUTIONS

YL and GWC designed the project. YL and CJ modified the 3D-CNN structure, YL performed the CNN training, image-derived phenotype construction, and downstream analysis. CP annotated the associated variants and genes, AH carried out the AMP-AD gene expression submodule enrichment analysis, AU and HSY performed RNA-Seq data analysis, KK and VP advised on data analysis. YL and GWC drafted the manuscript. YL, AH, and GWC revised the manuscript. GWC obtained the funding and was responsible for the study supervision. All authors read and approved the final manuscript.

DATA SHARING

Summary statistics from the GWAS based on the CNN-derived image phenotypes will be made available for download upon publication from <https://www.synapse.org/#!Synapse:syn21069604>. The Python codes for CNN fine-tuning and prediction are also available there. However, the ADNI MRI images, which are controlled data, are not allowed to accompany the Python codes.

URLS

ADNI: <http://www.adni-info.org>

ADNI database: <http://adni.loni.usc.edu/>

AIBL: <https://aibl.csiro.au/adni/index.html>

AIBL hosted at ADNI: <http://adni.loni.usc.edu/study-design/collaborative-%20studies/aibl/>

AMP-AD: <https://adknowledgeportal.synapse.org>

FUMA: <http://fuma.ctglab.nl/>

CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

REFERENCES

1. Jack CR, Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535-562.
2. Jansen IE, Savage JE, Watanabe K, et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat Genet*. 2019;51:404-413.
3. Kunkle BW, Grenier-Boley B, Sims R, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat Genet*. 2019;51:414-430.
4. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45:1452-1458.
5. Racine AM, Brickhouse M, Wolk DA, Dickerson BC. The personalized Alzheimer's disease cortical thickness index predicts likely pathology

- and clinical progression in mild cognitive impairment. *Assess Dis Monit*. 2018;10:301-310.
6. Rajagopalan V, Pioro EP. Disparate Voxel Based Morphometry (VBM) results between SPM and FSL softwares in ALS patients with frontotemporal dementia: which VBM results to consider?. *BMC Neurology*. 2015;15.
 7. Shen L, Kim S, Risacher SL, et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: a Study of the ADNI cohort. *Neuroimage*. 2010;53:1051-1063.
 8. Jo T, Nho K, Saykin AJ. Deep Learning in Alzheimer's Disease: diagnostic classification and prognostic prediction using neuroimaging data. *Front Aging Neurosci*. 2019;11.
 9. Mubeen AM, Asaei A, Bachman AH, et al. A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment. *J Neuroradiology*. 2017;44:381-387.
 10. Choi H, Jin KH, A.S.D.N. Initiative. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. *Behav Brain Res*. 2018;344:103-109.
 11. Li R, Zhang W, Suk HI, et al. Deep learning based imaging data completion for improved brain disease diagnosis. International Conference on Medical Image Computing and Computer-Assisted Intervention. 2014;17:305-312.
 12. Tien Duong V, Hyung-Jeong Y, Nguyen VQ, et al. Multimodal learning using convolution neural network and Sparse Autoencoder. in 2017 IEEE International Conference on Big Data and Smart Computing (Big-Comp). 2017.
 13. Liu S, Liu S, Cai W, et al. Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. *IEEE Trans Biomed Eng*. 2015;62:1132-1140.
 14. Hinton GE. Reducing the dimensionality of data with neural networks. *Science*. 2006;313:504-507.
 15. Baxter J. Theoretical models of learning to learn. *Learning to Learn*. 1998:71-94.
 16. Pratt L, Jennings B. A survey of connectionist network reuse through transfer. *Learning to Learn*. 1996:19-43.
 17. Hara K, Kataoka H, Satoh Y, Can spatiotemporal 3D CNNs retrace the history of 2D CNNs and ImageNet?, in IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018, IEEE.
 18. Segonne F. A hybrid approach to the skull stripping problem in MRI. *Neuroimage*. 2004.
 19. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998;17:87-97.
 20. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60:84-90.
 21. LeCun Y, Bottou L, Orr GB, Muller KR. Efficient BackProp. *Lect Notes Comput Sci*. 1998.
 22. Xie S, Girshick R, Dollar P, et al. Aggregated residual transformations for deep neural networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017.
 23. Hotelling H. Analysis of a complex of statistical variables into principal components. *J Educ Psychol*. 1933;24:417-441.
 24. Jackson DA. Stopping rules in principal components analysis: a comparison of heuristic and statistical approaches. *Ecology*. 1993;74:2204-2214.
 25. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol*. 1996;58:267-288.
 26. Dang C, Yassi N, Harrington KD, et al. Rates of age- and amyloid β -associated cortical atrophy in older adults with superior memory performance. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. 2019;11:566-575.
 27. Ellis KA, Bush AI, Darby D, et al. The Australian Imaging, Biomarkers and Lifestyle (AIBL) study of aging: methodology and baseline characteristics of 1112 individuals recruited for a longitudinal study of Alzheimer's disease. *Int Psychogeriatr*. 2009;21:672-687.
 28. Toledo JB, Arnold M, Kastenmuller G, et al. Metabolic network failures in Alzheimer's disease: a biochemical road map. *Alzheimers Dement*. 2017;13:965-984.
 29. Feng RN, Niu YC, Sun XW, et al. Histidine supplementation improves insulin resistance through suppressed inflammation in obese women with the metabolic syndrome: a randomised controlled trial. *Diabetologia*. 2013;56:985-994.
 30. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun*. 2017;8.
 31. Chen JA, Wang Q, Davis-Turak J, et al. A multiethnic genome-wide exome array study of Alzheimer disease, frontotemporal dementia, and progressive supranuclear palsy. *JAMA Neurol*. 2015;72:414-422.
 32. Neher E, Sakaba T. Multiple roles of calcium ions in the regulation of neurotransmitter release. *Neuron*. 2008;59:861-872.
 33. Leshchynska I, Liew HT, Shepherd C, et al. $A\beta$ -dependent reduction of NCAM2-mediated synaptic adhesion contributes to synapse loss in Alzheimer's disease. *Nat Commun*. 2015;6:8836.
 34. Yoshida H, Goedert M. Phosphorylation of microtubule-associated protein tau by AMPK-related kinases. *J Neurochem*. 2012;120:165-176.
 35. Rivero O, Seltzer MM, Sich S, et al. Cadherin-13, a risk gene for ADHD and comorbid disorders, impacts GABAergic function in hippocampus and cognition. *Transl Psychiatry*. 2015;5:e655-e655.
 36. Attie AD. High-maintenance proteins and hypertriglyceridemia. *Nat Genet*. 2007;39:1424-1425.
 37. Nishitsuji K, Hosono T, Uchimura K, Michikawa M. Lipoprotein lipase is a novel amyloid beta (A β)-binding protein that promotes glycosaminoglycan-dependent cellular uptake of A β in astrocytes. *J Biol Chem*. 2011;286:6393-6401.
 38. Mochida S, Maslen SL, Skehel M, Hunt T. Greatwall phosphorylates an inhibitor of protein phosphatase 2A that is essential for mitosis. *Science*. 2010;330:1670-1673.
 39. Grarup N, Moltke I, Andersen MK, et al. Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes. *Nat Genet*. 2018;50:172-174.
 40. Vrotsos EG, Kolattukudy PE, Sugaya K. MCP-1 involvement in glial differentiation of neuroprogenitor cells through APP signaling. *Brain Res Bull*. 2009;79:97-103.
 41. Arnold M, Nho K, Kueider-Paisley A, et al. The Alzheimer's Disease Metabolome: effects of Sex and APOE ϵ 4 genotype. *bioRxiv*. 2019:585455.
 42. Shen L, Thompson Pm Fau-Potkin SG, Potkin Sg Fau - Bertram L, et al. Genetic analysis of quantitative phenotypes in AD and MCI: imaging, cognition and biomarkers. *Brain Imaging Behav*. 2014;8:183-207.
 43. Lee JJ, Wedow R, Okbay A, et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat Genet*. 2018;50:1112-1121.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Li Y, Haber A, Preuss C, et al. Transfer learning-trained convolutional neural networks identify novel MRI biomarkers of Alzheimer's disease progression. *Alzheimer's Dement*. 2021;13:e12140. <https://doi.org/10.1002/dad2.12140>