의학박사 학위논문

Deep learning-based approach for the prediction of post-stroke dementia using brain FDG PET

뇌 FDG PET 을 이용한 딥러닝 기반 뇌졸중 후 치매 예측 연구

2022 년 8 월

서울대학교 대학원

의학과 핵의학 전공

이 리 리

Deep learning-based approach for the prediction of post-stroke dementia using brain FDG PET

지도교수 천 기 정

이 논문을 의학박사 학위논문으로 제출함
2022년 4월

서울대학교 대학원
의학과 핵의학 전공
이 리 리

이리리의 박사 학위논문을 인준함
2022년 7월

위 원 장 이 재 성 (인)

부위원장 천 기 정 (인)

위 원 권 형 민 (인)

위 원 최 홍 윤 (인)

위 원 석 주 원 (인)

Abstract

Deep learning-based approach for the prediction of post-stroke dementia using brain FDG PET

Reeree Lee

Department of Nuclear Medicine

The Graduate School

Seoul National University

Post-stroke cognitive impairment can affect up to one-third of stroke survivors. Since cognitive function greatly contributes to patients' quality of life, an objective quantitative biomarker for early prediction of dementia after stroke is required. Brain [^{18}F]fluorodeoxyglucose (FDG) positron emission tomography (PET) has been widely used for evaluating cognitive function with typical hypometabolic patterns. Here, a deep-learning (DL)-based signature using brain FDG PET was developed to objectively evaluate post-stroke dementia. Additionally, an association between DL-derived cognitive signature and gene co-expression network analysis (WGCNA)-based gene expression signature was evaluated to elucidate the imaging phenotype-genomics relation regarding the cognitive function.

A DL model was built to differentiate Alzheimer's disease (AD) from normal controls (NC) using brain FDG PET from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. The model was directly transferred to a prospectively enrolled cohort of patients with stroke to differentiate patients with dementia from those without dementia. The accuracy of the model was evaluated by the area under

the curve values of receiver operating characteristic curves (AUC-ROC). The distribution of DL-based features and brain regions that the model weighted for classification was visualized. Correlations between cognitive signature from the DL model and clinical variables were evaluated, and survival analysis for post-stroke dementia was performed in patients with stroke.

Microarray gene expression data from blood samples of the ADNI whole genome sequencing (WGS) cohort participants was used for WGCNA. After preprocessing of gene expression profile, weighted correlation network analysis was performed and gene modules, clusters of highly interconnected genes, were identified. Then, relationships between module gene expression and clinical features, including DL-derived cognitive signature, were estimated to determine relevant modules. To apprehend the underlying biological meaning of the relevant modules, pathway/process enrichment analyses were applied.

The classification of AD vs. NC subjects was performed with AUC-ROC of 0.94 (95% confidence interval [CI], 0.89–0.98). The transferred model discriminated stroke patients with dementia (AUC-ROC = 0.75). The score of cognitive decline signature from the DL-model was positively correlated with age, neutrophil–lymphocyte ratio and platelet-lymphocyte ratio and negatively correlated with body mass index in patients with stroke. The cognitive decline score was an independent risk factor for dementia following stroke after adjustment for other key variables.

Total of 24,198 genes were divided into 14 modules on WGCNA with hierarchical clustering algorithm. Among 14 modules, black, greenyellow, pink, red, tan, and brown modules showed a significant correlation with DL-based cognitive signature. Among them, black, greenyellow, and brown modules were significantly correlated with dementia status, as well. And these three modules were also associated with amyloid deposition, TAU/PTAU, and risk factors for stroke. On enrichment analyses, most of the enriched ontology terms in black and brown modules were related to inflammation, leukocyte, especially neutrophil, while greenyellow module was associated with lymphocyte, B-cell activation.

The DL-based cognitive signature using FDG PET was successfully transferred to an independent stroke cohort. It is suggested that DL-based cognitive evaluation using FDG PET could be utilized as an objective biomarker for post-stroke dementia. Furthermore, this study confirmed that inflammatory

condition measured by gene expression profile of peripheral blood as well as complete blood counts are deeply related to the DL-based imaging phenotype of cognitive function and risk factors for stroke.

Keywords: Post-stroke dementia, Post-stroke cognitive impairment, [^{18}F]FDG, Deep learning, gene co-expression network

Student number: 2019-32281

Contents

List of Figures

List of Tables

Introduction

Post-stroke dementia (PSD)

Post-stroke dementia (PSD) or post-stroke cognitive impairment in patients with stroke can affect up to one-third of stroke survivors (1). Physical disability after stroke onset tends to improve; however, cognitive decline generally worsens over time for unclear reasons (2). Since cognitive function is greatly responsible for patients' functional outcomes and quality of life, early detection of cognitive decline is expected to help reduce the social and economic burden. Therefore, an objective quantitative biomarker for predicting dementia after stroke is required.

Brain FDG PET for dementia evaluation

Brain [^{18}F]fluorodeoxyglucose (FDG) positron emission tomography (PET) is a functional imaging modality that reflects glucose metabolism in the brain. Patterns of brain FDG PET related to dementia, particularly Alzheimer's disease (AD), have also been widely studied, with representative findings of diffuse and symmetric decreased uptake in the neocortical association areas, medial temporal lobe and posterior cingulate cortex (3). These hypometabolic patterns in the brain are not specific to AD. Hypometabolism in the posterior cingulate, precuneus and prefrontal areas is common in various types of dementia, including vascular dementia and AD (4). This supports the common pathophysiology of cognitive impairment affected by degeneration of neurons in several types of dementia. Furthermore, vascular disease and AD share common risk factors, including hypertension, obesity, diabetes and atherosclerosis, with a bidirectional relationship (5, 6). In this regard, in patients with stroke, evidence of cognitive impairment is likely to be detected via the brain FDG PET pattern of AD.

The necessity of applying a deep learning model for PSD prediction

However, since patients with stroke often already have a decrease in metabolism/blood flow in the affected brain region, it is challenging to predict dementia by recognizing the FDG PET pattern of dementia through visual inspection. Additionally, in stroke patients with significant physical impairment, it may be difficult to assess limitations in daily living autonomy related to cognitive function (7). Thus, there is a considerable need for an imaging biomarker to objectively and non-invasively evaluate cognitive function in such patients. To this end, a brain metabolic signature based on a deep convolutional neural network (CNN) model was applied to predict dementia using FDG PET. The bottleneck of actively applying deep learning to medical imaging modalities is a limitation of data. To overcome this, transfer learning was adopted. Transfer learning can utilize the learned knowledge of feature maps, from trained model to untrained dataset which is even different domain of images (8, 9). It is hypothesized that the CNN model that classifies AD extracts cognitive function-related patterns from the brain FDG PET and could transfer those to other disorders to reflect cognitive function. The brain metabolic signature representing cognitive impairment was derived from the model trained using a large dataset of FDG PET from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and directly transferred (10, 11) to a PET dataset of patients with stroke. Whether the suggested model could predict not only AD but also dementia after a stroke in survivors was investigated. It is clarified that Part I. of this thesis has been published in peer-review journal (12).

Investigation of the imaging phenotype–genomics association

Additionally, a correlation between imaging phenotypes of cognition derived from deep CNN model and gene expression signature using co-expression network analysis was evaluated. Standard analysis of differential expression compares gene expression by dichotomizing given conditions (e.g., imaging phenotypes), requiring pre-defined hypothesis or prioritization, in which loss of information may occur.

Co-expression networks have been found to be useful to describe the pairwise relationships among gene transcripts by exploring system-level intercorrelation of genes, based on application of graph theory, and vigorous statistical methods (13, 14). Investigating the imaging phenotype–genomics association may provide insight into cognition-related genomic events at cellular level, and furnish useful details with respect to the diagnosis, prognosis, and treatment of cognitive impairment.

Methods

Patients for deep learning model

The Cerebral Atherosclerosis Research with Positron Emission Tomography (CARPET) is a prospective registry for understanding the pathophysiology of cerebral atherosclerosis by applying FDG PET in patients with acute cerebral infarction or transient ischemic attack. This study was reviewed and approved by the Institutional Review Board of Chung-Ang University Hospital (C2015061) and all subjects signed an informed consent form. The current study was registered in the Clinical Research Information Service (registration no.: KCT0002462) as a part of the International Clinical Trials Registry Platform of the World Health Organization supported by the Korea Centers for Disease Control and Prevention. All eligible patients underwent brain computed tomography angiography (CTA) at admission. Patients with carotid atherosclerosis ≥ 50% on brain CTA were included. To understand the relationship between cerebral atherosclerosis burden and hematopoietic organ activities, stroke patients with mild carotid stenosis or without carotid atherosclerosis concurrently were also included. Exclusion criteria were as follows: patients with overt cancer or autoimmune diseases, advanced renal impairment with an estimated glomerular filtration rate < 30 mL/min/1·73 m 2, uncontrolled diabetes mellitus, or other unstable medical conditions. The enrolled patients underwent a comprehensive stroke etiology workup, including brain magnetic resonance imaging, cardiac evaluation and bone mineral density evaluation using dual-energy X-ray absorptiometry. Finally, a total of 110 patients with stroke (mean age, 72.1 ± 9.9 years; female: male, 43: 67) were included in this study, and 13 patients with pre-stroke (11.8%) and 19 with post-stroke (17.3%) dementia were classified as stroke patients with dementia (29.1%). Pre-stroke dementia was designated when a patient had been diagnosed with dementia before index stroke. Post-stroke dementia was diagnosed when a patient or caregiver complained of continued cognitive decline hampering everyday life, and objective neuropsychological tests including, but not limited to, mini-mental status examination (MMSE)

and clinical dementia rating (CDR) confirmed decreased cognitive function at least 6 months after index stroke (12).

Baseline FDG PET images acquired from participants recruited in the ADNI (N = 693; 292 AD, and 401 normal controls [NC]) were used to train a model (http://adni.loni.usc.edu). Demographics of the individuals are listed in Table 1.

Table 1. Patient characteristics

	ADNI cohort		Stroke cohort		
	AD	NC	Pre-stroke dementia	Post-stroke dementia	Non-dementia
Subject number	292	401	13	19	78
Age (y, mean ± SD)	74.9 ± 8.0	73.8 ± 6.0	79.5 ± 4.8	75.5 ± 6.7	70.1 ± 10.5
Sex (F:M)	119:173	204:197	8:5	6:13	29:49
MMSE (mean ± SD)	23.2 ± 2.2	29.0 ± 1.2	9.8 ± 8.1	13.1 ± 8.7	NA

SD, standard deviation; AD, Alzheimer's disease; NC, normal controls

FDG PET Image acquisition and processing

Once a patient was stabilized after an index stroke event, whole-body FDG PET/CT was conducted using a combined scanner (Gemini TF 16, Philips Medical Systems, Cleveland, OH, USA). All patients fasted for at least 6 h, and blood glucose levels were confirmed to be < 150 mg/dL. After the intravenous injection of 259–370 MBq (7–10 mCi) of FDG, patients waited in a quiet room for approximately 60 min with their eyes open. A low-dose CT scan (120 kVp, 50 mAs) was performed first for attenuation correction and anatomical localization, and PET images were acquired from the vertex to the proximal thigh for 1 min per bed position (6–7 bed positions for a patient). The size of field of view of PET images was 576 × 576 (mm). Images were reconstructed with BLOB-OS-TF algorithm (three iterations, 33 subsets) and the reconstructed matrix size was 144 × 144 with final voxel size of 4 × 4 mm^2. Then, the images were spatially normalized to MNI space using SPM8 (University College of London, UK) with additional smoothing with an 8-mm Gaussian filter. Finally, the spatially normalized images were 79 × 95 × 68 matrix size.

Deep CNN model architecture and training

The architecture of the proposed network is shown in Figure 1. This architecture is the result which showed best accuracy in multiple model experiments. Four 3D convolutional layers used rectified linear unit (ReLU) activation as an activation function. A global average pooling layer (GAP) summarized the feature maps (10 × 12 × 9), which were extracted from convolutional layers, into a 128-dimensional vector. The vectors were finally connected to an output layer with sigmoid function as an activation function. If a dense layer is used instead of the GAP, the feature maps are flattened and their spatial information is lost, which prevents from applying class activation map (CAM) method, to be described later. Therefore, the output of the GAP layer was connected to the last layer in this study.

FDG PET images of AD (N = 292) and NC (N = 401) participants were used to train the CNN model. And the trained model was tested by the accuracy on the internal validation set (N = 70,

randomly selected for internal validation). Adam optimizer with a learning rate of 0.001, binary crossentropy loss function, and batch size of 16 were used for model training. Iterative training was stopped by monitoring crossentropy loss function and accuracy of the internal validation set (12).

Data preparation, CNN modeling, and model experiments were conducted in Python (version 3.7.7.) using a GPU-enabled Google Colaboratory environment, Keras (version 2.4.0) and TensorFlow (version 2.4.1) frameworks. The modeling codes can be found in https://colab.research.google.com/drive/11Aff2AnbZGLgaV4AN8cXj4RuUIHm08Z1?usp=sharing.

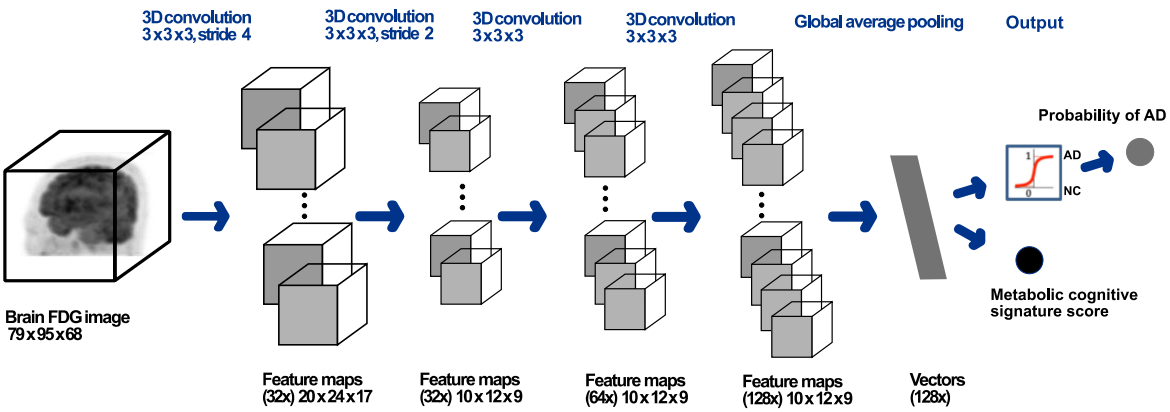


Figure 1. The architecture of the network.

The model contains four three-dimensional convolutional layers with 3 × 3 × 3 convolutional filters, and features of PET images are hierarchically extracted. A total of 128 feature maps of size 10 × 12 × 9 are vectorized by the global average pooling layer and connected to output for differentiating AD from NC.

Metabolic cognitive signature based on a deep CNN model

The accuracy of the model was evaluated by the area under the curve (AUC) values of receiver operating characteristic (ROC) curves (AUC-ROC). A ROC curve to differentiate between AD and NC in the internal validation set was drawn. The model was transferred, and the AUC-ROC was measured to test accuracy for discriminating stroke patients with dementia (pre-stroke or post-stroke) from those without dementia.

A nonlinear activation function is inevitable for classification of DL model. However, the current study was focused on the representative value of cognition rather than classification of AD vs. NC. For this reason, the metabolic cognitive signature score was obtained from the output of the CNN model without the activation function (sigmoid function) of the last layer (Figure 1). Thus, a one-dimensional score vector, named metabolic cognitive signature score, was obtained from a given PET volume.

Model visualization and interpretation

To visualize the CNN-based features, a heatmap and parametric t-distributed stochastic neighbor embedding (t-SNE) model were used. t-SNE method is a non-linear dimensionality reduction technique, aimed at preserving as much meaningful structure of the high-dimensional data as possible in the low-dimensional space. Using this technique, very similar datapoints can be preserved in low-dimensional presentation, which is practically impossible in traditional linear mapping methods (e.g., principal component analysis [PCA] or multidimensional scaling [MDS]). In t-SNE, the pairwise distances between points of the data are converted into probability and points with similar characteristics are assigned a higher probability. The probability distribution is constructed using Gaussian distribution in the high dimensional space, while t-Student distribution in the lower dimensional space. Finally, the dissimilarity between probability distributions of pairwise distance in the lower dimensional space is minimized by using Kullback-Leibler divergence (15). In this study, 128-dimensional vectors from global average pooling layer were inputted to the t-SNE model that was carried out by using scikit-learn

package, *TSNE* function to intuitively visualize the distribution of CNN derived features.

Class activation map (CAM) method was used to interpret decisions of the model. This method projects back the weights of the output layer on to the convolutional feature maps which allows for identifying which regions of an image are being used for discrimination by generating heatmap (16). Global average pooling spatially averages output feature maps of the last convolutional layer (in this study, the fourth convolutional layer). A weighted sum of the feature maps of the fourth convolutional layer is used to generate a CAM. By simply upsampling the CAM to the size of the original FDG PET volume, the image regions most relevant to AD can be identified. CAM visualization was performed with information extracted from the model using *tensorflow. keras.function*, and *get_weights* (for feature maps values, and weights, respectively) functions.

Gene expression data preprocessing

Microarray gene expression data from blood samples of the 744 ADNI whole genome sequencing (WGS) cohort participants was obtained from https://utilities.loni.usc.edu/download/files/genetic/c5992db7-4650-4ca7-9f3e-27931ed9b80c/adni/ADNI_Gene_Expression_Profile.zip and imputed for WGCNA. The details of the processing methods for microarray gene expression dataset can be found in https://utilities.loni.usc.edu/download/files/genetic/c5992db7-4650-4ca7-9f3e-27931ed9b80c/adni/ADNI_Microarry_Gene_Expression_Methods_Final_20150427.pdf. Of 49,386 transcripts, 24,198 genes with a log-2 transformed mean expression level greater than 4 were selected for the WGCNA to remove the lowly expressed genes that are not relevant to the study by visualizing the distribution of gene expression level for each sample (before filtering, Figure 2a; after filtering, Figure 2b). Among 744 participants, 639 subjects underwent brain FDG PET on the same visit as blood sampling for gene expression profiling. The metabolic cognitive signature score was obtained from brain FDG PET images of those 639 participants to evaluate a correlation between DL-based imaging phenotype for cognitive function and WGCNA-based gene expression signature.

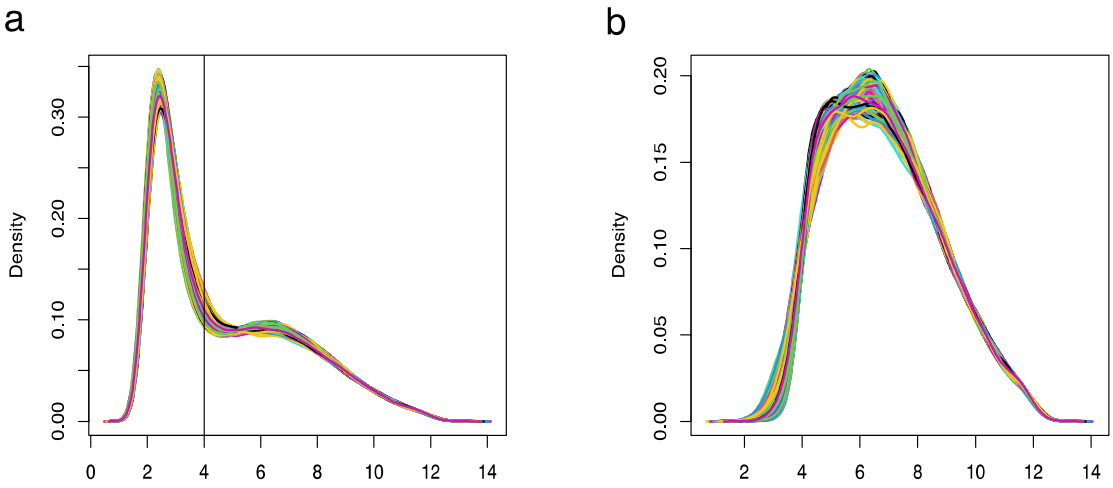


Figure 2. Density plots of log-2 transformed mean expression level of ADNI WGS cohort.

Of 49,386 transcripts, 24,198 genes with a log-2 transformed mean expression level greater than 4 were selected for the WGCNA to remove the lowly expressed genes that are not relevant to the study. The distributions of before (a) and after (b) filtering gene expression level for each sample are visualized.

Weighted correlation network

Weighted correlation network analysis was performed using R package, WGCNA (ver. 1.70-3) (14). Firstly, the co-expression similarity between gene i and j was defined as the correlation coefficient, S_{ij} = cor(i, j). Using a soft-thresholding procedure, the co-expression similarity is transformed into the weighted adjacency matrix, $a_{ij} = \text{power}(S_{ij}, \beta) = S_{ij}^{\beta}$ ($\beta >= 1$), with allowing to weigh by a continuous values between 0 and 1 rather than the binary number 0 or 1. By raising correlation to a power, disparity between strong and weak correlations can be amplified. To choose a power term, scale-free topology fitting method was tested with ranging the power value from 1 to 20. A common characteristic of real world networks is a small-world property of which few nodes that are highly connected to other nodes (k) in the network. Scale-free networks are a type of network characterized by the existence of large hubs. A scale-free network is one with a power-law degree distribution, the probability that a node is connected with k other node decays as a power law $p(k) \sim k^{-\gamma}$ (17). Because there is a trade-off between mean connectivity (maintaining mean number of connections) and maximizing scale-free topology fit, the lowest power at which the saturation of scale-free topology above 0.80 was picked as recommended (18). Then, the adjacency matrix was transformed into a topological overlap matrix to make use of a topological information as well as gene expression information (19, 20), and fed into unsupervised hierarchical clustering to identify gene modules, clusters of highly interconnected genes.

Identification of key modules related to the DL-based signature and clinical features

To incorporate clinical information into the co-expression network, gene significance (GS) measures, defined as correlation coefficient between clinical trait and gene expression value across samples, were used.

After the modules were identified, the module eigengene (ME), which is a representative value characterizing each module, was defined by the first principal component within each module. To

identify relevant modules, relationships between module gene expression and clinical features were estimated using the correlation between MEs and clinical variables. Modules with high correlation significance may represent pathways related to the clinical characteristics (14).

Enrichment analyses for key modules

Pathway/process enrichment analysis was applied to understand the underlying biological meaning of the modules by loading gene information of each modules on Metascape (http://metascape.org) with several ontology categories including GO Biological Process/Molecular functions/Cellular Components, Hallmark Gene Sets, Reactome Gene Sets, KEGG Pathway, Canonical Pathways, and BioCarta Gene Sets. The analysis applies standard accumulative hypergeometric statistical test to identify ontology terms. All resultant terms with a p-value < 0.01, a minimum hit count of 3, and an enrichment factor > 1.5 were automatically clustered into groups based on their similarities (Kappa similarity > 0.3) to reduce the degree of redundancy of ontology terms. The most enriched pathway was chosen as the representative term of the cluster (21).

Statistical Analysis

Values are expressed as mean ± standard deviation. Pearson's correlation test was performed to evaluate the correlations between the metabolic cognitive signature score and various clinical features in patients with stroke. A general linear model was used to compare the scores between stroke patients with dementia and those without dementia with age, neutrophil–lymphocyte ratio (NLR) and body mass index (BMI) as cofactors. Survival was calculated from the date of visit due to index stroke to that of the occurrence of dementia or the last follow-up visit at the hospital. The log-rank test and Cox regression were used for survival analysis. The specific cut-off values for clinical parameters and the metabolic cognitive signature score in patients with stroke were determined using the Cantal and

O'Quigley method (22). The data were analyzed using the R program (version 3.4.5, R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was set at $P < 0.05$.

Results

Part I. Development of DL-based cognitive signature

Model training and accuracies

Figure 3 shows the results of model training. The model fitting process was stopped after 120 epochs by monitoring the loss function and accuracy of internal validation dataset. The accuracy of the model to distinguish between AD and NC was assessed by the AUC-ROC and was 0.94 (95% CI, 0.89–0.99) in the internal validation set (N = 70). The model was directly transferred to differentiate stroke patients with dementia from those without dementia. The AUC-ROC for this transferred model was 0.75 (95% CI, 0.64–0.85) (Figure 4).

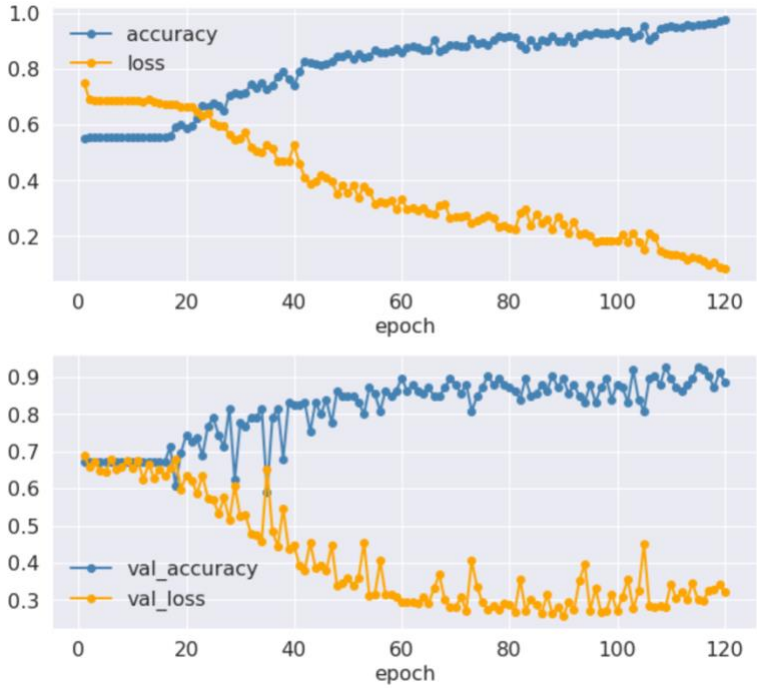


Figure 3. Results of model training process.

During the fitting process, accuracies tend to increase continuously, while loss values (crossentropy) tend to decrease in both training and internal validation datasets. The process is automatically stopped after 120 epochs.

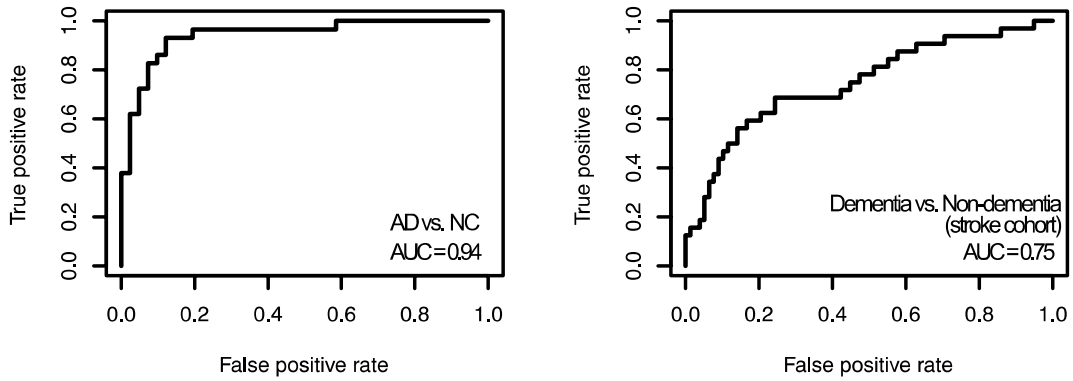


Figure 4. Accuracies of deep CNN models measured by AUC-ROC.

The AUC for differentiating AD from NC is 0.94 (95% confidence interval [CI], 0.89–0.99) for the internal validation set. The model is directly transferred to differentiating stroke patients with dementia from those without dementia with an AUC of 0.75 (95% CI, 0.64–0.85)

Visualization of CNN-based features and model interpretation

To visualize participants according to the similarity of brain FDG PET features extracted from the deep CNN model, those were projected onto the two-dimensional axes using parametric t-SNE. Each point corresponds to individual PET data, and data with similar metabolic characteristics are indicated by near points (Figure 5). All PET images from the ADNI and stroke cohorts were plotted, and dementia and non-dementia were clustered and located on the right and left sides, respectively (Figure 5a). When patients with stroke were highlighted, a similar clustering pattern (blue and red dots in the right and left sides, respectively) was observed (Figure 5b), and the metabolic cognitive signature score derived from the model was greater in the right upper area than in the left lower area (Figure 5c). Moreover, most patients with post-stroke dementia were concentrated in the right upper area, while those with pre-stroke dementia showed a scattered distribution (Figure 5d).

Regions related to dementia were visualized. CAMs generated heatmaps of brain metabolic features associated with dementia, as determined by the model. Heatmaps were drawn on individual FDG PET images of patients with stroke (Figure 6a-b), and averaged CAMs of AD patients in the ADNI cohort (Figure 7a) and PSD patients in the stroke cohort (Figure 7b) were generated.

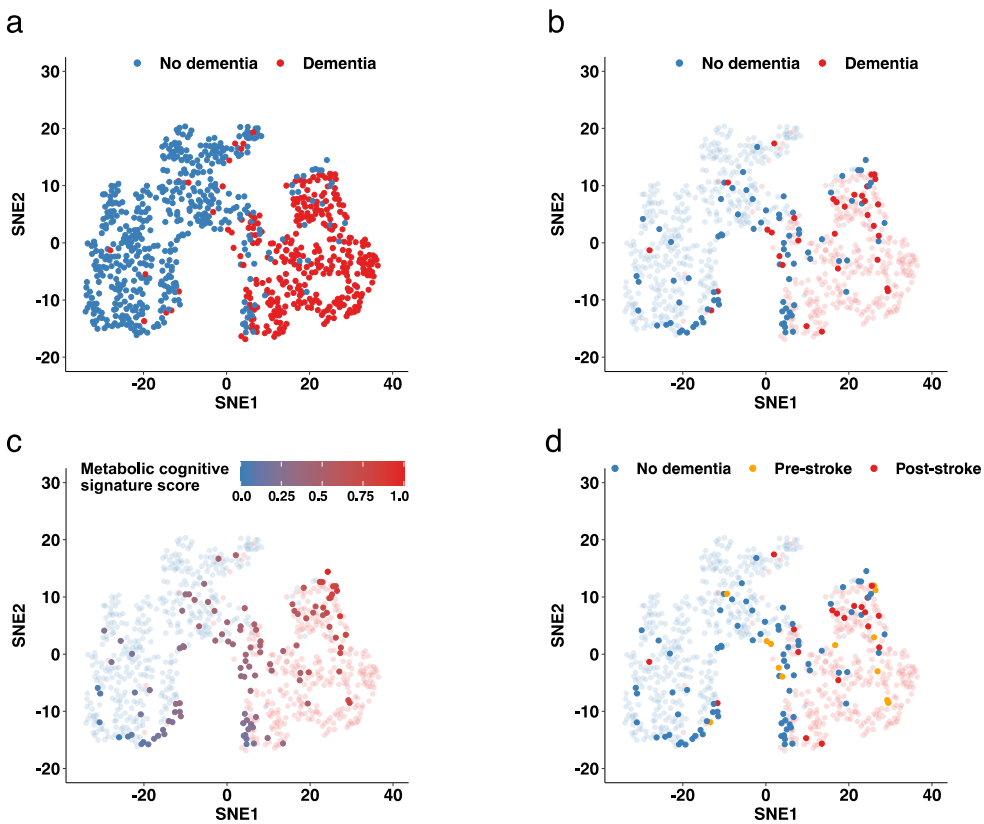


Figure 5. Visualization of 128 features extracted from four convolutional layers on two-dimensional axes. Each point represents individual PET data, and near points indicate data with similar brain metabolic characteristics. (a) When all PET images (ADNI and stroke cohorts) are inputted, dementia (red) and non-dementia (blue) are clustered, located on the right and left side, respectively. (b) A similar clustering pattern in the highlighted patients with stroke is noted. (c) The metabolic cognitive signature score derived from the model is greater in the right upper area than in the left lower area. (d) The distribution of most patients with post-stroke dementia (red) is concentrated in the right upper area, while that of patients with pre-stroke dementia (orange) is scattered.

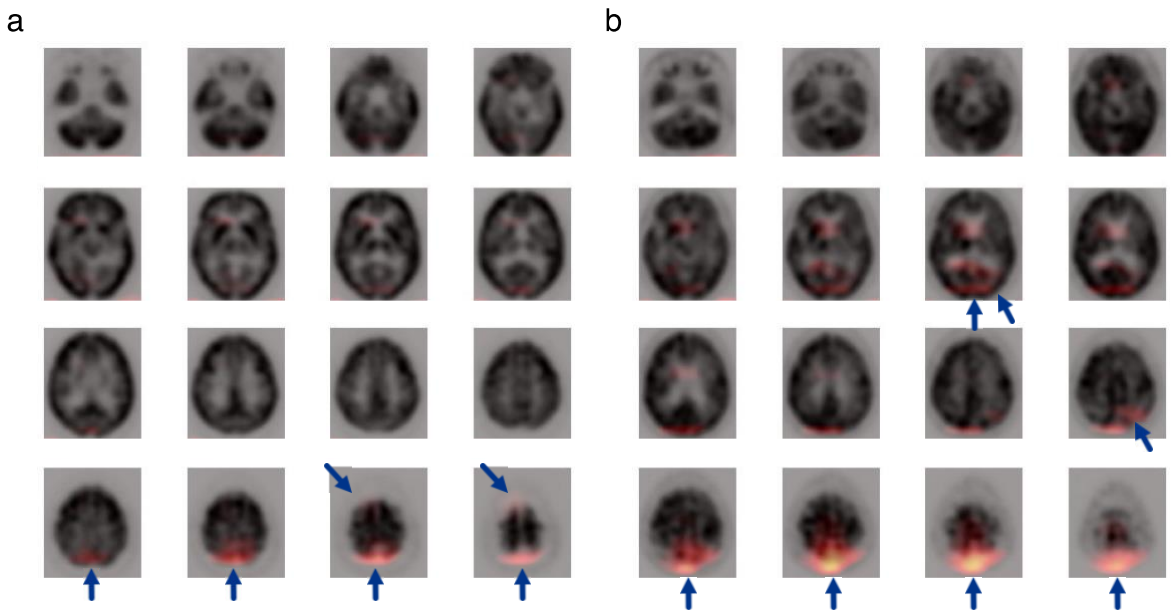


Figure 6. Representative brain FDG PET images of two post-stroke dementia patients with heatmaps showing regions associated with cognitive decline.

Although metabolic cognitive signature scores (a, 0.629; b, 0.661) are similar to each other, the presumed affected areas are different. The model weighted on bilateral parietal and right frontal cortices (a) and on bilateral parietal, left temporal and posterior cingulate cortices (b).

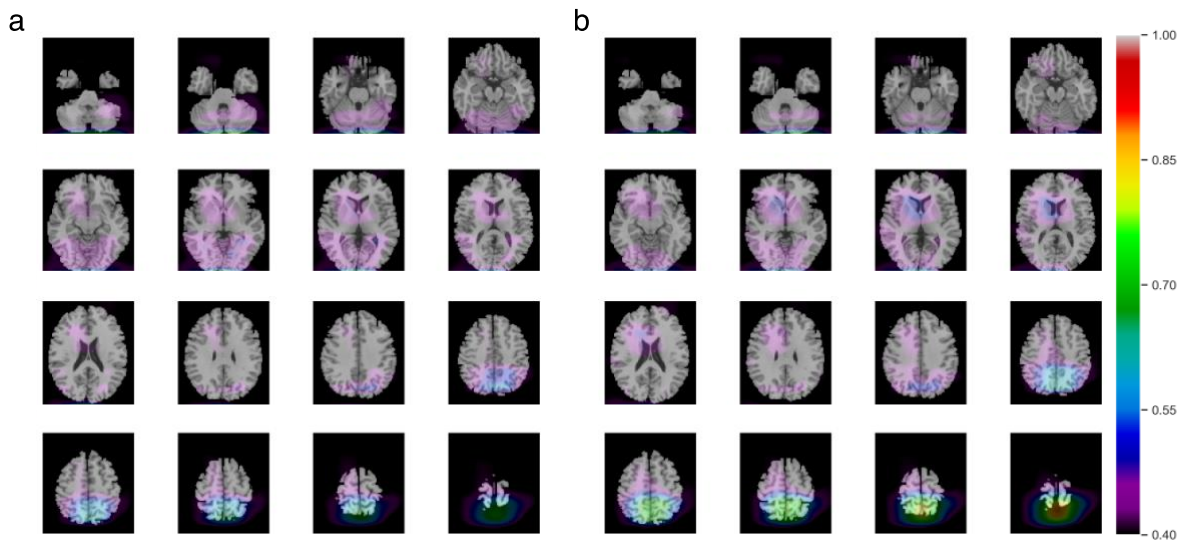


Figure 7. Averaged class activation map results in patients with AD and PSD.

The averaged brain regions weighted by the model exhibit more severe asymmetry in images of patients with AD (a) than those of patients with PSD (b). AD, Alzheimer's disease; PSD, post-stroke dementia

Metabolic cognitive signature based on CNN correlated with clinical features

Associations between metabolic cognitive signatures and clinical features were evaluated. Age, NLR and platelet-lymphocyte ratio (PLR) of patients with stroke were significantly correlated with higher metabolic cognitive signature scores ($R = 0.22$, $P = 0.02$, Figure 8a; $R = 0.32$, $P < 0.001$, Figure 8b; and $R = 0.29$, $P = 0.003$, Figure 8c, respectively). BMI was negatively correlated with metabolic cognitive signature score ($R = -0.3$, $P = 0.002$, Figure 8d). The score was significantly higher in stroke patients with dementia than in those without dementia, with age, NLR and BMI as cofactors (0.6 ± 0.2 vs. 0.4 ± 0.2, $P < 0.001$, Figure 8e).

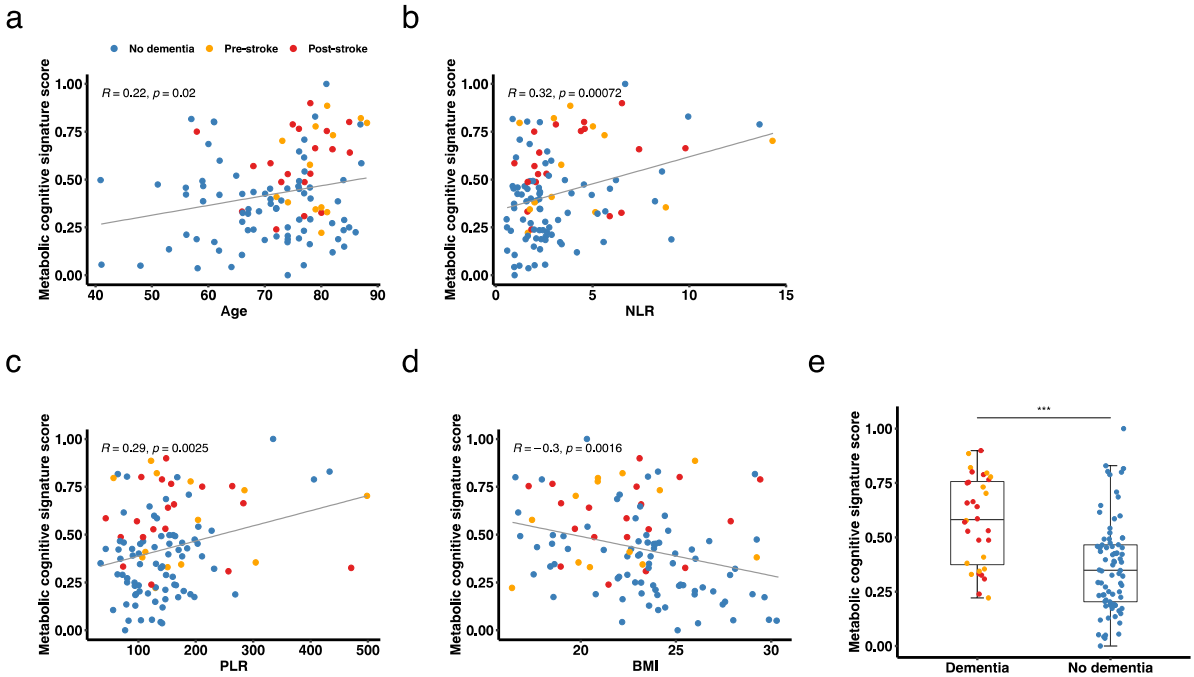


Figure 8. Correlations between the metabolic cognitive signature score and clinical features.

Age (a), NLR (b) and PLR (c) of patients with stroke are significantly correlated with a higher score. BMI is negatively correlated with score (d). The *p*-values are calculated using Pearson correlation analysis. (e) A significant difference is observed between dementia and non-dementia stroke patients, with age, NLR and BMI as cofactors in the general linear model (*** $P < 0.001$).

Survival analysis for post-stroke dementia

The median follow-up period was 18 months (interquartile range, 9 − 32 months), and the median overall survival for post-stroke dementia was not reached. The optimal cut-off value for the metabolic cognitive signature score was 0.48, as determined by the Cantal and O'Quigley method. A score of less than 0.48 yielded a significant survival advantage over the score of ≥ 0.48 (Figure 9). On univariable analysis using Cox proportional hazards regression, the metabolic cognitive signature score, age, LNR and BMI were significant factors for PSD. On multivariable analysis, the metabolic cognitive signature score was found to be independent of age, NLR and BMI. The results indicated that only the metabolic cognitive signature score increased the risk of PSD (hazard ratio, 10.12; 95% CI, 3.3 − 31.02; $P < 0.001$) after adjustment for other key variables (Table 2).

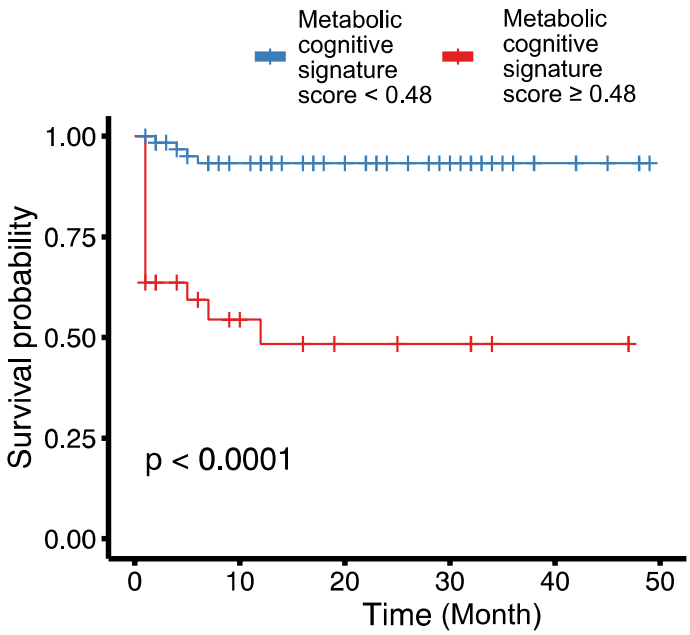


Figure 9. Kaplan–Meier plot and log-rank test.

The metabolic cognitive signature score less than 0.48 (blue) yields a significant survival advantage over the score of ≥ 0.48 (red).

Table 2. Cox proportional hazard model for determining risk factors of dementia after stroke

	Univariable analysis		Multivariable analysis (backward deletion)	
	Hazard ratio (95% CI)	*P*	Hazard ratio (95% CI)	*P*
Age (≥ 71 y vs. < 71 y)	3.8 (1.11 – 13.03)	**0.034**	2.54 (0.73 – 8.86)	0.064
Sex (women vs. men)	0.86 (0.32 – 2.25)	0.751	NA	NA
NLR (≥ 1.94 vs. < 1.94)	3.45 (1 – 11.86)	**0.049**	3.92 (1.13 – 13.66)	0.077
BMI (≥ 23.6 vs. < 23.6)	0.29 (0.1 – 0.87)	**0.028**	0.46 (0.15 – 1.42)	Eliminated
Diabetes (positive vs. negative)	1.71 (1.09 – 4.26)	0.248	NA	NA
Hypertension (positive vs. negative)	1.8 (0.42 – 7.8)	0.431	NA	NA
Metabolic cognitive signature score (≥ 0.48 vs. < 0.48)	10.90 (3.59 – 33.09)	**< 0.001**	10.12 (3.3 – 31.02)	**< 0.001**

Part II. Gene expression signature related to DL-based cognitive signature

Conventional differential gene expression analysis

Before constructing weighted correlation network, conventional differential gene expression was performed with pre-processed microarray data. With threshold of 0.48 for DL-based cognitive signature, no differentially expressed gene was confirmed to show false discovery rate (FDR)-adjusted *p*-values less than 0.05. Table 3 and Table 4 listed 10 top-ranked genes by descending FDR-adjusted *p*-values and descending log fold change (logFC), respectively.

Construction of weighted correlation network

Four outliers were detected in the microarray dataset of ADNI WGS cohort by sample clustering of the distance between samples, thus a total of 24,198 genes from 740 samples were used to construct a weighted correlation network.

Then, a scale-free topology method determined the soft-thresholding power $\beta = 8$, lowest possible β that leads to an approximately scale-free network topology (Figure 10). Finally, 24,198 genes were divided into 14 modules using the hierarchical clustering algorithm (Figure 11); 1,031 genes in the black, 4,403 genes in the blue, 3,282 genes in the brown, 1,143 genes in the green, 114 genes in the greenyellow, 3,664 genes in the grey, 193 genes in the magenta, 571 genes in the pink, 147 genes in the purple, 1,118 genes in the red, 31 genes in the salmon, 98 genes in the tan, 6,829 genes in the turquoise, and 1,573 genes in the yellow modules. The genes that were not grouped into a module fell into the grey module, and were removed from subsequent analyses.

Table 3. 10 top-ranked genes by descending FDR-adjusted *p*-values

Probe Set	Symbol	LogFC	AveExpr	*P*	FDR adjusted *P*
11717482_at	PDHA1	0.09983678	6.127200	4.430370E-05	0.4841542
11741086_x_at	APLP2	-0.06372404	10.811135	4.909628E-05	0.4841542
11718158_a_at	NMI	-0.10769525	8.469821	9.089896E-05	0.4841542
11739526_a_at	LPP	-0.11243598	7.934156	1.020640E-04	0.4841542
11743431_at	LYRM9	0.08813719	7.574373	1.117055E-04	0.4841542
11736631_a_at	APOBEC3H	-0.09818174	4.449792	1.516659E-04	0.5477919
11730791_at	SVIP	0.10954025	5.076255	1.886217E-04	0.5839458
11730788_a_at	APLP2	-0.05379736	10.822330	2.639093E-04	0.5869283
11741085_a_at	APLP2	-0.08425409	8.782874	2.727038E-04	0.5869283
11718896_x_at	SPAG9	-0.13226716	8.588381	2.784574E-04	0.5869283

LogFC, estimate of the log2-fold-change corresponding to the effect or contrast; AveExpr, average log2-expression for the probe over all arrays and channels; FDR, False discovery rate

Table 4. 10 top-ranked genes by descending absolute LogFCs

Probe Set	Symbol	LogFC	AveExpr	*P*	FDR adjusted *P*
11756083_x_at	ENSG00000206239	0.8459395	6.445175	0.0037434283	0.6235441
11716411_x_at	RPS4Y1	-0.6905994	7.621648	0.0188327911	0.6235441
11757733_s_at	XIST	0.6668309	6.180074	0.0390940015	0.6354166
11757857_s_at	XIST	0.6589034	6.223420	0.0245057450	0.6257225
11754194_s_at	XIST	0.6548379	5.831351	0.0317420346	0.6334085
11726814_x_at	KDM5D	-0.5190415	5.746248	0.0219669687	0.6235441
11724075_a_at	DDX3Y	-0.5139931	5.359906	0.0165533320	0.6235441
11720807_x_at	EIF1AY	-0.4990122	5.599486	0.0333203140	0.6347627
11745012_a_at	KDM5D	-0.4655602	5.040483	0.0191147760	0.6235441
11725295_s_at	USP9Y	-0.4393577	4.083355	0.0104322307	0.6235441

LogFC, estimate of the log2-fold-change corresponding to the effect or contrast; AveExpr, average log2-expression for the probe over all arrays and channels; FDR, False discovery rate

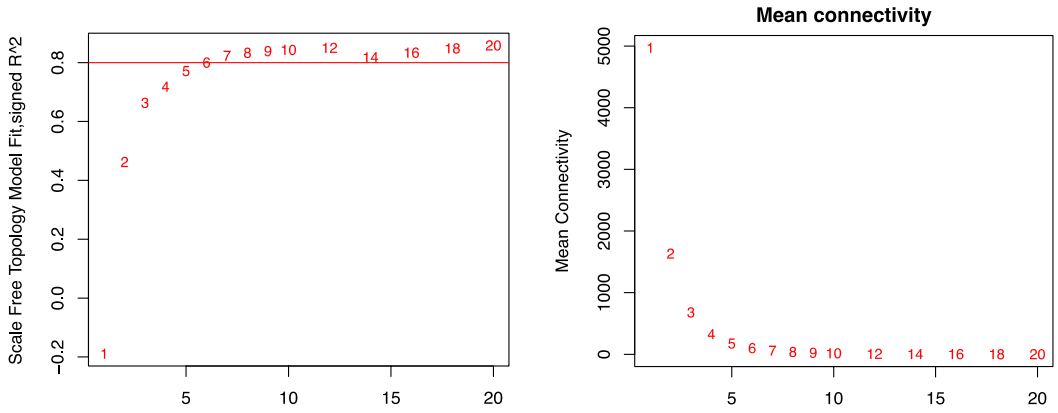


Figure 10. Scale free topology model fitting and mean connectivity by raising the power value from 1 to 20.

Since there is a trade-off between scale-free topology model fit and a high mean number of connections, the lowest power at which the saturation of scale-free topology above 0.80 is 8 was picked to prevent loss of information.

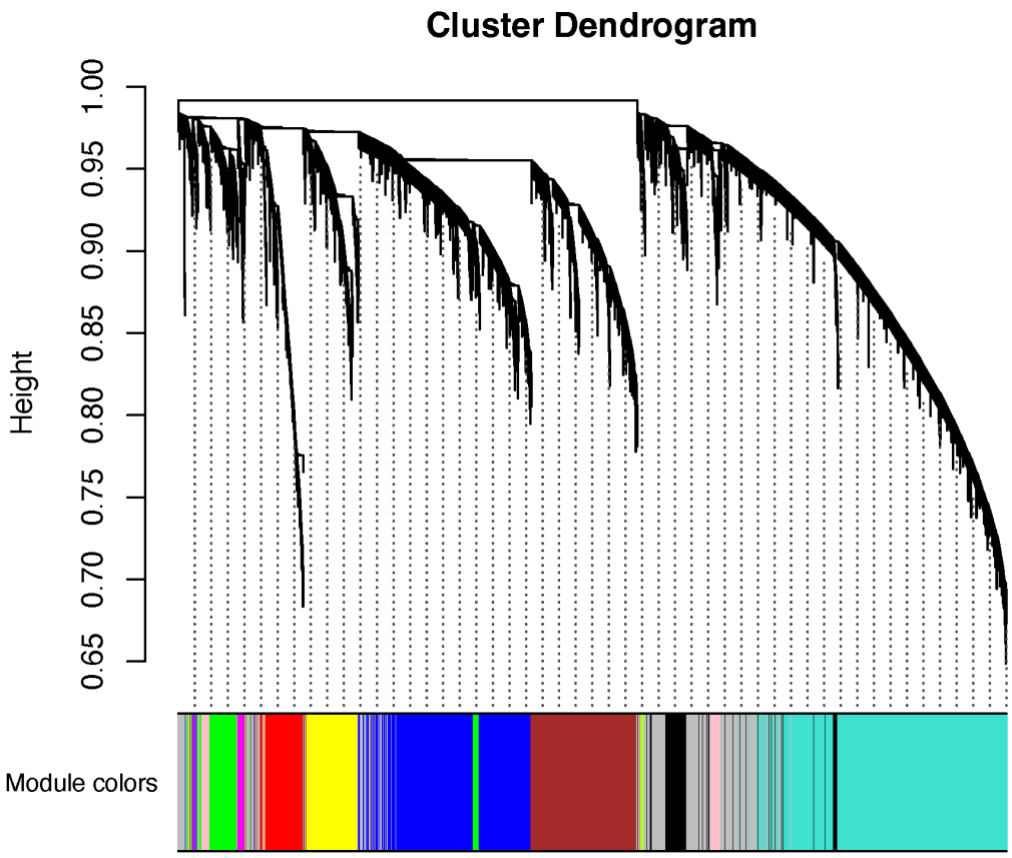


Figure 11. Hierarchical cluster dendrogram to divide genes into modules.

24,198 genes are divided into 14 modules; black, blue, brown, green, greenyellow, grey, magenta, pink, purple, red, salmon, tan, turquoise, yellow modules by hierarchical clustering algorithm.

Identification of module–trait relationships and key modules

GS for DL-based cognitive signature of each module was observed to be high as a positive value in the black and brown modules and as a negative value in the greenyellow module (Figure 12). Modules of which ME showed a significant correlation with DL-based cognitive signature were black, greenyellow, pink, red, tan, and brown modules (Figure 13). Among them, black, greenyellow, and brown modules were significantly correlated with dementia status, as well (black, $R = 0.079$, $P = 0.03$; greenyellow, $R = -0.19$, $P < 0.001$; brown, $R = 0.095$, $P = 0.009$, Figure 13). Whether ME, the first principal component, is eligible for the representative value of each module, was confirmed by the scatter plot and there was a significant positive correlation between the ME and GS for DL-based cognitive signature in all three modules (black, $R = 0.42$, $P < 0.001$; greenyellow, $R = 0.62$, $P < 0.001$; brown, $R = 0.066$, $P < 0.001$, Figure 14). Since the purpose of this study is to discover a gene expression signature related to cognitive function, these three modules were used for further analysis. Black and brown modules, which showed a positive correlation with the metabolic signature and dementia status, significantly correlated with amyloid deposition evaluated by AV45 PET, serum TAU, and PTAU. Additionally, the brown module was associated not only with the patient's age, but also with the history of hypertension and cardiovascular disorder, which are risk factors for stroke. On the contrary, the greenyellow module, which showed a negative correlation with the metabolic signature and dementia status, also negatively correlated with age, history of hypertension, metabolic disease, and amyloid deposit (Figure 13).

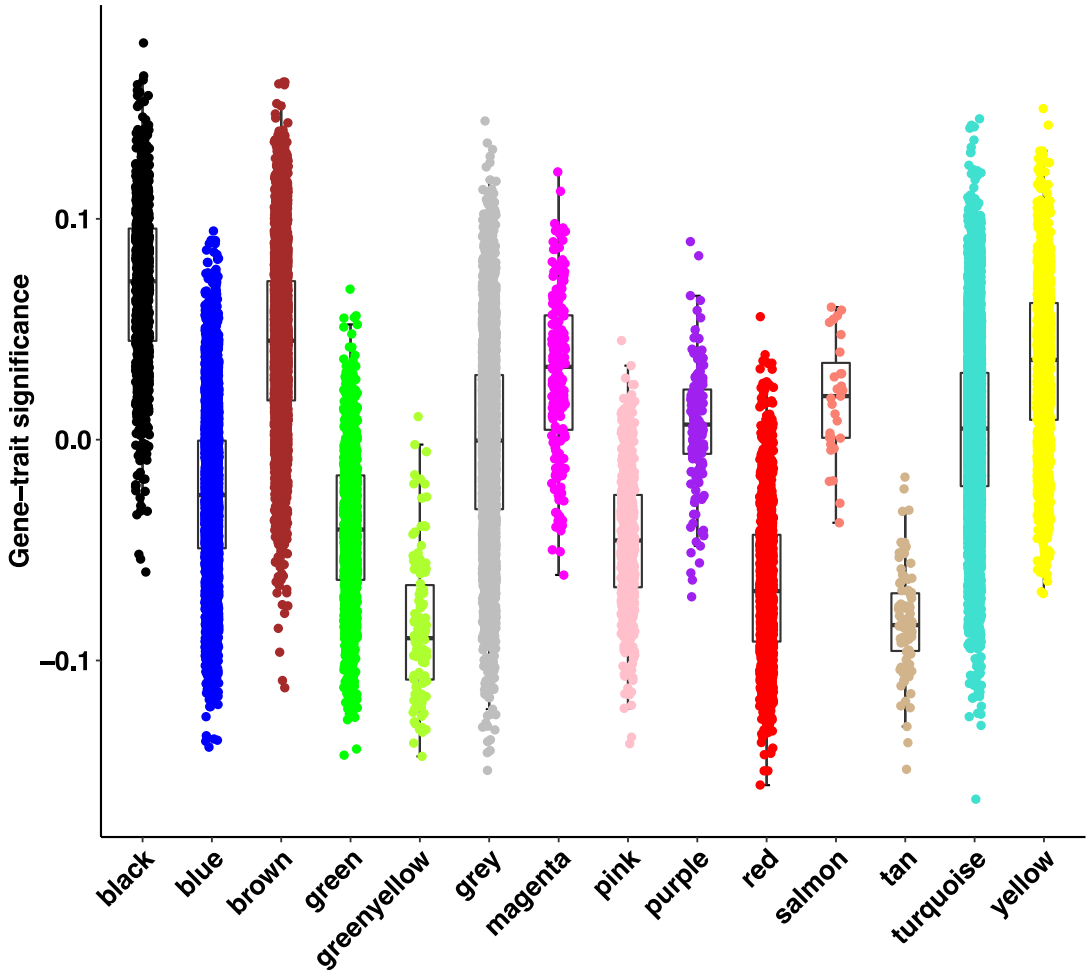


Figure 12. Gene − DL-based cognitive signature significance of each module.

Gene significance for DL-based cognitive signature is high as a positive value in the black and brown modules and as a negative value in the greenyellow module.

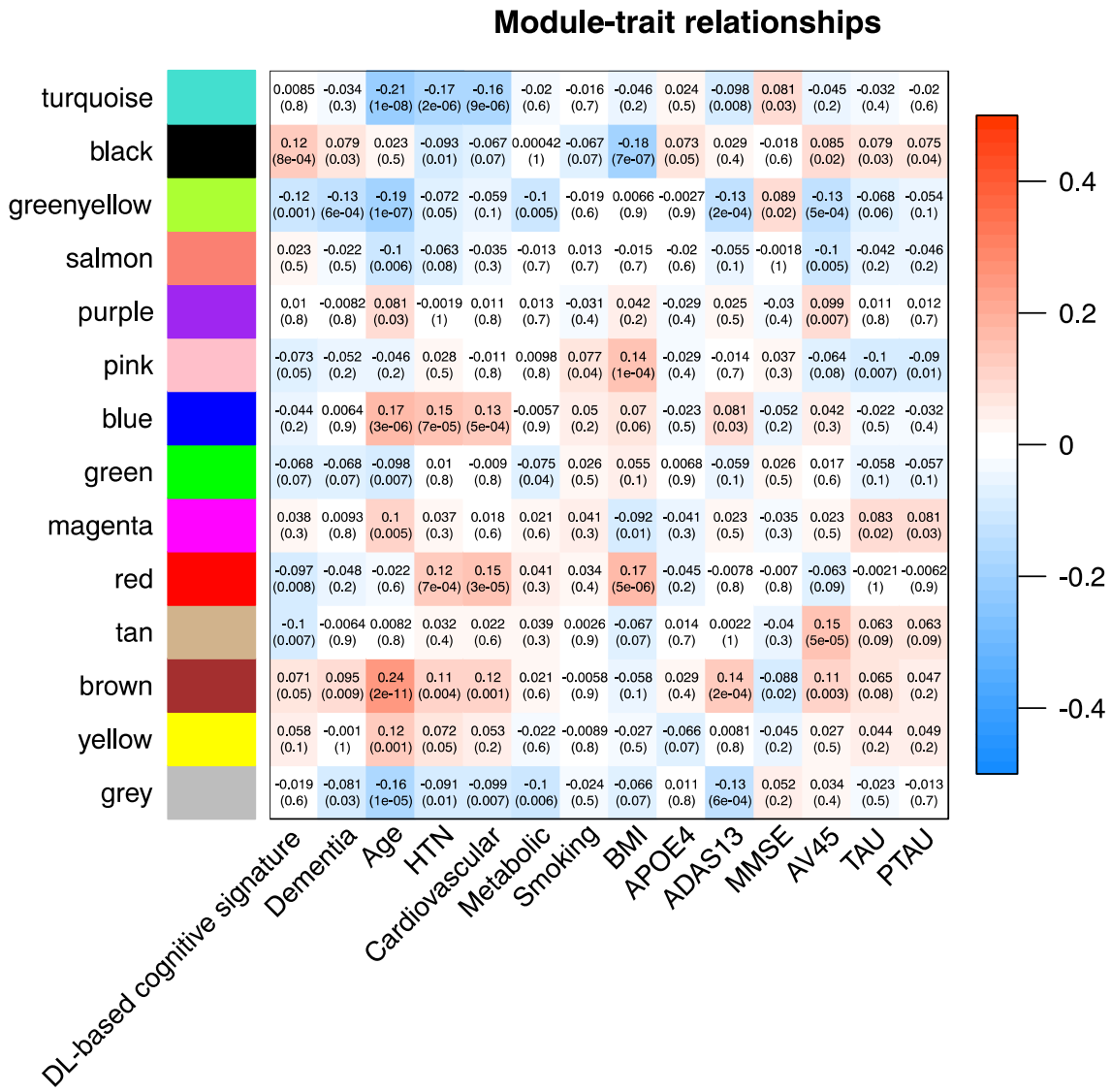


Figure 13. Heatmap presenting module – clinical features relationships.

The black, greenyellow, and brown modules are significantly correlated with both DL-based cognitive signature and dementia status. The black and brown modules significantly correlate with amyloid deposition evaluated by AV45 PET, serum TAU, and PTAU. Furthermore, the brown module is associated with age, history of hypertension and cardiovascular disorder. The greenyellow module negatively correlates with age, history of hypertension, metabolic disease, and amyloid deposit.

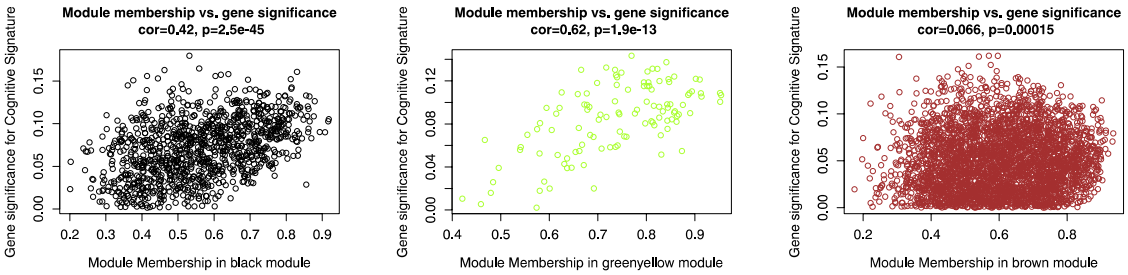


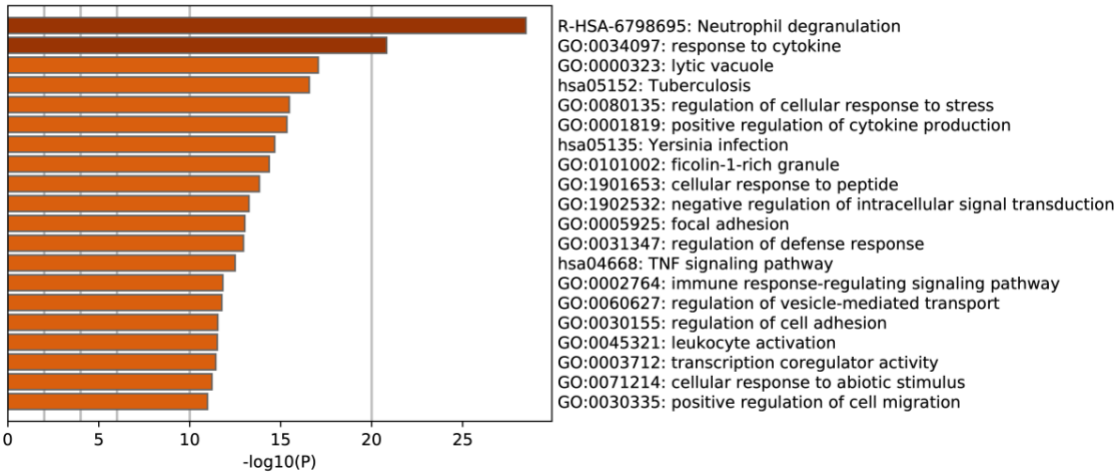
Figure 14. Module eigengene (membership) vs. gene significance for DL-based cognitive signature of black, greenyellow, and brown modules.

There is a significant positive correlation between the ME and GS for DL-based cognitive signature in all three modules. Therefore, the MEs of the three modules are considered to be representative value for gene expression in each module.

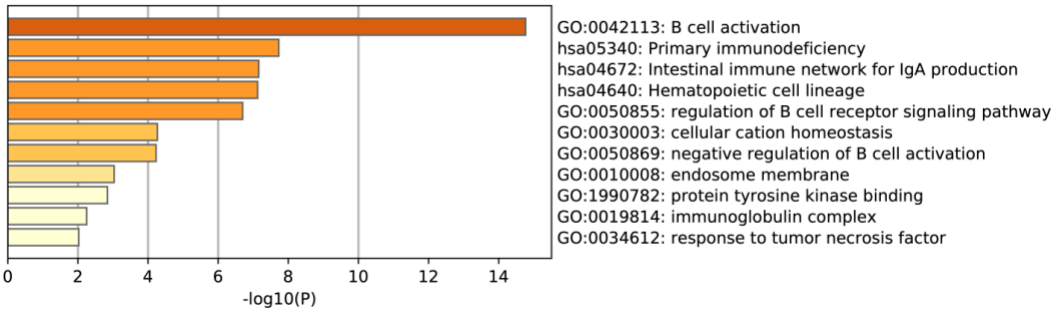
Enrichment analyses of genes in the black, greenyellow, and brown modules

Enrichments analyses were conducted for genes of black, greenyellow, and brown modules which showed significant correlation with CNN-derived metabolic cognitive signature and dementia status. The 20 top-score clusters of terms were visualized in the Figure 15a, b, and c (black, greenyellow, and brown, respectively) by descending *p*-values. Most of the enriched ontology terms in black and brown modules were related to inflammation, leukocyte, especially neutrophil (Figure 15a and c), while greenyellow module was associated with lymphocyte, B-cell activation (Figure 15b).

a


R-HSA-6798695: Neutrophil degranulation
GO:0034097: response to cytokine
GO:0000323: lytic vacuole
hsa05152: Tuberculosis
GO:0080135: regulation of cellular response to stress
GO:0001819: positive regulation of cytokine production
hsa05135: Yersinia infection
GO:0101002: ficolin-1-rich granule
GO:1901653: cellular response to peptide
GO:1902532: negative regulation of intracellular signal transduction
GO:0005925: focal adhesion
GO:0031347: regulation of defense response
hsa04668: TNF signaling pathway
GO:0002764: immune response-regulating signaling pathway
GO:0060627: regulation of vesicle-mediated transport
GO:0030155: regulation of cell adhesion
GO:0045321: leukocyte activation
GO:0003712: transcription coregulator activity
GO:0071214: cellular response to abiotic stimulus
GO:0030335: positive regulation of cell migration
0
5
10
15
20
25
-log10(P)

b


GO:0042113: B cell activation
hsa05340: Primary immunodeficiency
hsa04672: Intestinal immune network for IgA production
hsa04640: Hematopoietic cell lineage
GO:0050855: regulation of B cell receptor signaling pathway
GO:0030003: cellular cation homeostasis
GO:0050869: negative regulation of B cell activation
GO:0010008: endosome membrane
GO:1990782: protein tyrosine kinase binding
GO:0019814: immunoglobulin complex
GO:0034612: response to tumor necrosis factor
0
2
4
6
8
10
12
14
-log10(P)

c

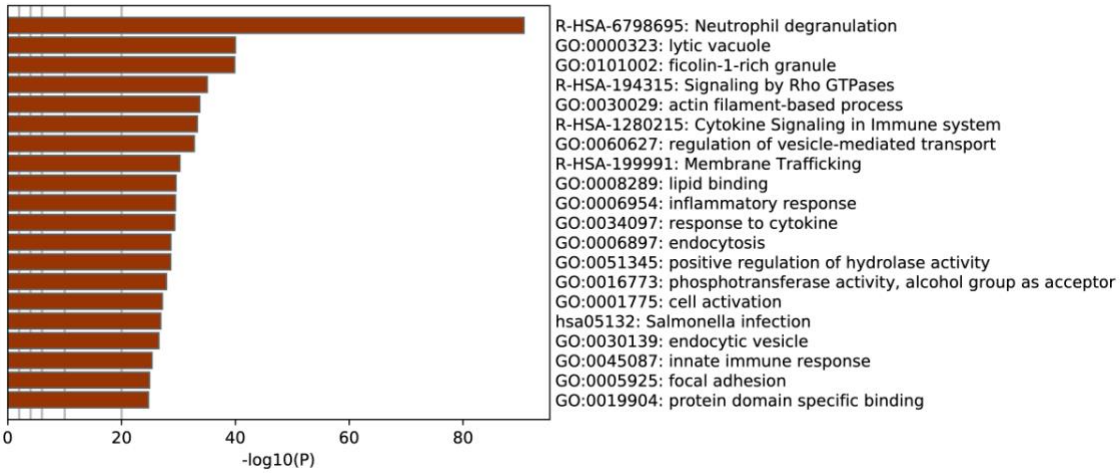

R-HSA-6798695: Neutrophil degranulation
GO:0000323: lytic vacuole
GO:0101002: ficolin-1-rich granule
R-HSA-194315: Signaling by Rho GTPases
GO:0030029: actin filament-based process
R-HSA-1280215: Cytokine Signaling in Immune system
GO:0060627: regulation of vesicle-mediated transport
R-HSA-199991: Membrane Trafficking
GO:0008289: lipid binding
GO:0006954: inflammatory response
GO:0034097: response to cytokine
GO:0006897: endocytosis
GO:0051345: positive regulation of hydrolase activity
GO:0016773: phosphotransferase activity, alcohol group as acceptor
GO:0001775: cell activation
hsa05132: Salmonella infection
GO:0030139: endocytic vesicle
GO:0045087: innate immune response
GO:0005925: focal adhesion
GO:0019904: protein domain specific binding
0
20
40
60
80
-log10(P)

Figure 15. Enrichments analyses for genes of black, greenyellow, and brown modules.

The 20 top-score clusters of terms of black (a), greenyellow (b), and brown (c) are visualized by descending *p-value*s. Most enriched ontology terms are related to the neutrophil degranulation in black and brown modules and B-cell activation in greenyellow module.

Discussion

In this study, it was suggested that a metabolic cognitive signature predicts cognitive outcomes in patients with stroke. The transferred model trained by patients with AD could differentiate dementia from non-dementia in patients with stroke. When CNN-derived similarity of FDG PET patterns was visualized in the two-dimensional space, PET images of post-stroke dementia were found to be clustered, while those of pre-stroke dementia were not. Furthermore, it was found that the metabolic cognitive signature score from the model was related to patients' clinical features and was the only independent risk factor for the occurrence of post-stroke dementia. Additionally, the DL-based cognitive signature derived from the current study was related to the subjects' inflammatory condition, which was measured by gene expression profile of peripheral blood as well as complete blood count.

FDG PET imaging has been used in a large number of studies to investigate brain glucose metabolism in neurodegenerative diseases. And recently, several studies have adopted deep learning for the PET modality to find evidence of neurodegenerative diseases (8, 23–29). However, no published study has evaluated PSD using FDG PET or deep learning in patients with stroke. The metabolic cognitive signature developed in the current study was able to distinguish patients with pre-/post-stroke dementia after learning the FDG PET pattern of cognitive dysfunction in patients with AD.

To aid the interpretation of the CNN model that has an intrinsic black-box problem, the current study visualized CNN-derived features using dimension reduction and CAM methods. Through dimension reduction of 128 features, which are outputs of the global average pooling layer of the model, the parametric t-SNE could visualize the distribution of all participants' cognitive states based on glucose metabolism. As expected, individuals with dementia and non-dementia were separately clustered in the ADNI and stroke cohort dataset, respectively. However, the distribution of stroke cohort is relatively located on borders and clustered compared to that of ADNI cohort because stroke cohort dataset is unseen data (out-of-distribution samples) for the model. However, the overall distribution patterns are similar to each other (right dementia,

left non-dementia), and no outlier was identified. Notably, most patients with post-stroke dementia were clustered with a greater metabolic cognitive signature score, while those with pre-stroke dementia were dispersed across the two-dimensionally embedded plot.

This difference suggests a difference in brain metabolism and neuronal dysfunction in PSD compared with pre-stroke dementia. Owing to the relatively small number of patients with pre-and post-stroke dementia, further studies with a larger cohort are warranted to evaluate the metabolic differences between the two subgroups. By generating a heatmap from the CAM, this study could visualize which brain region contributes the most to decision of the model for classification. It was observed that the extent and location of the regions associated with cognitive decline differed among patients.

The current study demonstrated the clinical feasibility of deep learning-based cognitive signature by evaluating its relationship with stroke survivors' clinical variables and identifying it as an independent risk factor for post-stroke dementia. It is a well-recognized fact that older age is associated with cognitive decline. Similarly, a higher metabolic cognitive signature score was significantly correlated with a higher age of patients with stroke. It was also found that a higher metabolic cognitive signature score was associated with higher LNR and PLR, which suggests that the inflammatory process might play a role in cognitive dysfunction.

The metabolic cognitive signature score exhibited a significant difference between stroke patients with dementia and those without dementia, with age, NLR and BMI as cofactors in the general linear model. Moreover, in the multivariable Cox regression analysis, a high metabolic cognitive signature score was the only significant risk factor for post-stroke dementia. These findings suggest that FDG PET can be a useful screening tool for cognitive dysfunction in patients with stroke. Additionally, the metabolic cognitive signature might be used as an objective biomarker for cognitive impairment in clinical trials of cerebrovascular diseases as well as neurodegenerative disorders.

Dementia after a stroke often consists of a mixture of vascular insults and neurodegenerative processes (30). The processes associated with neurodegenerative and ischemic vascular disease precipitate overlapping pathogenic and molecular changes that eventually prompt neuronal damage and cognitive

impairment (31). Therefore, rather than distinguishing dementia as vascular vs. neurodegenerative, it is worth noting that the capture of metabolic alterations in the brain, which are derived from complex pathological processes, might predict dementia in patients with stroke by using brain FDG PET images taken at the stroke onset. Although there is no established medical practice for preventing cognitive impairment after stroke, patients at risk of cognitive dysfunction assessed by objective deep CNN-based biomarkers could benefit from intensified secondary prevention schemes, including medical interventions and lifestyle modification.

Conventional differential gene expression analysis, which was performed prior to WGCNA application, compared logFCs in groups dichotomized by the threshold of DL-based cognitive signature of 0.48, found no differentially expressed gene. This may be due to the loss of information and is a limitation of conventional methods. Then, WGCNA was applied to investigate the relationship between DL-based imaging cognitive signature and gene expression signature of peripheral blood. The widely used unweighted network defines the network adjacency of gene-expression similarity using hard thresholding, which does not reflect the continuous nature of the co-expression of the gene expression profile. In contrast, the weighted network introduced in this study could prevent informational loss by taking continuous variables between 0 and 1 for the adjacency matrix by using soft thresholding.

A total of 13 modules were identified as a result of WGCNA analysis, of which black, brown, and greenyellow modules showed significant relationships with both DL-based cognitive signature and dementia status. Unlike other modules, the three modules were also related to amyloid deposition, TAU/PTAU, and risk factors for stroke; thus, they can be considered to be modules related to post-stroke dementia. In the enrichment analysis of these three modules, the black/brown modules, which showed a positive correlation with DL-based cognitive signature and dementia status, were associated with degranulation of neutrophil; and the greenyellow module, which showed a negative correlation with DL-based cognitive signature and dementia status, exhibited marked enrichment in lymphocyte activation. These results are consistent with the positive correlation between DL-based cognitive signature and NLR in the stroke cohort of this study.

Several investigations have demonstrated that sustained inflammation is a critical characteristic of neurodegenerative diseases (31–34). Furthermore, inflammation and/or inflammatory signaling is associated with risk factors for AD, including age, cardiovascular diseases and metabolic diseases (35). In this study, the CNN-derived cognition-related imaging phenotype was significantly correlated with patients' inflammatory condition (specifically, high neutrophil and low-lymphocyte) which was identified not only in the complete blood counts but also in the gene expression profile of peripheral blood.

This study has several limitations. First, since this was a post hoc analysis conducted to confirm the recurrence of vascular events after stroke, the current study did not perform systematic monitoring of cognitive function in all patients, but in patients diagnosed with dementia before stroke or suspected dementia after stroke. Second, all FDG PET images of ADNI/stroke cohort used for training/test of the model had been spatially normalized, and the image of stroke cohort with wide range of metabolic defects caused by ischemic burden may give rise to inaccuracy in the spatial normalization process which used brain template. Third, despite successful transfer learning, there is still an issue of data size. However, as the similarity of the institution's dataset to the ADNI dataset is outstanding compared with that of natural image databases, such as ImageNet, the CNN model appears to appropriately transfer the image pattern of AD learned from the ADNI dataset to the stroke dataset.

Case review regarding the survival analysis

Representative images of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) for PSD diagnosis were examined by setting the threshold of the metabolic cognitive signature score as 0.48, as applied by the Log rank test and Cox regression analysis. The spatially normalized FDG PET images of the highest score among positive cases with a score of 0.48 or higher and the lowest score among negative cases were visualized, respectively. As confirmed in the Figure 7a, the averaged CAM of ADNI cohort, a symmetric hypometabolism in the bilateral parieto-temporal lobe was

confirmed in the TP (Figure 16a) and FP (Figure 16b) cases as well. On the other hand, in the TN (Figure 16c) case, the glucose metabolism of the brain was preserved without a corresponding pattern. Interestingly, although Figure 16e is the same TN case, right cerebral hemisphere showed asymmetric metabolic decrease, and in this patient, acute recurrence of stroke developed 16 days later. In image of FN case (Figure 16d) where model tested negative but PSD occurred, asymmetric hypometabolism also appeared in the right cerebral hemisphere, while the metabolism of the left parietal lobe was preserved. Taken together, since this model learned the metabolic pattern of AD patients in ADNI cohort, TP/FP cases can be considered closer to the AD-like pattern, and TN cases, particularly stroke recurrence case and FN case can be deemed closer to the vascular dementia-like pattern. However, as confirmed by survival analysis, the model successfully predicted post-stroke dementia using the brain FDG PET obtained at the onset of stroke, and FP cases may be high-risk patients who may develop PSD in further follow-up.

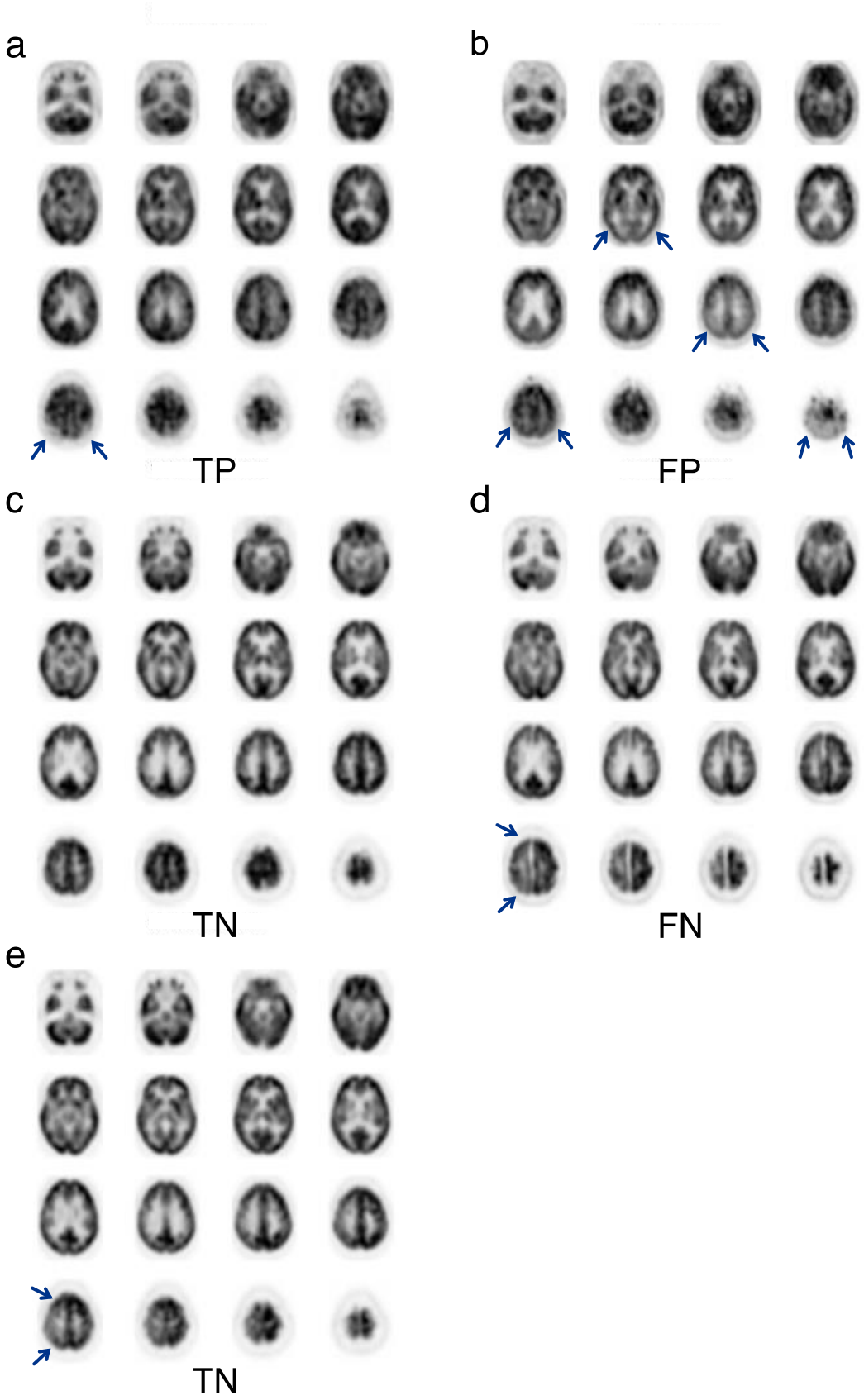

a
TP
b
FP
c
TN
d
FN
e
TN

Figure 16. Spatially normalized FDG PET images of TP, FP, TN, FN cases

(a) TP case, DL-based score of 0.899; a 78-year-old woman with both pICA and dICA stenosis presented with left MCA infraction. 18 months after the stroke onset, post-stroke dementia developed.

(b) FP case, DL-based score of 1.0; a 81-year-old man with left pICA mild stenosis presented with BA occlusion. The follow-up period was 9 months.

(c) TN case, DL-based score of 0.043; a 64-year-old woman with right MCA occlusion hospitalized due to transient ischemic attack. The follow-up period was 13 months.

(d) FN case, DL-based score of 0.239; a 72-year-old man with right pICA occlusion presented with right MCA infarction. Post-stroke dementia developed after 5 months of stroke onset.

(e) TN case, DL-based score of 0.055; a 41-year-old man with right MCA occlusion presented with right MCA infarction. 16 days later, he underwent stroke recurrence. The follow-up period was 5 months.

TP, true positive; FP, false positive; TN, true negative; FN, false negative; DL, deep learning; pICA, proximal internal carotid artery; dICA, distal internal carotid artery; BA, basilar artery; MCA, middle cerebral artery

Conclusions

The proposed CNN model differentiating AD from NC was successfully transferred to an independent stroke cohort. The deep learning-based cognitive signature is associated with clinical variables in patients with stroke and is an independent risk factor for dementia following stroke. Furthermore, this study confirmed that inflammatory condition measured by gene expression profile of peripheral blood as well as complete blood counts are deeply related to the DL-based imaging phenotype of cognitive function. CNN-based cognitive evaluation using FDG PET may be utilized as an objective biomarker for cognitive dysfunction in patients with cerebrovascular diseases as well as neurodegenerative disorders.

References

1. Gorelick PB, Nyenhuis D. Stroke and Cognitive Decline. JAMA. 2015;314:29-30.

2. Levine DA, Galecki AT, Langa KM et al. Trajectory of Cognitive Decline After Incident Stroke. JAMA. 2015;314:41-51.

3. Mosconi L. Brain glucose metabolism in the early and specific diagnosis of Alzheimer's disease. FDG-PET studies in MCI and AD. Eur J Nucl Med Mol Imaging. 2005;32:486-510.

4. Shivamurthy VK, Tahari AK, Marcus C, Subramaniam RM. Brain FDG PET and the diagnosis of dementia. AJR Am J Roentgenol. 2015;204:W76-85.

5. Jin Y, Di Legge S, Ostbye T, Feightner JW, Hachinski V. The reciprocal risks of stroke and cognitive impairment in an elderly population. Alzheimer's & Dementia. 2006;2:171-178.

6. Vinters HV, Zarow C, Borys E et al. Review: Vascular dementia: clinicopathologic and genetic considerations. Neuropathol Appl Neurobiol. 2018;44:247-266.

7. Ihle-Hansen H, Thommessen B, Wyller TB et al. Incidence and subtypes of MCI and dementia 1 year after first-ever stroke in patients without pre-existing cognitive impairment. Dement Geriatr Cogn Disord. 2011;32:401-407.

8. Choi H, Kim YK, Yoon EJ, Lee JY, Lee DS, Alzheimer's Disease Neuroimaging I. Cognitive signature of brain FDG PET based on deep learning: domain transfer from Alzheimer's disease to Parkinson's disease. Eur J Nucl Med Mol Imaging. 2020;47:403-412.

9. Xiuli L, Hao Z, Xiaolu Z, Hao L, Guotong X. Exploring transfer learning for gastrointestinal bleeding detection on small-size imbalanced endoscopy images. Annu Int Conf IEEE Eng Med Biol Soc. 2017;2017:1994-1997.

10. Oquab M, Bottou L, Laptev I, Sivic J. Learning and Transferring Mid-level Image Representations Using Convolutional Neural Networks. 2014 IEEE Conference on Computer Vision and Pattern Recognition. 20141717-1724.

11. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. 20143320-3328.

12. Lee R, Choi H, Park KY, Kim JM, Seok JW. Prediction of post-stroke cognitive impairment using brain FDG PET: deep learning-based approach. Eur J Nucl Med Mol Imaging. 2022;49:1254-1262.

13. Geschwind DH, Konopka G. Neuroscience in the era of functional genomics and systems biology. Nature. 2009;461:908-915.

14. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

15. Van der Maaten L, Hinton G. Visualizing data using t-SNE. Journal of machine learning research. 2008;9

16. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning Deep Features for Discriminative Localization. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016

17. Albert R, Barabási A-L. Statistical mechanics of complex networks. Reviews of Modern Physics. 2002;74:47-97.

18. Zhang. A General Framework for Weighted Gene Co-expression Network Analysis.

19. Li A, Horvath S. Network neighborhood analysis with the multi-node topological overlap measure. Bioinformatics. 2007;23:222-231.

20. Yip AM, Horvath S. Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinformatics. 2007;8:22.

21. Zhou Y, Zhou B, Pache L et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun. 2019;10:1523.

22. Contal C, O'Quigley J. An application of changepoint methods in studying the effect of age on survival in breast cancer. Computational Statistics & Data Analysis. 1999;30:253-270.

23. Choi H, Jin KH, Alzheimer's Disease Neuroimaging I. Predicting cognitive decline with deep learning of brain metabolism and amyloid imaging. Behav Brain Res. 2018;344:103-109.

24. Ding Y, Sohn JH, Kawczynski MG et al. A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using (18)F-FDG PET of the Brain. Radiology. 2019;290:456-464.

25. Huang Y, Xu J, Zhou Y, Tong T, Zhuang X, Alzheimer's Disease Neuroimaging I. Diagnosis of Alzheimer's Disease via Multi-Modality 3D Convolutional Neural Network. Front Neurosci. 2019;13:509.

26. Kim JY, Suh HY, Ryoo HG et al. Amyloid PET Quantification Via End-to-End Training of a Deep Learning. Nucl Med Mol Imaging. 2019;53:340-348.

27. Liu H, Nai YH, Saridin F et al. Improved amyloid burden quantification with nonspecific estimates using deep learning. Eur J Nucl Med Mol Imaging. 2021;48:1842-1853.

28. Liu M, Cheng D, Wang K, Wang Y, Alzheimer's Disease Neuroimaging I. Multi-Modality Cascaded Convolutional Neural Networks for Alzheimer's Disease Diagnosis. Neuroinformatics. 2018;16:295-308.

29. Yee E, Popuri K, Beg MF, Alzheimer's Disease Neuroimaging I. Quantifying brain metabolism from FDG-PET images into a probability of Alzheimer's dementia score. Hum Brain Mapp. 2020;41:5-16.

30. Brainin M, Tuomilehto J, Heiss WD et al. Post-stroke cognitive decline: an update and perspectives for clinical research. Eur J Neurol. 2015;22:229-38, e13.

31. Amor S, Peferoen LA, Vogel DY et al. Inflammation in neurodegenerative diseases--an update. Immunology. 2014;142:151-166.

32. Amor S, Puentes F, Baker D, van der Valk P. Inflammation in neurodegenerative diseases. Immunology. 2010;129:154-169.

33. Chen WW, Zhang X, Huang WJ. Role of neuroinflammation in neurodegenerative diseases (Review). Mol Med Rep. 2016;13:3391-3396.

34. Glass CK, Saijo K, Winner B, Marchetto MC, Gage FH. Mechanisms underlying inflammation in neurodegeneration. Cell. 2010;140:918-934.

35. Kinney JW, Bemiller SM, Murtishaw AS, Leisgang AM, Salazar AM, Lamb BT. Inflammation as a central mechanism in Alzheimer's disease. Alzheimers Dement (N Y). 2018;4:575-590.

국문초록

뇌 FDG PET 을 이용한 딥러닝 기반 뇌졸중 후 치매 예측 연구

이리리

서울대학교 대학원

핵의학 전공

뇌졸중후 치매는 뇌졸중을 겪은 환자의 약 3 분의 1 에서 발생한다. 인지 기능은 환자의 삶의 질에 지대한 영향을 끼치므로, 뇌졸중 후 치매의 조기 예측을 위한 객관적이고 정량화 가능한 바이오마커가 필요하다. 뇌 플루오데옥시글루코스(F-18 fluorodeoxyglucose, FDG) 양전자방출촬영(positron emission tomography, PET)은 그 전형적인 대사감소 패턴으로 인지 기능 평가에 널리 이용되어 왔다. 본 연구에서는, 뇌졸중 후 치매를 평가하기 위해 뇌 FDG PET 를 이용한 딥러닝 기반의 시그니처를 개발했다. 또한, 딥러닝 유래 인지기능 시그니처와 유전자 공발현

네트워크 분석(gene co-expression network analysis, WGCNA) 기반 유전자 발현 시그니처간의 연관성을 평가하여 인지 기능과 관련된 영상표현형-유전자형 관계를 설명하였다.

알츠하이머병 신경영상화 이니셔티브 (Alzheimer's Disease Neuroimaging Initiative, ADNI) 데이터베이스의 뇌 FDG PET 를 이용해 알츠하이머병과 정상 대조군을 구별하도록 딥러닝 모델을 개발했다. 그 모델은 곧바로 전향적으로 모집된 뇌졸중 환자 코호트로 전이(transfer)되어, 치매가 있는 환자와 없는 환자를 구별하였다. 모델의 정확도는 수신자 조작 특성 곡선(Receiver operating characteristic curve, ROC)의 곡선 아래 면적(area under the curve, AUC)으로 평가하였다. 딥러닝 기반의 특징의 분포와 모델이 분류를 위해 가중치를 준 뇌의 영역을 시각화 하였다. 딥러닝 모델로부터 나온 인지 시그니처와 임상 변수들 간의 상관관계를 평가했으며, 뇌졸중 환자를 대상으로 뇌졸중 후 치매에 대한 생존분석을 수행했다.

ADNI 의 whole genome sequencing (WGS) 코호트 참가자의 혈액 샘플에서 얻은 마이크로어레이 유전자 발현(microarray gene expression) 데이터를 WGCNA 에 이용하였다. 유전자 발현 프로파일을 전처리 한 후 가중 상관관계 네트워크 분석(weighted correlation network analysis)을 수행하였고, 상호 연관성이 높은 유전자의 클러스터인 유전자 모듈(module)을 확인하였다. 그런 다음, 모듈 유전자 발현과 딥러닝 유래 인지 시그니처를 포함한 임상 특성들 사이의 관계를 평가하여 인지기능 영상표현형과 관련된 모듈을 결정하였다. 관련된 모듈의 내재적인 생물학적 의미를 파악하기 위해 pathway/process enrichment 분석을 적용하였다.

알츠하이머병과 정상 대조군 참가자를 분류하는 AUC-ROC 는 0.94 (95% 신뢰수준, 089-0.98)로 확인되었다. 전이된 모델은 치매를 가진 뇌졸중 환자를 식별하였다(AUC-ROC = 0.75). 딥러닝 모델의 인지기능 감소 시그니처의 점수는 연령, 호중구-림프구 비율 및 혈소판-림프구 비율과 양의 상관관계가 있었으며, 뇌졸중 환자의 체질량 지수와는 음의 상관관계를 보였다. 인지기능 감소 점수는 다른 주요 변수에 대한 보정 후에도 뇌졸중 후 치매에 대한 독립적인 위험인자로 확인되었다.

총 740 명의 참가자의 혈액 샘플의 24,198 개의 유전자가 WGCNA 에서 계층적 클러스터링

알고리즘을 통해 14 개의 모듈로 나뉘었다. 14 개 모듈 중 검은색, 연두색, 분홍색, 빨간색, 황갈색, 갈색 모듈은 딥러닝 기반 인지 시그니처와 유의한 상관관계를 보였다. 이 가운데 검은색, 연두색, 갈색 모듈은 참가자의 치매 상태와도 유의한 상관관계가 있었다. 또한 이 세 모듈은 참가자의 뇌 아밀로이드 침착, TAU/PTAU 농도, 그리고 뇌졸중의 위험인자들과 관계가 있었다. Enrichment 분석에서, 검은색과 갈색 모듈의 풍부한 온톨로지 용어(enriched ontology term)의 대부분은 염증, 백혈구, 특히 호중구와 관련이 있는 반면, 연두색 모듈은 림프구, B 세포 활성화와 관련이 있었다.

FDG PET 를 사용한 딥러닝 기반 인지 시그니처는 독립적인 뇌졸중 코호트로 성공적으로 전이되었다. FDG PET 를 이용한 딥러닝 기반 인지기능 평가가 뇌졸중 후 치매의 객관적인 바이오마커로 활용될 수 있을 것이라 제안한다. 나아가, 본 연구에서는 일반 혈액 검사(complete blood count) 뿐만 아니라 말초혈액의 유전자 발현 프로파일에서 측정된 염증 상태가 인지기능의 딥러닝 기반 영상표현형 및 뇌졸중 위험인자와 깊은 관련이 있음을 확인하였다.

주요어 : 뇌졸중 후 치매, 뇌졸중 후 인지 장애, [^{18}F]FDG, 딥러닝, 유전자 공발현 네트워크

학번 : 2019-32281